

**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ**  
**ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ**

---

**ΑΝΑΛΥΤΙΚΑ ΜΟΝΤΕΛΑ  
ΑΞΙΟΛΟΓΗΣΗΣ ΠΙΣΤΩΤΙΚΟΥ  
ΚΙΝΔΥΝΟΥ ΣΕ ΠΛΑΤΦΟΡΜΕΣ  
PEER-TO-PEER LENDING**

---

Υπό  
**ΠΑΥΛΟΥ ΓΕΩΡΓΙΟ**

*Χανιά*  
**2018**

## Περιεχόμενα

Ευχαριστίες .....	3
Εισαγωγή.....	4
Κεφάλαιο 1.....	5
1.1    Γενικά για το FinTech.....	5
1.2    Νέες πλατφόρμες χρηματοδότησης CrowdFunding .....	10
1.3    Ορισμός του Peer-to-Peer Lending.....	11
1.4    Η σημαντικότητα του Peer-to-Peer Lending στο FinTech .....	11
1.5    Πλεονεκτήματα και μειονεκτήματα.....	12
1.6    Παρουσίαση Πλατφορμών.....	14
Κεφάλαιο 2.....	20
Κεφάλαιο 3.....	27
3.1    Παρουσίαση δεδομένων .....	27
3.2    Στατιστική Ανάλυση .....	35
3.3    Μεθοδολογία.....	37
3.4    Αποτελέσματα.....	39
Κεφάλαιο 4.....	42
Βιβλιογραφία.....	44

*Στον αδερφό μου Αντώνη και τους φίλους  
μου Κατερίνα και Μιχάλη.*

# ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της διπλωματικής μου εργασίας θα ήθελα πρωτίστως να ευχαριστήσω τον Καθηγητή της Σχολής Μηχανικών Παραγωγής και Διοίκησης κ. Μιχάλη Δούμπο για την εμπιστοσύνη που μου έδειξε αναθέτοντας μου την εκπόνηση της συγκεκριμένης εργασίας. Επίσης θα ήθελα να τον ευχαριστήσω για όλη τη βοήθεια που μου παρείχε κατά την διάρκεια της εκπόνησης της διπλωματικής μου.

Θα ήθελα επίσης να ευχαριστήσω τα μέλη της εξεταστικής επιτροπής που με τίμησαν με την παρουσία τους στην παρουσίαση της διπλωματικής μου εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω τους γονείς μου και τον αδερφό μου Αντώνη που με στήριξαν όλα αυτά τα χρόνια που σπούδασα στο Πολυτεχνείο Κρήτης και ήταν πάντα δίπλα μου να μου δίνουν κουράγιο, δύναμη.

Τέλος, θα ήθελα να ευχαριστήσω όλους τους φίλους μου, έναν-έναν ξεχωριστά για την ηθική τους υποστήριξη όλα αυτά τα χρόνια και περισσότερο από όλους την φίλη μου Κατερίνα που ήταν εκεί όποτε την χρειαζόμουν να με στηρίξει σε όλες μου τις αποφάσεις.

# ΕΙΣΑΓΩΓΗ

Στη σύγχρονη εποχή, οι συνεχείς εξελίξεις στο χώρο της τεχνολογίας έχουν επιφέρει σημαντικά αποτελέσματα στο χώρο του FinTech. Μία από αυτές ήταν και η δημιουργία ηλεκτρονικών ιστοσελίδων χρηματοδότησης Peer-to-Peer Lending με βάση την οποία παραλείπεται ο ενδιάμεσος μεσολαβητής όπως γίνεται στα κλασσικά χρηματοπιστωτικά ιδρύματα και η διεπαφή με τη διαδικασία θα επιτυγχάνεται μέσω σύγχρονων ηλεκτρονικών πλατφορμών. Με την εξέλιξη της συγκεκριμένης καινοτομίας, που δημιουργήθηκε για πρώτη φορά το 2005 στο Ηνωμένο Βασίλειο, ήρθαν στην επιφάνεια ορισμένες ανησυχίες ως προς τους κινδύνους που ελλοχεύουν στο χώρο αυτό. Ένας από αυτούς, είναι η μελέτη και αξιολόγηση του πιστοληπτικού κινδύνου με τη χρήση μεθόδων μηχανικής μάθησης.

Αυτό είχε ως αποτέλεσμα τη δημιουργία ενός μεγάλου φάσματος διεθνής βιβλιογραφίας με βάση την οποία ερευνώνται και αναλύονται νέοι συνδυαστικοί μέθοδοι γύρω από το συγκεκριμένο σκοπό. Όπως έχει τονιστεί καμία μέθοδος μηχανικής μάθησης δεν είναι καλύτερη σε γενικό επίπεδο από την άλλη, διότι κάθε μια μελετά με διαφορετικό τρόπο τα δεδομένα του εκάστοτε προβλήματος. Έτσι οι περισσότερες έρευνες επικεντρώνονται στη σύγκριση των μεθόδων αυτών ή και στο συνδυασμό αυτών για επίτευξη καλύτερης ακρίβειας.

Στη συγκεκριμένη εργασία επιδιώκεται, η καταγραφή των πιο σημαντικών μεταβλητών που δίνονται από την Peer-to-Peer Lending πλατφόρμα, prosper.com, η επεξήγησή τους και η χρήση επτά δημοφιλών μεθόδων μηχανικής μάθησης (LDA, QDA, LR, SVM, Classification Tree, Boosted Trees και Bugged Trees) με σκοπό την ταξινόμηση 106160 καταγραφών δανείων σε εξυπηρετούμενα και μη εξυπηρετούμενα, τον υπολογισμό της πιθανότητας ταξινόμησης τους σε κάθε μια από αυτές τις κατηγορίες και εν τέλει η σύγκριση των επτά μεθόδων με σκοπό της κατάταξής τους ως προς την ακρίβεια της προβλεπτικής τους ικανότητας.

Με αυτό τον τρόπο στο πρώτο μέρος της αναφοράς γίνεται μια πλήρης ανάλυση του θεωρητικού υποβάθρου του τομέα του Peer-to-Peer Lending μέσα από την οποία αναδεικνύεται η σημαντικότητα της και οι θετικές επιδράσεις της στον ευρύτερο τομέα του FinTech. Στο δεύτερο μέρος, παρουσιάζεται μια βιβλιογραφική ανασκόπηση που αφορά μέρος των εργασιών που έχουν εκπονηθεί γύρω από αυτή την προβληματική καθώς και ένας συγκεντρωτικός πίνακας της ανασκόπησης αυτής. Στο τρίτο μέρος, γίνεται η ανάλυση των δεδομένων που θα χρησιμοποιηθούν στη συγκεκριμένη εργασία και της αναλυτικής θεωρίας γύρω από τις μεθόδους που θα εξάγουν τα συμπεράσματα της εργασίας καθώς και η παρουσίαση του πειραματικού μέρους. Τέλος, στο τελευταίο μέρος θα γίνει η παρουσίαση των συμπερασμάτων που προέκυψαν καθώς και ιδέες για επιπλέον έρευνα και μελέτη.

# ΚΕΦΑΛΑΙΟ 1

Ξεκινώντας την ανάλυση του θέματος της εργασίας, θα παρατεθεί το βασικό θεωρητικό υπόβαθρο έτσι ώστε να προσεγγιστεί σφαιρικά το θέμα και να γίνει απολύτως κατανοητό. Πιο αναλυτικά, στην πρώτη υποενότητα του κεφαλαίου θα αναλυθούν ορισμένες βασικές έννοιες σχετικά με τον ευρύτερο χώρο του FinTech όπου κατατάσσεται και η θεματολογία των πρωτοπόρων ιδεών όπως του Crowdfunding και του Peer-to-Peer Lending. Στη συνέχεια, θα παρατεθεί το γενικό σκεπτικό του όρου Crowdfunding σε μια κατηγορία του οποίου ανήκει και το Peer-to-Peer Lending. Έπειτα, θα οριστεί το Peer-to-Peer Lending όπου είναι και ο βασικός στόχος του κεφαλαίου, η σημαντικότητα του στον κλάδο του FinTech, τα βασικά πλεονεκτήματα και μειονεκτήματά του και τέλος θα παρουσιαστούν οι πιο δημοφιλής πλατφόρμες Peer-to-Peer Lending ώστε να αναδειχτούν και στην πράξη οι βασικές μορφές του.

## 1.1 ΓΕΝΙΚΑ ΓΙΑ ΤΟ FINTECH

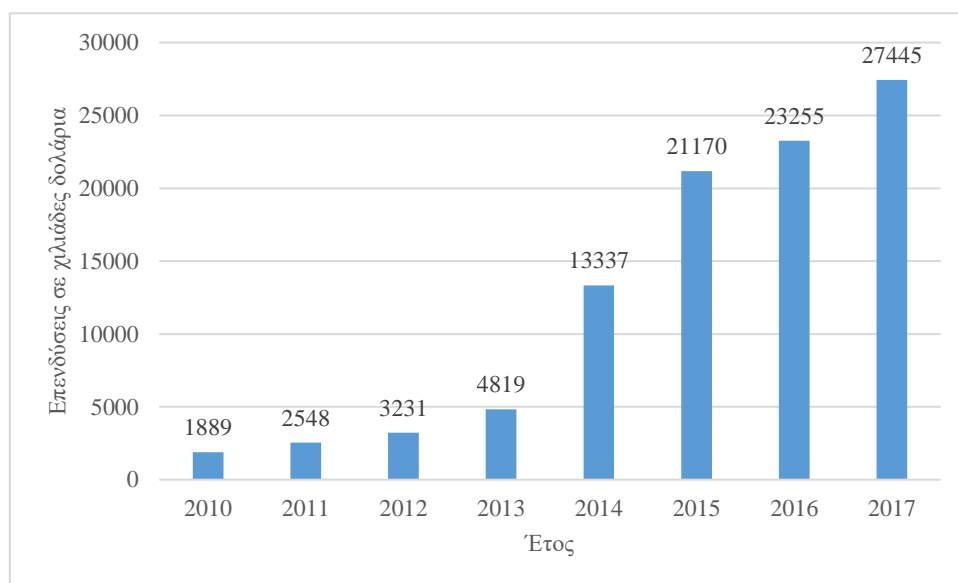
Ο όρος FinTech, αποτελεί ένα ακρωνύμιο των λέξεων Financial και Technology, δηλαδή χρηματοοικονομικά και τεχνολογία. Στην αρχή του 21<sup>ου</sup> αιώνα, δημιουργήθηκε ώστε να περιγράψει οποιαδήποτε καινοτομία ή χρήση της τεχνολογίας στο χρηματοοικονομικό τομέα, που συμβάλουν, στην καλύτερη διεκπεραίωση των κλασσικών χρηματοπιστωτικών διαδικασιών, αλλά και γενικά στη διευκόλυνση της επαφής των επιχειρήσεων και καταναλωτών με το χώρο αυτό.

Με την εξέλιξη της τεχνολογίας, οι επιστήμονες αναζητούσαν τρόπους απλοποίησης και αυτοματοποίησης της λειτουργίας του χρηματοπιστωτικού τομέα, σκοπεύοντας στην ευρύτερη ανάπτυξη του τομέα, τη μείωση του χρόνου, και του κόστους διεκπεραίωσης των συναλλαγών. Με αυτό τον τρόπο, δημιουργήθηκε ένα ολόκληρο σύστημα ηλεκτρονικής υποστήριξης, με τη βοήθεια ειδικών λογισμικών και ηλεκτρονικών υπολογιστών, που αντικατέστησε τις παρωχημένες μεθόδους υπολογισμών και λήψης αποφάσεων. Στη συνέχεια, με τη ραγδαία ανάπτυξη του FinTech, ξεκίνησε και η ταχύτατη ένταξη νέων καινοτομιών-προϊόντων, όπως για παράδειγμα των κρυπτονομισμάτων, προκειμένου να γίνει ακόμα πιο προσιτή η αλληλεπίδραση των καταναλωτών, επενδυτών και επιχειρήσεων με τον χώρο, να υπάρξει περισσότερη και καλύτερη διαχείριση της πληροφορίας και να διερευνηθούν νέοι τρόποι επενδύσεων και λήψης δανείων πέρα από τους κλασσικούς και ήδη υπάρχοντες.

Έτσι, παρόλο που στην αρχή το FinTech κινούνταν γύρω από το πλαίσιο του κλάδου των χρηματοπιστωτικών ιδρυμάτων, τώρα η χρήση του διευρύνεται και στοχεύει επιπλέον στις επιχειρήσεις αλλά και σε μεμονωμένα άτομα που θέλουν να ενημερώνονται για τις καθημερινές τους συναλλαγές (ATM και εφαρμογές στα κινητά τηλέφωνα) και να βρίσκουν νέους τρόπους ευελιξίας στην καθημερινότητά τους. Στο χώρο του FinTech

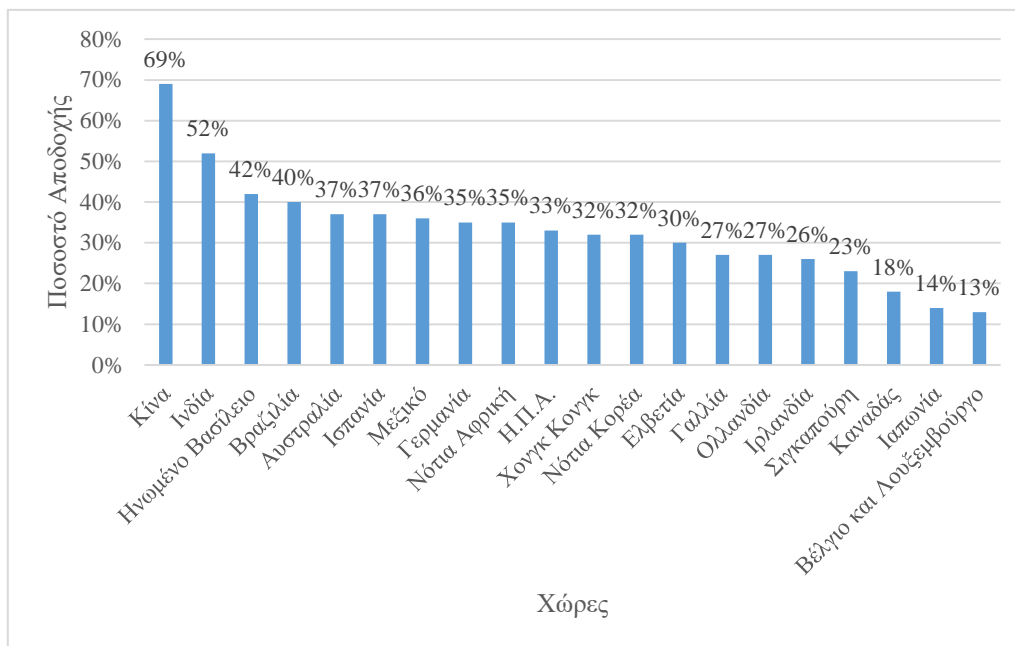
λοιπόν ανήκουν και οι πλατφόρμες CrowdFunding και Peer-to-Peer Lending που θα αναλυθούν περαιτέρω στα υπόλοιπα υποκεφάλαια.

Σύμφωνα με δεδομένα που παραχωρεί η εταιρεία Accenture, προκύπτει το διάγραμμα 1.1 στο οποίο φαίνεται ο όγκος των επενδύσεων που πραγματοποιούνται στον ευρύτερο κλάδο του FinTech σε παγκόσμια κλίμακα. Πιο αναλυτικά, από το 2010 φαίνεται να έχουν επενδυθεί περίπου 100 δισεκατομμύρια δολάρια σε τέτοιου είδους επιχειρήσεις. Με τα παρακάτω στοιχεία γίνεται ευδιάκριτη, η συνεχής ανάπτυξη του κλάδου αυτού, η σημαντικότητα του στο χώρο των επενδύσεων και ταυτόχρονα, φαίνεται η ιδιαίτερη έμφαση που δίνει το καταναλωτικό κοινό σε τέτοιου είδους νέα εγχειρήματα.



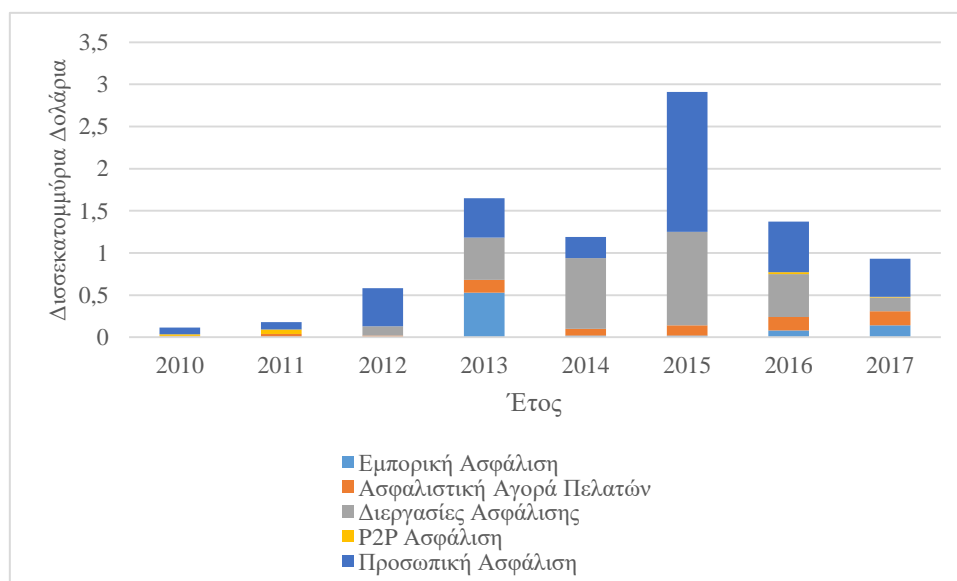
Διάγραμμα 1.1: Επενδύσεις στον κλάδο του FinTech ανά έτος.

Με βάση μια εκτεταμένη έρευνα της εταιρείας Ernst and Young, δημιουργήθηκε ένας παγκόσμιος δείκτης μέτρησης της ανταπόκρισης του FinTech (ο δείκτης FinTech Adoption), σε 20 αγορές ανά τον κόσμο, προσπαθώντας με στατιστικά στοιχεία να ποσοτικοποιήσει το μέγεθος της εξάπλωσης του Fintech και το μέγεθος ανταπόκρισης των ατόμων απέναντι στα παράγωγά του. Η έρευνα πραγματοποιήθηκε το 2017 και περιλάμβανε 22000 ερωτώμενους βγάζοντας ως γενικό συμπέρασμα ότι κατά μέσο όρο ένας στους τρεις που ανήκουν στο ψηφιακά ενεργό καταναλωτικό κοινό χρησιμοποιούν υπηρεσίες του Fintech σε σύγκριση με την ίδια έρευνα που είχε πραγματοποιήσει το 2015, όπου η χρήση του FinTech γινόταν από ένα άτομο στους επτά που ανήκουν στο ίδιο ενεργό καταναλωτικό κοινό. Στο διάγραμμα 1.2 παρουσιάζονται τα ποσοστά απήχησης του κλάδου ανά χώρα το έτος 2017. Όπως γίνεται φανερό το κατά μέσο όρο ποσοστό αποδοχής του FinTech και της φιλοσοφίας του είναι το 33%, ποσοστό που ξεπερνούν οι μισές τουλάχιστον χώρες του δείγματος και οι υπόλοιπες το προσεγγίζουν.



Διάγραμμα 1.2: Η απήχηση του FinTech σε 20 χώρες.

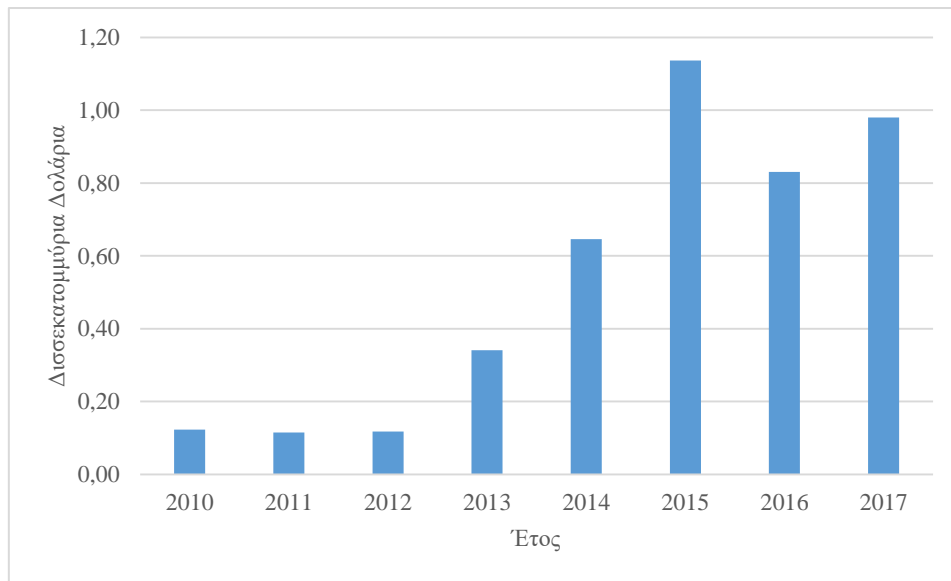
Στο σημείο αυτό, αξίζει να σημειωθούν και τα πιο βασικά παράγωγα της γενικότερης χρήσης του FinTech καθώς και τα ποσά χρηματοδότησης τους παγκοσμίως ώστε να γίνει κατανοητή η βαρύτητα που προσδίδει η αγορά τα τελευταία 7 χρόνια. Στα διαγράμματα 1.3-1.6 παρουσιάζονται 4 βασικοί τομείς επιχειρηματικής δραστηριότητας όπου στον καθένα φαίνονται οι επενδύσεις σε δισεκατομμύρια δολάρια που γίνονται ανά έτος από το 2010 έως και το 2017 από μια εκτεταμένη στατιστική έρευνα που εκπόνησε το ερευνητικό τμήμα της εταιρείας Deloitte με πειραματικά δεδομένα της επιχείρησης Venture Scanner.



Διάγραμμα 1.3: Χρηματοδοτήσεις στον τομέα των Κεφαλαιαγορών και των Τραπεζικών Υπηρεσιών ανά έτος.

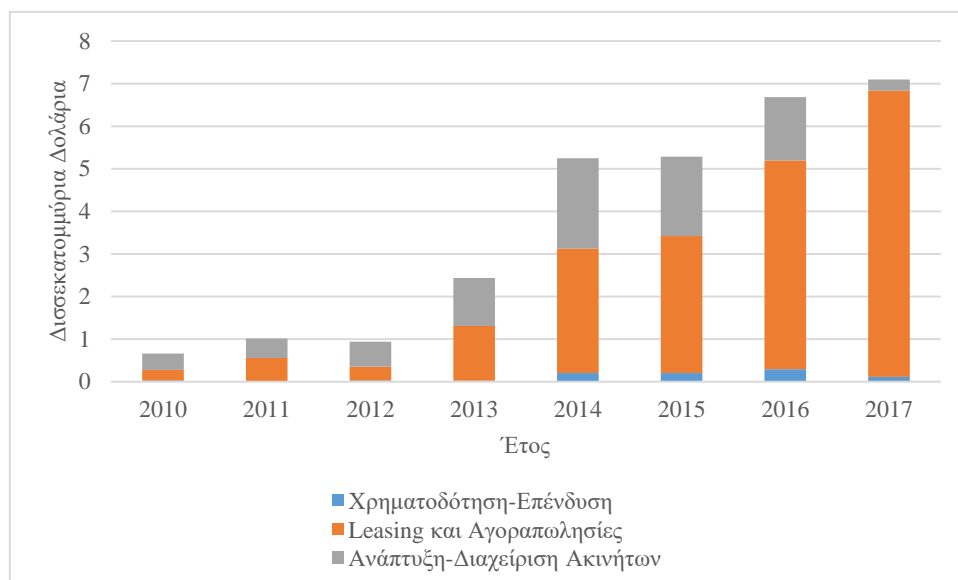


Πιο αναλυτικά, στο διάγραμμα 1.3 παρουσιάζονται τα ποσά χρηματοδοτήσεων στο χώρο των κεφαλαιαγορών και γενικότερα των τραπεζικών υπηρεσιών όπου διαπιστώνεται μια αυξημένη τάση χρηματοδότησης προς τις καταθέσεις και τους δανεισμούς καθώς και τις τραπεζικές διεργασίες. Αυτός όπως αναλύθηκε και παραπάνω ήταν ο πρωταρχικός στόχος της ανάπτυξης του Fintech ώστε να εξομαλυνθούν ορισμένες πολύπλοκες και χρονοβόρες διαδικασίες στις απαραίτητες αυτές λειτουργίες της σημερινής οικονομικής κοινότητας.



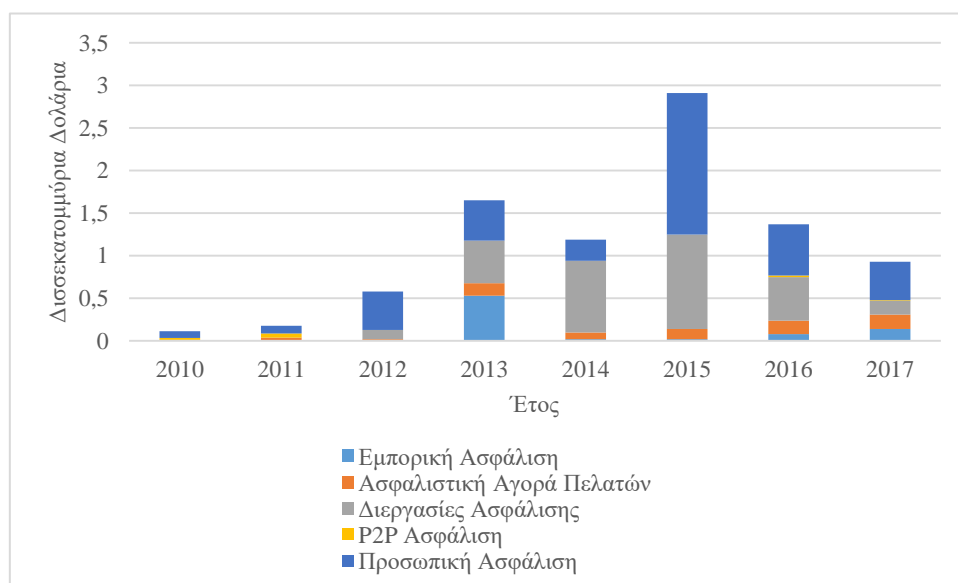
Διάγραμμα 1.4: Χρηματοδοτήσεις στον τομέα της Διαχείρισης των Επενδύσεων ανά έτος.

Στο διάγραμμα 1.4 γίνεται εξίσου εμφανής και η έμφαση που δίνεται στο χώρο της διαχείρισης των επενδύσεων. Έτσι, με γνώμονα τη συνεχή εξέλιξη του FinTech και των προτερημάτων που μπορεί να προσδώσει στο συγκεκριμένο τομέα, το επενδυτικό κοινό φαίνεται να δίνει ιδιαίτερο κύρος και προσοχή στο συγκεκριμένο διεθνή κλάδο της επιχειρηματικής βιομηχανίας.



Διάγραμμα 1.5: Χρηματοδοτήσεις στον τομέα των Ακινήτων ανά έτος.

Μια έντονη επιχειρηματική δραστηριότητα καταγράφεται και στο χώρο των ακινήτων όπως γίνεται αντιληπτή και στο διάγραμμα 1.5. Πιο συγκεκριμένα, οι χρηματοδοτήσεις ανά έτος φαίνεται να επικεντρώνονται στην ευρύτερη ανάπτυξη του τομέα της διαχείρισης των ακινήτων αλλά και πιο συγκεκριμένα στις αγοραπωλησίες και το Leasing.



Διάγραμμα 1.6: Χρηματοδοτήσεις στον τομέα των Ασφαλίσεων ανά έτος.

Ένα τελευταίο σημείο που αξίζει να καταγραφεί είναι, και η επιχειρηματική δραστηριότητα στον τομέα των ασφαλίσεων. Στο διάγραμμα 1.6 γίνεται εμφανές ότι ανά έτος οι χρηματοδοτήσεις στον συγκεκριμένο τομέα κινούνται κυρίως γύρω από την

ευρύτερη ανάπτυξη της προσωπικής ασφάλισης καθώς και τις απαραίτητες διεργασίες ασφάλισης.

## 1.2 ΝΕΕΣ ΠΛΑΤΦΟΡΜΕΣ ΧΡΗΜΑΤΟΔΟΤΗΣΗΣ CROWDFUNDING

Στα πλαίσια ανάπτυξης του FinTech και της πολύτιμης βοήθειας που προσφέρει το διαδίκτυο και οι πλατφόρμες δικτύωσης, η επιστημονική κοινότητα προσπάθησε να αντικαταστήσει τη κυριότερη χρησιμότητα των τραπεζών που είναι οι καταθέσεις χρηματικών ποσών και η δανειοδότηση. Ταυτόχρονα, οι συνεχόμενες μεταβολές στις οικονομίες, οι διαρκείς ανησυχίες των ανθρώπων σχετικά με το χρήμα και την κάλυψη των αναγκών τους και τα μεγάλα ποσοστά αποτυχίας των μικρομεσαίων επιχειρήσεων λόγω έλλειψης χρηματοδότησης, έφεραν στο παρασκήνιο νέες καινοτομίες στους τρόπους χρηματοδότησης και ιδιαίτερα στο Ηνωμένο Βασίλειο

Με αυτά τα ιδεολογικά θεμέλια, δημιουργήθηκε η ιδέα του CrowdFunding ή αλλιώς της χρηματοδότησης από το πλήθος. Η έννοια της ιδέας αυτής, περιλαμβάνει πολλά πεδία ανάλογα τη μορφή που έχει, αλλά στην πιο απλή της εκδοχή, είναι η χρηματοδότηση νέων ιδεών, καινοτομιών, έργων, από πλήθος ατόμων που πραγματοποιείται με την πρόσβαση και την αλληλεπίδραση σε αντίστοιχους διαδικτυακούς ιστότοπους.

Στο σημείο αυτό θα παρουσιαστούν οι τέσσερις μορφές των πλατφορμών CrowdFunding και θα αναλυθούν περαιτέρω.

- **Χρηματοδότηση δωρεάς ή ανταμοιβής:** Σε αυτή την κατηγορία, ανήκουν οι δωρεές που γίνονται από το πλήθος σε ιδέες, καινοτομίες και επιχειρήσεις χωρίς κάποια συγκεκριμένη οικονομική ανταμοιβή με την προσδοκία ότι κάποια στιγμή στο μέλλον να λάβουν κάτι συμβολικό ως αντίκτυπο της προσφοράς τους.
- **Royalty CrowdFunding:** Στη συγκεκριμένη περίπτωση, οι υποστηρικτές που χρηματοδοτούν το εγχείρημα, εισπράτουν ως οικονομικό αντάλλαγμα ποσοστά των εσόδων του έργου ή τα επιχειρήσεις όταν δημιουργηθούν εισοδηματικό κεφάλαιο. Η συγκεκριμένη κατηγορία CrowdFunding, αναφέρεται κυρίως σε δικαιώματα εκμετάλλευσης προϊόντων που προκύπτουν ανάλογα τη φύση και τα έσοδα του προϊόντος.
- **Χρηματοδότηση Δανεισμού:** Πρόκειται για την πλέον σημαντική μορφή CrowdFunding καθώς δεν έχει περιορισμούς στο κοινό που αναφέρεται. Όπως γίνεται κατανοητό, από την ονομασία της κατηγορίας αυτής εδώ το κοινό δανείζει κεφάλαια και λαμβάνει τόκους, όπως ακριβώς γίνεται και με τις τράπεζες. Αυτή η μορφή δανεισμού εμπεριέχει διάφορες υποκατηγορίες όπως το Peer-to-Peer Lending που θα αναλυθεί αναλυτικά στο επόμενο υποκεφάλαιο και ένας πιο έμμεσος δανεισμός που πραγματοποιείται με την έκδοση ομολόγων.
- **Equity Crowdfunding:** Στην περίπτωση της χρηματοδότησης με μετοχικά κεφάλαια το πλήθος χρηματοδοτεί την εταιρεία ή την επιχείρηση με

οικονομικό αντάλλαγμα τη συμμετοχή των ιδίων στη ίδια την εταιρεία με τη μορφή μετοχικών κεφαλαίων.

### **1.3 ΟΡΙΣΜΟΣ ΤΟΥ PEER-TO-PEER LENDING**

Στη διεθνή βιβλιογραφία θα βρει κανείς αρκετούς ορισμούς που περιγράφουν τη φιλοσοφία της έννοιας του όρου Peer-to-Peer Lending οι οποίοι κατά βάση καταλήγουν στο εξής συμπέρασμα:

“Peer-to-Peer Lending ή αλλιώς δανεισμός από άτομο σε άτομο ονομάζεται η διαδικασία δανεισμού από ένα άτομο ή μια ομάδα ατόμων για συγκεκριμένο σκοπό χωρίς να εμπλέκονται διαμεσολαβητές όπως στην κλασσική διαδικασία των χρηματοπιστωτικών ιδρυμάτων δηλαδή τις τράπεζες. Η διαδικασία αυτή επιτυγχάνεται, με τη χρήση εξειδικευμένων πλατφορμών δικτύωσης που εμπεριέχουν συγκεκριμένα και ανεπτυγμένα εργαλεία πρόβλεψης και ελέγχου όλου του μηχανισμού της διαδικασίας (Wang G. et al., 2015).”

Πιο αναλυτικά, η λήψη δανείου μέσω της διαδικασίας του Peer-to-Peer Lending μπορεί να ταυτιστεί ως ένα βαθμό, με τη διαδικασία του πλειστηριασμού ενός αγαθού στην πλειονότητα της χρήσης του. Το άτομο που χρειάζεται το δάνειο κάνει εγγραφή στην εξειδικευμένη πλατφόρμα Peer-to-Peer Lending δίνοντας τα στοιχεία του δανείου που ζητά από το κοινό, το λόγο για τον οποίο θέλει να λάβει το δάνειο καθώς και τα απαραίτητα οικονομικά του στοιχεία ώστε οι ενδιαφερόμενοι να μπορούν να αξιολογήσουν τη φερεγγυότητά του. Στη συνέχεια, ο πιθανός δανειολήπτης, θέτει ένα αρχικό επιτόκιο στο οποίο θεωρεί ότι μπορεί να ανταπεξέλθει μελλοντικά και έτσι ορίζεται το αρχικό επιτόκιο δανεισμού. Έπειτα, οι επενδυτές αξιολογούν με τη βοήθεια των δεικτών φερεγγυότητας το δάνειο και επιλέγουν αν θα επενδύσουν ή όχι στο δάνειο αυτό, το ποσό που θα επενδύσουν καθώς και το επιτόκιο το οποίο θεωρούν αποδεκτό. Τέλος, γίνεται χρήση του λειτουργικού συστήματος που υποστηρίζει η κάθε πλατφόρμα, προκειμένου να επιτευχθεί μια διαδικασία αντίστροφου ταιριάσματος του επιτοκίου του δανειολήπτη και των επενδυτών, ώστε να οριστεί ένα επιτόκιο δανεισμού που θα καλύπτει το ποσό δανείου που ζητήθηκε στην αρχή και θα είναι βιώσιμο.

### **1.4 Η ΣΗΜΑΝΤΙΚΟΤΗΤΑ ΤΟΥ PEER-TO-PEER LENDING ΣΤΟ FINTECH**

Με το ξεκίνημα της πρωτοποριακής ιδέας του Peer-to-Peer Lending από το Ηνωμένο Βασίλειο με την πλατφόρμα του Zopa.com, οι βλέψεις των επενδυτών άρχισαν να στρέφονται ολοένα και περισσότερο σε νέες εναλλακτικές μορφές επενδυτικών λύσεων. Χρησιμοποιώντας, τις εξελίξεις της τεχνολογίας ο χώρος του CrowdFunding έγινε πιο διαδεδομένος στο ευρύ κοινό φέρνοντας το Ηνωμένο Βασίλειο πρώτο στους δείκτες των εναλλακτικών χρηματοοικονομικών καινοτομιών.

Πιο συγκεκριμένα, με βασική δομή την αξιοποίηση των ήδη υφιστάμενων διαδικασιών χρηματοδότησης μέσω τραπεζών και τη ραγδαία ανάπτυξη της τεχνολογίας, τα βλέμματα στράφηκαν στην εξέλιξη των ήδη υφιστάμενων τρόπων αξιολόγησης ενός

δανείου και τις τεχνικές θεσμοθέτησης κριτηρίων ως προς την πιστοληπτική ικανότητα των δανειζόμενων. Έτσι, δημιουργήθηκαν νέα μοντέλα που προβλέπουν την αξιοπιστία ενός δανείου με ανάλογους δείκτες φερεγγυότητας για τον κάθε δανειζόμενο παραγκωνίζοντας τα αυστηρά όρια των τραπεζών.

Έπειτα, στην προσπάθεια δόμησης αυτής της νέας εναλλακτικής διαδικασίας δημιουργήθηκαν ηλεκτρονικές πλατφόρμες εύχρηστες και φιλικές προς το ευρύ κοινό. Με την ένταξη ευφυών συστημάτων ανάλυσης και διαχείρισης δεδομένων μειώθηκε σε μεγάλο βαθμό ο χρόνος που απαιτείται για τη σύναψη ενός δανείου και η πολυπλοκότητα της διαδικασίας δομώντας έτσι ένα πιο προσιτό περιβάλλον αλληλεπίδρασης μεταξύ επενδυτή και δανειολήπτη.

Ταυτόχρονα, σε τέτοιου είδους πλατφόρμες όπως είναι φανερό, απουσιάζει αισθητά το στοιχείο ελέγχου και ταυτοποίησης των απαραίτητων δικαιολογητικών καθώς και των προσωπικών χαρακτηριστικών κάθε δανειολήπτη. Έτσι, με κύριο στόχο τη διαφάνεια και τη διασφάλιση των προσωπικών δεδομένων κάθε πλατφόρμα Peer-to-Peer Lending διαθέτει εξειδικευμένο λογισμικό που ελέγχει τη γνησιότητα των στοιχείων του κάθε χρήστη και τα πιστοποιεί ελαχιστοποιώντας ταυτόχρονα και τον κίνδυνο κακόβουλων ενεργειών. Δημιουργεί δηλαδή συγκεκριμένες δικλίδες ασφαλείας ως προς τα προσωπικά δεδομένα των χρηστών ένα από τα πολλά χαρακτηριστικά που τη διαφοροποιεί από τις κλασσικές ιστοσελίδες κοινωνικής δικτύωσης.

## 1.5 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΜΕΙΟΝΕΚΤΗΜΑΤΑ

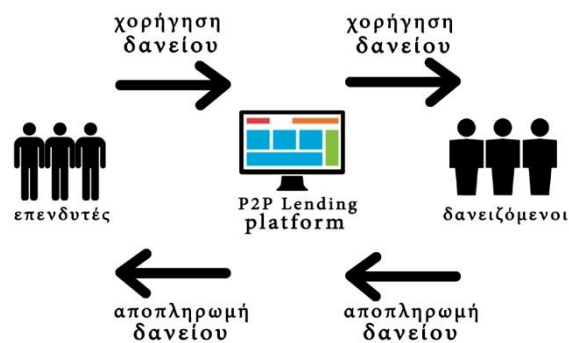
Αναλύοντας την έννοια του Peer-to-Peer Lending, αξίζει να σημειωθούν τα πιο βασικά πλεονεκτήματα και μειονεκτήματα ώστε να γίνει κατανοητή η σημασία αυτής της καινοτομίας αλλά και τονιστούν ορισμένα αρνητικά ζητήματα όπου χρειάζονται προσοχή και βελτίωση.

Αρχικά, όπως τονίστηκε και στην υποενότητα 1.3, το πιο σημαντικό πλεονέκτημα του Peer-to-Peer Lending, είναι η έλλειψη μεσαζόντων (Σχήμα 1.1) όπως γίνεται στην κλασσική διαδικασία χρηματοδότησης που πρεσβεύει το τραπεζικό σύστημα (Σχήμα 1.2). Πιο συγκεκριμένα, στην κλασσική χρηματοδότηση από τράπεζες, το πιστωτικό ίδρυμα αποτελεί το ρόλο του μεσάζοντα που λαμβάνει τις καταθέσεις των επενδυτών και τις χρησιμοποιεί ως κεφάλαια για την κάλυψη αιτήσεων χορήγησης δανείων. Η ιδέα του Peer-to-Peer Lending καταργεί αυτή τη γέφυρα επικοινωνίας μεταξύ δανειολήπτη και πιστωτή κάνοντας πιο ευέλικτη τη διαδικασία.

Βασικό πλεονέκτημα του Peer-to-Peer Lending είναι οι απολαβές των επενδυτών αλλά και τα χαμηλά επιτόκια δανεισμού. Πιο συγκεκριμένα, ο ορισμός του επιτοκίου δανεισμού προκύπτει από τους δείκτες φερεγγυότητας του δανειζόμενου (όσο πιο φερέγγυος ο δανειολήπτης τόσο χαμηλότερο το επιτόκιο δανεισμού στο οποίο διαπραγματεύεται) και οριστικοποιείται μέσα από μια διαδικασία διαπραγμάτευσης μεταξύ των επενδυτών και των δανειοληπτών. Τα επιτόκια δανεισμού όμως, είναι συνήθως πιο χαμηλά από ότι τα αντίστοιχα τραπεζικά δάνεια και ταυτόχρονα οι απολαβές των επενδυτών μεγαλύτερες αφού υπάρχει η έλλειψη μεσαζόντων. Αυτό

οφείλεται, ωστόσο και στο χαμηλό κόστος λειτουργίας των πλατφορμών που είναι κατά πολύ μικρότερο από αυτό των τραπεζικών συστημάτων.

Ένα ακόμα σημαντικό πλεονέκτημα, αποτελεί η παροχή δανείου σε άτομα που δεν έχουν τη δυνατότητα να δανειοδοτηθούν από τράπεζες λόγω έλλειψης φερεγγυότητας και κάλυψης των αυστηρών κριτηρίων που τίθενται από τα τραπεζικά συστήματα. Η καινοτομία του Peer-to-Peer Lending καταργεί το συγκεκριμένο εμπόδιο, δίνοντας την ευκαιρία σε οποιοδήποτε άτομο ή μικρή επιχείρηση που επιθυμεί να λάβει δάνειο ταιριάζοντας δανειζόμενους με δανειστές που έχουν το περιθώριο να επενδύσουν σε δάνεια μεγαλύτερου κινδύνου.



Σχήμα 1.1: Διαδικασία Δανεισμού μέσω πλατφόρμας Peer-to-Peer Lending.



Σχήμα 1.2: Διαδικασία δανεισμού μέσω Τράπεζας.

Από την άλλη πλευρά, το νέο αυτό εγχείρημα εμπεριέχει αρκετούς κινδύνους που ελλοχεύουν κυρίως από την πλευρά των επενδυτών, καθώς οποιαδήποτε επένδυση ενέχει τον κίνδυνο της ζημίας.

Πιο συγκεκριμένα, η πλειοψηφία των χρηστών που ζητούν δάνεια από πλατφόρμες τύπου Peer-to-Peer Lending, δεν έχουν τη δυνατότητα να λάβουν χρηματοδότηση μέσω του τραπεζικού συστήματος. Είναι φανερό, ότι τέτοιοι χρήστες έχουν αρκετά μειωμένη πιστοληπτική ικανότητα. Έτσι, υπάρχει πάντα η πιθανότητα να μην καταβληθεί όλο το ποσό από την πλευρά των δανειζόμενων και τότε οι δανειστές να ζημιωθούν από αυτή τους την επένδυση. Ταυτόχρονα, σε τέτοιου είδους συμφωνίες δεν υπάρχουν χρηματικές εγγυήσεις σε περιπτώσεις μη αποπληρωμής του δανείου. Αυτό είναι και το βασικό μειονέκτημα του Peer-to-Peer Lending.

Επιπρόσθετα, το ίδιο μειονέκτημα προκύπτει και στην περίπτωση χρεοκοπίας της ίδιας της πλατφόρμας Peer-to-Peer Lending. Στις περισσότερες των περιπτώσεων, η εγγραφή σε τέτοιου είδους πλατφόρμες είναι αρκετά μικρή και το ποσοστό κέρδους που λαμβάνουν από κάθε δάνειο είναι σχεδόν μηδαμινό. Αυτό έχει ως αποτέλεσμα, αρκετές πλατφόρμες Peer-to-Peer Lending να μην έχουν αρκετά κέρδη ώστε να μπορούν να υποστηρίξουν τη λειτουργία τους και έτσι χρεοκοπούν σαν νεοσύστατες μικρομεσαίες επιχειρήσεις.

Τέλος, χρήστες με χαμηλή πιστοληπτική αξιοπιστία που κατέχουν υψηλούς δείκτες επικινδυνότητας, αποτελούν επιλογές επένδυσης χαμηλότερης κατάταξης από άτομα με αυξημένη αξιοπιστία. Οι επενδυτές επιλέγουν στόχους με απόδοση και κίνδυνο που μπορούν να ανταπεξέλθουν. Αυτό σημαίνει ότι, υπάρχει ο κίνδυνός να μην γίνει η αντιστοίχιση του πλήρες ποσού που ζητάνε οι πιθανοί δανειολήπτες, με το ποσό που προσφέρεται από τους δανειστές και έτσι να μην επιτευχθεί η συμφωνία του δανείου τους.

## 1.6 ΠΑΡΟΥΣΙΑΣΗ ΠΛΑΤΦΟΡΜΩΝ

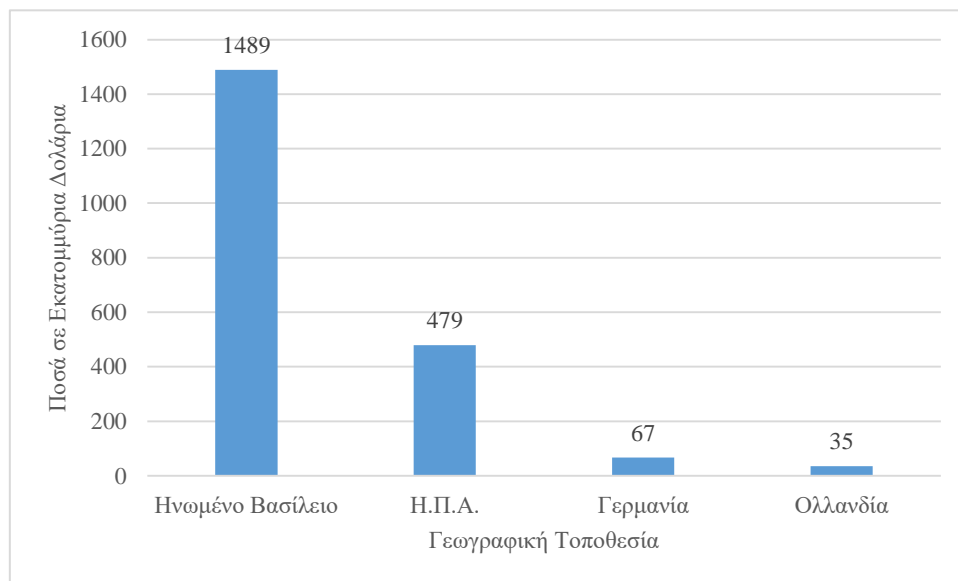
Σε αυτή την ενότητα θα γίνει παρουσίαση των πιο βασικών πλατφορμών ανά γεωγραφική περιοχή του Peer-to-Peer Lending με στόχο να αναδειχθούν και στην πράξη οι βασικές μορφές ως προς την αξιοποίηση της ιδέας του Peer-to-Peer Lending. Πιο συγκεκριμένα, θα παρατεθούν οι πιο δημοφιλείς πλατφόρμες που αποτελούν τις εναλλακτικές επιλογές ενός δανειολήπτη και ενός επενδυτή ταυτόχρονα. Ταυτόχρονα, θα παρατείνονται ορισμένα στατιστικά διαγράμματα, ώστε να γίνει πιο κατανοητή η ευρεία εξάπλωση της συγκεκριμένης καινοτομίας.

### Ηνωμένο Βασίλειο

- **Zopa.com:** Η συγκεκριμένη ιστοσελίδα αποτελεί την πρώτη ηλεκτρονική πλατφόρμα Peer-to-Peer Lending η οποία ξεκίνησε το 2005 από το Ηνωμένο Βασίλειο. Αποτελεί την πιο δημοφιλή ιστοσελίδα ηλεκτρονικού δανεισμού. Η φιλοσοφία της κινείται γύρω από το εμπορικό κομμάτι που λειτουργεί και μια τράπεζα. Σε αυτή την περίπτωση, δεν υπάρχει το κομμάτι του πλειστηριασμού για το επιτόκιο δανείου αλλά αντιστοίχια δανειζόμενου με πλήθος δανειστών που κινούνται στο ίδιο μήκος κύματος

επιτοκίου. Έχει μέγιστη διάρκεια περάτωσης συμφωνίας 24 ώρες, δέσμευση χρημάτων για τους επενδυτές 3 έως 5 χρόνια και 1% του ποσοστού των χρημάτων τους ως ποινή σε περίπτωση που εγκαταλείψουν πριν από το συμφωνηθέν χρονικό περιθώριο. Επιπρόσθετα, τα δάνεια προς υλοποίηση είναι της τάξης των £1000 έως £25000 με τους επενδυτές να έχουν ελάχιστο ποσό κατάθεσης τα £10. Τέλος, αξίζει να σημειωθεί ότι θέτει ένα επιπλέον μηνιαίο αντίτιμο περάτωσης του δανείου το οποίο επιστρέφεται στην λήξη του συμβολαίου. Τέλος, αξίζει να σημειωθεί ότι το zora.com έχει δανείσει μέχρι στιγμής £3.41 δισεκατομμύρια σε δανειολήπτες στο Ηνωμένο Βασίλειο καθώς και τους τελευταίους 12 μήνες έχει δανείσει σχεδόν £996 εκατομμύρια στοιχεία που προκύπτουν από την ίδια την πλατφόρμα ηλεκτρονικής δανειοδότησης zora.com.

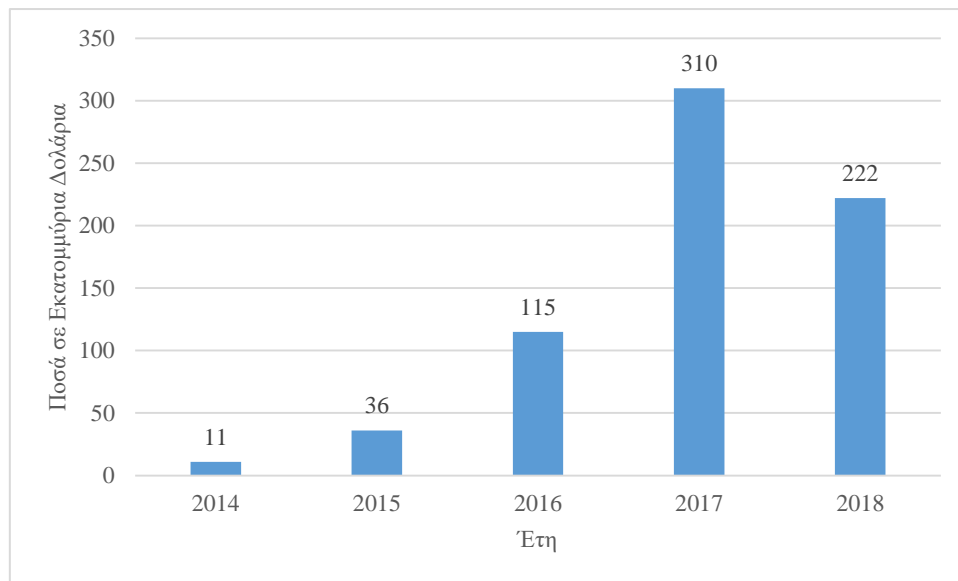
- **Funding Circle.com:** Μετά τη δημιουργία της P2P Lending αγοράς και πιο συγκεκριμένα την έναρξη λειτουργίας της πλατφόρμας Zora.com η αμέσως επόμενη πιο γνωστή πλατφόρμα στο Ηνωμένο Βασίλειο είναι η Funding Circle.com η οποία ιδρύθηκε το 2010. Η φιλοσοφία της ήταν επικεντρωμένη στο κοινό των μικρομεσαίων επιχειρήσεων θέλοντας να διευρύνει τους ορίζοντες της σε πολλά γεωγραφικά επίπεδα. Στο καθοριστικό κομμάτι του επιτοκίου δανεισμού χρησιμοποιεί τους δείκτες επικινδυνότητας (όπως αναλύθηκαν παραπάνω) με τη διαφορά ότι οι δανειζόμενοι καθορίζουν το ποσό δανεισμού και οι δανειστές καθορίζουν το ποσό που θα δανείσουν και το επιτόκιο τους που προκύπτει μέσω πλειστηριασμού και των δεικτών επικινδυνότητας. Στο διάγραμμα 1.7 αναγράφονται τα ποσά υπό τη διαχείριση του fundingcircle.com ανά γεωγραφική τοποθεσία όπως αυτά αναφέρονται στην ετήσια αναφορά του ίδιου του ιστοτόπου.



Διάγραμμα 1.7: Ποσά σε Εκατομμύρια Δολάρια που διαχειρίστηκε το 2017 η πλατφόρμα fundingcircle.com ανά γεωγραφική περιοχή.

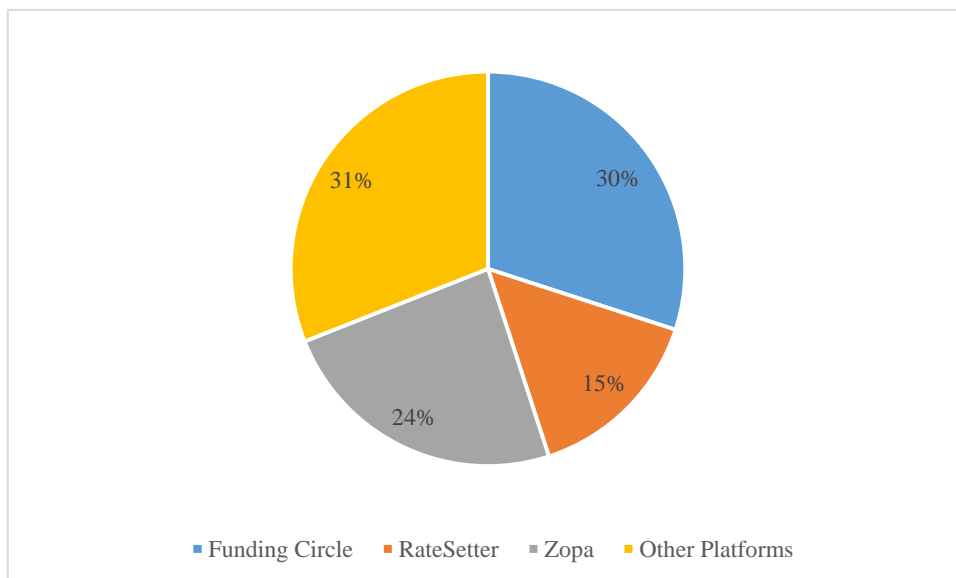


- **Rate Setter:** Ο συγκεκριμένος ιστότοπος λειτουργεί με ένα διαφορετικό τρόπο από ότι η πλειοψηφία των υπολοίπων αφού δε βασίζει τη δημιουργία συμφωνίας στους δείκτες επικινδυνότητας αφού δέχεται μόνο δανειολήπτες που ανήκουν κάτω από το μέσω όρο πιστωτικού κινδύνου. Καλύπτει ένα ευρύ φάσμα κοινού που απευθύνεται και ως μέτρο καθορισμού του επιτοκίου λαμβάνει τη διάρκεια του προβλεπόμενου δανείου. Στο διάγραμμα 1.8 παρουσιάζονται τα ποσά των δανείων που εκπληρώθηκαν μέχρι και σήμερα στη συγκεκριμένη ιστοσελίδα όπως αυτά προκύπτουν από το χαρτοφυλάκιο της.



Διάγραμμα 1.8: Ποσά σε Εκατομμύρια Δολάρια που εκπόνησε η ιστοσελίδα ratesetter.com ανά έτος μέχρι και σήμερα.

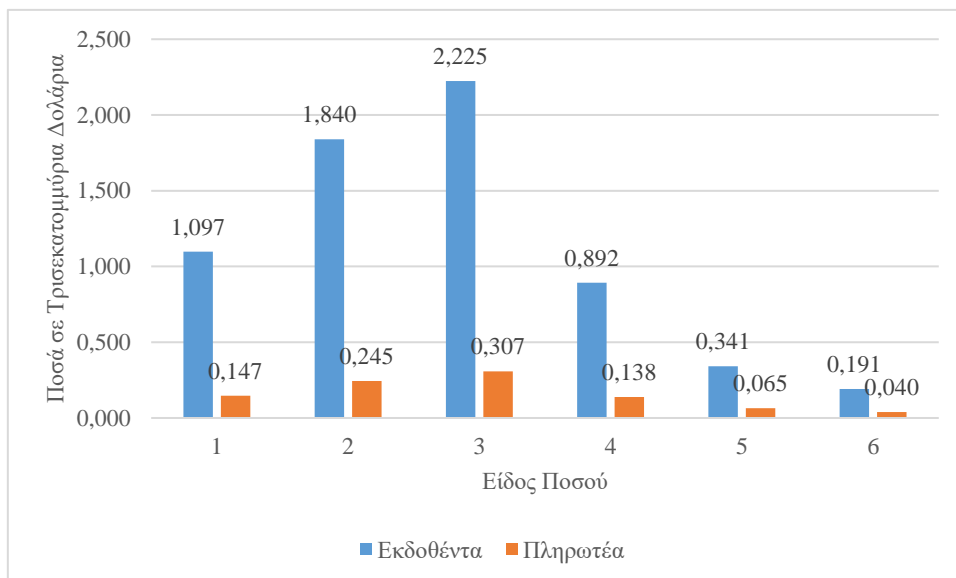
Τέλος, στο διάγραμμα 1.9 φαίνεται το μερίδιο της αγοράς που ικανοποιούν σε ηλεκτρονικά δάνεια οι πλατφόρμες κοινωνικής δικτύωσης τύπου Peer-to-Peer Lending στο Ηνωμένο Βασίλειο κατά το έτος 2017 σύμφωνα με την έρευνα του Jordan Stodart στην ετήσια αναφορά για το orcamoney.com.



Διάγραμμα 1.9: Μερίδιο αγοράς πλατφορμών Peer to Peer Lending στο Ηνωμένο Βασίλειο.

#### Η.Π.Α.

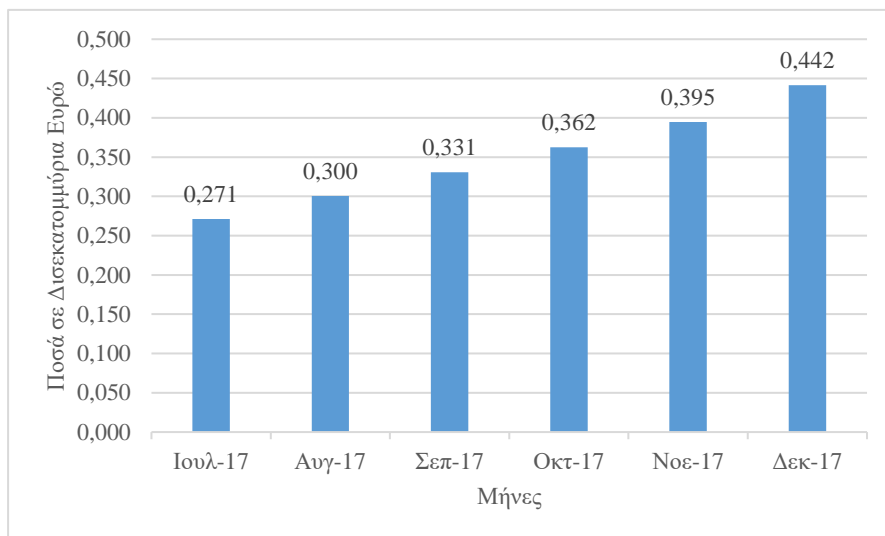
- **Prosper.com:** Αποτελεί την πρώτη πλατφόρμα Peer-to-Peer Lending που ιδρύθηκε στις Η.Π.Α. το 2005. Στην αρχή της λειτουργίας της πλατφόρμας, η φιλοσοφία της ομοιάζε με τον ορισμό που δόθηκε για την έννοια του Peer-to-Peer Lending. Στη συνέχεια όμως, υιοθέτησε διαφορετική τακτική. Με αρωγό τα εμπειρικά δεδομένα που κατέχει και με βάση την πιστοληπτική ικανότητα των αιτούντων δημιουργεί η ίδια το επιτόκιο δανεισμού, διαθέτοντας και δυο σταθερά ποσοστά για τις ακραίες περιπτώσεις του εντελώς αξιόπιστου πιθανού δανειολήπτη και του εντελώς αναξιόπιστου της τάξης του 6% και 30% αντίστοιχα. Τέλος, η διάρκεια επίτευξης συμφωνίας διαρκεί έως και 14 ημέρες καθώς και τα ποσά που πραγματεύονται για δανεισμό κυμαίνονται από \$2000 έως \$35000.
- **Lending Club.com:** Η συγκεκριμένη πλατφόρμα ιδρύθηκε το 2006 και αποτελεί μέχρι και σήμερα την πιο δημοφιλή πλατφόρμα Peer-to-Peer Lending. Επικεντρώνεται κυρίως σε καταναλωτικά και φοιτητικά δάνεια και επιπλέον, όπως οι περισσότερες πλατφόρμες τέτοιου είδους στις Η.Π.Α. επικεντρώνονται κυρίως στην καταναλωτική πίστη του κοινού στο οποίο στοχεύουν. Σε αυτό το σημείο αξίζει να σημειωθούν τα ποσά των δανείων που έχουν εκδοθεί το έτος 2017 καθώς και τα ποσά που έχουν ήδη πληρωθεί το ίδιο έτος ανά δείκτη φερεγγυότητας στο διάγραμμα 1.10 από στοιχεία που προσφέρει η ίδια πλατφόρμα του lendingclub.com.



Διάγραμμα 1.10: Ποσά ηλεκτρονικών στην πλατφόρμα Lending Club κατά το έτος 2017.

## Ευρώπη

- Mintos.com:** Πρόκειται για μια πλατφόρμα Peer-to-Peer Lending που ιδρύθηκε το 2015 και έχει ως έδρα τη Λετονία. Στην προκειμένη περίπτωση, η ιστοσελίδα κατηγοριοποιεί τα δάνεια ανά γεωγραφική περιοχή και είδος προσφέροντας στον επενδυτή τη δυνατότητα επιλογής σε μια μεγάλη αγορά από τέτοιου είδους προϊόντα. Αυτή είναι και η βασική διαφορά της συγκεκριμένης πλατφόρμας με τις υπόλοιπες. Αξίζει να σημειωθεί επίσης, ότι η συγκεκριμένη ιστοσελίδα διαθέτει μια ιδιαίτερη διαφορετική από τις υπόλοιπες πλατφόρμες πολιτική σε σχέση με τα δάνεια που δεν εκπληρώνονται. Πιο συγκεκριμένα δέχεται η πλειοψηφία των δανείων της εξασφαλίζεται με την εφαρμογή υποθήκης περιουσιακών στοιχείων, μια πολιτική που συναντάται κυρίως σε τραπεζικά συστήματα. Το mintos.com έχοντας ως διάρκεια δραστηριοποίησης 3 περίπου χρόνια παρέχει δεδομένα για τα δάνεια που έχουν οριστικοποιηθεί στη συγκεκριμένη πλατφόρμα έως τώρα. Στο διάγραμμα 1.11 φαίνονται τα συνολικά ποσά από τα δάνεια που έχουν χρηματοδοτηθεί στο τελευταίο εξάμηνο του έτους 2017 σε ευρώ.



Διάγραμμα 1.11: Ποσά δανείων που έχουν χρηματοδοτηθεί στην πλατφόρμα mintos.com το τελευταίο εξάμηνο του 2017.

### Άλλες Χώρες

Στην Κίνα οι εξελίξεις είναι το ίδιο ραγδαίες με τις Η.Π.Α. αλλά λιγότερο ελεγχόμενες ως προς την επικινδυνότητά του. Πιο αναλυτικά, σύμφωνα με τα στατιστικά στοιχεία του wdzj.com το Μάρτιο του 2017 υπήρχαν περίπου 9 χιλιάδες πλατφόρμες Peer-to-Peer Lending όπου δραστηριοποιούνταν στο συγκεκριμένο χώρο. Παρόλα αυτά, το ξεκίνημα έγινε με την Pairaidai το 2007. Η ηλεκτρονική δανειοδότηση στην Κίνα όμως ξεκίνησε από τα ήδη καθιερωμένα χρηματοπιστωτικά ιδρύματα που εξελίχθηκαν με την ένταξη τους στο διαδίκτυο και από άλλους ιδιωτικούς φορείς που εντάχθηκαν αργότερα στο διαδίκτυο περιλαμβάνοντας μέσα σε αυτούς και τις πλατφόρμες του Peer-to-Peer Lending. Στους δυο αυτούς τους κλάδους όμως φαίνεται ότι υπερέχει ο δεύτερος αφού υπερέχει στην ευελιξία ως προς τις συναλλαγές που πραγματοποιεί και τις καινοτομίες που εντάσσει ολοένα και περισσότερο στους κόλπους του. Οι πλατφόρμες αυτές όμως, ενέχουν περισσότερους κινδύνους στο κομμάτι της μη εκπλήρωσης δανείων καθώς δεν εποπτεύονται πλήρως από τις ρυθμιστικές αρχές της Κίνας.

Τέλος, ένα ξεχωριστό παράδειγμα που θα πρέπει επίσης να περιληφθεί στο κομμάτι αυτό είναι η περίπτωση του kiva.org το οποίο αποτελεί διεθνή πλατφόρμα ηλεκτρονικής χρηματοδότησης. Πρόκειται για μια διαφορετική περίπτωση πλατφόρμας από την προηγούμενη αφού αφορά έναν μη κερδοσκοπικό οργανισμό. Αυτό συνεπάγεται ότι, κανένα ποσοστό δε χορηγείται στον φορέα αυτόν ως επιπλέον κέρδος. Το Kiva, καλύπτει τα έξοδα συντήρησής του μέσω εθελοντικών δωρεών αφού έχει φιланθρωπικό σκοπό. Πιο συγκεκριμένα, οι δανειστές χορηγούν χρήματα σε εγχειρήματα, ιδέες και μεμονωμένα άτομα που έχουν ανάγκη χρηματοδότησης για καλύτερες συνθήκες διαβίωσης. Αυτό, επιτυγχάνεται με ένα δίκτυο συνεργατών του Kiva, μέσω του οποίου γίνεται και η αποπληρωμή του δανείου με σχεδόν μηδενικό επιτόκιο το οποίο καλύπτει ορισμένες υποχρεώσεις της υλοποίησης του δανείου από τους συνεργάτες του Kiva. Τονίζεται όμως ότι, αυτή η διαδικασία αφορά δάνεια και όχι δωρεές αφού τα χρηματικά ποσά επιστρέφοντα στο πλήρες ποσοστό τους και αρκετές φορές με κάποιο συμβολικό επιτόκιο.

# ΚΕΦΑΛΑΙΟ 2

## ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Στη συνέχεια της θεωρητικής ανάλυσης, που πραγματοποιήθηκε στο πρώτο κεφάλαιο, προκειμένου να γίνει σαφής η σημαντικότητα της ανάπτυξης του FinTech και πιο συγκεκριμένα της εφαρμογής του Peer-to-Peer Lending, γίνεται μετάβαση στο δεύτερο κεφάλαιο της εργασίας, όπου αναλύεται μέσα από μια ανασκόπηση στη διεθνή βιβλιογραφία σχετικά με το τι κέντρισε το ενδιαφέρον του ερευνητικού τομέα στο συγκεκριμένο χώρο για περαιτέρω ανάλυση. Με την εξάπλωση της εφαρμογής του Peer-to-Peer Lending, το ενδιαφέρον στράφηκε στην αξιοπιστία που παρέχουν οι νέοι αυτοί ιστότοποι αλλά και στους κινδύνους που ελλοχεύουν μέσα από τις συναλλαγές που πραγματοποιούνται. Στο επιστημονικό άρθρο των Pokorná και Sponer (2016) αναφέρεται ότι το κύριο μειονέκτημα της επικεντρώνεται στην ασυμμετρία της πληροφόρησης που υπάρχει μεταξύ πιστωτών και δανειοληπτών, η οποία δυσχεραίνει την αξιολόγηση της πιστοληπτικής αξιοπιστίας των δανειζόμενων από την πλευρά των πιστωτών.

Επικεντρώνοντας έτσι διάφορα μειονεκτήματα που πηγάζουν από τις συναλλαγές μέσα από τις συγκεκριμένες ιστοσελίδες, έγινε μια προσπάθεια ανάλυσης των παραγόντων που επηρεάζουν την αξιοπιστία ενός δανείου και πιο συγκεκριμένα της πιθανότητα αθέτησης του αλλά και άλλων μέσων που μπορούν να ελαχιστοποιήσουν αυτά τα σημαντικά μειονεκτήματα.

Από πιο παλιές έρευνες όπως αυτή των Stiglitz και Weiss (1981) είχε παρατηρηθεί η αύξηση της ασυμμετρίας της πληροφόρησης που είχαν οι τράπεζες παρά το γεγονός ότι διέθεταν αναλυτικό ιστορικό για τους πελάτες τους. Ο Narain (1992) ήταν ο πρώτος που εισήγαγε την εφαρμογή των Survival Analysis για να περιγράψει τον πιστωτικό κίνδυνο στα τραπεζικά ιδρύματα και ώστε να αναπτυχθούν πιο διεξοδικά οι μέθοδοι πρόβλεψης αθέτησης και επικινδυνότητας ενός δανειολήπτη, ώστε να διευκολύνεται περισσότερο η λήψη αποφάσεων για τη δανειοδότηση ενός πελάτη από οποιοδήποτε χρηματοπιστωτικό ίδρυμα.

Με τη βοήθεια του μεγάλου όγκου δεδομένων που παρέχονται από τέτοιου είδους συναλλαγές πριν ακόμα από την ανάπτυξη του Peer to Peer Lending, με τη βοήθεια των τεχνικών εκμάθησης μηχανών το ενδιαφέρον στράφηκε ολοένα και περισσότερο στη σύγκριση μεθόδων ταξινόμησης και παλινδρόμησης που είναι απαραίτητα για τη δημιουργία μοντέλων πρόβλεψης. Στην έρευνα του οι Bellotti και Crook (2009) έκαναν μια σύγκριση ανάμεσα στις πολύ δημοφιλείς μεθόδους Support Vector Machine, Logistic Regression, και K-Nearests Neighbours δημιουργώντας 10 που εμπεριέχουν 100 επιτυχημένες και 100 αποτυχημένες υποθέσεις δανείων η κάθε μία από δεδομένα που λήφθηκαν από χρηματοπιστωτικό ίδρυμα και θέτοντας ένα διαφορετικό κατώφλι για κάθε μέθοδο για τη λήψη απόφασης του αποτελέσματος. Τέλος με τη βοήθεια της

καμπύλης ROC μετριέται η απόδοση των αποτελεσμάτων για τους αλγόριθμους, επιτυγχάνοντας με αυτόν τον τρόπο και τη μεταξύ τους σύγκριση και με τη βοήθεια της μεθόδου μέγιστης πιθανοφάνειας προσεγγίστηκαν οι πιο σημαντικές μεταβλητές για τα μοντέλα. Αποδείχτηκε με αυτό τον τρόπο ότι η μέθοδος SVM είναι εξίσου αποδοτική με τις υπόλοιπες δυο δημοφιλής μεθόδους και ικανή να δείχνει ποιοι παράγοντες επηρεάζουν την πιθανότητα αθέτησης ενός δανείου.

Ως προς το χώρο του Peer to Peer Lending ακολούθησε μια ίδια πορεία ως προς τον πιστωτικό κίνδυνο και την αξιολόγηση της πιστοληπτικής ικανότητας. Πιο συγκεκριμένα, το επιστημονικό ενδιαφέρον προσελκύστηκε αρχικά στην αντιστοιχία της χρηματοδότησης του Peer to Peer Lending με τα δάνεια των τραπεζών. Μια κλασσική προσέγγιση είναι και αυτή του Byanjankar (2017) η οποία λαμβάνοντας δεδομένα από τον ιστότοπο bondora.com, χρησιμοποιώντας μια ανάλυση επιβίωσης προβλέπει την πιθανότητα επιτυχίας ενός δανείου σε διαφορετικές χρονικές περιόδους πετυχαίνοντας να οριοθετήσει με αυτόν τον τρόπο συγκεκριμένα χρονικά διαστήματα οι οποίες θεωρούνται πιο ασφαλής για την ολοκλήρωση ενός δανείου. Χρησιμοποιώντας τον εκτιμητή Kaplan Meier συνδυαστικά με το αναλογικό πρότυπο κινδύνου Cox, εξήγαγε την καμπύλη επιβίωσης σε διάστημα εμπιστοσύνης 95% από το οποίο προκύπτει ότι ένας μέσος πελάτης έχει πιθανότητα επιβίωσης 0.5 σε χρονικό ορίζοντα 24 μηνών. Επίσης, απέδειξε ότι οι νέοι πελάτες έχουν διάμεσο χρονικό διάστημα 21 μηνών το οποίο είναι μικρότερο σε σύγκριση με τους παλιούς πελάτες που έχουν διάμεσο διάστημα επιβίωσης τους 37 μήνες. Επίσης, προκύπτει ότι, οι νέοι πελάτες έχουν υψηλότερο ποσοστό αθέτησης από τους παλιούς πελάτες. Ορισμένες στατιστικές αναλύσεις ανάλογα με το φύλο και τη χώρα έδειξαν ότι οι γυναίκες έχουν διάμεσο διάστημα επιβίωσης 30 μήνες σε αντίθεση με τους άντρες που έχουν 22 μήνες ενώ οι πελάτες της Εσθονίας έχουν διάμεσο διάστημα επιβίωσης τους 46 μηνών έναντι της Σλοβακίας που έχουν 7 μήνες. Στο τέλος της ανάλυσης, με τη βοήθεια του μοντέλου αναλογικού κινδύνου Cox προκύπτει ο πίνακας που δείχνει την επίδραση των μεταβλητών που χρησιμοποιεί η ανάλυση στην πιθανότητα αθέτησης ενός δανείου καταλήγοντας στο γεγονός ότι η μεταβλητή που δείχνει τη χώρα του πελάτη έχει τη μεγαλύτερη επιρροή σε ένα πτωχευμένο δάνειο.

Στην έρευνα που πραγματοποίησαν οι Cinca et al. (2015) παρουσίασαν ορισμένες μεταβλητές από τα δεδομένα που χρησιμοποιεί η ιστοσελίδα Lending Club, εξετάζοντας 24449 περιπτώσεις δανείων ώστε να εξάγουν συμπεράσματα σχετικά με τις μεταβλητές που επηρεάζουν περισσότερο ή λιγότερο την αθέτηση ενός δανείου. Συμπερασματικά, χρησιμοποιώντας στατιστικές αναλύσεις όπως τον δείκτη συσχέτισης Pearson, το συντελεστή συσχέτισης Point-Biserial και τις ελέγχους  $\chi^2$  κατέληξαν σε ορισμένες συσχετίσεις μεταξύ των μεταβλητών που χρησιμοποιήθηκαν. Τα πιο επικίνδυνα δάνεια είναι τα δάνεια μικρών επιχειρήσεων, τα πιο ασφαλή στη συγκεκριμένη περίπτωση τα γαμήλια δάνεια καθώς επίσης τα πιο ασφαλή δάνεια προέρχονται από άτομα που έχουν δηλώσει κατάσταση σπιτιού με υποθήκη και ιδιοκτησία. Έπειτα, απέδειξαν ορισμένες υποθέσεις, που εξάγουν σημαντικά αποτελέσματα όπως ότι όσο μεγαλύτερο το επιτόκιο τόσο μεγαλύτερη και η πιθανότητα αθέτησης και ότι ο σκοπός του δανείου και τα χαρακτηριστικά του δανειζόμενου παίζουν ρόλο στην πιθανότητα αθέτησης, ενώ η διάρκεια και το ποσό του δανείου όχι. Στη συνέχεια με τη βοήθεια της παλινδρόμησης τύπου Cox εξήχθησαν ποσοτικά και διαγραμματικά συμπεράσματα για 33 μεταβλητές

όπου δίνουν ποσοτικά σε μια δεδομένη χρονική στιγμή κατά πόσο μικρότερος ή υψηλότερος είναι ο κίνδυνος αθέτησης όταν μια μεταβλητή μεταβάλλεται ενώ οι υπόλοιπες είναι σταθερές. Τέλος, δημιουργώντας 7 μοντέλα λογιστικής παλινδρόμησης αποδείχτηκε ότι, η μεταβλητή grade που χρησιμοποιείται από την ίδια την ιστοσελίδα για την βαθμολόγηση των δανειοληπτών ως προς την επικινδυνότητά τους, αποτελεί τη μεταβλητή με τη μεγαλύτερη προβλεπτική ικανότητα.

Στην έρευνα του οι Yum et al. (2012), χρησιμοποιώντας δεδομένα από την πλατφόρμα porfunding.com, και την τεχνική ψήφων αξιοπιστίας μεταξύ δανειστών, που εφαρμόζεται στη συγκεκριμένη πλατφόρμα, εκμεταλλευόμενη ως φιλοσοφία της συλλογικής ευφυΐας, δημιούργησαν δυο κατηγορίες υποθέσεων και με τη βοήθεια του αλγορίθμου της λογιστικής παλινδρόμησης κατέληξαν στα ακόλουθα συμπεράσματα για την πιστοληπτική αξιοπιστία των πιθανών δανειοληπτών. Συγκεκριμένα, η ανάλυση έδειξε ότι, οι πιστωτές δίνουν περισσότερη έμφαση στην καταγραφή των ψήφων εμπιστοσύνης όταν οι πιθανοί δανειολήπτες δεν έχουν υπάρχον ιστορικό και αυτό έχει θετική επίδραση στην επιτυχία του δανείου που αιτούνται, ενώ στην περίπτωση που οι δανειολήπτες αποδεικνύουν την αξιοπιστία τους με τις ήδη υπάρχουσες συναλλαγές τους τότε οι πιστωτές δεν δίνουν σημασία στη συλλογική νοημοσύνη. Ταυτόχρονα, οι πολλές αιτήσεις δανείων λειτουργούν αρνητικά στη χορήγηση δανείου, δηλαδή θεωρούνται ως αρνητικό χαρακτηριστικό αξιοπιστίας, ενώ στην περίπτωση που ο πιθανός δανειολήπτης έχει επενδύσει έστω και μια φορά, προσδίδει θετικά χαρακτηριστικά ως προς την αξιοπιστία του. Από την άλλη, αξίζει να σημειωθεί ότι ο αριθμός των αιτήσεων δανείου που προκύπτει από το κοινό που έχει ήδη εκπληρώσει κάποια συναλλαγή στο παρελθόν, δηλαδή έχει ήδη κάποιο ιστορικό έχει θετική επίδραση στο αποτέλεσμα μιας μελλοντικής συναλλαγής, ενώ ο όγκος των χρημάτων που τους έχουν χορηγηθεί στο παρελθόν δε δίνει κάποιο στοιχείο για την πιθανότητα επιτυχής έκβασης ενός μελλοντικού δανείου. Στη συγκεκριμένη έρευνα, ο αριθμός καθυστερήσεων σε προηγούμενες συναλλαγές καθώς και η διάρκεια καθυστέρησης δε δείχνουν να έχουν κάποια επιρροή στην πιθανότητα επιτυχίας ενός μέλλοντος δανείου. Τέλος, οι πρόωρες αποπληρωμές έχουν θετικό αντίκτυπο στην επιτυχία ενός μελλοντικού δανείου όπως και οι πληροφορίες μέσω ενός συστήματος ερωτήσεων- απαντήσεων που προκύπτουν κατά την περίοδο μιας καθυστερημένης πληρωμής.

Στη μεταπτυχιακή διατριβή της Polena (2017), λαμβάνοντας δεδομένα 23 μεταβλητών από τον ιστότοπο του Lending Club, συγκρίνονται δέκα αλγόριθμοι ταξινόμησης (Logistic Regression, Linear Discriminant Analysis, Support Vector Machine, Artificial Neural Network, k- Nearest Neighbors, Naïve Bayes και Bayesian Network, Classification and Regression Tree και Random Forest) οι οποίοι εξετάζονται ως προς την προβλεπτική τους ικανότητα με 6 διαφορετικές μετρήσεις (Percentage Correctly Classified, Kolmogorov-Smirnov Statistic, Brier Score, Area Under Curve, Partial Gini Index και H-Measure ). Τα αποτελέσματα στα οποία καταλήγει η έρευνα, κατατάσσουν ως καλύτερες μεθόδους τη λογιστική παλινδρόμηση, τα τεχνητά νευρωνικά δίκτυα και την γραμμική διακριτική ανάλυση. Στη συνέχεια της ταξινόμησης, κατατάσσεται η μέθοδος Random Forests, ενώ τις τελευταίες θέσεις προτίμησης κατέχουν οι υπόλοιπες μέθοδοι.

Στην εκτεταμένη έρευνα των Lessmann et al. (2015), εξετάζονται 41 αλγόριθμοι ταξινόμησης οι οποίοι συγκρίθηκαν σε 8 σύνολα δεδομένων αξιολόγησης με 6 μέτρα προβλεπτικής ικανότητας. Τα αποτελέσματα που προέκυψαν ήταν αρκετά ενδιαφέροντα αφού αποδείχτηκε ότι μερικοί μη δημοφιλείς αλγόριθμοι ταξινόμησης παρείχαν πιο ακριβή αποτελέσματα από άλλους πιο διαδεδομένους όπως η λογιστική παλινδρόμηση, πιο συγκεκριμένα η μέθοδος Artificial Neural Network αποδείχτηκε καλύτερη από τη Logistic Regression.

Σε μια διαφορετική σκοπιά οι Tsai Et al. (2014), χρησιμοποίησαν δεδομένα του ιστοτόπου Lending Club και έγινε σύγκριση μεταξύ 4 αλγορίθμων ταξινόμησης (Modified Logistic Regression, LibSVM, Naïve Bayes και Random Forest). Με τη βοήθεια της μεθόδου Principal Component Analysis, με τη μείωση του συντελεστή αθέτησης των δανειοληπτών και την επίτευξη της υψηλότερης προβλεπτικής ικανότητας, προέκυψαν τα εξής συμπεράσματα. Στην αρχή επετεύχθη μια σημαντική και μεγάλη ακρίβεια ως προς τον αλγόριθμο της λογιστικής παλινδρόμησης σε σύγκριση με τους υπόλοιπους 3. Τέλος, τονίστηκε ότι η μεγάλη ακρίβεια ήταν αποτέλεσμα της αύξησης ενός συντελεστή ποιότητος που είχαν θέσει εξαρχής πράγμα που επιβεβαίωσε το αποτέλεσμα.

Από την άλλη πλευρά η εργασία των Puro et al. (2010), λαμβάνοντας δεδομένα από τον ιστότοπο Prosper.com, εστιάζει στην ποσοτικοποίηση-αξιολόγηση των στρατηγικών που χρειάζεται να κάνουν οι δανειολήπτες ώστε να λάβουν χρηματοδότηση. Αρχικά με μια σύντομη ανάλυση συσχετίσεων εξάγεται το συμπέρασμα ότι οι μεταβλητές αρχικό επιτόκιο (που αρχικοποιείται από τον αιτούμενο δανείου), αιτούμενο ποσό, βαθμός πιστοληπτικής ικανότητας, δείκτης ποσό δανείου προς εισόδημα και ο αριθμός των καθυστερήσεων που αναφέρονται στο φάκελο πίστωσης των δανειοληπτών (μεταβλητές που ο δανειζόμενος μπορεί να μεταβάλλει) είναι οι πιο σημαντικές. Πιο αναλυτικά, αποδεικνύεται η θετική σχέση μεταξύ του αρχικού επιτοκίου που ζητά ο δανειολήπτης και του τελικού επιτοκίου του δανείου που έρχονται όμως σε αντιπαράθεση με την πιθανότητα επιτυχούς εξασφάλισης δανείου. Έτσι με τη βοήθεια της λογιστικής παλινδρόμησης δημιουργείται ένα μοντέλο-εργαλείο υποστήριξης αποφάσεων που εξετάζει τους δύο αυτούς όρους και ποσοτικοποιεί τη βέλτιστη λύση ανάλογα με το πρότυπο κάθε δανειολήπτη ή το αρχικό επιτόκιο αντίστοιχα.

Σε μια εκτεταμένη έρευνα των Serrano-Cinca et al. (2016), αποφεύγεται η επικέντρωση στην πιθανότητα αθέτησης και το ενδιαφέρον εστιάζεται στην πρόβλεψη της αναμενόμενης κερδοφορίας με τη βοήθεια του εσωτερικού βαθμού απόδοσης με στόχο τη δημιουργία ενός συστήματος υποστήριξης αποφάσεων για επενδύσεις στο Peer-to-Peer Lending με βάση το κέρδος. Η έρευνα χρησιμοποιώντας δεδομένα από την ιστοσελίδα του Lending Club και με τη βοήθεια μιας διερευνητικής ανάλυσης, που αποτελείται από μια πολυπαραγοντική ανάλυση τη μέθοδο CHAID και την τεχνική δέντρων αποφάσεων, οδήγησε στη δημιουργία ενός συστήματος λήψης αποφάσεων για την πρόβλεψη του εσωτερικού βαθμού απόδοσης του δανείου (εξαρτημένη μεταβλητή). Έπειτα, παρατίθενται οι παράγοντες που επηρεάζουν τον εσωτερικό βαθμό απόδοσης που τελικά είναι το επιτόκιο του δανειολήπτη, το χρέος του δανειολήπτη και ο σκοπός του δανείου. Τέλος, με το δέντρο αποφάσεων τύπου CHAID διαπιστώνεται ότι, η αγορά του Peer-to-Peer Lending δεν είναι μια πλήρως αποδοτική αγορά.



Στην έρευνα των Feis et al . (2016), οι ερευνητές με δεδομένα από την ιστοσελίδα του Lending Club προσπάθησαν να δημιουργήσαν ένα καινοτόμο μοντέλο πρόβλεψης αθέτησης δανείων με τον οποίο θα κατασκευαστεί μια καινούργια συμπληρωματική πλατφόρμα πάνω στο ήδη υπάρχον interface της υφιστάμενης ιστοσελίδας ώστε να λειτουργεί επικοδομητικά στον τρόπο κατάταξης των δανείων σε βαθμίδες αξιοπιστίας. Πιο συγκεκριμένα, η πλατφόρμα Lending Club λαμβάνοντας το ιστορικό των αιτούμενων δανείων κατατάσσει τις αιτήσεις σε βαθμίδες αξιοπιστίας οι οποίες αντιστοιχούν σε ένα συγκεκριμένο επιτόκιο έτσι για τη βαθμίδα A1(πιο ασφαλές δάνειο) το επιτόκιο θα είναι 5,32% και για τη βαθμίδα G5 (το λιγότερο ασφαλές δάνειο) το επιτόκιο θα είναι 30,99%. Η έρευνα αυτή εστιάζει στη δημιουργία ενός συνεχούς φάσματος επιτοκίων στην κάθε βαθμίδα λαμβάνοντας υπόψιν την υπολογισμένη πλέον αξία του δανείου σε δολάρια. Στην αρχή έγινε χρήση στατιστικών μεθοδολογιών όπως το δείκτη συσχέτισης Pearson για την εύρεση των μεταβλητών που ήταν αναγκαίες για τον υπολογισμό της πιθανότητας αθέτησης και έπειτα με τη μέθοδο Scikit Cross Validation και με τα μοντέλα μηχανικής μάθησης όπως η Logistic Regression, τη Linear Discriminant Analysis, τα Support Vector Machines και τα Ensemble Methods δημιούργησαν το προβλεπτικό μοντέλο κατάταξης των δανείων. Τέλος, δημιουργείται η επιπλέον προσθήκη της πλατφόρμας σε γλώσσα Python με την υποστήριξη του μοντέλου τιμολόγησης της αξίας του δανείου το οποίο προκύπτει από μια τροποποιημένη μορφή του μοντέλου Discounted Cash Flow και το οποίο αξιολογείται με τον δείκτη Sharpe ο οποίος εντέλει προκύπτει στην τιμή 3,41 έναντι της τιμής, που προέκυψε από την έρευνα μικρών δανείων μικρής αξίας με διάφορες αγορές, που είναι 0,3.

Στην τελευταία έρευνα της βιβλιογραφικής ανασκόπησης οι Malekipirbazari και Aksakalli (2015) το ενδιαφέρον επικεντρώθηκε στην εξαγωγή συμπερασμάτων για το υπολογισμό του κινδύνου των πιθανών δανειοληπτών. Όπως και στις προηγούμενες έρευνες που αναφέρθηκαν έτσι και σε αυτή έγινε καταγραφή των μεταβλητών της πλατφόρμας του Lending Club και στατιστική ανάλυση με υπολογισμούς όπως το Information Gain και τα Correlations όπου έγινε και η τελική καταγραφή των σημαντικών μεταβλητών του μαθηματικού μοντέλου. Στη συνέχεια, αναπτύσσεται το μαθηματικό μοντέλο που έχει σα σκοπό την εύρεση της καλύτερης εκτίμησης της πιθανότητας εύρεσης ενός φερέγγυου δανειολήπτη και έπειτα αυτό εκτελείται από τις μεθόδους Random Forest, Logistic Regression, k-Nearest Neighbors και Support Vector Machines. Τέλος, για την ταξινόμηση των μεθόδων που αναφέρθηκαν, ως προς την προβλεπτική του ικανότητα, χρησιμοποιήθηκε η μεθοδολογία των k-folders Cross Validation 5 φορές και έπειτα προέκυψαν οι προβλεπτικοί δείκτες overall classification accuracy, AUROC, RMSE, και ο πίνακας confusion, όπου προέκυψε ότι η μέθοδος Random Forests υπερέχει των υπολοίπων μεθόδων

Στη συνέχεια παρουσιάζεται ο συγκεντρωτικός πίνακας 2.1 ο οποίος περιέχει τις έρευνες που αναλύθηκαν στο παραπάνω τμήμα του κεφαλαίου ώστε να γίνεται πιο ευδιάκριτη η ανάλυση της βιβλιογραφικής ανασκόπησης.

	<u>Συγγραφέας</u>	<u>Λεδομένα</u>	<u>Σκοπός</u>	<u>Μέθοδοι</u>	<u>Αποτελέσματα</u>
1	Bellotti-Grook (2009)	Δεδομένα τραπεζικού συστήματος	Σύγκριση μεθόδου Support Vector Machine με άλλες δημοφιλής μεθόδους	SVM, Logistic Regression, k-Nearest Neighbours, μέθοδος μέγιστης πιθανοφάνειας	Τα αποτελέσματα των μεθόδων συμπίπτουν στην επιλογή των σημαντικών χαρακτηριστικών για το μοντέλο πρόβλεψης και οι μέθοδοι είναι εξίσου αποδοτικοί στην ταξινόμηση πετυχημένων και μη δανείων
2	Byanjankar (2017)	Bondora.com	Ανάλυση πιστωτικού κινδύνου, πρόβλεψη πιθανότητας επιβίωσης ενός δανείου σε διάφορα χρονικά διαστήματα	Εκτιμητής Kaplan Meier-Αναλυτικό πρότυπο κινδύνου Cox	Ανάπτυξη προγνωστικού μοντέλου για τον υπολογισμό της πιθανότητας πρόβλεψης σε συγκεκριμένα χρονικά διαστήματα
3	Cinca et al. (2015)	Lending Club	Προσδιορισμός των μεταβλητών που παίζουν σημαντικό ρόλο στην ταξινόμηση ενός δανείου σε ελαττωματικά ή όχι και προσδιορισμός της πιθανότητας αθέτησης ενός δανείου	Στατιστικές αναλύσεις όπως: Δείκτης συσχέτισης Pearson, συντελεστής συσχέτισης Point-Biserial, έλεγχοι $X^2$ , Cox Regression και Logistic Regression	Προέκυψε πίνακας των μεταβλητών που σχετίζονται με την πιθανότητα αθέτησης καταλήγοντας στο ότι η μεταβλητή grade αποτελεί τη μεταβλητή με τη μεγαλύτερη προβλεπτική ικανότητα
4	Yum et al. (2012)	Popfunding.com	Προσέγγιση της φιλοσοφίας λήψης απόφασης στον χώρο του Peer-to-Peer Lending και πως αυτή αντικατοπτρίζεται από τα στοιχεία των πιθανών δανειοληπτών	Logistic Regression	Στην περίπτωση που οι αιτούντες δεν έχουν άλλο υπάρχον ιστορικό στην πλατφόρμα οι πιστωτές στηρίζονται στην καταγραφή ψήφων εμπιστοσύνης. Τέλος, προκύπτουν ορισμένες υποθέσεις που λειτουργούν είτε θετικά ή αρνητικά στην επιτυχία ενός μελλοντικού δανείου
5	Polena (2017)	Lending Club	Σύγκριση και κατάταξη μεθόδων ως προς την προβλεπτική τους ακρίβεια με 6 δείκτες προβλεπτικής ικανότητας	8 αλγόριθμοι ταξινόμησης και 6 δείκτες προβλεπτικής ικανότητας: Percentage Correctly Classified, Kolmogorov-Smirnov Statistic, Brier Score, Area under Curve, Partial Gini Index και H-measure	Οι καλύτερες μέθοδοι ταξινόμησης είναι οι (με σειρά από την καλύτερη-στη λιγότερο καλύτερη) Logistic Regression, Artificial Neural Network και Linear Discriminant Analysis νάλυση ενώ οι λιγότερο καλές είναι οι C and Regression Trees και οι Random Forest
6	Lessmann et al. (2015)	Australian credit (AC) and German credit (GC) from the UCI Library (Lichman, 2013) and the Th02 data set	Σύγκριση και κατάταξη μιας μεγάλης γκάμας δημοφιλών μεθόδων κατάταξης	41 αλγόριθμοι ταξινόμησης και 6 δείκτες προβλεπτικής ικανότητας: Percentage Correctly Classified, Area Under Curve, Partial Gini Index, H-measure, Brief Score και Kolmogorov-Smirnov Statistic	Πλήρης κατάταξη των 41 αλγορίθμων η οποία δείχνει ότι η μέθοδος των Artificial Neural Network είναι καλύτερη από την Logistic Regression

		from Thomas, et al. (2002)			
7	Tsai et al. (2014)	Lending Club	Σύγκριση και ταξινόμηση 4 αλγορίθμων ταξινόμησης με γνώμονα τη μείωση του συντελεστή αθέτησης	Principal Component Analysis και 4 αλγόριθμοι ταξινόμησης Modified Logistic Regression Naïve Bayes, Random forest, Support Vector Machine,	Επίτευξη της βέλτιστης προβλεπτικής ακρίβειας των μεθόδων ταξινόμησης και κατάταξη της μεθόδου Logistic Regression
8	Puro et al. (2010)	Prosper.com	Ανάπτυξη ενός Συστήματος Υποστήριξης Αποφάσεων που θα ποσοτικοποιεί-αξιολογεί τις στρατηγικές των επενδυτών σε πλατφόρμες Peer-to-Peer Lending	Ανάλυση συσχετίσεων με το δείκτη Pearson και χρήση της μεθόδου ταξινόμησης Logistic Regression	Δημιουργία ενός ΣΥΑ όπου βρίσκει τη βέλτιστη στρατηγική του επενδυτή με γνώμονα το αρχικό επιτόκιο που αρχικοποιεί ο δανειολήπτης και το τελικό επιτόκιο
9	Cinca και Nieto (2016)	Lending Club	Ανάπτυξη ενός Συστήματος Υποστήριξης Αποφάσεων που θα επικεντρώνεται στην πρόβλεψη της αναμενόμενης κερδοφορίας από την πλευρά των επενδυτών	Μέθοδος CHAID και Τεχνική Δέντρων Αποφάσεων	Το επιτόκιο, ο σκοπός του δανείου και το χρέος του δανειολήπτη είναι οι μεταβλητές που επηρεάζουν περισσότερο τον εσωτερικό βαθμό απόδοσης του δανείου. Με το συγκερασμό αυτών των μεταβλητών δημιουργείται ένα ΣΥΑ το οποίο επικεντρώνεται στη βελτιστοποίηση της αναμενόμενης κερδοφορίας του δανείου
10	Feis et al. (2016)	Lending Club	Δημιουργία προσθήκης στην υφιστάμενη πλατφόρμα με σκοπό την κατάταξη των δανείων σε συνεχές φάσμα βαθμολόγησης της αξιοπιστίας	P-Value, Scikit Cross validation, Logistic Regression, Support Vector Machines, Linear Discriminant Analysis, Ensemble Methods, Discounted Cash Flow Model, Sharpe Ratio	Δημιουργία μοντέλου τιμολόγησης της αξίας των δανείων με τη βοήθεια του υπολογισμού της πιθανότητας αθέτησης των δανείων και υποστήριξη του μοντέλου σε κώδικα Python
11	Malekipirbazari και Aksakalli (2015)	Lending Club	Προσπάθεια αναγνώρισης των επιτυχημένων δανείων με δεδομένα που αυξάνουν την προβλεπτική ακρίβεια των μεθόδων και σύγκριση των 4 μεθόδων ταξινόμησης μεταξύ τους με τη βοήθεια ενός παραγόμενου μαθηματικού μοντέλου	Random Forest, k-Nearest Neighbors, Logistic Regression, Support Vector Machines	Υπεροχή της μεθόδου Random Forest έναντι των υπόλοιπων μεθόδων ταξινόμησης ως προς την προβλεπτική τους ακρίβεια. Τέλος αποδείχτηκε ότι τα αποτελέσματα που βασίζονται στη μέθοδο Random Forest είναι πιο αξιόπιστα από τις μεταβλητές που χρησιμοποιεί η ιστοσελίδα για λήψη αποφάσεων (FICO Score και Grade)

Πίνακας 2.1: Συγκεντρωτικός πίνακας βιβλιογραφικής ανασκόπησης.

# ΚΕΦΑΛΑΙΟ 3

## ΑΝΑΠΤΥΞΗ ΜΟΝΤΕΛΩΝ ΑΞΙΟΛΟΓΗΣΗΣ ΤΟΥ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

Στο παρόν κεφάλαιο θα πραγματοποιηθεί η μετάβαση από το θεωρητικό υπόβαθρο και την ανάλυση της διεθνούς βιβλιογραφίας, στο πειραματικό μέρος της εργασίας. Πιο αναλυτικά, η διάρθρωση του κεφαλαίου αυτού θα έχει ως εξής: στο πρώτο μέρος θα γίνει η παρουσίαση του δείγματος και η περιγραφή των μεταβλητών που παρέχει η πλατφόρμα για την καταγραφή του ιστορικού των δανειστών. Στο δεύτερο μέρος θα περιγραφεί η στατιστική ανάλυση και οι μεθοδολογίες που χρειάστηκαν ώστε να προκύψουν οι μεταβλητές που θα χρησιμοποιηθούν στο ερευνητικό κομμάτι της εργασίας. Στο τρίτο μέρος θα αναλυθούν οι μεθοδολογίες κατάταξης, οι οποίες χρησιμοποιήθηκαν ώστε να επιτευχθεί η πρόβλεψη της κατάστασης ενός δανείου. Στο τέταρτο μέρος θα παρουσιαστούν τα αποτελέσματα του πειράματος και στο τελευταίο μέρος θα γίνει η ανάλυση και ο σχολιασμός των αποτελεσμάτων της μεθοδολογικής προσέγγισης.

### 3.1 ΠΑΡΟΥΣΙΑΣΗ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα που χρησιμοποιήθηκαν για το πειραματικό μέρος της εργασίας, αποτελούν στοιχεία συνολικά 106160 καταχωρήσεων και πιο συγκεκριμένα 91926 επιτυχημένων και 14234 μη επιτυχημένων δανείων και 80 μεταβλητών, που επεξηγούν την κάθε περίπτωση δανείου, τα οποία παρέχει η πλατφόρμα prosper.com. Παρακάτω, στον πίνακα 3.1 παρουσιάζεται το όνομα της κάθε μεταβλητής και δίπλα μια συνοπτική εξήγηση καθώς και το εύρος τιμών της, σε όποια μεταβλητή ορίζεται αυτό. Για να γίνει πιο εύχρηστη η αξιοποίηση των δεδομένων, στην αρχή του πειραματικού μέρους έγινε μια αναδιάρθρωση ώστε να οριστικοποιηθεί ποια δεδομένα μπορούν να χρησιμοποιηθούν και πως ορισμένες μεταβλητές μπορούν να εκφραστούν με πιο κατανοητό τρόπο για το πρόβλημα.

A/A	<u>Ονομασία Μεταβλητής</u>	<u>Επεξήγηση-Εύρος Τιμών</u>
1	Listing Key	Ξεχωριστός κωδικός για κάθε καταχώρηση δανείου
2	Listing Number	Μοναδικός κωδικός αριθμός που προσδιορίζει την καταχώρηση του δανείου στο κοινό όπως εμφανίζεται στην ιστοσελίδα
3	Listing Creation Date	Ημερομηνία καταχώρησης του δανείου
4	Credit Grade	Βαθμός πιστοληπτικής ικανότητας της καταχώρησης που εκχωρήθηκε τη στιγμή που εισήχθη η καταχώρηση. Ισχύει μόνο για τις καταχωρήσεις που έγιναν μέχρι το 2009.-Παίρνει τιμές AA,AB,B, έως και HR, όπου AA η καταχώρηση με την καλύτερη πιστοληπτική ικανότητα
5	Term	Η διάρκεια του δανείου η οποία εκφράζεται σε μήνες
6	Loan Status	Η τρέχουσα κατάσταση δανείου. Παίρνει τιμές όπως Ακυρωμένο, χρεωμένο έως και ολοκληρωμένο
7	Closed Date	Ημερομηνία Κλεισίματος. Αποτελεί την ημερομηνία που το δάνειο λαμβάνει τελική κατάσταση. Παίρνει τιμές μόνο για τα δάνεια με κατάσταση ολοκληρωμένο, χρεωμένο, αθετημένο και ακυρωμένο
8	Borrower APR	Ετήσιο πραγματικό επιτόκιο. Είναι το ετήσιο επιτόκιο που χρεώνεται ο δανειολήπτης για ένα δάνειο και εκφράζεται ως ποσοστό που αντιπροσωπεύει το πραγματικό ετήσιο κόστος κεφαλαίων κατά τη διάρκεια ενός δανείου.
9	Borrower Rate	Το επιτόκιο του δανειολήπτη για το συγκεκριμένο δάνειο
10	Lender Yield	Η απόδοση του δανειστή για το συγκεκριμένο δάνειο το οποίο προκύπτει από το επιτόκιο του δανείου μειωμένο κατά το τέλος εξυπηρέτησης
11	Estimated Effective Yield	Η εκτιμώμενη αποτελεσματική απόδοση του δανείου η οποία προκύπτει από το επιτόκιο του δανειολήπτη αφαιρώντας το επιτόκιο των τελών εξυπηρέτησης και τους εκτιμώμενους μη εισπραχθέντες τόκους σε περίπτωση αθέτησης και προσθέτοντας τα εκτιμώμενα εισπραχθέντα τέλη καθυστερήσεων. Τιμές της αποτελεσματικής απόδοσης λαμβάνουν οι καταχωρήσεις μετά το 2009
12	Estimated Loss	Η εκτιμώμενη ζημία η οποία εκφράζει την ζημία που προκύπτει σε περίπτωση αθέτησης του δανείου. Τιμές της εκτιμώμενης ζημίας λαμβάνουν οι καταχωρήσεις μετά το 2009
13	Estimated Return	Η εκτιμώμενη απόδοση είναι μεταβλητή που καταχωρείται της στιγμή της καταχώρησης του δανείου και προκύπτει από τη διαφορά της αποτελεσματικής απόδοσης και του εκτιμώμενου ποσοστού ζημίας. Τιμές αυτής της μεταβλητής λαμβάνουν καταχωρήσεις που έγιναν μετά το 2009
14	Prosper Rating (Numeric)	Δείκτης αξιολόγησης πιστοληπτικής ικανότητας (αριθμητικός) ο οποίος εκφράζει με νούμερα τη μεταβλητή Credit Grade αλλά για τα δάνεια που καταχωρήθηκαν μετά το 2009. Λαμβάνει τιμές 1 για το βαθμό HR έως και 7 για το AA
15	Prosper Rating (Alpha)	Δείκτης αξιολόγησης πιστοληπτικής ικανότητας (Alpha). Η ίδια μεταβλητή με τη μεταβλητή 14 η οποία εκφράζεται με αλφαριθμητικούς χαρακτήρες αντί για αριθμούς. Λαμβάνει δηλαδή τις τιμές AA έως HR και είναι για τα δάνεια που καταχωρήθηκαν μετά το 2009
16	Prosper Score	Ένα προσαρμοσμένο σκορ κινδύνου το οποίο λαμβάνει τιμές από το 1 έως και το 10 με το 10 να είναι το λιγότερο επικίνδυνο. Ισχύει για δάνεια που καταχωρήθηκαν μετά το 2009
17	Listing Category (Numeric)	Η μεταβλητή που αρχικοποιεί ο δανειολήπτης για το σκοπό του δανείου. Παίρνει αριθμητικές τιμές ανάλογα με το σκοπό που περιλαμβάνεται στη λίστα. Για παράδειγμα το 1 αντιστοιχεί στο σκοπό σταθεροποίηση χρέους και το 20 για δάνεια γάμου
18	Borrower State	Η μεταβλητή που δηλώνει την πολιτεία του δανειολήπτη

19	Occupation	Το επάγγελμα που αρχειοποίησε ο δανειολήπτης τη στιγμή της καταχώρησης του δανείου
20	Employment Status	Η επαγγελματική κατάσταση του δανειολήπτη τη στιγμή που καταχωρεί το δάνειο. Διακρίνεται σε μερική απασχόληση, πλήρη απασχόληση και άλλα
21	Employment Status Duration	Η διάρκεια της επαγγελματικής κατάστασης του δανειολήπτη τη στιγμή της καταχώρησης. Εκφράζεται σε μήνες
22	Is Borrower Homeowner	Λογική μεταβλητή που δείχνει αν ο δανειολήπτης είναι ιδιοκτήτης σπιτιού ή όχι
23	Currently In Group	Λογική μεταβλητή που δείχνει αν ο δανειολήπτης την ώρα που καταχωρεί το δάνειο ανήκει σε κάποια άλλη ομάδα καταχώρησης
24	Group Key	Κωδικός αριθμός της ομάδας στην οποία ανήκει ο δανειολήπτης. Λαμβάνει τιμές μόνο όταν ο δανειολήπτης ανήκει σε κάποια ομάδα καταχώρησης και αυτή η ομάδα έχει 2 άτομα και πάνω ως μέλη
25	Date Credit Pulled	Ημερομηνία που ελήφθησαν δεδομένα για το προφίλ της συγκεκριμένης καταχώρησης πίστωσης
26	Credit Score Range Low	Η χαμηλότερη αξία που αντιπροσωπεύει το εύρος του πιστωτικού αποτελέσματος του δανειολήπτη όπως αυτή παρέχεται από έναν οργανισμό αξιολόγησης της πιστοληπτικής ικανότητας του καταναλωτή
27	Credit Score Range Upper	Η ανώτερη αξία που αντιπροσωπεύει το εύρος του πιστωτικού αποτελέσματος του δανειολήπτη όπως αυτή παρέχεται από έναν οργανισμό αξιολόγησης της πιστοληπτικής ικανότητας του καταναλωτή
28	First Recording Credit Line	Η ημερομηνία που άνοιξε το πρώτο πιστωτικό όριο για τον δανειολήπτη
29	Current Credit Lines	Αριθμός των τρεχουσών πιστωτικών ορίων του δανειολήπτη έως και τη στιγμή όπου έγινε η καταγραφή των δεδομένων για το συγκεκριμένο δάνειο
30	Open Credit Lines	Αριθμός των ανοιχτών διαθέσιμων πιστωτικών ορίων του δανειολήπτη έως και τη στιγμή όπου έγινε η καταγραφή για το συγκεκριμένο δάνειο
31	Total Credit Lines Past 7 Years	Συνολικός αριθμός πιστωτικών ορίων τα τελευταία 7 χρόνια μέχρι και τη στιγμή της καταγραφής των δεδομένων για το συγκεκριμένο δάνειο
32	Open Revolving Accounts	Αριθμός ανοικτών ανακυκλούμενων λογαριασμών τη στιγμή που ελήφθησαν τα δεδομένα για το συγκεκριμένο προφίλ πίστωσης
33	Open Revolving Monthly Payment	Μηνιαία Πληρωμή σε ανακυκλούμενους λογαριασμούς τη στιγμή που ελήφθησαν τα δεδομένα για το συγκεκριμένο προφίλ πίστωσης
34	Inquires Last 6 Months	Αριθμός ερωτήσεων τους τελευταίους 6 μήνες κατά τη στιγμή όπου ελήφθησαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
35	Total Inquires	Συνολικός αριθμός ερωτήσεων κατά τη στιγμή που ελήφθησαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
36	Current Delinquencies	Αριθμός καταχωρήσεων όπου ο συγκεκριμένος δανειολήπτης ήταν παραβατικός ως προς τις υποχρεώσεις του μέχρι και τη στιγμή όπου ελήφθησαν τα δεδομένα για τον συγκεκριμένο δανειολήπτη
37	Amount Delinquent	Ποσό σε δολάρια τα οποία αντιστοιχούν στις παραβάσεις του δανειολήπτη τη στιγμή όπου εξήχθησαν τα δεδομένα για το συγκεκριμένο δανειολήπτη
38	Delinquencies past 7 Years	Αριθμός παραβάσεων του συγκεκριμένου δανειολήπτη έως τη στιγμή που εξήχθησαν τα δεδομένα για το συγκεκριμένο προφίλ
39	Public Records Past 10 Years	Αριθμός δημοσίων εγγράφων τα τελευταία 10 χρόνια έως τη στιγμή όπου ελήφθησαν τα δεδομένα για τον συγκεκριμένο δανειολήπτη
40	Public Records Last 12 months	Αριθμός δημοσίων εγγράφων τους τελευταίους 12 μήνες έως τη στιγμή όπου ελήφθησαν τα δεδομένα για τον συγκεκριμένο δανειολήπτη

41	Revolving Credit Balance	Δολάρια ανακυκλούμενης πίστωσης τη στιγμή που καταγράφηκαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
42	Bankcard Utilization	Το ποσοστό της διαθέσιμης ανακυκλούμενης πίστωσης που χρησιμοποιείται το χρόνο κατά τον οποίο καταγράφηκαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
43	Available Bankcard Credit	Η συνολική διαθέσιμη πίστωση μέσω τραπεζικής κάρτας κατά τη στιγμή που καταγράφηκε το συγκεκριμένο προφίλ πίστωσης
44	Total Trades	Αριθμός των συνολικών συναλλαγών που πραγματοποιήθηκαν κατά τη στιγμή που ελήφθησαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
45	Trades Never Delinquent (Percentage)	Αριθμός συναλλαγών που δεν έχουν παραβιαστεί ποτέ κατά τη στιγμή που καταγράφηκε το συγκεκριμένο προφίλ πίστωσης
46	Trades Opened Last 6 Months	Αριθμός των συναλλαγών που πραγματοποιήθηκαν τους τελευταίους 6 μήνες κατά τη στιγμή που ελήφθησαν τα δεδομένα για τη συγκεκριμένη καταχώρηση
47	Debt To Income Ratio	Ο δείκτης χρέους προς έσοδα του δανειολήπτη κατά το χρόνο που καταγράφηκαν τα δεδομένα για το συγκεκριμένο δάνειο. Σε περίπτωση που αυτή η σχέση δεν είναι διαθέσιμη η μεταβλητή λαμβάνει την τιμή 0. Στην περίπτωση που υπερβαίνει την τιμή 1000% επιστρέφεται η τιμή 1001%. Η ανώτατη τιμή της είναι το 10,01%
48	Income Range	Το εύρος εισοδήματος στο οποίο ανήκει ο δανειολήπτης το χρόνο της δημιουργίας της καταχώρησής του. Λαμβάνει τιμές από 0 έως 100000+ δολάρια σε διαστήματα με εύρος 25000 δολαρίων
49	Income Verifiable	Η λογική μεταβλητή όπου δανειολήπτης δηλώνει την απαιτούμενη τεκμηρίωση για το εισόδημά του
50	Stated Monthly Income	Το σταθερό μηνιαίο εισόδημα που δηλώνει ο δανειολήπτης κατά τη χρονιά όπου καταχώρησε την αίτησή του
51	Loan Key	Μοναδικός κωδικός για την καταχώρηση κάθε δανείου
52	Total Prosper Loans	Συνολικός αριθμός δανείων που έχει ο δανειολήπτης ήδη στην πλατφόρμα Prosper μέχρι τη στιγμή της κατοχύρωσης της αίτησής του
53	On Time Prosper Payments	Αριθμός πληρωμών που πραγματοποιήθηκαν στην ώρα τους σε δάνεια της πλατφόρμας Prosper που είχε ο δανειολήπτης μέχρι τη στιγμή της συγκεκριμένης κατοχύρωσης του. Ο αριθμός αυτός θα είναι 0 στην περίπτωση που ο δανειολήπτης δεν είχε προηγούμενα δάνεια
54	Prosper Payments Less 1 Month Late	Αριθμός πληρωμών που πραγματοποίησε ο δανειολήπτης σε δάνεια που ήταν λιγότερο από ένα μήνα καθυστερημένα κατά τη στιγμή που κατοχύρωσε την αίτησή του. Η τιμή αυτή θα είναι 0 αν ο δανειολήπτης δεν είχε προηγούμενα δάνεια
55	Prosper Payments 1 Month Plus Late	Αριθμός πληρωμών που πραγματοποίησε ο δανειολήπτης σε δάνεια που ήταν περισσότερο από ένα μήνα καθυστερημένα κατά τη στιγμή που κατοχύρωσε την αίτησή του. Η τιμή αυτή θα είναι 0 αν ο δανειολήπτης δεν είχε προηγούμενα δάνεια
56	Prosper Principal Borrowed	Το σύνολο του κεφαλαίου που δανείστηκε ο δανειολήπτης από δάνεια της πλατφόρμας Prosper το χρόνο που πραγματοποίησε την κατοχύρωση της αίτησής του
57	Prosper Principal Outstanding	Ποσό που εκκρεμεί στα δάνεια της πλατφόρμας Prosper κατά το χρόνο της επικύρωσης της αίτησης του δανειολήπτη
58	Scorex Change At Time Of Listing	Η μεταβολή του πιστωτικού αποτελέσματος του δανειολήπτη κατά τη στιγμή που κατοχύρωσε την αίτησή του. Αυτή θα είναι σε σχέση με το τελευταίο δάνειο από την πλατφόρμα Prosper που είχε ο δανειολήπτης. Άρα η μεταβολή δε θα υπάρχει σε περίπτωση που ο δανειολήπτης δεν είχε προηγούμενα δάνεια
59	Loan Current Days Delinquent	Ο αριθμός των ημερών παραβατικότητας του δανειολήπτη

60	Loan First Defaulted Cycle Number	Η περίοδος που το δάνειο θεωρείται χρεωμένο. Αν το δάνειο δε χρεωθεί τότε η τιμή θα έχει μηδενική τιμή
61	Loan Months Since Origination	Αριθμός μηνών από την έναρξη του δανείου
62	Loan Number	Μοναδική αριθμητική τιμή που σχετίζεται με το δάνειο
63	Loan Original Amount	Το αρχικό ποσό του δανείου
64	Loan Origination Date	Ημερομηνία από την οποία ξεκινά το δάνειο
65	Loan Origination Quarter	Το τρίμηνο κατά το οποίο ξεκινά το δάνειο
66	Member Key	Ο μοναδικός κωδικός που συνδέεται με τον δανειολήπτη. Είναι ουσιαστικά το αναγνωριστικό που χρησιμοποιεί ο ίδιος στη διεπαφή του με την πλατφόρμα
67	Monthly Loan Payment	Η προγραμματισμένη μηνιαία πληρωμή του δανείου
68	LP Customer Payments	Προπληρωμένες σωρευτικές ακαθάριστες πληρωμές που πραγματοποίησε ο δανειολήπτης για το δάνειο. Αν το δάνειο είχε χρεωθεί θα αποκλείσει τυχόν ανακτήσεις
69	LP Customer Principal Payment	Προπληρωμένες σωρευτικές ακαθάριστες πληρωμές κεφαλαίου που πραγματοποίησε ο δανειολήπτης για το δάνειο. Αν το δάνειο είχε χρεωθεί θα αποκλείσει τυχόν ανακτήσεις
70	LP Interest and Fees	Προπληρωμένοι σωρευτικοί τόκοι και τέλη που καταβάλλονται από τον δανειολήπτη. Εάν το δάνειο έχει χρεωθεί αυτή η τιμή θα αποκλείσει τυχόν ανακτήσεις
71	LP Service Fees	Αθροιστικές αμοιβές υπηρεσιών που καταβάλλονται από τους επενδυτές που έχουν επενδύσει στο συγκεκριμένο δάνειο
72	LP Collection Fees	Αθροιστικές αμοιβές είσπραξης που καταβάλλονται από τους επενδυτές που έχουν επενδύσει στο συγκεκριμένο δάνειο
73	LP Gross Principal Loss	Το ακαθάριστο χρεωμένο ποσό του δανείου
74	LP Net Principal Loss	Το κεφάλαιο που παραμένει ανεκμετάλλευτο μετά από τυχόν ανακτήσεις
75	LP Non Principal Recovery Payments	Η κύρια συνιστώσα σύμφωνα με την τρέχουσα πολιτική για πληρωμές ανακτήσεις ,ε ακόλουθη σειρά: Τέλη, τόκοι και κεφάλαια
76	Percent Funded	Το ποσοστό χρηματοδότησης της συγκεκριμένης καταγραφής
77	Recommendations	Αριθμός συστάσεων που είχε ο δανειολήπτης τον χρόνο κατά τον οποίο πραγματοποίησε την καταγραφή της αίτησής του
78	Investment From Friends Count	Αριθμός φιλικών προσώπων του δανειολήπτη που πραγματοποίησαν χρηματοδότηση στο συγκεκριμένο δάνειο
79	Investment From Friends Amount	Ποσό σε δολάρια από επενδύσεις που προήλθαν από φιλικά πρόσωπα του δανειολήπτη
80	Investors	Ο αριθμός των επενδυτών που χρηματοδότησε το συγκεκριμένο δάνειο

Πίνακας 3.1: Πίνακας μεταβλητών που χρησιμοποιεί η πλατφόρμα prosper.com.



Πιο συγκεκριμένα, όπως θα αναφερθεί και στη συνέχεια την εξαρτημένη μεταβλητή του προβλήματος θα αποτελεί η μεταβλητή κατάστασης του δανείου, η οποία πλέον θα περιγράφεται ως λογική μεταβλητή και θα περιέχει τιμές 0 για τα επιτυχημένα δάνεια και 1 για τα μη επιτυχημένα δάνεια. Αυτό όμως προϋποθέτει να γίνει διαχωρισμός ως προς ποια δάνεια θα είναι επιτυχημένα και ποια όχι. Έτσι, τα δάνεια τα οποία στα αρχικά δεδομένα χαρακτηρίζονταν ως ολοκληρωμένα, τρέχοντα, σε διαδικασία τελευταίας πληρωμής και σε κατάσταση καθυστέρησης 1-90 ημερών (κόκκινα δάνεια) θα ενταχθούν στην κατηγορία των πετυχημένων δανείων και θα περιγράφονται με τον αριθμό 0. Οι καταστάσεις που περιγράφουν τις υπόλοιπες περιπτώσεις δανείων δηλαδή χρεωμένο, ακυρωμένο, σε κατάσταση αθέτησης και κατάσταση καθυστέρησης 91- 120+ ημέρες θα κατατάσσονται στην κατηγορία των μη επιτυχημένων δανείων και θα περιγράφονται από τον αριθμό 1.

Στη συνέχεια θα γίνει μια εξίσου σημαντική ενέργεια εκκαθάρισης στα δεδομένα ως προς τις μεταβλητές που περιγράφουν την πιστοληπτική ικανότητα μια καταχώρησης ενός δανείου. Πιο συγκεκριμένα, η μεταβλητή credit grade, η οποία περιγράφει την πιστοληπτική ικανότητα των δανείων πριν το 2009 μετατράπηκε σε μια αριθμητική κλίμακα από το 1 έως το 7, όπου το 1 αντιστοιχεί στα δάνεια με credit grade HR και το 7 στα δάνεια με credit grade AA. Στη συνέχεια, η μεταβλητή αυτή συγχωνεύτηκε με τη μεταβλητή Prosper Rating (numeric), η οποία περιγράφει την πιστοληπτική ικανότητα των δανείων μετά το 2009, και τέλος η μεταβλητή Prosper Rating (Alpha) διαγράφεται από τη λίστα, αφού περιγράφεται ήδη αλλά με διαφορετικό τρόπο από τη μεταβλητή Prosper Rating (numeric). Συγχωνεύοντας αυτές τις 2 μεταβλητές δημιουργείται πλέον η μεταβλητή Prosper Rating, η οποία περιγράφει πλέον την πιστοληπτική ικανότητα των δανείων σε όλο το φάσμα χρόνου καταγραφής των δεδομένων.

Επιπρόσθετα, η μεταβλητή Income Range, που περιγράφει το εύρος του εισοδήματος του δανειολήπτη μετατράπηκε σε βαθμονομημένη κλίμακα από το 1 έως το 6, όπου το 1 αναφέρεται σε εισόδημα 0 δολαρίων, το 2 σε εύρος εισοδήματος 1 έως 25000 δολάρια, το 3 σε εύρος εισοδήματος 25001 έως 50000 δολάρια έως και το 7 που αντιστοιχεί σε εύρος εισοδήματος 100000 και άνω δολάρια.

Επιπλέον, για τη στατιστική ανάλυση χρειάζεται μια επιπλέον εκκαθάριση στα δεδομένα των καταχωρήσεων και ιδιαίτερα στα κενά στοιχεία του πίνακα. Με άλλα λόγια, θα γίνει διαγραφή ορισμένων γραμμών του πίνακα. Για παράδειγμα, οι καταγραφές που δεν εμπεριείχαν κάποια τιμή στη μεταβλητή Prosper Rating δεν είχε νόημα να συμπεριληφθούν στην έρευνα μιας και οι προαναφερθείσες έρευνες στηρίζουν ότι μεταβλητές τέτοιου είδους έχουν την υψηλότερη προβλεπτική ικανότητα στην επίλυση προβλημάτων τέτοιου είδους. Έτσι στις μεταβλητές Prosper Score, Listing Category και Occupation μόνο για τη στατιστική ανάλυση, όπου υπήρχαν κενά στοιχεία οι καταχωρήσεις διαγράφονται. Αυτό είχε ως αποτέλεσμα, τα στοιχεία να μειωθούν στις 83486 καταχωρήσεις για την ανάλυση αυτή.

Στο τέλος αυτής της ενότητας παρουσιάζεται ο πίνακας 3.2, στον οποίο αναφέρονται ονομαστικά οι μεταβλητές, οι οποίες σύμφωνα με την ανάλυση της βιβλιογραφίας έχουν υψηλή σημαντικότητα, αφού μπορούν να χρησιμοποιηθούν ως προβλεπτικοί παράγοντες ως προς τη φερεγγυότητα ενός δανείου, ή αλλιώς στην ποσοτικοποίηση της πιθανότητας επιτυχίας του.

A/A	<u><b>Ονομασία Μεταβλητής</b></u>
1	Prosper Rating
2	Term
3	Borrower APR
4	Borrower Rate
5	Prosper Score
6	Listing Category
7	Borrower State
8	Occupation
9	Employment Status
10	Is Borrower Homeowner
11	Credit Score Range Lower
12	Credit Score Range Upper
13	Current Credit Lines
14	Open Credit Lines
15	Total Credit Lines past 7 Years
16	Open Revolving Account
17	Open Revolving Monthly Payment
18	Inquires Last 6 Months
19	Total Inquires
20	Current Delinquencies

21	Amount Delinquent
22	Delinquencies last 7 Years
23	Public Records last 10 Years
24	Public Records last 12 Months
25	Revolving Credit Balance
26	Bankcard Utilization
27	Available Bankcard Credit
28	Total Trades
29	Trades Never Delinquent (Percentage)
30	Trades Opened Lasts 6 Months
31	Debt to Income Ratio
32	Income Range
33	Involve Verifiable
34	Stated Monthly Income
35	Loan original Amount
36	Monthly Loan Payment
37	Recommendations
38	Investors

*Πίνακας 3.2: Παρουσίαση Μεταβλητών όπου θα χρησιμοποιηθούν στην ανάλυση του προβλήματος.*

## 3.2 ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ

Εφόσον ορίστηκαν οι σημαντικές μεταβλητές του προβλήματος στο συγκεκριμένο κομμάτι του κεφαλαίου, θα γίνει η προσέγγιση της στατιστικής μεθοδολογίας με βάση τους μη παραμετρικούς ελέγχους Kruskal-Wallis και τους ελέγχους  $X^2$ , που θα χρησιμοποιηθούν, καθώς και τα αποτελέσματα που προέκυψαν συνοδευόμενα από συγκεντρωτικούς πίνακες που θα βοηθήσουν στην κατανόηση της ανάλυσης.

Η στατιστική ανάλυση και το πειραματικό μέρος της εργασίας πραγματοποιήθηκαν με τη βοήθεια του προγραμματιστικού περιβάλλοντος της Matlab και τα αποτελέσματα διακρίνονται στους πίνακες 3.3 και 3.4.

Πιο αναλυτικά, για τις ποσοτικές μεταβλητές χρησιμοποιείται ο μη παραμετρικός έλεγχος Kruskal-Wallis τα αποτελέσματα του οποίου φαίνονται στον πίνακα 3.3. Σε αυτή την κατηγορία μεταβλητών υπολογίζεται η στατιστική διαφορά που υπάρχει μεταξύ των δύο κατηγοριών δανείου. Επίσης, στις στήλες 3 έως 8 παρουσιάζονται η μέση τιμή η διάμεσος και η τυπική απόκλιση για κάθε κατηγορία δανείου (0 για τα εξυπηρετούμενα και 1 για τα μη εξυπηρετούμενα δάνεια). Για τις ποιοτικές μεταβλητές χρησιμοποιείται ο έλεγχος  $X^2$ , ο οποίος ελέγχει αν υπάρχει σημαντική στατιστική διαφορά μεταξύ των δυο κατηγοριών δανείου, τα αποτελέσματα του οποίου παρουσιάζονται στον πίνακα 3.4.

Όπως γίνεται αντιληπτό από τους πίνακες 3.3 και 3.4 όλες οι μεταβλητές ποιοτικές και ποσοτικές είναι στατιστικά σημαντικές σε ένα επίπεδο εμπιστοσύνης 99%. Μόνο η μεταβλητή Debt to Income Ratio είναι σημαντική σε επίπεδο εμπιστοσύνης 95% ( $p$ -value=0,02). Με άλλα λόγια, μηδενική υπόθεση των δύο μη παραμετρικών ελέγχων, σύμφωνα με την οποία δεν υπάρχει διαφορά μεταξύ των μεταβλητών και της κατάστασης των δανείων μπορεί να απορριφθεί.

Μεταβλητή	$X^2$	p-Value	Mean (0)	Median (0)	Std (0)	Mean (1)	Median (1)	Std (1)
BorrowerAPR	3088,9	0	0,21	0,20	0,08	0,26	0,26	0,08
BorrowerRate	4504,5	0	0,19	0,18	0,07	0,23	0,24	0,07
CreditScoreRangeLower	3794,3	0	696,08	700,00	53,19	659,83	660,00	65,76
CreditScoreRangeUpper	3794,3	0	715,08	719,00	53,19	678,83	679,00	65,76
CurrentCreditLines	791,6	0	10,48	10,00	5,38	9,28	8,00	5,85
OpenCreditLines	1159,3	0	9,44	9,00	4,97	8,08	7,00	5,21
TotalCreditLinespast7years	272,7	0	27,27	26,00	13,39	25,63	23,00	14,68
OpenRevolvingAccounts	1292,0	0	7,34	6,00	4,54	6,11	5,00	4,63
OpenRevolvingMonthlyPayment	1538,7	0	422,26	299,00	436,79	353,23	190,00	504,13
InquiriesLast6Months	3865,8	0	1,09	1,00	1,71	2,61	1,00	3,65
TotalInquiries	2978,4	0	4,74	4,00	4,69	8,62	6,00	9,32
CurrentDelinquencies	1956,1	0	0,36	0,00	1,26	1,06	0,00	2,78

AmountDelinquent	1855,0	0	919,34	0,00	7106,25	1411,61	0,00	7502,92
DelinquenciesLast7Years	549,6	0	3,73	0,00	9,51	5,18	0,00	11,40
PublicRecordsLast10Years	204,4	0	0,29	0,00	0,66	0,41	0,00	0,93
PublicRecordsLast12Months	351,2	0	0,01	0,00	0,12	0,04	0,00	0,23
RevolvingCreditBalance	1353,5	0	17925,96	9020,00	32983,69	15508,35	5196,00	32519,93
BankcardUtilization	17,3	0	0,56	0,60	0,31	0,57	0,63	0,37
AvailableBankcardCredit	2306,0	0	11834,23	4544,50	20364,07	7175,70	1801,50	15288,84
TotalTrades	618,4	0	23,55	22,00	11,73	21,22	19,00	12,58
TradesNeverDelinquent	1971,5	0	0,90	0,95	0,14	0,82	0,88	0,20
TradesOpenedLast6Months	751,7	0	0,75	0,00	1,03	1,11	1,00	1,44
DebtToIncomeRatio	5,3	0,02081641	0,24	0,21	0,42	0,32	0,21	0,90
StatedMonthlyIncome	2223,6	0	5849,58	5000,00	8044,82	4524,56	3813,33	3554,47
LoanOriginalAmount	1786,7	0	8856,92	7400,00	6329,05	6737,79	5000,00	5512,59
MonthlyLoanPayment	870,9	0	284,56	240,12	191,77	244,68	174,55	191,44
Recommendations	659,6	0	0,04	0,00	0,33	0,10	0,00	0,40
Investors	2554,1	0	77,62	39,00	103,41	105,21	68,00	110,69

Πίνακας 3.3: Πίνακας στατιστικής ανάλυσης ποσοτικών μεταβλητών.

Μεταβλητές	X <sup>2</sup>	p-Value
Prosper Rating	3875,699787	0
Term	1809,694931	0
ProsperScore	12000,85816	0
ListingCategory	7805,81942	0
BorrowerState	1551,792256	0
Occupation	1364,165513	0
EmploymentStatus	8708,455901	0
IsBorrowerHomeowner	101,8782116	0
IncomeRange	2355,642638	0
IncomeVerifiable	150,422632	0

Πίνακας 3.4: Πίνακας στατιστικής ανάλυσης ποιοτικών μεταβλητών

### 3.3 ΜΕΘΟΔΟΛΟΓΙΑ

Το πειραματικό κομμάτι βασίζεται στη χρήση μεθόδων και εργαλείων μηχανικής μάθησης για την πρόβλεψη της κατάστασης των δανείων (εξυπηρετούμενα ή μη εξυπηρετούμενα), για τον υπολογισμό της πιθανότητας ένας πελάτης να ανήκει σε μία από τις δύο ομάδες δανείων και στον υπολογισμό 5 δεικτών ακρίβειας με τις οποίες θα αξιολογούνται οι 7 μέθοδοι ως προς την προβλεπτική τους ικανότητα.

Σε πρώτη φάση, μελέτες που χρησιμοποιούν μεθοδολογίες και εργαλεία μηχανικής μάθησης, χρησιμοποιούν ταυτόχρονα εργαλεία εξόρυξης δεδομένων που βοηθούν στην αξιολόγηση και την εγκυρότητα των μοντέλων machine learning. Μια στατιστική διαδικασία που βοηθά στην επίλυση του παραπάνω προβλήματος, είναι το Cross-Validation. Συνοπτικά, στη διαδικασία του cross-validation τα δεδομένα του προβλήματος (στη συγκεκριμένη περίπτωση τις καταχωρήσεις των 106161 δανείων) χωρίζονται σε ομάδες δεδομένων εκπαίδευσης και ελέγχου. Στη συγκεκριμένη περίπτωση εφαρμόστηκε η διαδικασία του 2- fold cross-validation στην οποία το δείγμα χωρίζεται σε δύο ίσα μέρη. Το ένα μέρος θα χρησιμοποιηθεί ως δείγμα εκπαίδευσης και το άλλο ως δείγμα επαλήθευσης. Η διαδικασία αυτή επαναλαμβάνεται 5 φορές (5×2-fold cross-validation) για την εξαγωγή πιο ασφαλών αποτελεσμάτων.

Σε δεύτερη φάση, θα γίνει η παρουσίαση των μεθόδων μηχανικής μάθησης που θα χρησιμοποιηθούν στο πειραματικό μέρος και θα περιλαμβάνουν τη γραμμική διακριτική ανάλυση, την τετραγωνική διακριτική ανάλυση, τη λογιστική παλινδρόμηση, τις μηχανές διανυσμάτων υποστήριξης, τα δέντρα ταξινόμησης και συνδυαστικά μοντέλα δέντρων ταξινόμησης.

**Γραμμική Διακριτική Ανάλυση (Linear Discriminant Analysis-LDA):** Η μέθοδος ταξινόμησης LDA δημιουργήθηκε από τον Fisher (1936) και είναι πιο διαδομένη μέθοδος ταξινόμησης. Σκοπός της είναι, η υλοποίηση μιας συνάρτησης η οποία θα περιγράφεται ως ο γραμμικός συνδυασμός των χαρακτηριστικών που οδηγούν στο συμπέρασμα ως προς την ταξινόμηση κάθε καταχώρησης σε μια ομάδα. Χρησιμοποιεί, συνεχείς ανεξάρτητες μεταβλητές και μια κατηγορική εξαρτημένη μεταβλητή.

**Τετραγωνική Διακριτική Ανάλυση (Quadratic Discriminant Analysis-QDA):** Η μέθοδος ταξινόμησης QDA δημιουργήθηκε από τον Smith (1947) και είναι μια εναλλακτική προσέγγιση της LDA. Προϋποθέτει κάθε κλάση ταξινόμησης να περιγράφεται από την κατανομή Gauss και η κύρια διαφορά της από την LDA είναι η υπόθεση ότι κάθε κατηγορία διαθέτει διαφορετικό πίνακα συνδιακύμανσης. Στο μαθηματικό μοντέλο διαθέτει μια παραπάνω τάξη από αυτή της LDA και είναι σημαντική γιατί είναι ικανή παρέχει πιο ακριβή μη γραμμικά όρια ταξινόμησης.

**Λογιστική Παλινδρόμηση (Logistic Regression-LR):** Πρόκειται για μια μέθοδο που προβλέπει το αποτέλεσμα μιας κατηγορικής μεταβλητής δυαδικού συνήθως χαρακτήρα με τη βοήθεια ενός λογιστικού μοντέλου. Το μοντέλο αυτό αποτελεί γραμμικό συνδυασμό των προβλεπτικών μεταβλητών του προβλήματος και επιστρέφει το λογάριθμο της πιθανότητας μια περίπτωση να ανήκει σε μια ομάδα έκβασης. Αξίζει να σημειωθεί ότι οι παράμετροι του λογιστικού μοντέλου, υπολογίζονται με τη βοήθεια της μεθόδου μέγιστης πιθανοφάνειας.

**Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines-SVM):** Η μέθοδος SVM θεωρείται ως μιας από της πιο πετυχημένες μεθόδους ταξινόμησης η οποία βασίζεται στην ύπαρξη μιας γραμμικής διαχωριστικότητας που υπάρχει ανάμεσα στις δύο ομάδες κατηγοριοποιήσεων που μελετώνται. Ουσιαστικά, προσπαθεί να μεγιστοποιήσει την απόσταση μεταξύ των δυο κλάσεων ορίζοντας έτσι ένα βέλτιστο υπερεπίπεδο που διαχωρίζει τις δύο κλάσεις.

**Δέντρα Ταξινόμησης (Classification Tree):** Η μέθοδος αυτή βασίζει την ταξινόμηση ενός δυαδικού προβλήματος στη δημιουργία ενός δέντρου απόφασης όπου οι καταστάσεις του μεταφράζονται ως εσωτερικοί κόμβοι και η τελική απόφαση ως φύλλα. Χρησιμοποιώντας τα χαρακτηριστικά του προβλήματος και τις προϋποθέσεις που τα διέπουν, διασπώντας το δέντρο σε κλαδιά, επιλέγεται το μικρότερο κόστος ακρίβειας στην κάθε περίπτωση και έτσι λαμβάνεται η τελική απόφαση.

**Ενισχυμένα Δέντρα (Boosted Trees):** Η μέθοδος αυτή λειτουργεί ως προσθετική στη μέθοδο ταξινόμησης με δέντρα. Λαμβάνει αποφάσεις από μια ακολουθία βασικών δέντρων και συνδυάζοντας την, καταλήγει στην τελική πρόβλεψη. Είναι ιδιαίτερα σημαντική αφού σε σύγκριση με τις κλασσικές μεθόδους λήψης αποφάσεων με χρήση δέντρων διαχειρίζεται καλύτερα τα δεδομένα και σε αντίθεση με κλασσικά γραμμικά μοντέλα λαμβάνει υπόψιν του και τη μη γραμμική αλληλεπίδραση μεταξύ των μεταβλητών του προβλήματος.

**«Σάκου» Δέντρων (Bagged Trees):** Το bagging όπως και το boosting είναι μέθοδος που βασίζονται σε ήδη υπάρχοντες ταξινομητές με σκοπό να δώσουν μια καλύτερη πρόβλεψη. Έτσι στη μεθοδολογία των bagged trees για να βελτιωθεί η

ταξινόμηση συνδυάζονται ταξινομήσεις τυχαίων παραγόμενων εκπαιδευτικών σετ δεδομένων. Η μεθοδολογία αυτή φαίνεται να μπορεί να εφαρμοστεί βοηθητικά αυξάνοντας την απόδοση ορισμένων μεθόδων απόδοσης αλλά και σε ορισμένες περιπτώσεις να την υποβαθμίσει.

### 3.4 ΑΠΟΤΕΛΕΣΜΑΤΑ

Στο σημείο αυτό θα αναλυθούν τα αποτελέσματα που προέκυψαν από την χρήση των μοντέλων αξιολόγησης ως προς τις εκτιμήσεις που δίνουν και θα παρουσιαστούν συγκριτικοί πίνακες που θα βοηθήσουν στην συγκριτική αξιολόγηση των μεθόδων ως προς την ακρίβειά τους.

Τα πρώτα αποτελέσματα έχουν να κάνουν με την επικύρωση των αποτελεσμάτων της ταξινόμησης. Από κάθε υποσύνολο δεδομένων (partition) εξάγεται για κάθε μια από τις 7 μεθόδους ένα διάνυσμα  $106160 \times 1$  το οποίο δείχνει σε ποια κατηγορία ανήκει η κάθε μια από τα 106160 καταγεγραμμένα δάνεια.

Έπειτα, προκύπτουν για κάθε μέθοδο οι πιθανότητες, η κάθε καταχώρηση να ανήκει στην ομάδα δανείων 1 δηλαδή τα μη εξυπηρετούμενα δάνεια. Στη συγκεκριμένη ανάλυση, εξάγονται και οι πιθανότητες η κάθε καταχώρηση να ανήκει στην ομάδα με κωδικό αριθμό 0-δηλαδή τα εξυπηρετούμενα δάνεια οι οποίες είναι συμπληρωματικές με τις πιθανότητες με τις προηγούμενες.

Τα τελευταία δεδομένα δίνουν τους δείκτες ακρίβειας που εξάγονται από την ανάλυση του πειραματικού μέρους. Πιο συγκεκριμένα, για κάθε υποσύνολο, προκύπτει ένας πίνακας όπου περιέχει 7 γραμμές-μία για κάθε μέθοδο και 5 στήλες μια για κάθε δείκτη ακρίβειας. Η πρώτη στήλη δίνει την ακρίβεια της συγκεκριμένης μεθόδου για την πρώτη ομάδα δανείων, η δεύτερη στήλη για τη δεύτερη ομάδα δανείων, η τρίτη στήλη τη συνολική ακρίβεια προβλέψεων ταξινόμησης των δανείων, η τέταρτη στήλη την μέση ακρίβεια της πρώτης και της δεύτερης στήλης και η πέμπτη και τελευταία στήλη την πιθανότητα ένας πελάτης να μην είναι φερέγγυος, η οποία εκφράζεται με την τιμή AUROC (Area Under an Receiver Operation Curve). Αυτό το είδος αποτελεσμάτων θα βοηθήσει στην ταξινόμηση των μεθόδων που χρησιμοποιήθηκαν ως προς την ακρίβεια των αποτελεσμάτων τους. Ο πίνακας 3.5 δίνει τις τιμές του δείκτη AUROC σε κάθε υποσύνολο που χρησιμοποιήθηκε και στη συνέχεια μια 6<sup>η</sup> στήλη η οποία παρουσιάζει τη μέση τιμή ακρίβειας για κάθε αλγόριθμο.

Μέθοδος	part1	part2	part3	part4	part5	Μέση Τιμή
LDA	0,84425	0,84472	0,84543	0,84473	0,84553	0,84493
QDA	0,75032	0,74749	0,74838	0,74842	0,75190	0,74930
LR	0,84642	0,84626	0,84663	0,84579	0,84724	0,84647
SVM	0,73188	0,73577	0,73057	0,73899	0,74418	0,73628
Classification Tree	0,79902	0,80116	0,80192	0,81485	0,79846	0,80308



Boosted Trees	0,83873	0,83955	0,83765	0,83726	0,83959	0,83856
Bagged Trees	0,85366	0,85324	0,85383	0,85330	0,85304	0,85342

Πίνακας 3.5: Πίνακας δεικτών AUROC για κάθε υποσύνολο και κάθε μέθοδο.

Όπως γίνεται αντιληπτό από τη μέση τιμή του υπολογισμού του δείκτη AUROC η μέθοδος που δίνει καλύτερα αποτελέσματα είναι η Bagged Trees. Επίσης η μέση τιμή εξάγει το συμπέρασμα για μια γενική κατάταξη των αλγορίθμων ως προς την ακρίβειά τους με βάση την τιμή του AUROC και αυτή φαίνεται στον πίνακα 3.6.

Κατάταξη	Μέθοδος
1η	Bagged Trees
2η	LR
3η	LDA
4η	Boosted Trees
5η	Classification Tree
6η	QDA
7η	SVM

Πίνακας 3.6: Πίνακας κατάταξης από την ακριβέστερη προς τη λιγότερο ακριβή μέθοδο με κριτήριο την μέση τιμή του δείκτη AUROC.

Στον πίνακα 3.7 θα γίνει σύγκριση των μεθόδων που χρησιμοποιήθηκαν με βάση τη μέση τιμή των τιμών ακρίβειας για κάθε κατηγορία δανείων. Σε αυτή τη συγκεκριμένη σύγκριση προκύπτει επίσης ένας πίνακας κατάταξης, ο πίνακας 3.8 με πλήρη κατάταξη των μεθόδων που χρησιμοποιήθηκαν. Διαπιστώνεται λοιπόν ότι, όπως στην προηγούμενη κατάταξη των αλγορίθμων έτσι και σε αυτή την περίπτωση η μέθοδος Bagged Trees φαίνεται να είναι η πιο αξιόπιστη.

Μέθοδος	part1	part2	part3	part4	part5	Μέση Τιμή
LDA	0,63472	0,76983	0,77069	0,76990	0,76987	0,74300
QDA	0,70759	0,70592	0,70565	0,70611	0,70664	0,70638
LR	0,77010	0,76967	0,77124	0,77032	0,77087	0,77044
SVM	0,69476	0,69770	0,69202	0,70131	0,70559	0,69828
ClassificationTree	0,75977	0,75933	0,76360	0,76064	0,75836	0,76034
Boosted Trees	0,76015	0,76118	0,76099	0,75949	0,76209	0,76078
Bagged Trees	0,77543	0,77470	0,77513	0,77465	0,77534	0,77505

Πίνακας 3.7: Πίνακας μέσης τιμής ακρίβειας για κάθε κατηγορία δανείων για κάθε μέθοδο και κάθε υποσύνολο.

Κατάταξη	Μέθοδος
1	Bagged Trees
2	LR
3	Boosted Trees
4	Classification Tree
5	LDA
6	QDA
7	SVM

Πίνακας 3.8: : Πίνακας κατάταξης από την ακριβέστερη προς τη λιγότερο ακριβή μέθοδο με κριτήριο την μέση τιμή της μέσης ακρίβειας.

# ΚΕΦΑΛΑΙΟ 4

## ΣΥΜΠΕΡΑΣΜΑΤΑ

Σκοπός της συγκεκριμένης εργασίας ήταν η αξιολόγηση πιστοληπτικού κινδύνου στην πλατφόρμα Peer-to-Peer Lending με τη χρήση 7 μεθόδων μηχανικής μάθησης και τελικά η αξιολόγηση των μεθόδων αυτών ως προς την ακρίβεια των αποτελεσμάτων τους.

Αρχικά, παρουσιάζοντας ένα εκτενές θεωρητικό υπόβαθρο γύρω από τον τομέα του Peer-to-Peer Lending δημιουργώντας ένα πλαίσιο γύρω από την γνώση βασικών εννοιολογιών που το πλαισιώνουν και ορισμένων στατιστικών στοιχείων που αποδεικνύουν τη σημαντικότητά του δημιουργήθηκε η ανάγκη μελέτης για το πώς προσεγγίζεται ένα από τα βασικά χαρακτηριστικά του, δηλαδή η αξιολόγηση της πιστοληπτικής ικανότητας ενός δανειολήπτη.

Πιο αναλυτικά, με τη βοήθεια των δεδομένων που ελήφθησαν από την πλατφόρμα prosper.com έγινε χρήση συγκεκριμένων μεταβλητών που καταγράφουν το ιστορικό 106160 καταχωρήσεων δανείων. Έπειτα, πραγματοποιήθηκε διαχωρισμός του δείγματος σε σετ δεδομένων εκπαίδευσης και επαλήθευσης ώστε να ταξινομηθούν οι καταχωρήσεις αυτές σε εξυπηρετούμενα δάνεια και μη εξυπηρετούμενα να υπολογιστεί η πιθανότητα ένταξης τους σε μία από αυτές τις ομάδες και να προσεγγιστεί η ακρίβεια της πρόβλεψης των μεθόδων με 5 δείκτες ακρίβειας.

Πραγματοποιώντας αυτή την ανάλυση προέκυψαν ορισμένα συμπεράσματα κυρίως ως προς την ακρίβεια των μεθόδων που χρησιμοποιήθηκαν (Linear Discriminant Analysis, Quadratic Discriminant Analysis, Logistic Regression, Gaussian Support Vector Machines, Classification Tree, Boosted Trees και Bagged Trees). Τα αποτελέσματα υπολογισμού δείχνουν πως η μέθοδος που έδωσε τα καλύτερα αποτελέσματα ήταν η Bagged Trees και αυτή που έδωσε τα λιγότερο ακριβή η μέθοδος των SVM

Παραλληλίζοντας τη μεθοδολογική προσέγγιση με τη διεθνή βιβλιογραφία που αναλύθηκε σε προηγούμενο κεφάλαιο, τα αποτελέσματα που εξάγονται από τη συγκεκριμένη εργασία εξάγουν παρόμοια συμπεράσματα αφού κατατάσσουν με βάση την ακρίβεια προβλεπτικής ικανότητας δημοφιλής μεθόδους μηχανικής μάθησης.

Φτάνοντας στο τέλος θα μπορούσαν να παρουσιαστούν ορισμένες ενδιαφέρουσες ιδέες για επιπλέον έρευνα και μελέτη:

- Λήψη δεδομένων από άλλες πλατφόρμες Peer-To-Peer Lending και κατασκευή μοντέλου με την ίδια προσεγγιστική πορεία χρησιμοποιώντας τις ίδιες μεθόδους μηχανικής μάθησης και ίδιες ή παρόμοιες μεταβλητές

που προσεγγίζουν τα ιστορικά δεδομένα όπως αυτές της συγκεκριμένης έρευνας.

- Δημιουργία και μελέτη της ίδιας έρευνας με την αξιοποίηση μόνο ποσοτικών μεταβλητών με τη μέθοδο Principal Component Analysis που θα προσεγγίζει πανομοιότυπα αποτελέσματα και θα εξάγει συμπεράσματα με την κατάταξη των μεθόδων μηχανικής μάθησης ως προς την ακρίβεια της προβλεπτικής τους ικανότητας.
- Επέκταση της συγκεκριμένης μεθοδολογίας με περισσότερες μεθόδους και τις ίδιες μεταβλητές με σκοπό την αξιολόγησή τους ως προς την ακρίβεια τους.

# ΒΙΒΛΙΟΓΡΑΦΙΑ

- Atz, U., Bholat, D., (2016). "Peer-to-Peer Lending and Financial Innovation in the United Kingdom", *Staff Working Paper No.598, Bank of England*
- Bellotti T., Crook J., (2009), "Support vector machines for credit scoring and discovery of significant features", *Expert Systems with Applications*, 36, 3302-3308.
- Byanjankar, A., (2018). "Predicting credit risk in Peer-to-Peer lending with survival analysis, in: 2017 IEEE Symposium Series on Computational Intelligence", *SSCI 2017 - Proceedings. Institute of Electrical and Electronics Engineers Inc.*, 1–8.
- Everett, C.R., (2014). "Origins and Development of Credit-Based Crowdfunding", *SSRN Electronic Journal*.
- Famish, C., (2015), "Peer pressure- How peer-to-peer lending platforms are transforming the consumer lending industry", *PwC Consumer Finance*.
- Feis, A., Mehta V., Morris, S., Solitario J., Graaf C., (2016). "P2P Loan Selection Team Members", Stanford University MS&E 448.
- Fisher R.A., (1936). "The use of multiple measurements in taxonomic problems." *Annals of Eugenics* 7: 179–188
- Lessmann S., Baesens B., Seow H.V., Thomas L., (2015), "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research", *European Journal of Operational Research*, 247(1), 124-136.
- Malekipirbazari, M., Aksakalli, V., (2015). "Risk assessment in social lending via random forests", *Expert Systems with Applications* 42, 4621-4631.
- Milne A., Parboteeah P., (2016), "The Business Models and Economics of Peer-to-Peer Lending", *European Credit Research Institute*, 17.
- Narain, B., (1991), "Survival Analysis and the Credit Granting Decision", in: Thomas, L.C., Crook, J.N. and Edelman, D.B., Eds., *Credit Scoring and Credit Control*, OUP, Oxford, 109-121.
- Pokorná, M., Sponer, M., (2016). Social Lending and Its Risks. *Procedia - Social and Behavioral Sciences* 220, 330–337.
- Polena M., (2017). "Performance Analysis of Credit Scoring Models on Lending Club Data", *Faculty of Social Sciences Institute of Economic Studies*.
- Puro L., Teich J., Wallenius H., Wallenius J. (2010), "Borrower Decision Aid for people-to-people lending", *Decision Support Systems*, 49(1), 52-60.
- Serrano-Cinca, C., Gutiérrez-Nieto B., (2016), "The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending", *Decision Support Systems*, 89, 113-122.

- Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). “Determinants of default in P2P lending”. *PLoS ONE*, 10(10).
- Smith W.L., (1947). Taxation and finance. *National Municipal Review*, 36(4), 227-230.
- Stiglitz, E., Weiss, A., (1981), “Credit Rationing in Markets with Imperfect Information”, *The Economic American Review*, 71(3), 393-410.
- Tasca, P., Aste, T., Pelizzon, L., Perony, N. (Eds.), (2016). “Banking Beyond Banks and Money”, *Springer International Publishing*.
- Torre, M.L., Mango, F., 2005. “Social Lending in Europe: Structures, Regulation and Pricing Models”, in: G. Radoselovics (ed.), *Crisis, Risk and Stability in Financial Markets*, Palgrave Macmillan Studies in Banking and Financial Institutions.
- Tsai K., Ramiah S., Singh S., (2014), “Peer Lending Risk Predictor”, *Stanford University CS229*.
- Wang, H., Greiner, M., Anderson, J.E., Wang, H., Greiner, M., (2009). People to People Lending: The Emerging e-Commerce Transformation of a Financial Market 182–195.
- Wang, J.G., Xu, H., Ma, J., (2015). “Financing the underfinanced: Online lending in China, Financing the Underfinanced: Online Lending in China”, *Springer Berlin Heidelberg*.
- Yum, H., Lee, B., Chae, M., (2012). From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. *Electronic Commerce Research and Applications*, 11, 469–483.
- Γκρούγια Ε., (2013). “Δέντρα ταξινόμησης και Bagging Μέθοδοι για την κατασκευή μοντέλων”, Εθνικό Μετσόβιο Πολυτεχνείο.
- Δούμπος Μ., Ζοπουνίδης Κ. (2001). “Πολυκριτήριες τεχνικές ταξινόμησης-Θεωρία και Εφαρμογές”. *Αθήνα: Κλειδάριθμος*.
- Κυρίτσης Κ., (2014). “Νευρωνικά Δίκτυα και Μηχανές Διανυσματικής Υποστήριξης”, Πανεπιστήμιο Πατρών.
- Ρεθεμιωτάκη Ε., (2016). “Υπολογιστικές μεθοδολογίες για την πρόβλεψη της πτώχευσης τραπεζών”, Πολυτεχνείο Κρήτης.

## Διαδικτυακή Βιβλιογραφία

- <https://www.investopedia.com/terms/f/fintech.asp>
- <https://www.maxmag.gr/epichirimatikotita/fintech-to-mellon-tou-chrimatistiriou/>
- <https://fleximize.com/articles/006773/crowdfunding-and-peer-to-peer-lending>
- <https://www.cbinsights.com/research/alternative-lending-timeline-expert-intelligence/>

<http://www.crowdready.com.au/news-1/2015/6/equity-crowdfunding-chronology>  
<https://www.fundingcircle.com/uk/businesses/peer-to-peer-vs-crowdfunding/>  
<https://fleximize.com/articles/006773/crowdfunding-and-peer-to-peer-lending>  
<http://www.crowdready.com.au/news-1/2015/6/equity-crowdfunding-chronology>  
<https://www.cbinsights.com/research/alternative-lending-timeline-expert-intelligence/>  
<https://eclass.aueb.gr/modules/document/file.php/MISC216/crowdfunding.pdf>  
<https://medium.com/the-crowdpolicy-collection/crowdfunding-in-greece-522752ed86ff>  
<https://dspace.lib.uom.gr/bitstream/2159/19706/6/PazoniAikateriniMsc2016.pdf>  
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2234765](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2234765)  
<http://www.nber.org/chapters/c12946.pdf>  
<https://onlinelibrary.wiley.com/doi/epdf/10.1111/jems.12093>  
<https://www.fundable.com/crowdfunding101/types-of-crowdfunding>  
<https://www.gobankingrates.com/loans/personal/5-perks-peer-to-peer-lending/>  
<https://www.investopedia.com/articles/financial-theory/08/peer-to-peer-lending.asp>  
<https://learn.off3r.com/advantages-disadvantages-p2p-lending/>  
<https://www.quora.com/What-are-the-advantages-and-disadvantages-of-P2P-lending>  
<https://www.ey.com/Publication/vwLUAssets/ey-fintech-adoption-index-2017/%24FILE/ey-fintech-adoption-index-2017.pdf>  
<https://www.pwc.com/gx/en/industries/financial-services/assets/pwc-global-fintech-report-2017.pdf>  
<https://newsroom.accenture.com/news/global-venture-capital-investment-in-fintech-industry-set-record-in-2017-driven-by-surge-in-india-us-and-uk-accenture-analysis-finds.htm>  
<https://www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/us-dcfs-fintech-by-the-numbers-web.pdf>  
<https://www.orcamoney.com/blog/peer-to-peer-lending-review>  
<http://peersociallending.com/investing/peer-to-peer-lending-sites-16-of-the-worlds-best/>  
<https://www.fundingcircle.com/uk/impact>  
<https://www.ratesetter.com/invest/investing-with-us>  
<https://marcoschwartz.com/best-p2p-platforms>  
<https://www.mintos.com/en/statistics/>  
<https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/understanding-alternative-finance/>

<https://icorating.com/upload/whitepaper/zHjM1Q8Yo9b47DHjQIGV67trd3A5H9CkPf18DgeN.pdf>