



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ  
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Ανασκόπηση Πολυπαραμετρικών Συστημάτων  
Υποστήριξης Κλινικών Αποφάσεων για την  
Πρόγνωση και Κλινική Διαχείριση του Καρκίνου**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αγγελική Σ. Παξινού

**Επιβλέπων :** Διονύσιος-Δημήτριος Κουτσούρης

Καθηγητής Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2014





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ  
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Ανασκόπηση Πολυπαραμετρικών Συστημάτων  
Υποστήριξης Κλινικών Αποφάσεων για την  
Πρόγνωση και Κλινική Διαχείριση του Καρκίνου**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Αγγελική Σ. Παξινού**

**Επιβλέπων :** Διονύσιος-Δημήτριος Κουτσούρης  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 1<sup>η</sup> Σεπτεμβρίου 2014.

.....  
Δ. Δ. Κουτσούρης  
Καθηγητής Ε.Μ.Π.

.....  
Μ. Χαρίτου  
Ερευνήτρια Β' ΕΠΙΣΕΥ-ΕΜΠ

.....  
Α. Πρέντζα  
Επ. Καθηγήτρια Παν. Πειραιώς

Αθήνα, Σεπτέμβριος 2014

.....  
Αγγελική Σ. Παξινού

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Αγγελική Σ. Παξινού, 2014.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

## Περίληψη

Ο καρκίνος αποτελεί μάστιγα της εποχής μας, με τις στατιστικές να τον τοποθετούν ως τη δεύτερη αιτία θανάτου παγκοσμίως. Αν και η έγκαιρη ανίχνευση και διάγνωση αποτελούν τα σημαντικότερα βήματα για την αντιμετώπισή του, η πρόγνωση του καρκίνου παραμένει εξίσου σημαντική. Η πρόγνωση περιλαμβάνει την πρόβλεψη της πιθανότητας εμφάνισης καρκίνου, την πρόβλεψη υποτροπής των ασθενών και την πρόβλεψη επιβίωσης (προσδόκιμο ζωής). Τα παραπάνω αποτελέσαν αντικείμενο μελέτης σε πολλές δημοσιευμένες έρευνες τα τελευταία χρόνια. Οι συγγραφείς-ερευνητές επιχείρησαν, χρησιμοποιώντας μεθόδους μηχανικής μάθησης και αλγόριθμους ταξινόμησης, να αναπτύξουν ολοκληρωμένα συστήματα υποστήριξης αποφάσεων, τα οποία μπορούν να παράγουν μια πρόβλεψη για την κατάσταση ενός ασθενούς με ικανοποιητική, συνήθως, ακρίβεια. Η εν λόγω πρόβλεψη μπορεί να βοηθήσει τους θεράποντες ιατρούς στη λήψη αποφάσεων, εξατομικευμένων για κάθε ασθενή, καθιστώντας τα συστήματα υποστήριξης κλινικών αποφάσεων χρήσιμα κλινικά εργαλεία.

Στην παρούσα διπλωματική εργασία, έγινε μια προσπάθεια να συγκεντρωθούν και να παρουσιαστούν μελέτες που περιλαμβάνουν τέτοιου είδους συστήματα για διαφορετικούς τύπους καρκίνου, να αξιολογηθούν τα αποτελέσματά τους και να εξαχθούν κάποια συμπεράσματα όσον αφορά τη δομή των συστημάτων υποστήριξης κλινικών αποφάσεων, τα δεδομένα που χρησιμοποιήθηκαν για τα σύνολα εκπαίδευσης και επαλήθευσης, τις τεχνικές ταξινόμησης και την απόδοσή τους.

Για την πραγματοποίηση της έρευνάς μας, εξετάστηκαν 11 διαφορετικοί τύποι καρκίνου (καρκίνος του μαστού, καρκίνος του προστάτη, καρκίνος του τραχήλου της μήτρας, καρκίνος του πνεύμονα, καρκίνος του παχέος εντέρου, λευχαιμία, καρκίνος του στόματος, καρκίνος των ωοθηκών, καρκίνος του θυρεοειδούς, καρκίνος του ήπατος, καρκίνος των οστών), καθώς και μια μελέτη πρόβλεψης της αποτελεσματικότητας της χημειοθεραπείας, και καταγράφηκε για κάθε επιμέρους μελέτη ο κλινικός σκοπός (πχ διάγνωση, πρόβλεψη υποτροπής κτλ), τα δεδομένα που χρησιμοποιήθηκαν (πχ κλινικά, απεικονιστικά, γονιδιακά), οι μέθοδοι ταξινόμησης (πχ τεχνητά νευρωνικά δίκτυα, Μπεϋζιανά δίκτυα) και τα γενικά αποτελέσματα (έξοδος του συστήματος και παράμετροι ακρίβειας, ευαισθησίας, ειδικότητας κτλ). Όλα τα παραπάνω παρουσιάζονται σε έναν συγκεντρωτικό πίνακα στο τέλος της μελέτης ώστε να καταστεί δυνατή η επισκόπησή της και να τεθούν υπό συζήτηση συμπεράσματα, περιορισμοί και μελλοντικές συστάσεις.

Αποδείχθηκε ότι η πλειονότητα των συστημάτων υποστήριξης κλινικών αποφάσεων επέδειξε εντυπωσιακά αποτελέσματα πρόβλεψης, γεγονός που ενισχύει την πεποίθησή μας ότι η ενσωμάτωση των συστημάτων αυτών σε κλινικές και νοσοκομειακές εφαρμογές είναι απαραίτητη και θα αποδειχθεί μεγάλης σημασίας στα επόμενα χρόνια.

**Λέξεις κλειδιά :** Συστήματα υποστήριξης κλινικών αποφάσεων, Διάγνωση και πρόγνωση καρκίνου, Μηχανική μάθηση, Αλγόριθμοι ταξινόμησης, Αναγνώριση προτύπων, Τεχνητά νευρωνικά δίκτυα, Μπεϋζιανά δίκτυα

## Abstract

Cancer has proven to be the plague of our time, as it is being placed by the statistics as the second cause of death worldwide. Although early cancer detection and diagnosis are major steps for fighting the disease, cancer prognosis remains equally important. Prognosis relates with the prediction of the likelihood of cancer, as well as the prediction of a patient's relapse or survival (life expectancy). All the above have been the object in several recently published studies. The researchers have attempted to develop integrated clinical decision support systems by using machine learning methods and classification algorithms. These systems are able to produce an accurate prediction of the patient's outcome, which may help clinicians in personalized decision-making.

The aim of this diploma thesis is the collection and presentation of studies involving such systems as the ones mentioned above for different types of cancer, the assessment of their results and the extraction of several conclusions regarding the structure of clinical decision support systems, the data used for their creation, training and validation, the classification techniques used and the systems' performance.

In order to conduct our study, we examined 11 different cancer types (breast cancer, prostate cancer, cervical cancer, lung cancer, colorectal cancer, leukemia, oral cancer, ovarian cancer, thyroid cancer, liver cancer, skeletal cancer), as well as a study regarding the effectiveness of chemotherapy. For each single study, we listed the study's clinical endpoint (eg diagnosis, recurrence), the data used (eg clinical, imaging, genomic), the classification methods (eg artificial neural networks, Bayesian networks) and the system's results (output, accuracy, sensitivity, specificity). All the above are presented in a collective table at the end of the study in order to facilitate the overview of our work and discuss conclusions, limitations and future recommendations.

It has been proved that the majority of the systems discussed here have shown quite impressive results in terms of prediction, a fact reinforcing our belief that clinical decision support systems will be a valuable integral part of clinical practice in the years to come.

**Keywords:** Clinical decision support systems, Cancer diagnosis and prognosis, Machine learning, Classification algorithms, Pattern Recognition, Artificial neural networks, Bayesian networks

## Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Εργαστήριο Βιοϊατρικής Τεχνολογίας της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ του Εθνικού Μετσόβιου Πολυτεχνείου.

Η δομή της εργασίας έχει ως εξής :

Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στην έννοια των συστημάτων υποστήριξης κλινικών αποφάσεων, περιγράφονται τα βήματα της διαδικασίας λήψης αποφάσεων και η δομή των συστημάτων με βάση πέντε βασικούς άξονες (περιεχομένου, γνώσης, υποστήριξης αποφάσεων, παροχής πληροφοριών, ροής εργασιών).

Στο δεύτερο κεφάλαιο προσεγγίζονται οι έννοιες της μηχανικής μάθησης, της αναγνώρισης προτύπων και της ταξινόμησης. Διασαφηνίζονται όροι όπως ο χώρος και η επιλογή των χαρακτηριστικών, ενώ αναλύονται διάφοροι αλγόριθμοι ταξινόμησης, τους οποίους θα συναντήσουμε στο επόμενο κεφάλαιο της εργασίας.

Το τρίτο κεφάλαιο πραγματεύεται την εφαρμογή της μηχανικής μάθησης και των συστημάτων υποστήριξης κλινικών αποφάσεων στην έρευνα για τον καρκίνο. Παρουσιάζονται εκτενώς επιμέρους μελέτες για κάθε μία από τις κατηγορίες καρκίνου και επισημαίνονται βασικά σημεία κάθε μελέτης που χρήζουν προσοχής.

Στο τέταρτο κεφάλαιο λαμβάνει χώρα μια επισκόπηση της μελέτης που διεξήχθη μέσω ενός συγκεντρωτικού πίνακα, ενώ καταγράφονται τα συμπεράσματα, οι περιορισμοί και οι συστάσεις για μελλοντικές έρευνες.

## Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή κ. Δ. Κουτσούρη για την ανάθεση του θέματος της παρούσας εργασίας και για την εμπιστοσύνη που έδειξε στο πρόσωπό μου. Οφείλω να αποδώσω ιδιαίτερες ευχαριστίες στον Υ.Δ. Παναγιώτη Μπούντρη για την πολύτιμη βοήθεια του και το χρόνο που μου αφιέρωσε κατά τη συνεργασία μας, για τη θέρμη που μου μετέδωσε πάνω στο αντικείμενο της Βιοϊατρικής και για την καθοδήγηση που μου προσέφερε σε όλη την πορεία εκπόνησης αυτής της διπλωματικής εργασίας. Ευχαριστώ ιδιαίτερα το μέλος της τριμελούς επιτροπής Δρα. Μαρία Χαρίτου για τις χρήσιμες συμβουλές της και την υποστήριξή της για την περάτωση της παρούσας εργασίας. Επίσης, ευχαριστώ την κα. Ανδριάννα Πρέντζα, μέλος της τριμελούς επιτροπής και Επίκουρη Καθηγήτρια του Πανεπιστημίου Πειραιώς. Θα ήθελα, τέλος, να ευχαριστήσω τους ανθρώπους του περιβάλλοντός μου για τη στήριξή τους.

Αγγελική Σ. Παξινού,  
Αθήνα, Σεπτέμβριος 2014



## Πίνακας περιεχομένων

<b>Κεφάλαιο 1 – Εισαγωγή</b> .....	<b>Error! Bookmark not defined.</b>
<b>1.1 Εισαγωγή στην έννοια των συστημάτων υποστήριξης κλινικών αποφάσεων</b> .....	<b>17</b>
<b>1.2 Λήψη κλινικών αποφάσεων</b> .....	<b>17</b>
1.2.1 Ανάλυση κλινικών αποφάσεων (επίλυση προβλημάτων) .....	18
<b>1.3 Δομή των συστημάτων υποστήριξης κλινικών αποφάσεων</b> .....	<b>22</b>
1.3.1 Περιεχόμενο .....	23
1.3.2 Πηγές δεδομένων.....	24
1.3.3 Μηχανισμοί λήψης αποφάσεων.....	26
1.3.4 Παροχή πληροφοριών προς τον χρήστη.....	27
1.3.5 Ροή εργασιών .....	28
<b>1.4 Παράγοντες επιτυχίας των συστημάτων CDSS</b> .....	<b>28</b>
<b>1.5 Αναφορές Κεφαλαίου</b> .....	<b>32</b>
<b>Κεφάλαιο 2 – Μηχανική μάθηση, Αναγνώριση προτύπων και Ταξινόμηση</b> .....	<b>33</b>
<b>2.1 Εισαγωγή</b> .....	<b>33</b>
<b>2.2 Επιλογή χαρακτηριστικών</b> .....	<b>35</b>
<b>2.3 Αλγόριθμοι μηχανικής μάθησης</b> .....	<b>40</b>
2.3.1 Έυρεση κοντινότερων γειτόνων .....	40
2.3.2 Μηχανές διανυσματικής υποστήριξης.....	41
2.3.3 Δέντρα αποφάσεων .....	43
2.3.4 Απλός Μπεϋζιανός ταξινομητής .....	44
2.3.5 Μπεϋζιανά δίκτυα.....	45
2.3.6 Τεχνητά νευρωνικά δίκτυα.....	47
2.3.7 Γενετικοί αλγόριθμοι.....	52
2.3.8 Ασαφής λογική .....	55
<b>2.4 Μέτρα απόδοσης</b> .....	<b>57</b>
2.4.1 Ευαισθησία και ειδικότητα .....	57
2.4.2 Απόδοση και ακρίβεια .....	58
2.4.3 Θετική και αρνητική προγνωστική αξία .....	59
2.4.4 Καμπύλη ROC .....	60
<b>2.5 Αναφορές Κεφαλαίου</b> .....	<b>61</b>
<b>Κεφάλαιο 3 – Εφαρμογή τεχνικών μηχανικής μάθησης και συστημάτων CDSS για την βελτίωση της διάγνωσης, πρόγνωσης και κλινικής διαχείρισης του καρκίνου</b> .....	<b>62</b>

3.1	Εισαγωγή.....	62
3.2	Μελέτες σχετικά με τον καρκίνο του μαστού .....	64
3.3	Μελέτες σχετικά με τον καρκίνο του προστάτη .....	77
3.4	Μελέτες σχετικά με τον καρκίνο του τραχήλου της μήτρας .....	87
3.5	Μελέτες σχετικά με τον καρκίνο του πνεύμονα .....	102
3.6	Μελέτες σχετικά με τον καρκίνο του παχέος έντερου .....	105
3.7	Μελέτες σχετικά με την λευχαιμία.....	109
3.8	Μελέτες σχετικά με τον καρκίνο του στόματος.....	114
3.9	Μελέτες σχετικά με τον καρκίνο των ωοθηκών .....	117
3.10	Μελέτες σχετικά με τον καρκίνο στο θυρεοειδή .....	120
3.11	Μελέτες σχετικά με τον καρκίνο του ήπατος.....	124
3.12	Μελέτες σχετικά με τον καρκίνο των οστών.....	126
3.13	Μελέτες σχετικά με την χημειοθεραπεία .....	128
3.14	Αναφορές Κεφαλαίου .....	131
	<b>Κεφάλαιο 4 – Συμπεράσματα και επισκόπηση μελέτης .....</b>	<b>133</b>
4.1	Συμπεράσματα, περιορισμοί και συστάσεις.....	133
4.2	Αναφορές Κεφαλαίου .....	137

## Ευρετήριο εικόνων

Εικόνα 1 : Τυπικό σύστημα ταξινόμησης .....	34
Εικόνα 2 : Διαδικασία ταξινόμησης.....	35
Εικόνα 3 : Χώρος χαρακτηριστικών 3 διαστάσεων και σύμπλεγμα υποχώρων. Οι υποχώροι 2 διαστάσεων $X_1$ , $X_2$ και $X_3$ αποκτώνται από το χώρο 3 διαστάσεων $X$ με απλή προβολή. Οι μονάδες στο γράφημα είναι άνευ σημασίας.....	36
Εικόνα 4 : Συσχέτιση της απόδοσης του ταξινομητή με τον αριθμό χαρακτηριστικών.....	37
Εικόνα 5 : Η «κατάρτα της διαστασιμότητας» - χώρος μεταβλητών εισόδου με μία, δύο και τρεις διαστάσεις.....	38
Εικόνα 6 : Διαχωριστικά υπερεπίπεδα για το SVM. Στο ένα υπερεπίπεδο διακρίνεται η μικρή απόσταση από το κοντινότερο από τα θετικά και αρνητικά παραδείγματα. Στο άλλο υπερεπίπεδο (σι) φαίνεται η μέγιστη απόσταση μεταξύ των δύο κλάσεων η οποία λαμβάνεται τελικά.....	42
Εικόνα 7 : (a) Ο γραμμικός ταξινομητής διαχωρίζει τα σημεία στο επίπεδο. (b) Στην περίπτωση αυτή, ο γραμμικός ταξινομητής επιτυγχάνει μέγιστο διαχωρισμό (μέγιστο περιθώριο) .....	43
Εικόνα 8 : Παράδειγμα ενός δέντρου αποφάσεων (εδώ αφορά καρκίνο του προστάτη) .....	44
Εικόνα 9 : Απεικόνιση ενός Μπεϋζιανού δικτύου.....	46
Εικόνα 10 : Απεικόνιση ενός απλοϊκού Μπεϋζιανού δικτύου .....	46
Εικόνα 11 : Τεχνητό νευρωνικό δίκτυο .....	48
Εικόνα 12 : Πολυεπίπεδος αισθητήρας με δύο κρυφά επίπεδα .....	50
Εικόνα 13 : Ένα πιθανοτικό νευρωνικό δίκτυο .....	52
Εικόνα 14 : Διάγραμμα ροής απλού γενετικού αλγορίθμου.....	54
Εικόνα 15 : Σχηματικό διάγραμμα αρχιτεκτονικής ασαφών συστημάτων.....	56
Εικόνα 16 : Δομή του δικτύου Bayes που εκπαιδεύτηκε. Οι κύκλοι με τις ετικέτες αναπαριστούν τους κόμβους και τα βέλη (τόξα) αναπαριστούν τις σχέσεις εξάρτησης. Ca++ = αποτιτανώσεις, FHx = οικογενειακό ιστορικό καρκίνου του μαστού, h/o = ιστορικό, HRT = θεραπεία ορμονικής υποκατάστασης (καρκίνος του μαστού, 1η μελέτη).....	66
Εικόνα 17 : Το συνδυαστικό προγνωστικό σύστημα που ανέπτυξαν οι συγγραφείς (καρκίνος του μαστού, 2η μελέτη).....	68
Εικόνα 18 : Το προγνωστικό σύστημα που βασίζεται στο νευρωνικό δίκτυο και στην μονάδα κατωφλίου για τη διαδικασία λήψης αποφάσεων (καρκίνος του μαστού, συμπληρωματική μελέτη) .....	71
Εικόνα 19 : Η εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes ( $p$ ), η πιθανότητα που λαμβάνεται από τον κανόνα απόφασης Bayes ( $p_{net}$ ) και η εκ των προτέρων πιθανότητα ταξινόμησης και ( $p(C2)$ ) για κάθε χρονικό διάστημα υπό μελέτη (καρκίνος του μαστού, συμπληρωματική μελέτη) .....	72

Εικόνα 20 : Η εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes (BCP), η σωστή πιθανότητα ταξινόμησης που λαμβάνεται με το προτεινόμενο σύστημα νευρωνικών δικτύων (NNCP) και η εκ των προτέρων πιθανότητα ταξινόμησης και (PCP) για κάθε χρονικό διάστημα υπό μελέτη (καρκίνος του μαστού, 2η μελέτη).....	72
Εικόνα 21 : Κατανομή των εκτιμήσεων του TNΔ της πενταετούς πιθανότητας υποτροπής ασθενών που έχουν ταξινομηθεί σύμφωνα με το τροποποιημένο σύστημα TNM (καρκίνος του μαστού, 4η μελέτη).....	76
Εικόνα 22 : Παράδειγμα νευρωνικού δικτύου (βιοψία καρκίνου του προστάτη, εισαγωγής καρκίνου του προστάτη) .....	77
Εικόνα 23 : Καμπύλες ROC για το tPSA (συνεχής γραμμή), το %fPSA (τελείες) και το TNΔ (διακεκομμένη γραμμή). Το γράφημα Α αντιστοιχεί στο εύρος 2 – 4 μg/L, το γράφημα Β αντιστοιχεί στο εύρος 4.1 – 10 μg/L, το γράφημα Γ αντιστοιχεί στο εύρος 10.1 – 20 μg/L και το γράφημα Δ αντιστοιχεί στο εύρος 2 – 10 μg/L (καρκίνος του προστάτη, 1η μελέτη).....	82
Εικόνα 24 : Ανάλυση καμπυλών ROC των συστημάτων υποστήριξης κλινικών αποφάσεων που χρησιμοποίησαν τα μοντέλα SVM και TNΔ. Η τιμή της περιοχής κάτω από τη ROC καμπύλη (AUC) του μοντέλου SVM ήταν ανώτερη από αυτή του TNΔ (καρκίνος του προστάτη, 2η μελέτη).....	84
Εικόνα 25 : Ανάπτυξη πέντε επιπέδων της υποδομής του συστήματος υποστήριξης κλινικών αποφάσεων (καρκίνος του προστάτη, 3η μελέτη).....	86
Εικόνα 26 : Εφαρμογή της κλινικής προόδου μέσω συστήματος κλινικής ποσοτικοποιημένης πληροφορίας (clinical infometrics system, καρκίνος του προστάτη, 3η μελέτη) .....	86
Εικόνα 27 : Επισκόπηση του συστήματος υποστήριξης κλινικών αποφάσεων (CDSS, καρκίνος του τραχήλου της μήτρας, 1η μελέτη) .....	88
Εικόνα 28 : Σχηματικό διάγραμμα που παρουσιάζει κάποιες από τις έννοιες στις αναφορές των τεστ Παπανικολάου οι οποίες απαιτούνται για την εφαρμογή των εθνικών κατευθυντήριων γραμμών για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου (καρκίνος του τραχήλου της μήτρας, 1η μελέτη).....	89
Εικόνα 29 : Διάγραμμα ροής για τον κανόνα κατευθυντήριων γραμμών (καρκίνος του τραχήλου της μήτρας, 1η μελέτη) .....	90
Εικόνα 30 : Γραφική δομή της στατικής έκδοσης του μοντέλου πληθυσμιακού ελέγχου του Πίτσμπουργκ για τον καρκίνο του τραχήλου της μήτρας (Pittsburgh Cervical Cancer Screening Model – PCCSM, καρκίνος του τραχήλου της μήτρας, 2η μελέτη).....	92
Εικόνα 31 : Προβλέψεις σχετικού κινδύνου για ιστοπαθολογικά προκαρκινικά στάδια ή διηθητικός καρκίνος του τραχήλου της μήτρας στα 2 έτη (ταξινομημένα από τον υψηλότερο στο χαμηλότερο κίνδυνο) με διαφορετικά τρέχοντα αποτελέσματα για τα τεστ Παπανικολάου και του ιού των ανθρώπινων κονδυλωμάτων (καρκίνος του τραχήλου της μήτρας, 2η μελέτη).....	93
Εικόνα 32 : Πιθανότητα να μην υπάρχουν ιστοπαθολογικά προκαρκινικά στάδια ή διηθητικός καρκίνος του τραχήλου της μήτρας εντός 3 ετών (καρκίνος του τραχήλου της μήτρας, 2η μελέτη).....	94

<i>Εικόνα 33 : Μία απλοποιημένη εκδοχή του μοντέλου PCCSM (καρκίνος του τραχήλου της μήτρας, συμπληρωματική μελέτη).....</i>	<i>95</i>
<i>Εικόνα 34 : Χρονικές εκτιμήσεις κινδύνου προκαρκινικού σταδίου του τραχήλου σε διάστημα 15 ετών (καρκίνος του τραχήλου της μήτρας, συμπληρωματική μελέτη) .....</i>	<i>96</i>
<i>Εικόνα 35 : Σχηματική απεικόνιση του Μαρκοβιανού μοντέλου (καρκίνος του τραχήλου της μήτρας, 3η μελέτη).....</i>	<i>97</i>
<i>Εικόνα 36 : Το σχηματικό διάγραμμα της δομής του LVQ-NN (καρκίνος του τραχήλου της μήτρας, 4η μελέτη).....</i>	<i>101</i>
<i>Εικόνα 37 : Τεχνητό νευρωνικό δίκτυο (καρκίνος του πνεύμονα).....</i>	<i>104</i>
<i>Εικόνα 38 : Αποτελέσματα του μοντέλου 5-ετούς επιβίωσης με χρήση ANN (A) και LR (B). Τα σκιασμένα κουτιά υποδεικνύουν ασθενείς που έχουν επιβιώσει, ενώ τα κενά κουτιά ασθενείς που έχουν αποβιώσει (καρκίνος του πνεύμονα).....</i>	<i>105</i>
<i>Εικόνες 39 &amp; 40 : Καμπύλη ROC για το νευρωνικό δίκτυο (αριστερά) και καμπύλη ROC για τη λογιστική παλινδρόμηση (δεξιά) (καρκίνος του παχέος εντέρου, 2η μελέτη) .....</i>	<i>109</i>
<i>Εικόνα 41 : Αρχιτεκτονική πιθανοτικού νευρωνικού δικτύου (λευχαιμία, 2η μελέτη) .....</i>	<i>113</i>
<i>Εικόνα 42 : Διάγραμμα ροής του προτεινόμενου συστήματος υποστήριξης απόφασης (καρκίνος του στόματος).....</i>	<i>115</i>
<i>Εικόνα 43 : Καμπύλες ROC για τα βέλτιστα σχήματα ταξινόμησης βασισμένα σε κλινικά, απεικονιστικά και γονιδιωματικά δεδομένα (καρκίνος του στόματος).....</i>	<i>117</i>
<i>Εικόνα 44 : Αρχιτεκτονική του FALCON-AART (καρκίνος των ωοθηκών).....</i>	<i>119</i>
<i>Εικόνα 45 : Μοντέλο δικτύου Μπεϋζιανής πίστης : Παθολογική διάγνωση στους θυρεοειδικούς όζους (Καλοήθεις έναντι Κακοηθών) (καρκίνος του θυρεοειδούς).....</i>	<i>123</i>
<i>Εικόνα 46 : Καμπύλη ROC για την πρόβλεψη καρκίνου. Η ευαισθησία απεικονίζεται στον άξονα y και η ειδικότητα στον άξονα x (καρκίνος του θυρεοειδούς).....</i>	<i>123</i>
<i>Εικόνα 47 : Διάφορα μοντέλα καμπύλης ανάπτυξης για την πρόβλεψη παράγοντα κινδύνου (καρκίνος του ήπατος).....</i>	<i>125</i>
<i>Εικόνες 48 &amp; 49 : Διαγράμματα ευαισθησίας και ακρίβειας του συνόλου επικύρωσης των μοντέλων ταξινόμησης της ηπατικής νόσου (καρκίνος του ήπατος).....</i>	<i>126</i>
<i>Εικόνες 50 &amp; 51 : Τρίμηνο μοντέλο BBN με μεταγενέστερες κατανομές που απεικονίζονται ως ποσοστά (%) του πληθυσμού εκπαίδευσης και δωδεκάμηνο μοντέλο BBN (καρκίνος των οστών) .....</i>	<i>128</i>
<i>Εικόνα 52 : Διαδικασία επιλογής χαρακτηριστικών του γενετικού αλγορίθμου (χημειοθεραπεία).....</i>	<i>130</i>

## Ευρετήριο Πινάκων

Πίνακας 1 : Μέθοδοι υπολογισμού αποστάσεων.....	41
Πίνακας 2 : Οι μεταβλητές που χρησιμοποιήθηκαν στο δίκτυο Bayes και οι αντίστοιχες τιμές τους (καρκίνος του μαστού, 1η μελέτη).....	65
Πίνακας 3 : Αποτελέσματα δοκιμών της ανάλυσης ROC για το προγνωστικό μοντέλο που βασίζεται στα νευρωνικά δίκτυα και τα δέντρα αποφάσεων (καρκίνος του μαστού, 2η μελέτη). ....	69
Πίνακας 4 : Περίληψη των δεδομένων των ασθενών, μέσος όρος, τυπική απόκλιση και διακύμανση (καρκίνος του μαστού, συμπληρωματική μελέτη). ....	70
Πίνακας 5 : Αριθμός των ασθενών, η εκ των προτέρων πιθανότητα υποτροπής, η πιθανότητα που λαμβάνεται από τον κανόνα απόφασης Bayes και το τυπικό σφάλμα (STD) για κάθε χρονικό διάστημα υπό μελέτη (καρκίνος του μαστού, συμπληρωματική μελέτη). .	71
Πίνακας 6 : Τύπος των χαρακτηριστικών που χρησιμοποιούνται στην επεξεργασία δεδομένων (καρκίνος του μαστού, 3η μελέτη).....	73
Πίνακας 7 : Επιλεγμένοι αλγόριθμοι ταξινόμησης και σχετικές τιμές (καρκίνος του μαστού, 3η μελέτη). ....	74
Πίνακας 8 : Προγνωστική ακρίβεια των προβλέψεων του μοντέλου για το εσωτερικό και το εξωτερικό σύνολο επικύρωσης (καρκίνος του μαστού, 4η μελέτη). ....	75
Πίνακας 9 : Διακριτική ικανότητα του μοντέλου του TNΔ έναντι του τροποποιημένου συστήματος TNM σε ασθενείς από τα δύο σύνολα επικύρωσης (καρκίνος του μαστού, 4η μελέτη). ....	76
Πίνακας 10 : Άρθρα που έχουν εκδοθεί στο πεδίο της χρήσης των TNΔ στον καρκίνο του προστάτη (δεν περιλαμβάνονται reviews, καρκίνος του προστάτη, εισαγωγή). ....	80
Πίνακας 11 : Ειδικότητες σε δοσμένες τιμές ευαισθησίας για το tPSA, το %fPSA και το TNΔ σε όλα τα εύρη PSA που ερευνούνται (καρκίνος του προστάτη, 1η μελέτη). ....	81
Πίνακας 12 : Ευαισθησίες σε δοσμένες τιμές ειδικότητας για το tPSA, το %fPSA και το TNΔ σε όλα τα εύρη PSA που ερευνούνται (καρκίνος του προστάτη, 1η μελέτη). ....	82
Πίνακας 13 : Μέσες και τυπικές αποκλίσεις των παραμέτρων εισόδου στα σύνολα εκπαίδευσης και δοκιμών (καρκίνος του προστάτη, 2η μελέτη).....	84
Πίνακας 14 : Κατανομή των περιπτώσεων στον πίνακα αποφάσεων (καρκίνος του τραχήλου της μήτρας, 1η μελέτη).....	91
Πίνακας 15 : Προβλέψεις του μοντέλου και τιμές επικύρωσης για την εμφάνιση του καρκίνου του τραχήλου, κατανομή σε στάδια FIGO και κίνδυνος για τη διάρκεια ζωής λόγω του καρκίνου του τραχήλου (καρκίνος του τραχήλου της μήτρας, 3η μελέτη). ....	98
Πίνακας 16 : Απόλυτες περιπτώσεις και περιπτώσεις και θάνατοι λόγω του καρκίνου του τραχήλου που θα μπορούσαν να έχουν προληφθεί και το προσδόκιμο ζωής με τον προσυμπτωματικό έλεγχο τεστ ΠΑΠ σε χρονικά διαστήματα 1, 2, 3 και 5 ετών (καρκίνος του τραχήλου της μήτρας, 3η μελέτη). ....	99

Πίνακας 17 : Οι μεταβλητές που χρησιμοποιήθηκαν για την εκπαίδευση του ταξινομητή LVQ-NN (καρκίνος του τραχήλου της μήτρας, 4η μελέτη). .....	100
Πίνακας 18 : Απόδοση του ταξινομητή LVQ-NN για τα σύνολα εκπαίδευσης, δοκιμών και στο σύνολο και κυτταρολογική απόδοση (καρκίνος του τραχήλου της μήτρας, 4η μελέτη). .....	102
Πίνακας 19 : Πίνακας μεταβλητών εισόδου του προγνωστικού μοντέλου (καρκίνος του πνεύμονα).....	103
Πίνακας 20 : Κλινικοπαθολογικές μεταβλητές που χρησιμοποιήθηκαν στην εκπαίδευση και επικύρωση των τεχνητών νευρωνικών δικτύων(καρκίνος του παχέος εντέρου,1η μελέτη).106	106
Πίνακας 21 : Απόδοση των τεχνητών νευρωνικών δικτύων που εκπαιδεύτηκαν και επικυρώθηκαν σε δεδομένα του 1ου ιατρικού κέντρου (καρκίνος του παχέος εντέρου, 1η μελέτη). .....	107
Πίνακας 22 : Απόδοση των ιατρών και του προηγουμένως εκπαιδευμένου 12μηνου τεχνητού νευρωνικού δικτύου σε βάση δεδομένων από το 2ο ιατρικό κέντρο (καρκίνος του παχέος εντέρου, 1η μελέτη). .....	107
Πίνακας 23 : Παθολογικές μεταβλητές που χρησιμοποιήθηκαν στα μοντέλα λογιστικής παλινδρόμησης και νευρωνικού δικτύου (καρκίνος του παχέος εντέρου, 2η μελέτη). .....	108
Πίνακες 24 & 25 : Πίνακας υποψήφιων χαρακτηριστικών και η τιμή MI που λαμβάνουν σε σχέση με το αποτέλεσμα, Πίνακας πειραμάτων με διαφορετικά σύνολα χαρακτηριστικών (λευχαιμία, 1η μελέτη). .....	111
Πίνακες 26 & 27 : Πίνακας εκτίμησης κινδύνου υποτροπής (α) από το προτεινόμενο σύστημα και (β) από το πρωτόκολλο BFM95 συγκριτικά με το προτεινόμενο μοντέλο (λευχαιμία, 1η μελέτη). .....	111
Πίνακας 28 : Αποτελέσματα ταξινόμησης διασταυρούμενης επικύρωσης 10 φορών των τεσσάρων μεθόδων επιλογής χαρακτηριστικών του ταξινομητή PNN (λευχαιμία, 2η μελέτη). .....	113
Πίνακας 29 : Αποτελέσματα ταξινόμησης διασταυρούμενης επικύρωσης 10 φορών των τεσσάρων μεθόδων επιλογής χαρακτηριστικών του PNN με διαφορετικό μέγεθος υποσυνόλων χαρακτηριστικών (λευχαιμία, 2η μελέτη). .....	114
Πίνακας 30 : Τα 50 σημαντικότερα γονίδια για κάθε μία από τις πέντε βάσεις δεδομένων όπως προκύπτουν από την ανάλυση της στατιστικής σήματος προς θόρυβο (λευχαιμία, 2η μελέτη). .....	114
Πίνακες 31, 32, 33 : Συγκεντρωτικοί πίνακες κλινικών, απεικονιστικών και γονιδιωματικών χαρακτηριστικών (καρκίνος του στόματος). .....	116
Πίνακας 34 : Λίστα χαρακτηριστικών που προκύπτουν από αιματολογικές αναλύσεις (καρκίνος των ωοθηκών).....	119
Πίνακας 35 : Πίνακας σύγκρισης διαφόρων μεθόδων ταξινόμησης όσον αφορά την πρόβλεψη του καρκίνου των ωοθηκών με χρήση των αιματολογικών εξετάσεων (καρκίνος των ωοθηκών). .....	120
Πίνακας 36 : Πίνακας χαρακτηριστικών του υπό μελέτη πληθυσμού (καρκίνος του θυρεοειδούς).....	122

*Πίνακας 37 : Χαρακτηριστικά ασθενών που συγκεντρώθηκαν σε αυτή την μελέτη (χημειοθεραπεία).....129*

*Πίνακας 38 : Συγκεντρωτικός πίνακας των μελετών που εξετάστηκαν όπου περιγράφεται ο τύπος του καρκίνου, ο σκοπός κάθε συστήματος, τα δεδομένα που χρησιμοποιήθηκαν, οι τεχνικές ταξινόμησης, τα γενικά αποτελέσματα και η βιβλιογραφία. ....134*



# Κεφάλαιο 1 – Εισαγωγή

## **1.1 Εισαγωγή στην έννοια των συστημάτων υποστήριξης κλινικών αποφάσεων**

Η ραγδαία ανάπτυξη των σύγχρονων τεχνολογιών πληροφορικής έχει δημιουργήσει νέες τάσεις και προκλήσεις στην εξέλιξη των υπολογιστικών συστημάτων, με σκοπό την υποβοήθηση κλινικών αποφάσεων θεραπευτικού ή διαγνωστικού χαρακτήρα με βάση εξατομικευμένα μοντέλα ασθενών [1].

Τα συστήματα υποστήριξης κλινικών αποφάσεων (*Clinical Decision Support Systems – CDSS*) είναι ηλεκτρονικά συστήματα που έχουν σχεδιαστεί για να βοηθήσουν τους επαγγελματίες υγειονομικής περίθαλψης στη λήψη κλινικών αποφάσεων για μεμονωμένους ασθενείς [2]. Με άλλα λόγια, πρόκειται για ενεργά συστήματα γνώσης τα οποία χρησιμοποιούν δύο ή περισσότερα στοιχεία των δεδομένων των ασθενών για να «παράγουν» εξατομικευμένες ή μη συμβουλές.

Αυτά τα είδη του λογισμικού χρησιμοποιούν τις σχετικές γνώσεις, κανόνες μέσα σε μια βάση γνώσεων και σχετικά κλινικά δεδομένα για τη βελτίωση της λήψης κλινικών αποφάσεων σε θέματα όπως η πρόληψη, βραχεία και χρόνια φροντίδα, πρακτικές συνταγογράφησης κ.ά. Ένα σύστημα υποστήριξης κλινικών αποφάσεων συσχετίζει δεδομένα σχετικά με τα χαρακτηριστικά των ασθενών με μια αξιόπιστη βάση γνώσεων για να καθοδηγήσει έναν κλινικό ιατρό σε ειδικές για τον ασθενή εκτιμήσεις ή συστάσεις. Οι κλινικοί ιατροί, το προσωπικό υγειονομικής περίθαλψης ή οι ασθενείς μπορούν να εισάγουν χειροκίνητα τα χαρακτηριστικά των ασθενών στα συστήματα πληροφορικής ή εναλλακτικά μπορούν να αναζητηθούν τα ηλεκτρονικά ιατρικά αρχεία για την ανάκτηση των χαρακτηριστικών του ασθενούς. Αυτά τα είδη των συστημάτων υποστήριξης απόφασης επιτρέπουν στους ιατρούς να εντοπίζουν και να επιλέγουν την πλέον κατάλληλη θεραπεία. Η παρεχόμενη υποστήριξη αποφάσεων βασίζεται σε διαδικασίες εξελιγμένης αξιολόγησης αποτελεσμάτων και αλγορίθμους που χρησιμοποιούν βάσεις δεδομένων για να αναζητήσουν τις τελευταίες εξελίξεις σχετικά με τις βέλτιστες πρακτικές.

Ανεξάρτητα από το πώς επιλέγουμε να ορίσουμε ένα σύστημα υποστήριξης κλινικών αποφάσεων, πρέπει να δεχθούμε ότι το πεδίο των παραπάνω συστημάτων είναι ραγδαίως αναπτυσσόμενο και δεν υπόκειται σε κανόνες. Αν τα εν λόγω συστήματα είναι κακώς σχεδιασμένα και έχουν αξιολογηθεί ανεπαρκώς, μπορούν να προκαλέσουν βλάβη, έχουν όμως και μια τεράστια δυναμική ώστε να ωφελήσουν, ειδικά στην απόδοση της παροχής υγειονομικής περίθαλψης και την ποιότητα της περίθαλψης των ασθενών.

## **1.2 Λήψη κλινικών αποφάσεων**

Υπάρχουν μερικές βασικές κατηγορίες που πρέπει να ληφθούν υπόψη στην προσπάθεια για άριστη λήψη αποφάσεων : α) ακριβή δεδομένα, β) εφαρμογή της γνώσης, γ) κατάλληλες δεξιότητες επίλυσης προβλημάτων.

Κατ' αρχήν, τα δεδομένα των ασθενών πρέπει να είναι επαρκή για τη λήψη μιας έγκυρης απόφασης. Το πρόβλημα προκύπτει όταν ο ιατρός έρχεται αντιμέτωπος με ένα καταγιστικό όγκο συγκεκριμένων ή μη δεδομένων, τα οποία δεν μπορεί να επεξεργαστεί ικανοποιητικά. Ως εκ τούτου, είναι σημαντικό να εκτιμηθεί τότε τα πρόσθετα στοιχεία συγχέουν παρά αποσαφηνίζουν την περίπτωση ενός ασθενούς.

Για παράδειγμα, μια συνήθης συνθήκη για ένα τέτοιο πρόβλημα είναι οι μονάδες εντατικής θεραπείας, όπου οι επαγγελματίες πρέπει να επεξεργάζονται μεγάλες ποσότητες δεδομένων από διαφορετικές θρόνες, να είναι ενήμεροι της κλινικής κατάστασης των ασθενών, του ιστορικού τους, των πιθανών χρόνιων ασθενειών, των φαρμάκων που λαμβάνουν και των ανεπιθύμητων ενεργειών τους κ.ά. και επιπρόσθετα να παίρνουν μια σωστή απόφαση σχετικά με την πορεία δράσης. Η ποιότητα των διαθέσιμων στοιχείων είναι αντίστοιχης σημασίας. Τα όργανα μέτρησης και παρακολούθησης πρέπει να είναι όσο το δυνατόν περισσότερο τεχνολογικά ακριβή, καθώς τα ανακριβή δεδομένα θα μπορούσαν να έχουν σοβαρές δυσμενείς επιπτώσεις στις αποφάσεις που αφορούν την περίθαλψη των ασθενών. Η γνώση που χρησιμοποιείται στη διαδικασία λήψης αποφάσεων πρέπει να είναι ακριβής και σύγχρονη. Είναι υψίστης σημασίας ο κλινικός γιατρός που αποφασίζει να έχει ένα ευρύ φάσμα ιατρικών γνώσεων και πρόσβαση σε πηγές πληροφόρησης, χάρη στις οποίες μπορεί συνεχώς να αναθεωρεί και να επικυρώνει τις γνώσεις του. Για να λαμβάνει ο ασθενής την κατάλληλη φροντίδα, ο γιατρός πρέπει να γνωρίζει τις τελευταίες εξελίξεις στον τομέα της συγκεκριμένης περίπτωσης. Είναι στο χέρι του κλινικού ιατρού να κάνει πράξη τις αποδεδειγμένες θεραπείες από τα ερευνητικά έργα. Τα συστήματα υποστήριξης κλινικών αποφάσεων χρειάζονται αναλογικά μια εκτεταμένη, καλά δομημένη και σύγχρονη πηγή γνώσης ώστε να εξυπηρετούν κατάλληλα τον κλινικό ιατρό.

Προ πάντων, απαιτούνται καλές δεξιότητες επίλυσης προβλημάτων για την αξιοποίηση των διαθέσιμων δεδομένων και της γνώσης. Οι κλινικοί ιατροί που αποφασίζουν πρέπει να θέτουν κατάλληλους στόχους για κάθε έργο, να γνωρίζουν πώς να αιτιολογούν κάθε στόχο και να λαμβάνουν υπόψη τους το κόστος και το όφελος της θεραπείας και της διάγνωσης. Επιπλέον, δεν πρέπει να παραβλέπουμε το γεγονός ότι οι έμπειροι ιατροί βασίζονται εκτενώς στην προσωπική τους εμπειρία. Με την ενσωμάτωση των εξατομικευμένων δεδομένων του ασθενή στην εφαρμόσιμη βάση γνώσεων και τις αντίστοιχες κατευθυντήριες γραμμές, το σύστημα υποστήριξης κλινικών αποφάσεων μπορεί να βελτιώσει την ποιότητα της περίθαλψης με την ενίσχυση της διαδικασίας λήψης αποφάσεων. Προκειμένου να είμαστε σε θέση να κατασκευάσουμε ένα εφαρμόσιμο σύστημα CDSS, είναι επιτακτική ανάγκη να υπάρχει μια ευρύτερη βάση κατανόησης της λήψης ιατρικών αποφάσεων, όπως συμβαίνει στο φυσικό περιβάλλον. Ο σχεδιασμός των συστημάτων CDSS, χωρίς την κατανόηση των γνωστικών διαδικασιών που διέπουν την ιατρική λογική και ανάλυση αποφάσεων, αποτελεί ρίσκο όσον αφορά την αποτελεσματικότητα και τη δυνατότητα εφαρμογής τους στην καθημερινή ροή εργασιών.

### **1.2.1 Ανάλυση κλινικών αποφάσεων (επίλυση προβλημάτων)**

Η ανάλυση αποφάσεων αποτελεί μια ποσοτική αξιολόγηση των αποτελεσμάτων που προκύπτουν από ένα σύνολο επιλογών σε μια συγκεκριμένη κλινική κατάσταση και είναι η μεθοδολογία που επιλέγεται συνήθως για την επίλυση κλινικών προβλημάτων.

Αυτή η διαδικασία είναι συχνά έμμεση και λαμβάνει χώρα μέσα στο εσωτερικό αλγορίθμων που έχει αναπτύξει ο κλινικός ιατρός με την πάροδο του χρόνου. Η ανάλυση αποφάσεων, απαιτώντας μια συγκεκριμένη δομή μοντέλου και την αξιολόγηση διαφόρων πιθανοτήτων και τιμών αποτελεσμάτων, καθιστά τη διαδικασία λήψης αποφάσεων πιο σαφή ώστε να υπόκειται εύκολότερα σε εξέταση, συζήτηση και εκ νέου επαλήθευση. Τα μοντέλα αποφάσεων χρησιμοποιούνται συχνά ως αναλυτικά εργαλεία για τη διεξαγωγή ανάλυσης κόστους – αποτελεσματικότητας, αφού αυτή η μεθοδολογία μπορεί να χρησιμοποιηθεί για την εύρεση της αναμενόμενης τιμής των προβλεπόμενων αποτελεσμάτων.

Το φάσμα των κλινικών προβλημάτων που είναι κατάλληλα για ανάλυση αποφάσεων είναι τεράστιο. Τέτοια προβλήματα επικεντρώνονται σε μια συγκεκριμένη απόφαση, όπου υπάρχουν διλήμματα, με την έννοια ότι μία από τις επιλογές δεν είναι κατ' ανάγκη ανώτερη. Για παράδειγμα, ένα διαγνωστικό τεστ ίσως εγκυμονεί κάποιο κίνδυνο, αλλά αυτό το αντισταθμίζει με πιο κατάλληλη θεραπεία και καλύτερη έκβαση για τον ασθενή όταν η θεραπεία κατευθύνεται από τα αποτελέσματα του εν λόγω τεστ. Η λήψη αποφάσεων σπανίως είναι μια σαφής διαδικασία και ως εκ τούτου συνεπάγεται αναπόφευκτους συμβιβασμούς. Σε γενικές γραμμές, η ανάλυση αποφάσεων αναπτύσσεται σε :

- ✚ Παροχή βοήθειας στη διαδικασία λήψης κλινικών αποφάσεων για ένα συγκεκριμένο ασθενή,
- ✚ Εκτίμηση των βέλτιστων στρατηγικών για τις κατηγορίες των ασθενών με ειδικά κλινικά χαρακτηριστικά σε δεδομένες καταστάσεις,
- ✚ Εκτιμήσεις της σύνδεσης κλινικών και οικονομικών αποτελεσμάτων (ανάλυση κόστους – αποτελεσματικότητας) για τη διαμόρφωση ζητημάτων που αφορούν πολιτική που ακολουθείται στην υγεία,
- ✚ Παροχή εκτιμήσεων των αναμενόμενων αποτελεσμάτων σε καταστάσεις όπου οι κλασσικές μέθοδοι όπως οι τυχαιοποιημένες μελέτες είναι είτε αδύνατες ή ανέφικτες.

Υπάρχουν πολλά πλεονεκτήματα της χρήσης της ανάλυσης αποφάσεων για τη διερεύνηση επιλογών που αφορούν ένα μεμονωμένο ασθενή. Η τοποθέτηση του προβλήματος σε ένα σαφές, αναλυτικό πλαίσιο αναγκάζει τους ιατρούς να καθιστούν σαφείς τις υποθέσεις τους. Η ανάλυση αποφάσεων μπορεί να ενσωματώσει άμεσα θέματα σχετικά με την ποιότητα ζωής και το πώς ο συγκεκριμένος ασθενής εκτιμά διαφορετικά αποτελέσματα. Ένα παράδειγμα μιας τέτοιας ανάλυσης αποφάσεων ζυγίζει τους κινδύνους και τα οφέλη της μαστεκτομής (αφαίρεση μαστού) και της ωοθηκεκτομής (αφαίρεση ωοθηκών) για ασθενείς με μεταλλάξεις τύπου BRCA1 ή BRCA2. Πρέπει να συνδυαστούν διάφορα είδη δεδομένων – εκτιμήσεων των αυξημένων κινδύνων του καρκίνου του μαστού και των ωοθηκών για γυναίκες με μεταλλάξεις, η γενική συχνότητα εμφάνισης του καρκίνου και οι κίνδυνοι που επιβάλλουν οι χειρουργικές επεμβάσεις. Η ανάλυση σε ένα συγκεκριμένο χρονικό σημείο αποδεικνύει, χρησιμοποιώντας τα καλύτερα διαθέσιμα σύγχρονα στοιχεία, ότι το προσδόκιμο ζωής για γυναίκες με τέτοιες μεταλλάξεις αυξήθηκε κατά 5,3 έτη με τη μαστεκτομή και 1,7 έτη με την ωοθηκεκτομή.

Η ανάπτυξη και η κατασκευή ενός μοντέλου ανάλυσης αποφάσεων ακολουθεί μια λογική σειρά βημάτων (επίλυση προβλημάτων). Τυχόν προβλήματα ή σφάλματα σε κάθε βήμα μπορούν να μεταβάλλουν τα τελικά αποτελέσματα. Έτσι, είναι σημαντική η σωστή συμμόρφωση σε κάθε βήμα, τόσο από την άποψη του ερευνητή που διεξάγει μια ανάλυση, όσο και ενός ιατρού που ερμηνεύει τα αποτελέσματα. Τα βήματα είναι [2]:

- **Προσδιορισμός της ερώτησης** – Η διαδικασία καθορισμού του πεδίου εφαρμογής και των ορίων της συγκεκριμένης κλινικής κατάστασης που πρόκειται να αναλυθεί. Με άλλα λόγια, σε αυτό το στάδιο πρέπει να αναλυθούν τα νέα δεδομένα που συνθέτουν ένα πρόβλημα και να εντοπιστεί το πρόβλημα, όπως εμφανίζεται.
- **Δόμηση του κλινικού προβλήματος** – Με απλά λόγια, σημαίνει την κατασκευή ενός μοντέλου λήψης αποφάσεων που να αντιπροσωπεύει τις σχετικές συνιστώσες του προβλήματος. Πρέπει να ερευνηθεί ποιες θα είναι οι εναλλακτικές λύσεις και να δημιουργηθεί μια λίστα εναλλακτικών υποθέσεων. Στο κλινικό περιβάλλον, αυτές μπορεί να είναι πιθανές θεραπευτικές επιλογές για μια συνολική διάγνωση. Η μαθηματική αναπαράσταση ενός μοντέλου απόφασης λέγεται «δέντρο απόφασης» και αποτελείται από αρκετά διαφορετικά στοιχεία. Στοιχεία που συνδυάζονται σε δέντρα περιλαμβάνουν ένα ποσό λεπτομερειών που ορίζονται αυθαίρετα. Οι λεπτομέρειες οφείλουν να είναι όσο το δυνατόν περισσότερο κλινικά ρεαλιστικές σε σχέση με τη δυνατότητα κατασκευής, επικύρωσης και παρουσίασης του μοντέλου. Όσο περισσότερες λεπτομέρειες απαιτούνται, τόσο πιο δύσκολη καθίσταται η κατασκευή, επικύρωση και παρουσίαση του μοντέλου. Ωστόσο, όσο λιγότερο λεπτομερές είναι η ανάλυση και μεγαλύτερος ο αριθμός των απλουστευμένων παραδοχών, τόσο πιο ευάλωτη είναι η ανάλυση ως προς την έλλειψη κλινικά σημαντικών στοιχείων.
- **Εκτίμηση των σχετικών πιθανοτήτων** – Όταν δομηθεί ένα δέντρο απόφασης, πρέπει να προσδιοριστούν οι αριθμητικές τιμές των διαφόρων πιθανοτήτων. Υπάρχουν πολλές πηγές δεδομένων που μπορούν να χρησιμοποιηθούν για αυτούς τους προσδιορισμούς. Αν και υπάρχει μια γενικά αποδεκτή ιεραρχία της ποιότητας της μελέτης (τυχαιοποιημένη ελεγχόμενη δοκιμή > μελέτη υποομάδων > διοικητικές βάσεις δεδομένων > συστηματικές ανασκοπήσεις), τέτοιες αξιολογήσεις δεν είναι πάντα χρήσιμες για την ανάλυση απόφασης, δεδομένου ότι το συγκεκριμένο είδος μελέτης δεν μπορεί να συμβάλει στην εκτίμηση μιας δεδομένης παραμέτρου. Για παράδειγμα, μια τυχαιοποιημένη ελεγχόμενη δοκιμή αποτελεί εξαιρετική πηγή σύγκρισης μιας θεραπείας έναντι μιας άλλης, αλλά είναι ελλιπής πηγή δεδομένων σχετικά με τη συχνότητα μιας συγκεκριμένης ασθένειας. Έτσι, είναι σημαντικό να προσαρμόσουμε την πηγή δεδομένων στον τύπο των δεδομένων που απαιτούνται. Ωστόσο, ενδέχεται να μην υπάρχουν αρκετά στοιχεία που να αντιστοιχίζουν σχετικές πιθανότητες σε υποψήφιες λύσεις ή ακόμη και για να διαρθρωθεί το κλινικό πρόβλημα και να βρεθούν κατάλληλες εναλλακτικές λύσεις. Έτσι η διαδικασία υποχρεούται να μεταρρυθμιστεί με μια νέα έρευνα ώστε να συμπληρωθεί το παρόν «χάσμα γνώσης» (*knowledge gap*). Δυνητικά μπορούν να συμβούν πολλές επαναλήψεις

αυτού του κύκλου συγκέντρωσης στοιχείων και παραγωγής υποθέσεων μέχρι ο κλινικός γιατρός που λαμβάνει τις αποφάσεις να έχει αρκετή διαύγεια για να προχωρήσει με τη διαδικασία. Η ανάλυση κλινικών αποφάσεων αποτελεί κατά συνέπεια θεμελιώδες βήμα για την περαιτέρω έρευνα, για τη συλλογή αποδεικτικών στοιχείων και επαναξιολόγηση του προβλήματος – όλα με σκοπό την εύρεση κατάλληλων λύσεων που θα βοηθήσουν στην αντιμετώπιση της κλινικής κατάστασης. Τα συστήματα CDSS μπορούν να προγραμματιστούν με τέτοιο τρόπο ώστε να αναδιατάσσονται για ένα συγκεκριμένο πρόβλημα, να αναζητούν σχετικές αποδείξεις, ακόμη και να δημιουργούν καινούρια γνώση.

- **Υπολογισμός των τιμών των αποτελεσμάτων** – Διαφορετικές λύσεις σε ένα πρόβλημα αντιστοιχούν σε διαφορετικά αποτελέσματα. Η δομή του προβλήματος καθορίζει το μέτρο του συγκεκριμένου αποτελέσματος που θα χρησιμοποιηθεί. Για παράδειγμα, αν ο θάνατος ήταν το πιθανό αποτέλεσμα μιας ή περισσότερων στρατηγικών, το προσδόκιμο ζωής θα ήταν ένα κατάλληλο μέτρο αποτελέσματος. Η πιο σημαντική πτυχή στην εκχώρηση αποτελεσμάτων είναι ότι μετρώνται με τις ίδιες μονάδες σε όλους τους κλάδους. Ένα χρήσιμο χαρακτηριστικό της ανάλυσης αποφάσεων είναι ότι ένα δεδομένο μοντέλο μπορεί να αξιολογηθεί με χρήση διαφορετικών μέτρων αποτελεσμάτων. Για παράδειγμα, εκτός από την επιβίωση, ο κλινικός γιατρός ίσως θέλει να παρακολουθήσει την επίδραση διαφορετικών θεραπειών στο ρυθμό του εγκεφαλικού επεισοδίου, έμφραγμα του μυοκαρδίου, πνευμονική εμβολή, κτλ, με διάφορες θεραπευτικές επιλογές. Το μοντέλο απόφασης μπορεί να αναλυθεί χρησιμοποιώντας οποιαδήποτε από αυτά τα αποτελέσματα.
- **Ανάλυση του δέντρου** – Ο προτιμώμενος τύπος ανάλυσης ενός προβλήματος χρησιμοποιεί την στρατηγική που μεγιστοποιεί την αναμενόμενη τιμή του αποτελέσματος. Για παράδειγμα, αν το αποτέλεσμα που ενδιαφέρει ήταν το προσδόκιμο ζωής, η ανάλυση απόφασης θα είχε την μορφή «ο μέσος όρος του προσδόκιμου ζωής με τη στρατηγική A είναι 11,3 έτη έναντι 8,6 ετών που αντιστοιχούν στη στρατηγική B και ως εκ τούτου η στρατηγική A είναι η βέλτιστη στρατηγική». Ένα τέτοιο δέντρο αναπαριστά την επιλογή ανάμεσα σε δύο θεραπείες, όπου η μία προτιμάται όσον αφορά το προσδόκιμο ζωής.
- **Δοκιμή των παραδοχών του μοντέλου** – Τα αποτελέσματα που προέκυψαν από την ανάλυση απόφασης εξαρτώνται από την ακρίβεια των δεδομένων που χρησιμοποιούνται για την εκτίμηση των πιθανοτήτων και των αποτελεσμάτων. Είναι σπάνια η περίπτωση οι εκτιμήσεις να είναι γνωστές με απόλυτη βεβαιότητα. Ένα από τα βασικά πλεονεκτήματα των μοντέλων ανάλυσης αποφάσεων είναι η ικανότητά τους να ελέγχουν γρήγορα τις υποθέσεις και τα εισερχόμενα δεδομένα και να επικυρώνουν το μοντέλο απόφασης εκτελώντας ανάλυση ευαισθησίας. Η απάντηση του μοντέλου μπορεί να συγκριθεί με την «αληθή» απάντηση για να επικυρωθεί η δομή του μοντέλου, αξιολογώντας ένα δέντρο με χρήση παραμέτρων των οποίων η τιμή είναι γνωστή εκ των προτέρων. Για παράδειγμα, για την επιλογή ανάμεσα σε μια πιο αποτελεσματική (αλλά πιο ριψοκίνδυνη) χειρουργική θεραπεία και σε μια λιγότερο αποτελεσματική (αλλά ασφαλέστερη) ιατρική θεραπεία, η ανάλυση ευαισθησίας που αξιώνει μηδενικό ποσοστό θνησιμότητας για τη χειρουργική

επέμβαση πρέπει να συνηγορήσει υπέρ του χειρουργείου από τη στιγμή που δεν υπάρχει κίνδυνος επιδείνωσης.

- **Ερμηνεία των αποτελεσμάτων** – Η ερμηνεία των αποτελεσμάτων στην ανάλυση απόφασης είναι συχνά περίπλοκο έργο. Η ικανότητα της ανάλυσης αποφάσεων να διερευνά πώς η βέλτιστη στρατηγική σε μια συγκεκριμένη κλινική κατάσταση αλλάζει με διακύμανση στις υποθέσεις (και κατά συνέπεια εντοπίζει περιοχές με ανάγκες για περισσότερα δεδομένα) είναι συχνά ένα από τα πιο ισχυρά χαρακτηριστικά αυτού του τύπου ανάλυσης. Πολλές λεπτομέρειες της ανάλυσης θα πρέπει να εξετάζονται πριν από τη χρήση των αποτελεσμάτων για την αλλαγή της πρακτικής : a) Ο πληθυσμός των ασθενών θα πρέπει να συμπίπτει με τον αριθμό των ασθενών που παρακολουθεί ο κλινικός ιατρός b) Ο αναγνώστης πρέπει να εκτιμήσει την ισχύ του αποτελέσματος (ανάλυση ευαισθησίας). Αν η ανάλυση ευαισθησίας δείχνει ότι η βέλτιστη επιλογή είναι ισχυρά εξαρτώμενη από μια δεδομένη παράμετρο, πρέπει να αναπτυχθούν ακριβή μέτρα εκτίμησης της παραμέτρου.

Αν οι δημιουργοί των συστημάτων CDSS μπορούν να δημιουργήσουν αποτελεσματικά συστήματα που θα ενσωματωθούν με επιτυχία στη ροή εργασίας, παρέχοντας στους ιατρούς την απαραίτητη υποστήριξη στην ικανότητα λήψης αποφάσεων, θα πρέπει να είναι έτοιμοι να δεχθούν τα γενικά βήματα στη διαδικασία λήψης κλινικών αποφάσεων, όπως περιγράφονται παραπάνω. Τα αποτελεσματικά σχεδιασμένα και σωστά εφαρμοσμένα συστήματα CDSS μπορούν να βελτιώσουν σημαντικά την απόδοση των εργαζομένων στον τομέα της υγείας και την έκβαση των ιατρικών περιστατικών συμβάλλοντας θετικά στην ποιότητα της υγειονομικής περίθαλψης.

### **1.3 Δομή των συστημάτων υποστήριξης κλινικών αποφάσεων**

Τα συστήματα CDSS διαφέρουν πολύ στον σχεδιασμό τους. Οι βασικές αρχές της δομής και του σχεδιασμού έχουν επίσης αλλάξει σημαντικά κατά την τελευταία δεκαετία. Τα διάφορα χαρακτηριστικά των CDSS σχετίζονται ή έχουν άμεσες επιπτώσεις στην κλινική αποτελεσματικότητα, λειτουργικότητα, πρόληψη σφάλματος, δυνατότητα αποδοχής από τον κλινικό κόσμο, φορητότητα του συστήματος, σχέση κόστους-αποτελεσματικότητας κτλ. Είναι επομένως σημαντικό να χαρακτηρίσουμε τα συστήματα CDSS έτσι ώστε να γίνει καλύτερα κατανοητή η ποικιλομορφία τους. Η γνώση της ταξινόμησης, συνοδευόμενη από μια ιδέα των γενικών διαδικασιών λήψης κλινικών αποφάσεων όπως περιγράφεται παραπάνω, αποτελούν ένα ισχυρό σύνολο βασικών στοιχείων που είναι χρήσιμα για τους σχεδιαστές και αξιολογητές των συστημάτων CDSS. Σε αυτό το κεφάλαιο, γίνεται μια προσπάθεια χαρακτηρισμού των συστημάτων CDSS, συνδυάζοντας διαφορετικές πηγές για τη δημιουργία μιας συνολικής κατάταξης που καταγράφει τα βασικά στοιχεία του σχεδιασμού και της λειτουργίας των συστημάτων αυτών. Τα συστήματα CDSS μπορούν να ταξινομηθούν με βάση 5 άξονες [2]:

### 1.3.1 Περιεχομένο

#### α) Κλινικό περιβάλλον

- Περιβάλλον νοσηλείας
- Εξωτερικά ιατρεία

#### β) Κλινικό έργο

- Βοήθεια κατά τη διάγνωση.  
Με βάση τα στοιχεία του ασθενούς και της γνωστικής βάσης του συστήματος, το σύστημα CDSS παρέχει πιθανές διαγνώσεις. Η διαγνωστική βοήθεια μπορεί να συνδυαστεί με σύνθετα συστήματα ανάκτησης δεδομένων, όπως το ηλεκτροκαρδιογράφημα. Επιδιώκει να εντοπίσει «τι είναι αλήθεια» για έναν συγκεκριμένο ασθενή.
- Σχολιασμός θεραπείας και συμβουλευτική.  
Αυτή η λειτουργία αξιολογεί τη θεραπεία, αναζητά ασυνέπειες, λάθη, παραπομπές για αλληλεπιδράσεις φαρμάκων και εμποδίζει τη συνταγογράφηση αλλεργιογόνων φαρμάκων. Έχει αποδειχθεί ότι η αναγκαιότητα ενός ιατρού να παρέχει μια δήλωση - δικαιολόγηση για συγκεκριμένο λόγο, αν δεν ακολουθήσει τις συστάσεις, αυξάνει σημαντικά την κλινική σημασία του συστήματος CDSS. Τα συστήματα CDSS μπορούν να χρησιμοποιήσουν πρωτόκολλα και τεκμηριωμένες οδηγίες, σε συνδυασμό με τα στοιχεία του ασθενή, ώστε να παρέχουν το βέλτιστο πλάνο θεραπείας και να βοηθήσουν στην υλοποίησή του. Αυτού του είδους τα συστήματα απαντούν στην ερώτηση «τι πρέπει να κάνουμε» με έναν ασθενή και συχνά συνδυάζονται με συστάσεις για περαιτέρω διαγνωστικές διαδικασίες (πχ ποιες εξετάσεις πρέπει να γίνουν, ακτινογραφίες, αξονική κτλ). Το εν λόγω λογισμικό μπορεί να δημιουργήσει πρόσθετα ερωτήματα ώστε να παρέχει ακόμη πιο συγκεκριμένες συμβουλές για περαιτέρω θεραπεία (και διάγνωση).
- Δοσολογία του φαρμάκου ή συνταγογράφηση.  
Τα συστήματα CDSS έχουν τη δυνατότητα να μειώνουν επίπεδα των τοξικών φαρμάκων, τα ιατρικά λάθη, να αλλάζουν τη συνταγογράφηση σύμφωνα με τις συστάσεις και να μειώνουν το χρόνο επίτευξης θεραπευτικού ελέγχου. Αν το σύστημα συνδεθεί με το ηλεκτρονικό ιατρικό αρχείο (*Electronic Medical Record – EMR*), μπορεί να εμποδίσει τη συνταγογράφηση φαρμάκων που προκαλούν αλλεργικές αντιδράσεις. Τα εν λόγω συστήματα έχουν γίνει ευρέως αποδεκτά, δεδομένου ότι μπορούν να ενταχθούν πλήρως στη ρουτίνα του ιατρού και παρέχουν αυτόματη ηλεκτρονική μετάδοση στα φαρμακεία. Σε γενικές γραμμές, η συνταγογράφηση φαρμάκων αποτελεί ένα από τα πιο συνηθισμένα καθήκοντα του ιατρού, αλλά και ένα από τα συχνότερα ιατρικά καθήκοντα στα οποία εφαρμόζονται τα συστήματα CDSS.
- Επιλογή εξετάσεων.
- Ειδοποιήσεις και υπενθυμίσεις.  
Ένα εξειδικευμένο σύστημα που είναι ενσωματωμένο σε μια συσκευή παρακολούθησης ή σε ένα σύστημα πληροφοριών υγειονομικής περίθαλψης

(πχ σύστημα πληροφοριών εργαστηρίου, ηλεκτρονικό ιατρικό αρχείο) μπορεί να παρέχει ήχο σε πραγματικό χρόνο, ειδοποιήσεις εικόνας ή αφής μέσω διαφόρων εργαλείων επικοινωνίας (πχ e-mail, sms, pager-τηλεειδοποιητής). Τα συστήματα υπενθύμισης έχουν σχεδιαστεί έτσι ώστε να υπενθυμίζουν στους γιατρούς τα καθήκοντα ζωτικής σημασίας που πρέπει να γίνουν πριν από ένα συγκεκριμένο γεγονός (πχ νηστεία πριν από την ενδοσκόπηση, όχι αντιπηκτικά πριν από εγχείριση στην κοιλιακή χώρα).

- Ανάκτηση πληροφοριών.  
Ανάκτηση σχετικών πληροφοριών μέσω του παγκόσμιου ιστού ή ολοκληρωμένες βάσεις δεδομένων.
- Αναγνώριση εικόνας και ερμηνεία.  
Οι κλινικές εικόνες από την αξονική, μαγνητική τομογραφία, από αγγειογραφίες κλπ μπορούν στις μέρες μας να ερμηνεύονται εν μέρει αυτόματα. Είναι σημαντικό το γεγονός ότι τα συστήματα CDSS λειτουργούν ως βασικό εργαλείο ανίχνευσης, όπου το λογισμικό επισημαίνει κρίσιμες εικόνες που απαιτούν την ιδιαίτερη προσοχή του ιατρού.
- Πρόληψη.
- Πληθυσμιακός έλεγχος (Screening).
- Εξειδικευμένο εργαστηριακό σύστημα.
- Διαχείριση χρόνιων ασθενειών.

### 1.3.2 Πηγές δεδομένων

Αυτός ο άξονας ασχολείται με τις πηγές, την ποιότητα και την προσαρμογή των γνώσεων και των δεδομένων του συστήματος CDSS.

#### α) Κλινική πηγή γνώσης.

Μπορεί να προέρχεται από πηγές υψηλής ποιότητας (πχ τυχαιοποιημένες ελεγχόμενες μελέτες, συστηματικές ανασκοπήσεις, εθνικές ή κοινωνικές κατευθυντήριες γραμμές) και/ή από τη συμμετοχή των γιατρών που τελικά χρησιμοποιούν το σύστημα.

#### β) Πηγή δεδομένων.

Τα δεδομένα που αφορούν ένα συγκεκριμένο ασθενή μπορούν να ανακτηθούν από την ηλεκτρονική καταχώρηση της εισόδου, ιατρικά μέσα (πχ συσκευή μέτρησης αρτηριακής πίεσης), ηλεκτρονικό ιατρικό αρχείο ή από άλλο αρχείο αποθήκευσης δεδομένων. Τα δεδομένα μπορούν επίσης να συγκεντρωθούν από ένα έντυπο που περιλαμβάνει διάγραμμα ή από ένα άτομο. Σε αυτή την περίπτωση, τα δεδομένα πρέπει να εισάγονται στο σύστημα με χρήση ενός διαμεσολαβητή εισαγωγής τους. Αυτό το χαρακτηριστικό επηρεάζει την πιθανότητα υιοθέτησης του συστήματος CDSS στην πράξη σε μεγάλο βαθμό, καθώς έχει αποδειχθεί ότι προτιμάται η αυτόματη ηλεκτρονική παροχή δεδομένων στο σύστημα (από ηλεκτρονικό ιατρικό αρχείο).

#### γ) Διαμεσολαβητής εισαγωγής δεδομένων.



Πρόκειται για ένα κλινικό ιατρό που εισάγει τα δεδομένα στην πηγή δεδομένων (όπως περιγράφεται παραπάνω). Ενδιάμεσοι φορείς μπορούν να είναι και οι ίδιοι οι ασθενείς.

#### δ) Κωδικοποίηση δεδομένων.

Για διάφορους λόγους (πχ χρηματοδότηση, επιδημιολογία), προτιμάται η χρήση ενός σχήματος κωδικοποίησης ευρείας αποδοχής (αναφέρουμε ενδεικτικά ICD-10, SNOMED). Προφανώς τα δεδομένα θα μπορούσαν επίσης να είναι σε μορφή απλού κειμένου.

#### ε) Προσαρμογή δεδομένων.

Όσο πιο στοχευμένες συστάσεις παράγει ένα σύστημα CDSS για ένα συγκεκριμένο ασθενή, προσαρμοσμένες στην ηλικία, το φύλο, την ταυτόχρονη διάγνωση κτλ, τόσο μεγαλύτερες πιθανότητες έχει να είναι κλινικά αποτελεσματικό και χρήσιμο.

#### στ) Ενημερωμένος και σύγχρονος μηχανισμός.

Όπως προαναφέρθηκε, η βάση γνώσεων πρέπει να είναι σύγχρονη και συνεχώς ανανεώσιμη.

Μπορούμε να διακρίνουμε τα συστήματα CDSS σε 2 βασικές κατηγορίες με βάση τη διαχείριση των πηγών δεδομένων τους: τα **knowledge based** και τα **non-knowledge based** συστήματα.

Τα **knowledge based** αφορούν συστήματα τα οποία βασίζονται αποκλειστικά σε μια βάση δεδομένων και σε ένα μηχανισμό εξαγωγής συμπερασμάτων από τη βάση αυτή. Αποτελούνται κυρίως από τρία μέρη – τη βάση γνώσεων (knowledge base), το μηχανισμό εξαγωγής συμπερασμάτων (inference engine) και το μηχανισμό επικοινωνίας (interaction mechanism). Περιέχουν εξειδικευμένες ιατρικές γνώσεις για πολύ συγκεκριμένα γεγονότα και καθήκοντα και έχουν την ικανότητα να ερμηνεύουν τα εισαγόμενα δεδομένα χάρη στις διάφορες πηγές που προαναφέρθηκαν. Αυτά τα συστήματα συνήθως χρησιμοποιούν τη γνώση σε μορφή κανόνων AN – TOTE και πιθανολογικές συσχετίσεις ανάμεσα στα συγκεντρωθέντα στοιχεία. Ο μηχανισμός συμπερασμάτων βρίσκεται στο επίκεντρο των knowledge-based συστημάτων – συνδυάζει και συσχετίζει τους κανόνες της βάσης γνώσεων με τα δεδομένα του ασθενή. Κατά βάση, ο μηχανισμός εξαγωγής συμπερασμάτων αλληλεπιδρά με τις δοσμένες πληροφορίες για να εξάγει νέα συμπεράσματα.

Τα **non-knowledge based** συστήματα χρησιμοποιούν την αρχή της μηχανικής μάθησης στη μορφή νευρωνικών δικτύων, γενετικών αλγορίθμων κτλ, όπου οι υπολογιστές «μαθαίνουν» από την εμπειρία του παρελθόντος και/ή βρίσκουν μοτίβα στα κλινικά δεδομένα του ασθενή. Χρησιμοποιούν μεν τη βάση δεδομένων (βάση γνώσης) για την εκπαίδευσή τους, αλλά μετά από τη διαδικασία μάθησης επιδεικνύουν ένα είδος νοημοσύνης, «μαθαίνοντας» ουσιαστικά πως θα κρίνουν νέα περιστατικά με βάση αντίστοιχα παλαιότερα.

### 1.3.3 Μηχανισμοί λήψης αποφάσεων

Η αντιμετώπιση της διαδικασίας λήψης μιας κατάλληλης απόφασης είναι μάλλον η πιο σημαντική διάσταση του συστήματος CDSS.

#### α) Μέθοδοι συλλογισμού – λήψης αποφάσεων

Μερικοί μηχανισμοί συλλογισμού των συστημάτων CDSS είναι :

- ◆ **Συστήματα rule-based** (συστήματα που βασίζονται σε κανόνες). Ένα σύστημα rule-based χρησιμοποιεί διαφορετικές εξειδικευμένες βάσεις γνώσεων που μπορούν να αξιολογηθούν ως κανόνες AN – ΤΟΤΕ (κανόνες παραγωγής συμπερασμάτων). Ένα τέτοιο σύστημα αποτελεί παράδειγμα της εφευρετικής προσέγγισης στην οποία μεμονωμένες λογικές δηλώσεις σε μορφή κανόνων παραγωγής αποκτώνται παρατηρώντας ανθρώπινους εμπειρογνώμονες ή με τον απολογισμό τους και μετά συνδυάζονται σε μια προσπάθεια να μιμηθούν τη λογική διαδικασία των εμπειρογνώμωνων.  
Αυτή η προσέγγιση χρησιμοποιήθηκε πρώτα στο πείραμα MYCIN (Shortliffe 1976) με στόχο την επιλογή κατάλληλης αντιμικροβιακής θεραπείας για έναν ασθενή.
- ◆ **Νευρωνικά δίκτυα**. Το τεχνητό νευρωνικό δίκτυο είναι ένα προσαρμοστικό, μη βασισμένο σε βάση γνώσης σύστημα CDSS που χρησιμοποιεί την μηχανική μάθηση για να «εκπαιδεύεται» από τις εμπειρίες και να αναγνωρίζει μοτίβα σε κλινικές πληροφορίες.
- ◆ **Δίκτυο Bayes**. Το δίκτυο αυτό είναι ένα τυπικό σύστημα λήψης αποφάσεων βασισμένο σε πηγή γνώσης (γνωστό και ως δίκτυο belief ή αιτιώδες πιθανολογικό δίκτυο) που δείχνει πιθανολογικές σχέσεις μεταξύ συνόλων μεταβλητών – ασθένειες και συμπτώματα, βασισμένο σε εξαρτώμενη πιθανότητα σύμφωνα με το θεώρημα Bayes. Πρόκειται για ένα δίκτυο με ρητή απαίτηση οι σχέσεις να είναι αιτιολογικές. Ένα τέτοιο δίκτυο βοηθά στη διαμόρφωση της εξέλιξης μιας ασθένειας με την πάροδο του χρόνου και της αλληλεπίδρασης μεταξύ ασθενειών. Ένα μεγάλο εμπόδιο αποτελεί ωστόσο το γεγονός ότι η ιατρική γνώση δύσκολα καθορίζει ποιο είναι το αποτέλεσμα και ποια η αιτία.
- ◆ **Συστήματα που βασίζονται σε μοντέλα**. Το τελευταίο επίτευγμα είναι μοντέλα εξειδικευμένα για κάθε ασθενή (patient specific modeling).
- ◆ **Λογική κατάσταση**. Ο λογικός συλλογισμός παίρνει αποφάσεις ανάλογα με την αξία μιας δεδομένης μεταβλητής. Τα αποτελέσματα της διαδικασίας λήψης αποφάσεων διαφέρουν αν η μεταβλητή είναι εντός ή εκτός των καθορισμένων ορίων.
- ◆ **«Εξόρυξη» δεδομένων και μηχανική μάθηση**. Αυτές οι μέθοδοι βασίζονται σε πιθανοτική λήψη αποφάσεων σύμφωνα με τη βάση δεδομένων του συστήματος. Οι ιδανικές βάσεις δεδομένων πρέπει να είναι μεγάλες και καλά κατασκευασμένες, ώστε να επιτρέπουν την ακριβή ανάκτηση δεδομένων ασθενών, παρόμοιων με τον τρέχοντα ασθενή. Η ανάλυση της απόκρισης

αυτών των ασθενών σε διάφορες θεραπείες χρησιμοποιείται για την επιλογή της καλύτερης θεραπείας για τον τρέχοντα ασθενή.

- ◆ **Γενετικοί αλγόριθμοι.** Ως μέθοδος που δε βασίζεται σε βάση γνώσης, χρησιμοποιούν επαναληπτικές διαδικασίες για να αναδιατάσσονται και να παρέχουν τη βέλτιστη λύση με βάση τα δεδομένα του ασθενή.

#### β) Επείγουσα κλινική κατάσταση.

Πρόκειται για την παροχή υποστήριξης αποφάσεων για αποφάσεις που πρέπει να ληφθούν επείγοντως. Το σύστημα CDSS πρέπει πρώτα να προβλέπει τη λειτουργία σε θέματα με σαφή κλινική προτεραιότητα, σύμφωνα με την αρχή «θεραπεύετε πρώτα ό,τι σκοτώνει πρώτο». Αυτό το χαρακτηριστικό οδηγεί σε καλύτερα αποτελέσματα για τους ασθενείς και καλύτερη απόδοση των γιατρών.

#### γ) Σαφήνεια των συστάσεων.

Οι χρήστες είναι πιο πιθανό να ακολουθήσουν σαφείς συστάσεις που παρέχουν συγκεκριμένη πορεία δράσης.

#### δ) Απαίτηση απάντησης.

Ο ιατρός – χρήστης ενδέχεται να υποχρεωθεί να αιτιολογήσει τον τρόπο που ανταποκρίθηκε στη σύσταση που του παρείχε το σύστημα CDSS. Αυτό θα μπορούσε να γίνει με μια μορφή αναγνώρισης της σύστασης, με μια δήλωση για το ποια εναλλακτική δράση ελήφθη και μια επεξήγηση της μη συμμόρφωσης στις συστάσεις.

### **1.3.4 Παροχή πληροφοριών προς τον χρήστη**

Αυτός ο άξονας ασχολείται με την παράδοση πρόσφατα παραχθέντων πληροφοριών στο χρήστη.

#### α) Μορφή παράδοσης πληροφοριών.

Σε χαρτί, ηλεκτρονικά (μέσω διαδικτύου ή ενσωματωμένες στο ηλεκτρονικό ιατρικό αρχείο), μέσω άλλων τεχνολογιών – τηλέφωνο, τηλεϊδοποιητής, e-mail κτλ.

#### β) Τρόπος παράδοσης.

Οι συστάσεις μπορούν να παραδοθούν χωρίς συναίνεση σε μορφή ειδοποίησης, υπενθύμισης ή βελτιστοποίησης της αίτησης. Στην πρώτη περίπτωση, όταν πρόκειται για παθητικό σύστημα, ο γιατρός πρέπει να κάνει μια επιπλέον προσπάθεια, να αναγνωρίζει τότε η συμβουλή θα ήταν χρήσιμη, να «απευθύνεται» στο πρόγραμμα και να εισάγει δεδομένα αναζητώντας διαγνωστική ή θεραπευτική αξιολόγηση. Τα λεγόμενα «συστήματα ώθησης» που παρέχουν αυτόματα συστάσεις, μπορεί να είναι πιο αποτελεσματικά και ουσιαστικά να χρησιμοποιούνται περισσότερο. Παίζουν ενεργό ρόλο στην παροχή υποστήριξης αποφάσεων ως υποπροϊόν των δραστηριοτήτων διαχείρισης δεδομένων (πχ παρακολούθηση, εποπτεία ηλεκτρονικού ιατρικού

αρχείου). Η λογική απόφασης τους συστήματος είναι κατά κάποιο τρόπο ενσωματωμένη στη βάση δεδομένων του ασθενή που έχει ήδη δημιουργηθεί από διάφορες πηγές και παρέχει αποτελέσματα της ανάλυσης αποφάσεων χωρίς να χρειάζεται επιπλέον προσπάθεια από γιατρό. Ένα σημαντικό σημείο που πρέπει να ληφθεί υπόψη εδώ είναι πως θα αποφευχθεί η λεγόμενη κόπωση εσφαλμένου συναγερμού (*'alarm fatigue'*), όταν ο γιατρός κατακλύζεται από μικρές ασυμφωνίες που σημειώνονται διαφορετικά αλλά είναι ευρέως κατανοητές.

#### γ) Ολοκλήρωση δράσης.

Είναι επιτακτική ανάγκη το σύστημα CDSS να παρέχει τη δυνατότητα στον υπεύθυνο λήψης της απόφασης να εφαρμόζει τις προτεινόμενες ενέργειες με ευκολία. Για παράδειγμα, το λογισμικό μπορεί, ενώ παρέχει οδηγίες για τη θεραπεία, να παρέχει επίσης απευθείας σύνδεση στο τμήμα σχεδιασμού της θεραπείας, στο ηλεκτρονικό ιατρικό αρχείο. Η δράση της αλλαγής της θεραπείας θα πρέπει να ολοκληρώνεται μέσα σε λίγα βήματα, για παράδειγμα «τσεκάροντας» ένα σημείο. Η ολοκλήρωση της δράσης αναμφίβολα συνεισφέρει στην ευρύτερη αποδοχή και χρηστικότητα των συστημάτων CDSS.

#### δ) Διαθεσιμότητα εξήγησης.

Πρόκειται για μια λειτουργία όπου το σύστημα παρέχει μια επεξήγηση των συστάσεων του μέσω συνδέσεων με τεκμηριωμένα άρθρα, βιβλία ή απευθείας στη βάση γνώσεων.

### **1.3.5 Ροή εργασιών**

Τα συστήματα CDSS μπορούν να θεωρηθούν ως μια διαδικασία, ωστόσο στον τομέα της περίθαλψης πρόκειται ουσιαστικά για μια τεχνολογική παρέμβαση, που θα μπορούσε να λειτουργήσει και ως διαταραχή. Τα συστήματα που δρουν συνδυαστικά με τη ροή εργασίας του εκάστοτε ιατρικού κέντρου είναι πιθανό να παρουσιάσουν αύξηση της χρήσης τους και να αποδειχθούν πιο αποτελεσματικά στη βελτιστοποίηση της απόδοσης των γιατρών.

## **1.4 Παράγοντες επιτυχίας των συστημάτων CDSS**

Παρά το γεγονός ότι τα ηλεκτρονικά συστήματα CDSS ήταν συνεχώς σε εξέλιξη από τη δεκαετία του 1970, το αντίκτυπο τους στην καθημερινή κλινική πρακτική δεν ήταν τόσο ισχυρό όσο αναμενόταν. Τα πιθανά οφέλη της χρήσης ηλεκτρονικών συστημάτων υποστήριξης αποφάσεων στην κλινική πράξη εμπίπτουν σε τρεις μεγάλες κατηγορίες [2]:

1. *Βελτίωση της ασφάλειας των ασθενών* (μείωση της λανθασμένης φαρμακευτικής αγωγής και των ανεπιθύμητων παρενεργειών, βελτιστοποίηση της παραγγελίας φαρμάκων και ιατρικών εξετάσεων)
2. *Βελτίωση της ποιότητας της περίθαλψης* (σωστή κατανομή του χρόνου των γιατρών στη φροντίδα των ασθενών, αύξηση της εφαρμογής των κλινικών κατευθυντήριων

γραμμών, ενθάρρυνση της χρήσης των πιο πρόσφατων κλινικών ευρημάτων, βελτιωμένη τεκμηρίωση και ικανοποίηση των ασθενών)

3. *Βελτίωση της αποτελεσματικότητας της ιατρικής περίθαλψης* (μείωση του κόστους μέσω ταχύτερης επεξεργασίας των εντολών, μείωση των δοκιμών επανάληψης, μείωση ανεπιθύμητων ενεργειών και αλλαγή των μοτίβων στη συνταγογράφηση, ευνοώντας φθηνότερα αλλά εξίσου αποτελεσματικά αντίγραφα προϊόντα – γενόσημα)

Η ανάπτυξη των συστημάτων CDSS είναι μια δύσκολη διαδικασία, η οποία μπορεί να οδηγήσει σε αποτυχία παρά την ύπαρξη θεωρητικών γνώσεων πάνω στο θέμα. Η κατανόηση των βαθύτερων αιτιών που οδηγούν είτε σε επιτυχία είτε σε αποτυχία μπορεί να συμβάλει στη βελτίωση της αποτελεσματικότητας της ανάπτυξης και εφαρμογής του συστήματος CDSS στην καθημερινή πρακτική. Διάφορες αστοχίες μπορεί να προέρχονται από διάφορες φάσεις ανάπτυξης και υλοποίησης : αποτυχία στην τεχνική ολοκλήρωση ενός κατάλληλου συστήματος, αποτυχία του συστήματος να γίνει αποδεκτό από τους χρήστες και να ενσωματωθεί στο περιβάλλον του χρήστη.

Εκτιμάται ότι το 45% των ηλεκτρονικών ιατρικών πληροφοριακών συστημάτων αποτυγχάνουν λόγω του χρήστη, έστω και αν τα συστήματα αυτά είναι τεχνολογικά συνεχή [3]. Κάποιοι λόγοι που δικαιολογούν ένα τόσο υψηλό ποσοστό αποτυχίας ίσως προκύπτουν από ανεπαρκή ικανότητα του υπολογιστικού συστήματος, περιορισμένη επαγγελματική αυτονομία, έλλειψη συνειδητοποίησης των μακροπρόθεσμων οφελών από τη χρήση των συστημάτων CDSS και έλλειψη επιθυμίας να αλλάξει η καθημερινή ροή εργασίας. Υπάρχει επίσης σαφής απόδειξη ότι οι υπηρεσίες των CDSS δεν χρησιμοποιούνται πάντα όταν υπάρχουν, καθώς υπεράριθμες ειδοποιήσεις των συστημάτων παρακάμπτονται ή αγνοούνται από τους γιατρούς.

Παρά τα προβλήματα και τις αδυναμίες που μπορεί να συνοδεύουν τα συστήματα CDSS, έχει αποδειχθεί ότι τα συστήματα αυτά βελτιώνουν την επιλογή φαρμάκων και τις συστάσεις δοσολογίας τους, μειώνουν σοβαρά λάθη φαρμακευτικής αγωγής επισημαίνοντας πιθανές αντιδράσεις των φαρμάκων, αλλεργίες και εντοπίζοντας την επικάλυψη θεραπειών. Με αυτό τον τρόπο, ενισχύουν την παροχή υπηρεσιών προληπτικής φροντίδας και τη συμμόρφωση στα συνιστώμενα πρότυπα περίθαλψης.

Πρόσφατες μελέτες δείχνουν ότι υπάρχουν κάποια χαρακτηριστικά των συστημάτων CDSS που είναι ζωτικής σημασίας για την επιτυχία τους [4,5]:

- ❖ Τα συστήματα CDSS θα πρέπει να παρέχουν υποστήριξη στη λήψη αποφάσεων αυτόματα ως μέρος της εργασίας των ιατρών, καθώς τα συστήματα όπου οι ιατροί απαιτείται να αναζητήσουν συμβουλές χειροκίνητα δεν έχουν αποδειχθεί ιδιαίτερα επιτυχημένα.
- ❖ Η υποστήριξη αποφάσεων θα πρέπει να παρέχεται στο χρόνο και τόπο της λήψης αποφάσεων. Αν ο γιατρός υποχρεούται να διακόψει το φυσιολογικό μοτίβο φροντίδας του ασθενούς για να μετακινηθεί σε ένα ξεχωριστό σταθμό εργασιών ή να ακολουθήσει πολύπλοκες, χρονοβόρες διαδικασίες εκκίνησης, είναι πιθανό ένα τέτοιο σύστημα να μην είναι καλοδεχούμενο.

- ❖ Τα συστήματα που αποτελούν ένα ολοκληρωμένο δομικό στοιχείο χαρτογράφησης ή παραγγελιών έχουν σημαντικά περισσότερες πιθανότητες να επιτύχουν από τα μεμονωμένα συστήματα. Σε γενικές γραμμές, το στοιχείο υποστήριξης της απόφασης θα πρέπει να ενσωματωθεί σε ένα μεγαλύτερο υπολογιστικό σύστημα που ήδη αποτελεί μέρος της ρουτίνας των ιατρών – χρηστών, καθιστώντας έτσι την υποστήριξη αποφάσεων ένα υποπροϊόν των απλών πρακτικών εργασίας.
- ❖ Θεωρείται ότι τα υπολογιστικά συστήματα πλεονεκτούν σε σχέση με τα συστήματα που βασίζονται σε χαρτί.
- ❖ Τα συστήματα θα πρέπει να παρέχουν συστάσεις και όχι απλώς μια εκτίμηση της κατάστασης του ασθενή. Για παράδειγμα, το σύστημα συνιστά ότι ο γιατρός πρέπει να συνταγογραφήσει διουρητικά σκευάσματα για κάποιον ασθενή και δεν αναγνωρίζει απλά ότι ο ασθενής έχει καρδιολογικό πρόβλημα.
- ❖ Το σύστημα CDSS πρέπει να ζητά από τον γιατρό να καταγράψει ένα λόγο μη συμμόρφωσης στις συμβουλές του συστήματος (ζητείται από το γιατρό να δικαιολογήσει την απόφαση με ένα λόγο, πχ «ο ασθενής αρνήθηκε»).
- ❖ Θα πρέπει να προωθεί τη δράση και όχι την απραξία των γιατρών.
- ❖ Δεν υπάρχει ανάγκη για πρόσθετη εισαγωγή κλινικών δεδομένων. Λόγω της προσπάθειας των γιατρών που απαιτείται για την εισαγωγή νέων δεδομένων του ασθενή, τείνουν να αποφεύγουν τη διαδικασία, που είναι απαραίτητη για νέα υποστήριξη αποφάσεων. Τα συστήματα θα πρέπει να αποκτούν νέα δεδομένα αυτόματα (πχ ανάκτηση δεδομένων από το ηλεκτρονικό ιατρικό αρχείο).
- ❖ Το σύστημα θα πρέπει να είναι εύκολο στη χρήση και την πλοήγηση, πχ με γρήγορη πρόσβαση και ελάχιστα βήματα μέσω υπολογιστή για τις επιθυμητές πληροφορίες.
- ❖ Το χρονοδιάγραμμα και η συχνότητα των συστάσεων έχουν μεγάλη σημασία. Για παράδειγμα, αν υπάρχουν πάρα πολλά μηνύματα, αυτά ενδεχομένως να αγνοηθούν και κατά συνέπεια να χαθούν σημαντικές πληροφορίες. Η χρονική στιγμή είναι επίσης πολύ σημαντική – οι ειδοποιήσεις δεν θα πρέπει να εμφανίζονται σε ακατάλληλες στιγμές και να διακόπτουν τη ροή εργασίας.
- ❖ Η παρουσίαση των δεδομένων ή των πληροφοριών στα συστήματα CDSS δεν θα έπρεπε να είναι πολύ πυκνή ή το κείμενο πολύ μικρό. Οι ερευνητές προτείνουν επίσης τη χρήση εικονιδίων που αναβοσβήνουν για σημαντικά καθήκοντα ή τη ρύθμιση των αλληλεπιδράσεων ανάλογα με τον επείγοντα χαρακτήρα τους.
- ❖ Τα αποτελέσματα της υποστήριξης αποφάσεων πρέπει να παρέχονται τόσο στους γιατρούς όσο και στους ασθενείς. Έρευνες έχουν δείξει θετικά αποτελέσματα τέτοιων δράσεων, επειδή προτρέπουν τους γιατρούς να συζητούν τις επιλογές θεραπείας με τους ασθενείς και κατά συνέπεια να τους κάνει να αισθάνονται ότι συμμετέχουν στη θεραπεία τους.
- ❖ Περιοδική ανατροφοδότηση σχετικά με τη συμμόρφωση του γιατρού στη λήψη αποφάσεων του συστήματος.

Αυτό που έχουν κοινό όλα τα παραπάνω χαρακτηριστικά είναι ότι καθιστούν πιο εύκολο για τους γιατρούς να υλοποιούν το σύστημα CDSS μέσα στη ροή εργασιών τους, ώστε να είναι πιο εύκολο στη χρήση. Ένα αποτελεσματικό σύστημα CDSS πρέπει να ελαχιστοποιεί την προσπάθεια να λαμβάνει και να ενεργεί με τις συστάσεις του συστήματος. Οι ιατροί θεωρούν επίσης ότι είναι πολύ πρακτικό το σύστημα CDSS να μπορεί

να στηρίζει τις αποφάσεις που λαμβάνει όντας συνδεδεμένο με άλλες πηγές γνώσης μέσω του διαδικτύου. Κατά τη γνώμη τους, η ασφάλεια και οι ειδοποιήσεις για την αλληλεπίδραση των φαρμάκων ήταν το πιο χρήσιμο χαρακτηριστικό. Πάνω από όλους τους οργανωσιακούς παράγοντες όμως, όπως είναι η διαθεσιμότητα ηλεκτρονικών υπολογιστών στον τομέα της υγειονομικής περίθαλψης και η τεχνική τελειότητα του CDSS, το υλικό και λογισμικό είναι ζωτικής σημασίας για την εφαρμογή.

Ο Kawamoto (2005), τέλος, στο άρθρο του υποδηλώνει ότι η αποτελεσματικότητα του CDSS παραμένει κυρίως αμετάβλητη όταν οι συστάσεις τους συστήματος δηλώνονται πιο ισχυρά και όταν η απόδειξη που υποστηρίζει τις προτροπές του συστήματος επεκτείνεται και περιλαμβάνει δεδομένα συγκεκριμένου οργανισμού [5]. Ομοίως, η αποτελεσματικότητα και η λειτουργικότητα παραμένουν αναλλοίωτες όταν οι συστάσεις γίνονται πιο συγκεκριμένες. Είναι ενδιαφέρον το γεγονός ότι τα συστήματα CDSS δεν επέτυχαν τα αναμενόμενα αποτελέσματα όταν προσελήφθησαν τοπικοί γιατροί στη διαδικασία ανάπτυξης του συστήματος, ούτε όταν δόθηκαν βιβλιογραφικές αναφορές για την υποστήριξη του συστήματος στο να δοθούν συστάσεις [2].

Συνοψίζοντας, κατά την ανάπτυξη των συστημάτων CDSS, υπάρχουν παράγοντες που πρέπει να ληφθούν υπόψη, πέρα από το λογισμικό και το περιεχόμενο. Θεμελιώδη ζητήματα περιλαμβάνουν τη διαθεσιμότητα και προσβασιμότητα του υλικού, επαρκή τεχνική υποστήριξη και εκπαίδευση στη χρήση του συστήματος, το επίπεδο της ενσωμάτωσης του συστήματος στην κλινική ροή εργασίας και την καταλληλότητα και την επικαιρότητα των ιατρικών μηνυμάτων που παρέχονται.

## 1.5 Αναφορές Κεφαλαίου

- [1] Κολοστούμπης Γ., Μακρυγιαννάκη Κ. (2012), «Συστήματα Υποστήριξης Κλινικών Αποφάσεων : Ένα Χρήσιμο Εργαλείο στην Κλινική Πράξη», Διεπιστημονική Φροντίδα Υγείας, Τόμος 4, Τεύχος 1
- [2] Dejan Dinevski, Uroš Bele, Tomislav Šarenac, Uroš Rajkovič and Olga Šušteršič (2011), 'Clinical Decision Support Systems, Telemedicine Techniques and Applications', Prof. Georgi Graschew (Ed.), ISBN: 978-953-307-354-5, InTech, Available from: <http://www.intechopen.com/books/telemedicine-techniques-andapplications/clinical-decision-support-systems>
- [3] Zheng, K., R. Padman, et al. (2005). "Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system." Int J Med Inform 74(7-8): 535-543.
- [4] Shortliffe, E. H. and J. J. Cimino (2006) 'Biomedical informatics : computer applications in health care and biomedicine', New York, Springer
- [5] Kensaku Kawamoto, Caitlin A Houlihan, E Andrew Balas, David F Lobach (2005), 'Improving clinical practice using clinical decision support systems : a systematic review of trials to identify features critical to success', BMJ 330(7494):765



# Κεφάλαιο 2 - Μηχανική Μάθηση,

## Αναγνώριση Προτύπων και

## Ταξινόμηση

### 2.1 Εισαγωγή

Πριν προχωρήσουμε στην ανάλυση των αλγορίθμων ταξινόμησης, κρίνεται απαραίτητη μια εισαγωγή σε βασικές έννοιες που αφορούν τα συστήματα υποστήριξης κλινικής απόφασης, όπως η μηχανική μάθηση, η αναγνώριση προτύπων και η ταξινόμηση. Πρόκειται για έννοιες αλληλένδετες μεταξύ τους, τις οποίες θα επιχειρήσουμε να αποσαφηνίσουμε παρακάτω.

Η *μηχανική μάθηση (machine learning)* είναι μια περιοχή της τεχνητής νοημοσύνης, η οποία περιλαμβάνει τον σχεδιασμό αλγορίθμων για τη δημιουργία ενός αυτόματου συστήματος που θα αποκτά γνώση βασιζόμενο σε εμπειρικά δεδομένα. Ένας αρκετά γενικός ορισμός που θα μπορούσε να δοθεί για την μηχανική μάθηση είναι ο εξής : «Ένα πρόγραμμα υπολογιστή μαθαίνει από την εμπειρία  $E$  ως προς κάποια κλάση εργασιών  $T$  και μέτρο απόδοσης  $P$ , αν η απόδοσή του σε εργασίες από το  $T$ , όπως μετρείται από το  $P$ , βελτιώνεται μέσω της εμπειρίας  $E$ » [1,2].

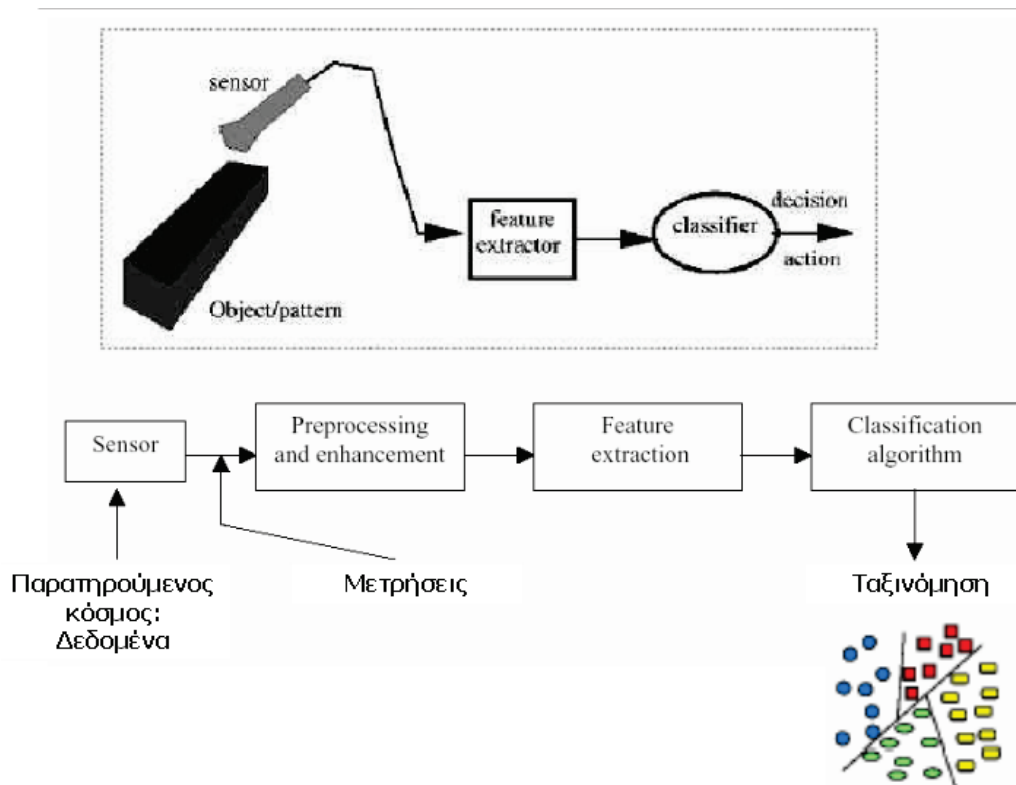
Δεν θα μπορούσε να διατυπωθεί βέβαια ο ισχυρισμός ότι η επιστήμη έχει φτάσει σε σημείο να ερμηνεύσει τη διαδικασία μάθησης του ανθρώπου, έχει όμως επιτευχθεί η δημιουργία αλγορίθμων οι οποίοι μπορούν να αυτοματοποιήσουν την κατασκευή ευφυών συστημάτων χρησιμοποιώντας δεδομένα εκπαίδευσης. Εδώ πρέπει να διευκρινιστεί επίσης ότι εξόρυξη γνώσης σημαίνει χρήση μεθόδων μηχανικής μάθησης σε μεγάλες βάσεις δεδομένων.

Συνοπτικά, η μηχανική μάθηση χρησιμοποιείται κυρίως για τρεις σκοπούς :

- **Πρόβλεψη τάσεων και συμπεριφορών.**
- **Αναγνώριση προτύπων.** Οι τυποποιημένες μορφές ανάμεσα στα δεδομένα μπορούν να χρησιμοποιηθούν για να αποκαλύψουν την ύπαρξη ενός γεγονότος, μιας δραστηριότητας. Οι άνθρωποι έχουν την ικανότητα να ταυτοποιούν πραγματικά δεδομένα χρησιμοποιώντας τις αισθήσεις τους και την αντιληπτική τους ικανότητα προκειμένου να λάβουν τις κατάλληλες αποφάσεις ώστε να επιβιώσουν στο περιβάλλον τους. Μία μηχανή, όπως ένας ηλεκτρονικός υπολογιστής, πρέπει να εκπαιδευθεί κατάλληλα ώστε να αναγνωρίζει πρότυπα (patterns) και να τα κατηγοριοποιεί αυτόματα σε κατηγορίες. Η διαδικασία αυτή είναι γνωστή ως μέθοδος μάθησης (*learning methodology*) και τα χρησιμοποιούμενα δεδομένα ως δεδομένα εκπαίδευσης (*training data*). Ανάλογα με την εφαρμογή, γίνεται κατάταξη των αντικειμένων σε κλάσεις με τη βοήθεια αλγορίθμων ταξινόμησης. Εφαρμογές της αναγνώρισης προτύπων αξιοποιούνται σε

πολλά επιστημονικά πεδία, όπως στην ιατρική (συστήματα CDSS, ανάλυση δεδομένων DNA και άλλες εφαρμογές της βιοπληροφορικής) και την επιστήμη υπολογιστών (υπολογιστική όραση, αναγνώριση χαρακτήρων ή φωνής, νευρωνικά δίκτυα, εξόρυξη δεδομένων και ανάκτηση γνώσης, τεχνητή νοημοσύνη).

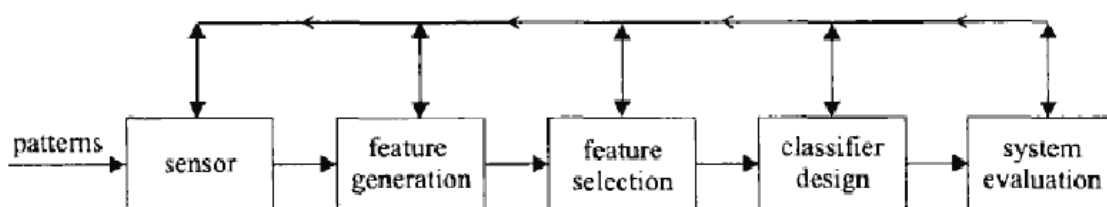
- Ταξινόμηση.** Το πρόβλημα της ταξινόμησης (classification) συνιστάται στην εκμάθηση μιας συνάρτησης στόχου ικανής να αντιστοιχεί άγνωστα αντικείμενα σε προκαθορισμένο σύνολο κατηγοριών. Στόχος της διαδικασίας μάθησης είναι η δημιουργία ενός μοντέλου πρόβλεψης της μη παρατηρούμενης ιδιότητας που είναι το διακριτικό (όνομα/τιμή) της κατηγορίας ενός αγνώστου αντικειμένου, βάσει των τιμών των χαρακτηριστικών που το προσδιορίζουν. Τυπικά παραδείγματα εφαρμογών ταξινόμησης αποτελούν η αυτόματη ταξινόμηση μηνυμάτων ηλεκτρονικού ταχυδρομείου από το χρήστη, ο διαχωρισμός καλοηθών από κακοήθεις όγκους μέσω παρατήρησης ιατρικών εξετάσεων κ.ά. Για την αντιμετώπιση του συγκεκριμένου προβλήματος, προϋποτίθεται η συγκέντρωση ενός αριθμού από αντιπροσωπευτικά για κάθε κατηγορία παραδείγματα (στιγμιότυπα του προβλήματος), τα οποία έχουν ταξινομηθεί από τον άνθρωπο – εκπαιδευτή. Στη συνέχεια, τα εν λόγω στιγμιότυπα του προβλήματος (σώμα δεδομένων εκπαίδευσης) δίνονται σε έναν αλγόριθμο μηχανικής μάθησης, ο οποίος κατόπιν της παρατήρησης των χαρακτηριστικών τους παράγει ως έξοδο ένα εκτιμώμενο μοντέλο, ικανό να αντιστοιχίζει άγνωστα στιγμιότυπα στις προδιαγεγραμμένες κατηγορίες. Το τελικό προϊόν της εκπαίδευσης καλείται *ταξινομητής (classifier)*.



**Εικόνα 1 :** Τυπικό σύστημα ταξινόμησης.

Τα προβλήματα μηχανικής μάθησης κατατάσσονται σε τρεις κύριες κατηγορίες ανάλογα με την έξοδο που επιθυμούμε να δώσει ο αλγόριθμος που χρησιμοποιείται. Έτσι, διακρίνουμε τις κατηγορίες της *μάθησης με επίβλεψη (supervised learning)*, της *μάθησης χωρίς επίβλεψη (unsupervised learning)* και της *ενισχυτικής μάθησης (reinforcement learning)*. Στην πρώτη κατηγορία, ανήκουν οι περιπτώσεις αλγορίθμων που δημιουργούν μια συνάρτηση τέτοια, ώστε να αντιστοιχίζει μια συγκεκριμένη είσοδο σε κάποια αυστηρά καθορισμένη έξοδο. Κάτι τέτοιο επιτυγχάνεται μέσω ενός συνόλου εκπαίδευσης αποτελούμενου από παραδείγματα ζευγών εισόδου και επιθυμητής εξόδου. Η μάθηση χωρίς επίβλεψη είναι μια μέθοδος που ενδιαφέρεται κυρίως να εκτιμήσει μια συνάρτηση κατανομής για το σύνολο εκπαίδευσης, βάσει συμπερασμάτων που εξαγει από αυτό, ενώ ασχολείται και με προβλήματα ομαδοποίησης. Εδώ δε γνωρίζουμε εξαρχής τη σωστή κατηγορία για κάποιο σύνολο παραδειγμάτων, αλλά προσπαθούμε να εκτιμήσουμε έναν κατάλληλο διαμερισμό των δεδομένων σε ομάδες. Εξαιτίας αυτού, τα προβλήματα αυτού του είδους είναι πιο δύσκολα. Στην περίπτωση, τέλος, της ενισχυτικής μάθησης, ο αλγόριθμος αποδίδει μια επιβράβευση (θετική ή και αρνητική), μετά από σειρά αποφάσεων. Σκοπός της είναι να μεγιστοποιήσει αυτή την επιβράβευση, η οποία εξαρτάται από το σύνολο αυτών των αποφάσεων, καθώς κάτι τέτοιο αποδεικνύει ότι αυτές μας οδήγησαν πολύ κοντά στον στόχο που μας ενδιαφέρει [1,2].

Επανερχόμαστε τώρα στη διαδικασία ταξινόμησης. Όπως απεικονίζεται στο παρακάτω διάγραμμα, αρχικά υπολογίζονται τα χαρακτηριστικά, στη συνέχεια επιλέγεται ο κατάλληλος αριθμός αυτών που θα χρησιμοποιηθούν στον αλγόριθμο ταξινόμησης, σχεδιάζεται ο ταξινομητής και τέλος το σύστημα αξιολογείται προκειμένου να ελεγχθεί ο ρυθμός του σφάλματος ταξινόμησης.

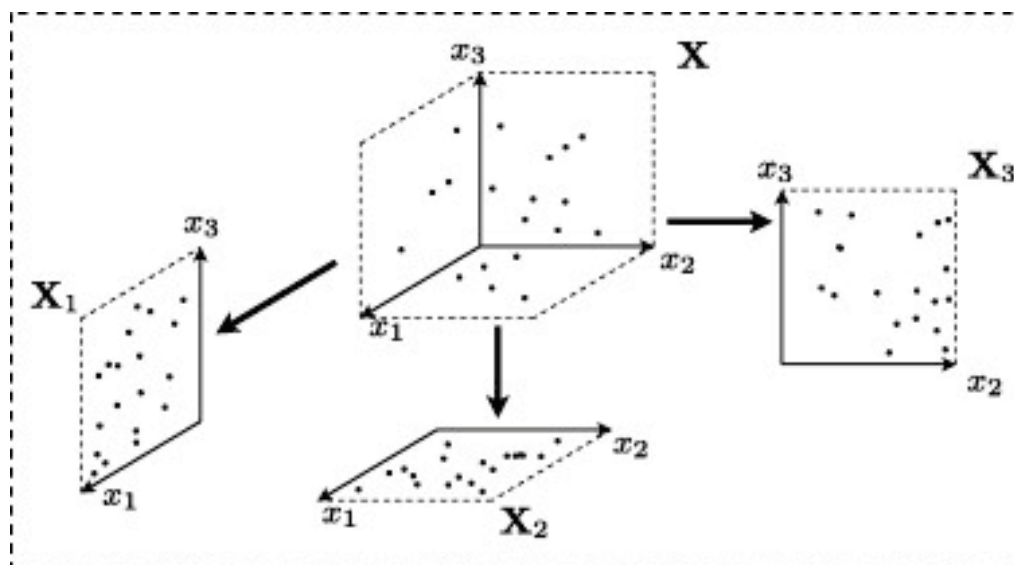


**Εικόνα 2 :** Διαδικασία ταξινόμησης.

## 2.2 Επιλογή Χαρακτηριστικών

Πριν προχωρήσουμε στην ανάλυση της επιλογής χαρακτηριστικών, κρίνεται σκόπιμη μια εισαγωγή στην έννοια του *χώρου χαρακτηριστικών (feature space)*. Στην μηχανική μάθηση και την αναγνώριση προτύπων, ο όρος *χώρος χαρακτηριστικών* αναφέρεται σε έναν ασαφή χώρο, όπου κάθε δείγμα προτύπου αναπαριστάται ως ένα σημείο στο χώρο  $n$  – διαστάσεων [3]. Η διάσταση του χώρου καθορίζεται από τον αριθμό των χαρακτηριστικών που χρησιμοποιούνται για να περιγράψουν τα μοτίβα. Παρόμοια δείγματα ομαδοποιούνται, γεγονός που επιτρέπει τη χρήση της εκτίμησης πυκνότητας για την εύρεση προτύπων.

Το διάγραμμα του χώρου χαρακτηριστικών είναι ένα διαδραστικό γράφημα του χώρου χαρακτηριστικών (ή ενός υποχώρου αυτού, αν υπάρχουν περισσότερα από τρία χαρακτηριστικά). Κάθε άξονας αντιπροσωπεύει ένα χαρακτηριστικό του μοντέλου, ενώ η θέση των σημείων στο διάγραμμα δείχνει τις τιμές των χαρακτηριστικών αυτών. Εκτός από τις τιμές των χαρακτηριστικών, τα σημεία στο διάγραμμα μεταφέρουν περαιτέρω πληροφορίες, όπως για παράδειγμα το σχήμα, το οποίο δείχνει το διαμέρισμα στο οποίο ανήκει ένα σημείο ή το χρώμα/σκίαση του σημείου, που είναι ενδεικτικό της αξίας του στόχου για την εν λόγω περίπτωση.



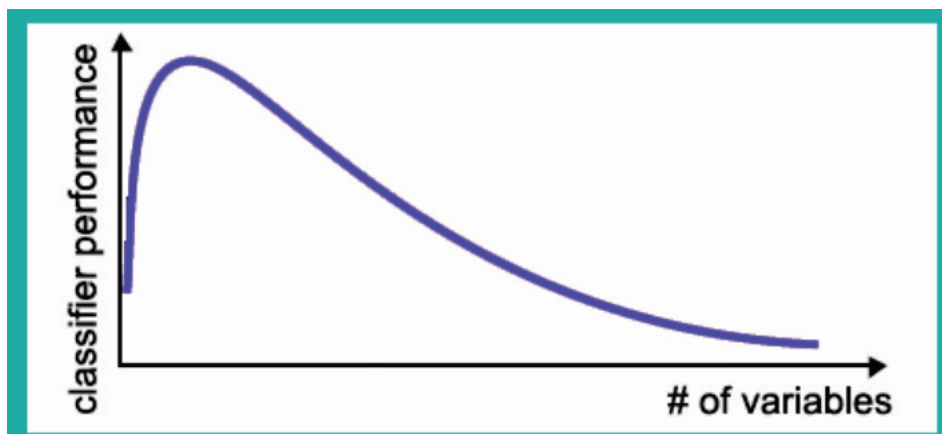
**Εικόνα 3 :** Χώρος χαρακτηριστικών 3 διαστάσεων και σύμπλεγμα υποχώρων. Οι υποχώροι 2 διαστάσεων  $X_1$ ,  $X_2$  και  $X_3$  αποκτώνται από το χώρο 3 διαστάσεων  $X$  με απλή προβολή. Οι μονάδες στο γράφημα είναι άνευ σημασίας.

Όπως θα δούμε παρακάτω, αυτό που μας ενδιαφέρει είναι η μείωση του αριθμού των χαρακτηριστικών στον ελάχιστο δυνατό, γεγονός που διευκολύνει και την απεικόνιση του χώρου χαρακτηριστικών.

Η επιλογή των χαρακτηριστικών είναι το τμήμα του συστήματος με την μεγαλύτερη ευαισθησία στο είδος των δεδομένων [3]. Ο τύπος του αλγορίθμου που θα χρησιμοποιηθεί, καθώς και οι παράμετροι του, θα επηρεάσουν σε μεγάλο βαθμό το τελικό αποτέλεσμα, καθώς σε αυτό το στάδιο επιλέγονται οι εισόδους του ταξινομητή. Άρα, όσο κατάλληλα κατασκευασμένος και αν είναι ο τελευταίος, αν δεν του δοθεί η έγκυρη πληροφόρηση που θα προέλθει από αυτό το στάδιο, δεν πρόκειται να αποδώσει σωστά.

Ένας ταξινομητής εκμεταλλεύεται τα χαρακτηριστικά εκείνα που καθιστούν τα δεδομένα εισόδου διαχωρίσιμα μεταξύ τους, ώστε να κατανεμηθούν μετέπειτα στις αντίστοιχες κλάσεις. Αν και σε πρώτη ανάγνωση, φαντάζει εύλογο το γεγονός ότι όσο περισσότερα χαρακτηριστικά διαθέτει ο ταξινομητής, τόσο καλύτερη θα είναι η επίδοσή του, εντούτοις αύξηση του αριθμού των χαρακτηριστικών πάνω από ένα όριο μειώνει την απόδοση του ταξινομητή, όπως απεικονίζεται σχηματικά παρακάτω. Κατά συνέπεια, επιδιώκεται μείωση των χαρακτηριστικών σε έναν επαρκή ελάχιστο αριθμό, ενώ

δεδομένου ότι το πλήθος των χαρακτηριστικών μπορεί να είναι της τάξης των δεκάδων ή εκατοντάδων, η επιλογή των καταλληλότερων κάνει τη διαφορά στην απόδοση του ταξινομητή.

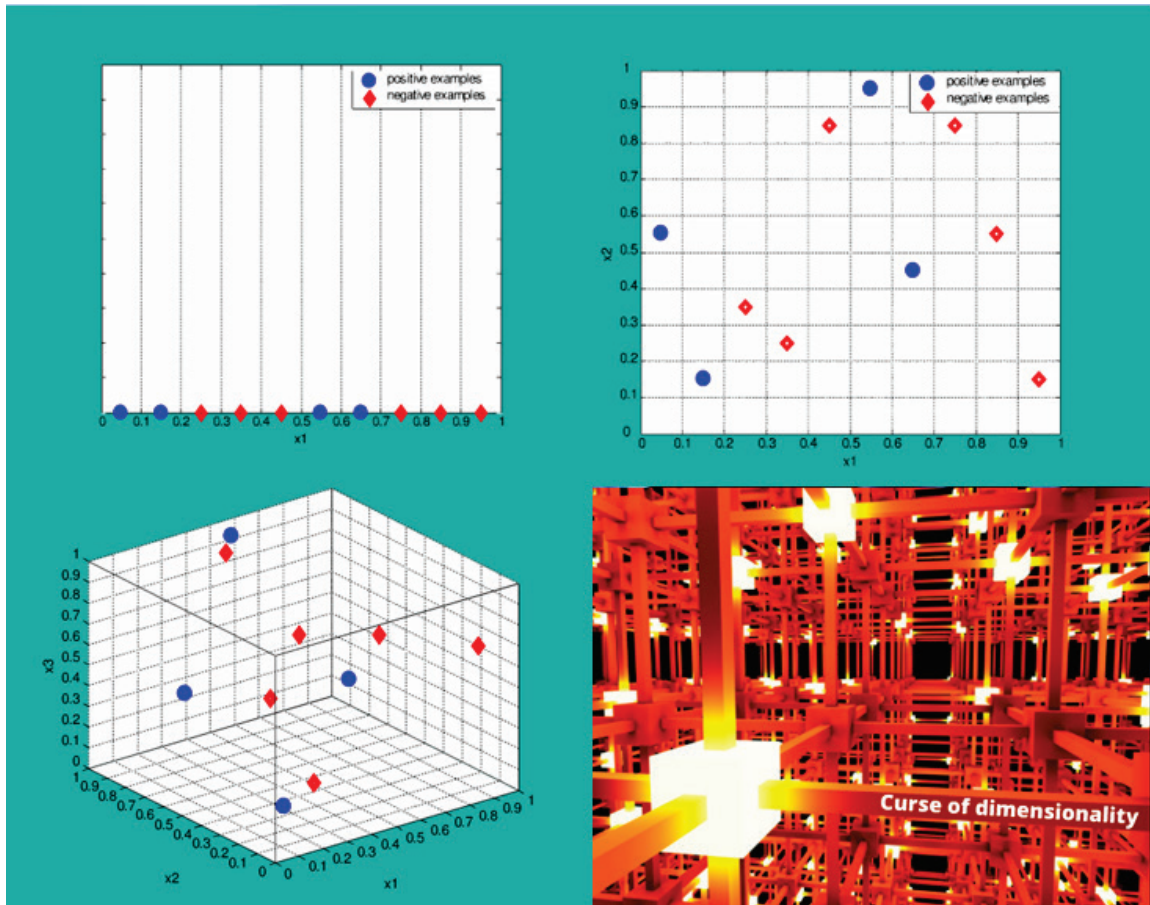


**Εικόνα 4 :** Συσχέτιση της απόδοσης του ταξινομητή με τον αριθμό χαρακτηριστικών.

Η μείωση της υπολογιστικής πολυπλοκότητας αποτελεί ίσως τον πιο σημαντικό λόγο μείωσης του αριθμού των χαρακτηριστικών.

Ένας άλλος λόγος που πρέπει να επισημανθεί είναι ότι, παρόλο που δύο χαρακτηριστικά μπορεί να φέρουν σημαντική πληροφορία κατηγοριοποίησης όταν αντιμετωπίζονται ξεχωριστά, όταν συνδυαστούν ενδέχεται να αποδίδουν λίγα οφέλη, λόγω της υψηλής αμοιβαίας πληροφορίας που διαμοιράζεται μεταξύ τους. Κατά συνέπεια, η πολυπλοκότητα αυξάνει χωρίς ουσιαστικό κέρδος.

Αξίζει να σημειώσουμε επιπροσθέτως ότι η μείωση των χαρακτηριστικών επιβάλλεται και από τη λεγόμενη «κατάρρα της διαστασιμότητας» (*curse of dimensionality*). Αυτό το φαινόμενο εμφανίζεται όταν ο αριθμός των χαρακτηριστικών που είναι διαθέσιμα στον σχεδιαστή του ταξινομητή είναι πολύ μεγάλος (δηλαδή σχεδόν πάντα στα πραγματικά προβλήματα), ενώ πρακτικά σημαίνει την εκθετική αύξηση των απαιτούμενων τιμών εκπαίδευσης με την αύξηση του αριθμού διαστάσεων του προβλήματος (δηλαδή του αριθμού των χρησιμοποιούμενων χαρακτηριστικών). Στη συνήθη περίπτωση, οι τιμές εκπαίδευσης που έχει στη διάθεσή του ο σχεδιαστής είναι λίγες. Όπως παρατηρούμε στις παρακάτω εικόνες, αυξάνοντας τις διαστάσεις, τα σημεία στο χώρο εισόδου κατανέμονται αραιά, με αποτέλεσμα τη φτωχή απεικόνιση της αντιστοίχισης στις μεταβλητές εξόδου:



**Εικόνα 5 :** Η «κατάρρα της διαστασιμότητας» - χώρος μεταβλητών εισόδου με μία, δύο και τρεις διαστάσεις.

Επιπλέον αυτών, ο αριθμός των χαρακτηριστικών επηρεάζει και τις ιδιότητες γενίκευσης του ταξινομητή, δηλαδή το πόσο καλά θα συμπεριφερθεί σε πρότυπα εισόδου που δεν έχει ξανασυναντήσει, αφού έχει ολοκληρώσει την εκπαίδευσή του. Πιο συγκεκριμένα, όσο μεγαλύτερος είναι ο λόγος των προτύπων εκπαίδευσης  $N$  προς τον αριθμό των ελεύθερων παραμέτρων ταξινόμησης  $I$ , τόσο καλύτερες είναι οι ιδιότητες γενίκευσης του τελικού ταξινομητή. Όμως ο αριθμός των χαρακτηριστικών είναι βασικό συστατικό των παραμέτρων ταξινόμησης. Συνεπώς ένας μεγάλος αριθμός χαρακτηριστικών μεταφράζεται άμεσα σε έναν μεγάλο αριθμό παραμέτρων ταξινόμησης (πχ τα συναπτικά βάρη σε ένα νευρωνικό, τα βάρη σε ένα γραμμικό ταξινομητή). Άρα, καθότι ο αριθμός  $N$  των προτύπων εκπαίδευσης είναι συγκεκριμένος και συνήθως περιορισμένος, το να διατηρήσουμε τον αριθμό των χαρακτηριστικών όσο το δυνατόν μικρότερο, είναι σε απόλυτη συμφωνία με την επιθυμία μας να σχεδιάσουμε ταξινομητές με υψηλό λόγο  $N/I$  και συνεπώς με καλές ιδιότητες γενίκευσης. Τέλος, ένα σημαντικό βήμα στη σχεδίαση ενός συστήματος ταξινόμησης είναι η απόδοση του σταδίου αξιολόγησης, στο οποίο εκτιμάται η πιθανότητα του σφάλματος ταξινόμησης. Η πιθανότητα αυτή βελτιώνεται (μειώνεται) όσο ο λόγος αυτός αυξάνεται.

Για τους λόγους αυτούς, επιθυμούμε τη μείωση των χαρακτηριστικών εισόδου του ταξινομητή στον απολύτως απαραίτητο. Αυτό επιτυγχάνεται με μια σειρά από αλγορίθμους

που ονομάζονται γενικά αλγόριθμοι επιλογής χαρακτηριστικών (Feature Selection Algorithms - FSA).

Στόχος ενός αλγορίθμου επιλογής χαρακτηριστικών είναι η εύρεση του βέλτιστου υποσυνόλου χαρακτηριστικών εισόδου, που θα τροφοδοτηθεί στη συνέχεια στον ταξινομητή. Ο αριθμός των επιλεγμένων χαρακτηριστικών που αποτελούν το βέλτιστο υποσύνολο, καθώς και η ταυτότητα τους, δηλαδή το ποια χαρακτηριστικά είναι αυτά, είναι τα δύο ζητήματα που καλείται να αντιμετωπίσει ο εκάστοτε αλγόριθμος. Το πόσο καλά θα τα καταφέρει, εξαρτάται πάντα από το είδος της εργασίας, καθώς ένας καλός αλγόριθμος για μια περίπτωση είναι πιθανό να μη λειτουργεί καλά σε κάποια άλλη και αντιστρόφως. Γι' αυτό εξάλλου υπάρχει και μεγάλη ποικιλία.

Γενικά πάντως, τείνουμε να διευκολύνουμε τη διαδικασία με το να προεπιλέγουμε τον αριθμό των επιθυμητών χαρακτηριστικών του τελικού υποσυνόλου. Συνεπώς ο αλγόριθμος δημιουργείται για να δώσει απάντηση στην εξής ερώτηση : «Δεδομένου ενός αριθμού χαρακτηριστικών, πώς μπορεί κάποιος να επιλέξει τα πιο σημαντικά από αυτά ώστε να μειώσει τον αριθμό τους και συγχρόνως να διατηρήσει όσο το δυνατόν περισσότερη διακριτική πληροφορία φέρουν;»

Η διαδικασία αυτή καλείται, όπως προαναφέρθηκε, επιλογή χαρακτηριστικών ή μείωση. Πρέπει να τονιστεί πως το βήμα αυτό είναι πολύ σημαντικό. Αν επιλεγούν χαρακτηριστικά με μικρή διακριτική ικανότητα, η απόδοση του ταξινομητή δεν θα είναι ικανοποιητική. Από την άλλη πλευρά, μεγάλη διακριτική ικανότητα συνεπάγεται αύξηση της πολυπλοκότητας στον ταξινομητή. Κατά συνέπεια, επιδιώκεται επιλογή χαρακτηριστικών με μεγάλη απόσταση μεταξύ διαφορετικών κλάσεων αλλά με μικρή διακύμανση στο εσωτερικό μιας κλάσης. Σε μια πιο ποσοτική περιγραφή, αυτό σημαίνει ότι επιζητούμε τα χαρακτηριστικά να παίρνουν απομακρυσμένες τιμές όταν ανήκουν σε διαφορετικές κλάσεις, αλλά κοντινές όταν ανήκουν στην ίδια κλάση.

Λαμβάνοντας υπόψη τα παραπάνω, το πρόβλημα της επιλογής χαρακτηριστικών υπό την μορφή επιβλεπόμενης επαγωγικής μάθησης διατυπώνεται ως εξής [3]:

Δεδομένου ενός συνόλου χαρακτηριστικών, επέλεξε και παρουσίασε ένα υποσύνολο με βάση μία από τις 3 κατώθι προσεγγίσεις :

- Το υποσύνολο με συγκεκριμένο μέγεθος που βελτιστοποιεί ένα κριτήριο επιλογής.
- Το υποσύνολο με το μικρότερο μέγεθος που ικανοποιεί ένα συγκεκριμένο περιορισμό στην μέτρηση του κριτηρίου επιλογής (βρίσκεται δηλαδή πάνω ή κάτω από ένα κατώφλι ανάλογα με το κριτήριο).
- Το υποσύνολο με τον καλύτερο συνδυασμό των δύο παραπάνω περιπτώσεων (μέγεθος – συγκεκριμένος περιορισμός), η οποία αποτελεί και τη γενικότερη περίπτωση.

Ο γενικευμένος στόχος προς επίτευξη είναι η βελτίωση του ταξινομητή, είτε υπό την έννοια της ταχύτητας, είτε της δυνατότητας γενίκευσης ή της απλότητας της απεικόνισης. Αυτό συντελεί στην καλύτερη κατανόηση των αποτελεσμάτων που προκύπτουν, στην μείωση του μεγέθους του αποθηκευτικού χώρου που χρειάζεται για τη λειτουργία του,



καθώς επίσης και στην μείωση του θορύβου που δημιουργείται από άσχετα ή περιττά χαρακτηριστικά, με τελικό όφελος να αποβληθεί τελικά όλη η περιττή πληροφορία.

Οι FSA μπορούν να κατηγοριοποιηθούν ανάλογα με το είδος της εξόδου που δίνουν :

- ♦ Στους αλγορίθμους εκείνους που αποδίδουν τα χαρακτηριστικά σε μια γραμμική σειρά (σημαντικότητας).
- ♦ Στους αλγορίθμους εκείνους που δίνουν ένα υποσύνολο των αρχικών χαρακτηριστικών.

## 2.3 Αλγόριθμοι μηχανικής μάθησης

Στο σημείο αυτό, κρίνεται σκόπιμη η αναφορά σε αλγορίθμους μηχανικής μάθησης και ταξινόμησης που θα συναντήσουμε παρακάτω, ώστε να αποσαφηνιστούν βασικές έννοιες και οι μεταξύ τους διαφορές.

### 2.3.1 Εύρεση Κοντινότερων Γειτόνων ( $k$ – Nearest Neighbors)

Στην αναγνώριση προτύπων, ο αλγόριθμος  $k$ -NN είναι μια μέθοδος κατηγοριοποίησης αντικειμένων με βάση τα  $k$  κοντινότερα σε αυτά πρότυπα στο χώρο των χαρακτηριστικών [4]. Ο αλγόριθμος  $k$ -NN είναι ένα είδος μάθησης βασισμένο σε στιγμιότυπα, όπου η συνάρτηση προσεγγίζεται μόνο τοπικά και όλοι οι υπολογισμοί αναβάλλονται μέχρι τη στιγμή της ταξινόμησης. Αποτελεί έναν από τους απλούστερους αλγορίθμους της μηχανικής μάθησης, καθώς ένα αντικείμενο κατηγοριοποιείται με βάση την πλειοψηφία των γειτόνων του, με το αντικείμενο να οδεύει προς την κατηγορία που υπερिशύχει ανάμεσα στους  $k$  κοντινότερους γείτονές του (όπου  $k$  είναι ένας θετικός, γενικά μικρός ακέραιος, προκαθορισμένος από το χρήστη). Εάν  $k=1$ , τότε το αντικείμενο απλά τοποθετείται στην κατηγορία του κοντινότερου γείτονα. Αυτό που πρέπει να καθοριστεί συνεπώς είναι η τιμή του  $k$  και η απόσταση που θα θεωρήσουμε. Η βέλτιστη τιμή του  $k$  εξαρτάται από τα δεδομένα. Γενικά, μεγαλύτερες τιμές του  $k$  μειώνουν την επίδραση του θορύβου στην κατηγοριοποίηση, αλλά κάνουν πιο αδρές τις διαχωριστικές γραμμές μεταξύ των κατηγοριών. Μια καλή τιμή για το  $k$  μπορεί να βρεθεί χρησιμοποιώντας διάφορες ευριστικές τεχνικές.

Η ίδια μέθοδος μπορεί να χρησιμοποιηθεί και για αναδρομή, απλά με το να αναθέσουμε την τιμή της ιδιότητάς του προτύπου στη μέση τιμή των τιμών των  $k$  κοντινότερων γειτόνων του. Μπορεί να είναι ωφέλιμη η χρήση βαρών στην συνεισφορά των γειτόνων, ώστε οι κοντινοί γείτονες να συνεισφέρουν περισσότερο στην μέση τιμή απ' ότι οι πιο απομακρυσμένοι. Ένας τυπικός τρόπος να πραγματοποιηθεί αυτό είναι να δοθεί σε γείτονα ένα βάρος  $1/d$ , όπου  $d$  είναι η απόσταση από το ζητούμενο πρότυπο. Αυτό αποτελεί μια γενίκευση της γραμμικής παρεμβολής.

Σε αυτό το σημείο πρέπει να ξεκαθαρίσουμε ότι ως γείτονες θεωρούνται εκείνα τα πρότυπα για τα οποία γνωρίζουμε την πραγματική τους κατηγορία, ενώ από αυτό το υποσύνολο προέρχονται τα αποτελέσματα της μεθόδου. Αυτό μπορούμε να πούμε πως αποτελεί το υποσύνολο εκπαίδευσης του αλγορίθμου, παρόλο που δεν πληροί τις



αυστηρές προδιαγραφές τις τυπικής θεώρησης των συνόλων εκπαίδευσης (όπως για παράδειγμα την ύπαρξη συγκεκριμένου βήματος εκπαίδευσης).

Η φάση εκπαίδευσης συνίσταται από τον προσδιορισμό του υποσυνόλου εκπαίδευσης (γειτόνων) και την αποθήκευση των ανάλογων προτύπων και των αντίστοιχων πληροφοριών κατηγοριοποίησης τους. Στη φάση κατηγοριοποίησης λοιπόν, ένα άγνωστο πρότυπο κατηγοριοποιείται με το να τοποθετείται στην κατηγορία που ανήκει η πλειοψηφία από τους  $k$  κοντινότερους γείτονές του.

Ο αλγόριθμος είναι ευαίσθητος στην τοπική δομή των δεδομένων. Το υποσύνολο εκπαίδευσης αποτελείται από διανύσματα στο πολυδιάστατο χώρο χαρακτηριστικών, κάθε ένα εκ των οποίων έχει και μια «ταμπέλα» που δίνει την πληροφορία για την κατηγορία στην οποία ανήκει.

Συνήθως χρησιμοποιείται η ευκλείδεια απόσταση ως μέθοδος υπολογισμού των αποστάσεων, ωστόσο είναι εφαρμόσιμη μόνο σε συνεχείς μεταβλητές. Για τις διακριτές μεταβλητές μπορούμε να χρησιμοποιήσουμε την απόσταση Hamming ή άλλες αντίστοιχες μεθόδους, όπως για παράδειγμα :

<b>Ευκλείδεια απόσταση</b>	$d(x, y) = \left( \sum_{i=1}^m  x_i - y_i ^2 \right)^{1/2}$
<b>Manhattan απόσταση</b>	$d(x, y) = \sum_{i=1}^m  x_i - y_i $
<b>Chebyshev απόσταση</b>	$d(x, y) = \max_{i=1, m}  x_i - y_i $

**Πίνακας 1 :** Μέθοδοι υπολογισμού αποστάσεων.

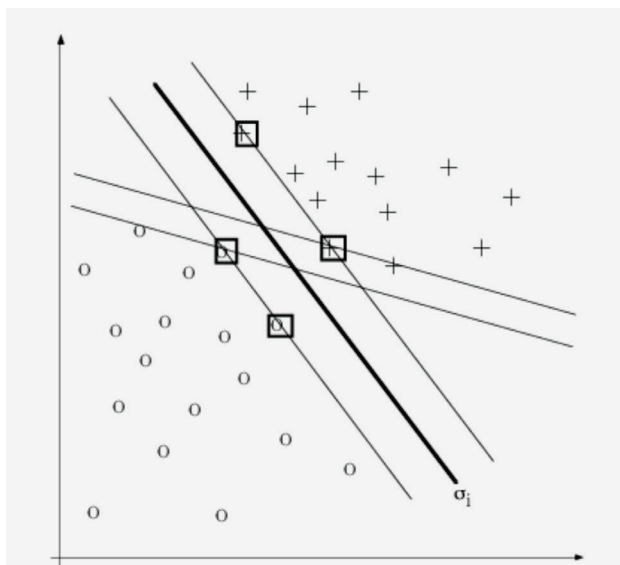
Συνοψίζοντας, θα λέγαμε ότι το κύριο πλεονέκτημα του παραπάνω αλγορίθμου είναι ότι προσεγγίζει ευκολότερα μια πολύπλοκη συνάρτηση - στόχο σε σχέση άλλες μεθόδους, ενώ προγραμματίζεται εύκολα. Επίσης, για μικρές αλλαγές στα δεδομένα εκπαίδευσης δεν παρατηρούνται μεγάλες αλλαγές στα αποτελέσματα ταξινόμησης. Όμως το μεγάλο μειονέκτημά του είναι το υπολογιστικό κόστος κατά την κατηγοριοποίηση των νέων στοιχείων και η αργή ταξινόμησή τους. Υπάρχει, επιπλέον, δυσκολία στην εύρεση του κατάλληλου  $k$ , ενώ οι κατηγορίες με τα πιο συχνά πρότυπα τείνουν να κυριαρχούν στην πρόβλεψη του αγνώστου προτύπου, καθώς είναι πιθανότερο να εμφανιστούν μέσα στους  $k$  κοντινότερους γείτονες, ειδικά αν το  $k$  δεν έχει μικρές τιμές.

### 2.3.2 Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines)

Οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines – SVMs) αποτελούν έναν αλγόριθμο ταξινόμησης προτύπων, ο οποίος βασίζεται στη στατιστική θεωρία μάθησης και αναπτύχθηκε από τον Vapnik [5,6]. Αποτελούν ένα είδος συγκερασμού

γραμμικών μοντέλων και της εύρεσης κοντινότερων γειτόνων. Για την ταξινόμηση, οι μηχανές διανυσμάτων υποστήριξης αναζητούν ένα υπερεπίπεδο (hyperplane) στο χώρο των πιθανών εισόδων. Αυτό το υπερεπίπεδο δημιουργείται ώστε να διαχωρίσει τα θετικά παραδείγματα από τα αρνητικά, επιδιώκοντας την μέγιστη απόσταση από το κοντινότερο θετικό και αρνητικό παράδειγμα. Με άλλα λόγια, το πρόβλημα θα μπορούσε να περιγραφεί ως εξής :

«Δοθέντος ενός συνόλου εκπαίδευσης, η μηχανή διανυσμάτων υποστήριξης αναζητά την κατασκευή ενός υπερεπιπέδου στην περιοχή των εισόδων, ως επιφάνεια απόφασης, έτσι ώστε ο διαχωρισμός θετικών και αρνητικών παραδειγμάτων να μεγιστοποιείται».



**Εικόνα 6 :** Διαχωριστικά υπερεπίπεδα για το SVM. Στο ένα υπερεπίπεδο διακρίνεται η μικρή απόσταση από το κοντινότερο από τα θετικά και αρνητικά παραδείγματα. Στο άλλο υπερεπίπεδο ( $\sigma_1$ ) φαίνεται η μέγιστη απόσταση μεταξύ των δύο κλάσεων η οποία λαμβάνεται τελικά.

Προφανώς, υπάρχουν πολλά υπερεπίπεδα που διαχωρίζουν τα θετικά και αρνητικά παραδείγματα, όμως αυτό που μας ενδιαφέρει περισσότερο είναι το βέλτιστο υπερεπίπεδο διαχωρισμού (optimal hyperplane), που είναι και μοναδικό. Για να πετύχουμε όμως μεγαλύτερη ταχύτητα στις μηχανές διανυσμάτων υποστήριξης από τις άλλες μεθόδους, πρέπει να γίνει επιλογή μικρού αριθμού στιγμιότυπων και από τις δύο κλάσεις, όταν αναφερόμαστε σε πρόβλημα με δύο κλάσεις.

Σημειώνεται ότι η μέθοδος αυτή ξεκινάει από προβλήματα με δύο κλάσεις στην έξοδο και μπορεί να γενικευτεί σε πολύ περισσότερες. Η απόσταση ανάμεσα στις δυο κλάσεις ονομάζεται περιθώριο (margin) και ένας καλός αλγόριθμος πρέπει να ελαχιστοποιεί το εμπειρικό λάθος (empirical risk), ενώ ταυτόχρονα μεγιστοποιεί την απόσταση διαχωρισμού (margin) :

$$R_{emp}[g] = \frac{1}{n} \sum_{i=1}^l (g(x_i), y_i)$$

Για τον λόγο αυτό ονομάζεται και ταξινομητής μεγάλου περιθωρίου (large margin classifier). Για την ταξινόμηση δεδομένων που ανήκουν σε επίπεδο μεγαλύτερο των δύο διαστάσεων (πολυδιάστατα δεδομένα που ανήκουν στο  $R^n$ ), εκμεταλλευόμαστε τις ιδιαίτερες ικανότητες των μηχανών αυτών. Κύριος στόχος είναι ο διαχωρισμός αυτών των δεδομένων με τη βοήθεια ενός υπερεπιπέδου  $n-1$  διαστάσεων, δηλαδή τη χρήση ενός γραμμικού ταξινομητή. Υπάρχουν πολλοί γραμμικοί ταξινομητές που μπορούν να βρουν ένα υπερεπίπεδο για το διαχωρισμό των δεδομένων, αλλά ο αλγόριθμος SVM διαφέρει, όπως αναφέραμε, διότι υπολογίζει το μέγιστο υπερεπίπεδο (ταξινομητής μέγιστου περιθωρίου).



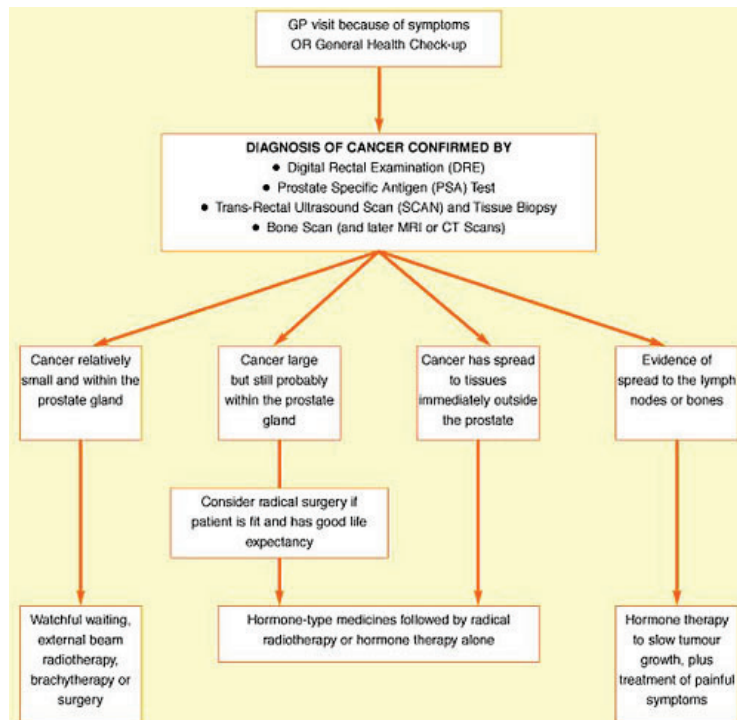
**Εικόνα 7 :** (a) Ο γραμμικός ταξινομητής διαχωρίζει τα σημεία στο επίπεδο. (b) Στην περίπτωση αυτή, ο γραμμικός ταξινομητής επιτυγχάνει μέγιστο διαχωρισμό (μέγιστο περιθώριο).

Αναφέρουμε, τέλος, ότι υπάρχουν δύο κατηγορίες εύρεσης του υπερεπιπέδου που αφορούν τα στιγμιότυπα, οι οποίες διαφοροποιούνται για γραμμικώς διαχωρίσιμα πρότυπα και για μη γραμμικώς διαχωρίσιμα πρότυπα.

### 2.3.3 Δέντρα Αποφάσεων (Decision Trees)

Στη θεωρία λήψης αποφάσεων, τα δεδομένα ενός προβλήματος μπορούν να αναπαρασταθούν γραφικά [7]. Αυτή η γραφική αναπαράσταση ονομάζεται δέντρο αποφάσεων (decision tree). Το δέντρο αποφάσεων, γνωστό επίσης και ως δίκτυο ροής αποφάσεων ή διάγραμμα αποφάσεων, αποτελεί ένα υποστηρικτικό εργαλείο λήψης αποφάσεων που χρησιμοποιεί μια γραφική απεικόνιση όμοια της μορφής δέντρου, συμπεριλαμβάνοντας όλες τις πιθανές αποφάσεις και όλα τα πιθανά αποτελέσματα.

Τα δέντρα αποφάσεων ως αλγόριθμοι ταξινόμησης μπορούν να προβλέπουν και να κατηγοριοποιούν μελλοντικές καταστάσεις, βασιζόμενα σε ένα σύνολο κανόνων απόφασης (decision rules). Η βασική ιδέα των δέντρων αποφάσεων είναι ο διαχωρισμός των δεδομένων αναδρομικά σε υποσύνολα ώστε κάθε ένα από αυτά να περιέχει ομοειδείς καταστάσεις της μεταβλητής της οποίας η τιμή πρέπει να προβλεφθεί. Σε κάθε σημείο όπου το δέντρο διαχωρίζεται σε κλάδους, εκτιμώνται όλα τα χαρακτηριστικά εισόδου προκειμένου να βρεθεί η επίδρασή τους στην μεταβλητή εξόδου. Έτσι, κάθε μονοπάτι του δέντρου συνιστά και ένα κανόνα απόφασης. Ένα παράδειγμα δέντρου αποφάσεων απεικονίζεται παρακάτω :



**Εικόνα 8 :** Παράδειγμα ενός δέντρου αποφάσεων (αφορά καρκίνο του προστάτη).

Για ένα δεδομένο σύνολο εκπαίδευσης, ο στόχος είναι η κατασκευή του βέλτιστου δέντρου αποφάσεων, ειδικότερα αυτού που επιτυγχάνει μια καλή ισορροπία μεταξύ πολυπλοκότητας και ακρίβειας, δηλαδή μεταξύ του συνολικού αριθμού των κόμβων και της ικανότητας ταξινόμησης. Η διαδικασία κατασκευής στοχεύει στο να επιλεγούν τα κατάλληλα χαρακτηριστικά ανάμεσα στα υποψήφια (ο αριθμός των επιλεγμένων χαρακτηριστικών πρέπει να είναι αρκετά μικρότερος από το συνολικό αριθμό των υποψήφια ιδιοτήτων) και να κατασκευαστεί το δέντρο απόφασης με βάση αυτές τις κατάλληλες ιδιότητες.

Βασικά πλεονεκτήματα της μεθόδου αυτής είναι η ταχύτητα με την οποία κατασκευάζεται το δέντρο και η ευκολία στον τρόπο ερμηνείας τους, καθώς η λογική με την οποία έχει κατασκευαστεί το δέντρο είναι εμφανής στο τελικό διάγραμμα. Ένα άλλο σημαντικό πλεονέκτημα είναι ότι ο αλγόριθμος εκπαίδευση λαμβάνει υπόψη του μόνο εκείνες τις μεταβλητές εισόδου που είναι καθοριστικές για την εξαγωγή ακριβούς διάγνωσης, αγνοώντας τις υπόλοιπες. Η διαδικασία εκπαίδευσης είναι αναδρομική και διαχωρίζει το αρχικό σύνολο σε υποσύνολα, ενώ τερματίζει όταν πλέον έχει κατασκευαστεί όλο το δέντρο απόφασης, κάτι που καθορίζεται από κατάλληλα κριτήρια τερματισμού.

### 2.3.4 Απλός Μπεϋζιανός Ταξινομητής (Naïve Bayesian Classifier)

Ο απλός ταξινομητής Bayes αποτελεί έναν από τους πιο αποδοτικούς επαγωγικούς ταξινομητές μηχανικής μάθησης που στηρίζεται σε στατιστικά στοιχεία (κατανομές πιθανότητας) [8]. Βασίζεται στην υπόθεση ότι η επίδραση της τιμής ενός χαρακτηριστικού σε μια δεδομένη κλάση είναι ανεξάρτητη από τις τιμές των υπόλοιπων χαρακτηριστικών. Αυτή η συνθήκη είναι γνωστή και ως υπό συνθήκη ανεξαρτησία (conditional

independence). Βασιζόμενοι σε αυτό και στην πιθανοτική φύση του μοντέλου, μπορούμε να εκπαιδεύσουμε ταξινομητές πολύ αποδοτικά σε περιβάλλον επιβλεπόμενης μάθησης. Πρακτικά, για κάθε δείγμα ο ταξινομητής θα υπολογίσει την κλάση που έχει την μεγαλύτερη πιθανότητα να ανήκει με βάση τα παραδείγματα που επεξεργάστηκαν στη διαδικασία εκπαίδευσης. Ένα πολύ σημαντικό πλεονέκτημα είναι ότι χρειάζεται λίγα μόνο παραδείγματα για την μάθηση και επειδή έχουμε υποθέσει την ανεξαρτησία των χαρακτηριστικών εισόδου, πρέπει να υπολογίσουμε μόνο τη διασπορά των μεταβλητών και όχι την μήτρα συσχέτισης.

Η στατιστική θεωρία του Bayes αποτελεί μια θεμελιώδη προσέγγιση στο πρόβλημα ταξινόμησης προτύπων, η οποία βασίζεται στην παραδοχή ότι τα προβλήματα αποφάσεων τίθενται σε πιθανοτικούς όρους και ότι όλες οι τιμές των πιθανοτήτων είναι γνωστές.

Η ποσότητα  $P$  που περιγράφει έναν απλό ταξινομητή Bayes για ένα σύνολο παραδειγμάτων εκφράζει την πιθανότητα να είναι  $c$  η τιμή της εξαρτημένης μεταβλητής  $C$  με βάση τις τιμές  $x = (x_1, x_2, \dots, x_n)$  των χαρακτηριστικών  $X = (X_1, X_2, \dots, X_n)$  και δίνεται από τη σχέση :

$$P(c|x) = P(c) * \prod_i P(x_i|c)$$

όπου τα χαρακτηριστικά  $X_i$  θεωρούνται ανεξάρτητα μεταξύ τους.

Ο υπολογισμός της παραπάνω ποσότητας για ένα σύνολο  $N$  παραδειγμάτων γίνεται με βάση τις σχέσεις :

- $P(c) = N(c)/N$
- $P(x_i|c) = N(x_i, c)/N(c)$  , για χαρακτηριστικό  $X_i$  με διακριτές τιμές
- $P(x_i|c) = g(x_i, \mu_c, \sigma_c^2)$  , για χαρακτηριστικό  $X_i$  με αριθμητικές τιμές

όπου  $N(c)$  ο αριθμός των παραδειγμάτων που έχουν στην εξαρτημένη μεταβλητή την τιμή  $c$ ,  $N(x_i, c)$  είναι ο αριθμός των παραδειγμάτων που έχουν στο χαρακτηριστικό  $X_i$  και στην εξαρτημένη μεταβλητή τιμές  $x_i$  και  $c$  αντίστοιχα, και  $g(x_i, \mu_c, \sigma_c^2)$  είναι η συνάρτηση πυκνότητας πιθανότητας Gauss με μέσο όρο  $\mu_c$  και διασπορά  $\sigma_c^2$  για το χαρακτηριστικό  $X_i$ .

### 2.3.5 Μπεϋζιανά Δίκτυα (Bayesian Networks)

Τα Μπεϋζιανά δίκτυα (*Bayesian networks*), που καλούνται επίσης και δίκτυα πίστης (*belief networks*) ή αιτιατά δίκτυα, συνιστούν άκυκλα κατευθυνόμενα γραφήματα που μοντελοποιούν πιθανοτικές επιρροές και σχέσεις μεταξύ μεταβλητών [9,10]. Μόλις περιγραφούν, αυτές οι σχέσεις επιτρέπουν την ανάπτυξη μιας δομής  $n$ -διαστάσεων ή ενός μοντέλου που κωδικοποιεί όλα τα χαρακτηριστικά, συμπεριλαμβανομένων και των αποτελεσμάτων σε ένα ενιαίο ιεραχικό δίκτυο. Το γραφικό μέρος ενός Μπεϋζιανού δικτύου αποτελεί τη δομή του μοντελοποιημένου προβλήματος, ενώ οι τοπικές αλληλεπιδράσεις μεταξύ γειτονικών κατανομών πιθανότητας ποσοτικοποιούνται με τη χρήση πραγματικών δεδομένων του συστήματος. Ειδικότερα, κάθε κόμβος στο γράφημα αντιπροσωπεύει μια



### 2.3.6 Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks)

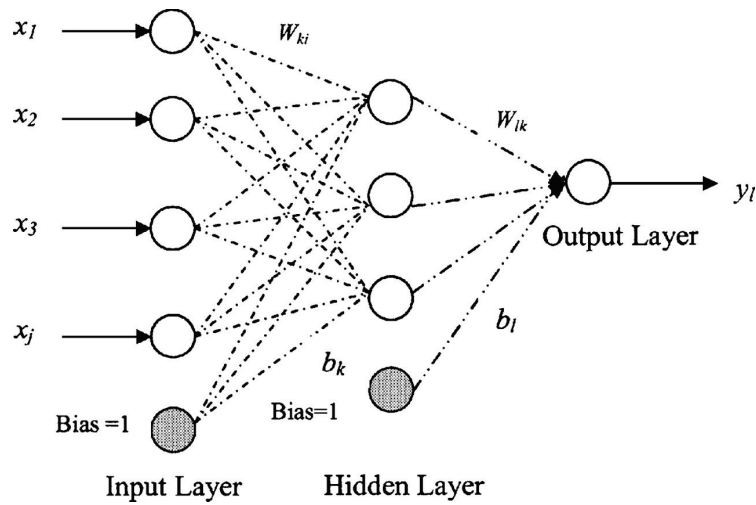
Τα Τεχνητά Νευρωνικά Δίκτυα (*Artificial Neural Networks – ANNs*) [11] αποτελούν πεδίο έρευνας τα τελευταία 40 χρόνια. Τα πρώτα υπολογιστικά, εκπαιδευσιμα δίκτυα αναπτύχθηκαν το 1959 από τον Rosenblatt για την αντιμετώπιση απλών γραμμικών προβλημάτων, ενώ οι πρώτες μη γραμμικές δυνατότητες τους καταγράφηκαν το 1959 από τον Werbos.

Πιο συγκεκριμένα, τα τεχνητά νευρωνικά δίκτυα είναι υπολογιστικές μέθοδοι που εκτελούν πολυπαραγοντική ανάλυση, εμπνευσμένα από τα δίκτυα των βιολογικών νευρώνων. Στη βιολογία, τα νευρικά δίκτυα αποτελούνται από ένα μεγάλο αριθμό νευρωνικών κυτάρων που είναι ευρέως διασυνδεδεμένα μεταξύ τους. Κάθε νευρωνικό κύτταρο μπορεί να παράγει ένα ηλεκτροχημικό σήμα. Μέσα από ένα πολύπλοκο ιστό κλάδων, γνωστών και ως δενδριτών, το νευρωνικό κύτταρο μπορεί να αλληλεπιδράσει με άλλους νευρώνες που μπορεί να βρίσκονται πλησιέστερα ή πιο μακρινά από αυτό. Υπάρχει επίσης ένα δίκτυο με δομή διακλάδωσης εξόδου, οι αποκαλούμενοι νευρίτες, οι οποίοι χρησιμοποιούνται για την εκτέλεση ενός συγκεκριμένου μηνύματος. Οι αλληλεπιδράσεις μεταξύ των νευριτών και των δενδριτών γειτονικών ή απομακρυσμένων νευρώνων, μέσω συνάψεων, βοηθά στην ερμηνεία του σήματος. Επιπλέον, η έλευση ενός σήματος συνεπάγεται μια απάντηση και προσεκτική ρύθμιση της μεταφοράς του από τις συνάψεις, κάτι που πιστεύεται ότι αποτελεί το πρωταρχικό βήμα για την μάθηση.

Ως νευρωνικά δίκτυα, τα τεχνητά νευρωνικά δίκτυα περιέχουν επίσης επίπεδα απλών σημείων (κόμβων) των δεδομένων, που αλληλεπιδρούν μέσω προσεκτικά σταθμισμένων γραμμών σύνδεσης, έτσι ώστε να μπορούν να επεξεργάζονται ένα αποτέλεσμα. Το ισοζύγιο των βαρών των γραμμών αυτών επιτυγχάνεται από μια διαδικασία εκπαίδευσης των δεδομένων εισόδου. Φυσικά, τα σύγχρονα ΤΝΔ έχουν αρκετά απλούστερη αρχιτεκτονική, με πολύ λιγότερους κόμβους ή διασυνδέσεις από το πραγματικό νευρικό σύστημα. Σημειώνουμε ότι υπάρχουν τρεις τύποι νευρώνων – κόμβων :

- a) Οι νευρώνες εισόδου (*input neurons*). Αυτοί δεν επιτελούν κανέναν υπολογισμό, μεσολαβούν απλώς ανάμεσα στις εισόδους του δικτύου και στους υπολογιστικούς νευρώνες.
- b) Οι νευρώνες εξόδου (*output neurons*), οι οποίοι διοχετεύουν στο περιβάλλον τις τελικές αριθμητικές εξόδους του δικτύου.
- c) Οι υπολογιστικοί ή κρυφοί νευρώνες (*hidden neurons*). Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν κάθε είσοδό τους με το αντίστοιχο συναπτικό βάρος και υπολογίζουν το ολικό άθροισμα των γινομένων. Το άθροισμα αυτό τροφοδοτείται ως όρισμα στη συνάρτηση ενεργοποίησης (*activation function*), την οποία υλοποιεί εσωτερικά κάθε κόμβος. Η τιμή που λαμβάνει η συνάρτηση για το εν λόγω όρισμα είναι και η έξοδος του νευρώνα για τις τρέχουσες εισόδους και τα βάρη.

Παρακάτω απεικονίζεται η δομή ενός τεχνητού νευρωνικού δικτύου :



**Εικόνα 11 :** Τεχνητό νευρωνικό δίκτυο.

Αν  $x_{ki}$  είναι η  $i$ -οστή είσοδος του  $k$ -νευρώνα,  $w_{ki}$  το  $i$ -οστό συναπτικό βάρος του  $k$ -νευρώνα και  $\Phi(\cdot)$  η συνάρτηση ενεργοποίησης του νευρωνικού δικτύου, τότε η έξοδος  $y_k$  του  $k$ -νευρώνα δίνεται από την εξίσωση :

$$y_k = \Phi\left(\sum_{i=0}^N x_{ki} * w_{ki}\right)$$

Στον  $k$ -οστό νευρώνα υπάρχει ένα συναπτικό βάρος  $w_{k0}$  με ιδιαίτερη σημασία, το οποίο καλείται πόλωση ή κατώφλι (bias/threshold). Η τιμή της εισόδου είναι πάντα η μονάδα,  $x_{k0} = 1$ . Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός. Για την συνάρτηση ενεργοποίησης μπορεί να χρησιμοποιηθεί ένα πλήθος συναρτήσεων αλλά έχει επικρατήσει η σιγμοειδής συνάρτηση:

$$S(t) = \frac{1}{1+e^{-t}}$$

Σε αυτό το σημείο, προκειμένου να γίνει κατανοητή στην πράξη η λειτουργία ενός τεχνητού νευρωνικού δικτύου, πρέπει να επανεξεταστεί η διαδικασία λήψης αποφάσεων στην Ιατρική και ειδικότερα στην Ογκολογία. Πιο συγκεκριμένα, η λήψη αποφάσεων βασίζεται σε ένα είδος μάθησης και στη συνέχεια παράγει ένα αποτέλεσμα. Η διαδικασία μάθησης της λήψης αποφάσεων αποτελείται από τη συλλογή των δεδομένων που ερμηνεύονται και κατ'επέκταση σταθμίζονται με βάση την ιατρική γνώση και εμπειρία. Με αυτόν τον τρόπο, πραγματοποιείται προσεκτική «εκπαίδευση» για την ανάλυση αποφάσεων, ώστε να παραχθεί ένα αποτέλεσμα.



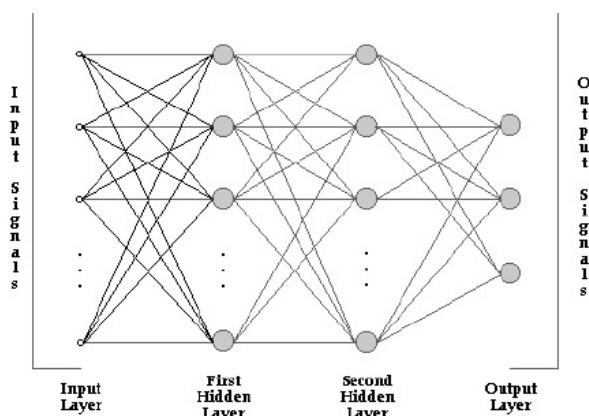
Όταν εξετάζεται έναν νέος ασθενής, συλλέγονται καθαρά δεδομένα (για παράδειγμα χρήσιμες τιμές του ορού του αίματος, ιστολογικές και απεικονιστικές λεπτομέρειες κτλ), προκειμένου να αξιολογηθεί η κατάστασή του, να γίνει μια πρόγνωση και στη συνέχεια να ληφθεί μια απόφαση με βάση την ιατρική γνώση και εμπειρία. Ο θεράπων ιατρός πρέπει να αποφασίσει με βάση την προοπτική του ασθενούς όσον αφορά την εξέλιξη της ασθένειας αλλά και την εξατομικευμένη αναγνώριση προτύπων.

Ομοίως, στα ΤΝΔ υπάρχει μια διαδικασία μάθησης. Είναι σημαντικό να σημειωθεί ότι κατά την εκπαίδευση ενός νευρωνικού δικτύου, πρέπει να χρησιμοποιούνται τρία μη επικαλυπτόμενα σύνολα δεδομένων. Συνήθως, τα δεδομένα από έναν ενιαίο πληθυσμό διαιρούνται τυχαία σε τρία υποσύνολα : το σύνολο εκπαίδευσης, το σύνολο δοκιμών και το σύνολο ελέγχου (επαλήθευσης). Το σύνολο εκπαίδευσης χρησιμοποιείται ώστε τα σήματα εισόδου των δεδομένων να τροφοδοτούνται σε ένα ή περισσότερα επίπεδα νευρώνων μέσω σταθμισμένων αλληλεπιδράσεων μεταξύ αυτών και να προσαρμόζονται τα βάρη τους. Η διαδικασία εκπαίδευσης ολοκληρώνεται από ένα πρόγραμμα εκμάθησης που επιλέγει περιπτώσεις από μια βάση δεδομένων προκειμένου να προσαρμοστούν οι αλληλεπιδράσεις μεταξύ των επιπέδων, ώστε να παραχθεί ένα αποτελέσμα που είναι κοντά στα πραγματικά αποτελέσματα της ίδιας βάσης δεδομένων. Η διαδικασία εκπαίδευσης συνεχίζεται μέχρι να ελαχιστοποιηθεί η πιθανότητα λάθους και τερματίζεται όταν δεν υπάρχει κάποια ένδειξη περαιτέρω μείωσης της πιθανότητας λάθους. Το σύνολο δοκιμών των δεδομένων χρησιμοποιείται ως εκ τούτου για να αποφασιστεί πότε θα σταματήσει η εκπαίδευση. Με αυτό το επίτευγμα του ελάχιστου σφάλματος, το ΤΝΔ αποφεύγει την υπερεκπαίδευση. Ένα παράδειγμα τέτοιου τύπου σφάλματος είναι η αποτυχία γενίκευσης, δηλαδή για παράδειγμα η αποτυχία να αναγνωρίζονται παρόμοια μοτίβα που δεν παρουσιάζονται με τον ίδιο ακριβώς τρόπο. Ένα τρίτο βήμα αποτελεί, τέλος, η διαδικασία επαλήθευσης, για την οποία δεσμεύεται ένα τρίτο σύνολο δεδομένων. Το σύνολο επαλήθευσης ή επικύρωσης είναι ανεξάρτητο και δεν χρησιμοποιείται καθόλου κατά την εκπαίδευση ή τον έλεγχο και επομένως μπορεί να θεωρηθεί μια πραγματική πρόβλεψη της απόδοσης του νευρωνικού δικτύου.

Μια επιτυχημένη εκπαίδευση μπορεί να οδηγήσει στην ερμηνεία των δεδομένων από το ΤΝΔ, με έναν ώριμο και συχνά αυθεντικό τρόπο. Με την υπολογιστική ισχύ που υπάρχει σήμερα, τα ΤΝΔ μπορούν να χειριστούν με ευκολία μη γραμμικά φαινόμενα οποιουδήποτε τύπου, χωρίς να απαιτούνται γραμμικές σχέσεις που αντικατοπτρίζουν απλούς συσχετισμούς. Κατά συνέπεια, ένα ΤΝΔ έχει τη δυνατότητα να συσχετίσει διαφορετικούς παράγοντες πρόβλεψης, να εντοπίσει κρυφές αλληλεπιδράσεις μεταξύ μεταβλητών, να προβλέψει ένα αποτέλεσμα για έναν ασθενή ή ένα σύνολο ασθενών, να ταξινομήσει τους ασθενείς σε ομάδες υψηλού κινδύνου ή να προσεγγίσει μια λειτουργία και να ολοκληρώσει ένα γνωστό μοτίβο.

**Πολυεπίπεδος Αισθητήρας (Multilayer Perceptron, MLP)** [12] : Ένα σύνθετο παράδειγμα τεχνητού νευρωνικού δικτύου όπου οι νευρώνες είναι δομημένοι σε επίπεδα, το οποίο θα συναντήσουμε αρκετά συχνά παρακάτω, είναι ο πολυεπίπεδος αισθητήρας (*multilayer perceptron*). Πιο συγκεκριμένα, το τροποποιημένο μοντέλο του απλού αισθητήρα ονομάζεται πολυεπίπεδος αισθητήρας. Ο απλός αισθητήρας είναι ένα δίκτυο με δύο επίπεδα. Το πρώτο επίπεδο απαρτίζεται από τις εισόδους του δικτύου, δεν έχει

νευρώνες και επομένως δε γίνεται καμία επεξεργασία πληροφορίας σε αυτό, ενώ το δεύτερο επίπεδο αποτελεί το επίπεδο εξόδου του δικτύου. Στον πολυεπίπεδο αισθητήρα, μεταξύ των επιπέδων εισόδου και εξόδου, μεσολαβούν ένα ή και περισσότερα επίπεδα, τα λεγόμενα κρυφά επίπεδα (*hidden layers*). Ένα παράδειγμα τέτοιου δικτύου με δύο κρυφά επίπεδα φαίνεται στην παρακάτω εικόνα :



**Εικόνα 12 :** Πολυεπίπεδος αισθητήρας με δύο κρυφά επίπεδα.

Η ροή της πληροφορίας σε ένα τέτοιο δίκτυο γίνεται πάντα από τα αριστερά προς τα δεξιά, ενώ δεν υπάρχει κανένας βρόχος ανάδρασης. Θεωρούμε επίσης ότι οι νευρώνες σε κάθε επίπεδο αλληλεπιδρούν μόνο με εκείνους τους νευρώνες που ανήκουν στα άμεσα γειτονικά τους επίπεδα. Ειδικότερα, το πρώτο κρυφό επίπεδο δέχεται τις τιμές του επιπέδου εισόδου, τα αποτελέσματα του πρώτου κρυφού επιπέδου περνάνε στο δεύτερο κρυφό, του οποίου τα αποτελέσματα στη συνέχεια περνάνε στο επίπεδο εξόδου. Ένα δίκτυο πολυεπίπεδο αισθητήρα, στο οποίο υπάρχουν όλες οι επιτρεπτές συνδέσεις μεταξύ των νευρώνων, ονομάζεται πλήρως συνδεδεμένο (*fully connected*), αλλιώς ονομάζεται μερικά συνδεδεμένο (*partially connected*).

Η πρόκληση που καλείται να αντιμετωπίσει ο πολυεπίπεδος αισθητήρας κατά την επίλυση μη γραμμικών διαχωρίσιμων προβλημάτων συνίσταται στην εύρεση ενός τρόπου να ενισχυθεί η συμμετοχή των νευρώνων που συμβάλλουν θετικά στην κατεύθυνση του σωστού αποτελέσματος και αντίθετα να αποδυναμωθούν οι δεσμοί εκείνων που συμβάλλουν αρνητικά. Αυτό ακριβώς επιτυγχάνεται με τον αλγόριθμο οπισθοδιάδοσης του σφάλματος (*error back-propagation algorithm*), ο οποίος χρησιμοποιείται πολύ συχνά για την εκπαίδευση του πολυεπίπεδο αισθητήρα.

**Αλγόριθμος οπισθοδιάδοσης του σφάλματος (*error back-propagation algorithm*) [12] :** Έστω ένας πολυεπίπεδος αισθητήρας με ένα ή περισσότερα κρυφά επίπεδα,  $m$  τιμές εισόδου και  $m_o$  νευρώνες στο επίπεδο εξόδου, όπου οι νευρώνες λειτουργούν με κάποια παραγωγίσιμη συνάρτηση ενεργοποίησης. Έστω επίσης ότι, για να εκπαιδύσουμε το δίκτυο αυτό, έχουμε στη διάθεσή μας ένα σύνολο  $N$  γνωστών παραδειγμάτων εισόδων – επιθυμητών εξόδων.

Ο στόχος μας είναι να βρούμε έναν κατάλληλο αλγόριθμο για να εκπαιδύσουμε το δίκτυο αυτό, σύμφωνα με το σύνολο των δεδομένων. Σε κάθε βήμα της εκπαίδευσης, οι

τιμές των συναπτικών βαρών μεταβάλλονται και η διαδικασία τερματίζεται, όταν κρίνουμε ότι το δίκτυο έχει εκπαιδευτεί σε ικανοποιητικό βαθμό.

Η βασική ιδέα του αλγόριθμου οπισθοδιάδοσης είναι να προσαρμόζονται τα συναπτικά βάρη σε κάθε βήμα της επανάληψης, έτσι ώστε η συνάρτηση κόστους να μειώνεται. Στην πραγματικότητα, η προσαρμογή των βαρών γίνεται σύμφωνα με τα σφάλματα που υπολογίζονται σε κάθε βήμα της διαδικασίας, για κάθε δεδομένο παράδειγμα. Ο μέσος όρος της μεταβολής των βαρών σε όλο το σύνολο των σετ δεδομένων που προκύπτει με αυτόν τον τρόπο αποτελεί μία εκτίμηση της μεταβολής που θα προέκυπτε αν ελαχιστοποιούσαμε τη συνάρτηση κόστους όλου του συνόλου.

Κατά την εκπαίδευση, επομένως, ενός πολυεπίπεδου αισθητήρα με τον αλγόριθμο οπισθοδιάδοσης, σε κάθε βήμα της επανάληψης, λαμβάνουν χώρα τα παρακάτω :

#### 1<sup>ο</sup> στάδιο – ευθύ πέρασμα

Τροφοδοτούμε το δίκτυο με το  $n$ -οστό σετ τιμών εισόδου. Οι τιμές αυτές περνάνε στο πρώτο κρυφό επίπεδο, οι νευρώνες του οποίου τις επεξεργάζονται και δίνουν τις τιμές εξόδου. Οι τελευταίες περνάνε στο δεύτερο κρυφό επίπεδο – αν υπάρχει – κ.ο.κ. και τελικά το επεξεργασμένο σήμα φτάνει στο επίπεδο εξόδου, το οποίο μας δίνει τις τελικές εξόδους του δικτύου. Συγκρίνουμε τις τιμές της τελικής εξόδου με το επιθυμητό αποτέλεσμα και υπολογίζουμε το σφάλμα για κάθε μία από αυτές.

#### 2<sup>ο</sup> στάδιο – ανάποδο πέρασμα

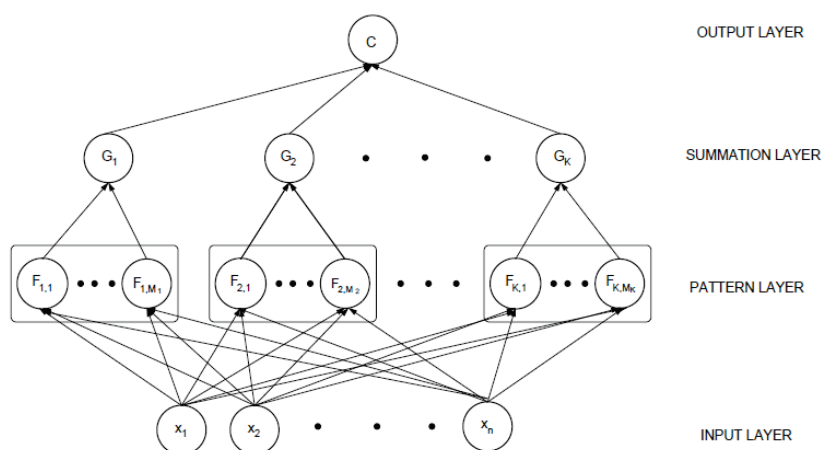
Έχοντας υπολογίσει τα σφάλματα στην έξοδο του δικτύου, υπολογίζουμε τις τοπικές βαθμίδες για το επίπεδο εξόδου. Βάσει των τελευταίων, υπολογίζουμε στη συνέχεια τις τοπικές βαθμίδες του τελευταίου κρυφού επιπέδου, έπειτα αυτές του προτελευταίου κρυφού επιπέδου – αν υπάρχει – κ.ο.κ., μέχρι να υπολογιστούν και αυτές του πρώτου κρυφού επιπέδου. Γνωρίζοντας τις τοπικές βαθμίδες για κάθε νευρώνα του δικτύου, υπολογίζουμε τη διόρθωση σε κάθε συναπτικό βάρος.

Τα δύο αυτά στάδια επαναλαμβάνονται, μέχρι τα αποτελέσματα του δικτύου να συμπίπτουν με τα επιθυμητά με ικανοποιητικό βαθμό ακρίβειας.

**Πιθανοτικά Νευρωνικά Δίκτυα (Probabilistic Neural Networks) [13,14]** : Μια άλλη σημαντική κατηγορία τεχνητών νευρωνικών δικτύων είναι τα πιθανοτικά νευρωνικά δίκτυα (ΠΝΔ), τα οποία προτάθηκαν το 1990 από τον Sprech. Τα ΠΝΔ αποτελούν μια κλάση νευρωνικών δικτύων που συνδυάζει τα καλύτερα χαρακτηριστικά των στατιστικών μεθόδων αναγνώρισης προτύπων και των νευρωνικών δικτύων πρόσθιας τροφοδότησης. Μπορούν να θεωρηθούν ως «έξυπνες μνήμες», καθώς κάθε πρότυπο εκπαίδευσης αποθηκεύεται στο δίκτυο σαν νευρώνας. Έτσι, ο χρόνος εκπαίδευσής τους είναι αισθητά μικρότερος από τον αντίστοιχο χρόνο εκπαίδευσης των υπόλοιπων νευρωνικών δικτύων, ενώ παράγουν αποτελέσματα με εκ των υστέρων πιθανότητες Bayes (Bayes posterior probabilities). Το τίμημα των επιθυμητών αυτών χαρακτηριστικών είναι οι αυξημένες απαιτήσεις σε μνήμη και σε χρόνο εκτέλεσης για την αναγνώριση άγνωστων προτύπων.

Τα ΠΝΔ είναι επιβλεπόμενα (supervised) νευρωνικά δίκτυα και χρησιμοποιούνται ευρέως στα πεδία της αναγνώρισης προτύπων, των μη γραμμικών απεικονίσεων και της εκτίμησης των πιθανοθεωρητικών μεγεθών κατά την ταξινόμηση των αντικειμένων σε κλάσεις. Επίσης, σχετίζονται στενά με τον κανόνα ταξινόμησης Bayes (Bayes classification rule). Ένα σημαντικό χαρακτηριστικό αυτών των δικτύων είναι ο καινοτόμος τρόπος θεώρησης της δομής του δικτύου με όρους συναρτήσεων πυκνότητας πιθανότητας. Η κλασική διαδικασία εκπαίδευσης ενός ΠΝΔ απαιτεί μονάχα ένα πέρασμα των προτύπων εκπαίδευσης, γεγονός που καθιστά τη διαδικασία αυτή πολύ γρήγορη σε σχέση με τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης.

Η δομή ενός ΠΝΔ είναι παρόμοια με αυτή ενός νευρωνικού δικτύου πρόσθιας τροφοδότησης. Όμως, η τοπολογία των ΠΝΔ περιορίζεται σε τέσσερα επίπεδα, το επίπεδο εισόδου (input layer), το επίπεδο προτύπων (pattern layer), το επίπεδο άθροισης (summation layer) και το επίπεδο εξόδου (output layer), όπως απεικονίζεται στην παρακάτω εικόνα :



**Εικόνα 13 :** Ένα πιθανοτικό νευρωνικό δίκτυο.

Ένα διάνυσμα εισόδου  $X$  τροφοδοτείται στους  $n$  νευρώνες του επιπέδου εισόδου και περνά στο επίπεδο προτύπων. Οι νευρώνες του επιπέδου προτύπων είναι χωρισμένοι σε  $K$  ομάδες, μία για κάθε μια από τις κλάσεις στις οποίες θα ταξινομηθούν τα πρότυπα. Το επίπεδο άθροισης του δικτύου υπολογίζει την προσέγγιση των δεσμευμένων συναρτήσεων πυκνότητας πιθανότητας των κλάσεων (conditional class probability functions), συνδυάζοντας τις ήδη υπολογισμένες πυκνότητες πιθανότητας. Στη συνέχεια, το διάνυσμα προτύπου  $X$  ταξινομείται στην κλάση που αντιστοιχεί στο νευρώνα άθροισης με την μεγαλύτερη έξοδο. Σημαντική επίδραση, τέλος, στην απόδοση ενός ΠΝΔ έχει η παράμετρος εξάπλωσης (spread parameter) του δικτύου, καθώς βελτιστοποίηση αυτής της παραμέτρου δύναται να βελτιώσει κατά πολύ την απόδοση του ΠΝΔ.

### 2.3.7 Γενετικοί Αλγόριθμοι (Genetic Algorithms)

Στις αρχές του 1970, ο John Holland, ένας από τους ανθρώπους που πρωτοεισήγαγαν την έννοια του εξελικτικού υπολογισμού, παρουσίασε την έννοια των γενετικών αλγορίθμων [15]. Ο στόχος του ήταν να δείξει στους υπολογιστές πώς να κάνουν ότι έκανε

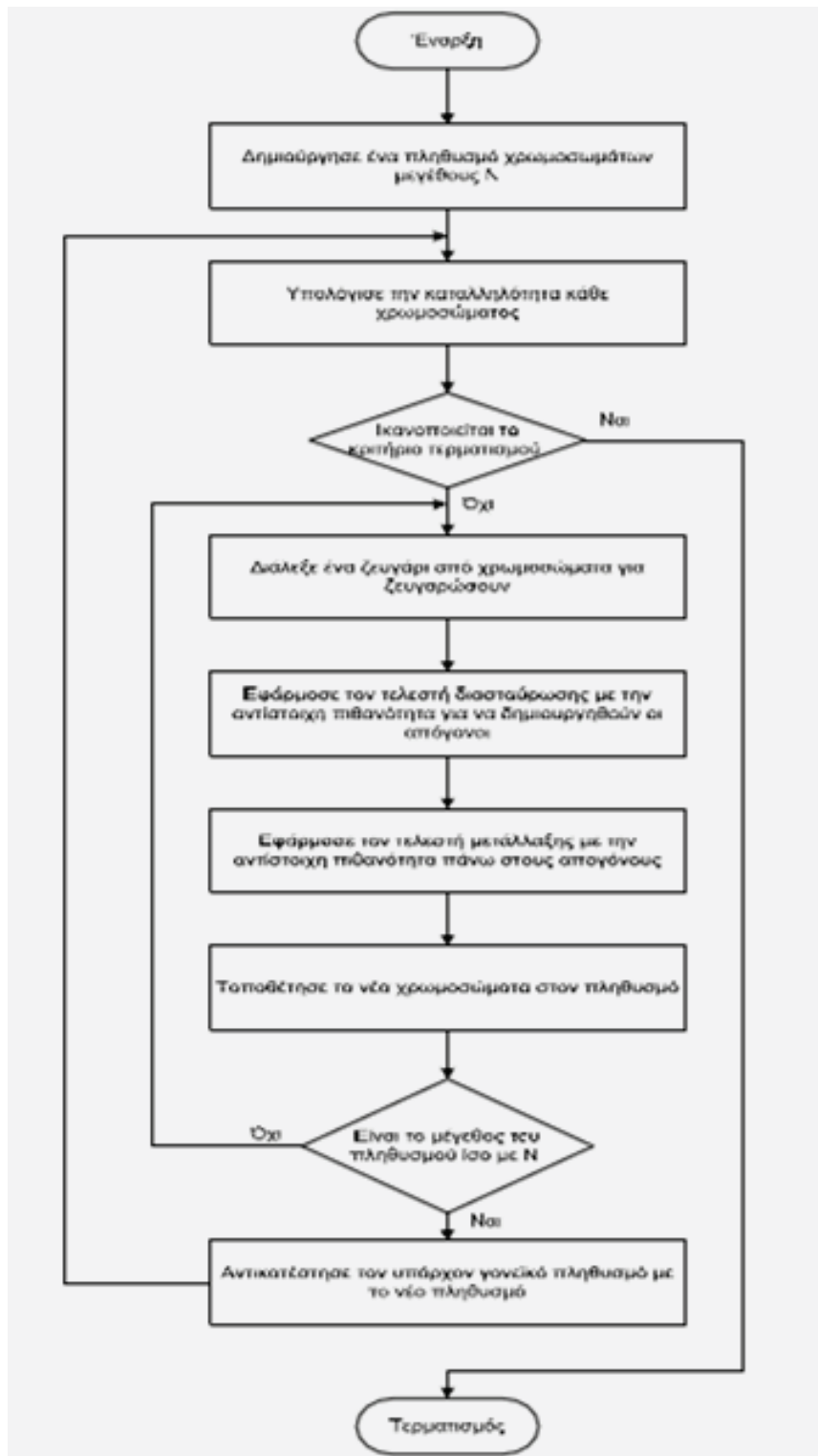
η φύση. Ως μηχανικός υπολογιστών, ο Holland ασχολήθηκε με αλγόριθμους που επεξεργάζονται σειρές δυαδικών ψηφίων. Θεωρούσε αυτούς τους αλγόριθμους ως μια ασαφή μορφή της φυσικής εξέλιξης. Οι γενετικοί αλγόριθμοι του Holland μπορούν να αναπαρασταθούν από μια σειρά βημάτων που οδηγούν από τον ένα πληθυσμό τεχνητών χρωμοσωμάτων σε ένα νέο. Χρησιμοποιεί τη μέθοδο της φυσικής επιλογής και τεχνικές εμπνευσμένες από τη γενετική, γνωστές ως διασταύρωση και μετάλλαξη. Κάθε χρωμόσωμα αποτελείται από έναν αριθμό γονιδίων, και κάθε γονίδιο αναπαρίσταται με 0 ή 1.

Η φύση έχει την ικανότητα να προσαρμόζεται και να μαθαίνει χωρίς υποδείξεις. Με άλλα λόγια, η φύση βρίσκει τα καλύτερα χρωμοσώματα «τυφλά». Οι γενετικοί αλγόριθμοι κάνουν ακριβώς το ίδιο. Δύο μηχανισμοί συνδέουν ένα γενετικό αλγόριθμο με το πρόβλημα που καλείται να λύσει: **η κωδικοποίηση και η αξιολόγηση.**

Στην εργασία του Holland, η κωδικοποίηση επιτελείται με την αναπαράσταση των χρωμοσωμάτων ως αλληλουχίες 0 και 1. Παρόλο που έχουν προταθεί και αρκετές άλλες τεχνικές κωδικοποίησης, καμία δε δουλεύει τέλεια σε όλες τις περιπτώσεις. Οι σειρές ψηφίων είναι η δημοφιλέστερη τεχνική.

Μια συνάρτηση αξιολόγησης  $J$  χρησιμοποιείται για τη μέτρηση της απόδοσης του χρωμοσώματος, ή της καταλληλότητας για το συγκεκριμένο πρόβλημα που καλείται να λύσει (μια συνάρτηση αξιολόγησης σε ένα ΓΑ παίζει τον ίδιο ρόλο που παίζει το περιβάλλον στη φυσική εξέλιξη). Ο ΓΑ χρησιμοποιεί μια μέτρηση της καταλληλότητας των επιμέρους χρωμοσωμάτων που θα επιτελέσουν την αναπαραγωγή. Καθώς η αναπαραγωγή λαμβάνει χώρα, ο τελεστής διασταύρωσης ανταλλάσει τμήματα δύο απλών χρωμοσωμάτων ώστε να δημιουργήσει τον απόγονο, και ο τελεστής μετάλλαξης μεταβάλλει την τιμή ενός τυχαίου γονιδίου του νέου χρωμοσώματος. Σαν αποτέλεσμα, ύστερα από ένα αριθμό επιτυχημένων αναπαραγωγών, τα λιγότερο κατάλληλα χρωμοσώματα εξαφανίζονται, ενώ αυτά που είναι ικανά να επιβιώσουν σταδιακά κυριαρχούν στον πληθυσμό. Μπορεί να είναι μια απλή προσέγγιση, ωστόσο ακόμα και οι λιγότερο περίτεχνοι μηχανισμοί αναπαραγωγής έχουν την ικανότητα να ανταπεξέρχονται σε απαιτητικά προβλήματα.

Σε μια προσπάθεια να κατανοήσουμε τις αρχές λειτουργίας τους, θα λέγαμε ότι οι γενετικοί αλγόριθμοι είναι μια κατηγορία στοχαστικών ευρετικών αλγορίθμων βασισμένων στη βιολογική εξέλιξη. Δεδομένου ενός σαφώς καθορισμένου προβλήματος προς λύση και μιας δυαδικής σειριακής αναπαράστασης για τις υποψήφιες λύσεις (κωδικοποίηση), ένας βασικός γενετικός αλγόριθμος μπορεί να αναπαρασταθεί όπως στο παρακάτω σχήμα :



**Εικόνα 14 :** Διάγραμμα ροής απλού γενετικού αλγορίθμου.

Ένας γενετικός αλγόριθμος ακολουθεί τα εξής βασικά βήματα :

1. Αναπαράστησε την περιοχή μεταβλητών του προβλήματος ως χρωμόσωμα καθορισμένου μεγέθους, επέλεξε το μέγεθος του πληθυσμού των χρωμοσωμάτων  $N$ , την πιθανότητα διασταύρωσης  $p_c$  και την πιθανότητα μετάλλαξης  $p_m$ .

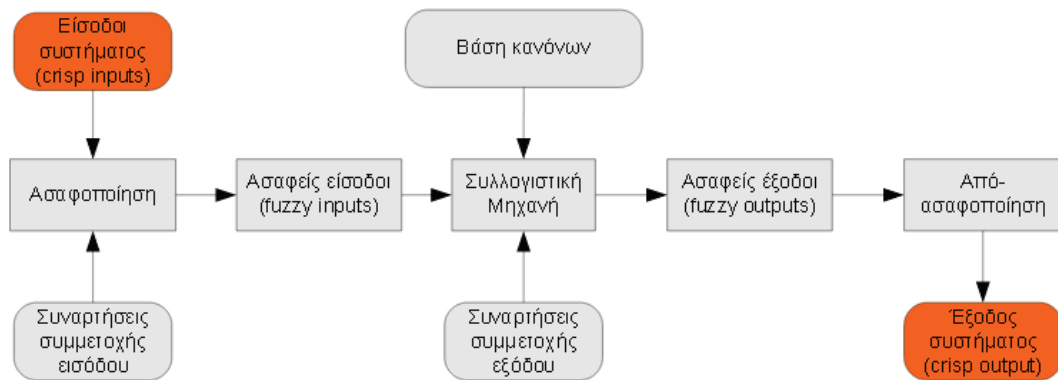
2. Όρισε μια συνάρτηση καταλληλότητας για τη μέτρηση της απόδοσης του κάθε επιμέρους χρωμοσώματος του χώρου του προβλήματος. Η συνάρτηση καταλληλότητας θεμελιώνει τη βάση πάνω στην οποία θα γίνει η επιλογή των χρωμοσωμάτων που θα ταιριάξουν κατά την αναπαραγωγική διαδικασία.
3. Δημιούργησε με τυχαίο τρόπο έναν αρχικό αριθμό χρωμοσωμάτων μεγέθους  $N : x_1, x_2, \dots, x_N$
4. Υπολόγισε την καταλληλότητα κάθε επιμέρους χρωμοσώματος :  $f(x_1), f(x_2), \dots, f(x_N)$
5. Διάλεξε ένα ζευγάρι από χρωμοσώματα για να ζευγαρώσουν από τον υπάρχοντα πληθυσμό. Τα γονεϊκά χρωμοσώματα επιλέγονται με βάση μια πιθανότητα σχετική με την καταλληλότητα τους. Τα κατά πολύ καταλληλότερα χρωμοσώματα έχουν μεγαλύτερη πιθανότητα να επιλεγούν σε σχέση με τα λιγότερο κατάλληλα.
6. Δημιούργησε ένα ζευγάρι από χρωμοσώματα-απογόνους με την εφαρμογή γενετικών τελεστών διασταύρωσης και μετάλλαξης.
7. Τοποθέτησε τα χρωμοσώματα απογόνους στο νέο πληθυσμό.
8. Επανάλαβε το βήμα 5 μέχρι το μέγεθος του νέου πληθυσμού να γίνει ίσο με το μέγεθος του αρχικού πληθυσμού  $N$ .
9. Αντικατέστησε το αρχικό (γονεϊκό) χρωμοσωματικό πληθυσμό με το νέο (απόγονο) πληθυσμό.
10. Πήγαινε στο βήμα 4, και επανάλαβε τη διαδικασία μέχρι να ικανοποιηθεί το κριτήριο τερματισμού.

Όπως βλέπουμε, ένας γενετικός αλγόριθμος αναπαριστά, εκτός των άλλων, και μια επαναληπτική διαδικασία. Κάθε επανάληψη ονομάζεται **γενεά**. Ένας τυπικός αριθμός γενεών για ένα απλό ΓΑ μπορεί να κυμαίνεται από 50 έως πάνω από 500. Το σύνολο των γενεών ονομάζεται **τρέξιμο**. Στο τέλος του αναμένεται να βρούμε ένα ή περισσότερα κατάλληλα χρωμοσώματα, ανάλογα με τις επιθυμίες μας.

### 2.3.8 Ασαφής Λογική (Fuzzy Logic)

Η ασαφής λογική [16] είναι ένα ευρύ επιστημονικό πεδίο που δημιουργήθηκε από την ανάγκη για παράκαμψη της παραδοσιακής δυαδικής λογικής, στην οποία υπάρχουν μόνο οι καταστάσεις του αληθούς ή του ψευδούς. Μέσω της χρήσης ασαφούς λογικής είναι δυνατός ο ορισμός των βαθμών αληθείας, οι οποίοι μετρούν το κατά πόσο κάποιο στοιχείο συμμετέχει σε ένα ασαφές σύνολο. Οι συναρτήσεις συμμετοχής ορίζουν την κατανομή των τιμών του βαθμού αληθείας (κατά πόσο συμμετέχει ένα στοιχείο σε ένα ασαφές σύνολο) που μπορεί να πάρει κάποιο στοιχείο του συνόλου όπου το πεδίο τιμών είναι το  $[0,1]$ . Στην ασαφή λογική, συνήθως τα ασαφή σύνολα χρησιμοποιούν λεκτικές μεταβλητές που υπάρχουν και στην ανθρώπινη γλώσσα. Κατά αυτόν τον τρόπο, επιτυγχάνεται η απευθείας κωδικοποίηση έμπειρης γνώσης σε διάφορες εφαρμογές.

Η γενική αρχιτεκτονική των ασαφών συστημάτων περιλαμβάνει τέσσερις μονάδες, οι οποίες αναφέρονται συνοπτικά παρακάτω. Ένα γενικό σχηματικό διάγραμμα της αρχιτεκτονικής παρατίθεται στο παρακάτω σχήμα:



**Εικόνα 15 :** Σχηματικό διάγραμμα αρχιτεκτονικής ασαφών συστημάτων.

Η βάση ασαφών κανόνων (fuzzy rule base) ή ασαφής βάση γνώσης είναι μια βάση της μορφής IF - THEN.

Η ασαφής συλλογιστική μηχανή (inference machine) ή μηχανισμός εξαγωγής ασαφών συμπερασμάτων αποτελεί τον πυρήνα του ασαφούς συστήματος και περιέχει τη λογική λήψης αποφάσεων. Περιέχει, συνήθως, εκτός από τους ασαφείς (γλωσσικούς) κανόνες και ένα τμήμα βάσης αριθμητικών δεδομένων τα οποία απαιτούνται για τη διαδικασία εξαγωγής συμπερασμάτων.

Μια μονάδα ασαφοποίησης (ασαφοποιητική μονάδα ή fuzzification unit) αντιστοιχεί τα καθορισμένα δεδομένα εισόδου στα ασαφή σύνολα που έχουν οριστεί από το σύστημα. Η μονάδα ασαφοποίησης εκτελεί τις παρακάτω εργασίες :

- Μετράει (παραλαμβάνει) τις (μη ασαφείς) τιμές των εισόδων του συστήματος.
- Απεικονίζει τις περιοχές μεταβολής των τιμών εισόδου σε κατάλληλα υπερσύνολα αναφοράς.
- Ασαφοποιεί τις εισερχόμενες τιμές των εισόδων, δηλαδή τις μετατρέπει σε ασαφή ή γλωσσική μορφή.

Η μονάδα από-ασαφοποίησης (από-ασαφοποιητική μονάδα ή defuzzification unit) αντιστοιχεί τα ασαφή συμπεράσματα ή αποφάσεις που έχουμε ως ασαφή σύνολα σε μια σαφώς καθορισμένη («ντετερμινιστική») μορφή μεταβλητών εξόδου. Απεικονίζει τις περιοχές μεταβολής τιμών εισόδου σε κατάλληλα υπερσύνολα αναφοράς και από-ασαφοποιεί τα αποτελέσματα που δίνει η ασαφής συλλογιστική μηχανή, δηλαδή τα μετατρέπει σε ντετερμινιστική (μη ασαφή) μορφή για περαιτέρω χρήση από επόμενα συστήματα ή διεργασίες απόφασης. Ενδεικτικά αναφέρουμε τις κυριότερες μεθόδους από-ασαφοποίησης : μεγίστου, μέσου όρου των μεγίστων, κέντρου βάρους, ύψους, τροποποιημένη μέθοδος ύψους.

Η ασαφής λογική ενδείκνυται για πολύ πολύπλοκες διαδικασίες, όταν δεν υπάρχει απλό μαθηματικό μοντέλο, σε υψηλής μη-γραμμικότητας προβλήματα και εάν πρόκειται να γίνει επεξεργασία έμπειρης γνώσης (γλωσσικά διατυπωμένης).



## 2.4 Μέτρα απόδοσης

Στην παρούσα ενότητα θα επεξηγηθούν συνοπτικά βασικές έννοιες οι οποίες θα χρησιμοποιηθούν εκτενώς στη συνέχεια, σχετικά με την αξιολόγηση των αποδόσεων των διαφόρων τεχνικών μηχανικής μάθησης και των συστημάτων CDSS. Οι χρησιμοποιούμενες έννοιες ως στατιστικά μέτρα απόδοσης έχουν ως εξής:

### 2.4.1 Ευαισθησία και ειδικότητα

Ευαισθησία (sensitivity ή recall rate ή power): Εκφράζει την ικανότητα ενός test ή ενός συστήματος ταξινόμησης να αναγνωρίζει σωστά τους ασθενείς. Μετρά δηλαδή το ποσοστό των αληθώς θετικών αποτελεσμάτων του test στο σύνολο των πραγματικών θετικών αποτελεσμάτων (το ποσοστό των πραγματικών ασθενών ανθρώπων που σωστά έχουν προσδιοριστεί από το test ως ασθενείς).

Ειδικότητα (specificity): Εκφράζει την ικανότητα του test ή ενός συστήματος ταξινόμησης να αναγνωρίζει σωστά εκείνους που δεν είναι ασθενείς. Μετρά δηλαδή το ποσοστό των αληθώς αρνητικών αποτελεσμάτων του test στο σύνολο των πραγματικών αρνητικών αποτελεσμάτων (το ποσοστό των πραγματικά υγιών ανθρώπων που σωστά έχουν προσδιοριστεί από το test ως υγιείς).

Στη βέλτιστη λοιπόν περίπτωση θα είχε επιτευχθεί 100% ευαισθησία (δηλαδή θα είχαν προβλεφθεί όλοι οι άνθρωποι από την ομάδα των ασθενών ως ασθενείς) και 100% ειδικότητα (δηλαδή θα είχαν προβλεφθεί όλοι οι άνθρωποι από την ομάδα των υγιών ως υγιείς).

Ανάλογα με τη σχέση των αποτελεσμάτων και της πραγματικής κατάστασης του ατόμου μπορεί να ληφθούν τα εξής αποτελέσματα:

- Αληθώς θετικό (true positive-TP): Ασθενής αναγνωρίστηκε σωστά ως ασθενής.
- Ψευδώς θετικό (false positive-FP): Υγιής αναγνωρίστηκε λανθασμένα ως ασθενής.
- Αληθώς αρνητικό (true negative-TN): Υγιής αναγνωρίστηκε σωστά ως υγιής.
- Ψευδώς αρνητικό (false negative-FN): Ασθενής αναγνωρίστηκε λανθασμένα ως υγιής.

Οι τύποι που περιγράφουν μαθηματικά τα δύο αυτά μέτρα απόδοσης είναι:

$$\text{Ευαισθησία} = \frac{\text{αληθώς θετικά}}{\text{άθροισμα όλων όσων έχουν την ασθένεια}}$$

$$\text{Ειδικότητα} = \frac{\text{αληθώς αρνητικά}}{\text{άθροισμα όλων όσων δεν έχουν την ασθένεια}}$$

Προκύπτει λοιπόν το συμπέρασμα ότι ένα θετικό αποτέλεσμα με υψηλή ειδικότητα χρησιμοποιείται για να επιβεβαιώσει την ασθένεια. Αντίστοιχα, ένα αρνητικό αποτέλεσμα με υψηλή ευαισθησία χρησιμοποιείται για να αποκλείσει την ασθένεια.

Μέσω των μέτρων της ειδικότητας και της ευαισθησίας μπορούν να προσδιοριστούν διάφορες πιθανότητες προσδιορισμού απόδοσης όπως οι εξής:

- Πιθανότητα ψευδώς θετικών αποτελεσμάτων False Positive Rate or Fraction = FPF =  $\alpha = FP / (FP + TN) = 1 - \text{specificity}$
- Πιθανότητα ψευδώς αρνητικών αποτελεσμάτων False Negative Rate or Fraction = FNF =  $\beta = FN / (TP + FN) = 1 - \text{sensitivity} = 1 - \text{power}$
- Πιθανότητα αληθώς θετικών αποτελεσμάτων True Positive Rate or Fraction = TPF =  $\text{sensitivity} = \text{power}$
- Πιθανότητα αληθώς αρνητικών αποτελεσμάτων True Negative Rate or Fraction = TNF =  $\text{specificity}$ .

#### 2.4.2 Απόδοση και ακρίβεια

Απόδοση (accuracy): Αναφέρεται στη διαφορά μεταξύ του μέσου όρου μιας σειράς μετρήσεων από μία τιμή η οποία είναι αποδεκτή ως η αληθής (true) ή ορθή (correct) τιμή της μετρούμενης ποσότητας. Αποτελεί με άλλα λόγια το βαθμό του πόσο κοντά στις πραγματικές τιμές μιας ποσότητας είναι οι μετρούμενες τιμές της (βαθμός προσέγγισης πραγματικότητας).

Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\text{Απόδοση} = \frac{\text{αληθώς θετικά} + \text{αληθώς αρνητικά}}{\text{άθροισμα όλων όσων εξετάστηκαν}}$$

Στην περίπτωση που είναι γνωστή η ευαισθησία και η ειδικότητα, η απόδοση μπορεί να υπολογιστεί από τον τύπο:

$$\text{Απόδοση} = (\text{ευαισθησία}) * (\text{επιπολασμός}) + (\text{ειδικότητα}) * (1 - (\text{επιπολασμός}))$$

(όπου με τον όρο επιπολασμός ή prevalence αποδίδεται η συχνότητα εμφάνισης της νόσου στον πληθυσμό).

Ακρίβεια (precision): Το συγκεκριμένο μέτρο εκφράζει την προσέγγιση της συμφωνίας μεταξύ των επαναλαμβανόμενων αποτελεσμάτων της μεθόδου και το βαθμό στον οποίο επαναλαμβανόμενες μετρήσεις (υπό τις ίδιες συνθήκες) δίνουν ίδια αποτελέσματα. Μπορεί να περιγραφεί ως η ποσότητα που μετρά τη διασπορά (dispersion) των αποτελεσμάτων όταν η αναλυτική μεθοδολογία επαναλαμβάνεται σε ένα δείγμα. Η διασπορά των αποτελεσμάτων προκαλείται από διάφορες τυχαίες πηγές και θα βρίσκεται γύρω από την

αναμενόμενη τιμή του αποτελέσματος εάν δεν υπάρχει συστηματικό σφάλμα. Πρόκειται ουσιαστικά για ένα μέτρο της σωστής πραγματοποίησης της δοκιμής, αντανακλώντας την ποιότητα του προσωπικού και των διαδικασιών που εφαρμόζονται.

Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\text{Ακρίβεια} = \frac{\text{αληθώς θετικά}}{\text{αληθώς θετικά} + \text{ψευδώς θετικά}}$$

Υποσύνολα της ακρίβειας θεωρούνται η επαναληψιμότητα (repeatability) και η αναπαραγωγιμότητα (reproducibility), με την πρώτη να αφορά το μέτρο της διασποράς των αποτελεσμάτων διαδοχικών ελέγχων στο ίδιο δείγμα, που εκτελούνται κάτω από τις ίδιες συνθήκες, δηλ. ίδια μέθοδος ελέγχου, ίδιος αναλυτής, ίδια συσκευή, ίδιο εργαστήριο και βραχύ χρονικό διάστημα, και τη δεύτερη να αφορά το μέτρο της διασποράς μεταξύ των αποτελεσμάτων που λαμβάνονται με την ίδια μέθοδο στο ίδιο δείγμα, κάτω από διαφορετικές συνθήκες, δηλ. διαφορετικός αναλυτής, διαφορετικές συσκευές, διαφορετικές παρτίδες αντιδραστηρίων, διαφορετικούς χρόνους. Ανάλογα λοιπόν με τις εκάστοτε συνθήκες του εργαστηριακού ελέγχου, υφίστανται η ενδοεργαστηριακή αναπαραγωγιμότητα ή ενδιάμεση πιστότητα (στο ίδιο εργαστήριο), η διεργαστηριακή αναπαραγωγιμότητα (σε διαφορετικά εργαστήρια), η εντός προσδιορισμού επαναληψιμότητα (ίδιο τμήμα δείγματος) και η μεταξύ προσδιορισμών επαναληψιμότητα (διαφορετικά τμήματα του ίδιου δείγματος).

### 2.4.3 Θετική και αρνητική προγνωστική αξία

Θετική προγνωστική αξία (Positive predictive value - PPV): Εκφράζει την αναλογία των ασθενών που σύμφωνα με το test ή το σύστημα ταξινόμησης είναι θετικοί και έχουν στην πραγματικότητα την ασθένεια. Το συγκεκριμένο μέτρο απαντά στο ερώτημα «αν ένα άτομο διαγνωσθεί θετικό, ποια είναι η πιθανότητα να έχει πραγματικά την ασθένεια;». Δείχνει λοιπόν ουσιαστικά πόσο καλό είναι το διαγνωστικό test στο να επιβεβαιώνει την ασθένεια (θετικό αποτέλεσμα). Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\text{Θετική προγνωστική αξία} = \frac{\text{αληθώς θετικά}}{\text{αληθώς θετικά} + \text{ψευδώς θετικά}}$$

Αρνητική προγνωστική αξία (Negative predictive value - NPV): Εκφράζει την αναλογία των ασθενών που σύμφωνα με το test ή το σύστημα ταξινόμησης είναι αρνητικοί και δεν έχουν στην πραγματικότητα την ασθένεια. Το συγκεκριμένο μέτρο απαντά στο ερώτημα «αν ένα άτομο διαγνωσθεί αρνητικό, ποια η πιθανότητα να μην έχει την ασθένεια;». Δείχνει λοιπόν ουσιαστικά πόσο καλό είναι το διαγνωστικό test στο να απορρίπτει την ασθένεια (θετικό αποτέλεσμα). Η μαθηματική απόδοση του συγκεκριμένου μέτρου περιγράφεται από τη σχέση:

$$\text{Αρνητική προγνωστική αξία} = \frac{\text{αληθώς αρνητικά}}{\text{αληθώς αρνητικά} + \text{ψευδώς αρνητικά}}$$

#### 2.4.4 Καμπύλη ROC

Οι καμπύλες ROC (Receiver Operating Characteristic curves) απεικονίζουν (με μια καμπύλη) τους συνδυασμούς της αναλογίας των ψευδών θετικών περιπτώσεων (1-specificity) (άξονας Χ) και της ευαισθησίας (άξονας Υ) για όλες τις τιμές του ελέγχου που παρατηρούνται στο δείγμα.

Εναλλακτικά μπορεί να χρησιμοποιηθούν οι πιθανότητες false positive fraction-FPF στο Χ-άξονα και true positive fraction-TPF στον Υ-άξονα.

Σημαντική ποσότητα στις καμπύλες ROC αποτελεί η περιοχή κάτω από την καμπύλη (AUC – area under curve), της οποίας το εμβαδόν συνδέεται με την πιθανότητα η τιμή του test για ένα ασθενή να είναι μεγαλύτερη από την τιμή του test για ένα άτομο που δεν έχει την ασθένεια.

Στην περίπτωση ενός τέλει διαγνωστικού test ισχύει AUC=1 και ενός μη πληροφοριακού test ισχύει AUC=0.5. Άλλα χαρακτηριστικά μεγέθη της καμπύλης είναι η μέγιστη κάθετη απόσταση (maximum vertical distance), δηλαδή η απόσταση της ROC curve από τη διαγώνιο γραμμή (όσο μεγαλύτερη είναι η απόσταση αυτή, τόσο πιο πληροφοριακό είναι το διαγνωστικό test) και το σημείο συμμετρίας (symmetry point), στο οποίο ισχύει ότι sensitivity = specificity.

Κλείνοντας, τέλος, σημειώνουμε ότι τα μέτρα απόδοσης των συστημάτων ταξινόμησης υπολογίζονται συνήθως με 2 τεχνικές :

- a) με την τεχνική *holdout validation*, σύμφωνα με την οποία χωρίζουμε τα δεδομένα σε train και test sets, εκπαιδεύουμε το σύστημα με το train set και στη συνέχεια το δοκιμάζουμε στο test set, από το οποίο προκύπτουν τα μέτρα απόδοσης
- b) με την τεχνική *k-fold cross validation* ή διασταυρούμενη επικύρωση k φορών. Στην περίπτωση μας, η πιο συχνά χρησιμοποιούμενη τεχνική είναι η διασταυρούμενη επικύρωση 10 φορών, η οποία λειτουργεί ως εξής : η βάση δεδομένων των ασθενών διαιρείται σε 10 σύνολα περίπου ίσου μεγέθους και ίσες κατανομές υποτροπιάζοντων και μη υποτροπιάζοντων ασθενών. Κάθε ένα από τα 10 τυχαία υποσύνολα δεδομένων χρησιμεύει ως σύνολο δοκιμής για το προγνωστικό μοντέλο που εκπαιδεύεται με τα υπόλοιπα 9 σύνολα. Η συνολική ακρίβεια πρόβλεψης του συστήματος αξιολογείται ως μέσος όρος των 10 δοκιμών.

## 2.5 Αναφορές Κεφαλαίου

- [1] MacKay, D. J. C. (2003) 'Information Theory, Inference, and Learning Algorithms', Cambridge University Press ISBN 0-521-64298-1
- [2] Mitchell, T. (1997) 'Machine Learning', McGraw Hill, ISBN 0-07-042807-7
- [3] Theodoridis S., Koutroumpas K. (2006) 'Pattern recognition', New York, NY:Academic, 3<sup>rd</sup> edition, pp.837, ISBN: 0-12-369531-7
- [4] Altman, N. S. (1992) 'An introduction to kernel and nearest-neighbor nonparametric regression', The American Statistician 46 (3): 175–185, doi:10.1080/00031305.1992.10475879
- [5] V.N. Vapnik (1998) 'Statistical Learning Theory', New York: John Wiley & Sons
- [6] J.C.Burges (1998) 'A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery', Vol.2, pp.121-167
- [7] Κωνσταντίνος Γ. Κλαψινός (2009), «Αλγόριθμοι ταξινόμησης με χρήση ασαφών δένδρων απόφασης», σελ.31-37, Διπλωματική εργασία, τμήμα Η.Μ.Μ.Υ, Α.Π.Θ.
- [8] Harry Zhang, 'The Optimality of Naïve Bayes', Faculty of Computer Science, University of New Brunswick Fredericton, New Brunswick, Canada
- [9] Forsberg JA, Eberhardt J, Boland PJ, Wedin R, Healey JH (2011) 'Estimating Survival in Patients with Operable Skeletal Metastases: An Application of a Bayesian Belief Network', PLoS ONE 6(5): e19956. doi:10.1371/journal.pone.0019956
- [10] R. Marshall Austin, Agnieszka Onisko, Marek J. Druzdzal (2009), 'The Pittsburgh Cervical Cancer Screening Model – A Risk Assessment Tool', page 2
- [11] Theodore Anagnostou, Mesut Remzi, Michael Lykourinas, Bob Djavan (2003), 'Artificial Neural Networks for Decision-Making in Urologic Oncology', pages 1-3, European Urology
- [12] Τσουχνίκα Μαρία, «Νευρωνικά δίκτυα και εφαρμογές», σελ.12-17, 14-23
- [13] A. C. Tan and D. Gilbert (2003) 'An empirical comparison of supervised machine learning techniques in bioinformatics' In Proceedings of the 1<sup>st</sup> Asia Pacific Bioinformatics Conference (APBC 2003), pages 219-222
- [14] Κωνσταντίνος Ε. Παρσόπουλος (2004), «Αλγόριθμοι υπολογιστικής νοημοσύνης για αριθμητική βελτιστοποίηση», σελ.232-233, Διδακτορική διατριβή, Πανεπιστήμιο Πατρών, τμήμα Μαθηματικών
- [15] Eiben, A. E. et al (1994) 'Genetic algorithms with multi-parent recombination', PPSN III: Proceedings of the International Conference on Evolutionary Computation. The Third Conference on Parallel Problem Solving from Nature: 78–87. ISBN 3-540-58484-6
- [16] Novák, V., Perfilieva, I. and Močkoř, J. (1999) 'Mathematical principles of fuzzy logic Dodrecht', Kluwer Academic. ISBN 0-7923-8595-0

# Κεφάλαιο 3 – Εφαρμογή τεχνικών μηχανικής μάθησης και συστημάτων CDSS για την βελτίωση της διάγνωσης, πρόγνωσης και κλινικής διαχείρισης του καρκίνου

## **3.1 Εισαγωγή**

Παρουσιάστηκαν νωρίτερα βασικές έννοιες που αφορούν τα συστήματα υποστήριξης κλινικής απόφασης, όπως είναι η μηχανική μάθηση και η χρήση μεθόδων της στην κατασκευή ευφυών συστημάτων που αξιοποιούνται σε πολλά επιστημονικά πεδία. Διαπιστώνουμε ότι η εφαρμογή της μηχανικής μάθησης στην έρευνα του καρκίνου δεν αποτελεί κάτι καινούριο. Για παράδειγμα, τα τεχνητά νευρωνικά δίκτυα και τα δέντρα αποφάσεων χρησιμοποιούνται στην ανίχνευση και τη διάγνωση του καρκίνου εδώ και περίπου 20 χρόνια. Πλέον οι μέθοδοι μηχανικής μάθησης χρησιμοποιούνται σε ευρύ φάσμα εφαρμογών, που ποικίλλουν από την ανίχνευση και ταξινόμηση όγκων μέσω ακτινών X και CRT εικόνων, μέχρι την κατάταξη των κακοηθειών από πρωτεομικές και γονιδιωματικές αναλύσεις. Σύμφωνα με τα τελευταία στατιστικά στοιχεία της PubMed (βάση δεδομένων όπου αναρτώνται επιστημονικές αναλύσεις), έχουν δημοσιευθεί περισσότερες από 1500 μελέτες πάνω στο θέμα της μηχανικής μάθησης και τον καρκίνο [1]. Ωστόσο, η συντριπτική πλειοψηφία των εργασιών αυτών ασχολείται με τη χρήση μεθόδων μηχανικής μάθησης στην αναγνώριση, ταξινόμηση, ανίχνευση ή χαρακτηρισμό των όγκων και άλλων κακοηθειών από ιατρικές εικόνες. Με άλλα λόγια, η μηχανική μάθηση έχει χρησιμοποιηθεί κυρίως ως βοήθημα στη διάγνωση και ανίχνευση του καρκίνου στην ιατρική απεικονιστική. Είναι σχετικά πρόσφατη η προσπάθεια των ερευνητών του καρκίνου να εφαρμόσουν τη μηχανική μάθηση στην πρόβλεψη και πρόγνωση του καρκίνου. Κατά συνέπεια, ο όγκος της βιβλιογραφίας στον τομέα της μηχανικής μάθησης και πρόγνωσης του καρκίνου είναι σχετικά μικρός, αν και τείνει να αυξάνει με το πέρασμα των χρόνων.

Οι βασικοί στόχοι της πρόβλεψης και πρόγνωσης του καρκίνου διαφοροποιούνται από τους στόχους της ανίχνευσης και διάγνωσης. Στην πρόβλεψη/πρόγνωση του καρκίνου μας ενδιαφέρουν τρία σημεία εστίασης :

1. Η πρόβλεψη της ευαισθησίας στον καρκίνο (πχ εκτίμηση κινδύνου – risk assessment),
2. Η πρόβλεψη της υποτροπής, και
3. Η πρόβλεψη της επιβίωσης (προσδόκιμο ζωής).

Στην πρώτη περίπτωση, προσπαθούμε να προβλέψουμε την πιθανότητα ανάπτυξης ενός τύπου καρκίνου πριν από την εμφάνιση της νόσου. Στη δεύτερη περίπτωση, προσπαθούμε να προβλέψουμε την πιθανότητα επανεμφάνισης του καρκίνου μετά την φαινομενική αντιμετώπιση της ασθένειας. Στην τρίτη περίπτωση, προσπαθούμε να προβλέψουμε ένα αποτέλεσμα (προσδόκιμο ζωής, επιβιωσιμότητα, εξέλιξη, ευαισθησία στη φαρμακευτική αγωγή κατά του όγκου) μετά τη διάγνωση της ασθένειας. Στις δύο τελευταίες περιπτώσεις, η επιτυχία της προγνωστικής πρόβλεψης εξαρτάται προφανώς εν μέρει από την επιτυχία ή την ποιότητα της διάγνωσης. Ωστόσο, η πρόγνωση μιας νόσου μπορεί να επιτευχθεί μόνο μετά από την ιατρική διάγνωση και η προγνωστική πρόβλεψη πρέπει να λάβει υπόψη της πολλά περισσότερα από μια απλή διάγνωση.

Πράγματι, η πρόγνωση του καρκίνου συνήθως περιλαμβάνει πολλούς ιατρούς διαφορετικών ειδικοτήτων που χρησιμοποιούν διαφορετικά υποσύνολα βιοδεικτών, άλλων δεδομένων όπως η ηλικία και η γενική υγεία του ασθενή, της τοποθεσίας και του τύπου του καρκίνου, όπως και του βαθμού και του μεγέθους του όγκου. Συνήθως ο θεράπων ιατρός πρέπει να ενσωματώνει προσεκτικά όλες τις ιστολογικές (που βασίζονται στα κύτταρα), κλινικές (που βασίζονται στον ασθενή) και δημογραφικές (που βασίζονται στον πληθυσμό) πληροφορίες ώστε να καταλήξει σε μια λογική πρόγνωση. Αυτό παραμένει δύσκολο καθήκον, ακόμα και για τον πιο εξειδικευμένο ιατρό. Παρόμοια προβλήματα αντιμετωπίζουν επίσης τόσο οι ιατροί όσο και οι ασθενείς όταν πρόκειται για θέματα πρόληψης του καρκίνου και πρόβλεψη της ευαισθησίας στον καρκίνο. Το οικογενειακό ιστορικό, η ηλικία, η διατροφή, το βάρος (παχυσαρκία), συνήθειες υψηλού κινδύνου (κάπνισμα, μεγάλη κατανάλωση αλκοόλ) και η έκθεση σε περιβαλλοντικές καρκινογόνες ουσίες (ακτινοβολία UV, ραδόνιο, αμίαντος, PCBs) παίζουν σημαντικό ρόλο στην πρόβλεψη του κινδύνου που διατρέχει κάποιος να αναπτύξει καρκίνο. Δυστυχώς, αυτές οι συμβατικές κλινικές, περιβαλλοντικές και συμπεριφορικές παράμετροι δεν παρέχουν γενικά αρκετές πληροφορίες ώστε να γίνει ισχυρή πρόβλεψη ή πρόγνωση. Ιδανικά, αυτό που χρειάζεται είναι μερικές πολύ συγκεκριμένες μοριακές πληροφορίες είτε σχετικά με τον όγκο, ή σχετικά με το γενετικό υλικό του ασθενή.

Με την ταχεία ανάπτυξη των γονιδιωματικών (DNA, αλληλουχίας, μικροσυστοιχιών), πρωτεομικών (τσιπ πρωτεϊνών, συστοιχίες ιστών, ανοσο-ιστολογία) και απεικονιστικών (fMRI, PET, micro-CT) τεχνολογιών, μπορεί να αποκτηθεί εύκολα αυτού του είδους η μοριακή κλίμακας πληροφορία για τους ασθενείς ή για τους όγκους. Έχει αποδειχθεί ότι οι μοριακοί βιοδείκτες, όπως οι σωματικές μεταλλάξεις σε ορισμένα γονίδια, η εμφάνιση ή η έκφραση συγκεκριμένων πρωτεϊνών του όγκου και το χημικό περιβάλλον του όγκου (ανοξικό, υποξικό) αποτελούν πολύ ισχυρούς προγνωστικούς δείκτες. Πιο πρόσφατα αποδείχθηκε ότι οι συνδυασμοί ή τα μοτίβα πολλαπλών μοριακών βιοδεικτών συμβάλλουν περισσότερο στην πρόβλεψη σε σχέση με τεστ ή ενδείξεις ενός μεμονωμένου συστατικού. Αν αυτά τα μοριακά μοτίβα συνδυαστούν με κλινικά δεδομένα μακρο-κλίμακας (τύπος του όγκου, κληρονομικοί παράγοντες, παράγοντες κινδύνου), η αξιοπιστία και η ακρίβεια της πρόγνωσης και πρόβλεψης του καρκίνου βελτιώνεται ακόμη περισσότερο. Ωστόσο, όσο ο αριθμός των παραμέτρων που προσμετράμε αυξάνει, το ίδιο συμβαίνει και με την πρόκληση του να κατανοήσουμε όλη αυτήν την πληροφορία.

Στο παρελθόν, η εξάρτησή μας από πληροφορίες μακρο-κλίμακας (όγκος, αθηνής, πληθυσμός και περιβαλλοντικά δεδομένα) διατηρούσε γενικά τον αριθμό των μεταβλητών αρκετά μικρό, ώστε οι τυπικές στατιστικές μέθοδοι ή ακόμα και η διαίσθηση του ιατρού να μπορούσε να χρησιμοποιηθεί για την πρόβλεψη του κινδύνου ανάπτυξης καρκίνου και των αποτελεσμάτων. Ωστόσο, με τις σημερινές υψηλής απόδοσης διαγνωστικές και απεικονιστικές τεχνολογίες κατακλυζόμαστε από δεκάδες ή και εκατοντάδες μοριακές, κυτταρικές και κλινικές παραμέτρους. Σε αυτές τις περιπτώσεις, η ανθρώπινη διαίσθηση και οι τυποποιημένες στατιστικές γενικά δε λειτουργούν. Αντ' αυτού, πρέπει συνεχώς να καταφεύγουμε σε μη παραδοσιακές, εντατικά υπολογιστικές προσεγγίσεις όπως η μηχανική μάθηση. Η χρήση των υπολογιστών (και της μηχανικής μάθησης) στη διάγνωση, πρόβλεψη και πρόγνωση ασθενειών αποτελεί δείγμα μιας αυξανόμενης τάσης προς την εξατομικευμένη, προληπτική ιατρική. Αυτή η τάση προς την προγνωστική ιατρική είναι σημαντική, όχι μόνο για τους ασθενείς (όσον αφορά τον τρόπο ζωής και τις αποφάσεις ποιότητας ζωής), αλλά και για τους ιατρούς (στη λήψη αποφάσεων για τη θεραπεία), για τους οικονομολόγους της υγείας και τους σχεδιαστές στρατηγικών (στην υλοποίηση μεγάλης κλίμακας στρατηγικών πρόληψης ή θεραπείας του καρκίνου).

Στη συνέχεια, ακολουθεί παρουσίαση από δημοσιευμένες μελέτες που χρησιμοποιούν μεθόδους μηχανικής μάθησης στην μορφή ολοκληρωμένων και αυτόνομων συστημάτων υποστήριξης κλινικής απόφασης για την βελτίωση της διάγνωσης, πρόγνωσης και κλινικής διαχείρισης του καρκίνου. Επιχειρούμε να εντοπίσουμε και να καταγράψουμε τους αλγορίθμους μηχανικής μάθησης που χρησιμοποιούνται (τους οποίους έχουμε αναλύσει νωρίτερα), τα είδη των δεδομένων εκπαίδευσης που εντάσσονται, τα είδη των καταληκτικών προβλέψεων που πραγματοποιούνται, και τη συνολική απόδοση των παρουσιαζόμενων συστημάτων.

### **3.2 Μελέτες σχετικά με τον καρκίνο του μαστού (Breast Cancer)**

Στην κατηγορία αυτή έχουμε συμπεριλάβει τέσσερις μελέτες. Σκοπός της πρώτης μελέτης είναι η διάγνωση του καρκίνου του μαστού, η δεύτερη και η τέταρτη στοχεύουν στην πρόβλεψη της υποτροπής, ενώ η τρίτη μελέτη έχει ως στόχο την επιλογή της κατάλληλης θεραπείας. Ακολουθούν οι μελέτες και ο τρόπος ανάπτυξής τους.

- 1. Πιθανοτικό υπολογιστικό μοντέλο που έχει αναπτυχθεί από κλινικά δεδομένα της Εθνικής Βάσης Δεδομένων Μαστογραφίας για την ταξινόμηση μαστογραφικών ευρημάτων [2]**

Στόχος της μελέτης αυτής είναι να προσδιορίσει κατά πόσον ένα δίκτυο Bayes που έχει εκπαιδευτεί σε μια μεγάλη βάση δεδομένων, η οποία περιλαμβάνει παράγοντες κινδύνου (από τα δημογραφικά στοιχεία των ασθενών) και ευρήματα που έχουν εντοπιστεί από ακτινολόγους σε διαδοχικές κλινικές εξετάσεις μαστογραφίας, δύναται να ξεπεράσει την απόδοση των ακτινολόγων στην ταξινόμηση των μαστογραφικών ευρημάτων ως καλοηθών ή κακοηθών.



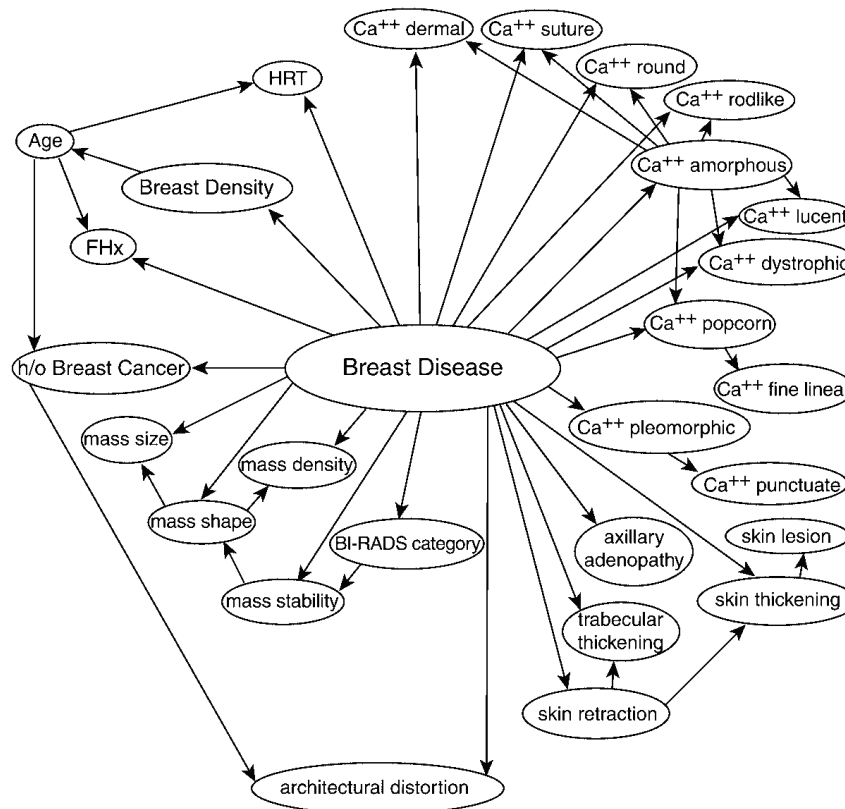
Αρχικά, συγκεντρώθηκαν 48744 συναπτές ομαδοποιημένες εξετάσεις μαστογραφίας, διαγνωστικές και προσυμπτωματικού ελέγχου, από 18269 ασθενείς, από τον Απρίλιο του 1999 έως το Φεβρουάριο του 2004. Στη συνέχεια, τα ευρήματα των μαστογραφιών αντιστοιχήθηκαν σε ένα κρατικό μητρώο καρκίνου, το οποίο λειτούργησε ως πρότυπο αναφοράς. Το δίκτυο Bayes εκπαιδεύτηκε και ελέγχθηκε, με χρήση διασταυρούμενης επικύρωσης 10 φορές, προκειμένου να εκτιμηθεί ο κίνδυνος εμφάνισης καρκίνου του μαστού χρησιμοποιώντας ως δεδομένα/μεταβλητές δημογραφικούς παράγοντες κινδύνου (ηλικία, οικογενειακό και προσωπικό ιστορικό καρκίνου του μαστού και χρήση θεραπείας ορμονικής υποκατάστασης) και ήδη υπάρχοντα μαστογραφικά ευρήματα που έχουν καταγραφεί στο αρχείο υποβολής εκθέσεων απεικόνισης του μαστού και δεδομένων συστήματος (Breast Imaging Reporting and Data System lexicon). Οι μεταβλητές που χρησιμοποιήθηκαν, καθώς και οι τιμές που αντιστοιχούν σε αυτές, παραθέτονται αναλυτικά στον παρακάτω πίνακα :

Variable	Instances
Age (y)	<45, 45–50, 51–54, 55–60, 61–64, >65
Hormone therapy*	None, <5 y, >5 y
Personal history of breast cancer	No, yes
Family history of breast cancer†	None, minor, major
Breast density class‡	1, 2, 3, 4
Mass shape	Circumscribed, ill defined, microlobulated, spiculated, cannot discern
Mass stability	Decreasing, stable, increasing, cannot discern
Mass margins	Oval, round, lobular, irregular, cannot discern
Mass density	Fat, low, equal, high, cannot discern
Mass size	None, small (<3 cm), large (≥3 cm)
Lymph node	Present, not present
Asymmetric density	Present, not present
Skin thickening	Present, not present
Tubular density	Present, not present
Skin retraction	Present, not present
Nipple retraction	Present, not present
Skin thickening	Present, not present
Trabecular thickening	Present, not present
Skin lesion	Present, not present
Axillary adenopathy	Present, not present
Architectural distortion	Present, not present
Calcification(s)	
Popcorn	Present, not present
Milk	Present, not present
Rodlike	Present, not present
Eggshell	Present, not present
Dystrophic	Present, not present
Lucent	Present, not present
Dermal	Present, not present
Round	Scattered, regional, clustered, segmental, linear ductal
Punctate	Scattered, regional, clustered, segmental, linear ductal
Amorphous	Scattered, regional, clustered, segmental, linear ductal
Pleomorphic	Scattered, regional, clustered, segmental, linear ductal
Fine linear	Scattered, regional, clustered, segmental, linear ductal
BI-RADS category	0, 1, 2, 3, 4, 5

**Πίνακας 2 :** Οι μεταβλητές που χρησιμοποιήθηκαν στο δίκτυο Bayes και οι αντίστοιχες τιμές τους.

Πιο αναλυτικά, στο δίκτυο Bayes που χρησιμοποιήθηκε, οι μεταβλητές αναπαρίστανται ως «κόμβοι», οι οποίοι περιλαμβάνουν έναν αριθμό από πιθανές καταστάσεις και αποθηκεύουν πιθανότητες που συνδέονται με κάθε κατάσταση, ενώ τα κατευθυνόμενα βέλη υποδηλώνουν τις σχέσεις εξάρτησης μεταξύ των μεταβλητών. Η

διαδικασία εκπαίδευσης που ακολουθήθηκε συνίστατο στον καθορισμό των πιθανοτήτων μέσα σε κάθε κόμβο, όπως και στη σύνδεση των κόμβων μεταξύ τους. Μετά την εκπαίδευσή του, το δίκτυο Bayes υπολόγισε την εκ των υστέρων πιθανότητα (post test probability – πιθανότητα μετά από τη διεξαγωγή τεστ) κακοήθειας για κάθε μαστογραφικό εύρημα με χρήση της δομής και των πιθανοτήτων που αντλήθηκαν από τα δεδομένα. Σημειώνουμε ότι το δίκτυο Bayes εκπαιδεύτηκε με χρήση ενός αλγορίθμου που λέγεται *αφελής ταξινομητής Bayes επαυξημένος με δέντρο (tree augmented naive Bayes)*, ενώ για τη δοκιμή/αξιολόγησή του χρησιμοποιήθηκε, όπως αναφέρθηκε νωρίτερα, μια συνήθης τεχνική μηχανικής μάθησης που λέγεται διασταυρούμενη επικύρωση 10 φορές.



**Εικόνα 16 :** Δομή του δικτύου Bayes που εκπαιδεύτηκε. Οι κύκλοι με τις ετικέτες αναπαριστούν τους κόμβους και τα βέλη (τόξα) αναπαριστούν τις σχέσεις εξάρτησης.

$Ca^{++}$  = αποτιτανώσεις, FHx = οικογενειακό ιστορικό καρκίνου του μαστού, h/o = ιστορικό, HRT = θεραπεία ορμονικής υποκατάστασης.

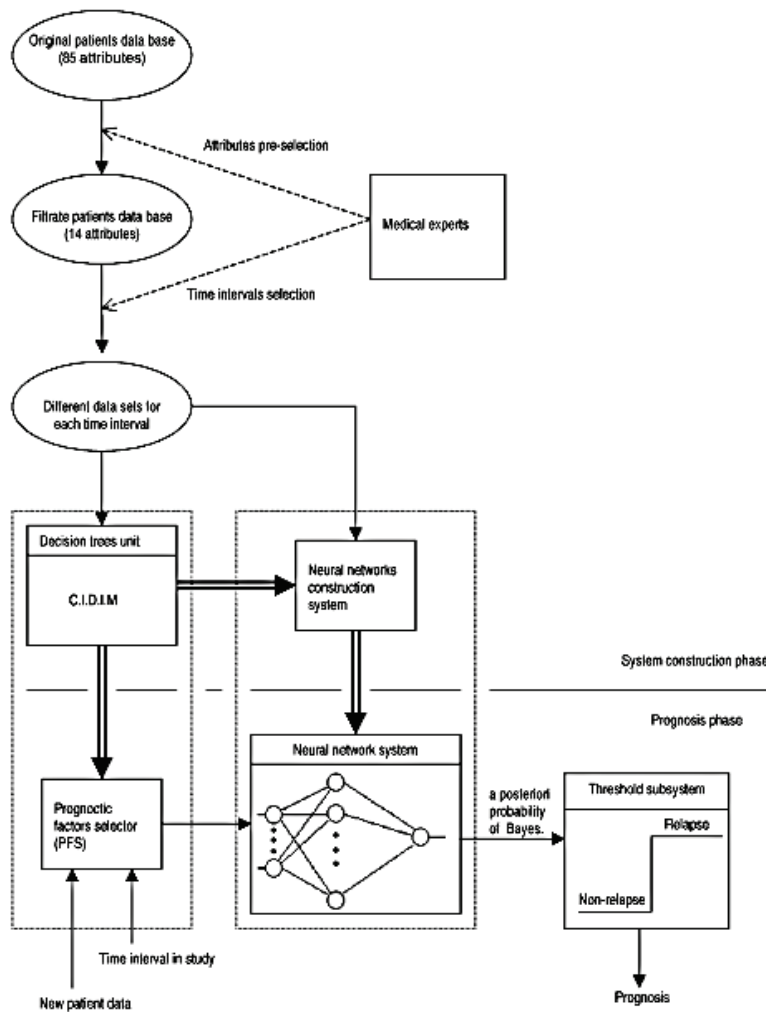
Η σύγκριση της απόδοσης των συνολικά οκτώ ακτινολόγων που κλήθηκαν να ερμηνεύσουν τα ευρήματα σε σχέση με το δίκτυο Bayes αξιολογήθηκε με χρήση παραμέτρων όπως η περιοχή AUC κάτω από την καμπύλη ROC, η ευαισθησία (sensitivity) και η ειδικότητα (specificity). Το δίκτυο Bayes υπερέβη σημαντικά την απόδοση των ακτινολόγων από την άποψη της AUC (0,960 έναντι 0,939,  $P = 0,002$ ), ευαισθησίας (90,0% έναντι 85,3%,  $P < 0,001$ ) και ειδικότητας (93,0% έναντι 88,1%,  $P < 0,001$ ).

Οι συγγραφείς καταλήγουν στο ότι, βάσει των μελλοντικά συγκεντρωμένων μεταβλητών, το υπό αξιολόγηση δίκτυο Bayes μπορεί να προβλέψει την πιθανότητα εμφάνισης καρκίνου του μαστού και να υπολογίσει με ακρίβεια τον κίνδυνο κακοήθειας, ενώ υπερβαίνει την απόδοση των ακτινολόγων που ερμηνεύουν τα μαστογραφικά ευρήματα. Κατά συνέπεια, τα δίκτυα Bayes μπορούν να βοηθήσουν τους ακτινολόγους να βελτιώσουν την ερμηνεία των ευρημάτων της μαστογραφίας και να διευκολύνουν την ταξινόμησή τους σε καλοήθεις και κακοήθεις παθήσεις του μαστού.

## **2. Ένα συνδυαστικό μοντέλο νευρωνικού δικτύου και δέντρων αποφάσεων για την πρόγνωση υποτροπής καρκίνου του μαστού [3]**

Η πρόβλεψη της κλινικής έκβασης ασθενών μετά από χειρουργική επέμβαση για καρκίνο του μαστού παίζει σημαντικό ρόλο σε ιατρικά καθήκοντα όπως η διάγνωση και ο σχεδιασμός της θεραπείας. Διαφορετικοί προγνωστικοί παράγοντες για την έκβαση του καρκίνου του μαστού δείχνουν να είναι σημαντικοί δείκτες πρόβλεψης για τη συνολική επιβίωση, αλλά πιθανώς αποτελούν μέρος μιας ευρύτερης εικόνας που περιλαμβάνει πολλούς παράγοντες. Οι εκτιμήσεις επιβίωσης συνήθως υλοποιούνται από τους κλινικούς ιατρούς με χρήση στατιστικών τεχνικών ανάλυσης επιβίωσης. Υπό αυτή την έννοια, τα τεχνητά νευρωνικά δίκτυα φαίνεται να είναι ένα ισχυρό εργαλείο για την ανάλυση βάσεων δεδομένων όπου υπάρχουν πολύπλοκες μη γραμμικές αλληλεπιδράσεις ανάμεσα στα δεδομένα εισόδου και στις εκτιμήσεις που πρέπει να γίνουν.

Στην εργασία αυτή, οι συγγραφείς παρουσιάζουν ένα εργαλείο υποστήριξης αποφάσεων για την πρόγνωση υποτροπής του καρκίνου του μαστού που συνδυάζει ένα νέο αλγόριθμο *TDIDT* (που σημαίνει *δέντρο απόφασης επαγωγής από πάνω προς τα κάτω*, ενώ γίνεται έλεγχος της επαγωγής με την μέθοδο διαίρεσης του δείγματος, *control of induction by sample division method – CIDIM*), ώστε να επιλεγούν οι πιο σχετικοί προγνωστικοί παράγοντες για την ακριβή πρόγνωση του καρκίνου του μαστού, με ένα σύστημα που αποτελείται από διαφορετικές τοπολογίες νευρωνικών δικτύων που δέχονται ως είσοδο τις επιλεγμένες μεταβλητές προκειμένου να επιτευχθεί καλή πιθανότητα σωστής ταξινόμησης. Επιπλέον, προτείνεται μια νέα μέθοδος για την εκτίμηση του βέλτιστου σφάλματος Bayes χρησιμοποιώντας το παράδειγμα του νευρωνικού μοντέλου, καθώς επίσης και μια νέα μεθοδολογία για την επεξεργασία των «λογοκρινμένων» δεδομένων όταν ο χρόνος παρακολούθησης των ασθενών διακριτοποιείται σε διαφορετικά χρονικά διαστήματα. Μια αναπαράσταση του τρόπου λειτουργίας του περιγραφόμενου συστήματος απεικονίζεται παρακάτω :



**Εικόνα 17 :** Το συνδυαστικό προγνωστικό σύστημα που ανέπτυξαν οι συγγραφείς.

Αρχικά, συλλέχθηκαν και καταγράφηκαν κλινικο-παθολογικά δεδομένα 1035 ασθενών με καρκίνο του μαστού από την Ιατρική Ογκολογική Υπηρεσία του Πανεπιστημιακού Νοσοκομείου της Μάλαγας στην Ισπανία στο διάστημα 1990-2000. Οι συγγραφείς επέλεξαν να απορρίψουν τις περιπτώσεις ασθενών που περιείχαν ελλιπείς τιμές δεδομένων για κάθε διάστημα παρακολούθησης, μέσω ενός κανόνα ταξινόμησης ανάλογα με την κατάσταση επιβίωσης και το χρονικό διάστημα για κάθε ασθενή.

Στη συνέχεια, από τα παραπάνω δεδομένα, προεπιλέχθηκαν 14 ανεξάρτητες μεταβλητές εισόδου για τον αλγόριθμο CIDIM ως πιθανοί παράγοντες κινδύνου. Ως αλγόριθμος που συμπεριλαμβάνεται στα δέντρα αποφάσεων, ο CIDIM εκτελεί προσαρμοστικό «κλάδεμα» με έλεγχο πρόβλεψης, μειώνοντας σημαντικά τον αριθμό των κανόνων και βελτιώνοντας την επιλογή των πιο σημαντικών χαρακτηριστικών που θα εισαχθούν ως παράμετροι εισόδου στο σύστημα του νευρωνικού δικτύου. Σημειώνουμε ότι με τη χρήση του CIDIM παράγονται μικρότερα δέντρα από αυτά που παράγουν άλλοι αλγόριθμοι.

Όσον αφορά το σύστημα των νευρωνικών δικτύων, αυτό υπολογίζει ένα σύνολο χαρακτηριστικών από τον επιλογέα προγνωστικών παραγόντων παράγοντας μια τιμή που αντιστοιχεί στην εκ των υστέρων πιθανότητα υποτροπής για τον εξεταζόμενο ασθενή. Η βάση δεδομένων διαιρείται σε δύο ανεξάρτητα σύνολα, το σύνολο εκπαίδευσης (80% της

βάσης δεδομένων) που χρησιμοποιείται για την εκπαίδευση του νευρωνικού δικτύου και το σύνολο δοκιμών (20% της βάσης δεδομένων) που επικυρώνει την προβλεπτική του ικανότητα. Για την αξιολόγηση του προτεινόμενου μοντέλου, χρησιμοποιήθηκε μια συνήθης τακτική διασταυρούμενης επικύρωσης 10 φορές.

Τα αληθή και ψευδή θετικά και αρνητικά, η ευαισθησία, η ειδικότητα και οι αρνητικές προγνωστικές αξίες υπολογίστηκαν για κάθε πρόβλεψη και αναγνωρίστηκε το σημείο της καμπύλης ROC που ελαχιστοποιεί το συνολικό σφάλμα για κάθε χρονικό διάστημα της παρακολούθησης των ασθενών. Τα αποτελέσματα παρουσιάζονται συγκεντρωτικά στον παρακάτω πίνακα :

Time interval	False negative	False positive	Positive predictive value	Negative predictive value	Sensitivity	Specificity
$I_1$ (0-10)	12.4	21.4	0.722	0.983	0.818	0.972
$I_2$ (10-20)	9.8	77.2	0.772	0.984	0.887	0.965
$I_3$ (20-30)	14.3	26.5	0.665	0.976	0.787	0.957
$I_4$ (30-40)	15.1	22.7	0.721	0.970	0.796	0.957
$I_5$ (40-50)	8.6	25.7	0.648	0.980	0.846	0.945
$I_6$ (50-60)	6.1	16.3	0.699	0.985	0.861	0.961
$I_7$ (>60)	5.9	15.0	0.674	0.986	0.840	0.965

**Πίνακας 3 :** Αποτελέσματα δοκιμών της ανάλυσης ROC για το προγνωστικό μοντέλο που βασίζεται στα νευρωνικά δίκτυα και τα δέντρα αποφάσεων.

Οι συγγραφείς υποστηρίζουν ότι τα αποτελέσματα δείχνουν ότι το προτεινόμενο σύστημα είναι ένα χρήσιμο εργαλείο που μπορεί να χρησιμοποιηθεί από τους ιατρούς όταν χρειάζεται να ψάχνουν μεγάλες βάσεις δεδομένων αναζητώντας ανεπαίσθητα μοτίβα σε προγνωστικούς παράγοντες, ενώ παράλληλα μπορεί να βοηθήσει την επιλογή κατάλληλων συμπληρωματικών θεραπειών για ένα συγκεκριμένο ασθενή. Μάλιστα, το σύστημα μπορεί να αποδειχθεί ιδιαίτερα χρήσιμο όταν τα δεδομένα παρουσιάζουν ένα σημαντικό αριθμό χαρακτηριστικών με ελλειπείς τιμές, όταν απαιτείται η ακρίβεια πρόβλεψης αλλά και πρόσθετη γνώση για τους σημαντικότερους προγνωστικούς παράγοντες για κάθε χρονικό διάστημα και όταν η σημασία των προγνωστικών παραγόντων δεν είναι η ίδια με την πάροδο του χρόνου παρακολούθησης των ασθενών.

Οι ίδιοι συγγραφείς είχαν επιχειρήσει και σε προηγούμενη μελέτη τους να αναπτύξουν ένα μοντέλο νευρωνικού δικτύου για την πρόγνωση του πρώιμου καρκίνου του μαστού [4]. Πιο συγκεκριμένα, προσπάθησαν να θέσουν το μέγιστο όριο ακρίβειας πρόβλεψης με βάση την προσέγγιση της εκ των υστέρων πιθανότητας Bayes μέσω νευρωνικών δικτύων εμπρόσθιας τροφοδότησης (*feed-forward*) τριών επιπέδων. Το αποτέλεσμα αυτό εφαρμόστηκε σε διαφορετικά χρονικά διαστήματα παρακολούθησης ασθενών προκειμένου να επιτευχθεί η καλύτερη ακρίβεια πρόβλεψης για τη σωστή πιθανότητα ταξινόμησης της υποτροπής των ασθενών μετά από χειρουργική επέμβαση. Και σε αυτή την περίπτωση, χρησιμοποιήθηκαν τα κλινικο-παθολογικά δεδομένα 1035 ασθενών από την Ιατρική Ογκολογική Υπηρεσία του Πανεπιστημιακού Νοσοκομείου της Μάλαγας στην Ισπανία στο διάστημα 1990-2000. Ερευνήθηκαν διαφορετικές τοπολογίες δικτύων και παράμετροι μάθησης για την καλύτερη ακρίβεια πρόβλεψης, ενώ τα πραγματικά αποτελέσματα έδειξαν ότι, μετά τη διαδικασία εκπαίδευσης, το τελικό μοντέλο κρίθηκε κατάλληλο ώστε να

παράγει προβλέψεις για την πιθανότητα υποτροπής σε διαφορετικά διαστήματα παρακολούθησης.

Στην εργασία αυτή, οι συγγραφείς σχεδίασαν διάφορα συστήματα με διαφορετικούς συνδυασμούς προγνωστικών παραγόντων ώστε να καταλήξουν στο σύνολο των μεταβλητών που επιτυγχάνει καλύτερη ακρίβεια πρόβλεψης (μέγεθος του όγκου, ηλικία του ασθενούς, αριθμός των λεμφογενών όζων κάτω από τις μασχάλες, αριθμός των κύσεων, ηλικία της πρώτης κύησης και ηλικία της εμμηναρχής). Σημειώνουμε ότι στην μεταγενέστερη μελέτη που διεξήγαγαν, οι συγγραφείς χρησιμοποίησαν τον αλγόριθμο CIDIM για να καταλήξουν στο βέλτιστο συνδυασμό των μεταβλητών. Ο παρακάτω πίνακας δείχνει τον μέσο όρο, τη διακύμανση και την τυπική απόκλιση των μεταβλητών που αντιπροσωπεύουν τους προγνωστικούς παράγοντες :

*Table 1.* Summary of patient data: means, standard deviation and ranges.

Prognostic factors	Minimum	Maximum	Range	Mean	STD
Age	24	89	65	54,59	12,27
Menarchy age	9	20	11	12,94	1,69
First pregnancy age	15	42	27	25,04	4,59
No. of Axillary lymph nodes	0	34	34	2,52	4,29
Tumor size	0	230	230	31,15	21,03
No. of Pregnancies	1	18	17	3,48	1,96

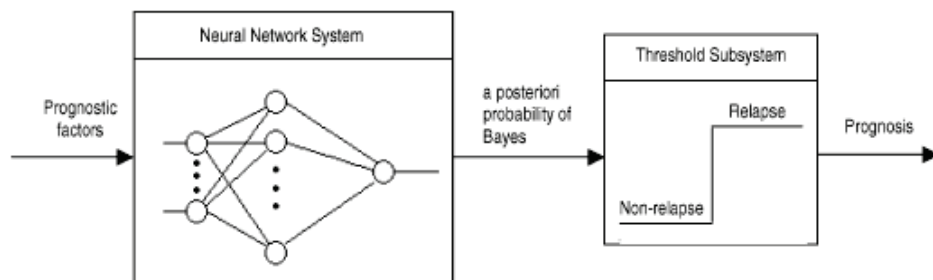
**Πίνακας 4 :** Περίληψη των δεδομένων των ασθενών, μέσος όρος, τυπική απόκλιση και διακύμανση.

Και σε αυτήν την περίπτωση, έγινε μια εκτίμηση του βέλτιστου σφάλματος Bayes και δόθηκε ο προγεγιστικός κανόνας απόφασης Bayes, ενώ τα αποτελέσματα του προτεινόμενου συστήματος για τα διαφορετικά χρονικά διαστήματα του χρόνου παρακολούθησης των ασθενών απεικονίζονται στον παρακάτω πίνακα. Το νευρωνικό δίκτυο υπολογίζει ένα σύνολο χαρακτηριστικών από το αρχείο του ασθενούς δίνοντας μια τιμή που αντιστοιχεί στην εκ των υστέρων πιθανότητα υποτροπής :

Time interval (# months)	Number of patients	A priori probability of relapse ( $p(C_1)$ )	$p_{net}$	STD error
$I_1$ (0–10)	845	6,75%	0.956	0.0071
$I_2$ (10–20)	741	10,66%	0.945	0.0084
$I_3$ (20–30)	681	6,90%	0.962	0.0073
$I_4$ (30–40)	600	6,33%	0.963	0.0077
$I_5$ (40–50)	520	4,36%	0.972	0.0072
$I_6$ (50–60)	466	5,36%	0.964	0.0086
$I_7$ (>60)	466	7,08%	0.943	0.0107

**Πίνακας 5 :** Αριθμός των ασθενών, η εκ των προτέρων πιθανότητα υποτροπής, η πιθανότητα που λαμβάνεται από τον κανόνα απόφασης Bayes και το τυπικό σφάλμα (STD) για κάθε χρονικό διάστημα υπό μελέτη.

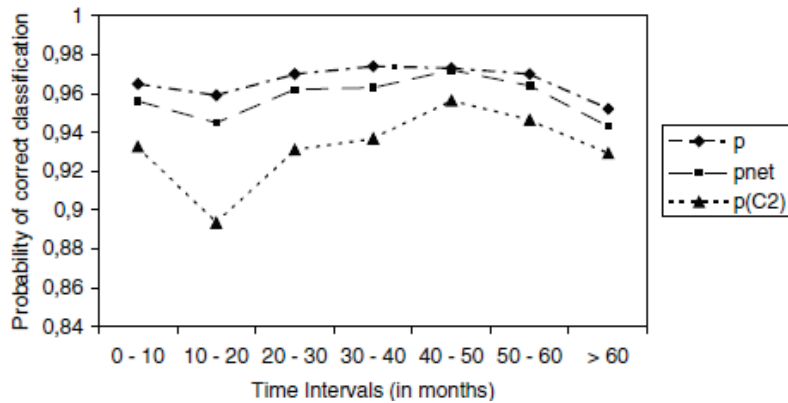
Το σύστημα αποτελείται από ένα νευρωνικό δίκτυο (πολυεπίπεδος αισθητήρας) και μια μονάδα κατωφλίου (*threshold unit*) για την εφαρμογή της διαδικασίας λήψης αποφάσεων, όπως φαίνεται στο παρακάτω σχήμα. Πρόκειται για πολύ απλούστερο σύστημα σε σχέση με αυτό που περιγράφηκε παραπάνω, το οποίο ανέπτυξαν μετεγενέστερα οι συγγραφείς :



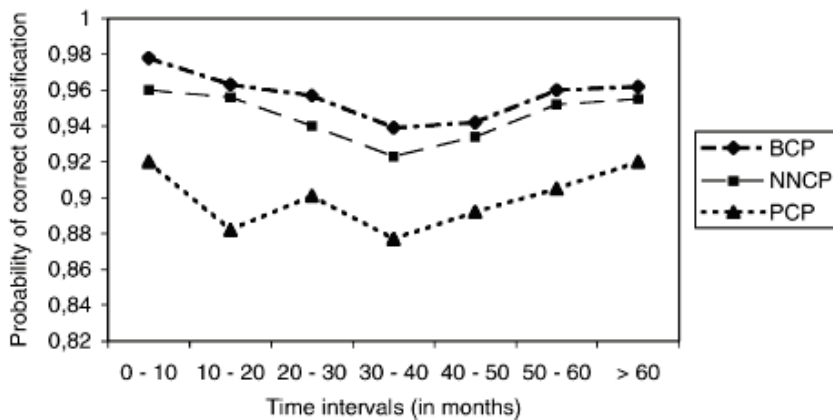
**Εικόνα 18 :** Το προγνωστικό σύστημα που βασίζεται στο νευρωνικό δίκτυο και στην μονάδα κατωφλίου για τη διαδικασία λήψης αποφάσεων.

Σε γενικές γραμμές, στη συνέχεια, ακολουθήθηκαν παρόμοια βήματα με αυτά που περιγράφηκαν στην ανάλυση της μεταγενέστερης μελέτης των συγγραφέων, όσον αφορά τη διάκριση της βάσης δεδομένων σε σύνολο εκπαίδευσης και επικύρωσης και τη διαδικασία ταξινόμησης ανάλογα με την κατάσταση επιβίωσης και το χρονικό διάστημα για κάθε ασθενή. Για την αξιολόγηση της πιθανότητας σωστής ταξινόμησης του παρόντος προτεινόμενου συστήματος, η εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes, η εκ των προτέρων πιθανότητα ταξινόμησης και η πιθανότητα που λαμβάνεται από τον κανόνα απόφασης Bayes απεικονίστηκαν ταυτόχρονα για κάθε χρονικό διάστημα στο παρακάτω γράφημα. Κρίνουμε σκόπιμο σε αυτό το σημείο να αντιπαραβάλλουμε τα διαγράμματα τόσο για το παρόν σύστημα, όσο και για το μεταγενέστερο ώστε να καταδείξουμε τις διαφορές ανάμεσα στα δύο συστήματα :





**Εικόνα 19 :** Η εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes ( $p$ ), η πιθανότητα που λαμβάνεται από τον κανόνα απόφασης Bayes ( $p_{net}$ ) και η εκ των προτέρων πιθανότητα ταξινόμησης και ( $p(C2)$ ) για κάθε χρονικό διάστημα υπό μελέτη (**αρχική μελέτη**).



**Εικόνα 20 :** Η εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes (BCP), η σωστή πιθανότητα ταξινόμησης που λαμβάνεται με το προτεινόμενο σύστημα νευρωνικών δικτύων (NNCP) και η εκ των προτέρων πιθανότητα ταξινόμησης και (PCP) για κάθε χρονικό διάστημα υπό μελέτη (**μεταγενέστερη μελέτη**).

Αυτό που επιχειρούν τελικά οι συγγραφείς σε αυτή την μελέτη είναι η δημιουργία ενός προγνωστικού μοντέλου που βασίζεται στον κανόνα απόφασης Bayes, το οποίο παράγει προβλέψεις για την πιθανότητα υποτροπής σε διαφορετικά διαστήματα της παρακολούθησης με ένα μικρό ποσοστό λάθους. Επιπλέον, προτείνουν μία μέθοδο βασισμένη σε νευρωνικό δίκτυο για την εκτίμηση της σωστής πιθανότητας ταξινόμησης Bayes, βασισμένη στην προσέγγιση της εκ των υστέρων πιθανότητας από μια κατανομή που χρησιμοποιεί τον πολυεπίπεδο αισθητήρα. Το επόμενο στάδιο και ο επόμενος στόχος της μελέτης που ακολούθησε ήταν ο συνδυασμός ενός αλγορίθμου TDIDT (δέντρο απόφασης επαγωγής από πάνω προς τα κάτω) με το σύστημα νευρωνικού δικτύου που προτείνεται εδώ, καθώς ο αλγόριθμος TDIDT θα επιλέξει τους πιο σχετικούς προγνωστικούς παράγοντες σε διαφορετικά χρονικά διαστήματα, ενώ το νευρωνικό δίκτυο θα λάβει ως είσοδο αυτές τις μεταβλητές προκειμένου να επιτευχθεί μια καλή πιθανότητα σωστής ταξινόμησης.



### 3. Ένα εργαλείο λογισμικού για τον προσδιορισμό των θεραπευτικών μεθόδων καρκίνου του μαστού με χρήση της προσέγγισης εξόρυξης δεδομένων [5]

Στην εργασία αυτή, οι θεραπευτικές μέθοδοι καρκίνου του μαστού προσδιορίζονται χρησιμοποιώντας την εξόρυξη δεδομένων (*data mining*). Για το σκοπό αυτό, αναπτύσσεται λογισμικό προκειμένου να βοηθήσει τον ογκολόγο να υποδείξει την εφαρμογή των θεραπευτικών μεθόδων για ασθενείς με καρκίνο του μαστού. Οι περισσότερες θεραπείες που υφίστανται είναι τυποποιημένες και περιλαμβάνουν χειρουργική επέμβαση, θεραπεία με ακτινοβολία, χημειοθεραπεία, ορμονοθεραπεία, αγωγή με ταξιμοφαίνη (στεροειδές φάρμακο).

Αρχικά, χρησιμοποιούνται 462 δεδομένα ασθενών με καρκίνο του μαστού που έχουν ληφθεί από το Ογκολογικό Νοσοκομείο της Άγκυρας για να προσδιοριστούν οι μέθοδοι θεραπείας για νέους ασθενείς. Αυτή η βάση δεδομένων υφίσταται επεξεργασία από ένα εργαλείο εξόρυξης δεδομένων που λέγεται Weka, δηλαδή αφαιρούνται τα περιττά χαρακτηριστικά (πχ ημερομηνία που ο ασθενής ήρθε στο νοσοκομείο), άλλα σύνθετα χαρακτηριστικά χωρίζονται σε επιμέρους απλούστερα, μετονομάζονται κ.ά. Σημειώνουμε ότι το 80% των δεδομένων απαρτίζει το σύνολο εκπαίδευσης και τα υπόλοιπα δεδομένα το σύνολο δοκιμών, ενώ τα χαρακτηριστικά εισόδου και εξόδου που χρησιμοποιούν οι συγγραφείς απεικονίζονται στον παρακάτω πίνακα :

Using type	Attribute name	Using type	Attribute name
Input	age	Output	
	menopause		
	grade		
	lymphinv		ht
	t_tnm		tamox
	n_tnm		kt
	m_tnm		rt
	stage		
	er		
	pr		
	surgery		

**Πίνακας 6 :** Τύπος των χαρακτηριστικών που χρησιμοποιούνται στην επεξεργασία δεδομένων.

Στη συνέχεια, κατά τη διαδικασία αξιολόγησης του μοντέλου, αλγόριθμοι ταξινόμησης εφαρμόζονται ένας προς έναν στη βάση δεδομένων και τα αποτελέσματα συγκρίνονται ως προς το ποσοστό ακρίβειας (percentage accuracy) και την τιμή ρίζας μέσου τετραγωνικού σφάλματος (RMSE value), για κάθε ένα από τα τέσσερα χαρακτηριστικά εξόδου του Πίνακα 6 (ht, tamox, kt, rt), ώστε να βρεθεί η κατάλληλη θεραπευτική μέθοδος. Με άλλα λόγια, ένα ανεπτυγμένο πρόγραμμα λογισμικού που καλείται 'Treatment Assistant' ή αλλιώς «Βοηθός Θεραπείας» χρησιμοποιεί διαφορετικούς αλγορίθμους (ενδεικτικά αναφέρουμε IB1, Πολυεπίπεδος Αισθητήρας και Πίνακας Αποφάσεων) για να ανακαλύψει ποιος δίνει το καλύτερο αποτέλεσμα για κάθε χαρακτηριστικό εξόδου, χρησιμοποιώντας ένα περιβάλλον εργασίας με τίτλο Java Netbeans. Οι καλύτερες επιδόσεις εμφανίζονται για τους αλγορίθμους του παρακάτω πίνακα :

Output attribute name	Algorithm name	Percentage accuracy	RMSE
ht	lazy.IB1	94.6237	0.2319
tamox	functions.MultiLayerPerceptron	92.0000	0.2603
kt	rules.DecisionTable	97.7778	0.1184
rt	functions.MultiLayerPerceptron	95.2381	0.1474

**Πίνακας 7 :** Επιλεγμένοι αλγόριθμοι ταξινόμησης και σχετικές τιμές.

Οι θεραπευτικές μέθοδοι καθορίζονται για την μετεγχειρητική λειτουργία των ασθενών με καρκίνο του μαστού με τη χρήση αυτού του ανεπτυγμένου λογισμικού εργαλείου. Από τους αλγορίθμους που επικράτησαν, ο IB1 αποτελεί έναν απλό αλγόριθμο που χρησιμοποιεί την κοντινότερη απόσταση του παραδείγματος εκπαίδευσης από το παράδειγμα δοκιμής ώστε να τα κατατάξει στην ίδια κλάση, ο Πολυεπίπεδος Αισθητήρας αποτελεί μια δομή τεχνητού νευρωνικού δικτύου πρόσθιας (feedforward) τροφοδότησης εκπαιδευμένου με τον τυπικό αλγόριθμο ανάστροφης διάδοσης (backpropagation), ενώ ο Πίνακας Αποφάσεων λειτουργεί ανάλογα με τα δέντρα αποφάσεων για πολύ σύνθετα ή εκτεταμένα σύνολα συνθηκών.

Συμπερασματικά, τα αποτελέσματα της μελέτης, όπως απεικονίζονται στον Πίνακα 7, είναι για την ορμονοθεραπεία η έξοδος IB1 (ακρίβεια 94.6237%), για την ακτινοθεραπεία και την ταμοξιφαίνη η έξοδος Πολυεπίπεδος Αισθητήρας (ακρίβεια 95.2381% και 92% αντίστοιχα) και για τη χημειοθεραπεία η έξοδος του αλγορίθμου Πίνακα Αποφάσεων (ακρίβεια 97.7778%). Εν κατακλείδι, αυτή η εργασία δείχνει ότι η προσέγγιση της εξόρυξης δεδομένων μπορεί να αποτελέσει ένα χρήσιμο εργαλείο για ιατρικές εφαρμογές, ιδιαίτερα στο στάδιο απόφασης της θεραπείας, ενώ οι γιατροί μπορούν να λάβουν αποφάσεις σε σύντομο χρονικό διάστημα. Ωστόσο, ένας περιορισμός της μελέτης αυτής είναι ο σχετικά μικρός αριθμός των δεδομένων των ασθενών και τα ελλιπή στοιχεία κάποιων χαρακτηριστικών, ενώ δεν υπάρχει εξωτερική επικύρωση του συστήματος.

#### **4. Ένα προγνωστικό μοντέλο που παράγει ποσοτικές εκτιμήσεις της πιθανότητας υποτροπής ασθενών με καρκίνο του μαστού [6]**

Μέχρι πριν μια δεκαετία, το σύστημα σταδιοποίησης TNM (*Tumor-Node-Metastasis*) αποτελούσε το συνήθη τρόπο εκτίμησης της πρόγνωσης ασθενών με καρκίνο του μαστού. Ωστόσο, το σύστημα αυτό δεν εκμεταλλεύεται τις πληροφορίες που προκύπτουν από τους δείκτες της βιολογικής «επιθετικότητας» του καρκίνου του μαστού και είναι σαφώς μη ικανοποιητικό για τη διαδικασία λήψης αποφάσεων της βέλτιστης θεραπείας και την παροχή συμβουλών στους ασθενείς. Για το λόγο αυτό, οι συγγραφείς ανέπτυξαν ένα προγνωστικό μοντέλο, βασισμένο σε λίγες προγνωστικές μεταβλητές που εξετάζονται συχνά, το οποίο παράγει ποσοτικές εκτιμήσεις του κινδύνου υποτροπής μεμονωμένων ασθενών με καρκίνο του μαστού.

Αρχικά, οι συγγραφείς χρησιμοποίησαν δεδομένα που αφορούσαν 2441 από 2990 διαδοχικούς ασθενείς με καρκίνο του μαστού οι οποίοι υποβλήθηκαν σε χειρουργική εκτομή στο διάστημα 1974-2002 (σημειώνουμε ότι οι πληροφορίες για αυτούς τους 2441

ασθενείς ήταν ολοκληρωμένες), ώστε να αναπτύξουν ένα τεχνητό νευρωνικό δίκτυο (ΤΝΔ) για την πρόβλεψη της πιθανότητας υποτροπής σε διάστημα πέντε ετών. Οι προγνωστικές μεταβλητές που χρησιμοποιήθηκαν ήταν οι εξής : ηλικία του ασθενούς, μέγεθος του όγκου, αριθμός των μεταστάσεων κάτω από την μασχάλη, οιστρογόνα και επίπεδα του υποδοχέα προγεστερόνης, κλάσμα S-φάσης και πλοειδία του όγκου (δηλαδή ο αριθμός πολλαπλασιών των απλών χρωμοσωμάτων ενός κυττάρου), ενώ ορισμένες από τις παραπάνω μεταβλητές μετασχηματίστηκαν λογαριθμικά προκειμένου να μειωθεί η ασυμμετρία στην κατανομή των τιμών και να σταθεροποιηθεί η διακύμανση. Επιπλέον οι συγγραφείς χώρισαν τυχαία τη βάση δεδομένων σε τρία υποσύνολα περίπου ίσου μεγέθους, σχηματίζοντας ένα σύνολο εκπαίδευσης, ένα σύνολο δοκιμών και ένα σύνολο εσωτερικής επικύρωσης. Οι επιδόσεις του μοντέλου αξιολογήθηκαν όσον αφορά τη διακριτική ικανότητα και την ποσοτική ακρίβεια, ενώ οι προβλέψεις επικυρώθηκαν και σε ένα ανεξάρτητο σύνολο 310 ασθενών από έναν ιατρικό φορέα άλλης χώρας (πιο συγκεκριμένα από την Ιταλία, ενώ η αρχική βάση δεδομένων προήλθε από το Πανεπιστήμιο του Τέξας).

Το μοντέλο τεχνητού νευρωνικού δικτύου που χρησιμοποιήθηκε ήταν ο πολυεπίπεδος αισθητήρας με ένα κρυφό επίπεδο πέντε στοιχείων επεξεργασίας, ενώ χρησιμοποιήθηκε και ο αλγόριθμος οπισθοδιάδοσης λάθους (*back-propagation-of-error algorithm*) για να προσαρμόσει τα βάρη σύνδεσης κατά τη διαδικασία μάθησης. Τόσο για το εσωτερικό σύνολο επικύρωσης, όσο και για το ανεξάρτητο σύνολο εξετάστηκε η περιοχή κάτω από την καμπύλη ROC, μέσω της εκτίμησης μέγιστης πιθανότητας. Η προγνωστική ακρίβεια των εκτιμήσεων του μοντέλου απεικονίζεται στον παρακάτω πίνακα όπου οι ασθενείς έχουν χωριστεί στα γκρουπ ANN1 – ANN8 αυξανόμενου κινδύνου :

Risk group	Prediction range	Internal validation set [Az <sup>a</sup> 0.728 (SD = 0.03)]			External validation set [Az 0.732 (SD = 0.04)]		
		Average prediction	Observed relapse rate (SE)	n	Average prediction	Observed relapse rate (SE)	n
ANN1	0.00–0.15	0.133	0.110 (0.03)	172	0.130	0.121 (0.06)	16
ANN2	0.16–0.25	0.200	0.170 (0.03)	290	0.210	0.231 (0.05)	78
ANN3	0.26–0.35	0.290	0.245 (0.05)	122	0.296	0.294 (0.03)	88
ANN4	0.36–0.45	0.404	0.385 (0.07)	72	0.415	0.450 (0.08)	39
ANN5	0.46–0.55	0.494	0.465 (0.08)	64	0.510	0.530 (0.09)	36
ANN6	0.56–0.65	0.583	0.549 (0.08)	34	0.589	0.578 (0.10)	23
ANN7	0.66–0.80	0.712	0.698 (0.08)	39	0.725	0.865 (0.09)	15
ANN8	0.81–1.00	0.903	0.914 (0.09)	21	0.911	0.869 (0.09)	15

<sup>a</sup> Az = area-under-the-ROC-curve

**Πίνακας 8 :** Προγνωστική ακρίβεια των προβλέψεων του μοντέλου για το εσωτερικό και το εξωτερικό σύνολο επικύρωσης.

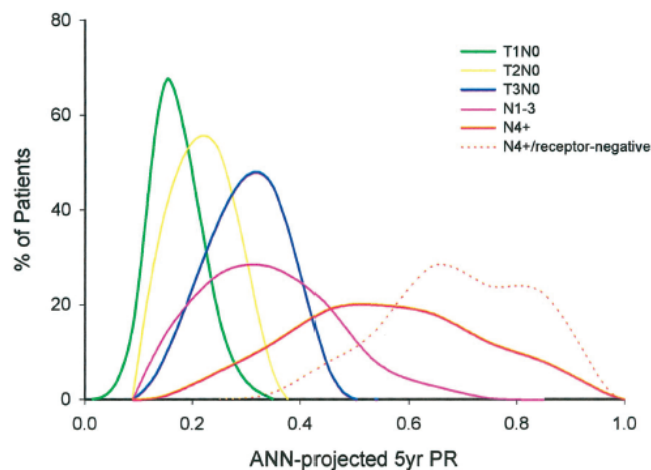
Στη συνέχεια, οι ασθενείς από τα δύο σύνολα επικύρωσης συνενώθηκαν σε ένα σύνολο (n=1124), προκειμένου να συγκριθεί η ικανότητα πρόβλεψης του μοντέλου με αυτή του συστήματος ταξινόμησης TNM. Σημειώνουμε ότι το σύστημα TNM τροποποιήθηκε ώστε να λαμβάνεται υπόψη ο αριθμός των μεταστάσεων στους λεμφαδένες της μασχάλης, με αποτέλεσμα τη δημιουργία πέντε κλάσεων : T1N0, T2N0, T3N0, N1-3 (ασθενείς με 1-3 μεταστάσεις), N41 (ασθενείς με 4 ή περισσότερες μεταστάσεις). Προέκυψε ότι το ΤΝΔ κατάφερε να διαχωρίσει τους ασθενείς ανάλογα με τον κίνδυνο υποτροπής καλύτερα από το σύστημα TNM (P = 0.0015), ενώ η ποσοτική ακρίβεια των εκτιμήσεων του μοντέλου ήταν ακριβής, γεγονός που αποδεικνύεται από τα αποτελέσματα του παρακάτω πίνακα :

	Az <sup>a</sup> (SD)	z	P
ANN	0.726 (0.02)	-3.1778	0.0015
TNM	0.677 (0.02)		

<sup>a</sup> Az, area-under-the-ROC-curve; z, z statistic.

**Πίνακας 9 :** Διακριτική ικανότητα του μοντέλου του TNΔ έναντι του τροποποιημένου συστήματος TNM σε ασθενείς από τα δύο σύνολα επικύρωσης.

Για να αποσαφηνιστούν οι διαφορές μεταξύ του μοντέλου TNΔ και του συστήματος ταξινόμησης TNM, στο παρακάτω γράφημα παριστάνεται γραφικά η κατανομή των προβλέψεων του TNΔ για ασθενείς που ανήκουν στις πέντε κλάσεις του τροποποιημένου συστήματος TNM. Αποδεικνύεται ότι ο κίνδυνος υποτροπής που προέκυψε από το μοντέλο που ανέπτυξαν οι συγγραφείς ποίκιλε σε μεγάλο βαθμό μέσα στην ίδια κλάση TNM, ιδιαίτερα για ασθενείς με τέσσερις ή περισσότερες λεφογενείς μεταστάσεις :

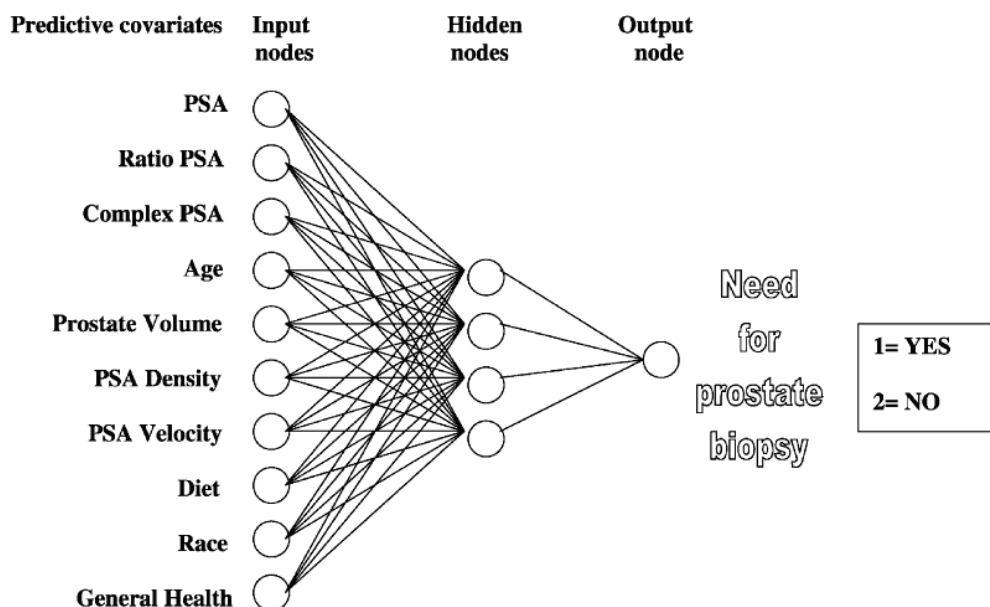


**Εικόνα 21 :** Κατανομή των εκτιμήσεων του TNΔ της πενταετούς πιθανότητας υποτροπής ασθενών που έχουν ταξινομηθεί σύμφωνα με το τροποποιημένο σύστημα TNM.

Τέλος, οι συγγραφείς καταλήγουν στο συμπέρασμα ότι το μοντέλο τους παράγει καλύτερα αποτελέσματα για την πρόγνωση σε σχέση με την ταξινόμηση TNM και είναι σε θέση να εντοπίσει τους ασθενείς με εντυπωσιακά διαφορετικούς κινδύνους υποτροπής μέσα σε κάθε κλάση TNM. Το πιο σημαντικό όμως σε αυτή την μελέτη είναι η ύπαρξη του συνόλου εξωτερικής επικύρωσης, κάτι που δεν συναντάται συχνά. Το σύστημα ωστόσο επιδέχεται περαιτέρω βελτίωση μέσω της χρήσης λεπτομερών πληροφοριών για τη θεραπεία, της χρήσης νέων βιολογικών μεταβλητών και της δημιουργίας μοντέλων σε υποσύνολα ασθενών που θέτουν ειδικά θεραπευτικά ερωτήματα.

### 3.3 Μελέτες σχετικά με τον καρκίνο του προστάτη (Prostate Cancer)

Όσον αφορά τη δημοσιευμένη βιβλιογραφία για τον καρκίνο του προστάτη, παραθέτουμε συνοπτικά μια σειρά από μελέτες προκειμένου να διερευνηθεί το κλινικό όφελος των τεχνητών νευρωνικών δικτύων σε σχέση με τις συμβατικές στατιστικές μεθόδους στην Ουρολογική Ογκολογία (*Urologic Oncology*) [7]. Ο στόχος των μελετών που συμπεριλήφθηκαν ήταν η ακριβέστερη διάγνωση και πρόγνωση, ιδιαίτερα στον καρκίνο του προστάτη. Όπως έχει ήδη αναφερθεί, τα τεχνητά νευρωνικά δίκτυα αποτελούν υπολογιστικές μεθόδους που εκτελούν πολυπαραγοντικές αναλύσεις, εμπνευσμένα από τα δίκτυα των βιολογικών νευρώνων, ενώ βρίσκουν εφαρμογή σε πολλούς τομείς της Ιατρικής και της Ουρολογίας. Ένα παράδειγμα των διασυνδέσεων ανάμεσα στους κόμβους ενός νευρωνικού δικτύου που αποφαινεται για την αναγκαιότητα ή μη βιοψίας καρκίνου του προστάτη απεικονίζεται παρακάτω :



**Εικόνα 22 :** Παράδειγμα νευρωνικού δικτύου (βιοψία καρκίνου του προστάτη).

Κρίνεται σκόπιμη η διάκριση των εφαρμογών των τεχνητών νευρωνικών δικτύων στον καρκίνο του προστάτη σε τρεις ενότητες : (1) Έγκαιρη διάγνωση και πληθυσμιακός έλεγχος, (2) Κατάταξη σε στάδια (σταδιοποίηση), (3) Εξέλιξη.

Όσον αφορά την έγκαιρη διάγνωση και τον πληθυσμιακό έλεγχο για τον καρκίνο του προστάτη, αξιολογήθηκαν τέσσερις μελέτες. Προτού τις παραθέσουμε, πρέπει να διασαφηνίσουμε την έννοια του ειδικού προστατικού αντιγόνου, γνωστού ως PSA (Prostate-Specific Antigen), το οποίο έχει αναγνωριστεί παγκοσμίως ως προς τη χρησιμότητά του στην έγκαιρη διάγνωση καρκίνου του προστάτη. Το διάστημα τιμών PSA από 4 έως 10 ng/ml καλείται «γκρι» περιοχή λόγω της επικάλυψης ασθενών με καρκίνο του προστάτη και ασθενών με καλή ή υπερπλασία ή φλεγμονή, ενώ έχει αποδειχθεί και η ύπαρξη ενός σημαντικού ποσοστού ασθενών που διαγνώστηκαν με καρκίνο του προστάτη έχοντας τιμή PSA από 2,5 έως 4 ng/ml.

Στην μελέτη Babaian et al. [7], οι συγγραφείς ανέπτυξαν έναν αλγόριθμο βασισμένο σε ένα τεχνητό νευρωνικό δίκτυο που χρησιμοποίησε μεταβλητές όπως η ηλικία, το ολικό PSA, το ελεύθερο PSA και η κρεατινίνη κινάση, ενώ περιελάμβανε τρία ανεξάρτητα εκπαιδευμένα δίκτυα. Ο καρκίνος εντοπίστηκε στο 24,5% των ασθενών, δηλαδή στους 37 από τους 151 άνδρες της μελέτης με τιμές PSA 2,5-4 ng/ml, ενώ η αυαισθησία του δικτύου ήταν 92%. Οι συγγραφείς υποστηρίζουν ότι αν το σύστημά τους χρησιμοποιούνταν αντί του δείκτη ελεύθερου/ολικού PSA, θα σημειωνόταν μείωση των περιττών βιοψιών της τάξης του 39%.

Στην μελέτη Djavan et al. [7], αναπτύχθηκαν δύο νευρωνικά δίκτυα για ασθενείς με επίπεδα PSA 2,5-4 ng/ml και 4-10 ng/ml ώστε να προβλεφθεί η ύπαρξη καρκίνου του προστάτη, ενώ η προγνωστική τους ακρίβεια συγκρίνεται με αυτή των συμβατικών μονοπαραγοντικών στατιστικών αναλύσεων του ολικού PSA, της αναλογίας PSA κ.ά. Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι και τα δύο νευρωνικά δίκτυα υπερέχουν των συμβατικών παραμέτρων PSA, οδηγώντας σε λιγότερες βιοψίες, χωρίς όμως να παρουσιάζουν στατιστικά σημαντική διαφορά με τη λογιστική παλινδρόμηση. Στην περίπτωση που μια σειρά βιοψιών είναι αρνητική αλλά ενυπάρχει ακόμη η υποψία καρκίνου, οι Djavan et al. υποστηρίζουν ότι το μοντέλο τους βρήκε ένα ισχυρό προγνωστικό μοτίβο για την έκβαση επαναλαμβανόμενων βιοψιών, συνδυάζοντας μεμονωμένους κλινικούς και βιοχημικούς δείκτες και επιτυγχάνοντας ειδικότητα 68% σε υψηλά επίπεδα ευαισθησίας.

Στην μελέτη Finne et al. [7], οι συγγραφείς επιχείρησαν να συγκρίνουν έναν πολυεπίπεδο αισθητήρα (Multilayer Perceptron – MLP), την μέθοδο λογιστικής παλινδρόμησης και την μονοπαραγοντική ανάλυση αναλογίας PSA ως προς τη διαγνωστική τους απόδοση στην πρόβλεψη του αποτελέσματος βιοψίας, με τον πολυεπίπεδο αισθητήρα να επιδεικνύει καλύτερη ακρίβεια σε σχέση με τα άλλα δύο μοντέλα.

Στην μελέτη Ronco and Fernandez [7], οι συγγραφείς απέδειξαν ότι ένα νευρωνικό δίκτυο που συνδυάζει διάφορες κλινικές, βιοχημικές και υπερηχογραφικές μεταβλητές επιτυγχάνει υψηλότερη θετική και αρνητική προγνωστική αξία σε σχέση με τη λογιστική παλινδρόμηση όσον αφορά την πρόβλεψη καρκίνου, αλλά μικρότερη ακρίβεια. Αρκετές παρόμοιες μελέτες απεικονίζονται στον **Πίνακα 10**.

Όσον αφορά τη δεύτερη κατηγορία, δηλαδή την κατάταξη σε στάδια, συμπεριλαμβάνονται τέσσερις μελέτες. Οι Partin et al. [7] ανέπτυξαν τα πιο ευρέως διαδεδομένα μοντέλα για την πρόβλεψη του παθολογικού σταδίου της νόσου, τα οποία χρησιμοποιούν το κλινικό στάδιο, τη βιοψία Gleason Sum και τις τιμές PSA προ θεραπείας προκειμένου να εντοπίσουν το στάδιο σε περιπτώσεις εντοπισμένου καρκίνου.

Στην μελέτη Murphy et al. [7], χρησιμοποιήθηκαν διάφοροι προγνωστικοί παράγοντες σταδιοποίησης προκειμένου να μελετηθούν οι ικανότητες σταδιοποίησης ενός TNΔ. Αποδείχθηκε ότι από τους παράγοντες που εξετάστηκαν, τα επίπεδα PSA, το σπινθηρογράφημα οστών και μια σάρωση που λέγεται Prostateint αποτέλεσαν σημαντικές μεταβλητές πρόβλεψης της κατάστασης με ακρίβεια 50% σε σχέση με την ακρίβεια της τάξης του 20% της παραδοσιακής ανάλυσης παλινδρόμησης.

Σε μία παρόμοια μελέτη των Batuello et al. [7], χρησιμοποιήθηκαν το κλινικό στάδιο, η βιοψία Gleason Sum και τα επίπεδα PSA ως μεταβλητές εισόδου σε ένα ΤΝΔ προκειμένου να προβλεφθεί η εξάπλωση στους λεμφαδένες και με αυτό τον τρόπο οι συγγραφείς κατάφεραν να ερμηνεύσουν εξατομικευμένα τον κίνδυνο εξάπλωσης στους λεμφαδένες.

Οι Han et al. [7] χρησιμοποίησαν μεταβλητές εισόδου που περιελάμβαναν κλινικά και παθολογικά στοιχεία ασθενών μετά από ριζική προστατεκτομή, προκειμένου να αναπτύξουν έναν πολυεπίπεδο αισθητήρα και να δοκιμάσουν τη προγνωστική τους αξία στη σταδιοποίηση της νόσου. Συνολικά, το ΤΝΔ ξεπέρασε τα νομογράμματα που αναφέρθηκαν στην πρώτη μελέτη των Partin et al. [7] στην πρόβλεψη του παθολογικού σταδίου τη στιγμή της χειρουργικής επέμβασης και ήταν πιο ακριβές ως προς την ευασθησία και την ειδικότητα.

Όσον αφορά την τελευταία κατηγορία, δηλαδή την εξέλιξη/υποτροπή του καρκίνου, συμπεριλήφθηκαν έξι μελέτες, χωρίς ωστόσο να παρατίθενται αναλυτικά τα αποτελέσματα. Οι Mattfeldt et al. [7] σύγκριναν δύο γκρουπ 20 ασθενών (με ή χωρίς εξέλιξη της νόσου) ίδιας ηλικίας, προεγχειρητικού PSA και διάρκειας παρακολούθησης, ενώ το ΤΝΔ προέβλεψε την εξέλιξη του καρκίνου με χρήση τριών μόνο μεταβλητών (βαθμός Gleason, βαθμός WHO, διάμετρος όγκου) με ελπιδοφόρα αποτελέσματα. Οι Han et al. [7] κατέγραψαν την ικανότητα ενός ΤΝΔ να κατατάσσει τους ασθενείς ανάλογα με το βαθμό Gleason, ενώ οι Ziada et al. [7] σημείωσαν καλό αποτέλεσμα για το δικό τους ΤΝΔ. Οι Potter et al. [7] χρησιμοποίησαν τρία «γενετικά» κατασκευασμένα ΤΝΔ ανάλογα με τις μεταβλητές που θα χρησιμοποιούνταν κατά περίπτωση (παθολογία και ηλικία, DNA, όλες οι μεταβλητές), τα οποία μεγιστοποίησαν την απόδοσή τους στα σύνολα δοκιμών σε σχέση με την ανάλυση λογιστικής παλινδρόμησης. Οι Porter et al. [7] βρήκαν την μεγαλύτερη (80%) περιοχή κάτω από την καμπύλη στην ανάλυση ROC για το δικό τους μοντέλο νευρωνικού δικτύου. Τέλος, οι Naguib et al. [7] κατάφεραν, χρησιμοποιώντας ένα ΤΝΔ, να ταξινομήσουν σωστά το 80% των ασθενών ως προς την έκβασή τους σύμφωνα με 6 συμβατικούς παράγοντες (ηλικία, στάδιο, θεραπεία κτλ) και 2 πειραματικούς δείκτες (ενδεικτικά αναφέρουμε bcl-2, p53).

Ακολουθεί ο **Πίνακας 10** με τα συγκεντρωτικά αποτελέσματα των μελετών που προαναφέρθηκαν :



Reference	Application	Sensitivity (%)	Specificity (%)	Accuracy (%)
[14]	Diagnosis	92	62	
[15]	Diagnosis	95	59–67	0.87–0.91 <sup>a</sup>
[16]	Diagnosis	90	46	56
[18]	Diagnosis	95	68	0.91 <sup>a</sup>
[27]	Diagnosis			87
[28]	Diagnosis	90	97	
[29]	Diagnosis			0.95 <sup>a</sup>
[30]	Diagnosis	81	92	90
[31]	Diagnosis			80
[32]	Diagnosis	72	78	77
[33]	Diagnosis	90	+32–44 <sup>b</sup>	
[46]	Diagnosis	95	38	
[19]	Staging	95	48 <sup>c</sup>	
[20]	Staging			0.81 <sup>a</sup>
[21]	Staging			0.88 <sup>a</sup>
[24]	Staging	79	81	
[34]	Staging	94	69	
[39]	Staging	81–100	72–75	
[35]	Imaging			87–99
[36]	Imaging	79		
[37]	Imaging			82
[22]	Prognosis			85
[23]	Prognosis	31	95 <sup>d</sup>	
[24]	--/--			80
[25]	Prognosis			80
[26]	Prognosis	74	78	0.80 <sup>a</sup>
[27]	Prognosis			90
[38]	Prognosis			80

<sup>a</sup> Area under ROC curve.  
<sup>b</sup> This percentage shows the enhancement of ANN in terms of specificity, compared to PSA, percent free PSA specificity values.  
<sup>c</sup> Refers as to nodal positivity, while traditional regression analysis is slightly above 20%.  
<sup>d</sup> Refers as to prediction of five-year recurrence, while Logistics Regression using all input variables shows 16%.

**Πίνακας 10 :** Άρθρα που έχουν εκδοθεί στο πεδίο της χρήσης των ΤΝΔ στον καρκίνο του προστάτη (δεν περιλαμβάνονται reviews).

Η παραπάνω δημοσιευμένη βιβλιογραφία πάνω στον καρκίνο του προστάτη μετρά ήδη μία δεκαετία και δεν μπορεί εκ των πραγμάτων να συμπεριλάβει σύγχρονα μοντέλα που επιδεικνύουν εξαιρετικά αποτελέσματα. Αυστηρά με βάση τα παραπάνω, θα μπορούσαμε να διατυπώσουμε με επιφύλαξη το συμπέρασμα ότι τα νευρωνικά δίκτυα και οι στατιστικές μέθοδοι καλό θα ήταν να χρησιμοποιούνται με συμπληρωματικό τρόπο κατά την ερμηνεία των ιατρικών δεδομένων, καθώς ανακύπτουν προβλήματα όπως η ποιότητα του ιατρικού ιστορικού, η ανάγκη για κοινά πρότυπα, οι βάσεις δεδομένων κατάλληλης ποιότητας.

Στην συνέχεια ακολουθεί η ανάλυση τριών επιλεγμένων μελετών που αφορούν τον καρκίνο του προστάτη. Σκοπός των δύο πρώτων μελετών ήταν η διάγνωση του καρκίνου του προστάτη, ενώ η τρίτη μελέτη που περιγράφει ένα knowledge-based σύστημα παρουσιάζει ποσοτικοποιημένες πληροφορίες για τον συγκεκριμένο τύπο καρκίνου.



## 1. Πολυκεντρική αξιολόγηση ενός Τεχνητού Νευρωνικού Δικτύου για την αύξηση του ποσοστού ανίχνευσης καρκίνου του προστάτη και την μείωση των περιττών βιοψιών [8]

Σε όσα αναφέρθηκαν νωρίτερα, κάναμε λόγο για την κλινική σημασία του ειδικού προστατικού αντιγόνου (PSA) στη διάγνωση του καρκίνου του προστάτη. Έχει αποδειχθεί ότι το ποσοστό του ελεύθερου ειδικού προστατικού αντιγόνου (%fPSA) βελτιώνει την ειδικότητα για τη διάγνωση του καρκίνου του προστάτη σε σχέση με το ολικό PSA (tPSA). Για το λόγο αυτό, οι συγγραφείς διεξήγαγαν μία μελέτη προκειμένου να αξιολογήσουν τη διαγνωστική αξία ενός τεχνητού νευρωνικού δικτύου που βασίζεται στο εν λόγω ποσοστό (%fPSA) σε άνδρες με συγκεντρώσεις tPSA μεταξύ 2 και 20 µg/L, για την ανίχνευση των ασθενών με αυξημένο κίνδυνο θετικής βιοψίας για τον καρκίνο του προστάτη.

Μελετήθηκαν 1188 άνδρες (ηλικίας από 40 έως 89 ετών) από 6 διαφορετικά νοσοκομεία με καρκίνο του προστάτη ή καλοήγη προστάτη στο διάστημα 1996-2001. Χρησιμοποιήθηκε ένα προσφάτως ανεπτυγμένο ΤΝΔ με δεδομένα εισόδου όπως : το ολικό PSA, το %fPSA, η ηλικία του ασθενούς, ο όγκος του προστάτη και η δακτυλική εξέταση του ορθού (digital rectal evaluation – DRE) προκειμένου να υπολογιστεί ο κίνδυνος της παρουσίας καρκίνου του προστάτη εντός διαφορετικών περιοχών tPSA (2 – 4, 4.1 – 10, 2 – 10, 10.1 – 20 και 2 – 20 µg/L) σε ειδικότητα ή σημείο αποκοπής ευαισθησίας της τάξης του 90% και 95%, ανάλογα με τη συγκέντρωση tPSA. Τέλος, χρησιμοποιήθηκαν η ανάλυση ROC και οριακοί υπολογισμοί για την εκτίμηση της διαγνωστικής βελτίωσης του ΤΝΔ σε σύγκριση με το ποσοστό %fPSA.

Αποδείχθηκε ότι στο χαμηλό εύρος tPSA (2 – 4 µg/L), το ΤΝΔ ανίχνευσε το 72% και 65% των καρκίνων σε ειδικότητες 90% ή 95% αντίστοιχα. Στο εύρος tPSA 4 – 10 µg/L, το ΤΝΔ ανίχνευσε το 90% και 95% των καρκίνων με ειδικότητες 62% και 41% αντίστοιχα. Η χρήση του ΤΝΔ στο εύρος tPSA 2 – 10 µg/L ενίσχυσε την ειδικότητα του %fPSA κατά 20 – 22%, μειώνοντας έτσι τον αριθμό των περιττών βιοψιών. Τα αποτελέσματα της μελέτης επεικονίζονται αναλυτικά στους παρακάτω πίνακες :

tPSA range, µg/L	Sensitivity, %	Specificity, <sup>a</sup> %			Cutoffs <sup>b</sup>	
		tPSA	%fPSA	ANN	%fPSA	ANN
2–4	90	25 (19–32)	20 (15–27)	38 (31–45)	23.7	0.04
	95	19 (14–25)	14 (9.2–19)	22 (17–29)	26.8	–0.04
4.1–10	90	20 (16–25)	34 (29–40)	62 (57–67)	17.3	0.39
	95	10 (7.3–14)	21 (16–26)	41 (36–47)	19.4	0.27
10.1–20	90	12 (7.2–20)	40 (31–49)	62 (53–71)	15.1	0.45
	95	5.7 (2.3–12)	26 (19–35)	39 (30–48)	18.3	0.38
2–10	90	30 (26–34)	31 (27–35)	53 (49–58)	19.4	0.27
	95	21 (17–25)	13 (10–16)	33 (29–37)	24.1	0.03
2–20	90	31 (28–35)	34 (31–38)	61 (57–65)	17.6	0.24
	95	22 (19–25)	23 (2–26)	35 (32–39)	21.6	0.02

<sup>a</sup> Specificities and their 95% confidence intervals (in parentheses) at sensitivity cutoffs of 90% and 95% for all investigated tPSA ranges.

<sup>b</sup> Cutoffs for %fPSA and the values obtained with the ANN at the given sensitivities of 90% and 95%.

**Πίνακας 11 :** Ειδικότητες σε δοσμένες τιμές ευαισθησίας για το tPSA, το %fPSA και το ΤΝΔ σε όλα τα εύρη PSA που ερευνούνται.

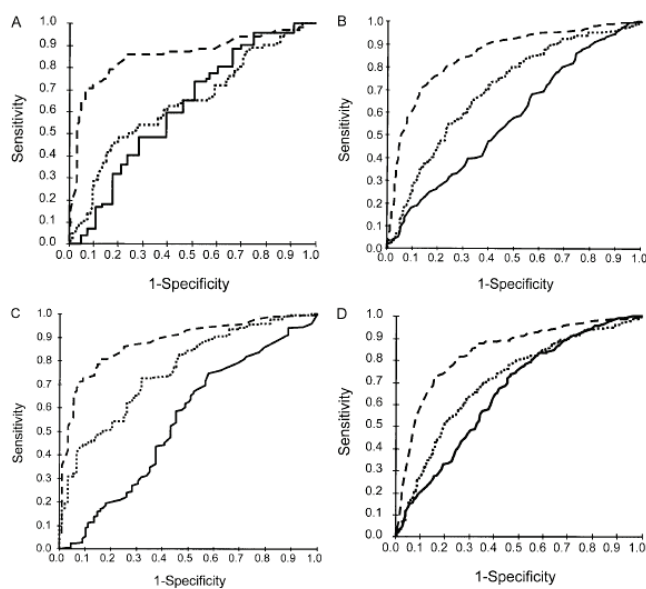
tPSA range, μg/L	Specificity, %	Sensitivity, <sup>a</sup> %			Cutoffs <sup>b</sup>	
		tPSA	%fPSA	ANN	%fPSA	ANN
2–4	90	6.9 (2.8–14)	28 (19–38)	72 (62–81)	9.6	0.32
	95	4.2 (1.2–11)	9.7 (4.7–18)	65 (55–75)	6.5	0.8
4.1–10	90	19 (15–22)	26 (23–30)	61 (57–66)	7.4	0.77
	95	7.5 (5.4–10)	13 (10–16)	48 (44–53)	5.6	0.83
10.1–20	90	5.5 (3.4–8.3)	44 (39–49)	73 (68–77)	7.2	0.72
	95	2.5 (1.2–4.8)	31 (26–36)	55 (50–60)	6.0	0.93
2–10	90	21 (18–24)	27 (23–30)	59 (56–64)	8.1	0.85
	95	14 (11–17)	14 (11–17)	35 (31–39)	6.0	0.89
2–20	90	26 (23–29)	32 (29–35)	64 (61–67)	7.5	0.77
	95	9.2 (7.5–11)	20 (18–23)	45 (42–48)	6.0	0.91

<sup>a</sup> Sensitivities and their 95% confidence intervals (in parentheses) at specificity cutoffs of 90% and 95% for all investigated tPSA ranges.

<sup>b</sup> Cutoffs for %fPSA and the values obtained with the ANN at the given specificities of 90% and 95%.

**Πίνακας 12 :** Ευαισθησίες σε δοσμένες τιμές ειδικότητας για το tPSA, το %fPSA και το TNΔ σε όλα τα εύρη PSA που ερευνούνται.

Παραθέτουμε επίσης και τις καμπύλες ROC για κάθε ένα από τα τέσσερα διαστήματα :



**Εικόνα 23 :** Καμπύλες ROC για το tPSA (συνεχής γραμμή), το %fPSA (τελείες) και το TNΔ (διακεκομμένη γραμμή). Το γράφημα A αντιστοιχεί στο εύρος 2 – 4 μg/L, το γράφημα B αντιστοιχεί στο εύρος 4.1 – 10 μg/L, το γράφημα Γ αντιστοιχεί στο εύρος 10.1 – 20 μg/L και το γράφημα Δ αντιστοιχεί στο εύρος 2 – 10 μg/L.

Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι ένα τεχνητό νευρωνικό δίκτυο που βασίζεται στο ποσοστό %fPSA και περιλαμβάνει επίσης κλινικές πληροφορίες από τη δακτυλική εξέταση του ορθού (digital rectal evaluation – DRE) και μετρήσεις του όγκου του προστάτη μπορεί να επιτύχει αυξημένη ακρίβεια στην ανίχνευση του καρκίνου του προστάτη σε σχέση με αυτή που αποκτάται μόνο από τη χρήση του %fPSA, συντελώντας έτσι στην περαιτέρω μείωση των περιττών βιοψιών στο διάστημα tPSA 2 – 10 μg/L και στην μείωση των επαναλαμβανόμενων βιοψιών στο διάστημα tPSA 10.1 – 20 μg/L. Ωστόσο, το παρόν TNΔ πρέπει να επικυρωθεί μελλοντικά και σε άλλους πληθυσμούς.

## **2. Προ-επεμβατική πρόβλεψη προχωρημένου καρκίνου του προστάτη χρησιμοποιώντας συστήματα υποστήριξης κλινικών αποφάσεων : Σύγκριση ακρίβειας ανάμεσα στις μηχανές υποστήριξης διανυσμάτων και στα τεχνητά νευρωνικά δίκτυα [9]**

Ο σκοπός της παρούσα μελέτης ήταν η ανάπτυξη μοντέλων μηχανής υποστήριξης διανυσμάτων (*Support Vector Machine – SVM*) και τεχνητού νευρωνικού δικτύου (*TNΔ*) για την προεγχειρητική πρόβλεψη προχωρημένου καρκίνου του προστάτη με χρήση παραμέτρων που αποκτήθηκαν από βιοψίες του προστάτη με γνώμονα το διορθικό υπερηχογράφημα (*transrectal ultrasound – TRUS*), καθώς και η σύγκριση της ακρίβειας ανάμεσα στα δύο μοντέλα. Στόχος δηλαδή, εδώ, των συστημάτων υποστήριξης κλινικών αποφάσεων είναι η πρόβλεψη του παθολογικού σταδίου της ασθένειας.

Αρχικά, οι 532 ασθενείς, ηλικίας 39-80 ετών, που συμπεριλήφθησαν στην μελέτη και υποβλήθηκαν σε βιοψίες προστάτη και προστατεκτομές για καρκίνο του προστάτη στο διάστημα 2003-2008, χωρίστηκαν τυχαία σε ομάδες εκπαίδευσης (n=300 ασθενείς) και ομάδες δοκιμών (n=232 ασθενείς).

Από τα δεδομένα στην ομάδα εκπαίδευσης (training group), κατασκευάστηκαν δύο συστήματα υποστήριξης κλινικών αποφάσεων (SVM και TNΔ) με παραμέτρους εισόδου όπως : ηλικία, επίπεδο ειδικού προστατικού αντιγόνου/PSA, ψηφιακή εξέταση ορθού και πέντε παραμέτρους βιοψίας, και δεδομένο εξόδου την πιθανότητα για προχωρημένο καρκίνο του προστάτη [ $>pT3a$ ]. Από τα δεδομένα στην ομάδα δοκιμών (testing group), αξιολογήθηκε η ακρίβεια των δεδομένων εξόδου.

Το μοντέλο του τεχνητού νευρωνικού δικτύου που ανέπτυξαν οι συγγραφείς αποτελείτο από αρχιτεκτονική αισθητήρα τριών στρωμάτων, με ένα επίπεδο εισόδου 8 κόμβων, ένα κρυφό επίπεδο 10 κρυφών κόμβων και ένα επίπεδο εξόδου με ένα κόμβο. Το TNΔ εκπαιδεύτηκε χρησιμοποιώντας τα δεδομένα του συνόλου εκπαίδευσης (n=300) ώστε να παράγει τιμή ίση με 1 για προχωρημένο καρκίνο του προστάτη (μεγαλύτερο από pT3a) και τιμή -1 για περιορισμένο καρκίνο του προστάτη, ενώ οι συγγραφείς μετέτρεψαν, εφαρμόζοντας μια σιγμοειδή συνάρτηση, την τιμή εξόδου του TNΔ σε πιθανότητα να έχει ο ασθενής προχωρημένο καρκίνο (πιθανότητα σε ποσοστιαίες μονάδες).

Όσον αφορά το μοντέλο SVM, οι συγγραφείς χρησιμοποίησαν το λογισμικό LIBSVM των Chang και Lin για το πρόβλημα δύο κατηγοριών (προχωρημένου ή περιορισμένου καρκίνου), μοντέλο που παρουσιάζει θεωρητική βέλτιστη διαχωριστικότητα όσον αφορά τα δεδομένα εκπαίδευσης. Το μοντέλο SVM έλαβε οκτώ δεδομένα εισόδου από το σύνολο εκπαίδευσης και μία τιμή εξόδου για κάθε ασθενή, ενώ εκπαιδεύτηκε ώστε να παράγει αντίστοιχα αποτελέσματα με το TNΔ.

Οι περιοχές κάτω από την καμπύλη (AUC) λειτουργικών χαρακτηριστικών δέκτη (ROC) υπολογίστηκαν για να γίνει σύνοψη των συνολικών επιδόσεων των δύο συστημάτων και πραγματοποιήθηκε μια σύγκριση των καμπυλών ROC ( $p<0.05$ ), της ακρίβειας, της ευαισθησίας και της ειδικότητας τους.

Ο Πίνακας 1 που ακολουθεί δείχνει τις μέσες και τυπικές αποκλίσεις των συνόλων εκπαίδευσης και δοκιμών :

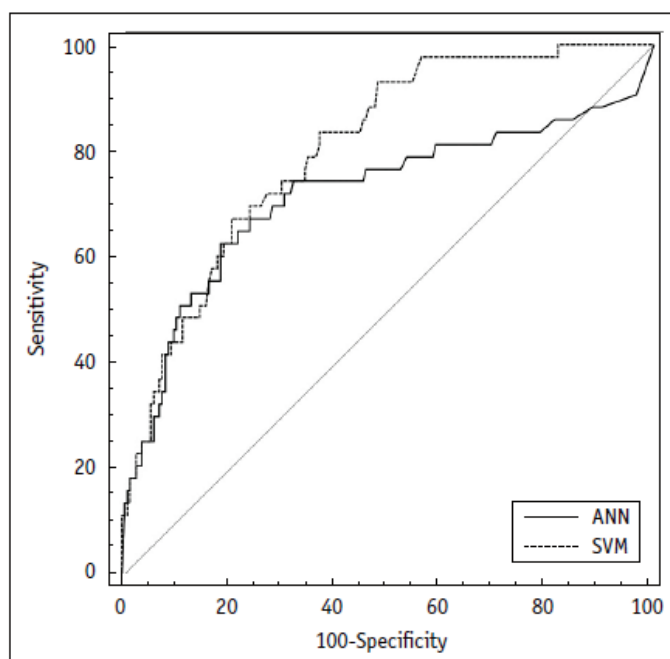
**Table 1. Mean and Standard Deviations of Input Parameters in Training and Test Groups**

	Mean ( $\pm$ SD)		
	Training Group (n = 300)	Test Group (n = 232)	Total (n = 532)
Age (Year)	65.2 (6.5)	64.3 (6.8)	64.8 (6.6)
Total PSA (ng/mL)	11.97 (13.26)	11.88 (21.23)	11.93 (17.13)
Percentages of DRE-positive (%)	82 (27)	53 (23)	135 (25)
Mean number of positive cores	4.5 (3.3)	3.4 (3.0)	4.0 (3.2)
Mean percentage of positive biopsy core (%)	34.7 (25)	26.6 (23)	31.2 (25)
Mean total linear cancer length (cm)	2.23 (2.89)	1.45 (2.38)	1.90 (2.71)
Mean percentage cancer length (%)	26.6 (18)	20.5 (15)	24.0 (17)
Mean maximum cancer core length (cm)	0.64 (0.44)	0.47 (0.39)	0.57 (0.43)
Percentages of advanced prostate cancer (%)	94 (31)	62 (27)	156 (29)

**Note.**— DRE = digital rectal examination, PSA = prostate specific antigen, SD = standard deviation

**Πίνακας 13 :** Μέσες και τυπικές αποκλίσεις των παραμέτρων εισόδου στα σύνολα εκπαίδευσης και δοκιμών.

Όσον αφορά την απόδοση των συστημάτων, η περιοχή κάτω από την καμπύλη AUC της μηχανής υποστήριξης διανυσμάτων και του τεχνητού νευρωνικού δικτύου ήταν 0.805 και 0.719 αντίστοιχα (στατιστική διαφορά  $p = 0.020$ ) στην προεγχειρητική πρόβλεψη προχωρημένου καρκίνου του προστάτη, με τιμές ευαισθησίας, ειδικότητας και ακρίβειας για το SVM ίσες με 67%, 79% και 77% αντίστοιχα, ενώ για το ΤΝΔ οι παραπάνω τιμές ισούσαν με 63%, 81% και 78%. Το παρακάτω γράφημα απεικονίζει τις καμπύλες ROC για τα δύο μοντέλα :



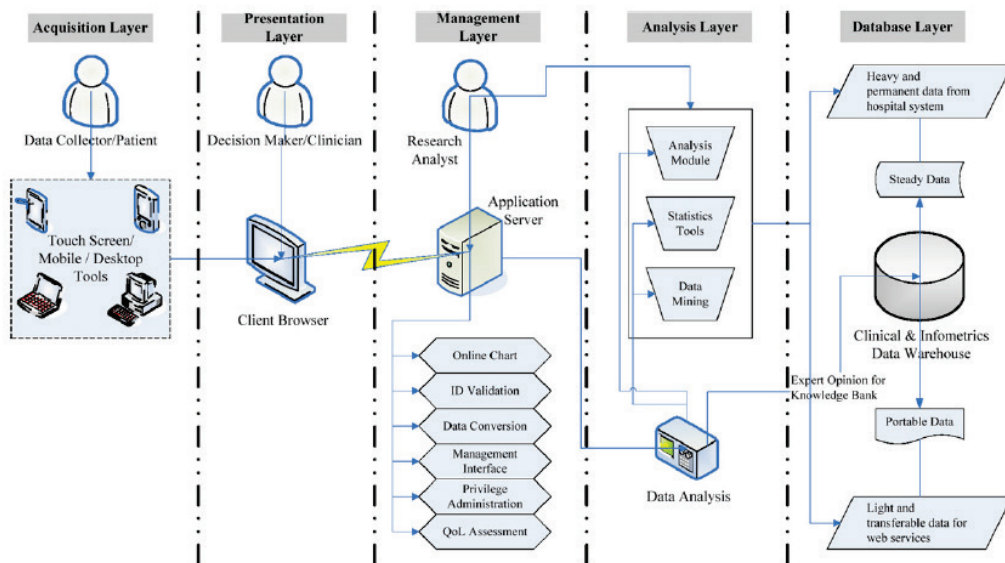
**Εικόνα 24 :** Ανάλυση καμπυλών ROC των συστημάτων υποστήριξης κλινικών αποφάσεων που χρησιμοποίησαν τα μοντέλα SVM και ΤΝΔ. Η τιμή της περιοχής κάτω από τη ROC καμπύλη (AUC) του μοντέλου SVM ήταν ανώτερη από αυτή του ΤΝΔ.

Με βάση τα παραπάνω στοιχεία, οι συγγραφείς καταλήγουν στο ότι, αν και η απόδοση του SVM αποδεικνύεται στατιστικά ανώτερη από αυτή του ΤΝΔ στην προεγχειρητική πρόβλεψη προχωρημένου καρκίνου του προστάτη, και τα δύο συστήματα, παρά τους περιορισμούς που ενυπάρχουν (όπως η έλλειψη εξωτερικής επικύρωσης και η υποκειμενική «ματιά» των αναλυτών που συνέλεξαν τα στοιχεία), αποτελούν χρήσιμα εργαλεία για τον καρκίνο του προστάτη.

**3. Ένα διαδικτυακό σύστημα αξιολόγησης πραγματικού χρόνου με μοντελοποιημένη αρχιτεκτονική που στοχεύει στην ποσοτικοποιημένη κλινική πληροφορία σε αναφορές προερχόμενες από ασθενείς με καρκίνο του προστάτη [10]**

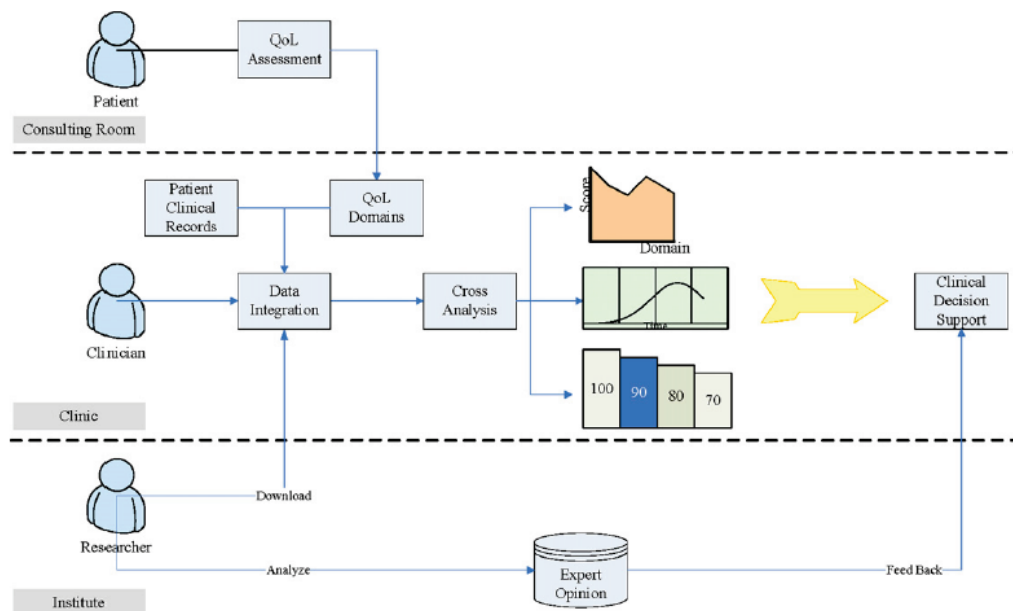
Ο σκοπός της παρούσας μελέτης ήταν η δημιουργία ενός διαδικτυακού συστήματος υποστήριξης αποφάσεων και υγείας πραγματικού χρόνου, με την ενσωμάτωση της μοντελοποιημένης αρχιτεκτονικής και των υπηρεσιών του διαδικτύου στη νέα τεχνολογία πληροφοριών για ποσοτικοποιημένες κλινικές μετρήσεις πάνω σε αναφορές προερχόμενες από ασθενείς (*patient reported outcome – PRO*) και στην ποιότητα ζωής (*quality of life – QOL*) ασθενών με καρκίνο του προστάτη. Ουσιαστικά, δεν πρόκειται για ένα υπολογιστικά «έξυπνο» σύστημα που ακολουθεί μια διαδικασία μάθησης ώστε να παράγει ένα επιθυμητό αποτέλεσμα, αλλά για ένα knowledge-based σύστημα, δηλαδή ένα υπολογιστικό πρόγραμμα που χρησιμοποιεί μια βάση δεδομένων για την επίλυση προβλημάτων.

Το περιβάλλον εργασίας που αφορά τον ασθενή επιτεύχθηκε μέσω μιας προγραμματιστικής τεχνικής μετατροπής δεδομένων που ονομάζεται “*object-relational mapping*” και μιας βάσης κλινικών δεδομένων ώστε να προκύψουν μετρήσεις της ποιότητας ζωής και ιατρικές ποσοτικοποιημένες μετρήσεις μέσω διαδικτύου, ενσωματώνοντας μια ποικιλία νοσοκομειακών πληροφοριακών συστημάτων. Τα στοιχεία που οι ασθενείς καλούνταν να παρέχουν στο σύστημα περιελάμβαναν μεταξύ άλλων τη φυσική, γνωστική και συναισθηματική τους κατάσταση, την κούραση, τον πόνο, τη ναυτία, τη σεξουαλική λειτουργία, συμπτώματα του ουροποιητικού συστήματος και του εντέρου, στοιχεία σχετικά με τη θεραπεία. Ο σχεδιασμός της υποδομής έγινε με πέντε βασικά επίπεδα ώστε να οργανωθεί η ροή των δεδομένων της απευθείας διαδικτυακής αξιολόγησης και τα κλινικά δεδομένα για υποστήριξη αποφάσεων πραγματικού χρόνου. Όπως φαίνεται στο παρακάτω σχήμα, από τα δεξιά προς τα αριστερά, το σύστημα σχεδιάστηκε ενσωματώνοντας τη βάση δεδομένων, την ανάλυση, τη διαχείριση, την παρουσίαση και τα επίπεδα πρόσβασης σε αυτό, κάτω από μια μοντελοποιημένη αρχιτεκτονική :



**Εικόνα 25 :** Ανάπτυξη πέντε επιπέδων της υποδομής του συστήματος υποστήριξης κλινικών αποφάσεων (CDSS).

Με βάση το σχεδιασμό του συστήματος, οι ιατροί και οι ερευνητές ήταν σε θέση να αποκτήσουν πρόσβαση στα δεδομένα αφότου οι ασθενείς είχαν ολοκληρώσει τις αξιολογήσεις τους στο ειδικά διαμορφωμένο περιβάλλον εργασίας, ενώ οι ιατροί θα μπορούσαν παράλληλα να συγκρίνουν το ιατρικό αρχείο του ασθενούς με άλλα που περιλαμβάνουν παρόμοιους δείκτες, μέσω διαγραμμάτων ή αριθμητικών στοιχείων. Το παρακάτω διάγραμμα απεικονίζει τη λειτουργική διαδικασία μέσω της οποίας ο ιατρός δύναται να συγκρίνει τη συνολική πληροφορία για τη θεραπεία και την ποιότητα ζωής ενός ασθενή κατά τη διάρκεια του κύκλου θεραπείας :



**Εικόνα 26 :** Εφαρμογή της κλινικής πρόδου μέσω συστήματος κλινικής ποσοτικοποιημένης πληροφορίας (clinical infometrics system).

Συμπερασματικά, θα λέγαμε ότι οι συγγραφείς διαμόρφωσαν μια προκαταρκτική τράπεζα γνώσεων που μπορεί να προσφέρει διαδικτυακή καθοδήγηση στη λήψη αποφάσεων. Παρά κάποια ζητήματα που προκαλούν προβληματισμό (όπως η εμπειρία των χρηστών στον τρόπο λειτουργίας του συστήματος ή η ηλεκτρονική ασφάλεια των προσωπικών δεδομένων), οι συγγραφείς ισχυρίζονται ότι το διαδικτυακό σύστημα που δημιούργησαν μπορεί να βοηθήσει τους κλινικούς ιατρούς να συλλέγουν και ταυτόχρονα να αναλύουν αναφορές προερχόμενες από ασθενείς (*patient reported outcome – PRO*) και την ποιότητα ζωής των ασθενών (*quality of life – QOL*) ώστε να ενισχυθεί η επικοινωνία μεταξύ ιατρών και ασθενών με καρκίνο του προστάτη και να βελτιωθεί η ποιότητα της υγειονομικής περίθαλψης.

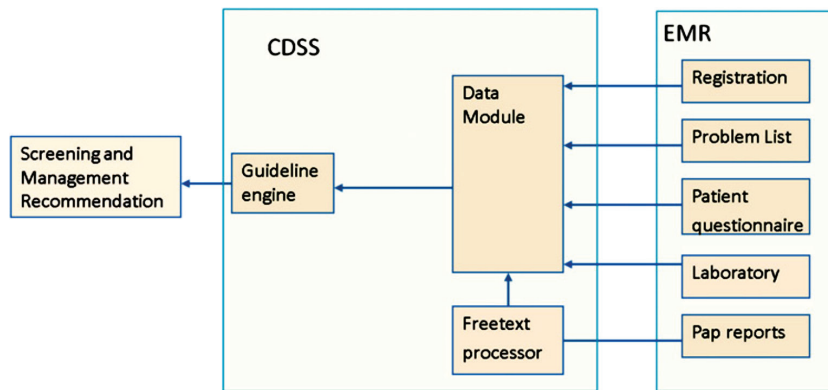
### **3.4 Μελέτες σχετικά με τον καρκίνο του τραχήλου της μήτρας (Cervical cancer)**

Στην κατηγορία που αφορά τον καρκίνο του τραχήλου της μήτρας, έχουμε συμπεριλάβει και εξετάσει τέσσερις μελέτες. Το πρώτο σύστημα που αναλύθηκε ήταν ένα knowledge-based σύστημα με στόχο τη λήψη αποφάσεων για κάθε ασθενή, το δεύτερο και τέταρτο σύστημα είχαν ως κλινικό σκοπό τη διάγνωση προκαρκινικών σταδίων ή/και καρκίνου, ενώ η τρίτη μελέτη αφορούσε την πρόβλεψη της επιβίωσης.

#### **1. Υποστήριξη κλινικών αποφάσεων με αυτοματοποιημένη επεξεργασία κειμένου για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου της μήτρας [11]**

Αντικείμενο της παρούσας μελέτης αποτέλεσε η ανάπτυξη ενός μηχανογραφημένου συστήματος υποστήριξης κλινικών αποφάσεων (*clinical decision support system, CDSS*) για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου της μήτρας που θα μπορεί να ερμηνεύσει αναφορές ελευθέρου κειμένου από το τεστ Παπανικολάου. Πρόκειται για ένα knowledge-based σύστημα, δηλαδή ένα σύστημα που παράγει αποτελέσματα βασιζόμενο σε βάσεις δεδομένων.

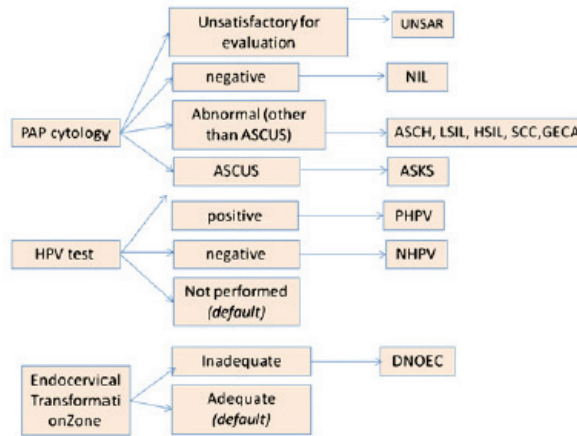
Η μελέτη αυτή διεχρήχθη στην Κλινική Mayo, ενώ το σύστημα CDSS που αναπτύχθηκε αποτελείτο από τρεις μονάδες και δύο κανόνες : έναν κανόνα ελευθέρου κειμένου για την ερμηνεία των αναφορών τεστ Παπανικολάου και έναν κανόνα κατευθυντήριων γραμμών. Η δομή του μοντέλου απεικονίζεται στο παρακάτω σχήμα :



**Εικόνα 27 :** Επισκόπηση του συστήματος υποστήριξης κλινικών αποφάσεων (CDSS). Υπάρχουν τρεις μονάδες : η μηχανή κατευθυντήριων γραμμών, η μονάδα δεδομένων και ο επεξεργαστής ελεύθερου κειμένου. Η μονάδα δεδομένων αναζητά πληροφορίες των ασθενών από το ηλεκτρονικό ιατρικό αρχείο της Κλινικής Μαγο, διατηρεί τις πληροφορίες σε μια μορφή που είναι αποδεκτή από την μηχανή δεδομένων και εξαρτάται από τον επεξεργαστή ελεύθερου κειμένου να ερμηνεύσει τις αναφορές Παπανικολάου σε ελεύθερη μορφή. Η μηχανή κατευθυντήριων γραμμών και ο επεξεργαστής ελεύθερου κειμένου είναι ουσιαστικά βασισμένα σε κανόνες.

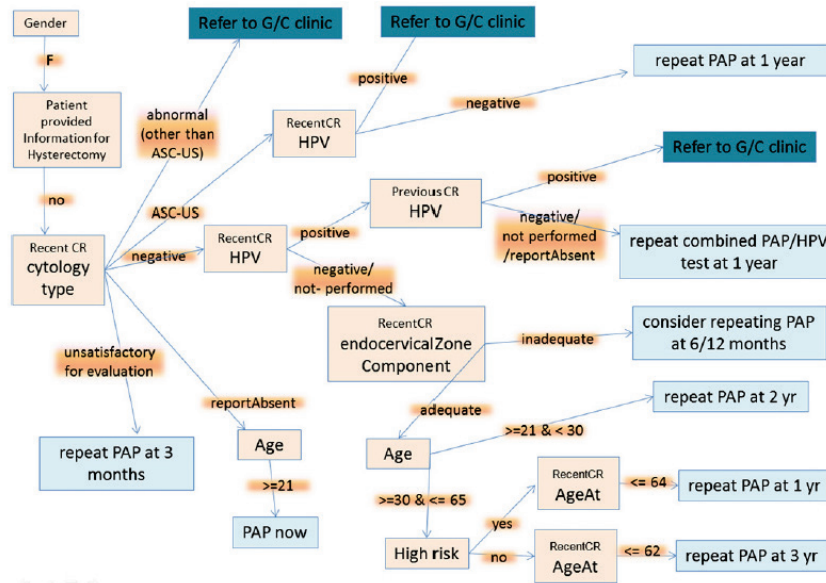
Ο επεξεργαστής ελεύθερου κειμένου περιλαμβάνει δύο στοιχεία που ενεργούν σειριακά : ένα λεκτικό αναλυτή και έναν μεταφραστή. Ο λεκτικός αναλυτής μετατρέπει το κείμενο σε σήματα και ο μεταφραστής επεξεργάζεται τα σήματα ώστε να παράγει την έξοδο. Ο κανόνας ελεύθερου κειμένου έχει, κατά συνέπεια, δύο σύνολα κανόνων : ένα λεκτικό αναλυτή και έναν μεταφραστή. Στην προκειμένη περίπτωση, οι κανόνες του λεκτικού αναλυτή αναπτύχθηκαν από την ανάλυση ενός συνόλου 49293 αναφορών τεστ Παπανικολάου από την Κλινική Μαγο στο διάστημα 2006-2011. Αναγνωρίστηκαν επαναλαμβανόμενα μοτίβα λέξεων, αντιστοιχήθηκαν στο πρότυπο του παθολογικού τμήματος για την παραγωγή των αναφορών και συντάχθηκε η παθολογική αναφορά από την μετάφραση των κωδικών με πρότυπα κειμένου. Τα αναγνωρισμένα μοτίβα λέξεων αντιστοιχήθηκαν ουσιαστικά στις έννοιες που μεταφέρονται στην αναφορά του τεστ Παπ. Στη συνέχεια, το σύνολο των κανόνων του μεταφραστή που καλείται να ερμηνεύσει τις αναγνωρισμένες έννοιες κατασκευάζεται με τη συμβολή της ιατρικής γνώμης, ακόμα και όταν αυτή υπονοείται, ενώ υπάρχουν και κανόνες για τον έλεγχο λαθών ώστε να διασφαλιστεί η λογική συνέπεια των αναφορών. Το παρακάτω σχηματικό διάγραμμα απεικονίζει συνοπτικά τον τρόπο εφαρμογής όσων αναφέρθηκαν :





**Εικόνα 28 :** Σχηματικό διάγραμμα που παρουσιάζει κάποιες από τις έννοιες στις αναφορές των τεστ Παπανικολάου οι οποίες απαιτούνται για την εφαρμογή των εθνικών κατευθυντήριων γραμμών για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου. Συντομεύσεις : ASC-H, υψηλού βαθμού πλακώδη ενδοεπιθηλιακή αλλοίωση, ASC-US, άτυπα πλακώδη κύτταρα απροσδιορίστου σημασίας (ASKS), DNOEC, συστατικό ζώνης ανεπαρκούς μετασχηματισμού, GECA, αδενική ανωμαλία των επιθηλιακών κυττάρων, HPV, ιός των ανθρωπίνων θηλωμάτων, HSIL, ενδοεπιθηλιακές βλάβες που κατηγοριοποιούνται ως υψηλού βαθμού, LSIL, ενδοεπιθηλιακές βλάβες που κατηγοριοποιούνται ως χαμηλού βαθμού, NHPV, αρνητικό τεστ HPV, PHPV, θετικό τεστ HPV, SCC, καρκίνωμα πλακωδών κυττάρων, UNSAAR, μη ικανοποιητική για την αξιολόγηση, NIL, αρνητική για ενδοεπιθηλιακή αλλοίωση ή κακοήθεια.

Ο κανόνας κατευθυντήριων γραμμών κατασκευάστηκε χρησιμοποιώντας εθνικές οδηγίες για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου που έχουν θεσπιστεί από διάφορους εθνικούς οργανισμούς (όπως η Αμερικανική Αντικαρκινική Εταιρεία, η Υπηρεσία Πρόληψης Ηνωμένων Πολιτειών κτλ). Οι εν λόγω οδηγίες ήταν διατυπωμένες σε μορφή μη υπολογιστικά ερμηνεύσιμης φυσικής γλώσσας, γι' αυτό και αναπτύχθηκε ένα διάγραμμα ροής ώστε να αναπαριστά τη γνώση που περιέχεται στις κατευθυντήριες γραμμές, το οποίο απεικονίζεται παρακάτω :



**Εικόνα 29 :** Διάγραμμα ροής για τον κανόνα κατευθυντήριων γραμμών. ASC-US, άτυπα πλακώδη κύτταρα απροσδιόριστης σημασίας, G/C, γυναικολογική κλινική, CR, έκθεση κυτταρολογίας του τραχήλου (Παπ), HPV, ανθρώπινα θηλώματα, Pap, τεστ Παπανικολάου.

Για την ολοκληρωμένη αξιολόγηση του συστήματος, το CDSS απέκτησε πρόσβαση στο ηλεκτρονικό ιατρικό αρχείο προκειμένου να παράγει τις βέλτιστες συστάσεις, εξειδικευμένες για κάθε ασθενή και για συγκεκριμένη χρονική στιγμή. Όσον αφορά την αξιολόγηση, οι συστάσεις που προήλθαν από το CDSS για 74 ασθενείς εξετάστηκαν και από έναν ιατρό, ο οποίος είχε παράγει, ανεξάρτητα από το σύστημα, αντίστοιχες προβλέψεις.

Η αξιολόγηση απέδειξε ότι το CDSS εξήγε τις βέλτιστες συστάσεις προσυμπτωματικού ελέγχου για 73 από τους 74 ασθενείς που εξετάστηκαν, αναγνώρισε δύο περιπτώσεις για γυναικολογική παραπομπή, οι οποίες παραλήφθηκαν από τον ιατρό και τον βοήθησε να τροποποιήσει τις συστάσεις του σε έξι περιπτώσεις. Η περίπτωση αποτυχίας προέκυψε επειδή η εξέταση για τον ιό των ανθρώπινων θηλωμάτων (HPV) εκτελέστηκε μερικές φορές ξεχωριστά από το τεστ Παπ και τα αποτελέσματα αυτά καταγράφηκαν από ένα εργαστηριακό σύστημα το οποίο δεν ερευνηθήκε από το CDSS. Ακολούθως, το CDSS αναβαθμίστηκε ώστε να αναζητά τα αποτελέσματα HPV που παρέλειψε στην προηγούμενη περίπτωση και παράγει τις βέλτιστες συστάσεις για όλες τις 74 περιπτώσεις που εξετάστηκαν. Η κατανομή των περιπτώσεων στα πιθανά μονοπάτια αποφάσεων παρουσιάζεται στον παρακάτω πίνακα :

History of hysterectomy	Last cytology	Last HPV	Last E2C	Age, years	High risk	Previous HPV	Recommendation	No of cases
No	Negative	Negative	Inadequate				Consider repeating Pap at 6/12 months	5
No	Negative	Positive				Report absent	Repeat combined Pap-HPV test at 12 months	4
No	Report absent			≥ 21			Pap now	3
Yes							No Pap needed	1
No	ASC-US	Negative					Repeat Pap at 12 months	9
No	Negative	Negative	Adequate	≥ 30 and ≤ 62	No		Repeat Pap at 3 years	9
No	Negative	Positive				Not performed	Repeat combined Pap-HPV test at 12 months	9
No	Unsatisfactory for evaluation						Repeat Pap at 3 months	3
No	Negative	Negative	Adequate	> 64			No Pap needed	4
No	Negative	Negative	Adequate	≥ 30 and ≤ 64	Yes		Repeat Pap at 1 year	1
No	Negative	Negative	Adequate	≥ 21 and < 30			Repeat Pap at 2 years	1
No	Negative	Not performed	Adequate	≥ 21 and < 30			Repeat Pap at 2 years	2
No	Negative	Positive				Negative	Repeat combined Pap-HPV test at 12 months	5
No	Negative	Positive				Positive	Refer to G/N	4
No	Abnormal (other than ASC-US)						Refer to G/N	6
No	Report absent				< 21		No Pap needed	3
No	ASC-US	Positive					Refer to G/N	5

Each row indicates the set of variables that correspond to a decision path in the flowchart. Some cells in the table are blank, which indicates that those variables were not required to be evaluated for the particular decision path. ASC-US, atypical squamous cells of undetermined significance; E2C, endocervical zone component; G/N, gynecology; HPV, human papillomavirus; Pap, Papanicolaou.

#### Πίνακας 14 : Κατανομή των περιπτώσεων στον πίνακα αποφάσεων.

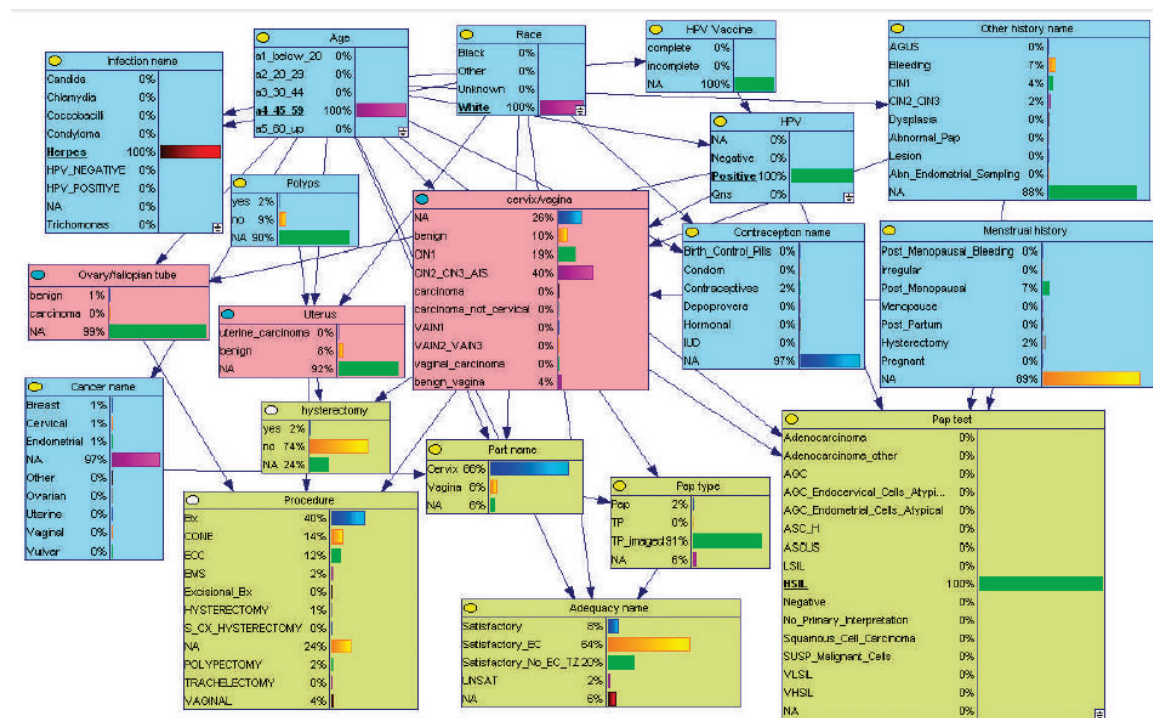
Αυτό που οι ίδιοι οι συγγραφείς αναγνωρίζουν ως περιορισμό στο σύστημά τους ήταν η ύπαρξη ενός μόνο ιδρύματος που συμπεριλήφθηκε στην μελέτη, χωρίς να είναι γνωστό το δυναμικό του CDSS σε άλλα συστήματα με διαφορετικά μοτίβα λέξεων στις αναφορές των τεστ Παπ και άλλα διαθέσιμα δεδομένα ασθενών. Επιπλέον, μόνο ένας εμπειρογνώμονας συμμετείχε στην μελέτη, γεγονός που δημιουργεί την ανάγκη για περαιτέρω επικύρωση του συστήματος και από άλλους ειδικούς.

Οι συγγραφείς υποστηρίζουν, τέλος, ότι κατασκεύασαν ένα ακριβές και αποτελεσματικό σύστημα CDSS για τον πληθυσμιακό έλεγχο του καρκίνου του τραχήλου της μήτρας, δεδομένων των τυποποιημένων εκθέσεων του τεστ Παπ και τη διαθεσιμότητα ρητών κατευθυντήριων γραμμών. Συνολικά, η μελέτη καταδεικνύει ότι το ελεύθερο κείμενο στο ηλεκτρονικό ιατρικό αρχείο μπορεί να χρησιμοποιηθεί αποτελεσματικά μέσω της επεξεργασίας φυσικής γλώσσας (*natural language processing, NLP*), ώστε να αναπτυχθούν εργαλεία υποστήριξης κλινικών αποφάσεων.

#### 2. Το μοντέλο πληθυσμιακού ελέγχου του Πίτσμπουργκ για τον καρκίνο του τραχήλου της μήτρας (The Pittsburgh Cervical Cancer Screening Model – PCCSM) : ένα εργαλείο εκτίμησης κινδύνου [12]

Η αξιολόγηση του πληθυσμιακού ελέγχου για τον καρκίνο του τραχήλου της μήτρας έχει εξελιχθεί σε εξαιρετικά πολύπλοκη διαδικασία με την καθιέρωση του εμβολιασμού για τον ιό των ανθρώπινων θηλωμάτων (*human papillomavirus – HPV*) και τις νεότερες τεχνολογίες ανίχνευσης που έχουν εγκριθεί από την Αμερικάνικη Υπηρεσία Τροφίμων και Φαρμάκων. Ο στόχος, λοιπόν, των συγγραφέων στην παρούσα μελέτη είναι η δημιουργία ενός μοναδικού μοντέλου πληθυσμιακού ελέγχου για τον καρκίνο του τραχήλου της μήτρας στο Πίτσμπουργκ (*PCCSM*), το οποίο θα προσδιορίζει ποσοτικά τον κίνδυνο των ιστοπαθολογικών προκαρκινικών σταδίων του τραχήλου (τραχηλική ενδοεπιθηλιακή νεοπλασία CIN 2, CIN 3 και επιτόπιο αδενοκαρκίνωμα *in situ*) και του καρκίνου του τραχήλου σε ένα περιβάλλον χρησιμοποιώντας κατά κύριο λόγο νέες τεχνολογίες ανίχνευσης.

Πιο αναλυτικά, το μοντέλο PCCSM είναι ένα δυναμικό Μπεϋζιανό δίκτυο (*dynamic Bayesian network*) που αποτελείται από 19 μεταβλητές οι οποίες είναι διαθέσιμες στο εργαστηριακό σύστημα πληροφόρησης και περιλαμβάνουν δεδομένα του ιστορικού του ασθενούς (ηλικία, φυλή, ιστορικό λοιμώξεων και καρκίνου, τα πιο πρόσφατα δεδομένα του εμβολιασμού για τον ιό HPV κτλ), αποτελέσματα του τεστ Παπανικολάου, αποτελέσματα υψηλού κινδύνου για τον HPV (θετικά, αρνητικά ή μη διαθέσιμα), δεδομένα χειρουργικών επεμβάσεων και ιστοπαθολογικά αποτελέσματα. Η γραφική δομή του μοντέλου, δηλαδή οι αλληλεπιδράσεις μεταξύ των διαμορφωμένων μεταβλητών, βασίστηκε στη δημοσιευμένη βιβλιογραφία, ενώ μια στατική έκδοσή του μοντέλου PCCSM απεικονίζεται στο παρακάτω σχήμα :

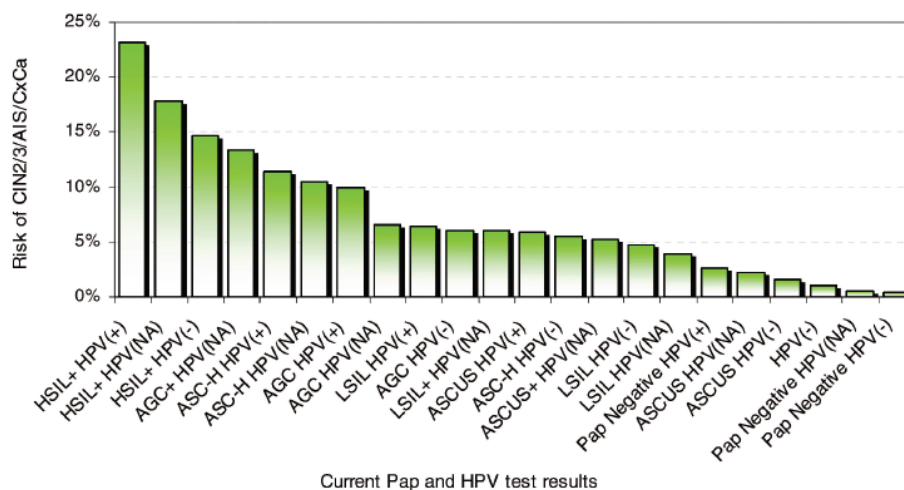


**Εικόνα 30 :** Γραφική δομή της στατικής έκδοσης του μοντέλου πληθυσμιακού ελέγχου του Πίτσμπουργκ για τον καρκίνο του τραχήλου της μήτρας (*Pittsburgh Cervical Cancer Screening Model – PCCSM*). Οι συντομεύσεις αντιστοιχούν σε παραμέτρους του μοντέλου.

Για την κατασκευή και την εκπαίδευση του μοντέλου, χρησιμοποιήθηκαν αποτελέσματα από 375441 αρχεία ασθενών στο διάστημα 2005 – 2008, ενώ για τη δοκιμή του μοντέλου χρησιμοποιήθηκαν επιπρόσθετα δεδομένα από 45930 ασθενείς που συλλέχθηκαν κατά τη διάρκεια ενός πεντάμηνου διαστήματος. Το μοντέλο δημιουργήθηκε και δοκιμάστηκε με τη χρήση μιας μηχανής εξαγωγής συμπερασμάτων και ενός περιβάλλοντος ανάπτυξης που δημιουργήθηκαν στο Εργαστήριο Συστημάτων Αποφάσεων του Πανεπιστημίου του Πίτσμπουργκ.

Όσον αφορά τα αποτελέσματά του, το μοντέλο PCCSM συγκρίνει ποσοτικά τον κίνδυνο για τα ιστοπαθολογικά επαληθεύσιμα στάδια CIN 2, CIN 3, επιτόπιου αδενοκαρκινώματος και καρκίνου του τραχήλου σε εξεταζόμενους ασθενείς για κάθε κατηγορία τρέχοντος αποτελέσματος κυτταρολογικής εξέτασης και για κάθε αποτέλεσμα HPV. Τα αποτελέσματα

των προβλέψεων του συστήματος αφορούν διαφορετικές χρονικές περιόδους, που ποικίλλουν από τη στιγμή της εξέτασης έως και 3 χρόνια. Για κάθε τρέχον αποτέλεσμα κυτταρολογικής εξέτασης, τα αποτελέσματα της εξέτασης HPV επηρεάζουν τον κίνδυνο, αν και ο βαθμός της κυτταρολογικής ανωμαλίας παραμένει ο μεγαλύτερος θετικός προγνωστικός δείκτης του κινδύνου. Το παρακάτω σχήμα δείχνει τις ποσοστικές προβλέψεις σχετικού κινδύνου στα δύο έτη :

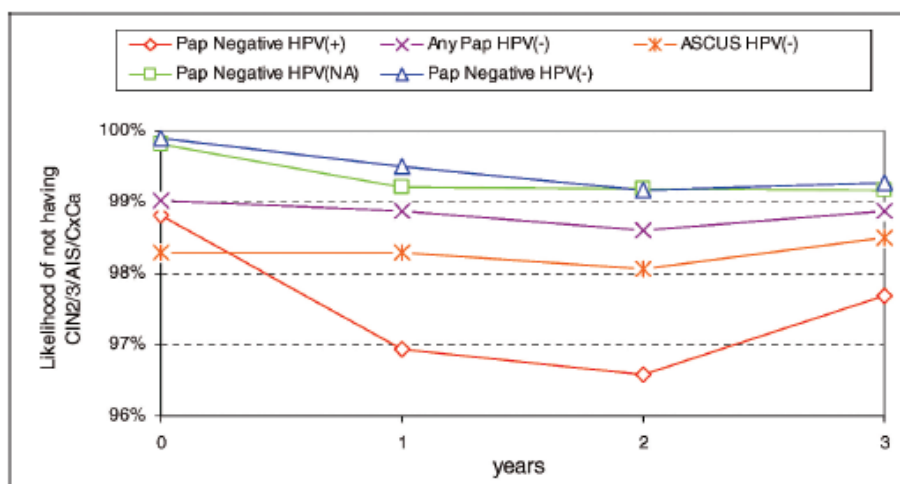


**Εικόνα 31 :** Προβλέψεις σχετικού κινδύνου για ιστοπαθολογικά προκαρκινικά στάδια ή διηθητικός καρκίνος του τραχήλου της μήτρας στα 2 έτη (ταξινομημένα από τον υψηλότερο στο χαμηλότερο κίνδυνο) με διαφορετικά τρέχοντα αποτελέσματα για τα τεστ Παπανικολάου και του ιού των ανθρωπίνων κονδυλωμάτων. Συντομεύσεις : AGC+, άτυπα αδενικά κύτταρα ή χειρότερα (AGC, ASC-H, HSIL, ύποπτα ή καρκίνος), ASC-H +, άτυπα πλακώδη κύτταρα, δεν μπορούν να αποκλειστούν HSIL ή χειρότερα (ASC-H, HSIL, ύποπτα ή καρκίνος), ASCUS +, άτυπα πλακώδη κύτταρα απροσδιορίστης σημασίας ή χειρότερα (ASCUS και όλα τα άλλα μη φυσιολογικά αποτελέσματα κυτταρολογικής εξέτασης), CIN2/CIN3/AIS/CxCa, προκαρκινικά στάδια ή διηθητικός καρκίνος του τραχήλου της μήτρας, HPV, ιός των ανθρωπίνων κονδυλωμάτων, HPV(+), θετικά αποτελέσματα για τον ιό HPV, HPV(-), αρνητικά αποτελέσματα για τον ιό HPV, HPV(NA), μη διαθέσιμα αποτελέσματα για τον ιό HPV, HSIL +, υψηλού βαθμού πλακώδης ενδοεπιθηλιακή αλλοίωση ή χειρότερα (HSIL, ύποπτη ή καρκίνος), LSIL+, πλακώδης ενδοεπιθηλιακή αλλοίωση χαμηλού βαθμού ή χειρότερα (LSIL, AGC, ASC-H, HSIL, ύποπτη ή καρκίνος), Pap, τεστ Παπανικολάου.

Στη συνέχεια, το μοντέλο PCCSM χρησιμοποιήθηκε για να αξιολογήσει τη συσχέτιση του πρότερου ιατρικού ιστορικού των ασθενών με τον κίνδυνο προκαρκινικού ή καρκινικού σταδίου. Αποδείχθηκε ότι το προγενέστερο ιστορικό μεταβάλλει τον κίνδυνο για CIN 2, CIN 3, επιτόπιο αδενοκαρκίνωμα και καρκίνο του τραχήλου για ασθενείς με συνήθη αποτελέσματα κυτταρολογικών εξετάσεων και τεστ HPV. Το μοντέλο PCCSM χρησιμοποιήθηκε επίσης για παράγει προβλέψεις αρνητικού κινδύνου με την εκτίμηση της πιθανότητας της απουσίας ιστοπαθολογικών CIN 2, CIN 3, επιτόπιου αδενοκαρκινώματος και καρκίνου του τραχήλου σε ασθενείς που υποβάλλονται στον προσυμπτωματικό έλεγχο.



Το παρακάτω σχήμα απεικονίζει τις προβλέψεις του μοντέλου για την πιθανότητα να μην υπάρχει ιστοπαθολογική διάγνωση CIN 2, CIN 3, επιτόπιου αδενοκαρκινώματος και καρκίνου του τραχήλου σε περιόδους από 1 έως 3 έτη για 5 ομάδες ασθενών με διαφορετικούς συνδυασμούς αποτελεσμάτων χαμηλού κινδύνου για τις εξετάσεις προσυμπτωματικού ελέγχου :



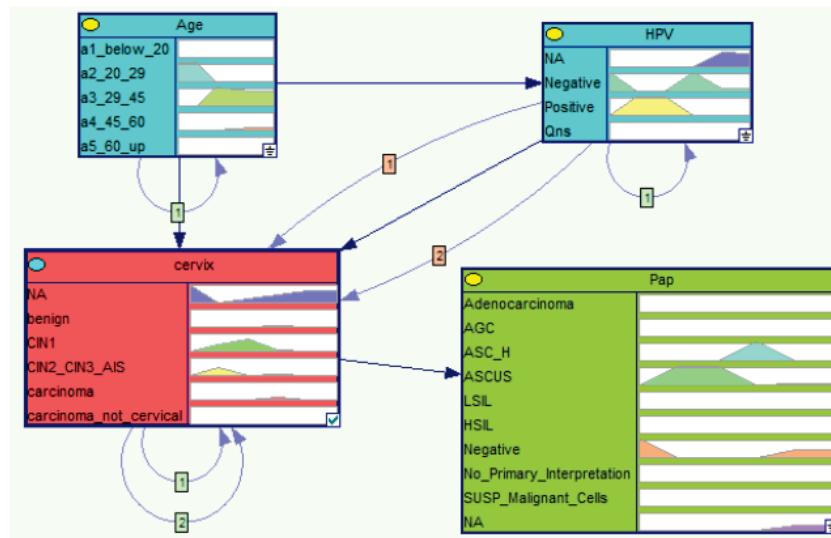
**Εικόνα 32 :** Πιθανότητα να μην υπάρχουν ιστοπαθολογικά προκαρκινικά στάδια ή διηθητικός καρκίνος του τραχήλου της μήτρας εντός 3 ετών. Συντομεύσεις : ASCUS, άτυπα πλακώδη κύτταρα απροσδιόριστης σημασίας, CIN2/3/AIS/CxCa, προκαρκινικά στάδια και διηθητικός καρκίνος του τραχήλου, HPV, ιός των ανθρωπίνων κονδυλωμάτων, HPV(+), θετικά αποτελέσματα για τον ιό HPV, HPV(-), αρνητικά αποτελέσματα για τον ιό HPV, HPV(NA), μη διαθέσιμα αποτελέσματα για τον ιό HPV, LSIL, πλακώδης ενδοεπιθηλιακή αλλοίωση χαμηλού βαθμού, Pap, τεστ Παπανικολάου.

Εν κατακλείδι, το μοντέλο PCCSM αποτελεί ένα δυναμικό Μπεϋζιανό δίκτυο που υπολογίζει ποσοτικές εκτιμήσεις του κινδύνου του καρκίνου του τραχήλου της μήτρας για ασθενείς που υποβάλλονται σε προσυμπτωματικό έλεγχο του τραχήλου. Διαρκώς ενημερωμένο σε σχέση με τα τρέχοντα δεδομένα του συστήματος, το μοντέλο PCCSM συνιστά ένα νέο εργαλείο για την παρακολούθηση του κινδύνου της νόσου του τραχήλου στην περίοδο που ακολουθεί μετά τον εμβολιασμό και ένα εξατομικευμένο μέσο βοήθειας στην κλινική διαχείριση και λήψη αποφάσεων.

Οι ίδιοι συγγραφείς προχώρησαν μεταγενέστερα σε μια νέα συμπληρωματική μελέτη με τίτλο «Εφαρμογή των δυναμικών Μπεϋζιανών δικτύων στον πληθυσμιακό έλεγχο για τον καρκίνο του τραχήλου της μήτρας» [13]. Όπως αναφέρθηκε νωρίτερα, το μοντέλο πληθυσμιακού ελέγχου του Πίτσμπουργκ για τον καρκίνο του τραχήλου της μήτρας (*Pittsburgh Cervical Cancer Screening Model – PCCSM*) [12] είναι ένα δυναμικό Μπεϋζιανό δίκτυο που συνδυάζει δύο πηγές δεδομένων : τη γνώμη των ειδικών και τα αντικειμενικά δεδομένα του νοσοκομείου, εκτιμώντας τον κίνδυνο προκαρκινικών σταδίων και διηθητικού καρκίνου του τραχήλου της μήτρας σε διάφορα χρονικά διαστήματα. Αυτές οι ποσοτικές εκτιμήσεις του κινδύνου είναι χρήσιμες για τον καθορισμό της βέλτιστης χρονικής στιγμής για επανεξέταση του ασθενούς και αποτελούν το πρώτο βήμα για τον προγραμματισμό εξατομικευμένης επαναξιολόγησης. Οι συγγραφείς υποστηρίζουν επίσης

ότι, αν και υπάρχουν πολλές δημοσιευμένες μελέτες που περικλείουν εκτιμήσεις κινδύνου προκαρκινικών σταδίων και διηθητικού καρκίνου του τραχήλου της μήτρας (Catle et al., 2007, Khan et al., 2005, Ronco et al., 2006) [13], αυτές βασίζονται δεδομένα που προέρχονται από κλινικές μελέτες και αφορούν την τρέχουσα κατάσταση του ασθενούς, ενώ δεν περιλαμβάνουν κανένα αρχείο του ιστορικού τους, σε αντίθεση με το μοντέλο PCCSM.

Όσον αφορά τη δομή του μοντέλου, οι συγγραφείς υπέθεσαν ότι κάθε ασθενής υπέφερε στο παρελθόν από μία μόνο λοίμωξη και αυτό διότι το 98% των περιπτώσεων στη βάση δεδομένων ήταν περιπτώσεις μιας μοναδικής λοίμωξης. Μια απλοποιημένη δομή του μοντέλου απεικονίζεται στο παρακάτω σχήμα :



**Εικόνα 33 :** Μία απλοποιημένη εκδοχή του μοντέλου PCCSM.

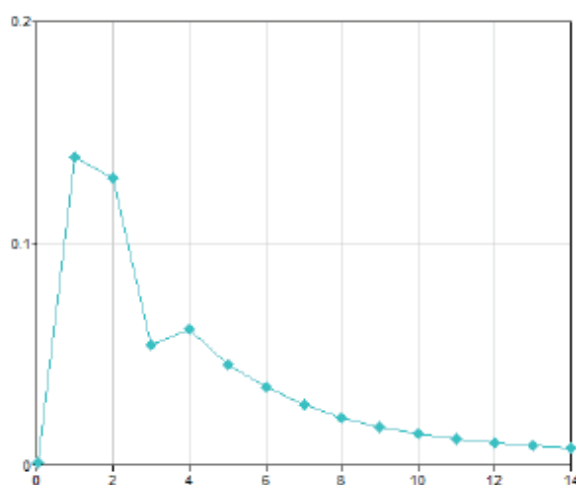
Τα δυναμικά τόξα που παρατηρούμε στο παραπάνω σχήμα αντιπροσωπεύουν αλλαγές με την πάροδο του χρόνου σε μεταβλητές όπως η ηλικία, το τεστ HPV και ο Τράχηλος. Οι μονοψήφιοι αριθμοί πάνω στα τόξα υποδηλώνουν τη χρονική καθυστέρηση της επιρροής, δηλαδή για παράδειγμα ένα τόξο που έχει επισημανθεί με το 1 (μονάδα) ανάμεσα στις μεταβλητές τεστ HPV και Τράχηλος υποδηλώνει μια επιρροή που διαρκεί ένα χρονικό βήμα. Ουσιαστικά, το μοντέλο κωδικοποιεί την ακόλουθη δεσμευμένη κατανομή σε σχέση με την μεταβλητή Τράχηλος :

$$Pr(Cervix_t | Age_t, HPV_t, HPV_{t-1}, HPV_{t-2}, Cervix_{t-1}, Cervix_{t-2})$$

Το χρονικό βήμα που επιλέχθηκε για το μοντέλο ήταν το ένα έτος, που συμβαδίζει με τις κατευθυντήριες γραμμές στις ΗΠΑ για τη συχνότητα διεξαγωγής του τεστ Παπ. Για τη διαδικασία εκπαίδευσης, χρησιμοποιήθηκαν 393531 δεδομένα ασθενών από το Πανεπιστήμιο του Πίτσμπουργκ (2005-2008), ενώ για τη διαδικασία δοκιμών χρησιμοποιήθηκαν, όπως αναφέρει η αρχική μελέτη, 45930 δεδομένα ασθενών ηλικίας από 12 έως 95 ετών, προκειμένου να εκτιμηθεί ο κίνδυνος προκαρκινικών σταδίων και διηθητικού καρκίνου του τραχήλου της μήτρας για κάθε ασθενή. Τα ελλιπή δεδομένα που

ενδέχεται να υπήρχαν λόγω της μη ολοκληρωμένης παρακολούθησης ενός ασθενή αναπαραστάθηκαν ως μια πρόσθετη κατάσταση ή παραλήφθηκαν σύμφωνα με τη διαδικασία μοντελοποίησης με Μπεϋζιανά δίκτυα.

Αυτό που προκαλεί ιδιαίτερο ενδιαφέρον είναι το παρακάτω σχήμα, το οποίο αποτυπώνει ποσοτικές εκτιμήσεις κινδύνου για προκαρκινικό στάδιο σε μια χρονική περίοδο 15 ετών για μία μόνο περίπτωση ασθενούς. Όπως φαίνεται λοιπόν, ο ασθενής διατρέχει τον μεγαλύτερο κίνδυνο προκαρκινικού σταδίου του τραχήλου ανάμεσα στο 1<sup>ο</sup> και στο 3<sup>ο</sup> έτος μετά από την αρχική εξέταση, κίνδυνος ο οποίος μειώνεται μετά από τον 4<sup>ο</sup> χρόνο :



**Εικόνα 34 :** Χρονικές εκτιμήσεις κινδύνου προκαρκινικού σταδίου του τραχήλου σε διάστημα 15 ετών.

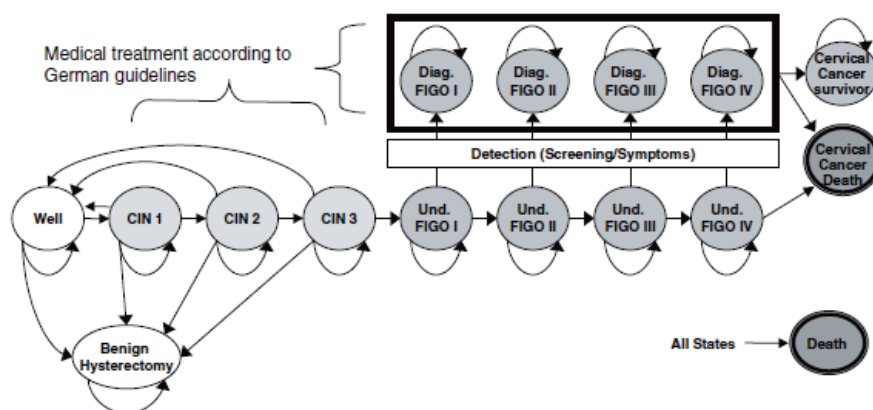
Όσον αφορά, τέλος, την επικύρωση του μοντέλου, οι συγγραφείς επικεντρώθηκαν σε μεμονωμένες περιπτώσεις ασθενών, επαληθεύοντας τις εκτιμήσεις του μοντέλου με τη διαίσθηση των εμπειρογνομόνων, αν και απαιτείται περαιτέρω επίσημη επικύρωση, κάτι που αποτελεί επόμενο στόχο των συγγραφέων. Ωστόσο, βάσει των αποτελεσμάτων, το μοντέλο αυτό μπορεί να αποτελέσει ένα χρήσιμο κλινικό εργαλείο.

**3. Το γερμανικό μοντέλο πληθυσμιακού ελέγχου για τον καρκίνο του τραχήλου της μήτρας : ανάπτυξη και επικύρωση ενός αναλυτικού μοντέλου λήψης αποφάσεων για τον πληθυσμιακό έλεγχο για τον καρκίνο του τραχήλου της μήτρας στη Γερμανία [14]**

Στην παρούσα μελέτη, οι συγγραφείς επιδίωξαν να αναπτύξουν και να επικυρώσουν ένα αναλυτικό μοντέλο λήψης αποφάσεων για τη φυσική πορεία του καρκίνου του τραχήλου της μήτρας στα πλαίσια της γερμανικής υγειονομικής περίθαλψης, καθώς και να το εφαρμόσουν στον πληθυσμιακό έλεγχο για τον καρκίνο του τραχήλου της μήτρας. Σημειώνουμε ότι, σύμφωνα με στατιστικά στοιχεία, κάθε χρόνο στη Γερμανία προσβάλλονται από καρκίνο του τραχήλου περίπου 6600 γυναίκες.



Για το λόγο αυτό, αναπτύχθηκε ένα Μαρκοβιανό μοντέλο (*Marcov model*) 16 καταστάσεων, το οποίο αντικατοπτρίζει τις τρέχουσες γερμανικές προδιαγραφές για τον πλυθησιακό έλεγχο, τη διαγνωστική παρακολούθηση (*follow-up*) και τη θεραπεία σχετικά με τον καρκίνο του τραχήλου της μήτρας και τους προπομπούς του. Το εν λόγω Μαρκοβιανό μοντέλο προσομοιώνει τη φυσική διαδικασία της καρκινογένεσης στον τράχηλο. Σύμφωνα με το μοντέλο, μια υποθετική ομάδα γυναικών ηλικίας 15 ετών και άνω κινείται σε ετήσιους κύκλους μέσα από διαφορετικές καταστάσεις υγείας με βάση τις κυτταρολογικές και ιστολογικές εξετάσεις του τραχήλου κατά τη διάρκεια της ζωής της. Οι γυναίκες ενδέχεται να παραμείνουν στην ίδια κατάσταση υγείας, να εξελιχθούν ή να επανέλθουν σε άλλη κατάσταση, όπως επίσης να πεθάνουν λόγω του καρκίνου ή από άλλα αίτια. Οι προκαρκινικές αλλοιώσεις μπορούν να ανιχνευθούν μόνο από τον προσυμπτωματικό έλεγχο, ενώ οι περιπτώσεις διεισδυτικού καρκίνου είτε από τον προσυμπτωματικό έλεγχο, ή από την ανάπτυξη των συμπτωμάτων. Οι 16 καταστάσεις της υγείας του μοντέλου απεικονίζονται και περιγράφονται αναλυτικά στο παρακάτω σχήμα :



**Εικόνα 35 :** Σχηματική απεικόνιση του Μαρκοβιανού μοντέλου. Οι 16 καταστάσεις υγείας που περιλαμβάνει είναι οι εξής : καμία αλλοίωση του τραχήλου (καλή), καλοήθης υστερεκτομή (*Benign hysterectomy*), ήπια τραχηλική ενδοεπιθηλιακή νεοπλασία (*CIN 1*), μέτρια τραχηλική ενδοεπιθηλιακή νεοπλασία (*CIN 2*), επιτόπια ενδοεπιθηλιακή νεοπλασία τραχήλου/ καρκίνωμα *in situ* (*CIN 3*), αδιάγνωστος διηθητικός καρκίνος του τραχήλου FIGO στάδια I-IV (*Und.FIGO I-IV*), διαγνωσμένος διηθητικός καρκίνος του τραχήλου FIGO στάδια I-IV (*Diag.FIGO I-IV*), επιζώντες από τον καρκίνο του τραχήλου 5 έτη μετά από τη διάγνωση και τη θεραπεία (*Cervical Cancer Survivor*), θάνατος εξαιτίας του καρκίνου του τραχήλου (*Cervical Cancer Death*) και θάνατος από άλλες αιτίες (*Death*).

Όσον αφορά την ανάπτυξη του μοντέλου, τα δεδομένα για την εξέλιξη της ασθένειας και την επιβίωση του καρκίνου του τραχήλου ελήφθησαν από τη βιβλιογραφία και τα γερμανικά μητρώα καρκίνου, ενώ η ακρίβεια του τεστ Παπανικολάου (*Pap test*) βασίστηκε σε μετα-αναλύσεις.

Οι συγγραφείς πραγματοποίησαν εσωτερική και εξωτερική επικύρωση του μοντέλου συγκρίνοντας τα αποτελέσματα του με παρατηρηθέντα επιδημιολογικά δεδομένα μη

ελεγμένων γυναικών από διαφορετικά γερμανικά μητρώα καρκίνου (δηλαδή άλλα δεδομένα χρησιμοποιήθηκαν για την εσωτερική και άλλα για την εξωτερική επικύρωση του μοντέλου). Το μοντέλο παρήγε προβλέψεις για το προσδόκιμο ζωής, την επίπτωση των διαπιστωμένων περιπτώσεων καρκίνου του τραχήλου της μήτρας, τους κινδύνους για τη διάρκεια ζωής λόγω του καρκίνου του τραχήλου και τη θνησιμότητα.

Τα αποτελέσματα συνοψίζονται ως εξής : το μοντέλο προέβλεψε κίνδυνο για τη διάρκεια ζωής λόγω του καρκίνου του τραχήλου της τάξης του 3.0% και θνησιμότητα της τάξης του 1.0%, με μέγιστη συχνότητα εμφάνισης του καρκίνου 84/100000 (δηλαδή 84 περιπτώσεις καρκίνου του τραχήλου ανά 100000 γυναίκες) στην ηλικία των 51 ετών. Τα αποτελέσματα αυτά ήταν παρόμοια με τα δεδομένα που παρατηρήθηκαν από τα γερμανικά μητρώα καρκίνου, τα στοιχεία της γερμανικής βιβλιογραφίας και αποτελέσματα από άλλα διεθνή μοντέλα. Οι προβλέψεις του μοντέλου απεικονίζονται στον παρακάτω πίνακα :

	Model prediction German Cervical Cancer Screening Model	Internal validation Observed German data: Common Cancer Registry (1964–1968)	External validation Observed German data: Cancer Registry Saarland (1968–1970), Brenner <i>et al.</i> , <sup>39</sup> Gustafsson <i>et al.</i> <sup>9</sup>	Cross-model validation International literature data: Myers <i>et al.</i> <sup>15</sup>
Peak age (years) CIN				
CIN 1	25			28
CIN 2	38			42 <sup>a</sup>
CIN 3	48			42 <sup>a</sup>
Peak age (years) cervix carcinoma	51	55–59	47; 53	48
Peak incidence cervix carcinoma (per 100000)	84	112	107–112	81
Total incidence cervix carcinoma (per 100000)	46	46		
Distribution cervix carcinoma FIGO stages (%)				
FIGO I	38.8	38.8	38	46.4
FIGO II	31.6	31.6	27	27.0
FIGO III	24.1	24.1	35 <sup>b</sup>	18.1
FIGO IV	5.45	5.44		8.5
Lifetime risk cervix carcinoma (%)	3.0		3.3 <sup>c</sup>	2.5

a: CIN 2/3 aggregated data  
b: FIGO III + IV aggregated data  
c: Women until age 85 years

**Πίνακας 15 :** Προβλέψεις του μοντέλου και τιμές επικύρωσης για την εμφάνιση του καρκίνου του τραχήλου, κατανομή σε στάδια FIGO και κίνδυνος για τη διάρκεια ζωής λόγω του καρκίνου του τραχήλου.

Το μοντέλο εφαρμόστηκε επίσης προκειμένου να αξιολογηθούν τα μακροπρόθεσμα θετικά κλινικά αποτελέσματά του ελέγχου μέσω τεστ ΠΑΠ. Πράγματι, αποδείχθηκε ότι ο ετήσιος πληθυσμιακός έλεγχος μέσω του τεστ ΠΑΠ θα μπορούσε να αποτρέψει το 98.7% των διαγιγνωσκόμενων κρουσμάτων καρκίνου και το 99.6% των θανάτων εξαιτίας του καρκίνου του τραχήλου σε γυναίκες που συμμορφώθηκαν πλήρως με τον προσυμπτωματικό έλεγχο και τη θεραπεία. Ωστόσο, η παράταση του διαστήματος προσυμπτωματικών ελέγχων από ένα σε δύο, τρία ή πέντε έτη είχε ως αποτέλεσμα την

μειωμένη αποτελεσματικότητα του προσυμπτωματικού ελέγχου, γεγονός που αποδεικνύουν τα στοιχεία του παρακάτω πίνακα :

	No screening	Pap screening			
		1 year	2 years	3 years	5 years
Absolute cases per 100 000					
Cervical cancer cases	3032	38	180	310	699
Cervical cancer deaths	1004	4	23	53	128
Prevented cases per 100000 <sup>a</sup>					
Cervical cancer cases	–	2994	2851	2676	2333
Cervical cancer deaths	–	1000	981	951	876
Life expectancy					
Absolute life years	65.80	66.06	66.06	66.05	66.02
Life days gained <sup>a</sup>	–	94.2	91.7	88.0	79.5

a: Versus no screening

**Πίνακας 16 :** Απόλυτες περιπτώσεις και περιπτώσεις και θάνατοι λόγω του καρκίνου του τραχήλου που θα μπορούσαν να έχουν προληφθεί και το προσδόκιμο ζωής με τον προσυμπτωματικό έλεγχο τεστ ΠΑΠ σε χρονικά διαστήματα 1, 2, 3 και 5 ετών.

Οι συγγραφείς υποστηρίζουν, τέλος, ότι το μοντέλο αυτό μπορεί να αποτελέσει ένα εργαλείο για την αξιολόγηση της μακροπρόθεσμης αποτελεσματικότητας διαφορετικών προσυμπτωματικών ελέγχων και στρατηγικών για τον καρκίνο του τραχήλου της μήτρας. Ωστόσο, το μοντέλο θέτει κάποιους περιορισμούς λόγω της αβεβαιότητας και της διαθεσιμότητας ή μη τιμών των παραμέτρων, της προσαρμογής διεθνών μελετών στα επιδημιολογικά δεδομένα της Γερμανίας και της ετερογένειας των χαρακτηριστικών του πληθυσμού, η οποία δεν αποτυπώνεται πλήρως (για παράδειγμα δε λήφθηκε υπόψη η ετερογένεια του HPV στον πληθυσμό). Παρόλα αυτά, το Γερμανικό Μοντέλο Πληθυσμιακού Ελέγχου για τον Καρκίνο του Τραχήλου της Μήτρας δύναται να αναπτύξει κατευθυντήριες γραμμές για την αξιολόγηση νέων τεχνολογιών, όπως το τεστ HPV ή ο εμβολιασμός κατά του HPV.

#### **4. Προσδιορισμός των γυναικών για παραπομπή σε κολποσκόπηση μέσω Νευρωνικών Δικτύων : Μία προκαταρκτική μελέτη που βασίζεται στην κυτταρολογική εξέταση υγρών (liquid-based cytology – LBC) και στους μοριακούς βιοδείκτες [15]**

Στόχος της παρούσας μελέτης ήταν η διερεύνηση του δυναμικού ενός ταξινομητή νευρωνικού δικτύου με εκπαιδευόμενο καταμεριστή διανυσμάτων (*learning vector quantizer neural network, LVQ-NN*), όσον αφορά διάφορες διαγνωστικές μεταβλητές που χρησιμοποιούνται στο σύγχρονο εργαστήριο κυτταροπαθολογίας, καθώς και η κατασκευή ενός αλγορίθμου που μπορεί να διευκολύνει την ταξινόμηση μεμονωμένων περιπτώσεων για προκαρκινικά στάδια ή καρκίνο του τραχήλου της μήτρας.

Όπως υποστηρίζουν οι συγγραφείς, μέχρι και σήμερα εκλείπουν ακριβείς μέθοδοι διαλογής (*triage*) των ασθενών που παρουσιάζουν κυτταρολογικές εξετάσεις τόσο χαμηλού, όσο και υψηλού βαθμού. Για το λόγο αυτό, πιστεύεται ότι ανερχόμενες τεχνολογίες και βιολογικοί δείκτες όπως ο προσδιορισμός του γονοτύπου HPV DNA, οι εξετάσεις E6&7 mRNA και η ανοσοχρώση της πρωτεΐνης p16 ενδέχεται να βελτιώσουν την ακρίβεια της τελικής διάγνωσης σε σχέση με αυτή που προέρχεται μόνο από τα κυτταρολογικά αποτελέσματα.

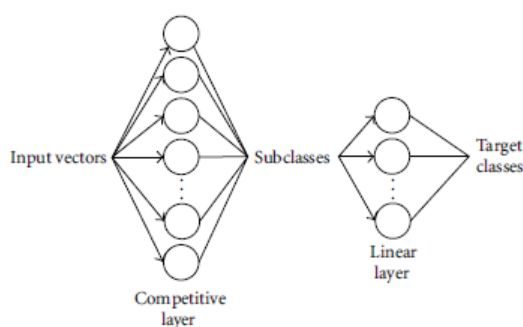
Η πολυκεντρική διαγνωστική μελέτη που διεξήγαγαν οι συγγραφείς έλαβε χώρα στο Αττικό Πανεπιστημιακό Νοσοκομείο και στο Πανεπιστημιακό Νοσοκομείο Ιωαννίνων στο διάστημα 2007-2010. Από τις 1258 γυναίκες που συμμετείχαν στην μελέτη, λήφθηκε ένα δείγμα κυτταρολογικής εξέτασης, ένα μέρος του οποίου υποβλήθηκε σε κυτταρολογική αξιολόγηση και το υπόλοιπο υποβλήθηκε σε τεστ HPV DNA, E6/E7 HPV mRNA και ανοσοκυτταρικής έκφρασης της πρωτεΐνης p16 (αναφέρουμε συνοπτικά πως έχει αποδειχθεί ότι η κατάσταση έκφρασης της πρωτεΐνης p16 σχετίζεται με τη δυνατότητα των ανθρώπινων θηλωμάτων να προκαλέσουν αλλοιώσεις στον τράχηλο της μήτρας). Τα δεδομένα, αφού επεξεργάστηκαν πριν την είσοδό τους στο μοντέλο LVQ-NN, ταξινομήθηκαν σε δύο κατηγορίες : CIN 2 ή χειρότερα (μέτρια τραχηλική ενδοεπιθηλιακή νεοπλασία – θετικό γκρουπ) και CIN 1 ή καλύτερα (ήπια τραχηλική ενδοεπιθηλιακή νεοπλασία – αρνητικό γκρουπ). Οι μισές περιπτώσεις χρησιμοποιήθηκαν για την εκπαίδευση του LVQ-NN, ενώ οι υπόλοιπες περιπτώσεις (σύνολο δοκιμών) χρησιμοποιήθηκαν για την επικύρωση του. Τα δεδομένα που χρησιμοποιήθηκαν για την μελέτη απεικονίζονται στον παρακάτω πίνακα :

Variable name	Variable description	Variable value range
Cytological diagnosis	The result of the cytological examination expressed according to Bethesda system	6: SCC or ADENOCA, 5: HSIL, 4: ASC-H, 3: LSIL, 2: ASCUS, 1: WNL
Binary HPV test	The result of the HPV DNA test	Has a value of 0 if no HPV subtype was found and 1 if there was found even a single subtype
A6, A11, A16, A18, A26, A31, A33, A35, A39, A40, A42, A43, A44, A45, A51, A52, A53, A54, A56, A58, A59, A61, A62, A66, A68, A70, A71, A72, A73, A81, A82, A83, A84, A85, A89	The existence of individual subtypes according to the HPV typing examination	For each individual subtype: 1 if it was found and 0 if it was not found
Number of HPV subtypes	The number of HPV subtypes found by the HPV DNA test	Expressed as number of identified subtypes
High-risk HPV subtypes	The number of high risk and very high risk subtypes found by the HPV DNA test	Expressed as number of identified high risk or very high risk subtypes
N16	The result of the E6/E7 mRNA test for HPV subtype 16	0 if negative and 1 if positive
N18	The result of the E6/E7 mRNA test for HPV subtype 18	0 if negative and 1 if positive
N31	The result of the E6/E7 mRNA test for HPV subtype 31	0 if negative and 1 if positive
N33	The result of the E6/E7 mRNA test for HPV subtype 33	0 if negative and 1 if positive
Flow test result	The result of the identification of E6/E7 mRNA expression of high risk HPV using FLOW cytometry technique	0 if negative (positive expression <1.5%) and 1 if positive (positive expression >1.5)
p16 result	The result of the P16 immunocytochemical examination	0 if negative and 1 if positive
Histological examination	The outcome of the histological examination	1 for positive or 0 for negative

**Πίνακας 17 :** Οι μεταβλητές που χρησιμοποιήθηκαν για την εκπαίδευση του ταξινομητή LVQ-NN.

Όσον αφορά τον τρόπο λειτουργίας του ταξινομητή LVQ-NN, τα διαθέσιμα δεδομένα χωρίζονται στα σύνολα εκπαίδευσης και επικύρωσης, ενώ η κατηγορία στην οποία ανήκουν είναι γνωστή εκ των προτέρων. Κατά τη φάση εκπαίδευσης, τα δεδομένα χρησιμοποιούνται

μαζί με την κατανεμημένη κλάση τους και ο ταξινομητής εκπαιδεύεται από το συγκεκριμένο σύνολο δεδομένων. Αρχικά, ο ταξινομητής LVQ-NN δημιουργεί κατατμήσεις στο χώρο χαρακτηριστικών, κάθε μία από τις οποίες χαρακτηρίζεται από ένα διάνυσμα στο κέντρο της (codebook vector), η κλάση του οποίου χαρακτηρίζει την κλάση ολόκληρης της κατάτμησης. Ένα πέρασμα όλων των διανυσμάτων εκπαίδευσης στον ταξινομητή αρχικοποιεί τα διανύσματα codebook, εφαρμόζοντας με αυτόν τον τρόπο τον αλγόριθμο εκπαίδευσης (απαιτούνται 50-200 περάσματα). Κατά τη διάρκεια της φάσης δοκιμών, αξιολογείται το εκπαιδευόμενο δίκτυο LVQ. Κάθε άγνωστη περίπτωση που παρουσιάζεται στο δίκτυο αναπαριστάται από ένα διάνυσμα δεδομένων και εκτιμάται ποιο διάνυσμα codebook βρίσκεται πλησιέστερα σε αυτό το διάνυσμα δεδομένων, ούτως ώστε η κατάτμησή του να θεωρηθεί της ίδιας κλάσης με την κατάτμηση του διανύσματος codebook. Το σχηματικό διάγραμμα της δομής του LVQ-NN δίνεται από το παρακάτω σχήμα :



**Εικόνα 36 :** Το σχηματικό διάγραμμα της δομής του LVQ-NN.

Για την εκτίμηση της απόδοσης του LVQ-NN, υπολογίστηκαν παράμετροι όπως η ευαισθησία, η ειδικότητα, η θετική και αρνητική προγνωστική αξία, καθώς και η συνολική ακρίβεια σε σχέση με την κυτταρολογική εξέταση με σύγκριση των περιοχών κάτω από την καμπύλη ROC. Από τις 1258 περιπτώσεις που αναλύθηκαν, η κυτταρολογική εξέταση αναγνώρισε σωστά το 72.90% των περιπτώσεων με CIN 2 ή χειρότερα και το 97.37% των περιπτώσεων με CIN 1 ή καλύτερα, με συνολική ακρίβεια της τάξης του 94.36%. Αντιστοίχως, η εφαρμογή του LVQ-NN στο σύνολο δοκιμών επέτρεψε σωστή ταξινόμηση για το 84.62% των περιπτώσεων με CIN 2 ή χειρότερα και το 97.64% των περιπτώσεων με CIN 1 ή καλύτερα, με συνολική ακρίβεια 96.03%. τα συγκεντρωτικά αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα :

	LVQ training set	LVQ test set	LVQ training and test sets	Cytology ASC-H+ versus CIN2+	Cytology ASCUS+ versus CIN2+
Sensitivity	85.71%	84.62%	85.16%	72.90%	97.42%
Specificity	98.37%	97.64%	98.01%	97.37%	76.97%
PPV	88.00%	83.54%	85.71%	79.58%	37.28%
NPV	98.01%	97.82%	97.92%	96.24%	99.53%
FPR	1.63%	2.36%	1.99%	2.63%	23.03%
FNR	14.29%	15.38%	14.84%	27.10%	2.58%
OA	96.82%	96.03%	96.42%	94.36%	79.49%

**Πίνακας 18 :** Απόδοση του ταξινομητή LVQ-NN για τα σύνολα εκπαίδευσης, δοκιμών και στο σύνολο και κυτταρολογική απόδοση. Συντομώσεις και χρήσιμες έννοιες : Sensitivity – ευαισθησία, Specificity – ειδικότητα, PPV – θετική προγνωστική αξία, NPV – αρνητική προγνωστική αξία, OA – συνολική ακρίβεια.

Για την περαιτέρω αξιολόγηση της αξιοπιστίας του συστήματος ακολούθησαν πέντε νέα πειράματα, κατά τα οποία η βάση δεδομένων διαιρέθηκε εκ νέου σε «νέα» σύνολα εκπαίδευσης και επικύρωσης, ενώ η σταθερότητα της απόδοσης των «νέων» πέντε ταξινομητών εκτιμήθηκε κάθε φορά μετά από επανεκπαίδευση.

Αποδείχθηκε ότι το LVQ-NN παρουσιάζει απόδοση ανώτερη της κυτταρολογικής εξέτασης για την ανίχνευση υψηλού βαθμού αλλοιώσεων σε μικτό πληθυσμό (γυναίκες που υποβάλλονται σε πληθυσμιακό έλεγχο και κολποσκόπηση). Συνεπώς, η χρήση του μαζί με την κυτταρολογική εξέταση και τους προτεινόμενους βιοδείκτες βελτιώνει σημαντικά τη σωστή ταξινόμηση των προκαρκινικών αλλοιώσεων του τραχήλου της μήτρας και/ή του καρκίνου και μπορεί να διευκολύνει τη διάγνωση και την αντιμετώπιση των ασθενών. Οι συγγραφείς εκτιμούν επίσης ότι η απόδοση του προτεινόμενου συστήματος θα ήταν υψηλότερη αν συνδυάζονταν οι προτεινόμενοι βιοδείκτες που σχετίζονται με το HPV με άλλα επιδημιολογικά δεδομένα (δημογραφικά κτλ) και με αποτελέσματα κολποσκόπησης και κυτταρολογικής εξέτασης. Στόχοι τους σε μελλοντικές μελέτες αποτελούν η εκτίμηση της απόδοσης του LVQ-NN σε έναν αυστηρά καθορισμένο πληθυσμό αποτελούμενο από γυναίκες με επιχρίσματα LSIL (χαμηλού βαθμού ενδοεπιθηλιακών αλλοιώσεων πλακώδους), συνδυασμός των βιοδεικτών με άλλους δείκτες για μείωση του κόστους, μελέτη της σχέσης αποτελεσματικότητας – κόστους και σαφέστατα εκτίμηση και επικύρωση των αποτελεσμάτων του μοντέλου σε μεγαλύτερες βάσεις δεδομένων.

### 3.5 Μελέτες σχετικά με τον καρκίνο του πνεύμονα (Lung Cancer)

Όσον αφορά τον καρκίνο του πνεύμονα, μελετήθηκε ένα σύστημα τεχνητού νευρωνικού δικτύου με στόχο την πρόβλεψη της επιβίωσης, ενώ τα αποτελέσματά του συγκρίθηκαν με αυτά της λογιστικής παλινδρόμησης. Ο τίτλος της μελέτης που εξετάστηκε ήταν : «Εφαρμογή προγνωστικών μοντέλων σε ασθενείς με μη μικροκυτταρικό καρκίνο του πνεύμονα με χρήση τεχνητών νευρωνικών δικτύων σε σύγκριση με την μέθοδο λογιστικής παλινδρόμησης» [16].



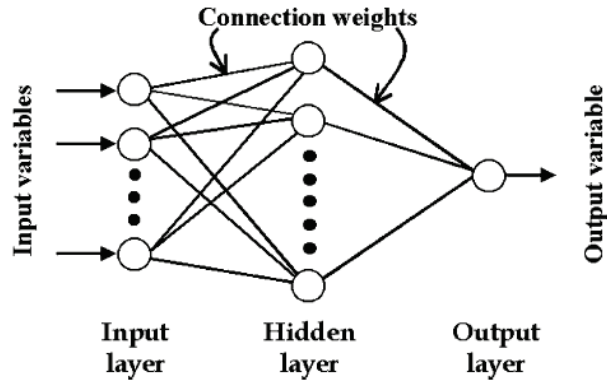
Όπως υποστηρίζουν οι συγγραφείς, ο καρκίνος του πνεύμονα (μικροκυτταρικός καρκίνος του πνεύμονα και μη μικροκυτταρικός καρκίνος του πνεύμονα) αποτελεί την κύρια αιτία θανάτου οφειλόμενο σε καρκίνο τόσο στην Ιαπωνία, όπου λαμβάνει χώρα η παρούσα μελέτη, όσο και στις χώρες του δυτικού κόσμου. Ωστόσο, είναι δύσκολο να προβλεφθεί με ακρίβεια η έκβαση κάθε μεμονωμένου ασθενή με μη μικροκυτταρικό καρκίνο του πνεύμονα (non-small-cell lung cancer/NSCLC) χρησιμοποιώντας συμβατικές στατιστικές μεθόδους, όπως η λογιστική παλινδρόμηση, και συνήθεις κλινικο-παθολογικές μεταβλητές, διότι ενυπάρχουν περιορισμοί. Στόχος, λοιπόν, της μελέτης αυτής είναι η πρόβλεψη της πρόγνωσης μεμονωμένων ασθενών, με χρήση τεχνητών νευρωνικών δικτύων (artificial neural networks – ANN), ώστε να επιτευχθεί καλύτερη κλινική διαχείριση των ασθενών με βάση τον αναμενόμενο κίνδυνο.

Για το λόγο αυτό, οι συγγραφείς κατασκεύασαν ένα προγνωστικό μοντέλο που συμπεριέλαβε 125 ασθενείς με NSCLC, με μέσο όρο ηλικίας 63,3 έτη, οι οποίοι υποβλήθηκαν σε δυνητικά θεραπευτική εκτομή από το 1986 έως το 1988. Επιπλέον, χρησιμοποιήθηκαν 48 ασθενείς που υποβλήθηκαν σε εγχείριση το 1996 για την επικύρωση του μοντέλου. Το μοντέλο συμπεριέλαβε 17 πιθανές μεταβλητές εισόδου, εκ των οποίων οι 12 είναι κλινικο-παθολογικές μεταβλητές (ηλικία, φύλο, δείκτης καπνίσματος, μέγεθος του όγκου, παράγοντας  $p$  – υπεζωκοτική συμμετοχή,  $pT$ ,  $pN$ , παθολογικό στάδιο, ιστολογικός τύπος σύμφωνα με τη νέα ταξινόμηση του Παγκόσμιου Οργανισμού Υγείας), ενώ οι υπόλοιπες 5 αποτελούν ανοσοϊστοχημικές μεταβλητές (ποσοστό  $p27$ , ένταση  $p27$ ,  $p53$ , κυκλίνη D1, ρετινοβλάστωμα (RB)) :

Input variables	Data type	Value
Age	Continuous	36–82
Sex	Categorical	1: male, 2: female
Smoking index	Continuous	0–2350 cigarettes×year
Tumor size	Continuous	36–110 mm
p factor	Categorical	1: p0, 2: p1, 3: p2, 4: p3
T factor	Categorical	1: T1, 2: T2, 3: T3, 4: T4
N factor	Categorical	1: N0, 2: N1, 3: N2
Pathological stage	Categorical	1: I, 2: II, 3: IIIa, 4: IIIb
Histological types adeno	Categorical	1: yes, 2: no
Histological types squamous	Categorical	1: yes, 2: no
Histological types adenosquamous	Categorical	1: yes, 2: no
Histological types large	Categorical	1: yes, 2: no
p27 percentage	Categorical	1: <5%, 2: 5–30%, 3: 31–60%, 4: 61%<
p27 intensity	Categorical	1: negative, 2: decreased, 3: normal, 4: increased
p53	Categorical	1: negative, 2: positive
Cyclin D1	Categorical	1: negative, 2: positive
RB	Categorical	1: negative, 2: positive

**Πίνακας 19 :** Πίνακας μεταβλητών εισόδου του προγνωστικού μοντέλου.

Αξίζει να σημειωθεί ότι στην παρούσα μελέτη χρησιμοποιήθηκαν μεταβλητές, όπως η κυκλίνη, το ρετινοβλάστωμα κ.ά., οι οποίες από προηγούμενη μελέτη των συγγραφέων αποδείχθηκαν σημαντικοί προγνωστικοί παράγοντες. Το νευρωνικό δίκτυο ANN που χρησιμοποιήθηκε συμπεριέλαβε τρία επίπεδα (εισόδου, εξόδου και κρυφό) και σχεδιάστηκε όπως απεικονίζεται παρακάτω :



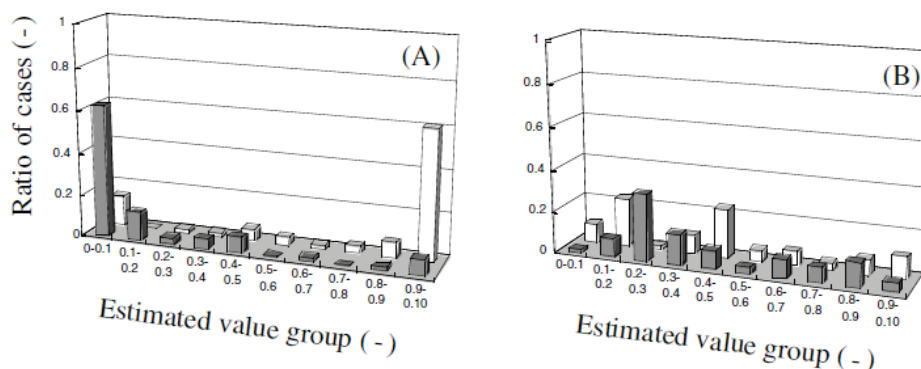
**Εικόνα 37 :** Τεχνητό νευρωνικό δίκτυο.

Το επίπεδο εξόδου αποτελείται από μία μόνο μονάδα, που αντιπροσωπεύει την επιβίωση ή μη των ασθενών, ενώ αρχικά στα βάρη σύνδεσης αντιστοιχήθηκαν τυχαία τιμές μεταξύ του 0 και του 1. Για τη βελτιστοποίηση του μοντέλου, κρίθηκε απαραίτητη η μείωση του αριθμών των επιπέδων εισόδου και κρυφού, χρησιμοποιώντας την μέθοδο αύξησης παραμέτρων (parameter-increasing method/PIM).

Στη συνέχεια, προκειμένου να εξεταστεί η ευελιξία του νευρωνικού δικτύου, διεξήχθη διασταυρούμενη επικύρωση, όπου τα σύνολα δεδομένων χωρίστηκαν σε 5 ομάδες, οι οποίες χρησιμοποιήθηκαν όλες με τον ίδιο τρόπο ως δεδομένα αξιολόγησης για να αξιολογηθεί η ακρίβεια και η αναλογία απόφασης (judgment ratio - JR). Παράλληλα, κατασκευάστηκε ένα συμβατικό στατιστικό μοντέλο με χρήση λογιστικής παλινδρόμησης για τον έλεγχο του νευρωνικού δικτύου.

Χρησιμοποιώντας το προκύπτον μοντέλο ANN, ήταν δυνατή η πρόβλεψη σε 104 από τους 125 ασθενείς (83% αναλογία απόφασης (judgment ratio - JR) και ακρίβεια πρόβλεψης της επιβίωσης για 5 χρόνια ίση με 87%). Αντίστοιχα, η αναλογία απόφασης JR και η ακρίβεια πρόβλεψης επιβίωσης στο μοντέλο λογιστικής παλινδρόμησης (LR) ήταν 37% και 78% αντίστοιχα. Επιπλέον, το τεχνητό νευρωνικό δίκτυο επέδειξε καλύτερες επιδόσεις στην πρόβλεψη της επιβίωσης για 1 ή 3 έτη, σε σχέση με τη λογιστική παλινδρόμηση. Σε αυτές τις περιπτώσεις, η μέθοδος αύξησης παραμέτρων PIM επέλεξε ως μεταβλητές την ένταση p27 και την κυκλίνη D1 για το τριετές μοντέλο επιβίωσης και την μεταβλητή p53 για το μονοετές μοντέλο επιβίωσης, πέραν από τις κλινικοπαθολογικές μεταβλητές. Τελικά, ακόμα και στο ανεξάρτητο σύνολο δεδομένων επικύρωσης των 48 ασθενών που υποβλήθηκαν σε χειρουργική επέμβαση δέκα χρόνια αργότερα, το παρόν μοντέλο ANN κατάφερε να προβλέψει την έκβαση των ασθενών στα πέντε χρόνια με αναλογία απόφασης JR και ακρίβεια πρόβλεψης ίσες με 81% και 77% αντίστοιχα. Τα παρακάτω γραφήματα παρουσιάζουν μια άμεση σύγκριση των αποτελεσμάτων του νευρωνικού δικτύου ANN και του μοντέλου λογιστικής παλινδρόμησης LR όσον αφορά την πενταετή επιβίωση των ασθενών :





**Εικόνα 38 :** Αποτελέσματα του μοντέλου 5-ετούς επιβίωσης με χρήση ANN (A) και LR (B). Τα σκιασμένα κουτιά υποδεικνύουν ασθενείς που έχουν επιβιώσει, ενώ τα κενά κουτιά ασθενείς που έχουν αποβιώσει.

Παραθέτοντας αποτελέσματα από παλαιότερες μελέτες πάνω σε ασθενείς με NSCLC, οι συγγραφείς ισχυρίζονται ότι όχι μόνο τα τεχνητά νευρωνικά δίκτυα αποτελούν δυναμικά χρησιμοποιότερο εργαλείο από τις συμβατικές στατιστικές μεθόδους πρόβλεψης της επιβίωσης των ασθενών με NSCLC, αλλά και ότι η συμπερίληψη των σχετικών μοριακών δεικτών ως μεταβλητών εισόδου στο παρόν μοντέλο (γεγονός που διαφοροποιεί την μελέτη αυτή από άλλες αντίστοιχες) ενισχύει την προβλεπτική του ικανότητα.

### 3.6 Μελέτες σχετικά με τον καρκίνο του παχέος εντέρου (Colorectal cancer)

Στην κατηγορία που αφορά τον καρκίνο του παχέος εντέρου συμπεριλήφθηκαν δύο μελέτες συστημάτων που περιλαμβάνουν τεχνητά νευρωνικά δίκτυα με στόχο την πρόβλεψη της επιβίωσης.

1. Τεχνητά νευρωνικά δίκτυα που εφαρμόζονται για την πρόγνωση αποτελέσματος σε ασθενείς με καρκίνο του παχέος εντέρου από διαφορετικά ιατρικά κέντρα [17]

Επί του παρόντος, παραμένει εξαιρετικά δύσκολο να προβλεφθεί το προσδόκιμο ζωής ενός μεμονωμένου ασθενή που έχει χειρουργηθεί για καρκίνο του παχέος εντέρου. Οι μόνες διαθέσιμες τεχνικές εκτίμησης της επιβίωσης βασίζονται σε μεγάλες ομάδες εντός του πληθυσμού ως σύνολο (δηλαδή στατιστικά στοιχεία πληθυσμού). Ωστόσο, η μέση επιβίωση του πληθυσμού δεν αποτελεί χρήσιμο προγνωστικό παράγοντα για έναν μεμονωμένο ασθενή, καθώς αυτός ενδέχεται να αποκλίνει από συνήθη συστήματα ταξινόμησης (για παράδειγμα, το σύστημα ταξινόμησης για τον καρκίνο παχέος εντέρου του Βρετανού παθολόγου Dukes που ορίζει 4 στάδια καρκίνου με διαφορετικά ποσοστά πενταετούς επιβίωσης, από το στάδιο Dukes' A – 90% πενταετής επιβίωση έως το στάδιο Dukes' D - <5% πενταετής επιβίωση δεν ενδείκνυται για όλες τις περιπτώσεις). Για το λόγο αυτό, οι συγγραφείς εφάρμοσαν τεχνητά νευρωνικά δίκτυα για την ανάλυση δεδομένων

ασθενών με καρκίνο του παχέος εντέρου στην προσπάθειά τους να επιτύχουν ακριβείς προβλέψεις επιβίωσης για μεμονωμένα περιστατικά.

Οι συγγραφείς αποφάσισαν να χρησιμοποιήσουν ένα πλήρως συνδεδεμένο πολυεπίπεδο δίκτυο εμπροσθόδρομης τροφοδοσίας, όπου ο αριθμός των μονάδων στο επίπεδο εισόδου καθορίζεται από τον αριθμό των τιμών των δεδομένων εισόδου, ενώ χρησιμοποιούνται ποικίλος αριθμός ενδιάμεσων επιπέδων και ένα μόνο επίπεδο εξόδου. Στην μελέτη συμπερηλήφθηκαν δεδομένα από την πενταετή παρακολούθηση 334 ασθενών (εκ των οποίων επιλέχθηκαν τυχαία 284 ασθενείς για εκπαίδευση και 50 για επικύρωση) που έλαβαν θεραπεία για τον καρκίνο του παχέος εντέρου προκειμένου να εκπαιδευτούν και να επικυρωθούν 6 νευρωνικά δίκτυα που έχουν σχεδιαστεί για την πρόβλεψη θανάτου εντός 9, 12, 15, 18, 21 και 24 μηνών. Χρησιμοποιήθηκαν 42 κλινικοπαθολογικές μεταβλητές, οι οποίες απεικονίζονται στον παρακάτω πίνακα, ενώ ο κανόνας μάθησης που επιλέχθηκε για τα νευρωνικά δίκτυα ήταν η *αντίστροφη διάδοση σφάλματος (back propagation of error)*, που ρυθμίζει τις εσωτερικές παραμέτρους του δικτύου κατά τη διάρκεια των επαναλαμβανόμενων κύκλων ώστε να μειωθεί το συνολικό σφάλμα.

**Clinicopathological factors used to train and validate the artificial neural networks**

<b>Patient dependent</b>	<b>Clinician dependent</b>
Age	Liver metastases on preoperative ultrasound scan
Sex	Liver metastases at operation
Abdominal pain	Operation type
Altered bowel habit	Surgeon rank
Weight loss	Tumour fixity
Obstruction	Involvement of adjacent organs
Palpable mass	Peritoneal involvement
Mucus per rectum	Tumour perforation
Bleeding per rectum	Resection judged curative
Anaemia	Anastomosis type
Family history	Blood transfusion
History of inflammatory bowel disease	Adjuvant therapy
	Dukes' stage

**Πίνακας 20 :** Κλινικοπαθολογικές μεταβλητές που χρησιμοποιήθηκαν στην εκπαίδευση και επικύρωση των τεχνητών νευρωνικών δικτύων.

Το μοντέλο 12μηνης πρόβλεψης που εκπαιδεύτηκε προηγουμένως εφαρμόστηκε στην συνέχεια στα δεδομένα διетуούς παρακολούθησης 92 ασθενών από διαφορετικό ιατρικό κέντρο, των οποίων η έκβαση παρέμεινε κρυφή από την ερευνητική ομάδα. Δεν ακολούθησε περαιτέρω εκπαίδευση του νευρωνικού δικτύου, ενώ τα αποτελέσματα του μοντέλου συγκρίθηκαν με εκείνα δύο χειρουργών του παχέος εντέρου, οι οποίοι είχαν στη διάθεσή τους τα ίδια δεδομένα.

Τα αποτελέσματα της μελέτης κατέδειξαν ότι το σύνολο των έξι νευρωνικών δικτύων ήταν σε θέση να επιτύχουν συνολική ακρίβεια μεγαλύτερη από 80% στην πρόβλεψη θανάτου για μεμονωμένους ασθενείς στο πρώτο ιατρικό κέντρο εντός 9, 12, 15, 18, 21 και 24 μηνών. Η μέση ευαισθησία και ειδικότητα ήταν 60% και 88%. Όταν το νευρωνικό δίκτυο του 12μηνου μοντέλου εφαρμόστηκε στα δεδομένα του δεύτερου ιατρικού κέντρου, επιτεύχθηκε συνολική ακρίβεια 90%, σε σύγκριση με την ακρίβεια των χειρουργών που βρέθηκε ίση με 79% και 75% αντίστοιχα. Εξετάστηκαν επίσης οι τιμές της θετικής και αρνητικής προγνωστικής αξίας, καθώς και το ποσοστό πιθανότητας που δείχνει την αξία

μιας δοκιμής για αυξημένη βεβαιότητα για την πρόβλεψη ή τη διάγνωση. Οι παρακάτω πίνακες παρουσιάζουν συγκεντρωτικά τα στοιχεία που προαναφέρθηκαν :

Survival period (months)	Overall accuracy (%) (95% CI)	Sensitivity (%)	Specificity (%)	Mortality (%)	Positive predictive value (%)	Negative predictive value (%)	Likelihood ratio (95% CI)
9	82 (72-90)	42	93	21	61	86	6.0 (1.5-21)
12	81 (70-91)	52	90	24	63	86	5.4 (1.7-16)
15	81 (70-91)	60	89	28	68	85	5.6 (2.0-16)
18	81 (70-91)	65	87	29	68	86	5.2 (2.0-13)
21	81 (70-92)	71	86	32	70	86	4.9 (2.0-12)
24	80 (69-91)	71	85	35	71	85	4.7 (2.0-11)

Table 3: Performance of artificial neural networks trained and validated on data from institution I

**Πίνακας 21 :** Απόδοση των τεχνητών νευρωνικών δικτύων που εκπαιδεύτηκαν και επικυρώθηκαν σε δεδομένα του 1<sup>ου</sup> ιατρικού κέντρου.

Predictor	Overall accuracy (%) (95% CI)	Sensitivity (%)	Specificity (%)	Mortality (%)	Positive predictive value (%)	Negative predictive value (%)	Likelihood ratio (95% CI)
Clinician 1	79 (71-87)	33	82	7	11	95	1.9 (0.56-6.5)
Clinician 2	75 (66-84)	67	76	7	16	97	2.7 (1.4-5.4)
Artificial neural network	90 (84-96)	67	92	7	36	98	8.2 (3.3-20)

Table 4: Performances of clinicians and previously trained 12-month artificial neural network on data set from institution II

**Πίνακας 22 :** Απόδοση των ιατρών και του προηγουμένως εκπαιδευμένου 12μηνου τεχνητού νευρωνικού δικτύου σε βάση δεδομένων από το 2<sup>ο</sup> ιατρικό κέντρο.

Οι συγγραφείς καταλήγουν στο συμπέρασμα ότι τα νευρωνικά δίκτυα ήταν σε θέση να προβλέψουν την έκβαση μεμονωμένων ασθενών με καρκίνο του παχέος εντέρου με πολύ μεγαλύτερη ακρίβεια σε σχέση με τις τρέχουσες διαθέσιμες κλινικοπαθολογικές μεθόδους. Η διαπίστωση αυτή έχει επιπτώσεις στην επιλογή των ασθενών για συμπληρωματική θεραπεία, ανεξάρτητα από τα στατιστικά στοιχεία που αναλογούν στο στάδιο του όγκου. Το σημαντικότερο όμως καινοτόμο βήμα της παρούσας μελέτης ήταν ότι, αφού εκπαιδεύτηκαν σε δεδομένα από ένα ίδρυμα, τα νευρωνικά δίκτυα ήταν σε θέση να προβλέψουν με ακρίβεια την έκβαση των ασθενών από ένα άλλο ασυσχέτιστο ιατρικό κέντρο, αν και απαιτούνται μεγαλύτερες βάσεις δεδομένων για την απόδειξη αυτού του ισχυρισμού.

## **2. Εφαρμογή των Νευρωνικών Δικτύων στην πρόβλεψη επιβίωσης ασθενών με καρκίνο του παχέος εντέρου [18]**

Η πρόβλεψη της έκβασης ασθενών με καρκίνο του παχέος εντέρου που έχουν υποβληθεί σε χειρουργική επέμβαση αποτελεί σημαντικό μέλημα διότι σχεδόν το 50% των ασθενών που υποβάλλονται σε μια δυνητικά θεραπευτική εκτομή θα αντιμετωπίσουν υποτροπή. Έχει ήδη αναφερθεί ότι οι παρούσες προγνωστικές μέθοδοι όπως η μέθοδος του Dukes (ταξινόμηση 4 σταδίων με διαφορετικά ποσοστά πενταετούς επιβίωσης που περιγράφηκε νωρίτερα) ή η εφαρμογή της ταξινόμησης σύμφωνα με το σύστημα TNM (Tumor, Nodes, Metastasis – Περιλαμβάνονται 4 βασικά στάδια που περιλαμβάνουν επιμέρους υποστάδια, ενώ το ευρύτερο στάδιο του καρκίνου αναφέρεται ως αριθμός I,II, III, IV. Είναι προφανές ότι ένας μεγαλύτερος αριθμός υποδεικνύει έναν πιο προχωρημένο καρκίνο και πιθανώς μια χειρότερη έκβαση) είναι υπερβολικά ευρείες και απαιτείται περαιτέρω επεξεργασία για την πρόβλεψη που αφορά υποομάδες υψηλού κινδύνου.

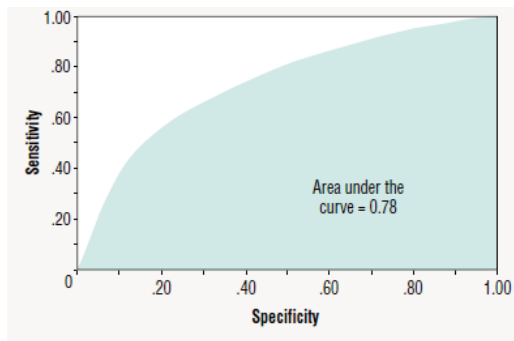
Μια πιο καινοτόμα προσέγγιση είναι να καθοριστεί ένα φαινότυπο (δηλαδή ένα σύνολο χαρακτηριστικών που εκδηλώνει ένας οργανισμός μια δεδομένη χρονική στιγμή) που σχετίζεται με την υποτροπή. Οι συγγραφείς επιχείρησαν λοιπόν στην παρούσα μελέτη να συγκρίνουν δύο μεθόδους ανάλυσης τέτοιου είδους δεδομένων, έναν πιο παραδοσιακό όπως η λογιστική παλινδρόμηση και έναν πιο σύγχρονο όπως ένα νευρωνικό δίκτυο. Αναλύθηκαν παθολογικά δεδομένα για 403 ασθενείς (οι οποίοι αποτελούν μια ομοιογενή ομάδα, με την ίδια διάγνωση στο ίδιο στάδιο της νόσου) από μια μεγάλη κλινική μελέτη, ενώ το αποτέλεσμα που προέκυψε από την ανάλυση ήταν η υποτροπή ή μη της νόσου. Οι παθολογικές μεταβλητές που χρησιμοποιήθηκαν και η ερμηνεία τους παρατίθενται στον παρακάτω πίνακα :

Variable	Description
Age	Age at randomization
Site	Site of tumor (cecum, ascending, transverse, descending or sigmoid colon, rectum)
Size	Greatest diameter of tumor
Grade	Grade of tumor (well, moderately, or poorly differentiated)
Depth of Penetration	Muscularis propria, subserosa, or serosa
Vascular Invasion	Presence of extramural vascular invasion
Lymphocytic Response	Evidence of lymphatic response in tumor
Infiltrative Growth Margin	Presence of an infiltrative growth margin of the tumor
Mucinous Tumor	Presence of > 50% mucin
Sex	Male or female

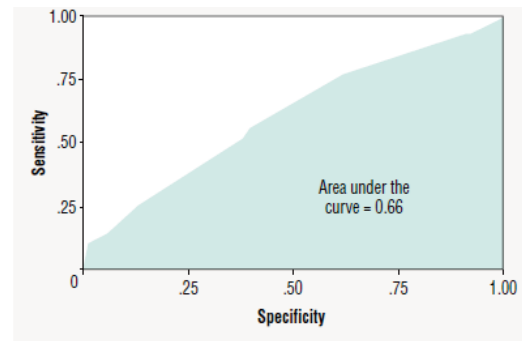
**Πίνακας 23 :** Παθολογικές μεταβλητές που χρησιμοποιήθηκαν στα μοντέλα λογιστικής παλινδρόμησης και νευρωνικού δικτύου.

Ο τύπος νευρωνικού δικτύου που επέλεξαν να χρησιμοποιήσουν οι συγγραφείς ήταν ο πολυεπίπεδος αισθητήρας (*Multilayer Perceptron – MLP*), καθώς παρουσιάζει ισχυρό πλεονέκτημα στην ταξινόμηση των μεταβλητών εισόδου σε δύο διακριτές ομάδες εξόδου (πχ ζωντανοί ή νεκροί). Το νευρωνικό δίκτυο παρείχε αποτέλεσμα για την ROC AUC (περιοχή κάτω από την καμπύλη) της τάξης του 78% για τα δεδομένα εκπαίδευσης και δοκιμής, γεγονός που υποδηλώνει καλή γενίκευση παρά το ότι δεν χρησιμοποιήθηκε ένα ανεξάρτητο σύνολο επικύρωσης. Το σύστημα επέδειξε επίσης ευαισθησία ίση με 95% και για την αναγνώριση θανάτου και για την αναγνώριση επιβίωσης και ειδικότητα ίση με 32% για την εύρεση θανάτου και 37% για την εύρεση επιβίωσης αντίστοιχα. Αξίζει να σημειωθεί ότι το νευρωνικό δίκτυο δεν παρέλειψε κάποιες μεταβλητές ως μικρής σημασίας, αλλά αντίθετα τις διατήρησε όλες. Οι πιο σημαντικές μεταβλητές ωστόσο ως προς τη συμβολή τους στο μοντέλο ήταν η αγγειακή εισβολή, το διηθητικό πρότυπο ανάπτυξης, η περιοχή του όγκου και το φύλο.

Στη συνέχεια, τα αποτελέσματα του νευρωνικού δικτύου συγκρίθηκαν με την ανάλυση λογιστικής παλινδρόμησης όσον αφορά τα γραφήματα καμπύλης ROC (λειτουργικών χαρακτηριστικών δέκτη). Το μοντέλο λογιστικής παλινδρόμησης έδωσε ένα αποτέλεσμα της τάξης του 66%, ενώ το νευρωνικό δίκτυο, όπως αναφέρθηκε, 78%, καθώς το λογιστικό μοντέλο ήταν σε θέση να ταξινομήσει σωστά μόνο το 80% των δεδομένων που εισήχθησαν :



Abbreviation: ROC = receiver operating characteristic



Abbreviation: ROC = receiver operating characteristic

**Εικόνες 39 & 40 :** Καμπύλη ROC για το νευρωνικό δίκτυο (αριστερά) και καμπύλη ROC για τη λογιστική παλινδρόμηση (δεξιά).

Το συμπέρασμα που προκύπτει από αυτή την μελέτη είναι ότι το νευρωνικό δίκτυο υπερσχύει της λογιστικής παλινδρόμησης. Απαιτείται, ωστόσο, περαιτέρω ανάλυση με χρήση ένα μεγαλύτερο μέγεθος δείγματος ασθενών ( $n > 1000$ ), βελτιωμένα στατιστικά μοντέλα και ένα πιο εκλεπτυσμένο νευρωνικό δίκτυο. Θα μπορούσαμε να ισχυριστούμε ότι επειδή έχει παρέλθει μια δεκαετία από τότε που διεξήχθη η παρούσα έρευνα, το μοντέλο που χρησιμοποίησαν οι συγγραφείς ήταν απλό και όχι ιδιαίτερα εξελιγμένο σε σχέση με τα σημερινά δεδομένα, αν και επέδειξε ικανοποιητικά αποτελέσματα για την εποχή του.

### 3.7 Μελέτες σχετικά με την λευχαιμία (Leukemia)

Στην περίπτωση της λευχαιμίας, εξετάστηκαν δύο μελέτες. Η πρώτη είχε στόχο την επιλογή έντασης της θεραπείας στην οξεία λεμφοβλαστική παιδική λευχαιμία, ενώ η δεύτερη στόχευε στη διάγνωση της λευχαιμίας.

#### 1. Νέο εργαλείο υποστήριξης αποφάσεων για την επιλογή έντασης της θεραπείας στην οξεία λεμφοβλαστική παιδική λευχαιμία [19]

Η οξεία λεμφοβλαστική λευχαιμία (*Acute Lymphoblastic Leukemia – ALL*) αποτελεί τον πιο συνηθισμένο τύπο καρκίνου στην παιδική ηλικία, ιδιαίτερα σε παιδιά κάτω των 15 ετών. Σε αυτόν τον τύπο καρκίνου, η θεραπεία διαμορφώνεται ανάλογα με τον κίνδυνο υποτροπής. Ως εκ τούτου, μια κατάλληλη εκτίμηση αυτού του κινδύνου αποτελεί τον πιο σημαντικό παράγοντα για τον καθορισμό της στρατηγικής που αφορά τη θεραπεία.

Στην παρούσα εργασία, προτείνεται ένα νέο εργαλείο υποστήριξης αποφάσεων για τη βελτίωση της επιλογής της έντασης της θεραπείας στην οξεία λεμφοβλαστική παιδική λευχαιμία. Στην έρευνα που διεξήγαγαν οι συγγραφείς, συμμετείχαν 158 παιδιά με οξεία λεμφοβλαστική λευχαιμία από τη Βραζιλία, ενώ η πλειοψηφία των περιπτώσεων νοσηλεύτηκαν τη δεκαετία 1993 – 2004 στα δύο κύρια Πανεπιστημιακά Νοσοκομεία του Ρίο ντε Τζανέιρο. Τα πρωτόκολλα BFM290 και BFM95, δύο ευρέως χρησιμοποιούμενα πρωτόκολλα ταξινόμησης που ακολουθήθηκαν αρχικά για τη θεραπεία των ασθενών, διαχώρισαν τους ασθενείς σε ομάδες χαμηλού, μεσαίου και υψηλού κινδύνου με σχεδόν

πανομοιότυπες ομάδες χαμηλού και μεσαίου κινδύνου λόγω της χαμηλής διακριτικής τους ικανότητας.

Η βάση δεδομένων περιελάμβανε αρχικά 189 πιθανά χαρακτηριστικά εισόδου που θα μπορούσαν να ταξινομηθούν ως εξής : (α) κλινικά δεδομένα, όπως ηλικία της διάγνωσης, φύλο, φυλή, αναιμία, πυρετός, αιμορραγία κτλ, (β) εργαστηριακά δεδομένα, όπως αριθμός ερυθρών και λευκών αιμοσφαιρίων, αιματοκρίτης, αιμοπετάλια κτλ και (γ) δεδομένα σχετικά με την απόκριση στη θεραπεία, όπως το σύνολο των βλαστικών κυττάρων την 8<sup>η</sup> ημέρα της θεραπείας ή η πλήρης ύφεση την 33<sup>η</sup> ημέρα της θεραπείας. Μία από τις προκλήσεις που αντιμετώπισαν οι συγγραφείς ήταν, όπως αναμένετο, η κατάλληλη επιλογή των χαρακτηριστικών. Για το σκοπό αυτό, ακολούθησε προεπεξεργασία των υποψήφιων μεταβλητών εισόδου ούτως ώστε να εξαληφθούν οι περιττές ή πλεονάζουσες μεταβλητές και να αξιολογηθεί η διακριτική τους ικανότητα (δηλαδή η εξάρτηση μεταβλητών εισόδου-αποτελέσματος με διμεταβλητό πίνακα συνάφειας και διτετράγωνη στατιστική) και η αξία της θετικής πρόβλεψης (δηλαδή η πιθανότητα υποτροπής ενός ασθενός σε σχέση με την ύπαρξη ή μη μιας μεταβλητής). Με αυτόν τον τρόπο, οι μεταβλητές μειώθηκαν σε 17 από τις 189 που ήταν αρχικά και ο αριθμός των ασθενών που πληρούσαν τα κριτήρια σε 144 από τους 158. Στη συνέχεια, οι συγγραφείς εφάρμοσαν μια μεθοδολογία βασισμένη σε αμοιβαία πληροφορία και ένα νευρωνικό δίκτυο ώστε από τις 17 προεπιλεγμένες μεταβλητές να επιλεγούν αυτές που παρουσιάζουν μεγάλη συσχέτιση με το αποτέλεσμα και ταυτόχρονα μικρή με τις υπόλοιπες μεταβλητές του συνόλου (*MI : measure of (in)dependence*) ώστε να προκύψει μια αποτελεσματική εκτίμηση του κινδύνου.

Για την εκτίμηση της απόδοσης του προτεινόμενου συστήματος επιλέχθηκαν τα εξής μεγέθη :

- 1) το ποσοστό υψηλού κινδύνου στην ομάδα υποτροπής (RHR : rate of high-risk). Όσο υψηλότερο το ποσοστό αυτό, τόσο καλύτερη η απόδοση, καθώς αντιπροσωπεύει τη σωστή αναγνώριση των ασθενών υψηλού κινδύνου.
- 2) το ποσοστό του υπερεκτιμημένου κινδύνου (ROR : rate of overestimated risk), που αντιπροσωπεύει τους ασθενείς που κρίθηκαν ως υψηλού κινδύνου από το παρόν μοντέλο και όχι από το BFM95.
- 3) το ποσοστό του υποτιμημένου κινδύνου (RSR : rate of subestimated risk), δηλαδή το ποσοστό αυτών που εκτιμήθηκαν ως χαμηλού κινδύνου από το μοντέλο και όχι ως υψηλού κινδύνου από το BFM95.

Στους παρακάτω πίνακες παρουσιάζονται τα 17 προεπιλεγμένα χαρακτηριστικά που αναφέρθηκαν νωρίτερα και μια επισκόπηση των πειραμάτων με διαφορετικά σύνολα μεταβλητών :



Ordering	Variable	MI <sup>(*)</sup>
1	White cell blood count at diagnosis (WBC)	0.197
2	Age	0.086
3	T-immunophenotype	0.032
4	Blast count cell in blood at 8 <sup>th</sup> day of treatment	0.036
5	Gender	0.024
6	FAB classification	0.000
7	Race	0.012
8	Malaise	0.004
9	Fever	0.004
10	Bone pain	0.002
11	Bleeding	0.002
12	Infectious focus	0.001
13	Pain	0.001
14	Hemoglobin	0.069
15	Lymphonodomegaly	0.001
16	Spleen	0.033
17	Liver	0.063

	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
<b>All 17 status</b>	WBC Age	WBC Age	WBC Age	WBC Age	WBC Age
	T-immunophenotype Blasts blood count in 8 <sup>th</sup> day of treatment	T-immunophenotype Blasts blood count in 8 <sup>th</sup> day of treatment	T-immunophenotype Blasts blood count in 8 <sup>th</sup> day of treatment	T-immunophenotype Blasts blood count in 8 <sup>th</sup> day of treatment	T-immunophenotype Blasts blood count in 8 <sup>th</sup> day of treatment
	Gender	Gender	Gender	Gender	Gender
	FAB classification	FAB classification			Race
	Race				
<b>RHR (%)</b> <i>the higher the better</i>	91	87	81	87	98
<b>ROR (%)</b> <i>the lower the better</i>	47	31	32	32	21
<b>RSR (%)</b> <i>the lower the better</i>	4	0	2	0	0

**Πίνακες 24 & 25 :** Πίνακας υποψήφιων χαρακτηριστικών και η τιμή MI που λαμβάνουν σε σχέση με το αποτέλεσμα, Πίνακας πειραμάτων με διαφορετικά σύνολα χαρακτηριστικών.

Στη σύγκριση των αποτελεσμάτων του εν λόγω μοντέλου με το πρωτόκολλο BFM95, παρατηρήθηκε ότι μεταξύ των ασθενών που παρουσίασαν υποτροπή, ένα ποσοστό της τάξης του 98,2% χαρακτηρίστηκε ως υψηλού κινδύνου από την προτεινόμενη μεθοδολογία, γεγονός που αποδεικνύει την υπεροχή του συστήματος σύμφωνα με τα στοιχεία εκτίμησης κινδύνου των παρακάτω πινάκων :

Risk estimated by the model	Non-relapsed patients	Relapsed patients	Total	BFM Risk ↓	Model Risk ↓	Non-relapsed patients	Relapsed patients	Total
Low	60 (67.3%)	1 (1.8%)	61 (42%)	Low	Low	20 (22.2%)	0 (0%)	20 (13.9%)
High	30 (33.3%)	53 (98.2%)	83 (58%)		High	4 (4.4%)	5 (9.3%)	9 (6.3%)
Total	90 (100%)	54 (100%)	144 (100%)	Intermed.	Low	31 (34.4%)	1 (1.9%)	32 (22.2%)
					High	15 (16.7%)	23 (42.6%)	38 (26.4%)
				High	Low	9 (10.0%)	0 (0%)	9 (6.3%)
					High	11 (12.2%)	25 (46.3%)	36 (25.0%)
				Total		90 (100%)	54 (100%)	144 (100%)

**Πίνακες 26 & 27 :** Πίνακας εκτίμησης κινδύνου υποτροπής (α) από το προτεινόμενο σύστημα και (β) από το πρωτόκολλο BFM95 συγκριτικά με το προτεινόμενο μοντέλο.

Αξίζει να σημειώσουμε, τέλος, ότι το μοντέλο που προτείνουν οι συγγραφείς έρχεται να καλύψει μια βασική αδυναμία του πρωτοκόλλου BFM95, η οποία έγκειται στο σύνολο των ασθενών μεσαίου κινδύνου, καθώς μια σημαντική μερίδα του συνόλου αυτού που υποτροπίασε θα χαρακτηριζόταν ως υψηλού κινδύνου από το παρόν μοντέλο. Έτσι, από κλινικής άποψης, η προτεινόμενη μεθοδολογία επιτρέπει τον καλύτερο εντοπισμό της ομάδας των ασθενών που θα πρέπει να αντιμετωπίζονται με πιο εντατική χημειοθεραπεία.

## **2. Εφαρμογή πιθανοτικών νευρωνικών δικτύων στην πρόβλεψη της λευχαιμίας και των εμβρυονικών νεοπλασμάτων του κεντρικού νευρικού συστήματος [20]**

Έχει ήδη διαπιστωθεί και διατυπωθεί ότι η ακριβής διάγνωση και ταξινόμηση των καρκινικών υποτύπων αποτελούν το βασικό ζήτημα για τη βέλτιστη αντιμετώπιση των ασθενών με καρκίνο. Αρκετές μελέτες δείχνουν ότι η αναγνώριση και ταξινόμηση διαφόρων

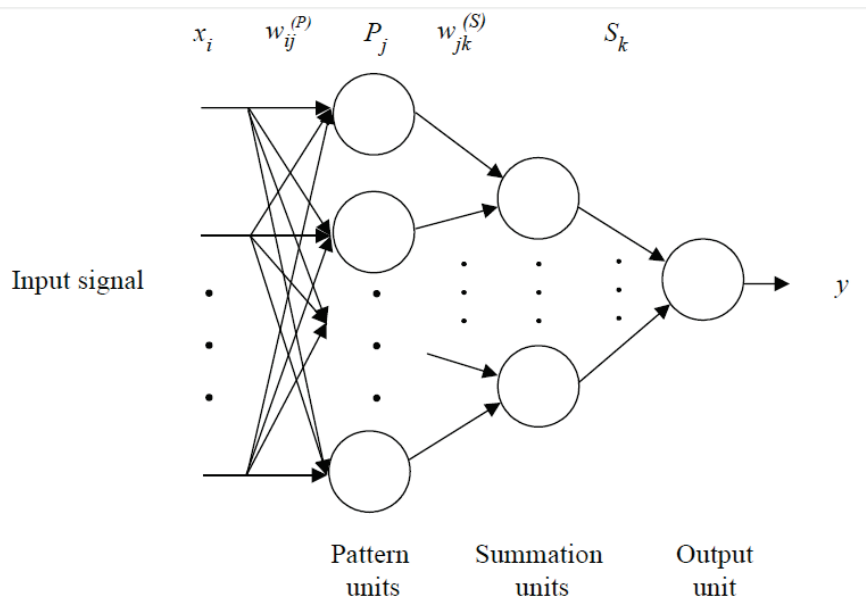
τύπων καρκίνου μπορεί να πραγματοποιηθεί με υψηλή ακρίβεια, ευαισθησία και ειδικότητα από τα χαρακτηριστικά γονιδιακής έκφρασης βασισμένα σε μικροσυστοιχίες με τη χρήση τεχνητών νευρωνικών δικτύων.

Στην παρούσα εργασία, έλαβε χώρα μια ολοκληρωμένη μελέτη για να διερευνηθεί η ικανότητα των πιθανοτικών νευρωνικών δικτύων (*Probabilistic Neural Networks – PNN*) με χρήση μιας μεθόδου επιλογής χαρακτηριστικών, της λεγόμενης στατιστικής σήματος προς θόρυβο (*signal-to-noise statistic*), κατά την εφαρμογή της ταξινόμησης καρκινικών υποτύπων. Η στατιστική σήματος προς θόρυβο, η οποία αντιπροσωπεύει τη συσχέτιση με τη διάκριση της κλάσης, χρησιμοποιείται για την επιλογή των γενετικών δεικτών και την μείωση της διάστασης των δειγμάτων των δεδομένων για το πιθανοτικό νευρωνικό δίκτυο.

Οι συγγραφείς χρησιμοποιούν για τη βάση δεδομένων που αφορά τη λευχαιμία ένα σύνολο εκπαίδευσης 38 δειγμάτων σε συνδυασμό με ένα σύνολο δοκιμών αποτελούμενο από 34 δείγματα (σύνολο 72 δειγμάτων) και τρεις βάσεις δεδομένων για τα εμβρυονικά νεοπλάσματα του κεντρικού νευρικού συστήματος αποτελούμενες από 42 εμβρυονικά νεοπλάσματα του κεντρικού νευρικού συστήματος, 34 δείγματα μορφολογίας μυελοβλαστώματος και 60 δείγματα μυελοβλαστώματος που έχουν υποστεί θεραπεία. Κάθε δείγμα της βάσης δεδομένων αποτελείται από έναν εντυπωσιακά μεγάλο αριθμό χαρακτηριστικών γονιδιακής έκφρασης, εκ των οποίων πρέπει να επιλεχθούν αυτά που παράγουν αποτελεσματικότερη πρόβλεψη. Όπως ήδη αναφέρθηκε, επιλέχθηκε από τους συγγραφείς η στατιστική σήματος προς θόρυβο ως μέθοδος επιλογής χαρακτηριστικών, καθώς συγκρινόμενη με τρεις άλλες δημοφιλείς μεθόδους όπως η διτετράγωνη στατιστική, η μέθοδος αρωγής και η μέθοδος βασισμένη σε συσχετισμό, αποδίδει καλύτερα. Η εξίσωση ορίζεται ως  $F = \left| \frac{\mu_1 - \mu_2}{\sigma_1 + \sigma_2} \right|$ , όπου το  $\mu_1$  υποδηλώνει το μέσο επίπεδο έκφρασης και το  $\sigma_1$  την τυπική απόκλιση έκφρασης για τα δείγματα στην κλάση 1, ενώ τα  $\mu_2$  και  $\sigma_2$  ορίζονται αντίστοιχα για την κλάση 2. Η παραπάνω εξίσωση επιχειρεί να συγκεντρώσει τα γονίδια που χαρακτηρίζονται από ευρύτερο διαχωρισμό τάξης και μικρότερη διασπορά γύρω από το μέσον.

Οι συγγραφείς επιλέγουν να χρησιμοποιήσουν ένα πιθανοτικό νευρωνικό δίκτυο επειδή μπορεί να διαχειριστεί πολλαπλές βάσεις δεδομένων μικρού σχετικά μεγέθους αλλά με μεγάλο διαχωρισμό τάξης αποδίδοντας ικανοποιητικά. Η αρχιτεκτονική του εν λόγω δικτύου παρουσιάζεται στο παρακάτω σχήμα :





**Εικόνα 41 :** Αρχιτεκτονική πιθανοτικού νευρωνικού δικτύου.

Όπως υποστηρίζουν οι συγγραφείς, τα πειραματικά αποτελέσματα αποδεικνύουν πως η συσχέτιση του PNN με τη στατιστική σήματος προς θόρυβο μπορεί να επιτύχει καλύτερα αποτελέσματα ταξινόμησης για δύο τύπους οξείας λευχαιμίας και τρεις κατηγορίες εμβρυονικών νεοπλασμάτων του κεντρικού νευρικού συστήματος με ικανοποιητική ταχύτητα υπολογισμού, όπως αποδεικνύουν οι παρακάτω πίνακες, ενώ το PNN επιτυγχάνει ακρίβεια 100%.

Feature selection method	Data set			
	Leukemia (%)	Embryonal CNS tumor (%)	Medulloblastoma morphology (%)	Medulloblastoma treatment outcome (%)
Signal-to-noise statistic	95.4	84	94.1	69.2
$\chi^2$ statistic	94	79.9	92.3	55.8
Relief	93	70.7	92.6	60.7
CFS	94.6	73.1	88.6	60.3

**Πίνακας 28 :** Αποτελέσματα ταξινόμησης διασταυρούμενης επικύρωσης 10 φορές των τεσσάρων μεθόδων επιλογής χαρακτηριστικών του ταξινομητή PNN.

Number of genes	Data set			
	Leukemia (%)	Embryonal CNS tumor (%)	Medulloblastoma morphology (%)	Medulloblastoma treatment outcome (%)
10	95.2	63.7	86.8	63.8
50	95.3	80.5	89.8	68.0
100	95.4	84.0	86.7	69.2
200	93.0	86.1	84.5	68.9
500	92.8	80.7	82.9	66.2

**Πίνακας 29 :** Αποτελέσματα ταξινόμησης διασταυρούμενης επικύρωσης 10 φωνών των τεσσάρων μεθόδων επιλογής χαρακτηριστικών του PNN με διαφορετικό μέγεθος υποσυνόλων χαρακτηριστικών.

Παραθέτουμε στη συνέχεια τα 50 σημαντικότερα γονίδια που επιλέχθηκαν από τη στατιστική σήματος προς θόρυβο τα οποία συνδέονται πιο στενά με κάθε τύπο όγκου :

Data set	Top 50 genes
Leukemia (ALL/AML)	ZyxinX95735, DF D compon, CD33 CD33 a, CST3 Cystat, PRG1 Proteo, CTSD Cathep, APLP2 Amylo, PPGB Protec, LYN V-yes-1, GLUTATHIONE, FAH Fumaryl, ATP6C Vacuo, MYL1 Myosin, SPTAN1 Spec, IGB Immunog, CD19 geneM8, Terminal tr, GB DEF PROTEASOME, CCND3 Cycli, MB-1 geneU0, TCF3 Transc
Embryonal CNS multi-type tumors	D13900_atMi, D83735_atAd, M12125_atSk, D26070_atT
Medulloblastoma morphology	D14530_at40, HG3543-HT37, X53331_atMG, X65724_atND, X67951_atPA, U63842_atNe, X64330_atAT, HG1980-HT20
Medulloblastoma treatment outcome	U08998_atTA

**Πίνακας 30 :** Τα 50 σημαντικότερα γονίδια για κάθε μία από τις στέσσες βάσεις δεδομένων όπως προκύπτουν από την ανάλυση της στατιστικής σήματος προς θόρυβο.

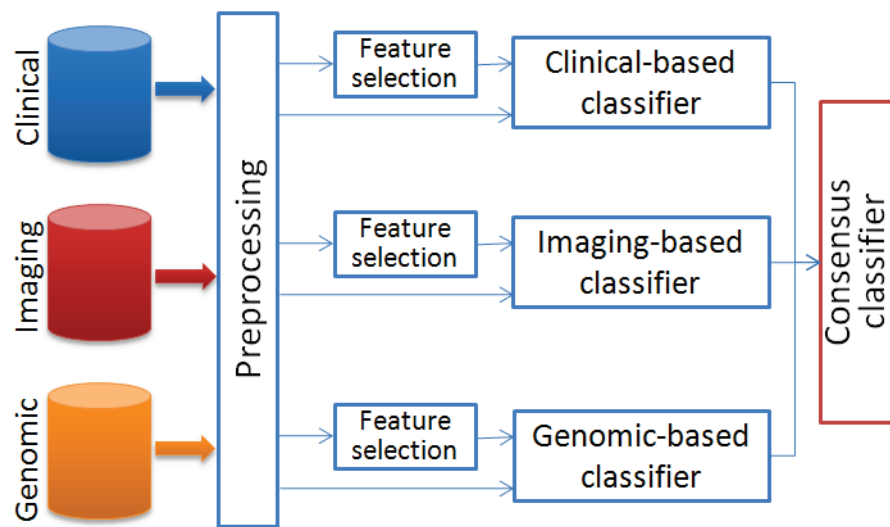
Συμπερασματικά, θα λέγαμε ότι το πιθανοτικό νευρωνικό δίκτυο αποδεικνύεται ταχύτερο από άλλα νευρωνικά δίκτυα, όπως το MLP, όταν το μέγεθος της βάσης δεδομένων κυμαίνεται στην κλίμακα δεκάδων έως εκατοντάδων, διότι δεν απαιτούνται επαναλαμβανόμενες διαδικασίες στην εκπαίδευση του PNN, ενώ επιπλέον το μοντέλο παρέχει νέες γνώσεις πάνω στην ανάλυση και ερμηνεία των δεδομένων μικροσυστοιχιών.

### 3.8 Μελέτες σχετικά με τον καρκίνο του στόματος (Oral cancer)

Όσον αφορά τον καρκίνο του στόματος, αναλύθηκε μία μελέτη που πραγματοποιήθηκε ένα πολυπαραμετρικό σύστημα υποστήριξης αποφάσεων για την πρόβλεψη επανεμφάνισης του καρκίνου του στόματος [21].

Το πλακώδες καρκίνωμα του στόματος (oral squamous cell carcinoma – OSCC) αποτελεί το κυρίαρχο νεόπλασμα της περιοχής της κεφαλής και του αυχένα, ενώ χαρακτηρίζεται από ιδιαίτερα επιθετική φύση με αρκετά δυσμενή πρόγνωση. Στην παρούσα εργασία, οι συγγραφείς διαμορφώνουν ένα σύστημα υποστήριξης αποφάσεων, το οποίο ενσωματώνει μια πληθώρα ετερόκλητων δεδομένων (κλινικά, απεικονιστικά και γονιδιώματος), καλύπτοντας ως εκ τούτου όλες τις εκδηλώσεις της νόσου. Πρωταρχικός στόχος των συγγραφέων ήταν ο προσδιορισμός των παραγόντων που υπαγορεύουν την εξέλιξη του πλακώδους καρκινώματος του στόματος και ακολούθως η πρόβλεψη των πιθανών υποτροπών (τοπικών ή μεταστατικών) της νόσου με χρήση ενός συστήματος ταξινόμησης.

Το παρακάτω σχήμα απεικονίζει το διάγραμμα ροής της προτεινόμενης από τους συγγραφείς μεθοδολογίας. Αρχικά, όλες οι πηγές δεδομένων (κλινικά, απεικονιστικά και γονιδιώματος) υπόκεινται σε προεπεξεργασία, προκειμένου να βελτιωθεί η ποιότητα δεδομένων εισόδου που τροφοδοτούνται στα επόμενα βήματα. Στη συνέχεια, τα δεδομένα εισόδου είτε τροφοδοτούνται άμεσα για ταξινόμηση ή χρησιμοποιείται ένας αλγόριθμος επιλογής χαρακτηριστικών προκειμένου να παραληφθούν ενδεχομένως περιττά χαρακτηριστικά. Εκτός από τους επί μέρους ταξινομητές για κάθε τύπο δεδομένων, υλοποιείται και ένας «συναινετικός» ταξινομητής που συνδυάζει τις προαναφερθείσες προβλέψεις με συμπληρωματικό τρόπο ώστε να προκύψουν πιο ακριβή αποτελέσματα.



**Εικόνα 42 :** Διάγραμμα ροής του προτεινόμενου συστήματος υποστήριξης απόφασης.

Στην μελέτη συμπεριλήφθησαν 41 ασθενείς από κλινικές της Ιταλίας και της Ισπανίας, οι οποίοι διαγνώστηκαν με πλακώδες καρκίνωμα του στόματος και υποβλήθηκαν σε θεραπεία που οδήγησε σε πλήρη ύφεση της νόσου. Από αυτούς, οι 13 παρουσίασαν υποτροπή, ενώ οι υπόλοιποι 28 απαλλάχθηκαν από τη νόσο. Τα κλινικά δεδομένα που συμπεριέλαβαν οι συγγραφείς αποτελούνται από το ιατρικό ιστορικό του ασθενούς, την παθολογία και τα δεδομένα παραγόντων κινδύνου, τα δεδομένα απεικόνισης περιέχουν πληροφορίες εξαγόμενες από εικόνες αξονικής και μαγνητικής τομογραφίας του αρχικού όγκου και των γειτονικών λεμφαδένων της περιοχής της κεφαλής και του αυχένα, ενώ τα γονιδιωματικά δεδομένα περιλαμβάνουν πληροφορίες γονιδιακής έκφρασης από το δείγμα του όγκου. Από τα γονιδιωματικά δεδομένα, που ανέρχονται σε 450015 σε αριθμό, διατηρούνται τα 40 σημαντικότερα, αφού επαληθευτούν μέσω διαφορετικών αλγορίθμων ταξινόμησης. Τα προαναφερθέντα δεδομένα παρουσιάζονται συγκεντρωτικά στους παρακάτω πίνακες :

CLINICAL FEATURES				IMAGING FEATURES		
Ecog Status	Mobile Prosthesis	BMI	Grade Of Differentiation	Contrast Take-Up Rate	Necrosis	Side
Weight	Dental Cusps	Substance Exposition	Surgical Margins	Minor Axis Bigger than 10mm	Central Necrosis	Side Relative to Tumor
Height	Galvanic Current	Precancerous Lesions	Martinez-Gimeno Score	Extra Nodal Spreading	Bone Infiltration	Cluster
Diabetes	Oral Hygiene	Duration	Anneroths Mod Score	Shape Deviation	Carotid Infiltration	Number of Lymph Nodes
Allergies	Infection	Immunosuppressor Treatments Presence	D2_40Stain	Texture	Cutaneous Invasion	Number of Lymph Nodes Bigger than 3
Cholesterol	Type Of Infection	Immuno Duration	P53_STAIN	Water Content	Site	
Hypertension	Physical Agents	Immuno Type	P16Ink4aStain	GENES PINPOINTED AS MOST SIGNIFICANT		
Family History Of Malignance	Type Of Physical Agent	Tumor Maximum Diameter	EGFR Stain	LPO	TMC5	AI916628
Smoker	Diet Deficit	Tumor Thickness	CyclinD1Stain	MSLN	ROPN1	PIGR
Smoking Habits	Fe Haematic Concentration	Depth Of Invasion	Ki67Stain	CAPN13	AGR2	C20orf114
Quantity Per Day	Phummer Vinson	Basaloid Features	HPV_DNA	GLYATL2	SCGB1D1	CP
Smoking For	Hb Haematic Concentration	Lympho Plasmacytic Reaction	T Staging	CB959193	LOC440335	C10orf81
Ex Smoker	B12 Vitamins Haematic Concentration	Lympho Plasmacytic Invasion	N Staging	CLDN22	THC2339617	VTCN1
Quitted Smoking	A Vitamins Haematic Concentration	Perineural Invasion	M Staging	BCMP11	UPK1B	SCGB2A1
Alcohol	E Vitamins Haematic Concentration	Degree Of Cells Keratinisation		C20orf85	CRISP2	MSMB
Drinking Habits	Folati	Nuclear Pleomorphism		SCGB2A2	CHST9	FOXA1
Mechanical Trauma	Eating Habits	Number of Mitoses per 10HPF		SLC34A2	PROM1	C10orf81
						CA946373
						CLDN8
						CTAG1A
						SCGB3A1
						LOC63928
						OLFM4
						KCNJ16
						LOC124220
						PIP
						STATH

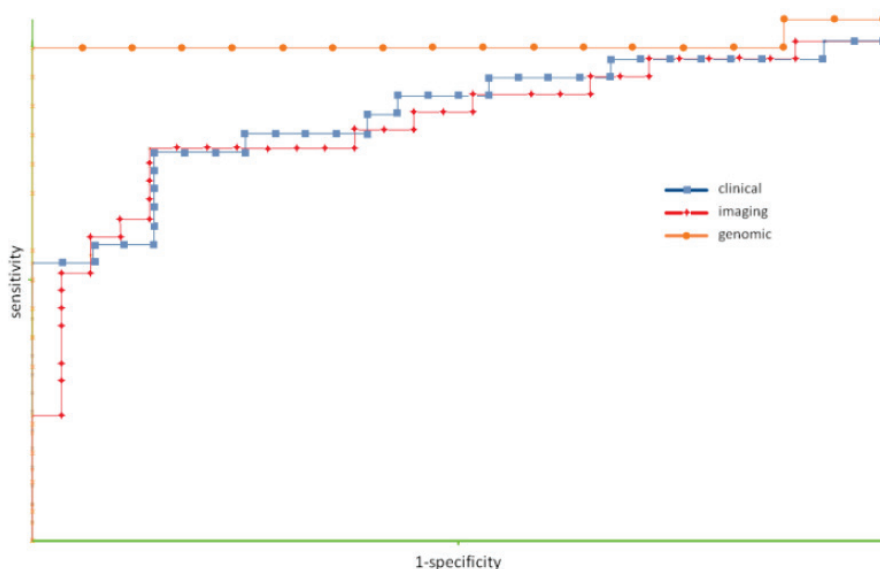
Note: Features in bold have >90% missing values

**Πίνακες 31, 32, 33** : Συγκεντρωτικοί πίνακες κλινικών, απεικονιστικών και γονιδιωματικών χαρακτηριστικών.

Αφότου οριστικοποιηθούν οι βάσεις δεδομένων, εξετάζεται η απόδοση πέντε αλγορίθμων ταξινόμησης όσον αφορά τη διάκριση των ασθενών που παρουσιάζουν ή όχι υποτροπή. Ειδικότερα, οι συγγραφείς εφαρμόζουν Μπεϋζιανά Δίκτυα (Bayesian Networks – BNs), Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANNs), Μηχανές Υποστήριξης Διανυσμάτων (Support Vector Machines – SVMs), Δέντρα Αποφάσεων (Decision Trees – DTs) και Τυχαία Δάση (Random Forests – RFs), τα οποία έχουν όλα χρησιμοποιηθεί σε ένα ευρύ φάσμα προβλημάτων στη βιοϊατρική μηχανική και έχουν αναλυθεί νωρίτερα. Σημειώνουμε μόνο ότι τα Τυχαία Δάση εμφανίστηκαν σχετικά πρόσφατα και χρησιμοποιούν υποσύνολα χαρακτηριστικών για να κατασκευάσουν πολλαπλά δέντρα αποφάσεων, κάθε ένα από τα οποία συμμετέχει με την ψήφο του στην προβλεπόμενη

κατηγορία, έτσι ώστε το τελικό αποτέλεσμα να προκύπτει από αυτό που συλλέγει την πλειοψηφία των μεμονωμένων ψήφων.

Στη συνέχεια, ακολούθησε διασταυρούμενη επικύρωση 10 φορών και σύγκριση των παραμέτρων ευαισθησίας, ειδικότητας, ακρίβειας και της Καμπύλης Λειτουργίας Δέκτη Χαρακτηριστικών (Receiver Operating Characteristic – ROC) για κάθε μία από τις τρεις πηγές δεδομένων και τους πέντε προαναφερθέντες αλγορίθμους ταξινόμησης, με και χωρίς επιλογή των χαρακτηριστικών. Με βάση λοιπόν τα συνολικά αποτελέσματα, παρατηρήθηκε ότι η δυνατότητα διάκρισης των γονιδιωματικών δεδομένων είναι κατά κάποιο τρόπο ανώτερη σε σχέση με τα κλινικά και απεικονιστικά δεδομένα.



**Εικόνα 43 :** Καμπύλες ROC για τα βέλτιστα σχήματα ταξινόμησης βασισμένα σε κλινικά, απεικονιστικά και γονιδιωματικά δεδομένα.

Εκτός από τους μεμονωμένους ταξινομητές, οι συγγραφείς εφάρμοσαν ένα επιπλέον σύστημα ταξινόμησης, το οποίο χρησιμοποιεί έναν πλειοψηφικό αλγόριθμο λήψης αποφάσεων που συνδυάζει τις καλύτερες επιδόσεις των συστημάτων ταξινόμησης κάθε πηγής δεδομένων προκειμένου να ληφθεί η συναινετική απόφαση. Μετά από επικύρωση, αποδεικνύεται ότι η προτεινόμενη προσέγγιση επιτυγχάνει απόλυτη διάκριση (ακρίβεια : 100%) μεταξύ των ασθενών με και χωρίς υποτροπή σε σχέση με τα αποτελέσματα άλλων μεθοδολογιών που παρουσιάζονται στη βιβλιογραφία, αν και κρίνεται αναγκαία περαιτέρω αξιολόγηση.

### 3.9 Μελέτες σχετικά με τον καρκίνο των ωοθηκών (Ovarian cancer)

Η μελέτη που αφορά τον καρκίνο των ωοθηκών είχε ως στόχο τη διάγνωση του καρκίνου με χρήση ασαφούς νευρωνικού δικτύου συμπληρωματικής εκμάθησης [22].

Όπως έχει ήδη επισημανθεί σε όλους τους προαναφερθέντες τύπους καρκίνου, η έγκαιρη διάγνωση είναι ύψιστης σημασίας και για την μείωση του υψηλού ποσοστού θνησιμότητας του καρκίνου των ωοθηκών, που αποτελεί την τέταρτη αιτία θανάτου οφειλόμενου σε καρκίνο σε γυναίκες σε Αμερική και Καναδά. Δυστυχώς, όπως υποστηρίζουν οι συγγραφείς στην παρούσα μελέτη, τα τρέχοντα εργαλεία ανίχνευσης δεν παρουσιάζουν υψηλή ευαισθησία στην ανίχνευση του καρκίνου. Επιπρόσθετα, οι νέες τεχνικές, όπως η μικρο-συστοιχία δεοξυριβονουκλεϊκού οξέος και τα πρωτεομικά δεδομένα, είναι δύσκολο να αναλυθούν λόγω της υψηλής διαστασιμότητας, ενώ οι συμβατικές μέθοδοι, όπως η εξέταση αίματος, δεν είναι ούτε ευαίσθητες, ούτε εξειδικευμένες.

Προκύπτει, κατά συνέπεια, η ανάγκη για νέες, υπολογιστικά έξυπνες μεθόδους. Εργαζόμενοι σε αυτήν την κατεύθυνση, οι συγγραφείς προτείνουν ένα λειτουργικό μοντέλο ανθρώπινης αναγνώρισης προτύπου γνωστό ως ασαφές νευρωνικό δίκτυο συμπληρωματικής εκμάθησης (*complementary learning fuzzy neural network – CLFNN*) για να ενισχύσουν τις υπάρχουσες διαγνωστικές μεθόδους. Πιο συγκεκριμένα, προτείνεται ένα ασαφές δίκτυο ελέγχου προσαρμοστικής μάθησης με μια θεωρία προσαρμοστικού συντονισμού, που ονομάζεται FALCON-AART και αποτελεί ένα τύπο του CLFNN, για να καταδείξει την απόδοση του. Σε αντίθεση με τις συμβατικές μεθόδους υπολογιστικής νοημοσύνης, το CLFNN εκμεταλλεύεται την πλευρική αναστολή μεταξύ θετικών και αρνητικών δειγμάτων. Με πιο απλά λόγια, ένα θετικό δείγμα (κακοήθης περίπτωση) θα ενεργοποιήσει ταυτόχρονα τους θετικούς κανόνες και θα αναστείλει τους αρνητικούς κανόνες όταν παρουσιάζονται στο σύστημα, οδηγώντας ως εκ τούτου σε ένα θετικό συμπέρασμα. Επιπλέον, το σύστημα είναι εξοπλισμένο με αυτόνομη εγκατάσταση παραγωγής κανόνων, ενώ η διαδικασία συλλογισμού που παρέχει είναι παρόμοια με του ανθρώπου.

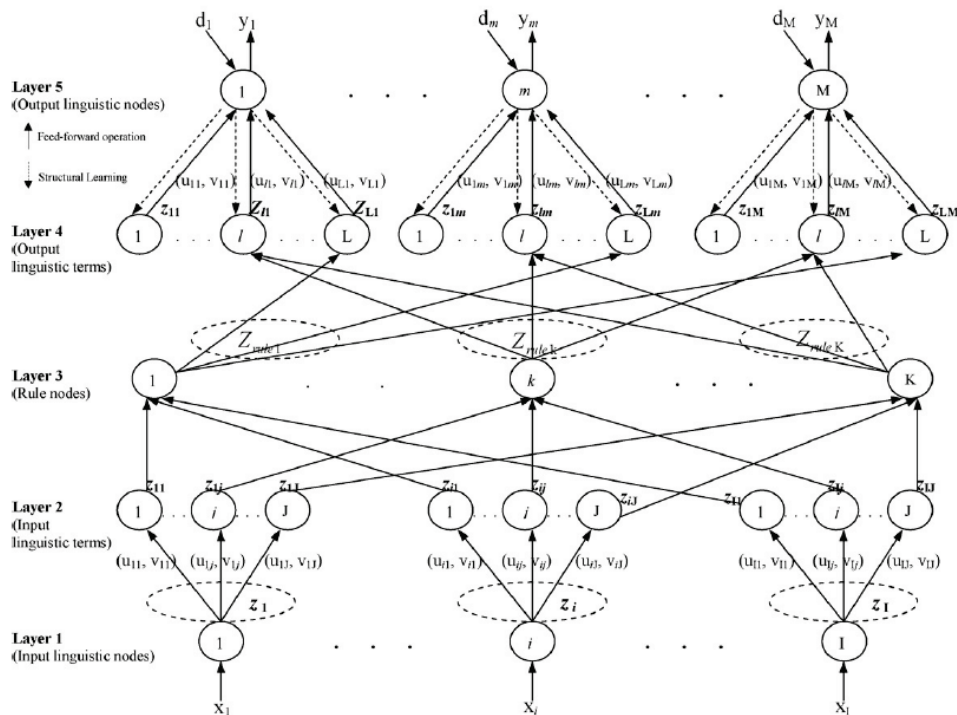
Οι συγγραφείς χρησιμοποιούν τρεις βάσεις δεδομένων στην παρούσα μελέτη :

- I. Μικρο-συστοιχίες γονιδιακής έκφρασης, που περιέχουν 30 δείγματα από όγκους των ωοθηκών και 24 φυσιολογικά δείγματα, καθένα από τα οποία περιλαμβάνει 1536 χαρακτηριστικά.
- II. Αιματολογικές αναλύσεις 172 ασθενών με χαρακτηριστικά που παρατίθενται στον πίνακα που ακολουθεί.
- III. Πρωτεομικά φάσματα, που προέρχονται από ένα σύστημα φασματογράφου μάζας και περιλαμβάνουν 91 υγιή δείγματα ασθενών και 162 δείγματα ασθενών με καρκίνο των ωοθηκών.

$x_1$ : age	$x_{16}$ : D-dimer
$x_2$ : packed cell volume (PCV)	$x_{17}$ : tissue plasminogen activator (tPA) activity
$x_3$ : hemoglobin (Hgb)	$x_{18}$ : tPA antigen
$x_4$ : $\beta$ -Thromboglobulin (TG)	$x_{19}$ : urokinase-like plasminogen activator (uPA) activity
$x_5$ : reaction time of thrombelastography (TEG-r)	$x_{20}$ : uPA antigen
$x_6$ : coagulation time of TEG (TEG-K)	$x_{21}$ : uPA receptor
$x_7$ : maximum amplitude of TEG (TEG-MA)	$x_{22}$ : PA inhibitor 1 (PAI-1) activity
$x_8$ : fibrinogen	$x_{23}$ : PAI-1 antigen
$x_9$ : factor VII	$x_{24}$ : PAI-2 antigen
$x_{10}$ : von Willebrand-factor (VWF)	$x_{25}$ : protein C antigen
$x_{11}$ : thrombin-antithrombin complex (T/AT)	$x_{26}$ : protein S antigen
$x_{12}$ : prothrombin fragment 1 & 2 (F1 + 2)	$x_{27}$ : cancer antigen-125 (CA-125)
$x_{13}$ : antithrombin III (A-TH3) activity	$x_{28}$ : cancer-associated serum antigen (CASA)
$x_{14}$ : A-TH3 antigen	$x_{29}$ : tissue factor pathway inhibitor (TFPI)
$x_{15}$ : plasminogen	$x_{30}$ : blood platelet

**Πίνακας 34 :** Λίστα χαρακτηριστικών που προκύπτουν από αιματολογικές αναλύσεις.

Όπως αναφέρθηκε ήδη, το FALCON-AART ανήκει στην κλάση του CLFNN και δημιουργεί αυτόνομα κανόνες στην μορφή που περιγράφεται από την εξίσωση : «Αν το  $x_1$  είναι Α και το  $x_2$  είναι Β, τότε το  $\psi_1$  είναι Γ και το  $\psi_2$  είναι Δ». Ειδικότερα, αποτελείται από πέντε στρώματα που έχουν αντιστοιχιστεί στα στοιχεία του ασαφούς κανόνα, ενώ η αρχιτεκτονική του καταδεικνύεται στο παρακάτω σχήμα :



**Εικόνα 44 :** Αρχιτεκτονική του FALCON-AART.

Όσον αφορά την πειραματική διαδικασία που ακολουθείται με το FALCON-AART, κάθε δείγμα της βάσης δεδομένων διαιρείται σε θετικές (κοκοήθη) και αρνητικές (καλοήθη και φυσιολογικά) κλάσεις, ενώ δημιουργούνται για κάθε βάση δεδομένων τρία σύνολα εκπαίδευσης και ελέγχου διασταυρούμενης επικύρωσης. Στη συνέχεια, το σύστημα εκπαιδεύεται και αξιολογείται με χρήση των συνόλων εκπαίδευσης/ελέγχου και το FALCON-ART συγκρίνεται με δημοφιλή εργαλεία όπως ο αφελής ταξινομητής Bayes, οι μηχανές υποστήριξης διανυσμάτων, η μέθοδος του κοντινότερου γείτονα κ.ά. όσον αφορά τις τιμές της ειδικότητας και της ευαισθησίας. Παραθέτουμε ενδεικτικά ένα πίνακα



σύγκρισης του FALCON-AART με άλλες μεθόδους ταξινόμησης με χρήση των αιματολογικών εξετάσεων :

Method	Sensitivity (%)	Specificity (%)	Accuracy (%)	Training time (s)	Number of rules
NB	73.20	60.00	67.60	0.02	NA
RBF	70.70	50.00	62.00	0.10	NA
ANFIS	92.00	55.00	74.30	2.60	10
SVM	100.00	33.30	60.00	0.05	NA
C4.5	100.00	50.00	78.90	20.00	4
kNN	87.80	43.30	69.00	0.01	NA
MLP	90.20	50.00	73.20	1.90	NA
FALCON-AART	95.00	57.00	78.90	0.50	5

**Πίνακας 35 :** Πίνακας σύγκρισης διαφόρων μεθόδων ταξινόμησης όσον αφορά την πρόβλεψη του καρκίνου των ωοθηκών με χρήση των αιματολογικών εξετάσεων.

Καταλήγοντας, οι συγγραφείς υποστηρίζουν ότι τα προτεινόμενα σύστημα επέδειξε την υπεροχή του ως υπολογιστικό σύστημα υποστήριξης αποφάσεων για τη διάγνωση του καρκίνου των ωοθηκών. Γρήγορη εκπαίδευση, απλή παραγωγή ασαφών κανόνων και υψηλή ακρίβεια είναι μερικά από τα πλεονεκτήματα του FALCON-AART σε αυτήν την μελέτη, που ξεπερνούν σε ακρίβεια τα συμβατικά τεχνητά νευρωνικά δίκτυα και τις στατιστικές μεθόδους. Το σημαντικότερο πλεονέκτημά του όμως είναι η δυνατότητα να παράγει συμπληρωματικούς ασαφείς κανόνες για τη συλλογιστική του διαδικασία που προσεγγίζουν στενά την ανθρώπινη διαδικασία συλλογισμού, καθιστώντας έτσι την απόφαση διάγνωσης ακριβής και συνεπής. Έτσι, το ασαφές νευρωνικό δίκτυο συμπληρωματικής εκμάθησης αποτελεί ένα πολλά υποσχόμενο εργαλείο για την υποστήριξη κλινικών αποφάσεων.

### 3.10 Μελέτες σχετικά με τον καρκίνο στο θυρεοειδή (Thyroid nodules)

Κλινικός σκοπός της μοναδικής μελέτης που εξετάστηκε όσον αφορά τον καρκίνο του θυρεοειδούς με τίτλο «Ανάπτυξη ενός μοντέλου κλινικών αποφάσεων για θυρεοειδικούς όζους» ήταν η διάγνωση του καρκίνου [23].

Όπως επισημαίνουν οι συγγραφείς στην παρούσα μελέτη, οι θυρεοειδικοί όζοι αντιπροσωπεύουν ένα σύνθετο πρόβλημα που χρήζει ιατρικής περίθαλψης. Ένα ποσοστό της τάξης του 4 έως 7% του ενήλικου πληθυσμού στις Ηνωμένες Πολιτείες (10 – 18 εκατομμύρια άνθρωποι) έχει ένα ψηλαφητό όζο στο θυρεοειδή, ωστόσο η πλειονότητα των όζων (>95%) αυτών είναι καλοήθεις. Αν και η παρακέντηση με λεπτή βελόνα (fine needle aspiration – FNA) παραμένει το πιο αποδοτικό και ακριβές διαγνωστικό εργαλείο για θυρεοειδικούς όζους στην πράξη (εξετάζονται επίσης ενδελεχώς το επίπεδο θυρεοτροπίνης στον ορό/TSH, η φυσική κατάσταση και ο υπέρηχος θυρεοειδούς), πάνω από το 20% των ασθενών που υποβάλλονται σε παρακέντηση παρουσιάζουν ακαθόριστη κυτταρολογική εξέταση (νεόπλασμα ωοθυλακίων) με συχνότητα συσχετιζόμενου κινδύνου κακοήθειας της



τάξης του 20 – 30%. Για αυτούς τους ασθενείς απαιτείται θυρεοειδική λοβεκτομή/αφαίρεση του μεσαίου τμήματος του θυρεοειδή καθαρά για το σκοπό της επίτευξης μιας οριστικής διάγνωσης. Δεδομένου ότι η πλειοψηφία (70 – 80%) αυτών των ασθενών έχουν καλοήγη χειρουργική παθολογία, η θυρεοειδεκτομή σε αυτούς διεξάγεται κυρίως με διαγνωστική πρόθεση. Επομένως, κρίνεται απαραίτητη η ανάπτυξη κλινικών μοντέλων πρόβλεψης του κινδύνου κακοήθειας για την υποστήριξη αποφάσεων θεραπείας σε ασθενείς με θυρεοειδικούς όζους, ώστε να μειωθεί η συχνότητα περιττών διαγνωστικών χειρουργικών εξετάσεων. Ο πρωταρχικός στόχος, λοιπόν, των συγγραφέων σε αυτή την μελέτη είναι να καθορίσουν το κατά πόσον είναι εφικτό ένα προγνωστικό Μπεϋζιανό μοντέλο να βοηθήσει τον κλινικό ιατρό στην ερμηνεία διαγνωστικών πληροφοριών.

Οι συγγραφείς επιχείρησαν, χρησιμοποιώντας μια πλατφόρμα μηχανικής μάθησης, να κατασκευάσουν ένα δίκτυο Μπεϋζιανής πίστης, το οποίο θα μπορούσε να εφαρμοστεί σε ένα συγκεκριμένο πληθυσμό ασθενών υπό μελέτη που εκπροσωπείται από σχετικές κλινικές μεταβλητές (πχ ηλικία του ασθενούς, φύλο, χαρακτηριστικά υπερήχου και εμπέδησης και κυτταρολογική εξέταση FNA), με σκοπό τη δημιουργία ενός εργαλείου εκτίμησης κινδύνου για κακοήθεια.

Αρχικά, συλλέχθηκαν και αναλύθηκαν δεδομένα από μια ολοκληρωμένη ομαδική μελέτη σε μια περίοδο τεσσάρων χρόνων με τη συμμετοχή 216 ασθενών με θυρεοειδικούς όζους που υποβλήθηκαν σε υπέρηχο (ultrasound – US), σάρωση ηλεκτρικής αντίστασης (electrical impedance scanning – EIS) και κυτταρολογική εξέταση παρακέντησης με λεπτή βελόνα (fine needle aspiration cytology – FNA) πριν από τη θυρεοειδεκτομή. Βρέθηκε μάλιστα ότι οι 109 ασθενείς θα έπρεπε να υποβληθούν σε διαγνωστική θυρεοειδεκτομή λόγω ασαφούς κυτταρολογικής εξέτασης. Τα κλινικά, τα βασιζόμενα σε εικόνες, τα κυτταρολογικά και τα παθολογικά χαρακτηριστικά του πληθυσμού των ασθενών που χρησιμοποιήθηκαν στο μοντέλο απεικονίζονται στον παρακάτω πίνακα :

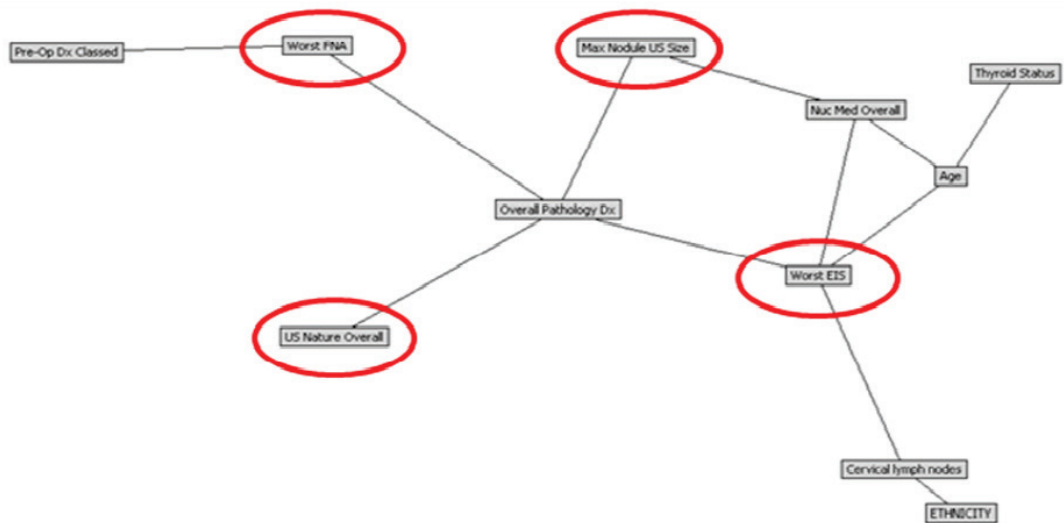
Patient Characteristics	No.	%
<b>Gender</b>		
Male	46	21.3
Female	170	78.7
Patient age, years	mean $\pm$ SD = 47.1 $\pm$ 15.9	
Median (range)	47 (18 – 85)	
<b>Disease Characteristics</b>		
Serum TSH	mean $\pm$ SD = 2.6 $\pm$ 6.8	
<b>Pre-operative thyroid status</b>		
Euthyroid	174	80.6
Hyperthyroid	26	12
Hypothyroid	16	7.4
Dominant thyroid nodule size (cm)	mean $\pm$ SD = 2.8 $\pm$ 1.6	
<b>Dominant thyroid nodule size</b>		
<2 cm	78	35.9
2 – 4 cm	101	46.5
>4 cm	38	17.5
<b>Imaging Characteristics</b>		
<b>Thyroid Scintigraphy</b>		
Cold	21	58.3
Warm	3	8.3
Hot	12	33.3
<b>Thyroid Ultrasound</b>		
Simple cyst	4	1.9
Complex cyst	8	3.7
Mixed	32	14.8

Solid	No.	%
<b>Thyroid EIS Level of Suspicion</b>		
I: Definitely benign	22	10.1
II: Highly unlikely to be malignant	49	22.5
III: Unlikely to be malignant	21	9.6
IV: Likely to be malignant	64	29.5
V: Highly likely to be malignant	62	28.4
<b>Pathological Characteristics</b>		
<b>Fine needle aspiration cytology</b>		
Inadequate	6	2.8
Not done	9	4.2
Negative	30	13.9
Positive	62	28.7
Indeterminate	109	50.4
<b>Surgical Histopathology</b>		
Benign	106	49.1
Malignant	110	50.9

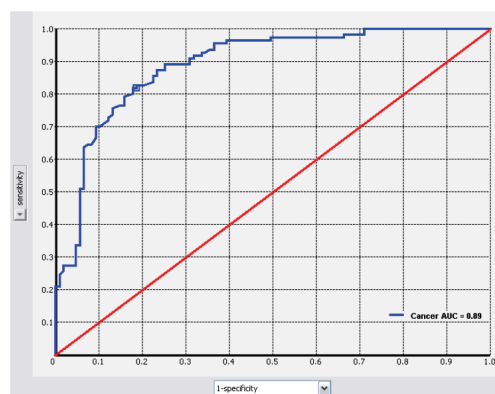
**Πίνακας 36 :** Πίνακας χαρακτηριστικών του υπό μελέτη πληθυσμού.

Το Μπεϋζιανό μοντέλο που σχεδιάστηκε για την πρόβλεψη της κακοήθειας των θυρεοειδικών όζων βασίστηκε σε πολυμεταβλητές σχέσεις εξάρτησης ανάμεσα σε ανεξάρτητες συμμεταβλητές, ενσωματώνοντας σχέσεις «γονέα – παιδιού» μεταξύ των κόμβων στον κατευθυνόμενο γράφο και παρέχοντας «εκ των υστέρων» εκτιμήσεις πιθανοτήτων (posterior probabilities). Στη συνέχεια, εκτελέστηκε διασταυρούμενη επικύρωση 10 φορών για να εκτιμηθεί το σφάλμα του ταξινομητή, όπου το σύνολο των δεδομένων διαχωρίστηκε τυχαία σε δέκα ξεχωριστά και μοναδικά σύνολα εκπαίδευσης και δοκιμών. Το σύνολο εκπαίδευσης συμπεριέλαβε το 90% των αρχείων, ενώ το σύνολο δοκιμών το 10% των αρχείων. Υπολογίστηκε η καμπύλη δέκτη λειτουργίας χαρακτηριστικών (ROC curve), η περιοχή κάτω από την καμπύλη (AUC), η θετική και η αρνητική προγνωστική τιμή (PPV-positive predictive value, NPV-negative predictive value) για τον προσδιορισμό της αρτιότητας (ευρωστίας) του μοντέλου στην πρόβλεψη της κακοήθειας των θυρεοειδικών όζων. Παρακάτω αποδίδεται σχηματικά το Μπεϋζιανό μοντέλο :



**Εικόνα 45 :** Μοντέλο δικτύου Μπεϋζιανής πίστης : Παθολογική διάγνωση στους θυρεοειδικούς όζους (Καλοήθειες έναντι Κακοηθιών). Η δομή του μοντέλου καθορίζει τέσσερις βασικούς προγνωστικούς δείκτες της ιστοπαθολογίας του θυρεοειδικού όζου (κόκκινοι κύκλοι) : κυτταρολογική εξέταση παρακέντησης με λεπτή βελόνα (FNA), μέγιστο μέγεθος όζου (όπως καθορίζεται από τον υπέρηχο), σάρωση ηλεκτρικής αντίστασης (EIS) και χαρακτηριστικά υπερήχου (US) του όζου.

Το μέγεθος του θυρεοειδικού όζου, η κυτταρολογική εξέταση FNA, τα χαρακτηριστικά της σάρωσης EIS και των υπερήχων US επέδειξαν υψηλή ικανότητα πρόβλεψης της κακοήθειας. Η διασταυρούμενη επικύρωση του μοντέλου που κατασκευάστηκε με την Μπεϋζιανή ανάλυση δικτύου προέβλεψε αποτελεσματικά την κακοήθεια [AUC = 0,88 (95%CI: 0,82–0,94)] των θυρεοειδικών όζων. Οι θετικές και οι αρνητικές προγνωστικές τιμές του μοντέλου ήταν 83% (95%CI: 76%–91%) και 79% (95%CI: 72%–86%), αντίστοιχα, ενώ η ευαισθησία και η ειδικότητα βρέθηκαν ίσες με 82% (95%CI: 74%–91%) και 77% (95%CI: 68%–86%).



**Εικόνα 46 :** Καμπύλη ROC για την πρόβλεψη καρκίνου. Η ευαισθησία απεικονίζεται στον άξονα y και η ειδικότητα στον άξονα x.

Συμπεραίνοντας, θα μπορούσαμε να πούμε ότι ένα ολοκληρωμένο προγνωστικό μοντέλο αποφάσεων, βασισμένο σε ένα δίκτυο Μπεϋζιανής πίστης, που ενσωματώνει εύκολα ανακτήσιμες μετρήσεις θυρεοειδικών όζων είναι σημαντικό για την κλινική πράξη,

καθώς προβλέπει ικανοποιητικά την κακοήθεια των θυρεοειδικών όζων. Ωστόσο, το μοντέλο αυτό, όπως και τα περισσότερα που συναντάμε, απαιτεί περαιτέρω επικύρωση σε προβλεπόμενες κλινικές δοκιμές, η οποία αν κριθεί επιτυχημένη καθιστά το παραπάνω μοντέλο χρήσιμο κλινικό εργαλείο.

### **3.11 Μελέτες σχετικά με τον καρκίνο του ήπατος (Liver cancer)**

Όσον αφορά τον καρκίνο του ήπατος, μελετήθηκε ένα σύστημα που εξέτασε την πρόβλεψη της εμφάνισης του εν λόγω καρκίνου. Τίτλος της παρούσας μελέτης ήταν ο εξής : «Ανάλυση δεδομένων πληθυσμιακού ελέγχου για μοντέλο πρόβλεψης ηπατικής νόσου με χρήση καμπύλης ανάπτυξης» [24].

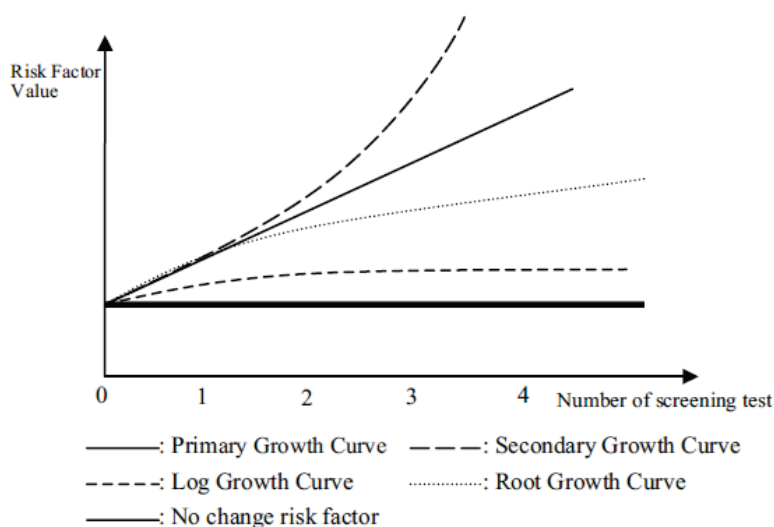
Στην μελέτη αυτή, ο πρώτος στόχος της έρευνας των συγγραφέων ήταν να εκτιμήσουν τα επακόλουθα της ασθένειας του ήπατος και να διαμορφώσουν ένα μοντέλο πρόβλεψης της εμφάνισης της εν λόγω ασθένειας, με βάση τους ανθρώπους που συμμετείχαν στο τεστ πληθυσμιακού ελέγχου στο διάστημα 1994 έως 2001 στην Κορέα. Δεύτερη επιδίωξη αποτέλεσε η διαμόρφωση του μοντέλου πρόβλεψης της ασθένειας του ήπατος εκτιμώντας τους παράγοντες κινδύνου της νόσου μέσω της καμπύλης ανάπτυξης, όταν οι ασθενείς υποβληθούν σε τεστ πληθυσμιακού ελέγχου στο μέλλον. Αυτό το μοντέλο πρόβλεψης μπορεί να παίζει σημαντικό ρόλο, με την έννοια της πρόληψης, σαν μια πρωταρχική διάγνωση του καρκίνου του ήπατος. Η μέθοδος ανάλυσης σε αυτή την μελέτη είναι η υπόθεση του παράγοντα κινδύνου των αιτιών της ηπατικής νόσου με χρήση της καμπύλης ανάπτυξης.

Το τεστ πληθυσμιακού ελέγχου αποτελεί μια εξέταση που στοχεύει στην κατανόηση της διάγνωσης σε πρώιμο στάδιο της ασθένειας ή της τρέχουσας κατάστασης της υγείας, εξετάζοντας τους παράγοντες κινδύνου κάθε ασθένειας, εν προκειμένη του καρκίνου του ήπατος. Τα δεδομένα που χρησιμοποιήθηκαν στο τεστ πληθυσμιακού ελέγχου (παράγοντες κινδύνου για καρκίνο του ήπατος, τοπικά και διεθνώς) ήταν η ηλικία, το φύλο, ο δείκτης μάζας σώματος, λίστες εξετάσεων των λειτουργιών του ήπατος, η αλφαφετοπρωτεΐνη, ο διαβήτης, η διατροφική συμπεριφορά, η επαγγελματική ενασχόληση και τα τεστ υπερήχων, τα οποία μπορούν να ανιχνεύσουν διαταραχές όπως όγκους ή κίρρωση του ήπατος που είναι σημαντικές εξαρτημένες μεταβλητές στο μοντέλο πρόβλεψης.

Ως μοντέλα ταξινόμησης για την πάθηση του ήπατος στην παρούσα μελέτη χρησιμοποιήθηκαν τα εξής : λογιστική παλινδρόμηση, δέντρα αποφάσεων και νευρωνικά δίκτυα. Σε αυτό το σημείο, κρίνουμε σκόπιμο να αναφερθούμε στη λογιστική παλινδρόμηση ως μια ευρέως χρησιμοποιημένη στατιστική τεχνική μοντελοποίησης. Πιο συγκεκριμένα, η λογιστική παλινδρόμηση χρησιμοποιείται κυρίως για να περιγράψει τη σχέση της πιθανότητας (π.χ. βελτίωσης ασθένειας) ενός χαρακτηριστικού με διάφορους παράγοντες (θεραπεία, φάρμακο A,B,..., φύλο, κτλ). Το μοντέλο της λογιστικής παλινδρόμησης μοιάζει με το αντίστοιχο της απλής παλινδρόμησης με τη διαφορά ότι η μεταβλητή απόκρισης (εξαρτημένη μεταβλητή) είναι δίτιμη, δηλαδή περιγράφει το αποτέλεσμα ως «επιτυχία-αποτυχία», π.χ. βελτίωση – μη βελτίωση κατάστασης ασθενούς

ύστερα από θεραπεία. Όσον αφορά την μελέτη που εξετάζεται, σημειώνουμε ότι στο μοντέλο ταξινόμησης χρησιμοποιήθηκαν δεδομένα ασθενών οι οποίοι υποβλήθηκαν στο τεστ πληθυσμιακού ελέγχου περισσότερες από δύο φορές, ενώ τα παλαιότερα δεδομένα αποτέλεσαν το σύνολο εκπαίδευσης και τα πιο πρόσφατα το σύνολο επικύρωσης. Οι μεταβλητές που χρησιμοποιήθηκαν για τη δόμηση του μοντέλου ήταν εξαρτημένες μεταβλητές που παράχθηκαν από τα αποτελέσματα των τεστ υπερήχων (δυαδικά), ενώ οι ανεξάρτητες μεταβλητές ήταν ολόκληρες οι λίστες του τεστ πληθυσμιακού ελέγχου.

Για την καμπύλη ανάπτυξης ακολουθήθηκε η εξής διαδικασία : ο αριθμός των ασθενών που υποβλήθηκαν σε τεστ πληθυσμιακού ελέγχου περισσότερες από τέσσερις φορές ήταν 1030 (σύνολο εκπαίδευσης για το μοντέλο πρόβλεψης), εκ των οποίων οι 430 (σύνολο επικύρωσης για το μοντέλο πρόβλεψης) υποβλήθηκαν και σε πέμπτο τεστ. Η εκτίμηση της πέμπτης δοκιμής προέκυψε από την ανάλυση της καμπύλης ανάπτυξης, η οποία βασίστηκε στην τιμή των μεταβλητών γνωστών ως παραγόντων κινδύνου της ασθένειας του ήπατος μέχρι την τέταρτη δοκιμή. Στην καμπύλη ανάπτυξης, συμπεριλήφθηκαν μόνο οι παράγοντες κινδύνου με σημαντικές αλλαγές. Με βάση αυτούς που υποβλήθηκαν στο τεστ τέσσερις ή περισσότερες φορές, η τιμή του παράγοντα κινδύνου μπορεί να συναχθεί με το να υποβληθούν στο τεστ στο οποίο έχει ενσωματωθεί το μοντέλο πρωτοβάθμιας καμπύλης ανάπτυξης, δευτεροβάθμιας καμπύλης ανάπτυξης, λογαριθμικής καμπύλης ανάπτυξης, καμπύλης ανάπτυξης ρίζας κτλ, οι οποίες απεικονίζονται στο παρακάτω γράφημα :

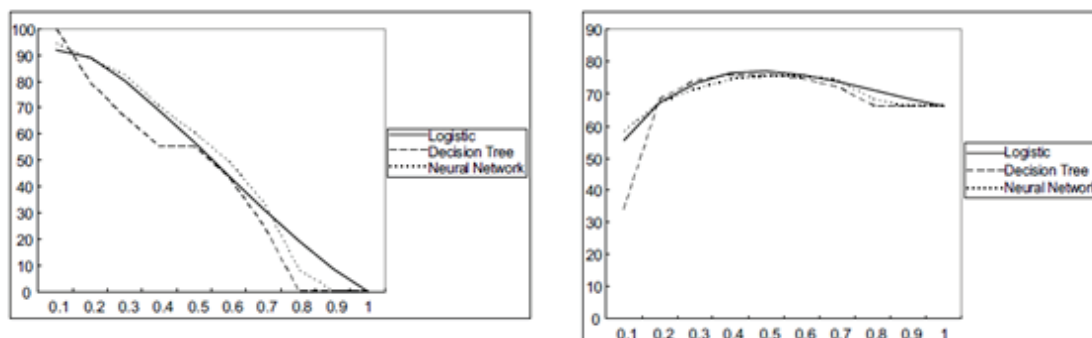


**Εικόνα 47 :** Διάφορα μοντέλα καμπύλης ανάπτυξης για την πρόβλεψη παράγοντα κινδύνου.

Μέσα από τις διάφορες μορφές των μοντέλων καμπύλης ανάπτυξης, μπορεί να προβλεφθεί η πέμπτη φορά και η ακρίβεια της πρόβλεψης αυτής μπορεί να μετρηθεί ως ποσοστό σφάλματος. Συνολικά, η λογαριθμική καμπύλη ανάπτυξης παρουσιάζει μικρότερο ποσοστό σφάλματος ανάμεσα στην πραγματική τιμή και την εκτιμώμενη.

Όσον αφορά τους παράγοντες κινδύνου της ηπατικής νόσου, ο πιο σημαντικός παράγοντας αποδεικνύεται η τιμή λειτουργίας BioAlb, όπως επίσης και η εικόνα της ηπατικής λειτουργίας, ο δείκτης μάζας σώματος, τα λιπίδια του ορού, τα στοιχεία ελέγχου

του διαβήτη. Εν γένει, διαπιστώθηκε ότι όλη η λίστα του τεστ πληθυσμιακού προσυμπτωματικού ελέγχου μπορεί να αποτελέσει σημαντικό παράγοντα κινδύνου. Οι συγγραφείς ισχυρίζονται επίσης μέσω της έρευνάς τους ότι όσο συχνότερα υποβάλλεται κάποιος σε τεστ πληθυσμιακού προσυμπτωματικού ελέγχου, τόσο μικρότερο είναι το ποσοστό εμφάνισης διαταραχής του ήπατος ( $\chi^2 - test\ statistic = 67,13, DF = 3, P - value = 0,0001$ ).



**Εικόνες 48 & 49 :** Διαγράμματα ευαισθησίας και ακρίβειας του συνόλου επικύρωσης των μοντέλων ταξινόμησης της ηπατικής νόσου.

Διερευνώντας τις επιπτώσεις της ασθένειας του ήπατος υπό την οπτική γωνία του μέλλοντος, αυτή η μελέτη επιχειρήσε να εκτιμήσει τους παράγοντες κινδύνου μέσω διάφορων αναλύσεων καμπυλών ανάπτυξης και συμπεριέλαβε χρήση λογιστικής παλινδρόμησης, δέντρων αποφάσεων και νευρωνικών δικτύων. Στην περίπτωση του νευρωνικού δικτύου με χρήση εκτίμησης καμπύλης ανάπτυξης της μορφής  $X_{i(5)} = a_i + \beta_i * T + \varepsilon_{iT}$ , η ακρίβεια πρόβλεψης της ασθένειας του ήπατος ήταν 72,55% και η ευαισθησία ήταν 78,62%. Από την άλλη πλευρά, στην περίπτωση του μοντέλου πρόβλεψης ηπατικής νόσου με χρήση εκτιμήσεων πρόσφατων δεδομένων πληθυσμιακού ελέγχου, η ακρίβεια ήταν 72,09% και η ευαισθησία ήταν 71,72%. Αυτά τα ποσοστά είναι χαμηλότερα από την προηγούμενη ανάλυση που περιελάμβανε καμπύλη ανάπτυξης. Συμπεραίνουμε, εν κατακλείδι, ότι σε διαφορετικά μοντέλα πρόβλεψης στα οποία χρησιμοποιήθηκε καμπύλη ανάπτυξης, η τιμή της ευαισθησίας βελτιώθηκε.

### 3.12 Μελέτες σχετικά με τον καρκίνο των οστών (Skeletal cancer)

Η κατηγορία που αφορά τον καρκίνο των οστών συμπεριέλαβε μία μελέτη με στόχο την πρόβλεψη επιβίωσης των ασθενών, και πιο συγκεκριμένα την «Εκτίμηση προσδόκιμου ζωής σε ασθενείς με χειρουργήσιμη σκελετική μετάσταση : Μια εφαρμογή δικτύου Μπεϋζιανής πίστης (Bayesian Belief network)», όπως είναι και ο τίτλος της μελέτης [25].

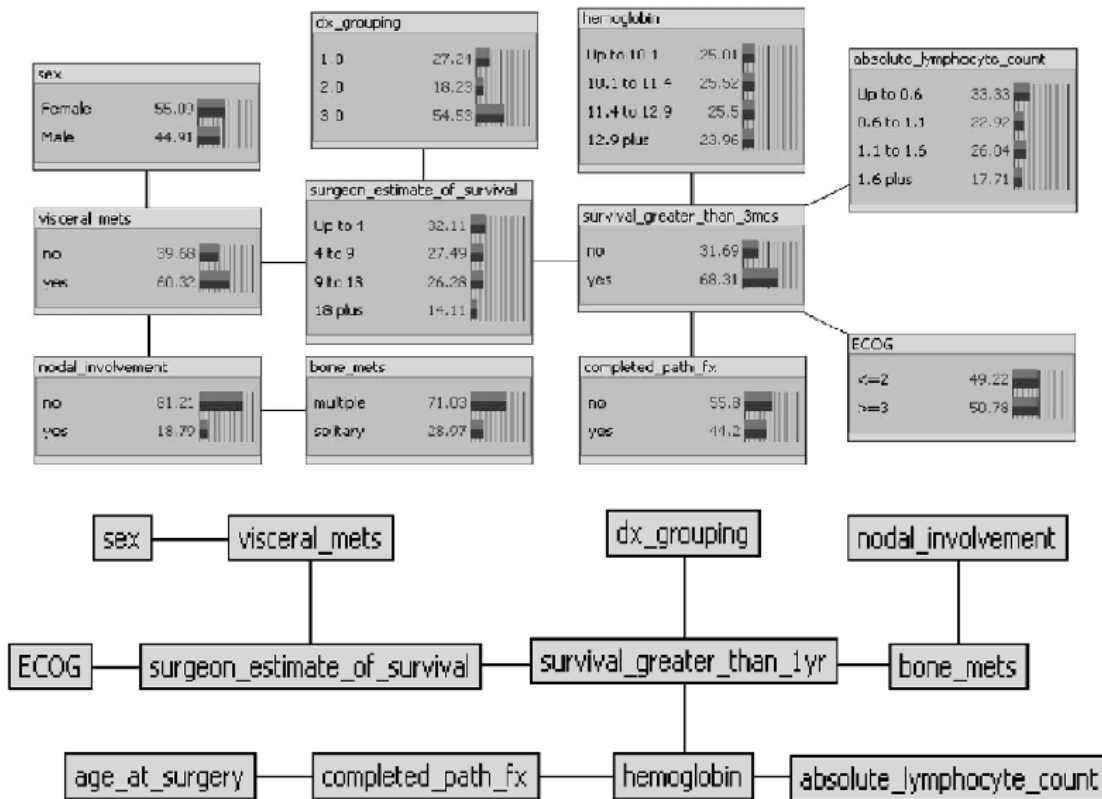
Στην παρούσα εργασία, αντικείμενο μελέτης αποτελούν οι ασθενείς με μεταστατικό καρκίνο που επηρεάζει τα άκρα. Οι ακριβείς εκτιμήσεις του προσδόκιμου ζωής τους είναι

σημαντικές για τη διαμόρφωση των προσδοκιών τόσο των ασθενών και της οικογένειας τους, αλλά και του ίδιου του ιατρού. Κλινικά, η απόφαση για τον αν θα πρέπει να χειρουργηθούν οι ασθενείς με σκελετική μετάσταση, καθώς και η επιλογή της χειρουργικής διαδικασίας, βασίζεται στην ατομική εκτίμηση επιβίωσης του ασθενή.

Επί του παρόντος, δεν υπάρχουν αξιόπιστες μέθοδοι για την εκτίμηση επιβίωσης σε αυτή την μερίδα ασθενών. Η Μπεϋζιανή ταξινόμηση (Bayesian classification), που περιλαμβάνει την μοντελοποίηση του δικτύου Μπεϋζιανής πίστης (Bayesian Belief network – BBN) είναι μια στατιστική μέθοδος που διερευνά υποθετικές, πιθανοτικές σχέσεις μεταξύ των μεταβλητών για την εκτίμηση της πιθανότητας ενός αποτελέσματος χρησιμοποιώντας τα παρατηρούμενα δεδομένα. Έτσι, τα μοντέλα BBN χρησιμοποιούνται με αυξανόμενη συχνότητα σε μια ποικιλία διαγνώσεων για να κωδικοποιήσουν πολύπλοκα κλινικά δεδομένα σε προγνωστικά μοντέλα. Ο σκοπός των συγγραφέων στην μελέτη αυτή είναι να καθορίσουν κατά πόσον είναι εφικτή η δημιουργία ταξινομητών Bayes με σκοπό την εκτίμηση της επιβίωσης σε ασθενείς που υποβάλλονται σε χειρουργική επέμβαση για μεταστάσεις στα οστά του άξονα του σώματος (όπως κρανίο, σπονδυλική στήλη, πλευρά και στέρνο – axial skeleton) και στα οστά των άκρων και τα θωρακικά (appendicular skeleton).

Οι συγγραφείς ερεύνησαν τη βάση διαχείρισης δεδομένων ασθενών του Αντικαρκινικού Κέντρου Memorial Sloan-Kettering, αναζητώντας όλους τους ασθενείς που υποβλήθηκαν σε χειρουργική επέμβαση για σκελετικές μεταστάσεις στο διάστημα 1999 έως 2003. Στη συνέχεια, ανέπτυξαν και εκπαίδευσαν ένα BBN μοντέλο μηχανικής μάθησης προκειμένου να εκτιμήσει την επιβίωση σε μήνες, χρησιμοποιώντας 15 υποψήφια χαρακτηριστικά. Τα χαρακτηριστικά αυτά επιλέχθηκαν με βάση την τρέχουσα κλινική ή ιστορική συσχέτισή τους με την επιβίωση ασθενών με χειρουργικά αντιμετωπίσιμη σκελετική μετάσταση και περιλαμβάνουν την ηλικία τη στιγμή της χειρουργικής επέμβασης, τη φυλή, το φύλο, τον εκτιμώμενο ρυθμό σπειραματικής διήθησης, τη συγκέντρωση ασβεστίου στον ορό, τη συγκέντρωση αλβουμίνης στον ορό, την ένδειξη για χειρουργική επέμβαση (επικείμενο ή ολοκληρωμένο παθολογικό κάταγμα), τον αριθμό των μεταστάσεων στα οστά (μεμονωμένες ή πολλαπλές), την εκτίμηση του χειρουργού για την επιβίωση (μετεχειρητικά, σε μήνες), την παρουσία ή απουσία μεταστάσεων στα σπλάχνα, την παρουσία ή απουσία μεταστάσεων στους λεμφαδένες, προγενέστερη χημειοθεραπεία, την προεγχειρητική αιμογλοβίνη (κατά την εισαγωγή, πριν την μετάγγιση, εάν υπάρχει), την απόλυτη μέτρηση λεμφοκυττάρων και την αρχική ογκολογική διάγνωση. Αξίζει να σημειωθεί ότι εκπαιδεύτηκαν ξεχωριστοί ταξινομητές της εκτιμώμενης επιβίωσης για 3 και 12 μήνες, διότι, τα δίκτυα BBN, ως παραμετρική μεθοδολογία μοντελοποίησης, δεν είναι κατάλληλα στο να παρέχουν διακριτές εκτιμήσεις στο χρόνο, ενώ θεωρήθηκε αυτά τα χρονικά σημεία διαφοροποιούν τη λήψη αποφάσεων για τη χειρουργική διαδικασία.

Από τους 189 ασθενείς που τελικά επιλέχθηκαν, το 90 % αποτέλεσε το σύνολο εκπαίδευσης και το 10 % το σύνολο δοκιμής του δικτύου BBN. Σημειώνεται ότι οι προγνωστικοί δείκτες πρώτου βαθμού διαφοροποιήθηκαν ανάμεσα στα τρίμηνα και στα δωδεκάμηνα μοντέλα, όπως απεικονίζεται και παρακάτω :



**Εικόνες 50 & 51 :** Τρίμηνο μοντέλο BBN με μεταγενέστερες κατανομές που απεικονίζονται ως ποσοστά (%) του πληθυσμού εκπαίδευσης και δωδεκάμηνο μοντέλο BBN.

Πραγματοποιήθηκε διασταυρούμενη επικύρωση 10 φορές (ten-fold cross-validation) και ανάλυση της χαρακτηριστικής καμπύλης λειτουργίας (receiver operating characteristics curve – ROC) ώστε να αξιολογηθεί η ακρίβεια και η αρτιότητα των τελικών τρίμηνων και δωδεκάμηνων μοντέλων. Μετά τη διασταυρούμενη επικύρωση, η περιοχή AUC κάτω από την καμπύλη ROC ήταν 0,85 (95% CI: 0,80-0,93) για τρίμηνη πιθανότητα επιβίωσης και 0,83 (95% CI: 0,77-0,90) για δωδεκάμηνη πιθανότητα επιβίωσης.

Οι συγγραφείς ισχυρίζονται ότι ανέπτυξαν επιτυχώς ένα άρτιο, ακριβές, πιθανοτικό αφελές BBN μοντέλο χρησιμοποιώντας τα παρατηρούμενα δεδομένα για την εκτίμηση επιβίωσης εξατομικευμένα σε ασθενείς με εγχειρήσιμη σκελετική μετάσταση. Ωστόσο, πρέπει να αναφέρουμε ότι δεν έγιναν συγκρίσεις με άλλα υπάρχοντα μοντέλα, ενώ η μέθοδος αυτή απαιτεί περαιτέρω ανάπτυξη και πρέπει να επικυρωθεί εξωτερικά σε άλλους, ανεξάρτητους πληθυσμούς ασθενών ώστε να αξιολογηθεί επαρκώς.

### 3.13 Μελέτες σχετικά με την χημειοθεραπεία (Chemotherapy)

Η μελέτη που εξετάστηκε και αφορά την χημειοθεραπεία είχε ως στόχο τη λήψη απόφασης για την επαναχορήγηση ή μη φαρμακευτικής αγωγής όταν έχει προηγηθεί



ανεπιθύμητη αντίδραση. Τίτλο της παρούσας μελέτης αποτέλεσε ο εξής : «Πειραματική μελέτη πάνω στην ανάπτυξη ενός εργαλείου υποστήριξης αποφάσεων για την καθοδήγηση εκ νέου χορήγησης χημειοθεραπευτικού μέσου μετά από σοβαρή ανεπιθύμητη αντίδραση στο φάρμακο» [26].

Μέχρι στιγμής, δεν υπάρχουν τυποποιημένες κατευθυντήριες γραμμές για τη σύσταση επαναχορήγησης ενός χημειοθεραπευτικού φαρμάκου σε έναν ασθενή μετά από ένα επεισόδιο σοβαρής ανεπιθύμητης ενέργειας, ενώ ίσως είναι απαραίτητο να του χορηγηθεί εκ νέου λόγω έλλειψης εναλλακτικών θεραπειών. Η απόφαση για το κατά πόσον πρέπει το φάρμακο να επαναχορηγηθεί στον ασθενή βασίζεται στην εμπειρία του ιατρού και είναι υποκειμενική σε μεγάλο βαθμό. Αυτό, λοιπόν, που επιδιώκουν οι συγγραφείς στην μελέτη αυτή είναι να αναπτύξουν ένα εργαλείο υποστήριξης αποφάσεων, με χρήση μεθόδων εξόρυξης δεδομένων, προκειμένου να βοηθήσουν τους γιατρούς σε αυτή τη διαδικασία λήψης αποφάσεων.

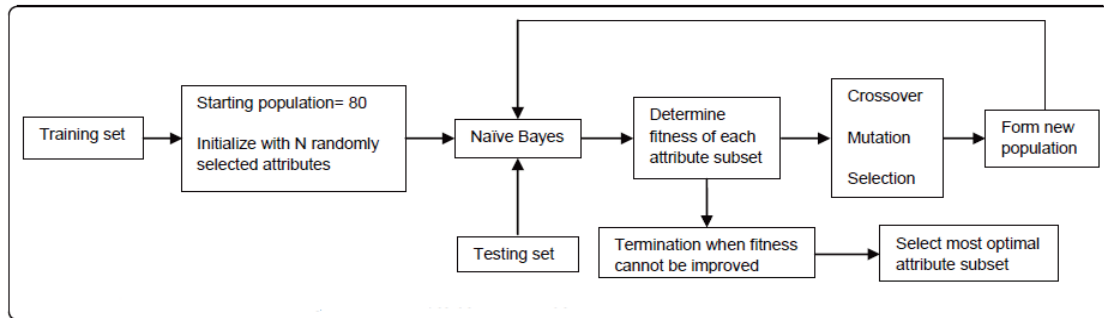
Τα κριτήρια ένταξης των ασθενών σε αυτή την μελέτη είναι : (1) να έχουν υποβληθεί σε χημειοθεραπεία στο Εθνικό Κέντρο Καρκίνου της Σιγκαπούρης, (2) να έχουν υποστεί σοβαρές ανεπιθύμητες ενέργειες και (3) να τους έχει χορηγηθεί εκ νέου το φάρμακο. Από 854 ασθενείς που νοσηλεύτηκαν, οι 81 αντιμετώπισαν σοβαρές ανεπιθύμητες ενέργειες και από αυτούς οι 46 ασθενείς βρέθηκαν να πληρούν και τα τρία κριτήρια ένταξης.

Αρχικά, χρησιμοποιήθηκε ένας γενετικός αλγόριθμος ως μέθοδος επιλογής χαρακτηριστικών για να εντοπιστούν οι προγνωστικοί δείκτες (χαρακτηριστικά) που μπορούν να διακρίνουν τους ασθενείς ανάλογα με την εμφάνιση ή μη ανεπιθύμητων ενεργειών μετά από επαναχορήγηση του φαρμάκου. Ο γενετικός αλγόριθμος δύναται να εντοπίσει αρκετούς συνδυασμούς προγνωστικών δεικτών που είναι εξίσου καλοί. Στο γενετικό αλγόριθμο εισήχθησαν 53 στο σύνολο χαρακτηριστικά, εκ των οποίων κάποια συγκεντρώθηκαν από τους 46 ασθενείς και τα υπόλοιπα κατασκευάστηκαν. Τα χαρακτηριστικά των ασθενών που συγκεντρώθηκαν απεικονίζονται στον παρακάτω πίνακα :

Attributes			
<i>Demographics</i>	- Age	- Gender	- Ethnicity
	- Weight		
<i>Medical Conditions</i>	- Drug allergy	- Cancer type	- Cancer malignancy
	- Comorbidities	- Symptoms of the ADRs	- Hospitalization for prior ADRs
	- ADRs affected organ systems	- Onset of ADRs	
<i>Medications</i>	- Number of cycles	- Chemotherapeutic drug	- Chemotherapeutic drug class
	- Number of doses	- Concurrent medications	
	- Dose reduction on rechallenge	- Rechallenge on same day	
<i>Laboratory parameters</i>	- White blood cell	- Red blood cell	- Eosinophil
	- Neutrophil	- Platelet	- Monocyte
	- Lymphocyte	- Serum creatinine	- Alkaline Phosphatase
	- Alanine aminotransferase	- Aspartate aminotransferase	- Basophil
			- Albumin

**Πίνακας 37 :** Χαρακτηριστικά ασθενών που συγκεντρώθηκαν σε αυτή την μελέτη.

Μια επισκόπηση της διαδικασίας επιλογής χαρακτηριστικών του γενετικού αλγορίθμου παρουσιάζεται παρακάτω :



**Εικόνα 52 :** Διαδικασία επιλογής χαρακτηριστικών του γενετικού αλγορίθμου.

Στη συνέχεια, αναπτύχθηκε ένας αφελής ταξινομητής Bayes για να αναλύσει τις διαφορές στα προφίλ των ασθενών με θετική ή αρνητική ανταπόκριση στην επαναχορήγηση και για να προβλέψει την κατάσταση των ασθενών σε περίπτωση επαναχορήγησης. Για τη δημιουργία του μοντέλου χρησιμοποιήθηκαν 35 περιπτώσεις ασθενών που εμφανίστηκαν στο διάστημα 2004 – 2008, ενώ για την εξωτερική επικύρωσή του χρησιμοποιήθηκαν 11 ασθενείς. Επιλέχθηκε ο αφελής ταξινομητής Bayes διότι παράγει σχετικά καλή απόδοση ταξινόμησης και είναι απλός στην εφαρμογή.

Για την εκτίμηση της απόδοσης του μοντέλου αξιολογήθηκαν οι παράμετροι AUC (area under curve – περιοχή κάτω από την καμπύλη), η ευαισθησία και η ειδικότητα. Ως αποτέλεσμα, 8 χαρακτηριστικά των ασθενών (ηλικία, χημειοθεραπευτικό φάρμακο, επίπεδο αλβουμίνης, επίπεδο ερυθρών αιμοσφαιρίων, επίπεδο αιμοπεταλίων, μη φυσιολογικό επίπεδο λευκών αιμοσφαιρίων, μη φυσιολογικό επίπεδο αλκαλικής φωσφατάσης και μη φυσιολογικό επίπεδο αλανινικής αμινοτρασφεράσης) ταυτοποιήθηκαν ως κλινικοί προγνωστικοί δείκτες για την κατάσταση του ασθενή σε σχέση με την επαναχορήγηση του φαρμάκου. Αποδείχθηκε επίσης ότι το μοντέλο του αφελούς ταξινομητή Bayes είχε AUC ίση με 0,767, η οποία συγκρινόμενη με AUC από άλλες μελέτες που περιλαμβάνουν συστήματα CDSS κρίθηκε ικανοποιητική και συνεπώς το μοντέλο μπορεί να συνεισφέρει στη διαδικασία λήψης κλινικών αποφάσεων αφότου οι γιατροί εντοπίσουν μια ομάδα ασθενών που θα υποβληθούν ξανά σε θεραπεία. Επιπλέον, με τιμή κατωφλίου ίση με 0.8, η ευαισθησία και η ειδικότητα βρέθηκαν ίσες με 67% και 80 %, αντίστοιχα.

Αξίζει να σημειώσουμε ότι λόγω του περιορισμένου μεγέθους του συνόλου επικύρωσης (11 ασθενείς) απαιτείται μια πιο εκτεταμένη επικύρωση του μοντέλου προτού μπορέσει να εγκριθεί για κλινική χρήση σε συχνή βάση. Μια άλλη αδυναμία του μοντέλου αποτελεί επίσης η ασθενής συσχέτιση των επιλεγμένων χαρακτηριστικών με την κατάσταση επαναπρόκλησης των ασθενών, η οποία βασίζεται στο ότι το μοντέλο Bayes αποδίδει ικανοποιητικά για τα επιλεγμένα χαρακτηριστικά. Εν κατακλείδι όμως, μόλις επικυρωθεί, το μοντέλο μπορεί να χρησιμοποιηθεί για να βοηθήσει τους ιατρούς στο να αποφασίσουν για το αν θα χορηγήσουν εκ νέου το φάρμακο στους ασθενείς, προσδιορίζοντας αν η αρχική τους εκτίμηση για την κατάσταση του ασθενούς είναι ακριβής.

### 3.14 Αναφορές Κεφαλαίου

- [1] Joseph A. Cruz, David S. Wishart (2006), 'Application of Machine Learning in Cancer Prediction and Prognosis', University of Alberta, Canada
- [2] Elizabeth S. Burnside, Jesse Davis, Jagpreet Chhatwal, Oguzhan Alagoz, Mary J. Lindstrom, Berta M. Geller, Benjamin Littenberg, Katherine A. Shaffer, Charles E. Kahn, C. David Page (2009), 'Probabilistic Computer Model Developed from Clinical Data in National Mammography Database Format to Classify Mammographic Findings', Radiology Volume 251, Number 3
- [3] J.M. Jerez-Aragones et al. (2002), 'A combined neural network and decision trees model for prognosis of breast cancer relapse', Elsevier
- [4] J.A. Gomez-Ruiz et al. (2004), 'A Neural Network Based Model for Prognosis Of Early Breast Cancer', Kluwer Academic Publishers
- [5] Cakir et al. (2009), 'A software tool for determination of breast cancer treatment methods using data mining approach', Springer Science & Business Media
- [6] Laurentiis et al. (1999), 'A prognostic model that makes quantitative estimates of probability of relapse for breast cancer patients'
- [7] Theodore Anagnostou, Mesut Remzi, Michael Lykourinas, Bob Djavan (2003), 'Artificial Neural Networks for Decision-Making in Urologic Oncology', European Urology
- [8] Stephan et al. (2002), 'Multicenter evaluation of an artificial neural network to increase the prostate cancer detection rate and reduce unnecessary biopsies', Clinical Chemistry 48:8
- [9] Sang Youn Kim et al. (2011), 'Pre-Operative Prediction of Advanced Prostatic Cancer Using Clinical Decision Support Systems: Accuracy Comparison between Support Vector Machine and Artificial Neural Network', Korean Journal of Radiology
- [10] Hsueh-Chun Lin et al. (2009), 'A real time online assessment system with modeled architecture on clinical infometrics for patient reported outcomes of prostate cancer', Elsevier Ireland
- [11] Waghlikar et al. (2012), 'Clinical decision support with automated text processing for cervical cancer screening', Mayo Clinic
- [12] Austin et al. (2009), 'The Pittsburgh Cervical Cancer Screening Model – A Risk Assessment Tool'
- [13] Onisko Et al., 'Application of Dynamic Bayesian Networks to Cervical Cancer Screening', Intelligent Information Systems 9999, ISBN 666-666-666
- [14] Siebert et al. (2006), 'The German Cervical Cancer Screening Model: development and validation of a decision-analytic model for cervical cancer screening in Germany'
- [15] Karakitsos et al. (2012), 'Identification of Women for Referral to Colposcopy by Neural Networks: A Preliminary Study Based on LBC and Molecular Biomarkers', Hindawi Publishing Corporation, Journal of Biomedicine and Biotechnology
- [16] Hanai et al. (2003), 'Prognostic models in patients with non-small-cell lung cancer using artificial neural networks in comparison with logistic regression'

- [17]** Bottaci et al. (1997), 'Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions', *The Lancet*
- [18]** Grumett et al. (2003), 'Neural Networks in the Prediction of Survival in Patients with Colorectal Cancer', *Clinical Colorectal Cancer*, Vol. 2, No. 4, 239-244, 2003
- [19]** Pedreira et al. (2009), 'New Decision Support Tool for Treatment Intensity Choice in Childhood Acute Lymphoblastic Leukemia', *IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, VOL. 13, NO. 3, MAY 2009
- [20]** Huang et al. (2004), 'Application of Probabilistic Neural Networks to the Class Prediction of Leukemia and Embryonal Tumor of Central Nervous System', *Neural Processing Letters* 19: 211–226, 2004
- [21]** Exarchos et al. (2011), 'Multiparametric Decision Support System for the Prediction of Oral Cancer Reoccurrence'
- [22]** Tuan Zea Tan et al. (2008), 'Ovarian cancer diagnosis with complementary learning fuzzy neural network', *Elsevier Artificial Intelligence in Medicine* (2008) 43, 207—222
- [23]** Stojadinovic et al. (2009), 'Development of a clinical decision model for thyroid nodules', *BMC surgery*
- [24]** Young Sun Kim et al. (2003), 'Screening test data analysis for liver disease prediction model using growth curve', *Biomedicine & Pharmacotherapy* 57 (2003) 482–488
- [25]** Forsberg JA, Eberhardt J, Boland PJ, Wedin R, Healey JH (2011) 'Estimating Survival in Patients with Operable Skeletal Metastases: An Application of a Bayesian Belief Network'. *PLoS ONE* 6(5): e19956. doi:10.1371/journal.pone.0019956
- [26]** Pei Yi Loke et al. (2011), 'Pilot study on developing a decision support tool for guiding re-administration of chemotherapeutic agent after a serious adverse drug reaction', <http://www.biomedcentral.com/1471-2407/11/319>

# Κεφάλαιο 4 – Συμπεράσματα και επισκόπηση μελέτης

## **4.1 Συμπεράσματα, περιορισμοί και συστάσεις**

Οι μελέτες που περιγράφονται στις προηγούμενες σελίδες αποτελούν μερικά μόνο παραδείγματα για το πώς θα πρέπει να λειτουργούν τα σύγχρονα συστήματα υποστήριξης κλινικών αποφάσεων και το πώς θα πρέπει να καταγράφονται, επικυρώνονται και αξιολογούνται οι μέθοδοι και τα αποτελέσματα. Προφανώς, υπάρχουν και δημοσιεύονται συνεχώς πολλά παραδείγματα εξίσου καλών μελετών με εξίσου εντυπωσιακά αποτελέσματα [1]. Εν γένει, στην παρούσα εργασία έγινε μια προσπάθεια να συμπεριληφθούν μελέτες για όσο το δυνατόν περισσότερους τύπους καρκίνου, προκειμένου να αξιολογηθεί η δυναμική των συστημάτων υποστήριξης κλινικών αποφάσεων σε ετερόκλητες προκλήσεις. Ως επί το πλείστον, ως κλινικοί σκοποί των μελετών που αναλύσαμε καθορίζονται η διάγνωση της νόσου, η πρόβλεψη της επιβίωσης και η πρόβλεψη της υποτροπής των ασθενών. Μια γενική επισκόπηση των μελετών που περιγράφηκαν παρουσιάζεται στον παρακάτω συγκεντρωτικό πίνακα, ώστε να καταστεί πιο εύκολη η σύγκριση των τεχνικών και των αποτελεσμάτων τους:

Τύπος καρκίνου	Σκοπός	Δεδομένα εισόδου (Πλήθος ασθενών, Τύπος δεδομένων)	Τεχνικές	Γενικά αποτελέσματα	Βιβλιογραφία
καρκίνος του μαστού	διάγνωση	18629 ασθενείς, δημογραφικοί παράγοντες & εξετάσεις μαστογραφίας	Μπεϋζιανό δίκτυο	πιθανότητα κακοήθειας, AUC=0.960, ευαισθησία=90%, ειδικότητα=93%	Burnside et al., 2009
καρκίνος του μαστού	υποτροπή	1035 ασθενείς, κλινικο-παθολογικά δεδομένα	αλγόριθμος TDIT, σύστημα νευρωνικών δικτύων	πρόβλεψη κακοήθειας, ευαισθησία=0.787-0.887, ειδικότητα=0.945-0.972	Aragones et al., 2002
καρκίνος του μαστού	επιλογή θεραπείας	462 ασθενείς, κλινικο-παθολογικά δεδομένα	αλγόριθμος IB1, πολυεπίπεδος αισθητήρας, πίνακας αποφάσεων	ορμονοθεραπεία/ακρίβεια IB1=94.6237%, ακτινοθεραπεία/ακρίβεια πολυεπίπεδου αισθητήρα=95.2381%, ταμοξιφαίνη/ακρίβεια πολυεπίπεδου αισθητήρα=92%, χημειοθεραπεία/ακρίβεια πίνακα αποφάσεων=97.7778%	Cakir et al., 2009
καρκίνος του μαστού	υποτροπή	2441 ασθενείς, κλινικο-παθολογικά δεδομένα	πολυεπίπεδος αισθητήρας	5ετής πιθανότητα υποτροπής, AUC=0.726	Laurentiis et al., 1999
καρκίνος του προστάτη	διάγνωση	1188 άνδρες, κλινικο-παθολογικά δεδομένα & δακτυλική εξέταση ορθού	TND	ανίχνευση καρκίνου εντός διαφορετικών περιοχών tPSA, μέγιστη ακρίβεια 95% για ειδικότητα 41% και tPSA 4-10μg/L	Stephan et al., 2002
καρκίνος του προστάτη	διάγνωση	532 άνδρες, παράμετροι βιοψιών προστάτη & διορθικού υπερηχογραφήματος	μηχανή υποστήριξης διανυσμάτων (SVM), TND (πολυεπίπεδος αισθητήρας)	AUC του SVM=0.805, AUC του TND=0.719	Sang Youn Kim et al., 2011
καρκίνος του προστάτη (knowledge-based)	ποσοτικοποιημένες κλινικές πληροφορίες				Hsueh-Chun Lin et al., 2009
καρκίνος του τραχήλου της μήτρας (knowledge-based)	λήψη αποφάσεων	49293 αναφορές τεστ Παπ, εφαρμογή σε 74 ασθενείς		βέλτιστες συστάσεις και για τους 74 ασθενείς	Waghlikar et al., 2012
καρκίνος του τραχήλου της μήτρας	διάγνωση	375441 γυναίκες, 19 μεταβλητές (κλινικές, αποτελέσματα τεστ Παπ, χειρουργικές επεμβάσεις)	δυναμικό Μπεϋζιανό δίκτυο	πιθανότητα (%) μη ύπαρξης προκαρκινικών σταδίων/καρκίνου σε 1, 2, 3 έτη	Austin et al., 2009
καρκίνος του τραχήλου της μήτρας	επιβιωσιμότητα	τεστ Παπ, δεδομένα εξέλιξης ασθένειας	Μαρκοβιανό μοντέλο	κίνδυνος για τη διάρκεια ζωής 3%, θνησιμότητα 1%	Siebert et al., 2006
καρκίνος του τραχήλου της μήτρας	διάγνωση	1258 γυναίκες, κυτταρολογική εξέταση υγρών & μοριακοί βιοδείκτες	νευρωνικό δίκτυο με καταμεριστή διανυσμάτων (LVQ-NN)	ανίχνευση και ταξινόμηση προκαρκινικών αλλοιώσεων και καρκίνου, συνολική ακρίβεια 96.03%	Karakitsos et al., 2012
καρκίνος του πνεύμονα	επιβιωσιμότητα	125 ασθενείς, 17 κλινικοπαθολογικές & ανοσοϊστοχημικές μεταβλητές	TND	πρόβλεψη επιβίωσης για 5 έτη, ακρίβεια=87%	Hanai et al., 2003
καρκίνος του παχέος εντέρου	επιβιωσιμότητα	334 ασθενείς, 42 κλινικοπαθολογικές μεταβλητές	TND εμπροσθόδρομης τροφοδοσίας	πρόβλεψη επιβίωσης, ακρίβεια=80%, ευαισθησία=60%, ειδικότητα=88%	Bottaci et al., 1997
καρκίνος του παχέος εντέρου	επιβιωσιμότητα	403 ασθενείς, παθολογικές μεταβλητές	TND (πολυεπίπεδος αισθητήρας)	υποτροπή ή μη της νόσου, πρόβλεψη επιβίωσης με AUC=78%, ευαισθησία=95%	Grumett et al., 2003
λευχαιμία	επιλογή έντασης θεραπείας	158 ασθενείς, 17 μεταβλητές (κλινικές, εργαστηριακές, απόκριση στη θεραπεία)	TND	κίνδυνος υποτροπής σε σχέση με το πρωτόκολλο BFM95, χαμηλός 42% και υψηλός 58%	Pedreira et al., 2009
λευχαιμία	διάγνωση	72 ασθενείς, 50 χαρακτηριστικά γονιδιακής έκφρασης	πιθανοτικό νευρωνικό δίκτυο	ταξινόμηση τύπων λευχαιμίας σε σχέση με τη στατιστική σήματος προς θόρυβο, ακρίβεια 100%	Huang et al., 2004
καρκίνος του στόματος	υποτροπή	41 ασθενείς, δεδομένα κλινικά, απεικονιστικά & γονιδιώματος	Μπεϋζιανά δίκτυα, TND, SVM, δέντρα αποφάσεων, τυχαία δάση & πλειοψηφικός αλγόριθμος	διάκριση ασθενών με ή χωρίς υποτροπή, ακρίβεια πλειοψηφικού αλγορίθμου 100%	Exarchos et al., 2011
καρκίνος των ωοθηκών	διάγνωση	3 βάσεις δεδομένων (μικροσυστοιχίες γονιδιώματος, αιματολογικές αναλύσεις, πρωτεομικά φάσματα)	ασαφές νευρωνικό δίκτυο συμπληρωματικής εκμάθησης	διάγνωση καρκίνου, ακρίβεια=78.9%, ευαισθησία=95%, ειδικότητα=57%	Tuan Zea Tan et al., 2008
καρκίνος του θυρεοειδούς	διάγνωση	216 ασθενείς, κλινικο-παθολογικά, κυτταρολογικά & απεικονιστικά δεδομένα	δίκτυο Μπεϋζιανής πίστης	πρόβλεψη κακοήθειας, AUC=0.88, ευαισθησία=82%, ειδικότητα=57%	Stojadinovic et al., 2009
καρκίνος του ήπατος	πρόβλεψη εμφάνισης	1030 ασθενείς, κλινικο-παθολογικές μεταβλητές (παράγοντες κινδύνου για τον καρκίνο του ήπατος)	λογιστική παλινδρόμηση, δέντρα αποφάσεων, νευρωνικό δίκτυο (ND)	ακρίβεια πρόβλεψης καρκίνου, ακρίβεια=72.55%, ευαισθησία=78.62% για το ND με καμπύλη ανάπτυξης	Young Sun Kim et al., 2003
καρκίνος των οστών	επιβιωσιμότητα	189 ασθενείς, 15 κλινικο-παθολογικά χαρακτηριστικά	δίκτυο Μπεϋζιανής πίστης	επιβίωση για 3 & 12 μήνες, 3μηνο μοντέλο με AUC=0.85, 12μηνο μοντέλο με AUC=0.83	Forsberg et al., 2011
χημειοθεραπεία	απόφαση επαναχορήγησης φαρμάκου	46 ασθενείς, δεδομένα δημογραφικά, εργαστηριακά, ιατρικής κατάστασης, φαρμακευτικής αγωγής	επιλογή χαρακτηριστικών μέσω γενετικού αλγορίθμου, αφελής Μπεϋζιανός ταξινομητής	απόφαση επαναχορήγησης φαρμάκου μετά από αλλεργική αντίδραση, AUC=0.767, ευαισθησία=67%, ειδικότητα=80%	Pei Yi Loke et al., 2011

**Πίνακας 38 :** Συγκεντρωτικός πίνακας των μελετών που εξετάστηκαν όπου περιγράφεται ο τύπος του καρκίνου, ο σκοπός κάθε συστήματος, τα δεδομένα που χρησιμοποιήθηκαν, οι τεχνικές ταξινόμησης, τα γενικά αποτελέσματα και η βιβλιογραφία.

Σε αυτό το σημείο πρέπει να επισημανθεί ότι δεν πραγματοποιούνται όλες οι μελέτες με την ίδια αυστηρότητα ή προσοχή στη λεπτομέρεια. Το να είναι σε θέση κάποιος να αναγνωρίσει πιθανά προβλήματα είτε στον πειραματικό σχεδιασμό, στην επικύρωση ή στις μαθησιακές τεχνικές, είναι σημαντικό όχι μόνο για όσους επιδιώκουν να αναπτύξουν αντίστοιχα συστήματα, αλλά και για αυτούς που χρειάζεται να αναλύσουν διαφορετικές μελέτες ή να αξιολογήσουν διαφορετικές επιλογές μηχανικής μάθησης.

Ένα από τα πιο συνηθισμένα προβλήματα που παρατηρήθηκαν μεταξύ των μελετών ήταν και η έλλειψη προσοχής όσον αφορά το μέγεθος των συνόλων εκπαίδευσης, δοκιμών και επικύρωσης. Μια ελάχιστη απαίτηση για κάθε σύστημα ταξινόμησης είναι η ύπαρξη μιας αρκετά μεγάλης βάσης δεδομένων που θα μπορεί να χωριστεί σε ξεχωριστά σύνολα εκπαίδευσης και δοκιμών ή θα υποβάλλεται σε κάποια λογικής μορφής διασταυρούμενη επικύρωση *n*-φορών για μικρότερα σύνολα δεδομένων. Συνήθως η διασταυρούμενη επικύρωση 10 φορών, που είδαμε να χρησιμοποιείται πολύ συχνά, επαρκεί για την επικύρωση σχεδόν κάθε αλγορίθμου μάθησης. Αυτό το είδος της αυστηρής εσωτερικής επικύρωσης είναι κρίσιμο για τη δημιουργία ενός ισχυρού συστήματος που θα μπορεί να χειρίζεται με συνέπεια νέα δεδομένα. Πέρα όμως από τη συνήθη πρακτική της εσωτερικής επικύρωσης, είναι ιδιαίτερα χρήσιμη η διενέργεια μιας δοκιμής επικύρωσης χρησιμοποιώντας μια εξωτερική πηγή δεδομένων. Η εξωτερική επικύρωση αποτελεί ένα σημαντικό λογικό έλεγχο και επίσης βοηθάει στην ελαχιστοποίηση οποιασδήποτε προκατάληψης που μπορεί να επιβληθεί από ειδικές για τον ασθενή πρακτικές κλινικών μετρήσεων. Φυσικά, αυτό το σύνολο εξωτερικής επικύρωσης, αν υπάρχει, πρέπει επίσης να είναι επαρκώς μεγάλο μεγέθους ώστε να εξασφαλίζεται η αναπαραξιμότητα του συστήματος.

Όπως έχει επανειλημμένα επισημανθεί, το μέγεθος ενός δεδομένου συνόλου εκπαίδευσης έχει πολλές επιπτώσεις όσον αφορά την αξιοπιστία, την αναπαραξιμότητα και την ακρίβεια του συστήματος. Ένα πρώτο συμπέρασμα είναι ότι για ένα μικρότερο μέγεθος δειγμάτων, σχεδόν κάθε μοντέλο είναι επιρρεπές σε υπερβολική εκπαίδευση. Η υπερβολική εκπαίδευση μπορεί να οδηγήσει σε παραπλανητικές τιμές ακρίβειας. Για παράδειγμα, στη μελέτη του πολυπαραμετρικού συστήματος υποστήριξης αποφάσεων για την πρόβλεψη επανανεμφάνισης του καρκίνου του στόματος (Exarchos et al. 2011), οι συγγραφείς πέτυχαν ακρίβεια διάκρισης των ασθενών με και χωρίς υποτροπή της τάξης του 100%. Ωστόσο, η βάση δεδομένων (σύνολα εκπαίδευσης και επικύρωσης) αποτελείτο από 41 μόνο ασθενείς, χωρίς να υπάρχει σύνολο εξωτερικής επικύρωσης. Το γεγονός λοιπόν ότι επετεύχθη ακρίβεια 100% είναι εντυπωσιακό, αλλά δημιουργεί κάποιες αμφιβολίες για την ευρωστία του μοντέλου. Σίγουρα η ύπαρξη ενός συνόλου εξωτερικής επικύρωσης θα ήταν επιθυμητή ώστε να ενισχυθεί ο ισχυρισμός της απόλυτης ακρίβειας.

Το μέγεθος της βάσης δεδομένων δεν είναι ο μόνος περιορισμός για την αποτελεσματικότητα των μελετών. Η ποιότητα της βάσης δεδομένων και η προσεκτική επιλογή χαρακτηριστικών είναι εξίσου σημαντικά. Για μεγάλες βάσεις δεδομένων, η εισαγωγή και η επαλήθευση των δεδομένων είναι ύψιστης σημασίας. Συχνά, η απρόσεκτη εισαγωγή δεδομένων μπορεί να οδηγήσει σε λάθη αντιστοίχισης τιμών σε μια συγκεκριμένη μεταβλητή. Για το λόγο αυτό είναι απαραίτητη η περαιτέρω επικύρωση και από εξωτερικούς πεπειραμένους αναλυτές.

Αντίστοιχης σημασίας με την ποιότητα των δεδομένων είναι και η ποιότητα των χαρακτηριστικών. Σίγουρα το υποσύνολο των χαρακτηριστικών που επιλέγεται για την εκπαίδευση ενός μοντέλου κάνει τη διαφορά ανάμεσα σε ένα ισχυρό, ακριβές μοντέλο και σε ένα ελαττωματικό και ανακριβές. Ιδανικά, τα χαρακτηριστικά θα πρέπει να επιλέγονται ώστε να είναι αναπαραξίμα και επακριβώς μετρήσιμα από ένα εργαστήριο (ή κλινική) σε ένα άλλο. Κατά κανόνα, τα καλύτερα χαρακτηριστικά είναι αυτά που αναπαράγονται απόλυτα και παγκοσμίως (ηλικία, φύλο, βάρος, ορισμένες μετρήσεις βιοδεικτών και αποτελέσματα εξετάσεων κτλ). Ακόμη όμως και με αυτά τα φαινομενικά ισχυρά χαρακτηριστικά, πρέπει να ληφθεί υπόψη το γεγονός ότι τα κλινικά σύνολα δεδομένων δεν είναι στατικές οντότητες. Επιπλέον, με την πάροδο του χρόνου, η σημασία ή η σχετικότητα αυτών των κλινικών μέτρων μπορεί να εξελιχθεί με αποτέλεσμα κάποια χαρακτηριστικά να προστεθούν, τροποποιηθούν ή διαγραφούν, κάτι που τόνισαν οι συγγραφείς σε διάφορες μελέτες (Aragones et al. 2002 , πρόβλεψη υποτροπής ασθενών με καρκίνο του μαστού). Ως εκ τούτου, ένα ταξινομητής πρέπει να είναι επίσης σε θέση να προσαρμόζεται σε διαφορετικά σύνολα χαρακτηριστικών με την πάροδο του χρόνου.

Ένα άλλο σημαντικό σημείο που πρέπει να τονίσουμε είναι η αξία της χρήσης πολλαπλών προγνωστικών μοντέλων που βασίζονται σε διαφορετικούς αλγόριθμους ταξινόμησης. Αν και τα τεχνητά νευρωνικά δίκτυα θεωρούνται προηγμένες μέθοδοι μηχανικής μάθησης, δεν αποτελούν πάντα το καλύτερο εργαλείο. Μερικές φορές, απλούστερες τεχνικές όπως ο αφελής ταξινομητής Bayes και τα δέντρα αποφάσεων μπορούν να επιδείξουν εξαιρετικά αποτελέσματα. Η αξιολόγηση της απόδοσης ενός ταξινομητή σε σχέση με άλλους είναι ζωτικής σημασίας για την επιλογή του βέλτιστου εργαλείου (για παράδειγμα, στη μελέτη των Sang Youn Kim et al. 2011 που αφορά τη διάγνωση του καρκίνου του προστάτη, έγινε σύγκριση των επιδόσεων ενός τεχνητού νευρωνικού δικτύου και μιας μηχανής υποστήριξης διανυσμάτων SVM, με τη μηχανή SVM να υπερσχύει ως διαγνωστικό εργαλείο). Είναι επίσης κρίσιμο να αποφασιστεί αν η προτεινόμενη μέθοδος είναι καλύτερη από ήδη υπάρχοντα συστήματα. Ιδανικά, κάθε νέο μοντέλο θα πρέπει να συγκρίνεται είτε με άλλα αντίστοιχα μοντέλα, ένα παραδοσιακό στατιστικό μοντέλο ή ένα καθιερωμένο προγνωστικό σχήμα όπως το σύστημα σταδιοποίησης TNM (όπως στην περίπτωση του προγνωστικού μοντέλου που παράγει ποσοτικές εκτιμήσεις της πιθανότητας υποτροπής ασθενών με καρκίνο του μαστού, όπου τα αποτελέσματα του συστήματος συγκρίνονται με αυτά που παράγει το σύστημα σταδιοποίησης TNM, Laurentiis et al. 1999).

Κλείνοντας, έχουμε την πεποίθηση ότι αν η ποιότητα των μελετών που δημοσιεύονται συνεχίζει να βελτιώνεται, αυξάνεται η πιθανότητα ενσωμάτωσης των υπολογιστικά έξυπνων συστημάτων σε κλινικές και νοσοκομειακές εφαρμογές. Προς αυτή την κατεύθυνση λοιπόν, οι λεπτομέρειες που αφορούν τη δομή των συστημάτων υποστήριξης κλινικών αποφάσεων, τα δεδομένα που χρησιμοποιήθηκαν για τα σύνολα εκπαίδευσης και επικύρωσης, ο διαχωρισμός των συνόλων, οι αλγόριθμοι και οι εφαρμογές τους πρέπει να περιγραφούν λεπτομερώς και να τεθούν στη διάθεση της επιστημονικής κοινότητας ώστε να επιτραπεί και σε άλλους ερευνητές η επαλήθευση και αναπαραγωγή των αποτελεσμάτων, αλλά και η χρήση τους στην κλινική πράξη.



## **4.2 Αναφορές Κεφαλαίου**

- [1] Joseph A. Cruz, David S. Wishart (2006), 'Application of Machine Learning in Cancer Prediction and Prognosis', University of Alberta, Canada