



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών

Εργαστήριο Τεχνητής Νοημοσύνης και
Συστημάτων Μάθησης, Τομέας
Τεχνολογίας Πληροφορικής και
Υπολογιστών

**Χρήση ευφυών τεχνικών για τη μελέτη παιχνιδιού
στρατηγικής με μερικός παρατηρήσιμο περιβάλλον**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΧΑΡΑΛΑΜΠΟΣ Χ. ΑΣΗΜΑΚΟΠΟΥΛΟΣ

Επιβλέπων : Ανδρέας-Γεώργιος Σταφυλοπάτης

Καθηγητής Ε.Μ.Π.

Συνεπιβλέπων : Γεώργιος Σιόλας

Ε.ΔΙ.Π. Ε.Μ.Π.

Αθήνα, Δεκέμβριος 2019



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών

Εργαστήριο Τεχνητής Νοημοσύνης και
Συστημάτων Μάθησης, Τομέας
Τεχνολογίας Πληροφορικής και
Υπολογιστών

**Χρήση ευφύων τεχνικών για τη μελέτη παιχνιδιού
στρατηγικής με μερικώς παρατηρήσιμο περιβάλλον**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΧΑΡΑΛΑΜΠΟΣ Χ. ΑΣΗΜΑΚΟΠΟΥΛΟΣ**

Επιβλέπων : Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπων : Γεώργιος Σιόλας
Ε.ΔΙ.Π. Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 17^η Δεκεμβρίου 2019.

.....
Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Παναγιώτης Τσανάκας
Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Στάμου
Αναπληρωτής Καθηγητής Ε.Μ.Π

Αθήνα, Δεκέμβριος 2019

.....
Χαράλαμπος Χ. Ασημακόπουλος

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

Copyright © Χαράλαμπος Χ. Ασημακόπουλος, 2019.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Στην παρούσα διπλωματική εργασία, μελετάται η μοντελοποίηση ενός μέρους του μερικώς παρατηρήσιμου παιχνιδιού καρτών Tichu. Ειδικότερα, με τη χρήση ευφών τεχνικών, μοντελοποιείται όλο το αρχικό στάδιο του παιχνιδιού, δηλαδή αυτό που αφορά τη δήλωση Grand Tichu. Αναλύονται, αρχικά, παλαιότερες έρευνες σε άλλα παιχνίδια με μερικώς παρατηρήσιμο περιβάλλον, καθώς και παλαιότερες έρευνες που έχουν εφαρμοστεί πάνω στο παιχνίδι Tichu, ή απλοποιήσεις αυτού. Στη συνέχεια, παρουσιάζεται η δική μας πρόταση ευφών πρακτόρων όσον αφορά το Tichu, για την οποία δημιουργούνται τα μοντέλα της λογιστικής παλινδρόμησης και της μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση. Αφού μελετώνται τα αποτελέσματα της εκπαίδευσης των μοντέλων αυτών, αυτό που θα χρησιμοποιηθεί για καθέναν από τους ευφείς πράκτορες μας. Τέλος, προτείνονται μελλοντικές μελέτες πάνω στο συγκεκριμένο αντικείμενο.

Λέξεις κλειδιά

Tichu, μερικώς παρατηρήσιμο περιβάλλον, μηχανική μάθηση, μηχανές διανυσμάτων υποστήριξης με παλινδρόμηση, λογιστική παλινδρόμηση, επιλογή χαρακτηριστικών

Abstract

The purpose of the current dissertation is the modeling of the decision-making process called Grand Tichu, which is a part of the partially-observable card game Tichu. Older researches regarding partially-observable games are analyzed, including researches and applications that are based on the game of Tichu or simplified versions of it. Additionally, our own smart Tichu agents are introduced, by use of the models which we create in this paper: a logistic regression and a support vector regression one. We select which of these models will be used for each one of our smart agents, depending on the results of the models' training. Last but not least, we share our thoughts about potential future work in relation to this field which could be picked up by other researchers.

Key words

Tichu, partially observable space, machine learning, support vector regression, logistic regression, feature extraction

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο του προπτυχιακού προγράμματος σπουδών της Σχολής Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου και σηματοδοτεί την ολοκλήρωση των σπουδών μου ενώ συγχρόνως αποτελεί το ερέθισμα για περαιτέρω έρευνα στο συγκεκριμένο αντικείμενο. Προτού όμως αναφερθώ στη περιγραφή της εργασίας και στα αποτελέσματα που προέκυψαν, θα ήθελα να ευχαριστήσω θερμά τους ανθρώπους οι οποίοι μέσω της συνεργασίας μας, συνέβαλαν σημαντικά στην ολοκλήρωση αυτής της εργασίας.

Αρχικά θα ήθελα να απευθύνω τις ευχαριστίες μου στον επιβλέποντα κ. Ανδρέα-Γεώργιο Σταφυλοπάτη, Καθηγητή Ε.Μ.Π ο οποίος μου προσέφερε τη δυνατότητα να εκπονήσω την διπλωματική μου σε ένα αντικείμενο ιδιαίτερα ελκυστικό και ενδιαφέρον για μένα και να διευρύνω τις επιστημονικές μου γνώσεις. Παράλληλα θα ήθελα να ευχαριστήσω τους κ.κ. Παναγιώτη Τσανάκα, Καθηγητή Ε.Μ.Π και Γεώργιο Στάμου, Αναπληρωτή Καθηγητή Ε.Μ.Π για την τιμή που μου έκαναν να είναι μέλη της επιτροπής εξέτασης της διπλωματικής εργασίας.

Επίσης οφείλω ιδιαίτερες ευχαριστίες στον συνεπιβλέποντα κ. Γεώργιο Σιόλα, ΕΔΠ Ε.Μ.Π για το χρόνο που αφιέρωσε και την θεμελιώδη του συνεισφορά στην εκπόνηση της συγκεκριμένης εργασίας. Η στήριξη του, επιστημονική και πνευματική, καθώς και η καθοδήγηση του σε όλη τη διάρκεια της πορείας αυτής συνέβαλαν τα μέγιστα στην επίτευξη ενός πολύ σημαντικού για εμένα στόχου. Η προθυμία του να με βοηθήσει μέσω της εμπειρίας και των γνώσεων του σε οποιαδήποτε δυσκολία συνάντησα στάθηκαν καθοριστικές και η συνεργασία μας θεωρώ πως ήταν άκρως επιτυχημένη και επικοινωνιακή.

Τέλος, με εξίσου μεγάλη θέρμη θέλω να αναφερθώ και να ευχαριστήσω την οικογένεια μου , η οποία με στήριξε όλα αυτά τα χρόνια, καθώς και τους φίλους και τους συμφοιτητές μου, οι οποίοι στάθηκαν δίπλα μου σε όλη τη διάρκεια της ακαδημαϊκής μου πορείας, ο καθένας με τον δικό του ξεχωριστό τρόπο, με ιδιαίτερη μνεία σε όσους με μύησαν στον κόσμο του παιχνιδιού Tichu το οποίο και εξετάζουμε στην εργασία αυτή.

Χαράλαμπος Χ. Ασημακόπουλος,

Πίνακας περιεχομένων

1) Εισαγωγή	12
2) Μερικώς παρατηρήσιμα παιχνίδια καρτών	16
i) Μαρκοβιανές διαδικασίες αποφάσεων	16
ii) Μερικώς Παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων	17
iii) Παιχνίδια που μπορούν να μοντελοποιηθούν ως Μερικώς Παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων	20
3) Το παιχνίδι Tichu	22
i) Τι είναι το Tichu;	22
ii) Το παιχνίδι	24
iii) Η τράπουλα	25
iv) Ειδικές κάρτες	26
v) Δυναμικότητα φύλλων	30
vi) Συνδυασμοί	31
vii) Δηλώσεις Grand Tichu και Tichu	33
viii) Ανταλλαγές	34
ix) Γενικοί Κανόνες	34
x) Πόντοι	35
4) Εφαρμογές τεχνητής νοημοσύνης και μηχανικής μάθησης για το Tichu	36
i) Το Mini Tichu και ευφυείς υλοποιήσεις του	36
ii) Χρήση ευφύων τεχνικών σε διάφορα κομμάτια του παιχνιδιού Tichu	55
iii) Εφαρμογές παιχνιδιού Tichu	65
5) Η προσέγγιση μας / Συλλογή δεδομένων	66
i) Διαβάθμιση δυναμικότητας παικτών και ομάδων στο Tichu με τη χρήση του συστήματος αξιολόγησης Elo	67
ii) Συλλογή δεδομένων	72
6) Η επιλογή των χαρακτηριστικών των ευφύων συστημάτων	73
i) Μοντελοποίηση των συνδυασμών Tichu	77

ii) Επιλογή επιπλέον χαρακτηριστικών	80
7) Η επιλογή και η εκπαίδευση των ευφών συστημάτων	82
i) Επιλογή ευφών συστημάτων	82
ii) Επιλογή επιπλέον μέτρων απόδοσης	83
iii) Μηχανές διανυσμάτων υποστήριξης	85
iv) Μηχανές διανυσμάτων υποστήριξης με παλινδρόμηση	92
v) Μοντέλο μηχανών διανυσμάτων υποστήριξης με παλινδρόμηση	90
vi) Λογιστική παλινδρόμηση	94
vii) Μοντέλο λογιστικής παλινδρόμησης	96
8) Αποτελέσματα των μοντέλων μετά την εκπαίδευση	98
i) Αποτελέσματα μοντέλου Μηχανής Διανυσμάτων Υποστήριξης με παλινδρόμηση	98
ii) Αποτελέσματα Μοντέλου Μηχανής Λογιστικής παλινδρόμησης	104
iii) Σύγκριση των αποτελεσμάτων των δύο μοντέλων	109
9) Επίλογος και προτάσεις για μελλοντικές μελέτες	115
Παράρτημα - Επιλογή χαρακτηριστικών με βάση τους επιτρεπτούς συνδυασμούς στο Tichu	118
Βιβλιογραφία	125

Πίνακας γραφημάτων

2-1 Μαρκοβιανή Διαδικασία Απόφασης στην ενισχυμένη μάθηση	16
3-1 Tichu logo	22
3-2 Τράπουλα Tichu	25
3-3 Ειδικές κάρτες του Tichu.....	26
3-4 Mahjong.....	26
3-5 Τα σκυλιά	28
3-6 Ο δράκος	29
3-7 Ο φοίνικας	29
4-1 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των καρτών του κόμβου	41
4-2 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου ..	41
4-3 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών	42
4-4 Απλή αξιολόγηση εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)	42
4-5 Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου	43
4-6 Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών	43
4-7 Αξιολόγηση με βάση τον αριθμό των καρτών του κόμβου εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)	44
4-8 Αξιολόγηση με βάση τον αριθμό των συνδυασμών του κόμβου εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών.....	44
4-9 Αξιολόγηση με βάση τον αριθμό των κρυφών συνδυασμών εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)	45
4-10 Συνολικές νίκες κάθε συνάρτησης αξιολόγησης.....	45
4-11 Βήματα διαδικασίας επιλογής MCTS.....	47
4-12 12 Ποσοστό νίκης χωρίς ντετερμινισμό	48
4-13 MCTS και MC πράκτορες εναντίον τυχαίων πρακτόρων με διαφορετικό αριθμό πιθανών κινήσεων χωρίς ντετερμινισμό.	50
4-14 Ποσοστό των σωστά επιλεγμένων συνδυασμών για τις άγνωστες κάρτες των υπολοίπων παικτών με και χωρίς ντετερμινισμό για MC και MCTS πράκτορες.	50
4-15 FSM για το παιχνίδι Mini Tichu όταν και όλοι οι παίκτες είναι ενεργοί σε έναν γύρο.....	54
4-16 Εξωτερικές κατανομές και το μέσο της παραμέτρου p	56
4-17 Εξωτερικές κατανομές και το μέσο της παραμέτρου p για ένα καλό και ένα κακό χέρι	57
4-18 Αποτελέσματα του εμβადού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα για το Grand Tichu	61

4-19 Χαρακτηριστικές καμπύλες όλων των μοντέλων με είσοδο 14 χαρακτηριστικά συν τις κάρτες του χεριού	61
4-20 Συντελεστές για τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη Grand Tichu	62
4-21 Αποτελέσματα του εμβადού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα για το Tichu	63
4-22 Χαρακτηριστικές καμπύλες όλων των μοντέλων με είσοδο 14 χαρακτηριστικά συν τις κάρτες του χεριού	63
4-23 Συντελεστές για τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη Tichu	63
4-24 Πιθανότητα επιτυχούς δήλωσης Grand Tichu	64
4-25 Εφαρμογή Tichu για iOS	65
7-1 Διαχωριστικό υπερεπίπεδο	86
7-2 Απόσταση μεταξύ κλάσεων	86
7-3 Διαχωριστικό υπερεπίπεδο με χαλαρές μεταβλητές	87
7-4 Κλάσεις γραμμικώς διαχωρίσιμες στο χώρο R^2 , αλλά μη γραμμικώς διαχωρίσιμες στο χώρο R^1	88
7-5 Παράδειγμα γραμμικού πυρήνα, όπου τα σημεία που περικλείονται με τετραγωνικό πλαίσιο αποτελούν διανύσματα υποστήριξης	89
7-6 Παράδειγμα διαχωρισμού μη-γραμμικού προβλήματος με χρήση SVMs γραμμικού (διακεκομμένη γραμμή) και πολυωνυμικού πυρήνα (συνεχής γραμμή)	90
7-7 Αριστερή υπό-εικόνα: Διαχωρισμός με ακτινωτό πυρήνα στον αρχικό χώρο. Δεξιά υπό-εικόνα: Γραμμικός διαχωρισμός σε χώρο υψηλότερης διάστασης με τη βοήθεια μετασχηματισμού ακτινωτού πυρήνα	90
7-8 Γράφημα ενός μονοδιάστατου μοντέλου μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση	92
8-1 Το ποσοστό ορθότητας δηλώσεων του μοντέλου SVR ανά τιμή κατωφλίου	99
8-2 Το ποσοστό ακρίβειας δηλώσεων του μοντέλου SVR ανά τιμή κατωφλίου	100
8-3 Το ποσοστό επιτυχούς ανάκλησης του μοντέλου SVR ανά τιμή κατωφλίου	101
8-4 Η βαθμολογία F1 του μοντέλου SVR ανά τιμή κατωφλίου	102
8-5 Η καμπύλη ROC του μοντέλου SVR	103
8-6 Το ποσοστό ορθότητας δηλώσεων του μοντέλου LR ανά τιμή κατωφλίου	104
8-7 Το ποσοστό ακρίβειας δηλώσεων του μοντέλου LR ανά τιμή κατωφλίου	105
8-8 Το ποσοστό επιτυχούς ανάκλησης του μοντέλου LR ανά τιμή κατωφλίου	106
8-9 Η βαθμολογία F1 του μοντέλου LR ανά τιμή κατωφλίου	107
8-10 Η καμπύλη ROC του μοντέλου LR	108
8-11 Σύγκριση ποσοστών ορθότητας δηλώσεων των μοντέλων SVR και LR ανά τιμή κατωφλίου ..	109
8-12 Σύγκριση ποσοστών ακρίβειας δηλώσεων των μοντέλων SVR και LR ανά τιμή κατωφλίου	110
8-13 Σύγκριση ποσοστών επιτυχούς ανάκλησης των μοντέλων SVR και LR ανά τιμή κατωφλίου ..	111
8-14 Σύγκριση βαθμολογίας F1 των μοντέλων SVR και LR ανά τιμή κατωφλίου	112
8-15 Σύγκριση καμπύλης ROC των μοντέλων SVR και LR ανά τιμή κατωφλίου	113

Κεφάλαιο 1

Εισαγωγή

Τεχνητή Νοημοσύνη (ΤΝ) είναι ο τομέας της επιστήμης των υπολογιστών, που ασχολείται με τη σχεδίαση ευφυών υπολογιστικών συστημάτων, δηλαδή συστημάτων ικανών για λειτουργίες που αποδίδονται σε ανθρώπινη νοημοσύνη. Τα τελευταία χρόνια έχει παρουσιαστεί ραγδαία αύξηση στη χρήση ευφυών τεχνικών για τη δημιουργία πρακτόρων σε διάφορους τομείς. Ειδικότερα, ο τομέας των παιχνιδιών, έχει γίνει αντικείμενο πολλών ερευνών και επενδύσεων και οι πράκτορες που έχουν δημιουργηθεί για διάφορα παιχνίδια, τόσο ηλεκτρονικά εν γένει όσο και επιτραπέζια σε ηλεκτρονική μορφή, παρουσιάζουν σταδιακά όλο και μεγαλύτερη επιτυχία.

Ο τομέας των παιχνιδιών αποτέλεσε ξεχωριστό πεδίο ερευνών της ΤΝ από τις απαρχές της εμφάνισης της. Ένα από τα πρώτα παραδείγματα, αποτελεί η ηλεκτρονική μορφή του παιχνιδιού Nim η οποία δημιουργήθηκε το 1951 και δημοσιεύτηκε το 1952. Οι πράκτορες που χρησιμοποιήθηκαν για την ηλεκτρονική μορφή του παιχνιδιού ήταν ικανοί να νικούν συστηματικά τους καλύτερους παίχτες του παιχνιδιού. (Grant, 1952) . Επίσης, το 1951, ο Christopher Strachey κατάφερε χρησιμοποιώντας τη μηχανή Ferranti Mark του πανεπιστημίου του Manchester, να χρησιμοποιήσει ευφυείς πράκτορες για το παιχνίδι checkers όπως αντίστοιχα και ο Dietrich Prinz για το σκάκι. ("A Brief History of Computing") Αυτές ήταν οι πρώτες απόπειρες χρήσης πρακτόρων σε παιχνίδια και ταυτόχρονα από τα πρώτα προγράμματα που γράφτηκαν ποτέ για υπολογιστή. Οι πράκτορες για το περιβάλλον του παιχνιδιού checkers και για το σκάκι, συνέχισαν να μελετώνται και να βελτιώνονται καταλήγοντας στη διάσημη ήττα του πρωταθλητή Garry Kasparov στο σκάκι από τη μηχανή Deep Blue της IBM το 1997. (McCorduck, 2004)

Βέβαια, τα πρώτα ηλεκτρονικά παιχνίδια που εμφανίστηκαν τις δεκαετίες του '60 και του '70 όπως το Spacewar!, το Pong και το Gotcha, έθεται αντιμέτωπους δύο παίκτες μεταξύ τους και άρα η χρήση πρακτόρων δεν είχε αρχικά εφαρμογή. Παιχνίδια που πρόσφεραν τη δυνατότητα σε έναν παίχτη να παίξει μόνος του εναντίον αντιπάλων-πρακτόρων άρχισαν να εμφανίζονται τη δεκαετία του '70 με πιο αξιοσημείωτα το Speed Race της Taito και τα Qwak και Pursuit της Atari , και καθιερώθηκαν πλέον το 1978 με τη μεγάλη επιτυχία του Space

Invaders. Έκτοτε, άρχισαν να δημιουργούνται παιχνίδια με όλο και πολυπλοκότερα μοντέλα ευφυών πρακτόρων.

Σήμερα, η παρουσία των ευφυών πρακτόρων είναι αισθητή σε όλα τα ηλεκτρονικά παιχνίδια και στις προσομοιώσεις επιτραπέζιων παιχνιδιών σε υπολογιστή και η μελέτη του τομέα αυτού έχει οδηγήσει στην εμφάνιση όλο και επιτυχέστερων πρακτόρων. Χαρακτηριστικά αναφέρουμε τη νίκη της ομάδας ευφυών πρακτόρων της εταιρίας OpenAI εναντίον των παγκόσμιων πρωταθλητών στο παιχνίδι e-sports Dota 2 της εταιρίας Valve (OpenAI, 2019) , καθώς και του AlphaGo που έχει αναπτύξει η DeepMind της Google που κατάφερε το 2016 να νικήσει τον Κορεάτη πρωταθλητή Lee Se-dol 4-1 σε ένα από τα περιπλοκότερα επιτραπέζια παιχνίδια στον κόσμο , το Go (Καρατζάς, 2019). Μάλιστα, η DeepMind βελτίωσε κι άλλο από τότε το πρόγραμμά της και η τελευταία εκδοχή της, το AlphaGo Zero, κατάφερε να νικήσει το αρχικό AlphaGo σε 100 συνεχόμενα ματς, οδηγώντας τον Lee Se-dol στην ανακοίνωση πως δε θα παίζει πλέον σε επαγγελματικό επίπεδο, αφού η TN είναι ουσιαστικά ανίκητη. Η ίδια εταιρεία κατάφερε πρόσφατα για πρώτη φορά να δημιουργήσει πράκτορες για ανταγωνιστικά παιχνίδια πολλαπλών παικτών (multiplayer shooter videogames) που απαιτούν ομαδική συνεργασία. Οι πράκτορες που εκπαιδευσε κατάφεραν να συνεργάζονται για να παίξουν με επιτυχία το κλασσικό βιντεοπαιγνίδι Quake III Arena, νικώντας έμπειρους ανθρώπους παίκτες. Επιπλέον, αποδείχθηκε ότι ένας ευφυής πράκτορας μπορεί να συνεργασθεί μια χαρά με έναν άνθρωπο παίκτη. (Team, 2019)

Αντιλαμβανόμαστε λοιπόν, με βάση τα παραπάνω ότι το πεδίο εφαρμογής των ευφυών πρακτόρων στα παιχνίδια είναι διαρκώς εξελίξιμο και ακόμη υπάρχουν κομμάτια και είδη που δεν έχουν πλήρως μελετηθεί με επιτυχία. Ένα τέτοιο πεδίο, όπου υπάρχει χώρος περαιτέρω εξέλιξης των ευφυών πρακτόρων που έχουν δημιουργηθεί είναι τα παιχνίδια με μερικώς παρατηρήσιμο περιβάλλον, παιχνίδια δηλαδή, στα οποία ένα κομμάτι της πληροφορίας είναι κρυφό από τον παίκτη κατά τη διάρκεια του παιχνιδιού. Τότε παρουσιάζονται δυσκολίες τόσο στη μοντελοποίηση όσο και στην εκπαίδευση πρακτόρων, μιας και τα μερικώς παρατηρήσιμα παιχνίδια έχουν πολύ μεγαλύτερο χώρο αναζήτησης από ότι τα παρατηρήσιμα παιχνίδια. Έχουν γίνει βέβαια προσπάθειες δημιουργίας πρακτόρων και σε τέτοιας φύσεως παιχνίδια. Πρόσφατα, η DeepMind, ανακοίνωσε ότι δημιούργησε την πρώτη τεχνητή νοημοσύνη για το Starcraft 2, κάτι το οποίο αποτελούσε μεγάλη πρόκληση επειδή ο παίκτης συχνά δεν βλέπει τα κομμάτια/μονάδες του αντιπάλου του (Τεχνητή νοημοσύνη της DeepMind έφτασε στο ανώτατο επίπεδο παικτών στο Starcraft 2, δημοφιλές

παιχνίδι των esports, 2019) . Επίσης, έχουν δημιουργηθεί πράκτορες για παιχνίδια όπως το μπριτζ και το πόκερ, παιχνίδια καρτών όπου ο παίκτης, έχει γνώση κατά βάση μόνο των καρτών που κρατάει στο χέρι του. Στο μπριτζ υπάρχουν τα προγράμματα Jack και Wbridge5 τα οποία κατατάσσονται αρκετά ψηλά στις εθνικές κατατάξεις παικτών μπριτζ. Το Jack μάλιστα το 2005 έπαιξε εναντίον 7 ζευγαριών παικτών υψηλού επιπέδου και κατάφερε να κερδίσει 3 από αυτά. (Wikipedia, 2019) Αντίστοιχα, στο πόκερ υπάρχουν διάφορα μοντέλα και είδη ευφυών πρακτόρων με αποτέλεσμα να υπάρχει ετήσιο τουρνουά (The Annual Computer Poker Competition) όπου αυτά ανταγωνίζονται μεταξύ τους και όπου δεν υπάρχει ως τώρα σταθερός νικητής. (Annual Computer Poker Competition)

Όσον αφορά λοιπόν τα παιχνίδια με μερικώς παρατηρήσιμο περιβάλλον , οι ευφυείς πράκτορες που έχουν δημιουργηθεί έχουν ακόμη μεγάλα περιθώρια βελτίωσης , γι' αυτό και στην παρούσα εργασία επιλέχθηκε ένα παιχνίδι της εν λόγω κατηγορίας για να μελετηθεί ως προς αυτή την κατεύθυνση, το Tichu. Το παιχνίδι καρτών Tichu αποτελεί ένα διαδομένο παιχνίδι αναρρίχησης, με μερικώς παρατηρήσιμο περιβάλλον. Το παιχνίδι αυτό παρουσιάζει μεγάλη πολυπλοκότητα και τεράστιο όγκο διαθέσιμων επιλογών ανά κίνηση και άρα αποτελεί ένα πολύ ενδιαφέρον πεδίο ερευνών για τη μοντελοποίηση πρακτόρων και την επιλογή χαρακτηριστικών των μοντέλων αυτών. Επίσης στο παιχνίδι αυτό συμμετέχουν τέσσερις παίκτες, χωρισμένοι σε δύο ομάδες. Η εισαγωγή της συνεργατικότητας ως παραμέτρου, περιπλέκει ακόμα παραπάνω τη μοντελοποίηση του εν λόγω παιχνιδιού.

Στην παρούσα διπλωματική εργασία επιχειρείται λοιπόν η μοντελοποίηση με χρήση ευφυών τεχνικών του παιχνιδιού αυτού. Η έρευνα μας επικεντρώθηκε σε ένα κομμάτι του παιχνιδιού Tichu, αυτό της δήλωσης Grand Tichu. Η δήλωση Grand Tichu είναι πολύ σημαντική για το παιχνίδι, καθώς εάν είναι επιτυχής δίνει το μεγαλύτερο δυνατό βαθμολογικό προβάδισμα σε μία ομάδα, ενώ εάν είναι ανεπιτυχής προκαλεί τη μεγαλύτερη δυνατή βαθμολογική απώλεια για την ομάδα αυτή. Σκοπός μας είναι η δημιουργία τριών πρακτόρων διαφορετικής επιθετικότητας, δηλαδή ενός συντηρητικού, ενός ουδέτερου και ενός επιθετικού πράκτορα.

Αρχικά, στο Κεφάλαιο 2, αναλύονται οι Μαρκοβιανές διαδικασίες αποφάσεων καθώς και ειδικότερα, οι μερικώς παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων. Αναφέρονται διάφορες έρευνες και εργασίες πάνω στη μοντελοποίηση μερικώς παρατηρήσιμων παιχνιδιών και εξετάζεται το πώς αυτά μπορούν να μοντελοποιηθούν ως Μερικώς Παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων. Στο Κεφάλαιο 3, γίνεται εκτενής

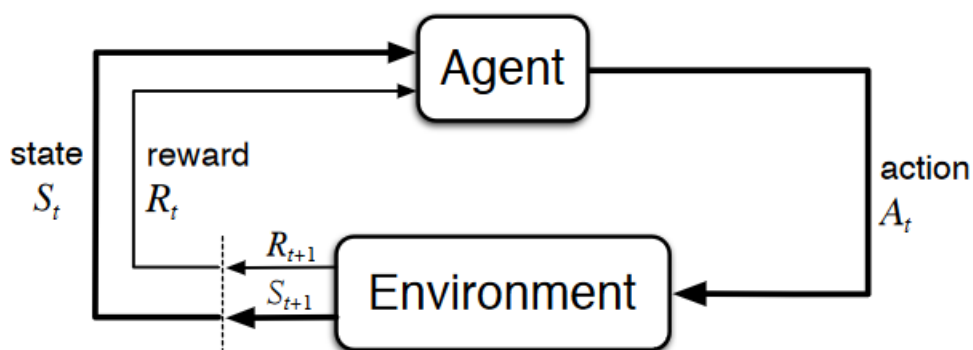
αναφορά στο παιχνίδι Tichu, τα χαρακτηριστικά και τους κανόνες του. Στο Κεφάλαιο 4, καταγράφονται τα αποτελέσματα ερευνών που έχουν γίνει πάνω σε διάφορα κομμάτια του παιχνιδιού Tichu μέχρι σήμερα και οι εφαρμογές του παιχνιδιού με χρήση ευφών πρακτόρων που είναι διαθέσιμες για το ευρύ κοινό. Στο Κεφάλαιο 5 γίνεται επεξήγηση του πλέον επικρατέστερου τρόπου βαθμολόγησης παιχνιδιών όπως το Tichu και το σκάκι, της βαθμολόγησης Elo, με βάση τον οποίο επιλέξαμε τους παίκτες τα δεδομένα των οποίων χρησιμοποιήσαμε για την εκπαίδευση των ευφών μοντέλων μας. Στο Κεφάλαιο 6, αναλύουμε τη διαδικασία εξόρυξης χαρακτηριστικών για τη μοντελοποίηση των δεδομένων μας με σκοπό τη δημιουργία ενός μοντέλου λήψης απόφασης για δήλωση Grand Tichu. Τα μοντέλα τα οποία επιλέξαμε είναι αυτά της λογιστικής παλινδρόμησης και της μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση. Στο κεφάλαιο 7, εξηγείται το μαθηματικό υπόβαθρο πίσω από τα μοντέλα αυτά, καθώς και οι διάφοροι παράμετροι αυτών. Στο κεφάλαιο 8, παρουσιάζουμε τα αποτελέσματα μετά την εκπαίδευση των μοντέλων αυτών. Γίνεται σύγκριση μεταξύ των δύο μοντέλων με βάση τα μέτρα απόδοσης που έχουμε επιλέξει και γίνεται επιλογή μοντέλων για τον κάθε ένα από τους 3 πράκτορες που θέλουμε να δημιουργήσουμε. Τέλος, προτείνονται διάφορες μελλοντικές μελέτες που μπορούν να διεξαχθούν χρησιμοποιώντας ως βάση τα αποτελέσματα της παρούσας εργασίας.

Κεφάλαιο 2

Μερικώς παρατηρήσιμα παιχνίδια καρτών

i) Μαρκοβιανές διαδικασίες αποφάσεων

Μια Μαρκοβιανή διαδικασία απόφασης (Markov Decision Process - MDP) είναι μια διακριτού χρόνου στοχαστική διαδικασία ελέγχου. Παρέχει ένα μαθηματικό πλαίσιο για τη μοντελοποίηση της λήψης αποφάσεων σε καταστάσεις όπου τα αποτελέσματα είναι εν μέρει τυχαία και εν μέρει υπό τον έλεγχο ενός υπεύθυνου για τη λήψη αποφάσεων. Τα MDPs ήταν γνωστά τουλάχιστον από τη δεκαετία του 1950 και χρησιμοποιούνται σε πολλούς κλάδους όπως στη ρομποτική, στον αυτόματο έλεγχο, στην οικονομία και στις κατασκευές. Το όνομα τους προέρχεται από τον Ρώσο μαθηματικό Αντρέι Μάρκοβ. (Zheng, 2019, p. 1)



2-1 Μαρκοβιανή Διαδικασία Απόφασης στην ενισχυμένη μάθηση, (Richard S. Sutton, 2015, p. 54)

Η Μαρκοβιανή διαδικασία απόφασης είναι μια τούπλα $\langle S, A, P, R, \gamma \rangle$ όπου:

- S είναι ένα αριθμήσιμο σύνολο καταστάσεων
- A είναι ένα αριθμήσιμο σύνολο ενεργειών
- P είναι μια μήτρα πιθανοτήτων μεταβατικών καταστάσεων
- R είναι μια συνάρτηση ανταμοιβής
- γ είναι ένας συντελεστής έκπτωσης ($\gamma \in [0,1]$). (Silver, 2015, p. 10)

Σε κάθε χρονικό βήμα, η διαδικασία βρίσκεται σε κάποια κατάσταση s , και ο υπεύθυνος λήψης αποφάσεων μπορεί να διαλέξει οποιαδήποτε ενέργεια a που είναι διαθέσιμη στην

κατάσταση s . Η διαδικασία απαντά στο επόμενο χρονικό βήμα μεταβαίνοντας τυχαία σε μια νέα κατάσταση s' , και δίνοντας στον υπεύθυνο λήψης αποφάσεων την αντίστοιχη ανταμοιβή, όπως αυτή προκύπτει από τη συνάρτηση ανταμοιβής $R_\alpha(s, s')$. (Richard S. Sutton, 2015, p. 54)

Η πιθανότητα η διαδικασία να μεταβεί στη νέα κατάσταση s' επηρεάζεται από την επιλεγμένη ενέργεια. Συγκεκριμένα, αυτή δίνεται από τη μήτρα πιθανοτήτων μεταβατικών καταστάσεων $P_\alpha(s, s')$. Επομένως, η επόμενη κατάσταση s' εξαρτάται από την τωρινή κατάσταση s και την ενέργεια του υπεύθυνου λήψης αποφάσεων a . Αλλά δοθέντων των s και a , είναι υπό όρους ανεξάρτητη από όλες τις προηγούμενες καταστάσεις και ενέργειες. Με άλλα λόγια, οι μεταβάσεις καταστάσεων μιας MDP ικανοποιούν την Μαρκοβιανή ιδιότητα (στη θεωρία των πιθανοτήτων και της στατιστικής, ο όρος Μαρκοβιανή ιδιότητα αναφέρεται στην ιδιότητα μιας στοχαστικής διαδικασίας να μην επηρεάζεται από τη μνήμη της). (Lagoudakis, 2017, p. 1311)

ii) Μερικώς παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων

Μια μερικώς παρατηρήσιμη Μαρκοβιανή διαδικασία αποφάσεων (Partially observable Markov decision process - POMDP) είναι μια γενίκευση των Μαρκοβιανών διαδικασιών αποφάσεων. Μία POMDP μοντελοποιεί τη διαδικασία απόφασης ενός πράκτορα στην οποία υποτίθεται ότι η δυναμική του συστήματος καθορίζεται από μία MDP, αλλά ο πράκτορας δε μπορεί να παρατηρήσει άμεσα την “κρυμμένη” κατάσταση. Αντ’ αυτού, πρέπει να διατηρεί μια κατανομή πιθανότητας όσον αφορά το σύνολο των πιθανών καταστάσεων, με βάση ένα σύνολο παρατηρήσεων και ένα σύνολο πιθανοτήτων παρατήρησης καθώς και την “κρυμμένη” MDP.

Το πλαίσιο των POMDP είναι αρκετά γενικό ώστε να μοντελοποιεί επιτυχώς μια ποικιλία διαδικασιών διαδοχικών αποφάσεων στον πραγματικό κόσμο. Οι εφαρμογές περιλαμβάνουν προβλήματα πλοήγησης ρομπότ, συντήρησης μηχανών και γενικότερες περιπτώσεις προγραμματισμού υπό αβεβαιότητα. Το γενικό πλαίσιο των MDP με ελλιπείς πληροφορίες για το χώρο καταστάσεων περιγράφηκε από τον Karl Johan Åström το 1965 για την περίπτωση διακριτού χώρου καταστάσεων και στη συνέχεια μελετήθηκε περαιτέρω από την

ερευνητική κοινότητα των επιχειρήσεων όπου δημιουργήθηκε το ακρωνύμιο POMDP. Αργότερα προσαρμόστηκε για προβλήματα τεχνητής νοημοσύνης και για προβλήματα αυτοματοποιημένου σχεδιασμού από τους Leslie P. Kaelbling και Michael L. Littman.

Μια ακριβής λύση σε μία POMDP αποδίδει τη βέλτιστη ενέργεια για κάθε πιθανή πεποίθηση πάνω στο σύνολο καταστάσεων. Η βέλτιστη ενέργεια μεγιστοποιεί (ή ελαχιστοποιεί) την αναμενόμενη ανταμοιβή (ή κόστος) του πράκτορα σε έναν πιθανώς άπειρο ορίζοντα. Η ακολουθία βέλτιστων ενεργειών είναι γνωστή και ως βέλτιστη πολιτική του πράκτορα. (Wikipedia, 2019)

Μια POMDP είναι μια MDP με κρυφές καταστάσεις. Είναι ένα κρυφό Μαρκοβιανό μοντέλο με ενέργειες.

Μια POMDP είναι μια τούπλα $\langle S, A, O, P, R, Z, \gamma \rangle$ όπου:

- S είναι ένα αριθμήσιμο σύνολο καταστάσεων
- A είναι ένα αριθμήσιμο σύνολο ενεργειών
- O είναι ένα αριθμήσιμο σύνολο παρατηρήσεων
- P είναι μια μήτρα πιθανοτήτων μεταβατικών καταστάσεων
- R είναι μια συνάρτηση ανταμοιβής
- Z είναι μια συνάρτηση παρατήρησης
- γ είναι ένας συντελεστής έκπτωσης ($\gamma \in [0,1]$).

Σε κάθε χρονική περίοδο, το περιβάλλον βρίσκεται σε κάποια κατάσταση $s \in S$. Ο πράκτορας επιλέγει μια ενέργεια $a \in A$, η οποία προκαλεί μετάβαση του περιβάλλοντος στην κατάσταση $s' \in S$ με πιθανότητα που δίνεται από τη μήτρα πιθανοτήτων μεταβατικών καταστάσεων $P(s' | s, a)$. Την ίδια στιγμή, ο πράκτορας λαμβάνει μια παρατήρηση $o \in O$ η οποία εξαρτάται από τη νέα κατάσταση του περιβάλλοντος (s') και από την ενέργεια που μόλις επιλέχθηκε (a) με πιθανότητα $O(o | s', a)$. Ολοκληρώνοντας, ο πράκτορας λαμβάνει μια ανταμοιβή r ίση με $R(s, a)$. Στη συνέχεια, η διαδικασία επαναλαμβάνεται. Ο στόχος για τον πράκτορα είναι να διαλέγει ανά κάθε χρονική στιγμή ενέργειες οι οποίες μεγιστοποιούν την εκτιμώμενη μελλοντική ανταμοιβή μετά έκπτωσης $E[\sum_{t=0}^{\infty} \gamma^t r_t]$, όπου r_t είναι η ανταμοιβή που λήφθηκε τη χρονική στιγμή t . Ο συντελεστής έκπτωσης γ καθορίζει το κατά

πόσον προτιμάμε τις άμεσες ανταμοιβές σε σχέση με τις ανταμοιβές που έρχονται σε βάθος χρόνου. Όταν $\gamma = 0$ ο πράκτορας ενδιαφέρεται μόνο για το ποια ενέργεια θα αποφέρει την υψηλότερη εκτιμώμενη άμεση ανταμοιβή, ενώ όταν $\gamma = 1$ ο πράκτορας ενδιαφέρεται για τη μεγιστοποίηση του εκτιμώμενου αθροίσματος των μελλοντικών ανταμοιβών. (Silver, Computer Science, UCL, p. 51)

Επειδή ο πράκτορας δε παρατηρεί άμεσα την κατάσταση του περιβάλλοντος, λαμβάνει αποφάσεις υπό αβεβαιότητα σχετικά με την πραγματική κατάσταση του περιβάλλοντος. Ωστόσο, μέσα από την αλληλεπίδραση με το περιβάλλον και τη λήψη παρατηρήσεων, ο πράκτορας μπορεί να ανανεώσει τις τιμές πίστης του στην πραγματική κατάσταση μέσω της ενημέρωσης της κατανομής πιθανότητας για την τρέχουσα κατάσταση. Αποτέλεσμα αυτής της ιδιότητας είναι ότι η βέλτιστη συμπεριφορά μπορεί συχνά να περιλαμβάνει ενέργειες οι οποίες διαλέγονται μόνο επειδή βελτιώνουν την εκτίμηση του πράκτορα για την τρέχουσα κατάσταση, επιτρέποντάς του έτσι να λαμβάνει καλύτερες αποφάσεις στο μέλλον.

Είναι χρήσιμο να συγκρίνουμε τον παραπάνω ορισμό με τον ορισμό της διαδικασίας λήψης αποφάσεων σε μία MDP. Μία MDP δεν περιλαμβάνει το σύνολο παρατηρήσεων, επειδή ο πράκτορας γνωρίζει πάντα με βεβαιότητα την τρέχουσα κατάσταση του περιβάλλοντος. Εναλλακτικά, μία MDP μπορεί να αναδιατυπωθεί ως μία POMDP ορίζοντας το σύνολο παρατηρήσεων ίσο με το σύνολο των καταστάσεων και προσδιορίζοντας τις ενδεχόμενες πιθανότητες παρατηρήσεων ώστε να επιλέγεται ντετερμινιστικά η παρατήρηση που αντιστοιχεί στην πραγματική κατάσταση. (Wikipedia, 2019)

iii) Παιχνίδια που μπορούν να μοντελοποιηθούν ως μερικώς παρατηρήσιμες Μαρκοβιανές διαδικασίες αποφάσεων

Στο (Ishii, 2005), οι Shin Ishii, Hajime Fujita et al. δημιουργούν μια φόρμουλα απόκτησης αυτόματης στρατηγικής για το παιχνίδι καρτών “Hearts” ως ένα πρόβλημα ενισχυτικής μάθησης. Το παιχνίδι αυτό μπορεί να μοντελοποιηθεί ως μία POMDP, αλλά λόγω του τεράστιου αριθμού πιθανών καταστάσεων είναι απαραίτητο να γίνουν περαιτέρω προσεγγίσεις ώστε να καταστεί η επίλυση του προβλήματος εφικτή από έναν υπολογιστή. Το άρθρο αυτό παρουσιάζει μια μέθοδο προσέγγισης που βασίζεται στην εκτίμηση των μη παρατηρήσιμων μεταβλητών κατάστασης και στην πρόβλεψη για τις ενέργειες των άλλων πρακτόρων και τα αποτελέσματα των προσομοιώσεων δείχνουν ότι η μέθοδος ενισχυτικής μάθησης μπορεί να εφαρμοστεί σε ένα τόσο δύσκολο πρόβλημα πολλαπλών πρακτόρων.

Στο (Peter Cowling, 2012), οι Cowling, Ward et al. χρησιμοποιούν τον αλγόριθμο Monte Carlo με χρήση δέντρου αναζήτησης (Monte Carlo Tree Search - MCTS) για μία παραλλαγή του παιχνιδιού καρτών Magic: The Gathering. Αυτό επιτυγχάνεται μέσα από μία ευρεία γκάμα προσεγγίσεων που χρησιμοποιούν ντετερμινισμό, όπου όλες οι κρυφές και τυχαίες πληροφορίες θεωρούνται γνωστές σε όλους τους παίκτες, σε συνδυασμό με τον αλγόριθμο MCTS και χρησιμοποιώντας ένα φάσμα επιπέδων πολυπλοκότητας, εξειδικευμένων γνώσεων και μείωση της ανταμοιβής ώστε να ενθαρρύνεται το πιο επιθετικό και άμεσο παιχνίδι. Εξετάζονται επίσης διάφορες στρατηγικές κλάδευσης προκειμένου να αυξηθούν οι πληροφορίες που αποκτώνται από κάθε ντετερμινισμό, παράλληλα με μεθόδους που αυξάνουν τη συνάφεια των τυχαίων επιλογών. Τα παραπάνω βοηθούν στην επίτευξη βελτιώσεων στη δυναμικότητα των πρακτόρων.

Στο (Fujii, 2008), το 2008, οι Fujii, Hashida et al. ερευνούν το καθεστώς που ισχύει για τα βιντεοπαιχνίδια ανταλλαγής καρτών και παρουσιάζουν ένα αυτόματο σύστημα απόκτησης στρατηγικής για τα παιχνίδια αυτά, με χρήση τεχνικής δειγματοληψίας, πρόβλεψη ενέργειας και συνάρτηση τιμής κατάστασης για τη λήψη μιας απόφασης σε ένα χώρο αναζήτησης με πολλές μη παρατηρήσιμες μεταβλητές. Οι προσομοιώσεις τους, στις οποίες ο πράκτορας του συστήματος αυτού έπαιξε εναντίον ενός πράκτορα που χρησιμοποιεί ένα μοντέλο βασισμένο σε κανόνες για τη λήψη αποφάσεων, έδειξαν ότι ο πράκτορας που παρουσιάζουν στο άρθρο τους είναι πιο ισχυρός και προσαρμόζεται ευκολότερα στην στρατηγική του αντιπάλου του.

Στο (Robiliard, 2014), οι Robilliard, Fonlupt et al. χρησιμοποιούν και αυτοί τον αλγόριθμο αναζήτησης δέντρου Monte-Carlo και συγκεκριμένα τον Upper Confidence Bounds Monte-Carlo Tree Search, ο οποίος έχει προσφέρει μεγάλες βελτιώσεις στους τεχνητούς πράκτορες πολλών παιχνιδιών. Εδώ ο αλγόριθμος χρησιμοποιείται σε ένα πιο περίπλοκο παιχνίδι το παιχνίδι 7 Wonders. Αυτό το παιχνίδι καρτών συγκεντρώνει πολλές ιδιότητες που είναι γνωστό πως προκαλούν δυσκολίες στην μοντελοποίηση του, όπως κρυμμένες πληροφορίες, πολλούς παίκτες και στοχαστικότητα. Περιλαμβάνει επίσης ένα στάδιο ανταλλαγής καρτών μεταξύ παικτών που χρησιμοποιεί συνδυαστική αναζήτηση για να αποφασιστεί ποιες αποφάσεις είναι επιτρεπτές. Επιπλέον, είναι δύσκολο να δημιουργηθεί μια αποτελεσματική λειτουργία αξιολόγησης αφού η αξία της κάθε κάρτας εξαρτάται σε μεγάλο βαθμό από το στάδιο του παιχνιδιού και τις αποφάσεις των υπολοίπων παικτών. Παρά τις προαναφερθείσες δυσκολίες τα αποτελέσματα είναι ενθαρρυντικά, αν και το συγκεκριμένο παιχνίδι δεν είναι τόσο συγγενές με τα κλασικά αφηρημένα παιχνίδια.

Στο (da Silva, 2018) , οι Alysson Ribeiro da Silva και Luis Fabricio Wanderley Goes μελετούν το ψηφιακό παιχνίδι συλλογής καρτών HearthStone ως ένα εν μέρει παρατηρήσιμο παιχνίδι με εναλλασσόμενους γύρους. Η κύρια πρόκληση είναι η δημιουργία στρατηγικών ικανών να υποδαυλίζουν τις κινήσεις του αντιπάλου. Από τη σκοπιά της Τεχνητής Νοημοσύνης ο χώρος όλων των πιθανών συνδυασμών δράσεων είναι τεράστιος και δυναμικός λόγω του αριθμού των καρτών, αλλά και της τυχαιότητας, γεγονός που καθιστά το σχεδιασμό αποδοτικών αυτόνομων πρακτόρων ένα δύσκολο πρόβλημα. Στο (da Silva, 2018) προτείνεται ένα προσαρμοστικό νευρωνικό δίκτυο εμπνευσμένο από το ασαφές ARAM και ARTMAP καθώς επίσης ένας νέος μηχανισμός για την ταξινόμηση και την πρόβλεψη πληροφοριών ώστε να αποφεύγεται η υπεργενίκευση του προβλήματος από αυτά τα δίκτυα. Επιπλέον, η προτεινόμενη λύση αξιολογήθηκε εναντίον ενός προσομοιωτή HearthStone κι ενός πράκτορα που χρησιμοποιεί αναζήτηση δέντρου Monte Carlo. Τα αποτελέσματα δείχνουν ότι η προτεινόμενη λύση κερδίζει 80% των φορών εναντίον γνωστών συνδυασμών καρτών και 70% εναντίον πιο σπανίων συνδυασμών καρτών.

Κεφάλαιο 3

Το παιχνίδι Tichu

- i) Τι είναι το Tichu;



3-1 Tichu logo. (Fata Morgana, 2013)

Το Tichu είναι ένα ομαδικό παιχνίδι συγκομιδής πόντων το οποίο παίζεται με κάρτες. Εμπεριέχει στοιχεία από το Bridge, το Poker και το Daifugō. Παίζεται από τέσσερις παίκτες οι οποίοι χωρίζονται σε δύο ομάδες των δύο παικτών. Υπάρχουν και παραλλαγές των τριών ή έξι παικτών, οι οποίες όμως δε θα μας απασχολήσουν. Σκοπός κάθε ομάδας είναι η συγκομιδή πόντων, καθώς η πρώτη ομάδα που θα φτάσει ένα προκαθορισμένο σύνολο πόντων (συνήθως τους 1000 πόντους) είναι και η νικητήρια ομάδα.

Το παιχνίδι αυτό προέρχεται από δύο κινέζικα παραδοσιακά παιχνίδια, καθώς είναι μια παραλλαγή του Dou Dizhu, Choi Dai Di ή Da Lao Er (δηλαδή του κινέζικου παιχνιδιού καρτών “big two”) σε συνδυασμό με το Zheng Fen (ενός άλλου κινέζικου παιχνιδιού συγκομιδής πόντων). (Wikipedia, 2018)

Ο ιδρυτής του Tichu είναι ο Urs Hostettler και το παιχνίδι εμφανίστηκε για πρώτη φορά το 1991. Ο Urs Hostettler τονίζει πως εμπνεύστηκε την ιδέα για το Tichu από ένα παιχνίδι που είδε να παίζουν στην πόλη της Nanjing στην Κίνα οδηγοί λεωφορείων. Σήμερα διανέμεται από την ελβετική εταιρία παιχνιδιών Fata Morgana. (Hostettler, 2004)

Εντάσσεται στην κατηγορία των Climbing Games (ιδιαίτερα διαδεδομένη κατηγορία παιχνιδιών στην Ασία και ειδικά στην Κίνα), αφού κάθε παίκτης στη σειρά του πρέπει να παίξει ένα υψηλότερο φύλλο ή συνδυασμό φύλλων από τον προηγούμενο παίκτη. Αν ο παίκτης αυτός δε μπορεί ή δεν επιθυμεί να “χτυπήσει” το προηγούμενο φύλλο ή συνδυασμό φύλλων μπορεί να δηλώσει “πάσο”. (Tabletopia, 2019)

Στην Ελλάδα το παιχνίδι αυτό γίνεται όλο και πιο δημοφιλές, κάτι το οποίο διαφαίνεται από το πλήθος των δημοσιεύσεων στο διαδίκτυο για το παιχνίδι αυτό σε ελληνικές ιστοσελίδες, καθώς και από τη συμμετοχή Ελλήνων παικτών σε διαδικτυακές πλατφόρμες παιχνιδιών όπου μπορούν να παίξουν Tichu διαδικτυακά με άλλους παίκτες. Η πιο δημοφιλής πλατφόρμα για Tichu στο διαδίκτυο είναι το γερμανικό (BrettSpielWelt), το οποίο φιλοξενεί χιλιάδες Έλληνες παίκτες Tichu (πάνω από το 50% των συνολικών παικτών Tichu στην εφαρμογή), καθώς και το (dod) , το οποίο αποτελεί τη μεγαλύτερη πλατφόρμα παιχνιδιών της ελληνικής κοινότητας και είναι και διαθέσιμο μέσα από το Facebook. Άλλες λιγότερο διαδεδομένες πλατφόρμες είναι το (Tichu) και το (Pro) (πρώην TichuIQ). Επίσης όλο και αυξανόμενη συμμετοχή παρατηρείται σε διάφορα τουρνουά Tichu που διοργανώνονται σε όλες τις πόλεις της Ελλάδας.

ii) Το παιχνίδι

Το Tichu είναι ένα παιχνίδι καρτών. Παίζεται από τέσσερις παίκτες οι οποίοι χωρίζονται σε δύο ομάδες των δύο παικτών. Στόχος της κάθε ομάδας είναι η συγκομιδή περισσότερων πόντων από την άλλη ομάδα. Όποια ομάδα φτάσει πρώτη στους χίλιους πόντους είναι η νικήτρια ομάδα. Αν και οι δύο ομάδες ξεπεράσουν τους χίλιους πόντους στον ίδιο γύρο, η ομάδα με τους περισσότερους πόντους κερδίζει. Στόχος κάθε παίκτη είναι να παίξει όλες τις κάρτες στο χέρι του. Σε κάθε γύρο το σύνολο των πόντων που δίνεται είναι 100 πόντοι, και αυτοί μοιράζονται στις δύο ομάδες ανάλογα με τη σειρά που έμειναν χωρίς φύλλα στα χέρια τους οι παίκτες. Ο παραπάνω κανόνες δεν ισχύει σε περίπτωση που ένας ή περισσότεροι παίκτες δηλώσει (είτε με δήλωση Grand Tichu, ή με δήλωση Tichu) πως πιστεύει ότι θα ξεφορτωθεί πρώτος όλες τις κάρτες του. Τότε, για μια επιτυχή δήλωση, δηλαδή αν ο παίκτης αυτός όντως καταφέρει να ξεφορτωθεί πρώτος το σύνολο των καρτών του, η ομάδα του, πέρα από το πως έχουν μοιραστεί οι 100 βαθμοί του γύρου, κερδίζει επιπλέον βαθμούς (200 για μια επιτυχημένη δήλωση Grand Tichu και 100 για μια επιτυχημένη δήλωση Tichu). Αντίστοιχα σε μια αποτυχημένη δήλωση η ομάδα του παίκτη χάνει βαθμούς (πάλι 200 και 100 βαθμούς ανάλογα με το αν έχει δηλωθεί Grand Tichu ή Tichu). Υπάρχει επίσης η περίπτωση οι παίκτες μιας ομάδας να ξεφορτωθούν τις κάρτες τους πριν και από τους δύο αντιπάλους τους. Στην περίπτωση αυτή η ομάδα τους παίρνει 200 αντί για 100 βαθμούς, ενώ η αντίπαλη ομάδα παίρνει 0 βαθμούς. Το παιχνίδι παίζεται δεξιόστροφα και η σειρά των παικτών στο τραπέζι είναι τέτοια που πάντα μετά από έναν παίκτη της μιας ομάδας παίζει παίκτης της αντίπαλης ομάδας. Σε κάθε γύρο τις κάρτες μοιράζει ένας παίκτης με δεξιόστροφη φορά, δίνοντας από μια κάρτα σε κάθε παίκτη ξεκινώντας από τον αμέσως επόμενο παίκτη. Η διαδικασία επαναλαμβάνεται μέχρι να έχουν μοιραστεί όλες οι κάρτες. Στον επόμενο γύρο μοιράζει ο επόμενος παίκτης. Κάθε γύρος χωρίζεται σε τέσσερα στάδια. Αυτά είναι: το στάδιο του μοιράσματος των πρώτων οκτώ φύλλων σε κάθε παίκτη, το στάδιο της δήλωσης του Grand Tichu, το στάδιο του μοιράσματος των 14 φύλλων και των ανταλλαγών και το παίξιμο του γύρου. (Tichurules Blogspot, 2012)

iii) Η τράπουλα



3-2 Τράπουλα Tichu, (Pillow Fights, 2016)

Η τράπουλα του Tichu αποτελείται από 56 κάρτες. Οι κάρτες χωρίζονται σε 52 κάρτες, που είναι αυτές που συναντάμε στα περισσότερα παιχνίδια καρτών και σε τέσσερις ειδικές κάρτες. Οι 52 κάρτες αυτές είναι οι αριθμοί 2,3,4,5,6,7,8,9,10 σε τέσσερα διαφορετικά χρώματα (κόκκινο, μαύρο, μπλε, πράσινο) και οι φιγούρες.

Φιγούρες είναι ο βαλές (Jack, απεικονίζεται ως J πάνω στην κάρτα), η ντάμα (Queen, απεικονίζεται ως Q πάνω στην κάρτα) και ο ρήγας (King, απεικονίζεται ως K πάνω στην κάρτα). Επίσης υπάρχει και ο άσος (Ace, απεικονίζεται ως A πάνω στην κάρτα). Οι φιγούρες και οι άσοι έχουν και αυτές τέσσερα διαφορετικά χρώματα (κόκκινο, μαύρο, μπλε, πράσινο), όπως συμβαίνει και με τους αριθμούς.

Οι τέσσερις ειδικές κάρτες του Tichu είναι το Mahjong, τα σκυλιά, ο φοίνικας και ο δράκος. Ο ρόλος τους θα εξηγηθεί στην επόμενη ενότητα.

Το κάθε χρώμα έχει και το δικό του σύμβολο πάνω στις κάρτες. Το κόκκινο έχει το αστέρι, το μαύρο έχει το σπαθί, το μπλε έχει τη παγόδα, ενώ το πράσινο έχει το διαμάντι.

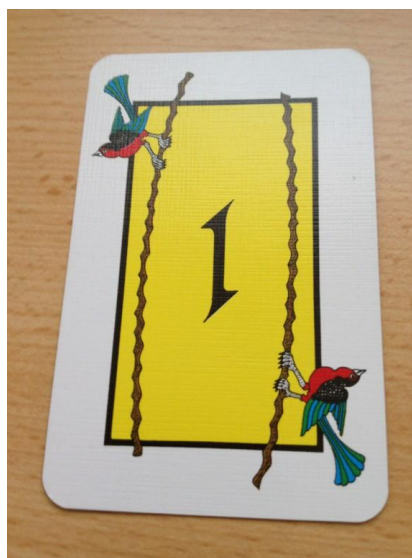
iv) Ειδικές Κάρτες



3-3 Ειδικές κάρτες του Tichu, (Lim, 2014)

Η τράπουλα Tichu εμπεριέχει τέσσερις ειδικές κάρτες, καθεμία από τις οποίες έχει συγκεκριμένες ιδιότητες που επηρεάζουν τη ροή του παιχνιδιού. Οι κάρτες αυτές είναι το Mahjong, τα σκυλιά, ο φοίνικας και ο δράκος. (Achro)

Το Mahjong



3-4 Mahjong, (Jesta, 2015)

Το Mahjong είναι το μικρότερο φύλλο της τράπουλας και δρα σαν το νούμερο ένα στο παιχνίδι. Μπορεί δηλαδή να παιχτεί είτε στην αρχή μιας κέντας πριν από το δύο (κέντα 1-5) είτε σαν το πιο αδύναμο μονό φύλλο. Ο παίκτης που έχει το Mahjong στα φύλλα του είναι αυτός που ξεκινάει το γύρο. Μπορεί να ρίξει οποιοδήποτε συνδυασμό θέλει (χωρίς απαραίτητα να ξεκινήσει με συνδυασμό που περιέχει το Mahjong).

Τη στιγμή που κάποιος παίκτης παίζει το Mahjong (είτε ως μονό φύλλο είτε ως κέντα) δικαιούται να ζητήσει ένα φύλλο το οποίο όμως δεν είναι μια άλλη ειδική κάρτα. Η επιθυμία του Mahjong πρέπει να εκπληρωθεί άμεσα, όταν βέβαια αυτό είναι δυνατόν.

Για παράδειγμα, έστω ότι το Mahjong παίζεται ως μονό φύλλο και ο παίκτης που το ρίχνει αποφασίζει να ζητήσει την κάρτα 3. Ο επόμενος παίκτης είναι υποχρεωμένος να ρίξει 3 εάν έχει έστω και ένα φύλλο 3 στη διάθεσή του. Εάν δεν έχει τότε μπορεί να ρίξει ότι θέλει ή να πάει πάσο. Ο τρίτος στη σειρά πρέπει τώρα αυτός να ρίξει 3 εάν έχει και μπορεί (δηλαδή εάν προηγουμένως δεν έχει πέσει φύλλο μεγαλύτερο του 3) και ούτω καθεξής. Σε περίπτωση που το Mahjong πέφτει σε κέντα και ζητείται ένα φύλλο τότε αυτό το φύλλο πρέπει να πέσει σε κέντα ίδιου αριθμού φύλλων, εάν φυσικά αυτό είναι δυνατόν.

Η επιθυμία του Mahjong παραμένει έως ότου είναι δυνατόν να εκπληρωθεί.

Να σημειώσουμε ότι οι βόμβες (συνδυασμός που θα αναφερθούμε στην ενότητα νί αυτού του κεφαλαίου) μπορούν να παιχτούν ανά πάσα στιγμή και πάνω σε οποιονδήποτε συνδυασμό πράγμα που σημαίνει ότι εάν κάποιος έχει βόμβα στο ζητούμενο φύλλο μπορεί να χρειαστεί να την παίξει. Στο προηγούμενο παράδειγμα, αν ο τρίτος παίκτης στη σειρά είχε μια βόμβα στο 3 (δηλαδή είχε στα φύλλα του και τα τέσσερα φύλλα 3), τότε θα ήταν αναγκασμένος να ρίξει τη βόμβα αν ο δεύτερος στη σειρά είχε παίξει φύλλο υψηλότερο του 3. Αλλιώς θα είχε την επιλογή του να παίξει ένα φύλλο 3 ή ολόκληρη τη βόμβα μαζί.

Τα σκυλιά



3-5 Τα σκυλιά, (Jesta, 2015)

Τα σκυλιά είναι η ειδική κάρτα η οποία μπορεί να παιχτεί από τον παίκτη που έχει «πάρει χέρι», δηλαδή αυτόν τον οποίο ορίζει τον συνδυασμό που πρόκειται να παιχτεί. Αντί λοιπόν να ορίσει έναν καινούριο συνδυασμό, μπορεί να ρίξει τα σκυλιά δίνοντας τη θέση του στο συμπαίκτη του, ο οποίος υποχρεούται πλέον να ορίσει αυτός το συνδυασμό που θα παιχτεί, όπως αυτός επιθυμεί.

Εάν ο συμπαίκτης έχει βγει τότε αυτός που ορίζει τον νέο συνδυασμό είναι ο επόμενος παίκτης από τη θέση όπου βρισκόταν ο συμπαίκτης του. Σε περίπτωση που και αυτός έχει βγει τότε το νέο συνδυασμό τον διαλέγει πάλι ο παίκτης που χρησιμοποίησε τα σκυλιά.

Τα σκυλιά είναι το μόνο φύλλο πάνω στο οποίο δεν μπορεί να χρησιμοποιηθεί μια βόμβα.

Ο δράκος



3-6 Ο δράκος, (Jesta, 2015)

Ο δράκος είναι το υψηλότερο μονό φύλλο και μπορεί να παιχτεί μόνο όταν παίζονται μονά φύλλα. Μπορεί να χτυπηθεί μόνο από βόμβα. Όταν κάποιος παίκτης παίρνει τη μπάζα με το δράκο, δηλαδή παίξει το δράκο και μετά όλοι οι υπόλοιποι παίκτες κάνουν πάσο, τότε τη μπάζα αυτή την δίνει σε έναν από τους αντίπαλους παίκτες, αυτόν που πιθανολογεί ότι θα βγει τελευταίος (σε αυτόν, δηλαδή, που θα μείνει μόνος με φύλλα πάνω του στο τέλος της παρτίδας).

Ο φοίνικας



3-7 Ο φοίνικας, (Jesta, 2015)

Ο φοίνικας χρησιμοποιείται ως μπαλαντέρ. Μπορεί να χρησιμοποιηθεί στη θέση οποιουδήποτε κλασικού φύλλου (όχι δηλαδή ειδικού) και να συμπληρώσει οποιονδήποτε συνδυασμό εκτός από βόμβα. Όταν παίζεται ως μονό φύλλο τότε η αξία του είναι μισή παραπάνω από το τελευταίο μονό φύλλο που έχει παιχτεί. Εάν δηλαδή παιχτεί πάνω από 6 τότε η αξία του είναι $6+\frac{1}{2}$, εάν παιχτεί πάνω από Άσσο τότε η αξία του είναι $A+\frac{1}{2}$. Ο φοίνικας δεν μπορεί να παιχτεί πάνω από τον Δράκο στα μονά φύλλα, και πάνω από την υψηλότερη δυνατή κλασική κέντα (10,J,Q,K,A).

ν) Δυναμικότητα Φύλλων

Πριν προχωρήσουμε στην παρουσίαση όλων των συνδυασμών που μπορούν να παίξουν οι παίκτες στο Tichu, θα αναλύσουμε τη δυναμικότητα των διαφόρων φύλλων.

Το Mahjong, θεωρείται ως το πιο αδύναμο μονό φύλλο (έστω δυναμικότητα 1) ή η πιο αδύναμη αρχή μιας κέντας και δε μπορεί να χρησιμοποιηθεί ως τίποτε άλλο.

Οι αριθμοί 2 έως 10 έχουν την δυναμικότητα που υποδηλώνει ο αριθμός τους. Ο Βαλές (J) είναι το αμέσως πιο δυνατό φύλλο από το 10 (έστω δυναμικότητα 11), η Ντάμα (Q) το αμέσως πιο δυνατό φύλλο από το βαλέ (έστω δυναμικότητα 12) ενώ ο Ρήγας (K) το αμέσως πιο δυνατό φύλλο από τη ντάμα (έστω δυναμικότητα 13). Ο Άσσος (A) είναι πιο δυνατός από τον ρήγα και γενικά το πιο δυνατό μη-ειδικό φύλλο (έστω δυναμικότητα 14). Ο φοίνικας σαν μονό φύλλο έχει δυναμικότητα $+0.5$ πάνω από το φύλλο που χτυπάει. Εάν δεν υπάρχει άλλο φύλλο κάτω όταν πέσει, τότε έχει δυναμικότητα 1.5. Ο φοίνικας δε μπορεί να χτυπήσει τον δράκο. Ο Δράκος είναι το πιο δυνατό μονό φύλλο (έστω δυναμικότητα 15).

vi) Συνδυασμοί

Υπάρχουν διαφόρων ειδών συνδυασμοί στο Tichu. Οι περισσότεροι από αυτούς συναντιούνται και σε άλλα παιχνίδια και είναι διαδεδομένοι στο ευρύ κοινό.

Οι συνδυασμοί αυτοί είναι οι εξής:

Μονοφυλλία: ένα μονό φύλλο. Η δυναμικότητα των φύλλων περιγράφηκε στην αμέσως προηγούμενη ενότητα. Ο δράκος μπορεί να παιχτεί μόνο σαν μονό φύλλο.

Διφυλλία (ζευγάρι): Δύο φύλλα μπορούν να παιχτούν σαν ένα ζευγάρι. Απαραίτητη προϋπόθεση είναι τα δύο φύλλα αυτά να είναι ίδια (σε δυναμικότητα), ή ένα από αυτά να είναι ο φοίνικας. Η δυναμικότητα των ζευγαριών προκύπτει από τη δυναμικότητα των φύλλων που αποτελούν το ζευγάρι (ο φοίνικας εδώ έχει τη δυναμικότητα του άλλου φύλλου του ζευγαριού, δηλαδή το ζευγάρι 2-Φοίνικας είναι πιο αδύναμο από το ζευγάρι 3-3).

Τριφυλλία (τριπλέτα): Τρία φύλλα μπορούν να παιχτούν σαν μια τριπλέτα. Απαραίτητη προϋπόθεση είναι τα τρία φύλλα αυτά να είναι ίδια (σε δυναμικότητα), ή ένα από αυτά να είναι ο φοίνικας. Η δυναμικότητα των τριπλετών προκύπτει από τη δυναμικότητα των φύλλων που αποτελούν την τριπλέτα (ο φοίνικας εδώ έχει τη δυναμικότητα των άλλων φύλλων της τριπλέτας, δηλαδή η τριπλέτα 9-9-Φοίνικας είναι πιο αδύναμη από την τριπλέτα 10-10-10).

Κέντα: Κέντα είναι πέντε ή παραπάνω φύλλα διαδοχικής δυναμικότητας. Η κέντα ορίζεται από το πλήθος των φύλλων της και τη δυναμικότητα του μεγαλύτερου της φύλλου. Μια κέντα έξι φύλλων (έστω 4-5-6-7-8-9) είναι πιο αδύναμη από την κέντα έξι φύλλων 9-10-J-Q-K-A. Μια κέντα n – φύλλων (όπου $n \geq 5$) μπορεί να νικηθεί μόνο από μια κέντα n – φύλλων όπου το μεγαλύτερο φύλλο της έχει υψηλότερη δυναμικότητα από το υψηλότερο φύλλο της πρώτης.

Full House: Το Full House είναι ένας συνδυασμός πέντε φύλλων που αποτελείται από μια τριπλέτα και ένα ζευγάρι. Προφανώς η τριπλέτα και το ζευγάρι είναι διαφορετικής δυναμικότητας. Η δυναμικότητα του Full House ορίζεται από τη δυναμικότητα της τριπλέτας

αυτού. Δηλαδή ένα Full House 4-4-4-2-2 είναι πιο δυνατό από ένα Full House 3-3-3-A-A, μιας και η τριπλέτα 4-4-4 είναι πιο δυνατή από την τριπλέτα 3-3-3.

Steps: Ένας παίκτης μπορεί να παίζει επίσης όσα συνεχόμενης δυναμικότητας ζευγάρια έχει σαν έναν κοινό συνδυασμό. Τα διαδοχικά ζευγάρια, όπως και οι κέντες, ορίζονται και αυτά από το πλήθος των ζευγαριών και τη δυναμικότητα του μεγαλύτερου ζευγαριού. Για παράδειγμα τα διαδοχικά ζευγάρια 6-6-7-7-8-8 είναι πιο αδύναμα από τα διαδοχικά ζευγάρια 7-7-8-8-9-9, αφού το 9 έχει υψηλότερη δυναμικότητα από το 8. Όπως και στις κέντες n το πλήθος διαδοχικά ζευγάρια μπορεί να νικηθεί μόνο από n το πλήθος διαδοχικά ζευγάρια όπου το μεγαλύτερο ζευγάρι έχει υψηλότερη δυναμικότητα από το μεγαλύτερο ζευγάρι των πρώτων.

Βόμβα: Η βόμβα είναι ένας ιδιαίτερος συνδυασμός στο Tichu, με την έννοια ότι μπορεί να παιχτεί ακόμα και εκτός σειράς του παίκτη (δηλαδή ακόμα και όταν δεν παίζει ο ίδιος κάποιον συνδυασμό) πάνω σε οποιοδήποτε φύλλο (εκτός από την ειδική κάρτα Σκυλιά) και κερδίζει οποιοδήποτε από τους παραπάνω συνδυασμούς - εκτός βέβαια από βόμβα υψηλότερης δυναμικότητας. Μια βόμβα αποτελείται είτε από τέσσερα ίδια δυναμικότητας φύλλα (για παράδειγμα 2-2-2-2 και ονομάζεται Βόμβα Καρέ) ή από μια κέντα φύλλων ίδιου χρώματος (για παράδειγμα 3-4-5-6-7 κόκκινου χρώματος). Η δυναμικότητα των βομβών ορίζεται με τον εξής τρόπο: μια βόμβα με τέσσερα ίδια φύλλα είναι λιγότερο δυνατή από οποιαδήποτε βόμβα κέντας ίδιου χρώματος, ενώ για τις βόμβες κέντας ίδιου χρώματος ίδιων φύλλων ισχύει ότι και για τις κανονικές κέντες (δηλαδή η βόμβα κέντα ίδιου χρώματος όπου έχει υψηλότερη δυναμικότητα το μεγαλύτερο φύλλο της υπερτερεί). Ανάμεσα σε μια βόμβα κέντα ίδιου χρώματος n φύλλων και μια βόμβα κέντα ίδιου χρώματος m φύλλων, όπου $m > n$, πιο δυνατή είναι η βόμβα με τα m φύλλα. Για παράδειγμα η βόμβα 2-2-2-2 είναι λιγότερη δυνατή από τη βόμβα 4-4-4-4, η οποία με τη σειρά της είναι λιγότερη δυνατή από τη βόμβα 8-9-10-J-Q ίδιου χρώματος, η οποία με τη σειρά της είναι λιγότερο δυνατή από τη βόμβα 10-J-Q-K-A ίδιου χρώματος, η οποία με τη σειρά της είναι λιγότερο δυνατή από τη βόμβα 5-6-7-8-9-10 ίδιου χρώματος.

vii) Δηλώσεις Grand Tichu και Tichu

Το Grand Tichu και το Tichu είναι οι δύο δηλώσεις με τις οποίες κάποιος παίκτης μπορεί να μαζέψει παραπάνω πόντους για την ομάδα του. Το ρίσκο το οποίο ενέχουν, μιας και σε περίπτωση που ο παίκτης που τα δηλώσει δε ξεφορτωθεί πρώτος τις κάρτες του η ομάδα του θα δεχτεί ποινή αρνητικών πόντων ίση με τους παραπάνω πόντους που θα έπαιρνε η ομάδα αν ξεφορτωνόταν πρώτος τις κάρτες του, σε συνδυασμό με το μεγάλο προβάδισμα το οποίο αυτές οι δηλώσεις μπορούν να δώσουν, καθιστούν τις δηλώσεις αυτές κομβικές για την πορεία του παιχνιδιού και την ανάδειξη της νικητήριας ομάδας.

Grand Tichu μπορεί να δηλώσει κάποιος παίκτης, όταν έχει δει μόνο τα οκτώ πρώτα του φύλλα. Από τη στιγμή λοιπόν που θα μοιραστούν από οκτώ φύλλα σε όλους τους παίκτες μέχρι τη στιγμή που θα δουν τα υπόλοιπα φύλλα τους, ένας ή και περισσότεροι παίκτες μπορούν να δηλώσουν Grand Tichu. Ο παίκτης που κάνει τη δήλωση Grand Tichu πιθανολογεί πως θα είναι ο πρώτος που θα μείνει χωρίς φύλλα. Οι πόντοι που προστίθενται ή αφαιρούνται στο τέλος του γύρου αν καταφέρει να μείνει πρώτος χωρίς φύλλα ή όχι αντίστοιχα είναι 200 πόντοι.

Κάθε παίκτης οποιαδήποτε στιγμή πριν παίξει το πρώτο του φύλλο (δηλαδή όσο έχει 14 φύλλα στο χέρι του) μπορεί να δηλώσει Tichu. Αυτό σημαίνει ότι πιθανολογεί πως θα είναι αυτός ο οποίος θα μείνει πρώτος χωρίς φύλλα. Εάν αυτό γίνει τότε η ομάδα του κερδίζει επιπλέον 100 πόντους. Σε αντίθετη περίπτωση η ομάδα του χάνει 100 πόντους.

Εδώ αξίζει να τονίσουμε ότι οι πόντοι του Grand Tichu και του Tichu είναι επιπλέον πόντοι που προστίθενται στους 100 πόντους του κάθε γύρου.

viii) Ανταλλαγές

Μετά τη δήλωση ή μη Grand Tichu και αφού μοιραστούν και στους τέσσερις παίκτες τα 14 φύλλα ακολουθεί η φάση των ανταλλαγών. Σε αυτήν τη φάση κάθε παίκτης είναι αναγκασμένος να δώσει από ένα φύλλο σε κάθε άλλον παίκτη (ένα στο συμπαίκτη του και από ένα σε κάθε αντίπαλό του) και να λάβει αντίστοιχα τρία φύλλα (ένα από το συμπαίκτη του και ένα από κάθε αντίπαλο του). Σε όλη αυτή τη διαδικασία ο κάθε παίκτης δεν έχει καμία πληροφορία για τα φύλλα που θα λάβει, μέχρις ότου δώσει φύλλα στους άλλους παίκτες.

ix) Γενικοί Κανόνες

Αφού τελειώσει η φάση των ανταλλαγών, ξεκινάει η φάση του παιχνιδιού. Αρχίζει πρώτος ο παίκτης που έχει το Mahjong. Μπορεί να παίξει (ή ρίξει) οποιοδήποτε συνδυασμό θέλει. Οι υπόλοιποι παίκτες παίζουν με δεξιόστροφη φορά και μπορούν να ρίξουν είτε υψηλότερο ίδιο συνδυασμό ή βόμβα ή να πουν πάσο. Ο παίκτης ο οποίος “κερδίζει” είναι αυτός ο οποίος θα πετάξει τελευταίος συνδυασμό και οι υπόλοιποι τρεις παίκτες θα πάνε πάσο. Τότε, ο παίκτης αυτός μαζεύει τα παιγμένα φύλλα στη μπάζα του (μόνο στη περίπτωση που κερδίσει χέρι με το Δράκο δίνει τη μπάζα αλλού, όπως εξηγήθηκε παραπάνω) και παίζει πρώτος οποιοδήποτε συνδυασμό θέλει. Στόχος του παιχνιδιού για κάθε παίκτη είναι να απαλλαγεί από όλα του τα φύλλα μαζεύοντας, ταυτόχρονα, όσο το δυνατόν παραπάνω πόντους. Ο γύρος τελειώνει είτε όταν οι δύο παίκτες της ίδιας ομάδας ξεφορτωθούν όλα τα φύλλα τους πριν και από τους δύο παίκτες της αντίπαλης ομάδας - είτε όταν τρεις από τους τέσσερις παίκτες ξεφορτωθούν όλα τα φύλλα τους.

Έστω ένα παράδειγμα, όπου οι παίκτες 1 και 3 είναι συμπαίκτες και οι παίκτες 2 και 4 είναι οι αντίπαλοί τους και ξεκινάει ο Παίκτης 2 είναι το εξής: ο Παίκτης 2 παίζει ζευγάρι 4-4, ο Παίκτης 3 παίζει 6-6, ο Παίκτης 4 παίζει 9-9, ο Παίκτης 1 παίζει Q-Q, ο Παίκτης 2 K-K και ο Παίκτης 3 παίζει A-φοίνικα (αντιστοιχεί σε ζευγάρι A-A). Οι παίκτες 4,1 και 2 πάνε πάσο διαδοχικά και ο Παίκτης 3 παίρνει όλα τα φύλλα που παίχθηκαν στη μπάζα του. Έπειτα, είναι σειρά του να ρίξει όποιον συνδυασμό θέλει.

χ) Πόντοι

Ο γύρος σταματάει σε δύο περιπτώσεις. Η πρώτη περίπτωση είναι όταν οι τρεις από τους τέσσερις παίκτες ξεφορτωθούν όλα τους τα φύλλα. Τότε μετριούνται οι πόντοι των φύλλων της συνολικής μπάζας της κάθε ομάδας.

Οι πόντοι από τις κάρτες στο Tichu είναι 100 και είναι οι παρακάτω:

Πεντάρι (5): 5 πόντοι

Δεκάρι (10): 10 πόντοι

Ρήγας (K): 10 πόντοι

Δράκος: 25 πόντοι

Φοίνικας: -25 πόντοι

Όλα τα υπόλοιπα φύλλα δεν έχουν πόντους.

Ο τελευταίος παίκτης δίνει τους πόντους που έχει στο χέρι του στην αντίπαλη ομάδα και τους πόντους που έχει στη μπάζα του στην ομάδα που ανήκει ο παίκτης που ξεφορτώθηκε πρώτος όλα του τα φύλλα (μπορεί δηλαδή να είναι και ο συμπαίκτης του).

Η δεύτερη περίπτωση είναι όταν βγουν πρώτος και δεύτερος οι δύο παίκτες μιας ομάδας (λέγεται και “ένα – δύο”). Τότε δεν μετριούνται οι πόντοι και η ομάδα αυτή παίρνει αυτόματα 200 πόντους ενώ η άλλη 0.

Κεφάλαιο 4

Εφαρμογές της Τεχνητής Νοημοσύνης και της Μηχανικής Μάθησης για το Tichu

Σε ότι αφορά το παιχνίδι Tichu, έχουν γίνει προσπάθειες δημιουργίας ευφών τεχνικών (χρήση τεχνητής νοημοσύνης και μηχανικής μάθησης) είτε στη λήψη αποφάσεων, είτε στην υλοποίηση ενός ευφούς παίκτη Tichu.

i) Το παιχνίδι Mini Tichu και ευφείς υλοποιήσεις του

Το παιχνίδι Mini Tichu (όπως ονομάστηκε από την Μάρθα Βλάχου-Κογχυλάκη) είναι μια απλοποιημένη μορφή του Tichu. Στο Mini Tichu δεν περιλαμβάνονται οι τέσσερις ειδικές κάρτες. Αυτό διευκολύνει την ανάπτυξη ευφών τεχνικών υλοποίησης του παιχνιδιού μιας και το πλήθος των διαφορετικών επιλογών μειώνεται. Μιας και δεν υπάρχει το φύλλο Mahjong για να καθορίσει ποιος παίζει πρώτος, αυτό γίνεται μέσω μιας τυχαίας συνάρτησης.

Για το παιχνίδι αυτό έχουν παρουσιαστεί δύο διαφορετικές προσεγγίσεις για τη δημιουργία ευφών πρακτόρων. Η μία είναι αυτή της Μάρθας Βλάχου-Κογχυλάκη στο (Βλάχου-Κογχυλάκη, 2012), ενώ η δεύτερη του Koen-Castelein στο (Castelein, 2017).

Ανάπτυξη Παιχνιδιού Στρατηγικής Μερικής Πληροφόρησης για το παιχνίδι Mini Tichu

Στο (Βλάχου-Κογχυλάκη, 2012) η διαδικασία δημιουργίας ενός ευφυούς πράκτορα ξεκινά με την εύρεση όλων των πιθανών συνδυασμών που μπορούν να προκύψουν από τις κάρτες τις οποίες έχει στο χέρι του τη στιγμή που καλείται να πάρει την απόφαση (δηλαδή να παίξει). Μέσα από αυτήν τη διαδικασία δημιουργείται ένα δέντρο αναζήτησης. Κάθε κόμβος του δέντρου αυτού αποτελείται από τρία στοιχεία. Πρώτο στοιχείο είναι οι κάρτες οι οποίες δεν ανήκουν σε κάποιο συνδυασμό. Δεύτερο στοιχείο είναι μια λίστα με συνδυασμούς που προκύπτουν από τις κάρτες που έχει στο χέρι του ο παίκτης και τρίτο ένας δείκτης στον πατέρα του κόμβου αυτού. Η δημιουργία του δέντρου γίνεται αναδρομικά, δηλαδή για κάθε κατάσταση (κάρτες τις οποίες έχει στο χέρι του εκείνη τη στιγμή ο παίκτης) δημιουργούνται τόσα παιδιά όσα το πλήθος των διαφορετικών λιστών συνδυασμών και καρτών που δεν ανήκουν σε κάποιο συνδυασμό, που μπορούν να δημιουργηθούν. Όταν η αναζήτηση ολοκληρωθεί διαγράφονται οι κόμβοι που είναι ίδιοι και στη συνέχεια οι κόμβοι βαθμολογούνται ανάλογα με την κεντρική αξία κάθε συνδυασμού, με το πλήθος των φύλλων των συνδυασμών, με το ειδικό βάρος των συνδυασμών και κανονικοποιούνται με βάση το πλήθος των συνδυασμών του κόμβου ή το πλήθος των καρτών που έχει ο παίκτης εκείνη τη στιγμή. Ένας τελευταίος τρόπος αξιολόγησης είναι η μέτρηση των πιθανών συνδυασμών που απομένουν στο παιχνίδι και το πόσο δυνατοί είναι οι συνδυασμοί του κόμβου σε σχέση με αυτούς.

Για τη φάση των ανταλλαγών χρησιμοποιείται ο εξής τρόπος εύρεσης των τριών φύλλων. Στο συμπαίκτη δίνεται το μεγαλύτερης δυναμικότητας φύλλο εκτός συνδυασμού και εάν αυτό δεν υπάρχει, τότε δίνεται απλώς το φύλλο με τη μεγαλύτερη δυναμικότητα. Για τα δύο φύλλα που δίνονται στους αντιπάλους, αυτά προκύπτουν από τους πιθανούς κόμβους που προκύπτουν από την αφαίρεση κάθε πιθανής δυάδας φύλλων δυναμικότητας από 2 έως και J (Βαλέ). Ο παίκτης τελικά επιλέγει το δέντρο στο οποίο ανήκει ο καλύτερος τερματικός κόμβος και δίνει στους αντιπάλους τις κάρτες που αφαιρέθηκαν για να καταλήξει σε αυτόν.

Το εάν ένας παίκτης θα παίξει ή όχι έναν επιτρεπόμενο συνδυασμό εξαρτάται από τη διαφορά που θα προκύψει ανάμεσα στον καλύτερο δυνατό τερματικό κόμβο και σε αυτόν που θα προκύψει άμα ο παίκτης παίξει τον συνδυασμό. Αυτό γίνεται πολλαπλασιάζοντας την

τιμή του καλύτερου κόμβου με έναν συντελεστή μικρότερο της μονάδας (πολλαπλασιαστής ανεκτικότητας). Εάν η αξία του ελαττωμένου κόμβου είναι μεγαλύτερη του κατωφλίου που προκύπτει από τον προηγούμενο πολλαπλασιασμό και ικανοποιούνται τα κριτήρια της συνεργατικότητας τότε ο συνδυασμός μπορεί να αφαιρεθεί. Ο πράκτορας, δηλαδή, παίζει αυτόν τον συνδυασμό. Σε διαφορετική περίπτωση, ο πράκτορας θα πάει πάσο.

Τα βάρη των διαφόρων συνδυασμών, η συνεργατικότητα μεταξύ των συμπαικτών (πότε δηλαδή απαγορεύεται ένας συμπαίκτης να παίζει πάνω από έναν συνδυασμό, όταν ο τελευταίος συνδυασμός που έχει παιχτεί είναι από το συμπαίκτη του) και το όριο ανεκτικότητας του πράκτορα είναι παράμετροι των οποίων οι τιμές ορίζονται αυθαίρετα και μπορούν να αλλάζουν ώστε να επιτευχθούν καλύτερα αποτελέσματα.

Τέλος μπορεί να επιλεγθεί ένας από τους παρακάτω τρόπους αξιολόγησης των κόμβων ανάμεσα στους: καμία αξιολόγηση (τυχαίος τρόπος παιχνίματος), απλή αξιολόγηση κόμβων, αξιολόγηση με βάση τον αριθμό των καρτών του κόμβου, αξιολόγηση με βάση τον αριθμό των συνδυασμών του κόμβου και αξιολόγηση με βάση τον αριθμό των κρυφών συνδυασμών.

Στη συνέχεια, για κάθε συνδυασμό (κέντα, τέσσερα όμοια φύλλα, full house, διαδοχικά ζευγάρια, ζευγάρια) το ειδικό βάρος ενός εκ των συνδυασμών άλλαξε από 0.2 έως και 1.6, ενώ τα βάρη των άλλων συνδυασμών έμειναν σταθερά στις αρχικοποιημένες τιμές τους όπως φαίνονται παρακάτω.

Συνδυασμός	Ειδικός Βάρος Συνδυασμού
Μονό φύλλο	0.5
Ζευγάρι φύλλων	0.4
Τριπλέτα φύλλων	0.3
Διαδοχικά Ζευγάρια φύλλων	0.4
Full House	0.5
Κέντα	0.5

Βόμβα Καρέ	4
Βόμβα Κέντα ίδιου χρώματος	4
Ρήγας	0.7
Άσσος	0.8

Έπειτα καταγράφηκε το πλήθος των εμφανίσεων του συνδυασμού αυτού στον καλύτερο κόμβο μέσα σε 500 διαφορετικά μοιράσματα φύλλων (παρτίδες). Η συνάρτηση αξιολόγησης των κόμβων που χρησιμοποιήθηκε ήταν η απλή:

$$\begin{aligned}
 \text{Αξία Κόμβου} = & \sum_{i=1}^{\# \text{ συνδυασμ \acute{o}ν}} \text{κεντρική αξία συνδυασμού} \times \# \text{ φύλλων συνδυασμού} \\
 & \times \text{ειδικό βάρος συνδυασμού}
 \end{aligned}$$

Για τον υπολογισμό της κεντρικής αξίας συνδυασμού χρησιμοποιήθηκε ο εξής πίνακας τιμών:

Συνδυασμός	Κεντρική Αξία Συνδυασμού
Μονό φύλλο	Αξία μονού φύλλου
Ζευγάρι φύλλων	Αξία ενός φύλλου του ζευγαριού
Τριπλέτα φύλλων	Αξία ενός φύλλου της τριπλέτας
Διαδοχικά Ζευγάρια φύλλων	Αξία ενός φύλλου από το υψηλότερο ζευγάρι
Full House	Αξία ενός φύλλου από την τριπλέτα του Full House
Κέντα	Αξία του υψηλότερου φύλλου της κέντας

Βόμβα Καρέ	Αξία ενός φύλλου από το καρέ
Βόμβα Κέντα ίδιου χρώματος	Αξία του υψηλότερου φύλλου της κέντας ίδιου χρώματος
Ρήγας	Αξία του Ρήγα
Άσσοι	Αξία του Άσου

Παρατηρήθηκε ότι όσο αυξάνεται ο συντελεστής τόσο η εμφάνιση του συνδυασμού στον καλύτερο κόμβο πλησιάζει την πιθανότητα εμφάνισης του συνδυασμού στο χέρι του παίκτη. Μπορεί με αυτόν τον τρόπο να επιλεγεί το βάρος που χρειάζεται ένας συνδυασμός για να είναι πάντα η επιλογή του πράκτορα, διαλέγοντας απλά την τιμή ειδικού βάρους συνδυασμού στην οποία η επιλογή του συνδυασμού στον καλύτερο κόμβο φτάνει ή τείνει να φτάσει την πιθανότητα εμφάνισης του συνδυασμού στο χέρι του παίκτη.

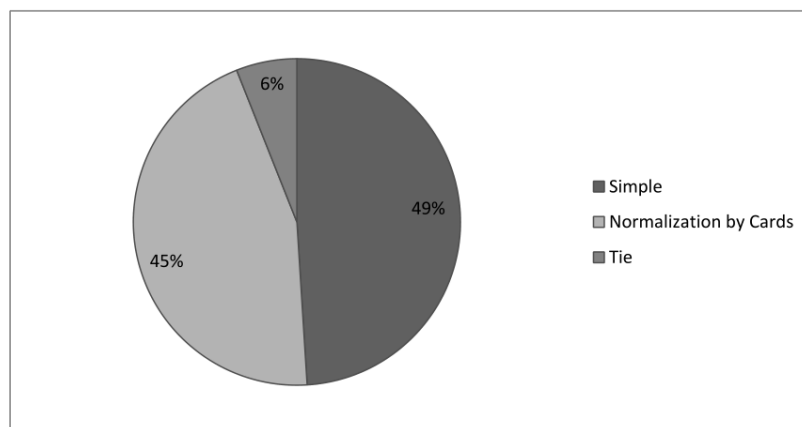
Να σημειωθεί ότι στις παραπάνω μετρήσεις υπάρχουν σφάλματα λόγω του μικρού πλήθους των παρτίδων που καταγράφηκαν. Συνεπώς, λόγω της μικρής πιθανότητας εμφάνισης του συνδυασμού κέντα-χρώμα τα αποτελέσματα που θα εμφανίζονταν για 500 μοιράσματα φύλλων θα είχαν μεγαλύτερη απόκλιση από την πραγματικότητα σε σχέση με τους άλλους συνδυασμούς οι οποίοι έχουν μεγαλύτερη πιθανότητα εμφάνισης. Για το λόγο αυτό δεν έχουν γίνει μετρήσεις για την εμφάνιση του συνδυασμού αυτού.

Μετά εξετάστηκε η επιθετικότητα του παίκτη. Το αποτέλεσμα που προέκυψε είναι ότι όσο μικρότερος είναι ο συντελεστής ανεκτικότητας του πράκτορα, τόσο πιο επιθετικός γίνεται ο πράκτορας (δηλαδή παίζει περισσότερες φορές από ότι πάει πάσο).

Τέλος μελετήθηκε το ποσοστό νικών δύο ομάδων οι οποίες έπαιξαν ως αντίπαλες χρησιμοποιώντας τους τρόπους αξιολόγησης που αναφέρθηκαν παραπάνω. Το μέγεθος του δείγματος ήταν 100 γύροι. Σε κάθε ομάδα οι δύο παίκτες χρησιμοποιούν τον ίδιο τρόπο αξιολόγησης των καρτών τους. Τα ονόματα των ομάδων δόθηκαν με βάση τον τρόπο αξιολόγησης που χρησιμοποιούσαν εκείνη τη στιγμή. Άρα έχουμε την ομάδα Simple η οποία χρησιμοποιεί την απλή αξιολόγηση κόμβων, την ομάδα Normalization by Cards η οποία

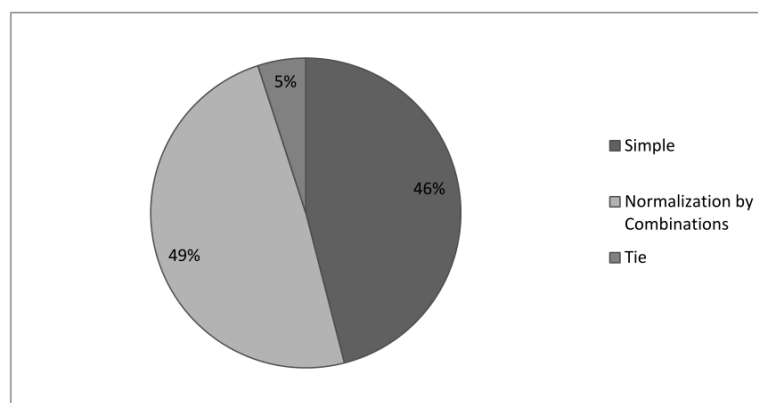
χρησιμοποιεί την αξιολόγηση με βάση τον αριθμό των καρτών του κόμβου, την Normalization by Combinations η οποία χρησιμοποιεί την αξιολόγηση με βάση τον αριθμό των συνδυασμών του κόμβου και την Normalization by Possibilities η οποία αξιολογεί τους κόμβους με βάση τον αριθμό των κρυφών συνδυασμών. Τα αποτελέσματα κάθε ομάδας παρουσιάζονται στα παρακάτω σχήματα.

- Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των καρτών του κόμβου



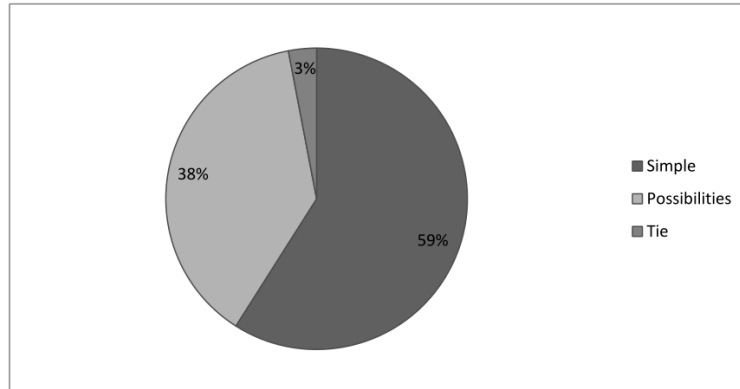
4-1 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των καρτών του κόμβου, (Βλάχου-Κογχυλάκη, 2012, p. 73)

- Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου



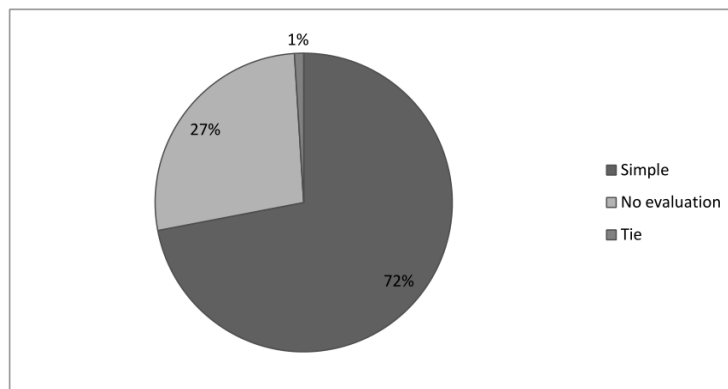
4-2 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου, (Βλάχου-Κογχυλάκη, 2012, p. 74)

- Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών



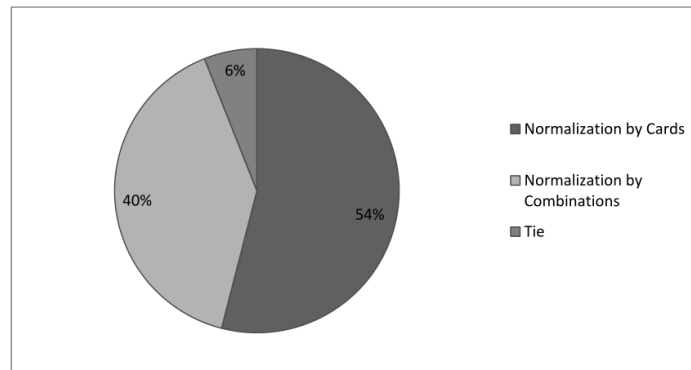
4-3 Απλή αξιολόγηση εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών, (Βλάχου-Κογχυλάκη, 2012, p. 74)

- Απλή αξιολόγηση εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)



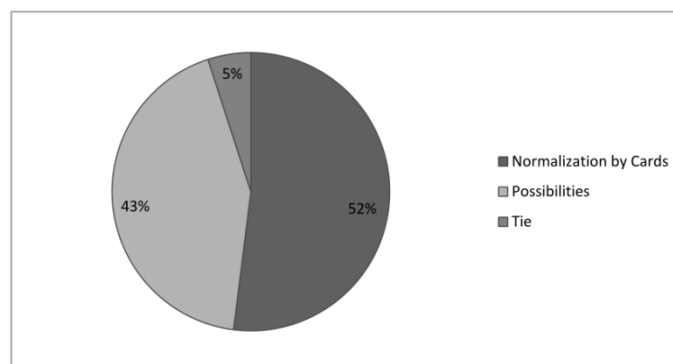
4-4 Απλή αξιολόγηση εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα), (Βλάχου-Κογχυλάκη, 2012, p. 75)

- Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου



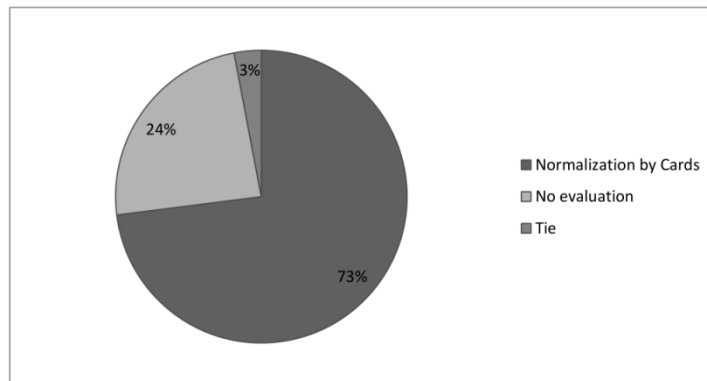
4-5 Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των συνδυασμών του κόμβου, (Βλάχου-Κογχυλάκη, 2012, p. 75)

- Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών



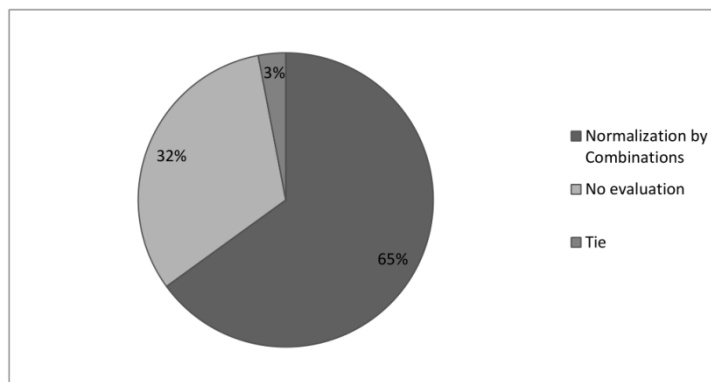
4-6 Αξιολόγηση με βάση τον αριθμό των καρτών εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών, (Βλάχου-Κογχυλάκη, 2012, p. 76)

- Αξιολόγηση με βάση τον αριθμό των καρτών του κόμβου εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)



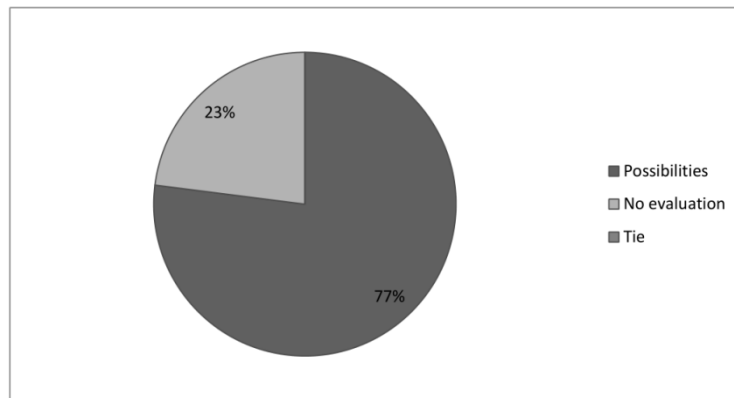
4-7 Αξιολόγηση με βάση τον αριθμό των καρτών του κόμβου εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα), (Βλάχου-Κογκυλάκη, 2012, p. 76)

- Αξιολόγηση με βάση τον αριθμό των συνδυασμών του κόμβου εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών



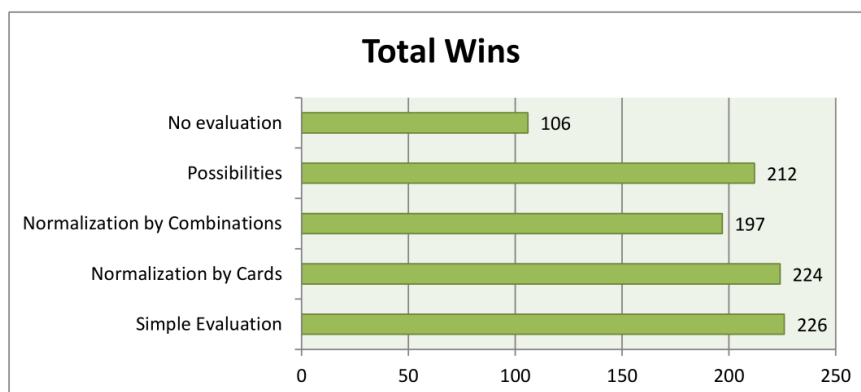
4-8 Αξιολόγηση με βάση τον αριθμό των συνδυασμών του κόμβου εναντίον Αξιολόγησης με βάση τον αριθμό των κρυφών συνδυασμών, (Βλάχου-Κογκυλάκη, 2012, p. 77)

- Αξιολόγηση με βάση τον αριθμό των κρυφών συνδυασμών εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα)



4-9 Αξιολόγηση με βάση τον αριθμό των κρυφών συνδυασμών εναντίον καμίας Αξιολόγησης (τυχαίου πράκτορα), (Βλάχου-Κογγυλάκη, 2012, p. 78)

- Συνολικές νίκες κάθε συνάρτησης αξιολόγησης



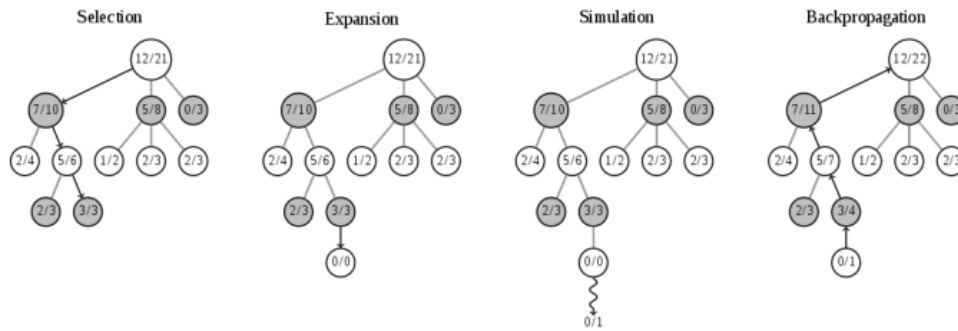
4-10 Συνολικές νίκες κάθε συνάρτησης αξιολόγησης

Ντετερμινιστικότητα στον αλγόριθμο αναζήτησης δέντρου Monte Carlo για το παιχνίδι καρτών Mini Tichu

Στο (Castelein, 2017) εξετάστηκε η χρήση διαφόρων πρακτόρων με βάση τον αλγόριθμο Monte Carlo για το παιχνίδι Mini Tichu. Τα είδη των πρακτόρων που χρησιμοποιήθηκαν ήταν ο τυχαίος πράκτορας, ο πράκτορας που χρησιμοποιεί τον αλγόριθμο Monte Carlo στη βασική του μορφή και ο πράκτορας που χρησιμοποιεί τον αλγόριθμο Monte Carlo με δένδρο αναζήτησης (Monte Carlo Tree Search – MCTS).

Ο πράκτορας ο οποίος χρησιμοποιεί τον αλγόριθμο Monte Carlo στη βασική του μορφή, χρησιμοποιεί τυχαίες προσομοιώσεις του παιχνιδιού και συγκρίνει τα αποτελέσματα των προσομοιώσεων αυτών για να βρει ένα καλό αποτέλεσμα. Αυτό γίνεται παρατηρώντας το αποτέλεσμα όλων των τυχαίων προσομοιώσεων του παιχνιδιού όταν γίνει μια κίνηση. Έπειτα επιλέγει την κίνηση η οποία έχει το μεγαλύτερο ποσοστό νικών. Εάν υπάρχουν περισσότερες από μια τέτοιες κινήσεις τότε επιλέγεται η πρώτη κίνηση που βρέθηκε. Αυτή η μέθοδος είναι πολύ καλή στο να βρίσκει καλές κινήσεις σε παιχνίδια, καθώς οι υπολογιστές μπορούν εύκολα να προσομοιώσουν εκατοντάδες παιχνίδια σε πολύ μικρό χρονικό διάστημα. Στο Tichu όπου το ποσοστό νικών δεν είναι τόσο σημαντικό, το σκορ μπορεί επίσης να χρησιμοποιηθεί ως παράμετρος. Επιλέχθηκε όμως το ποσοστό των νικών, γιατί εμπειρικά δίνει τα καλύτερα αποτελέσματα.

Ο πράκτορας ο οποίος χρησιμοποιεί τον αλγόριθμο MCTS, διαλέγει “ενδιαφέρουσες” κινήσεις προς αξιολόγηση. Αυτό επιτυγχάνεται με τη δημιουργία ενός δέντρου παιχνιδιών κατά τη διάρκεια των προσομοιώσεων και επιλέγοντας τους κόμβους στους οποίους θα συνεχιστεί η προσομοίωση. Χρησιμοποιούνται τα βήματα που φαίνονται παρακάτω.



4-11 Βήματα διαδικασίας επιλογής MCTS, (Castelein, 2017, p. 10)

Επιλογή (Selection): Κατά τη διαδικασία της επιλογής, ο αλγόριθμος ερευνά κάθε πιθανό παιδί σύμφωνα με ένα προκαθορισμένο πρωτόκολλο και επιλέγει ένα από αυτά. Κανονικά, αυτό το πρωτόκολλο εφαρμόζεται αναδρομικά, έως ότου καταλήξει σε ένα φύλλο. Επειδή όμως δεν υπάρχει τέλεια γνώση των καρτών των άλλων παικτών, πρέπει να σταματήσει η αναζήτηση ακόμα και όταν ένας κόμβος δεν έχει επεκταθεί πλήρως. Αυτό γίνεται, διότι δημιουργούνται νέες κάρτες για τους υπόλοιπους παίκτες και αυτοί μπορεί να έχουν νέες επιλογές παιχνιδιού διαθέσιμες. Από την άλλη, μπορεί κάποιες επιλογές παιχνιδιού να μην είναι διαθέσιμες και άρα εξετάζονται μόνο οι κόμβοι που είναι διαθέσιμοι αυτήν τη στιγμή με τις επιλεγμένες κάρτες για τους παίκτες.

Επέκταση (Expansion): Ο επιλεγμένος κόμβος επεκτείνεται πλήρως με όλες τις δυνατές επιλογές παιχνιδιού αυτή τη στιγμή.

Προσομοίωση (Simulation): Εάν ένας κόμβος δεν είναι τερματικός, τότε ξεκινά μια προσομοίωση ξεκινώντας από τον συγκεκριμένο κόμβο.

Διάδοση προς τα πίσω (Back propagation): Χρησιμοποιείται το αποτέλεσμα των προσομοιώσεων για να ανανεωθούν όλες οι τιμές των γονιών μέχρις ότου το δέντρο φτάσει στην τωρινή κατάσταση.

Τα παραπάνω βήματα επαναλαμβάνονται έως ότου αποφασιστεί ο τερματισμός του αλγορίθμου. Αυτό γίνεται, στην περίπτωση της διατριβής αυτής, μετά από ένα συγκεκριμένο αριθμό προσομοιώσεων. Είναι πολύ σημαντικό όσον αφορά το πρωτόκολλο επιλογής να εξισορροπηθεί η έρευνα των νέων κόμβων με το πόσο καλοί είναι οι ήδη ερευνημένοι κόμβοι.

Εδώ χρησιμοποιείται η συνθήκη η οποία προτάθηκε στο (Hyeong Soo Chang, 2015):

$$\frac{w}{n} + c * \sqrt{\frac{\ln t}{n}}, \text{ όπου } w_i \text{ είναι η συνολική ποσότητα των φορών που ο } i\text{-οστός κόμβος παιδί}$$

οδήγησε σε νίκη, n_i είναι η συνολική ποσότητα των προσομοιώσεων που έχουν γίνει για τον i -οστό κόμβο παιδί, c είναι η παράμετρος εξερεύνησης και είναι ίση με $\sqrt{2}$ για όλες τις δοκιμές και t είναι ο συνολικός αριθμός των προσομοιώσεων που έχουν γίνει για τον κόμβο πατέρα. Ο πρώτος όρος της συνθήκης εστιάζει στο ότι ένας κόμβος ο οποίος έχει κερδίσει συχνά είναι πιο πιθανό να επιλεγθεί. Ο δεύτερος όρος εστιάζει στο ότι εάν ένας κόμβος παιδί έχει παιχτεί λιγότερες φορές σε σχέση με άλλους κόμβους παιδιά, είναι πιο πιθανό να επιλεγθεί.

Ο ντετερμινισμός είναι ο τρόπος με τον οποίο καθορίζονται οι άγνωστες παράμετροι στην τωρινή κατάσταση παιχνιδιού. Για το Tichu, αυτό σημαίνει τον καθορισμό των καρτών των άλλων παικτών. Ένας πράκτορας είτε θα χρησιμοποιεί ντετερμινισμό, δηλαδή μία από τις μεθόδους που θα περιγράψουν παρακάτω, ή απλά θα καθορίζει τις άγνωστες κάρτες των υπολοίπων παικτών τυχαία (τυχαίος ντετερμινισμός). Για να βελτιωθεί ο τυχαίος ντετερμινισμός έγινε σύγκριση προηγούμενων παιχνιδιών που έχουν παιχτεί από αυτούς του πράκτορες με το τωρινό παιχνίδι. Αυτό γίνεται συγκρίνοντας τη κίνηση για την οποία θα γίνει η προσομοίωση με προηγούμενα παιχνίδια στα οποία παίχτηκε μια παρόμοια κίνηση. Αυτή η πληροφορία στη συνέχεια χρησιμοποιείται για να χωρίσει τις άγνωστες κάρτες με ημιτυχαίο τρόπο ανάμεσα στους υπόλοιπους παίκτες.

Για το λόγο αυτό αποθηκεύονται οι πληροφορίες που θα χρησιμοποιηθούν στα επόμενα παιχνίδια. Οι πληροφορίες αυτές είναι οι εξής: α) ο συνδυασμός των καρτών, β) η σειρά της κίνησης για τις πρώτες 90 σειρές (κάθε γύρος έχει περίπου 70 σειρές κίνησης και μετά από 90 σειρές οι κάρτες που έχουν μείνει είναι πολύ λίγες και άρα λίγες είναι και οι αξιόλογες πληροφορίες που παίρνουμε από αυτό το στοιχείο), γ) ποιος παίκτης έκανε αυτή την κίνηση και δ) αν είχε ή όχι ο πράκτορας τις κάρτες που ντετερμινιστικά επιλέχθηκαν ως οι άγνωστες κάρτες του.

Συνολικά έχουμε 7 συνδυασμούς x 90 σειρές κίνησης x 4 παίκτες x 2 πιθανές καταστάσεις (αν είχε ή όχι τις κάρτες αυτές ο πράκτορας) δηλαδή 5040 πόντους δεδομένων. Οι πόντοι αυτοί χρησιμοποιήθηκαν για την εξαγωγή ενός ποσοστού πιθανότητας ένας πράκτορας να έχει έναν συνδυασμό.

Για τις προσομοιώσεις οι όποιες έγιναν, κρατήθηκε σαν όριο η εξέταση 20 δυνατών κινήσεων για έναν πράκτορα ανά κατάσταση παιχνιδιού. Κάθε παίκτης θα μπορούσε να έχει διαφορετικό πράκτορα, αλλά χρησιμοποιήθηκε ο ίδιος πράκτορας και για τους δύο παίκτες της ίδιας ομάδας. Καθώς το Tichu έχει πολλούς γύρους, παρατηρήθηκαν μόνο γύροι και το εάν η ομάδα κέρδισε τον εκάστοτε συγκεκριμένο. Για όλα τα παρακάτω αποτελέσματα χρησιμοποιήθηκε ο μέσος όρος 1000 γύρων. Συγκρίθηκε το πλήθος των νικών σαν ποσοστό για κάθε μια από τις δύο ομάδες. Η φόρμουλα που χρησιμοποιήθηκε είναι η εξής: $100 \times \frac{\text{wins}}{\text{total games} - \text{ties}}$. Αποφασίστηκε να αφαιρεθούν οι ισοπαλίες γιατί είναι ένα πολύ μικρό ποσοστό των αποτελεσμάτων, γύρω στο 3%.

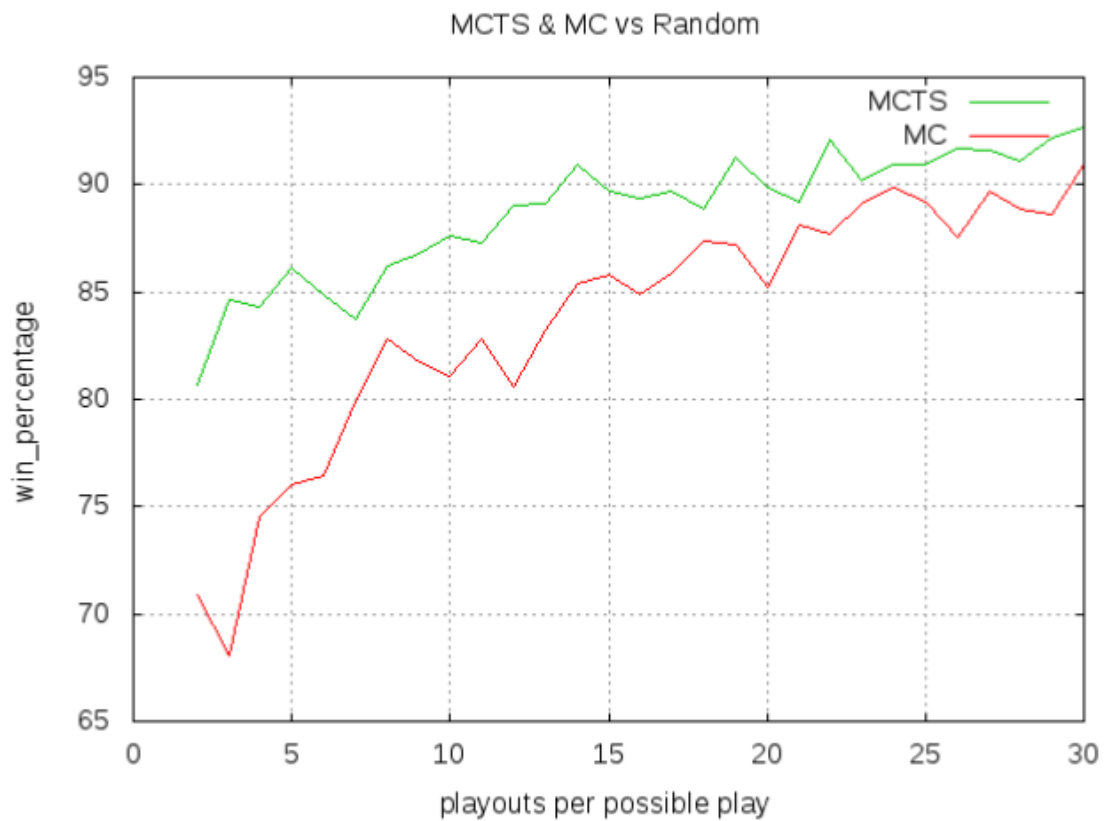
Στον παρακάτω πίνακα παρουσιάζονται τα αποτελέσματα των ομάδων που χρησιμοποιούν ένα είδος πρακτόρων όταν παίζουν ενάντια σε ομάδες που χρησιμοποιούν άλλο είδος πρακτόρων. Τα ποσοστά πάνω στη διαγώνιο του διαγράμματος πρέπει να συγκλίνουν ακριβώς στο 50% για περισσότερες προσομοιώσεις, μιας και οι δύο ομάδες χρησιμοποιούν τους ίδιους πράκτορες και οι κάρτες που μοιράζονται σε κάθε ομάδα αλλάζουν ανά κάθε γύρο και άρα καμία ομάδα δεν έχει κάποιο πλεονέκτημα.

	Random	MC	MCTS
Random	50.2	89.2	90.6
MC	10.8	50.4	50.6
MCTS	9.4	49.4	49.6

Table 6.1: Winrate without determinization (read as MC has a 89.2% chance to win against Random)

4-12 Ποσοστό νίκης χωρίς ντετερμινισμό, (Castelein, 2017, p. 14)

Για τα παραπάνω αποτελέσματα χρησιμοποιήθηκαν προσομοιώσεις 20 πιθανών επόμενων κινήσεων ανά κατάσταση παιχνιδιού. Στο παρακάτω διάγραμμα φαίνεται το πως η αλλαγή του αριθμού των πιθανών κινήσεων που εξετάζει ο αλγόριθμος MC (Monte Carlo) και MCTS (Monte Carlo Tree Search) επηρεάζει το ποσοστό νίκης ενάντια στους πράκτορες που παίζουν τυχαία. Παρατηρείται επίσης ένας κορεσμός μετά από έναν αριθμό πιθανών κινήσεων (πολύ μικρή βελτίωση).



4-13 MCTS και MC πράκτορες εναντίον τυχαίων πρακτόρων με διαφορετικό αριθμό πιθανών κινήσεων χωρίς ντετερμινισμό, (Castelein, 2017, p. 15)

Στη συνέχεια εξετάστηκε το ποσοστό των σωστά επιλεγμένων συνδυασμών για τις άγνωστες κάρτες των υπολοίπων παικτών. Πάνω από ένα δισεκατομμύριο επιλογές εξετάστηκαν και τα αποτελέσματα (των ντετερμινιστικών και των μη-ντετερμινιστικών MC και MCTS αλγορίθμων αντίστοιχα) εμφανίζονται στον παρακάτω πίνακα:

-	MC	MCTS
without determinization	87.3%	84.2%
with determinizaion	86.9%	88.2%

Table 6.2: Correctly predicted card combinations with and without determinization. Results were taken from over 1 billion predictions at various stages in the game.

4-14 Ποσοστό των σωστά επιλεγμένων συνδυασμών για τις άγνωστες κάρτες των υπολοίπων παικτών με και χωρίς ντετερμινισμό για MC και MCTS πράκτορες, (Castelein, 2017, p. 17)

Για τον αλγόριθμο Monte Carlo φάνηκε ότι η χρήση της τεχνικής ντετερμινισμού που περιγράφηκε παραπάνω είναι χειρότερη από την τυχαία μέθοδο. Για τον αλγόριθμο Monte Carlo με χρήση δέντρου αναζήτησης παρουσιάστηκε κάποια βελτίωση αν και μικρή. Για να επιβεβαιωθεί ότι η βελτίωση δεν είναι τυχαία και με την υπόθεση ότι αν δεν υπήρχε βελτίωση το ποσοστό νίκης του μη-ντετερμινιστικού αλγορίθμου MCTS εναντίον του ντετερμινιστικού αλγορίθμου MCTS θα ήταν 50%, τέθηκαν ως αντίπαλες δύο ομάδες χρησιμοποιώντας τον μη-ντετερμινιστικό αλγόριθμο MCTS και τον ντετερμινιστικό αλγόριθμο MCTS όπως προέκυψε μετά από 100000 παιχνίδια προπόνησης. Ο ντετερμινιστικός αλγόριθμος κέρδισε 24336 φορές και έχασε 24212 φορές. Η πιθανότητα να γίνει αυτό ενώ το ποσοστό νίκης είναι 50% δίνεται από τον τύπο $\sum_{i=k}^n \binom{n}{i} / 2^n$ όπου n είναι ο αριθμός των παιχνιδιών που παίχτηκαν χωρίς τις όποιες ισοπαλίες και k ο αριθμός των νικών του ντετερμινιστικού αλγορίθμου MCTS. Εάν το ποσοστό νίκης είναι 50% τότε το να νικήσει 24336 φορές και να χάσει 24212 φορές έχει πιθανότητα 28.8%. Άρα μπορούμε να υποθέσουμε ότι πιθανώς ο ντετερμινισμός με τις παραμέτρους που παρουσιάστηκαν παραπάνω έχει κάποιο θετικό αποτέλεσμα, άλλα αυτό είναι πολύ μικρό. Άρα ο ντετερμινισμός στη μορφή που περιγράφηκε παραπάνω δεν βελτιώνει αισθητά κανέναν από τους παραπάνω πράκτορες για το παιχνίδι Mini-Tichu.

Στη διατριβή αυτή έγινε και μια προσπάθεια για την υλοποίηση του αλγορίθμου Monte Carlo (αλλά όχι και του Monte Carlo με δέντρο αναζήτησης) και για το ολοκληρωμένο παιχνίδι του Tichu. Τα αποτελέσματα ήταν ενθαρρυντικά όταν μια ομάδα χρησιμοποιούσε πράκτορες με αλγόριθμο Monte Carlo εναντίον τυχαίων πρακτόρων (ποσοστό νικών 86% για την ομάδα που χρησιμοποιούσε αλγόριθμο MC). Η χρήση του μοντέλου ντετερμινισμού που περιγράφηκε παραπάνω δε πρόσφερε κάποια βελτίωση, όπως και για το Mini-Tichu.

Συνδυασμός Διαβούλευσης και Στρατηγικών Αντίδρασης για ευφυείς πράκτορες στο παιχνίδι Mini-Tichu

Οι Μάρθα Βλάχου-Κογχυλάκη και Σταύρος Βάσσος μελέτησαν στο (Martha Vlachou-Konchylaki, 2013) τους συνδυασμούς διαβούλευσης και στρατηγικών αντίδρασης για ευφυείς πράκτορες για το παιχνίδι Mini-Tichu.

Το παιχνίδι Mini-Tichu είναι ένα παιχνίδι μερικώς παρατηρήσιμο αφού κάθε παίκτης ξεκινά γνωρίζοντας μόνο τις δικές του κάρτες (το $\frac{1}{4}$ ολόκληρης της τράπουλας). Κάθε γύρος είναι απολύτως παρατηρήσιμος, αφού κάθε κίνηση αποκαλύπτει κρυμμένες κάρτες οι οποίες αφαιρούνται από το παιχνίδι. Επίσης, το παιχνίδι είναι πολύ δυναμικό μιας και υπάρχει ισχυρή αλληλεπίδραση μεταξύ των παικτών μιας και παίζουν σε ομάδα με έναν άλλον παίκτη και πρέπει συνέχεια να αξιολογούν τη θέση όλων των παικτών και να αλλάζουν τρόπο παιχνιδιού από αμυντικό σε επιθετικό και αντιστρόφως.

Λόγω του υψηλού βαθμού της άγνωστης πληροφορίας στην αρχή του παιχνιδιού και τον αριθμό των παικτών, ο χώρος των καταστάσεων που περιγράφει τις πιθανές εξελίξεις του παιχνιδιού είναι αχανής. Αυτό ουσιαστικά κάνει απαγορευτική τη χρήση τεχνικών αναζήτησης όπως δένδρα-AB ή του αλγορίθμου Minimax μέχρι τα τελευταία στάδια του παιχνιδιού, όπου πολλές κάρτες έχουν αφαιρεθεί. Από την άλλη, οι στρατηγικές αντίδρασης οι οποίες βασίζονται στην τωρινή κατάσταση του παιχνιδιού και στις τελευταίες κινήσεις που έχουν γίνει προσφέρουν μια διαφορετική προσέγγιση για τη λήψη αποφάσεων και έχουν φανεί χρήσιμες σε άλλα είδη παιχνιδιών με χαρακτήρες που δεν χειρίζονται οι παίκτες (NPC - non-player characters). Συγκεκριμένα, οι μηχανές πεπερασμένων καταστάσεων (Finite State Machines ή FSMs) δίνουν έναν απλό τρόπο προσδιορισμού μιας συμπεριφοράς αντιδράσεων. Ένα FSM ορίζεται από έναν πεπερασμένο αριθμό καταστάσεων, κάθε μια από τις οποίες εκφράζει έναν συγκεκριμένο τύπο προκαθορισμένων συμπεριφορών. Αλλαγές μεταξύ καταστάσεων συμβαίνουν μέσα από συνθήκες, όταν αυτές εκπληρούνται, που προκύπτουν από διακριτές ενέργειες ή περιστατικά κατά τη διάρκεια του παιχνιδιού.

Τα κομμάτια της προσέγγισης αυτής είναι τα εξής: α) ο προσδιορισμός όλων των δυνατών συνδυασμών καρτών για ένα δοσμένο χέρι, όπου μια δενδρική διάταξη δημιουργείται όταν

το χέρι χωρίζεται σε διάφορους συνδυασμούς καταλήγοντας στους τερματικούς κόμβους που περιέχουν όλους τους δυνατούς συνδυασμούς (για παράδειγμα, ένα χέρι 5 φύλλων 5 5 5 3 3 θα δημιουργήσει τρεις τερματικούς κόμβους, έναν με 5 5 3 3 ζευγάρια και 5 μόνο φύλλο, έναν με 3 3 ζευγάρι και 5 5 5 τριπλέτα και ένα Full House 5 5 5 3 3), β) υπολογισμός της δυναμικότητας ενός συγκεκριμένου συνδυασμού με βάση τις κάρτες που έχουν παιχτεί ήδη, γ) εξέταση του εάν ένας συνδυασμός μπορεί να “κερδίσει” (για παράδειγμα αν έχουν παιχτεί ήδη τρεις Άσσοι, τότε ένα ζευγάρι Ρηγάδων μπορεί να “κερδίσει” με απόλυτη βεβαιότητα στο παιχνίδι Mini-Tichu), δ) βαθμολόγηση των διαφόρων τερματικών κόμβων που προκύπτουν από την ανάλυση του χεριού, χρησιμοποιώντας σαν παραμέτρους τη δυναμικότητα των συνδυασμών, τον αριθμό των καρτών κάθε συνδυασμού και έναν εμπειρικό πολλαπλασιαστή για κάθε τύπο συνδυασμού και ε) χρησιμοποιώντας διαφορετικές τεχνικές παιχνιδιού ανάλογα με τις τελευταίες κινήσεις των υπολοίπων παικτών. Χρησιμοποιείται διαφορετική μηχανή πεπερασμένων καταστάσεων ανάλογα με τις διάφορες φάσεις του παιχνιδιού (για παράδειγμα, όταν απομένουν τρεις παίκτες με κάρτες στο χέρι τους). Οι καταστάσεις οι οποίες αντιστοιχούν στον πράκτορα τεχνητής νοημοσύνης καθορίζουν πληροφορίες σχετικά με τη στρατηγική του χρησιμοποιώντας όλα τα υπόλοιπα που αναφέρθηκαν παραπάνω. Οι υπόλοιπες καταστάσεις κάθε διαγράμματος χρησιμοποιούνται σαν καταχωρητές μνήμης για τις ενέργειες των υπολοίπων παικτών.

Παρακάτω φαίνεται ένα διάγραμμα του FSM που χρησιμοποιείται όταν παίζουν και οι τέσσερις παίκτες. Κάθε μετάβαση δείχνει την ενέργεια κάποιου παίκτη. Οι καταστάσεις Παίξε (Play) και Πήγαινε πάσο (Pass) δείχνουν συγκεκριμένους τρόπους παιχνιδιού που πρέπει να εκτελεστούν από τον πράκτορα τεχνητής νοημοσύνης. T, O1, O2 αντιστοιχούν στο συμπαίκτη, τον πρώτο και τον δεύτερο αντίπαλο του πράκτορα αντίστοιχα. Οι μεταβάσεις P1 και Pas αντιπροσωπεύουν κινήσεις παιχνιδιού και κινήσεις όπου ο παίκτης πήγε πάσο αντίστοιχα.

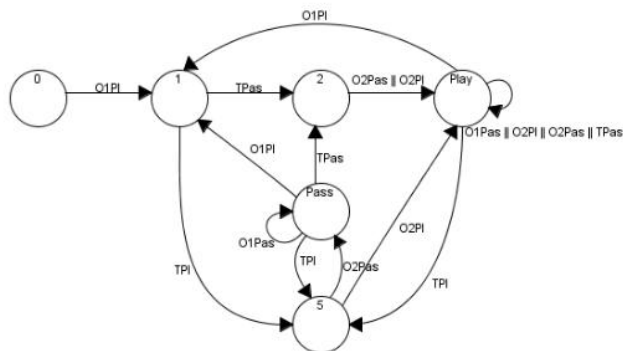


Figure 1: FSM when all four players are in the game

4-15 FSM για το παιχνίδι Mini Tichu όταν και όλοι οι παίκτες είναι ενεργοί σε έναν γύρο,
(Martha Vlachou-Konchylaki, 2013, p. 2)

Με το πέρας κάθε γύρου ένα νέο δένδρο αναζήτησης κατασκευάζεται και ο καλύτερος κόμβος επανεκλέγεται. Κατά τη διάρκεια του παιχνιδιού και ανάλογα με τη θέση που βρίσκεται ο πράκτορας θα μπορεί να παίξει αμυντικά (δηλαδή να παίξει έναν συνδυασμό μόνο αν κρατάει σταθερή τη μέση αξία του χεριού του), να πάει πάσο ή να παίξει επιθετικά (δηλαδή να παίξει με κύριο σκοπό να κερδίσει το συγκεκριμένο γύρο συνδυασμού).

Αυτή η προσέγγιση συμβάλλει στη δημιουργία ενός πολυδιάστατου και ικανού ευφυή πράκτορα που μπορεί και λαμβάνει υπόψη πολλά από τα χαρακτηριστικά του παιχνιδιού.

ii) Χρήση ευφρών τεχνικών σε διάφορα κομμάτια του παιχνιδιού Tichu

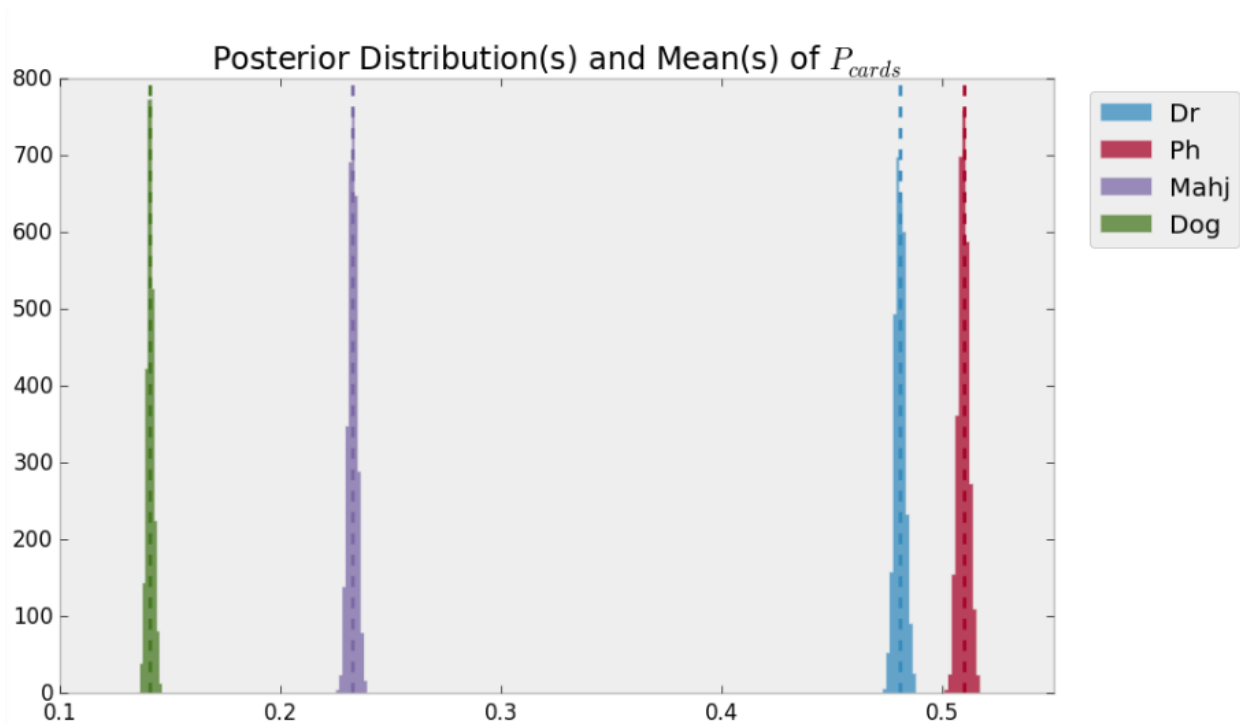
Bayes-ιανό μοντέλο Tichu για σύγκριση δυναμικότητας χεριών

Ο Paul Reid στο (Reid, 2015) χρησιμοποίησε ένα Bayes-ιανό στατιστικό μοντέλο για να συγκρίνει τη δυναμικότητα, δηλαδή την πιθανότητα του να ξεφορτωθεί ένας παίκτης όλες τις κάρτες από το χέρι του πρώτος σε έναν γύρο, διαφορετικών συνδυασμών καρτών Tichu. Η Bayes-ιανή φύση του μοντέλου επιτρέπει τη σύγκριση της πιθανότητας και της συνδεδεμένης αβεβαιότητας.

Χρησιμοποιήθηκαν 175000 χέρια Tichu όπου κάθε χέρι είναι οι 14 κάρτες που μοιράζονται σε έναν παίκτη πριν ξεκινήσει ένας γύρος. Τα χέρια αυτά είναι παρτίδες παιγμένα στο BrettSpielWelt και βρέθηκαν στο (Tichumania).

Η διαθέσιμη πληροφορία είναι οι κάρτες σε κάθε χέρι και εάν ο παίκτης που είχε αυτό το χέρι ξεφορτώθηκε όλες τις κάρτες του πρώτος ή όχι. Η Bayes-ιανή ανάλυση είναι ένα καλό εργαλείο για την εκτίμηση των πιθανοτήτων με αυτό το σύνολο δεδομένων, επειδή υπάρχουν πολλοί τύποι χεριών και τα δεδομένα είναι σχετικά αραιά και άρα η χρήση αυτού του δείγματος ως εκτιμητή θα οδηγήσει σε ακραία υψηλές και χαμηλές πιθανότητες για τους τύπους χεριών με λίγες παρατηρήσεις.

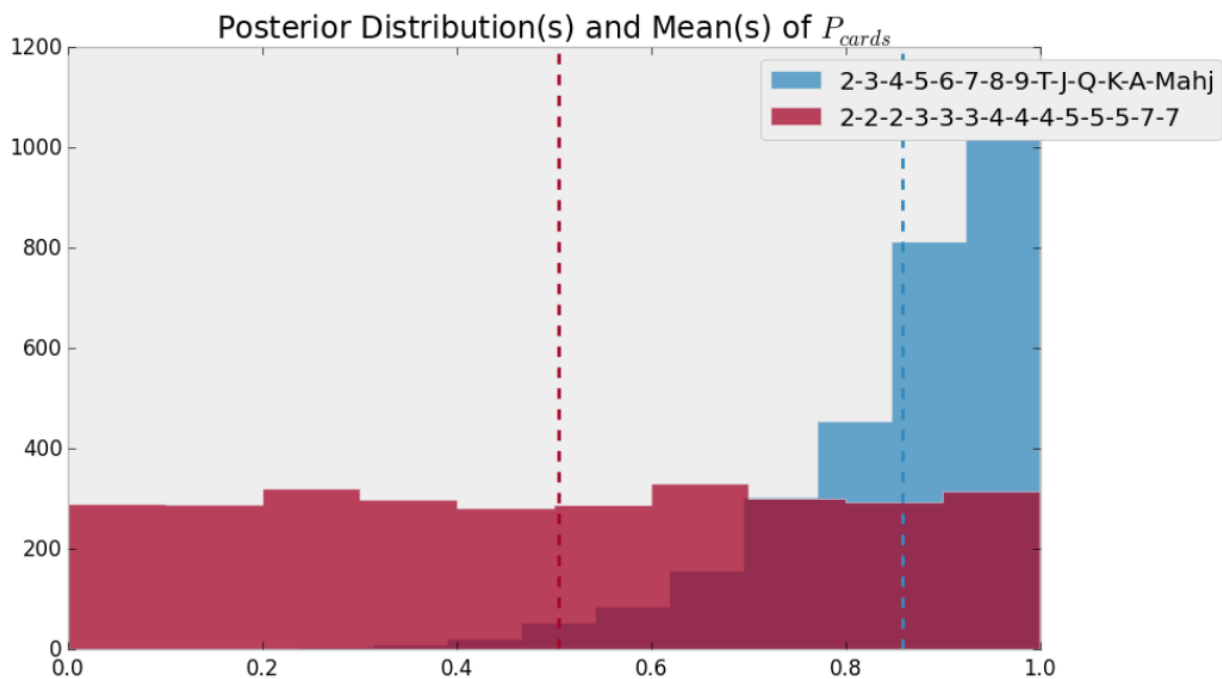
Το μοντέλο χρησιμοποιεί έναν αλγόριθμο Monte Carlo με Μαρκοβιανή Αλυσίδα για να εκτιμηθεί η σχετική δύναμη χεριών που περιέχουν κάθε ειδική κάρτα (Mahjong, Σκυλιά, Φοίνικας και Δράκος). Το να ξεφορτωθεί ο παίκτης πρώτος όλες τις κάρτες του μοντελοποιείται σαν μια τυχαία μεταβλητή Bernoulli, F , με παράμετρο p . Στη συνέχεια αρχικοποιείται η παράμετρος p για κάθε ένα από τα 4 φύλλα αυτά με τον ίδιο τρόπο (ομοιόμορφη κατανομή). Με εφαρμογή στο σύνολο των δεδομένων εξάγονται οι εξωτερικές κατανομές και το μέσο της παραμέτρου p για κάθε μία από τις κάρτες όπως φαίνεται παρακάτω:



4-16 Εξωτερικές κατανομές και το μέσο της παραμέτρου p , (Reid, 2015)

Παρατηρούμε ότι οι εξωτερικές κατανομές έχουν μικρό εύρος γύρω από το μέσο, λόγω των πολλών παρατηρήσεων που έχουμε για κάθε τύπο χεριού. Μιας και οι κατανομές δεν έχουν κοινά σημεία, είναι ξεκάθαρο ότι τα χέρια που περιέχουν την κάρτα Φοίνικας έχουν μεγαλύτερη εκτιμώμενη πιθανότητα να τα ξεφορτωθεί πρώτος ο παίκτης από ότι αυτά που περιέχουν την κάρτα Δράκος, ακόμα και αν η διαφορά είναι μικρή. Το μοντέλο μπορεί να χρησιμοποιηθεί για να συγκρίνει οποιοδήποτε αριθμό διαφορετικών χεριών και επειδή είναι Bayes-ιανό μοντέλο οι εκτιμήσεις γίνονται πιο αβέβαιες για μικρότερο αριθμό παρατηρήσεων.

Για παράδειγμα δύο πολύ σπάνια χέρια (Mahjong-2-3-4-5-6-7-8-9-10-J-Q-K-A και 2-2-2-3-3-3-4-4-4-5-5-5-7-7) εκ των οποίων το πρώτο αποτελεί μια κέντα 14 φύλλων και είναι σίγουρο ότι ο παίκτης που έχει αυτό το χέρι θα ξεφορτωθεί πρώτος τα φύλλα του, ενώ στο δεύτερο είναι αρκετά δύσκολο να συμβεί το ίδιο. Αλλά λόγω του τεράστιου πλήθους πιθανών συνδυασμών και της σπανιότητας των συγκεκριμένων χεριών δεν υπάρχουν αρκετές παρατηρήσεις ίδιες με αυτά τα δύο χέρια στο σύνολο δεδομένων που χρησιμοποιήθηκε. Έτσι, ενώ για το δυνατό χέρι (που έχει και μεγαλύτερο αριθμό δειγμάτων) η εξωτερική κατανομή είναι προς τη σωστή κατεύθυνση και το μέσο είναι κοντά, η κατανομή του κακού χεριού δεν είναι κατανομημένη σωστά όπως φαίνεται παρακάτω:



4-17 Εξωτερικές κατανομές και το μέσο της παραμέτρου p για ένα καλό και ένα κακό χέρι, (Reid, 2015)

Απόφαση για Δήλωση Tichu με χρήση Perceptron

Στο (Torres, 2016) μελετήθηκε η διαδικασία της απόφασης για δήλωση ή μη Tichu από έναν παίκτη Tichu με τη χρήση Perceptron. Για την εκπαίδευση του Perceptron χρησιμοποιήθηκαν όλες οι παρτίδες των 45 καλύτερων παικτών του BrettSpielWelt. Στη συνέχεια δοκιμάστηκαν πολλά είδη διαφορετικών χαρακτηριστικών όπως ο αριθμός των καρτών, ο αριθμός κάθε υψηλής κάρτας, ο συνολικός αριθμός των υψηλών καρτών, ο αριθμός των συνδυασμών χαμηλής, μεσαίας και υψηλής δυναμικότητας, ο αριθμός των Άσσων. Με τη χρήση των παραπάνω ο αλγόριθμος Perceptron είχε αποτελέσματα με 71.64% ακρίβεια και 55% F1-score, ενώ το Perceptron που χρησιμοποιεί το μέσο όρο των αποτελεσμάτων των χαρακτηριστικών (averaged Perceptron) επέστρεψε αποτελέσματα με 87.62% ακρίβεια και 65% F1-score.

Στη συνέχεια χρησιμοποιήθηκαν χαρακτηριστικά του τύπου υπάρχει ή δεν υπάρχει ένας συνδυασμός στο χέρι που εξετάζεται. Τα χαρακτηριστικά αυτά είναι τα εξής: υπάρχει

ζευγάρι, υπάρχει τριπλέτα, υπάρχουν συνεχόμενα ζευγάρια, υπάρχει Full House, υπάρχει Κέντα, υπάρχει Φοίνικας και Δράκος. Με τη χρήση μόνο αυτών των χαρακτηριστικών ο αλγόριθμος Perceptron επέστρεψε αποτελέσματα με 70.44% ακρίβεια και 55% F1-score, ενώ το Perceptron που χρησιμοποιεί το μέσο όρο των αποτελεσμάτων των χαρακτηριστικών (averaged Perceptron) επέστρεψε αποτελέσματα με 87.58% ακρίβεια και 65% F1-score.

Άλλα χαρακτηριστικά τα οποία προστέθηκαν στη συνέχεια είναι ο μέσος όρος δυναμικότητας του χεριού, που προκύπτει από το άθροισμα της δυναμικότητας όλων των φύλλων του χεριού διαιρεμένο από τον αριθμό των φύλλων του χεριού, και το μέσο όρο των συνδυασμών χαμηλής, μεσαίας και υψηλής δυναμικότητας δια 18. Τα δύο αυτά χαρακτηριστικά δεν προκάλεσαν κάποια βελτίωση στο αποτέλεσμα. Άλλο ένα χαρακτηριστικό που επιλέχθηκε είναι η διαφορά του αριθμού των άσπων από το άθροισμα των καρτών που είναι μικρότερες του 4 σε δυναμικότητα. Από όλα τα παραπάνω χαρακτηριστικά το χαρακτηριστικό το οποίο είχε τα καλύτερα αποτελέσματα από όλα ήταν ο αριθμός των καρτών. Αυτό το χαρακτηριστικό κρατάει σαν πληροφορία το πόσες από κάθε μία από τις κάρτες έχει το χέρι που εξετάζουμε αυτή τη στιγμή και άρα δίνει πολύ σημαντική διαχωριστική πληροφορία στο Perceptron.

Αν χρησιμοποιηθεί μόνο του έχει ως αποτέλεσμα 80.44% ακρίβεια και 62% F1-score με τον αλγόριθμο Perceptron, ενώ το Perceptron που χρησιμοποιεί το μέσο όρο των αποτελεσμάτων των χαρακτηριστικών (averaged Perceptron) επέστρεψε αποτελέσματα με 87.26% ακρίβεια και 64% F1-score.

Στη συνέχεια έγινε προσπάθεια με επιλογές διαφόρων υποσυνόλων των διαθέσιμων χαρακτηριστικών να επιτευχθεί καλύτερο αποτέλεσμα. Το καλύτερο αποτέλεσμα που προέκυψε ήταν 88.68% ακρίβεια και 68% F1-score.

Σαν μελλοντική βελτίωση προτείνεται η χρήση ενός Perceptron ψηφοφορίας, όπου στη σταθμισμένη αυτή ψηφοφορία θα συμμετέχουν με διαφορετικό βάρος όλα τα Perceptrons. Το βάρος θα προκύπτει ανάλογα με το πόσες συνεχόμενες φορές ένα Perceptron έχει δώσει το σωστό αποτέλεσμα στις προηγούμενες ψηφοφορίες.

Στρατηγική για Δηλώσεις Tichu/Grand Tichu στο παιχνίδι καρτών Tichu

Στο (Yang, 2018) γίνεται προσπάθεια να προβλεφτεί το αποτέλεσμα του παιχνιδιού με βάση τις αρχικές κάρτες που μοιράζονται. Συγκεκριμένα, ο Yang προσπαθεί να βρει την πιθανότητα να δηλωθεί επιτυχώς Tichu ή Grand Tichu με χρήση ενός εκπαιδευμένου μοντέλου για την πρόβλεψη της πιθανότητας του να πετύχει κάποιος παίκτης που έχει δηλώσει Tichu ή Grand Tichu να παίξει πρώτος όλες τις κάρτες του. Με πρόσβαση σε 84127 γύρους πραγματικών παιχνιδιών από τον online Tichu server (Online Tichu) στους οποίους δηλώθηκαν 20484 Grand Tichus και 38057 Tichus, γίνεται προσπάθεια να γίνει ομαδοποίηση κάθε χεριού με χρήση αλγορίθμων ομαδοποίησης όπως του απλού αλγορίθμου Bayes (Naive Bayes), του αλγορίθμου Logistic Regression, του αλγορίθμου τυχαίου δάσους και της τεχνικής της ενίσχυσης.

Ο απλός αλγόριθμος Bayes υποθέτει ότι όλα τα χαρακτηριστικά είναι ανεξάρτητα του αποτελέσματος (εάν η δήλωση ήταν επιτυχής ή όχι).

Ο αλγόριθμος Logistic Regression χρησιμοποιεί μια σιγμοειδή συνάρτηση για να καθορίσει τη πιθανότητα κάθε αποτελέσματος δοθέντος ενός σταθμισμένου βάρους όλων των χαρακτηριστικών.

Ο αλγόριθμος τυχαίων δασών (random forests algorithm) συνδυάζει πολλαπλά δέντρα αποφάσεων για να καταλήξει σε έναν εκτιμητή. Σε κάθε δέντρο απόφασης, χρησιμοποιούνται διαφορετικά δείγματα εκκίνησης και σε κάθε κόμβο απόφασης χρησιμοποιούνται διαφορετικά υποσύνολα χαρακτηριστικών. Αυτή η διαδικασία τείνει να μειώσει τις πιθανότητες του overfitting.

Η τεχνική της ενίσχυσης παίρνει μια σειρά από αδύναμους ομαδοποιητές και δίνει συγκεκριμένο βάρος σε καθέναν από αυτούς και τα αποτελέσματά τους, με σκοπό να φτιαχτεί ένας πιο δυνατό ομαδοποιητής.

Στα δείγματα που χρησιμοποιούνται για την εκπαίδευση των ομαδοποιητών χρησιμοποιείται η παραδοχή ότι εάν ένας παίκτης της ομάδας A δηλώσει Tichu αλλά ένας παίκτης της ομάδας B καταφέρει και ξεφορτωθεί όλες τις κάρτες του πρώτος, τότε το δείγμα καρτών του

παίκτη της ομάδας B θεωρείται σαν ένα δείγμα που δήλωσε Tichu και η δήλωση του ήταν επιτυχής.

Επίσης για τις δηλώσεις Grand Tichu χρησιμοποιήθηκαν σαν δείγματα μόνο τα χέρια στα όποια δηλώθηκε Grand Tichu.

Τα χαρακτηριστικά τα οποία χρησιμοποιήθηκαν ήταν ο αριθμός των Άσσων, εάν υπάρχει ο Φοίνικας στο χέρι που εξετάζεται, εάν υπάρχει ο Δράκος στο χέρι που εξετάζεται, εάν υπάρχουν τα σκυλιά στο χέρι που εξετάζεται, ο αριθμός των βομβών, ο αριθμός των τριπλετών, η υψηλότερη τριπλέτα στο χέρι που εξετάζεται, ο αριθμός των ζευγαριών, το υψηλότερο ζευγάρι στο χέρι που εξετάζεται, το μήκος της πιο μακριάς κέντας στο χέρι που εξετάζεται, η πιο υψηλή κέντα στο χέρι που εξετάζεται, ο αριθμός των συνεχόμενων ζευγαριών και ο αριθμός των μονόφυλλων δυναμικότητας μικρότερης από τη ντάμα.

Για το Grand Tichu, ο Yang κατέληξε ότι ο θόρυβος είναι αρκετά μεγάλος και το να ψάχνει ο αλγόριθμος για όλα αυτά τα χαρακτηριστικά δε βοηθάει αισθητά την ακρίβεια. Ο τρόπος με τον οποίο αξιολογήθηκαν τα διάφορα μοντέλα είναι το εμβαδόν της περιοχής (AUC) κάτω από τη χαρακτηριστική καμπύλη λειτουργίας του δέκτη (receiver operating characteristic curve - ROC). Η ROC καμπύλη προκύπτει από τη γραφική παράσταση του ποσοστού των αληθινά θετικών προβλέψεων προς το ποσοστό των ψευδώς θετικών προβλέψεων. Η καμπύλη αυτή δείχνει πως το μοντέλο λειτουργεί κάτω από διαφορετικά κατώφλια αποφάσεων. Εάν το μοντέλο αποφάσιζε τυχαία, το εμβαδόν της περιοχής αυτής θα ήταν 0.5, ενώ εάν το μοντέλο ήταν τέλειο θα ήταν 1. Η μοντελοποίηση του διάνυσματος του χεριού (raw hand) προβάλλει το χέρι που εξετάζεται σαν ένα διάνυσμα μήκους 56 από 0 ή 1, που υποδεικνύει αν το χέρι περιέχει μια κάρτα ή όχι. Το συμπιεσμένο διάνυσμα του χεριού (compressed hand) μοντέλο μετρά τον αριθμό των καρτών ανά δυναμικότητα φύλλου. Το minimal pattern μοντέλο περιέχει μόνο τέσσερα χαρακτηριστικά (αριθμό Άσσων, ύπαρξη Δράκου στο χέρι, ύπαρξη Φοίνικα στο χέρι και αριθμό βομβών) τα οποία είχαν το μεγαλύτερο συντελεστή από τα 13 χαρακτηριστικά που περιγράφηκαν παραπάνω. Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα του εμβαδού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα που περιγράφηκαν για την περίπτωση του Grand Tichu, καθώς και οι συντελεστές των τεσσάρων πιο σημαντικών χαρακτηριστικών (αριθμός άσσων, αν υπάρχει ο Δράκος στο χέρι, αν υπάρχει ο Φοίνικας στο χέρι, ο αριθμός των βομβών στο χέρι).

	LR	NB	AdaBoost	RF
Raw hand	0.619	0.605	0.591	0.612
Compressed hand	0.624	0.619	0.609	0.627
Minimal Pattern (4 features)	0.626	0.552	0.625	0.627
Full Pattern (13 features)	0.628	0.582	0.622	0.620
Full Pattern + Compressed hand	0.624	0.595	0.618	0.622

TABLE 1
AUC FOR DIFFERENT MODEL AND FEATURES COMBINATION. LR: LOGISTIC REGRESSION. NB: NAIVE BAYES. RF: RANDOM FOREST. IT SEEMS THAT THE MINIMAL FEATURE HAS CONTAINS MOST OF THE PREDICTIVE POWER, AND ALL MODELS HAVE SIMILAR PERFORMANCE.

4-18 Αποτελέσματα του εμβადού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα για το Grand Tichu, (Yang, 2018)

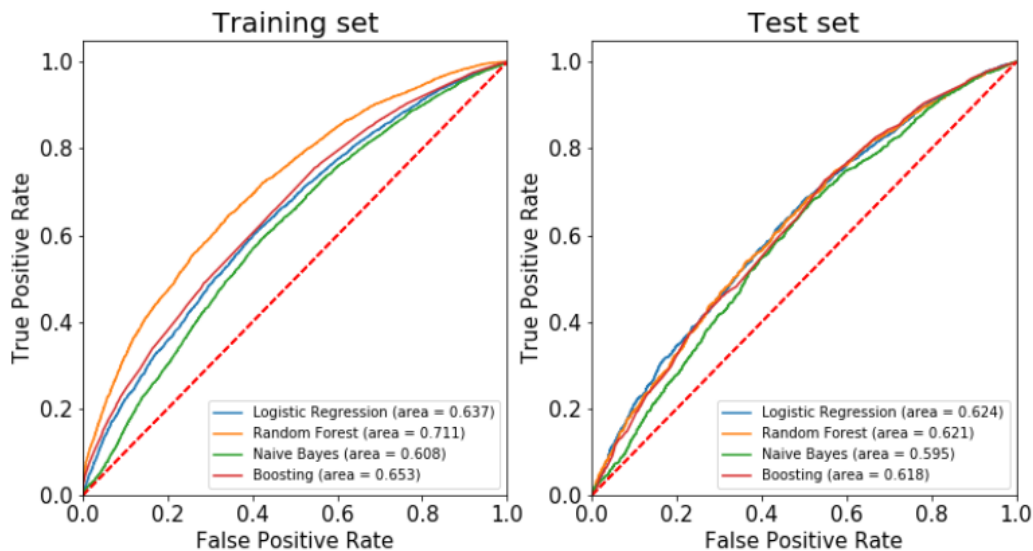


FIG. 2.— With input as 14 identified patterns plus card hand. Naive Bayes performs worse than the other models.

4-19 Χαρακτηριστικές καμπύλες όλων των μοντέλων με είσοδο 14 χαρακτηριστικά συν τις κάρτες του χεριού, (Yang, 2018)

	N_A	N_{dragon}	$N_{phoenix}$	N_{bomb}
Coeff	1.0	2.8	2.8	3.4
Approx	1	3	3	3

TABLE 2
COEFFICIENTS FOR THE 4 MOST IMPORTANT FEATURES FOR GRAND TICHU PREDICTION.

4-20 Συντελεστές για τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη Grand Tichu, (Yang, 2018)

Με στρογγυλοποίηση των συντελεστών των τεσσάρων πιο σημαντικών χαρακτηριστικών φαίνεται ότι ο Δράκος και ο Φοίνικας είναι περίπου τρεις φορές πιο σημαντικοί από έναν Άσσο στη δήλωση Grand Tichu. Τα παραπάνω αποτελέσματα οδηγούν στο ότι αν στα πρώτα 8 φύλλα ένας παίκτης έχει έναν άσσο, υπάρχει πάνω από 50% πιθανότητα να ξεφορτωθεί όλα τα φύλλα του πρώτος στο γύρο αυτό, αν δηλώσει Grand Tichu, δηλαδή πάνω από 50% πιθανότητα η δήλωση Grand Tichu να είναι επιτυχής.

Με την ίδια λογική δημιουργήθηκαν και τα αντίστοιχα μοντέλα για τη δήλωση Tichu. Μιας και πριν τη δήλωση Tichu ο παίκτης έχει στη διάθεση του πληροφορία για όλες τις κάρτες του και όχι μόνο για τις 8 που έχει στο Grand Tichu, τα ελάχιστα χαρακτηριστικά που χρησιμοποιήθηκαν είναι πλέον επτά, τα τέσσερα του Grand Tichu που περιγράφηκαν παραπάνω, το εάν ο παίκτης έχει τα σκυλιά στις 14 κάρτες του, ο αριθμός των κεντών που υπάρχει στο χέρι του καθώς και ο αριθμός των μονών φύλλων δυναμικότητας χαμηλότερης της ντάμας. Τα χαρακτηριστικά αυτά αποκτούν παραπάνω αξία στο Tichu μιας και η τυχαιότητα είναι λιγότερη, εξαιτίας των παραπάνω γνωστών καρτών. Στο Tichu παρατηρήθηκε πως η ακρίβεια των αποτελεσμάτων αυξήθηκε όταν χρησιμοποιήθηκαν τα επιπλέον χαρακτηριστικά. Τα 14 χαρακτηριστικά που περιγράφηκαν παραπάνω έχουν εξίσου καλό αποτέλεσμα με τα 14 χαρακτηριστικά συν το διάνυσμα των καρτών του χεριού σαν είσοδο στο μοντέλο. Τα 6 ελάχιστα χαρακτηριστικά τα οποία περιγράφονται παραπάνω έχουν ελάχιστα χειρότερα αποτελέσματα, αλλά καλύτερα από το διάνυσμα των καρτών του χεριού σαν είσοδο. Άλλη μια σημαντική παρατήρηση είναι πως το συμπιεσμένο διάνυσμα του χεριού (compressed hand) έχει καλύτερο αποτέλεσμα σε σύγκριση με το διάνυσμα του χεριού (raw hand). Αυτό πιθανώς προκαλείται από το ότι το διάνυσμα του χεριού αναπαριστά κάθε κάρτα με τιμές 0 και 1 και δεν διαφοροποιεί τη τιμή της κάρτας από το χρώμα της. Παρακάτω φαίνονται τα αποτελέσματα του εμβαδού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα που περιγράφηκαν για

την περίπτωση του Tichu, καθώς και οι συντελεστές των επτά πιο σημαντικών χαρακτηριστικών (αριθμός άσπων, αν υπάρχει ο Δράκος στο χέρι, αν υπάρχει ο Φοίνικας στο χέρι, ο αριθμός των βομβών στο χέρι, αν υπάρχει η κάρτα σκυλιά στο χέρι, ο αριθμός των κεντών στο χέρι καθώς και ο αριθμός των μονών φύλλων δυναμικότητας χαμηλότερης από αυτή της ντάμας).

	LR	NB	AdaBoost	RF
Raw hand	0.785	0.783	0.786	0.783
Compressed hand	0.791	0.789	0.804	0.802
Minimal Pattern (6 features)	0.827	0.827	0.82	0.827
Full Pattern (14 features)	0.834	0.74	0.832	0.837
Full Pattern + Compressed hand	0.837	0.80	0.833	0.835

TABLE 4
AUC FOR TICHU PREDICTION WITH DIFFERENT MODEL/FEATURE COMBINATIONS.

4-21 Αποτελέσματα του εμβαδού της περιοχής κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη για όλα τα μοντέλα για το Tichu, (Yang, 2018)

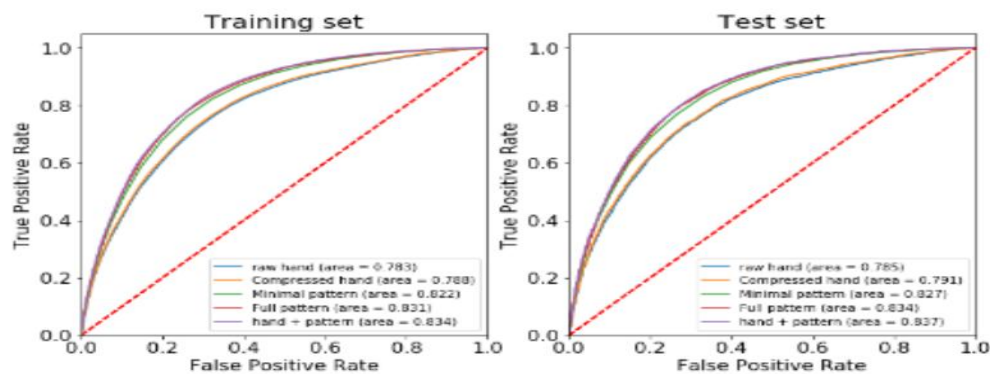


FIG. 3.— With input as 14 identified patterns plus card hand. Naive Bayes performs worse than the other models.

4-22 Χαρακτηριστικές καμπύλες όλων των μοντέλων με είσοδο 14 χαρακτηριστικά συν τις κάρτες του χεριού, (Yang, 2018)

	N_A	N_{dog}	N_{dragon}	$N_{phoenix}$	N_{bomb}	$N_{straight}$	N_{small}
Coeff	2.0	-1.9	5.9	5.7	4.7	1.0	-1.1
Approx	2	-2	6	6	5	1	-1

TABLE 3
COEFFICIENTS FOR THE 6 MOST IMPORTANT FEATURES FOR TICHU PREDICTION.

4-23 Συντελεστές για τα πιο σημαντικά χαρακτηριστικά για την πρόβλεψη Tichu, (Yang, 2018)

Με βάση τα αποτελέσματα προτείνονται επίσης μνημονικοί κανόνες σε κάθε παίκτη για τη δήλωση ή μη Grand Tichu και Tichu.

Τέλος, γίνεται μια προσπάθεια αποφυγής δήλωσης Grand Tichu όταν οι πιθανότητες να πει κάποιος Grand Tichu είναι να μην είναι πάνω από το 50%, αλλά υπάρχει αυξημένη πιθανότητα οι επόμενες 6 κάρτες να κάνουν πολύ αδύναμο το χέρι του παίκτη και άρα να μειωθούν εξαιρετικά οι πιθανότητες του να μπορεί ο παίκτης αυτός να ξεφορτωθεί όλα τα φύλλα του πρώτος. Έγιναν 10000 προσομοιώσεις χεριών Grand Tichu σε κάθε μία από τις οποίες έγιναν 100 προσομοιώσεις με 6 διαφορετικές κάρτες κάθε φορά για δήλωση Tichu. Λήφθηκε υπόψη η διαφορά των 100 πόντων που προκύπτει εάν μια δήλωση Grand Tichu θα ήταν επιτυχής, αλλά ο παίκτης δηλώσει απλό Tichu αντί για Grand Tichu και άρα προσφέρει 100 αντί για 200 πόντους στην ομάδα του. Ο Yang κατέληξε στο ότι ένα Grand Tichu πρέπει να δηλώνεται όταν η πιθανότητα η δήλωση του Grand Tichu να είναι επιτυχής υπερβαίνει το $0.5 \times P_{Tichu} + 0.25$, δηλαδή όταν η μπλε κουκίδα του παρακάτω σχήματος βρίσκεται πάνω από την κίτρινη κουκίδα του παρακάτω σχήματος.

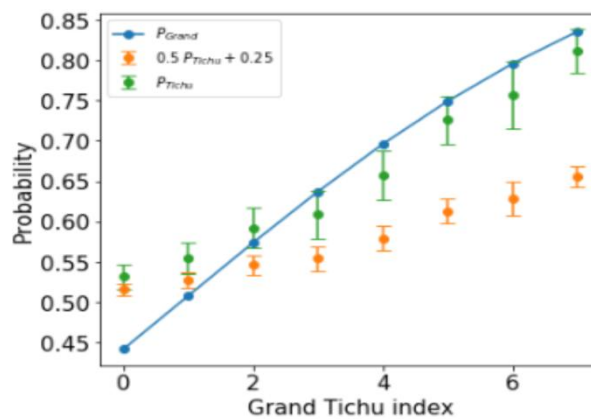


FIG. 5.— P_{grand} , $\langle P_{Tichu} \rangle$, and $0.5\langle P_{Tichu} \rangle + 0.25$ at different Grand Tichu index I_g

4-24 Πιθανότητα επιτυχούς δήλωσης Grand Tichu, (Yang, 2018)

iii) Εφαρμογές παιχνιδιού Tichu



4-25 Εφαρμογή Tichu για iOS, (Board Game Geek, 2011)

Υπάρχουν δύο εφαρμογές Tichu, όπου οι παίκτες μπορούν να παίξουν εναντίον ευφυών πρακτόρων που κάνουν κινήσεις χρησιμοποιώντας κάποιες ευφυείς τεχνικές. Δυστυχώς ο κώδικας που χρησιμοποιείται σε αυτές τις εφαρμογές δεν είναι διαθέσιμος στο ευρύ κοινό (open-source) και άρα δε μπορούμε να εξάγουμε συμπεράσματα για το ποιες μεθόδους ακολουθούν στη λήψη των αποφάσεων των παικτών. Οι εφαρμογές αυτές είναι οι (Holdren) για Android και (House Full of Games) για iOS.

Άλλη μια προσπάθεια για τη δημιουργία μιας εφαρμογής Tichu, όπου οι παίκτες μπορούν να παίξουν εναντίον ευφυών πρακτόρων που κάνουν κινήσεις χρησιμοποιώντας κάποιες ευφυείς τεχνικές έχει καταγραφεί και στο (René Schwarz, 2013). Εκεί οι René, Pfannkuch, et al., περιγράφουν την προσπάθεια δημιουργίας μιας εφαρμογής Tichu σε Android. Δυστυχώς δε γνωρίζουμε πολλά σχετικά με τις ευφυείς τεχνικές που χρησιμοποιούνται, παρά μόνο την μεγάλη υπολογιστική ισχύ που χρειάζεται για να υποστηριχτούν, αφού προκαλούνται προβλήματα απόδοσης σε πιο αδύναμες συσκευές.

Κεφάλαιο 5

Η προσέγγιση μας / Συλλογή δεδομένων

Στην παρούσα έρευνα μελετάται κυρίως το κομμάτι της δήλωσης Grand Tichu. Το Grand Tichu είναι το πιο σημαντικό στοιχείο του παιχνιδιού Tichu, διότι μια επιτυχής δήλωση Grand Tichu δίνει 200 πόντους στην ομάδα του παίκτη και αφού οι υπόλοιπες κάρτες δίνουν 100 πόντους ανά γύρο αρκούν συνήθως τέσσερις επιτυχείς δηλώσεις Grand Tichu για να αποκτήσει μια ομάδα 1000 πόντους και να ανακηρυχθεί νικήτρια. Σκοπός της εν λόγω έρευνας είναι με τη χρήση ευφών τεχνικών να δημιουργήσουμε δύο μοντέλα τα οποία θα αποφασίζουν εάν ο πράκτορας θα πρέπει να δηλώνει Grand Tichu ή όχι με τα 8 φύλλα που του έχουν μοιραστεί. Τα μοντέλα μας δεν είναι απλοί ομαδοποιητές, δηλαδή δεν απαντάνε με ένα ναι ή ένα όχι, αλλά επιστρέφουν τιμές μεταξύ του μηδέν και του ένα. Αυτό μας επιτρέπει να ορίσουμε και διάφορα είδη πρακτόρων (περισσότερο ή λιγότερο επιθετικούς πράκτορες). Ένας πράκτορας που είναι πιο επιθετικός δηλώνει Grand Tichu με χαμηλότερες τιμές από ότι ένας πιο συντηρητικός πράκτορας.

Για την εκπαίδευση των μοντέλων χρησιμοποιήθηκαν σαν δεδομένα παρτίδες Tichu οι οποίες έχουν παιχτεί online στη γερμανική πλατφόρμα παιχνιδιών (BrettSpielWelt). Η πλατφόρμα αυτή φιλοξενεί καθημερινά πολλά παιχνίδια Tichu τα αποτελέσματα των οποίων είναι διαθέσιμα στην ιστοσελίδα (Tichulog BrettSpielWelt). Τα αποτελέσματα αυτά εμφανίζονται σε μια πιο ωραιοποιημένη μορφή στην ιστοσελίδα (Tichumania).

Η επιλογή των παιχνιδιών δεν ήταν τυχαία. Επιλέχθηκαν παρτίδες στις οποίες συμμετέχουν παίκτες οι οποίοι είναι αναγνωρισμένοι στην κοινότητα του Tichu και έχουν διακρίσεις σε διάφορα τουρνουά που λαμβάνουν χώρα ανά καιρούς στην Ελλάδα και το σύστημα βαθμολόγησης Elo που χρησιμοποιεί το (Tichumania) για τον υπολογισμό του σχετικού επιπέδου δεξιότητας των παικτών σε παιχνίδια μηδενικού αθροίσματος όπως το Tichu.

i) Διαβάθμιση δυναμικότητας παικτών και ομάδων στο Tichu με τη χρήση του συστήματος αξιολόγησης Elo

Το παιχνίδι Tichu είναι ένα παιχνίδι με μηδενικό άθροισμα (zero-sum game). Αυτό σημαίνει ότι το κέρδος μιας ομάδας που αποτελείται από δύο παίκτες για μία νίκη αντισταθμίζεται ακριβώς από τις απώλειες της αντίπαλης ομάδας που χάνει το παιχνίδι. Το Tichumania χρησιμοποιεί το σύστημα αξιολόγησης Elo για να υπολογίσει το σχετικό επίπεδο δεξιοτήτων της κάθε ομάδας σε ένα παιχνίδι Tichu. Ονομάστηκε έτσι από το δημιουργό του Arpad Elo, έναν Ουγγρο-Αμερικανό καθηγητή φυσικής. Το σύστημα Elo πρωτοεμφανίστηκε ως ένα βελτιωμένο σύστημα βαθμολόγησης σκακιού σε σχέση με το σύστημα που χρησιμοποιούνταν προηγουμένως, το Harkness. Το σύστημα Elo χρησιμοποιείται όμως και ως σύστημα διαβάθμισης του ανταγωνισμού σε πολλά βιντεοπαιχνίδια, για το Αμερικάνικο ποδόσφαιρο, για το μπάσκετ, για το μπέιζμπολ, το πινγκ πονγκ, το Scrabble, για επιτραπέζια παιχνίδια όπως το “Διπλωματία” και για πολλά άλλα παιχνίδια.

Η διαφορά στις βαθμολογίες μεταξύ δύο παικτών ή στην περίπτωση του Tichu μεταξύ δύο ομάδων χρησιμεύει στο να προβλέπεται η έκβαση ενός παιχνιδιού. Δύο αντίπαλες ομάδες με ισοτιμία στο σύστημα αξιολόγησης Elo αναμένεται να κερδίσουν ίσο αριθμό παιχνιδιών. Μια ομάδα που έχει βαθμολογία 100 πόντους μεγαλύτερη από την αντίπαλη αναμένεται να κερδίσει το 64% των μεταξύ τους παιχνιδιών, ενώ εάν αυτή η διαφορά είναι 200 πόντοι, τότε αναμένεται να κερδίσει το 76%.

Η βαθμολογία Elo ενός παίκτη αντιπροσωπεύεται από έναν αριθμό ο οποίος αυξάνεται ή μειώνεται ανάλογα με το αποτέλεσμα των παιχνιδιών μεταξύ των παικτών που συμμετέχουν στην κατάταξη Elo. Στην περίπτωση του Tichu, η βαθμολογία Elo ενός παίκτη αυξάνεται ή μειώνεται ανάλογα με το αποτέλεσμα των παιχνιδιών μεταξύ των ομάδων. Εάν ο παίκτης βρίσκεται στην ομάδα που θα κερδίσει το παιχνίδι τότε η προσωπική του βαθμολογία Elo θα αυξηθεί, αλλιώς θα μειωθεί. Η διαφορά ανάμεσα στις βαθμολογίες Elo της νικητήριας και της ηττημένης ομάδας πριν το παιχνίδι, καθορίζει το πόσους πόντους θα χάσει ή θα κερδίσει μια ομάδα μετά από το παιχνίδι. Σε μια σειρά παιχνιδιών ανάμεσα σε μία ομάδα με υψηλή και σε μία ομάδα με χαμηλή βαθμολογία η ομάδα με την υψηλή βαθμολογία προβλέπεται να νικήσει παραπάνω φορές. Εάν η ομάδα με την υψηλή βαθμολογία κερδίσει, τότε μόνο λίγοι πόντοι θα παρθούν από την ομάδα με τη χαμηλή βαθμολογία. Αντίθετα, αν η ομάδα με τους

λίγους πόντους κερδίσει (παρά τα προγνωστικά) τότε θα κερδίσει πολλούς πόντους, τους οποίους θα χάσει η ομάδα με την υψηλή βαθμολογία. Η ομάδα με την χαμηλή βαθμολογία θα κερδίσει και λίγους πόντους στην περίπτωση ισοπαλίας (εάν στην εκδοχή αυτή του Tichu, επιτρέπεται η ισοπαλία, όπως γίνεται στο (BrettSpielWelt)). Αυτό σημαίνει ότι το σύστημα Elo αυτοδιορθώνεται και ότι οι παίκτες που έχουν χαμηλότερη βαθμολογία από το ταλέντο τους στο παιχνίδι, σε βάθος χρόνου, θα έχουν καλύτερα αποτελέσματα από αυτά που προβλέπει το σύστημα και έτσι θα συλλέξουν πόντους στη βαθμολογία μέχρι η βαθμολογία αυτή να αντικατοπτρίζει την πραγματική τους αξία. (Wikipedia, 2019)

Στο (BrettSpielWelt) η βαθμολογία Elo μιας ομάδας προκύπτει από το μέσο όρο της βαθμολογίας Elo των δύο παικτών που απαρτίζουν την ομάδα. Μετά το τέλος του παιχνιδιού η ομάδα που κερδίζει παίρνει πόντους ανάλογα με τη διαφορά στη βαθμολογία Elo μεταξύ αυτής και της αντίπαλης ομάδας πριν το μεταξύ τους παιχνίδι. Οι πόντοι αυτοί προστίθενται στο ατομικό Elo του κάθε παίκτη της νικητήριας ομάδας και αφαιρούνται από το ατομικό Elo κάθε παίκτη της ηττημένης ομάδας.

Κάθε παίκτης στο (BrettSpielWelt) ξεκινά με μια βαθμολογία 1500 Elo.

Η απόδοση ενός παίκτη δε μετριέται απόλυτα, αλλά συνάγεται από τις νίκες, τις ήττες και τις ισοπαλίες με τις αντίπαλες ομάδες (και κατ'επέκταση με τους αντίπαλους παίκτες). Οι βαθμολογίες των παικτών εξαρτώνται από τις βαθμολογίες των αντιπάλων τους και τα αποτελέσματα που σημειώνονται εναντίον τους. Η διαφορά στη βαθμολογία μεταξύ δύο ομάδων ορίζει μια εκτίμηση για το αναμενόμενο αποτέλεσμα στο μεταξύ τους παιχνίδι. Τόσο ο μέσος όρος όσο και η διασπορά των βαθμολογιών μπορούν να επιλεγούν αυθαίρετα. Ο Elo πρότεινε μια διαφορά 200 βαθμών στη βαθμολογία Elo στο σκάκι να ισοδυναμεί σε πιθανότητα 75% ο ισχυρότερος παίκτης (δηλαδή αυτός που έχει 200 βαθμούς υψηλότερη βαθμολογία Elo από τον αντίπαλο του) να κερδίσει το παιχνίδι. Στόχος της ομοσπονδίας σκακιού ήταν ο μέσος όρος της βαθμολογίας να είναι 1500, δηλαδή ένας μέσος παίκτης της λέσχης να έχει βαθμολογία Elo 1500.

Το αναμενόμενο αποτέλεσμα για έναν παίκτη είναι η πιθανότητα νίκης συν την μισή πιθανότητα της ισοπαλίας. Έτσι, ένας παίκτης με αναμενόμενο αποτέλεσμα 0.75 σημαίνει είτε 75% πιθανότητα νίκης, 0% πιθανότητα ισοπαλίας και 25% πιθανότητα ήττας ή 50% πιθανότητα νίκης, 0% πιθανότητα ήττας και 50% πιθανότητα ισοπαλίας. Η πιθανότητα ισοπαλίας, σε αντίθεση με το να προκαλέσει κάποιο συγκεκριμένο αποτέλεσμα, δεν

προσδιορίζεται στο σύστημα Elo. Αντιθέτως, μια ισοπαλία θεωρείται σαν μισή νίκη και μισή ήττα. Στην πράξη και δεδομένου ότι οι πραγματικές δυνατότητες κάθε παίκτη είναι άγνωστες, οι πιθανότητες νίκης υπολογίζονται χρησιμοποιώντας τις τρέχουσες βαθμολογίες Elo των παικτών ως εξής:

Εάν η ομάδα A έχει βαθμολογία Elo R_A και η ομάδα B έχει βαθμολογία Elo R_B , το εκτιμώμενο αποτέλεσμα για την ομάδα A είναι:

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}$$

Με παρόμοιο τρόπο το εκτιμώμενο αποτέλεσμα για την ομάδα B είναι:

$$E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}$$

Τα παραπάνω θα μπορούσαν επίσης να εκφραστούν σαν:

$$E_A = \frac{Q_A}{Q_A + Q_B} \text{ και } E_B = \frac{Q_B}{Q_A + Q_B}, \text{ όπου } Q_A = 10^{R_A/400} \text{ και } Q_B = 10^{R_B/400}.$$

Από τις παραπάνω εκφράσεις και παρατηρώντας πως και οι δύο εκφράσεις έχουν κοινό παρονομαστή είναι προφανές πως $E_A + E_B = 1$. Αυτό σημαίνει ότι μελετώντας μόνο τους αριθμητές, παρατηρούμε ότι το εκτιμώμενο αποτέλεσμα για την ομάδα A είναι $\frac{Q_A}{Q_B}$ φορές μεγαλύτερο από το εκτιμώμενο αποτέλεσμα για την ομάδα B. Μετά βλέπουμε ότι για κάθε 400 πόντους διαφορά στη βαθμολογία των δύο ομάδων, η ομάδα με το πλεονέκτημα (με τους περισσότερους δηλαδή πόντους) βλέπει το εκτιμώμενο αποτέλεσμα της να δεκαπλασιάζεται σε σχέση με το εκτιμώμενο αποτέλεσμα της αντίπαλης ομάδας. (Mark E. Glickman, 1996, pp. 66-70)

Όταν τα πραγματικά αποτελέσματα μιας ομάδας είναι καλύτερα από τα εκτιμώμενα αποτελέσματα αυτής, τότε το σύστημα βαθμολόγησης Elo εκλαμβάνει αυτά τα αποτελέσματα ως πειστήρια ότι η βαθμολογία Elo της ομάδας είναι πολύ χαμηλή και πρέπει να προσαρμοστεί προς τα πάνω. Παρομοίως, όταν τα πραγματικά αποτελέσματα μιας ομάδας είναι χειρότερα από τα εκτιμώμενα αποτελέσματα αυτής, τότε το σύστημα βαθμολόγησης Elo εκλαμβάνει αυτά τα αποτελέσματα ως πειστήρια ότι η βαθμολογία Elo της ομάδας είναι πολύ υψηλή και πρέπει να προσαρμοστεί προς τα κάτω. Η αρχική πρόταση του Elo, η οποία εξακολουθεί να χρησιμοποιείται ευρέως, ήταν μια απλή γραμμική προσαρμογή ανάλογα με το πόσο μια ομάδα υπερέβαινε ή ήταν κατώτερη του εκτιμώμενου αποτελέσματός της. Η μέγιστη δυνατή αναδιαμόρφωση του εκτιμώμενου αποτελέσματος μιας ομάδας ανά παιχνίδι ονομάζεται συντελεστής K και στο (Tichumania) ορίζεται στο $K = 15$. (Nate Silver, 2015)

Υποθέτοντας πως η ομάδα A είχε εκτιμώμενο αποτέλεσμα E_A αλλά πραγματικό αποτέλεσμα S_A η σχέση με την οποία ανανεώνεται η βαθμολογία της ομάδας είναι η εξής:

$$R'_A = R_A + K(S_A - E_A)$$

Στο (Tichumania) αυτή η ανανέωση γίνεται μετά από κάθε παιχνίδι. Για παράδειγμα έστω ότι δύο ομάδες, η μία (ομάδα A) με βαθμολογία Elo 1700 (έστω ότι αποτελείται από δύο παίκτες, όπου ο ένας έχει βαθμολογία Elo 1650 και ο άλλος 1750, άρα η βαθμολογία της ομάδας είναι ο μέσος όρος των βαθμολογιών Elo των παικτών) και η άλλη (ομάδα B) με βαθμολογία Elo 1400. Σε περίπτωση νίκης της ομάδας A οι νέες βαθμολογίες των ομάδων προκύπτουν ως εξής:

$$E_A = \frac{1}{1 + 10^{(1400-1700)/400}} = 0.849 \text{ και } E_B = \frac{1}{1 + 10^{(1700-1400)/400}} = 0.151$$

$$R'_A = R_A + K(S_A - E_A) = 1700 + 15(1 - 0.849) = 1702.265 \approx 1702$$

$$R'_B = R_B + K(S_B - E_B) = 1400 + 15(0 - 0.151) = 1397.735 \approx 1398$$

Σε περίπτωση νίκης της ομάδας B οι νέες βαθμολογίες των ομάδων προκύπτουν ως εξής:

$$R'_A = R_A + K(S_A - E_A) = 1700 + 15(0 - 0.849) = 1687.265 \approx 1687$$

$$R'_B = R_B + K(S_B - E_B) = 1400 + 15(1 - 0.151) = 1412.735 \approx 1413$$

Έτσι και η βαθμολογία Elo κάθε παίκτης της ομάδας θα αυξηθεί ή θα μειωθεί τόσο όσο αυξήθηκε ή μειώθηκε η βαθμολογία Elo της ομάδας.

Τα μεγαλύτερα μαθηματικά προβλήματα που προκύπτουν όσον αφορά την αρχική εργασία του Elo είναι η εύρεση του ακριβέστερου μοντέλου διανομής και της ακριβέστερης τιμής του συντελεστή K . Άλλο ένα πρόβλημα είναι ότι πολύ παίκτες ωθούνται στο να σταματήσουν να παίζουν συχνά προσπαθώντας έτσι να προστατεύσουν τη βαθμολογία Elo τους.

Το Elo χρησιμοποιείται σε πολλά παιχνίδια και αθλήματα, όπως στο Αμερικάνικο κολεγιακό ποδόσφαιρο, στο μπίιζμπολ, σαν ένα από τα συστήματα βαθμολόγησης στο τένις, από τη FIFA στο ποδόσφαιρο, στο χόκεϊ, στο ράγκμπι και στο Korfbal.

Επιτραπέζια παιχνίδια τα οποία χρησιμοποιούν τη βαθμολογία Elo είναι το Scrabble, το Lexulous, το τάβλι σε διαδικτυακούς server, το ευρωπαϊκό Go, το Magic: The Gathering, το Pokemon USA και τα παιχνίδια καρτών για το Star Trek και το Star Wars και άλλα.

Βιντεοπαιχνίδια τα οποία χρησιμοποιούν τη βαθμολογία Elo είναι το Overwatch, το Counter-Strike, το Guild Wars, το World of Warcraft, το League of Legends και άλλα. (Wikipedia, 2019)

ii) Συλλογή δεδομένων

Για τη συλλογή δεδομένων χρησιμοποιήθηκαν τα καταγεγραμμένα παιχνίδια μεταξύ παικτών στην πλατφόρμα Tichu του (BrettSpielWelt). Για διευκόλυνση μας χρησιμοποιήθηκε μία βάση δεδομένων την οποία διατηρούσε ο (Steinacher) και η οποία μας βοήθησε στο να ομαδοποιήσουμε τα παιχνίδια ανάλογα με τους παίκτες που συμμετέχουν σε αυτά. Οι παρτίδες οι οποίες επιλέχθηκαν ήταν αυτές όπου συμμετείχε τουλάχιστον ένας παίκτης με υψηλή βαθμολογία Elo, συν παρτίδες στις οποίες συμμετείχαν παίκτες με διακρίσεις στα τουρνουά Tichu, ανεξαρτήτως της βαθμολογίας Elo που είχαν σε περίπτωση που δεν ήταν αρκετά ενεργοί ώστε η βαθμολογία Elo τους να περιγράφει ικανοποιητικά τη δεξιότητα τους στο παιχνίδι Tichu. Επιλέχθηκαν συνολικά περίπου 1500 παιχνίδια για την εκπαίδευση των ευφυών συστημάτων και 100 παιχνίδια για τον έλεγχο της απόδοσης τους.

Από τα “χέρια” των παρτίδων των παιχνιδιών αυτών αφαιρέθηκαν τα “χέρια” του συμπαίκτη του παίκτη που δήλωσε Grand Tichu σε κάποιον γύρο. Αυτό γίνεται γιατί εάν δηλώσουν στον ίδιο γύρο Grand Tichu δύο συμπαίκτες τότε η ομάδα δε θα έχει κανένα πλεονέκτημα εάν ένας από τους δύο πετύχει και ξεφορτωθεί πρώτος όλες τις κάρτες του, μιας και η επιτυχής δήλωση δίνει 200 επιπλέον πόντους στην ομάδα, αλλά η ανεπιτυχής δήλωση αφαιρεί 200 πόντους από την ομάδα. Έτσι, όταν ένας παίκτης δηλώνει Grand Tichu, ο συμπαίκτης του δε δηλώνει Grand Tichu στον ίδιο γύρο ανεξαρτήτως του αν πιστεύει ότι το “χέρι” του είναι καλό για δήλωση Grand Tichu ή μη. Άρα αυτά τα “χέρια” δε βοηθούν στην αξιολόγηση της δήλωσης Grand Tichu ή όχι και αφαιρέθηκαν από τα δεδομένα μας.

Κεφάλαιο 6

Η επιλογή των χαρακτηριστικών των ευφυών συστημάτων

Στη μηχανική μάθηση, στην αναγνώριση προτύπων και στην επεξεργασία εικόνων, η εξαγωγή χαρακτηριστικών ξεκινά από ένα αρχικό σύνολο δεδομένων και παράγει κάποια χαρακτηριστικά που προσφέρουν μη περιττή πληροφορία, διευκολύνοντας τα ακόλουθα βήματα μάθησης και γενίκευσης, οδηγώντας, σε μερικές περιπτώσεις, σε καλύτερες ανθρώπινες ερμηνείες. Η εξαγωγή χαρακτηριστικών σχετίζεται με τη μείωση των διαστάσεων.

Όταν τα δεδομένα εισόδου σε έναν αλγόριθμο είναι πολύ μεγάλα για να υποστούν επεξεργασία και υπάρχει η υποψία ότι είναι περιττά (π.χ. η ίδια μέτρηση στα πόδια και τα μέτρα ή η επαναληπτικότητα των εικόνων που παρουσιάζονται ως pixels) τότε μπορούν να μετατραπούν σε ένα μειωμένο σύνολο των χαρακτηριστικών (ονομάζεται επίσης διάνυμα χαρακτηριστικών). Ο προσδιορισμός ενός υποσυνόλου των αρχικών χαρακτηριστικών ονομάζεται επιλογή χαρακτηριστικών. Οι επιλεγμένες λειτουργίες αναμένεται να περιέχουν τις σχετικές πληροφορίες από τα δεδομένα εισόδου, έτσι ώστε η επιθυμητή εργασία να μπορεί να εκτελεστεί με τη χρήση αυτής της μειωμένης αναπαράστασης αντί των πλήρων αρχικών δεδομένων.

Οι τεχνικές επιλογής χαρακτηριστικών χρησιμοποιούνται για τέσσερις λόγους: απλοποίηση των μοντέλων ώστε να διευκολυνθεί η ερμηνεία τους από ερευνητές / χρήστες, συντομότεροι χρόνοι εκπαίδευσης, για να αποφύγουμε την κατάρα των διαστάσεων, βελτιωμένη γενίκευση με τη μείωση του over fitting (τυπικά, μείωση της διακύμανσης).

Η κεντρική προϋπόθεση όταν χρησιμοποιείται μια τεχνική επιλογής χαρακτηριστικών είναι τα δεδομένα να περιέχουν ορισμένα χαρακτηριστικά που είναι είτε περιττά είτε άσχετα και μπορούν συνεπώς να αφαιρεθούν χωρίς να υποστούν πολλές απώλειες πληροφοριών. Οι έννοιες του περιττού και του άσχετου είναι διάφορες μεταξύ τους, καθώς ένα σχετικό

χαρακτηριστικό μπορεί να είναι περιττό με την παρουσία ενός άλλου σχετικού χαρακτηριστικού με το οποίο σχετίζεται έντονα. (Γιαννούλη, 2014)

Στην περίπτωση της δήλωσης του Grand Tichu, η οποία εξετάζεται στο πλαίσιο αυτής της διπλωματικής, κάθε παίκτης έχει 8 φύλλα τη στιγμή που καλείται να αποφασίσει αν θα δηλώσει ή όχι Grand Tichu. Το εάν θα δηλώσει Grand Tichu ή όχι εξαρτάται από πολλούς παράγοντες. Τα χαρακτηριστικά που εμείς αποφασίσαμε ότι είναι σημαντικά και κρατήσαμε είναι:

- Τα 8 φύλλα που έχει ο παίκτης.

Στο Tichu η τράπουλα αποτελείται από 56 φύλλα, άρα αυτά τα 8 φύλλα μπορούν να μοντελοποιηθούν με πολλούς τρόπους. Εμείς αποφασίσαμε να χωρίσουμε τα φύλλα σε 17 χαρακτηριστικά. Αυτά είναι:

1. Υπάρχει ή όχι το Mahjong στα 8 φύλλα: με τιμή 0 ή 1.
2. Υπάρχουν ή όχι τα Σκυλιά στα 8 φύλλα: με τιμή 0 ή 1.
3. Υπάρχει ή όχι ο Δράκος στα 8 φύλλα: με τιμή 0 ή 1.
4. Υπάρχει ή όχι ο Φοίνικας στα 8 φύλλα: με τιμή 0 ή 1.

Τα παραπάνω φύλλα εμφανίζονται μόνο μια φορά στην τράπουλα, σε αντίθεση με τα υπόλοιπα φύλλα (2 έως και Άσσο) που εμφανίζονται από 4 φορές το καθένα. Εδώ είχαμε την επιλογή να διαλέξουμε να θεωρήσουμε κάθε φύλλο ίδιας δυναμικότητας, για παράδειγμα το κόκκινο και το πράσινο δύο, ως διαφορετικά χαρακτηριστικά. Αποφασίσαμε όμως αντι για αυτό να αψηφήσουμε τη διαφορά στο χρώμα και έτσι για κάθε δυναμικότητα προκύπτει ένα χαρακτηριστικό - το πόσες φορές εμφανίζεται φύλλο αυτής της δυναμικότητας στο χέρι. Για λόγους κανονικοποίησης αποφασίσαμε όλα τα χαρακτηριστικά να έχουν τιμή από 0 μέχρι 1 και άρα:

5. Αριθμός Δυαριών: με τιμή $0.25 \times \#Δυαριών$

6. Αριθμός Τριαριών: με τιμή $0.25 \times \#Τριαριών$

7. Αριθμός Τεσσαριών: με τιμή $0.25 \times \#Τεσσαριών$

8. Αριθμός Πενταριών: με τιμή $0.25 \times \#Πενταριών$
9. Αριθμός Εξαριών: με τιμή $0.25 \times \#Εξαριών$
10. Αριθμός Εφταριών: με τιμή $0.25 \times \#Εφταριών$
11. Αριθμός Οχταριών: με τιμή $0.25 \times \#Οχταριών$
12. Αριθμός Εννιαριών: με τιμή $0.25 \times \#Εννιαριών$
13. Αριθμός Δεκαριών: με τιμή $0.25 \times \#Δεκαριών$
14. Αριθμός Βαλέδων: με τιμή $0.25 \times \#Βαλέδων$
15. Αριθμός Νταμών: με τιμή $0.25 \times \#Νταμών$
16. Αριθμός Ρηγάδων: με τιμή $0.25 \times \#Ρηγάδων$
17. Αριθμός Άσσων: με τιμή $0.25 \times \#Άσσων$

Στη συνέχεια σκεφτήκαμε πως η ύπαρξη κάποιων από τους συνδυασμούς του παιχνιδιού Tichu στο “χέρι” που εξετάζουμε πρέπει να είναι μέρος των χαρακτηριστικών. Αυτό συμβαίνει, διότι η ύπαρξη κάποιων δυνατών συνδυασμών στο χέρι ενός παίκτη αυξάνει την αξία του “χεριού” αυτού και μπορεί να τον ωθήσει σε μία δήλωση Grand Tichu. Οι πιθανοί συνδυασμοί είναι οι εξής: ζευγάρια, τριπλέτες, full house, συνεχόμενα ζευγάρια, κέντες, κέντες ίδιου χρώματος και καρτέ (4 ίδια φύλλα).

Τα ζευγάρια και οι τριπλέτες αποφασίσαμε ότι εκπροσωπούνται επαρκώς από τα χαρακτηριστικά Αριθμός Φύλλων Δυναμικότητας 2 έως Άσσο. Η πληροφορία για την ύπαρξη ζευγαριών ή τριπλετών ενός φύλλο ενθυλακώνεται μέσα στα χαρακτηριστικά αυτά, μιας και εάν ένα “χέρι” έχει 3 πεντάρια, τότε το χαρακτηριστικό 8 (Αριθμός Πενταριών) θα έχει την τιμή 0.75.

Αντίθετα, αποφασίσαμε να μην περιοριστούμε μόνο στα παραπάνω χαρακτηριστικά για τα καρέ (4 φύλλα ίδιας δυναμικότητας στο “χέρι”). Αυτό έγινε γιατί στο Tichu το καρέ αποτελεί μέρος μια ειδικής κατηγορίας συνδυασμών που ονομάζονται βόμβες. Οι βόμβες είναι οι πιο δυνατοί συνδυασμοί στο παιχνίδι και άρα αποφασίσαμε να δώσουμε ειδική μνεία στις βόμβες δημιουργώντας επιπλέον χαρακτηριστικά για την καταγραφή της εμφάνισης τους σε ένα “χέρι”.

i) Μοντελοποίηση των συνδυασμών Tichu

Συνεχόμενα Ζευγάρια

Τα συνεχόμενα ζευγάρια μοντελοποιούνται σαν μια λίστα χαρακτηριστικών. Το πρώτο χαρακτηριστικό της λίστας είναι το αν υπάρχουν συνεχόμενα ζευγάρια όπου το μεγαλύτερο ζευγάρι να είναι δυναμικότητας 3, ενώ το τελευταίο χαρακτηριστικό της λίστας είναι το αν υπάρχουν συνεχόμενα ζευγάρια όπου το μεγαλύτερο ζευγάρι να είναι δυναμικότητας Άσσου. Η μοντελοποίηση των παραπάνω από δύο συνεχόμενων ζευγαριών γίνεται ως εξής: βρίσκουμε το μεγαλύτερο ζευγάρι και στο χαρακτηριστικό του δίνουμε τιμή ίση με το πλήθος των συνεχόμενων ζευγαριών πριν από αυτό (δηλαδή μικρότερης δυναμικότητας). Η διαδικασία συνεχίζεται μέχρι να φτάσουμε στο συνεχόμενο ζευγάρι που έχει μόνο ένα ζευγάρι μικρότερης δυναμικότητας πριν από αυτό. Για παράδειγμα αν το “χέρι” ενός παίκτη είναι τα φύλλα (αγνοούμε τα χρώματα): 2-2-2-3-3-4-4-4-5-5-6-7-8-9 τότε το χαρακτηριστικό **Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Πεντάρια** έχει την τιμή 3, το χαρακτηριστικό **Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Τεσσάρια** έχει την τιμή 2 και το χαρακτηριστικό **Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Τριάρια** έχει την τιμή 1.

Όπως και πριν, σκοπός είναι να κανονικοποιήσουμε την τιμή των χαρακτηριστικών των συνεχόμενων ζευγαριών, ώστε να έχουν μέγιστη τιμή το 1. Παρατηρούμε ότι για το στάδιο της δήλωσης Grand Tichu κάθε “χέρι” αποτελείται από 8 φύλλα και άρα τα μέγιστα συνεχόμενα ζευγάρια είναι τέσσερα συνεχόμενα ζευγάρια. Με την μοντελοποίηση που περιγράψαμε παραπάνω τα τέσσερα συνεχόμενα ζευγάρια έχουν μέγιστη τιμή 3 (όπως ακριβώς φάνηκε και στο παραπάνω παράδειγμα που είχαμε τέσσερα συνεχόμενα ζευγάρια). Έτσι για την κανονικοποίηση των χαρακτηριστικών αυτών αρκεί η μοντελοποίηση με τον παραπάνω τρόπο και στη συνέχεια διαίρεση με το 3. Έτσι κανονικοποιούμε τα χαρακτηριστικά αυτά στο εύρος τιμών από 0 έως 1.

Τα χαρακτηριστικά που προκύπτουν μπορούν να βρεθούν στο τέλος της εργασίας.
(Παράρτημα - σελ. 118)

Κέντες

Οι κέντες (πέντε ή παραπάνω φύλλα διαδοχικής δυναμικότητας) μοντελοποιούνται και αυτές σαν μια λίστα χαρακτηριστικών. Το πρώτο χαρακτηριστικό της λίστας είναι το αν υπάρχουν πέντε συνεχόμενα φύλλα διαδοχικής δυναμικότητας όπου το μεγαλύτερο φύλλο να είναι δυναμικότητας 5, ενώ το τελευταίο χαρακτηριστικό της λίστας είναι το αν υπάρχουν πέντε συνεχόμενα φύλλα διαδοχικής δυναμικότητας όπου το μεγαλύτερο φύλλο να είναι δυναμικότητας Άσσου. Η μοντελοποίηση των κεντών γίνεται ως εξής: βρίσκουμε το μεγαλύτερο φύλλο και στο χαρακτηριστικό του δίνουμε τιμή ίση με το πλήθος των συνεχόμενων φύλλων πριν από αυτό (δηλαδή μικρότερης δυναμικότητας) τα οποία είναι υψηλότερα φύλλα κέντας αντίστοιχα συν ένα. Η διαδικασία συνεχίζεται μέχρι να φτάσουμε στο μεγαλύτερο φύλλο που έχει μόνο τέσσερα συνεχόμενα μικρότερης δυναμικότητας φύλλα πριν από αυτό. Για παράδειγμα αν το “χέρι” ενός παίκτη είναι τα φύλλα (αγνοούμε τα χρώματα): 2-3-4-5-6-7-8-9-J-J-K-K-Dr τότε το χαρακτηριστικό **Συνεχόμενης Δυναμικότητας Φύλλων με μεγαλύτερο φύλλο το Εννιά** έχει την τιμή 4, το χαρακτηριστικό **Συνεχόμενης Δυναμικότητας Φύλλων με μεγαλύτερο φύλλο το Οκτώ** έχει την τιμή 3, το χαρακτηριστικό **Συνεχόμενης Δυναμικότητας Φύλλων με μεγαλύτερο φύλλο το Επτά** έχει την τιμή 2 και το χαρακτηριστικό **Συνεχόμενης Δυναμικότητας Φύλλων με μεγαλύτερο φύλλο το Έξι** έχει την τιμή 1.

Όπως και πριν, θέλουμε να κανονικοποιήσουμε την τιμή των χαρακτηριστικών των συνεχόμενων ζευγαριών, ώστε να έχουν μέγιστη τιμή το 1. Παρατηρούμε ότι για το στάδιο της δήλωσης Grand Tichu κάθε “χέρι” αποτελείται από 8 φύλλα και άρα η μεγαλύτερη σε πλήθος φύλλων κέντα που μπορεί να υπάρχει είναι μια κέντα 8 φύλλων. Με την μοντελοποίηση που περιγράψαμε παραπάνω οι κέντες 8 φύλλων έχουν τιμή 4 (όπως ακριβώς φάνηκε και στο παραπάνω παράδειγμα που είχαμε μια κέντα 2 έως 9, δηλαδή 8 φύλλων). Έτσι για την κανονικοποίηση των χαρακτηριστικών αυτών αρκεί η μοντελοποίηση με τον παραπάνω τρόπο και στο τέλος διαίρεση με το 4. Έτσι κανονικοποιούμε τα χαρακτηριστικά αυτά στο εύρος τιμών από 0 έως 1. Να

σημειωθεί ότι επειδή είναι δυνατόν να χρησιμοποιηθεί και η ειδική κάρτα Mahjong για τη δημιουργία μιας κέντας Mahjong έως πέντε τα χαρακτηριστικά των κεντών ξεκινούν από μία κέντα με υψηλότερο φύλλο της κέντας να είναι το πέντε. Στο παράδειγμα που περιγράψαμε παραπάνω καλούμε τις κέντες οι οποίες είναι μικρότερου μήκους και με μικρότερη υψηλότερη κάρτα από το Εννιά, αλλά αποτελούν υποσύνολο της κέντας 2-9 (δηλαδή τις κέντες 2-8, 2-7, 2-6) ως υπο-κέντες.

Τα χαρακτηριστικά που προκύπτουν μπορούν να βρεθούν στο τέλος της εργασίας. (Παράρτημα – σελ. 118)

Full Houses

Τα Full Houses (πέντε φύλλα που είναι μια τριπλέτα και ένα ζευγάρι διαφορετικής δυναμικότητας) μοντελοποιούνται και αυτά σαν μια λίστα χαρακτηριστικών. Το πρώτο χαρακτηριστικό της λίστας είναι το αν υπάρχει ένα Full House όπου η δυναμικότητα της τριπλέτας να είναι 2 και άρα και η δυναμικότητα του Full House να είναι 2 μιας και η δυναμικότητα του Full House ισούται με τη δυναμικότητα της τριπλέτας αυτού, ενώ το τελευταίο χαρακτηριστικό της λίστας είναι το αν υπάρχει ένα Full House με δυναμικότητα Άσσου (δηλαδή με τριπλέτα στους Άσσους). Η μοντελοποίηση των Full Houses γίνεται ως εξής: βρίσκουμε την τριπλέτα φύλλων ίδιας δυναμικότητας και δίνουμε τιμή ίση με ένα στο αντίστοιχο χαρακτηριστικό. Για παράδειγμα αν το “χέρι” ενός παίκτη είναι τα φύλλα (αγνοούμε τα χρώματα): 2-3-4-4-4-6-6-9-9-J-J-K-K-Dr τότε το χαρακτηριστικό **Full House Δυναμικότητας 4** έχει την τιμή 1. Αν δεν υπήρχαν ζευγάρια στο “χέρι” (για παράδειγμα 2-3-4-4-4-6-7-8-9-10-J-Q-K-Dr) τότε η τιμή του χαρακτηριστικού **Full House Δυναμικότητας 4** θα ήταν ίση με μηδέν. Τα χαρακτηριστικά αυτά δε χρειάζονται περαιτέρω κανονικοποίηση για να έχουν τιμές στο εύρος τιμών από 0 έως 1 μιας και είναι δυαδικά χαρακτηριστικά (υπάρχουν ή δεν υπάρχουν).

Τα χαρακτηριστικά που προκύπτουν μπορούν να βρεθούν στο τέλος της εργασίας. (Παράρτημα – σελ. 118)

Βόμβες

Οι βόμβες (είτε καρέ ή κέντα ίδιο χρώμα) μοντελοποιούνται και αυτές σαν μια λίστα χαρακτηριστικών. Τα χαρακτηριστικά αυτά προκύπτουν από τα παραπάνω ως εξής: Βρίσκουμε χαρακτηριστικά για τις κέντες ίδιου χρώματος με τον τρόπο που περιγράψαμε παραπάνω για τις κανονικές κέντες απλά λαμβάνοντας υπόψη και το χρώμα. Στη συνέχεια δημιουργούμε άλλο ένα χαρακτηριστικό που είναι ο συνολικός αριθμός βομβών του “χεριού” όπου λαμβάνεται υπόψη και ο αριθμός των καρέ στο “χέρι” (μέχρι 2 για τη στιγμή της δήλωσης Grand Tichu) και το οποίο κανονικοποιείται και αυτό σε εύρος τιμών 0 έως 1.

Τα χαρακτηριστικά που προκύπτουν μπορούν να βρεθούν στο τέλος της εργασίας (βλ. Παράρτημα 1 σελ. 118)

ii) Επιλογή επιπλέων χαρακτηριστικών

Στο Tichu, όπως έχει αναφερθεί και παραπάνω, νικήτρια ομάδα είναι η ομάδα η οποία θα φτάσει πρώτη τους 1000 πόντους (ή η ομάδα που θα έχει τους παραπάνω πόντους άμα και οι δύο ομάδες φτάσουν τους 1000 πόντους στον ίδιο γύρο). Αυτό σημαίνει ότι κατά τη διάρκεια του παιχνιδιού ένας παίκτης μπορεί να είναι περισσότερο ή λιγότερο δεκτικός στο να δηλώσει Grand Tichu, ακόμα και με το ίδιο «χέρι», ανάλογα με τη φάση του παιχνιδιού και τη διαφορά στη βαθμολογία. Για παράδειγμα, αν οι πόντοι της ομάδας A είναι 500 και οι πόντοι της ομάδας B είναι 500 πριν τη δήλωση Grand Tichu σε έναν γύρο ένας παίκτης της ομάδας A θα είναι πιο διστακτικός στο να δηλώσει Grand Tichu από ότι θα ήταν εάν οι πόντοι της ομάδας A ήταν -100 και της ομάδας B 700, όπου θα θεωρούσε πιο δελεαστικό να ρισκάρει με σκοπό να μειώσει τη διαφορά των πόντων σε περίπτωση επιτυχούς δήλωσης. Από την άλλη, στο τελευταίο σενάριο ένας παίκτης της ομάδας B θα ήταν πιο διστακτικός στο να δηλώσει Grand Tichu γιατί μια αποτυχημένη δήλωση θα απομάκρυνε την ομάδα του πολύ από τον στόχο των 1000 πόντων. Με μία τέτοια διαφορά από την ομάδα A και με απόσταση μόλις 300 πόντων από τους 1000 ένα πιο συντηρητικό στυλ παιχνιδιού, όπου οι δηλώσεις Grand Tichu μειώνονται και αυξάνονται οι δηλώσεις Tichu ή ακόμα και καμία δήλωση είναι κάτι το οποίο περιμένουμε από τον παίκτη της ομάδας B.

Με βάση τα παραπάνω και την εμπειρία μας στο παιχνίδι Tichu, αποφασίσαμε να επιλέξουμε τη προσθήκη δύο χαρακτηριστικών για να περιγράψουν το σκορ των δύο ομάδων τη στιγμή που κάθε παίκτης καλείται να δηλώσει ή όχι Grand Tichu με ένα συγκεκριμένο “χέρι”. Τα χαρακτηριστικά αυτά είναι οι πόντοι της αντίπαλης ομάδας και της ομάδας του παίκτη αντίστοιχα. Αφού έχουμε κανονικοποιήσει όλα τα χαρακτηριστικά με τιμές 0 έως 1, κάναμε το ίδιο και εδώ θεωρώντας ως 0 το να βρίσκεται μία ομάδα στους -1000 πόντους, καθώς θεωρούμε ότι οτιδήποτε πιο χαμηλά από τους -1000 πόντους θα έχει το ίδιο αποτέλεσμα στο να επηρεάσει έναν παίκτη κατά τη διάρκεια δήλωσης Grand Tichu όσο οι -1000 πόντοι. Με το ίδιο σκεπτικό οι 1000 πόντοι αντιστοιχούν στο 1 και άρα τα δύο χαρακτηριστικά που προκύπτουν είναι:

$$94. \text{ Σκορ αντίπαλης ομάδας: } \frac{\# \text{Πόντων αντίπαλης ομάδας (min: -1000, max: 1000)} + 1000}{2000}$$

$$95. \text{ Σκορ δικής μας ομάδας: } \frac{\# \text{Πόντων δικής μας ομάδας (min: -1000, max: 1000)} + 1000}{2000}$$

Αυτά τα 95 χαρακτηριστικά είναι τα χαρακτηριστικά τα οποία χρησιμοποιούνται για την εκπαίδευση των ευφυών συστημάτων που θα περιγράψουμε στο κεφάλαιο που ακολουθεί.

Κεφάλαιο 7

Η επιλογή και η εκπαίδευση των ευφύων συστημάτων

i) Επιλογή ευφύων συστημάτων

Όπως έχει αναφερθεί παραπάνω, σκοπός μας είναι η δημιουργία κάποιων ομαδοποιητών που θα διαλέγουν τη δήλωση ή μη Grand Tichu δεδομένων των χαρακτηριστικών που περιγράφηκαν παραπάνω. Επίσης θέλουμε να λάβουμε ως παράμετρο την επιθυμία μας για τη δημιουργία τριών στυλ πρακτόρων ανάλογα με την επιθετικότητα τους. Έτσι, ένας πράκτορας θα δηλώνει Grand Tichu με μεγαλύτερη ευκολία, εάν είναι επιθετικός και με μεγαλύτερη δυσκολία εάν είναι διστακτικός. Συνεπώς, η αναζήτηση για την επιλογή ευφύων συστημάτων επικεντρώθηκε στην εύρεση συστημάτων τα οποία θα μας δίνουν ως αποτελέσματα βαθμονομημένες πιθανότητες, που μπορούν να ερμηνευτούν ως εμπιστοσύνη σε μια απόφαση. Με αυτόν τον τρόπο θα μπορούμε να δημιουργήσουμε διάφορους πράκτορες οι οποίοι για τιμές πάνω από διαφορετικά κατώφλια εμπιστοσύνης (τιμές εμπιστοσύνης ανάμεσα στο μηδέν και το ένα) θα δηλώνουν Grand Tichu, ενώ για τιμές κάτω από αυτά τα κατώφλια δε θα δηλώνουν Grand Tichu.

Επιλέξαμε να χρησιμοποιήσουμε δύο μοντέλα επιτηρούμενης μηχανικής μάθησης: το μοντέλο της Λογιστικής Παλινδρόμησης (Logistic Regression Model) και το μοντέλο των Μηχανών Διανυσμάτων Υποστήριξης με παλινδρόμηση (Support Vector Regression Model). Οι αλγόριθμοι και ο τρόπος λειτουργίας αυτών των δύο μοντέλων θα περιγράφουν αναλυτικά στη συνέχεια. Τα μοντέλα αυτά επιστρέφουν μια τιμή εμπιστοσύνης και άρα μπορούν να χρησιμοποιηθούν σαν ομαδοποιητές με τη χρήση ενός κατωφλίου. Κάθε τιμή πάνω από το επιλεγμένο κατώφλι οδηγεί σε δήλωση Grand Tichu, ενώ κάθε τιμή κάτω από το επιλεγμένο κατώφλι οδηγεί σε μη δήλωση Grand Tichu για ένα συγκεκριμένο χέρι 8 καρτών. Για την εκπαίδευση των παραπάνω μοντέλων χρησιμοποιήθηκαν τα ίδια δεδομένα εκπαίδευσης και τα ίδια δεδομένα δοκιμής.

ii) Επιλογή Μέτρων Απόδοσης

Για τον έλεγχο της απόδοσης των μοντέλων χρησιμοποιήθηκαν τέσσερα μέτρα απόδοσης. Κάθε μοντέλο, ανάλογα με την επιλογή του ή όχι να δηλώσει Grand Tichu σε ένα χέρι σε σχέση με το αν δήλωσε ή όχι Grand Tichu στο ίδιο χέρι ο αληθινός ανθρώπινος παίκτης, έχει τα εξής δυνατά αποτελέσματα για την επιλογή του: ορθώς θετική (true positive - TP), ψευδώς θετική (false positive - FP), ορθά αρνητική (true negative - TN) και ψευδώς αρνητική (false negative - FN) δήλωση. Δηλαδή σε περίπτωση που το μοντέλο προβλέψει ότι ένας παίκτης δεν έχει δηλώσει Grand Tichu και αυτός έχει δηλώσει τότε το αποτέλεσμα είναι ψευδώς αρνητική δήλωση του μοντέλου, μιας και θα έπρεπε να είναι θετική. Αν ο παίκτης δηλώσει Grand Tichu και το μοντέλο προβλέψει σωστά τη δήλωση αυτή τότε το αποτέλεσμα είναι μια ορθά θετική δήλωση.

Τα πέντε μέτρα απόδοσης των μοντέλων που χρησιμοποιήθηκαν είναι τα εξής (Παπαδάκης, 2016, σσ. 5-6) : ορθότητα δηλώσεων (accuracy), ακρίβεια δηλώσεων (precision), ανάκληση (recall), η βαθμολογία F1 (F1 score) και η χαρακτηριστική καμπύλη λειτουργίας του δέκτη (receiver operating characteristic curve - Καμπύλη ROC) όπου:

$$\text{ορθότητα δηλώσεων (accuracy): } \frac{\# \text{ ορθών προβλ έψεων}}{\# \text{ συνολικ ός αριθμ ός προβλ έψεων}} = \frac{\#(TP + TN)}{\#(TP + FP + TN + FN)}$$

$$\text{ακρίβεια δηλώσεων (precision): } \frac{\# \text{ ορθών προβλ έψεων}}{\# \text{ θετικών προβλ έψεων}} = \frac{\#TP}{\#(TP + FP)}$$

$$\text{ανάκληση (recall): } \frac{\# \text{ ορθών προβλ έψεων}}{\# \text{ ορθών θετικών και ψευδών αρνητικών προβλ έψεων}} = \frac{\#TP}{\#(TP + FN)}$$

$$\text{βαθμολογία F1 (F1 score): } 2 \times \frac{\text{ανάκληση} \times \text{ακρίβεια θετικών δηλώσεων}}{\text{ανάκληση} + \text{ακρίβεια θετικών δηλώσεων}} =$$
$$2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

Η βαθμολογία F1 δηλαδή προκύπτει από τον σταθμισμένο μέσο όρο της ακρίβειας δηλώσεων και της ανάκλησης. Επομένως, αυτή η βαθμολογία λαμβάνει υπόψη τόσο τα ψευδώς θετικά όσο και τα ψευδώς αρνητικά. Διαισθητικά δεν είναι τόσο εύκολο να

καταλάβει κανείς τη βαθμολογία F1, όσο είναι να καταλάβει την ακρίβεια των δηλώσεων, αλλά η βαθμολογία F1 είναι συνήθως πιο χρήσιμη από την ακρίβεια των δηλώσεων, ειδικά αν υπάρχει μια ανομοιογενής ταξική κατανομή.

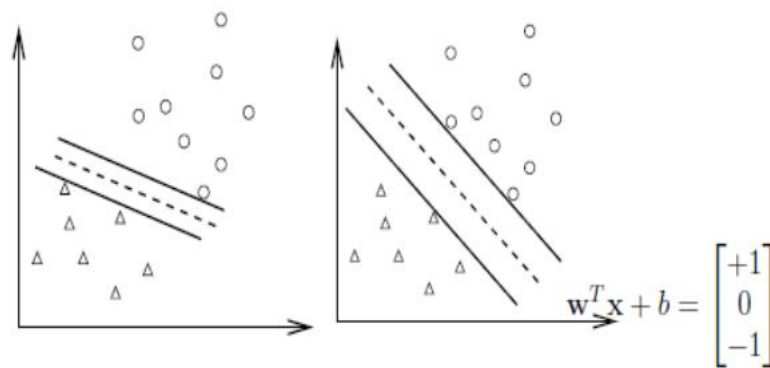
Η χαρακτηριστική καμπύλη λειτουργίας του δέκτη προκύπτει από τη γραφική παράσταση του ποσοστού ψευδών θετικών προβλέψεων με το ποσοστό των ορθών θετικών προβλέψεων για κάθε τιμή του κατωφλίου. Ένα «τέλειο» μοντέλο ταξινομεί σωστά όλες τις δηλώσεις, δηλαδή οι βαθμολογίες που αποδίδει στα χέρια όπου έγινε δήλωση Grand Tichu, είναι πάντα υψηλότερες από τις βαθμολογίες στα χέρια όπου δεν έγινε δήλωση Grand Tichu. Για το μοντέλο αυτό, η καμπύλη ROC είναι μια ευθεία παράλληλη στον άξονα x. Αντίθετα, για ένα ανεπαρκές μοντέλο η καμπύλη ROC είναι μια ευθεία γραμμή που ξεκινά από την αρχή των αξόνων. Τέλος, η καμπύλη ROC ενός ικανοποιητικού μοντέλου βρίσκεται στο ενδιάμεσο των δύο περιπτώσεων.

Κριτήριο για την αποτελεσματικότητα ενός μοντέλου λοιπόν, αποτελεί η μορφή της καμπύλης ROC και συγκεκριμένα το εμβαδόν της περιοχής κάτω από την καμπύλη (AUC, Area Under Curve). Προφανώς το εμβαδόν κάτω από την καμπύλη για ένα «τέλειο» μοντέλο θα λαμβάνει την τιμή 1, ενώ για ένα ανεπαρκές δεν θα υπερβαίνει το 0.5. Για ένα ικανοποιητικό μοντέλο το εμβαδόν κάτω από την καμπύλη κυμαίνεται μεταξύ του 0.5 και 1, με υψηλότερες τιμές να υποδεικνύουν υψηλότερη αποτελεσματικότητα. (Κυρίτσης, 2014, σσ. 69-70)

iii) Μηχανές Διανυσμάτων Υποστήριξης

Οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) αποτελούν ένα εργαλείο για την επεξεργασία δεδομένων και βασίζονται στην στατιστική θεωρία εκμάθησης. Αναπτύχθηκαν από τον Vapnik και την Cortes το 1995. Η αρχή των SVM μπορούμε να πούμε ότι έγινε με το (Vapnik V., 1979) πάνω στην θεωρία στατιστικής εκμάθησης (statistical learning theory) και ειδικά με την περαιτέρω ανάπτυξη αυτής της θεωρίας από τον Vapnik. Η αρχή της μεθόδου έγκειται στο διαχωρισμό κλάσεων (σύνολο σημείων με ίδια ιδιότητα) με επίπεδα (υπερεπίπεδα) και με βάση αυτόν το διαχωρισμό γίνεται η κατάταξη ενός νέου σημείου σε κάποια από τις κλάσεις που διαχωρίστηκαν. Η τεχνική αυτή άρχισε πολύ γρήγορα να αναπτύσσεται και από άλλους επιστήμονες σε άρθρα όπως στο (J. K. Anlauf, 1989) όπου πρότειναν τα υπερεπίπεδα διαχωρισμού με μέγιστο διάκενο, στο (T. Poggio, 1990) όπου μελετάται η χρήση των συναρτήσεων κελύφους και στο (Boser, 1992) όπου τα SVM περίπου με τη μορφή που τα γνωρίζουμε προτάθηκαν στο συνέδριο COLT. Η ανάπτυξη της μεθόδου συνεχίστηκε και στο (Cortes C., 1995), χάρη στους οποίους έχουμε την μελέτη που αναπτύσσει τους διαχωριστές “χαλαρού” διακένου (soft margin classifiers) όπως επίσης και την επέκταση των SVM στα προβλήματα προσέγγισης συνάρτησης (παλινδρόμηση - regression) (Vapnik, 1995).

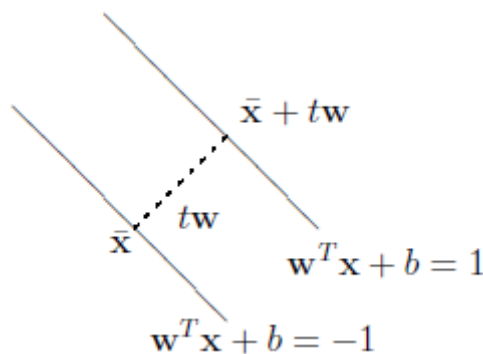
Η αρχική ιδέα των γραμμικών SVM ταξινομητών είναι να χρησιμοποιηθεί ένα γραμμικό διαχωριστικό υπερεπίπεδο για να δημιουργηθεί ένας ταξινομητής που θα διαχωρίζει τις δύο κλάσεις. Παίρνοντας ως δεδομένο ένα σύνολο διανυσμάτων $x_i, i = 1, \dots, l$ μήκους n , και ένα διάνυσμα y όπου το y_i έχει την τιμή 1, αν το x_i ανήκει στη κλάση ένα και -1 αν το x_i ανήκει στη κλάση δύο. Ο SVM ταξινομητής προσπαθεί να βρει το διαχωριστικό υπερεπίπεδο με το μεγαλύτερο περιθώριο μεταξύ των δύο κλάσεων, που μετράται κατά μήκος μιας κάθετης προς το υπερεπίπεδο γραμμής. Για παράδειγμα, στο παρακάτω σχήμα οι δύο κλάσεις θα μπορούσαν να διαχωριστούν πλήρως από μια διακεκομμένη γραμμή $w^T x + b = 0$. Ζητούμενο είναι να βρεθεί η γραμμή με το μεγαλύτερο περιθώριο, δηλαδή η απόσταση μεταξύ των δύο κλάσεων των δεδομένων εκπαίδευσης να είναι όσο το δυνατόν μεγαλύτερη. (Γιαννούλη, 2014, σσ. 47-48)



7-1 Διαχωριστικό υπερεπίπεδο, (Γιαννούλη, 2014, σ. 48)

Η απόσταση μεταξύ των $w^T x + b = 1$ και $w^T x + b = -1$ υπολογίζεται με τον ακόλουθο τρόπο:

1. Θεωρούμε ένα σημείο \underline{x} στο $w^T x + b = -1$
2. Έστω w το κάθετο στο υπερεπίπεδο $w^T x + b = -1$ διάνυσμα.
3. Ξεκινώντας από το \underline{x} και κινούμενοι κατά μήκος της κατεύθυνσης w υποθέτουμε ότι $\underline{x} + tw$ αγγίζει το επίπεδο $w^T x + b = 1$. Αφού $w^T(\underline{x} + tw) + b = 1$ και $w^T \underline{x} + b = -1$ έχουμε ότι $tw^T w = 2$

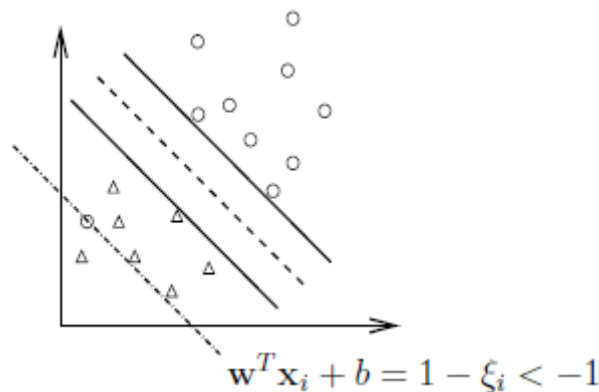


7-2 Απόσταση μεταξύ κλάσεων, (Γιαννούλη, 2014, σ. 48)

Οπότε η απόσταση (δηλαδή το μήκος του tw) είναι $\|tw\| = \frac{2\|w\|}{w^T w} = \frac{2}{\|w\|}$. Η ελαχιστοποίηση του $\frac{2}{\|w\|}$ είναι ισοδύναμη με τη μεγιστοποίηση του $\frac{w^T w}{2}$, επομένως έχουμε το ακόλουθο πρόβλημα: $\min_{w,b} \frac{1}{2} w^T w$ με $y_i(w^T x + b) \geq 1, i = 1, \dots, l$. Πρέπει λοιπόν τα δεδομένα της κλάσης ένα να είναι στη δεξιά πλευρά του $w^T x + b = 0$, ενώ τα δεδομένα της άλλης

κλάσης στην αριστερή. Σημειώνουμε ότι ο λόγος της μεγιστοποίησης της απόστασης μεταξύ $w^T x + b = \pm 1$ βασίζεται στον Vapnik.

Ωστόσο, πρακτικά προβλήματα μπορεί να μην είναι γραμμικά διαχωρίσιμα με τον τρόπο που περιγράφεται στην προηγούμενη παράγραφο. Έτσι οι Cortes και Vapnik στο (Cortes C., 1995) εισήγαγαν τις χαλαρές μεταβλητές $\xi_i, i = 1, \dots, l$ στο πρόβλημα ελαχιστοποίησης και άρα το πρόβλημα πλέον μετατρέπεται σε $\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i$ με $y_i(w^T x + b) \geq 1 - \xi_i, i = 1, \dots, l$, όπου “επιτρέπεται” στα δεδομένα εκπαίδευσης να μην είναι στη σωστή πλευρά του διαχωριστικού υπερεπίπεδου $w^T x + b = 0$. Αυτό συμβαίνει όταν $\xi_i \geq 1$ και ένα παράδειγμα είναι το παρακάτω σχήμα:



7-3 Διαχωριστικό υπερεπίπεδο με χαλαρές μεταβλητές, (Γιαννούλη, 2014, σ. 51)

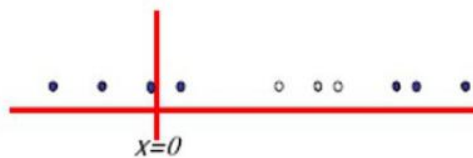
Το νέο πρόβλημα όπως περιγράφηκε είναι πάντα εφικτό αφού για κάθε (w, b) έχουμε $\xi_i \equiv \max(0, 1 - y_i(w^T x + b)), i = 1, \dots, l$ και οδηγούμαστε στο (w, b, ξ) που είναι εφικτή λύση.

Χρησιμοποιώντας αυτή τη ρύθμιση, μπορεί να ανησυχούμε ότι για γραμμικά διαχωρίσιμα δεδομένα, κάποια $\xi_i > 1$ και ως εκ τούτου τα αντίστοιχα δεδομένα είναι λάθος ταξινομημένα. Στην περίπτωση που τα περισσότερα δεδομένα, εκτός από κάποια θορυβώδη, μπορούν να διαχωριστούν από μια γραμμική συνάρτηση, θα θέλαμε το $w x + b$ να ταξινομεί σωστά την πλειοψηφία των σημείων. Έτσι, στην αντικειμενική συνάρτηση προσθέτουμε έναν όρο ποινής $C \sum_{i=1}^l \xi_i$, όπου $C > 0$ είναι μια παράμετρος ποινής. Για να έχουμε την αντικειμενική αξία όσο το δυνατόν μικρότερη, τα περισσότερα ξ_i πρέπει να είναι μηδέν, έτσι ώστε ο περιορισμός να παίρνει την αρχική του μορφή. Θεωρητικά μπορούμε να

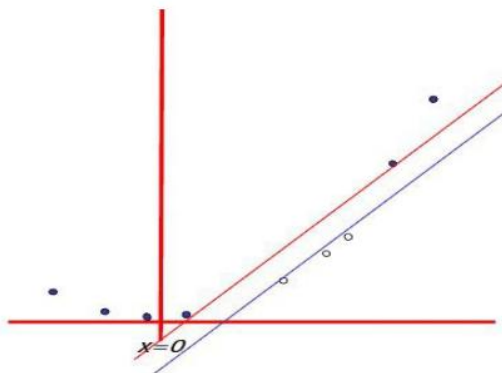
αποδείξουμε ότι αν τα δεδομένα είναι γραμμικά διαχωρίσιμα και το C είναι μεγαλύτερο από ένα συγκεκριμένο αριθμό όλα τα ξ_i είναι μηδέν. Δηλαδή για μεγάλο C , το περιθώριο είναι μικρό και άρα τα λάθη είναι μικρά, ενώ για μικρό C , το περιθώριο είναι μεγάλο και άρα τα λάθη είναι μεγάλα.

Δυστυχώς μια τέτοια ρύθμιση δεν είναι αρκετή, γιατί αν τα δεδομένα κατανέμονται με ένα μη γραμμικό τρόπο, χρησιμοποιώντας μόνο μια γραμμική συνάρτηση πολλά δεδομένα εκπαίδευσης είναι στη λάθος πλευρά του υπερεπιπέδου. Έτσι εμφανίζεται λάθος τοποθέτηση και η συνάρτηση απόφασης δεν αποδίδει καλά.

Για να τοποθετήσουμε τα δεδομένα κατάρτισης καλύτερα, μπορούμε να σκεφτούμε τη χρήση μιας μη γραμμικής καμπύλης. Άλλη μια προσέγγιση είναι να χαρτογραφήσουμε τα δεδομένα σε ένα χώρο υψηλότερων διαστάσεων, όπως φαίνεται στο επόμενο σχήμα.



Δύο κλάσεις που δεν είναι γραμμικώς διαχωρίσιμες στον R^1



Οι εικόνες των δύο κλάσεων στο χώρο R^2 είναι γραμμικώς διαχωρίσιμες.

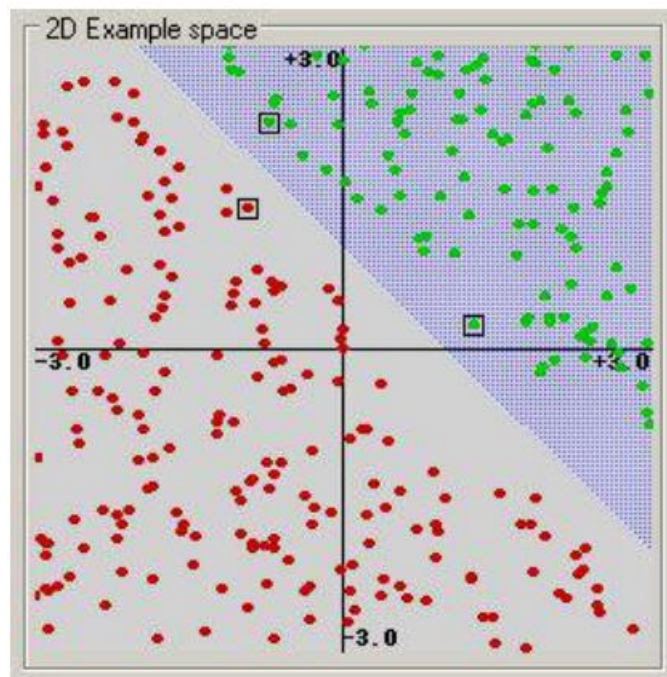
7-4 Κλάσεις γραμμικώς διαχωρίσιμες στο χώρο R^2 , αλλά μη γραμμικώς διαχωρίσιμες στο χώρο R^1 ,
(Γιαννούλη, 2014, σ. 52)

Έτσι, μη γραμμικές μηχανές διανυσμάτων υποστήριξης μετατρέπουν τον αρχικό χώρο εισόδου σε έναν υψηλότερων διαστάσεων χώρο των χαρακτηριστικών. Πιο συγκεκριμένα, το

δεδομένο εκπαίδευσης χαρτογραφείται σε ένα διάνυσμα σε ένα χώρο υψηλότερων διαστάσεων $\varphi(x) = [\varphi(x_1), \varphi(x_2), \dots]$. Σε αυτόν το χώρο υψηλότερων διαστάσεων, είναι πιο πιθανό τα δεδομένα να μπορούν να διαχωριστούν γραμμικά. (Γιαννούλη, 2014, σσ. 50-52)

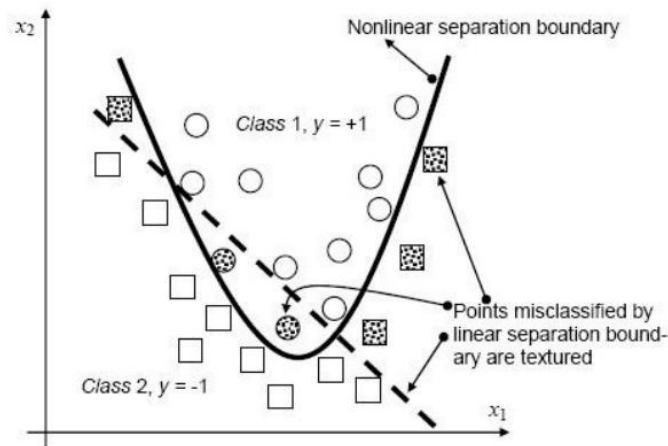
Οι πυρήνες ή συναρτήσεις πυρήνων που χρησιμοποιούνται στις μηχανές διανυσμάτων υποστήριξης για χαρτογράφηση είναι οι εξής:

1. Γραμμικός πυρήνας (Linear Kernel): $K(x, x) = x * x + \gamma$, όπου η παράμετρος γ καθορίζεται από το χρήστη. Αυτός ο τύπος πυρήνα έχει εφαρμογή μόνο σε γραμμικώς διαχωρίσιμα προβλήματα.



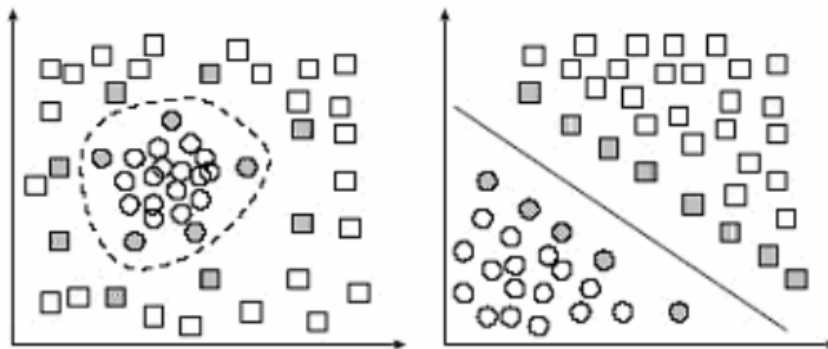
7-5 Παράδειγμα γραμμικού πυρήνα, όπου τα σημεία που περικλείονται με τετραγωνικό πλαίσιο αποτελούν διανύσματα υποστήριξης, (Γιαννούλη, 2014, σ. 56)

2. Πολυωνυμικός πυρήνας (Polynomial Kernel): $K(\mathbf{x}, \mathbf{x}') = (\mathbf{x}, \mathbf{x}' + 1)^p$, όπου η παράμετρος p καθορίζεται από τον χρήστη. Στο παρακάτω σχήμα φαίνεται η σαφώς μεγαλύτερη διαχωριστική ικανότητα του SVM πολυωνυμικού πυρήνα.



7-6 Παράδειγμα διαχωρισμού μη-γραμμικού προβλήματος με χρήση SVMs γραμμικού (διακεκομμένη) και πολυωνυμικού πυρήνα (συνεχής γραμμή), (Γιαννούλη, 2014, σ. 56)

3. Ακτινωτός πυρήνας (Radial Basis Function (RBF) ή Gaussian Kernel): $K(\mathbf{x}, \mathbf{x}') = e^{-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2s^2}}$, όπου η παράμετρος s καθορίζεται από τον χρήστη.



7-7 Αριστερή υπό-εικόνα: Διαχωρισμός με ακτινωτό πυρήνα στον αρχικό χώρο. Δεξιά υπό-εικόνα: Γραμμικός διαχωρισμός σε χώρο υψηλότερης διάστασης με τη βοήθεια μετασχηματισμού ακτινωτού πυρήνα, (Γιαννούλη, 2014, σ. 57)

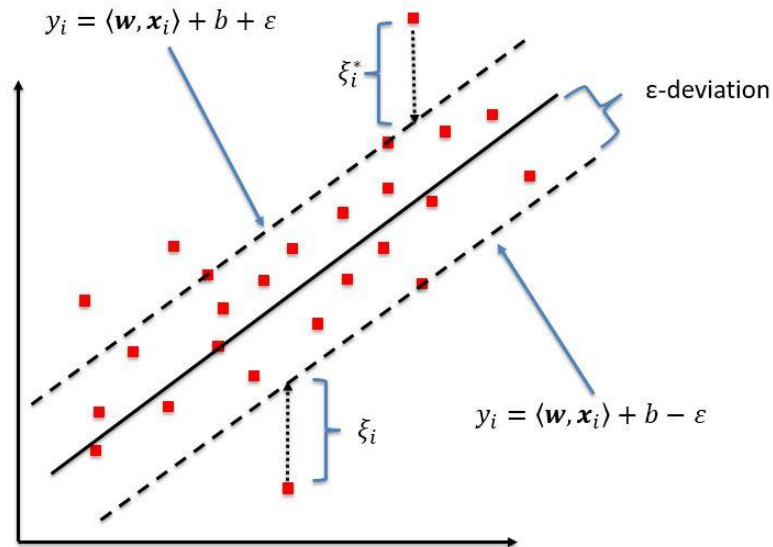
4. Σιγμοειδής πυρήνας (Sigmoid Kernel): $K(\mathbf{x}, \mathbf{x}') = \tanh(\kappa * \mathbf{x} * \mathbf{x}' + \theta)$, όπου κ, θ οι παράμετροι των πυρήνων. (Γιαννούλη, 2014, σσ. 54-57)

iv) Μηχανές Διανυσμάτων Υποστήριξης με παλινδρόμηση

Παρόμοια με την ταξινόμηση μοντέλων διανυσματικής υποστήριξης, τα μοντέλα παλινδρόμησης με διανύσματα υποστήριξης (Support Vector Regression models - SVR) έχουν έναν αριθμό ελεύθερων παραμέτρων που πρέπει να είναι συντονισμένες. Αυτές οι παράμετροι είναι οι πυρήνες K με τις αντίστοιχες παραμέτρους λ , η παράμετρος ε , και η σταθερά κόστους C . Οι τεχνικές που ισχύουν στις μηχανές διανυσμάτων υποστήριξης μεταφέρονται στα μοντέλα παλινδρόμησης με διανύσματα υποστήριξης, με τη διαφορά ότι η συνάρτηση απώλειας 0-1 αντικαθίσταται από μια συνάρτηση απώλειας που υπολογίζει πόσο καλά το μοντέλο ταιριάζει με τις παρατηρήσεις. Δηλαδή, αντί να μετράμε πόσο εσφαλμένη είναι η ταξινόμηση, μετράμε πόσο διαφορετική είναι η προβλεπόμενη τιμή από την παρατηρούμενη τιμή.

Η πιο κοινή εκτίμηση σφάλματος για τη συνάρτηση παλινδρόμησης είναι το μέσο τετραγωνικό σφάλμα (MSE). Ορίζουμε μια συνάρτηση απώλειας L_2 που υπολογίζει το τετράγωνο του υπολοίπου για μία παρατήρηση (\underline{x}, y) δοθέντος ενός μοντέλου \hat{f} , $L_2(y, \hat{f}(\underline{x})) = (y - \hat{f}(\underline{x}))^2$ και ενός συνόλου εκπαίδευσης για την παλινδρόμηση $D = \{(\underline{x}_1, y_1), (\underline{x}_2, y_2), \dots, (\underline{x}_l, y_l)\} \subset R^n \times R$ ορίζουμε το μέσο τετραγωνικό σφάλμα υπολογισμένο στο D , $MSE_D [\hat{f}_d[k, \lambda, \varepsilon, C]] = \frac{1}{l} \sum_{i=1}^l L_2(y_i, \hat{f}_d[k, \lambda, \varepsilon, C](\underline{x}_i)), (\underline{x}_i, y_i) \in D$. Σε αυτή την περίπτωση μπορούμε να ερμηνεύσουμε το μέσο τετραγωνικό σφάλμα ως τη μέση απώλεια L_2 του μοντέλου \hat{f}_d , από τα δεδομένα του συνόλου D . Θα πρέπει να αναφερθεί ότι μια άλλη δημοφιλής μέθοδος προσδιορισμού του σφάλματος ενός μοντέλου παλινδρόμησης είναι η ρίζα του μέσου τετραγωνικού σφάλματος (root-mean-squared error) που προέρχεται από το μέσο τετραγωνικό σφάλμα απλά με τη λήψη της τετραγωνικής του ρίζας. (Γιαννούλη, 2014, σσ. 72-73)

Στο παρακάτω σχήμα, προβάλλεται ένα μονοδιάστατο μοντέλο SVR. Μόνο τα σημεία πάνω και κάτω από τις διακεκομμένες γραμμές χρησιμοποιούνται για την διαδικασία πρόβλεψης.



7-8 Γράφημα ενός μονοδιάστατου μοντέλου μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση, (Tania Kleynhans, 2017, p. 8)

ν) Μοντέλο Μηχανών Διανυσμάτων Υποστήριξης με παλινδρόμηση

Για τη δημιουργία του μοντέλου Μηχανής Διανυσμάτων Υποστήριξης με παλινδρόμηση (Support Vector Regression Model - SVR) χρησιμοποιήθηκε η βιβλιοθήκη scikit learn της Python, και συγκεκριμένα η κλάση `sklearn.svm.SVR` (SciKit Learn SVM). Η κλάση αυτή παρέχει όλες τις απαραίτητες μεθόδους για την εκπαίδευση του μοντέλου με τη χρήση των δεδομένων εκπαίδευσης, αλλά και για τη χρησιμοποίησή του για εξαγωγή αποτελεσμάτων πάνω στα δεδομένα δοκιμής. Με βάση τις μελέτες των (Staelin, 2003) και (Chih-Wei Hsu, 2003) οι παράμετροι οι οποίες χρησιμοποιήθηκαν για να διαφοροποιήσουμε τα διάφορα μοντέλα Μηχανών Διανυσμάτων Υποστήριξης που χρησιμοποιούμε είναι οι υπερ-παράμετροι C και γ . Η παράμετρος C είναι η παράμετρος της ποινής του όρου λάθους και η παράμετρος γ είναι ο συντελεστής πυρήνα για τον πυρήνα που επιλέξαμε, δηλαδή τον πυρήνα Ακτινικής λειτουργίας βάσης (Radial Basis Function kernel – RBF kernel). Ένα γραμμικό SVR, δηλαδή ένα SVR χωρίς πυρήνα, έχει μόνο μια ουσιαστική διαφορά από το μοντέλο γραμμικής παλινδρόμησης – το ότι χρησιμοποιεί μόνο το υποσύνολο των δεδομένων

που βρίσκονται μακριά από τη γραμμή διαχωρισμού και ότι η βελτιστοποίηση του μοντέλου SVR δεν εξαρτάται από τις διαστάσεις του χώρου εισόδου.

Διαισθητικά, η παράμετρος γ καθορίζει πόσο μεγάλη είναι η επιρροή ενός μόνο παραδείγματος εκπαίδευσης. Χαμηλές τιμές του γ σημαίνουν μεγάλη επιρροή, ενώ υψηλές τιμές σημαίνουν μικρότερη επιρροή. Η παράμετρος γ λοιπόν μπορεί να θεωρηθεί ως το αντίστροφο της ακτίνας επιρροής των δειγμάτων που επιλέγονται από το μοντέλο ως διανύσματα υποστήριξης.

Η παράμετρος C ανταλλάσει τη σωστή ταξινόμηση των παραδειγμάτων εκπαίδευσης έναντι στη μεγιστοποίηση του περιθωρίου της συνάρτησης λήψης της απόφασης. Για μεγαλύτερες τιμές του C , ένα μικρότερο περιθώριο θα γίνει αποδεκτό αν η συνάρτηση λήψης της απόφασης είναι καλύτερη στη σωστή ταξινόμηση όλων των σημείων κατάρτισης. Μια χαμηλότερη τιμή της παραμέτρου C ενθαρρύνει ένα μεγαλύτερο περιθώριο, επομένως μία απλοποιημένη συνάρτηση λήψης αποφάσεων, μειώνοντας όμως την ακρίβεια του μοντέλου που εκπαιδεύουμε.

Παρατηρείται ότι για κάποιες ενδιάμεσες τιμές της παραμέτρου γ παίρνουμε εξίσου αποτελεσματικά μοντέλα όταν η παράμετρος C γίνεται πολύ μεγάλη. Δεν είναι απαραίτητο να επιβληθεί μεγαλύτερο περιθώριο. Η ακτίνα του πυρήνα RBF λειτουργεί από μόνη της ως καλός διαρθρωτικός κανονιστής. Στην πράξη όμως θα ήταν ενδιαφέρον να απλοποιηθεί η συνάρτηση λήψης αποφάσεων με μια μικρότερη τιμή της παραμέτρου C ώστε να ευνοηθούν τα μοντέλα που χρησιμοποιούν λιγότερη μνήμη και είναι πιο γρήγορα στο να επιλέγουν αποτελέσματα.

Με βάση τις δύο μελέτες που αναφέρθηκαν παραπάνω επιλέξαμε να γίνει η εκπαίδευση των μοντέλων με χρήση αναζήτησης δικτύου (grid search) με τις παραμέτρους C και γ να παίρνουν τιμές $\log_2 C \in \{-5, -4, \dots, 15\}$ και $\log_2 \gamma \in \{-15, -14, \dots, 3\}$.

Επίσης χρησιμοποιήθηκε η διαδικασία διασταυρούμενης επικύρωσης *10-φορών*, δηλαδή το δείγμα κατανεμήθηκε τυχαία σε 10 ισόποσα υπό-δείγματα. Από τα υπο-δείγματα αυτά, διατηρήθηκε ένα υπο-δείγμα ως δεδομένο επικύρωσης για τη δοκιμή του μοντέλου και τα υπόλοιπα 9 υπο-δείγματα ως δεδομένα εκπαίδευσης. Στη συνέχεια, η διαδικασία διασταυρούμενης επικύρωσης επαναλαμβάνεται 10 φορές, με καθένα από τα 10 υπο-

δείγματα να χρησιμοποιείται ακριβώς μια φορά ως δεδομένο επικύρωσης. Τα αποτελέσματα μπορούν στη συνέχεια να υπολογιστούν κατά μέσο όρο και να παραχθεί μια ενιαία εκτίμηση. Το πλεονέκτημα αυτής της μεθόδου έναντι της επανειλημμένης τυχαίας υπο-δειγματοληψίας είναι ότι όλες οι παρατηρήσεις χρησιμοποιούνται τόσο για την εκπαίδευση όσο και για την επικύρωση και κάθε παρατήρηση χρησιμοποιείται για την επαλήθευση ακριβώς μία φορά.

Το διάνυσμα των χαρακτηριστικών κανονικοποιήθηκε, ώστε κάθε χαρακτηριστικό να έχει τιμές ανάμεσα στο μηδέν και στο ένα όπως περιγράφηκε στο προηγούμενο κεφάλαιο.

Συνολικά εκπαιδεύτηκαν 399 Μοντέλα Μηχανών Διανυσμάτων Υποστήριξης με παλινδρόμηση, δηλαδή ένα για κάθε πιθανό συνδυασμό των υπερ-παραμέτρων C και γ .

Για κάθε συνδυασμό C, γ επιλέχθηκαν 2000 κατώφλια (από το 0 μέχρι το 1 με βήμα 0.0005). Τα κατώφλια αυτά αποτελούν την τιμή εμπιστοσύνης. Για μεγαλύτερες τιμές εμπιστοσύνης από αυτήν, το συγκεκριμένο μοντέλο ξεκινά να δηλώνει Grand Tichu, ενώ για τιμές εμπιστοσύνης μικρότερες αυτής το μοντέλο δεν αποφασίζει να δηλώσει Grand Tichu για το συγκεκριμένο χέρι.

vi) Λογιστική Παλινδρόμηση

Μία άλλη κατηγορία παλινδρόμησης, μη γραμμικής για τη λύση προβλημάτων δυαδικής εξαρτημένης μεταβλητής, είναι η λογιστική παλινδρόμηση. Σε αντίθεση με τη γραμμική παλινδρόμηση που προσαρμόζει μία ευθεία γραμμή στα δεδομένα, η λογιστική παλινδρόμηση προσαρμόζει μία συνάρτηση S στα δεδομένα.

Έστω $Y = \{0,1\}$ όπου $Y = 0$ αντιστοιχεί στη μία κατάσταση της μεταβλητής για παράδειγμα μη δήλωση Grand Tichu και $Y = 1$ αντιστοιχεί σε μία άλλη κατάσταση (δήλωση Grand Tichu). Έστω επίσης ότι p είναι η πιθανότητα να συμβεί $Y = 1$, δηλαδή $p = P(Y = 1)$. Τότε το μοντέλο της λογιστικής παλινδρόμησης επιχειρεί να λύσει την ακόλουθη εξίσωση:

$$\ln\left(\frac{p}{1-p}\right) = a + \beta \times X + \varepsilon, \text{ ή}$$

$$\left[\frac{p}{(1-p)} \right] = e^a e^B e^\varepsilon$$

Ο λόγος $\frac{p}{(1-p)}$ λέγεται λόγος των πιθανοτήτων της μεταβλητής Y , ενώ η κατανομή του νεπέριου λογάριθμου αυτής της ποσότητας κυμαίνεται στο διάστημα $[0,1]$.

Για να ελαχιστοποιηθούν οι παρατηρήσεις που ταξινομούνται εσφαλμένα, θα πρέπει να προβλέψουμε το $Y = 1$ όταν $p \geq 0.5$ και $Y = 0$ όταν $p \leq 0.5$. Αυτό σημαίνει ότι μπορούμε να πάρουμε έναν γραμμικό ταξινομητή με χρήση της λογιστικής παλινδρόμησης. Η λογιστική παλινδρόμηση όχι μόνο δείχνει που είναι το όριο μεταξύ των τάξεων, αλλά επίσης δείχνει ότι οι πιθανότητες των κλάσεων εξαρτώνται από την απόσταση από το όριο, με ένα συγκεκριμένο τρόπο και ότι πηγαίνουν πιο γρήγορα προς τα άκρα όταν το μέτρο του b είναι μεγαλύτερο.

Είναι αυτές οι δηλώσεις σχετικά με τις πιθανότητες που κάνουν τη λογιστική παλινδρόμηση κάτι παραπάνω από έναν ταξινομητή. Κάνει ισχυρότερες, λεπτομερέστερες προβλέψεις και μπορεί να εκπαιδευτεί με διαφορετικό τρόπο, αλλά ακόμα και αυτές οι ισχυρές προβλέψεις μπορεί να είναι λανθασμένες.

Η χρήση της λογιστικής παλινδρόμησης για την πρόβλεψη πιθανών κλάσεων είναι μια επιλογή που έχει να κάνει με την μοντελοποίηση, ακριβώς όπως είναι μια επιλογή μοντελοποίησης το να προβλέψεις ποσοτικές μεταβλητές χρησιμοποιώντας γραμμική παλινδρόμηση. Σε καμία από τις δύο περιπτώσεις η καταλληλότητα του μοντέλου δεν είναι εγγυημένη. Αρχικά ξεκινάμε θέτοντας το μοντέλο με σκοπό να έχουμε ένα αποτέλεσμα το οποίο να μπορούμε να μελετήσουμε και στο τέλος να ελέγξουμε εάν πραγματικά ταιριάζει με τα δεδομένα ή εάν έχει συστηματικά ελαττώματα.

Η λογιστική παλινδρόμηση είναι ένα από τα πιο συχνά χρησιμοποιούμενα εργαλεία στην εφαρμοσμένη στατιστική και τη διακριτή ανάλυση δεδομένων. Οι τέσσερις κύριοι λόγοι για τους οποίους συμβαίνει αυτό είναι η παράδοση, το ότι συχνά λειτουργεί εκπληκτικά καλά ως ταξινομητής, το ότι είναι στενά συνδεδεμένη με κατανομές “εκθετικής οικογένειας” οι οποίες εμφανίζονται σε πολλά πλαίσια στη στατιστική θεωρία και στη φυσική και άρα υπάρχουν πολλά προβλήματα που μπορούν να μετατραπούν σε λογιστική παλινδρόμηση και τέλος, η

ευκολία που προκύπτει ως προς την ανάλυση των πινάκων των λογαριθμικών αποδόσεων.
(Shalizi, 2013, pp. 231-235)

vii) Μοντέλο Λογιστικής Παλινδρόμησης

Για τη δημιουργία του μοντέλου Λογιστικής Παλινδρόμησης (Logistic Regression - LR) χρησιμοποιήθηκε η βιβλιοθήκη `scikit learn` της Python, και συγκεκριμένα η κλάση `sklearn.linear_model.Logistic Regression` (SciKit Learn Logistic Regression). Η κλάση αυτή παρέχει όλες τις απαραίτητες μεθόδους για την εκπαίδευση του μοντέλου με τη χρήση των δεδομένων εκπαίδευσης, αλλά και για τη χρησιμοποίησή του για εξαγωγή αποτελεσμάτων πάνω στα δεδομένα δοκιμής. Η παράμετρος C είναι η παράμετρος της αντίστροφης ισχύς κανονικοποίησης των αποτελεσμάτων (inverse regularization strength) και πρέπει να είναι μια θετική τιμή. Για μικρές τιμές του C , αυξάνουμε την ισχύ κανονικοποίησης και δημιουργούμε απλά μοντέλα τα οποία υπο-προσαρμόζονται στα δεδομένα. Για μεγάλες αξίες του C , χαμηλώνουμε την ισχύ κανονικοποίησης, κάτι το οποίο συνεπάγεται ότι το μοντέλο επιτρέπεται να αυξήσει την πολυπλοκότητά του και, ως εκ τούτου, να υπερ-προσαρμοστεί στα δεδομένα εκπαίδευσης (overfitting).

Επιλέξαμε να γίνει η εκπαίδευση των μοντέλων με χρήση αναζήτησης δικτύου (grid search) με την παράμετρο C να παίρνει τιμές $C \in \{0,1, \dots, 10\}$ και ανάλογα με την τιμή του C για την οποία το μοντέλο έχει την καλύτερη απόδοση να γίνεται επανεκπαίδευση σε ένα εύρος τιμών C γύρω από την τιμή αυτή. Η διαδικασία αυτή επαναλαμβάνεται έως ότου τα C γύρω από την τιμή να πάψουν πλέον να δίνουν καλύτερα αποτελέσματα.

Επίσης χρησιμοποιήθηκε η διαδικασία διασταυρούμενης επικύρωσης *10-φορών*, όπως αναλύθηκε παραπάνω για το μοντέλο της μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση.

Το διάνυσμα των χαρακτηριστικών κανονικοποιήθηκε ώστε κάθε χαρακτηριστικό να έχει τιμές ανάμεσα στο μηδέν και στο ένα όπως περιγράφηκε στο προηγούμενο κεφάλαιο.

Συνολικά εκπαιδεύτηκαν 50 Μοντέλα Λογιστικής Παλινδρόμησης, δηλαδή ένα για κάθε τιμή C μέχρι να μη βρεθεί καλύτερο αποτέλεσμα για διαφορετική τιμή C .

Για το μοντέλο που προέκυψε για τη συγκεκριμένη, “βέλτιστη” τιμή C επιλέχθηκαν 2000 κατώφλια (από το 0 μέχρι το 1 με βήμα 0.0005). Τα κατώφλια αυτά αποτελούν την τιμή εμπιστοσύνης. Για μεγαλύτερες τιμές εμπιστοσύνης από αυτήν, το συγκεκριμένο μοντέλο ξεκινά να δηλώνει Grand Tichu, ενώ για τιμές εμπιστοσύνης μικρότερες αυτής το μοντέλο αποφασίζει να μη δηλώσει Grand Tichu για το συγκεκριμένο χέρι.

Κεφάλαιο 8

Αποτελέσματα των μοντέλων μετά την εκπαίδευση

Ο κώδικας για την επεξεργασία των δεδομένων, την επιλογή και τον υπολογισμό των τιμών των χαρακτηριστικών την εκπαίδευση των μοντέλων και την εξαγωγή των αποτελεσμάτων, μαζί με τα δεδομένα εκπαίδευσης και τις γραφικές παραστάσεις των αποτελεσμάτων μπορεί να βρεθεί στο αποθετήριο (Asimakoroulos, 2019)

i) Αποτελέσματα μοντέλου Μηχανής Διανυσμάτων Υποστήριξης με παλινδρόμηση

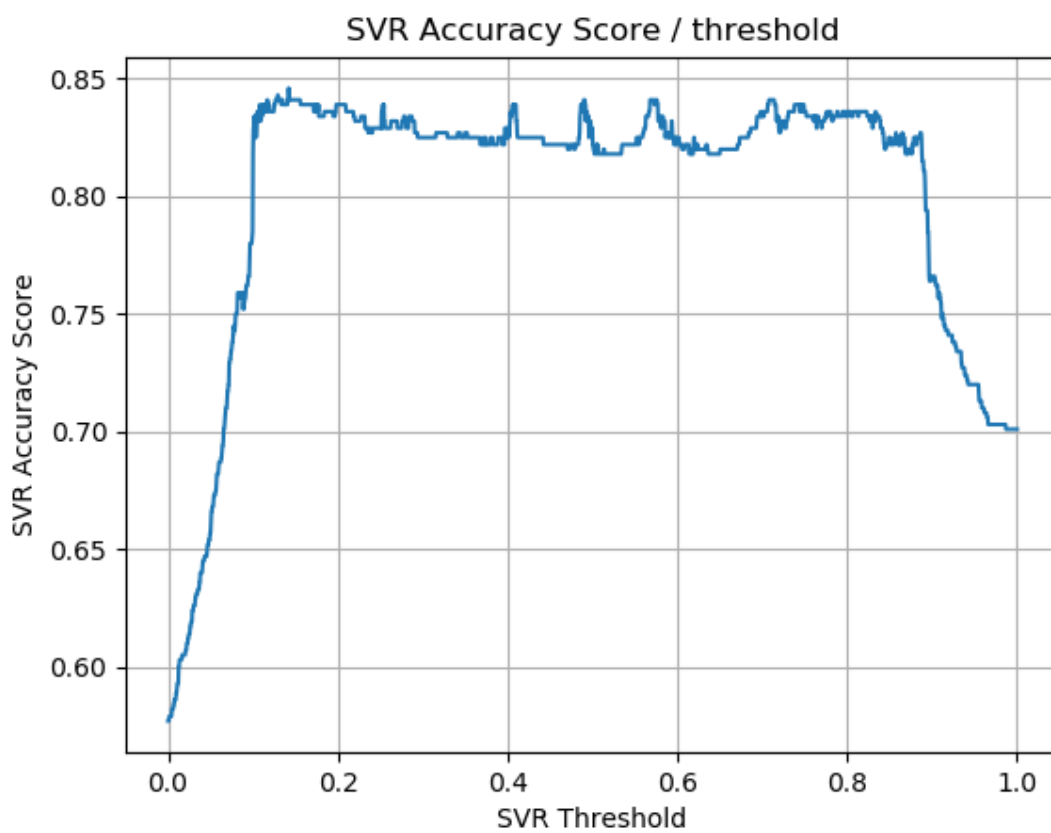
Για το Μοντέλο Μηχανής Διανυσμάτων Υποστήριξης με παλινδρόμηση, επιλέχθηκε για κάθε τιμή του κατωφλίου το μοντέλο SVR με τιμές (C, γ) για το οποίο το εκάστοτε μέτρο απόδοσης παίρνει τη μέγιστη τιμή.

Ο χρόνος που χρειάστηκε ώστε να εκπαιδευτεί το μοντέλο για όλες τις τιμές (C, γ) που επιλέχθηκαν με βάση τις προτάσεις ερευνητών που αναφέραμε παραπάνω ήταν πολύ μεγάλος. Αυτό οφείλεται γιατί το χρονικό κόστος για τον υπολογισμό μιας απόστασης είναι μεγάλο, λόγω του μη-γραμμικού πυρήνα που χρησιμοποιήσαμε (RBF kernel). Ο πυρήνας αυτός δημιουργεί έναν ομαδοποιητή όπου η εκμάθηση είναι βασισμένη σε στιγμιότυπα και άρα στη χειρότερη περίπτωση αν έχουμε n το πλήθος δεδομένα εκπαίδευσης, η ταξινόμηση ενός μόνο νέου στιγμιότυπου έχει υπολογιστική πολυπλοκότητα $O(n)$. (Norvig, 2003)

Παρακάτω παρουσιάζονται τα αποτελέσματα για τα μέτρα απόδοσης που περιγράφηκαν παραπάνω.

Ορθότητα Δηλώσεων (Accuracy)

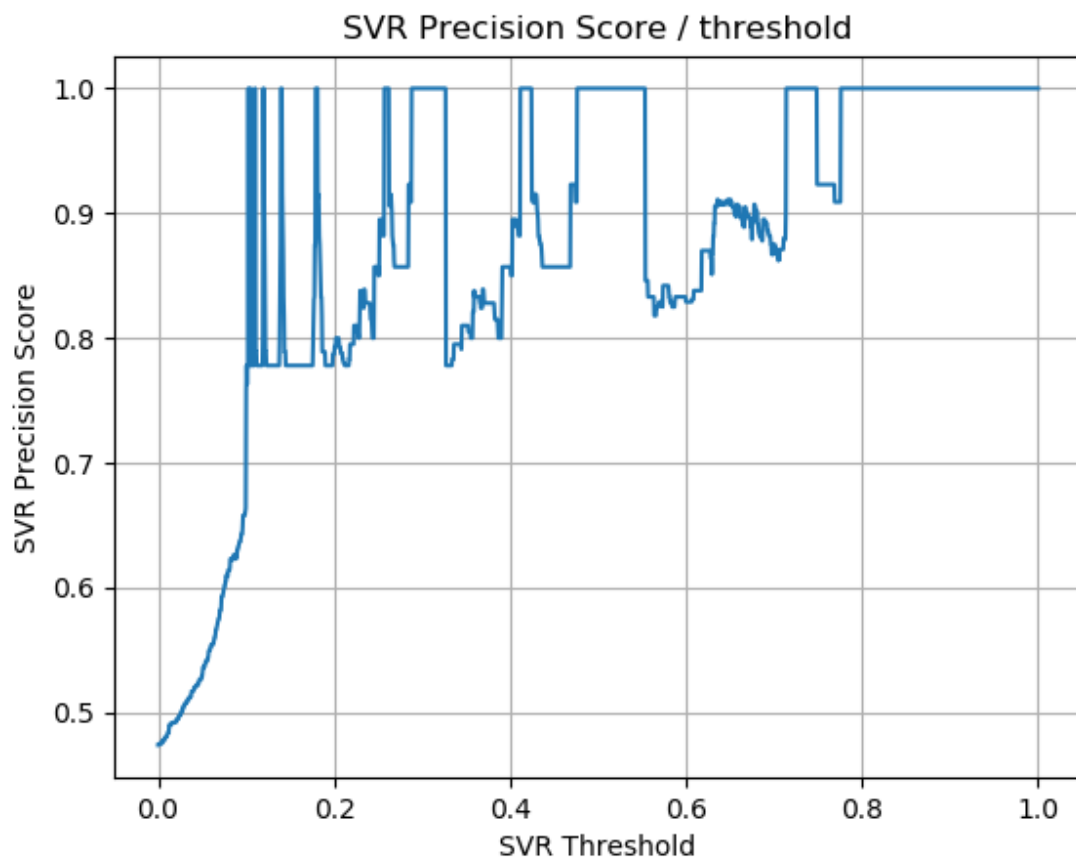
Το μοντέλο SVR παρουσιάζει υψηλή βαθμολογία στην ορθότητα δηλώσεων για τιμές κατωφλίου που δεν είναι ακραίες, δηλαδή βρίσκονται μακριά από το 0 και το 1. Για τιμές από 0.1 έως και 0.9 η ορθότητα δηλώσεων του μοντέλου SVR είναι πάνω από 80% με τη μέγιστη τιμή να είναι 84.58% για ($C = 2$, $\gamma = 0.25$, τιμή κατωφλίου: 0.142).



8-1 Το ποσοστό ορθότητας δηλώσεων του μοντέλου SVR ανά τιμή κατωφλίου

Ακρίβεια Δηλώσεων (Precision)

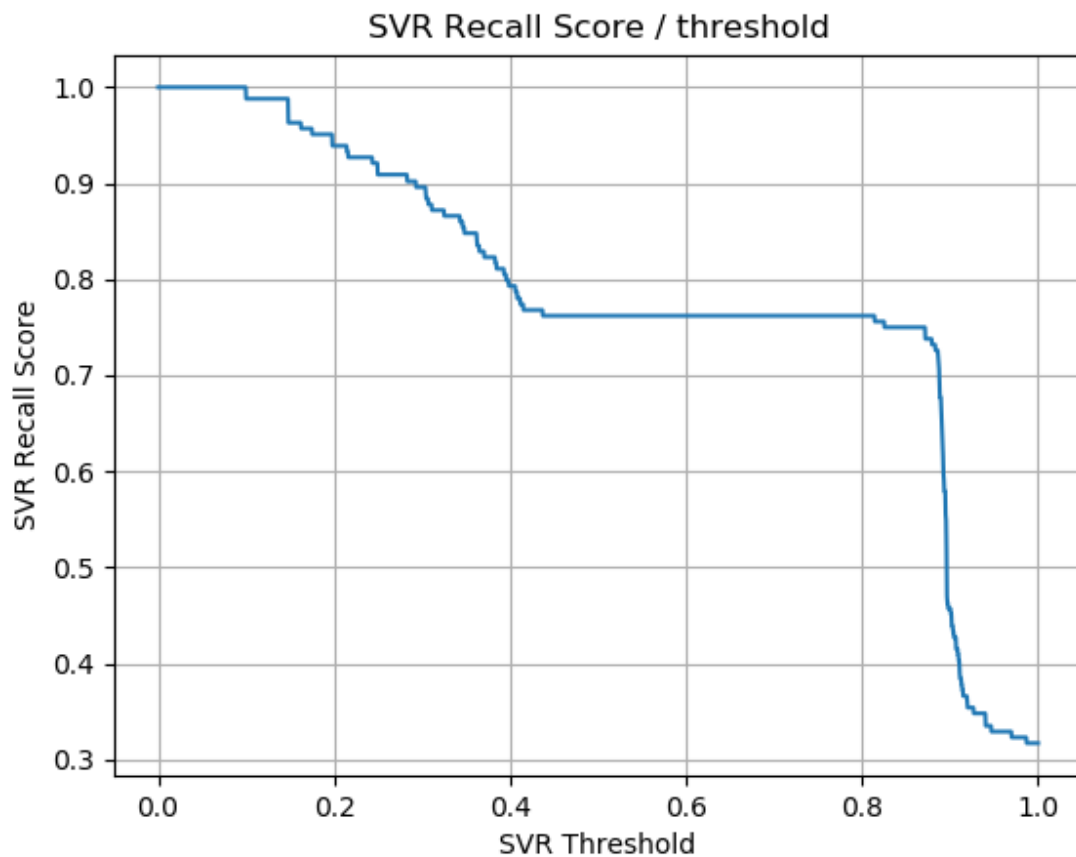
Το μοντέλο SVR παρουσιάζει υψηλή βαθμολογία στην ακρίβεια δηλώσεων για τιμές κατώφλιου που δεν είναι κοντά στο 0. Όσο το κατώφλι πλησιάζει την τιμή 1, παρατηρούμε αύξηση στην ακρίβεια των δηλώσεων. Οι τιμές 1 που παρουσιάζονται για χαμηλά κατώφλια εξηγούνται μιας και δεν πρόκειται για το ίδιο SVR μοντέλο ανά κατώφλι αλλά για διαφορετικά μοντέλα με διαφορετικές τιμές παραμέτρων C , γ ώστε να προκύπτει το καλύτερο δυνατό αποτέλεσμα ανά τιμή κατώφλιου. Η μέγιστη τιμή ακρίβειας δηλώσεων προκύπτει για πολλά κατώφλια και τιμές C , γ και είναι το ποσοστό 100%.



8-2 Το ποσοστό ακρίβειας δηλώσεων του μοντέλου SVR ανά τιμή κατώφλιου

Ακρίβεια Ανάκλησης (Recall)

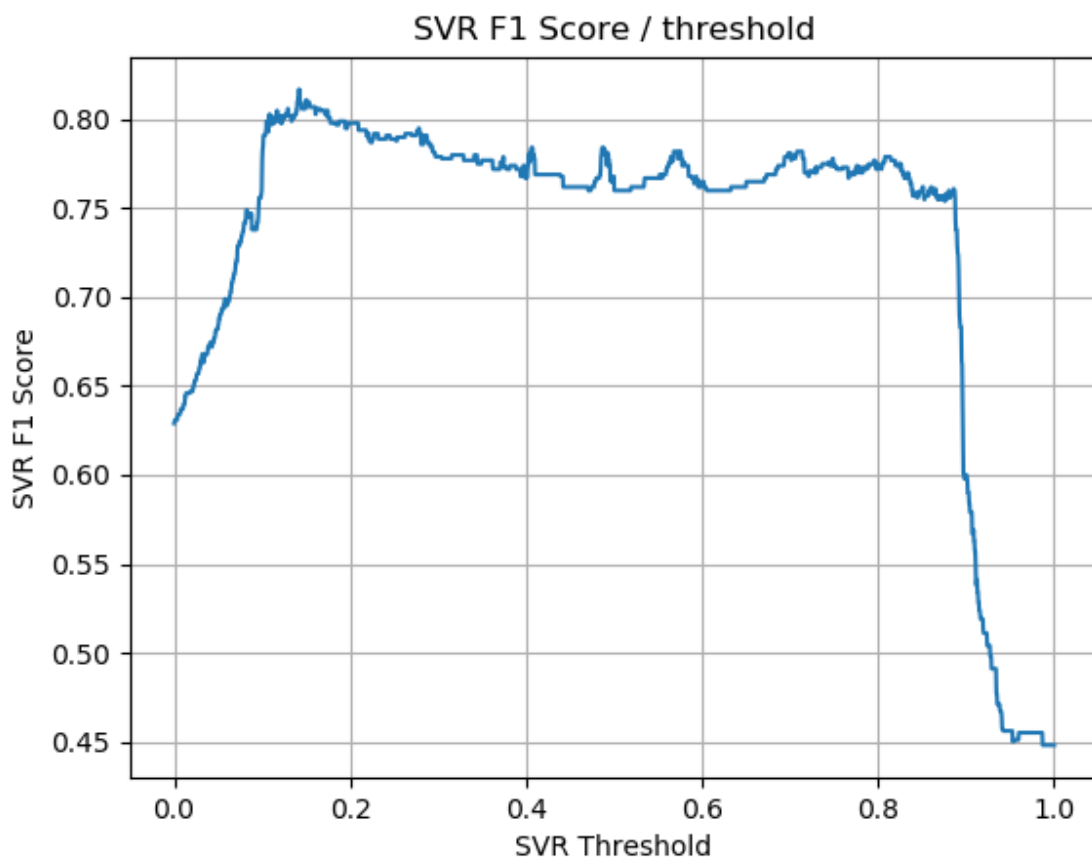
Το μοντέλο SVR παρουσιάζει υψηλή βαθμολογία στην ακρίβεια ανάκλησης για τιμές κατώφλιου που δεν είναι κοντά στο 1. Όσο το κατώφλι πλησιάζει την τιμή 0, παρατηρούμε αύξηση στην ακρίβεια ανάκλησης. Η μέγιστη τιμή ακρίβειας ανάκλησης προκύπτει για πολλά κατώφλια κοντά στο 0 και τιμές C , γ και είναι το ποσοστό 100%.



8-3 Το ποσοστό επιτυχούς ανάκλησης του μοντέλου SVR ανά τιμή κατώφλιου

Ακρίβεια Βαθμολογίας F1 (F1 Score)

Το μοντέλο SVR παρουσιάζει υψηλή βαθμολογία F1 για τιμές κατωφλίου που δεν είναι ακραίες, δηλαδή για τιμές κατωφλίου που δεν είναι κοντά στο 0 και στο 1. Όσο το κατώφλι πλησιάζει την τιμή 1, παρατηρούμε αισθητή μείωση στη βαθμολογία F1. Η μέγιστη τιμή ακρίβειας βαθμολογίας F1 προκύπτει για τιμή κατωφλίου 0.142 και τιμές $C = 2$, $\gamma = 0.25$ και είναι ίση με 81.67%.



8-4 Η βαθμολογία F1 του μοντέλου SVR ανά τιμή κατωφλίου

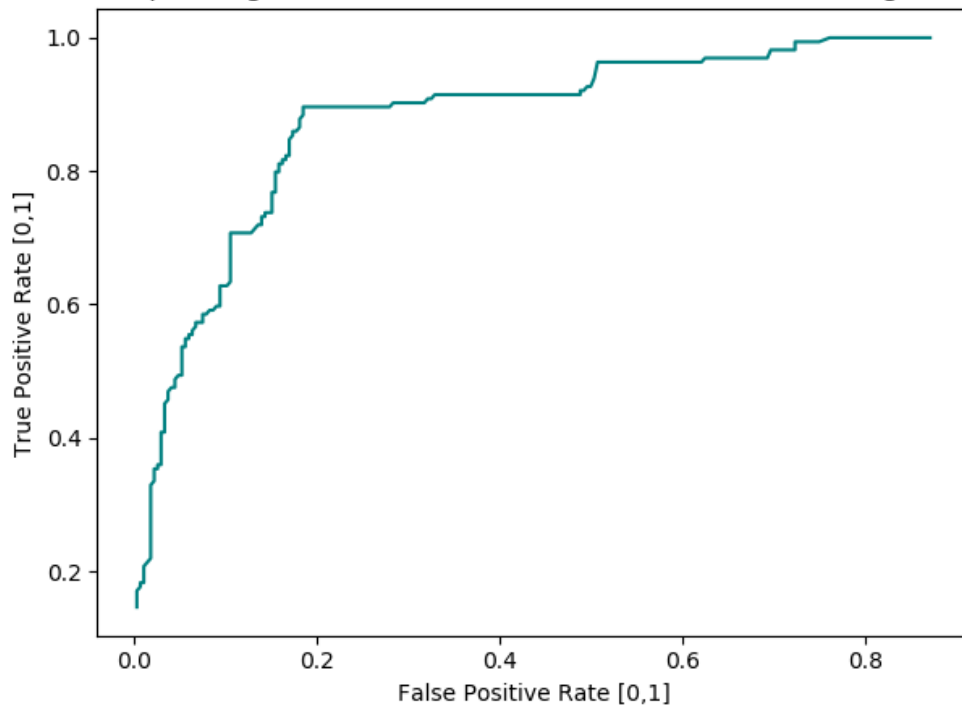
Οι μέγιστες τιμές που παρατηρήθηκαν είναι:

- Βαθμολογία F1: 81.67% για ($C = 2$, $\gamma = 0.25$, τιμή κατωφλίου: 0.142)
- Ορθότητα: 84.58% για ($C = 2$, $\gamma = 0.25$, τιμή κατωφλίου: 0.142)
- Ανάκληση: 100% για τιμή κατωφλίου 0
- Ακρίβεια Δηλώσεων: 100% για τιμή κατωφλίου 1

Καμπύλη ROC για ($C = 2$, $\gamma = 0.25$)

Για την καμπύλη ROC επιλέξαμε τις τιμές των υπερ-παραμέτρων C , γ για τις οποίες η βαθμολογία F1 παρουσίασε τη μέγιστη τιμή ($C = 2$, $\gamma = 0.25$). Παρατηρούμε ότι η καμπύλη ROC παρουσιάζει πολύ θετικά αποτελέσματα και το εμβαδόν κάτω από την καμπύλη ROC (Area Under Curve – AUC) έχει τιμή 75.85%.

Receiver Operating Characteristic Curve for SVR where $C=2$ and $\gamma=0.25$



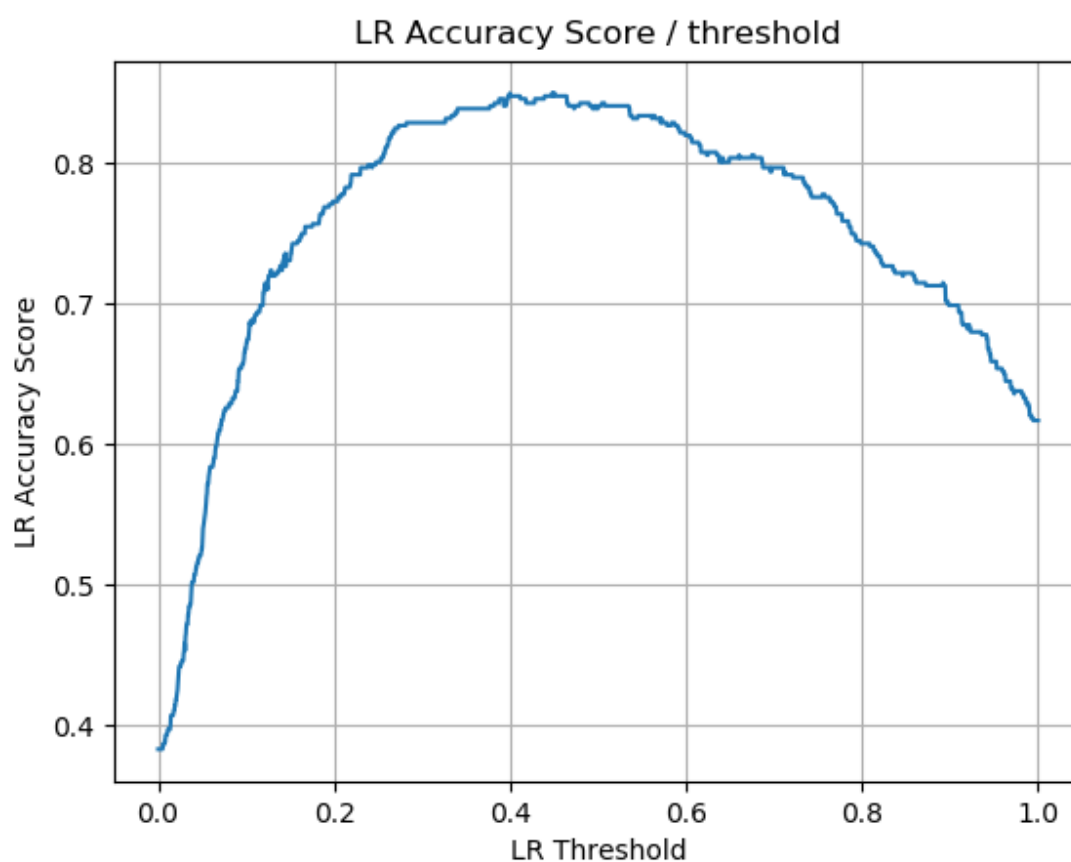
8-5 Η καμπύλη ROC του μοντέλου SVR

ii) Αποτελέσματα Μοντέλου Μηχανής Λογιστικής παλινδρόμησης

Για το Μοντέλο Μηχανής Λογιστικής παλινδρόμησης (LR), επιλέχθηκε το Μοντέλο με την τιμή C το οποίο παρουσίασε τα καλύτερα αποτελέσματα (8.3022). Η παραπάνω τιμή του C βρέθηκε μετά από βελτιστοποίηση με χρήση *Grid Search*. Παρακάτω παρουσιάζονται τα αποτελέσματα για τα μέτρα απόδοσης που περιγράφηκαν παραπάνω.

Ορθότητα Δηλώσεων (Accuracy)

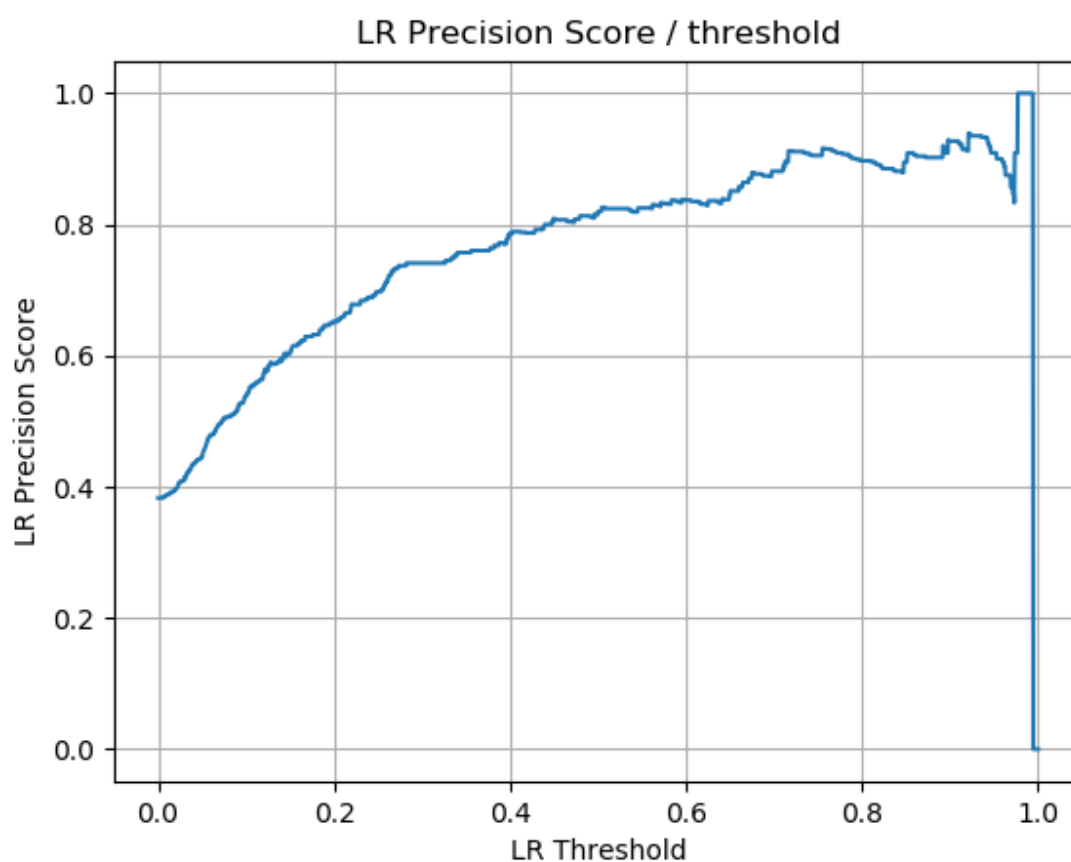
Το μοντέλο LR παρουσιάζει υψηλή βαθμολογία στην ακρίβεια δηλώσεων για τιμές κατωφλίου που δεν είναι ακραίες, δηλαδή βρίσκονται μακριά από το 0 και το 1. Για τιμές από 0.25 έως και 0.7 η ακρίβεια δηλώσεων του μοντέλου LR είναι πάνω από 80% με τη μέγιστη τιμή να είναι 85.05% για τιμή κατωφλίου ίση με 0.3997.



8-6 Το ποσοστό ορθότητας δηλώσεων του μοντέλου LR ανά τιμή κατωφλίου

Ακρίβεια Δηλώσεων (Precision)

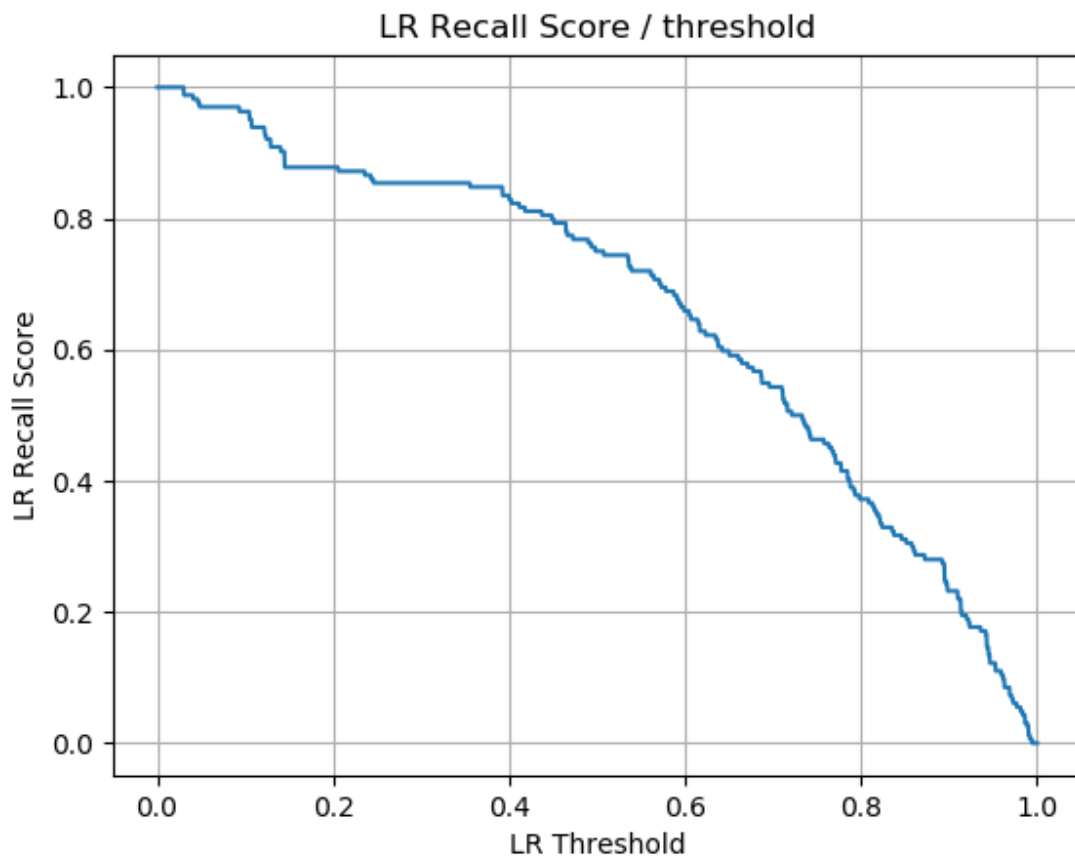
Το μοντέλο LR παρουσιάζει υψηλή βαθμολογία στην ακρίβεια δηλώσεων για τιμές κατωφλίου που βρίσκονται μακριά από το 0. Συγκεκριμένα, για τιμές από 0.42 έως και 1 η ακρίβεια δηλώσεων του μοντέλου LR είναι πάνω από 80% με τη μέγιστη τιμή να είναι 100% για πολλές τιμές κατωφλίου κοντά στο 1.



8-7 Το ποσοστό ακρίβειας δηλώσεων του μοντέλου LR ανά τιμή κατωφλίου

Ακρίβεια Ανάκλησης (Recall)

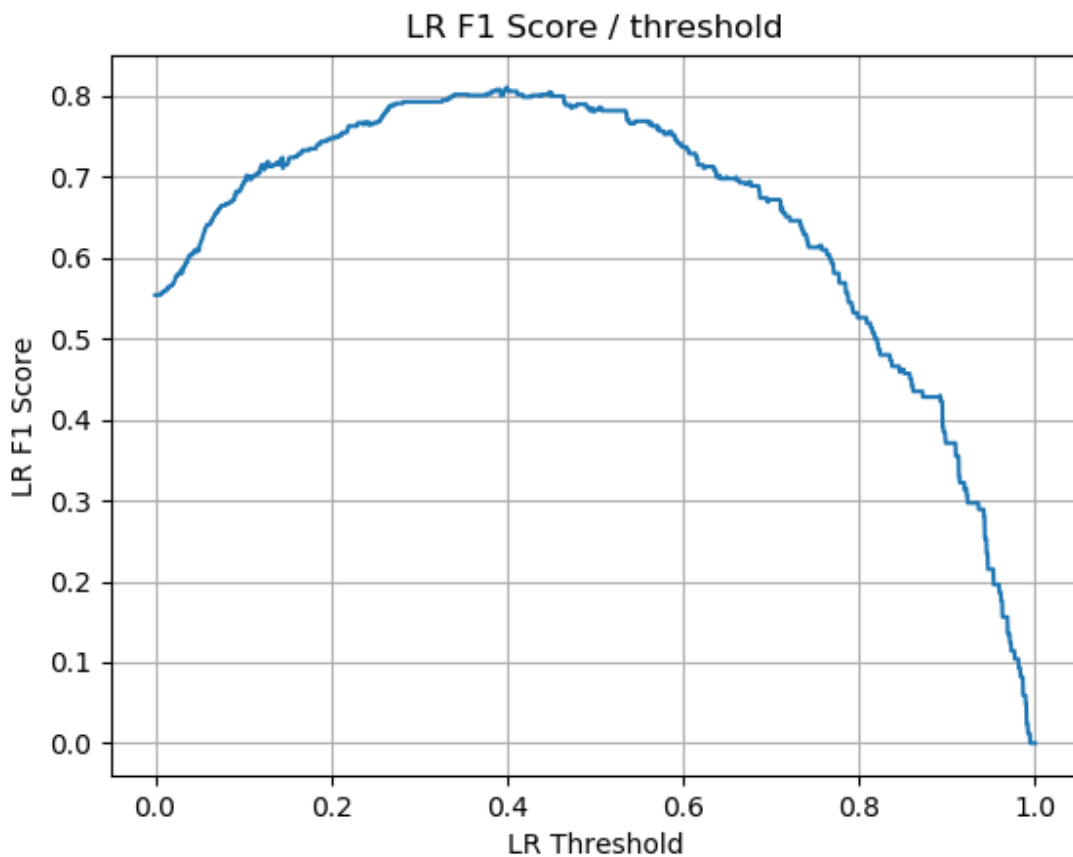
Το μοντέλο LR παρουσιάζει υψηλή βαθμολογία στην ακρίβεια ανάκλησης για τιμές κατωφλίου που δεν είναι κοντά στο 1. Όσο το κατώφλι πλησιάζει την τιμή 0, παρατηρούμε αύξηση στην ακρίβεια ανάκλησης. Η μέγιστη τιμή ακρίβειας ανάκλησης προκύπτει για πολλά κατώφλια κοντά στο 0 και είναι το ποσοστό 100%.



8-8 Το ποσοστό επιτυχούς ανάκλησης του μοντέλου LR ανά τιμή κατωφλίου

Ακρίβεια Βαθμολογίας F1 (F1 Score)

Το μοντέλο LR παρουσιάζει υψηλή βαθμολογία F1 για τιμές κατωφλίου που δεν είναι ακραίες, δηλαδή για τιμές κατωφλίου που δεν είναι κοντά στο 0 και στο 1. Όσο το κατώφλι πλησιάζει την τιμή 1, παρατηρούμε αισθητή μείωση στη βαθμολογία F1. Η μέγιστη τιμή ακρίβειας βαθμολογίας F1 προκύπτει για τιμή κατωφλίου 0.3997 και είναι ίση με 81.07%.



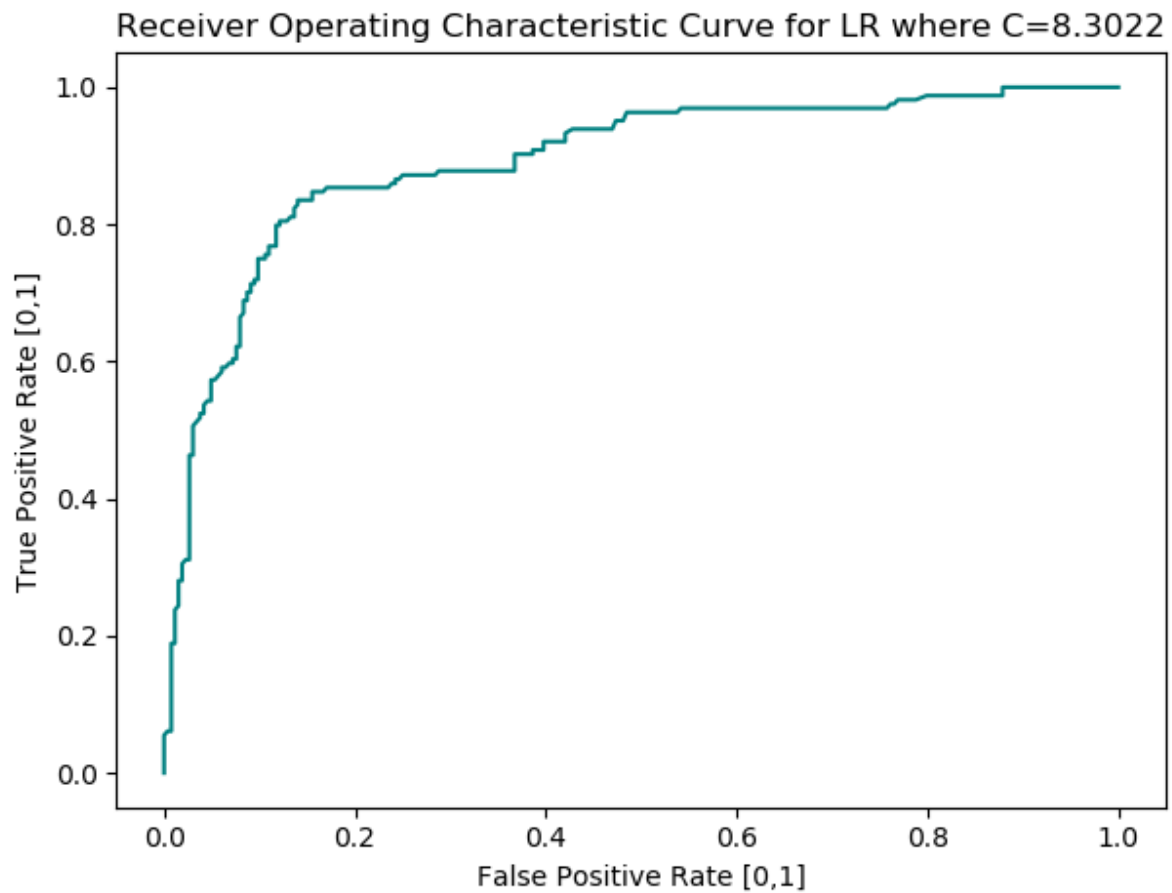
8-9 Η βαθμολογία F1 του μοντέλου LR ανά τιμή κατωφλίου

Οι μέγιστες τιμές που παρατηρήθηκαν είναι:

- Βαθμολογία F1: 81.07% για τιμή κατωφλίου: 0.3997
- Ορθότητα: 85.05% για τιμή κατωφλίου: 0.3997
- Ανάκληση: 100% για τιμή κατωφλίου 1
- Ακρίβεια Δηλώσεων: 100% για τιμή κατωφλίου 0

Καμπύλη ROC για C = 8.3022

Για την καμπύλη ROC του μοντέλου LR, παρατηρούμε ότι παρουσιάζει πολύ θετικά αποτελέσματα και το εμβαδόν κάτω από την καμπύλη ROC (Area Under Curve – AUC) έχει τιμή 89.13%.

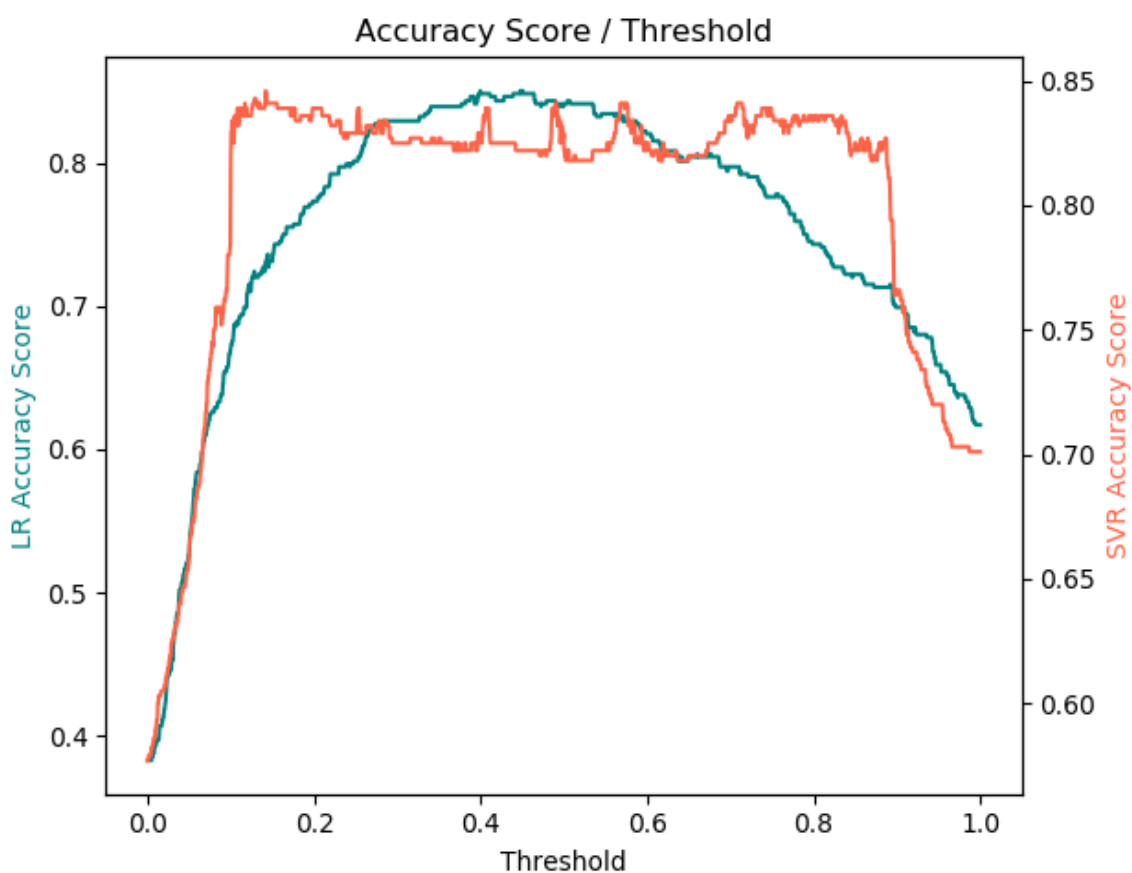


8-10 Η καμπύλη ROC του μοντέλου LR

iii) Σύγκριση αποτελεσμάτων των δύο μοντέλων

Ορθότητα Δηλώσεων (Accuracy)

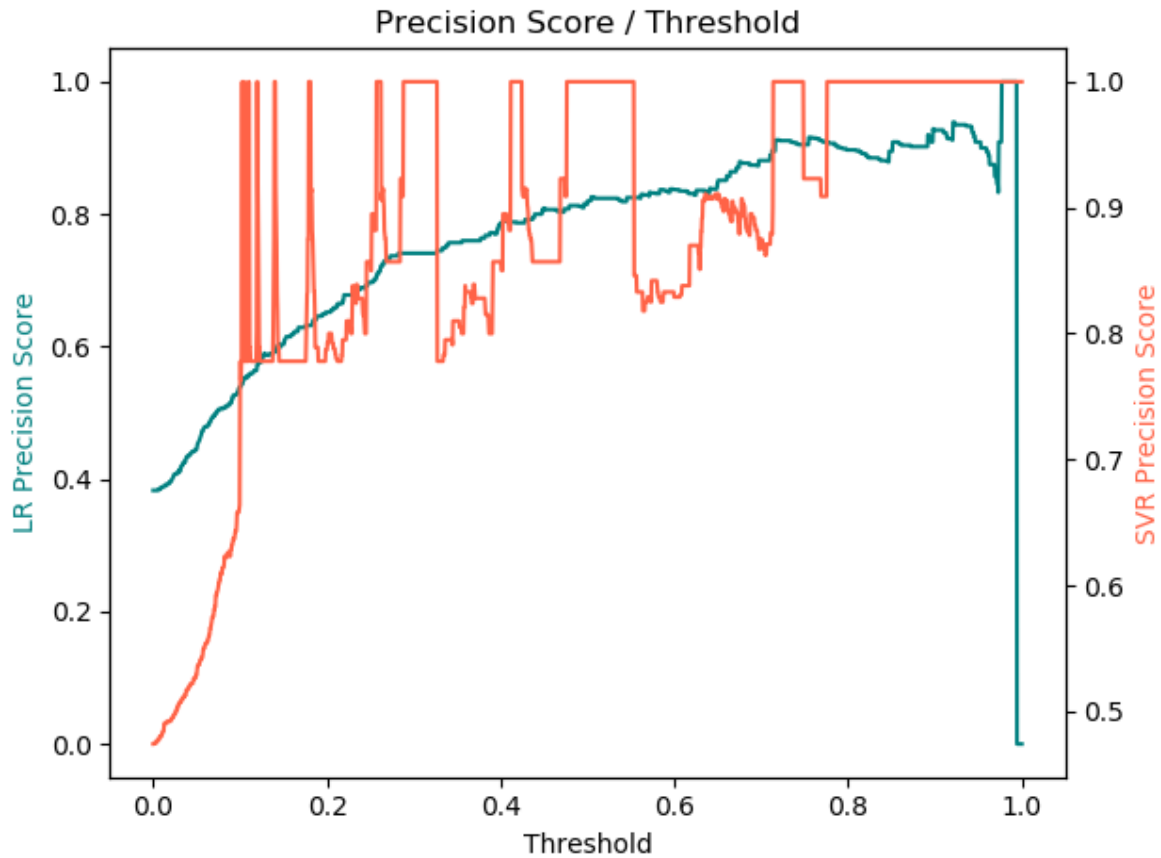
Εδώ παρατηρούμε ότι το μοντέλο SVR παρουσιάζει καλά αποτελέσματα σε ένα μεγαλύτερο εύρος τιμών κατωφλίου σε σχέση με το μοντέλο LR. Το μοντέλο LR παρουσιάζει καλύτερα αποτελέσματα για τιμές κατωφλίου μεταξύ 0.3 και 0.5 και μεταξύ 0.9 και 1, ενώ το μοντέλο SVR παρουσιάζει καλύτερα αποτελέσματα στα άλλα διαστήματα, με ιδιαίτερη διαφορά στα διαστήματα 0.1 έως 0.2 και 0.65 έως 0.85. Η μέγιστη τιμή της ορθότητας δηλώσεων του μοντέλου LR είναι ελάχιστα καλύτερη από αυτήν του μοντέλου SVR (85.05% έναντι 84.58%).



8-11 Σύγκριση ποσοστών ορθότητας δηλώσεων των μοντέλων SVR και LR ανά τιμή κατωφλίου

Ακρίβεια Δηλώσεων (Precision)

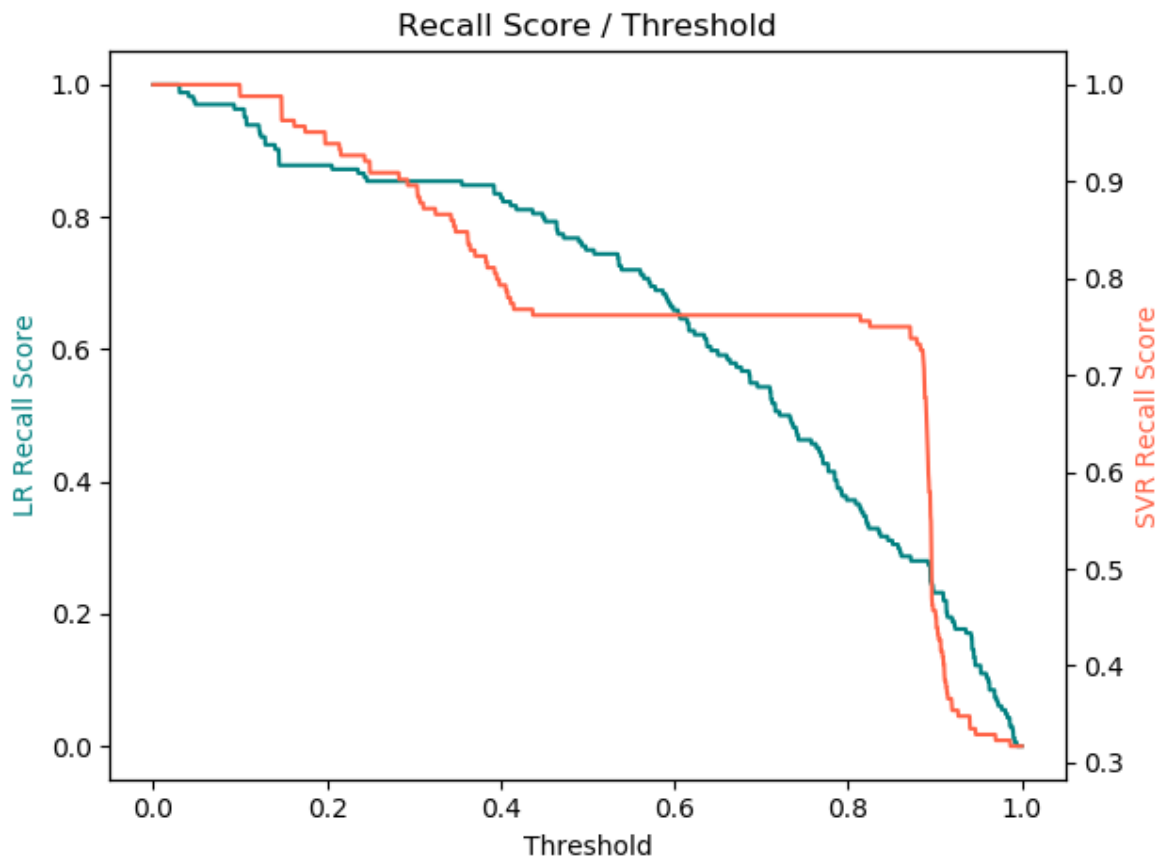
Εδώ παρατηρούμε ότι το μοντέλο LR παρουσιάζει καλύτερα αποτελέσματα όσο η τιμή κατωφλίου μεγαλώνει, ενώ το μοντέλο SVR λόγω της επιλογής διαφορετικού C , γ ανά τιμή κατωφλίου παρουσιάζει διακυμάνσεις στην ακρίβεια δηλώσεων.



8-12 Σύγκριση ποσοστών ακρίβειας δηλώσεων των μοντέλων SVR και LR ανά τιμή κατωφλίου

Ακρίβεια Ανάκλησης (Recall)

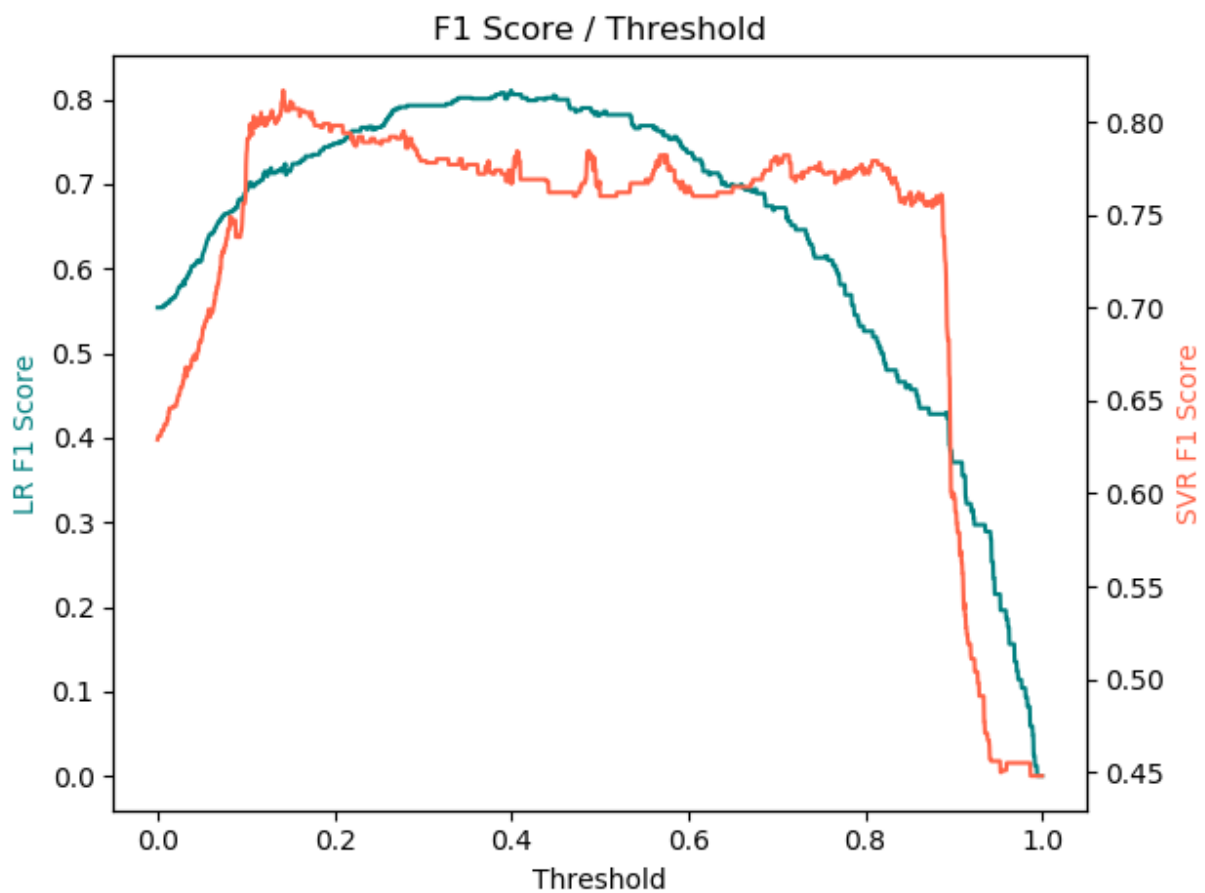
Εδώ παρατηρούμε ότι το μοντέλο LR παρουσιάζει καλύτερα αποτελέσματα σε σχέση με το μοντέλο SVR για τιμές κατωφλίου από 0.4 έως 0.6 και 0.9 έως 1, ενώ το μοντέλο SVR παρουσιάζει καλύτερα αποτελέσματα στις υπόλοιπες τιμές κατωφλίου, ιδιαίτερα στο διάστημα τιμών 0.6 έως 0.85.



8-13 Σύγκριση ποσοστών επιτυχούς ανάκλησης των μοντέλων SVR και LR ανά τιμή κατωφλίου

Ακρίβεια Βαθμολογίας F1 (F1-score)

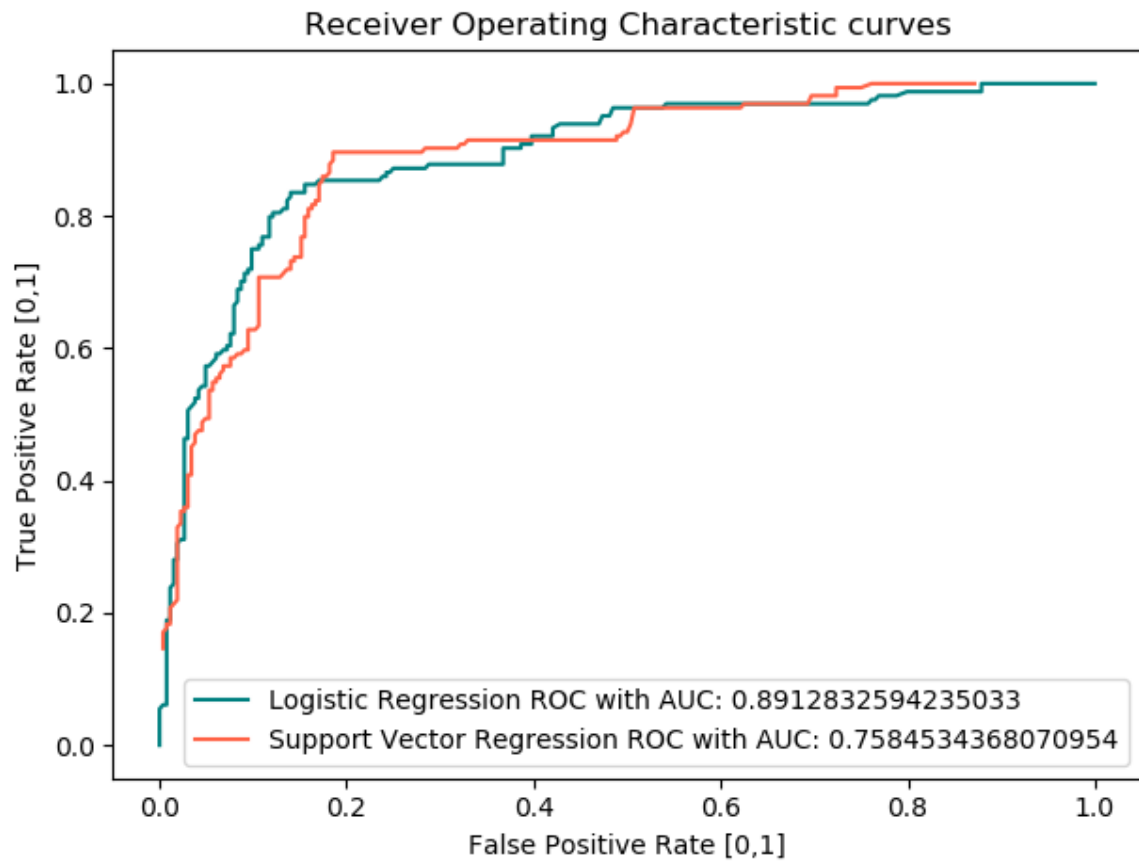
Εδώ παρατηρούμε ότι το μοντέλο LR παρουσιάζει καλύτερα αποτελέσματα σε σχέση με το μοντέλο SVR για τιμές κατωφλίου από 0.2 έως 0.6 και 0.9 έως 1, ενώ το μοντέλο SVR παρουσιάζει καλύτερα αποτελέσματα στις υπόλοιπες τιμές κατωφλίου, ιδιαίτερα στο διάστημα τιμών 0.65 έως 0.85. Η μέγιστη τιμή της βαθμολογίας F1 του μοντέλου SVR είναι ελάχιστα καλύτερη από αυτήν του μοντέλου LR (81.67% έναντι 81.07%).



8-14 Σύγκριση βαθμολογίας F1 των μοντέλων SVR και LR ανά τιμή κατωφλίου

Καμπύλη ROC

Η περιοχή κάτω από την καμπύλη ROC (AUC) είναι υψηλή και για τα δύο μοντέλα αλλά έχει καλύτερη τιμή για το μοντέλο γραμμικής παλινδρόμησης (89.13% έναντι 75.85%).



8-15 Σύγκριση καμπύλης ROC των μοντέλων SVR και LR ανά τιμή κατοφλίου

Λαμβάνοντας υπόψη την τεράστια διαφορά χρόνου που δαπανήθηκε για την εκπαίδευση των δύο μοντέλων (για το μοντέλο Μηχανής Διανυσμάτων Υποστήριξης χρειάστηκε τάξης μεγέθους περισσότερος χρόνος εκπαίδευσης από το μοντέλο Λογιστικής Παλινδρόμησης για τον ίδιο αριθμό δεδομένων εκπαίδευσης), θα δώσουμε μια μικρή υπεροχή στο μοντέλο Λογιστικής Παλινδρόμησης όσο αυξάνεται ο όγκος των δεδομένων εκπαίδευσης.

Ο σκοπός μας είναι η δημιουργία τριών διαφορετικών πρακτόρων οι οποίοι θα έχουν διαφορετικές τιμές ρίσκου στις επιλογές τους. Η τιμή ρίσκου προκύπτει από την τιμή κατωφλίου πάνω από την οποία ένας πράκτορας θα δηλώνει Grand Tichu. Για έναν πράκτορα που θέλουμε να είναι επιθετικός, δηλαδή να δηλώνει Grand Tichu για μικρές τιμές κατωφλίου και πάνω θα προτιμήσουμε το μοντέλο SVR που παρουσιάζει καλύτερα αποτελέσματα για χαμηλότερα κατώφλια. Για έναν πράκτορα που θέλουμε να είναι ουδέτερος, δηλαδή να δηλώνει Grand Tichu για τιμές κατωφλίου κοντά στο 0.5 θα προτιμήσουμε το μοντέλο LR, που παρουσιάζει καλύτερα αποτελέσματα για τα κατώφλια αυτά, ενώ για έναν συντηρητικό πράκτορα που δηλώνει Grand Tichu για μεγάλες τιμές κατωφλίου θα προτιμήσουμε το μοντέλο SVR που παρουσιάζει ελαφρώς καλύτερα αποτελέσματα για υψηλά κατώφλια.

Αν όμως θέλουμε να εκπαιδύσουμε ένα νέο μοντέλο με ελαφρώς διαφορετικά χαρακτηριστικά ή παραπάνω δεδομένα εκπαίδευσης, θεωρούμε ότι η διαφορά στα αποτελέσματα είναι πολύ μικρή για να δικαιολογήσει την τεράστια διαφορά στο χρόνο εκπαίδευσης μεταξύ των δύο μοντέλων και άρα θα επιλέγαμε το μοντέλο λογιστικής παλινδρόμησης.

Κεφάλαιο 9

Επίλογος και προτάσεις για μελλοντικές μελέτες

Στην παρούσα εργασία έγινε μια προσπάθεια χρήσης ευφών μοντέλων για τη λήψη αποφάσεων στο κομμάτι παιχνιδιού Tichu που αφορά τη δήλωση Grand Tichu. Η μοντελοποίηση των δεδομένων, η εξόρυξη χαρακτηριστικών από αυτά καθώς και η επιλογή και η εκπαίδευση των μοντέλων λήψης αποφάσεων αποτέλεσαν μεγάλες προκλήσεις. Το Tichu είναι ένα παιχνίδι για το οποίο είναι υπολογιστικά δύσκολο να δημιουργηθούν ευφείς πράκτορες και το οποίο δεν έχει ερευνηθεί σε βάθος από την επιστημονική κοινότητα. Η συνεργατικότητα μεταξύ των δύο παικτών που βρίσκονται στην ίδια ομάδα, η βαθμολογία καθώς και το πλήθος των επιτρεπτών συνδυασμών επιτρέπουν την επιλογή πλήθους διαφορετικών διανυσμάτων χαρακτηριστικών για την εκπαίδευση των μοντέλων, προσθέτοντας ταυτόχρονα μια επιπλέον δυσκολία στην επιλογή αυτών.

Η μεγαλύτερη δυσκολία που παρατηρήθηκε κατά τη διάρκεια της έρευνας μας, ήταν ο χρόνος που χρειάστηκε για την εκπαίδευση του μοντέλου μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση. Ο απαιτούμενος χρόνος ήταν τάξης μεγέθους παραπάνω από αυτόν που χρειάστηκε για το μοντέλο της λογιστικής παλινδρόμησης και απαγορευτικός, στα πλαίσια της εργασίας αυτής, για εκπαίδευση των μοντέλων με παραπάνω δεδομένα.

Τα αποτελέσματα τα οποία είχαν τα μοντέλα της λογιστικής παλινδρόμησης και της μηχανής διανυσμάτων υποστήριξης με παλινδρόμηση που ερευνήσαμε στην εργασία αυτή, μας άφησαν απολύτως ικανοποιημένους και ελπίζουμε να δώσουν ώθηση για μελλοντικές μελέτες πάνω στο παιχνίδι Tichu. Η δημιουργία τριών διαφορετικών πρακτόρων με βάση τη λήψη της πιθανότητας δήλωσης Grand Tichu από τα μοντέλα μας αποτελεί έναν ευφάνταστο τρόπο χρησιμοποίησης των μοντέλων αυτών, μιας και η έξοδος των μοντέλων αυτών είναι μια κατανομή πιθανότητας και όχι αποκλειστικά μια δυαδική τιμή.

Ανάμεσα στα δύο μοντέλα που εξετάσαμε στα πλαίσια της διατριβής, για έναν πράκτορα που θέλουμε να έχει πιο επιθετικό, αλλά και για έναν πράκτορα που θέλουμε να έχει πιο συντηρητικό τρόπο παιξίματος, προτείνουμε το μοντέλο μηχανής διανυσμάτων υποστήριξης.

Αντιθέτως για ένα παίκτη ουδέτερης επιθετικότητας προτείνουμε το μοντέλο λογιστικής παλινδρόμησης. Για εκπαίδευση νέων μοντέλων που αντλούν στοιχεία από τα όσα μελετήσαμε, προτείνουμε τη χρήση μοντέλων λογιστικής παλινδρόμησης, μιας και τα όποια μικρά πλεονεκτήματα παρουσιάζει σε διάφορες τιμές κατωφλίου το μοντέλο μηχανής διανυσμάτων δε δικαιολογούν την τεράστια διαφορά στο χρόνο εκπαίδευσης που απαιτείται σε σχέση με το μοντέλο λογιστικής παλινδρόμησης.

Στην αναζήτηση για παρατηρήσιμα παιχνίδια μερικής πληροφόρησης παρατηρήθηκε ότι ενώ υπάρχουν κάποιες έρευνες και προσπάθειες στον τομέα αυτό, υπάρχει μεγάλο περιθώριο για μελλοντικές μελέτες από ερευνητές. Έμφαση πρέπει να δοθεί στο πως αυτά τα παιχνίδια μπορούν να γίνουν υπολογιστικά επιλύσιμα, δηλαδή να μειωθεί η πολυπλοκότητα της μοντελοποίησης τους, είτε μέσω απλοποίησης των μοντέλων, είτε με κάποιον άλλον τρόπο.

Συγκεκριμένα, για το παιχνίδι Tichu πιθανές μορφές μελλοντικής έρευνας θα μπορούσαν να είναι:

- Η βελτίωση του μοντέλου μας λαμβάνοντας υπόψη εάν κάποιος αντίπαλος δήλωσε Grand Tichu στο ίδιο χέρι που καλείται το μοντέλο μας να αποφασίσει, εάν θα δηλώσει Grand Tichu ή όχι. Η δήλωση Grand Tichu από κάποιον αντίπαλο πρακτικά σημαίνει ότι ο αντίπαλος αυτός έχει με μεγάλη πιθανότητα ένα πολύ δυνατό χέρι και άρα θα πρέπει να δρα αρνητικά στη δήλωση Grand Tichu από τον πράκτορα μας.
- Η εύρεση της καλύτερης δυνατής συσχέτισης μεταξύ των 8 καρτών πριν τη δήλωση Grand Tichu με το να καταφέρει όντως ο παίκτης που δηλώνει Grand Tichu να βγαίνει πρώτος, δηλαδή η δήλωση να είναι επιτυχής. Πρέπει να γίνει μια μελέτη των χειρών οπού οδηγούν σε δήλωση Grand Tichu και αναλόγως με το εάν η δήλωση είναι επιτυχής ή όχι, να εφαρμοστεί μια ποινή ή επιβράβευση στο μοντέλο για την επιλογή της δήλωσης.
- Η δημιουργία ενός μοντέλου για το κομμάτι των ανταλλαγών καρτών μεταξύ των παικτών. Οι ανταλλαγές των καρτών είναι μια πολύ σύνθετη διαδικασία στην οποία οι παίκτες Tichu χρησιμοποιούν μνημονικούς κανόνες και αρκετό ή καθόλου ρίσκο ανάλογα με τον τύπο παίκτη. Έτσι θα μπορούσαμε και εδώ να έχουμε διαφορετικούς τύπους πρακτόρων ανάλογα με το ρίσκο που αυτοί θα παίρνουν.

- Η δημιουργία ενός μοντέλου για τη δήλωση ή όχι Tichu. Εδώ πλέον οι κάρτες για τις οποίες έχει πληροφορία το μοντέλο είναι οι 14 κάρτες του χεριού, συν οι 3 κάρτες των ανταλλαγών που ο πράκτορας έχει δώσει στους δύο αντιπάλους και στο συμπαίκτη. Εάν μάλιστα ο πράκτορας δεν παίζει πρώτος, τότε έχει και σαν επιπλέον πληροφορία τη στοίβα των καρτών που έχει παιχτεί, το ποιος παίκτης έχει ρίξει έναν συγκεκριμένο συνδυασμό φύλλων, το εάν ένας παίκτης έχει πει πάσο και άλλα.
- Η δημιουργία ευφών πρακτόρων για ολόκληρο το παιχνίδι Tichu.
- Καλύτερη ή διαφορετική εξαγωγή χαρακτηριστικών. Ιδανικά θα μπορούσαν να δημιουργηθούν πολλά μοντέλα με χρήση διαφορετικών χαρακτηριστικών και να επιλεγούν αυτά που παρουσιάζουν τα καλύτερα αποτελέσματα ή το βέλτιστο συνδυασμό υπολογιστικής πολυπλοκότητας και αποτελεσμάτων.

ΠΑΡΑΡΤΗΜΑ

Επιλογή χαρακτηριστικών με βάση τους επιτρεπτούς συνδυασμούς στο Tichu

18. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Τριάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Τριάρια} / 3$
19. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Τεσσάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Τεσσάρια} / 3$
20. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Πεντάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Πεντάρια} / 3$
21. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Εξάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Εξάρια} / 3$
22. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Εφτάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Εφτάρια} / 3$
23. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Οχτάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Οχτάρια} / 3$
24. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Εννιάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Εννιάρια} / 3$
25. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τα Δεκάρια: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Δεκάρια} / 3$
26. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τους Βαλέδες: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Βαλέδες} / 3$

27. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τις Ντάμες: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Ντάμες} / 3$
28. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τους Ρηγάδες: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Ρηγάδες} / 3$
29. Συνεχόμενα Ζευγάρια με μεγαλύτερο ζευγάρι τους Άσσους: με τιμή $1 \times \# \text{Συνεχομένων Ζευγαριών Μικρότερων από Άσσους} / 3$
30. Κέντα με μεγαλύτερο φύλλο το Πέντε: με τιμή $1 \times \# \text{Υποκεντών} / 4$
31. Κέντα με μεγαλύτερο φύλλο το Έξι: με τιμή $1 \times \# \text{Υποκεντών} / 4$
32. Κέντα με μεγαλύτερο φύλλο το Επτά: με τιμή $1 \times \# \text{Υποκεντών} / 4$
33. Κέντα με μεγαλύτερο φύλλο το Οκτώ: με τιμή $1 \times \# \text{Υποκεντών} / 4$
34. Κέντα με μεγαλύτερο φύλλο το Εννιά: με τιμή $1 \times \# \text{Υποκεντών} / 4$
35. Κέντα με μεγαλύτερο φύλλο το Δέκα: με τιμή $1 \times \# \text{Υποκεντών} / 4$
36. Κέντα με μεγαλύτερο φύλλο το Βαλέ: με τιμή $1 \times \# \text{Υποκεντών} / 4$
37. Κέντα με μεγαλύτερο φύλλο τη Ντάμα: με τιμή $1 \times \# \text{Υποκεντών} / 4$
38. Κέντα με μεγαλύτερο φύλλο το Ρήγα: με τιμή $1 \times \# \text{Υποκεντών} / 4$
39. Κέντα με μεγαλύτερο φύλλο τον Άσσο: με τιμή $1 \times \# \text{Υποκεντών} / 4$
40. Υπάρχει Full House Δυναμικότητας Δύο: με τιμή 0 ή 1.
41. Υπάρχει Full House Δυναμικότητας Τρία: με τιμή 0 ή 1.

42. Υπάρχει Full House Δυναμικότητας Τέσσερα: με τιμή 0 ή 1.
43. Υπάρχει Full House Δυναμικότητας Πέντε: με τιμή 0 ή 1.
44. Υπάρχει Full House Δυναμικότητας Έξι: με τιμή 0 ή 1.
45. Υπάρχει Full House Δυναμικότητας Επτά: με τιμή 0 ή 1.
46. Υπάρχει Full House Δυναμικότητας Οκτώ: με τιμή 0 ή 1.
47. Υπάρχει Full House Δυναμικότητας Εννιά: με τιμή 0 ή 1.
48. Υπάρχει Full House Δυναμικότητας Δέκα: με τιμή 0 ή 1.
49. Υπάρχει Full House Δυναμικότητας Βαλέ: με τιμή 0 ή 1.
50. Υπάρχει Full House Δυναμικότητας Ντάμας: με τιμή 0 ή 1.
51. Υπάρχει Full House Δυναμικότητας Ρήγα: με τιμή 0 ή 1.
52. Υπάρχει Full House Δυναμικότητας Άσσου: με τιμή 0 ή 1.
53. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Πέντε: με τιμή

$$1 \times \#Υποκεντών / 4$$
54. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Έξι: με τιμή

$$1 \times \#Υποκεντών / 4$$
55. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Επτά: με τιμή

$$1 \times \#Υποκεντών / 4$$

56. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Οκτώ: με τιμή

$$1 \times \#Υποκεντών / 4$$

57. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Εννιά: με τιμή

$$1 \times \#Υποκεντών / 4$$

58. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Δέκα: με τιμή

$$1 \times \#Υποκεντών / 4$$

59. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Βαλέ: με τιμή

$$1 \times \#Υποκεντών / 4$$

60. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο τη Ντάμα: με τιμή

$$1 \times \#Υποκεντών / 4$$

61. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο το Ρήγα: με τιμή

$$1 \times \#Υποκεντών / 4$$

62. Κέντα ίδιο χρώμα μπλε με μεγαλύτερο φύλλο τον Άσσο: με τιμή

$$1 \times \#Υποκεντών / 4$$

63. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Πέντε: με τιμή

$$1 \times \#Υποκεντών / 4$$

64. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Έξι: με τιμή

$$1 \times \#Υποκεντών / 4$$

65. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Επτά: με τιμή

$$1 \times \#Υποκεντών / 4$$

66. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Οκτώ: με τιμή

$$1 \times \#Υποκεντών / 4$$

67. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Εννιά: με τιμή

$$I \times \#Υποκεντών / 4$$

68. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Δέκα: με τιμή

$$I \times \#Υποκεντών / 4$$

69. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Βαλέ: με τιμή

$$I \times \#Υποκεντών / 4$$

70. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο τη Ντάμα: με τιμή

$$I \times \#Υποκεντών / 4$$

71. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο το Ρήγα: με τιμή

$$I \times \#Υποκεντών / 4$$

72. Κέντα ίδιο χρώμα κόκκινο με μεγαλύτερο φύλλο τον Άσσο: με τιμή

$$I \times \#Υποκεντών / 4$$

73. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Πέντε: με τιμή

$$I \times \#Υποκεντών / 4$$

74. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Έξι: με τιμή

$$I \times \#Υποκεντών / 4$$

75. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Επτά: με τιμή

$$I \times \#Υποκεντών / 4$$

76. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Οκτώ: με τιμή

$$I \times \#Υποκεντών / 4$$

77. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Εννιά: με τιμή

$$I \times \#Υποκεντών / 4$$

78. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Δέκα: με τιμή

$$I \times \#Υποκεντών / 4$$

79. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Βαλέ: με τιμή

$$I \times \#Υποκεντών / 4$$

80. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο τη Ντάμα: με τιμή

$$I \times \#Υποκεντών / 4$$

81. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο το Ρήγα: με τιμή

$$I \times \#Υποκεντών / 4$$

82. Κέντα ίδιο χρώμα πράσινο με μεγαλύτερο φύλλο τον Άσσο: με τιμή

$$I \times \#Υποκεντών / 4$$

83. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Πέντε: με τιμή

$$I \times \#Υποκεντών / 4$$

84. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Έξι: με τιμή

$$I \times \#Υποκεντών / 4$$

85. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Επτά: με τιμή

$$I \times \#Υποκεντών / 4$$

86. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Οκτώ: με τιμή

$$I \times \#Υποκεντών / 4$$

87. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Εννιά: με τιμή

$$I \times \#Υποκεντών / 4$$

88. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Δέκα: με τιμή

$$I \times \#Υποκεντών / 4$$

89. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Βαλέ: με τιμή

$$1 \times \#Υποκεντών / 4$$

90. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο τη Ντάμα: με τιμή

$$1 \times \#Υποκεντών / 4$$

91. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο το Ρήγα: με τιμή

$$1 \times \#Υποκεντών / 4$$

92. Κέντα ίδιο χρώμα μαύρο με μεγαλύτερο φύλλο τον Άσσο: με τιμή

$$1 \times \#Υποκεντών / 4$$

93. Συνολικός αριθμός βομβών στο χέρι: με τιμή από 0 έως 1

(κανονικοποιημένη)

Βιβλιογραφία

A Brief History of Computing. Ανάκτηση Δεκέμβριος, 2, 2019 από: alanturing.net

Achro. Ανάκτηση Νοέμβριος 4, 2019, από: <https://www.achro.gr/pdf/TICHU.pdf>

Annual Computer Poker Competition. Ανάκτηση Δεκέμβριος, 2, 2019 από:
<http://www.computerpokercompetition.org/>

Asimakopoulos, C. (2019, Δεκέμβριος 1). *GitHub Tichu Sensei*. Ανάκτηση Δεκέμβριος 2, 2019, από GitHub: <https://github.com/ch-asimakopoulos/Tichu-Sensei>

Board Game Geek. (2011, Ιούνιος 29). Ανάκτηση Νοέμβριος 6, 2019, από
<https://boardgamegeek.com/blogpost/3269/ios-review-mu-and-tichu>

Boser, B. G. (1992). A training algorithm for optimal margin classifiers. *In Proceedings of the 5th Annual Workshop on Computational Learning Theory*. ACM Press , σσ. 144-152.

BrettSpielWelt. Ανάκτηση Νοέμβριος 9, 2019, από: <https://www.brettspielwelt.de/>

Castelein, K. (2017, Ιούλιος 20). Determinization for Monte Carlo Tree Search in the Card Game Tichu.

Chih-Wei Hsu, C.-C. C.-J. (2003). *A Practical Guide to Support Vector Classification*. Ανάκτηση Νοέμβριος 19, 2019, από Department of Computer Science, National Taiwan University, Taipei 106, Taiwan: <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

Cortes C., V. V. (1995). Support-vector network. *Mach. Learn.* 20 , σσ. 273–297.

da Silva, W. G. (2018, Ιούνιος). HearthBot: An Autonomous Agent Based on Fuzzy ART Adaptive Neural Networks for the Digital Collectible Card Game HearthStone. *IEEE Transactions on Games, Vol. 10 Issue: 2* , σσ. 170-181.

dod. Ανάκτηση από: <https://dod.gr>

Fata Morgana. (2013). Ανάκτηση Νοέμβριος 3, 2019, από:
https://www.fatamorgana.ch/images/fatamorgana/tichu_box.jpg

Fujii, H. K. (2008, Δεκέμβριος 03). Strategy-acquisition system for video trading card game. *ACE '08 Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* , σσ. 175-182.

Grand, E. F. (1952, Αύγουστος 2). "The Talk of the Town - It". *The New Yorker*.

Holdren, C. (n.d.). *Tichu*. Ανάκτηση Νοέμβριος 6, 2019, από <https://play.google.com/store/apps/details?id=com.mystcmage.android.tichu>

Hostettler, U. (2004, Δεκέμβριος 12). *Tichu Fata Morgana*. Ανάκτηση Νοέμβριος 3, 2019, από Fata Morgana: <https://urs.fatamorgana.ch/tichu.html>

House Full of Games. (n.d.). Ανάκτηση Νοέμβριος 6, 2019, από <https://itunes.apple.com/us/app/tichu/id386800660?mt=8>

Hyeong Soo Chang, M. C. (2015). An adaptive sampling algorithm for solving Markov decision processes. *Operations Research*, 53(1) , σσ. 126-139.

Ishii, F. M. (2005, Μάϊος). A Reinforcement Learning Scheme for a Partially-Observable Multi-Agent Game. *Machine Learning, May 2005, Volume 59, Issue 1-2* , σσ. 31-54.

J. K. Anlauf, M. B. (1989, Δεκέμβριος). The AdaTron: An Adaptive Perceptron Algorithm. *EPL (Europhysics Letters)*, Volume 10, Number 7 .

Jesta. (2015, Σεπτέμβριος 7). Ανάκτηση Νοέμβριος 4, 2019, από: <https://www.jestatharogue.com/game-reviews/review-tichu/>

Lagoudakis, M. G. (2017). *Value Function Approximation, Encyclopedia of Machine Learning and Data Mining*.

Lim, A. (2014, Ιούλιος). Ανάκτηση Νοέμβριος 5, 2019, από <http://www.andreelim.com/2014/07/tichu.html>

Mark E. Glickman, A. C. (1996, Ιανουάριος 28). Rating the Chess Rating System. *American Chess Journal*, 3 , σσ. 59-102.

Martha Vlachou-Konchylaki, S. V. (2013, Μάιος 14). Combining Deliberation and Reactive Behavior for AI Players in the Mini-Tichu Card-game. *Foundation of Digital Games, 2013, Chania, Crete, Greece* , σ. 2.

McCorduck, Pamela (2004). *Machines Who Think* (2nd ed.) ISBN 1-56881-205-1 A. K. Peters, Ltd., pp. 480-483

Nate Silver, R. F.-B. (2015, Μάιος 21). *FiveThirtyEight*. Ανάκτηση Νοέμβριος 11, 2019, από: How we calculate the Elo rankings: <https://fivethirtyeight.com/features/how-we-calculate-nba-elo-ratings/>

(2003). Στο S. a. Norvig, *Artificial Intelligence: A Modern Approach, second edition* (σ. 733). Prentice Hall.

Online Tichu. Ανάκτηση Νοέμβριος 7, 2019, από: <https://onlinetichu.com>

OpenAI. (2019, Δεκέμβριος 1). *Dota 2 with Large Scale Deep Reinforcement Learning* Ανάκτηση Δεκέμβριος 2, 2019, από GitHub: <https://cdn.openai.com/dota-2.pdf>

Peter Cowling, C. W. (2012, Δεκέμβριος). Ensemble Determinization in Monte Carlo Tree Search for the Imperfect Information Card Game Magic: The Gathering. *IEEE Transactions on computational intelligence and AI in games, Vol. 4, No. 4, December 2012* , σσ. 241-257.

Reid, P. (2015, Μάρτιος 16). *Paul Hogan Reid, Wordpress*. Ανάκτηση Νοέμβριος 8, 2019, από: <https://paulhoganreid.wordpress.com/tag/tichu/>

René Schwarz, T. P. (2013, Μάιος 2). Developing an Android app for the card game Tichu. σσ. 1-11.

Richard S. Sutton, A. G. (2015). *Reinforcement Learning: An Introduction, Second edition*. Cambridge, Massachusetts: The MIT Press.

Robiliard, F. T. (2014). Monte-Carlo Tree Search for the Game of "7-Wonders". *Computer Games. CGW 2014. Communications in Computer and Information Science, vol 504. Springer* .

SciKit Learn Logistic Regression. Ανάκτηση Νοέμβριος 19, 2019, από: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

SciKit Learn SVM. Ανάκτηση Νοέμβριος 18, 2019, από SciKit Learn SVM SVR: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html>

Shalizi, C. (2013, Φεβρουάριος 27). *Undergraduate Advanced Data Analysis*. Ανάκτηση Νοέμβριος 22, 2019, από Carnegie Mellon University: <http://www.stat.cmu.edu/~cshalizi/uADA/13/lectures/ch12.pdf>

Silver, D. (2015, Μαΐου 13). *Computer Science, UCL*. Retrieved Νοέμβριος 1, 2019, από: http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching_files/MDP.pdf

Staelin, C. (2003, Νοέμβριος 10). Parameter selection for support vector machines. *HP Laboratories Israel HPL-2002-354 (R.1)*.

Steinacher, M. (n.d.). *Steinacher*. Ανάκτηση από: <https://steinacher.info/>

T. Poggio, F. G. (1990, Φεβρουαρίου 23). Regularization Algorithms for Learning that are Equivalent to Multilayer Networks. *science, New Series, Vol. 247, No. 4945.*, σσ. 978-982.

Tabletopia. (2019, Μάϊος 25). *Tabletopia*. Ανάκτηση Νοέμβριος 5, 2019, από: <https://tabletopia.com/games/tichu>

Tania Kleynhans, M. M. (2017, Μάϊος 3). Predicting Top-of-Atmosphere Thermal Radiance Using MERRA-2 Atmospheric Data with Deep Learning. *Proceedings of the SPIE Defense and Commercial Sensing, Anaheim, CA, USA, Volume 10178*.

Team, ΤοΒΗΜΑ (2019, Μάϊος 31). Η τεχνητή νοημοσύνη νικάει πια και παιχνίδια που απαιτούν συνεργασία. Ανάκτηση Δεκέμβριος 1, 2019 από: <https://www.tovima.gr/2019/05/31/science/i-texniti-noimosyni-nikaei-pia-kai-paixnidia-pou-apaitoun-synergasia/>

Tichu Online. Ανάκτηση από: <https://www.onlinetichu.com>

Tichu Pro. Ανάκτηση από: <https://www.tichu.pro/el/>

Tichulog BrettSpielWelt. Ανάκτηση Ιούλιος 22, 2017, από <http://tichulog.brettspielwelt.de/>

Tichumania. Ανάκτηση Νοέμβριος 5, 2019, από: <http://log.tichumania.de>

Tichurules Blogspot. (2012, Αύγουστος). Ανάκτηση Νοέμβριος 4, 2019, από: <http://tichurules.blogspot.com/>

Torres, A. (2016, Απρίλης 1). *Tichu Perceptron*. Ανάκτηση Νοέμβριος 5, 2019, από: <https://github.com/agustintorres/tichu-perceptron>

Vapnik V., C. A. (1979). *Theory of Pattern Recognition [in Russian]*. Μόσχα.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Μόσχα: Springer.

Wikipedia. (2019, Οκτωβρίου 17). Ανάκτηση Νοεμβρίου 1, 2019, από: https://en.wikipedia.org/wiki/Partially_observable_Markov_decision_process

Wikipedia. (2019, Νοέμβριος 13). Ανάκτηση Νοέμβριος 15, 2019, από Elo rating system: https://en.wikipedia.org/wiki/Elo_rating_system

Wikipedia. (2018, Μάρτιος 2). Ανάκτηση Νοέμβριος 3, 2019, από: <https://en.wikipedia.org/wiki/Tichu>

Wikipedia. (2019, Νοέμβριος 3). Ανάκτηση Δεκέμβριος 2, 2019, από: https://en.wikipedia.org/wiki/Computer_bridge#Computers_versus_humans

Yang, E. (2018, Δεκέμβριος 14). *Betting strategy for the card game Tichu*. Ανάκτηση Νοέμβριος 8, 2019, από: <https://pdfs.semanticscholar.org/aab0/c47f162d1a04377097baec1e6595f0e6bde6.pdf?>

Zheng, T. (2019, Σεπτέμβριος 1). <https://tianxiaozheng.weebly.com/>. Ανάκτηση Νοέμβριος 1, 2019, από: <https://tianxiaozheng.weebly.com/uploads/4/6/5/4/46548039/2-markov-decision-process.pdf>

Βένου, Θ. Μ. (2016, Φεβρουάριος 16). Ανάκτηση Νοέμβριος 4, 2019, από: <https://www.pillowfights.gr/anatomy-of-a-pillow/giati-oloi-trelainomaste-gia-tichu/>

Βλάχου-Κογχυλάκη, Μ. (2012, Οκτώβριος). Ανάπτυξη παιχνιδιού στρατηγικής μερικής πληροφόρησης.

Γιαννούλη, Δ. Π. (2014, Ιούλιος). Εφαρμογές των Μηχανών Διανυσματικής Υποστήριξης σε Προβλήματα Ταξινόμησης και Παλινδρόμησης. Διπλωματική Εργασία, Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών επιστημών, Εθνικό Μετσόβιο Πολυτεχνείο.

Κυρίτσης, Κ. (2014, Μάϊος). *nemertes.lis.upatras.gr*. Ανάκτηση Νοέμβριος 16, 2019, από Νευρωνικά Δίκτυα και Μηχανές Διανυσματικής Υποστήριξης:
<https://nemertes.lis.upatras.gr/jspui/bitstream/10889/8010/1/Master%20Thesis%20Kyritsis%20Konstantinos.pdf>

Παπαδάκης, Ε. (2016, Ιούνιος). *artemis.cslab.ece.ntua.gr*. Ανάκτηση Νοέμβριος 12, 2019, από Ανάλυση Συναισθήματος από Κείμενο με Τεχνικές Μηχανικής:
<http://artemis.cslab.ece.ntua.gr:8080/jspui/bitstream/123456789/13134/1/DT2016-0114.pdf>

Τεχνητή νοημοσύνη της DeepMind έφτασε στο ανώτατο επίπεδο παικτών στο Starcraft 2, δημοφιλές παιχνίδι των esports, (2019, 31, Οκτώβριος). Ανάκτηση Δεκέμβριος 1, 2019, από Naftemporiki: <https://www.naftemporiki.gr/story/1528619/texniti-noimosuni-tis-deepmind-eftase-sto-anotato-epipedo-paikton-sto-starcraft-2-dimofiles-paixnidi-ton-esports>