

**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ**

ΔΗΜΗΤΡΙΟΣ Φ. ΠΑΠΑΓΕΩΡΓΙΟΥ

Διπλωματική Εργασία

**ΔΙΑΔΙΚΑΣΙΕΣ ΔΙΑΚΡΙΤΟΠΟΙΗΣΗΣ
ΠΟΣΟΤΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΓΙΑ ΤΗΝ ΑΝΑΠΤΥΞΗ
ΣΥΣΤΗΜΑΤΩΝ ΕΚΤΙΜΗΣΗΣ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ**



ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: ΜΙΧΑΛΗΣ ΔΟΥΜΠΟΣ

ΧΑΝΙΑ 2004

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέπων καθηγητή κ. Μιχάλη Δούμπο για τον πολύτιμο χρόνο που αφιέρωσε, τη συνεχή συνεργασία, την καθοδήγηση και την πολύτιμη βοήθεια που μου προσέφερε καθ' όλη τη διάρκεια της εκπόνησης της παρούσας διπλωματικής εργασίας.

Τέλος θα ήθελα να ευχαριστήσω θερμά, όλους τους φίλους, τους συμφοιτητές, τους καθηγητές μου καθώς και την οικογένειά μου για την πολύτιμη βοήθεια και την ηθική υποστήριξη που μου προσέφεραν καθ' όλη τη διάρκεια της φοίτησής μου στο τμήμα Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης.

Περιεχόμενα

Εισαγωγή	2
Κεφάλαιο 1: Συστήματα εκτίμησης πιστωτικού κινδύνου	4
1.1 Ο χώρος του προβλήματος.....	4
1.2 Τι είναι ένα σύστημα εκτίμησης πιστωτικού κινδύνου	5
1.3 Βασικά χαρακτηριστικά ενός συστήματος εκτίμησης πιστωτικού κινδύνου	6
1.3.1 Κατηγορίες διαχωρισμού του πιστωτικού κινδύνου.....	7
1.3.2 Καθορισμός και σύνθεση παραγόντων που επηρεάζουν τον πιστωτικό	10
1.3.3 Τεχνικές ταξινόμησης	13
1.3.3.1 Στατιστικές και οικονομετρικές τεχνικές.....	13
1.3.3.2 Μη παραμετρικές τεχνικές.....	15
1.4 Χρησιμότητα των συστημάτων εκτίμησης πιστωτικού κινδύνου	23
1.4.1 Στοιχεία για τις πιστωτικές εξελίξεις στην Ελλάδα και τη συναλλακτική συμπεριφορά των ελληνικών επιχειρήσεων	24
1.5 Διαδικασία ανάπτυξης ενός συστήματος εκτίμησης πιστωτικού κινδύνου.....	28
1.6 Απαιτήσεις – κριτήρια ελέγχου συστημάτων εκτίμησης πιστωτικού κινδύνου	30
Κεφάλαιο 2: Ανάπτυξη πειράματος και αποτελέσματα	32
2.1 Εισαγωγή	32
2.2 Ανάπτυξη πειράματος.....	33
2.3 Εξεταζόμενες διαδικασίες διακριτοποίησης.....	35
2.3.1 Η διαδικασία των ποσοστιμώριων (percentiles cut-points)	35
2.3.2 Η απόσταση Kolmogorov-Smirnoff	35
2.3.3 Μέθοδος CART. Δένδρα ταξινόμησης και παλινδρόμησης.....	38
2.4 Λογιστική παλινδρόμηση.....	42
2.5 Ανάλυση αποτελεσμάτων	43
2.6 Αξιολόγηση των διαδικασιών	53
2.6.1 Τύποι σφαλμάτων	53
2.6.2 Καμπύλη R.O.C. (Receiver Operating Characteristics Graph).....	55
Κεφάλαιο 3: Συμπεράσματα	61
Παράρτημα	64
Βιβλιογραφία	69

ΕΙΣΑΓΩΓΗ

Η ανάλυση και διαχείριση των χρηματοοικονομικών κινδύνων είναι σήμερα ένα από τα σημαντικότερα πεδία έρευνας στο χώρο της χρηματοοικονομικής επιστήμης. Η ανάπτυξη τεχνικών για την εκτίμηση των κινδύνων αυτών έχει εξελιχθεί εντυπωσιακά κατά τη διάρκεια των τελευταίων ετών εξαιτίας κάποιων παγκόσμιων οικονομικών αλλαγών, οι οποίες έχουν μεταβάλλει δραστικά τις συνθήκες μέσα στις οποίες λειτουργούν οι επιχειρήσεις και οι οργανισμοί. Η παγκοσμιοποίηση της οικονομίας και των χρηματοοικονομικών αγορών, ο αυξανόμενος ανταγωνισμός μεταξύ των επιχειρήσεων, οι ραγδαίες τεχνολογικές και κοινωνικό-πολιτικές εξελίξεις, η αύξηση των πτωχεύσεων (επιχειρήσεων ή τραπεζών), η αυξανόμενη ζήτηση για δανεισμό κεφαλαίων κ.ά. έχουν οδηγήσει στην αυξημένη πολυπλοκότητα του επιχειρηματικού και οικονομικού περιβάλλοντος.

Ως παραδείγματα χρηματοοικονομικών κινδύνων μπορούν να αναφερθούν:

- ο κίνδυνος πτώχευσης μιας εταιρίας ή ενός πιστωτικού ιδρύματος.
- ο πιστωτικός κίνδυνος που εμπεριέχεται στο δανεισμό κεφαλαίου από μια τράπεζα σε μια εταιρία ή σε ιδιώτη.
- ο κίνδυνος μιας επένδυσης σε χρηματιστηριακούς τίτλους, αμοιβαία κεφάλαια, παράγωγα, ομόλογα κ.ά.
- ο κίνδυνος κρατών (δηλαδή η αξιολόγηση των οικονομικών τους επιδόσεων και της δανειοληπτικής τους ικανότητας).

Η πολυπλοκότητα λοιπόν αυτή, που διέπει την λήψη χρηματοοικονομικών αποφάσεων, καθιστά επιτακτική την ανάγκη οι χρηματοπιστωτικοί οργανισμοί και οι επιχειρήσεις να προχωρήσουν στο σχεδιασμό και στην εφαρμογή συγκεκριμένων μεθοδολογιών ανάλυσης και διαχείρισης κάθε μορφής χρηματοοικονομικού κινδύνου που εμπεριέχεται στις δραστηριότητές τους.

Συγκεκριμένα, η λήψη χρηματοοικονομικών αποφάσεων σχετικών με τον πιστωτικό κίνδυνο είναι από τα πιο πολύπλοκα και κακώς δομημένα προβλήματα. Λόγω της πολυδιάστατης φύσης του πιστωτικού κινδύνου και της ύπαρξης πολλαπλών παραγόντων που τον επηρεάζουν, οι εμπειρικές προσεγγίσεις του είναι πλέον ανεπαρκείς και η αναζήτηση βέλτιστων λύσεων, όσον αφορά την εκτίμησή του, είναι ανέφικτη.

Έτσι λοιπόν, σε παγκόσμιο επίπεδο, οι ερευνητές έχουν στρέψει την προσοχή τους στην ανάπτυξη συστημάτων εκτίμησης του πιστωτικού κινδύνου που εξάγουν ικανοποιητικές λύσεις και καθιστούν ικανά τα αρμόδια στελέχη των χρηματοπιστωτικών ιδρυμάτων να διαχειρίζονται ορθολογικά τον πιστωτικό κίνδυνο. Οι έρευνες αυτές βασίζονται στους επιστημονικούς κλάδους των εφαρμοσμένων μαθηματικών, της οικονομετρίας, της στατιστικής, της επιχειρησιακής έρευνας, της

τεχνητής νοημοσύνης και της πληροφορικής (σε συνδυασμό πάντα με τις θεωρητικές εξελίξεις του χώρου).

Η συγκεκριμένη μελέτη πραγματεύεται θέματα που αφορούν εξολοκλήρου τον πιστωτικό κίνδυνο των επιχειρήσεων και τα συστήματα εκτίμησής του. Ως πιστωτικός κίνδυνος ορίζεται η αδυναμία μιας επιχείρησης να εκπληρώσει τις υποχρεώσεις που της δημιουργεί ο δανεισμός (default risk). Ένα σύστημα εκτίμησης πιστωτικού κινδύνου είναι λοιπόν, σε γενικές γραμμές, ένα εργαλείο που βοηθάει τις τράπεζες να εκτιμήσουν τον πιστωτικό κίνδυνο που διέπει μια επικείμενη δανειοδότηση και να αποφανθούν αν η υπό εξέταση επιχείρηση, που έχει κάνει αίτηση για δανεισμό, πληρεί τις προϋποθέσεις ώστε να χρηματοδοτηθεί.

Πιο συγκεκριμένα, η μελέτη αυτή επικεντρώνεται στην ανάπτυξη και εφαρμογή κάποιων διαδικασιών προετοιμασίας των δεδομένων που δέχεται ένα σύστημα εκτίμησης πιστωτικού κινδύνου. Τα δεδομένα αυτά είναι κάποια χαρακτηριστικά της επιχείρησης και βάσει αυτών, έπειτα από την επεξεργασία τους, το σύστημα εξάγει την εκτίμηση του πιστωτικού κινδύνου που την χαρακτηρίζει. Σκοπός της μελέτης είναι η σύγκριση των διαδικασιών αυτών και η εξαγωγή χρήσιμων συμπερασμάτων όσον αφορά το πια μέθοδο προετοιμασίας των δεδομένων είναι πιο ακριβής. Δηλαδή, ποία τεχνική αν χρησιμοποιηθεί από ένα σύστημα εκτίμησης πιστωτικού κινδύνου θα εξάγει την ορθότερη πρόβλεψη.

Η εργασία αυτή χωρίζεται σε τρία κεφάλαια. Στο πρώτο κεφάλαιο πραγματοποιείται μια εκτεταμένη εισαγωγή στο χώρο των συστημάτων εκτίμησης πιστωτικού κινδύνου. Αρχικά αναλύονται κάποια από τα βασικά χαρακτηριστικά τους και γίνεται αναφορά στη χρησιμότητά τους. Έπειτα παρουσιάζονται κάποιες από τις πιο γνωστές τεχνικές που χρησιμοποιούνται για την κατασκευή τους και αναλύεται η διαδικασία ανάπτυξής τους.

Στο δεύτερο κεφάλαιο παρουσιάζεται το κύριο μέρος της μελέτης αυτής. Γίνεται μια εκτενής ανάλυση του πειράματος που πραγματοποιήθηκε. Δηλαδή παρουσιάζονται οι διαδικασίες προετοιμασίας των δεδομένων που αναπτύχθηκαν και εφαρμόστηκαν, καθώς και η διαδικασία που ακολουθήθηκε ώστε να γίνει εφικτή η σύγκριση των διαδικασιών αυτών. Στο τέλος του κεφαλαίου αυτού παρουσιάζονται τα αποτελέσματα της ανάλυσης.

Τέλος, στο τρίτο κεφάλαιο παρατίθενται χρήσιμα συμπεράσματα που προέκυψαν έπειτα από την ανάλυση των αποτελεσμάτων. Προτείνονται οι πιο ακριβείς, ως προς την αποτελεσματικότητά τους, από τις διαδικασίες που εφαρμόστηκαν και αναφέρονται κάποιες προοπτικές για την ανάπτυξη τέτοιων διαδικασιών που θα συμβάλλουν στην πιο ασφαλή αντιμετώπιση του προβλήματος της εκτίμησης του πιστωτικού κινδύνου.

ΚΕΦΑΛΑΙΟ 1

ΣΥΣΤΗΜΑΤΑ ΕΚΤΙΜΗΣΗΣ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

1.1 Ο χώρος του προβλήματος

Στον ευρύτερο οικονομικό χώρο στις μέρες μας, κάθε φυσικό πρόσωπο, επιχείρηση ή οργανισμός που συνάπτει οποιαδήποτε σχέση (ως επενδυτής, χρηματοδότης ή μέτοχος) με μια άλλη επιχείρηση ενδιαφέρεται για την ανάλυση των επιδόσεων και της βιωσιμότητας της επιχείρησης αυτής, με σκοπό την έγκαιρη και ορθή διάγνωση των προβλημάτων που πιθανόν αντιμετωπίζει. Οι χρηματοοικονομικοί ερευνητές έχουν προσεγγίσει το ζήτημα αυτό εξετάζοντας τις επιμέρους μορφές της χρηματοοικονομικής αποτυχίας (financial distress) μιας επιχείρησης, όπως η αθέτηση υποχρεώσεων προς τους πιστωτές της (default), η έλλειψη ρευστότητας και ο υψηλός δανεισμός (insolvency) και φυσικά η πτώχευση (bankruptcy). Οι επιδράσεις των τριών αυτών φαινομένων δημιουργούν σοβαρά προβλήματα σε όσους σχετίζονται άμεσα με τις επιχειρήσεις που τα παρουσιάζουν.

Οι διαπιστώσεις λοιπόν αυτές τονίζουν με τον καλύτερο τρόπο την αναγκαιότητα ανάπτυξης κατάλληλων διαδικασιών και συστημάτων ελέγχου των επιδόσεων και της βιωσιμότητας των επιχειρήσεων. Οι διαδικασίες αυτές στις μέρες μας έχουν αποδειχθεί χρήσιμα εργαλεία για τα χρηματοπιστωτικά ιδρύματα, τους επενδυτές και τις επιχειρήσεις διότι μέσω αυτών είναι σε θέση να αξιολογήσουν τις προοπτικές και την πορεία των επιχειρήσεων με τις οποίες συνδέονται.

Πιο συγκεκριμένα, όσον αφορά τα χρηματοπιστωτικά ιδρύματα η ανάπτυξη και εφαρμογή συστημάτων εκτίμησης του πιστωτικού κινδύνου που χαρακτηρίζει τις επιχειρήσεις, αποτελεί μια διαδικασία-ενέργεια ζωτικής σημασίας και αυτό γιατί η πιο καθοριστική αιτία οικονομικής αποτυχίας των περισσότερων τραπεζών είναι η απώλεια κεφαλαίων από χορηγούμενα δάνεια προς προβληματικές επιχειρήσεις.

Για το λόγο αυτό ολοένα και περισσότερες τράπεζες επικεντρώνουν την προσοχή τους στον τομέα ανάπτυξης τέτοιων συστημάτων διότι τα θεωρούν σημαντικά εργαλεία για τη διατήρηση της 'εσωτερικής πιστωτικής τους πειθαρχίας'. Επίσης έχουν προβεί και σε ουσιαστικές επενδύσεις για να αναβαθμίσουν τα ήδη υπάρχοντα συστήματα προς την κατεύθυνση της καλύτερης μέτρησης του κινδύνου.

Άλλος ένας σημαντικός λόγος για την ανάπτυξη τέτοιων συστημάτων, από τους χρηματοπιστωτικούς οργανισμούς, είναι μια σειρά αποφάσεων που ελήφθησαν από την Επιτροπή Τραπεζικής Επιθεώρησης της Βασιλείας (Basle Committee of Banking Supervisors) οι οποίες άλλαξαν τα διεθνή δεδομένα. Μέσω αυτών των αποφάσεων επετράπη στις τράπεζες να χρησιμοποιούν συστήματα αξιολόγησης του πιστωτικού κινδύνου και ρυθμίστηκε ένα πλαίσιο για τον έλεγχο και την απόδειξη της ορθής λειτουργίας τους (βλ. Παράρτημα).

1.2 Τι είναι ένα σύστημα εκτίμησης πιστωτικού κινδύνου

Ως πιστωτικός κίνδυνος μιας επιχείρησης, όπως προαναφέρθηκε, ονομάζεται ο κίνδυνος που προκύπτει από την αδυναμία εκπλήρωσης των υποχρεώσεών της, οι οποίες δημιουργούνται από την λήψη κάποιου δανείου. Δηλαδή μια επιχείρηση μπορεί να μην είναι ικανή να αποπληρώσει ένα δάνειο, να το αποπληρώσει μερικώς ή να καθυστερήσει την αποπληρωμή του. Ένα σύστημα εκτίμησης πιστωτικού κινδύνου (credit risk rating system) λοιπόν, είναι ένα εργαλείο το οποίο βοηθάει ένα χρηματοπιστωτικό οργανισμό στη λήψη απόφασης για την δανειοδότηση ή μη μιας επιχείρησης, παρέχοντας το βαθμό του πιστωτικού κινδύνου που διέπει την επικείμενη-υπό εξέταση δανειοδότηση. Ουσιαστικά είναι ένα σύστημα το οποίο εκτιμά την πιστοληπτική ικανότητα της εκάστοτε επιχείρησης.

Τα συστήματα όμως αυτά μπορούν να χρησιμοποιηθούν από τα πιστωτικά ιδρύματα και ως συστήματα ελέγχου των επιχειρήσεων που έχουν ήδη λάβει κάποιο δάνειο. Ο χρονικός ορίζοντας για τον οποίον ισχύει η εκτίμηση είναι συνήθως το ένα έτος διότι η διαδικασία της δανειοδότησης είναι δυναμική. Κατά την διάρκεια της δανειοδότησης κάποιες καταστάσεις μπορούν να μεταβάλλουν τις παραμέτρους του δανεισμού. Για παράδειγμα μπορεί μια επιχείρηση να είναι ικανή να δανειοδοτηθεί αλλά με το πέρασ του χρόνου, λόγω κάποιων δυσχερών συγκυριών, να μεταβληθεί η πιστοληπτική της ικανότητα. Μέσω των συστημάτων λοιπόν αυτών μπορεί να επαναπροσδιοριστεί ο κίνδυνος που διέπει την εκάστοτε δανειοδότηση.

Στο σημείο αυτό καλό θα είναι να διευκρινιστεί η έννοια του πιστωτικού κινδύνου, για να μην δημιουργείται σύγχυση στον αναγνώστη. Ο πιστωτικός κίνδυνος που χαρακτηρίζει μια επιχείρηση είναι ταυτόσημος με τον πιστωτικό κίνδυνο που διέπει την επικείμενη δανειοδότηση της επιχείρησης αυτής.

Ένα σύστημα εκτίμησης-αξιολόγησης του πιστωτικού κινδύνου μπορεί να θεωρηθεί και ως ένα σύστημα ταξινόμησης των επιχειρήσεων, βάσει του πιστωτικού κινδύνου που τις χαρακτηρίζει. Αυτό συμβαίνει διότι το σύστημα εξετάζει την εκάστοτε επιχείρηση που κάνει αίτηση για χρηματοδότηση (κάποια χαρακτηριστικά της) και ανάλογα με τον πιστωτικό κίνδυνο που της προσδίδει, μέσω κάποιας τεχνικής-διαδικασίας, την κατατάσσει σε κάποια κατηγορία. Η κάθε κατηγορία

φανερώνει το βαθμό του κινδύνου αθέτησης των υποχρεώσεων της επιχείρησης.

Τα πιο εξελιγμένα συστήματα παρέχουν επίσης και πληροφορίες σχετικά με το ύψος του χορηγούμενου κεφαλαίου που μπορεί να μην αποπληρωθεί όπως και για το είδος της δανειοδότησης που μπορεί να λάβει ο υπό εξέταση δανειολήπτης-επιχείρηση, ανάλογα την κατηγορία στην οποία θα ταξινομηθεί.

Τα συστήματα εκτίμησης πιστωτικού κινδύνου διαχωρίζονται σε:

- αυτά που αναπτύσσει και χρησιμοποιεί μια τράπεζα και ονομάζονται εσωτερικά συστήματα αξιολόγησης (internal rating systems).
- και σε αυτά που χρησιμοποιούν οι οργανισμοί αξιολόγησης επενδυτικών αποφάσεων (εξωτερικά συστήματα , external rating systems).

Η ειδοποιός διαφορά τους έγκειται στο γεγονός ότι στα εξωτερικά συστήματα ο οργανισμός που τα κατασκευάζει έχει ουδέτερο ρόλο στην επενδυτική απόφαση που αξιολογεί, διατηρεί ίσες αποστάσεις ανάμεσα στο δανειζόμενο και στο δανειστή. Οι τράπεζες όμως κατασκευάζουν συστήματα πιο υποκειμενικά προς την κατεύθυνση της ‘υποστήριξης’ των δικών τους συμφερόντων και αναγκών. Μια ακόμη διαφορά είναι ότι οι οργανισμοί που κάνουν αξιολογήσεις μπορεί να αποκαλύπτουν το αποτέλεσμα της αξιολόγησης αλλά δεν αποκαλύπτουν τη διαδικασία. Οι τράπεζες όμως γνωρίζουν λεπτομερώς την διαδικασία που ακολουθήθηκε, αφού αυτές ανέπτυξαν το σύστημα εκτίμησης και καθόρισαν τις παραμέτρους της απόφασης, και έτσι είναι σε θέση να υποστηρίξουν μια επικείμενη απόφαση πιο ολοκληρωμένα και πιο ουσιαστικά.

1.3 Βασικά χαρακτηριστικά ενός συστήματος εκτίμησης πιστωτικού κινδύνου

Ένα σύστημα εκτίμησης πιστωτικού κινδύνου διέπεται από κάποια βασικά χαρακτηριστικά, τα οποία διαδραματίζουν σημαντικό ρόλο στην αξιοπιστία των εκτιμήσεων που αυτό παρέχει. Τα πιο σημαντικά από αυτά, που αναλύονται σε αυτή την ενότητα, είναι ο αριθμός των κατηγοριών στις οποίες ταξινομούνται οι επιχειρήσεις βάσει του πιστωτικού τους κινδύνου, η τεχνική ταξινόμησης που χρησιμοποιείται καθώς και η εύρεση και ορθή ρύθμιση των παραμέτρων-παραγόντων που επηρεάζουν τον εκτιμώμενο πιστωτικό κίνδυνο.

1.3.1 Κατηγορίες διαχωρισμού του πιστωτικού κινδύνου

Το προφανέστερο βασικό χαρακτηριστικό ενός ισχυρού συστήματος είναι ο αριθμός κατηγοριών στον οποίο διαχωρίζει τον κίνδυνο. Ο αριθμός αυτός, που καθορίζεται από τους κατασκευαστές του συστήματος, εκφράζει την ικανότητα ενός συστήματος εκτίμησης να διακρίνει τον κίνδυνο.

Τα περισσότερα συστήματα εκτίμησης του πιστωτικού κινδύνου που χρησιμοποιούνταν από τις τράπεζες πριν μερικά χρόνια ταξινομούσαν τις επιχειρήσεις σε δύο κατηγορίες, σε αυτές που ήταν ικανές να λάβουν κάποιο δάνειο και σε αυτές που δεν πληρούσαν τις προϋποθέσεις. Τα συστήματα αυτά αποτελούσαν συμπληρωματικό εργαλείο στην απόφαση της δανειοδότησης. Κύριο ρόλο έπαιξε η κρίση και η εμπειρία του πιστωτικού αναλυτή της τράπεζας.

Έρευνες έδειξαν ότι τα συστήματα που ταξινομούν τις επιχειρήσεις σε περισσότερες κατηγορίες είναι πιο αξιόπιστα και βοηθούν με μεγαλύτερη αποτελεσματικότητα στην απόφαση δανειοδότησης μιας επιχείρησης. Αυτό συμβαίνει διότι γίνεται καλύτερη διάκριση του πιστωτικού κινδύνου. Ο βαθμός του κινδύνου που χαρακτηρίζει κάθε κατηγορία ορίζεται σε ένα στενότερο πεδίο τιμών από ότι σε ένα σύστημα που χρησιμοποιεί δύο κατηγορίες ταξινόμησης. Αυτό έχει ως αποτέλεσμα τα συστήματα αυτά να παρέχουν πιο σαφείς πληροφορίες για τον κίνδυνο αθέτησης των υποχρεώσεων μιας επιχείρησης, δηλαδή η κάθε κατηγορία διέπεται από κάποια πιο ευδιάκριτα χαρακτηριστικά όσον αφορά την πιστοληπτική ικανότητα των επιχειρήσεων που ανήκουν σε αυτή.

Οι τράπεζες λοιπόν, προχώρησαν στην ανάπτυξη συστημάτων στα οποία ο πιστωτικός κίνδυνος της υπό εξέταση δραστηριότητας (χορήγηση δανείου) μπορεί να ταξινομηθεί σε περισσότερες κατηγορίες. Οι κατηγορίες χωρίζονται σε αυτές:

- που ο εκτιμώμενος πιστωτικός κίνδυνος είναι χαμηλός (pass grade ή investment grade) και σε αυτές κατατάσσονται επιχειρήσεις που είναι ικανές για δανεισμό (πάντα με τον ανάλογο κίνδυνο).
- και σε αυτές που ο πιστωτικός κίνδυνος είναι υψηλός (below investment) και σε αυτές ανήκουν επιχειρήσεις που είναι επισφαλείς ή μη ικανές για χορήγηση δανείου.

Σε έρευνα που έγινε στις πενήντα μεγαλύτερες τράπεζες της Αμερικής και εξέτασε τα συστήματα αξιολόγησης (εκτίμησης πιστωτικού κινδύνου) που χρησιμοποιούν βρέθηκε ότι οι κατηγορίες διαχωρισμού των επιχειρήσεων κυμαίνονται από δύο έως έξι, όσον αφορά τις κατηγορίες χαμηλού κινδύνου (Treacy και Carey, 2000). Επίσης η έρευνα έδειξε ότι οι μεγαλύτερες σε εμπορική αξία τράπεζες χρησιμοποιούν συστήματα που κατατάσσουν τον πιστωτικό κίνδυνο σε πολλές κατηγορίες. Αυτό συμβαίνει διότι διαχειρίζονται υπέρογκα χρηματικά ποσά

και σαφώς θέλουν να είναι πιο σαφής, άρα και πιο ασφαλής, η εκτίμηση του κινδύνου.

Κύρια βάση για την ανάπτυξη τέτοιων συστημάτων αποτελεί η πιθανότητα ασυνέπειας (probability of default). Το σύστημα εκτίμησης του πιστωτικού κινδύνου προσδίδει στην επιχείρηση μια βαθμολογία που αφορά την αθέτηση των υποχρεώσεων της. Σύμφωνα με την βαθμολογία αυτή η επιχείρηση κατατάσσεται και στην ανάλογη κατηγορία.

Τα τελευταία χρόνια μάλιστα, οι τράπεζες σε συνεργασία με τους ειδικούς έχουν προβεί σε ενέργειες για την κατασκευή καινοτόμων συστημάτων εκτίμησης πιστωτικού κινδύνου. Τα συστήματα αυτά δεν παρέχουν μόνο την εκτίμηση του πιστωτικού κινδύνου που διέπει μια επικείμενη χρηματοδότηση αλλά και την εκτίμηση του κεφαλαίου που είναι πιθανόν να μην αποπληρωθεί. Τα συστήματα αυτά ονομάζονται «δύο κατευθύνσεων» (two-dimensional rating systems, Πίνακας 1.1). Δηλαδή συνδυάζουν την πιθανότητα αθέτησης των υποχρεώσεων του δανειζόμενου (probability of default - PD) με την αναλογία του δανείου που είναι πιθανόν να απολεσθεί στην περίπτωση της αθέτησης (loss in the event of default). Ο συνδυασμός αυτός (πολλαπλασιασμός των δύο συντελεστών) παρέχει την αναμενόμενη ζημία από τη χρηματοδότηση μιας επιχείρησης (expected loss).

ΠΙΝΑΚΑΣ 1.1: Σύστημα αξιολόγησης δύο κατευθύνσεων

Κατηγορία	Πιθανότητα ασυνέπειας (%)	Συντελεστής απώλειας κεφαλαίου σε περίπτωση ασυνέπειας (%)	Αναμενόμενη ζημία (%)
1 - Καθόλου Κίνδυνος	0.0	Έστω 30 κατά μέσο όρο	0.00
2 - Πολύ Χαμηλός Κίνδυνος	0.1		0.03
3 - Χαμηλός Κίνδυνος	0.3		0.09
4 - Μέσος Κίνδυνος	1		0.30
5 - Αποδεκτός Κίνδυνος	3		0.90
6 - Οριακός Κίνδυνος	6		1.80
7 - Μη Αποδεκτός Κίνδυνος	20		6.00
8 - Υψηλός Κίνδυνος	62		18.00
9 - Πολύ Υψηλός Κίνδυνος	100		30.00

Πηγή: Treacy και Carey, 2000

Επίσης υπάρχουν και μερικά συστήματα που έχουν τη δυνατότητα να συνδέουν τον εκτιμώμενο πιστωτικό κίνδυνο που προσδίδουν σε μια επιχείρηση με τα χαρακτηριστικά του δανείου που μπορεί να τους χορηγηθεί, όπως τη διάρκεια (μικρής ή μακράς διάρκειας) και τον τρόπο αποπληρωμής (μέσω αλληλόχρεων λογαριασμών ή δόσεων). Δηλαδή υπάρχει σύνδεση κάθε κατηγορίας ταξινόμησης των επιχειρήσεων, βάσει του εκτιμώμενου πιστωτικού κινδύνου που τις χαρακτηρίζει, με το είδος της δανειοδότησης.

Στον πίνακα 1.2 που ακολουθεί, παρουσιάζονται οι κατηγορίες ταξινόμησης των συστημάτων που χρησιμοποιούν δύο από τους μεγαλύτερους οργανισμούς αξιολόγησης επενδυτικών αποφάσεων παγκοσμίως, των Moody's και Standard & Poor's (external rating systems). Τα συστήματα εκτίμησης του πιστωτικού κινδύνου των οργανισμών αυτών ταξινομούν τις επιχειρήσεις σε περισσότερες κατηγορίες από ότι τα εσωτερικά συστήματα που χρησιμοποιούν συνήθως από έξι έως δέκα κατηγορίες ταξινόμησης (Krahnert και Weber, 2001).

ΠΙΝΑΚΑΣ 1.2**Κατηγορίες αξιολόγησης πιστωτικού κινδύνου των Moody's και Standard & Poor's**

	Moody's		Standard & Poor's	
	Κατηγορίες	Υποκατηγορίες	Κατηγορίες	Υποκατηγορίες
Κατηγορίες Χαμηλού Κινδύνου	Aaa		AAA	
	Aa	Aa1, Aa2, Aa3	AA	AA+, AA, AA-
	A	A1, A2, A3	A	A+, A, A-
	Baa	Baa1, Baa2, Baa3	BBB	BBB+, BBB, BBB-
Κατηγορίες Υψηλού Κινδύνου	Ba	Ba1, Ba2, Ba3	BB	BB+, BB, BB-
	B	B1, B2, B3	B	B+, B, B-
	Caa, Ca, C		CCC, CC, C	
	Αποτυχία		D	

Πηγή: Treacy και Carey, 2000

1.3.2 Καθορισμός και σύνθεση παραγόντων που επηρεάζουν τον πιστωτικό κίνδυνο

Αν ο αριθμός των κατηγοριών ταξινόμησης είναι το κριτήριο για το πόσο καλά ένα σύστημα διαφοροποιεί τον κίνδυνο, ο καθορισμός και η σύνθεση των παραγόντων που τον επηρεάζουν παίζει το σημαντικότερο ρόλο στην ορθότητα της εκτίμησης του πιστωτικού κινδύνου που το σύστημα παρέχει. Τα κριτήρια αυτά επιλέγονται για να δώσουν μια ολοκληρωμένη εικόνα της επιχείρησης που κάνει αίτηση για χρηματοδότηση. Είναι αναγκαίο λοιπόν να επιλεγούν και να αναλυθούν λεπτομερώς, ώστε όταν αυτά ενσωματωθούν στο σύστημα να το καθιστούν ικανό να εξασφαλίζει ασφάλεια και πλήρης πληροφόρηση όσον αφορά την εκτίμηση που αυτό παρέχει.

Η δυσκολότερη πτυχή λοιπόν στην κατασκευή ενός συστήματος εκτίμησης πιστωτικού κινδύνου είναι εκτός από την εύρεση των παραγόντων που επηρεάζουν το συνολικό πιστωτικό κίνδυνο και η σύνθεσή τους. Σε κάθε κριτήριο δηλαδή πρέπει να αποδοθεί ένα βάρος που θα υποδηλώνει το βαθμό συμμετοχής του κριτηρίου στη διαδικασία αξιολόγησης. Οι χρηματοοικονομικοί και πιστωτικοί αναλυτές πρέπει να αποδώσουν μέσα στο σύστημα με ορθότητα σε τι βαθμό κάθε παράγοντας επηρεάζει την εκτίμηση του κινδύνου ούτως ώστε αυτό να εξάγει ασφαλή αποτελέσματα.

Τα κριτήρια που επιλέγονται μπορούν να χωριστούν σε δύο κατηγορίες: στα ποσοτικά και στα ποιοτικά. Τα ποσοτικά είναι κυρίως η βαθμολογία των επιχειρήσεων σε κάποιους χρηματοοικονομικούς δείκτες. Οι κυριότεροι χρηματοοικονομικοί δείκτες (ποσοτικά κριτήρια) που χρησιμοποιούνται για την ανάπτυξη συστημάτων εκτίμησης πιστωτικού κινδύνου παρέχουν πληροφορίες σχετικά με [Crouhy et al. (2001), Ζοπουνίδης (2000)]:

1) τη ρευστότητα (liquidity ratios). Δείχνουν την ικανότητα της επιχείρησης να αντεπεξέρχεται σε κάποια ανεπιθύμητη εξέλιξη στη ροή των κεφαλαίων κίνησης (στις βραχυπρόθεσμες υποχρεώσεις της).

- Δείκτης Γενικής Ρευστότητας (current ratio)
(Κυκλοφορούν Ενεργητικό / Βραχυπρόθεσμες Υποχρεώσεις)
- Δείκτης Ειδικής Ρευστότητας (acid test ratio)
(Κυκλοφορούν Ενεργητικό – Αποθέματα / Βραχυπρόθεσμες Υποχρεώσεις)

2) τη φερεγγυότητα (solvency and capital measure ratios). Δείχνουν την ικανότητα της επιχείρησης να αντεπεξέρχεται στα χρέη της, την εισροή-εκροή κεφαλαίων (cash-flow) καθώς και το βαθμό συμμετοχής των ιδίων κεφαλαίων στη λειτουργία της.

- Δείκτης Συνολικής Ικανότητας Δανεισμού (debt to assets)
(Σύνολο Υποχρεώσεων / Σύνολο Ενεργητικού)
- Δείκτης Κάλυψης των Τόκων (number of times interest earned)
(Καθαρά Κέρδη προ τόκων και φόρων / Σύνολο τόκων)
- Δείκτης Κάλυψης Αναγκών σε Κεφάλαιο Κίνησης (working capital leverage ratio)
(Κεφάλαιο Κίνησης / Ανάγκες σε Κεφάλαιο Κίνησης)
- Δείκτης Ικανότητας Αποπληρωμής Υποχρεώσεων - Οικονομικής Αυτάρκειας
(Σύνολο Υποχρεώσεων / Περιθώριο Αυτοχρηματοδότησης από Ίδια Κεφάλαια)
- Δείκτης Ιδίων Κεφαλαίων προς Συνολικά (ratio of owner's equity to total assets)
(Ίδια Κεφάλαια / Συνολικά Κεφάλαια)
- Δείκτης Ιδίων Κεφαλαίων / Ξένα Κεφάλαια (ratio of owner's to total liabilities)
(Ίδια Κεφάλαια / Ξένα Κεφάλαια)

3) την αποδοτικότητα της (profitability). Παρέχουν μια εικόνα για την κερδοφορία της επιχείρησης. Επίσης δείχνουν πόσο αποτελεσματικά μια επιχείρηση διαχειρίζεται τους πόρους της ώστε να αυξήσει το κέρδος της και να επιτύχει τους στόχους της.

- Δείκτης Μικτού Κέρδους (gross profit margin)
(Πωλήσεις - Κόστος Πωλήσεων / Πωλήσεις)
- Δείκτης Καθαρού Κέρδους (net profit margin)
(Καθαρό Κέρδος / Πωλήσεις)
- Δείκτης Ταχύτητας Κυκλοφορίας Ενεργητικού (asset turnover ratio)
(Καθαρές Πωλήσεις / Σύνολο Ενεργητικού)
- Δείκτης Ταχύτητας Κυκλοφορίας Ιδίων Κεφαλαίων (equity turnover ratio)
(Πωλήσεις / Ίδια Κεφαλαία)

- Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων (return on equity)
(Καθαρά Λειτουργικά Κέρδη / Σύνολο Ιδίων Κεφαλαίων)
- Δείκτης Αποδοτικότητας Απασχολούμενων Κεφαλαίων (return on capital employed)
(Καθαρά Κέρδη / Συνολικά Απασχολούμενα Κεφάλαια)
- Δείκτης Οικονομικής Μόχλευσης (financial leverage ratio)
(Αποδοτικότητα Ιδίων Κεφαλαίων / Αποδοτικότητα Συνολικών Κεφαλαίων)

4) την επίδοση της διαχείρισης. Οι δείκτες της κατηγορίας αυτής βοηθούν στην αξιολόγηση της πολιτικής που ακολουθεί μια επιχείρηση στη διαχείριση των αποθεμάτων, των πιστώσεων που δίνει στους πελάτες και των πιστώσεων που παίρνει από τους προμηθευτές της.

- Δείκτης Ταχύτητας Εισπράξεως των Απαιτήσεων (receivable turnover ratio)
(Καθαρές Πωλήσεις / Μέσος Όρος Απαιτήσεων)
- Δείκτης Ταχύτ. Εξόφλησης Βραχ/μων Υποχ/ων. (trade creditors to purchases ratio)
(Αγορές ή Κόστος Πωληθέντων / Μέσο Όρο Βραχ/μων Υποχ/ων.)
- Δείκτης Ταχύτητας Κυκλοφορίας Αποθεμάτων (inventories turnover ratio)
(Κόστος Πωληθέντων / Μέσο Απόθεμα)

Τα ποιοτικά κριτήρια σχετίζονται με την ποιότητα διοίκησης της επιχείρησης, την οργάνωσή της, τη δομή λειτουργίας της, τη θέση της στην αγορά και στον κλάδο στον οποίο δραστηριοποιείται, στις προοπτικές ανάπτυξής της κ.ά. (Krahnert και Weber, 2001). Οι επιδόσεις μιας επιχείρησης στα ποιοτικά κριτήρια αξιολογούνται σε ποσοτικές κλίμακες.

Τα πρώτα συστήματα που κατασκευάστηκαν χρησιμοποιούσαν μόνο ποσοτικά κριτήρια (χρηματοοικονομικούς δείκτες). Όμως η μονομερής χρησιμοποίηση των δεικτών αυτών δέχθηκε από πολλούς ερευνητές κριτική ως προς την αποτελεσματικότητά τους να αποδώσουν ολοκληρωτικά την εικόνα μιας επιχείρησης. Για αυτό το λόγο σε πολλά συστήματα εκτίμησης πιστωτικού κινδύνου άρχισαν να χρησιμοποιούνται πλέον και ποιοτικά κριτήρια αξιολόγησης.

1.3.3 Τεχνικές ταξινόμησης

Ταξινόμηση ορίζεται ως η ένταξη ορισμένων αντικειμένων, στη συγκεκριμένη μελέτη επιχειρήσεων, σε κάποιες κατηγορίες. Στην αγγλική ορολογία χρησιμοποιούνται συνήθως οι όροι όπως discrimination (διάκριση), classification (ταξινόμηση) και sorting (διατεταγμένη ταξινόμηση). Η τεχνική ταξινόμησης που θα χρησιμοποιηθεί από ένα σύστημα παίζει σημαντικό ρόλο όσον αφορά την ορθότητα των αποτελεσμάτων που αυτό παρέχει διότι κάθε τεχνική επεξεργάζεται διαφορετικά τα εισερχόμενα δεδομένα και συνεπώς παρέχει διαφορετικά αποτελέσματα. Κάθε τεχνική ταξινόμησης διέπεται από κάποιες παραμέτρους. Η ενσωμάτωση των κριτηρίων που επηρεάζουν τον κίνδυνο, καθώς και ο τρόπος με τον οποίο συνθέτονται, στην τεχνική ταξινόμησης ρυθμίζει τις παραμέτρους αυτές και παρέχουν το μοντέλο ταξινόμησης το οποίο αποτελεί το κύριο 'σώμα' του συστήματος εκτίμησης (βλ. §1.5).

Οι τεχνικές ταξινόμησης που χρησιμοποιούνται για την ανάπτυξη συστημάτων εκτίμησης πιστωτικού κινδύνου μπορούν να χωριστούν σε δύο κατηγορίες:

- Στις στατιστικές και οικονομετρικές τεχνικές, οι οποίες αποτελούν τον 'παραδοσιακό' τρόπο αντιμετώπισης του προβλήματος της ταξινόμησης.
- Στις μη παραμετρικές τεχνικές οι οποίες έχουν αναπτυχθεί τα τελευταία χρόνια και χαρακτηριστεί ως καινοτόμες διότι είναι πιο αποτελεσματικές τεχνικές ανάπτυξης υποδειγμάτων ταξινόμησης.

1.3.3.1 Στατιστικές και οικονομετρικές τεχνικές

Η στατιστική είναι η παλαιότερη των επιστημών αντικείμενο της οποίας είναι η ανάλυση δειγμάτων με απώτερο σκοπό την εξαγωγή συμπερασμάτων επί του συνόλου. Ως ένα τέτοιο πρόβλημα αντιμετωπίζεται και η ταξινόμηση, στα πλαίσια της στατιστικής θεωρίας. Η πρώτη σημαντική μέθοδος που αναπτύχθηκε, από τον Fischer το 1936, είναι η γραμμική διακριτική ανάλυση (linear discriminant analysis) η οποία για δεκαετίες αποτελούσε την πλέον διαδεδομένη μεθοδολογία για την ανάπτυξη υποδειγμάτων ταξινόμησης. Μερικά χρόνια αργότερα, το 1947, ο Smith επέκτεινε τη μέθοδο της γραμμικής διακριτικής ανάλυσης και ανέπτυξε την τετραγωνική διακριτική ανάλυση (quadratic discriminant analysis) που θεωρήθηκε ως μια καταλληλότερη μορφή της.

Η διακριτική ανάλυση θεωρεί ότι κάθε επιχείρηση χαρακτηρίζεται από ένα διάνυσμα του οποίου στοιχεία είναι οι τιμές ανεξάρτητων μεταβλητών (χαρακτηριστικά των επιχειρήσεων), οι οποίες είναι κανονικά κατανομημένες μέσα στις δύο ομάδες ταξινόμησης (συνεπείς, μη συνεπείς επιχειρήσεις). Στόχος της

μεθόδου είναι ο υπολογισμός μιας γραμμικής σχέσης μεταξύ των μεταβλητών, ούτως ώστε να επιτευχθεί μεγιστοποίηση της διασποράς μεταξύ των δύο ομάδων ταξινόμησης και ελαχιστοποίηση της διασποράς εντός αυτών. Ο γραμμικός συνδυασμός των μεταβλητών προσδίδει στις επιχειρήσεις μια βαθμολογία. Αυτή η βαθμολογία συγκρίνεται με ένα οριακό αποτέλεσμα, το οποίο έχει υπολογιστεί βάσει μιας δεδομένης πιθανότητας ταξινόμησης και μιας εσφαλμένης πιθανότητας ταξινόμησης, και έτσι η επιχείρηση κατατάσσεται σε μια από τις δύο κατηγορίες ταξινόμησης. Πρώτος ο Altman, το 1968, εφήρμοσε τη διακριτική ανάλυση για την ανάπτυξη ενός υποδείγματος ταξινόμησης πτωχευμένων και υγιών επιχειρήσεων.

Στις δεκαετίες που ακολούθησαν δόθηκε ιδιαίτερη έμφαση στις οικονομετρικές μεθόδους ταξινόμησης. Οι πιο γνωστές από αυτές τις μεθόδους είναι το γραμμικό υπόδειγμα πιθανότητας (linear probability model), το λογιστικό υπόδειγμα πιθανότητας (logit probability model) και το κανονικό υπόδειγμα πιθανότητας (probit analysis). Οι τρεις αυτές τεχνικές είναι ουσιαστικά ειδικές μορφές της γνωστής στατιστικής παλινδρόμησης. Στόχος της στατιστικής παλινδρόμησης είναι ο εντοπισμός της συναρτησιακής μορφής που συνδέει μια εξαρτημένη μεταβλητή με ένα σύνολο ανεξάρτητων μεταβλητών. Για παράδειγμα, στη συγκεκριμένη μελέτη εξαρτημένη μεταβλητή είναι η ταξινόμηση των επιχειρήσεων και ανεξάρτητες μεταβλητές οι χρηματοοικονομικοί δείκτες (βλ. Κεφ.2 §2.2).

Το λογιστικό και το κανονικό υπόδειγμα πιθανότητας χρησιμοποιούν μη γραμμικές συναρτήσεις βάσει των οποίων υπολογίζεται η πιθανότητα μια επιχείρηση να ανήκει σε κάθε μια από τις υπό εξέταση κατηγορίες. Η διαφορά τους έγκειται στην μορφή της συνάρτησης που αναπτύσσεται. Πιο συγκεκριμένα, στο λογιστικό υπόδειγμα χρησιμοποιείται η γνωστή λογιστική συνάρτηση:

$$F(\alpha + \beta X_i) = \frac{1}{1 + e^{-\alpha - \beta X_i}}$$

Δηλαδή η πιθανότητα χρηματοοικονομικής αποτυχίας μιας επιχείρησης i , δεδομένου του διανύσματος μεταβλητών (χαρακτηριστικών της επιχείρησης) X_i , δίνεται από τη σχέση:

$$P(X_i, \beta) = F(\alpha + \beta X_i)$$

Με βάση αυτή την πιθανότητα κάθε επιχείρηση ταξινομείται στην ομάδα των συνεπών ή μη επιχειρήσεων έπειτα από τη σύγκριση με μια οριακή πιθανότητα, η οποία υπολογίζεται με στόχο την ελαχιστοποίηση των σφαλμάτων ταξινόμησης. Οι συντελεστές α και β (το διάνυσμα το οποίο περιέχει τους συντελεστές των μεταβλητών) υπολογίζονται χρησιμοποιώντας τεχνικές μέγιστης πιθανοφάνειας. Στο κανονικό υπόδειγμα η πιθανότητα αποτυχίας υπολογίζεται βάσει της αθροιστικής συνάρτησης πυκνότητας πιθανότητας της κανονικής κατανομής.

Οι πιο διαδεδομένες από τις προαναφερθείσες μεθόδους ανάπτυξης υποδειγμάτων ταξινόμησης είναι η διακριτική ανάλυση και το λογιστικό υπόδειγμα. Τα τελευταία χρόνια το λογιστικό υπόδειγμα έχει αντικαταστήσει τη διακριτική ανάλυση διότι δεν υπόκεινται σε στατιστικούς περιορισμούς. Ουσιαστικά όμως, και σε αυτό παρατηρούνται στατιστικές υποθέσεις, οι οποίες μπορεί να μην αναφέρονται στα εξεταζόμενα δεδομένα (όπως στη διακριτική ανάλυση), αλλά υφίστανται στην πιθανότητα εμφάνισης της κάθε επιλογής ταξινόμησης η οποία ακολουθεί μια συγκεκριμένη στατιστική κατανομή. Επιπλέον διάφορες έρευνες έχουν δείξει ότι τα υποδείγματα ταξινόμησης που αναπτύσσονται χρησιμοποιώντας το λογιστικό υπόδειγμα δεν είναι αποτελεσματικότερα σε σημαντικό βαθμό σε σχέση με αυτά που χρησιμοποιούν τη διακριτική ανάλυση.

Παρά την έντονη κριτική που έχουν δεχθεί οι στατιστικές και οικονομετρικές μέθοδοι ταξινόμησης, λόγω του ότι βασίζονται σε στατιστικές υποθέσεις, αποτελούν ακόμη και σήμερα σημαντικά εργαλεία για την ανάπτυξη συστημάτων εκτίμησης πιστωτικού κινδύνου. Ιδιαίτερα διαδεδομένη είναι στις μέρες μας και η χρήση τους σε συγκριτικές έρευνες οι οποίες έχουν ως στόχο την αξιολόγηση νέων αναπτυσσόμενων τεχνικών ταξινόμησης. [Πηγές: Altman και Saunders (1998), Δούμπος και Ζοπουνίδης (2001)].

1.3.3.2 Μη παραμετρικές τεχνικές

Το γεγονός ότι οι στατιστικές και οικονομετρικές τεχνικές ταξινόμησης χρησιμοποιούν στατιστικές υποθέσεις για να προσεγγίσουν τις ιδιότητες και τις σχέσεις που διέπουν τα δεδομένα ώθησε πολλούς ερευνητές στην ανάπτυξη μη παραμετρικών τεχνικών. Οι τεχνικές αυτές παρέχουν αυξημένη ευελιξία στον αποφασίζοντα αφού δεν υπόκεινται σε στατιστικές υποθέσεις και μπορούν να προσαρμόζονται ικανοποιητικά, ανάλογα με τα χρησιμοποιούμενα σύνολα δεδομένων, είτε ως γραμμικά υποδείγματα ταξινόμησης είτε ως μη-γραμμικά. Επιστημονικά πεδία όπως η προσομοίωση, η τεχνητή νοημοσύνη, η ασαφής λογική, η πολυκριτήρια ανάλυση αποφάσεων κ.ά., έχουν συμβάλει στην ανάπτυξη των τεχνικών αυτών. Παρακάτω παρατίθενται κάποια στοιχεία για τις πιο διαδεδομένες από τις τεχνικές αυτές.

Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα (neural networks) ή αλλιώς τεχνητά νευρωνικά δίκτυα (artificial neural networks) αναπτύχθηκαν από ερευνητές του χώρου της τεχνητής νοημοσύνης. Κύρια μεθοδολογική βάση για την ανάπτυξή τους υπήρξε η προσπάθεια προσομοίωσης του τρόπου λειτουργίας του ανθρώπινου εγκεφάλου όταν αυτός επεξεργάζεται ερεθίσματα τα οποία λαμβάνει λόγω των αισθητήρων του ανθρωπίνου σώματος. Η δομή του ανθρωπίνου εγκεφάλου αποτελείται από ένα πλήθος νευρώνων (neurons) οργανωμένων σε ένα ιδιαίτερα περίπλοκο δίκτυο. Κάθε νευρώνας μπορεί να δεχθεί ένα σήμα εισόδου που μπορεί να προέρχεται είτε από τη διέγερση των αισθητήρων του ανθρωπίνου σώματος είτε από άλλους νευρώνες. Το σήμα εισόδου υπόκεινται σε μια επεξεργασία, αποτέλεσμα της οποίας είναι η παραγωγή ενός σήματος εξόδου. Άρα λοιπόν όταν τα νευρωνικά δίκτυα χρησιμοποιούνται ως τεχνική ταξινόμησης το σήμα εισόδου είναι τα δεδομένα και σήμα εξόδου η ταξινόμηση.

Μια τυπική αρχιτεκτονική ενός τεχνητού νευρωνικού δικτύου περιλαμβάνει (Σχήμα 1.1):

- 1) Ένα επίπεδο εισόδου το οποίο αποτελείται από κόμβους, έναν για κάθε κριτήριο, π.χ. έναν για κάθε χρηματοοικονομικό δείκτη.
- 2) Ένα επίπεδο εξόδου το οποίο αποτελείται από έναν αριθμό κόμβων ο οποίος, όταν το νευρωνικό δίκτυο χρησιμοποιείται ως τεχνική ταξινόμησης, είναι ίσος με τις κατηγορίες ταξινόμησης π.χ. συνεπείς και ασυνεπείς επιχειρήσεις.
- 3) Μια σειρά ενδιάμεσων επιπέδων που και αυτά αποτελούνται από κόμβους. Γενικά δεν υπάρχει ένας συγκεκριμένος κανόνας που να καθορίζει τον αριθμό των κόμβων σε κάθε ενδιάμεσο επίπεδο. Αυτό συνήθως επιτυγχάνεται μέσω διαδικασιών δοκιμής και λάθους ή χρησιμοποιώντας νευρωνικά δίκτυα τα οποία μπορούν να προσαρμόσουν την αρχιτεκτονική τους ανάλογα με το εκάστοτε πρόβλημα που αντιμετωπίζουν (self-organizing neural networks). Σε προβλήματα ταξινόμησης έρευνες που έχουν πραγματοποιηθεί έδειξαν ότι συνήθως ένα ενδιάμεσο επίπεδο αρκεί για την εξαγωγή ικανοποιητικών αποτελεσμάτων. Επίσης το πλήθος των κόμβων στο ενδιάμεσο επίπεδο μπορεί να κυμαίνεται από q έως $2n+1$, όπου q το πλήθος των κατηγοριών και n ο αριθμός των κριτηρίων αξιολόγησης [Patuwo et al. (1993), Subramanian et al. (1993)]. Όλοι οι κόμβοι διαδοχικών επιπέδων συνδέονται μεταξύ τους, ενώ είναι δυνατή και η σύνδεση κόμβων που δεν ανήκουν σε διαδοχικά επίπεδα. Η σύνδεση γίνεται μέσω κάποιων βαρών (weight) τα οποία αναπαριστούν την ισχύ της σύνδεσης. Ο καθορισμός των βαρών αυτών επιτυγχάνεται μέσω διαδικασιών βελτιστοποίησης, με σκοπό την ελαχιστοποίηση των αποκλίσεων μεταξύ των αποτελεσμάτων του δικτύου και της δεδομένης ταξινόμησης (από το δείγμα εκπαίδευσης, βλ. § 1.5). Ως μέτρο των αποκλίσεων συνήθως χρησιμοποιείται το άθροισμα των τετραγώνων των σφαλμάτων.



Σχήμα 1.1: Η γενική μορφή ενός νευρωνικού δικτύου

Τα βασικά πλεονεκτήματα που παρέχουν τα τεχνητά νευρωνικά δίκτυα είναι η δυνατότητα παράλληλης επεξεργασίας των δεδομένων και κυρίως η δυνατότητα αναπαράστασης μη γραμμικών συμπεριφορών, η οποία τους επιτρέπει να προσεγγίζουν με μεγάλη ακρίβεια τη συνάρτηση με την οποία αυτά συνδέονται (Kosko, 1992). Όμως η συνάρτηση αυτή μπορεί να προσεγγίζεται, λόγω της δομής των νευρωνικών δικτύων, αλλά δε γίνεται ‘γνωστή’. Η λειτουργία των νευρωνικών δικτύων έχει χαρακτηριστεί ως ένα «μαύρο κουτί». Δηλαδή είναι δύσκολη η επεξήγηση των σχέσεων που διέπουν τις παραμέτρους των δικτύων αυτών (των κόμβων) και η συμβολή τους στο εξαγόμενο αποτέλεσμα. Επίσης ο αυξημένος υπολογιστικός φόρτος που απαιτείται για την ‘εκπαίδευσή’ τους είναι ακόμη ένα μειονεκτήματα τους.

Βάσει των μελετών που έχουν γίνει, οι τεχνικές ταξινόμησης που βασίζονται στα νευρωνικά δίκτυα παρέχουν καλύτερα αποτελέσματα από τις στατιστικές ή οικονομετρικές τεχνικές όταν η πολυπλοκότητα του προβλήματος αυξάνει (μεγάλος αριθμός κατηγοριών και κριτηρίων αξιολόγησης). Πάντως οι έρευνες που έχουν πραγματοποιηθεί με σκοπό τη διερεύνηση της αποτελεσματικότητας των νευρωνικών δικτύων ως εναλλακτικής τεχνικής ταξινόμησης έναντι των «παραδοσιακών» προσεγγίσεων που αναλύθηκαν προηγουμένως, δεν έχουν καταλήξει σε κάποιο ξεκάθαρο συμπέρασμα [Patuwo et al. (1993), Subramanian et al. (1993)].

Μηχανική Μάθηση

Η μηχανική μάθηση (machine learning) είναι ένας από τους σημαντικότερους χώρους της τεχνητής νοημοσύνης, ο οποίος αποσκοπεί στη διερεύνηση των μηχανισμών και των υπολογιστικών διαδικασιών μέσω των οποίων είναι δυνατή η εξαγωγή επιθυμητών αποτελεσμάτων χρησιμοποιώντας την υπάρχουσα εμπειρία και γνώση. Στην κατηγορία της μηχανικής μάθησης ανήκουν και τα έμπειρα συστήματα (expert systems).

Συνήθως οι μηχανισμοί αυτοί λαμβάνουν την μορφή ενός δένδρου απόφασης (Σχήμα 1.2) και αναπτύσσονται σύμφωνα με τον αλγόριθμο της επαγωγικής μάθησης. Από τους πιο διαδεδομένους αλγορίθμους είναι ο C4.5, ο οποίος αναπτύχθηκε από τον Quinlan (1993). Ο αλγόριθμος του δένδρου απόφασης αφού κατασκευαστεί, ελέγχεται μέσω του δείγματος εκπαίδευσης (βλ. § 1.5) και διορθώνεται, αν τυχόν υπάρχουν λάθη στην ταξινόμηση την οποία εξάγει. Δηλαδή, το δένδρο διαμορφώνεται έτσι ώστε να διέπεται από κάποιους κανόνες που θα οδηγήσουν στην ταξινόμηση των επιχειρήσεων. Κάθε κόμβος του δένδρου περιλαμβάνει κάποια κριτήρια αξιολόγησης (g_i), συνήθως χρηματοοικονομικούς δείκτες, τα οποία ελέγχονται από τις συνθήκες που βρίσκονται στα 'κλαδιά' του δένδρου. Τέλος στα 'φύλλα' του δένδρου βρίσκονται οι δύο κατηγορίες ταξινόμησης (C1=συνεπείς επιχειρήσεις, C2=ασυνεπείς επιχειρήσεις).

Βασικά πλεονεκτήματα των συστημάτων ταξινόμησης που χρησιμοποιούν τεχνικές βασισμένες στη μηχανική μάθηση είναι ότι μπορούν να επεξεργαστούν ποιοτικά δεδομένα και ότι είναι κατανοητή η μορφή του αναπτυσσόμενου υποδείγματος. Μια νέα μεθοδολογία στο χώρο της μηχανικής μάθησης είναι η θεωρία των προσεγγιστικών συνόλων (rough set theory) που αναπτύχθηκε από τον Pawlak το 1982 ως μια νέα προσέγγιση για τον εντοπισμό και την περιγραφή των συσχετίσεων των χαρακτηριστικών των υπό εξέταση αντικειμένων (για περαιτέρω πληροφορίες βλέπε Δούμπος και Ζοπουνίδης, 2001). Η τεχνική αυτή έχει ομοιότητες με την τεχνική των ασαφών συνόλων, η οποία παρουσιάζεται παρακάτω.

Σχήμα 1.2: Σχηματική αναπαράσταση ενός δένδρου απόφασης

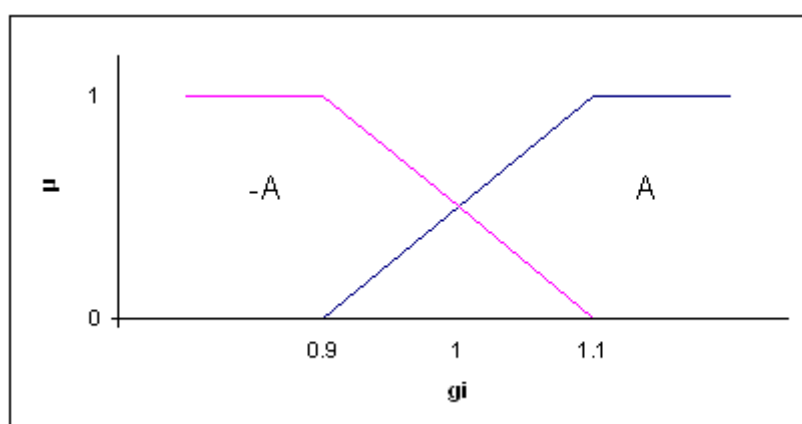
Ασαφής Λογική

Η θεωρία της ασαφούς λογικής (fuzzy set theory) η οποία αναπτύχθηκε από τον Zadeh το 1965, παρέχει τα κατάλληλα μέσα για τη μοντελοποίηση, αναπαράσταση και αντιμετώπιση προβλημάτων λήψης αποφάσεων τα οποία εμπεριέχουν κάποιο βαθμό ασάφειας. Η βασική έννοια της καινοτόμου αυτής τεχνικής είναι η έννοια του ασαφούς συνόλου. Ως ασαφές σύνολο θεωρείται ένα σύνολο αντικειμένων το οποίο δεν έχει σαφή όρια.

Στην περίπτωση ενός σαφούς ορισμένου συνόλου A , η πρόταση «η επιχείρηση x ανήκει στο σύνολο A » είναι είτε αληθής είτε ψευδής. Αντίθετα στην περίπτωση ενός ασαφούς συνόλου η παραπάνω δεν μπορεί απλά να χαρακτηριστεί ως αληθής ή ψευδής. Δεδομένης της έλλειψης σαφώς καθορισμένων ορίων για το ασαφές σύνολο η πρόταση δύναται να είναι εν μέρει αληθής (και αντίστοιχα εν μέρει ψευδής).

Στη θεωρία των ασαφών συνόλων η μοντελοποίηση του βαθμού ισχύος τέτοιων ασαφών εκτιμήσεων, υλοποιείται μέσω του ορισμού μιας συνάρτησης συμμετοχής (membership function). Η μορφή της συνάρτησης συμμετοχής καθορίζει το βαθμό συμμετοχής, ο οποίος κυμαίνεται στο διάστημα $[0, 1]$. Στο προαναφερθέν

παράδειγμα, ένας βαθμός συμμετοχής ίσος με τη μονάδα υποδεικνύει ότι η πρόταση είναι αληθής. Αντίστοιχα εάν είναι ίσος με μηδέν, τότε η πρόταση είναι ψευδής. Οποιοσδήποτε ενδιάμεσος βαθμός συμμετοχής (μεταξύ 0 και 1) συνεπάγεται ότι η πρόταση είναι εν μέρει αληθής. Όπως είναι προφανές, ο βαθμός συμμετοχής μιας επιχείρησης στο συμπληρωματικό υποσύνολο ενός ασαφούς συνόλου ισούται με τη μονάδα πλην το βαθμό συμμετοχής στο ασαφές σύνολο, υποδεικνύοντας τον βαθμό ισχύος της άρνησης της παραπάνω πρότασης. Το Σχήμα 1.3 παρουσιάζει ένα δείγμα της μορφής που μπορεί να έχει η συνάρτηση συμμετοχής μ της πρότασης «η επιχείρηση χ ανήκει στο σύνολο A βάσει του κριτηρίου αξιολόγησης g_i » και της άρνησης της (ως $-A$ συμβολίζεται το συμπληρωματικό σύνολο του συνόλου A).



Σχήμα 1.3: Παράδειγμα της συνάρτησης συμμετοχής

Προκειμένου να εξαχθεί ένα συνολικό συμπέρασμα σχετικά με το βαθμό συμμετοχής μιας επιχείρησης σε ένα ασαφές σύνολο, βάσει των κριτηρίων αξιολόγησης, χρειάζεται να γίνει η σύνθεση όλων των επιμέρους βαθμών συμμετοχής για κάθε κριτήριο αξιολόγησης. Αυτό επιτυγχάνεται με τη χρήση ορισμένων τελεστών όπως «ΚΑΙ» (ένωση), «Η» (τομή), ή κάποιου συνδυασμού τους.

Στην περίπτωση προβλημάτων ταξινόμησης κάθε κατηγορία μπορεί να θεωρηθεί ως ένα ασαφές σύνολο. Το αναπτυσσόμενο μοντέλο ταξινόμησης έχει τη μορφή ασαφών κανόνων απόφασης της μορφής:

$$\text{ΕΑΝ } (g_{j1} \text{ ανήκει } A_{1a}) \wedge (g_{j2} \text{ ανήκει } A_{2b}) \wedge \dots \wedge (g_{ji} \text{ ανήκει } A_{nc}) \text{ ΤΟΤΕ } X_n \text{ ανήκει } C_k,$$

όπου κάθε $A_{i(j)}$ αντιστοιχεί σε ένα ασαφές σύνολο οριζόμενο στο πεδίο τιμών του κάθε κριτηρίου αξιολόγησης g_i , X_n η εκάστοτε επιχείρηση (n ο αριθμός των επιχειρήσεων) και C_k οι κατηγορίες ταξινόμησης, συνεπείς και μη επιχειρήσεις ($k = 0$ ή 1). Η ισχύς κάθε επιμέρους συνθήκης εξετάζεται υπολογίζοντας το βαθμό

συμμετοχής της κάθε τιμής g_{ji} στο αντίστοιχο ασαφές σύνολο του κριτηρίου το οποίο εμφανίζεται στη στοιχειώδη συνθήκη (Δούμπος και Ζοπουνίδης, 2001).

Η ασαφής λογική τα τελευταία χρόνια έχει αρχίσει να εξετάζεται σε συνεργασία και με άλλες τεχνικές, όπως τα νευρωνικά δίκτυα (neurofuzzy systems), τα έμπειρα συστήματα (fuzzy ruled-based expert systems) και την πολυκριτήρια ανάλυση αποφάσεων με σκοπό την εξαγωγή σωστότερων αποτελεσμάτων.

Πολυκριτήρια Ανάλυση

Η πολυκριτήρια ανάλυση αποφάσεων (multicriteria decision aid) αποτελεί έναν εξελιγμένο κλάδο της επιχειρησιακής έρευνας, ο οποίος έχει γνωρίσει μεγάλη ανάπτυξη τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Βασικό ρόλο στην ανάπτυξη και διάδοσή της αποτέλεσε η λογική διαπίστωση ότι όταν κάποιος αντιμετωπίζει ένα πολύπλοκο πρόβλημα λήψεως μιας απόφασης χρειάζεται να εξετάσει όλες τις παραμέτρους του προβλήματος και τα κριτήρια που επηρεάζουν τη λήψη της απόφασης. Το πρόβλημα όμως που γεννάται είναι κατά ποίον τρόπο πρέπει να πραγματοποιηθεί η σύνθεση των παραμέτρων και κριτηρίων αυτών ώστε να επιτευχθεί η λήψη μιας σωστής απόφασης. Αντικείμενο λοιπόν της πολυκριτήριας ανάλυσης αποφάσεων είναι η αντιμετώπιση του προβλήματος αυτού.

Η κύρια διαφορά της πολυκριτήριας ανάλυσης από τις προηγούμενες τεχνικές που αναλύθηκαν δεν είναι τόσο η σύνθεση όλων των παραμέτρων ενός προβλήματος. Αυτή πραγματοποιείται και μέσω των προαναφερθησών μεθόδων. Η ειδοποιός διαφορά και βασικό χαρακτηριστικό γνώρισμα της είναι ότι ο αποφασίζων διαδραματίζει σημαντικό ρόλο στην σύνθεση αυτή, κατά κάποιον τρόπο 'προκαθορίζει' το αποτέλεσμα. Οι προηγούμενες τεχνικές που εξετάστηκαν προσδίδουν στον αποφασίζοντα ένα παθητικό ρόλο διότι δεν είναι σε θέση να ενσωματώσουν τις προτιμήσεις του επί του αποτελέσματος και κατά τη διαδικασία ανάπτυξή τους. Η επίτευξη του στόχου όμως αυτού είναι μια περίπλοκη διαδικασία.

Σύμφωνα με τον Roy, πρωτοπόρο επιστήμονα στο χώρο αυτό, η αντιμετώπιση πολυδιάστατων προβλημάτων λήψης αποφάσεων αποτελείται από τέσσερα στάδια. Το γενικό αυτό πλαίσιο που ανέπτυξε το 1985 αποτελεί τη φιλοσοφία όλων των μεθοδολογιών του χώρου.

Στο πρώτο στάδιο εντοπίζεται το σύνολο των εφικτών λύσεων και δυνατών αποφάσεων και καθορίζεται το αντικείμενο του προβλήματος. Στο δεύτερο στάδιο καθορίζονται όλοι οι παράγοντες (κριτήρια) οι οποίοι επιδρούν στο αποτέλεσμα της ανάλυσης των δεδομένων. Έπειτα, στο τρίτο στάδιο, καθορίζεται η μορφή του υποδείγματος σύνθεσης των κριτηρίων βάσει του οποίου θα αντιμετωπιστεί το πρόβλημα. Συνήθως το υπόδειγμα είναι μια συνάρτηση που αποτελείται από τα

κριτήρια και τους συντελεστές τους (βάρη). Δηλαδή μέσω αυτής μοντελοποιείται και αναπαριστάται η πολιτική του που ακολουθεί ο αποφασίζων. Ανάλογα με τη μέθοδο επεξεργασίας της συνάρτησης αυτής εξάγεται το επιθυμητό αποτέλεσμα. Τέλος στο τέταρτο στάδιο της διαδικασίας πραγματοποιούνται οι απαραίτητες διαδικασίες για να κατανοήσει ο αποφασίζων πως εξήχθησαν τα αποτελέσματα και να λάβει μια ικανοποιητική απόφαση.

Ο χώρος της πολυκριτήριας ανάλυσης είναι ιδιαίτερα ευρύς ως προς τη φύση των μεθοδολογικών προσεγγίσεων που έχουν αναπτυχθεί. Οι πολυκριτήριες προσεγγίσεις μπορούν να ομαδοποιηθούν στις εξής κατηγορίες [περαιτέρω πληροφορίες για τις κατηγορίες αυτές βλέπε Δούμπος και Ζοπουνίδης (2001)]:

- 1) Πολυκριτήριος μαθηματικός προγραμματισμός (multiobjective mathematical programming).
- 2) Πολυκριτήρια θεωρία χρησιμότητας (multiattribute utility theory).
- 3) Θεωρία των σχέσεων υπεροχής (outranking relations).
- 4) Αναλυτική-συνθετική προσέγγιση (preference disaggregation approach).

Η εφαρμογή τέτοιων τεχνικών μπορεί να οδηγήσει στην επίλυση προβλημάτων ταξινόμησης, κατάταξης (ranking) ή επιλογής. Αν για παράδειγμα το εξεταζόμενο πρόβλημα αφορά επιχειρήσεις, χρησιμοποιώντας την πολυκριτήρια ανάλυση αυτές μπορούν να ταξινομηθούν σε υγιείς ή μη, να καταταχθούν από τις καλύτερες προς τις χειρότερες ή να γίνει η επιλογή μιας από αυτές που ανταποκρίνεται καλύτερα στην επίλυση του υπό εξέταση προβλήματος.

Οι πρώτες προσπάθειες χρησιμοποίησης του μεθοδολογικού πλαισίου της πολυκριτήριας ανάλυσης στην επίλυση προβλημάτων ταξινόμησης έγιναν στις αρχές τις δεκαετίας του 1970. Έκτοτε ολοένα και περισσότεροι ερευνητές ασχολούνται με το πρόβλημα της ταξινόμησης χρησιμοποιώντας τέτοιες τεχνικές και ειδικότερα αυτές που υπάγονται στις τρεις τελευταίες κατηγορίες. Η πολυκριτήρια ανάλυση λοιπόν, ως μια νέα προσέγγιση στις τεχνικές ταξινόμησης έχει οδηγήσει στην πιο ολοκληρωμένη και ρεαλιστική αντιμετώπιση των προβλημάτων αυτών.

1.4 Χρησιμότητα των συστημάτων εκτίμησης πιστωτικού κινδύνου

Η ανάπτυξη και χρησιμοποίηση συστημάτων εκτίμησης πιστωτικού κινδύνου από τους χρηματοπιστωτικούς οργανισμούς παρέχει σημαντικά οφέλη σε αυτούς [Δούμπος και Ζοπουνίδης (2001), Treacy και Carey, (2000)]:

- ✓ Μειώνουν τις λανθασμένες αποφάσεις για χορήγηση δανείου. Άρα αυξάνεται η αποδοτικότητα τους.
- ✓ Εισάγουν μια κοινή βάση για την αξιολόγηση όλων των επιχειρήσεων που ζητούν χρηματοδότηση από ένα χρηματοπιστωτικό οργανισμό. Οι αιτήσεις πολλές φορές αξιολογούνται σε περιφερειακό και όχι σε κεντρικό επίπεδο, ιδιαίτερα σε περιπτώσεις όπου το ύψος του δανείου είναι περιορισμένο. Με αυτό τον τρόπο περιορίζουν την υποκειμενικότητα των πιστωτικών αναλυτών του χρηματοπιστωτικού οργανισμού με αποτέλεσμα την αποφυγή λανθασμένων αποφάσεων.
- ✓ Περιορίζουν το χρόνο και το κόστος αξιολόγησης των επιχειρήσεων. Οι πιστωτικοί αναλυτές πραγματοποιούν διεξοδικές διερευνήσεις αιτήσεων χρηματοδότησης μόνο σε εξεζητημένες περιπτώσεις.
- ✓ Αποτελούν χρήσιμο εργαλείο για τον καθορισμό του είδους του δανείου και του χρηματικού ύψους του, που μπορεί να χορηγηθεί στην εκάστοτε επιχείρηση.
- ✓ Διευκολύνουν τη διαδικασία παρακολούθησης και διαχείρισης των χορηγήσεων του χρηματοπιστωτικού οργανισμού. Όπως προαναφέρθηκε, χρονικός ορίζοντας για τον οποίον ισχύει η εκτίμηση των συστημάτων αυτών είναι συνήθως το ένα έτος διότι η διαδικασία της δανειοδότησης είναι δυναμική. Μέσω των συστημάτων αυτών μπορεί λοιπόν να επαναπροσδιοριστεί ο κίνδυνος που διέπει την εκάστοτε δανειοδότηση και έτσι γίνεται πιο εύκολος ο έλεγχος του χαρτοφυλακίου των τραπεζών. Δηλαδή, μέσω των συστημάτων η τράπεζα μπορεί να γνωρίζει ανά πάσα στιγμή το ύψος των κεφαλαίων που έχει χορηγήσει και να προσδιορίζει το επίπεδο του κινδύνου που διατρέχει.
- ✓ Βοηθούν στην ανάλυση και αύξηση της αποδοτικότητας των χορηγούμενων κεφαλαίων και επομένως και του χρηματοπιστωτικού οργανισμού διότι μια από τις κύριες πηγές εσόδων για μια τράπεζα είναι η είσπραξη τόκων. Εφόσον τα συστήματα αυτά βοηθούν στην εκτίμηση του κινδύνου των χορηγούμενων κεφαλαίων, ο εκάστοτε χρηματοπιστωτικός οργανισμός είναι σε θέση να γνωρίζει-εκτιμάει τη αποδοτικότητά τους. Μέσω αυτής της ανάλυσης μπορεί να προβεί σε διορθώσεις όσον αφορά τις υπάρχουσες ή μελλοντικές χορηγήσεις δανείων έτσι ώστε να αυξήσει τα κέρδη του από αυτές.

1.4.1 Στοιχεία για τις πιστωτικές εξελίξεις στην Ελλάδα και τη συναλλακτική συμπεριφορά των ελληνικών επιχειρήσεων

Για να γίνει κατανοητή η σημαντικότητα των συστημάτων αξιολόγησης του πιστωτικού κινδύνου για ένα πιστωτικό ίδρυμα στην παράγραφο αυτή παρατίθενται κάποια στοιχεία που αναδεικνύουν το μέγεθος του προβλήματος. Τα στοιχεία αυτά σχετίζονται με το μέγεθος των κεφαλαίων που χορηγούν οι ελληνικές τράπεζες προς τις επιχειρήσεις, με τη συναλλακτική συμπεριφορά των ελληνικών επιχειρήσεων καθώς και με τη μεταβολή της πιστοληπτικής τους ικανότητα.

Στον παρακάτω πίνακα παρατίθενται στοιχεία που αφορούν τη δανειοδότηση των επιχειρήσεων από τα χρηματοπιστωτικά ιδρύματα (Χ.Ι.). Παρατηρείται ότι οι τραπεζικές πιστώσεις προς τις επιχειρήσεις αυξήθηκαν κατά 9,6% το 2002, έναντι ανόδου κατά 18,5% το 2001 και 24,6% το 2000. Επίσης οι ελληνικές επιχειρήσεις χρωστούν 55 δισεκατομμύρια ευρώ στις τράπεζες λόγω δανείων που έχουν λάβει

ΠΙΝΑΚΑΣ 1.3
ΔΑΝΕΙΑ ΑΠΟ ΤΑ ΕΓΧΩΡΙΑ ΧΙ ΠΡΟΣ ΕΛΛΗΝΙΚΕΣ ΕΠΙΧΕΙΡΗΣΕΙΣ ΚΑΤΑ ΚΛΑΔΟ

	Υπόλοιπα 31.12.02 (εκατ. ευρώ)	Μεταβολές Υπολοίπων					
		Σε εκατ. ευρώ			Σε ποσοστά %		
		2000	2001	2002	2000	2001	2002
Επιχειρήσεις	55.012	8.370	7.838	4.813	24,6	18,5	9,6
1.Γεωργία	3.225	270	-161	-500	7,5	-4,1	-13,4
2.Βιομηχανία	14.364	1.342	791	1.749	12,8	6,7	13,9
3.Εμπόριο	15.671	3.156	3.150	147	34,2	25,5	0,9
4.Τουρισμός	2.903	207	357	732	12,9	19,7	33,7
5.Ναυτιλία	4.147	714	489	43	24,6	13,5	1,1
6. Χρημ/κές υπηρεσίες	2.869	427	1.018	1.169	167,7	149,3	68,8
7.Λοιπές	11.833	2.254	2.194	1.473	38,1	26,9	14,2

Πηγή: Τράπεζα της Ελλάδος (2003)

Γίνεται λοιπόν αντιληπτό, βάσει των κεφαλαίων που χορηγούνται προς τις επιχειρήσεις, ότι η μη πληρωμή ή η καθυστέρηση πληρωμής των δανείων αποτελούν σημαντικό πρόβλημα για τα χρηματοπιστωτικά ιδρύματα. Σύμφωνα με την Τράπεζα της Ελλάδος, κατά μέσο όρο, η καθυστέρηση άνω των τριών μηνών για το 2002 και 2003 ήταν 5% και 6%, αντίστοιχα επί των υπολοίπων. Επιπροσθέτως, σύμφωνα με τα διαθέσιμα στοιχεία που έχουν οι τράπεζες φέτος, μέχρι στιγμής το ποσοστό των καθυστερήσεων για πάσης φύσεως επιχειρηματικών δανείων ανέρχεται περίπου στο 7%.

Τα δυσμενή στοιχεία που συγκεντρώνονται από την εταιρία 'Τραπεζικά Συστήματα Πληροφοριών-Τειρεσίας Α.Ε.' και παρατίθενται στην Έκθεση του Διοικητή της Τραπεζής Ελλάδος για το έτος 2002, παρέχουν μια εικόνα της συναλλακτικής συμπεριφοράς των επιχειρήσεων (Πίνακας 1.4). Πιο συγκεκριμένα, ο συνολικός αριθμός των αιτήσεων πτώχευσης εμφανίζει σημαντική μείωση όλα τα έτη εκτός του 2000 που σημειώθηκε μια μικρή αύξηση, της τάξεως του 1,7%, στοιχείο θετικό για τις ελληνικές τράπεζες διότι ο κύριος λόγος πτώχευσης μιας επιχείρησης είναι η αδυναμία εκπλήρωσης των υποχρεώσεων της προς τους πιστωτές της.

Όσον αφορά τις διαταγές πληρωμής μπορεί ο αριθμός τους να αυξήθηκε το 2002 σε σχέση με το 2001 αλλά το χρηματικό ύψος μειώθηκε αισθητά, θετικό γεγονός και αυτό για τα ελληνικά χρηματοπιστωτικά ιδρύματα. Αντίθετα παρατηρείται αύξηση κατά 26,5% της συνολικής αξίας των ακάλυπτων επιταγών το 2002 έναντι του 2001. Δηλαδή αντιστράφηκε η πτωτική τους πορεία που είχε σημειωθεί τα τρία προηγούμενα έτη, εξέλιξη αρνητική διότι μια ακάλυπτη επιταγή σημαίνει αθέτηση ή αδυναμία εκπλήρωσης κάποιου χρέους.

ΠΙΝΑΚΑΣ 1.4
ΕΚΑΤΟΣΤΙΑΙΕΣ ΜΕΤΑΒΟΛΕΣ ΔΥΣΜΕΝΩΝ ΣΤΟΙΧΕΙΩΝ

	1997	1998	1999	2000	2001	2002
1. Αιτήσεις πτώχευσης						
Πλήθος	-20,1	-36,7	-1,0	-22,1	1,7	-37,1
Αξία	34,6	-45,4	82,9	-9,5	-22,1	-67,8
2. Διαταγές πληρωμής						
Πλήθος	19,4	-10,9	-5,2	-34,7	-26,7	21,9
Αξία	31,5	-12,2	7,5	-18,4	-18,5	-38,5
3. Ακάλυπτες επιταγές						
Αξία	16,4	1,4	-4,5	-18,4	-18,6	26,5

Πηγή: Τράπεζα της Ελλάδος (2003).

Η έρευνα που διεξήγαγε η ICAP, η μεγαλύτερη εταιρία οικονομικών πληροφοριών, εκδόσεων και συμβούλων επιχειρήσεων στην Ελλάδα, καταγράφει τη μεταβολή της πιστοληπτικής ικανότητας των ελληνικών Α.Ε. και Ε.Π.Ε. την πενταετία 1998-2002 και στους τρεις τομείς της οικονομίας. Οι επιχειρήσεις οι οποίες εξετάστηκαν στην έρευνα προέρχονται και από τους τρεις βασικούς τομείς δραστηριότητας της Ελληνικής οικονομίας, τη βιομηχανία, το εμπόριο και τις υπηρεσίες (δεν εξετάστηκαν οι Τράπεζες, Ασφάλειες, Εταιρείες Συμμετοχών, Factoring κ.λ.π). Το δείγμα αποτελούνταν από 20.936 επιχειρήσεις. Η κατάταξη των επιχειρήσεων έγινε σε τρεις βασικές περιοχές πιστοληπτικής ικανότητας: περιοχή χαμηλού πιστωτικού κινδύνου (Low Risk Zone), περιοχή μέσου πιστωτικού κινδύνου

(Medium Risk Zone) και στην περιοχή υψηλού πιστωτικού κινδύνου (High Risk Zone).

Εξετάζοντας το σύνολο των επιχειρήσεων, ανεξαρτήτως τομέα δραστηριότητας, τα κύρια ευρήματα της έρευνας, σε σχέση με τις μεταβολές της πιστοληπτικής ικανότητας είναι τα ακόλουθα: περίπου 3 από τις 4 επιχειρήσεις οι οποίες κατατάχθηκαν σε μία από τις τρεις περιοχές πιστωτικού κινδύνου παρέμειναν στην ίδια περιοχή ένα έτος μετά την αξιολόγησή τους (το 75,25%), Πίνακας 1.5. Επίσης ο αριθμός των επιχειρήσεων των οποίων αναβαθμίστηκε η πιστοληπτική ικανότητα είναι κατά 16,1% μεγαλύτερος αυτών των οποίων υποβαθμίστηκε. Αναβαθμίστηκε το 13% των επιχειρήσεων (2.726 επιχειρήσεις), ενώ υποβαθμίστηκε το 11,21% δηλαδή 2.348 επιχειρήσεις

Πίνακας 1.5: Μέση ετήσια μεταβολή πιστοληπτικής ικανότητας συνόλου επιχειρήσεων 1998-2002

Αξιολόγηση (t / t+1)	Περιοχή Χαμηλού Πιστωτικού Κινδύνου	Περιοχή Μέσου Πιστωτικού Κινδύνου	Περιοχή Υψηλού Πιστωτικού Κινδύνου	Σύνολα
Περιοχή Χαμηλού Πιστωτικού Κινδύνου	74.54% 3.914 επιχ.	24.07% 1.264 επιχ.	1.39% 73 επιχ.	100% 5.251 επιχ.
Περιοχή Μέσου Πιστωτικού Κινδύνου	13.54% 1.464 επιχ.	77.12% 8.340 επιχ.	9.34% 1.011 επιχ.	100% 10.815 επιχ.
Περιοχή Υψηλού Πιστωτικού Κινδύνου	1.18% 58 επιχ.	24.72% 1.204 επιχ.	74.10% 3.608 επιχ.	100% 4.870 επιχ.
			Σύνολο	20.936 επιχ.

Αντίστοιχα σε μέση τριετή μεταβολή (βλ. Πίνακα 1.6) παρατηρείται ότι ο αριθμός των επιχειρήσεων των οποίων αναβαθμίστηκε η πιστοληπτική ικανότητα είναι κατά 43,10% μεγαλύτερος αυτών των οποίων υποβαθμίστηκε. Δηλαδή στο σύνολο του δείγματος αναβαθμίστηκε το 21,68% των επιχειρήσεων (3.911 επιχειρήσεις) ενώ υποβαθμίστηκε το 15,15% (2.733 επιχειρήσεις).

Πίνακας 1.6: Μέση τριετής μεταβολή πιστοληπτικής ικανότητας συνόλου επιχειρήσεων 1998-2002

Αξιολόγηση (t / t+3)	Περιοχή Χαμηλού Πιστωτικού Κινδύνου	Περιοχή Μέσου Πιστωτικού Κινδύνου	Περιοχή Υψηλού Πιστωτικού Κινδύνου	Σύνολα
Περιοχή Χαμηλού Πιστωτικού Κινδύνου	65.91% 2.867 επιχ.	30.64% 1.333 επιχ.	3.45% 150 επιχ.	100% 4.350 επιχ.
Περιοχή Μέσου Πιστωτικού Κινδύνου	21.00% 1.963 επιχ.	65.63% 6.134 επιχ.	13.37% 1.250 επιχ.	100% 9.347 επιχ.
Περιοχή Υψηλού Πιστωτικού Κινδύνου	4.89% 212 επιχ.	40.05% 1.736 επιχ.	55.06% 2.387 επιχ.	100% 4.335 επιχ.
Σύνολο				18.032 επιχ.

(*Η διαφορά η οποία παρατηρείται στον αριθμό των επιχειρήσεων στους πίνακες οφείλεται στο γεγονός ότι κάποιες επιχειρήσεις δεν ήταν δυνατό να αξιολογηθούν τρία έτη μετά την αρχική τους αξιολόγηση είτε γιατί δεν ήταν διαθέσιμα τα οικονομικά στοιχεία τους είτε γιατί αδράνησαν ή πτώχευσαν.)

Οι μεταβολές των επιχειρήσεων όσον αφορά την πιστοληπτική ικανότητά τους είναι μια πραγματικότητα που καταδεικνύει τη σημαντικότητα των συστημάτων εκτίμησης του πιστωτικού κινδύνου. Ένα ισχυρό σύστημα πρέπει να προβλέπει ορθά τον πιστωτικό κίνδυνο που χαρακτηρίζει μια επιχείρηση. Τυχόν μεταβολή της πιστοληπτικής ικανότητας μιας επιχείρησης έχει αρνητικές επιπτώσεις για μια τράπεζα. Αν μια επιχείρηση ταξινομηθεί σε κατηγορία χαμηλού κινδύνου και με το πέρασ του χρόνου αποδειχθεί ότι ανήκει σε κατηγορία υψηλότερου κινδύνου η τράπεζα εκτίθεται σε κίνδυνο. Επίσης αν μια τράπεζα εκτιμηθεί ότι ανήκει σε κατηγορία υψηλού κινδύνου (δεν εγκριθεί η χορήγηση για την δανειοδότηση) και τελικά αποδειχθεί ότι ανήκει σε κατηγορία χαμηλότερου κινδύνου η τράπεζα χάνει ένα συνεπή πελάτη.

Πάντως, την τελευταία τετραετία, κυρίως μετά το 1999, σημειώθηκε σημαντική βελτίωση της ποιότητας του χαρτοφυλακίου των ελληνικών τραπεζών λόγω της πολιτικής που εφαρμόζουν υιοθετώντας διεθνή πρότυπα για τη διαχείριση των πιστωτικών κινδύνων. Όμως, όπως αναφέρεται και στην Έκθεση του Διοικητή της Τραπέζης Ελλάδος του 2002, η ταχεία πιστωτική επέκταση των τελευταίων ετών αυξάνει τη σημασία της εκτίμησης από τις τράπεζες του πιστωτικού κινδύνου, ο οποίος γενικά αποτελεί το σημαντικότερο κίνδυνο που αντιμετωπίζουν.

1.5 Διαδικασία ανάπτυξης ενός συστήματος εκτίμησης πιστωτικού κινδύνου

Η διαδικασία ανάπτυξης ενός συστήματος εκτίμησης πιστωτικού κινδύνου μπορεί να χωριστεί σε τρία στάδια [Σχήμα 1.5, Schumann και Liu (2002)]. Στο πρώτο στάδιο καθορίζεται το πρόβλημα το οποίο το σύστημα καλείται να ‘επιλύσει’ και συλλέγονται και προετοιμάζονται τα δεδομένα τα οποία θα χρησιμοποιήσει (problem definition & data preparation).

Τα δεδομένα (χαρακτηριστικά επιχειρήσεων) συλλέγονται βάσει των κριτηρίων που έχουν καθοριστεί από τους πιστωτικούς αναλυτές. Προέρχονται από εταιρίες για τις οποίες είναι γνωστό αν έχουν ανταποκριθεί στις υποχρεώσεις του δανεισμού ή όχι (δηλαδή γνωρίζουν οι κατασκευαστές του συστήματος την ακριβή ταξινόμηση των εταιριών αυτών). Οι ‘ποιότητα’ των δεδομένων πρέπει να είναι ικανοποιητική ούτως ώστε να γίνει καλύτερη εκπαίδευση του υπό ανάπτυξη συστήματος και εντέλει να κατασκευάζεται ένα αποτελεσματικό σύστημα. Η βάση δεδομένων που θα χρησιμοποιηθεί πρέπει να είναι αρκετά μεγάλη και να αντιστοιχεί σε ένα βάθος χρόνου. Με αυτό τον τρόπο κατά την εκπαίδευση του επιτυγχάνεται η ενσωμάτωση της σχέσεως του πιστωτικού κινδύνου των επιχειρήσεων με παραμέτρους όπως οι οικονομικές εξελίξεις που διαμορφώνουν το εκάστοτε επιχειρηματικό περιβάλλον.

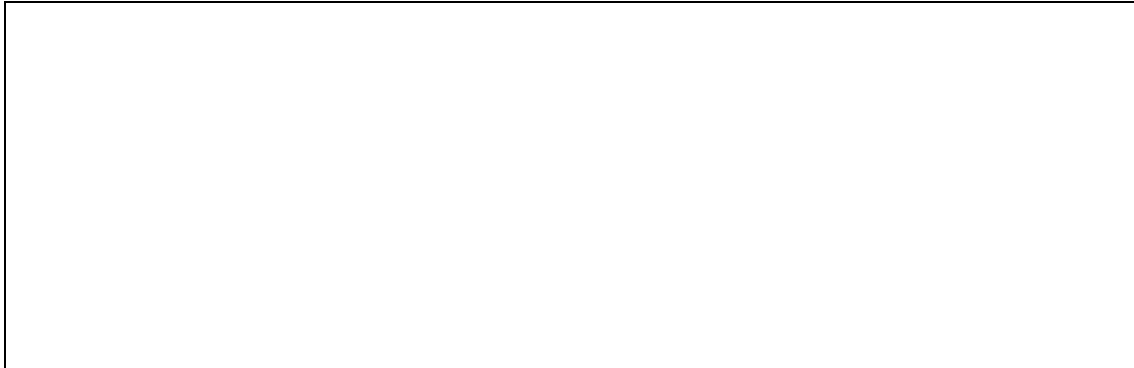
Η βάση δεδομένων χωρίζεται στο δείγμα εκπαίδευσης (training sample) και στο δείγμα ελέγχου (test sample). Οι προετοιμασίες τους έγκειται στον τρόπο με τον οποίο θα εισαχθούν στο σύστημα (βλ. Κεφ.2 §2.1 & §2.2).

Στο δεύτερο στάδιο (Σχήμα 1.4) επιλέγεται το μοντέλο με το οποίο θα αναλυθούν τα δεδομένα και πραγματοποιείται η ανάπτυξη του συστήματος (data analysis & model building). Οι παράμετροι του μοντέλου είναι τα κριτήρια μέσω των οποίων θα γίνει η ταξινόμηση και ο τρόπος με τον οποίο αυτά συνδέονται. Οι παράμετροι αυτοί ενσωματώνονται στην τεχνική ταξινόμησης που χρησιμοποιεί το σύστημα.

Το μοντέλο θα επεξεργαστεί τα δεδομένα (το δείγμα εκπαίδευσης) και σύμφωνα με αυτό θα γίνει η ταξινόμηση των επιχειρήσεων σε αυτές που μπορούν να χρηματοδοτηθούν και σε αυτές που δε μπορούν (εκτιμώμενη ταξινόμηση). Η μέθοδος ταξινόμησης που θα επιλεγεί για την ανάπτυξη του μοντέλου πρέπει να ελαχιστοποιεί τις διαφορές στην ταξινόμηση που θα προκύψει σε σχέση με τη δεδομένη ταξινόμηση του δείγματος εκπαίδευσης.

Το δείγμα εκπαίδευσης λοιπόν είναι αυτό που ‘ρυθμίζει’ τις παραμέτρους του συστήματος ταξινόμησης. Δηλαδή, αν τα αποτελέσματα που εξάγει το σύστημα, όταν σε αυτό εισέρχονται ως δεδομένα αυτά του δείγματος εκπαίδευσης, δεν είναι τα επιθυμητά οι παράμετροι του επαναπροσδιορίζονται. Για την περαιτέρω ανάλυση του υπό κατασκευή συστήματος, το μοντέλο υπόκειται σε επιπλέον έλεγχο μέσω των δειγμάτων ελέγχου. Έπειτα χρησιμοποιούνται κάποιες τεχνικές, συνήθως στατιστικές

και γραφικές, με την εφαρμογή των οποίων ελέγχεται πάλι η ορθότητα του συστήματος (ανάλυση των τεχνικών αυτών γίνεται στο Κεφ.2 §2.6).



Σχήμα 1.4: Σχηματική αναπαράσταση της διαδικασίας ανάπτυξης του μοντέλου ταξινόμησης

Το δεύτερο στάδιο, στην κατασκευή ενός τέτοιου συστήματος, ίσως είναι το σημαντικότερο και εγείρει σημαντικά ερωτήματα για τους ειδικούς. Ποία μέθοδος επεξεργασίας των δεδομένων πρέπει να χρησιμοποιηθεί (τεχνική ταξινόμησης), με ποίον τρόπο είναι καλύτερο να εισαχθούν τα δεδομένα, πώς πρέπει να καθοριστούν και να ρυθμιστούν οι παράμετροι του μοντέλου, πόσοι έλεγχοι του συστήματος πρέπει να γίνουν ώστε αυτό να εξάγει ασφαλή αποτελέσματα, ποιες μέθοδοι αξιολόγησης του πρέπει να εφαρμοστούν κ.ά.

Τέλος, στο τρίτο στάδιο αφού έχει ολοκληρωθεί η ανάπτυξη του συστήματος, αυτό εφαρμόζεται στην πράξη και αξιολογείται η 'ισχύς' του (model application & model validation). Μετά την ολοκλήρωση της διαδικασίας ανάπτυξης του μοντέλου ταξινόμησης και εφόσον αυτό κριθεί ικανοποιητικό όσον αφορά την ρύθμιση των παραμέτρων που το διέπουν και την αποτελεσματικότητά του μπορεί πλέον να χρησιμοποιηθεί και στην πράξη, για την ταξινόμηση επιχειρήσεων που κάνουν αίτηση για χρηματοδότηση.

Το σύστημα κατά την περίοδο εφαρμογής του σε πραγματικές καταστάσεις παρατηρείται από τους ειδικούς. Έπειτα από κάποιο χρονικό διάστημα συνήθως μεταβάλλονται κάποιες από τις παραμέτρους του σύμφωνα με τις διορθώσεις των ειδικών, έτσι ώστε να εξάγει ακόμα πιο ασφαλή αποτελέσματα. Γίνεται λοιπόν αντιληπτό ότι τα στάδια ενός συστήματος εκτίμησης πιστωτικού κινδύνου διέπονται από μια αμφίδρομη σχέση και ότι αυτό πρέπει να υπόκειται σε συνεχής ελέγχους και τροποποιήσεις έτσι ώστε να δίνει τη βέλτιστη δυνατή εκτίμηση.

Σχήμα 1.5: Στάδια της διαδικασίας ανάπτυξης ενός συστήματος εκτίμησης πιστωτικού κινδύνου

1.6 Απαιτήσεις – κριτήρια ελέγχου συστημάτων εκτίμησης πιστωτικού κινδύνου

Η ανάπτυξη και η χρήση συστημάτων εκτίμησης πιστωτικού κινδύνου από τα πιστωτικά ιδρύματα αποσκοπεί στην μείωση των απολεσθέντων κεφαλαίων από δανειοδοτήσεις. Η ‘ισχύς’ και η ορθότητά τους λοιπόν είναι ένα πρόβλημα που απασχολεί τις τράπεζες. Σύμφωνα όμως και με τα νέα διεθνή δεδομένα λόγω της εκδόσεως των διατάξεων της Επιτροπής Τραπεζικής Επιθεώρησης της Βασιλείας (βλ. Παράρτημα) είναι ένα θέμα που απασχολεί και την εποπτική αρχή που έχει οριστεί να τα ελέγχει.

Επομένως οι τράπεζες έχουν κάποιες απαιτήσεις από αυτά οι οποίες αποτελούν και κριτήρια ελέγχου της ορθότητας τους. Τα κριτήρια ελέγχου σχετίζονται κυρίως με τα βασικά χαρακτηριστικά ενός συστήματος που αναφέρθηκαν στην παράγραφο 1.3.2 και στον τρόπο με τον οποίο αναπτύσσεται το μοντέλο ταξινόμησης. Οι βασικές απαιτήσεις-κριτήρια ελέγχου μπορούν να χωριστούν στα ποιοτικά στοιχεία (quantitative aspects) και στα ποσοτικά (qualitative aspects) [Deutsche Bundesbank, (2003), Krahnen και Weber, (2001)]:

Τα ποσοτικά αφορούν, καταρχήν, τη διακριτική ικανότητα του συστήματος. Καλύτερη διάκριση του κινδύνου σημαίνει και πιο σαφής εκτίμηση του. Επιπλέον σημασία έχει και η ορθότητά των εκτιμήσεων που παρέχει το σύστημα. Ένα σύστημα προφανώς δεν μπορεί να είναι τέλειο. Αυτό που ενδιαφέρει το πιστωτικό ίδρυμα που το χρησιμοποιεί είναι αν οι αιτήσεις δανειοδότησης που εγκρίθηκαν (μέσω του συστήματος) μελλοντικά θα περιέχουν ένα μικρό βαθμό σφάλματος. Δηλαδή αν ελάχιστες επιχειρήσεις που εκτιμήθηκαν ως συνεπείς θα αποδειχθούν ασυνεπείς.

Τα ποιοτικά κριτήρια αφορούν τη διαδικασία που ακολουθήθηκε για την ανάπτυξη του συστήματος (μοντέλου ταξινόμησης). Δηλαδή αν έγινε σωστή επιλογή και σύνθεση των παραγόντων που επηρεάζουν τον κίνδυνο, αν χρησιμοποιήθηκε για

την ανάπτυξη του μοντέλου η κατάλληλη τεχνική ταξινόμησης, αν αυτό υποβλήθηκε σε αρκετές δοκιμές και αν η ‘ποιότητα’ των δεδομένων που χρησιμοποιήθηκαν για την ανάπτυξή του ήταν καλή. Επίσης βασική απαίτηση από ένα σύστημα είναι να είναι κατανοητά στους πιστωτικούς αναλυτές και στους αποφασίζοντες μιας επικείμενης δανειοδότησης τα αποτελέσματα τα οποία παράγει. Δηλαδή πρέπει να είναι εμφανής η επίδραση των μεταβλητών απόφασης στο εξαγόμενο αποτέλεσμα (Schumman και Liu, 2002).

ΚΕΦΑΛΑΙΟ 2

ΑΝΑΠΤΥΞΗ ΠΕΙΡΑΜΑΤΟΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

2.1 Εισαγωγή

Όπως προαναφέρθηκε, στη συγκεκριμένη μελέτη αναπτύσσονται και εφαρμόζονται διαδικασίες προετοιμασίας των δεδομένων που εισέρχονται σε ένα σύστημα εκτίμησης πιστωτικού κινδύνου. Ο τρόπος με τον οποίο προετοιμάζονται και εισέρχονται τα δεδομένα σε ένα σύστημα αποτελεί χαρακτηριστικό του, και το διαφοροποιεί ως προς τα αποτελέσματα που παράγει. Δηλαδή η διαδικασία αυτή παίζει σημαντικό ρόλο ως προς την ακρίβεια των εκτιμήσεων που παράγει ένα σύστημα.

Τα δεδομένα που χρησιμοποιήθηκαν για την ανάπτυξη του συγκεκριμένου πειράματος είναι α) οι επιδόσεις 1411 επιχειρήσεων σε έντεκα χρηματοοικονομικούς δείκτες και β) η ταξινόμηση των επιχειρήσεων ως συνεπείς και ασυνεπείς, η οποία αποτελεί την εξαρτημένη μεταβλητή του προβλήματος. Οι δείκτες αποτελούν τις ανεξάρτητες μεταβλητές του προβλήματος και είναι οι εξής:

1. Κέρδη προ τόκων και φόρων / Σύνολο ενεργητικού
2. Καθαρά κέρδη / Ίδια κεφάλαια
3. Πωλήσεις / Σύνολο ενεργητικού
4. Μικτά κέρδη / Σύνολο ενεργητικού
5. Καθαρά κέρδη / Κεφάλαιο κίνησης
6. Σύνολο υποχρεώσεων / Σύνολο ενεργητικού
7. Μακροπρόθεσμες υποχρεώσεις / (Μακροπρόθεσμες υποχρεώσεις + Ίδια κεφάλαια)
8. (Κυκλοφορούν ενεργητικό-Αποθέματα) / Βραχυπρόθεσμες υποχρεώσεις
9. Διαθέσιμα / Βραχυπρόθεσμες υποχρεώσεις
10. Βραχυπρόθεσμες υποχρεώσεις / Ίδια κεφάλαια
11. Σύνολο υποχρεώσεων / Κεφάλαιο κίνησης

Οι επιδόσεις των επιχειρήσεων στις 11 μεταβλητές (χρηματοοικονομικούς δείκτες) αποτελούν ποσοτικά χαρακτηριστικά της κάθε επιχείρησης και είναι συνεχής τιμές. Μια συνήθης τακτική, όταν αναπτύσσεται ένα σύστημα εκτίμησης πιστωτικού κινδύνου, είναι η ομαδοποίηση των πιθανών τιμών που πηγάζουν από τις μεταβλητές απόφασης-πρόβλεψης σε διαστήματα (Hand και Adams, 2000). Η τεχνική αυτή ονομάζεται διακριτοποίηση.

Μέσω κάποιων διαδικασιών η κάθε μεταβλητή χωρίζεται σε κάποιες κατηγορίες. Έτσι λοιπόν αντί να εισαχθούν στο σύστημα οι συνεχείς τιμές της υπό εξέταση επιχείρησης για κάθε μεταβλητή, εισάγονται οι αριθμοί που υποδηλώνουν την κατηγορία στην οποία ανήκει (για κάθε μεταβλητή) η υπό εξέταση επιχείρηση (Σχήμα 2.1).

	Δείκτης X	
	Διάστημα 1ο [-∞ , 0.1]	Διάστημα 2ο [0.1 , +∞]
Επιχείρηση A σκορ 0.05	1	
Επιχείρηση B σκορ 0.2		2

Σχήμα 2.1: Παράδειγμα διακριτοποίησης

Επομένως, οι διαδικασίες που αναπτύχθηκαν και εφαρμόστηκαν στη συγκεκριμένη μελέτη είναι διαδικασίες διακριτοποίησης ποσοτικών χαρακτηριστικών ή αλλιώς εύρεσης βέλτιστων διαστημάτων εντός των μεταβλητών. Το ερώτημα όμως που εγείρεται είναι πόσες κατηγορίες-διαστήματα πρέπει να αναπτυχθούν εντός των μεταβλητών ώστε το σύστημα να παρέχει ορθότερες εκτιμήσεις.

Οι εξεταζόμενες διαδικασίες στο συγκεκριμένο πείραμα είναι τρεις, παρουσιάζονται στην παράγραφο 2.3 και αποσκοπούν στη δημιουργία τριών, τεσσάρων και πέντε διαστημάτων εντός της κάθε μεταβλητής.

2.2 Ανάπτυξη πειράματος

Η διαδικασία που ακολουθήθηκε για την ανάπτυξη του πειράματος είναι οι εξής (Σχήμα 2.2). Αρχικά τα δεδομένα χωρίστηκαν στο δείγμα εκπαίδευσης που περιείχε 705 επιχειρήσεις και στο δείγμα ελέγχου που περιείχε 706 επιχειρήσεις. Η ομαδοποίηση των επιχειρήσεων στα δύο δείγματα έγινε τυχαία και εντός αυτών υπήρχαν συνεπείς και ασυνεπείς επιχειρήσεις. Το δείγμα εκπαίδευσης χρησιμοποιήθηκε για την ανάπτυξη των διαδικασιών διακριτοποίησης των ποσοτικών χαρακτηριστικών των επιχειρήσεων και το δείγμα εκπαίδευσης για την εξέτασή τους.

Δηλαδή χρησιμοποιώντας το δείγμα εκπαίδευσης οι διαδικασίες παρήγαγαν τα βέλτιστα σημεία διαχωρισμού-διαστήματα για την κάθε μεταβλητή. Έπειτα αυτά εφαρμόστηκαν στο δείγμα ελέγχου και έγινε η διακριτοποίηση των τιμών του. Στη συνέχεια μέσω της μεθόδου της λογιστικής παλινδρόμησης (logistic regression, βλ. § 2.4) εντοπίστηκε η συναρτησιακή μορφή που συνδέει την εξαρτημένη μεταβλητή

(δεδομένη ταξινόμηση) με τις ανεξάρτητες μεταβλητές (χρηματοοικονομικούς δείκτες). Σε αυτό το σημείο, στις επιχειρήσεις προσδόθηκαν κάποια βάρη, με τέτοιο τρόπο ούτως ώστε εντός των δειγμάτων να υπάρχει ίδιος αριθμός συνεπών και ασυνεπών επιχειρήσεων. Η συναρτησιακή σχέση αυτή εφαρμόστηκε στο δείγμα ελέγχου, του οποίου τα δεδομένα είχαν διακριτοποιηθεί, και παρήχθησαν εκτιμήσεις για τη δανειοληπτική ικανότητα των επιχειρήσεων. Οι εκτιμήσεις αυτές συγκρίθηκαν με τη δεδομένη ταξινόμηση των επιχειρήσεων και εξήχθησαν χρήσιμα συμπεράσματα για την αποτελεσματικότητα των εξεταζόμενων διαδικασιών.



Σχήμα 2.2: Βήματα ανάπτυξης πειράματος

Η όλη διαδικασία-πείραμα, δηλαδή ο διαχωρισμός των 1411 επιχειρήσεων στο δείγμα εκπαίδευσης και ελέγχου και η επεξεργασία των δειγμάτων αυτών μέσω των παραπάνω βημάτων, πραγματοποιήθηκε πέντε φορές. Αυτό συνέβη διότι μέσω της επανάληψης τους πειράματος είναι πιο ασφαλή και σαφή τα συμπεράσματα που εξάγονται όσον αφορά τη σύγκριση των μεθόδων διακριτοποίησης που εξετάστηκαν.

2.3 Εξεταζόμενες διαδικασίες διακριτοποίησης

2.3.1 Η διαδικασία των ποσοστιμόριων (percentiles cut-points)

Η τεχνική αυτή χωρίζει-ομαδοποιεί τις τιμές μιας συνεχούς μεταβλητής σε c διαστήματα με τέτοιο τρόπο ώστε $\frac{100}{c}\%$ των τιμών της συνεχούς μεταβλητής να ανήκει σε κάθε διάστημα. Ο διαχωρισμός γίνεται με την εύρεση της αντίστοιχης τιμής-εκατοστημορίου. Δηλαδή το x εκατοστημόριο (x percentile) είναι η τιμή πάνω από την οποία ανήκει το $(1-x)\%$ των υπό εξέταση περιπτώσεων και κάτω από αυτή το $x\%$.

Η διαδικασία είναι απλή διότι δε λαμβάνει υπόψη της τη σχέση των ανεξάρτητων μεταβλητών και της εξαρτημένης μεταβλητής. Αυτό μπορεί να οδηγήσει στην απώλεια σημαντικών πληροφοριών για ένα μοντέλο ταξινόμησης το οποίο χρησιμοποιεί τη συγκεκριμένη διαδικασία διακριτοποίησης. Παρά το γεγονός αυτό, τα μοντέλα αυτά είναι άμεσα υλοποιήσιμα και σε προηγούμενες έρευνες έχουν εξάγει ικανοποιητικά αποτελέσματα (Hand και Adams, 2000).

2.3.2 Η απόσταση Kolmogorov-Smirnoff

Η απόσταση των Kolmogorov-Smirnoff (K-S distance) χρησιμοποιήθηκε για πρώτη φορά το 1977 από τον Friedman για την κατασκευή δένδρων απόφασης. Η απόσταση είναι μια θεμελιώδεις έννοια που σκοπός της είναι να μετρήσει πόσο απέχουν δύο παρατηρήσεις, να ποσοτικοποιηθεί δηλαδή η διαφορά τους. Ένα δένδρο απόφασης που διέπεται από την τεχνική αυτή μπορεί να χρησιμοποιηθεί είτε ως τεχνική ταξινόμησης είτε ως διαδικασία διακριτοποίησης ποσοτικών χαρακτηριστικών.

Για διακριτοποίηση χαρακτηριστικών σε δύο διαστήματα η τεχνική λειτουργεί ως εξής (Utgoff και Clouse, 1996): για να χωριστούν οι τιμές μιας συνεχούς μεταβλητής x , στη προκειμένη μελέτη ενός χρηματοοικονομικού δείκτη, σε δύο διαστήματα, χρειάζεται να βρεθεί μια τιμή της μεταβλητής αυτής, έστω a , η οποία αποτελεί την τιμή διαχωρισμού. Το ένα διάστημα διακριτοποίησης θα αποτελείται από τις τιμές του $x < a$ και το άλλο από τις τιμές του $x > a$. Δηλαδή η τιμή a αποτελεί το πάνω όριο τιμών του ενός διαστήματος, $x \in [-\infty, a]$, και το κάτω του άλλου, $x \in [a, +\infty]$. Το ερώτημα όμως που προκύπτει είναι πώς βρίσκεται η τιμή αυτή.

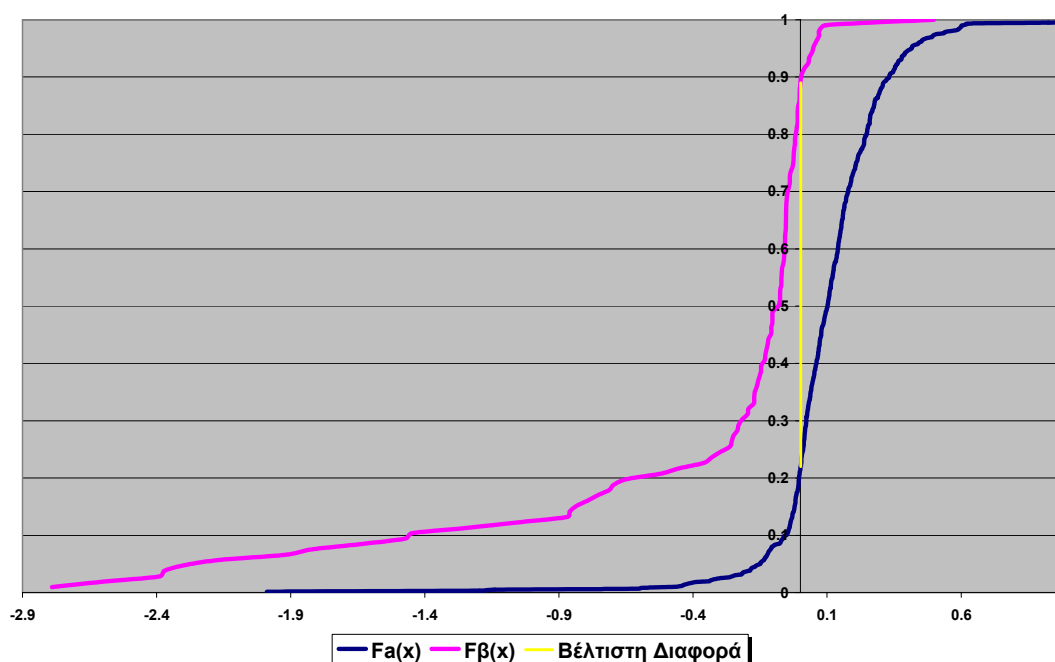
Έστω ότι είναι γνωστές οι συναρτήσεις κατανομής που χαρακτηρίζουν τις κατηγορίες των συνεπών και ασυνεπών επιχειρήσεων και αφορούν τη συνεχή

μεταβλητή x , $F_A(x)$ και $F_B(x)$ αντίστοιχα. Η βέλτιστη τιμή διαχωρισμού a μεγιστοποιεί την διαφορά $|F_A(a) - F_B(a)|$. Η εύρεση των συναρτήσεων κατανομών των δύο κατηγοριών των εταιριών γίνεται με εμπειρικό τρόπο.

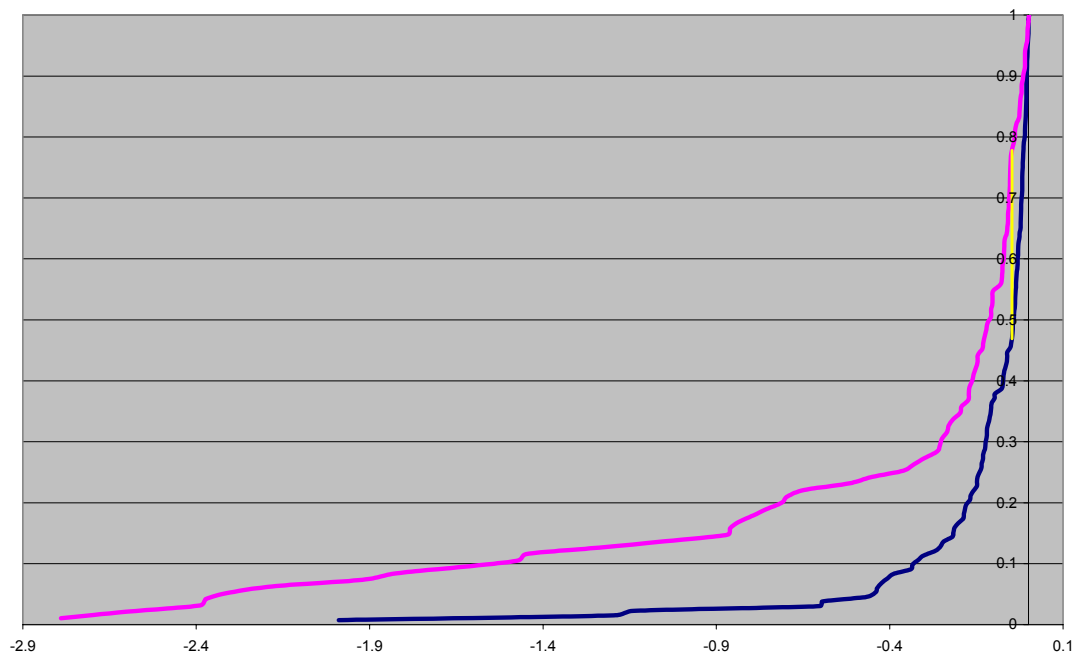
Ένα σχηματικό παράδειγμα της διαδικασίας αυτής δίνεται στο Σχήμα 2.3. Για τις δύο κατηγορίες των εταιριών τοποθετούνται στον άξονα τον x οι τιμές τους για τη συγκεκριμένη συνεχή μεταβλητή (είναι γνωστές από το δείγμα εκπαίδευσης), από τη μικρότερη στη μεγαλύτερη, και σχηματίζουν δύο κλιμακωτές συναρτήσεις που είναι οι προσεγγιστικές συναρτήσεις κατανομών.

Αν όμως το ερώτημα που τίθεται είναι η διακριτοποίηση ενός συνεχούς χαρακτηριστικού σε πάνω από δύο διαστήματα, τότε χρειάζεται να βρεθούν καινούργιες τιμές διαχωρισμού. Αφού βρεθεί η βέλτιστη τιμή διαχωρισμού a , που χωρίζει τις τιμές σε δύο διαστήματα, εφαρμόζεται πάλι η μέθοδος σε κάθε διάστημα ξεχωριστά και προκύπτουν επιπλέον σημεία διαχωρισμού (βλ. Σχήματα 2.4, 2.5).

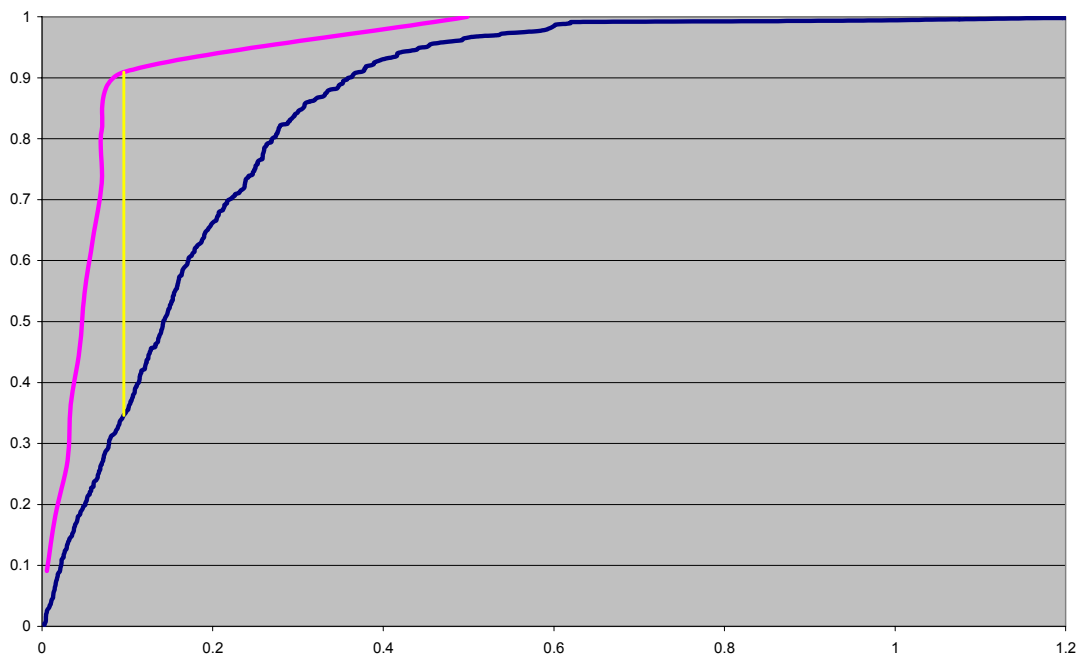
Η εύρεση των τιμών διαχωρισμού των συνεχών μεταβλητών (των έντεκα χρηματοοικονομικών δεικτών) χρησιμοποιώντας τη μέθοδο K-S πραγματοποιήθηκε στο λογισμικό Matlab.



Σχήμα 2.3: Εύρεση βέλτιστου σημείου διαχωρισμού



Σχήμα 2.4: Εύρεση επιπλέον σημείου διαχωρισμού από διάστημα $x < a$



Σχήμα 2.5: Εύρεση επιπλέον σημείου διαχωρισμού από διάστημα $x > a$

2.3.3 Μέθοδος CART. Δένδρα ταξινόμησης και παλινδρόμησης

Η μέθοδος CART (Δένδρα Ταξινόμησης και Παλινδρόμησης - Classification and Regression Trees) είναι μια μη παραμετρική τεχνική που αναπτύχθηκε για την ανάλυση προβλημάτων ταξινόμησης. Δηλαδή δύναται να παρουσιάσει την σχέση που συνδέει τις μεταβλητές απόφασης, για τον καθορισμό ενός αποτελέσματος ή μιας εξαρτημένης μεταβλητής. Αν η εξαρτημένη μεταβλητή είναι συνεχής τότε η μέθοδος CART παράγει ένα δένδρο παλινδρόμησης ενώ αν είναι διακριτή ένα δένδρο ταξινόμησης (Yohannes και Webb, 1999). Και στις δύο περιπτώσεις ο σκοπός της μεθόδου είναι να αποκαλύψει την δομή του προβλήματος εκτίμησης-ταξινόμησης. Οι πρώτες έρευνες που πραγματοποιήθηκαν για την ανάλυση χρηματοοικονομικών προβλημάτων ταξινόμησης χρησιμοποιώντας τη μέθοδο αυτή έγιναν από τους Breiman et al. το 1984 και Frydman et al. το 1985.

Η μέθοδος οδηγεί στην ανάπτυξη ενός δυαδικού δένδρου ταξινόμησης με βάση ένα δείγμα επιχειρήσεων, αξιοποιώντας τα χαρακτηριστικά τους, την πραγματική τους ταξινόμηση, τις αρχικές πιθανότητες κατάταξης και το κόστος κάθε μορφής εσφαλμένης ταξινόμησης (σφάλμα τύπου I και II, βλ. § 2.6). Το δένδρο αυτό αποτελείται από κόμβους σε κάθε έναν από τους οποίους αντιστοιχεί ένας κανόνας απόφασης. Οι κανόνες αυτοί έχουν συνήθως μονομεταβλητή μορφή. Πρόκειται, δηλαδή, για ένα χρηματοοικονομικό δείκτη και μία τιμή – όριο γι' αυτόν (Σχήμα 2.6).

Οι κανόνες είναι ενσωματωμένοι στον αλγόριθμο που χρησιμοποιεί το δένδρο και προσδιορίζονται μέσω κάποιων τεχνικών, οι οποίες εφαρμόζονται ώστε να παρέχει το δένδρο ακρίβεια στις εκτιμήσεις που εξάγει. Κυρίως επιλέγονται με κριτήριο την ελαχιστοποίηση του κόστους από τις εσφαλμένες ταξινομήσεις (Ζοπουνίδης, 1998).

$$R(t) = [C((0/1)) + C((1/0))] \pi_0 \pi_1 \frac{1}{p(t)} \frac{n_0(t)}{N_0} \frac{n_1(t)}{N_1},$$

όπου:

N_0, N_1 : ο συνολικός αριθμός επιχειρήσεων σε κάθε ομάδα (συνεπών – μη συνεπών) επιχειρήσεων.

$n_0(t), n_1(t)$: ο αριθμός των επιχειρήσεων της κάθε ομάδας που απομένουν στον κόμβο t.

$C(1/0)$: το κόστος εσφαλμένης ταξινόμησης μιας επιχείρησης, στην ομάδα 1 ενώ στην πραγματικότητα ανήκει στην ομάδα 0.

$C(0/1)$: το κόστος εσφαλμένης ταξινόμησης μιας επιχείρησης, στην ομάδα 0 ενώ στην πραγματικότητα ανήκει στην ομάδα 1.

π_0, π_1 : η *a priori* πιθανότητα για μια επιχείρηση να ανήκει στην ομάδα 0 ή στην ομάδα 1 αντίστοιχα.

$p(t)$: η πιθανότητα ταξινόμησης μιας επιχείρησης στον κόμβο t.

Σχήμα 2.6: Παράδειγμα δένδρου ταξινόμησης CART

Το δένδρο εκπαιδεύεται μέσω του δείγματος εκπαίδευσης και ελέγχεται για την αξιοπιστία του μέσω των δειγμάτων ελέγχου. Ο αλγόριθμος που χρησιμοποιείται για την ανάπτυξη δένδρων ταξινόμησης και παλινδρόμησης είναι ο αλγόριθμος της αναδρομικής διαφοροποίησης (recursive partitioning algorithm). Αρχικά ο αλγόριθμος χρησιμοποιεί κάποιες τεχνικές διαχωρισμού των μεταβλητών. Οι πιο σημαντικοί από αυτούς είναι το κριτήριο του Gini, ο δυαδικός κανόνας (twoing rule) και οι διαχωρισμοί γραμμικού συνδυασμού (linear combination splits) [βλ. Breiman et al. (1984), Yohannes και Webb (1999)]. Επιπλέον μπορούν να χρησιμοποιηθούν και άλλοι κανόνες διαχωρισμού ανάλογα με τις ανάγκες του εκάστοτε χρήστη. Έπειτα εφαρμόζεται σε κάθε κόμβο του δένδρου η συνάρτηση (impurity function) που ελέγχει κατά πόσον οι τιμές διαχωρισμού των κριτηρίων (κανόνες) είναι οι βέλτιστες, δηλαδή δημιουργούν ένα αξιόπιστο δένδρο ταξινόμησης. Επειδή ο τρόπος με τον οποίον μετράται η αξιοπιστία ποικίλει, η μορφή της συνάρτησης μπορεί να διαφέρει. Η συνάρτηση συνήθως έχει τη μορφή:

$$\Delta_i(s, t) = i(t) - p_L[i(t_L)] - p_R[i(t_R)] ,$$

όπου:

s = μια συγκεκριμένη τιμή διαχωρισμού.

p_L = η αναλογία των επιχειρήσεων στον κόμβο t που πηγαίνουν στο αριστερό ‘κλαδί’ του δένδρου, t_L .

p_R = η αναλογία των επιχειρήσεων στον κόμβο t που πηγαίνουν στο δεξί ‘κλαδί’ του δένδρου, t_R .

$i(t_L)$ = λανθασμένες εκτιμήσεις στο αριστερό ‘κλαδί’.

$i(t_R)$ = λανθασμένες εκτιμήσεις στο δεξί ‘κλαδί’.

Ο αλγόριθμος επιλέγει τη βέλτιστη τιμή διαχωρισμού για κάθε μεταβλητή. Η βέλτιστη τιμή είναι αυτή που ελαχιστοποιεί την παραπάνω συνάρτηση. Στη συνέχεια για κάθε τελικό κόμβο ορίζεται και η κατηγορία την οποία αντιπροσωπεύει σύμφωνα με τη συνάρτηση κόστος των λανθασμένων εκτιμήσεων.

Ο αλγόριθμος αναδρομικής διαφοροποίησης παρουσιάζεται ως ένας συνδυασμός μονομεταβλητής και πολυμεταβλητής προσέγγισης, παρουσιάζοντας χαρακτηριστικά και των δύο. Από τη μία, η κατάταξη των επιχειρήσεων γίνεται βάση ενός μοναδικού χαρακτηριστικού. Από την άλλη, όμως, τα δένδρα που δημιουργεί η μέθοδος είναι μια αλληλουχία κόμβων που είναι ουσιαστικά μια αλληλουχία χαρακτηριστικών για την ταξινόμηση των επιχειρήσεων.

Επειδή ο Α.Α.Δ. είναι μια επαναληπτική μέθοδος, δηλαδή συνεχίζει η διαδικασία στους κόμβους που δεν είναι τελικοί, ενδεχόμενο σημαντικό πρόβλημα είναι η εμφάνιση υπερπροσαρμογής (overfitting) του δένδρου στα δεδομένα. Η συνεχόμενη διαφοροποίηση είναι δυνατόν να οδηγήσει σε δένδρο στο οποίο κάθε τελικός κόμβος κατατάσσει-αντιπροσωπεύει μόνο μια επιχείρηση. Τα μεγάλα δένδρα παρουσιάζουν δύο μειονεκτήματα: 1) παρόλο που είναι υψηλής ακρίβειας, παρέχουν μη ικανοποιητικά αποτελέσματα όταν σε αυτά εισέρχονται καινούργια δεδομένα-επιχειρήσεις για ταξινόμηση (Steinberg και Colla, 1995) και 2) η διαδικασία που ακολουθείται είναι περίπλοκη και δυσνόητη. Η περιπλοκότητα του δένδρου ορίζεται από το πλήθος των τελικών του κόμβων.

Για να αποφευχθεί η δημιουργία περίπλοκων και μεγάλων δένδρων αφαιρούνται συνδεδεμένοι κόμβοι (pruning-‘κλάδεμα’). Αυτό γίνεται μέσω της τεχνικής cross validation. Η τεχνική χρησιμοποιώντας το δείγμα εκπαίδευσης παράγει δέκα καινούργια δένδρα που χαρακτηρίζονται από κάποια κόστη. Τα δέκα δένδρα κατατάσσονται σύμφωνα με τα κόστη αυτά και με τον αριθμό των τελικών τους κόμβων. Η τεχνική βρίσκει το δένδρο με το μικρότερο κόστος (minimum cost-tree) και το βέλτιστο δένδρο (optimum cost tree). Το βέλτιστο δένδρο είναι αυτό που το κόστος που το χαρακτηρίζει είναι το αμέσως μεγαλύτερο από το ελάχιστο κόστος, αλλά έχει μικρότερο αριθμό τελικών κόμβων από το δένδρο με το ελάχιστο κόστος (Breiman et al., 1984).

Στη συγκεκριμένη μελέτη η τεχνική CART χρησιμοποιήθηκε για τη διακριτοποίηση των χρηματοοικονομικών δεικτών. Εισάγοντας τις τιμές του δείκτη και την ταξινόμηση (από τις επιχειρήσεις του δείγματος εκπαίδευσης) η μέθοδος παρήγαγε για κάθε δείκτη τα βέλτιστα σημεία διαχωρισμού. Χρησιμοποιήθηκαν δύο δένδρα ταξινόμησης. Ένα μεγάλο δένδρο και ένα δένδρο που είχε προκύψει από την εφαρμογή της μεθόδου cross-validation. Ο αλγόριθμος αναπτύχθηκε στο λογισμικό Matlab.

2.4 Λογιστική παλινδρόμηση

Η λογιστική παλινδρόμηση (logistic regression) στην ουσία είναι γενίκευση της απλής γραμμικής παλινδρόμησης. Η βασική υπόθεση της λογιστικής προσέγγισης είναι ότι ο λογάριθμος των συναρτήσεων πυκνότητας πιθανότητας των δύο κατηγοριών είναι γραμμική συνάρτηση.

$$Y_i \sim \ln[p_i / (1-p_i)] = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$$

Οι κατηγορίες αφορούν την εξαρτημένη μεταβλητή Y η οποία είναι δυαδική ($Y=1$, αν η επιχείρηση είναι συνεπής και $Y=2$ αν είναι ασυνεπής). Η πιθανότητα η υπό εξέταση επιχείρηση-παρατήρηση να ανήκει στην κατηγορία των ασυνεπών ($Y=2$) δίδεται ως εξής:

$$p_i = \frac{e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}}} ,$$

όπου $\beta_0, \dots, \beta_p$ οι συντελεστές των μεταβλητών $X_{1i}, \dots, X_{pi}$. Αν λάβουμε $p_i \geq 0.5$ η υπό εξέταση επιχείρηση κατατάσσεται στην κατηγορία $Y=2$, αλλιώς εντάσσεται στην κατηγορία $Y=1$.

Η λογιστική παλινδρόμηση όμως εκτός από τεχνική ταξινόμησης μπορεί να χρησιμοποιηθεί και ως τεχνική μοντελοποίησης της σχέσεως μεταξύ της εξαρτημένης μεταβλητής και των ανεξάρτητων μεταβλητών. Δηλαδή μπορεί να παρέχει μια εκτίμηση του διανύσματος $[\beta_0, \beta_1, \dots, \beta_p]$ αν είναι γνωστές οι τιμές της εξαρτημένης μεταβλητής και των χαρακτηριστικών-μεταβλητών.

Στο πείραμα που αναπτύχθηκε η λογιστική παλινδρόμηση εφαρμόστηκε πρώτα στο δείγμα εκπαίδευσης για να βρεθεί το διάνυσμα των συντελεστών των μεταβλητών. Έπειτα γνωρίζοντας τις τιμές των συντελεστών των μεταβλητών (χρηματοοικονομικών δεικτών) εφαρμόστηκε στο δείγμα ελέγχου για να παράγει εκτιμήσεις για τις επιχειρήσεις που ανήκαν στο δείγμα αυτό. Το λογισμικό που χρησιμοποιήθηκε είναι το στατιστικό πακέτο SPSS.

2.5 Ανάλυση αποτελεσμάτων

Στην παράγραφο αυτή παρουσιάζονται και αναλύονται τα αποτελέσματα που εξήγαγαν οι διαδικασίες διακριτοποίησης. Τα αποτελέσματα είναι ο μέσος όρος των πεδίων ορισμού των υποδιαστημάτων που παρήγαγαν οι διαδικασίες (για τις πέντε φορές που πραγματοποιήθηκε το πείραμα) και η ανάλυση τους έγκειται στον έλεγχο της ορθότητας τους.

Όπως προαναφέρθηκε κάθε διαδικασία χρησιμοποιήθηκε για να χωρίσει τα χαρακτηριστικά σε τρία, τέσσερα και πέντε υποδιαστήματα. Για κάθε υποδιάστημα παρουσιάζεται και το πλήθος των συνεπών και ασυνεπών επιχειρήσεων (του δείγματος εκπαίδευσης) που ανήκουν σε αυτό.

Για να ελεγχθεί η ορθότητα των διαστημάτων στα οποία χωρίστηκε κάθε χαρακτηριστικό προσδιορίστηκε η ισχύς κάθε μεταβλητής (characteristic strength) μετά τη διακριτοποίησή της. Η ισχύς αυτή παρέχει μια εκτίμηση για τη δυνατότητα των κατασκευασμένων διαστημάτων να εξάγουν σωστές και ασφαλές προβλέψεις αν ενσωματωθούν από ένα σύστημα εκτίμησης πιστωτικού κινδύνου (χρησιμοποιηθούν για τη διακριτοποίηση των χαρακτηριστικών-κριτηρίων πρόβλεψης).

Ο προσδιορισμός της ισχύος γίνεται ως εξής:

$$\sum_i \left[\frac{(\% \Sigma_i - \% A \Sigma_i) * \left[\ln \left(\frac{\% \Sigma_i}{\% A \Sigma_i} \right) * 100 \right]}{10} \right]$$

όπου Σ_i = Συνεπείς επιχειρήσεις στο i υποδιάστημα και $A \Sigma_i$ = Ασυνεπείς στο i υποδιάστημα.

Μεταβλητές με ισχύ (σε απόλυτη τιμή) από 0 έως 30 κρίνονται ότι έχουν χαμηλή δυνατότητα πρόβλεψης (weak predictor). Αν η ισχύς κυμαίνεται (σε απόλυτη τιμή) από 30 έως 100 τότε το χαρακτηριστικό κρίνεται ότι έχει μέτρια δυνατότητα πρόβλεψης (moderate predictor) και αν είναι μεγαλύτερη του 100 τότε το χαρακτηριστικό έχει υψηλή προβλεπτική ικανότητα (strong predictor).

Στο πίνακα 2.1 που ακολουθεί, παρουσιάζεται ο αριθμός των χρηματοοικονομικών δεικτών που ανήκει σε κάθε κατηγορία ισχύος, για τις διαδικασίες διακριτοποίησης που εφαρμόστηκε. Παρατηρείται ότι η διαδικασία των ποσοστιμοριών και η προσέγγιση των Kolmogorov-Smirnov παρήγαγαν δείκτες με ικανοποιητική δυνατότητα πρόβλεψης. Αντιθέτως για τη διαδικασία CART, για όλες τις περιπτώσεις δημιουργίας υποδιαστημάτων, παρατηρείται ότι τουλάχιστον ένας δείκτης χαρακτηρίζεται από χαμηλή προβλεπτική ικανότητα.

Πίνακας 2.1: Προβλεπτική ισχύς δεικτών

Διαδικασία διακριτοποίησης	δείκτες με ισχυρή προβλεπτική ικανότητα	δείκτες με μέτρια προβλεπτική ικανότητα	δείκτες με χαμηλή προβλεπτική ικανότητα
percentiles (3 υποδ.)	2	9	-
percentiles (4 υποδ.)	3	8	-
percentiles (5 υποδ.)	6	5	-
K-S (3 υποδ.)	1	9	1
K-S (4 υποδ.)	1	9	-
K-S (5 υποδ.)	3	8	-
CART (3 υποδ.)	-	2	9
CART (4 υποδ.)	2	8	1
CART (5 υποδ.)	3	7	1
CART2	1	6	4

Άλλο ένα ‘κριτήριο’ που χρησιμοποιήθηκε για τον έλεγχο της ορθότητας των υποδιαστημάτων είναι και αυτό της μονοτονίας, υπολογίζοντας τον συντελεστή ποσοστό ασυνεπών / (1- ποσοστό ασυνεπών) για κάθε υποδιάστημα.

Επίσης σημαντικό γεγονός για την ορθότητα ενός υποδιαστήματος είναι και το εύρος του πεδίου ορισμού του. Ένα υποδιάστημα με στενό πεδίο ορισμού στην ουσία δεν υφίσταται. Δηλαδή αν η συγκεκριμένη διακριτοποίηση εφαρμοστεί σε νέα δεδομένα (επιχειρήσεις), στο υποδιάστημα αυτό θα μπορούν να ανήκουν λίγες επιχειρήσεις λόγω του μικρού του εύρους.

Διαδικασία Ποσοστιμόριων (percentiles)

Οι μέσοι όροι των πεδίων ορισμού (από τις πέντε φορές που πραγματοποιήθηκε το πείραμα) των υποδιαστημάτων που εξήγαγε η διαδικασία του ποσοστιαίου διαχωρισμού παρουσιάζονται στους Πίνακες 2.2, 2.3 και 2.4. Σχετικά με τη δημιουργία τριών υποδιαστημάτων παρατηρείται ότι το εύρος του πεδίου ορισμού του εβδόμου δείκτη για το δεύτερο υποδιάστημα είναι πολύ μικρό [0,0.041]. Επίσης όσον αφορά τη μονοτονία αυτή αλλάζει στον έβδομο δείκτη (1^ο υποδ. 0.23, 2^ο υποδ. 0.07, 3^ο υποδ. 0.19) και στον ενδέκατο δείκτη (0.04, 0.28, 0.23).

Πίνακας 2.2: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία percentiles (για 3 υποδιαστήματα)

Δείκτες	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				Ισχύς
	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	
1 ^{ος}	$[-\infty, 0.009]$	138	97	235	$[0.009, 0.141]$	226.8	8.2	235	$[0.141, +\infty]$	233.6	1.4	235	82.06
2 ^{ος}	$[-\infty, 0.001]$	129	106	235	$[0.001, 0.218]$	234	1	235	$[0.218, +\infty]$	235	0	235	$+\infty$
3 ^{ος}	$[-\infty, 0.683]$	158	77	235	$[0.683, 1.78]$	213	22	235	$[1.78, +\infty]$	227	8	235	52.11
4 ^{ος}	$[-\infty, 0.082]$	147.8	87.2	235	$[0.082, 0.358]$	220	15	235	$[0.358, +\infty]$	230.6	4.4	235	62.90
5 ^{ος}	$[-\infty, -0.05]$	139.2	95.8	235	$[-0.05, 0.275]$	224.2	10.8	235	$[0.275, +\infty]$	235	0	235	$+\infty$
6 ^{ος}	$[-\infty, 0.211]$	223	12	235	$[0.211, 0.662]$	219	16	235	$[0.662, +\infty]$	156	79	235	51.07
7 ^{ος}	$[-\infty, 0]$	253	57	310	$[0, 0.041]$	149	11	160	$[0.041, +\infty]$	197	38	235	43.03
8 ^{ος}	$[-\infty, 0.927]$	150.2	84.8	235	$[0.927, 2.84]$	224.4	10.6	235	$[2.84, +\infty]$	223.8	11.2	235	56.46
9 ^{ος}	$[-\infty, 0.125]$	162	73	235	$[0.125, 0.778]$	215.4	19.6	235	$[0.778, +\infty]$	221	14	235	47.29
10 ^{ος}	$[-\infty, 0.204]$	221	14	235	$[0.204, 1.372]$	207	28	235	$[1.372, +\infty]$	171	64	235	44.02
11 ^{ος}	$[-\infty, 0.473]$	225	10	235	$[0.473, 2.071]$	183	52	235	$[2.071, +\infty]$	191	44	235	44.68

Όσον αφορά την εφαρμογή της διαδικασίας για τη διακριτοποίηση των χαρακτηριστικών σε τέσσερα υποδιαστήματα, στενό πεδίου ορισμού παρατηρείται για τον έβδομο δείκτη στο δεύτερο υποδιάστημα [0,0.005]. Επίσης η μονοτονία αλλάζει στον έβδομο (0.23, 0.15, 0.05, 0.25) και στον ενδέκατο δείκτη (0.06, 0.07, 0.36, 0.22).

Τέλος, σχετικά με τη δημιουργία πέντε υποδιαστημάτων το εύρος του πεδίου ορισμού του δεύτερου υποδιαστήματος για τον πρώτο έβδομο δείκτη είναι πολύ μικρό και η μονοτονία μεταβάλλεται στον έβδομο, δέκατο και ενδέκατο δείκτη.

Πίνακας 2.3: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία percentiles (για 4 υποδιαστήματα)

	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				Υποδιάστημα 4				
Δείκτες	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Ισχύς
1 ^{ος}	$[-\infty, -0.019]$	90.2	86.8	177	$[-0.019, 0.068]$	159.8	16.2	176	$[0.068, 0.182]$	173.6	2.4	176	$[0.182, +\infty]$	174.8	1.2	176	109.46
2 ^{ος}	$[-\infty, -0.044]$	79.2	97.8	177	$[-0.044, 0.078]$	167.2	8.8	176	$[0.078, 0.344]$	176	0	176	$[0.344, +\infty]$	176	0	176	$+\infty$
3 ^{ος}	$[-\infty, 0.357]$	112	65	177	$[0.357, 1.21]$	145.8	30.2	176	$[1.21, 2.059]$	172.4	3.6	176	$[2.059, +\infty]$	168.2	7.8	176	76.88
4 ^{ος}	$[-\infty, 0.016]$	100.6	76.4	177	$[0.016, 0.212]$	153	23	176	$[0.212, 0.447]$	171.8	4.2	176	$[0.447, +\infty]$	173	3	176	88.88
5 ^{ος}	$[-\infty, -0.193]$	92.2	84.8	177	$[-0.193, 0.071]$	154.2	21.8	176	$[0.071, 0.434]$	176	0	176	$[0.434, +\infty]$	176	0	176	$+\infty$
6 ^{ος}	$[-\infty, 0.133]$	167.2	9.8	177	$[0.133, 0.42]$	168.8	7.2	176	$[0.42, 0.77]$	158	18	176	$[0.77, +\infty]$	104.4	71.6	176	72.17
7 ^{ος}	$[-\infty, 0]$	253	57.4	310.4	$[0, 0.005]$	37.2	5.4	42.6	$[0.005, 0.084]$	167.6	8.4	176	$[0.084, +\infty]$	140.6	35.4	176	59.07
8 ^{ος}	$[-\infty, 0.792]$	96.4	80.6	177	$[0.792, 1.516]$	166.8	9.2	176	$[1.516, 4.411]$	168.8	7.2	176	$[4.411, +\infty]$	166.4	9.6	176	80.49
9 ^{ος}	$[-\infty, 0.074]$	114.4	62.6	177	$[0.074, 0.3]$	152	24	176	$[0.3, 1.272]$	167.6	8.4	176	$[1.272, +\infty]$	164.4	11.6	176	65.28
10 ^{ος}	$[-\infty, 0.112]$	166.2	10.8	177	$[0.112, 0.572]$	166.8	9.2	176	$[0.572, 2.206]$	136.4	39.6	176	$[2.206, +\infty]$	122.4	53.6	176	59.98
11 ^{ος}	$[-\infty, 0.284]$	166.4	10.6	177	$[0.284, 1.096]$	164.4	11.6	176	$[1.096, 3.16]$	129	47	176	$[3.16, +\infty]$	143.8	32.2	176	61.45

Πίνακας 2.4: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία percentiles (για 5 υποδιαστήματα)

	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				Υποδιάστημα 4				Υποδιάστημα 5				
Δείκτες	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Ισχύς
1 ^{ος}	$[-\infty, -0.041]$	62.6	78.4	141	$[-0.041, 0.031]$	121	20	141	$[0.031, 0.111]$	134.2	6.8	141	$[0.111, 0.219]$	140.8	0.2	141	$[0.219, +\infty]$	139.8	1.2	141	152.24
2 ^{ος}	$[-\infty, -0.104]$	51.6	89.6	141.2	$[-0.104, 0.02]$	123.8	17	140.8	$[0.02, 0.154]$	141	0	141	$[0.154, 0.443]$	141	0	141	$[0.443, +\infty]$	141	0	141	$+\infty$
3 ^{ος}	$[-\infty, 0.172]$	85.4	55.6	141	$[0.172, 0.863]$	110.6	30.4	141	$[0.863, 1.563]$	128.4	12.6	141	$[1.563, 2.325]$	139.8	1.2	141	$[2.325, +\infty]$	134.2	6.8	141	101.03
4 ^{ος}	$[-\infty, 0.001]$	82	61.8	143.8	$[0.001, 0.14]$	106.2	32	138.2	$[0.14, 0.29]$	133	8	141	$[0.29, 0.529]$	138.8	2.2	141	$[0.529, +\infty]$	138.4	2.6	141	110.19
5 ^{ος}	$[-\infty, -0.381]$	72.6	68.4	141	$[-0.381, 0.002]$	104	37.6	141.6	$[0.002, 0.187]$	139.8	0.6	140.4	$[0.187, 0.563]$	141	0	141	$[0.563, +\infty]$	141	0	141	$+\infty$
6 ^{ος}	$[-\infty, 0.081]$	133.6	7.4	141	$[0.081, 0.283]$	134.8	6.2	141	$[0.283, 0.553]$	132.4	8.6	141	$[0.553, 0.848]$	123.8	17.2	141	$[0.848, +\infty]$	73.8	67.2	141	92.95
7 ^{ος}	$[-\infty, 0]$	253	57.4	310.4	$[0, 0.021]$	103.6	9	112.6	$[0.021, 0.126]$	135.4	5.6	141	$[0.126, +\infty]$	106.4	34.6	141	-	-	-	-	64.92
8 ^{ος}	$[-\infty, 0.627]$	64.2	76.8	141	$[0.627, 1.137]$	130.6	10.4	141	$[1.137, 2.125]$	135.6	5.4	141	$[2.125, 6.253]$	134.8	6.2	141	$[6.253, +\infty]$	133.2	7.8	141	104.82
9 ^{ος}	$[-\infty, 0.045]$	85.8	55.2	141	$[0.045, 0.184]$	118.2	22.8	141	$[0.184, 0.491]$	127.4	13.6	141	$[0.491, 1.981]$	135.8	5.2	141	$[1.981, +\infty]$	131.2	9.8	141	82.70
10 ^{ος}	$[-\infty, 0.069]$	132.4	8.6	141	$[0.069, 0.302]$	133.4	7.6	141	$[0.302, 1]$	130.2	10.8	141	$[1, 2.934]$	97.4	43.6	141	$[2.934, +\infty]$	105	36	141	78.96
11 ^{ος}	$[-\infty, 0.184]$	134.8	6.2	141	$[0.184, 0.69]$	136.2	4.8	141	$[0.69, 1.555]$	106.2	34.8	141	$[1.555, 4.348]$	104.8	36.2	141	$[4.348, +\infty]$	116.4	24.6	141	80.20

Διαδικασία Kolmogorov-Smirnoff (K-S)

Οι μέσοι όροι των πεδίων ορισμού των υποδιαστημάτων που εξήγαγε η διαδικασία των Kolmogorov-Smirnoff παρατίθενται στους Πίνακες 2.5, 2.6 και 2.7. Σχετικά με τη δημιουργία τριών υποδιαστημάτων παρατηρείται ότι τα πεδία ορισμού του δεύτερου, εβδόμου και ένατου δείκτη για το δεύτερο υποδιάστημα είναι πολύ μικρά. Επίσης όσον αφορά τη μονοτονία αυτή μεταβάλλεται στο δέκατο και ενδέκατο δείκτη, (0.07, 0.48, 0.30) και (0.04, 0.62, 0.23) αντίστοιχα.

Πίνακας 2.5: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία K-S (για 3 υποδιαστήματα)

Δείκτες	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				AS
	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	
1 ^{ος}	$[-\infty, 0.001]$	121.4	96.4	217.8	$[0.001, +\infty]$	477	10.2	487.2	-	-	-	-	39.14
2 ^{ος}	$[-\infty, -0.067]$	74.8	96	170.8	$[-0.067, -0.003]$	428	10.4	438.4	$[-0.003, +\infty]$	478	1	479	97.16
3 ^{ος}	$[-\infty, 0.615]$	153.2	76.2	229.4	$[0.615, 0.904]$	285.6	24.2	309.8	$[0.904, +\infty]$	399	15.5	414.5	53.22
4 ^{ος}	$[-\infty, 0.079]$	145.4	89.6	235	$[0.079, +\infty]$	453	17	470	-	-	-	-	35.29
5 ^{ος}	$[-\infty, -0.183]$	94.4	87.6	182	$[-0.183, 0.002]$	81.4	19	100.4	$[0.002, +\infty]$	422.6	0	422.6	$+\infty$
6 ^{ος}	$[-\infty, 0.806]$	508.4	35	543.4	$[0.806, 1.123]$	82	42.6	124.6	$[1.123, +\infty]$	13.3	48.4	61.7	32.69
7 ^{ος}	$[-\infty, 0.176]$	440	70.2	510.2	$[0.176, 0.206]$	129.6	25.2	154.8	$[0.206, +\infty]$	72	28	100	28.50
8 ^{ος}	$[-\infty, 0.283]$	19.4	50	69.4	$[0.283, 0.602]$	40.2	27.2	67.4	$[0.602, +\infty]$	538.8	29.4	568.2	31.00
9 ^{ος}	$[-\infty, 0.055]$	86.6	52.8	139.4	$[0.055, 0.105]$	239.4	34.2	273.6	$[0.105, +\infty]$	454	32.7	486.7	38.58
10 ^{ος}	$[-\infty, 1.022]$	400.2	26.6	426.8	$[1.022, 2.149]$	112.8	54.6	167.4	$[2.149, +\infty]$	142.3	42.4	184.7	32.82
11 ^{ος}	$[-\infty, 1.034]$	325	13	338	$[1.034, 2.072]$	80.4	49.6	130	$[2.072, +\infty]$	193	44	237	40.15

Όσον αφορά την εφαρμογή της διαδικασίας για τη διακριτοποίηση των χαρακτηριστικών σε τέσσερα υποδιαστήματα, στενό πεδίου ορισμού παρατηρείται για το δεύτερο και πέμπτο δείκτη στο τρίτο υποδιάστημα, $[-0.027, 0.002]$ και $[-0.006, 0.005]$ αντίστοιχα. Επίσης η μονοτονία αλλάζει στο δεύτερο, έβδομο και ενδέκατο δείκτη.

Τέλος, σχετικά με τη δημιουργία πέντε υποδιαστημάτων στενά πεδία ορισμού παρατηρούνται στον πρώτο δείκτη (3^ο υποδ.), πέμπτο δείκτη (3^ο υποδ.), και στον ένατο (2^ο και 4^ο υποδ.) και η μονοτονία μεταβάλλεται στον έβδομο, δέκατο και ενδέκατο δείκτη.

Διαδικασία CART (Δένδρα Παλινδρόμησης και Ταξινόμησης)

Οι μέσοι όροι των πεδίων ορισμού των υποδιαστημάτων που εξήγαγε η διαδικασία CART παρατίθενται στους Πίνακες 2.8, 2.9, 2.10 και 2.11. Σχετικά με τη δημιουργία τριών υποδιαστημάτων δεν έχει νόημα να εξεταστεί η μονοτονία και το εύρος των υποδιαστημάτων αφού όλοι οι δείκτες (εκτός του έβδομου) χωρίστηκαν μέσω της μεθόδου σε δύο υποδιαστήματα.

Πίνακας 2.8: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία CART (για 3 υποδιαστήματα)

Δείκτες	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				AS
	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	
1 ^{ος}	$[-\infty, 0.001]$	118.4	100.4	218.8	$[0.001, +\infty]$	460.8	25.4	486.2	-	-	-	-	26.09
2 ^{ος}	$[-\infty, -0.018]$	99.4	108.4	207.8	$[-0.018, +\infty]$	479.8	17.4	497.2	-	-	-	-	30.88
3 ^{ος}	$[-\infty, 0.963]$	204.6	95.4	300	$[0.963, +\infty]$	374.6	30.4	405	-	-	-	-	24.12
4 ^{ος}	$[-\infty, 0.081]$	143.2	94.2	237.4	$[0.081, +\infty]$	436	31.6	467.6	-	-	-	-	23.56
5 ^{ος}	$[-\infty, 0.001]$	167.8	111.8	279.6	$[0.001, +\infty]$	411.4	14	425.4	-	-	-	-	32.39
6 ^{ος}	$[-\infty, 0.922]$	542	64.2	606.2	$[0.922, +\infty]$	37.2	61.6	98.8	-	-	-	-	18.06
7 ^{ος}	$[-\infty, 0.002]$	262.4	68.8	331.2	$[0.002, 0.554]$	297.4	32.4	329.8	$[0.554, +\infty]$	11	33	44	29.02
8 ^{ος}	$[-\infty, 0.601]$	58.6	78	136.6	$[0.601, +\infty]$	520.6	47.8	568.4	-	-	-	-	20.27
9 ^{ος}	$[-\infty, 0.097]$	132.6	73.4	206	$[0.097, +\infty]$	446.6	52.4	499	-	-	-	-	18.63
10 ^{ος}	$[-\infty, 1.011]$	383.6	41.2	424.8	$[1.011, +\infty]$	195.6	84.6	280.2	-	-	-	-	21.30
11 ^{ος}	$[-\infty, 1.032]$	313.4	24	337.4	$[1.032, +\infty]$	265.8	101.8	367.6	-	-	-	-	26.32

Όσον αφορά την εφαρμογή της διαδικασίας για τη διακριτοποίηση των χαρακτηριστικών σε τέσσερα υποδιαστήματα, στενά πεδία ορισμού παρατηρούνται για τον 1^ο, 2^ο, 3^ο, 5^ο, και 6^ο δείκτη στο δεύτερο υποδιάστημα και για τον 1^ο, 3^ο, 5^ο, 6^ο, 7^ο και 8^ο δείκτη για το τρίτο υποδιάστημα. Επίσης η μονοτονία μεταβάλλεται σε όλους τους δείκτες εκτός του τετάρτου και του ένατου, αφού σε αυτούς δε γίνεται να μεταβληθεί διότι χωρίζονται μόνο σε δύο υποδιαστήματα.

Τέλος, σχετικά με τη δημιουργία πέντε υποδιαστημάτων στενά πεδία ορισμού παρατηρούνται σε όλους τους δείκτες εκτός του 4^{ου}, 7^{ου}, 9^{ου} και 11^{ου} και η μονοτονία μεταβάλλεται σε όλους τους δείκτες (εκτός 4^{ου} και 9^{ου} αφού χωρίζονται σε δύο υποδιαστήματα).

Πίνακας 2.9: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία CART (για 4 υποδιαστήματα)

	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				Υποδιάστημα 4				
Δείκτες	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	AS
1 ^{ος}	$[-\infty, 0.001]$	122.2	96.6	218.8	$[0.001, 0.07]$	338	8.4	346.4	$[0.07, 0.071]$	2	1.5	3.5	$[0.071, +\infty]$	343.5	2.5	346	84.36
2 ^{ος}	$[-\infty, -0.092]$	90.6	99.4	190	$[-0.092, -0.066]$	300.6	1.6	302.2	$[-0.066, -0.009]$	29.5	11.5	41	$[-0.009, +\infty]$	488.5	2.5	491	108.20
3 ^{ος}	$[-\infty, 0.963]$	209	91	300	$[0.963, 1.042]$	315.4	12.8	328.2	$[1.042, 1.086]$	8	3	11	$[1.086, +\infty]$	362	11	373	70.15
4 ^{ος}	$[-\infty, 0.081]$	147.4	90	237.4	$[0.081, +\infty]$	451	16.6	467.6	-	-	-	-	-	-	-	-	31.87
5 ^{ος}	$[-\infty, -0.104]$	116.6	94.6	211.2	$[-0.104, -0.046]$	117.8	3.6	121.4	$[-0.046, 0]$	27.25	10.25	37.5	$[0, +\infty]$	427.75	0.25	428	111.82
6 ^{ος}	$[-\infty, 0.878]$	543	41.6	584.6	$[0.878, 0.879]$	43.2	52.4	95.6	$[0.879, 0.943]$	31	1	32	$[0.943, +\infty]$	30	62	92	56.94
7 ^{ος}	$[-\infty, 0.002]$	269.2	62	331.2	$[0.002, 0.389]$	290.8	16.2	307	$[0.389, 0.681]$	34.2	20.6	54.8	$[0.681, +\infty]$	10.5	19.5	30	38.13
8 ^{ος}	$[-\infty, 0.601]$	59.4	77.2	136.6	$[0.601, 0.956]$	449	24.4	473.4	$[0.956, 0.964]$	1	2	3	$[0.964, +\infty]$	449	23	472	55.59
9 ^{ος}	$[-\infty, 0.097]$	221.8	79	300.8	$[0.097, +\infty]$	376.6	27.6	404.2	-	-	-	-	-	-	-	-	27.47
10 ^{ος}	$[-\infty, 0.006]$	326.8	21	347.8	$[0.006, 0.8]$	160.8	64.8	225.6	$[0.8, 1.059]$	354	21	375	$[1.059, +\infty]$	200	83	283	56.72
11 ^{ος}	$[-\infty, 0.825]$	262.6	10.6	273.2	$[0.825, 1.647]$	108.8	52.8	161.6	$[1.647, 4.021]$	129.75	14.75	144.5	$[4.021, +\infty]$	154	39.25	193.25	57.53

Πίνακας 2.10: Μέσες τιμές υποδιαστημάτων των χρηματοοικονομικών δεικτών που προέκυψαν από τη διαδικασία CART (για 5 υποδιαστήματα)

	Υποδιάστημα 1				Υποδιάστημα 2				Υποδιάστημα 3				Υποδιάστημα 4				Υποδιάστημα 5				
	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	Εύρος	Σ	ΑΣ	Σύν.	AS
1 ^{ος}	$[-\infty, 0.001]$	122.2	96.6	218.8	$[0.001, 0.07]$	338	8.4	346.4	$[0.07, 0.071]$	2	1.5	3.5	$[0.071, +\infty]$	343.5	2.5	346	-	-	-	-	84.36
2 ^{ος}	$[-\infty, -0.069]$	86.6	98.8	185.4	$[-0.069, -0.062]$	209.2	2.2	211.4	$[-0.062, -0.006]$	20.7	7.6	28.3	$[-0.006, +\infty]$	483.7	1.6	485.3	-	-	-	-	105.85
3 ^{ος}	$[-\infty, 0.963]$	209	91	300	$[0.963, 1.042]$	315.4	12.8	328.2	$[1.042, 1.086]$	8	3	11	$[1.086, +\infty]$	362	11	373	-	-	-	-	70.15
4 ^{ος}	$[-\infty, 0.081]$	147.4	90	237.4	$[0.081, +\infty]$	451	16.6	467.6	-	-	-	-	-	-	-	-	-	-	-	-	31.87
5 ^{ος}	$[-\infty, -0.104]$	116.6	94.6	211.2	$[-0.104, -0.046]$	117.8	3.6	121.4	$[-0.046, 0]$	27.3	10.2	37.5	$[0, +\infty]$	427.75	0.25	428	-	-	-	-	37.52
6 ^{ος}	$[-\infty, 0.533]$	399.8	28.8	428.6	$[0.533, 0.621]$	43.8	29.2	73	$[0.621, 0.622]$	44.3	9.4	53.7	$[0.622, 0.91]$	186.3	30.4	216.7	$[0.91, +\infty]$	41	62	103	47.54
7 ^{ος}	$[-\infty, 0.002]$	269.2	62	331.2	$[0.002, 0.389]$	290.8	16.2	307	$[0.389, 0.597]$	31.4	20.6	52	$[0.597, 0.681]$	7	0	7	$[0.681, +\infty]$	10.5	19.5	30	$+\infty$
8 ^{ος}	$[-\infty, 0.601]$	59.4	77.2	136.6	$[0.601, 0.921]$	353.6	19.8	373.4	$[0.921, 30.747]$	222.5	10.5	233	$[30.747, 85.191]$	226	13	239	$[85.19, +\infty]$	30	1	31	111.16
9 ^{ος}	$[-\infty, 0.097]$	135.4	70.6	206	$[0.097, +\infty]$	463	36	499	-	-	-	-	-	-	-	-	-	-	-	-	23.91
10 ^{ος}	$[-\infty, 0.552]$	245	15.6	260.6	$[0.552, 0.595]$	93	40	133	$[0.595, 1.236]$	139	10.3	149.3	$[1.236, 20.189]$	231.3	47.4	278.7	$[20.189, +\infty]$	95.5	41	136.5	50.04
11 ^{ος}	$[-\infty, 0.825]$	262.6	10.6	273.2	$[0.825, 1.647]$	108.8	52.8	161.6	$[1.647, 3.149]$	109.3	8.7	118	$[3.149, 5.984]$	149	42.5	191.5	$[5.984, +\infty]$	102	11	113	78.53

Όσον αφορά τη διακριτοποίηση των δεικτών μέσω της μεθόδου CART, χρησιμοποιώντας το βέλτιστο δένδρο με τους λιγότερους τελικούς κόμβους (το δένδρο που προέκυψε έπειτα από την εφαρμογή της τεχνικής cross-validation), επίσης δεν έχει νόημα να εξεταστεί η μονοτονία και το εύρος των υποδιαστημάτων αφού όλοι οι δείκτες (εκτός του πέμπτου και του έβδομου) χωρίστηκαν μέσω της μεθόδου σε δύο υποδιαστήματα.

Γενικότερα, για να διαχωρίσει η μέθοδος τους δείκτες σε τέσσερα και πέντε υποδιαστήματα ανέπτυξε μεγάλα δένδρα. Παρατηρείται ότι τα πεδία ορισμού των υποδιαστημάτων που παρήγαγαν αυτά τα δένδρα είναι πολύ μικρά. Γίνεται λοιπόν αντιληπτό, όπως προαναφέρθηκε και στην παράγραφο 2.3.3, ότι όταν αυτά τα διαστήματα χρησιμοποιηθούν για την ταξινόμηση νέων επιχειρήσεων θα εξάγουν μη ικανοποιητικά αποτελέσματα. Επίσης η μονοτονία μεταβάλλεται μεταξύ των υποδιαστημάτων σε αρκετούς δείκτες.

Το γεγονός ότι για μερικούς δείκτες η διαδικασία παρήγαγε λιγότερα υποδιαστήματα από ότι αναμενόταν οφείλεται στον τρόπο με τον οποίο αναπτύσσεται ένα δένδρο ταξινόμησης (στον αλγόριθμο του).

2.6 Αξιολόγηση των διαδικασιών

Στην ενότητα αυτή παρουσιάζονται τα αποτελέσματα της αξιολόγησης των διαδικασιών διακριτοποίησης που αναπτύχθηκαν. Η αξιολόγηση έγινε μέσω δύο μεθόδων που χρησιμοποιούνται ευρέως από τους ερευνητές για την ανάλυση των αποτελεσμάτων που εξάγουν τα συστήματα εκτίμησης του πιστωτικού κινδύνου.

Όπως προαναφέρθηκε, τα διαστήματα που προέκυψαν από τις μεθόδους διακριτοποίησης που αναπτύχθηκαν, εφαρμόστηκαν στο δείγμα ελέγχου. Έπειτα χρησιμοποιώντας την λογιστική παλινδρόμηση παρήχθησαν οι εκτιμήσεις για τις επιχειρήσεις του δείγματος αυτού. Δηλαδή, μέσω της λογιστικής παλινδρόμησης, κατασκευάστηκε ένα σύστημα εκτίμησης πιστωτικού κινδύνου που χρησιμοποιώντας τις προαναφερθείσες διαδικασίες διακριτοποίησης παρήγαγε κάποιες εκτιμήσεις.

Οι μέθοδοι αξιολόγησης συγκρίνουν τις εκτιμήσεις αυτές με τη δεδομένη. Άρα λοιπόν, στην ουσία στη συγκεκριμένη μελέτη χρησιμοποιούνται για να αξιολογήσουν τις διαδικασίες διακριτοποίησης που εφαρμόστηκαν.

2.6.1 Τύποι σφαλμάτων

Η μέθοδος αυτή συγκρίνει την εκτιμώμενη ταξινόμηση των επιχειρήσεων με τη δεδομένη. Ο πρώτος τύπος σφάλματος δίνει το ποσοστό των λανθασμένων εκτιμήσεων που αφορούν τις ασυνεπείς επιχειρήσεις. Δηλαδή πόσες από τις επιχειρήσεις που ήταν ασυνεπείς (από τη δεδομένη ταξινόμηση) τελικά εκτιμήθηκαν ως συνεπείς.

$$\text{Σφάλμα I} = \frac{\text{λαθος εκτιμωμενες ασυνεπεις}}{\text{συνολο ασυνεπων}}$$

Ο δεύτερος τύπος δίνει το ποσοστό των λανθασμένων εκτιμήσεων που αφορούν τις συνεπείς επιχειρήσεις. Δηλαδή πόσες από τις επιχειρήσεις που ήταν συνεπείς (από τη δεδομένη ταξινόμηση) τελικά εκτιμήθηκαν ως ασυνεπείς (Schumann και Liu, 2002).

$$\text{Σφάλμα II} = \frac{\text{λαθος εκτιμωμενες συνεπεις}}{\text{συνολο συνεπων}}$$

Ο πρώτος τύπος σφάλματος θεωρείται πιο σημαντικός. Αυτό συμβαίνει διότι για ένα πιστωτικό ίδρυμα έχει μεγαλύτερο κόστος η εκτίμηση μιας ασυνεπούς επιχείρησης ως συνεπής παρά μιας συνεπούς ως ασυνεπής. Στην πρώτη περίπτωση η

τράπεζα εκτίθεται στον πιστωτικό κίνδυνο ενώ στην δεύτερη απλά χάνει ένα συνεπή πελάτη.

Αν ένα πιστωτικό ίδρυμα παρουσιάζει υψηλές τιμές στο σφάλμα τύπου I σημαίνει ότι ακολουθεί μια ‘γενναιόδωρη’ πολιτική στις χορηγήσεις δανείων με αποτέλεσμα να διατρέχει μεγαλύτερο κίνδυνο να απολέσει μέρος των χορηγούμενων κεφαλαίων. Αν έχει υψηλή τιμή στο σφάλμα τύπου II σημαίνει ότι ακολουθεί μια αυστηρή πολιτική σε σχέση με τις χορηγήσεις δανείων. Αυτό έχει ως αποτέλεσμα να χάνει πελάτες και να μειώνεται η εμπορικότητά της (το μερίδιό της στην τραπεζική αγορά).

Στον πίνακα 2.12 που ακολουθεί παρουσιάζονται οι μέσοι όροι των σφαλμάτων εκτίμησης, που εξήχθησαν από τις πέντε φορές που έλαβε χώρα η διαδικασία, για κάθε μια από τις μεθόδους διακριτοποίησης. Στο Παράρτημα παρατίθεται ο πίνακας που παρουσιάζει τα σφάλματα που χαρακτηρίζουν κάθε διαδικασία διακριτοποίησης για κάθε επανάληψη του πειράματος.

Πίνακας 2.12: Μέσα σφάλματα διαδικασιών διακριτοποίησης

Διαδικασία διακριτοποίησης	Μέσο Σφάλμα I	Μέσο Σφάλμα II	Μέσος Όρος
percentiles (3 υποδιαστήματα)	6.82%	9.14%	7.98%
percentiles (4 υποδιαστήματα)	10.58%	5.22%	7.90%
percentiles (5 υποδιαστήματα)	10.06%	4.48%	7.27%
K-S (3 υποδιαστήματα)	8.82%	8.04%	8.43%
K-S (4 υποδιαστήματα)	10.90%	4.94%	7.92%
K-S (5 υποδιαστήματα)	12.76%	4.90%	8.83%
CART (3 υποδιαστήματα)	8.44%	10.70%	9.57%
CART (4 υποδιαστήματα)	13.30%	8.10%	10.70%
CART (5 υποδιαστήματα)	15.36%	7.14%	11.25%
CART2	10.70%	8.42%	9.56%

Όσον αφορά τον πρώτο τύπο σφάλματος παρατηρείται, για κάθε διαδικασία διακριτοποίησης, ότι όσο αυξάνεται ο αριθμός των διαστημάτων εντός των χαρακτηριστικών αυξάνει και η τιμή του σφάλματος. Η μικρότερη τιμή του σφάλματος για κάθε διαδικασία παρατηρείται κατά την δημιουργία τριών υποδιαστημάτων εντός των χαρακτηριστικών. Τα πιο ικανοποιητικά αποτελέσματα παρήγαγε η διαδικασία των ποσοστιμορίων (για 3 υποδιαστήματα).

Για το δεύτερο τύπο σφάλματος παρατηρείται ότι πολύ ικανοποιητικά αποτελέσματα παρήγαγαν οι διαδικασίες των ποσοστιμορίων και η K-S (για τέσσερα και πέντε υποδιαστήματα). Επίσης, παρατηρείται ότι ο δεύτερος τύπος σφάλματος λαμβάνει μικρότερες τιμές από τον πρώτο.

Σύμφωνα με το μέσο όρο των δύο σφαλμάτων παρατηρείται ότι η καλύτερη διαδικασία είναι αυτή των ποσοστιμορίων (για τη δημιουργία πέντε και τεσσάρων υποδιαστημάτων) ενώ ακολουθεί η K-S (για τέσσερα υποδιαστήματα).

Γενικότερα καλύτερη διαδικασία διακριτοποίησης ως προς την ορθότητα των εκτιμήσεων που παράγει (σύμφωνα με τους δύο τύπους σφάλματος) κρίνεται αυτή των ποσοστιμορίων, ικανοποιητικές η K-S και η CART2 ενώ η CART σε σχέση με τις υπόλοιπες δεν κρίνεται και τόσο ικανοποιητική.

2.6.2 Καμπύλη R.O.C. (Receiver Operating Characteristics Graph)

Μερικά μοντέλα ταξινόμησης παράγουν ως αποτελέσματα μια βαθμολογία για την κάθε επιχείρηση. Για προβλήματα ταξινόμησης σε δύο κατηγορίες, μια σημαντική γραφική τεχνική που απεικονίζει την αποτελεσματικότητα των συστημάτων ταξινόμησης που εξάγουν μια βαθμολογία των επιχειρήσεων είναι η καμπύλη ROC.

Για την κατασκευή της καμπύλης ROC οι επιχειρήσεις τοποθετούνται σε σειρά σύμφωνα με τις βαθμολογίες τους, από τη μικρότερη στη μεγαλύτερη. Το διάγραμμα ROC αναπαριστάται σε δύο άξονες. Ο άξονας x αναπαριστά το ποσοστό του σφάλματος τύπου II ενώ ο y αναπαριστά τη διαφορά $1 -$ το ποσοστό του σφάλματος τύπου I. Γενικότερα, ο άξονας y δείχνει το ποσοστό των σωστών εκτιμήσεων για την κατηγορία υψηλού κινδύνου (false alarm rate ή $1 - \text{specificity}$) και ο x το ποσοστό των λανθασμένων εκτιμήσεων για την κατηγορία χαμηλού πιστωτικού κινδύνου (hit rate ή recall ή sensitivity), (Fawcett, 2003).

Η κατασκευή της καμπύλης ROC πραγματοποιείται θεωρώντας ότι η ταξινόμηση των επιχειρήσεων γίνεται χρησιμοποιώντας ένα σημείο διαχωρισμού (cut-off point) στη βαθμολογία των επιχειρήσεων. Για διάφορες τιμές του σημείου διαχωρισμού υπολογίζονται τα σφάλματα τύπου I και II και σημειώνονται τα

αντίστοιχα σημεία πάνω στο διάγραμμα που προαναφέρθηκε ώστε να διαμορφωθεί η καμπύλη ROC.

Για ένα τέλειο σύστημα η καμπύλη ROC είναι παράλληλη στον άξονα x και ξεκινάει από το μέγιστο σημείο του άξονα y. Αυτό σημαίνει ότι το σύστημα αυτό ταξινομεί όλες τις επιχειρήσεις σωστά. Ένα σύστημα με καμπύλη την ευθεία γραμμή $y = x$ σημαίνει ότι ταξινομεί τις επιχειρήσεις σωστά κατά 50%. Τα πραγματικά συστήματα εκτίμησης πιστωτικού κινδύνου χαρακτηρίζονται από καμπύλες που βρίσκονται μεταξύ των δύο αυτών περιπτώσεων.

Συγκρίνοντας τις καμπύλες ROC δύο συστημάτων μπορεί να παρατηρηθεί η διαφορά στην αποτελεσματικότητά τους. Για μια δεδομένη τιμή του σφάλματος τύπου I το σύστημα με την καμπύλη που έχει τη μικρότερη τιμή σφάλματος τύπου II θα είναι καλύτερο. Ομοίως για μια δεδομένη τιμή σφάλματος τύπου II η καμπύλη με τη μικρότερη τιμή σφάλματος τύπου I θα απεικονίζει το πιο ακριβές σύστημα. Μόνο όταν δύο καμπύλες ταυτίζονται ή τέμνονται σε ένα σημείο (το οποίο καθορίζεται από το όριο διαχωρισμού) τα συστήματα τα οποία αντιπροσωπεύουν έχουν την ίδια απόδοση.

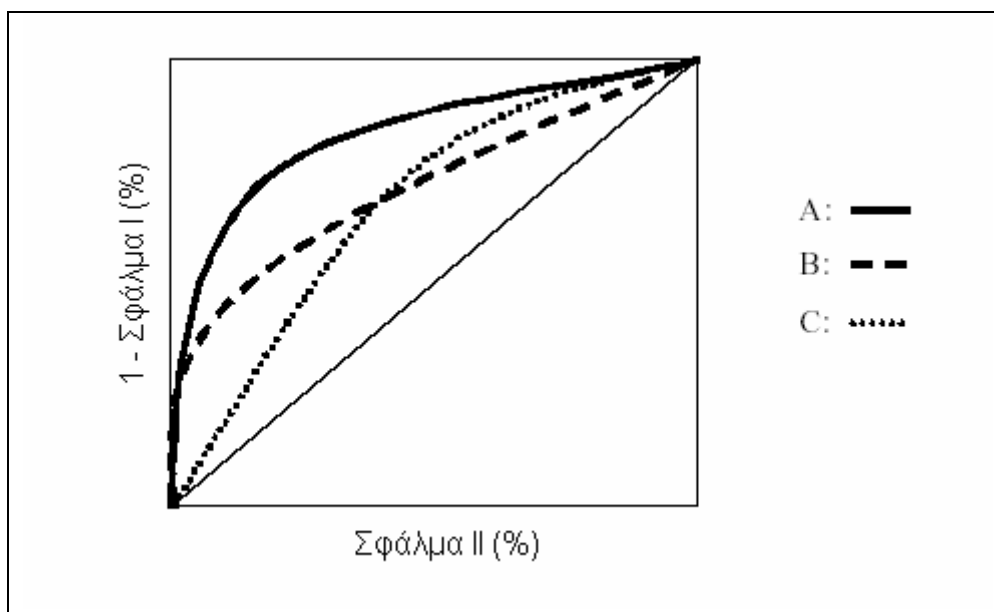
Ένα άλλο κριτήριο σύγκρισης της αποδοτικότητας συστημάτων είναι η περιοχή κάτω από την καμπύλη (Area Under Curve – A.U.C.). Το σύστημα με την καμπύλη που έχει τη μεγαλύτερη περιοχή κάτω από αυτήν είναι και το καλύτερο. Η περιοχή κάτω από την καμπύλη παρέχει ένα ποσοστό, το οποίο μπορεί να κυμαίνεται μεταξύ του 0 και του 1. Όσο μεγαλύτερο το ποσοστό αυτό τόσο πιο αποτελεσματικό είναι το σύστημα ταξινόμησης.

Στο σχήμα 2.7 το σύστημα με την καμπύλη A υπερτερεί των συστημάτων που αντιστοιχούν στις καμπύλες B και C. Σε μερικές περιπτώσεις όμως μπορεί δύο καμπύλες να έχουν ίδιο ποσοστό κάτω από την καμπύλη και η μια καμπύλη να είναι υψηλότερη σε κάποια σημεία και η άλλη σε κάποια άλλα (καμπύλες B και C). Το κριτήριο της περιοχής κάτω από την καμπύλη τότε δεν βοηθάει στην επιλογή του αποτελεσματικότερου συστήματος. Σε αυτή την περίπτωση καλύτερο σύστημα θεωρείται αυτό που η καμπύλη του είναι μεγαλύτερη στο πεδίο που θεωρείται ότι παρέχει καλύτερη λύση σύμφωνα με το υπάρχον πρόβλημα. Αν η καμπύλη B αντιστοιχεί σε ένα αποτελεσματικότερο σύστημα από ότι η καμπύλη C εξαρτάται από το εύρος των τιμών του ορίου διαχωρισμού που επιλέγεται. Η επιλογή του βέλτιστου ορίου γίνεται μέσω της συνάρτησης κόστους [Schumann και Liu (2002) και Ζοπουνίδης (1998)].

$$\text{Αναμενομενη Ζημια} = P_{\sigma} C_I \Sigma_I (\%) + P_{\alpha} C_{II} \Sigma_{II} (\%)$$

όπου P_{σ} , P_{α} είναι οι πιθανότητες μια επιχείρηση να είναι συνεπής / ασυνεπής, C_I , C_{II} είναι τα κόστη σφαλμάτων τύπου I και II και Σ_I , Σ_{II} οι τιμές των σφαλμάτων τύπου I

και II αντίστοιχα. Το βέλτιστο όριο είναι αυτό που ελαχιστοποιεί τη συνάρτηση κόστους, δηλαδή το αναμενόμενη ζημία από λανθασμένες εκτιμήσεις.



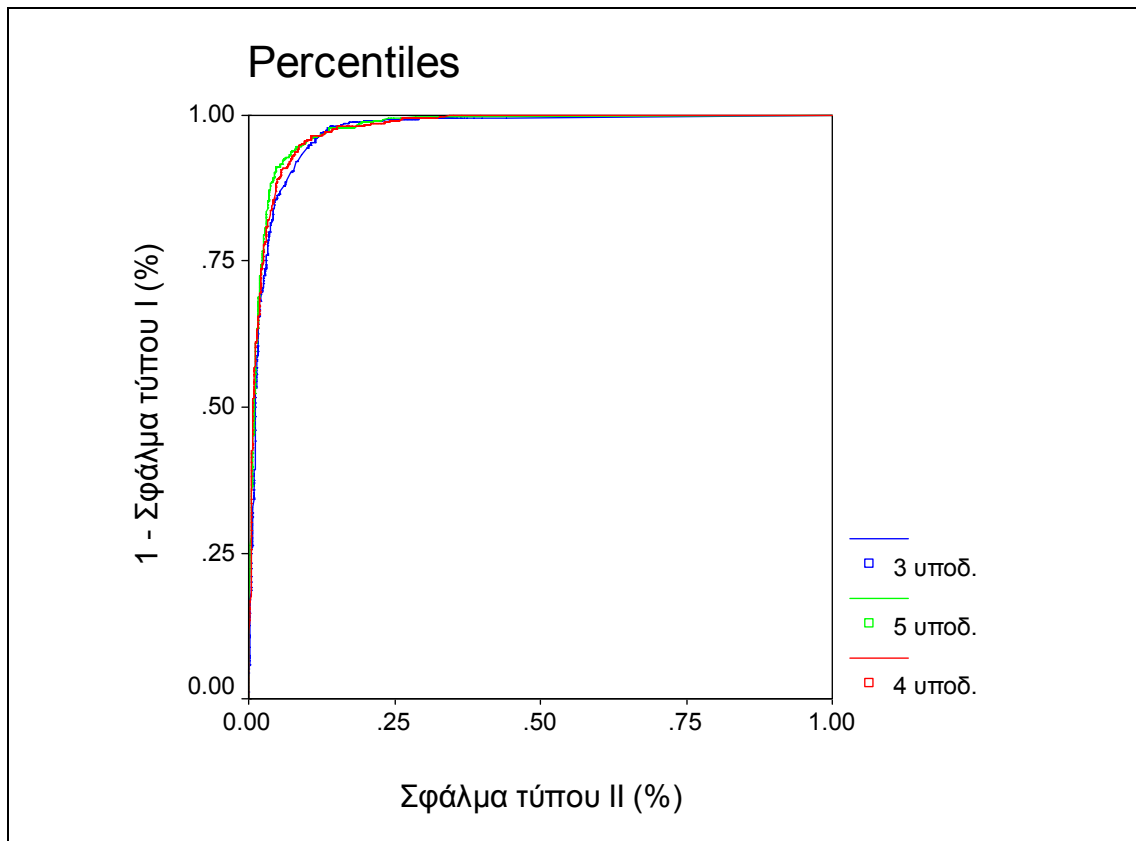
Σχήμα 2.7: Σύγκριση καμπυλών ROC

Παρακάτω παρατίθενται οι καμπύλες ROC για κάθε διαδικασία διακριτοποίησης που αναπτύχθηκε και εφαρμόστηκε στο δείγμα ελέγχου (Σχήματα 2.8, 2.9 και 2.10) καθώς και οι τιμές για την περιοχή κάτω από την καμπύλη των διαδικασιών (Πίνακας 2.13). Για την παραγωγή των καμπύλων χρησιμοποιήθηκαν οι βαθμολογίες που εξήγαγε η λογιστική παλινδρόμηση και από τις πέντε φορές που έλαβε χώρα η πειραματική διαδικασία.

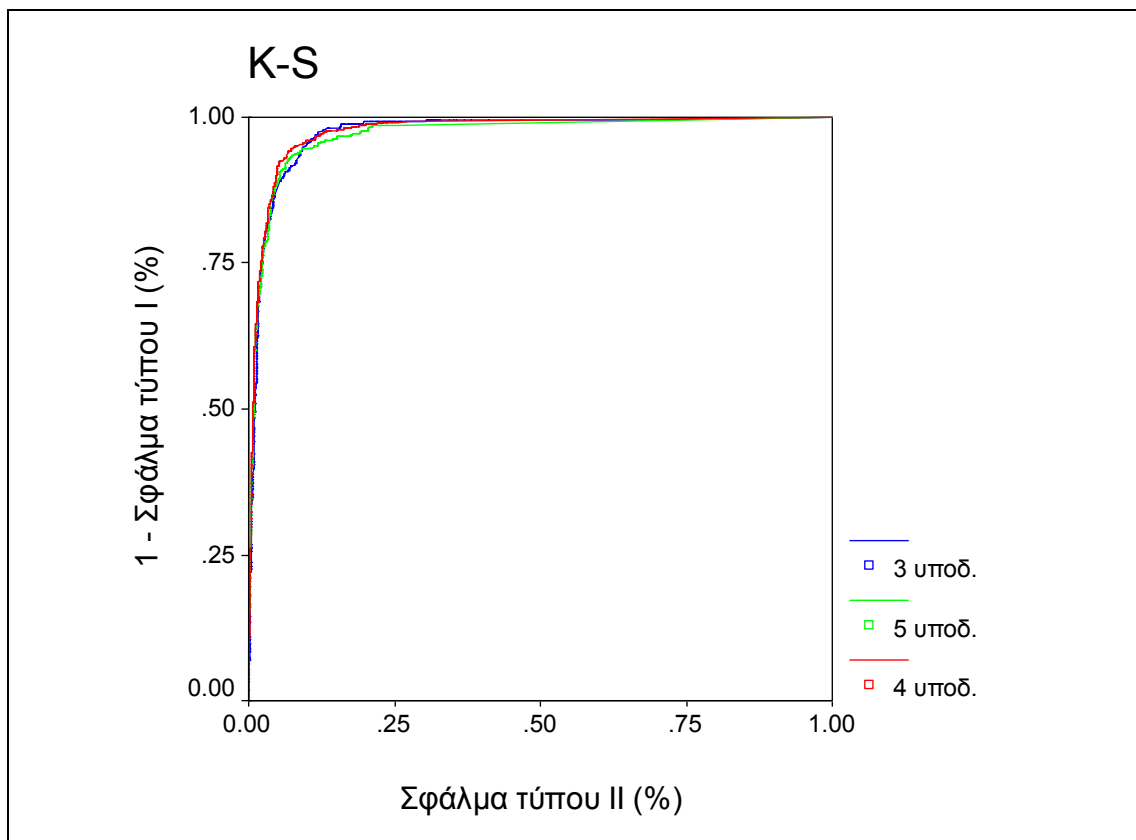
Πίνακας 2.13: Περιοχή κάτω από την καμπύλη

Διαδικασίες Διακριτοποίησης	Μέσο ποσοστό περιοχής κάτω από την καμπύλη
percentiles (3 υποδιαστήματα)	0.971
percentiles (4 υποδιαστήματα)	0.977
percentiles (5 υποδιαστήματα)	0.978
K-S (3 υποδιαστήματα)	0.974
K-S (4 υποδιαστήματα)	0.977
K-S (5 υποδιαστήματα)	0.971
CART (3 υποδιαστήματα)	0.966
CART (4 υποδιαστήματα)	0.957
CART (5 υποδιαστήματα)	0.951
CART2	0.966

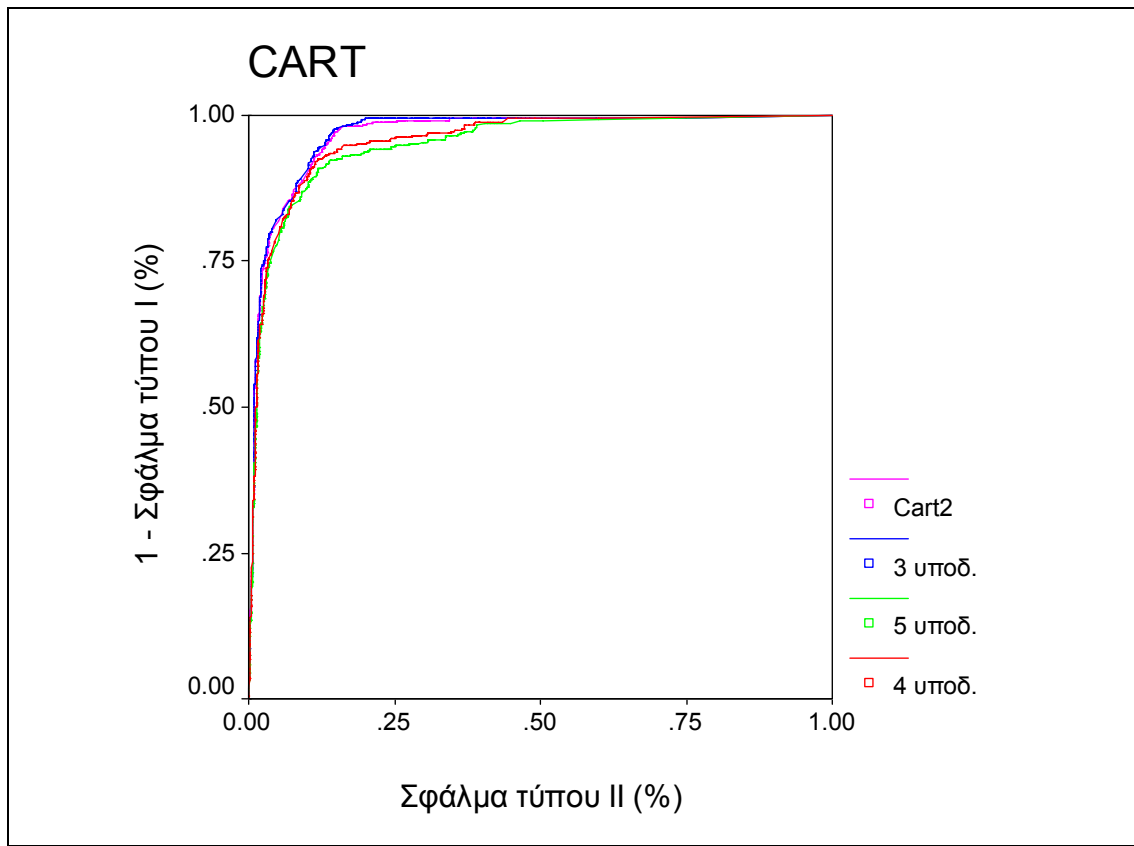
Παρατηρείται ότι όλες οι διαδικασίες παρουσιάζουν πολύ ικανοποιητικές τιμές όσον αφορά το κριτήριο της περιοχής κάτω από την καμπύλη. Καλύτερη από όλες είναι η διαδικασία των ποσοστιμορίων όσον αφορά τη δημιουργία πέντε υποδιαστημάτων εντός των χαρακτηριστικών και ακολουθούν η K-S και η διαδικασία των ποσοστιμορίων για τη δημιουργία τεσσάρων υποδιαστημάτων. Σύμφωνα και με αυτό το κριτήριο η διαδικασία CART υστερεί από τις άλλες παρά το γεγονός ότι και αυτή κρίνεται ικανοποιητική. Στο παράρτημα παρατίθεται το ποσοστό κάτω από την καμπύλη για όλες τις διαδικασίες διακριτοποίησης για κάθε επανάληψη του πειράματος.



Σχήμα 2.8: Καμπύλες ROC της διαδικασίας διακριτοποίησης percentiles



Σχήμα 2.9: Καμπύλες ROC της διαδικασίας διακριτοποίησης K-S



Σχήμα 2.10: Καμπύλες ROC της διαδικασίας διακριτοποίησης CART

ΚΕΦΑΛΑΙΟ 3

ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο τομέας της ανάλυσης και διαχείρισης χρηματοοικονομικών κινδύνων αναπτύσσεται συνεχώς στις μέρες μας. Λόγω των αλλαγών των οικονομικών συνθηκών που επικρατούν παγκοσμίως έχει εξελιχθεί σε ένα από τους πιο σημαντικούς χώρους έρευνας και απασχολεί τόσο τους ακαδημαϊκούς ερευνητές όσο και τους χρηματοπιστωτικούς αναλυτές.

Οι αυξανόμενες ανάγκες των ανθρώπων έχουν οδηγήσει τα πιστωτικά ιδρύματα στην παροχή χρηματοπιστωτικών προϊόντων όπως πιστωτικές κάρτες, δάνεια κάθε μορφής (στεγαστικά, καταναλωτικά, επιχειρηματικά) κ.ά. Η πιστωτική ‘βιομηχανία’ έχει γνωρίσει λοιπόν ταχεία ανάπτυξη και για το λόγο αυτό τα πιστωτικά ιδρύματα ενδιαφέρονται, ολοένα και περισσότερο, για την καλύτερη εκτίμηση των κινδύνων στους οποίους εκτίθενται. Για παράδειγμα το 1980 το κεφάλαιο από ακάλυπτες πιστώσεις ανερχόταν στα 55.1 δισεκατομμύρια δολάρια για τις αμερικάνικες τράπεζες ενώ το 2000 αυξήθηκε στα 633.2 δισεκατομμύρια δολάρια (Nargundkar και Priestley, 2003).

Στην προσπάθεια τους λοιπόν οι τράπεζες να μεγιστοποιήσουν την προσέγγιση συνεπών πελατών (επομένως και την αποδοτικότητά τους) στράφηκαν στην δημιουργία ενός ολοκληρωτικού πλαισίου (πολιτικής) που αφορά την διαχείριση των πελατειακών σχέσεων (Customer Relationship Management) και βοηθάει στην υποστήριξη αποφάσεων σχετικών με την χορήγηση χρηματικών κεφαλαίων πάσης φύσεως. Βασικό μέρος της στρατηγικής αυτής αποτελούν τα συστήματα εκτίμησης του πιστωτικού κινδύνου. Τα συστήματα αυτά εκτός από εργαλεία για την εκτίμηση του κινδύνου στον οποίο εκτίθεται μια τράπεζα είναι συνεπώς και εργαλεία που βοηθούν στην εποπτεία της κεφαλαιακής επάρκειάς της.

Για τους λόγους αυτούς βασική επιδίωξη όλων των εμπλεκομένων είναι ο έλεγχος και η βελτίωση της ακρίβειας των εκτιμήσεων που αυτά παρέχουν με απώτερο σκοπό την ανάπτυξη ‘εύρωστων’ συστημάτων αξιολόγησης (robust rating systems).

Στην κατεύθυνση αυτή κινείται και το νέο πλαίσιο αποδοχής και λειτουργίας των συστημάτων εκτίμησης πιστωτικού κινδύνου από τις τράπεζες που καθόρισε η Επιτροπή Τραπεζικής Επιθεώρησης της Βασιλείας (Basle Committee on Banking Supervision) και έχει ως βάση τις δαπάνες που απορρέουν από τον πιστωτικό κίνδυνο.

Αν μια τράπεζα διαθέτει ένα μεγάλο πλήθος πελατών μια μικρή αύξηση στο βαθμό ακρίβειας τους, μπορεί να παρέχει σημαντική αύξηση των κερδών του πιστωτικού χαρτοφυλακίου της. Άρα λοιπόν εκτός από την κατασκευή ενός

συστήματος εκτίμησης πιστωτικού κινδύνου ένα μεγάλο μέρος της εργασίας των ερευνητών και των αναλυτών σχετίζεται με το πεδίο της αξιολόγησης και βελτίωσης τους. Δηλαδή η ανάγκη να κριθεί η ποιότητα των συστημάτων έχει εξελιχθεί σε ιδιαίτερα σημαντική υπόθεση τα τελευταία χρόνια.

Η βελτίωση ενός συστήματος μπορεί να πραγματοποιηθεί κυρίως με τροποποιήσεις:

1. στις παραμέτρους του. Δηλαδή στην επιλογή των μεταβλητών απόφασης καθώς και στη σχέση που τις διέπει.
2. σε κάποια από τα βασικά χαρακτηριστικά του. Δηλαδή στην τεχνική ταξινόμησης που χρησιμοποιεί το σύστημα ή στις κατηγορίες στις οποίες διαχωρίζεται ο πιστωτικός κίνδυνος.
3. στον τρόπο προετοιμασίας των δεδομένων που θα εισαχθούν στο σύστημα.
4. στη διαδικασία ανάπτυξής του.

Το πείραμα που αναπτύχθηκε και παρουσιάστηκε στη συγκεκριμένη μελέτη αφορούσε στην βελτίωση των συστημάτων εκτίμησης πιστωτικού κινδύνου μέσω της ανάπτυξης και εφαρμογής συγκεκριμένων διαδικασιών προετοιμασίας των δεδομένων που εισέρχονται σε ένα σύστημα. Αναπτύχθηκαν και εφαρμόστηκαν κάποιες διαδικασίες διακριτοποίησης ποσοτικών χαρακτηριστικών, συγκεκριμένα χρηματοοικονομικών δεικτών. Σκοπός της ανάπτυξεως των διαδικασιών αυτών ήταν η σύγκρισή τους και η εύρεση της βέλτιστης εξ αυτών.

Η σύγκριση έγινε μέσω δύο διαδεδομένων τεχνικών που χρησιμοποιούνται σε όλες σχεδόν τις έρευνες για την αξιολόγηση συστημάτων που εκτιμούν τον πιστωτικό κίνδυνο. Της στατιστικής μεθόδου των δύο τύπων σφαλμάτων και της γραφικής τεχνικής των καμπυλών R.O.C.

Οι δύο αυτές τεχνικές εξήγαγαν παρόμοια αποτελέσματα. Δηλαδή ένα σύστημα στο οποίο θα ενσωματωθεί κάποια διαδικασία διακριτοποίησης που αναπτύχθηκε θα παράγει ικανοποιητικά αποτελέσματα. Στη συγκεκριμένη μελέτη το υποθετικό σύστημα που αναπτύχθηκε χρησιμοποιούσε ως τεχνική ταξινόμησης τη λογιστική παλινδρόμηση για να ταξινομήσει τις επιχειρήσεις σε συνεπείς και ασυνεπείς.

Πιο συγκεκριμένα όμως ένα σύστημα που θα χρησιμοποιεί τη μέθοδο των ποσοστιμορίων (percentiles) και ιδιαίτερα για τη δημιουργία πέντε υποδιαστημάτων εντός των μεταβλητών απόφασης θα είναι το πιο ακριβές. Έπειτα ακολουθεί η προσέγγιση των Kolmogorov-Smirnoff και μετά η μέθοδος των δένδρων ταξινόμησης και παλινδρόμησης (CART), κυρίως όταν αναπτύσσονται μικρά δένδρα.

Γενικότερα για τις δύο πρώτες διαδικασίες πιο ικανοποιητικά αποτελέσματα εξήχθησαν όταν αυτές χρησιμοποιήθηκαν για το διαχωρισμό των χαρακτηριστικών σε πέντε ή τέσσερα υποδιαστήματα. Όσον αφορά τη διαδικασία CART παρήγαγε καλύτερα αποτελέσματα όταν εφαρμόστηκε για τη δημιουργία τριών υποδιαστημάτων. Επίσης όλες οι διαδικασίες, όσον αφορά τον πρώτο τύπο

σφάλματος, παρήγαγαν καλύτερα αποτελέσματα όταν δημιουργούν τρία υποδιαστήματα εντός των χαρακτηριστικών.

Σίγουρα, οι αλλαγές στον τρόπο προετοιμασίας των δεδομένων θα δώσει περισσότερη ώθηση στο χώρο των συστημάτων εκτίμησης του πιστωτικού κινδύνου και θα οδηγήσει σε περαιτέρω ανάπτυξη του. Εκτός αυτού όμως στο μέλλον θα προκύψουν ενδιαφέροντα προβλήματα που θα αφορούν την εκτίμηση και διαχείριση των πιστωτικών κινδύνων, τα οποία θα πρέπει να επιλυθούν. Αδιαμφισβήτητα λοιπόν θα υπάρξει πρόοδος, όπως είχε προβλέψει και ο βρετανός Samuel Butler σχεδόν 100 χρόνια λέγοντας ότι «Όλη η πρόοδος είναι βασισμένη σε μια παγκόσμια έμφυτη επιθυμία κάθε οργανισμού να μεγιστοποιεί το κέρδος του».

Παράρτημα

Το θεσμικό πλαίσιο των συστημάτων εκτίμησης πιστωτικού κινδύνου

Η Επιτροπή Τραπεζικής Επιθεώρησης της Βασιλείας έχει προβεί στην έκδοση μιας σειράς οδηγιών που αποσκοπούν στην εποπτεία της κεφαλαιακής επάρκειας των τραπεζών. Δηλαδή εξέδωσε κάποιες διατάξεις που αφορούν την εποπτική αντιμετώπιση των χρηματοδοτήσεων επιχειρήσεων καθώς και την ευνοϊκότερη στάθμιση του πιστωτικού κινδύνου του χαρτοφυλακίου των τραπεζών. Η υλοποίηση των διατάξεων αυτών από τις τράπεζες επιβλέπεται από μια ορισμένη εποπτική αρχή.

Στην ουσία έθεσε ένα θεσμικό πλαίσιο αποδοχής και λειτουργίας των συστημάτων εκτίμησης πιστωτικού κινδύνου που χρησιμοποιούν οι τράπεζες. Μια βασική απαίτηση, σύμφωνα με τις διατάξεις που εκδόθηκαν, για ένα ορθό σύστημα εκτίμησης πιστωτικού κινδύνου είναι 'να διαφοροποιεί αποτελεσματικά τον κίνδυνο και να έχει αξιοπιστία στην εκτίμηση που παρέχει'. Βάσει του νέου αυτού πλαισίου, ο κίνδυνος κατηγοριοποιείται πιο σωστά και με αυτόν τον τρόπο προκύπτουν μικρότερες κεφαλαιακές απαιτήσεις. Δηλαδή ελέγχονται πιο σωστά οι δανειοδοτήσεις, άρα ή δεν χορηγούνται επισφαλή δάνεια είτε αυτά που εγκρίνονται χαρακτηρίζονται από έναν βαθμό ασφάλειας με αποτέλεσμα να μειώνονται οι κεφαλαιακές απαιτήσεις.

Τα νέα δεδομένα λοιπόν έχουν στρέψει τις τράπεζες προς την υιοθέτηση τεχνικών παρακολούθησης και διαχείρισης του πιστωτικού κινδύνου μέσω της ανάπτυξης εσωτερικών συστημάτων εκτίμησης πιστωτικού κινδύνου και στην βελτίωση των ήδη υπαρχόντων.

Προς τον σκοπό αυτό, αλλά και για την αντιμετώπιση σχετικών θεμάτων εποπτείας, η Τράπεζα της Ελλάδος έχει αρχίσει διαβουλεύσεις με ορισμένες τράπεζες ούτως ώστε να τεθούν οι βασικές κατευθύνσεις όσον αφορά την προσαρμογή των τραπεζών στη νέα μεθοδολογία διαχείρισης κινδύνων και εποπτικής αξιολόγησης.

Σύμφωνα με τη διάταξη της Οδηγίας 98/31/EK, η Τράπεζα της Ελλάδος παρέχει τη δυνατότητα στα πιστωτικά ιδρύματα να χρησιμοποιούν, υπό αυστηρά καθορισμένες προϋποθέσεις, εσωτερικά υποδείγματα διαχείρισης κινδύνου για τον υπολογισμό των κεφαλαιακών απαιτήσεων έναντι των κινδύνων της αγοράς, καθιστώντας αποτελεσματικότερη τη χρήση των ιδίων κεφαλαίων. (Τράπεζα της Ελλάδος, 2003)

Σφάλματα τύπου I και II διαδικασιών για κάθε επανάληψη

Μέθοδος	Επανάληψη	Σφάλμα I (%)	Σφάλμα II (%)	Σύνολο
percentiles (4 υποδιαστήματα)	1η	13.40%	5.40%	9.40%
	2η	9.30%	5.70%	7.50%
	3η	10.30%	5.80%	8.05%
	4η	11.00%	5.00%	8.00%
	5η	8.90%	4.20%	6.55%
	Μέσος Όρος	10.58%	5.22%	7.90%
percentiles (5 υποδιαστήματα)	1η	8.00%	5.40%	6.70%
	2η	9.30%	4.50%	6.90%
	3η	11.20%	5.10%	8.15%
	4η	13.80%	3.70%	8.75%
	5η	8.00%	3.70%	5.85%
	Μέσος Όρος	10.06%	4.48%	7.27%
percentiles (3 υποδιαστήματα)	1η	13.40%	5.40%	9.40%
	2η	7.40%	10.90%	9.15%
	3η	6.00%	9.80%	7.90%
	4η	4.60%	9.70%	7.15%
	5η	2.70%	9.90%	6.30%
	Μέσος Όρος	6.82%	9.14%	7.98%
K-S (4 υποδιαστήματα)	1η	8.90%	5.70%	7.30%
	2η	7.40%	4.00%	5.70%
	3η	14.70%	3.10%	8.90%
	4η	12.80%	4.20%	8.50%
	5η	10.70%	7.70%	9.20%
	Μέσος Όρος	10.90%	4.94%	7.92%
K-S (5 υποδιαστήματα)	1η	8.90%	5.70%	7.30%
	2η	13.00%	3.80%	8.40%
	3η	14.70%	3.20%	8.95%
	4η	16.50%	4.70%	10.60%
	5η	10.70%	7.10%	8.90%
	Μέσος Όρος	12.76%	4.90%	8.83%
K-S (3 υποδιαστήματα)	1η	7.10%	8.90%	8.00%
	2η	11.10%	6.90%	9.00%
	3η	8.60%	8.60%	8.60%
	4η	11.90%	5.20%	8.55%
	5η	5.40%	10.60%	8.00%
	Μέσος Όρος	8.82%	8.04%	8.43%

CART (4 υποδιαστήματα)	1η	7.10%	8.40%	7.75%
	2η	16.70%	9.20%	12.95%
	3η	13.80%	8.30%	11.05%
	4η	11.90%	7.50%	9.70%
	5η	17.00%	7.10%	12.05%
	Μέσος Όρος	13.30%	8.10%	10.70%
CART (5 υποδιαστήματα)	1η	8.00%	7.70%	7.85%
	2η	15.70%	9.00%	12.35%
	3η	20.70%	7.30%	14.00%
	4η	11.90%	7.50%	9.70%
	5η	20.50%	4.20%	12.35%
	Μέσος Όρος	15.36%	7.14%	11.25%
CART (3 υποδιαστήματα)	1η	8.00%	8.80%	8.40%
	2η	10.20%	10.40%	10.30%
	3η	9.50%	11.50%	10.50%
	4η	8.30%	11.20%	9.75%
	5η	6.20%	11.60%	8.90%
	Μέσος Όρος	8.44%	10.70%	9.57%
CART2	1η	8.90%	8.80%	8.85%
	2η	10.20%	10.40%	10.30%
	3η	9.50%	11.50%	10.50%
	4η	7.30%	11.20%	9.25%
	5η	6.20%	11.60%	8.90%
	Μέσος Όρος	8.42%	10.70%	9.56%

Περιοχή κάτω από την καμπύλη των διαδικασιών για κάθε επανάληψη

	Επανάληψη					
Διαδικασίες Διακριτοποίησης	1η	2η	3η	4η	5η	Μέσος όρος
percentiles (3 υποδιαστήματα)	97.40%	96.60%	96.90%	97.00%	97.90%	97.16%
percentiles (4 υποδιαστήματα)	97.20%	97.80%	97.70%	97.60%	98.20%	97.70%
percentiles (5 υποδιαστήματα)	97.80%	98.00%	97.40%	97.60%	98.90%	97.94%
K-S (3 υποδιαστήματα)	97.50%	97.00%	96.90%	97.00%	98.30%	97.34%
K-S (4 υποδιαστήματα)	97.80%	98.00%	97.10%	97.10%	97.40%	97.48%
K-S (5 υποδιαστήματα)	97.80%	96.50%	95.90%	96.70%	97.30%	96.84%
CART (3 υποδιαστήματα)	97.50%	96.10%	96.70%	96.10%	97.70%	96.82%
CART (4 υποδιαστήματα)	97.50%	95.70%	96.00%	96.00%	95.60%	96.16%
CART (5 υποδιαστήματα)	97.50%	95.70%	95.10%	96.00%	94.30%	95.72%
CART2	97.50%	96.10%	96.20%	96.20%	97.70%	96.74%

Βιβλιογραφία

Δούμπος, Μ. και Ζοπουνίδης, Κ. (2001), *Πολυκριτήριες Τεχνικές Ταξινόμησης*, Εκδόσεις Κλειδάριθμος, Αθήνα.

Ζοπουνίδης, Κ.(1998), *Ανάλυση και Διαχείριση Χρηματοοικονομικών Κινδύνων*, Εκδόσεις Κλειδάριθμος, Αθήνα.

Ζοπουνίδης, Κ. (2000), *Βασικές Αρχές Χρηματοοικονομικού Μάνατζμεντ*, Μεσογειακό Αγρονομικό Ινστιτούτο Χανίων, Χανιά.

Τράπεζα της Ελλάδος (2003), *Έκθεση του διοικητή για το έτος 2002*, Αθήνα.

Altman, E.I. (1968), “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy”, *Journal of Finance*, 23, 589-609.

Altman, E.I and Saunders, A. (1998), “Credit risk measurement: Developments over the last 20 years”, *Journal Banking and Finance*, 21, 1721-1742.

Breiman, L., Friedman J.H., Olsen R.A. and Stone C.J. (1984), *Classification and Regression Trees*, Pacific Grove, California.

Crouhy, M., Galai, D. and Mark, R. (2001), “Prototype risk rating systems”, *Journal of Banking and Finance*, 25, 47-95.

Deutsche Bundesbank (Monthly report September 2003), “Approaches to the validation of internal rating systems”.

Fawcett, T. (2003), “ROC graphs: Notes and practical considerations for data mining researchers”, HP Labs, Working paper HPL-2003-4.

Fisher, R.A. (1936), “The use of multiple measurements in taxonomic problems”, *Annals of Eugenics*, 7, 179-188.

Friedman, J.H. (1977), “A recursive partitioning decision rule for nonparametric classification”, *IEEE Transactions on Computers*, C-26, 404-408.

Frydman, H., Altman, E.I. and Kao, D.L. (1985), “Introducing recursive partitioning for financial classification: The case of financial distress”, *Journal of Finance*, XL/1, 269, 269-291.

Hand, D.J. and Adams, N.L. (2000), “Defining attributes for scorecard construction in credit scoring”, *Journal of Applied Statistics*, 27/5, 527-540.

Kosko, B. (1992), *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, New Jersey.

Krahnert, J.P. and Weber, M. (2001), "Generally accepted rating principles: A primer", *Journal of Banking and Finance*, 25, 3-23.

Nargundkar, S. and Priestley, J.L., "Assessment of evaluation methods for binary classification modeling", Proceedings of the 2003 Decision Science Institute National Conference, Washington, D.C., November 2003: pp.1-6.

Patuwo, E., Hu, M.Y. and Hung, M.S. (1993), "Two-group classification using neural networks", *Decision Sciences*, 24, 825-845.

Pawlak, Z. (1982), "Rough sets", *International Journal of Information and Computer Sciences*, 11, 341-356.

Quinlan, J.R. (1993), *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, Los Altos, California.

Roy, B. (1985), *Methologie Multicritere d'Aide a la Decision*, Economica, Paris.

Schumann, M. and Liu, Y. (02/2002), "The evaluation of classification models for credit scoring", *Institute fur Wirtschaftsinformatik*.

Smith, C. (1947), "Some examples of discrimination", *Annals of Eugenics*, 13, 272-282.

Steinberg, D. and Colla (1995), *CART: Tree-structured nonparametric data analysis*, San Diego, Calif., U.S.A: Salford Systems.

Subramanian, V., Hung, M.S. and Hu, M.Y. (1993), "An experimental evaluation of neural networks for classification", *Computers and Operation Research*, 20/7, 769-782.

Treacy, W.F. and Carey, M. (2000), "Credit risk rating systems at large US banks", *Journal of Banking and Finance*, 24, 167-201.

Utgoft, P.E. and Clouse, J.A. (1996), "A Kolmogorov-Smirnoff metric for decision tree induction", *Department of Computer Science, University of Massachusetts*, Technical Report, 96-3.

Yohannes, Y. and Webb, P. (1999), *Classification and regression trees, Cart: A user manual for identifying indicators of vulnerability to famine and chronic food insecurity*, International Food Policy Research Institute, Microcomputers in Policy Research 3, Washington D.C.

Zadeh, L.A., (1965) “Fuzzy sets”, *Information and Control*, 8, 338-353.

Πηγές από το διαδίκτυο:

- Εξέλιξη της πιστοληπτικής ικανότητας των ελληνικών επιχειρήσεων για τα έτη 1998 έως 2002. ICAP A.E. Τομέας Επιχειρηματικής Πληροφόρησης. Διεύθυνση Ανάπτυξης Υπηρεσιών Επιχειρηματικής Πληροφόρησης. (<http://www.icap.gr/>)