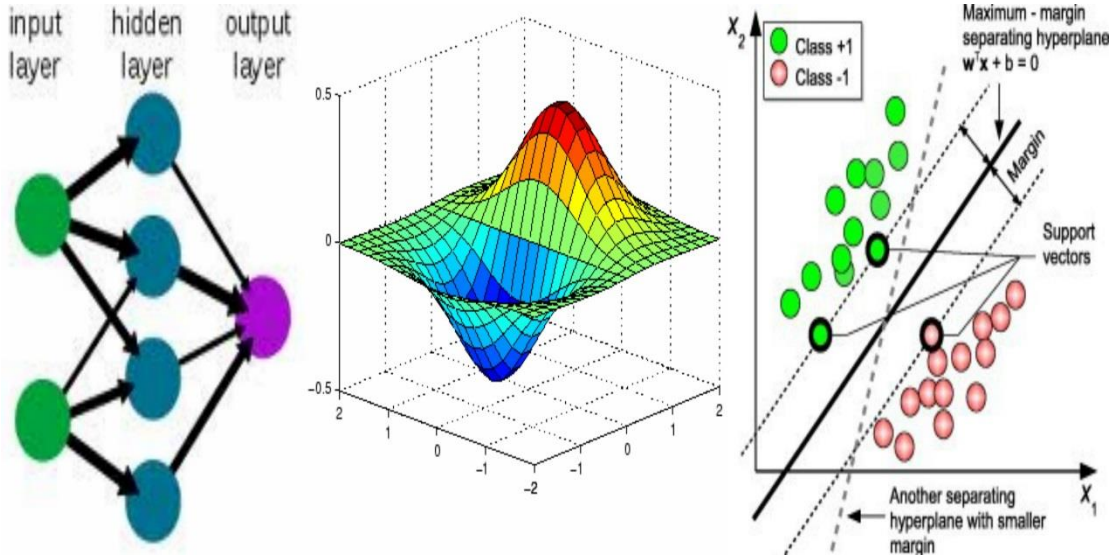
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΔΟΜΟΣΤΑΤΙΚΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:

ΥΠΟΛΟΓΙΣΤΙΚΑ ΜΟΝΤΕΛΑ ΜΕΙΩΜΕΝΗΣ ΚΛΙΜΑΚΑΣ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΠΟΛΙΤΙΚΟΥ ΜΗΧΑΝΙΚΟΥ

ΠΟΛΥΔΩΡΟΥ ΔΕΣΠΟΙΝΑ

Επιβλέπων: Επικ. Καθ. Λαγαρός Νίκος



Αθήνα, Ιούλιος 2013

ΕΥΧΑΡΙΣΤΙΕΣ

Ολοκληρώνοντας τη διπλωματική μου εργασία και ταυτόχρονα τις σπουδές μου στο τμήμα Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου νιώθω την ανάγκη να ευχαριστήσω κάποια άτομα που με βοήθησαν τόσο στην εκπόνηση της παρούσας εργασίας όσο και γενικότερα κατά τη διάρκεια του κύκλου σπουδών μου.

Κατ' αρχήν θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή της διπλωματικής μου εργασίας κ. Νίκο Λαγαρό που μου έδωσε τη δυνατότητα να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα και συγχρόνως μου προσέφερε τις επιστημονικές του συμβουλές και τη βοήθειά του. Ακόμη, ένα ευχαριστώ στον κ. Αριστείδιο Καρλαύτη για την παραχώρηση των δεδομένων σχετικά με το παράδειγμα του οδοστρώματος.

Επίσης, θέλω να ευχαριστήσω όλους τους καθηγητές στο τμήμα Πολιτικών Μηχανικών του ΕΜΠ για την ανεκτίμητη επιστημονική γνώση που μου μετέδωσαν κατά τη διάρκεια των σπουδών μου αλλά και για τη βοήθειά τους σε οποιοδήποτε θέμα χρειάστηκε.

Κυρίως, όμως, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στους γονείς μου για τη συνεχή στήριξή τους όλα αυτά τα χρόνια αλλά και γιατί χάρη σε αυτούς έχω καταφέρει να φτάσω ως εδώ. Τέλος, ένα μεγάλο ευχαριστώ στους φίλους μου για όλες τις εμπειρίες και τις όμορφες στιγμές που περάσαμε μαζί και οι οποίες θα με συνοδεύουν τα επόμενα χρόνια.

ABSTRACT

In modern times, civil engineers have to solve problems that are very demanding in computational power and runtime thus conventional methods fail to give a good approach. Today, due to the evolution of science, these conventional methods have been replaced by new computational tools which are cheaper, faster and their aim is to find the 'best' through a range of safe options. This is a change of direction made in recent decades not only in the field of civil engineering but also in many other scientific domains. Through this new direction, old methods which are time consuming, inefficient and difficult to implement will be abolished or used only in the appropriate problems.

So, a part of this new 'trend' is this diploma thesis, which focuses on the use of modern reduced scale computational models in the implement of civil engineers' problems. The assessment of these models is based on their performance in predicting and optimizing computational experiments. More specifically, it considers three tools for solving problems in different areas of civil engineering. Firstly, a theoretical analysis of each method is provided and then the results, obtained by solving numerical applications through these methods, are represented. The computational tools that have been selected are Neural Networks, Kriging and Support Vector Machines and the problems they have to deal with concern crane, liquefaction and pavement construction.

After the presentation of the results, the performance of each model is evaluated and final conclusions about their fitness and their general behavior to every problem are exported. The parameters that help us make the conclusions are firstly the accuracy of the prediction and secondly the configuration and execution time of the algorithms that were constructed for each method.

Περιεχόμενα

1. Εισαγωγή 5
2. Τεχνητά Νευρωνικά Δίκτυα 7
2.1 Εισαγωγή 7
2.2 Βασικές αρχές των Τεχνητών Νευρωνικών Δικτύων 9
2.2.1 Λειτουργία βιολογικού και τεχνητού νευρώνα 9
2.2.2 Συναρτήσεις μεταφοράς 11
2.2.3 Δίκτυα πρόσθιας τροφοδότησης 13
2.2.4 Εκπαίδευση των τεχνητών νευρωνικών δικτύων 16
2.3 Ο αλγόριθμος της πίσω διάδοσης 17
3. Μέθοδος προσομοίωσης Kriging 22
3.1 Εισαγωγή 22
3.2 Βασικές αρχές του Kriging 23
3.3 Λειτουργία του Kriging – Βασικές παράμετροι 27
3.3.1 Μοντέλα παλινδρόμησης 27
3.3.2 Μοντέλα συσχέτισης 28
3.3.3 Πειραματικός σχεδιασμός 31
4. Support Vector Machines 34
4.1 Εισαγωγή 34
4.2 Λειτουργία των SVM 35
4.3 Βασικές έννοιες γύρω από τα SVM 37
4.3.1 Διάσταση VC και ικανότητα των μηχανών εκμάθησης 37
4.3.2 Δυναμικότητα των υπερεπιπέδων 38
4.3.3 ΣΥΝΑΡΤΗΣΗ KERNEL 44
4.4 SVM ως μέθοδος παλινδρόμησης 45
4.4.1 Περιγραφή του προβλήματος 45
4.4.2 Κλασσική μορφή του προβλήματος παλινδρόμησης 46
5. Αριθμητικές εφαρμογές 48
5.1 Εισαγωγή 48
5.2 Μελέτη προβλημάτων 49
5.2.1 Γερανογέφυρα 49
5.2.1.1 Παρουσίαση του προβλήματος 49
5.2.1.2 Επίλυση με τη μέθοδο των Νευρωνικών Δικτύων 56
5.2.1.3 Επίλυση με τη μέθοδο Kriging 60

5.2.1.4 Επίλυση με τη μέθοδο SVM 62
5.2.2 Αξιολόγηση της πιθανότητας ρευστοποίησης αμμώδων εδαφών 64
5.2.2.1 Εισαγωγή στην έννοια της ρευστοποίησης 64
5.2.2.2 Επίλυση με τη μέθοδο ΤΩΝ Νευρωνικών Δικτύων 70
5.2.2.3 Επίλυση με τη μέθοδο Kriging 73
5.2.2.4 Επίλυση με τη μέθοδο SVM 74
5.2.3 Προσδιορισμός στοιχείων οδοστρώματος 75
5.2.3.1 Θεωρητικές έννοιες- Ανάλυση πειράματος 75
5.2.3.2 Επίλυση με τη μέθοδο των Νευρωνικών Δικτύων 81
5.2.3.3 Επίλυση με τη μέθοδο Kriging 91
5.2.3.4 Επίλυση με τη μέθοδο SVM 95
6. Συμπεράσματα 100
7. Βιβλιογραφία 102

1

ΕΙΣΑΓΩΓΗ

Στη σύγχρονη εποχή, πολλές φορές τα προβλήματα στα οποία καλείται να βρει λύση ο πολιτικός μηχανικός είναι από τη φύση τους πολύ απαιτητικά σε υπολογιστική ισχύ και χρόνο εκτέλεσης με αποτέλεσμα οι συμβατικές μέθοδοι να αδυνατούν να δώσουν λύση.

Σήμερα, ο σχεδιασμός αποτελεί το κύριο αντικείμενο που απασχολεί κάθε μηχανικό ο οποίος στοχεύει στη σχεδίαση ενός μηχανήματος ή μιας κατασκευής η οποία καλείται να αντιμετωπίσει κάποιες λειτουργικές ανάγκες. Μέχρι πριν από τρεις δεκαετίες η εκπλήρωση αυτού του στόχου ήταν δύσκολη ως και αδύνατη λόγω ανυπαρξίας μεθόδων που θα βοηθούσαν προς την κατεύθυνση αυτή. Οι μέθοδοι που χρησιμοποιούνταν ως τότε είχαν ως κύριο στόχο τη μονοδιάστατη έννοια του ‘υπέρ της ασφαλείας’ χωρίς να επιδιώκουν το ‘βέλτιστο’. Όμως, από τότε μέχρι σήμερα, μέσω της εξέλιξης της επιστήμης, οι συμβατικές μέθοδοι έχουν αντικατασταθεί από νέα υπολογιστικά εργαλεία τα οποία είναι φθηνότερα, ταχύτερα και στοχεύουν στην εύρεση του ‘βέλτιστου’ μέσω ενός συνόλου ασφαλών επιλογών. Πρόκειται για μια αλλαγή κατεύθυνσης που πραγματοποιείται τις τελευταίες δεκαετίες τόσο στον χώρο των πολιτικών μηχανικών όσο και σε άλλους επιστημονικούς κλάδους και στοχεύει στην κατάργηση των μεθόδων εκείνων, οι οποίες είναι πολλές φορές χρονοβόρες, μη αποδοτικές και δύσκολες στην εφαρμογή.

Μέρος, λοιπόν, αυτής της αλλαγής κατεύθυνσης που παρατηρείται τα τελευταία χρόνια αποτελεί και η παρούσα διπλωματική εργασία, η οποία μελετά τη χρήση σύγχρονων υπολογιστικών μοντέλων μειωμένης κλίμακας σε προβλήματα πολιτικού μηχανικού. Πρόκειται για μοντέλα τα οποία αξιολογούνται με βάση την απόδοσή τους στην πρόβλεψη αλλά και την αξιοποίηση υπολογιστικών πειραμάτων.

Πιο συγκεκριμένα, εξετάζονται τρία εργαλεία για την επίλυση προβλημάτων που αφορούν σε διαφορετικούς τομείς ενός πολιτικού μηχανικού. Αρχικά, γίνεται μια θεωρητική ανάλυση της κάθε μεθόδου και του τρόπου λειτουργίας της και στη συνέχεια, παρατίθενται τα αποτελέσματα που προέκυψαν από την επίλυση αριθμητικών εφαρμογών μέσω των συγκεκριμένων μεθόδων. Τα υπολογιστικά εργαλεία τα οποία έχουν επιλεχθεί είναι τα Νευρωνικά Δίκτυα, το Kriging και τα Support Vector Machines και τα προβλήματα στα οποία εξετάζονται αφορούν σε γερανογέφυρα, σε ρευστοποίηση εδάφους και σε κατασκευή οδοστρώματος.

Έπειτα από την παράθεση των αποτελεσμάτων, αξιολογείται η απόδοση κάθε μοντέλου και εξάγονται τα τελικά συμπεράσματα για την καταλληλόλητά τους αλλά και για τη γενικότερη συμπεριφορά τους σε κάθε πρόβλημα. Οι παράμετροι που βοηθούν στην εξαγωγή των συμπερασμάτων είναι τόσο η ακρίβεια με την οποία προσεγγίζουν τα υπολογιστικά προβλήματα όσο και ο χρόνος διαμόρφωσης και εκτέλεσης των αλγορίθμων που κατασκευάστηκαν για την κάθε μέθοδο.

2

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1 ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό γίνεται αναφορά στις βασικές αρχές λειτουργίας και στα βασικά χαρακτηριστικά των τεχνητών νευρωνικών δικτύων (ANN). Τα τεχνητά νευρωνικά δίκτυα αποτελούν ένα σύγχρονο υπολογιστικό εργαλείο το οποίο, μέσω κατάλληλης εκπαίδευσης, μπορεί να παράγει πολύ καλές προσεγγίσεις σε διαφόρων ειδών προβλήματα με αρκετά καλή ακρίβεια. Το μεγαλύτερο πλεονέκτημα τους είναι ότι απαιτούν πολύ μικρή υπολογιστική προσπάθεια, σε αντίθεση με τις συμβατικές μεθόδους οι οποίες κατά κύριο λόγο εκτελούν μεγάλο αριθμό υπολογιστικών πράξεων. Στόχος τους είναι η ταχεία επίλυση προβλημάτων μέσω μηχανισμών των οποίων η σύλληψη βασίζεται σε βιολογικά πρότυπα.

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από τη βιολογία και τα στοιχεία τα οποία τα αποτελούν συμπεριφέρονται κατά τρόπο ανάλογο των στοιχειωδών λειτουργιών των βιλογικών κυττάρων. Τα στοιχεία αυτά είναι οργανωμένα κατά τέτοιο τρόπο ώστε να προσομοιάζουν την ανατομία του ανθρώπινου εγκεφάλου. Μπορεί η ομοιότητά τους με τον ανθρώπινο εγκέφαλο να μην είναι και τόσο μεγάλη, ωστόσο καταφέρνουν να προσεγγίσουν ένα μεγάλο αριθμό χαρακτηριστικών της δομής αλλά και λειτουργίας του εγκεφάλου. Για παράδειγμα, μαθαίνουν χρησιμοποιώντας εμπειρία την οποία έχουν συσσωρεύσει, έχουν την ικανότητα γενίκευσης από προηγούμενα παραδείγματα σε νέα, μπορούν να επεξεργαστούν μια ομάδα δεδομένων και να ξεχωρίσουν από αυτή τα σημαντικότερα χαρακτηριστικά.

Η διαδικασία με την οποία λειτουργούν τα νευρωνικά δίκτυα (NN) - η οποία ονομάζεται εκπαίδευση (training)- είναι η εξής: ένα σύνολο τιμών αποτελούμενο από εισόδους ή από ζεύγη εισόδων-εξόδων (I/O pairs) επιδρά στο δίκτυο και βάσει

του συνόλου αυτού οι παράμετροι του δικτύου διορθώνονται έτσι ώστε να παραχθεί η επιθυμητή για το δίκτυο απόκριση. Στόχος των αλγορίθμων εκπαίδευσης είναι το δίκτυο να μην παρουσιάζει ευαισθησία στην απόκρισή του για μικρές αλλαγές στις εισόδους του. Είναι σημαντικό να επισημανθεί ότι από τη φύση τους τα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα της γενίκευσης, χωρίς όμως να χρησιμοποιούν κάτι ανάλογο της ανθρώπινης νοημοσύνης. Ορισμένα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα να αποσπούν τα ουσιώδη στοιχεία από ένα σύνολο δεδομένων εισαγωγής και να τα χρησιμοποιούν για να παράγουν κάτι που ήταν κρυμμένο μέσα σε αυτό το σύνολο. Για παράδειγμα, ένα δίκτυο μπορεί να εκπαιδευτεί σε μία ακολουθία από διάφορες μορφές του γράμματος *A*. Μετά την εκμάθηση όταν θα του παρουσιαστεί το γράμμα *A* σε μια αλλοιωμένη μορφή, το δίκτυο είναι ικανό να παράγει το γράμμα *A* στην σωστή του μορφή.

Τα βασικότερα πλεονεκτήματα των τεχνητών νευρωνικών δικτύων είναι:

(i) εκπαίδευση σε παράλληλο υπολογιστικό περιβάλλον
(ii) ικανότητα εκμάθησης
(iii) ικανότητα αναδημιουργίας αντικειμένων όταν είναι γνωστό ένα μόνο μέρος της περιγραφής τους
(iv) ικανότητα αποθήκευσης αντικειμένων και ανάκληση τους με βάση το περιεχόμενο και όχι τη διεύθυνση, σχετικιστική μνήμη (associative memory)
(v) ικανότητα προσομοίωσης συναρτήσεων

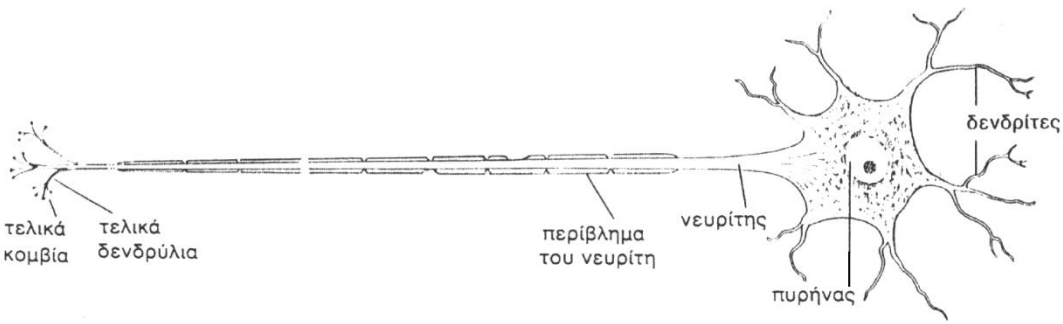
Η χρήση των νευρωνικών δικτύων φαίνεται ότι ξεκίνησε τη δεκαετία του ’40 και πιο συγκεκριμένα το 1943, όταν οι McCulloch – Pitts παρουσιάζουν το πρώτο μαθηματικό μοντέλο νευρωνικού δικτύου. Από τότε και ως τη δεκαετία του ’80 δημιουργούνται τα μοντέλα *perceptron* και *adaline* τα οποία αποτελούν επέκταση του αρχικού μοντελού και γίνονται κάποιες πρώτες εφαρμογές πάνω στην ευστάθεια των μοντέλων αυτών. Η χρησιμοποίηση τεχνικών τεχνητής νοημοσύνης όπως είναι τα νευρωνικά δίκτυα έχει αρχίσει να βρίσκει εφαρμογή σε προβλήματα υπολογιστικής μηχανικής την τελευταίες δεκαετίες. Τέτοιου είδους εφαρμογές είναι η ανάλυση αξιοπιστίας κατασκευών, όπου στόχος είναι η πρόβλεψη των αποτελεσμάτων της ανάλυσης καθώς και προβλήματα βέλτιστου σχεδιασμού, όπου οι απαιτούμενες τιμές των περιορισμών για κάθε νέο σχεδιασμό προέρχονται από την πρόβλεψη ενός κατάλληλα εκπαιδευμένου νευρωνικού δικτύου. Ένα σωστά εκπαιδευμένο νευρωνικό δίκτυο μπορεί να παράγει αποδεκτά, από πλευράς ακρίβειας, αποτελέσματα ενώ απαιτεί πάρα πολύ μικρή υπολογιστική προσπάθεια. Η ιδιότητα αυτή των NN αποτελεί και το βασικό τους πλεονέκτημα.

2.2 ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΩΝ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

2.2.1 Λειτουργία βιολογικού και τεχνητού νευρώνα

Τα τεχνητά νευρωνικά δίκτυα είναι εμπνευσμένα από τη βιολογία και έχουν σχεδιαστεί ώστε να μιμηθούν τα βασικά χαρακτηριστικά του βιολογικού νευρώνα. Το ανθρώπινο νευρικό σύστημα είναι χτισμένο από πολύπλοκα κύτταρα τα οποία ονομάζονται ν*ευρώνες*. Κατά μία εκτίμηση ο ανθρώπινος εγκέφαλος αποτελείται από 10^{11} νευρώνες που σχηματίζουν έναν τεράστιο αριθμό συνδέσεων (10^{15}) μεταξύ τους μέσω διαδρόμων επικοινωνίας που είναι πιθανό να έχουν μήκος μεγαλύτερο του ενός μέτρου. Κάθε νευρώνας μοιράζεται πολλά κοινά χαρακτηριστικά με άλλα κύτταρα στο σώμα αλλά εμφανίζει ορισμένες μοναδικές ικανότητες, όπως το να δέχεται, να επεξεργάζεται και να μεταφέρει ηλεκτροχημικά σήματα μέσω των νευρωνικών μονοπατιών τα οποία αποτελούν το σύστημα επικοινωνίας του εγκεφάλου. Κάθε τέτοιο ηλεκτροχημικό σήμα ονομάζεται νευρική ώση. Ένας τυπικός νευρώνας, στην πιο συνηθισμένη του μορφή, αποτελείται από το *σώμα* και τις α*ποφυάδες* του. Οι αποφυάδες είναι δύο ειδών: ο *νευρίτης* που είναι μια μακριά αποφυάδα, που το μήκος της μπορεί να είναι από λίγα δέκατα του χιλιοστομέτρου μέχρι πάνω από ένα μέτρο σε μερικές περιπτώσεις, και οι *δενδρίτες* που είναι πολλοί, κοντοί και έχουν διακλαδώσεις σαν δέντρο. Οι τελικές απολήξεις του νευρίτη λέγονται τελικά *δενδρύλλια*, στα άκρα των οποίων υπάρχουν τα τελικά κομβία (Σχήμα 2.1). Κάθε νευρώνας με τα τελικά κομβία μπορεί να έρχεται σε λειτουργική σύνδεση με τους δενδρίτες των γειτονικών νευρώνων. Η θέση της λειτουργικής αυτής σύνδεσης, μέσω της οποίας μεταβιβάζεται η νευρική ώση, λέγεται *σύναψη*. Από τη μεριά της σύναψης όπου έγινε η υποδοχή, οι είσοδοι κατευθύνονται στο σώμα του κυττάρου, όπου και αθροίζονται. Ορισμένες από τις εισόδους τείνουν να διεγείρουν το κύτταρο ενώ άλλες τείνουν να αναχαιτίσουν τη διέγερση του.

Για να προκληθεί νευρική ώση πρέπει το ερέθισμα να έχει τέτοια ένταση, ώστε να γίνει αναστροφή του δυναμικού ηρεμίας στον νευρώνα. Ένας νευρώνας βρίσκεται σε ηρεμία όταν δεν επιδρά πάνω του κάποιο ερέθισμα. Ερεθίσματα μικρότερης έντασης δεν προκαλούν νευρική ώση, ενώ εντονότερα ερεθίσματα δίνουν απλώς μεγαλύτερη συχνότητα νευρικών ώσεων.



Σχήμα 2.1 Βιολογικός νευρώνας

Ο τεχνητός νευρώνας, λοιπόν, βασίζεται στη μορφή του βιολογικού νευρώνα που μόλις αναπτύχθηκε και ακολουθεί τον παρακάτω τρόπο λειτουργίας: ένα σύνολο εισόδων, οι οποίες αντιστοιχούν στα ηλεκτροχημικά σήματα που αναφέρθηκαν στην περιγραφή του βιολογικού νευρώνα και προέρχονται από γειτονικούς νευρώνες, επιδρά επί του νευρώνα. Η επιρροή που έχει κάποιο σήμα σε κάποιο νευρώνα προσεγγίζεται με τον πολλαπλασιασμό της εισόδου (σήμα) με κάποια παράμετρο η οποία ονομάζεται συναπτικό βάρος (synaptic weight), το οποίο είναι ανάλογο της αντοχής της σύναψης του βιολογικού νευρώνα. Τα γινόμενα [βάρος]×[είσοδος] αθροίζονται ώστε να οριστεί το επίπεδο διέγερσης του νευρώνα. Οι είσοδοι που επιδρούν στον τεχνητό νευρώνα μπορούν να ενσωματωθούν σε ένα διάνυσμα **X**. Κάθε σήμα πολλαπλασιάζεται με ένα βάρος w_1, w_2,..., w_n πριν γίνει η άθροιση. Κάθε βάρος αντιστοιχεί στην "αντοχή" της κάθε βιολογικής σύνδεσης. Το σύνολο των βαρών σχηματίζει το διάνυσμα **W**. Η μορφή της εξίσωσης που περιγράφει την παραπάνω διαδικασία είναι η εξής:

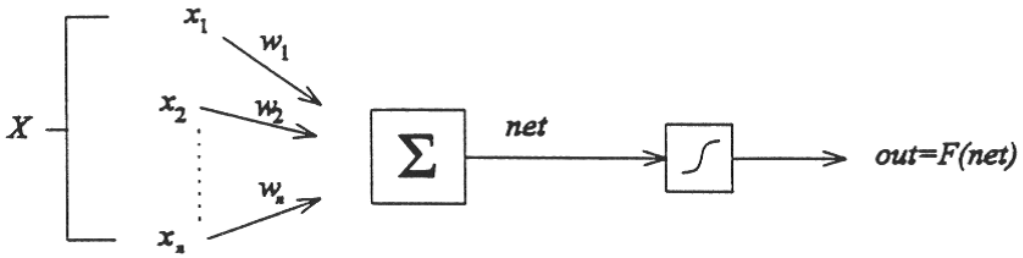
$$\mathrm{NET} = \mathrm{x}_1\mathrm{w}_1 + \mathrm{x}_2\mathrm{w}_2 + \ldots + \mathrm{x}_\mathrm{n}\mathrm{w}_\mathrm{n} = [x_1 \quad x_2 \quad . \quad . \quad x_n] \cdot \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ w_n \end{bmatrix} = \mathbf{x}^T\mathbf{w} \qquad (2.1)$$

Τα τεχνητά νευρωνικά δίκτυα έχουν αναπτυχθεί σε μεγάλη ποικιλία σχηματισμών-αρχιτεκτονικών. Οι βασικές αρχιτεκτονικές στις οποίες μπορούμε να κατατάξουμε τα τεχνητά νευρωνικά δίκτυα είναι δύο: *δίκτυα πρόσθιας τροφοδότησης* (feed-forward neural networks), στα οποία η ροή των σημάτων γίνεται προς μία κατεύθυνση και *δίκτυα πίσω τροφοδότησης ή δίκτυα με ανατροφοδότηση* (feed-

back or recurrent neural networks), στα οποία πραγματοποιείται επαναφορά της ροής σε μονάδες του δικτύου. Αν και οι δύο αυτοί τύποι αρχιτεκτονικών βασίζονται σε διαφορετική φιλοσοφία και υπακούουν σε διαφορετικές αρχές, χαρακτηρίζονται από μία κοινή ιδιότητα που έγκειται στην ικανότητα να μαθαίνουν μέσω της εκπαίδευσης. Η διαδικασία εκπαίδευσης μπορεί να γίνει με *επίβλεψη* (supervised) ή *χωρίς επίβλεψη* (unsupervised). Τόσο η αρχιτεκτονική όσο και η διαδικασία εκπαίσευσης είναι κάποια από τα θεματα που αναπτύσσονται παρακάτω.

2.2.2 Συναρτήσεις μεταφοράς

Το σήμα NET που παράγεται από την εξίσωση (2.1) περνά και από ένα επιπλέον στάδιο επεξεργασίας που ορίζεται από μία συνάρτηση, τη **συνάρτηση μεταφοράς** (transfer function), ώστε να παραχθεί τελικά η έξοδος του νευρώνα η οποία συμβολίζεται με OUT (Σχήμα 2.2). Η συνάρτηση μεταφοράς μπορεί να πάρει διάφορες μορφές.



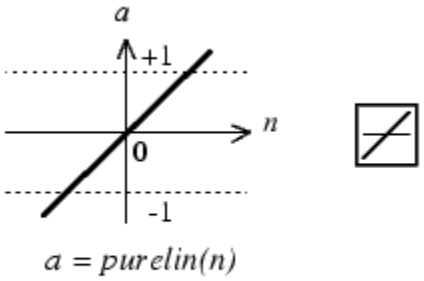
Σχήμα 2.2 Τεχνητός νευρώνας με τη συνάρτηση μεταφοράς

Όπως φαίνεται και στο σχήμα, το σήμα NET συμπιέζεται από την συνάρτηση F ώστε το τελικό σήμα OUT που παράγεται να μην υπερβαίνει κάποια χαμηλά όρια ανεξάρτητα της τιμής του σήματος NET. Σκοπός της συνάρτησης μεταφοράς είναι η προσομοίωση των καταστάσεων ηρεμίας (0 ή -1) και διέγερσης (1 ή +1) του βιολογικού νευρώνα. Κάποιες συναρτήσεις μεταφοράς που έχουν χρησιμοποιηθεί είναι:

(α) απλή γραμμική συνάρτηση (linear transfer function)

$$OUT = F(NET) = k \cdot NET \tag{2.2}$$

όπου k είναι ένας σταθερός αριθμός

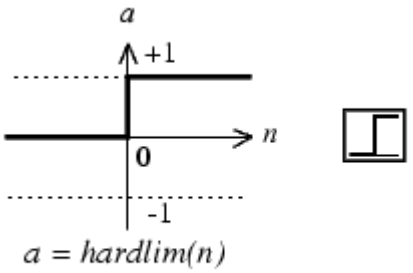


Linear Transfer Function

(β) Συνάρτηση κατωφλίου (hard-limit transfer function)

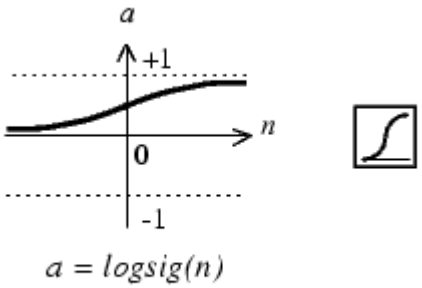
$OUT = 0$, αν η τιμή της εισόδου n είναι μικρότερη από 0

$OUT = 1$, αν η τιμή της εισόδου n είναι μεγαλύτερη ή ίση με το 0



Hard-Limit Transfer Function

(γ) συνάρτηση που προσομοιώνει με μεγαλύτερη ακρίβεια τα μη γραμμικά χαρακτηριστικά μεταφοράς των βιολογικών νευρώνων. Μια τέτοια συνάρτηση μεταφοράς που συχνά επιλέγεται είναι η λογιστική ή σιγμοειδής συνάρτηση (log-sigmoid function), όπως συχνά αποκαλείται, λόγω της μορφής του γραφήματός της που είναι σχήματος S.



Log-Sigmoid Transfer Function

Το σήμα εξόδου που θα παραχθεί από την σιγμοειδή συνάρτηση μεταφοράς έχει τη μορφή:

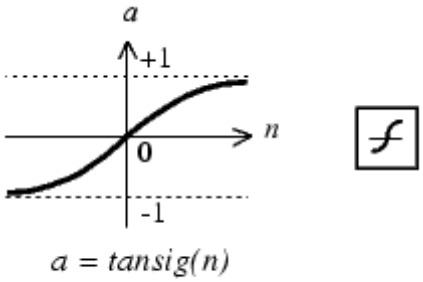
$$OUT = \frac{1}{1+e^{-NET}} = F(NET) \tag{2.3}$$

η οποία είναι αντισυμμετρική.

Η σιγμοειδής συνάρτηση συμπιέζει την ακτίνα του σήματος NET έτσι ώστε το αποτέλεσμα OUT να βρίσκεται μεταξύ του μηδέν και του ένα. Η συνάρτηση συμπίεσης παράγει την απαιτούμενη μη γραμμικότητα στον αλγόριθμο. Υπάρχουν πολλές συναρτήσεις που θα μπορούσαν να χρησιμοποιηθούν ως συναρτήσεις μεταφοράς για τον αλγόριθμο πίσω διάδοσης. Ο αλγόριθμος απαιτεί η συνάρτηση αυτή να είναι παντού παραγωγίσιμη γιατί η παράγωγος της συναρτήσεως μεταφοράς χρησιμοποιείται στο βήμα διόρθωσης των βαρών. Η σιγμοειδής συνάρτηση ικανοποιεί αυτήν την απαίτηση, και έχει και ένα πρόσθετο πλεονέκτημα παρέχοντας μια μορφή αυτόματου ελέγχου της κλίσης της καμπύλης. Για μικρές τιμές σημάτων η κλίση της καμπύλης εισόδων/εξόδων είναι απότομη, συνεπώς παράγεται μεγάλη κλίση, ενώ η κλίση της σιγμοειδούς μικραίνει καθώς το μέγεθος του σήματος γίνεται μεγαλύτερο.

Εναλλακτικά, μπορεί να χρησιμοποιηθεί η συνάρτηση υπερβολικής εφαπτομένης (tan-sigmoid function), η οποία έχει παρόμοιο γράφημα με την σιγμοειδή συνάρτηση και συχνά χρησιμοποιείται για τεχνητά δίκτυα με πολλά κρυμμένα επίπεδα. Έχει την έκφραση:

$$OUT = \tanh(NET) \tag{2.4}$$



Tan-Sigmoid Transfer Function

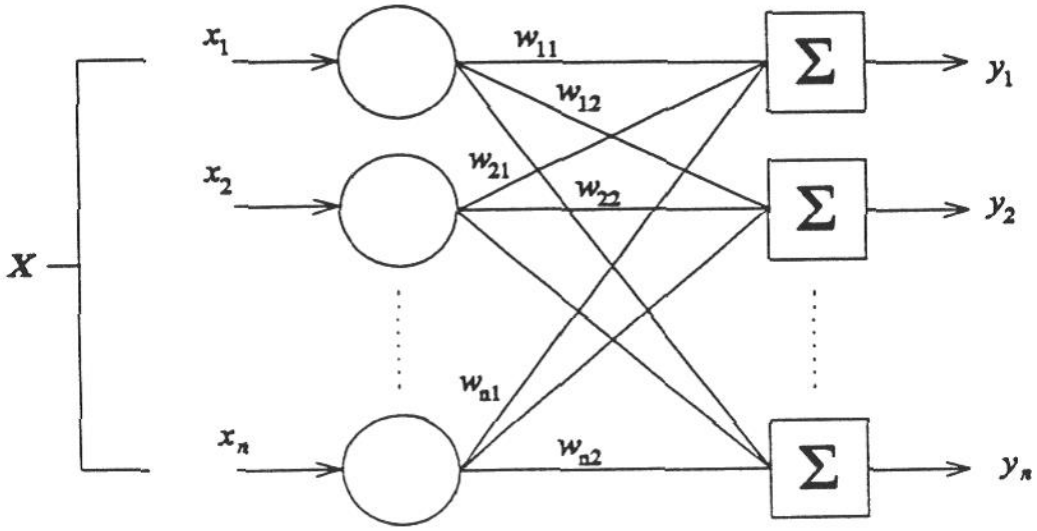
2.2.3 Δίκτυα πρόσθιας τροφοδότησης

Τα τεχνητά νευρωνικά δίκτυα αυτού του τύπου αποτελούνται από ένα ή περισσότερα *επίπεδα ή στοιβάδες* (layers) τα οποία περιέχουν τεχνητούς νευρώνες. Οι τεχνητοί νευρώνες, οι οποίοι ανήκουν σε γειτονικά επίπεδα συνδέονται μεταξύ τους μέσω των συναπτικών βαρών. Αυτού του τύπου οι αρχιτεκτονικές ονομάζονται *πρόσθιας τροφοδότησης* εφόσον οι έξοδοι κάθε επιπέδου τροφοδοτούν τους νευρώνες του αμέσως επόμενου επιπέδου. Τα μοντέλα Perceptron και Adaline αποτελούν τα πρώτα μοντέλα αυτού του τύπου αρχιτεκτονικής. Όμως, οι πραγματικές δυνατότητες των νευρωνικών δικτύων πρόσθιας τροφοδότησης

ανακαλύφθηκαν πρόσφατα με την εμφάνιση των πολυεπίπεδων νευρωνικών δικτύων. Ένα δίκτυο πρόσθιας τροφοδότησης μπορεί να προσομοιαστεί ως ένα σύστημα απεικόνισης ενός συνόλου εισόδων σε ένα σύνολο εξόδων κάτι το οποίο επιτυγχάνεται προσαρμόζοντας κατάλληλα το σύνολο των συναπτικών βαρών μέσω της διαδικασίας εκπαίδευσης.

- *Τεχνητό νευρωνικό δίκτυο ενός επιπέδου*

Οι υπολογιστικές δυνατότητες των νευρωνικών δικτύων προέρχονται από τη σύνδεση νευρώνων μεταξύ τους σε σχηματισμό δικτύου. Το πιο απλό δίκτυο είναι εκείνο το οποίο αποτελείται από μια ομάδα νευρώνων συνδεδεμένων μεταξύ τους σε ένα επίπεδο όπως φαίνεται στο Σχήμα 2.3.



Σχήμα 2.3 Τεχνητό νευρωνικό δίκτυο ενός επιπέδου

Οι κυκλικοί κόμβοι στα αριστερά της εικόνας χρησιμεύουν μόνο στην κατανομή των εισόδων στους νευρώνες, άρα δεν θεωρούνται ως επίπεδο. Η είσοδος κάθε νευρώνα είναι το άθροισμα των γινομένων [βάρος]×[είσοδος]. Πρακτικά τα τεχνητά και τα βιολογικά δίκτυα είναι δυνατό να μην έχουν πολλές από τις συνδέσεις τους ενεργές.

Τα βάρη του δικτύου σχηματίζουν το μητρώο **w**. Η διάσταση του μητρώου είναι m×n, όπου *m* είναι το πλήθος των εισόδων και *n* το πλήθος των νευρώνων. Για παράδειγμα το βάρος που συνδέει την τρίτη είσοδο με τον δεύτερο νευρώνα θα είναι το $w_{3,2}$. Οπότε, ο υπολογισμός του συνόλου των εξόδων των νευρώνων *NET*, για ένα επίπεδο είναι ένας απλός πολλαπλασιασμός μητρώων

$$\mathrm{NET} = \begin{bmatrix} x_1 & x_2 & . & . & x_m \end{bmatrix} \cdot \begin{bmatrix} w_{11} & w_{12} & . & . & . & w_{1n} \\ w_{21} & w_{22} & . & . & . & w_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ w_{m1} & w_{m2} & . & . & . & w_{mn} \end{bmatrix} \tag{2.5}$$

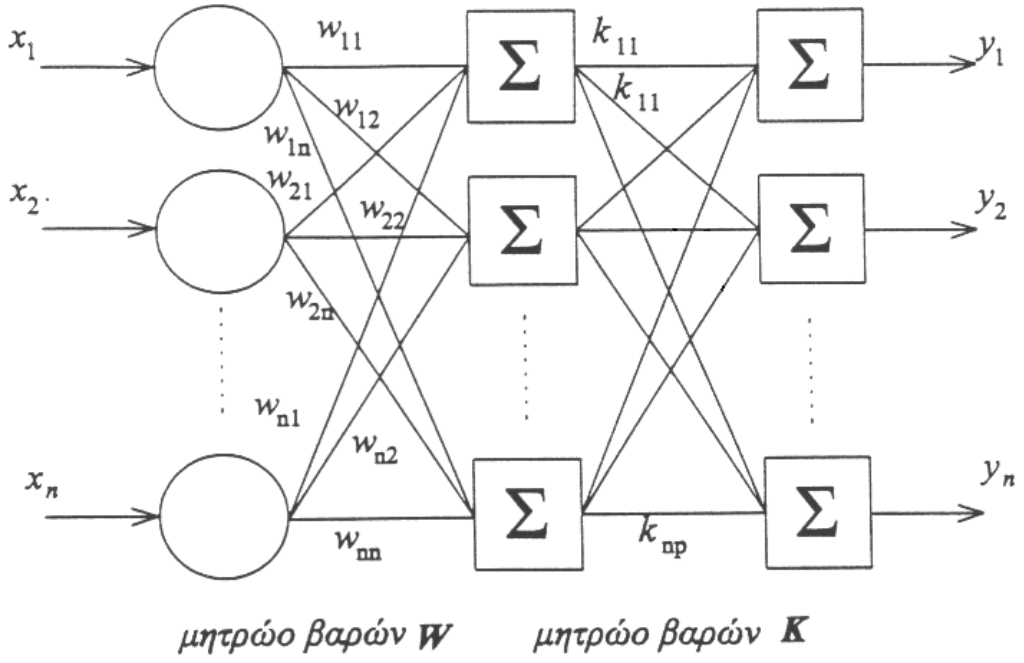
όπου οι έξοδοι του δικτύου είναι

$$OUT_i = F(\sum_{j=1}^{m} w_{j,i} \cdot x_j)$$

- *Τεχνητό νευρωνικό δίκτυο πολλών επιπέδων*

Σύντομα έγινε αντιληπτό ότι τα δίκτυα ενός επιπέδου δεν είναι δυνατό να δώσουν λύση σε μία μεγάλη ποικιλία προβλημάτων. Μεγαλύτερα, πιο πολύπλοκα δίκτυα προσφέρουν μεγαλύτερο δείκτη αξιοπιστίας. Η εισαγωγή των πολυεπίπεδων δικτύων έδωσε νέα ώθηση στον τομέα των νευρωνικών δικτύων. Ο κύριος λόγος ο οποίος απέτρεπε κατά το παρελθόν τη χρήση των πολυεπίπεδων δικτύων ήταν ότι δεν υπήρχε μία τεχνική κατάλληλη για την εκπαίδευση τέτοιων δικτύων. Ο αλγόριθμος της πίσω διάδοσης (**backpropagation**) που προτάθηκε από τους Rumelhart και McClelland χρησιμοποιήθηκε με επιτυχία για την εκπαίδευση πολυεπίπεδων δικτύων.

Τα πολυεπίπεδα δίκτυα σχηματίζονται από μία ομάδα επιπέδων κατάλληλα συνδεδεμένων μεταξύ τους. Οι έξοδοι ενός επιπέδου αποτελούν τις εισόδους του επόμενου επιπέδου. Στο Σχήμα 2.4 παρουσιάζεται ένα τέτοιο δίκτυο, το οποίο είναι δίκτυο με πλήρη σύνδεση εφόσον οι νευρώνες ενός επιπέδου είναι συνδεδεμένοι με όλους τους νευρώνες του επόμενου και του προηγούμενου επιπέδου. Τα επίπεδα των νευρώνων που βρίσκονται μεταξύ των επιπέδων εισόδου και εξόδου ονομάζονται κρυμμένα επίπεδα (hidden layers).



Σχήμα 2.4 Τεχνητό νευρωνικό δίκτυο δυο επιπέδων

2.2.4 Εκπαίδευση των τεχνητών νευρωνικών δικτύων

Ένα δίκτυο εκπαιδεύεται έτσι ώστε η παρουσίαση ενός νέου συνόλου εισόδων να παράγει ένα επιθυμητό ή τουλάχιστον συνεπές σύνολο εξόδων. Η εκπαίδευση επιτυγχάνεται με συνεχή εφαρμογή διανυσμάτων εισόδου και έλεγχο των διανυσμάτων εξόδου και συνεχή διόρθωση των συναπτικών βαρών του δικτύου σύμφωνα με μια προκαθορισμένη διαδικασία. Κατά τη διάρκεια της εκπαίδευσης, τα βάρη του δικτύου βαθμιαίως συγκλίνουν σε σταθερές τιμές έτσι ώστε κάθε διάνυσμα εισόδου να παράγει το επιθυμητό διάνυσμα εξόδου. Οι αλγόριθμοι εκπαίδευσης χωρίζονται σε δύο κατηγορίες, στους επιβλεπόμενης και μη επιβλεπόμενης εκπαίδευσης.

Όσον αφορά την πρώτη, τα στάδια από τα οποία αποτελείται είναι τα ακόλουθα: (i) επίδραση ενός διανύσματος εισόδου στο δίκτυο, (ii) υπολογισμός της εξόδου του δικτύου και σύγκριση με το αντίστοιχο διάνυσμα στόχου, (iii) επιστροφή στο δίκτυο του λάθους πρόβλεψης, (iv) διόρθωση των βαρών του δικτύου σύμφωνα με έναν αλγόριθμο ο οποίος στοχεύει στην ελαχιστοποίηση του σφάλματος πρόβλεψης.

Τα ζεύγη εκπαίδευσης εφαρμόζονται διαδοχικά και για κάθε ένα από αυτά υπολογίζεται το σφάλμα. Τα συναπτικά βάρη διορθώνονται με βάση το συνολικό σφάλμα του δικτύου για το σύνολο των ζευγών εκπαίδευσης. Η διαδικασία

διόρθωσης των βαρών συνεχίζεται μέχρι το λάθος για όλα τα μέλη του συνόλου εκπαίδευσης να φτάσει σε ένα αποδεκτά χαμηλό επίπεδο.

Η εκπαίδευση αυτού του είδους χρησιμοποιείται σπάνια καθώς είναι δύσκολο να εννοηθεί μηχανισμός εκμάθησης στον εγκέφαλο που θα συγκρίνει τις επιθυμητές με τις τρέχουσες εξόδους. Οπότε, κατά κύριο λόγο χρησιμοποιείται η εκπαίδευση χωρίς επίβλεψη που είναι μία καλύτερη προσέγγιση του βιολογικού αναλόγου. Αυτού του είδους η εκπαίδευση δεν απαιτεί κάποιο διάνυσμα στόχου για τη σύγκρισή του με τις προβλεπόμενες τιμές. Το σύνολο των ζευγών εκπαίδευσης αποτελείται μόνο από τα διανύσματα εισόδου. Οι αλγόριθμοι εκπαίδευσης διορθώνουν τα βάρη του δικτύου ώστε να παράγουν διανύσματα εξόδων που να είναι αποδεκτά. Δηλαδή, δύο παρόμοια διανύσματα εισόδου να παράγουν παρόμοια διανύσματα εξόδου.

2.3 Ο ΑΛΓΟΡΙΘΜΟΣ ΕΚΠΑΙΔΕΥΣΗΣ ΠΙΣΩ ΔΙΑΔΟΣΗΣ

Στα πρώτα στάδια εφαρμογής των NN τους ερευνητές δεν τους απασχολούσε η ανάπτυξη αλγορίθμων εκπαίδευσης πολυεπίπεδων νευρωνικών δικτύων. Όταν όμως διαπιστώθηκε ότι τα δίκτυα ενός επιπέδου είχαν περιορισμένες δυνατότητες τότε το ενδιαφέρον επεκτάθηκε και προς τα πολυεπίπεδα δίκτυα. Η ανακάλυψη του αλγορίθμου της πίσω διάδοσης κατόρθωσε να διευρύνει το πεδίο εφαρμογής των τεχνητών νευρωνικών δικτύων σημαντικά. Ο αλγόριθμος αυτός αποτελεί μία μέθοδο εκπαίδευσης πολυεπίπεδων τεχνητών νευρωνικών δικτύων πρόσθιας τροφοδότησης με επίβλεψη. Η εκπαίδευση μπορεί να θεωρηθεί ως ένα πρόβλημα ελαχιστοποίησης χωρίς περιορισμούς όπου τα βάρη είναι οι παράμετροι σχεδιασμού ενώ η συνάρτηση σφάλματος είναι η αντικειμενική συνάρτηση. Η μέθοδος της απότομης καθόδου είναι η πλέον ευρέως χρησιμοποιούμενη μέθοδος από τον αλγόριθμο της πίσω διάδοσης.

Στον αλγόριθμο της πίσω διάδοσης η εκπαίδευση διεξάγεται με τη διάδοση μέσω του δικτύου ενός συνόλου από εκπαιδευτικούς τύπους εισόδου. Το δίκτυο αυτό αποτελείται από ένα επίπεδο εισόδου, ένα ή περισσότερα κρυμμένα επίπεδα και ένα επίπεδο εξόδου. Κάθε επίπεδο αποτελείται από τις αντίστοιχες μονάδες του (στοιχεία επεξεργασίας, νευρώνες ή κόμβους) και συνδέσεις βάρους.

Οι είσοδοι $x_i,\ με\ i = 1, 2, \ldots, n$ οι οποίες γίνονται δεκτές από το επίπεδο εισόδου είναι ανάλογες των ηλεκτροχημικών συστημάτων που δέχονται οι νευρώνες στον ανθρώπινο εγκέφαλο. Στο απλούστερο μοντέλο, οι είσοδοι αυτές πολλαπλασιάζονται με τα βάρη των συνδέσεων και η ενεργής είσοδος $net_{p,j}$ στα στοιχεία επεξεργασίας είναι το σταθμισμένο άθροισμα των αρχικών εισόδων:

$$net_{p,j} = \sum_{i=1}^{n} w_{p,ij} net_{q,i} + b_{p,j} \tag{2.6}$$

όπου $w_{p,ij}$ αποτελεί το βάρος σύνδεσης ή το συναπτικό βάρος του επιπέδου p, το οποίο ξεκινάει από τον νευρώνα i του επιπέδου q (πηγή) και καταλήγει στον νευρώνα j του επιπέδου p (στόχος), $net_{q,i}$είναι η έξοδος που παράγεται στον νευρώνα j του επιπέδου q και $b_{p,j}$ είναι μια καθοδηγητική παράμετρος που χρησιμοποιείται για να εξομαλυνθεί (modulate) η έξοδος του στοιχείου. Οι είσοδοι x_i αντιστοιχούν στο $net_{q,i}$,$με\ i = 1,2,. . .,n$ με στην περίπτωση του επιπέδου εισόδου.

Εκπαίδευση

Αντικείμενο της εκπαίδευσης ενός δικτύου είναι ο προσδιορισμός των βαρών του δικτύου έτσι ώστε η εφαρμογή ενός συνόλου εισόδων να παράγει το επιθυμητό σύνολο εξόδων. Στη βασική έκδοση του αλγορίθμου πίσω διάδοσης χρησιμοποιείται ο *αλγόριθμος της απότομης καθόδου*. Η τοπική συνάρτηση σφάλματος για ένα τυχαίο p ζεύγος εκπαίδευσης μπορεί να μορφωθεί ως ακολούθως:

$$E_p = \frac{1}{2}\sum_{i=1}^{N_{out}}(tar_i{}^p - out_{k,i})^2 = \frac{1}{2}\sum_{i=1}^{N_{out}}(err_i{}^{(p)})^2 \qquad (2.7)$$

ενώ η καθολική συνάρτηση σφάλματος είναι η ακόλουθη

$$E = \sum_{p=1}^{N_p} E_p = \frac{1}{2}\sum_{p=1}^{N_p}\sum_{i=1}^{N_{out}}(tar_i{}^p - out_{k,i})^2 \qquad (2.8)$$

όπου $tar_i^{(p)}$ και $out_{k,i}$ είναι τα αποτελέσματα στόχου και τα αποτελέσματα πρόβλεψης, αντιστοίχως για τον νευρώνα i του επιπέδου εξόδου k, για το δείγμα εκπαίδευσης p. Οι τεταγμένες του διανύσματος στόχου βρίσκονται μεταξύ 0 και 1. Κατά συνέπεια, προκειμένου οι εξισώσεις 2.7, 2.8 να παράγουν λογικά αποτελέσματα, οι έξοδοι των εκπαιδευτικών διατάξεων που επιλέγονται, πρέπει να κανονικανοποιούνται στο ίδιο διάστημα τιμών.

Ο υπολογισμός του ελαχίστου της καθολικής συνάρτησης σφάλματος E γίνεται με τις παρακάτω δύο τεχνικές: Η πρώτη τεχνική ονομάζεται *on-line* κατά την οποία τα παραδείγματα εκπαίδευσης παρουσιάζονται σειριακά, συνήθως σε τυχαία σειρά διαδοχής. Για κάθε παράδειγμα, υπολογίζεται ξεχωριστά η τιμή της τοπικής συνάρτησης σφάλματος και βάσει αυτής γίνεται η τροποποίηση των τιμών των βαρών. Η δεύτερη τεχνική ονομάζεται *batch μέθοδος εκπαίδευσης*, βάσει της οποίας ελαχιστοποιείται η καθολική συνάρτηση σφάλματος E συσσωρεύοντας τις αλλαγές των βαρών για όλα τα δείγματα εκπαίδευσης πριν από την τελική διόρθωση των βαρών.

Η *on-line* τεχνική εκπαίδευσης δικτύων ακολουθεί τα βήματα: (i) επιλέγεται με τυχαίο τρόπο το επόμενο ζεύγος εκπαίδευσης από το σύνολο των ζευγών

εκπαίδευσης, το οποίο επιδρά στο δίκτυο, (ii) υπολογίζεται η έξοδος του δικτύου, (iii) υπολογίζεται το σφάλμα πρόβλεψης μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου, (iv) ελαχιστοποιείται το σφάλμα και διορθώνονται τα βάρη του δικτύου. Τα βήματα (i)-(iv) επαναλαμβάνονται για κάθε διάνυσμα του συνόλου εκπαίδευσης μέχρι το σφάλμα σε όλο το σύνολο να είναι κάτω από κάποιο αποδεκτό όριο.

Η *batch* τεχνική εκπαίδευσης δικτύων ακολουθεί τα βήματα: (i) διαδοχική επίδραση όλων των ζευγών στο δίκτυο, (ii) υπολογίζεται η έξοδος του δικτύου για κάθε ένα από αυτά τα ζεύγη εκπαίδευσης, (iii) υπολογίζεται το σφάλμα πρόβλεψης μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου, (iv) υπολογίζεται το μέγεθος της διόρθωσης για το σύνολο του σφάλματος και γίνεται ενημέρωση των βαρών.

Έλεγχος πορείας εκπαίδευσης

Υπάρχουν δύο κριτήρια βάση των οποίων ελέγχεται αν το δίκτυο έχει φτάσει σε ένα αποδεκτό επίπεδο πρόβλεψης ή όχι. Το πρώτο είναι το σφάλμα της τετραγωνικής ρίζας των μέσων τετραγώνων (Root Mean Square-RMS) το οποίο υπολογίζεται από τη σχέση

$$e_{RMS} = \sqrt{E} = \sqrt{\frac{1}{N_p N_{out}} \sum_{N_p} \sum_{i=1}^{N_{out}} (tar_i - out_i)^2} \tag{2.9}$$

όπου N_p είναι ο συνολικός αριθμός των ζευγών στο σύνολο εκπαίδευσης και N_{out} είναι ο αριθμός μονάδων στο επίπεδο εξόδου. Το μέγεθος e_{RMS} παρέχει ένα μέτρο της διαφοράς μεταξύ προβλεφθέντων και πραγματικών τιμών σε κάθε κύκλο εκπαίδευσης, για το σύνολο των ζευγών εκπαίδευσης. Αναφέρεται δηλαδή στο καθολικό σφάλμα του δικτύου.

Ένα δεύτερο κριτήριο ελέγχου της εκπαίδευσης είναι ο *βαθμός διόρθωσης* (correctness ratio) του συνόλου των ζευγών εκπαίδευσης, ο οποίος υπολογίζεται μέσω της σύγκρισης προβλεφθέντων-πραγματικών τιμών των εξόδων. Το δίκτυο θεωρείται “ακριβές” στην περίπτωση που οι προβλεπόμενες τιμές είναι κοντά στις επιθυμητές. Ο βαθμός διόρθωσης χαρακτηρίζει την τρέχουσα κατάσταση του δικτύου και είναι το ποσοστό των ζευγών του συνόλου εκπαίδευσης για τα οποία το δίκτυο εμφανίζεται να έχει “ακριβή” πρόβλεψη. Η βέλτιστη τιμή του βαθμού διόρθωσης είναι το 100%, δηλαδή η περίπτωση που για όλα τα ζεύγη εκπαίδευσης το δίκτυο εμφανίζει “σωστή” πρόβλεψη.

Όσον αφορά στη σχέση του e_{RMS} με τον βαθμό διόρθωσης δεν είναι πάντα αντιστρόφως ανάλογη. Είναι δυνατό να παρατηρηθεί κατάσταση του δικτύου με χαμηλή τιμή σφάλματος e_{RMS} το οποίο όμως να παρουσιάζει χαμηλή τιμή του

βαθμού διόρθωσης και συνεπώς ο βαθμός εκπαίδευσης του δικτύου να μην είναι αποδεκτός. Ας θεωρήσουμε ένα δίκτυο του οποίου τα περισσότερα πρότυπα οδηγούν σε τιμή τοπικού σφάλματος $err_{k,i} = tar_{k,i} - out_{k,i}$ κοντά στο μηδέν και τα υπόλοιπα κοντά στη μονάδα. Όπως αναφέρθηκε προηγουμένως, οι έξοδοι του δικτύου είναι κανονικοποιημένες στο διάστημα [0,1] οπότε στην περίπτωση που η τιμή του τοπικού σφάλματος είναι κοντά στη μονάδα η πρόβλεψη είναι λανθασμένη. Σε αυτή την περίπτωση η τιμή του βαθμού διόρθωσης δεν μπορεί να είναι υψηλή αν και το συνολικό σφάλμα e_{RMS} εμφανίζει χαμηλή τιμή. Αν θεωρηθεί ένα δεύτερο δίκτυο του οποίου το τοπικό σφάλμα είναι $err_{k,i} = tar_{k,i} - out_{k,i} \approx 0.1$ για όλα τα πρότυπα, τότε η τιμή του βαθμού διόρθωσης θα είναι κοντά στο 100%. Το φαινόμενο αυτό είναι συχνό για την περίπτωση μεγάλου πλήθους προτύπων εκπαίδευσης. Το φαινόμενο αυτό εμφανίζεται επίσης για την περίπτωση μη κατάλληλων αρχιτεκτονικών ή στην περίπτωση λανθασμένου συνδυασμού των παραμέτρων εκπαίδευσης και του momentum.

Όρος momentum

Οι πρώτοι που πρότειναν μία μέθοδο για τη μείωση του χρόνου εκπαίδευσης του αλγορίθμου πίσω διάδοσης καθώς και για την αύξηση της ευστάθειας του αλγορίθμου ήταν οι Rumelhart, Hinton και Williams το 1986. Η μέθοδος ονομάζεται *momentum backpropagation* και βασίζεται στην προσθήκη ενός όρου στην εξίσωση διόρθωσης των βαρών, που είναι ανάλογος κάποιου ποσοστού της διόρθωσης του προηγούμενου βήματος. Η διόρθωση των βαρών με τον όρο αυτό και για την περίπτωση του επιπέδου εξόδου k γράφεται

$$\Delta w_{k,ij}^{t+1} = \alpha \delta_{k,i} out_{p,j} + a\Delta w_{k,ji}^{t} \tag{2.10}$$

ενώ οι τρέχουσες τιμές των βαρών δίνονται από την εξίσωση

$$w_{k,ji}^{t+1} = w_{k,ji}^{t} + \Delta w_{k,ij}^{t+1} \tag{2.11}$$

όπου a είναι ο συντελεστής *momentum* ο οποίος εκλέγεται συνήθως κοντά στο 0.5 κατά την έναρξη της διαδικασίας ενώ στην περίπτωση που εντοπιστεί κάποιο τοπικό ελάχιστο αυξάνει μέχρι την τιμή 0.9.

Οι Sejnowski και Rosenberg το 1987 πρότειναν μία παρόμοια μέθοδο για τη βελτίωση του κλασικού αλγορίθμου της πίσω διάδοσης, σύμφωνα με την οποία το μέγεθος της διόρθωσης των βαρών γίνεται με τον ακόλουθο τρόπο

$$\Delta w_{k,ij}^{t+1} = (1 - a)\delta_{k,i} out_{p,j} + a\Delta w_{k,ji}^{t} \tag{2.12}$$

ενώ οι αλλαγές των βαρών υπολογίζονται από την σχέση

$$w_{k,ji}^{t+1}=w_{k,ji}^{t}+\alpha\Delta w_{k,ij}^{t+1} \quad (2.13)$$

όπου a είναι ένας συντελεστής με τιμή από το διάστημα [0, 1] όπως και το α.

Στην περίπτωση που ο συντελεστής a λάβει την τιμή 0 τότε η διόρθωση του βάρους περιλαμβάνει μηδενικό ποσοστό της προηγούμενης διόρθωσης. Αντίθετα, αν ο συντελεστής λάβει την τιμή 1, τότε η νέα αλλαγή αγνοείται και επαναλαμβάνεται η προηγούμενη διόρθωση. Ο συντελεστής η χρησιμεύει στη ρύθμιση του μεγέθους της μέσης αλλαγής των βαρών και είναι επιλογή του χρήστη.

3

ΜΕΘΟΔΟΣ ΠΡΟΣΟΜΟΙΩΣΗΣ KRIGING

3.1 ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό παρουσιάζονται οι βασικότερες μορφές του υπολογιστικού μοντέλου Kriging, που επιτρέπει τη δημιουργία μοντέλων για την προσέγγιση κάποιου υπολογιστικού πειράματος με πολύ καλή ακρίβεια, μειώνοντας σε μεγάλο βαθμό το χρόνο που απαιτείται για την προσομοίωση.

Μέχρι σήμερα, έχουν ερευνηθεί εκτενώς πολλές μέθοδοι οι οποίες βασίζονται στην παλινδρόμηση και την παρεμβολή και χρησιμοποιούνται κατά κόρον στη δημιουργία μεταμοντέλων. Λέγοντας μεταμοντέλο εννοούμε μια προσέγγιση της συνάρτησης μεταξύ δεδομένων εισόδου-εξόδου η οποία καθορίζεται από το υποκείμενο μοντέλο προσομοίωσης, το οποίο μπορεί να είναι είτε ντετερμινιστικό είτε τυχαίο. Οι περισσότερες μέθοδοι που έχουν χρησιμοποιηθεί, λοιπόν, για την παραγωγή μεταμοντέλου παρουσιάζουν το μειονέκτημα του ότι δε λαμβάνουν υπόψη την επίδραση της συσχέτισης μεταξύ των παραμέτρων σχεδιασμού και των αποτελεσμάτων που αυτές εξάγουν. Επίσης, επικεντρώνονται σε χαμηλής τάξης πολυωνυμική παλινδρόμηση οπότε ταιριάζουν κυρίως στα δεδομένα που αφορούν τοπικά και όχι παγκόσμια πειράματα.

Η μέθοδος Kriging, από την άλλη, παρέχει μια τεχνική δημιουργίας μεταμοντέλου η οποία επηρεάζεται από τη συσχέτιση αυτή που συμβαίνει κατά τη φάση της παραγωγής του μεταμοντέλου. Επιπρόσθετα, τα μοντέλα Kriging ταιριάζουν στα δεδομένα που λαμβάνονται για μεγαλύτερες πειραματικές περιοχές συγκριτικά με τις περιοχές που αντιστοιχούν σε μεταμοντέλα με χαμηλής τάξης πολυωνυμική παλινδρόμηση. Τα συγκεκριμένα μοντέλα χρησιμοποιούνται για πρόβλεψη έχοντας ως τελικούς στόχους την ανάλυση ευαισθησίας και τη βελτιστοποίηση.

Το Kriging αναπτύχθηκε αρχικά ώστε να εφαρμοστεί στη γεωστατιστική - για το λόγο αυτό είναι γνωστό και ως χωρική στατιστική - από το μεταλλειολόγο της Νοτίου Αφρικής Danie Krige. Στον τομέα των μαθηματικών χρησιμοποιήθηκε αργότερα από το Γάλλο μαθηματικό Georges Matheron. Έπειτα, εφαρμόστηκε και σε δεδομένα εισόδων- εξόδων (I/O) σε μοντέλα ντετερμινιστικής προσομοίωσης. Τα μοντέλα αυτά έχουν εισόδους k διαστάσεων, όπου k είναι ένας δεδομένος θετικός ακέραιος αριθμός (ενώ στη γεωστατιστική οι είσοδοι είναι μόνο δυο διαστάσεων). Οι πιο πρόσφατες δημοσιεύσεις περί τούτου είναι του Jones το 1998, του Simpson το 2001 και του Santner το 2003.

Το 2003, όμως, ο Jack P.C. Kleijnen και ο Van Beers ξεκίνησαν να εφαρμόζουν το Kriging και σε μοντέλα τυχαίας προσομοίωσης. Και παρότι χρησιμοποιείται ακόμη σπάνια σε τέτοιου είδους προβλήματα, η αποδοτικότητα που επιτυγχάνει στην ντετερμινιστική προσομοίωση δίνει ελπίδες ότι μπορεί να έχει εξίσου καλά αποτελέσματα και σε αυτόν τον τομέα.

3.2 ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΟΥ KRIGING

Γραμμική παλινδρόμηση

Αρχικά, επισημαίνονται οι διαφορές μεταξύ της γραμμικής παλινδρόμησης-ειδικά της χαμηλής τάξης πολυωνιμικής παλινδρόμησης- και του μοντέλου Kriging. Παρακάτω θεωρείται μια έξοδος του μοντέλου προσομοίωσης καθώς συνήθως το Kriging κάνει αυτήν την υπόθεση. Οπότε, η γενική μορφή του μοντέλου είναι:

$$w = s\,(d_1, \ldots, d_{k,} r_0) \tag{3.1}$$

όπου:

w είναι η έξοδος του μοντέλου

s υποδηλώνει τη μαθηματική συνάρτηση που ορίζεται έμμεσα από τον υπολογιστικό κώδικα που εφαρμόζει το μοντέλο προσομοίωσης

d_j $όπου\ j = 1, \ldots, k$ είναι οι τιμές σχεδιασμού για τις μεταβλητές εισόδου του προγράμματος προσομοίωσης

r_0 είναι το διάνυσμα των ψευδοτυχαίων αριθμών

Το μοντέλο παλινδρόμησης πρώτης τάξης για την εξίσωση (3.1) είναι:

$$y_{reg} = \beta_0 + \beta_1 d_1 + \ldots + \beta_\kappa d_\kappa + e_{reg} \tag{3.2}$$

όπου:

y_{reg} είναι η πρόβλεψη παλινδρόμησης για το αποτέλεσμα w της εξίσωσης(3.1). Ο δείκτης *reg* διακρίνει αυτό το μεταμοντέλο από το Kriging το οποίο παρουσιάζεται παρακάτω

$\beta = (\beta_0, \beta_1, . . . , \beta_\kappa)'$ είναι το διάνυσμα με τις παραμέτρους του μεταμοντέλου

e_{reg} είναι το σφάλμα του μεταμοντέλου

Η εξίσωση (3.2) είναι μια μορφή του γενικού γραμμικού μοντέλου παλινδρόμησης:

$$y_{reg} = X\beta + e_{reg} \tag{3.3}$$

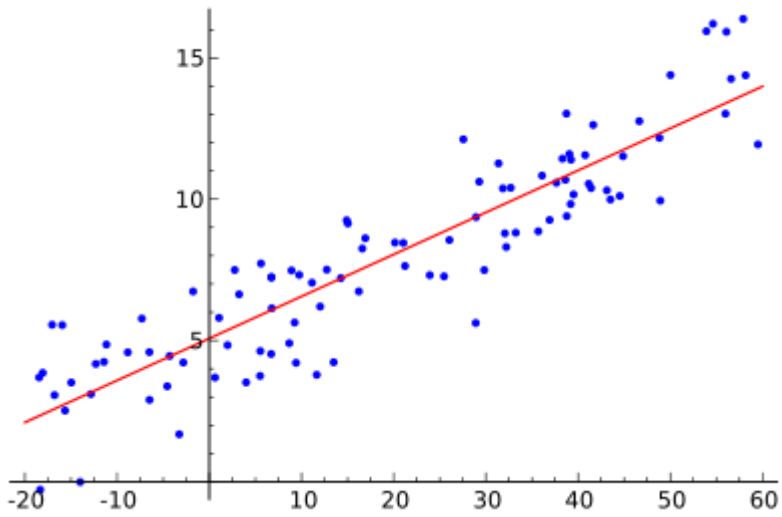
όπου:

y_{reg} υποδηλώνει το διάνυσμα n διαστάσεων με την πρόβλεψη παλινδρόμησης (προβλεπόμενα αποτελέσματα) όπου n είναι ο αριθμός των συνδυασμών προσομοίωσης

$X = (x_{ij})$ αντιστοιχεί στον πίνακα *n x q* των λεγόμενων παλινδρομητών. x_{ij} είναι η τιμή της μεταβλητής *j* στη δοκιμή *i*. Πολλές φορές ο πίνακας αυτός καλείται και πίνακας σχεδιασμού. Οι παλινδρομητές x_{ij} μπορούν να θεωρηθούν είτε τυχαίες μεταβλητές, τις οποίες απλώς έχουμε παρατηρήσει, είτε προκαθορισμένες σταθερές τιμές που τις επιλέγουμε εμείς. Και οι δύο ερμηνείες μπορεί να είναι κατάλληλες σε διαφορετικές περιπτώσεις.

β είναι το *q* διαστάσεων διάνυσμα παραμέτρων. Τα στοιχεία του ονομάζονται επίσης επιπτώσεις ή συντελεστές παλινδρόμησης. Η στατιστική εκτίμηση και η συμπερασματολογία στη γραμμική συσχέτιση επικεντρώνονται στην παράμετρο β.

e_{reg} περιλαμβάνει όλους τους άλλους παράγοντες που επηρεάζουν την εξαρτημένη μεταβλητή y_{reg} εκτός από τους παλινδρομητές x_{ij}. Η σχέση μεταξύ του όρου σφάλματος και των παλινδρομητών, για παράδειγμα, εάν συσχετίζονται, είναι ένα κρίσιμο βήμα για τη διαμόρφωση ενός γραμμικού μοντέλου παλινδρόμησης, καθώς θα καθορίσει τη μέθοδο που θα χρησιμοποιηθεί για την εκτίμηση.



Εικόνα 3.1 Παράδειγμα γραμμικής παλινδρόμησης

Ανάλυση μεθόδου Kriging για ντετερμινιστικές προσομοιώσεις

Στην ενότητα αυτή, η προσομοίωση y υποτίθεται ότι είναι μια ντετερμινιστική πραγματική συνάρτηση με μεταβλητές d-διαστάσεων $x = (x_1, \dots, x_d) \in D \subset \mathcal{R}^d$. Ως $X = \{x^{(1)}, \dots, x^{(n)}\}$ θεωρούνται τα σημεία στα οποία το y έχει ήδη εκτιμηθεί και οι τιμές του είναι $y = (y(x^1, \dots, x^n))^T$. Για κάθε $x \in D$, ο σκοπός του Kriging είναι να βρίσκει τη βέλτιστη πρόβλεψη Y_x μέσω ενός γραμμικού συνδυασμού των παρατηρήσεων y. Παρακάτω αναλύονται οι βασικότερες μορφές του Kriging και οι διαφορές τους από τη γραμμική παλινδρόμηση.

***Simple Kriging**: Από τη χωρική γραμμική παρεμβολή στη μέθοδο Gaussian*

Στο *Simple Kriging* (SK), η έξοδος του μοντέλου-δηλαδή το Y- θεωρείται ότι είναι το άθροισμα μιας γνωστής ντετερμινιστικής γραμμικής συνάρτησης μ : x ∈ $D \longrightarrow$ μ(x) ∈ $\mathcal{R}$ και μιας ολοκληρώσιμης συνάρτησης Z:

$$Y(x) = \mu(x) + Z(x) \tag{3.4}$$

Όπου η συνδιακύμανση C : (u, v) ∈ $D^2 \longrightarrow$ C(u, v) ∈ $\mathcal{R}$ μεταξύ των μεταβλητών της Z είναι γνωστή. Συνήθως , η συνάρτηση Z θεωρείται πολυώνυμο δευτέρου βαθμού έτσι ώστε $C(u, v) = \sigma^2 R(u - v; \psi)$όπου R η λεγόμενη συνάρτηση συσχέτισης με παραμέτρους ψ και σ² η λεγόμενη διακύμανση. Μπορούμε να θεωρήσουμε ότι, η καλύτερη πρόβλεψη για το Y_x , βασιζόμενη στις γνωστές παρατηρήσεις $Y(X)$, εξάγεται αν βρούμε τα κατάλληλα $\lambda * (x)$ τα οποία ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα MSE(x) = $E[\left(Y(x) - \lambda(x)^T Y(X)\right)^2]$. Αν ο C είναι αντιστρέψιμος, το διάγραμμα που απεικονίζει το MSE έχει κυρτότητα της μορφής U η οποία εξασφαλίζει την ύπαρξη αλλά και τη μοναδικότητα του $\lambda * (x)$. Έτσι, τα

βάρη για το SK δίνονται από τη σχέση $\lambda*(x) = C^{-1}c(x)$ όπου $C = (C(x^{(i)}, x^{(j)}))_{1\leq i,j\leq n}$ είναι ο πίνακας συνδιακύμανσης των $Y(X)$ και $c(x) = (C(x, x^{i}))_{1\leq i\leq n}$ είναι το διάνυσμα συνδιακύμανσης μεταξύ των $Y(x)$ και $Y(X)$. Αντικαθιστώντας το διάνυσμα $Y(X)$ με την πραγματική του τιμή y και το $\lambda(x)$ με $\lambda*(x)$στην έκφραση $\lambda(x)^T Y(X)$, έχουμε ως αποτέλεσμα τη λεγόμενη SK μέση πρόβλεψη στο x: $m_{SK}(x) = c(x)^T C^{-1} y$. Ομοίως, αντικαθιστώντας το βέλτιστο $\lambda*(x)$ στην έκφραση του MSE παίρνουμε τη διακύμανση στο x: ${s^2}_{SK}(x) = C(x,x) - c(x)^T C^{-1} c(x)$. Οπότε, έχουμε τις εξισώσεις:

$$m_{SK}(x) = \mu(x) + c(x)^T C^{-1}(y - \mu) \tag{3.5}$$

$${s^2}_{SK}(x) = C(x,x) - c(x)^T C^{-1} c(x) \tag{3.6}$$

όπου:

$\mu = \mu(X)$ αφορά στις μέσες τιμές των σημείων όπου το y έχει ήδη εκτιμηθεί και αποτελεί γνωστό διάνυσμα για το Simple Kriging .

Η συνάρτηση m_{SK} θεωρείται ότι παρεμβάλλει τα δεδομένα (X, y) και η συνάρτηση ${s^2}_{SK}$ θεωρείται πως είναι μη αρνητική και 0 στα σημεία σχεδιασμού του πειράματος. Επιπλέον, η διακύμανση του SK είναι ανεξάρτητη του y.

Ένα σημαντικό στοιχείο που αφορά στις εξισώσεις SK είναι ότι συμπίπτουν με τα κλασικά αποτελέσματα $Y(X) = y$ στην περίπτωση όπου η συνάρτηση Z ακολουθεί τη μέθοδο Gaussian.

Όπως αναφέρθηκε παραπάνω, στο Kriging προκειμένου να επιλεγούν οι βέλτιστες τιμές για τα βάρη, χρησιμοποιείται ένα κριτήριο το οποίο ελαχιστοποιεί το MSE της πρόβλεψης. Το κριτήριο, λοιπόν, που επιλέγεται είναι το Best Linear Unbiased Predictor (BLUP), σε αντίθεση με τη γραμμική παλινδρόμηση όπου χρησιμοποιείται το κριτήριο Sum of Squared Residuals για την εκτίμηση των παραμέτρων *β*.

Ordinary και Universal Kriging*: Όταν κάποιοι συντελεστές είναι άγνωστοι*

Υπάρχει περίπτωση η συνάρτηση $\mu(x)$ να μην είναι γνωστή, όπως ήταν παραπάνω, αλλά να γνωρίζουμε απλά ότι είναι της μορφής $\mu(x) = \Sigma_{j=1}^{p} \beta_j f_j(x)$ $(p \in \mathcal{N} - \{0\})$ όπου f_j είναι σταθερές συναρτήσεις και β_j είναι άγνωστοι πραγματικοί συντελεστές. Το *Universal Kriging* (UK) συνίσταται στο να καταλήγει στις καλύτερες γραμμικές προβλέψεις για το Y ,με βάση τις παρατηρήσεις $Y(X)$, έπειτα από εκτίμηση του διανύσματος $\beta = (\beta_1, \dots, \beta_p)^T$. Στην ιδιαίτερη περίπτωση, όπου οι συναρτήσεις f_j καταλήγουν σε μια μοναδική σταθερή συνάρτηση, το *Universal Kriging* αναφέρεται ως *Ordinary Kriging (*OK). Οι εξισώσεις για το UK είναι:

$$m_{UK}(x) = f(x)^T\hat{\beta} + c(x)^T C^{-1}(y - F\hat{\beta}) \tag{3.7}$$

$$s_{UK}{}^2(x) = s_{SK}{}^2(x) + (f(x)^T - c(x)^T C^{-1}F)^T (F^T C^{-1}F)^{-1}(f(x)^T - c(x)^T C^{-1}F) \tag{3.8}$$

όπου:

$\hat{\beta} = (F^T C^{-1}F)^{-1}F^T C^{-1}y$ και αποτελεί την καλύτερη γραμμική εκτίμηση του β.

Οι βασικές αρχές του UK είναι όμοιες με του SK, δηλαδή και τα δυο μοντέλα θεωρούν παρόμοια συμπεριφορά στην παρεμβολή με τη διακύμανση να μηδενίζεται για τα σημεία σχεδιασμού του πειράματος. Επιπλέον, το $m_{UK}(x)$ τείνει να είναι ίσο με $f(x)^T\hat{\beta}$ όταν εκλείπει η συνδιακύμανση $c(x)$, το οποίο είναι πιθανό να συμβεί όταν το x απέχει αρκετά από το X.

Το συγκεκριμένο μοντέλο (UK) περιέχει μια αβεβαιότητα- και ίσως ένα μεγαλύτερο σφάλμα- εξαιτίας των συντελεστών β, οι οποίοι μπορεί να μην εκτιμηθούν σωστά. Όπως και στην περίπτωση του SK, όμως, είναι δυνατό τα αποτελέσματα και αυτού του μοντέλου Kriging να είναι πολύ κοντά στην πραγματικότητα.

3.3 ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ KRIGING – ΒΑΣΙΚΕΣ ΠΑΡΑΜΕΤΡΟΙ

3.3.1 Μοντέλα παλινδρόμησης

Τα μοντέλα τα οποία παρέχει το Kriging είναι με πολυώνυμα βαθμού 0, 1 και 2. Πιο συγκεκριμένα, θεωρώντας ως x_n το στοιχείο n της μεταβλητής x, έχουμε τα εξής μοντέλα παλινδρόμησης:

- Σταθερό (Constant) , $p = 1$: $f_1(x) = 1$
- Γραμμικό (Linear), $p = n + 1$: $f_1(x) = 1,\ \ f_2(x) = x_1,\ \ .\ .\ .\ .\ , f_{n+1}(x) = x_n$
- Τετραγωνικό (Quadratic), $p = \frac{1}{2}(n + 1)(n + 2)$:

  $f_1(x) = 1$

  $f_2(x) = x_1,\ \ .\ .\ .\ .\ , f_{n+1}(x) = x_n$

  $f_{n+2}(x) = x_1{}^2,\ \ .\ .\ .\ .\ , f_{2n+1}(x) = x_1 x_n$

  $f_p(x) = x_n{}^2$

Οι αντίστοιχοι Ιακωβιανοί πίνακες είναι

Σταθερό: $J_f = [O_{nx1}]$

Γραμικό: $J_f = [O_{nx1} \quad I_{nxn}]$

Τετραγωνικό: $J_f = [O_{nx1} \quad I_{nxn} \quad H]$

Όπου Ο είναι ο πίνακας που περιέχει τις μηδενικές τιμές και ο πίνακας Η είναι:

Για $n = 2$: $H = \begin{bmatrix} 2x_1 & x_2 & 0 \\ 0 & x_1 & 2x_2 \end{bmatrix}$

Για $n = 3$: $H = \begin{bmatrix} 2x_1 & x_2 & x_3 & 0 & 0 & 0 \\ 0 & x_1 & 0 & 2x_2 & x_3 & 0 \\ 0 & 0 & x_1 & 0 & x_2 & 2x_3 \end{bmatrix}$

Το Kriging παρέχει εντολές μέσω των οποίων εφαρμόζονται τα είδη πολυωνύμων που αναφέρθηκαν. Έτσι, ανάλογα με τη σχέση που υπάρχει μεταξύ μεταβλητών εισόδου και εξόδου επιλέγεται το κατάλληλο μοντέλο παλινδρόμησης και κατά συνέπεια η αντίστοιχη εντολή.

3.3.2 Μοντέλα συσχέτισης

Η κύρια μορφή που έχουν τα μοντέλα συσχέτισης στη μέθοδο Kriging είναι:

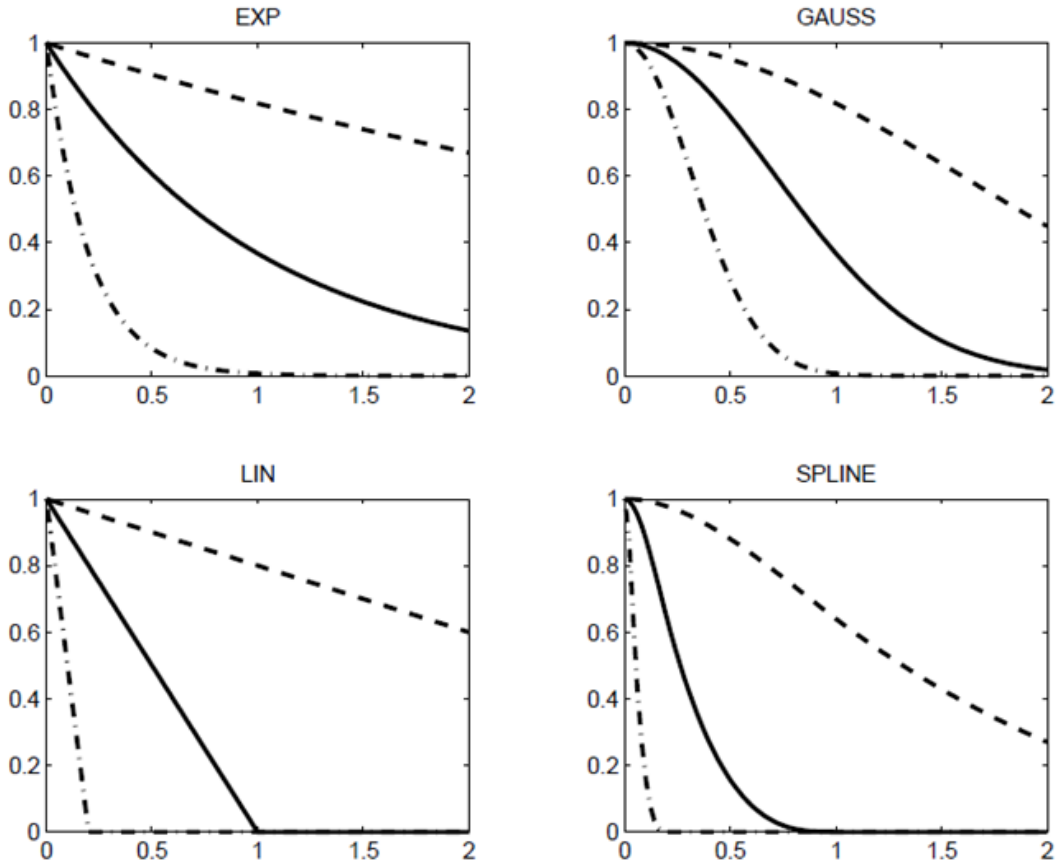
$$R(\theta, w, x) = \prod_{j=1}^{n} R_j(\theta, w_j - x_j) \tag{3.9}$$

Και, πιο συγκεκριμένα, η μέθοδος παρέχει 7 επιλογές της μορφής αυτής για τη λειτουργία συσχέτισης.

Πίνακας 3.1 Συναρτήσεις συσχέτισης $(\boldsymbol{d_j} = \boldsymbol{w_j} - \boldsymbol{x_j})$

ΟΝΟΜΑ	$R_j(\theta, d_j)$
EXP	$exp(-\theta_j \lvert d_j \rvert)$
EXPG	$exp\left(-\theta_j \lvert d_j \rvert^{\theta_{n+1}}\right) \; 0 \le \theta_{n+1} \le 2$
GAUSS	$\exp(-\theta_j {d_j}^2)$
LIN	$\max\{0, 1 - \theta_j \lvert d_j \rvert\}$
SPHERICAL	$1 - 1.5\xi_j + 0.5{\xi_j}^3, \qquad \xi_j = \min\{1, \theta_j \lvert d_j \rvert\}$
CUBIC	$1 - 3{\xi_j}^2 + 2{\xi_j}^3, \xi_j = \min\{1, \theta_j \lvert d_j \rvert\}$
SPLINE	$\varsigma(\xi_j), \xi_j = \theta_j \lvert d_j \rvert$

Μερικές από τις συναρτήσεις αυτές παρουσιάζονται παρακάτω, όπου φαίνεται ότι για όλες τις περιπτώσεις η συσχέτιση μειώνεται όσο αυξάνεται το $|d_j|$ και η μείωση αυτή είναι ακόμη πιο γρήγορη για μεγαλύτερες τιμές του θ_j.



Σχήμα 3.1 Συναρτήσεις συσχέτισης για $0 \leq d_j \leq 2$ και $\theta_j = 0.2, 1, 5$

Οι συναρτήσεις που αναφέρθηκαν στον πίνακα 3.1 μπορούν να χωριστούν σε δυο κατηγορίες, η πρώτη από τις οποίες περιέχει τις συναρτήσεις με παραβολική συμπεριφορά κοντά στην περιοχή των αξόνων (GAUSS, CUBIC, SPLINE), και η δεύτερη περιέχει εκείνες που παρουσιάζουν γραμμική συμπεριφορά κοντά στην αρχή των αξόνων (EXP, LIN, SPHERICAL). Η συνάρτηση EXPG μπορεί να λάβει και τις δυο μορφές, ανάλογα με την τιμή του θ_{n+1}. Όταν $\theta_{n+1} = 1$ παίρνει τη μορφή της εκθετικής συνάρτησης συσχέτισης (EXP) και όταν $\theta_{n+1} = 2$ λαμβάνει μορφή Gaussian.

Η επιλογή της κατάλληλης συνάρτησης σχετίζεται με το είδος του φαινομένου που μελετάται, για παράδειγμα αν πρόκειται για κάποιο φυσικό φαινόμενο τότε έχουμε γραμμική συμπεριφορά κοντά στην αρχή των αξόνων, οπότε οι συναρτήσεις exp, expg, linear και spherical ταιριάζουν καλύτερα. Επίσης, να σημειωθεί εδώ ότι για μεγάλη απόσταση από την αρχή των αξόνων η συσχέτιση είναι 0 σύμφωνα με τις

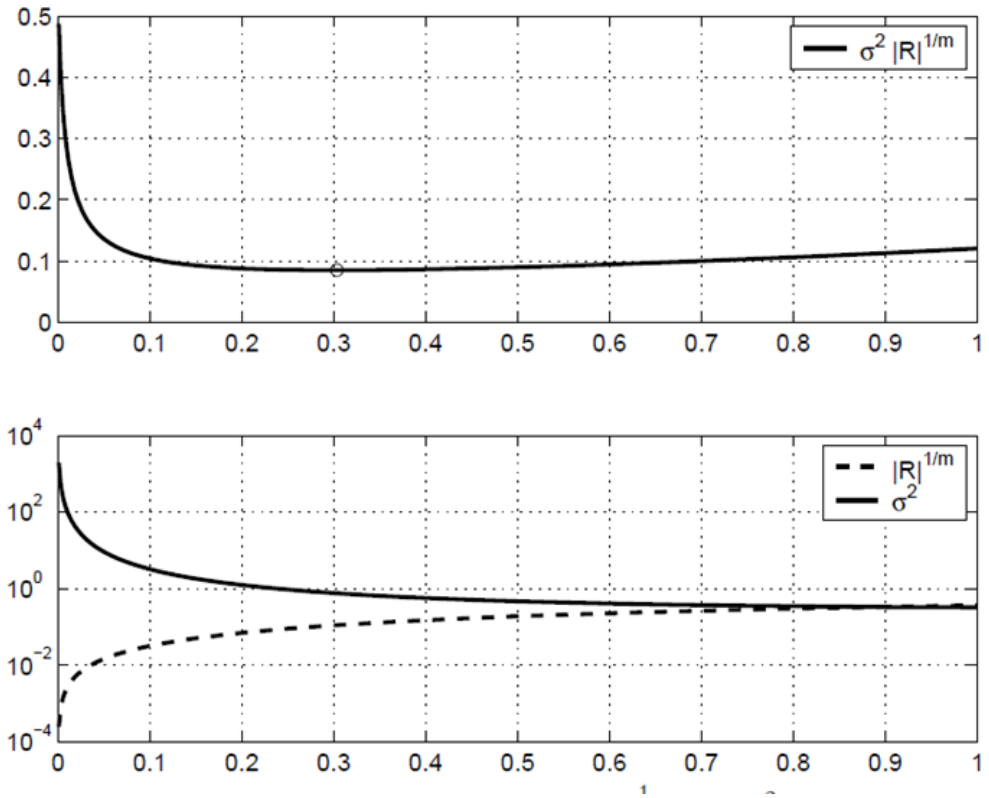
συναρτήσεις cubic, spline, linear και spherical ενώ για τις υπόλοιπες συναρτήσεις είναι ασύμπτωτη στον άξονα x (δηλαδή σχεδόν 0).

Όπως ήδη αναφέρθηκε, η μορφή μιας συνάρτησης συσχέτισης εξαρτάται από την παράμετρο θ και όπως φαίνεται και στο σχήμα 3.1 ανάλογα με την τιμή που παίρνει η παράμετρος αυτή, αλλάζει σχήμα και η συνάρτηση. Εφαρμόζοντας τη μέθοδο Gaussian, θεωρούμε πως οι βέλτιστοι συντελεστές θ^* είναι εκείνοι για τους οποίους ισχύει:

$$min_\theta\{\psi(\theta) \equiv |R|^{\frac{1}{m}}\sigma^2\} \tag{3.10}$$

όπου$|R|$ είναι η ορίζουσα της R

Ο ορισμός αυτός του θ^* αντιστοιχεί στην καλύτερη εκτίμηση πιθανότητας. Στο σχήμα παρακάτω απεικονίζεται η τυπική συμπεριφορά των παραμέτρων που συμμετέχουν στην επιλογή του θ^*.



Σχήμα 3.2 Τυπική συμπεριφορά των $\psi, |R|^{\frac{1}{m}}$ και σ^2 για $0 \leq \theta_j \leq 1$

Παρατηρούμε ότι το μέγεθος $|R|^{\frac{1}{m}}$ συνεχώς αυξάνεται σε όλο το διάστημα, κάτι το οποίο είναι αναμενόμενο, καθώς το R_θ είναι κοντά στη μονάδα για μεγάλα θ ενώ για μικρές τιμές του θ δεν ορίζεται. Για $\theta = 0$, ο πίνακας R είναι παντού 1. Σε περίπτωση που έχουμε τιμή του R η οποία δεν ορίζεται, τότε $\sigma^2 = \psi(\theta) =$ "∞".

3.3.3 Πειραματικός Σχεδιασμός

Ο πειραματικός σχεδιασμός τίθεται στο πλαίσιο αυτό ώστε να αποφασιστεί ποια από τα στοιχεία των εισόδων θα χρησιμοποιηθούν για να τρέξει ο ντετερμινιστικός κώδικας και θα είναι τα πιο κατάλληλα στο να ελέγξουν ή να μειώσουν τη στατιστική αβεβαιότητα της πρόβλεψης. Οι βασικότεροι αλγόριθμοι που εφαρμόζονται στο βήμα αυτό είναι το *Rectangular Grid* (ορθογώνιο πλέγμα) και το *Latin Hypercube Sampling* (υπερκύβος). Ο πρώτος αλγόριθμος παράγει ντετερμινιστικά σχέδια, ενώ ο δεύτερος παράγει σχέδια τα οποία βασίζονται σε τυχαίους αριθμούς.

Rectangular Grid

Ας υποθέσουμε ότι η περιοχή που μας ενδιαφέρει είναι η $D \in \mathcal{R}$, η οποία ορίζεται από τα σημεία $l_j \leq x_j \leq u_j$, $j = 1, \ldots, n$. Η απλούστερη κατανομή των σημείων σχεδιασμού ορίζεται από όλους τους διαφορετικούς συνδυασμούς των

$$s_j^{(i)} = l_j + k_j^{(i)} \frac{u_j - l_j}{\nu_j}, \qquad k_j^{(i)} = 0,1,\ldots,\nu_j$$

όπου οι $\{\nu_j\}$ είναι ακέραιοι αριθμοί. Αν όλα τα $\nu_j = \nu$, τότε ο αριθμός των σημείων σχεδιασμού είναι $(\nu + 1)^n$.

Η εντολή η οποία καλείται για το σχεδιασμό του ορθογωνίου πλέγματος είναι

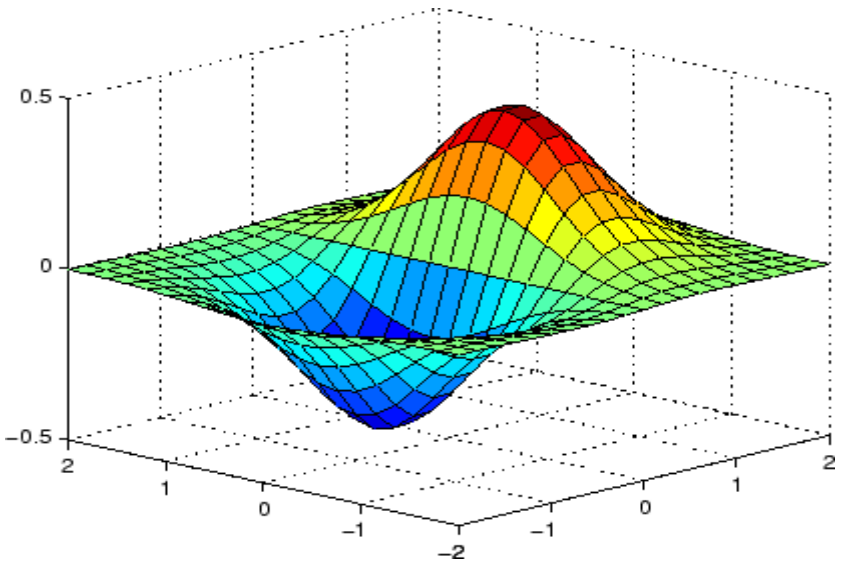
X = gridsamp (range, q)

όπου:

range είναι τα κάτω και άνω όρια των τιμών του πλέγματος

q είναι το βήμα με το οποίο αυξάνονται οι τιμές του πλέγματος

X είναι το επιθυμητό αποτέλεσμα, δηλαδή ο πίνακας με τις τιμές του ορθογωνίου πλέγματος



Σχήμα 3.3 Παράδειγμα απεικόνισης σε 3-D μέσω rectangular grid

Παρακάτω παρατίθεται ένα παράδειγμα της εντολής αυτής:

$X = gridsamp\ ([\ 1\ \ 1; 3\ \ 3], 3)$

Το οποίο δίνει ως αποτέλεσμα

$X =$

$$\begin{matrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{matrix}$$

$$\begin{matrix} 2 & 1 \\ 2 & 2 \\ 2 & 3 \end{matrix}$$

$$\begin{matrix} 3 & 1 \\ 3 & 2 \\ 3 & 3 \end{matrix}$$

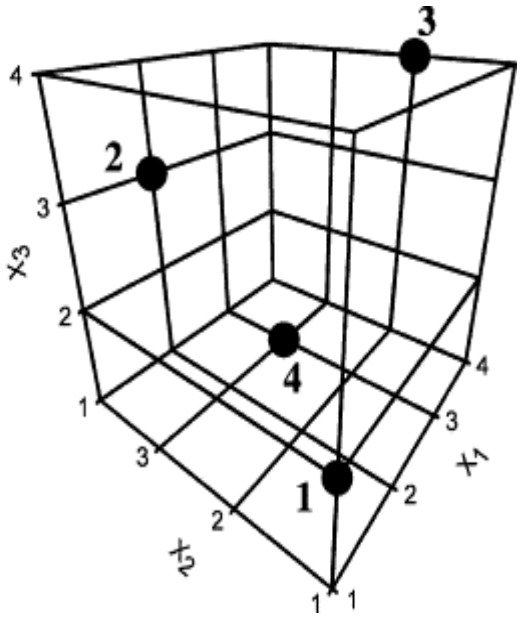
Latin Hypercube Sampling

Ο αλγόριθμος αυτός αποτελεί μια στρατηγική που χρησιμοποιείται συχνά για να παράγει ένα τυχαίο δείγμα από σημεία και που εξασφαλίζει ότι όλα τα σημεία του διανύσματος των μεταβλητών εισόδου αντιπροσωπεύονται. Ας εξετάσουμε την περίπτωση όπου θέλουμε να φτιάξουμε ένα δείγμα από m σημεία σε διάνυσμα n-διαστάσεων. Τα βήματα που ακολουθούμε είναι:

1) Χωρίζουμε το διάστημα κάθε μεταβλητής σε m μη επικαλυπτόμενα διαστήματα που έχουν ίση πιθανότητα (θεωρούμε ομοιόμορφη κατανομή ώστε τα διαστήματα να έχουν το ίδιο μέγεθος)
2) Διαλέγουμε τυχαία ένα σημείο από κάθε διάστημα της κάθε μεταβλητής
3) Φτιάχνουμε τυχαία ζεύγη με τα σημεία αυτά

Τα σημεία που εξάγει ο συγκεκριμένος αλγόριθμος βρίσκονται στο διάστημα (0,1) οπότε χρειάζεται κάποια προσαρμογή αν θέλουμε τα αποτελέσματα σε ένα άλλο διάστημα.

Point	x_1	x_2	x_3
1	1	1	2
2	2	4	3
3	4	2	4
4	3	3	1



Σχήμα 3.4 Παράδειγμα απεικόνισης σε μέσω latin hypercube

4

SUPPORT VECTOR MACHINES

4.1 ΕΙΣΑΓΩΓΗ

Η μέθοδος των Support Vector Machines (SVM) είναι μια σχετικά νέα μέθοδος εκμάθησης που χρησιμοποιείται για διμερή ταξινόμηση. Πρόκειται για μοντέλα εκμάθησης τα οποία, μέσω των κατάλληλων αλγορίθμων, μπορούν να αναλύουν δεδομένα και να αναγνωρίζουν τα πρότυπα που χρησιμοποιούνται για την ταξινόμηση και την ανάλυση παλινδρόμησης. Ένα βασικό μοντέλο SVM παίρνει ένα σύνολο δεδομένων εισόδου και προβλέπει, για κάθε δεδομένη είσοδο, ποια από τις δυο πιθανές κατηγορίες-κλάσεις αποτελεί την έξοδο. Έχοντας ένα δείγμα από παραδείγματα εκπαίδευσης, καθένα από τα οποία ανήκει σε μια από τις δυο κατηγορίες, ο αλγόριθμος εκπαίδευσης SVM παράγει ένα μοντέλο που εκχωρεί τα νέα παραδείγματα στη μία ή την άλλη κατηγορία. Η βασική, λοιπόν, ιδέα είναι να βρεθεί το υπερεπίπεδο εκείνο το οποίο θα διαχωρίζει τα δεδομένα του προβλήματος τέλεια στις δυο κλάσεις. Ωστόσο, δεδομένου ότι τα στοιχεία κάποιου παραδείγματος δεν είναι απαραίτητα γραμμικά διαχωρίσιμα, το μοντέλο SVM χρησιμοποιεί έναν πίνακα (kernel) που εισάγει τα δεδομένα σε έναν πολυδιάστατο χώρο και βοηθάει στο να διαχωρίζονται. Τυπικά, κάτι τέτοιο θα μπορούσε να δημιουργήσει υπολογιστικά προβλήματα ή υπερεκπάδευση (overfitting). Στη μέθοδο SVM, όμως, ο χώρος υψηλότερων διαστάσεων που χρησιμοποιείται δε χρειάζεται να αντιμετωπίζεται άμεσα, καθώς, όπως θα δούμε και παρακάτω, μόνο ο τύπος για το εσωτερικό γινόμενο σε αυτόν τον χώρο είναι απαραίτητος. Έτσι, εξαλείφονται τα προαναφερόμενα προβλήματα. Επιπλέον, η διάσταση VC του SVM (ένα μέτρο της πιθανότητας ενός συστήματος να αποδίδει καλά στα νεα δεδομένα, που θα αναλυθεί και παρακάτω) μπορεί να υπολογιστεί ρητά σε αντίθεση με άλλες μεθόδους εκμάθησης όπως τα νευρωνικά δίκτυα, για τα οποία δεν υπάρχει το αντίστοιχο μέτρο. Συνολικά, λοιπόν, η μηχανή εκμάθησης SVM είναι θεωρητικά

βάσιμη και πρακτικά επιτυχημένη, ακόμη και σε θέματα παλινδρόμησης, όπου το σύστημα εκπαιδεύεται ώστε να εξάγει μια αριθμητική τιμή αντί για ένα σκέτο ναι/όχι ανάλογα με το αν ανήκει το στοιχείο σε μια κλάση ή όχι.

Η μέθοδος που μελετάται στην ενότητα αυτή έχει εφαρμοστεί επισήμως για πρώτη φορά το 1995 από τον Vladimir Vapnik και τους συνεργάτες του, ενώ η παλαιότερη αναφορά στο αντικείμενο εντοπίζεται το 1979. Τα προβλήματα που παρατηρούνταν σε προηγούμενες υπολογιστικές μεθόδους και που τελικά οδήγησαν στην αρχική ανάπτυξη των SVM αφορούσαν σε διάφορους τομείς, όπως στη δυσκολία υπολογισμού της διακύμανσης των συντελεστών bias, στην ικανότητα ελέγχου και στην υπερπροσαρμογή (overfitting). Στις περισσότερες περιπτώσεις, η συγκεκριμένη μέθοδος δείχνει καλύτερη συμπεριφορά γενίκευσης από άλλες εναλλακτικές υπολογιστικές μεθόδους τόσο σε θέματα αναγνώρισης προτύπων όσο και στην εκτίμηση της παλινδρόμησης. Παρόλα αυτά, υπάρχουν φορές που τα μοντέλα SVM απαιτούν αρκετό χρόνο σε κάποια στάδια, πρόβλημα το οποίο μελετάται και από τους Burges, Osuna και Girosi.

Σε γενικά πλαίσια, για μια συγκεκριμένη εργασία εκμάθησης με δεδομένη πεπερασμένη ποσότητα από δεδομένα εκπαίδευσης, η καλύτερη απόδοση γενίκευσης μπορεί να επιτευχθεί αν δοθεί η σωστή ισορροπία ανάμεσα στην ακρίβεια που παρουσιάζεται στο συγκεκριμένο σύνολο εκπαίδευσης και στη δυναμικότητα της μεθόδου, δηλαδή στην ικανότητα της μεθόδου να μάθει (να προβλέψει) οποιοδήποτε δείγμα εκπαίδευσης χωρίς σφάλμα. Μηχανές με πολύ μεγάλη ή, αντίθετα, πολύ μικρή δυναμικότητα δεν μπορούν να γενικεύσουν καλά. Η εξερεύνηση και η μορφοποίηση αυτών των εννοιών αποτελούν τη βασική αρχή για τη μέθοδο της συγκεκριμένης ενότητας και έχουν οδηγήσει στην ανάπτυξη της θεωρίας γύρω από τη στατιστική εκμάθηση.

4.2 ΛΕΙΤΟΥΡΓΙΑ ΤΩΝ SVM

Υποθέτουμε ότι έχουμε l παραδείγματα για εκπαίδευση (παρατηρήσεις). Κάθε παρατήρηση αποτελείται από ένα ζεύγος: ένα διάνυσμα $x_i \in R^n, i = 1, \ldots, l$ και την αντίστοιχη πραγματική τιμή y_i που μας έχει δοθεί από σίγουρη πηγή. Στη μηχανή εκμάθησης η τιμή για το y_i θα είναι 1 αν το x_i ανήκει στη συγκεκριμένη κλάση και -1 αν όχι. Τώρα, ας υποθέσουμε ότι υπάρχει κάποια άγνωστη κατανομή πιθανότητας $P(x, y)$ από την οποία αντλούνται τα δεδομένα αυτά, δηλαδή τα δεδομένα θεωρούνται “iid” (independent and identically distributed), ανεξάρτητα σχεδιασμένα και ταυτόσημα κατανεμημένα. Κάνουμε την υπόθεση αυτή, η οποία είναι πιο γενική από το να αντιστοιχούσαμε ένα σταθερό y σε κάθε x, αφού επιτρέπει μια κατανομή του y για κάθε δεδομένο x.

Ας υποθέσουμε ακόμη ότι έχουμε μια μηχανή της οποίας η αποστολή είναι να μάθει την αντιστοιχία $x_i \rightarrow y_i$. Η μηχανή ορίζεται ουσιαστικά από ένα σύνολο πιθανών αντιστοιχιών $x \rightarrow f(x,a)$, όπου οι συναρτήσεις $f(x,a)$ χαρακτηρίζονται από τις ρυθμιζόμενες παραμέτρους α. Η μέθοδος θεωρείται ντετερμινιστική, δηλαδή, για μια δεδομένη είσοδο x και μια συγκεκριμένη τιμή της παραμέτρου α, δίνει πάντα την ίδια έξοδο $f(x,a)$. Μια συγκεκριμένη επιλογή της παραμέτρου α δημιουργεί αυτό που ονομάζουμε "εκπαιδευόμενη μηχανή ".

Το αναμενόμενο σφάλμα για μια εκπαιδευόμενη μηχανή είναι:

$$R(a) = \int \frac{1}{2}|y - f(x,a)|dP(x,y) \tag{4.1}$$

Να σημειωθεί ότι, όταν η πυκνότητα $p(x,y)$ υπάρχει, η ποσότητα $dP(x,y)$ μπορεί να γραφτεί $p(x,y)dxdy$. Αυτός είναι ο καλύτερος τρόπος για να εκφραστεί το πραγματικό μέσο σφάλμα, αλλά αν δεν έχουμε μια εκτίμηση του $P(x,y)$, δεν είναι χρήσιμο.

Η ποσότητα $R(a)$ ονομάζεται αναμενόμενο ρίσκο ή απλά ρίσκο, ενώ το εμπειρικό ρίσκο $R_{emp}(\alpha)$ ορίζεται ως το μετρημένο ποσοστό μέσου σφάλματος για το δείγμα εκπαίδευσης (για ένα δεδομένο, πεπερασμένο αριθμό παρατηρήσεων):

$$R_{emp}(\alpha) = \frac{1}{2l}\sum_{i=1}^{l}|y_i - f(x_i,a)| \tag{4.2}$$

Να σημειωθεί ότι δεν εμφανίζεται κατανομή πιθανότητας εδώ. $R_{emp}(\alpha)$ είναι ένας σταθερός αριθμός για μια συγκεκριμένη επιλογή της παραμέτρου α και για ένα δεδομένο δείγμα εκπαίδευσης $\{x_i, y_i\}$. Η ποσότητα $\frac{1}{2}|y_i - f(x_i,a)|$ καλείται απώλεια και εδώ μπορεί να πάρει μόνο τις τιμές 0 και 1. Διαλέγουμε μια τιμή n για την απώλεια τέτοια ώστε $0 \leq n \leq 1$, οπότε έχουμε το παρακάτω όριο (Vapnik,1995):

$$R(a) \leq R_{emp}(\alpha) + \sqrt{\left(\frac{h\left(\log\left(\frac{2l}{h}\right)+1-\log\left(\frac{n}{4}\right)\right)}{l}\right)} \tag{4.3}$$

όπου h είναι ένας μη αρνητικός ακέραιος αριθμός που ονομάζεται διάσταση Vapnik Chervonenkis (VC) και είναι ένα μέτρο της ικανότητας της μηχανής, όπως αναφέρθηκε και παραπάνω.

Τρία είναι τα σημεία-κλειδιά για το παραπάνω όριο. Πρώτον, το ρίσκο είναι ανεξάρτητο του $P(x,y)$ και απλώς υποθέτουμε ότι τόσο τα δεδομένα που χρησιμοποιούνται για εκπαίδευση όσο και εκείνα για τη δοκιμή έχουν σχεδιασθεί

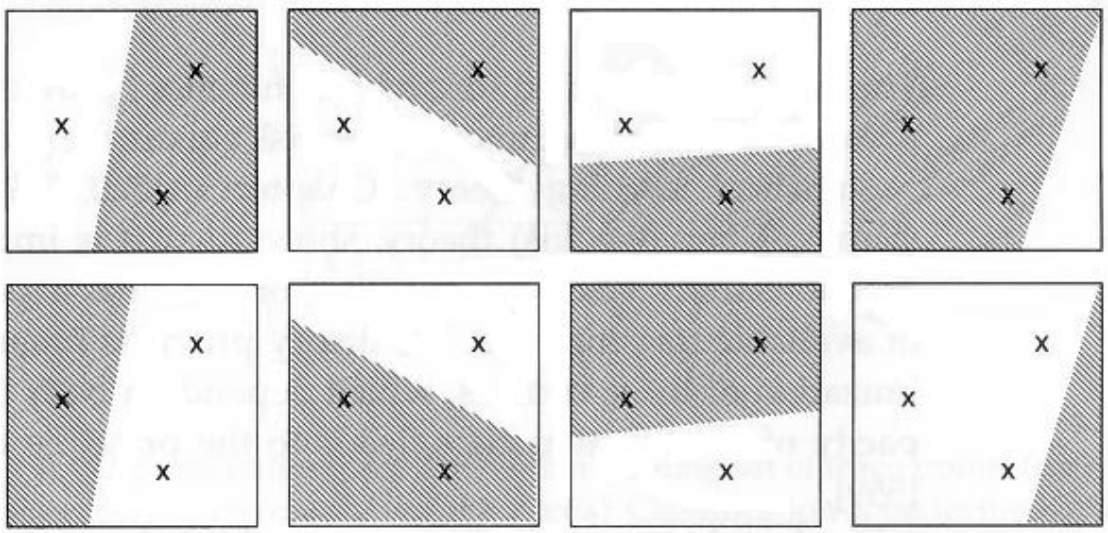
ανεξάρτητα σύμφωνα με κάποια κατανομή πιθανότητας $P(x, y)$. Δεύτερον, είναι πιθανό να μην μπορεί να υπολογιστεί η αριστερή πλευρά του ορίου. Τρίτον, αν γνωρίζουμε το στοιχείο h, μπορούμε εύκολα να υπολογίσουμε τη δεξιά πλευρά. Έτσι, έχοντας πολλές διαφορετικές μηχανές εκμάθησης και επιλέγοντας ένα συγκεκριμένο, αρκετά μικρό n, επιλέγουμε τη μηχανή εκείνη που δίνει το μικρότερο άνω όριο στο πραγματικό ρίσκο. Αυτή είναι η βασική μέθοδος για την επιλογή του κατάλληλου μηχανήματος (SVM) και αποτελεί την ουσιαστική ιδέα για την ελαχιστοποίηση του ρίσκου-κινδύνου κατασκευής. Έχοντας μια σταθερή οικογένεια μηχανών εκμάθησης από τις οποίες κάνουμε την επιλογή και ένα αρκετά μικρό άνω όριο για τουλάχιστον μία από τις μηχανές, τότε η συγκεκριμένη μηχανή αποτελεί σίγουρα τη μόνη λύση. Αν το όριο δεν είναι σφιχτό για καμία από τις μηχανές, τότε η δεξιά πλευρά μπορεί να δώσει κάποιες σημαντικές πληροφορίες σχετικά με το ποια μηχανή εκμάθησης ελαχιστοποιεί το πραγματικό ρίσκο.

4.3 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΓΥΡΩ ΑΠΟ ΤΑ SVM

4.3.1 Διάσταση VC και ικανότητα των μηχανών εκμάθησης

Η διάσταση VC για ένα δείγμα λειτουργιών $\{f(a)\}$ ορίζεται ουσιαστικά ως ο μέγιστος αριθμός των σημείων εκπαίδευσης που μπορούν να διαχωριστούν με όλους τους δυνατούς τρόπους μέσω των λειτουργιών $\{f(a)\}$. Για τα υπερεπίπεδα που ανήκουν στο R^n, η διάσταση VC είναι $n + 1$.

Ένα απλό παράδειγμα: θέλουμε να διαχωρίσουμε 3 σημεία τα οποία ανήκουν στο R^2 σε δύο κατηγορίες. Υπάρχουν 2^3=8 τρόποι για να γίνει αυτό. Για τα σημεία αυτά, οι 8 δυνατοί τρόποι μπορούν να πραγματοποιηθούν χρησιμοποιώντας διαχωριστικά υπερεπίπεδα. Με άλλα λόγια η λειτουργία της ταξινόμησης μπορεί να διαχωρίσει τα 3 αυτά σημεία κάτι το οποίο δε θα μπορούσε να επιτευχθεί για 4 σημεία, όπως και αν τα τοποθετούσαμε.



Εικόνα 4.1 Τρία σημεία στο R^2, διαχωρισμένα μέσω υπερεπιπέδων

Απλοποίηση του ορίου:

Σφάλμα δοκιμής $\leq$ Σφάλμα εκπαίδευσης + Πολυπλοκότητα του συνόλου των μοντέλων

Να σημειωθεί εδώ ότι πολλά όρια αυτής της μορφής έχουν αποδειχτεί (για διαφορετικά μέτρα ικανότητας). Φυσικά η επιλογή του μέτρου ικανότητας δεν είναι εύκολη. Αν επιλέξουμε υψηλή ικανότητα για τη λειτουργία $\{f(a)\}$, τότε το σφάλμα εκπαίδευσης είναι μικρό αλλά υπάρχει κίνδυνος υπερπροσαρμογής (overfitting). Αντίθετα, αν θεωρήσουμε ένα πολύ απλό σύνολο μοντέλων, τότε έχουμε χαμηλή πολυπλοκότητα αλλά δε θα επιτύχουμε μικρό σφάλμα εκπαίδευσης.

4.3.2 Δυναμικότητα των υπερεπιπέδων

- *Γραμμικά SVM*

Θεωρούμε υπερεπίπεδα στο R^n, τα οποία ως παραμέτρους έχουν το διάνυσμα w και μια σταθερά b και ικανοποιούν τη σχέση:

$$w \cdot x + b = 0 \tag{4.4}$$

Έχουμε, λοιπόν, ένα υπερεπίπεδο (w, b) το οποίο χωρίζει τα δεδομένα μας μέσω της συνάρτησης:

$$f(x) = sign(w \cdot x + b) \tag{4.5}$$

η οποία ταξινομεί σωστά τα δεδομένα εκπαίδευσης και άλλα δεδομένα που θα δοθούν ως δοκιμή, τα οποία δε θα τα έχει ξανασυναντήσει. Όμως, ένα συγκεκριμένο υπερεπίπεδο που αποτυπώνεται από τις παραμέτρους (w, b) εκφράζεται ισοδύναμα από όλα τα ζεύγη $\{\lambda w, \lambda b\}$ $\gamma\iota\alpha\ \lambda \in R^{+}$. Οπότε, ορίζουμε ως κανονικό υπερεπίπεδο εκείνο το οποίο χωρίζει τα στοιχεία από το υπερεπίπεδο κατά μια απόσταση τουλάχιστον ίση με 1. Οπότε, έχουμε τα υπερεπίπεδα εκείνα που ικανοποιούν τις σχέσεις:

$$x_i \cdot w + b \geq 1\ \acute{o}\tau\alpha\nu\ y_i = 1 \tag{4.6}$$

και

$$x_i \cdot w + b \leq -1\ \acute{o}\tau\alpha\nu\ y_i = -1 \tag{4.7}$$

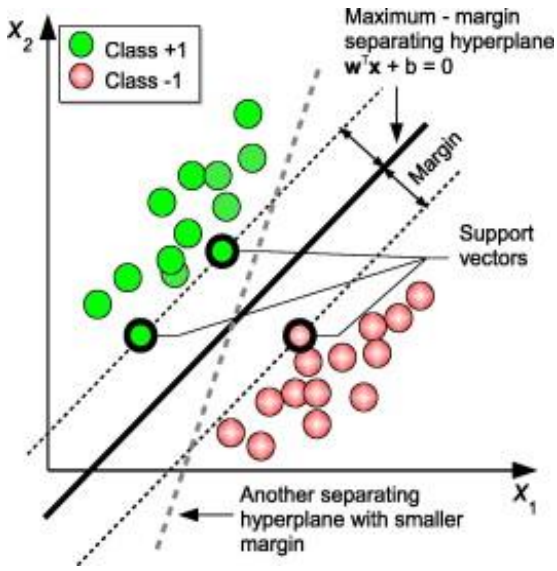
ή αλλιώς:

$$y_i(x_i \cdot w + b) \geq 1\ \ \forall i \tag{4.8}$$

Όλα αυτά τα υπερεπίπεδα έχουν “λειτουργική απόσταση” ≥ 1, όχι γεωμετρική ή Ευκλείδεια απόσταση(γνωστή και ως περιθώριο). Για ένα δεδομένο υπερεπίπεδο (w, b), όλα τα ζεύγη $\{\lambda w, \lambda b\}$ ορίζουν το ίδιο υπερεπίπεδο, αλλά το καθένα έχει διαφορετική λειτουργική απόσταση από ένα δεδομένο σημείο. Για να εξάγουμε τη γεωμετρική απόσταση μεταξύ του υπερεπιπέδου και του σημείου, πρέπει να κανονικοποιήσουμε μέσω του w. Η γεωμετρική, λοιπόν, απόσταση είναι:

$$d\big((w, b), x_i\big) = \frac{y_i(x_i \cdot w + b)}{\|w\|} \geq \frac{1}{\|w\|} \tag{4.9}$$

Το υπερεπίπεδο που θεωρείται το βέλτιστο είναι εκείνο που μεγιστοποιεί τη γεωμετρική απόσταση από το πλησιέστερο σημείο (margin), όπως απεικονίζεται και στην εικόνα 4.2 παρακάτω.



Εικόνα 4.2 Βέλτιστος διαχωρισμός στοιχείων μέσω υπερεπιπέδου SVM

Από την εξίσωση (4.9) βλέπουμε ότι για να καταλήξουμε στο κατάλληλο υπερεπίπεδο πρέπει να ελαχιστοποιήσουμε το $\|w\|$. Η κύρια μέθοδος για να το επιτύχουμε αυτό είναι μέσω των πολλαπλασιαστών Lagrange, όπως αναφέρουν και οι Vapnik, Burges σε σχετικές έρευνες, για τους οποίους ισχύει:

$$L_P \equiv \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} a_i y_i (x_i \cdot w + b) + \sum_{i=1}^{l} a_i$$

Έτσι, το πρόβλημα μετατρέπεται τελικά στο εξής:

Ελαχιστοποίηση του: $W(a) = -\sum_{i=1}^{l} a_i + \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j a_i a_j (x_i \cdot x_j)$ (4.10α)

ή αλλιώς μεγιστοποίηση του: $L_D = \sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j a_i a_j (x_i \cdot x_j)$ (4.10β)

Έγκειται στο: $\sum_{i=1}^{l} y_i a_i = 0, \quad 0 \le a_i \le C \ \ \forall i$ (4.11)

όπου a είναι το διάνυσμα αποτελούμενο από l μη αρνητικούς πολλαπλασιαστές Lagrange και C είναι μια σταθερά, που θα εξηγηθεί παρακάτω. Μπορούμε να ορίσουμε τον πίνακα $(H)_{ij} = y_i y_j (x_i x_j)$ και να εισάγουμε πιο συμπαγείς συμβολισμούς:

Ελαχιστοποίηση: $W(a) = -a^T 1 + \frac{1}{2} a^T H a$ (4.12)

Έγκειται στο: $a^T y = 0, \ 0 \le a \le C1$ (4.13)

Αυτό το πρόβλημα ελαχιστοποίησης είναι γνωστό και ως *τετραγωνικός προγραμματισμός - Quadratic Programming* (QP) και πολλές τεχνικές έχουν αναπτυχθεί για την επίλυσή του. Από την παραγώγιση των παραπάνω εξισώσεων, προκύπτει ότι το βέλτιστο υπερεπίπεδο μπορεί να εκφραστεί ως:

$$w = \sum_i a_i y_i x_i \tag{4.14}$$

από όπου φαίνεται πως το διάνυσμα w είναι απλώς ένας γραμμικός συνδυασμός των παραδειγμάτων εκπαίδευσης. Επίσης, αποδεικνύεται ότι:

$$a_i(y_i(w \cdot x_i + b) - 1) = 0 \ \ \forall i \tag{4.15}$$

το οποίο δείχνει πως όταν η λειτουργική απόσταση ενός δεδομένου είναι μεγαλύτερη από 1, δηλαδή όταν $y_i(w \cdot x_i + b) > 1$, τότε $a_i = 0$. Επομένως, μόνο τα πλησιέστερα σημεία επηρεάζουν το w. Τα δεδομένα για τα οποία $a_i > 0$, ονομάζονται διανύσματα υποστήριξης *(support vectors)*. Και είναι τα μόνα απαραίτητα για τον καθορισμό του βέλτιστου υπερεπιπέδου.

Υποθέτοντας πως έχουμε το βέλτιστο α, από το οποίο εξάγουμε το w, πρέπει ακόμη να καθορίσουμε τη σταθερά b προκειμένου να συγκεκριμενοποιήσουμε το υπερεπίπεδο. Για να γίνει αυτό, παίρνουμε οποιοδήποτε "θετικό" και "αρνητικό" support vector , x^+ και x^-, για τα οποία ισχύει:

$$w \cdot x^+ + b = +1$$

$$w \cdot x^- + b = -1$$

Λύνοντας το σύστημα των δυο αυτών εξισώσεων έχουμε:

$$b = -\frac{1}{2}(w \cdot x^+ + \ w \cdot x^-) \tag{4.16}$$

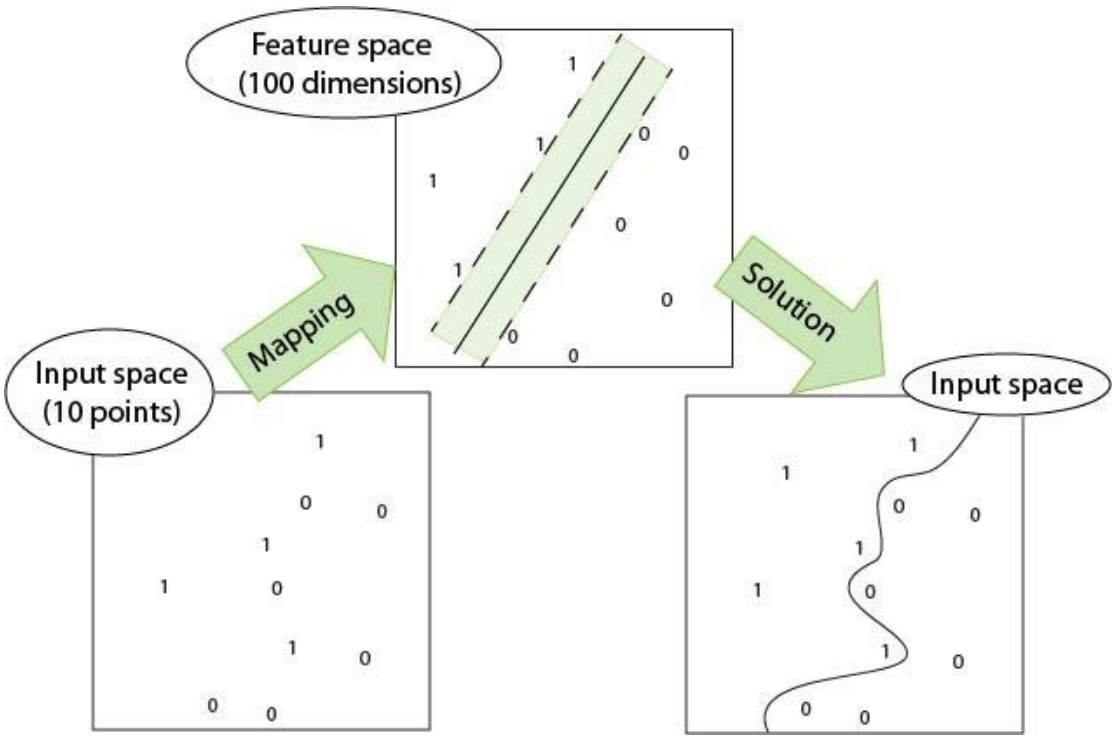
Όσον αφορά στον περιορισμό $a \leq C$, σημειώνεται ότι, αν θέσουμε $C = \infty$, τότε το βέλτιστο επίπεδό μας θα διαχωρίζει πλήρως τα δεδομένα οπότε όλα θα ανήκουν στην πραγματική τους κλάση. Αν θεωρήσουμε ένα πεπερασμένο C, τότε το υπερεπίπεδο είναι πιο ευέλικτο και προσπαθεί να ελαχιστοποιήσει το περιθώριο για κάθε παράδειγμα επιτρέποντας σε κάποια από τα δεδομένα να είναι ταξινομημένα με λάθος τρόπο. Πεπερασμένες τιμές για το C είναι χρήσιμες όταν τα δεδομένα δε διαχωρίζονται εύκολα.

- *Μη γραμμικά SVM*

Οι Boser, Gyron, Vapnik έδειξαν ότι μια αρκετά παλιά μέθοδος (Alizerman, 1964) μπορεί να χρησιμοποιηθεί για τη συγκεκριμένη κατηγορία δεδομένων. Να σημειωθεί εδώ, ότι ο μόνος τρόπος που εμφανίζονται τα δεδομένα στο πρόβλημα, εξισώσεις (4.10), (4.11), είναι με τη μορφή εσωτερικού γινομένου $x_i \cdot x_j$. Το πρώτο βήμα για την επίλυση του προβλήματός μας, είναι να απεικονίσουμε τα δεδομένα σε έναν άλλο Ευκλείδειο χώρο H μεγαλύτερων διαστάσεων (πιθανώς άπειρων) χρησιμοποιώντας μια χαρτογράφηση που ονομάζουμε Φ :

$$\Phi: R^d \rightarrow H \tag{4.17}$$

Οπότε, ο αλγόριθμος εκπαίδευσης εξαρτάται μόνο από τα στοιχεία του εσωτερικού γινομένου στον χώρο H, δηλαδή από τις συναρτήσεις τις μορφής $\Phi(x_i) \cdot \Phi(x_J)$. Αν υπήρχε κάποια συνάρτηση kernel K τέτοια ώστε $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, θα χρησιμοποιούσαμε απλώς τη συνάρτηση K στον αλγόριθμο και δε θα χρειαζόταν καν να γνωρίζουμε επακριβώς το Φ. Τώρα, στο χώρο που έχουμε μεταφέρει τα στοιχεία προσπαθούμε να βρούμε το υπερεπίπεδο που θα τα διαχωρίζει. Η πορεία που ακολουθούμε αλλά και μια μορφή της τελικής απεικόνισης παρουσιάζονται στις εικόνες 4.3,4.4 παρακάτω:



Εικόνα 4.3 Πορεία διαχωρισμού στοιχείων μέσω μη γραμμικών SVM

Ας υποθέσουμε τώρα, ότι τα δεδομένα του προβλήματος μας είναι διανύσματα στο χώρο R^2 και επιλέγουμε για $K(x_i, x_j) = (x_i \cdot x_j)^2$. Είναι εύκολο να βρούμε ένα χώρο H στον οποίο θα απεικονίσουμε τα δεδομένα μέσω της Φ, έτσι ώστε:

$$(x \cdot y)^2 = \Phi(x) \cdot \Phi(y)$$

Επιλέγουμε $H = R^3$ και

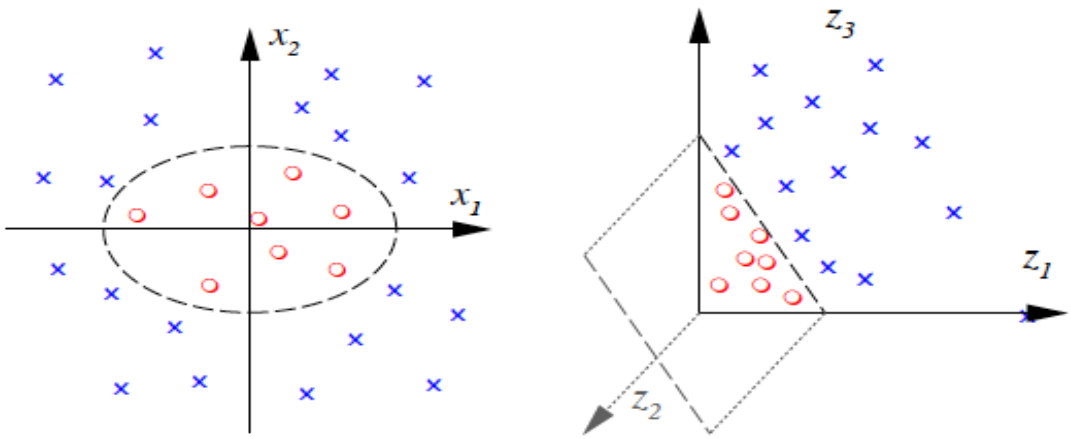
$$\Phi(x) = \begin{pmatrix} {x_1}^2 \\ \sqrt{2}x_1x_2 \\ {x_2}^2 \end{pmatrix}$$

Τόσο η συνάρτηση Φ, όσο και ο χώρος H δεν είναι μοναδικά για μια δεδομένη συνάρτηση K. Έτσι, για το παραπάνω πρόβλημα θα μπορούσε επίσης να επιλεχθεί ως H ο χώρος R^3 και

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} {x_1}^2 - {x_2}^2 \\ 2x_1x_2 \\ {x_1}^2 + {x_2}^2 \end{pmatrix}$$

ή ως H ο χώρος R^4 και

$$\Phi(x) = \begin{pmatrix} {x_1}^2 \\ x_1x_2 \\ x_1x_2 \\ {x_2}^2 \end{pmatrix}$$



Εικόνα 4.4 Διαχωρισμός στοιχείων μέσω μεταφοράς τους από R^2 σε R^3

4.3.3 ΣΥΝΑΡΤΗΣΗ KERNEL

Ένα πρόβλημα που συναντάται στον τετραγωνικό προγραμματισμό (QP) είναι όταν η διάσταση του $\Phi(x)$ είναι αρκετά μεγάλη οπότε υπάρχει δυσκολία στην εύρεση λύσης. Το θεώρημα Representer (Kimeldorf & Wahba, 1971) έδειξε ότι, για συγκεκριμένες περιπτώσεις SVM, ισχύει:

$$w = \sum_{i=1}^{m} a_i \Phi(x_i)$$
(4.18)

Έτσι, αντί να ψάχνουμε το βέλτιστο w μπορούμε να ψάχνουμε το βέλτιστο α.
Η εξίσωση που έχουμε να λύσουμε τώρα είναι:

$$f(x) = \sum_{i=1}^{m} a_i \Phi(x_i) \Phi(x_j) + b$$
(4.19)

και όπως αναφέρθηκε και παραπάνω η ποσότητα $K(x_i, x_j) = \Phi(x_i)\Phi(x_j)$ ονομάζεται συνάρτηση kernel. Οπότε, η εξίσωση γράφεται ως εξής:

$$f(x) = \sum_{i=1}^{m} a_i \Phi(x_i) \Phi(x_j) + b = \sum_{i=1}^{m} a_i \; K(x_i, x_j) + b$$
(4.20)

Οι βασικοί τύποι kernel που χρησιμοποιούνται καθώς και ο τρόπος που υπολογίζουν την ποσότητα $\Phi(x_i)\Phi(x_j)$ περιγράφονται στον ακόλουθο πίνακα (Πίνακας 4.1).

Πίνακας 4.1 Βασικοί τύποι Kernel

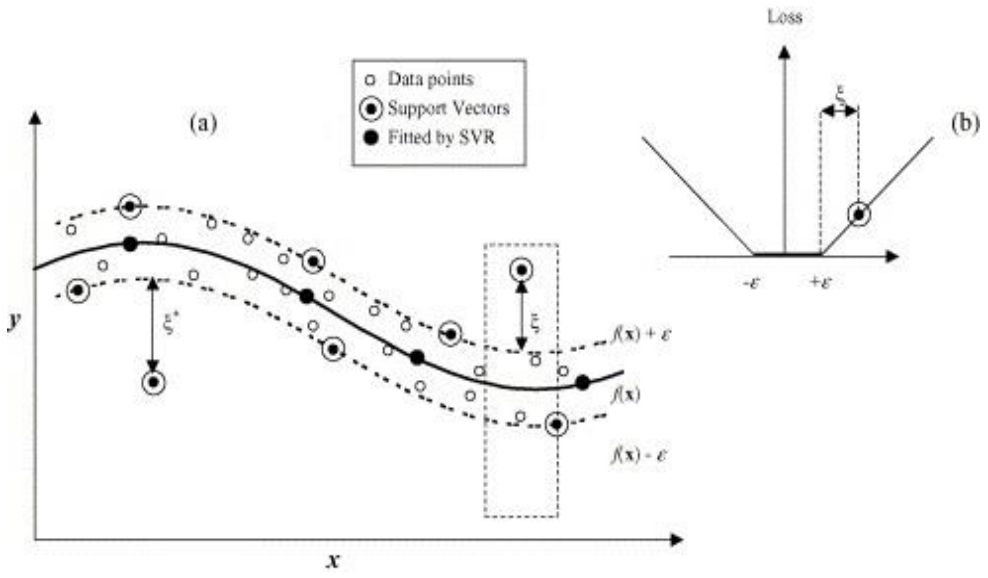
Τύπος Kernel	Περιγραφή
Γραμμικός	$K(x_1, x_2) = x_1 \cdot x_2$
Πολυώνυμο	$K(x_1, x_2) = (x_1 \cdot x_2 + 1)^p$, όπου p είναι ο βαθμός του πολυωνύμου
Gausian RBF	$K(x_1, x_2) = \exp(-\frac{\lVert x_1 - x_2 \rVert^2}{2\sigma^2})$

4.4 SVM ΩΣ ΜΕΘΟΔΟΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

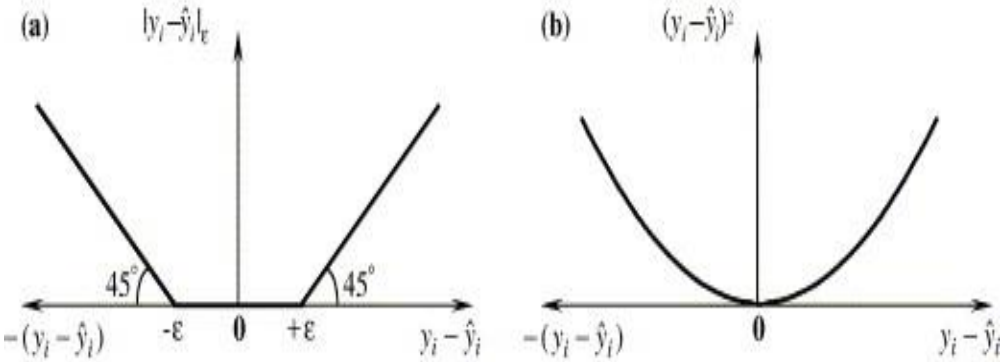
4.4.1 Περιγραφή του προβλήματος

Η μέθοδος SVM μπορεί να εφαρμοστεί τόσο σε προβλήματα ταξινόμησης, με τον τρόπο που αναλύσαμε παραπάνω, όσο και σε περιπτώσεις παλινδρόμησης. Η διαδικασία που ακολουθείται σε αυτήν την περίπτωση έχει αρκετές ομοιότητες με τη διαδικασία που ακολουθείται και για την ταξινόμηση, καθώς τα κύρια στοιχεία που χαρακτηρίζουν τον αλγόριθμο εύρεσης μέγιστου περιθωρίου (margin) παραμένουν τα ίδια. Τα δεδομένα και εδώ μεταφέρονται σε έναν χώρο μεγαλύτερων διαστάσεων στον οποίο προσπαθούμε να βρούμε την κατάλληλη συνάρτηση παλινδρόμησης. Η ικανότητα του συστήματος επηρεάζεται από παραμέτρους οι οποίες είναι ανεξάρτητες από τη διάσταση του χώρου.

Το πρόβλημα της παλινδρόμησης μπορεί να περιγραφεί ως εξής:
Έχοντας ένα δείγμα εκπαίδευσης $\{x_i, t_i\}\, με\ i = 1, \ldots, n,$ όπου x_i είναι τα διανύσματα εισόδου και t_i οι αντίστοιχες τιμές-στόχοι, ο στόχος είναι να βρεθεί η κατάλληλη συνάρτηση $g(x)$ που θα προσεγγίζει με τον καλύτερο τρόπο τη σχέση μεταξύ των δεδομένων και θα μπορεί να χρησιμοποιηθεί αργότερα για να εξάγει κάποιο αποτέλεσμα από μια νέα είσοδο. Οποιοσδήποτε αλγόριθμος παλινδρόμησης χρησιμοποιεί μια συνάρτηση απώλειας (loss function), η οποία περιγράφει τον τρόπο με τον οποίο η εκτιμώμενη συνάρτηση παρεκκλίνει από την πραγματική. Πολλές μορφές τέτοιων συναρτήσεων έχουν βρεθεί, αλλά οι κυριότερες είναι η *ε-insensitive* και η *quadratic loss function.* Η πρώτη έχει την ιδιαιτερότητα ότι αγνοεί το σφάλμα που εντοπίζεται εντός της ορισμένης απόστασης από την πραγματική τιμή (περιοχή από –ε μέχρι +ε στην εικόνα 4.5). Η δεύτερη (εικόνα 4.6) δεν αγνοεί κανένα σφάλμα και έχει το πλεονέκτημα της μικρότερης απαιτούμενης μνήμης. Οπότε, πολλές φορές προτιμάται για την κατασκευή ενός αλγορίθμου παλινδρόμησης.



Εικόνα 4.5 ε-insensitive loss function



Εικόνα 4.6 ε-insensitive και quadratic loss function

4.4.2 Κλασσική μορφή του προβλήματος παλινδρόμησης

Θεωρούμε ότι η συνάρτηση παλινδρόμησης στο χώρο m-διαστάσεων, όπου έχουν μεταφερθεί τα δεδομένα, έχει γραμμική μορφή. Οπότε έχουμε:

$$g(x) = w \cdot x + b$$

όπως είχαμε και στην περίπτωση της ταξινόμησης

Σκοπός είναι να βρούμε τη βέλτιστη λύση, που επιτυγχάνεται αν ελαχιστοποιηθεί το w. Οπότε έχουμε:

Ελαχιστοποίηση του : $\frac{1}{2}\|w\|^2$

Έγκειται στο : $\begin{cases} t_i - (w \cdot y + b) \leq \varepsilon \\ (w \cdot y + b) - t_i \leq \varepsilon \end{cases}$

Έχει γίνει η παραδοχή ότι υπάρχει συνάρτηση που να ικανοποιεί την παραπάνω συνθήκη, δηλαδή το πρόβλημα βελτιστοποίησης έχει λύση. Μερικές φορές, ωστόσο, αυτό δεν μπορεί να συμβεί ή εμείς οι ίδιοι επιθυμούμε να επιτρέπονται κάποια σφάλματα. Για το λόγο αυτό, μπορούν να εισαχθούν κάποιες χαλαρές μεταβλητές ζ_i, ζ_i^* για να αντιμετωπίσουν τους ανέφικτους περιορισμούς του προβλήματος βελτιστοποίησης. Έτσι, οι παραπάνω εξισώσεις γίνονται:

Ελαχιστοποίηση του : $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\zeta_i + \zeta_i^*)$

Έγκειται στο: $\begin{cases} t_i - (w \cdot y + b) \leq \varepsilon + \zeta_i \\ (w \cdot y + b) - t_i \leq \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0 \end{cases}$

Το πρόβλημα μπορεί να λυθεί ευκολότερα στη δυαδική του μορφή (dual problem), όπου:

$$w = \sum_{i=1}^{n} (a_i - a_i{}^*) y_i$$

και η λύση είναι:

$$g(x) = \sum_{i=1}^{n} (a_i - a_i{}^*)(y_i, y) + b$$

Κατά μία έννοια, η πολυπλοκότητα της αναπαράστασης μιας συνάρτησης παλινδρόμησης είναι ανεξάρτητη της διάστασης του χώρου εισόδου Y και εξαρτάται μόνο από τον αριθμό των SVM. Επιπλέον, ο πλήρης αλγόριθμος μπορεί να περιγραφεί μέσω του εσωτερικού γινομένου μεταξύ των δεδομένων. Ακόμα και κατά την εκτίμηση της συνάρτησης $g(x)$, η τιμή του w δε χρειάζεται να υπολογίζεται επακριβώς.

Στην περίπτωση μη γραμμικής παλινδρόμησης χρησιμοποιούμε τις συναρτήσεις Kernel, με τρόπο ανάλογο όπως και στην ταξινόμηση.

5

ΑΡΙΘΜΗΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

5.1 ΕΙΣΑΓΩΓΗ

Στα πλαίσια της παρούσας εργασίας θα μελετηθούν τρία διαφορετικά υπολογιστικά μοντέλα μειωμένης κλίμακας για την επίλυση τριών προβλημάτων τα οποία σχετίζονται με κατασκευές πολιτικού μηχανικού. Πιο συγκεκριμένα, τα υπολογιστικά μοντέλα που θα χρησιμοποιήσουμε είναι τα Νευρωνικά Δίκτυα, το Kriging και τα Support Vector Machines και τα προβλήματα που καλούνται να επιλύσουν τα παραπάνω μοντέλα αφορούν σε γερανογέφυρα, σε οδόστρωμα και σε ρευστοποίηση εδάφους. Κάθε ένα από αυτά τα προβλήματα θα το προσεγγίσουμε και με τα τρία μοντέλα και θα καταλήξουμε σε συμπεράσματα για την απόδοση των μοντέλων καθώς και για τη βέλτιστη λύση σε κάθε περίπτωση. Σε κάθε πρόβλημα εκτός από τη θεωρητική ανάλυση, παρουσιάζονται και τα αποτελέσματα που προέκυψαν από την επίλυσή τους μέσω των παραπάνω μεθόδων. Πιο συγκεκριμένα, αναφέρονται αναλυτικά τα χαρακτηριστικά των αλγορίθμων που κατασκευάστηκαν για κάθε περίπτωση και επιπλέον έχει γίνει μια προσπάθεια να απεικονίσουμε και διαγραμματικά τη συμπεριφορά των μοντέλων έτσι ώστε να είναι εφικτή η σύγκριση των μεθόδων στην ίδια κλίμακα και η εξαγωγή των τελικών συμπερασμάτων. Το πρόγραμμα που χρησιμοποιήθηκε για τη λειτουργία των μοντέλων είναι η Matlab.

5.2 ΜΕΛΕΤΗ ΠΡΟΒΛΗΜΑΤΩΝ

5.2.1 ΓΕΡΑΝΟΓΕΦΥΡΑ

5.2.1.1 Παρουσίαση του προβλήματος

Οι γερανοί αποτελούν τον καλύτερο τρόπο για την παροχή βαριάς εγκατάστασης ανύψωσης και καλύπτουν σχεδόν ολόκληρη την περιοχή ενός κτιρίου. Οι γερανογέφυρες είναι το πιο σημαντικό σύστημα χειρισμού υλικών για τα βαρέα φορτηγά. Το πρωταρχικό καθήκον των εναέριων γερανών είναι να χειριστούν και να μεταφέρουν βαριά ωφέλιμα φορτία από τη μία θέση στην άλλη και μια από τις πιο χαρακτηριστικές λειτουργίες τους είναι σε ναυπηγεία. Τα εξαρτήματα που συνθέτουν μια γερανογέφυρα Τα χαρακτηριστικά του σχεδιασμού τους ποικίλλουν ανάλογα με τις κύριες λειτουργικές προδιαγραφές τους δηλαδή τον τύπο της κινήσεως του γερανού δομής, το βάρος και τον τύπο της θέσης του φορτίου του γερανού, τα γεωμετρικά χαρακτηριστικά και τις περιβαλλοντικές συνθήκες. Δεδομένου ότι οι διαδικασίες σχεδιασμού των γερανών είναι ιδιαίτερα συνδεδεμένες με αυτά τα στοιχεία, οι περισσότερες προσπάθειες και χρόνος που δαπανάται είναι για την ερμηνεία και την εφαρμογή των διαθέσιμων προτύπων σχεδιασμού. Υπάρχουν πολλές δημοσιευμένες μελέτες σχετικά με τη διαρθρωτική συνιστώσα και τις τάσεις, την ασφάλεια υπό στατική φόρτιση και τη δυναμική συμπεριφορά των γερανών. Βασικός τρόπος εύρεσης των μετατοπίσεων και των αναπτυσσόμενων τάσεων στις περισσότερες μελέτες είναι η ανάλυση αυτών των δομών με τη μέθοδο των πεπερασμένων στοιχείων. Στη συγκεκριμένη εργασία τα στοιχεία αυτά –τα οποία έχουν βρεθεί αρχικά χρησιμοποιώντας το λογισμικό FEM90- προσπαθούμε να τα προσεγγίσουμε με μεθόδους λιγότερο χρονοβόρες χωρίς όμως να μειώνεται η αποτελεσματικότητα και αποδοτικότητα. Ας δούμε όμως πρώτα ποιες είναι οι διατάξεις και τα συστήματα που συντελούν μια γερανογέφυρα:

- Σιδηροτροχιές κύλισης γερανογέφυρας

- Δυο σιδηροτροχιές ελάχιστης διατομής A45 (DIN 536-1:1991-09) ή ορθογωνικής διατομής τουλάχιστον 50 x 30 mm σε περίπτωση κατά την οποία η γερανογέφυρα κινείται σε μεταλλικές γερανοδοκούς
- Χαλύβδινα ελάσματα έδρασης σιδηροτροχιών πλάτους τουλάχιστον 150 mm και πάχους 10 mm
- Εξαρτήματα συγκράτησης σιδηροτροχιών επί των ελασμάτων έδρασης
- Τερματικοί προσκρουτήρες

- Κύριος φορέας γερανογέφυρας (απλής ή διπλής κατατομής)

Δοκοί από τυποποιημένες διατομές μορφοσιδήρου ή ολόσωμες δοκοί κιβωτοειδούς διατομής από χαλυβδοελάσματα ποιότητας S235JR / S275Jr / S355JR

- Διάταξη κύλισης γερανογέφυρας

- Δυο πλαγιομορφές (φορείς κεφαλής γερανογέφυρας) κατασκευασμένοι από τυποποιημένες διατομές χάλυβα ή από σύνθετη κιβωτοειδή διατομή από στραντζαρισμένο χαλυβδοέλασμα ποιότητας S235JR / S275Jr / S355JR με τοχούς κύλισης
- Δυο τροχοί κύλισης ανά πλαγιοφορέα με δυο ακραίες στεφάνες εκ των οποίων ο ένας είναι κινητήριος
- Ζεύγος ηλεκτρομειωτήρων δυο ταχυτήτων με ηλεκτομαγνητική πέδη
- Τερματικοί ελαστικοί προσκρουτήρες επί των πλαγιοφορέων
- Οριακοί διακόπτες διαδρομής κύλισης

- Διάταξη ανύψωσης

- Πλήρες ηλεκτροκίνητο βαρούλκο
- Συρματόσχοινο
- Διάταξη συστήματος ανάρτησης φορτίου (διάταξη τροχαλίας – απλού ανοικτού αγκίστρου με ασφάλεια)

- Διάταξη κύλισης βαρούλκου

- Φορείο έδρασης βαρούλκου: πλαίσιο από τυποποιημένες διατομές χάλυβα με τροχούς κύλισης και ηλεκτρομειωτήρα κίνησης
- Ηλεκτρομειωτήρας δυο ταχυτήτων με ηλεκτρομαγνητική πέδη
- Τροχοί κύλισης επί των σιδηροτροχιών του κυρίως φορέα της γερανογέφυρας με δυο ακραίες στεφάνες. Τροχοί κύλισης μονόπλευρης στεφάνης στην περίπτωση βαρούλκου αναρτημένου από το κάτω πέλμα του κυρίως φορέα απλής κατανομής της γερανογέφυρας

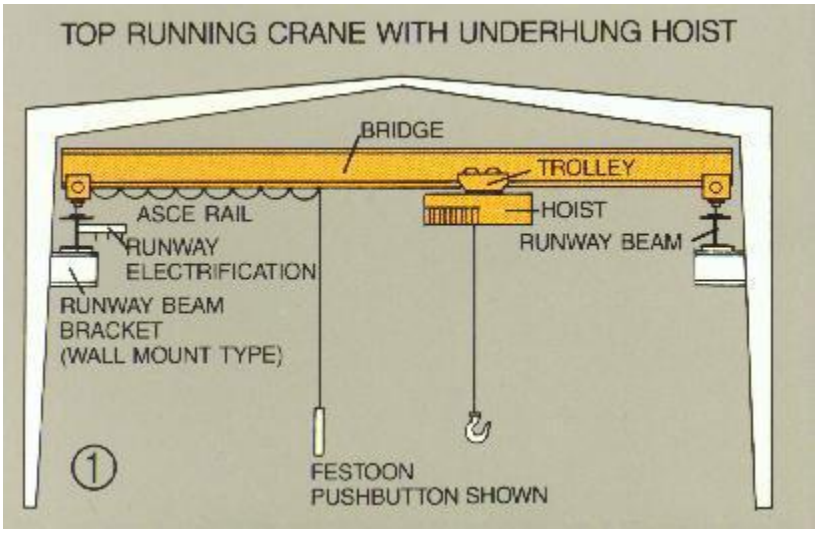
- Σύστημα τροφοδοσίας και ελέγχου

- Ροηφόρος γραμμή με τα εξαρτήματα σύνδεσης και στήριξης της και τον αντίστοιχο ψηκτροφορέα ρευματοληψίας
- Τυποποιημένες μεταλλικές επιψευδαργυρωμένες ράγες διατομής C με τα εξαρτηματά τους και τα τροχήλατα φορεία συγκράτησης και κύλισης των καλωδιώσεων

- Συρόμενα εύκαμπτα καλώδια τύπου σειρίδας για την τροφοδοσία των εγκατεστημένων κινητήρων στη γερανογέφυρα και μεταφορά των ηλ. Σημάτων του κυκλώματος ελέγχου
- Χειριστήριο ενσύρματο ή ασύρματο για τον τηλεχειρισμό της γερανογέφυρας από το έδαφος
- Πίνακας ελεγχου κίνησης

- Συστήματα ασφαλείας γερανογέφυρας

- Σύστημα έναντι της υπερφόρτωσης με ρύθμιση στο ωφέλιμο φορτίο ανύψωσης
- Ηχητική και οπτική διακοπτόμενη σήμανση



Εικόνα 5.1 Συστατικά γερανογέφυρας

Τα στοιχεία που συνθέτουν τη γερανογέφυρα που μελετάται στην εργασία αυτή έχουν επιλεχθεί από τους καταλόγους GH Cranes SA , γνωστό κατασκευαστή συστημάτων ανύψωσης. Η βελτιστοποίησή της έχει γίνει λαμβάνοντας υπόψη τους περιορισμούς που επιβάλλουν οι Ευρωκώδικες. Πιο συγκεκριμένα, ένας πρώτος σχεδιασμός πειραμάτων διεξήχθη χρησιμοποιώντας το κουτί-Behnken κάτι το οποίο έδωσε ως αποτέλεσμα ένα εκτεταμένο σύνολο των φορέων σχεδιασμού μαζί με τις απαντήσεις τους. Στη συνέχεια, αυτό το σύνολο χρησιμοποιήθηκε ως πρώτη ύλη για τα τρία υπολογιστικά μοντέλα που αναλύονται σε αυτό το κεφάλαιο.

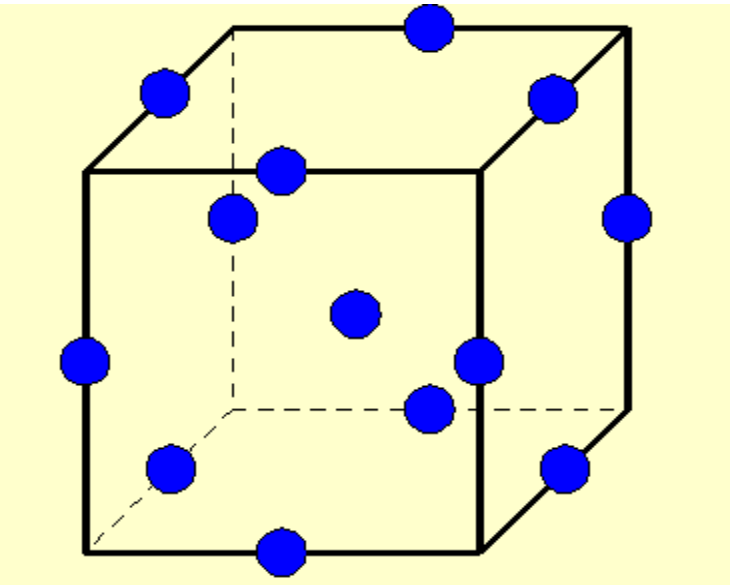
- **Σχεδιασμός Πειραμάτων**

Ο σχεδιασμός πειραμάτων προήλθε γύρω στο 1920 από τον Ronald A. Fisher, έναν Βρετανό επιστήμονα ο οποίος μελέτησε και πρότεινε μια πιο συστηματική προσέγγιση προκειμένου να μεγιστοποιηθεί η γνώση που αποκτήθηκε από τα πειραματικά δεδομένα. Πριν από αυτό, η παραδοσιακή προσέγγιση ήταν να δοκιμάζεται ένας παράγοντας κάθε φορά, που σήμαινε ότι κατά την πειραματική φάση ο πρώτος παράγοντας μεταφερόταν ενώ οι άλλοι παράγοντες παρέμεναν σταθεροί. Μετά εξεταζόταν ο επόμενος παράγοντας και ούτω καθεξής. Με τη χρήση αυτής της παλιάς τεχνικής, πολλές αξιολογήσεις συνήθως απαιτούνταν για να παρθούν επαρκείς πληροφορίες πράγμα το οποίο κατέληξε να είναι μια χρονοβόρα διαδικασία. Η προσέγγιση που προτείνεται από τον Ronald Fisher ξεπέρασε τις παραδοσιακές προσεγγίσεις, καθώς λάμβανε υπόψη όλες τις μεταβλητές ταυτόχρονα, άλλαζε περισσότερες από μία μεταβλητή κάθε φορά και έτσι έδινε τις πιο σχετικές πληροφορίες με την ελάχιστη προσπάθεια.

Από τότε, ο σχεδιασμός των πειραμάτων έχει καθιερωθεί ως μια σημαντική μεθοδολογία που μεγιστοποιεί τη γνώση που αποκτήθηκε από τα πειραματικά δεδομένα με τη χρήση τοποθέτησης σημείων στο χώρο. Η μεθοδολογία αυτή παρέχει ένα ισχυρό εργαλείο για το σχεδιασμό και την ανάλυση πειραμάτων, αφού εξαλείφει περιττές παρατηρήσεις και μειώνει το χρόνο και τους πόρους για να γίνουν τα πειράματα.

- **Κουτί Behnken**

Το κουτί σχεδιασμού Behnken είναι ένα ανεξάρτητο τετραγωνικό σχέδιο υπό την έννοια ότι δεν περιλαμβάνει ένα ενσωματωμένο εργοστασιακό ή κλασματικό παραγοντικό σχεδιασμό. Σε αυτόν τον σχεδιασμό οι συνδυασμοί βρίσκονται στα μεσαία σημεία των ακμών του χώρου διαδικασίας και στο κέντρο. Αυτά τα σχέδια είναι περιστρέψιμα και απαιτούν 3 επίπεδα από κάθε παράγοντα.



Εικόνα 5.2 Κουτί Behnken

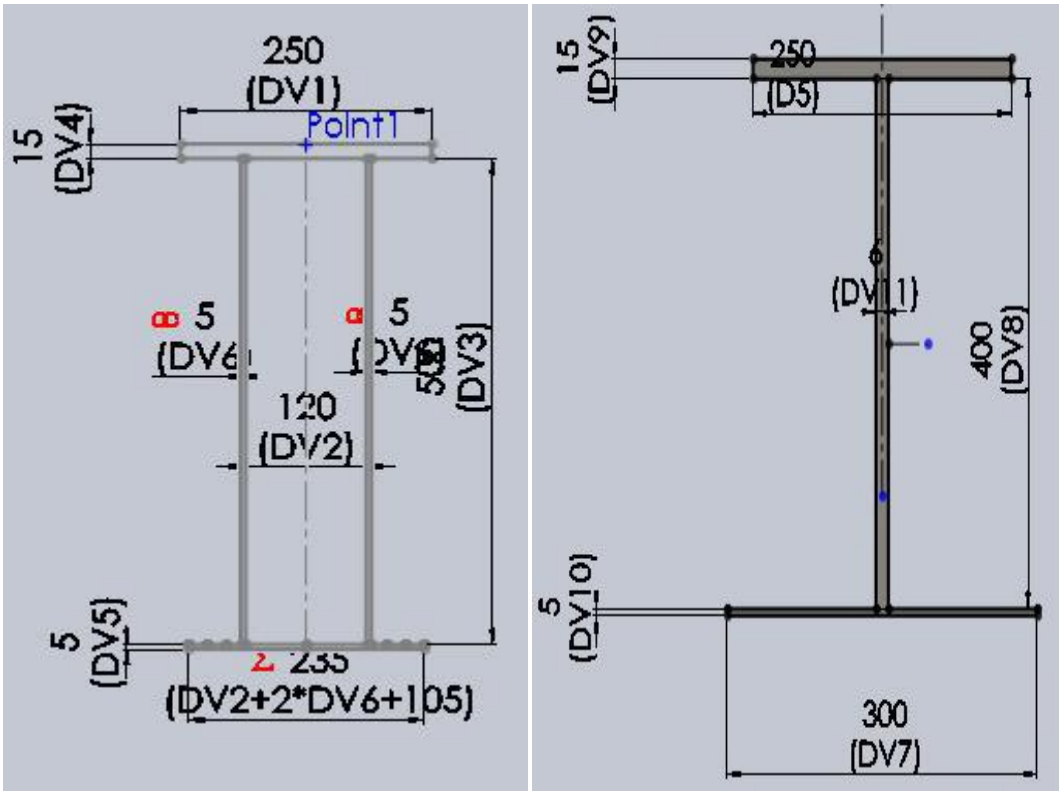
Η γεωμετρία αυτού του σχεδιασμού υποδηλώνει μια σφαίρα εντός του χώρου διεργασίας έτσι ώστε η επιφάνεια της σφαίρας προεξέχει μέσα από κάθε πρόσωπο με την επιφάνεια της σφαίρας εφαπτόμενη στο μέσο της κάθε ακμής του χώρου.

Σε πειραματικό σχέδιο, ένα σχέδιο Behnken Box είναι ένα είδος του σχεδιασμού επιφάνειας απόκρισης που έχει σημεία: το κέντρο του μέσου της κάθε πλευράς.

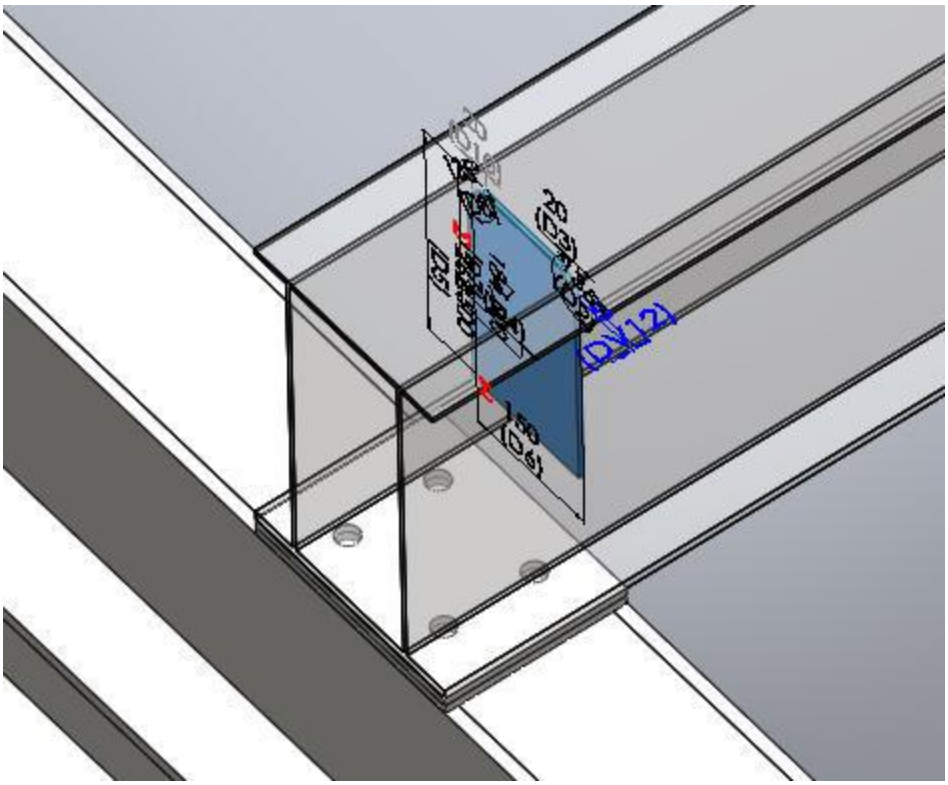
Δεν υπάρχουν γωνιακά σημεία. Ο σχεδιασμός είναι σφαιρικός, όλα τα σημεία (εκτός από το κεντρικό σημείο), βρίσκονται πάνω σε μία σφαίρα ακτίνας $\sqrt{2}$.

Οι μεταβλητές σχεδιασμού που υπολογίστηκαν με την παραπάνω διαδικασία είναι:

- DV1: πλάτος άνω πέλματος φορέα
- DV2: απόσταση κορμών φορέα
- DV3: ύψος φορέα
- DV4: πάχος άνω πέλματος φορέα
- DV5: πάχος κάτω πέλματος φορέα
- DV6: πάχος κορμών φορέα
- DV7: πλάτος κάτω πέλματος δοκού
- DV8: ύψος δοκού
- DV9: πάχος άνω πέλματος δοκού
- DV10: πάχος κάτω πέλματος δοκού
- DV11: πάχος κορμού δοκού
- DV12: πάχος ενισχυτικών



Εικόνα 5.3 Μεταβλητές σχεδιασμού



Εικόνα 5.4 Μεταβλητή σχεδιασμού DV12

Το πλάτος του κάτω πέλματος του φορέα υπολογίζεται αυτόματα για κάθε επανάληψη από τη σχέση DV2+2*DV6+105 όπου 105mm είναι η ελάχιστη απόσταση μεταξύ της εξωτερικής πλευράς του κορμού και του κάτω πέλματος σύμφωνα με τους καταλόγους GH. Επίσης, το πλάτος του άνω πέλματος των δοκών θεωρείται 250mm που συνίσταται επίσης από τους καταλόγους GH.

Μέσω, λοιπόν, του σχεδιασμού πειραμάτων που αναλύθηκε παραπάνω ο οποίος πραγματοποιήθηκε στο περιβάλλον του CosmosWorks ελήφθησαν 193 τιμές των μεταβλητών αυτών για τη σχεδίαση solid μαζί με τις απαντήσεις τους οι οποίες είναι η αναπτυσσόμενη τάση Von Mises, οι κάθετες και οριζόντιες μετατοπίσεις και η μάζα που είναι η αντικειμενική συνάρτηση του προβλήματος βελτιστοποίησης.

Τα σύνολα των μεταβλητών σχεδιασμού και οι έξοδοί τους είναι τα στοιχεία που χρησιμοποιούνται για την εκπαίδευση των υπολογιστικών μοντέλων που ακολουθούν στις παρακάτω ενότητες και κάποιες ενδεικτικές τιμές τους παρουσιάζονται στον ακόλουθο πίνακα, όπου οι διαστάσεις είναι σε mm (5.1).

Πίνακας 5.1 Μεταβλητές σχεδιασμού για τη γερανογέφυρα

DV	DV1	DV2	DV3	DV4	DV5	DV6	DV7	DV8	DV9	DV10	DV11	DV12
Iter1	250	100	350	10	5	7.5	250	300	15	10	7.5	10
Iter2	400	100	350	10	5	7.5	350	300	15	10	7.5	10
Iter3	250	200	350	10	5	7.5	350	300	15	10	7.5	10
Iter4	400	200	350	10	5	7.5	250	300	15	10	7.5	10
Iter5	250	100	350	10	15	7.5	250	300	15	10	7.5	10
Iter6	400	100	350	10	15	7.5	350	300	15	10	7.5	10
Iter7	250	200	350	10	15	7.5	350	300	15	10	7.5	10
Iter8	400	200	350	10	15	7.5	250	300	15	10	7.5	10
Iter9	400	200	350	10	15	7.5	350	300	15	10	7.5	10
Iter10	250	200	350	10	15	7.5	250	300	15	10	7.5	10
Iter11	400	100	350	10	15	7.5	250	300	15	10	7.5	10
Iter12	250	100	350	10	15	7.5	350	300	15	10	7.5	10
Iter13	400	200	350	10	5	7.5	350	300	15	10	7.5	10
Iter14	250	200	350	10	5	7.5	250	300	15	10	7.5	10
Iter15	325	100	350	10	5	7.5	250	300	15	10	7.5	10
Iter16	325	100	350	10	5	7.5	350	300	15	10	7.5	10
Iter17	325	100	200	10	10	5	300	200	15	10	7.5	10
Iter18	325	200	200	10	10	5	300	400	15	10	7.5	10
Iter19	325	100	500	10	10	5	300	400	15	10	7.5	10
Iter20	325	200	500	10	10	5	300	400	15	10	7.5	10

5.2.1.2 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Στο κεφάλαιο αυτό ο στόχος είναι να ερευνηθεί η ικανότητα των Νευρωνικών Δικτύων (NN) να προβέψουν τις ακριβείς εξόδους ανάλυσης που είναι αναγκαίες για την βελτιστοποίηση. Αυτό επιτυγχάνεται με κατάλληλη εκπαίδευση του NN η οποία περιλαμβάνει τις ακόλουθες εργασίες: (i) επιλογή του κατάλληλου συνόλου-δείγματος εκπαίδευσης (training set), (ii) επιλογή κατάλληλης αρχιτεκτονικής του δικτύου και (iii) επιλογή των κατάλληλων τιμών των χαρακτηριστικών παραμέτρων όπως ο ρυθμός μάθησης και ο όρος momentum.

Ο κύριος περιορισμός ενός αλγορίθμου εκπαίδευσης NN είναι το γεγονός ότι η αποτελεσματικότητά του εξαρτάται από το ρυθμό μάθησης, τον όρο momentum και την αρχιτεκτονική του δικτύου. Δυστυχώς, δεν υπάρχουν γενικοί κανόνες για την επιλογή αυτών των παραμέτρων. Συνήθως εξαρτώνται από τον τύπο του προβλήματος και την εμπειρία του σχεδιαστή και μερικές φορές μπορεί να επιλεγούν με μία προσέγγιση δοκιμής και σφάλματος.

Ένας ακόμα σημαντικός παράγοντας για τη λειτουργία των NN είναι και η συνάρτηση μεταφοράς, η οποία αποτελεί το τελικό στάδιο πριν παραχθεί τελικά η έξοδος του νευρώνα. Στον αλγόριθμό μας έχει επιλεγεί η συγμοειδής συνάρτηση μεταφοράς η οποία δίνει αποτελέσματα μεταξύ του 0 και 1. Οπότε για να παράγει αξιόλογα αποτελέσματα οι τιμές εξόδου θα πρέπει να βρίσκονται εντός της ίδιας κλίμακας. Επίσης το πιο σωστό είναι να μορφοποιήσουμε και τις τιμές εισόδου ώστε να βρίσκονται μεταξύ 0 και 1.

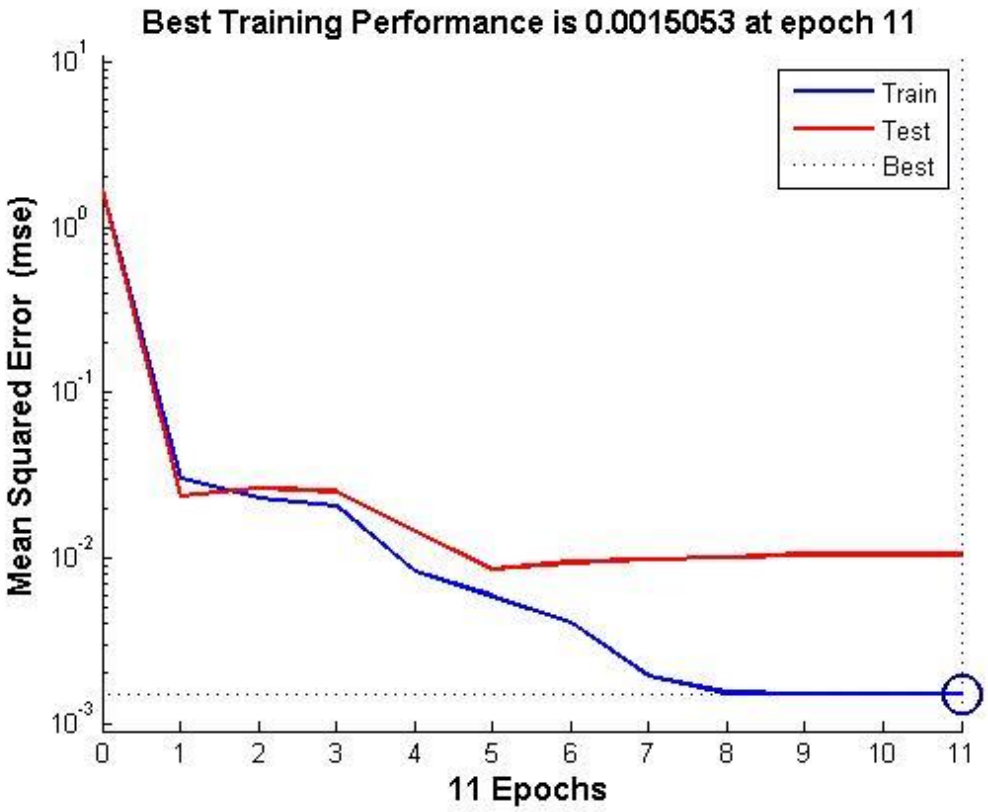
Κρυμμένα επίπεδα	2
Κόμβοι σε κάθε επίπεδο	50
Συνάρτηση μεταφοράς	Logsig
Data division function	Dividerand
Training function	Trainlm
Performance function	Mse
Derivative function	Defaultderiv
Maximum number of training epochs	300
Maximum training time	Infinity
Minimum performance value	0
Minimum gradient magnitude	1e-10
Maximum Mu	1e+10

Έπειτα λοιπόν από τον προσδιορισμό των παραπάνω παραμέτρων και μετά από μια σειρά δοκιμών καταλήξαμε στην τελική μορφή δικτύου, η οποία μας έδωσε την πιο καλή προσέγγιση εξόδων (αναπτυσσόμενων τάσεων και μετατοπίσεων). Ως κριτήριο για την επιλογή της βέλτιστης λύσης θεωρήθηκε τόσο το σφάλμα που υπολογιζεται

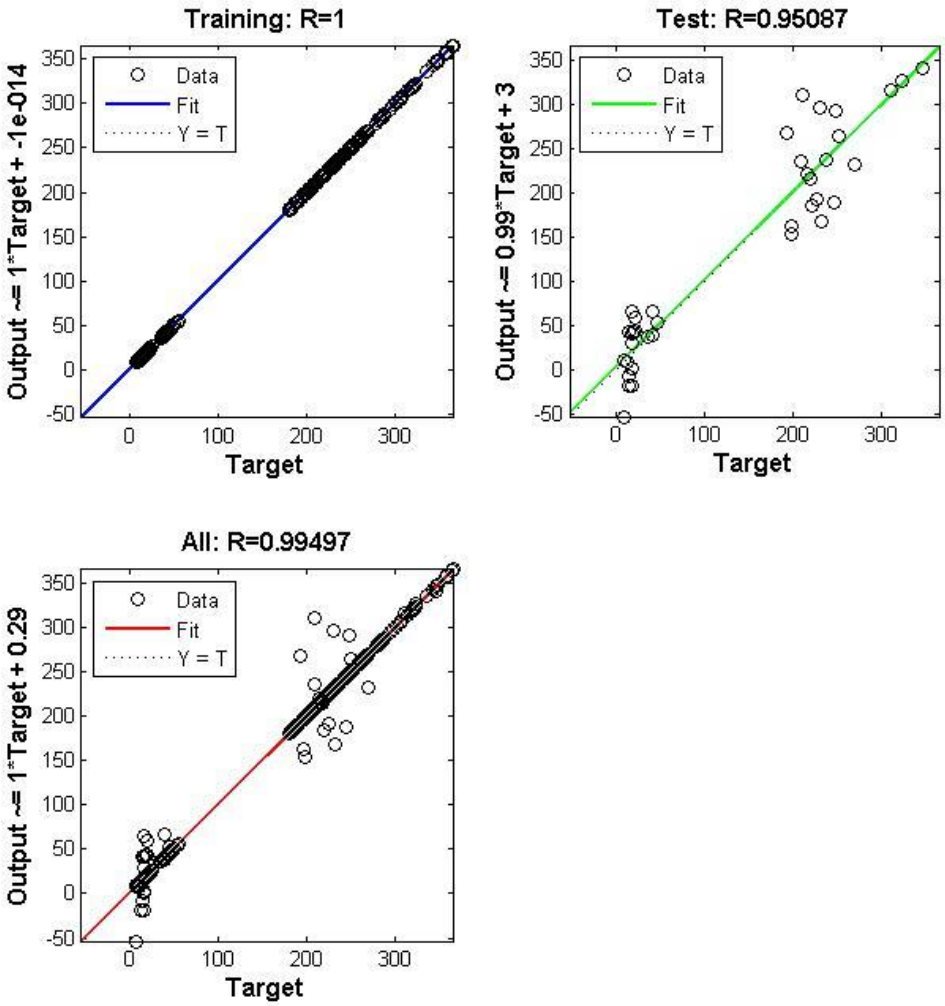
μέσω της σχέσης MSE= $(TARi,j-OUTi,j)^2$ όσο και η μορφή των διαγραμμάτων παλινδρόμησης. Το δίκτυο το οποίο παρουσιάζει τη βέλτιστη συμπεριφορά έχει τα χαρακτηριστικά του παραπάνω πίνακα.

Η τιμή στην ελάχιστη κλίση έχει αλλάξει από την προκαθορισμένη καθώς η τελευταία προκαλούσε πρόωρο τερματισμό στην εκπαίδευση του δικτύου και συνεπώς δεν επέτρεπε στο δίκτυο να μας δώσει τα κατάλληλα αποτελέσματα.

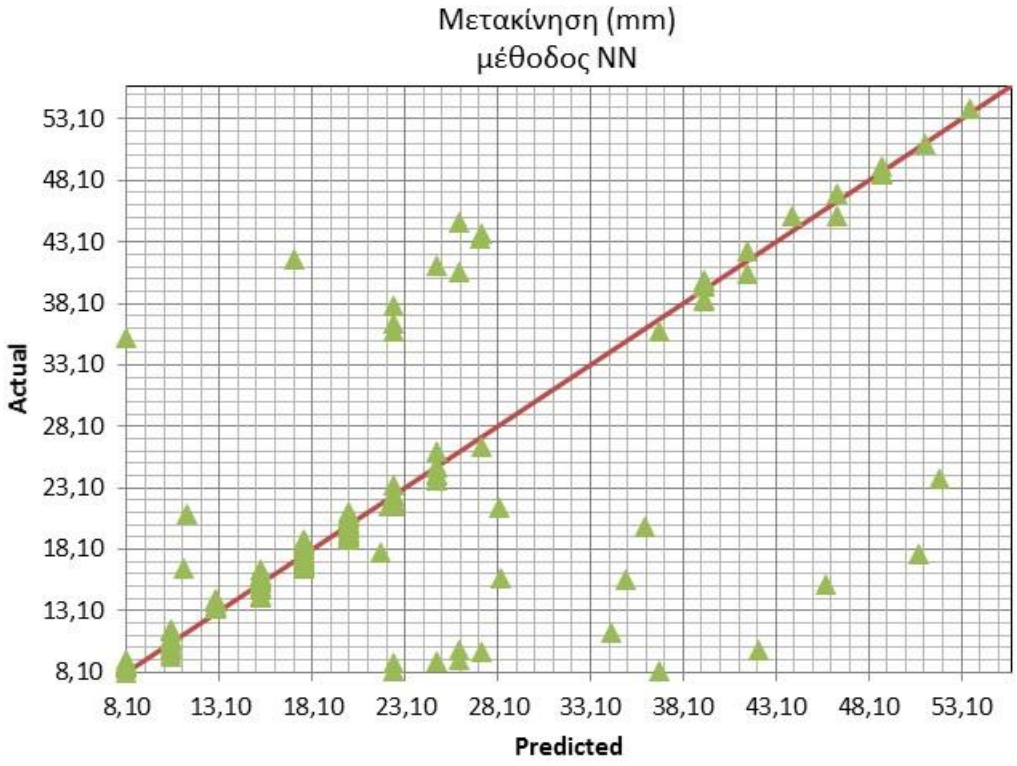
Η συμπεριφορά του δικτύου φαίνεται στα παρακάτω γραφήματα στα οποία παρουσιάζονται τόσο το σφάλμα του νευρωνικού δικτύου όσο και η παλινδρόμηση όπως προκύπτουν από το τρέξιμο του αλγορίθμου. Επιπλέον, παρατίθενται και διαγράμματα στα οποία απεικονίζεται η απόκλιση των προβλέψεων από τις πραγματικές τιμές, για κάθε μέγεθος χωριστά.



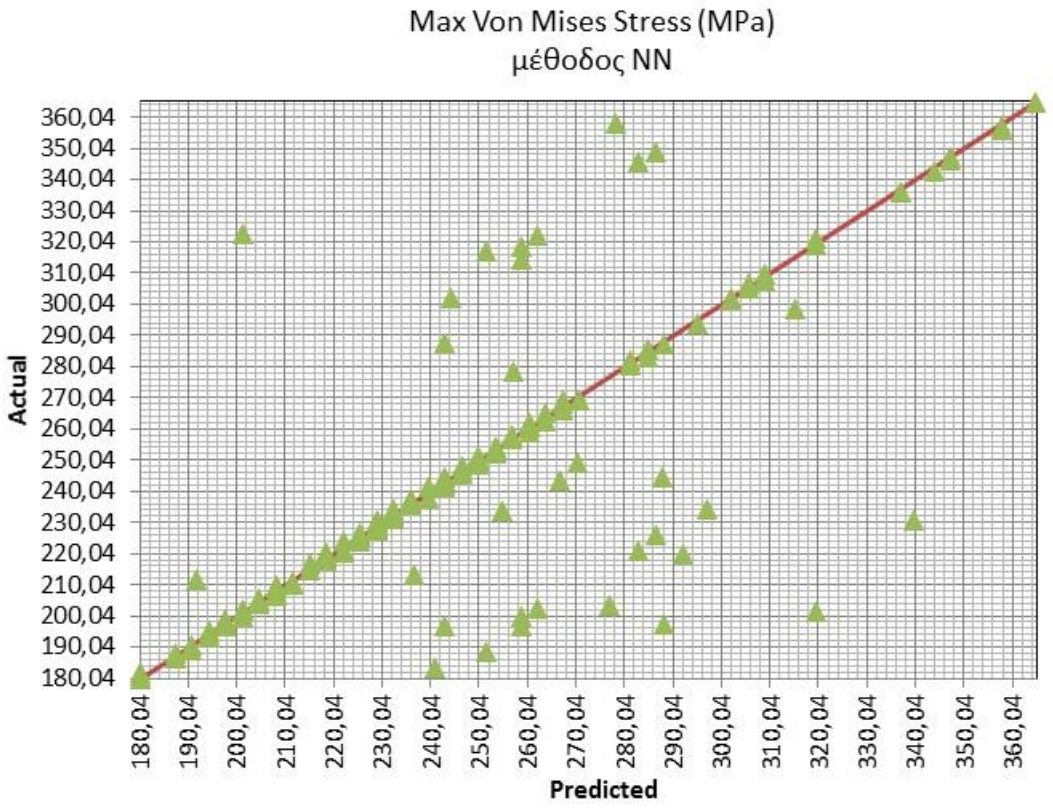
Σχήμα 5.1 Σφάλμα τελικού δικτύου στο training και testing set



Σχήμα 5.2 Παλινδρόμηση τελικού δικτύου



Σχήμα 5.3 Απεικόνιση των αποτελεσμάτων του NN για τη μετακίνηση της δοκού



Σχήμα 5.4 Απεικόνιση των αποτελεσμάτων του NN για την αναπτυσσόμενη τάση Von Mises

5.2.1.3 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ KRIGING

Στο κεφάλαιο αυτό στόχος μας είναι να ερευνήσουμε κατά πόσο η μέθοδος Kriging μπορεί να βοηθήσει στην επίλυση του υπολογιστικού πειράματος της γερανογέφυρας. Πιο συγκεκριμένα, εξετάζονται οι προσεγγίσεις που μας δίνει το μοντέλο αυτό για τη μετατόπιση της δοκού αλλά και την αναπτυσσόμενη τάση Von Mises.

Προκειμένου να φτιάξουμε τον κατάλληλο αλγόριθμο, ο οποίος θα δώσει την καλύτερη εκτίμηση για τις μεταβλητές μας, πρέπει να καθορίσουμε τις τιμές που θα δώσουμε στις παραμέτρους με βάση τις οποίες λειτουργεί το Kriging. Αυτές είναι η συνάρτηση παλινδρόμησης, η συνάρτηση συσχέτισης καθώς και ο συντελεστής ϑ, που αποτελεί παράμετρο της συνάρτησης συσχέτισης, μαζί με τα άνω και κάτω όριά του. Είναι φανερό, λοιπόν, πως η διαδικασία εύρεσης του κατάλληλου αλγορίθμου είναι αρκετά απλή εφόσον καθορίζεται από αρκετά λίγες παραμέτρους. Επίσης, ένα θετικό στοιχείο της συγκεκριμένης μεθόδου είναι ότι για κάθε τρέξιμο του αλγορίθμου τα αποτελέσματα είναι τα ίδια. Πρόκειται, δηλαδή, για μια ντετερμινιστική μέθοδο, σε αντίθεση με τη μέθοδο των Νευρωνικών δικτύων η

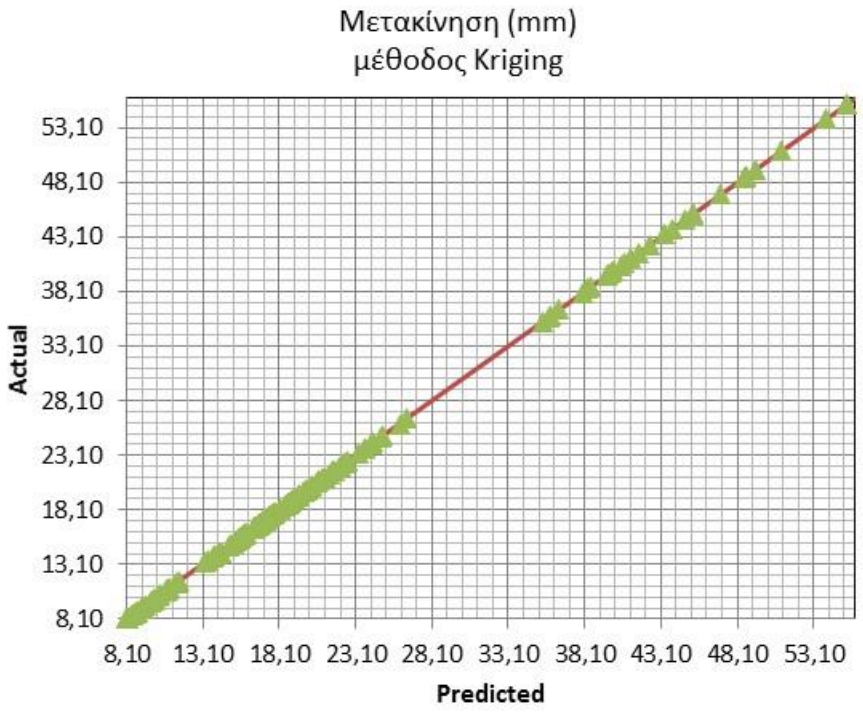
οποία δίνει διαφορετικά αποτελέσματα κάθε φορά που τρέχουμε έναν αλγόριθμο, άσχετα αν όλες οι παράμετροι διατηρούν τις ίδιες τιμές.

Οι μεταβλητές εισόδων και εξόδων παραμένουν οι ίδιες, όπως και στην επίλυση με Νευρωνικά δίκτυα. Έχουμε, δηλαδή 12 μεταβλητές εισόδου με 193 τιμές και 2 αντίστοιχες μεταβλητές εξόδου.

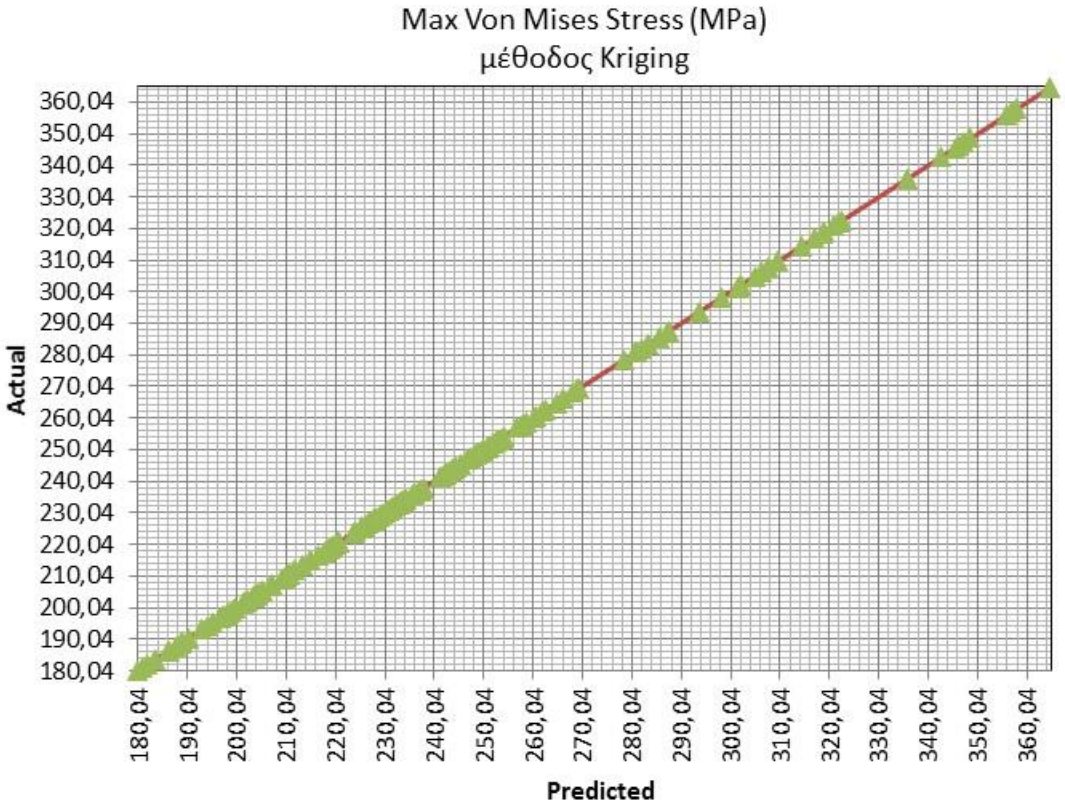
Ο αλγόριθμος που χρησιμοποιήθηκε, λοιπόν, στο συγκεκριμένο πρόβλημα περιέχει τα εξής στοιχεία:

Συνάρτηση παλινδρόμησης	POLY1
Συνάρτηση συσχέτισης	GAUSS
Παράμετρος θ	10
Κάτω όριο του ϑ	1e-1
Άνω όριο του ϑ	20
Πειραματικός σχεδιασμός	Latin Hypercube

Τα αποτελέσματα που παίρνουμε απεικονίζονται στα παρακάτω διαγράμματα, χωριστά για κάθε μέγεθος, και με τρόπο τέτοιο ώστε να φαίνεται η απόκλιση από τις πραγματικές τιμές.



Σχήμα 5.5 Απεικόνιση των αποτελεσμάτων του Kriging για τη μετακίνηση της δοκού



Σχήμα 5.6 Απεικόνιση των αποτελεσμάτων του Kriging για την αναπτυσσόμενη τάση Von Mises

5.2.1.4 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ SVM

Στην ενότητα αυτή, εξετάζεται η τρίτη και τελευταία μέθοδος για την προσέγγιση του προβλήματος της γερανογέφυρας. Όπως έχει αναφερθεί και στο κεφάλαιο 3, η μέθοδος SVM εφαρμόζεται τόσο σε περιπτώσεις ταξινόμησης όσο και σε περιπτώσεις παλινδρόμησης. Στο πρόβλημα της γερανογέφυρας μας ενδιαφέρει η δεύτερη χρήση τους, αφού σκοπός μας δεν είναι να διαχωρίσουμε τα δεδομένα μας σε κλάσεις αλλά να βρούμε τη σχέση που αναπτύσσεται μεταξύ τους. Πιο συγκεκριμένα, επιθυμούμε να προβλέπονται σωστά οι έξοδοι για όλα τα δεδομένα στοιχεία εισόδου και αν τυχόν δοθούν νέες είσοδοι στο μοντέλο να μπορούν να εξαχθούν με ικανοποιητική ακρίβεια οι αντίστοιχες έξοδοι.

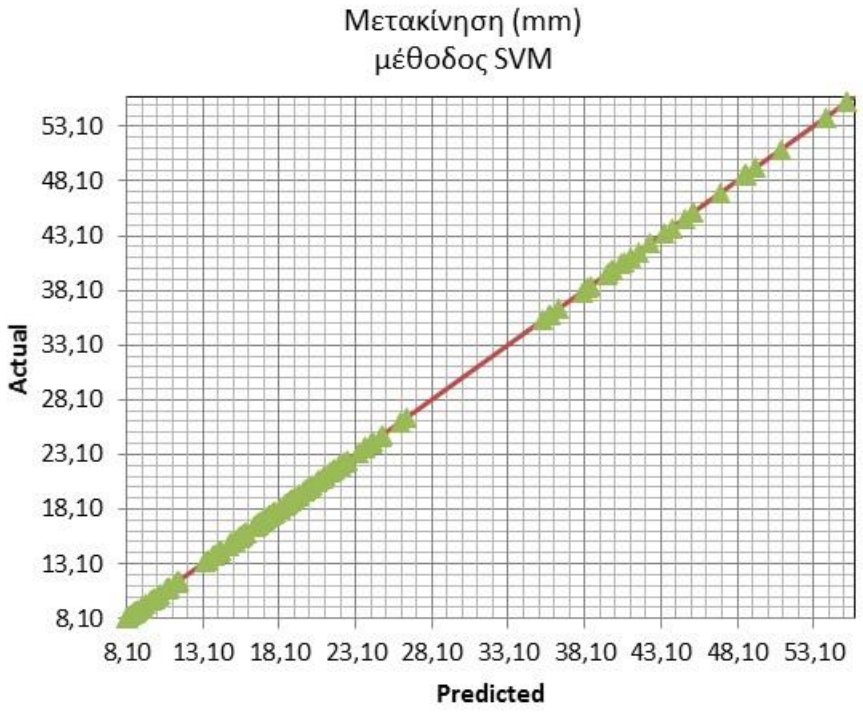
Τα βήματα που ακολουθούνται για την κατασκευή του αλγορίθμου του προβλήματος είναι τα ακόλουθα:

1. Διαβάζονται και κανονικοποιούνται τα δεδομένα
2. Εκπαιδεύονται τα SVM
3. Εξάγονται τα αποτελέσματα

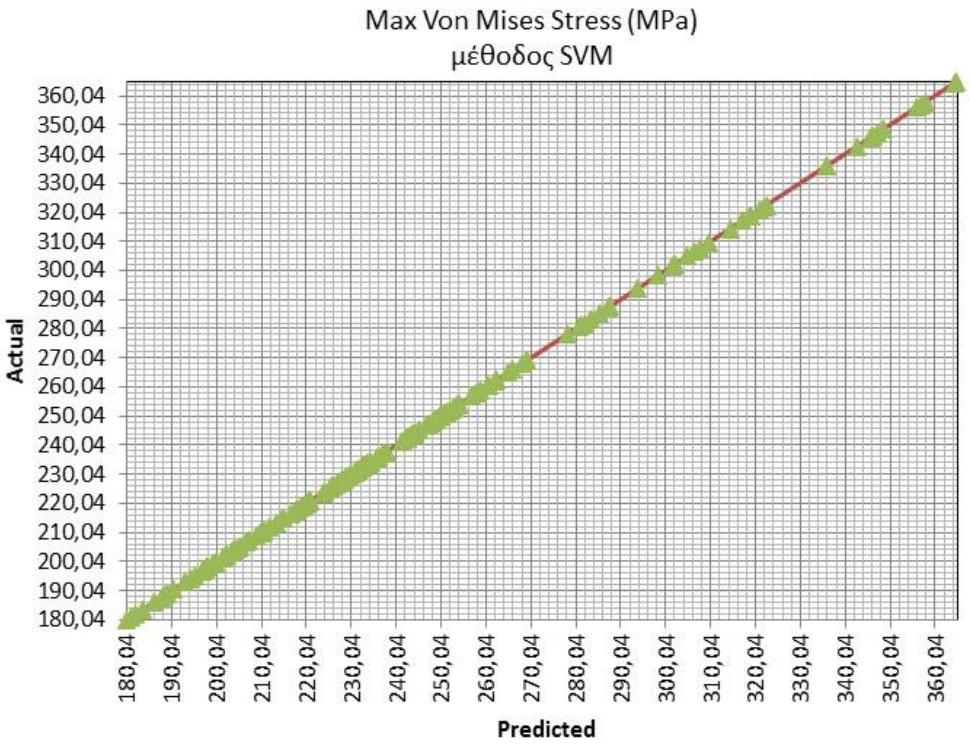
Οι παράμετροι που κατά κύριο λόγο επηρεάζουν την εκπαίδευση του αλγορίθμου είναι η μορφή κλιμάκωσης, ο τύπος Kernel, η παράμετρος p1 της συνάρτησης Kernel, η παράμετρος C, η συνάρτηση σφάλματος (loss function) και η αντίστοιχη παράμετρος e (insensitivity). Κάθε έξοδος του προβλήματος (μετακίνηση, τάση Von Mises) εξετάζεται χωριστά, αλλά, μιας και στη συγκεκριμένη περίπτωση χρησιμοποιήθηκαν ίδιες τιμές για όλες τις παραμέτρους και στις δυο εξόδους, συνοψίζουμε σε έναν πίνακα τα χαρακτηριστικά των αλγορίθμων.

Μορφή κλιμάκωσης	Anisotropic
Τύπος Kernel	Gausian RBF
Παράμετρος p1	1
Παράμετρος C	Inf
Loss function	Quadratic
Insensitivity e	0

Η ακρίβεια με την οποία προβλέπουν οι αλγόριθμοι τις τιμές για τη μετακίνηση αλλά και για την αναπτυσσόμενη τάση της δοκού παρουσιάζονται μέσω σχετικών διαγραμμάτων παρακάτω (Σχήματα 5.7,5.8).



Σχήμα 5.7 Αποτελέσματα της μεθόδου SVM για τη μετακίνηση της δοκού



Σχήμα 5.8 Αποτελέσματα της μεθόδου SVM για την αναπτυσσόμενη τάση Von Mises της δοκού

5.2.2 ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΠΙΘΑΝΟΤΗΤΑΣ ΡΕΥΣΤΟΠΟΙΗΣΗΣ ΑΜΜΩΔΩΝ ΕΔΑΦΩΝ

5.2.2.1 Εισαγωγή στην έννοια της ρευστοποίησης

Μη συνεκτικοί κορεσμένοι εδαφικοί σχηματισμοί, όταν υπόκεινται σε άμεση φόρτιση κάτω από αστράγγιστες συνθήκες, τείνουν προς συμπύκνωση το οποίο σε συνδυασμό με την αδυναμία μεταβολής του όγκου τους οδηγεί σε αύξηση της πίεσης του νερού των πόρων τους με ταυτόχρονη μείωση έως και μηδενισμό της διατμητικής τους αντοχής. Κατά την παραπάνω διαδικασία μετατρέπεται η κατάσταση αυτών των εδαφικών στρωμάτων από τη στερεά στη ρευστή φάση, δηλαδή προκαλείται ρευστοποίηση αυτών.

Βασικό ρόλο στην πρόκληση ρευστοποίησης διαδραματίζουν οι συνθήκες κάτω από τις οποίες οι εδαφικοί σχηματισμοί συναντώνται στη φύση. Πιο επιρρεπείς στην υγροποίηση είναι λεπτόκοκκοι μη συνεκτικοί σχηματισμοί όπως για παράδειγμα τα αμμώδη εδάφη. Όμως, πέρα από το είδος του εδαφικού σχηματισμού υπάρχουν και κάποια επιπλέον κριτήρια τα οποία καθιστούν ένα εδαφικό υλικό επιδεκτικό προς

ρευστοποίηση. Τα σημαντικότερα, λοιπόν, κριτήρια επιδεκτικότητας προς ρευστοποίηση ενός εδαφικού σχηματισμού είναι τα εξής:

- Γεωλογική ηλικία εδαφικών σχηματισμών
- Γεωμορφολογικά κριτήρια
- Στάθμη υδροφόρου ορίζοντα
- Σχετική πυκνότητα
- Κοκκομετρική σύσταση και φυσικά χαρακτηριστικά των εδαφών
- Ιστορικά περιστατικά

Η ρευστοποίηση ή υγροποίηση των εδαφών είναι ένα είδος αστοχίας του εδάφους που σχετίζεται κυρίως με τους σεισμούς. Η αύξηση της πίεσης των πόρων μη συνεκτικών κορεσμένων εδαφικών σχηματισμών, η οποία προκαλείται από δυναμικού τύπου δονήσεις (σεισμικές ταλαντώσεις κλπ) υπό αστράγγιστες συνθήκες, μπορεί να επιφέρει τη μείωση έως και την απώλεια της διατμητικής αντοχής του υλικού δημιουργώντας το φαινόμενο της ρευστοποίησης. Κατά τη ρευστοποίηση, οι κόκκοι της άμμου απλά αιωρούνται στο νερό χωρίς να ασκούνται μεταξύ τους δυνάμεις, δηλαδή η συμπεριφορά του εδάφους είναι παρόμοια με τη συμπεριφορά ενός υγρού.

Τα φαινόμενα ρευστοποίησης διακρίνονται με βάση το μηχανισμό γένεσης σε δύο γενικές κατηγορίες: την εδαφική ροή (flow failure) και την ανακυκλική κινητικότητα (cyclic mobility). Μολονότι η εδαφική ροή παρουσιάζει λιγότερες πιθανότητες εμφάνισης από την ανακυκλική κινητικότητα, οι επιπτώσεις της είναι αρκετά πιο σοβαρές. Η εδαφική ροή συναντάται συνήθως σε χαλαρά μη συνεκτικά ιζήματα τα οποία είναι τοποθετημένα σε πρανή με σημαντική κλίση και στα οποία η διατμητική τάση η οποία απαιτείται για τη στατική ισορροπία της εδαφικής μάζας είναι μεγαλύτερη από την υπολειπόμενη διατμητική αντοχή του εδάφους (Kramer, 1996).
Οι εφαρμοσμένες ανακυκλιζόμενες τάσεις οδηγούν τον εδαφικό σχηματισμό σε αστάθεια καθώς η αντοχή του μειώνεται με αποτέλεσμα οι υπάρχουσες στατικές φορτίσεις να προκαλέσουν στη συνέχεια την δημιουργία εδαφικής ροής. Ένα τέτοιο μηχανισμό μονοτονικής φόρτισης αποτελεί η διάβρωση του πόδα ενός πρανούς ή αντίστοιχα η φόρτιση της κεφαλής του. Το συγκεκριμένο φαινόμενο χαρακτηρίζεται από την απότομη εμφάνιση, την ταχύτατη κίνηση και τη μεγάλη σε κλίμακα απόσταση στην οποία μπορεί να κινηθούν τα υλικά του ρευστοποιημένου εδάφους. Χαρακτηριστικά παραδείγματα εμφάνισης εδαφικής ροής αποτελούν οι αστοχίες των φραγμάτων Sheffield και Lower San Fernando (Olson, 2001).
Η ανακυκλική κινητικότητα (cyclic mobility) η οποία χαρακτηρίζεται από μεγαλύτερη συχνότητα εμφάνισης, μπορεί να προκληθεί σε περιοχές χωρίς ή με μικρή κλίση (έως 3ο), ενώ οι επιπτώσεις αυτής χαρακτηρίζονται ελαφριές έως πολύ σοβαρές. Στην παρούσα έρευνα θα ασχοληθούμε κυρίως με τη δημιουργία φαινομένων

ανακυκλικής κινητικότητας και τις προκαλούμενες από αυτήν εδαφικές παραμορφώσεις και μετατοπίσεις. Σε αντίθεση με την εδαφική ροή, φαινόμενα ρευστοποίησης λόγω ανακυκλικής κινητικότητας προκαλούνται όταν η στατική διατμητική τάση είναι μικρότερη της διατμητικής αντοχής του εδαφικού σχηματισμού.

- **Διαδικασία ρευστοποίησης**

Όταν ένας χαλαρός κοκκώδης εδαφικός σχηματισμός υπόκειται σε ανακυκλική δόνηση, όπως στην περίπτωση ενός σεισμού, τότε επιδεικνύει τάση συμπύκνωσης δηλαδή μεταβολής του όγκου. Συγκεκριμένα, οι κόκκοι από τους οποίους αποτελείται το εδαφικό υλικό μετακινούνται έχοντας ως στόχο την μείωση των μεταξύ τους κενών. Όμως, η παρουσία νερού σε αυτά τα κενά, στην περίπτωση κορεσμένων χαλαρών εδαφών σε αστράγγιστες συνθήκες, εμποδίζει αυτήν την μετατόπιση. Αυτή η πίεση που μεταφέρεται στο νερό των πόρων, λόγω των αστράγγιστων συνθηκών δεν μπορεί να εκτονωθεί με συνέπεια την απότομη αύξηση της πίεσης των πόρων u.
Η συνεχιζόμενη αύξηση της u αντιστοιχεί σε μείωση της διατμητικής αντοχής του εδάφους έως και την πλήρη απώλεια αυτής σύμφωνα με τον νόμο του Coulomb:
$\tau = \sigma_v' \ \varepsilon\varphi\varphi'$
όπου $\sigma_v' = \sigma - u$, σ: τάση, σ_v': ενεργή τάση, φ': ενεργός (δρώσα) γωνία τριβής και u: πίεση των πόρων. Ο όρος της συνοχής c δεν υπάρχει διότι αναφερόμαστε σε μη συνεκτικά εδάφη.

Στο στάδιο αυτό, το οποίο ονομάζεται *ολική ρευστοποίηση*, η πίεση του νερού των πόρων, το οποίο δεν έχει απεγκλωβιστεί από το εδαφικό στρώμα, ανταποκρίνεται στην πίεση που ασκούν τα υπερκείμενα στρώματα διατηρώντας τον όγκο του σχηματισμού. Η εκτόνωση της πίεσης του νερού των πόρων στη συνέχεια θα προκαλέσει την αύξηση της πυκνότητας του εδάφους και την ταυτόχρονη μείωση του όγκου του. Η ποσότητα του νερού που παρέμεινε εγκλωβισμένη στο ίζημα, θα διοχετευτεί προς την επιφάνεια λόγω της μεγάλης υδραυλικής κλίσης παρασέρνοντας εδαφικό υλικό. Η διάρκεια αυτής της ροής προς την επιφάνεια εξαρτάται από το χρόνο που χρειάζεται για να αποκατασταθεί η υδραυλική ισορροπία στην εδαφική στήλη. Έτσι λοιπόν, είναι σύνηθες φαινόμενο η συνέχιση αυτής της ροής και μετά το τέλος της σεισμικής δόνησης.

- **Προσέγγιση του προβλήματος ρευστοποίησης**

Κατά τις τελευταίες δεκαετίες, έχουν αναπτυχθεί πολλές μελέτες που εξετάζουν τα φαινόμενα υγροποίησης και έχουν προσφέρει κάποιες τροποποιήσεις σε παλαιότερες μεθόδους. Οι πιο δημοφιλείς προσεγγίσεις οι οποίες χρησιμοποιούν τον πρότυπο δείκτη διείσδυσης (SPT) και δείκτη (CPT) για τον προσδιορισμό του συντελεστή ασφαλείας προσφέρουν ευκολία εφαρμογής σε πολλές περιπτώσεις, ειδικά για αμμώδη εδάφη. Έχουν όμως το μειονέκτημα του ότι χρησιμοποιούν έναν περιορισμένο αριθμό παραμέτρων ενώ τα φαινόμενα ρευστοποίησης περιλαμβάνουν πολλές σεισμικές αλλά και εδαφικές/γεωτεχνικές παραμέτρους. Στο παράδειγμα αυτό θα χρησιμοποιηθεί η σχέση των G. Sen και E. Akyol (2009), η οποία παρουσιάζεται και επεξηγείται παρακάτω:

$$LI = -5.13 \cdot SSSR_{7.5}{}^{4.39} + 2.29 \cdot \ln(r_d{}^{1.6} + 1) + 9.91 \cdot D_{50}{}^{1.31} \cdot SSSR_{7.5}{}^{1.40} - P1 \cdot \ln\left(D_{50}{}^{6.38} + 1\right) - 0.06 \cdot \ln(q_c{}^{2.62} + 1) \cdot r_d{}^{5.11} - P2 \cdot \ln\left(D_{50}{}^{7.74} + 1\right) \cdot GWT^{4.48} - 0.88 \quad (5.1)$$

Όπου P1=0 όταν $\frac{\sigma_{vo}}{\sigma'_{vo}} \leq 0.838$, ενώ *P1=8,97 όταν* $\frac{\sigma_{vo}}{\sigma'_{vo}} > 0{,}838$

Και *P2=0 όταν* $\frac{GWT}{z} \leq 0.555$, ενώ *P2=8,97 όταν* $\frac{GWT}{z} > 0{,}555$

Οι μεταβλητές που περιέχονται στο πρόβλημα καθώς και οι κατανομές τους παρουσιάζονται στον ακόλουθο πίνακα (5.2):

Πίνακας 5.2 Μεταβλητές του προβλήματος ρευστοποίησης

Μεταβλητή	Περιγραφή	Τύπος κατανομής	Παράμετροι κατανομής
a_{max}	Μέγιστη επιτάχυνση εδάφους	Εκθετική	λ=3,407
σ_{v0}	Ολική γεωστατική τάση εδάφους	Ομοιόμορφη	α=22,60, β=296,30
σ_{v0}'	Ενεργός γεωστατική τάση εδάφους	Ομοιόμορφη	α=13,90, β=227,50
q_{c1}	Διορθωμένη αντίσταση αιχμής CTP	Γάμμα	α=1,4133, β=5404,0
D_{50}	Μέση διάμετρος κόκκου	Γάμμα	α=2,6641, β=0,0608
GWT	Βάθος υδροφόρου Ορίζοντα	Ομοιόμορφη	α=0,20, β=8,40
z	Βάθος δοκιμής CTP	Ομοιόμορφη	α=0,59, β=15,10
$SSSSR$	Συντελεστής σεισμικής διατμητικής τάσης	Εκθετική	λ=4,1255
$SSSR_{7.5}$	Διορθωμένος συντελεστής σεισμικής διατμητικής τάσης	Γάμμα (3P)	α=3,0968, β=0,04965, γ=0,06745
q_c	Αντίσταση αιχμής CTP	Γάμμα	α=4,2756, β=0,06865

Στο κεφάλαιο αυτό στόχος μας είναι να ερευνηθεί η ικανότητα των τριών υπολογιστικών μοντέλων να προβλέψουν τις ακριβείς εξόδους που θα μας δείχνουν αν υπάρχει κίνδυνος ρευστοποίησης ή όχι. Πιο συγκεκριμένα, κάθε μοντέλο θα λαμβάνει ως εισόδους τις μεταβλητές που παρουσιάζονται στον παραπάνω πίνακα και ως έξοδο θα δίνει την τιμή για την παράμετρο LI. Αν η τιμή αυτή είναι μεγαλύτερη από 0.5, τότε υπάρχει σημαντική πιθανότητα να συμβεί υγροποίηση. Αν είναι μικρότερη, τότε δεν υπάρχει τέτοιος κίνδυνος. Για ευκολία τιμές οι οποίες είναι μεγαλύτερες από 0.5 θεωρούνται 1 και οι μικρότερες 0.

Η βάση δεδομένων περιλαμβάνει ίσο αριθμό υγροποιημένων και μη-υγροποιημένων εδαφών τα οποία είναι συνολικά 242 περιπτώσεις, από τις οποίες το 90% χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου και 10% ως δοκιμή. Ο διαχωρισμός των 242 στοιχείων στις δυο αυτές κατηγορίες βασίζεται σε τυχαία επιλογή.

Κάποιες ενδεικτικές τιμές για τις 10 μεταβλητές που περιέχονται στο πρόβλημα, οι οποίες επισημάνθηκαν και παραπάνω, παρουσιάζονται στον πίνακα 5.3. Η πρώτη στήλη δείχνει αν υπάρχει κίνδυνος ρευστοποίησης ή όχι και οι υπόλοιπες στήλες αφορούν στις παραμέτρους οι οποίες επηρεάζουν την συμπεριφορά του εδάφους και συντελούν στην υγροποίησή του.

Πίνακας 5.3 Ενδεικτικές τιμές για τις μεταβλητές του προβλήματος

LI	a_{max}	σ_{v0}	σ_{v0}'	q_{c1}	D_{50}	GWT	z	Cq	r_d	$SSSSR$	$SSSR_{7.5}$	q_c
1	0.2	111.8	54.3	3280	0.062	0.2	5.9	1.35	0.93	0.25	0.26	2.430
1	0.16	108.9	66.2	1860	0.07	2.5	6	1.24	0.93	0.15	0.14	1.500
1	0.16	108.9	66.2	2350	0.07	2.5	6	1.24	0.93	0.15	0.14	1.895
1	0.24	115	83.6	830	0.16	2.9	6.1	1.11	0.93	0.2	0.19	0.748
1	0.5	113.7	103.3	680	0.055	4.7	5.8	0.99	0.93	0.33	0.34	0.687
1	0.4	37.3	28.4	2780	0.12	1.1	2	1.67	0.98	0.33	0.34	1.665
1	0.2	93.2	51	1490	0.07	0.7	5	1.38	0.94	0.22	0.23	1.080
1	0.15	55.9	41.2	810	0.08	1.5	3	1.53	0.96	0.13	0.13	0.529
1	0.6	62.8	44.5	2560	0.11	2.1	4	1.45	0.95	0.52	0.46	1.766
1	0.29	130.5	91.2	12380	0.26	3	6.5	1.06	0.92	0.25	0.23	11.679
1	0.4	55.9	42.2	3770	0.14	1.6	3	1.48	0.96	0.33	0.34	2.547
1	0.24	120.6	81.8	1340	0.197	2.4	6.4	1.12	0.92	0.21	0.2	1.196
0	0.24	140.9	100.1	13080	0.35	2.7	6.9	1.01	0.92	0.2	0.19	1.295
0	0.24	131.9	117.1	5060	0.244	5.5	7	0.92	0.92	0.16	0.15	5.500
0	0.5	194.5	154.2	1370	0.07	5.8	9.9	0.78	0.88	0.36	0.3	1.756
0	0.16	84.3	45.1	11340	0.3	0.5	4.5	1.45	0.95	0.18	0.18	7.821
0	0.23	56.1	47.1	13960	0.32	2	3.1	1.42	0.96	0.17	0.18	9.831
0	0.29	177.1	113.3	17000	0.27	3	9.5	0.94	0.89	0.26	0.24	1.809
0	0.25	70.4	53.1	6680	0.1	1.8	3.5	1.36	0.96	0.21	0.15	4.912
0	0.24	118.4	88.5	20440	0.253	2.7	5.8	1.08	0.93	0.19	0.19	1.893
0	0.6	62.8	44.5	10110	0.08	2.1	4	1.45	0.95	0.52	0.46	6.972
0	0.1	99	56.9	3240	0.1	1	5.3	1.32	0.94	0.11	0.11	2.4550
0	0.24	140.9	100.1	13080	0.35	2.7	6.9	1.01	0.92	0.2	0.19	12.950
0	0.5	260.3	212.5	7310	0.4	8.4	13.3	0.62	0.84	0.33	0.28	11.790

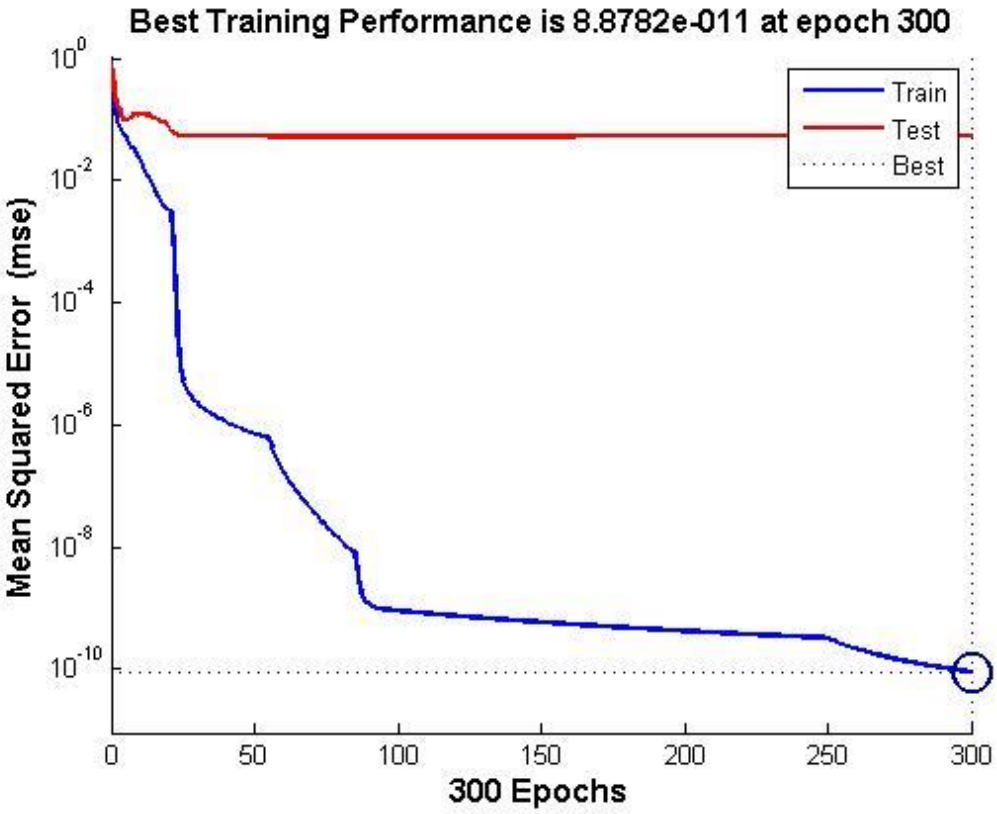
5.2.2.2 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Όπως και στο πρόβλημα της γερανογέφυρας, έτσι και εδώ προσπαθούμε να βρούμε το κατάλληλο νευρωνικό δίκτυο που θα δίνει την καλύτερη προσέγγιση στις ήδη υπολογισμένες τιμές μας, οι οποίες αφορούν τον κίνδυνο ρευστοποίησης εδάφους. Το δίκτυο στο οποίο καταλήγουμε, λοιπόν, έπειτα από αρκετές επαναλήψεις έχει τα εξής χαρακτηριστικά:

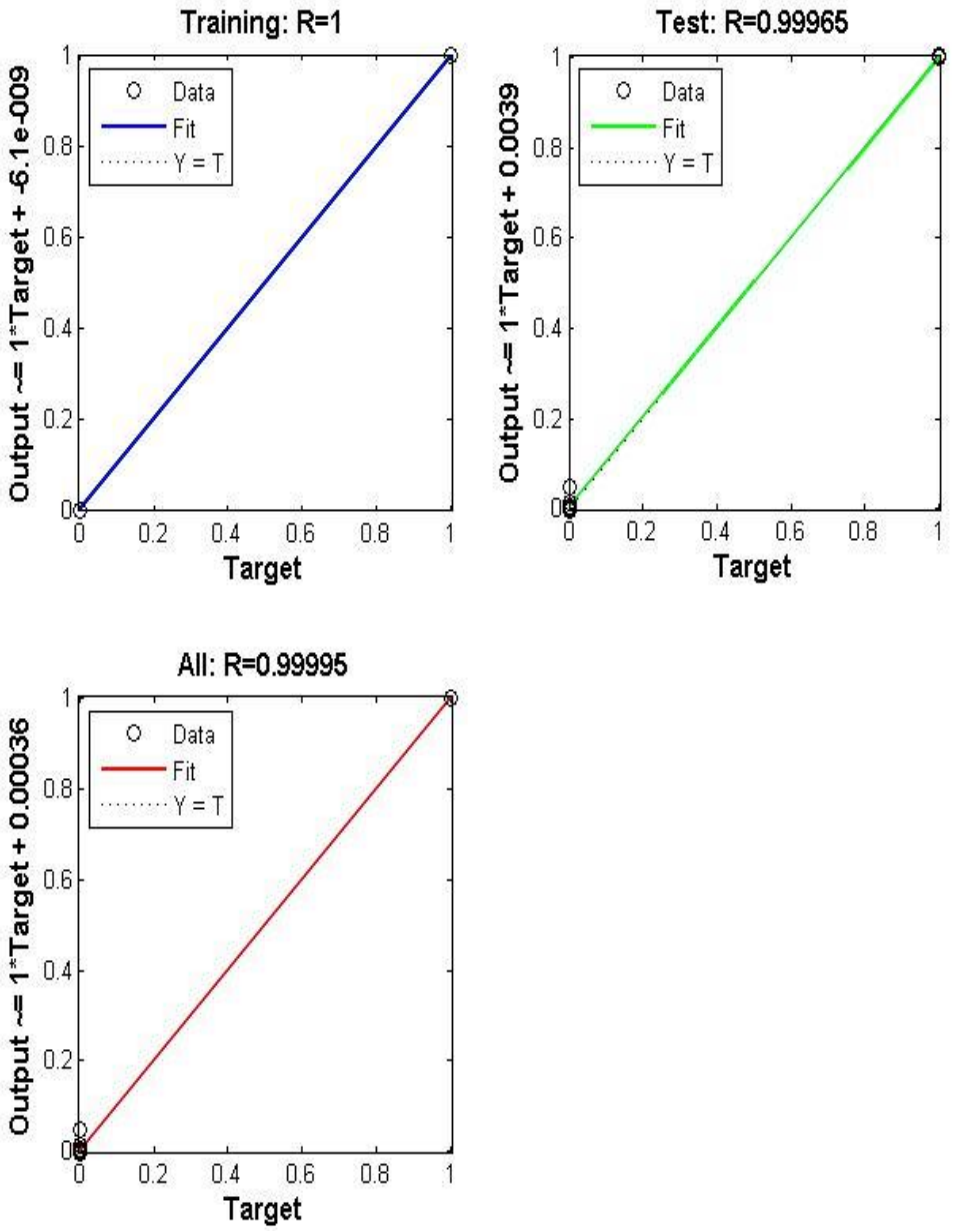
Κρυμμένα επίπεδα	2
Κόμβοι σε κάθε επίπεδο	20
Συνάρτηση μεταφοράς	Logsig
Data division function	Dividerand
Training function	Trainlm
Performance function	Mse
Derivative function	Defaultderiv
Maximum number of training epochs	300
Maximum training time	Infinity
Minimum performance value	0
Minimum gradient magnitude	1e-10
Maximum Mu	1e+10

Όπως αναφέρθηκε και στο πρώτο πρόβλημα, της γερανογέφυρας, η τιμή στην ελάχιστη κλίση έχει μειωθεί από την προκαθορισμένη για να βοηθήσει στην καλύτερη εκπαίδευση του δικτύου.

Η συμπεριφορά του δικτύου φαίνεται στα παρακάτω γραφήματα (Σχήματα 5.9, 5.10), στα οποία όπως και στο πρώτο μας πρόβλημα παρουσιάζονται τόσο το σφάλμα του νευρωνικού δικτύου όσο και η παλινδρόμηση. Όπως αποδεικνύουν και τα διαγράμματα η μέθοδος NN προβλέπει τα τελικά αποτελέσματα με ακρίβεια κοντά στο 100%.



Σχήμα 5.9 Σφάλμα τελικού δικτύου στο training και testing set



Σχήμα 5.10 Παλινδρόμηση τελικού δικτύου

5.2.2.3 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ KRIGING

Στην ενότητα αυτή το αντικείμενο της ρευστοποίησης προσεγγίζεται από τη μέθοδο προσομοίωσης Kriging. Όπως και στη μέθοδο των NN, έτσι και εδώ, κατασκευάζεται ο κατάλληλος αλγόριθμος για την περίπτωση, αφού πρώτα καθοριστούν οι τιμές για κάθε παράμετρο που επηρεάζει τη μέθοδο. Οι μεταβλητές που επηρεάζουν τη ρευστοποίηση είναι οι 12 που έχουν ήδη αναφερθεί ($\alpha_{max,}\sigma_{v0}, \sigma'_{v0}, q_{c1}, D_{50}, GWT, z, C_q, r_d, SSSSR, SSSR_{7.5}, q_c$) και χρησιμοποιούνται ομοίως και εδώ ως είσοδοι του μοντέλου και η έξοδος είναι μια, η οποία καθορίζει το αν υπάρχει κίνδυνος ρευστοποίησης ή όχι. Ο τελικός αλγόριθμος έχει τα εξής χαρακτηριστικά:

Συνάρτηση παλινδρόμησης	POLY1
Συνάρτηση συσχέτισης	GAUSS
Παράμετρος θ	10
Κάτω όριο του ϑ	1e-1
Άνω όριο του ϑ	20
Πειραματικός σχεδιασμός	Latin Hypercube

Τα αποτελέσματα που εξάγει η μέθοδος βρίσκονται πολύ κοντά στα πραγματικά, κάτι που αποδεικνύεται και από το σφάλμα της μεθόδου το οποίο είναι σε όλες τις θέσεις σχεδόν μηδενικό. Κάποια ενδεικτικά αποτελεσματα παρουσιάζονται και στον πίνακα 5.4.

Πίνακας 5.4 Σφάλμα της μεθόδου Kriging σε ενδεικτικές θέσεις

ΘΕΣΗ	ΣΦΑΛΜΑ
1	2,87E-15
2	9,56E-16
3	9,56E-16
4	9,56E-16
5	1,43E-15
6	1,43E-15
7	2,87E-15
8	1,43E-15
9	1,43E-15
10	2,87E-15
11	2,87E-15
12	9,56E-16
13	2,87E-15
14	1,43E-15
15	2,87E-15

5.2.2.4 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ SVM

Στην ενότητα αυτή, εξετάζεται η καταλληλότητα της μεθόδου SVM για την προσέγγιση του προβλήματος της ρευστοποίησης εδάφους. Η διαφορά εδώ σε σχέση με την περίπτωση της γερανογέφυρας, όπου και πάλι εφαρμόστηκε η ίδια μέθοδος, είναι ότι επιθυμούμε ταξινόμηση των δεδομένων και όχι εύρεση συνάρτησης παλινδρόμησης. Έχοντας ως εισόδους τις μεταβλητές που επηρεάζουν την ρευστοποίηση, θέλουμε να διαχωρίσουμε τα στοιχεία σε δυο κλάσεις ανάλογα με το αν υπάρχει κίνδυνος ρευστοποίησης ή όχι. Έτσι, κατασκευάζεται ο κατάλληλος αλγόριθμος με τα χαρακτηριστικά που παρουσιάζονται στον ακόλουθο πίνακα. Τα βήματα που ακολουθούνται κατά την κατασκευή του είναι τα ίδια όπως και στη γερανογέφυρα, δηλαδή διάβασμα και κανονικοποίηση των δεδομένων, εκπαίδευση των SVM και εξαγωγή των αποτελεσμάτων. Στην περίπτωση ταξινόμησης οι παράμετροι που καθορίζουν την εκπαίδευση του αλγορίθμου είναι η μορφή της κλιμάκωσης, ο τύπος Kernel, η παράμετρος p1 της συνάρτησης Kernel και τέλος η παράμετρος C.

Μορφή κλιμάκωσης	Anisotropic
Τύπος Kernel	Gausian RBF
Παράμετρος p1	1
Παράμετρος C	Inf

Από τα αποτελέσματα της μεθόδου, τα οποία έδειξαν πως τα δεδομένα διαχωρίστηκαν άριστα στις δυο κλάσεις κατανοούμε πως και η μέθοδος SVM λειτουργεί πολύ αποδοτικά στο πρόβλημα της ρευστοποίησης. Οπότε για το συγκεκριμένο παράδειγμα και οι τρεις μέθοδοι θεωρούνται κατάλληλες καθώς προσεγγίζουν τα επιθυμητά αποτελέσματα 100%.

5.2.3 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΣΤΟΙΧΕΙΩΝ ΟΔΟΣΤΡΩΜΑΤΟΣ

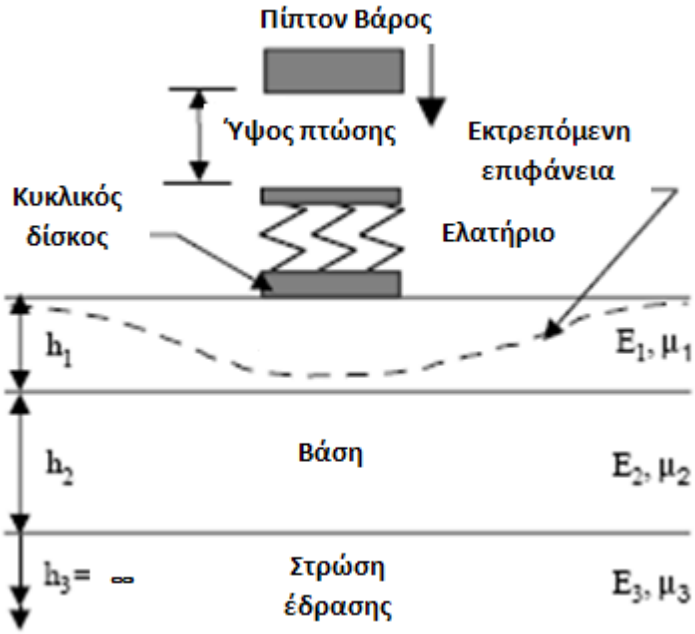
5.2.3.1 Θεωρητικές έννοιες- Ανάλυση πειράματος

Οι μεγάλοι αυτοκινητόδρομοι κατασκευάζονται κυρίως με τη μορφή εύκαμπτων οδοστρωμάτων τα οποία διαθέτουν καλύτερα υλικά στην κορυφή και κατώτερης ποιότητας υλικά στις κάτω στρώσεις. Εντούτοις, όσο καλής ποιότητας και να είναι τα υλικά που χρησιμοποιούνται σε κάποιο οδόστρωμα, οι επανειλημμένες εφαρμογές των φορτίων του οχήματος, οι καιρικές συνθήκες και κάποιοι ακόμη παράγοντες μειώνουν την λειτουργικότητά του. Για το λόγο αυτό, ένα πρόγραμμα συντήρησης θα πρέπει να συσταθεί για να αποφασίσει πότε και πού να πραγματοποιούνται εργασίες συντήρησης. Η πιο δύσκολη πτυχή είναι να καθορίσει ο μηχανικός την υπολειπόμενη διάρκεια ζωής του οδοστρώματος. Προκειμένου να καθοριστεί το υπόλοιπο ζωής, θα πρέπει να εξεταστούν οι ιδιότητες του οδοστρώματος σε κάθε στρώση του, δηλαδή το μέτρο ελαστικότητας, ο λόγος του Poisson, και το πάχος της στρώσης. Στις μέρες μας, υπάρχουν πολλές αξιόπιστες μέθοδοι για την αξιολόγηση, το σχεδιασμό, τη διαχείριση και τη συντήρηση των οδοστρωμάτων. Οι μη καταστρεπτικές μέθοδοι αποτελούν την πιο συνήθη διαδικασία για τη συλλογή επί τόπου στοιχείων, τα οποία έπειτα από κατάλληλη επεξεργασία μας δίνουν τα αποτελέσματα που επιθυμούμε. Είναι σημαντικό ότι οι μη καταστρεπτικές μέθοδοι χρησιμοποιούνται σε εύκαμπτα οδοστρώματα προκειμένου να συμβάλουν στην αποφυγή της καταστροφής τους.

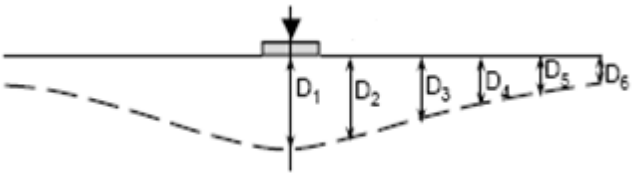
Στο πείραμα που πραγματοποιήθηκε για την εργασία αυτή, χρησιμοποιήθηκε το Παραμορφωσιμόμετρο Πίπτοντος Βάρους (Falling Weight Deflectometer), το οποίο αποτελεί ένα από τα πιο εξελιγμένα συστήματα εκτίμησης της φέρουσας ικανότητας ενός οδοστρώματος. Το FWD αποτελεί μία μη καταστρεπτική πειραματική μέθοδο (NDT) και χρησιμοποιείται για την καταγραφή στοιχείων του οδοστρώματος με σκοπό την εκτίμηση των μηχανικών χαρακτηριστικών του (Λοΐζος, 2008).

Η λειτουργία του FWD στηρίζεται στην κάθετη πτώση ενός βάρους που αφήνεται να πέσει υπό την επίδραση της βαρύτητας. Το βάρος αυτό πέφτοντας προσκρούει σε μια ειδικά σχεδιασμένη επιφάνεια και δημιουργεί μια παλμική φόρτιση. Η φόρτιση μεταβιβάζεται σε έναν κυκλικό δίσκο που βρίσκεται σε επαφή με το οδόστρωμα. Η διάρκεια της φορτίσεως είναι συνήθως 25-30 msec ή 45-60 msec ανάλογα με τη συσκευή και το υλικό που επιβάλλεται. Το μεγαλύτερο πλεονέκτημα της μεθόδου αυτής αποτελεί η προσομοίωση με την πραγματική φόρτιση του οδοστρώματος από τη διέλευση ενός τροχού.

Η φόρτιση που επιβάλλει το παραμορφωσιμόμετρο είναι δυναμική. Αυτό σημαίνει πως κάθε φορά σε οποιοδήποτε σημείο, μόνο ένα μέρος του υλικού επηρεάζεται από τη φόρτιση (Ullitz, 1998). Κατά τη διάρκεια της φόρτισης που προκαλεί το FWD στο οδόστρωμα καταγράφονται οι μέγιστες ελαστικές υποχωρήσεις (Di) κάτω από το φορτίο και σε ορισμένες αποστάσεις από αυτό. Στις παρακάτω εικόνες απεικονίζεται μία σχηματική περιγραφή της λειτουργίας του FWD καθώς και οι εν λόγω μετατοπίσεις.



Εικόνα 5.5: Μέτρηση με το FWD



Εικόνα 5.6: Καταγραφή ελαστικών υποχωρήσεων

Παράλληλα, πραγματοποιήθηκαν συστηματικές μετρήσεις της θερμοκρασίας, καθώς αυτή επηρεάζει την ανάπτυξη των παραμενουσών παραμορφώσεων. Συγκεκριμένα μετρήθηκε η θερμοκρασία στην επιφάνεια του οδοστρώματος με χρήση υπέρυθρων, η οποία κρίνεται αναγκαία λόγω της σημαντικής επίδρασης της θερμοκρασίας στα μηχανικά χαρακτηριστικά των υλικών των στρώσεων ενός οδοστρώματος.

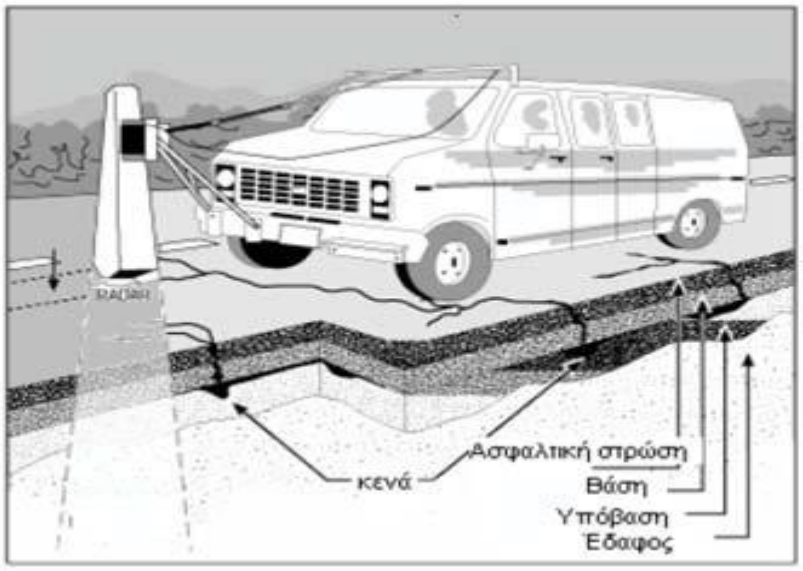
Ο βασικός στόχος της μη καταστρεπτικής μεθόδου αξιολόγησης μέσω των μετρήσεων FWD είναι να εκτιμηθούν οι χαρακτηριστικές παράμετροι των στρώσεων, όταν δίνεται η κατανομή των παραμορφώσεων. Αυτό ονομάζεται backcalculation πρόβλημα και συνήθως είναι αρκετά πολύπλοκο. Οι μετατοπίσεις που λαμβάνονται από την FWD χρησιμοποιούνται για την εκτίμηση των οι ιδιοτήτων των στρώσεων, οι οποίες είναι όπως αναφέρθηκε και παραπάνω το μέτρο ελαστικότητας, ο λόγος Poisson και τα πάχη. Δυστυχώς, η μέθοδος backcalculation δε δίνει τόσο ακριβή στοιχεία για τις παραμέτρους που μας ενδιαφέρουν παρά την ύπαρξη πολλών backcalculation προγραμμάτων που χρησιμοποιούν διαφορετικές διαδικασίες και αλγόριθμους.

Τα πάχη των στρώσεων που αποτελούν ένα οδόστρωμα είναι μια από τις σημαντικότερες παραμέτρους που χρησιμοποιούνται στο σύστημα διαχείρισης οδοστρωμάτων (PMS). Τα δεδομένα για τα πάχη χρησιμοποιούνται για την αξιολόγηση της κατάστασης ενός οδοστρώματος, για τις προβλέψεις της συμπεριφοράς του, για την επιλογή της κατάλληλης στρατηγικής συντήρησης και αποκατάστασης καθώς και ως είσοδος για την επικάλυψη του πάχους σχεδιασμού.

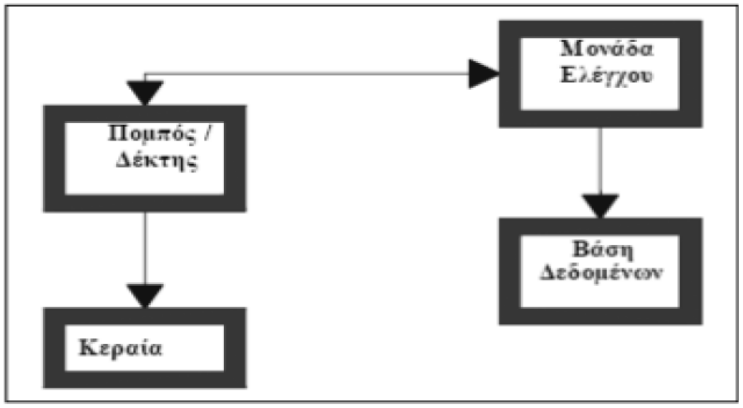
Παλαιότερα, χρησιμοποιούνταν παραδοσιακές μέθοδοι για τον υπολογισμό των παχών σε υφιστάμενα οδοστρώματα, όπως για παράδειγμα η γραμμική ελαστική θεωρία ή το ισοδύναμο πάχος στρώσεων. Όμως με αυτές τις μεθόδους οι τιμές που προέκυπταν δεν ήταν και τόσο ρεαλιστικές. Οι προσεγγίσεις αυτές επικεντρώνονταν σε στατιστικές φόρμουλες οι οποίες βασίζονταν κυρίως σε αναλύσεις παλινδρόμησης για να προβλέψουν την απόδοση ενός οδοστρώματος και ίσχυαν μόνο για συγκεκριμένες συνθήκες. Μια άλλη μέθοδος είναι η χρήση μη καταστρεπτικών τεχνικών όπως για παράδειγμα το γεωραντάρ, η οποία όμως είναι δαπανηρή και απαιτεί άτομα με ειδικές γνώσεις για την ερμηνεία των δεδομένων. Έτσι, σήμερα χρησιμοποιείται κυρίως η μέθοδος backcalculation η οποία όμως είναι μια επίπονη διαδικασία, καθώς απαιτεί υψηλό βαθμό ικανοτήτων και η ποιότητα των αποτελεσμάτων εξαρτάται από το άτομο που χρησιμοποιεί τη μέθοδο. Ένα βασικό μειονέκτημα είναι η έλλειψη μιας συγκεκριμένης αντιμετώπισης για όλα τα δεδομένα που προέρχονται από τις μετρήσεις με τη διαδικασία του FWD. Η συγκεκριμένη μέθοδος απαιτεί τη θεώρηση ενός μέγιστου ορίου ακρίβειας ενώ τα πιο ρεαλιστικά μοντέλα μειώνουν τον αριθμό των συστηματικών σφαλμάτων.

Στα πλαίσια του πειράματος που διεξήχθη για την εργασία αυτή, πραγματοποιήθηκαν μετρήσεις προκειμένου να προσδιοριστούν τα πάχη των επιμέρους στρώσεων του οδοστρώματος. Ο προσδιορισμός της στρωματογραφίας του οδοστρώματος πραγματοποιήθηκε μέσω ενός συστήματος γεωραντάρ (Ground Penetration Radar: GPR). Το σύστημα GPR είναι ένα υψηλής τεχνολογίας μη καταστρεπτικό σύστημα επιτόπου ελέγχου (NDT), το οποίο καταγράφει συνεχόμενα

τη στρωματογραφία του οδοστρώματος. Η καταγραφή αυτή οδηγεί στην εκτίμηση των παχών των επιμέρους στρώσεων ανά πολύ μικρές αποστάσεις με σημαντική ακρίβεια. Σημαντικό πλεονέκτημα του συγκεκριμένου συστήματος είναι η δυνατότητα λειτουργίας σε συνήθεις ταχύτητες κυκλοφορίας. Έτσι, δεν υφίσταται διακοπή της κυκλοφορίας των οχημάτων τη στιγμή που πραγματοποιείται η μέτρηση και δεν χρειάζεται η λήψη ειδικών μέτρων για την οδική ασφάλεια. Η θεμελιώδης αρχή λειτουργίας του βασίζεται στην ηλεκτρομαγνητική θεωρία, καθώς η αναγνώριση του εσωτερικού του εδάφους γίνεται με ηλεκτρομαγνητικά κύματα. Το σκαρίφημα της εικόνας παρακάτω απεικονίζει τη διάταξη του συστήματος βάσει του οποίου γίνεται η διερεύνηση του οδοστρώματος. Ένα τυπικό σύστημα γεωραντάρ αποτελείται από τη διάταξη πομπού/δέκτη, την κεραία ή το συνδυασμό κεραιών, τη μονάδα ελέγχου και τη βάση δεδομένων όπου εισάγονται τα στοιχεία συλλογής. Αρχικά το GPR τοποθετείται κοντά στην υπό μελέτη επιφάνεια και μακριά από αντικείμενα που θα μπορούσαν να επηρεάσουν το ηλεκτρομαγνητικό σήμα (π.χ. μεταλλικές επιφάνειες). Τα ραδιοκύματα κατευθύνονται στο έδαφος. Όταν συναντούν υλικό με διαφορετικά ηλεκτρομαγνητικά χαρακτηριστικά, αφενός μέρος τους ανακλάται πίσω στο δέκτη και αφετέρου μεταβάλλεται η ταχύτητα (v) που τα διαπερνά. Το σύστημα καταγράφει το χρόνο που χρειάζονται τα ηλεκτρομαγνητικά κύματα να διαπεράσουν ένα συγκεκριμένο υλικό και να επιστρέψουν στο δέκτη. Επίσης, σημαντική είναι και η συμβολή της ταχύτητας των κυμάτων μέσα στα διάφορα υλικά. Η ταχύτητα αυτή εξαρτάται από τη διηλεκτρική σταθερά των υλικών. Ο χρόνος σε συνδυασμό με την ταχύτητα χρησιμεύουν για τον υπολογισμό της αντίστοιχης απόστασης που διανύουν τα κύματα.



Εικόνα 5.7: Σκαρίφημα της διάταξης συστήματος GPR σε οδούς



Εικόνα 5.8: Μονάδες λειτουργίας ενός συστήματος GPR

Στο πείραμα που έγινε για την εν λόγω εργασία μελετήθηκε το οδόστρωμα του αυτοκινητόδρομου Κόρινθος-Ελευσίνα και εκτιμήθηκαν με τις διαδικασίες που αναφέρθηκαν παραπάνω τα πάχη των στρώσεων αλλά και οι βυθίσεις σε πολλά σημεία του δρόμου. Το φορτίο που ασκήθηκε ήταν 80, 60 και 40KN και οι παραμορφώσεις μετρήθηκαν σε 9 σημεία της επιφάνειας του οδοστρώματος. Στους παρακάτω πίνακες παρουσιάζονται κάποιες ενδεικτικές τιμές για τις βυθίσεις (σε mm) στις 9 θέσεις του οδοστρώματος για φόρτιση 80,60 και 40KN.

α. Φορτίο 80KN

Χ.Θ.	D1	D2	D3	D4	D5	D6	D7	D8	D9
24600	189	102	82	50	25	24	19	16	12
24800	164	73	55	32	16	15	11	12	8
25000	227	119	89	42	15	17	13	12	9
25200	216	73	50	31	14	13	9	9	5
25400	180	77	56	32	17	13	9	8	5
25600	155	100	92	62	32	25	17	17	11
28600	181	89	76	61	38	29	21	20	14
28800	195	111	98	71	41	27	15	11	6
29000	137	45	33	25	15	17	13	12	8
29200	179	85	75	62	41	36	26	22	16
32800	255	156	129	77	41	28	19	16	12
33000	322	198	150	82	49	33	24	20	15
33600	321	181	132	72	42	33	24	20	15
34400	258	127	93	60	36	28	20	17	14

β. Φορτίο 60KN

Χ.Θ.	D1	D2	D3	D4	D5	D6	D7	D8	D9
34400	199	96	71	47	27	22	15	13	11
35000	234	123	89	59	35	30	22	18	15
39800	104	82	77	52	27	22	15	14	12
40400	222	160	145	89	51	36	26	23	21
40600	268	201	172	105	54	51	41	36	31
41800	109	54	49	44	33	30	26	24	13
42000	101	38	33	31	17	21	19	19	16
45800	70	37	35	28	14	9	5	4	3
54000	139	76	66	45	22	21	15	14	12
54200	150	88	77	50	28	22	16	15	12
55000	191	118	97	60	29	28	20	18	15
60000	81	49	44	33	25	19	14	11	7
68000	104	59	54	50	36	35	29	27	22
70000	145	61	50	40	20	22	16	15	9
72000	114	51	46	43	29	28	24	24	18
74000	192	120	11	89	55	51	39	33	24

γ. Φορτίο 40KN

Χ.Θ.	D1	D2	D3	D4	D5	D6	D7	D8	D9
34400	145	68	53	34	14	13	8	9	7
35000	168	83	59	39	20	18	12	13	9
39800	160	112	94	44	12	9	4	7	4
40400	174	124	111	65	31	22	15	15	12
40600	201	145	125	72	29	32	25	24	22
41800	76	36	33	31	17	20	16	17	15
42000	70	24	23	22	7	14	11	14	9
45800	50	27	27	21	6	5	2	3	2
54000	86	45	40	27	9	12	9	8	6
54200	96	55	49	33	18	16	12	11	8
55000	119	73	61	38	21	18	13	12	10
60000	56	35	32	24	15	13	10	8	6
68000	70	37	35	34	19	21	17	17	12
70000	100	42	36	28	13	14	10	10	5
72000	82	35	32	31	15	19	15	15	11
74000	133	80	76	60	36	31	23	22	14
78000	149	62	35	13	3	11	9	9	6

5.2.3.2 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Στο συγκεκριμένο κεφάλαιο προτείνεται ως μια λύση η χρήση τεχνητού νευρωνικού δικτύου (ANN) το οποίο βασίζεται στις αρχές της μεθόδου backcalculation. Προκειμένου να αναπτυχθεί ένα τέτοιο μοντέλο είναι απαραίτητο πρώτα από όλα να υπάρχει μια ισχυρή βάση δεδομένων για να εκπαιδευτεί και να δοκιμαστεί το δίκτυο. Στο κεφάλαιο αυτό, λοιπόν, κατασκευάζεται ένα νευρωνικό δίκτυο το οποίο έχει ως εισόδους τις βυθίσεις που παρουσιάστηκαν στον πίνακα και ως εξόδους τα πάχη των στρώσεων του οδοστρώματος. Εξαιτίας του αρκετά μεγάλου δείγματος που διατίθεται για την συγκεκριμένη εφαρμογή χρειάστηκαν πολλές προσπάθειες μέχρι να βρεθεί το βέλτιστο δίκτυο που να δίνει αποτελέσματα πολύ κοντά στα αναμενόμενα. Ένα μειονέκτημα του δείγματος ήταν η διακύμανση ανάμεσα στα διάφορα στοιχεία, κάτι το οποίο όμως παρουσίασε μεγάλη βελτίωση μετά την κανονικοποίηση των δεδομένων. Σίγουρα, όμως, όσες δοκιμές και αν έγιναν για να βρεθεί το κατάλληλο δίκτυο, η όλη διαδικασία είναι αναμφίβολα λιγότερο χρονοβόρα από το να εκτιμούμε τα πάχη με πειραματικές μεθόδους και επί τόπου μετρήσεις. Παρακάτω παρουσιάζονται τα χαρακτηριστικά καθώς και οι μορφές των δικτύων τα οποία έδειξαν την καλύτερη συμπεριφορά.

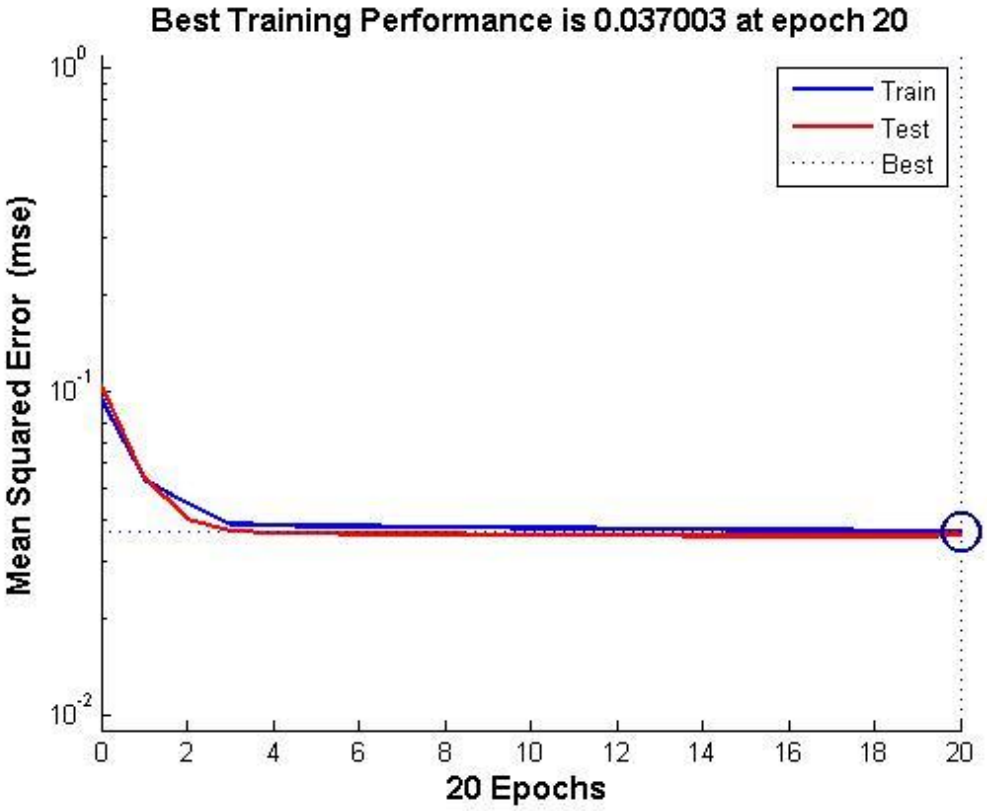
α. Φορτίο 80KN

Κρυμμένα επίπεδα	2
Κόμβοι σε κάθε επίπεδο	50
Συνάρτηση μεταφοράς	Logsig
Data division function	Dividerand
Training function	Trainlm
Performance function	Mse
Derivative function	Defaultderiv
Maximum number of training epochs	20
Maximum training time	Infinity
Minimum performance value	0
Minimum gradient magnitude	1e-05
Maximum Mu	1e+10

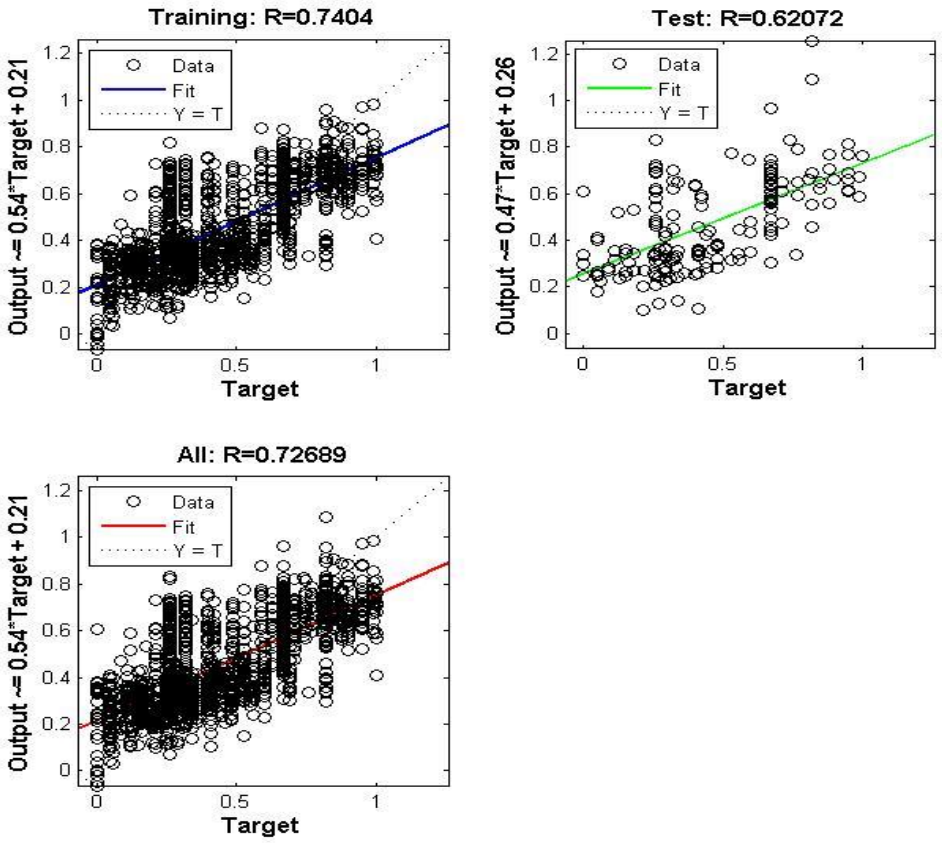
Η τιμή στο μέγιστο αριθμό κύκλων που θα κάνει το δίκτυο τέθηκε τόσο μικρή γιατί πέρα από τους 20 κύκλους παρατηρήθηκε overfitting δηλαδή υπερβολική εκπαίδευση και το σφάλμα στο testing set αυξανόταν. Το ίδιο συνέβη και για τις υπόλοιπες τιμές φόρτισης που παρουσιάζονται παρακάτω.

Για κάθε φόρτιση αποτυπώνονται και τα αντίστοιχα διαγράμματα τα οποία βοηθούν στην αξιολόγηση της μεθόδου και στη σύγκριση με τα υπόλοιπα δυο

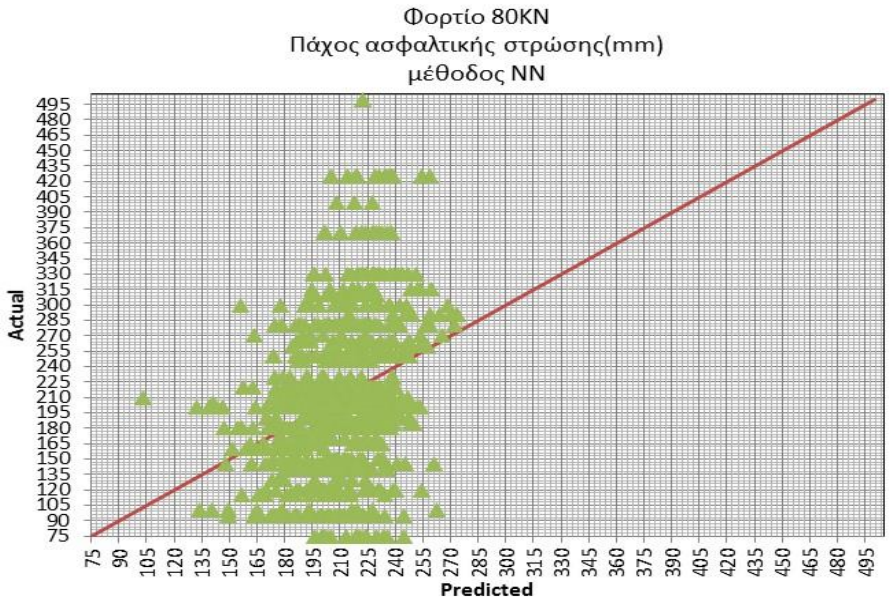
μοντέλα που θα παρουσιαστούν στις υπόλοιπες ενότητες. Για την εξαγωγή των τελικών συμπερασμάτων κύριο κριτήριο είναι η απόκλιση των προβλεπόμενων τιμών από τις πραγματικές, γι' αυτό και τα διαγράμματα που κατασκευάστηκαν για κάθε αλγόριθμο έχουν μορφή τέτοια ώστε να απεικονίζουν την απόκλιση αυτή.



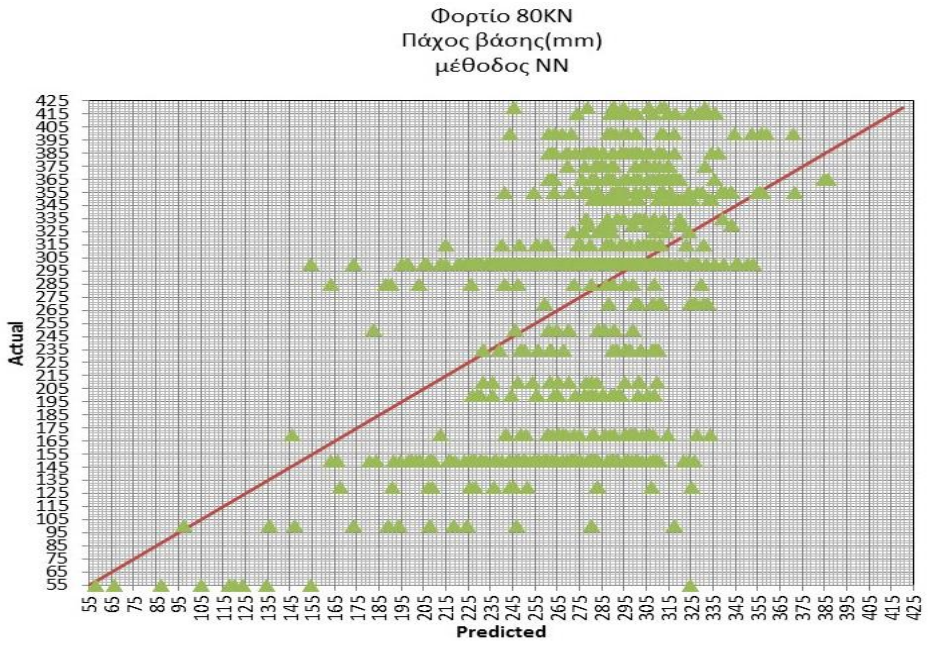
Σχήμα 5.11 Σφάλμα τελικού δικτύου στο training και testing set



Σχήμα 5.12 Παλινδρόμηση τελικού δικτύου



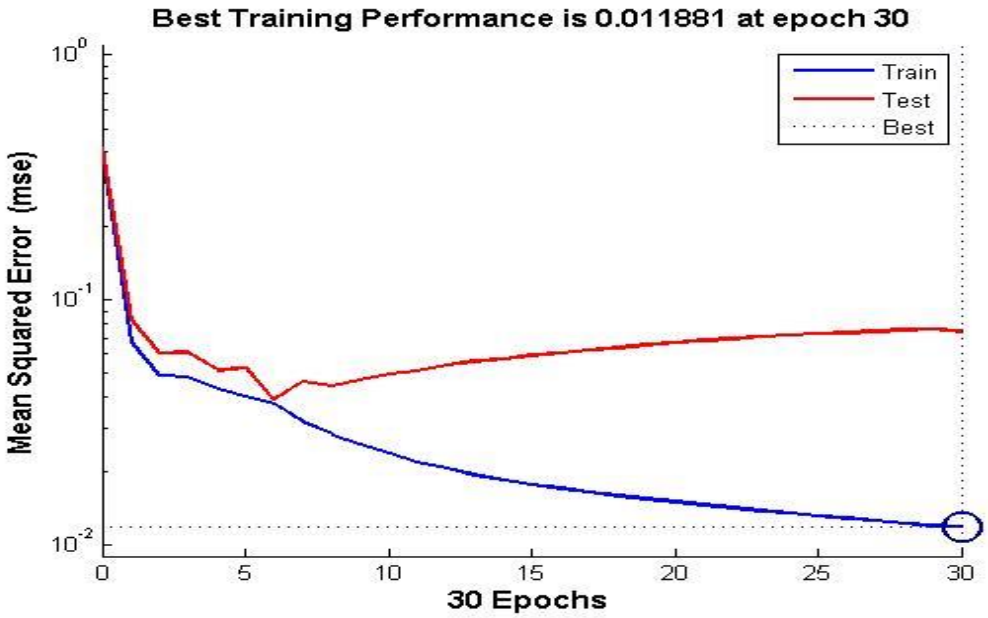
Σχήμα 5.13 Απεικόνιση των αποτελεσμάτων του NN για το πάχος ασφαλτικής στρώσης (80KN)



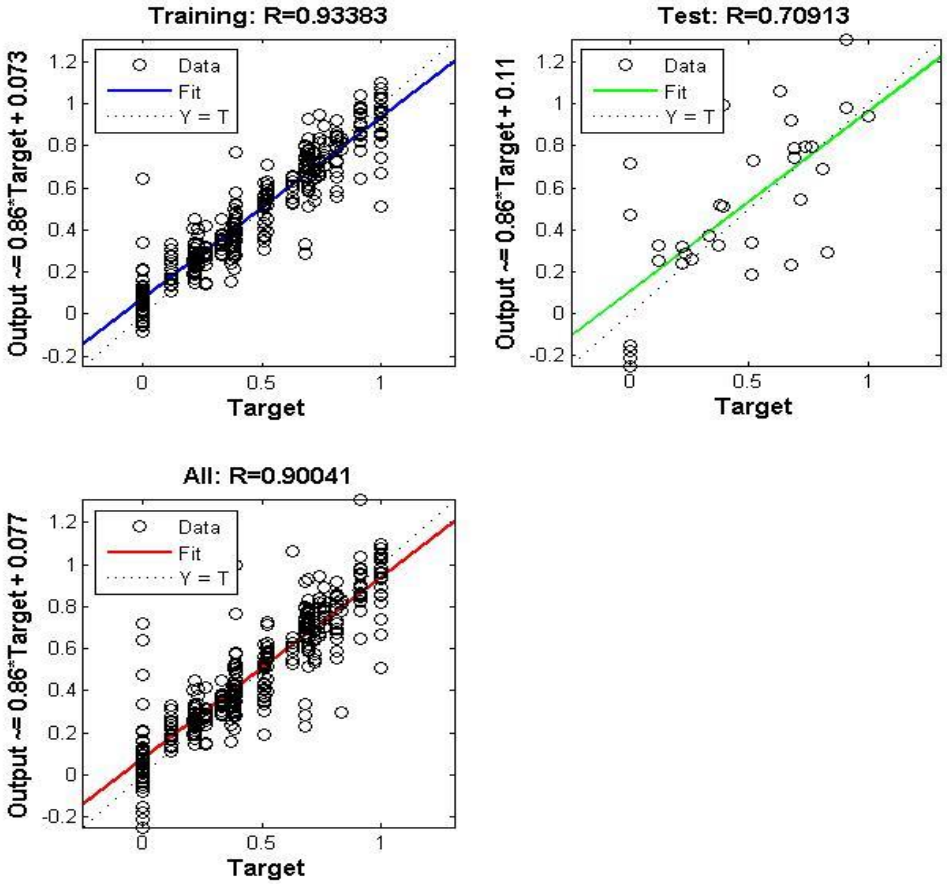
Σχήμα 5.14 Απεικόνιση των αποτελεσμάτων του NN για το πάχος βάσης (80KN)

β. Φορτίο 60KN

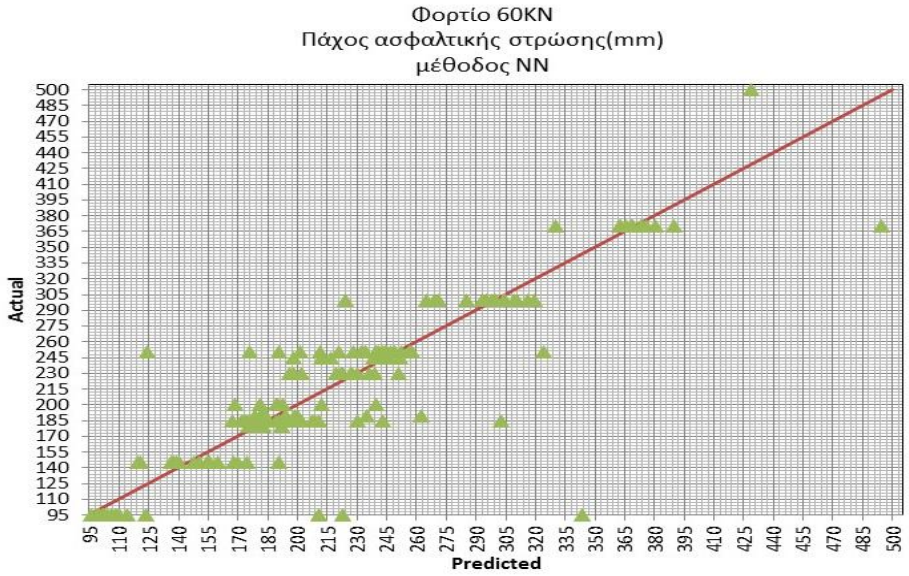
Κρυμμένα επίπεδα	2
Κόμβοι σε κάθε επίπεδο	20
Συνάρτηση μεταφοράς	Logsig
Data division function	Dividerand
Training function	Trainlm
Performance function	Mse
Derivative function	Defaultderiv
Maximum number of training epochs	30
Maximum training time	Infinity
Minimum performance value	0
Minimum gradient magnitude	1e-05
Maximum Mu	1e+10



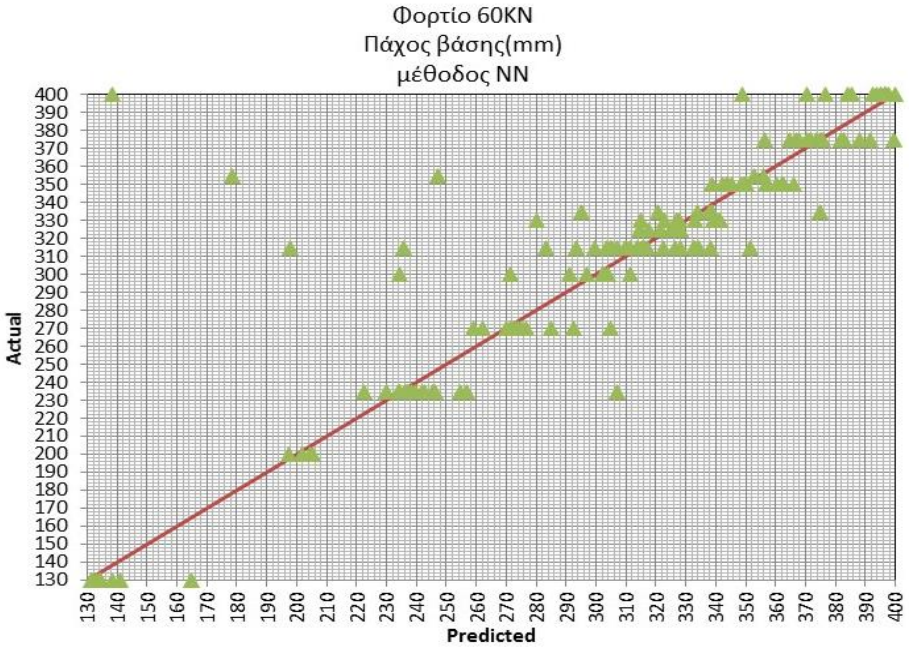
Σχήμα 5.15 Σφάλμα τελικού δικτύου στο training και testing set



Σχήμα 5.16 Παλινδρόμηση τελικού δικτύου



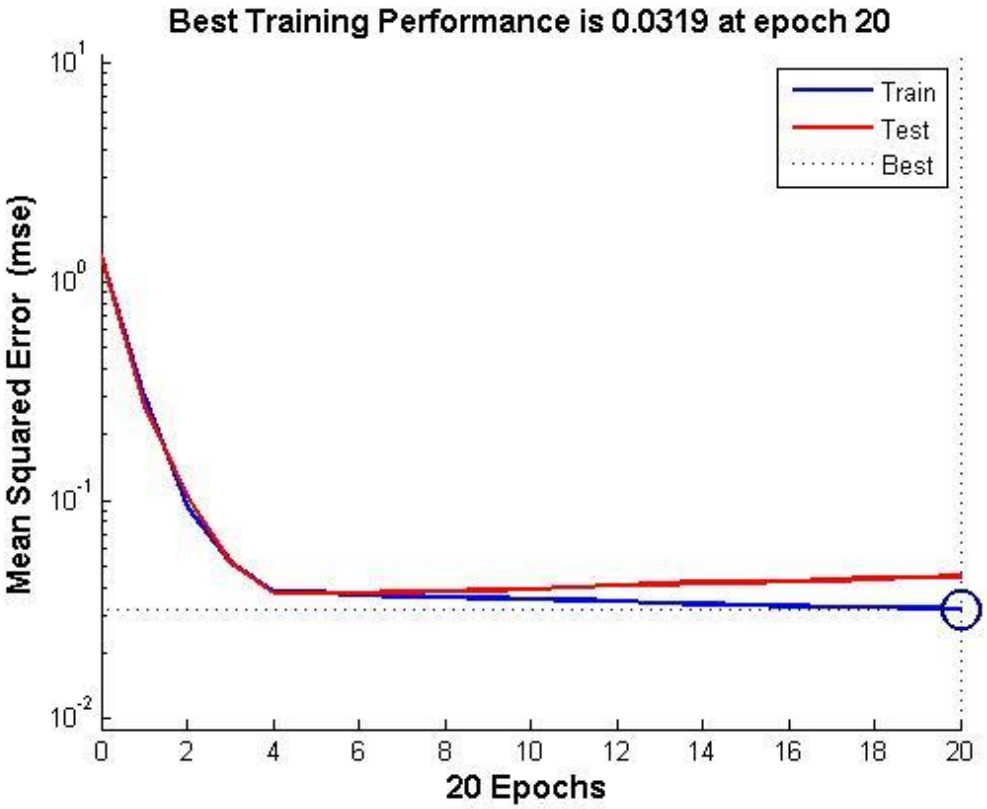
Σχήμα 5.17 Απεικόνιση των αποτελεσμάτων του NN για το πάχος ασφαλτικής στρώσης (60KN)



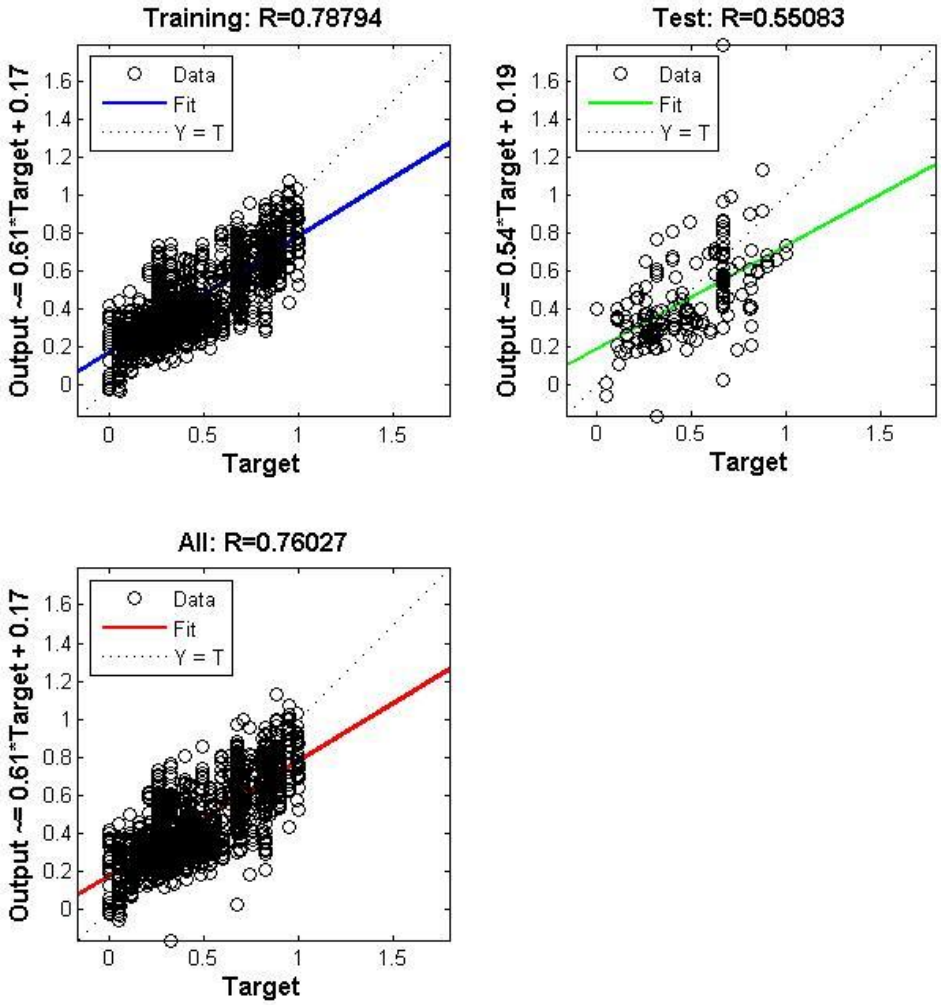
Σχήμα 5.18 Απεικόνιση των αποτελεσμάτων του NN για το πάχος βάσης (60KN)

γ. Φορτίο 40KN

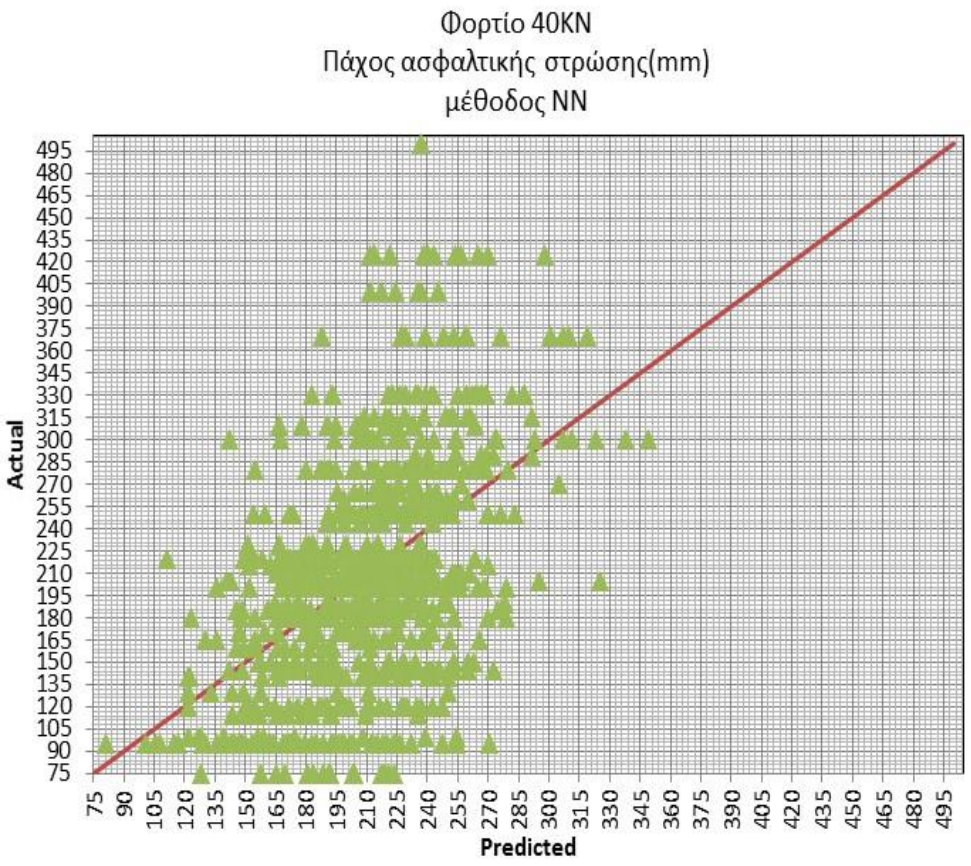
Κρυμμένα επίπεδα	2
Κόμβοι σε κάθε επίπεδο	20
Συνάρτηση μεταφοράς	Logsig
Data division function	Dividerand
Training function	Trainlm
Performance function	Mse
Derivative function	Defaultderiv
Maximum number of training epochs	20
Maximum training time	Infinity
Minimum performance value	0
Minimum gradient magnitude	1e-05
Maximum Mu	1e+10



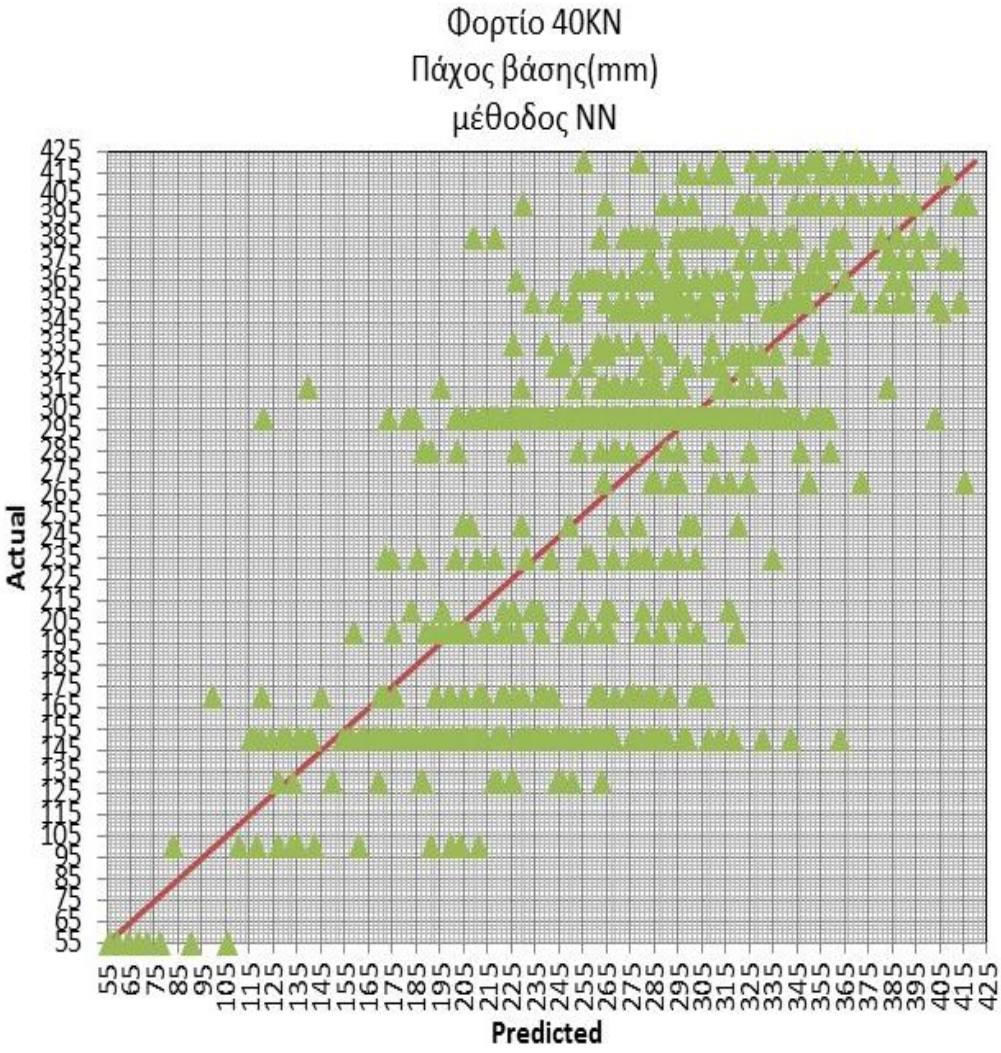
Σχήμα 5.19 Σφάλμα τελικού δικτύου στο training και testing set



Σχήμα 5.20 Παλινδρόμηση τελικού δικτύου



Σχήμα 5.21 Απεικόνιση των αποτελεσμάτων του NN για το πάχος ασφαλτικής στρώσης (40KN)



Σχήμα 5.22 Απεικόνιση των αποτελεσμάτων του NN για το πάχος βάσης (40KN)

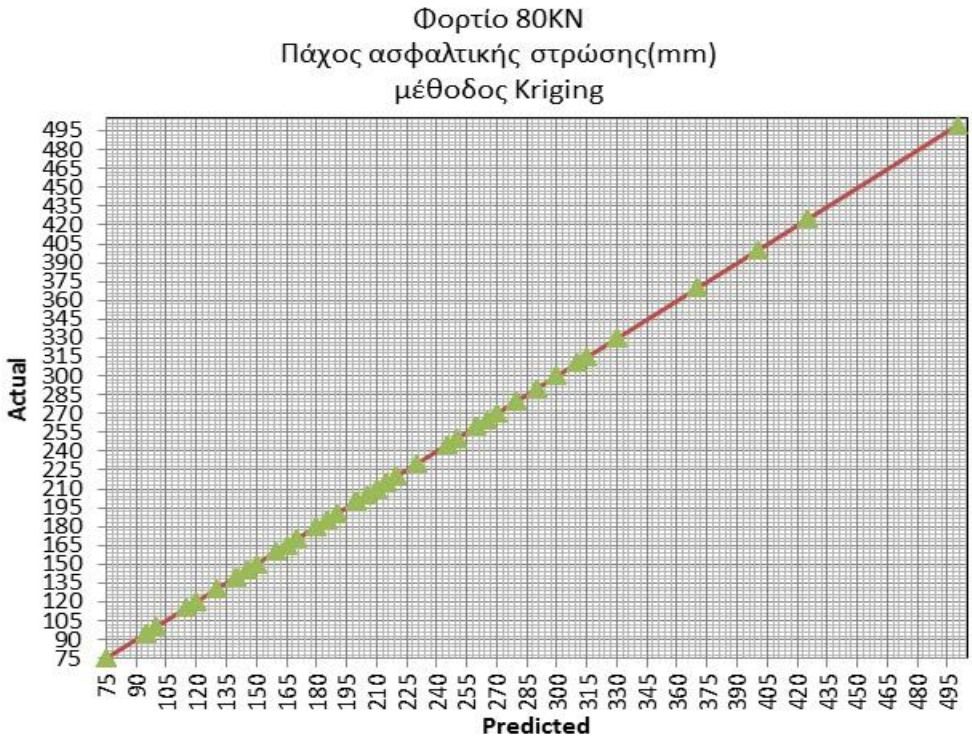
5.2.3.3 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ KRIGING

Ως δεύτερη μέθοδος για τον προσδιορισμό των στοιχείων του οδοστρώματος προτείνεται η μέθοδος Kriging, η οποία παρουσιάζει εντυπωσιακά καλύτερα αποτελέσματα από την πρώτη μέθοδο και δε δείχνει κάποια δυσλειτουργία ακόμη και αν στο πείραμα συμπεριλαμβάνονται πολλά δεδομένα και με μεγάλη διακύμανση. Η στρατηγική που ακολουθήθηκε για την εύρεση του κατάλληλου μοντέλου σε καθεμιά από τις τρεις φορτίσεις είναι η ίδια όπως και τα χαρακτηριστικά των αλγορίθμων που τελικά εφαρμόστηκαν. Οπότε, μπορούμε να τα συνοψίσουμε σε έναν πίνακα:

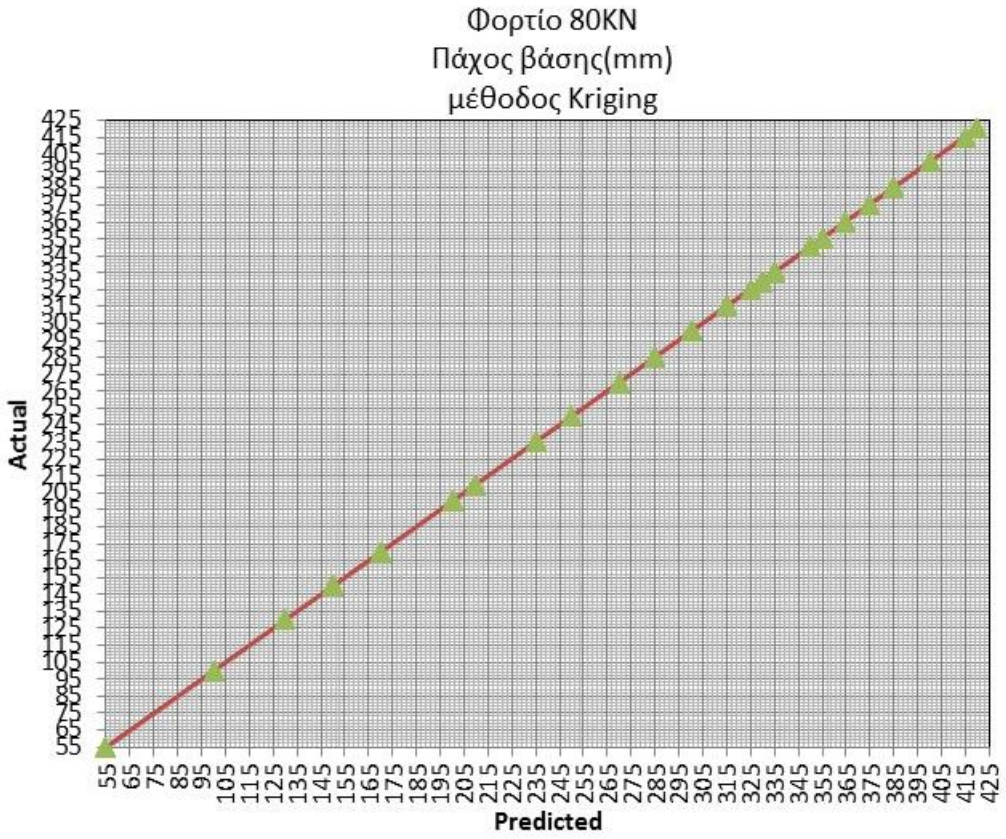
Συνάρτηση παλινδρόμησης	POLY1
Συνάρτηση συσχέτισης	GAUSS
Παράμετρος θ	10
Κάτω όριο του ϑ	1e-1
Άνω όριο του ϑ	20
Πειραματικός σχεδιασμός	Latin Hypercube

Τα αποτελέσματα κάθε αλγορίθμου παρουσιάζονται στα διαγράμματα που ακολουθούν:

α. Φορτίο 80KN

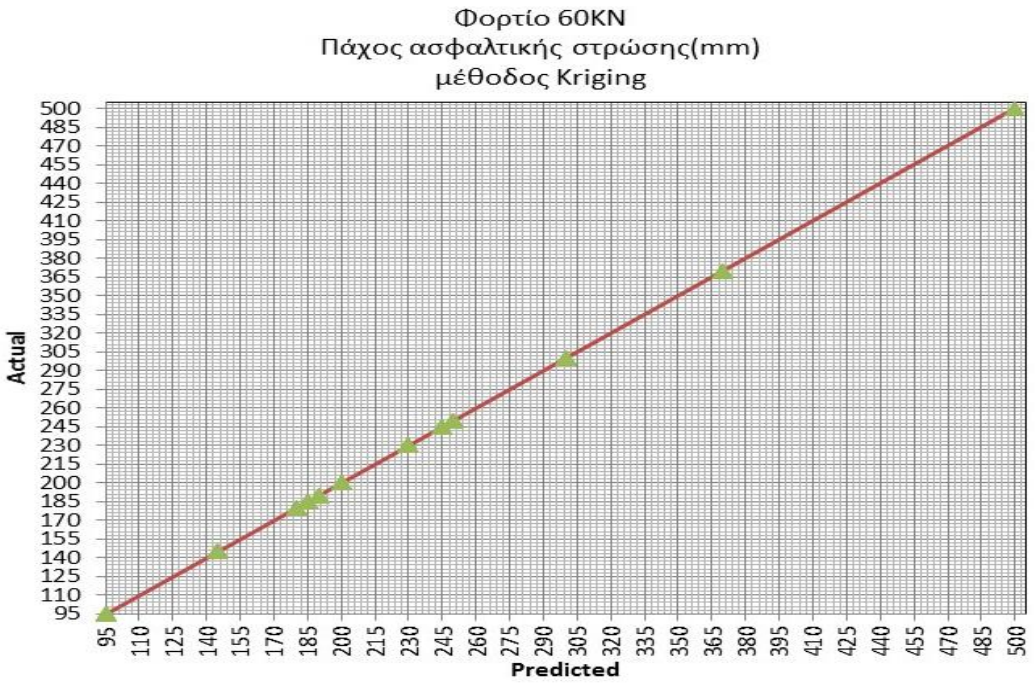


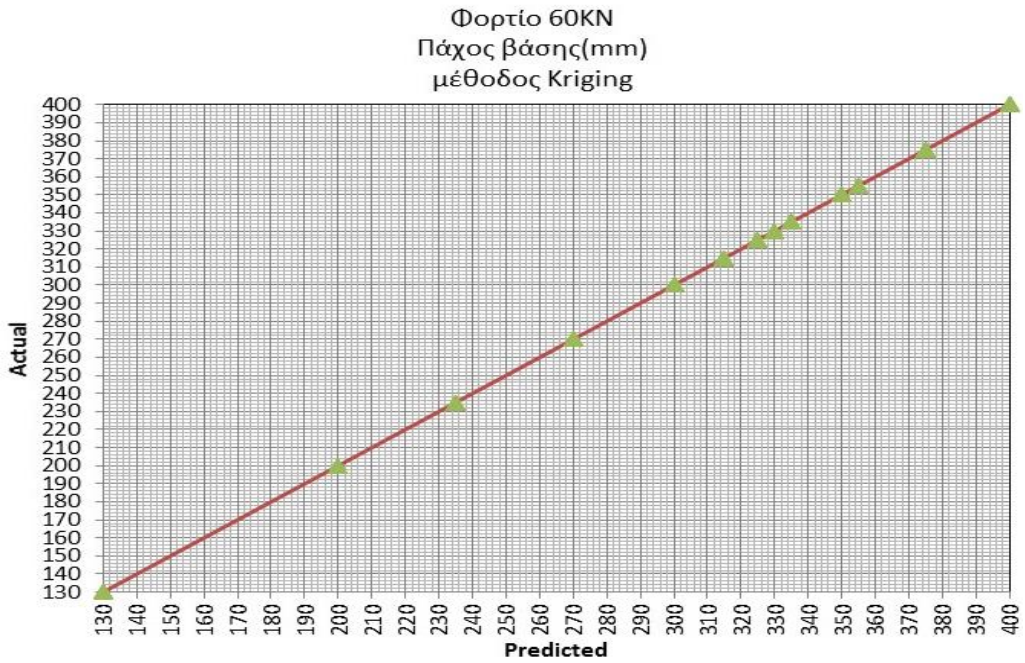
Σχήμα 5.23 Μέθοδος Kriging για το πάχος ασφαλτικής στρώσης (80KN)



Σχήμα 5.24 Μέθοδος Kriging για το πάχος βάσης (80KN)

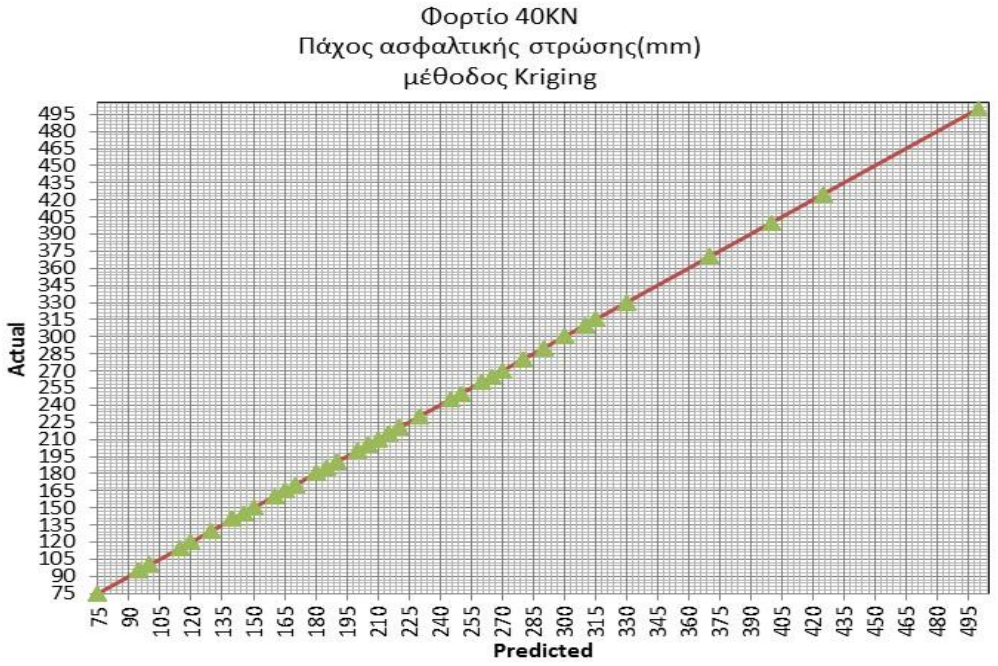
β. Φορτίο 60KN

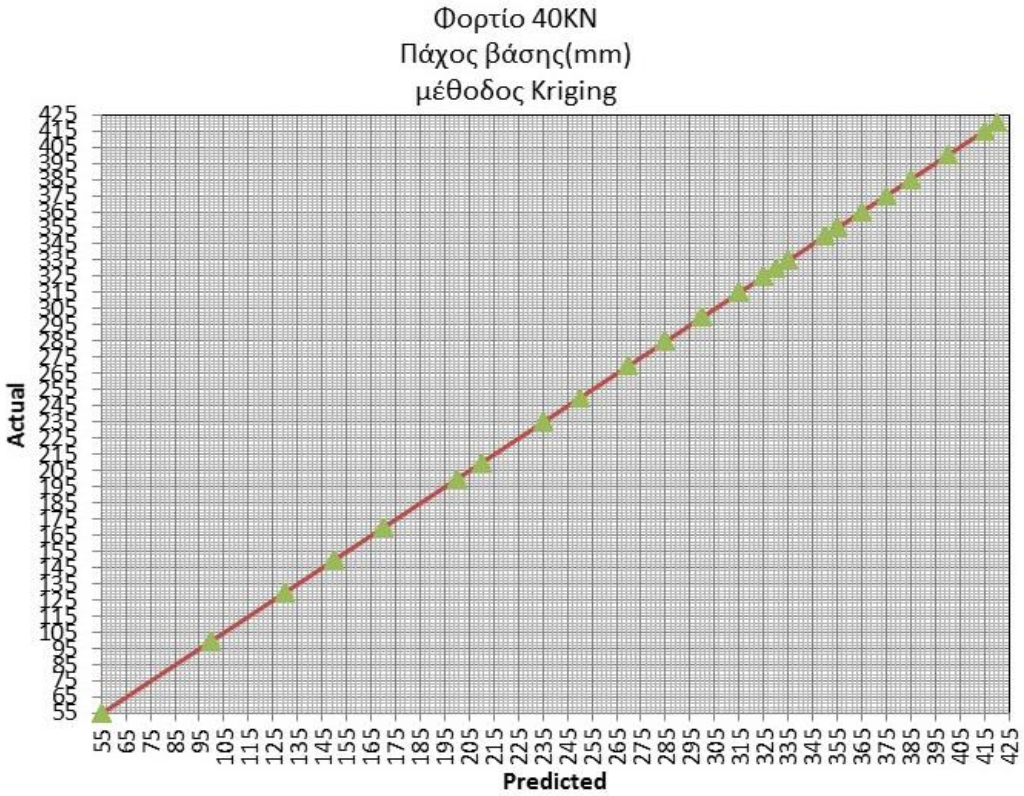




Σχήμα 5.25 Μέθοδος Kriging για πάχος ασφαλτικής στρώσης και βάσης (60KN)

γ. Φορτίο 40KN





Σχήμα 5.26 Μέθοδος Kriging για πάχος ασφαλτικής στρώσης και βάσης (40KN)

5.2.3.4 ΕΠΙΛΥΣΗ ΜΕ ΤΗ ΜΕΘΟΔΟ SVM

Ως τρίτη και τελευταία μέθοδος θεωρείται αυτή των SVM η οποία, όπως και στην περίπτωση της γερανογέφυρας, χρησιμοποιείται για να βρεθεί η κατάλληλη συνάρτηση παλινδρόμησης που αποτυπώνει τη σχέση μεταξύ των δεδομένων του προβλήματος. Μέσω αυτής, μπορούν να προβλεφθούν οι τιμές για τις εξόδους του προβλήματός μας τόσο για τις παρατηρούμενες εισόδους όσο και για νέες που θα μπορούσαν να δοθούν στο μέλλον. Τα βήματα για την κατασκευή του αλγορίθμου είναι τα γνωστά 3, όπως έχουν επισημανθεί και στις δυο προηγούμενες εφαρμογές των SVM (γερανογέφυρα, ρευστοποίηση εδάφους). Όσον αφορά στις παραμέτρους, είναι ίδιες με την πρώτη εφαρμογή καθώς και εδώ πρόκειται για παλινδρόμηση και οι τιμές που δόθηκαν σε αυτές παρουσιάζονται στους παρακάτω τρεις πίνακες ανάλογα με τη φόρτιση που ασκείται στο οδόστρωμα.

α. Φορτίο 80KN

Μορφή κλιμάκωσης	Anisotropic
Τύπος Kernel	Gausian RBF
Παράμετρος p1	0.3
Παράμετρος C	Inf
Loss function	Quadratic
Insensitivity e	0

β. Φορτίο 60KN

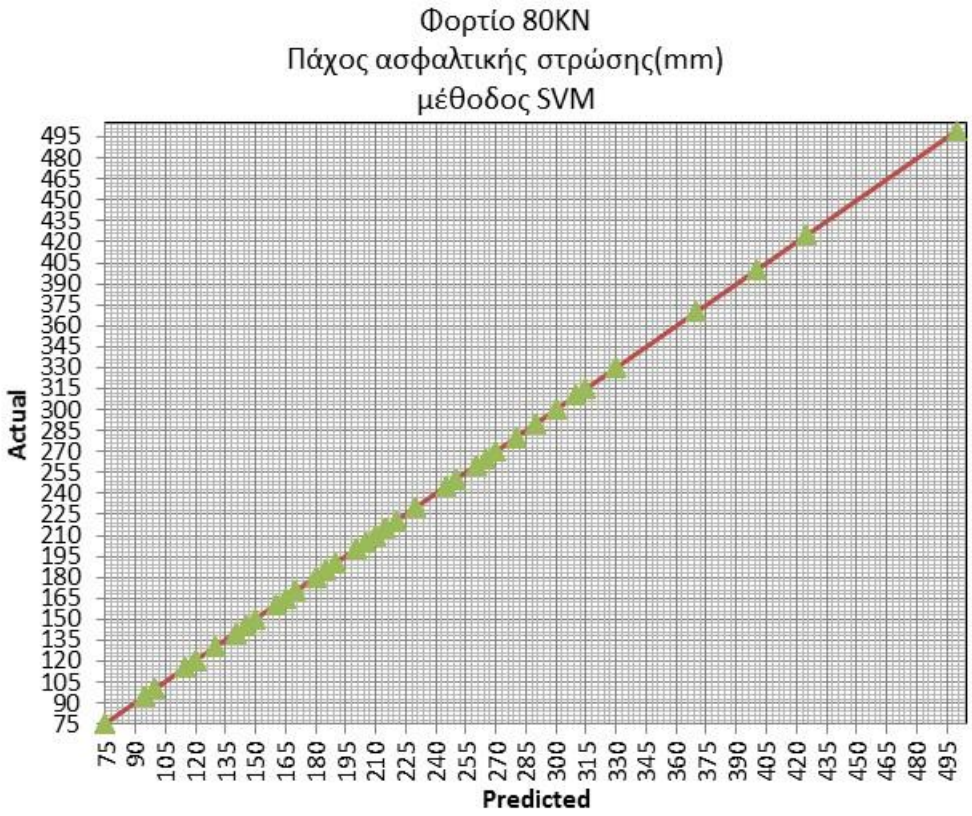
Μορφή κλιμάκωσης	Anisotropic
Τύπος Kernel	Gausian RBF
Παράμετρος p1	1
Παράμετρος C	Inf
Loss function	Quadratic
Insensitivity e	0

γ. Φορτίο 40KN

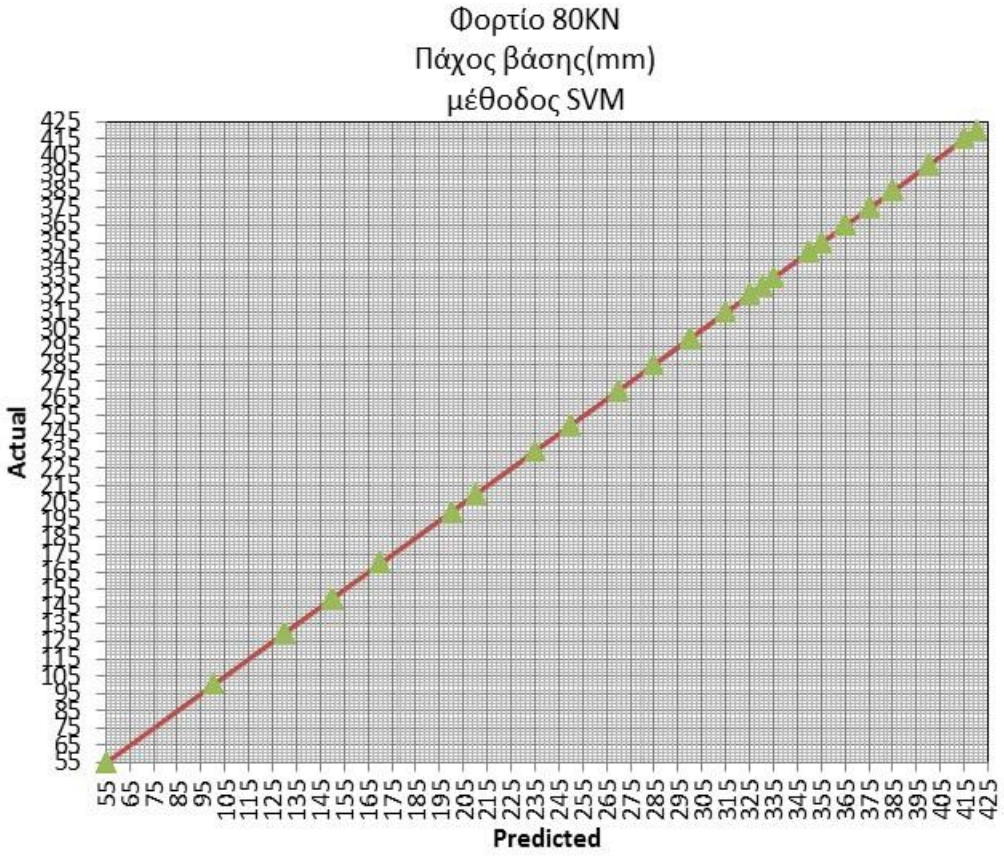
Μορφή κλιμάκωσης	Anisotropic
Τύπος Kernel	Gausian RBF
Παράμετρος p1	0.5
Παράμετρος C	Inf
Loss function	Quadratic
Insensitivity e	0

Στη συνέχεια παρουσιάζονται τα διαγράμματα για κάθε αλγόριθμο ανάλογα με τη φόρτιση και την έξοδο στην οποία αντιστοιχούν (πάχος ασφαλτικής στρώσης, πάχος βάσης). Μέσω αυτών, μπορεί κανείς να αντιληφθεί ότι η συμπεριφορά των αλγορίθμων είναι πολύ καλή, καθώς η απόκλιση από τις πραγματικές τιμές είναι σχεδόν μηδενική.

α. Φορτίο 80KN

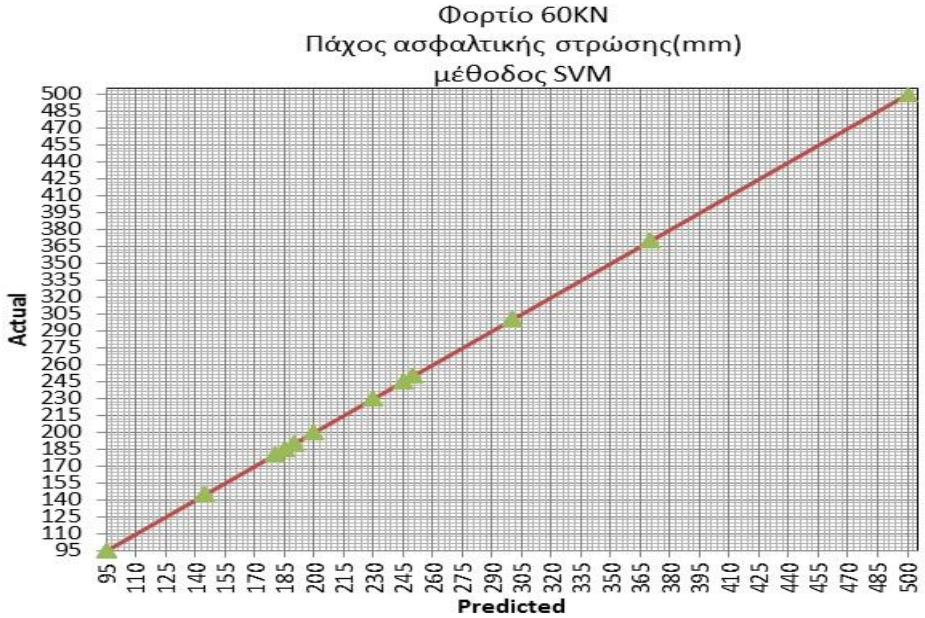


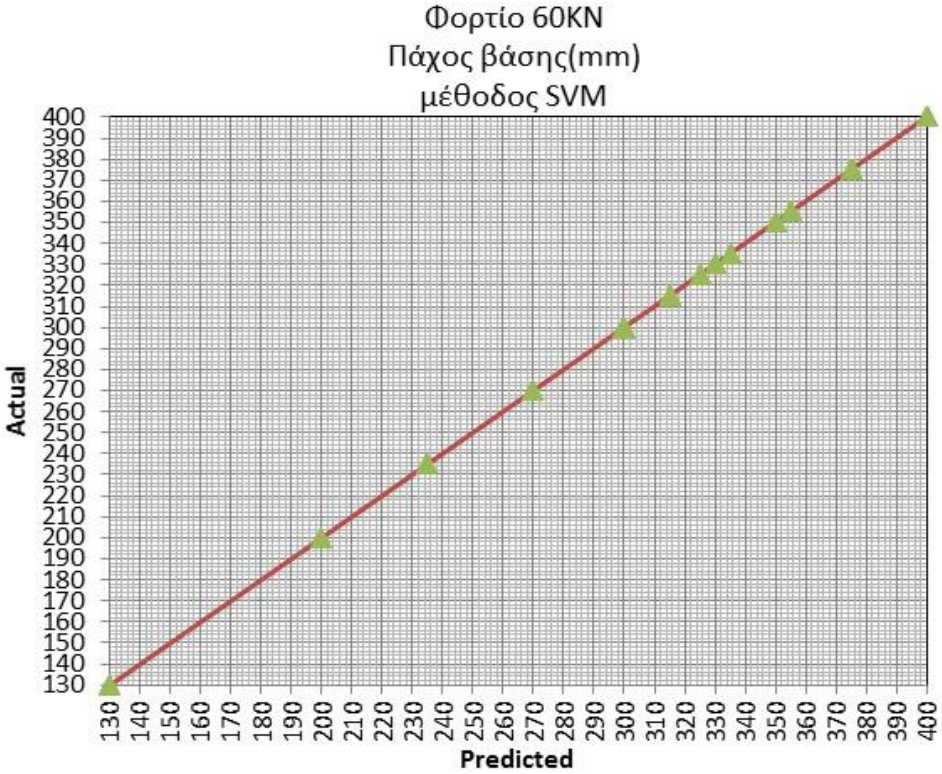
Σχήμα 5.27 Μέθοδος SVM για το πάχος ασφαλτικής στρώσης (80KN)



Σχήμα 5.28 Μέθοδος SVM για το πάχος βάσης (80KN)

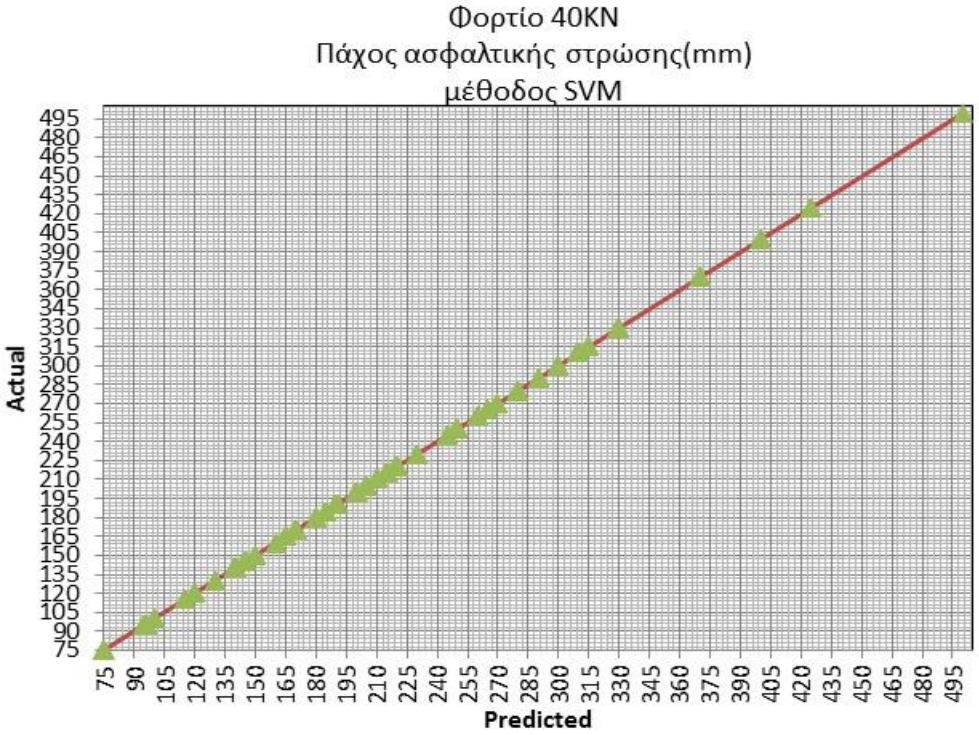
β. Φορτίο 60KN

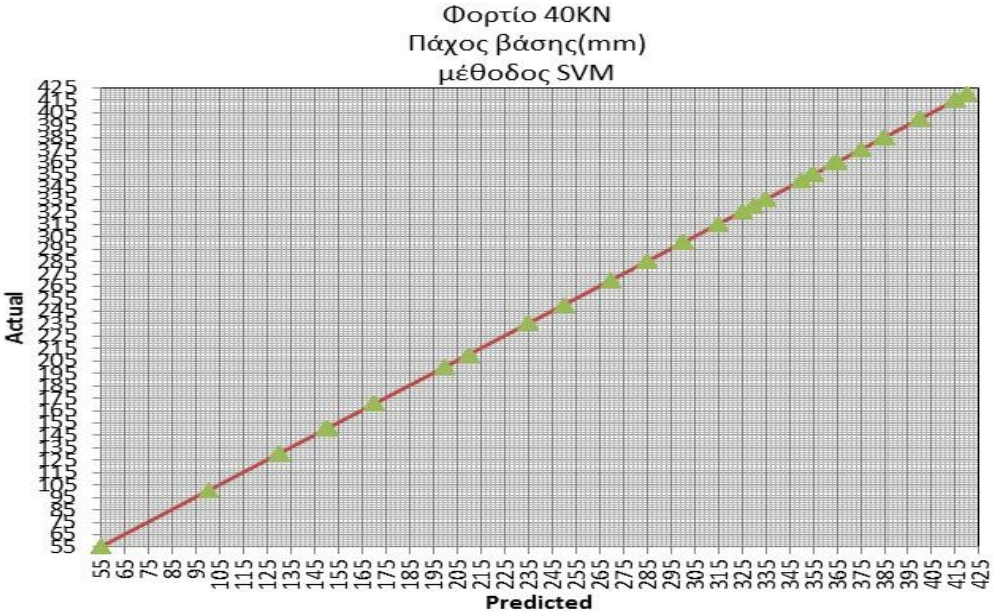




Σχήμα 5.29 Μέθοδος SVM για πάχος ασφαλτικής στρώσης και βάσης (60KN)

γ. Φορτίο 40KN





Σχήμα 5.30 Μέθοδος SVM για πάχος ασφαλτικής στρώσης και βάσης (40KN)

6

ΣΥΜΠΕΡΑΣΜΑΤΑ

Μετά την επίλυση των εφαρμογών που αναφέρθηκαν στο προηγούμενο κεφάλαιο ακολουθούν τα συμπεράσματα σχετικά με τη διαδικασία κατασκευής, το χρόνο εκτέλεσης, την ακρίβεια και την καταλληλότητα κάθε μεθόδου για κάθε πρόβλημα.

Στην εξαγωγή των συμπερασμάτων αυτών βοηθούν τα αποτελέσματα που προκύπτουν από την εκτέλεση των αλγορίθμων στο πρόγραμμα της Matlab, τα οποία στη συνέχεια αποτυπώθηκαν σε διαγράμματα για να είναι ευκολότερο να απεικονιστεί η απόκλισή τους από τις πραγματικές τιμές αλλά και η σύγκριση των μεθόδων μεταξύ τους.

Όσον αφορά στο ποια μέθοδος αποδείχτηκε καλύτερη, αυτή παρατηρήθηκε πως είναι η μέθοδος Kriging. Και αυτό γιατί, τόσο η διατύπωση των μοντέλων βάση των αλγορίθμων της μεθόδου Kriging ήταν αρκετά εύκολη αλλά και η ακρίβεια με την οποία προσέγγισε όλα τα προβλήματα ήταν άριστη. Το πλεονέκτημα το οποίο την κάνει να ξεχωρίζει από τις υπόλοιπες είναι ότι δε δείχνει ιδιαίτερη ευαισθησία στην αλλαγή των παραμέτρων, οπότε δε χρειάζεται να γίνουν πολλές προσπάθειες για να καταλήξουμε στον κατάλληλο αλγόριθμο. Επιπρόσθετα, πρόκειται για μια μέθοδο με πολύ μικρό χρόνο εκτέλεσης ανεξάρτητα με τον όγκο των δεδομένων που της δίνεται. Τέλος, τα εξαγόμενα αποτελέσματα παρουσιάζουν απόκλιση σχεδόν μηδενική από τα πραγματικά, κάτι το οποίο αποτελεί βασική προϋπόθεση για να χαρακτηριστεί μια εναλλακτική υπολογιστική μέθοδος εύστοχη.

Ως δεύτερη καλύτερη μέθοδος σε αποδοτικότητα θεωρείται αυτή των Support Vectors Machines, η οποία από άποψη ακρίβειας προβλέψεων αποδείχθηκε σχεδόν άριστη, αλλά παρουσίασε κάποια ευαισθησία στις αλλαγές των παραμέτρων. Στην περίπτωση της γερανογέφυρας, της ρευστοποίησης εδάφους και του οδοστρώματος με φόρτιση 60KN, η κατασκευή των αλγορίθμων ήταν αρκετά

εύκολη και τα αποτελέσματα ήταν τα επιθυμητά. Αντίθετα, στο παράδειγμα του οδοστρώματος με φόρτιση 40 και 80KN, η προσέγγιση του προβλήματος ήταν ελαφρώς δυσκολότερη καθώς οι αλγόριθμοι δεν έδειχναν καλή συμπεριφορά για τις προτεινόμενες (default) τιμές των παραμέτρων οπότε ήταν απαραίτητο να δοθούν αρκετές δοκιμαστικές τιμές σε αυτές μέχρι να επιτευχθεί το βέλτιστο αποτέλεσμα.

Όσον αφορά στη μέθοδο των Νευρωνικών Δικτύων αποτελεί το λιγότερο κατάλληλο εργαλείο τόσο από άποψη διαδικασίας κατασκευής όσο και από άποψη προσέγγισης των πραγματικών τιμών. Παρατηρήθηκε ότι, όσο μικρότερος είναι ο όγκος των δεδομένων αλλά και η διακύμανσή τους, τόσο καλύτερα μπορεί να λειτουργήσει η μέθοδος. Οπότε, στην περίπτωση της γερανογέφυρας, της ρευστοποίησης εδάφους και του οδοστρώματος με φόρτιση 60KN τα αποτελέσματα ήταν αρκετά κοντά στις πραγματικές εξόδους, ενώ στη φόρτιση οδοστρώματος με 40 και 80KN η προσέγγιση είχε μεγαλύτερη απόκλιση από τις επιθυμητές τιμές. Κύριο μειονέκτημα της μεθόδου αποτελεί το γεγονός ότι θεωρεί κάποιες τυχαίες τιμές για τα βάρη κάθε φορά που εκτελεί έναν αλγόριθμο, οπότε εξάγει διαφορετικά αποτελέσματα για διάφορα τρεξίματα του αλγορίθμου παρότι εμείς έχουμε σταθερές τις τιμές για τις παραμέτρους. Επιπλέον, εξαιτίας του ότι η μέθοδος εξαρτάται από πολλές παραμέτρους, είναι περίπλοκο και χρονοβόρο να βρεθεί η βέλτιστη προσέγγιση του προβλήματος. Τέλος, σε περίπτωση που απαιτούνται πολλοί κόμβοι για κάθε επίπεδο, ο χρόνος εκτέλεσης του αλγορίθμου είναι πολύ μεγαλύτερος σε σύγκριση με τις δυο άλλες μεθόδους.

7

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Aly Farag and Refaat M.Mohamed *Regression Using Support Vector Machines: Basic Foundations,* 2004

[2] Bruce Ankenman, Barry L. Nelson, Jeremy Staum *Stochastic Kriging for Simulation Metamodeling,* 2009

[3] Christopher J.C. Burges *A Tutorial on Support Vector Machines for Pattern Recognition,* 1998

[4] Dustin Boswell *Introduction to Support Vector Machines,* 2002

[5] G. Sen and E. Akyol *A genetic-algorithm approach for assessing the liquefaction potential of sandy soils,* 2010

[6] Howard Demuth Mark Beale *Neural Network Toolbox User's Guide*, 2002

[7] Jack P.C Kleijnen *Kriging metamodeling in simulation: a review,* 2007

[8] Kimeldorf, Wahba *Some results on Tchebycheffian spline functions,* 1971

[9] Mehmet Saltan, Volkan Emre Uz, Bekir Aktas *Artificial neural networks–based backcalculation of the structural properties of a typical flexible pavement,* 2012

[10] Olivier Roustant, David Ginsbourger, Yves Deville *DiceKriging, DiceOptim: Two R Packages for the Analysis of Computer Experiments by Kriging-Based Metamodeling and Optimization*, 2012

[11] Panagiota Kouli *Optimization of an Overhead Crane with Metaheuristics,* Postgraduate Dissertation, National Technical University of Athens, 2010.

[12] Serdal Terzi, Mehmet Saltan, Mustafa Karaşahin *Backcalculation of pavement layer thickness using data mining,* 2012

[13] Søren N. Lophaven, Hans Bruun Nielsen,Jacob Søndergaard *DACE: A MATLAB Kriging Toolbox,* 2002

[14] V.Vapnik *The Nature of Statistical Learning Theory*, 1995

[15] Λαγαρός Νίκος *Βελτιστοποίηση κατασκευών με χρήση εξελικτικών αλγορίθμων και νευρωνικών δικτύων*, Διδακτορική Διατριβή, ΕΜΠ, 2001

[16] Παπαθανασίου Γεώργιος *Φαινόμενα ρευστοποίησης εδαφών στον Ελληνικό χώρο,* Διδακτορική Διατριβή, ΑΠΘ, 2006

[17] Πηλιούνης Γεώργιος *Ανάλυση αξιοπιστίας γεωτεχνικών προβλημάτων με χρήση μετευριστικών μεθόδων βελτιστοποίησης*, Διπλωματική Εργασία, Σχολή Πολιτικών Μηχανικών, ΕΜΠ, 2011

[18] Τσαϊμού Χριστίνα *Πρόβλεψη παραμενουσών παραμορφώσεων ασφαλτικών οδοστρωμάτων*, Διπλωματική Εργασία, ΕΜΠ, 2012