

Διπλωματική Εργασία

«ΟΜΟΤΙΜΟΣ ΔΑΝΕΙΣΜΟΣ
(PEER-TO-PEER LENDING) ΚΑΙ Η
ΠΡΟΒΛΕΨΗ ΤΗΣ ΠΙΘΑΝΟΤΗΤΑΣ
ΑΘΕΤΗΣΗΣ ΤΩΝ ΔΑΝΕΙΩΝ»

της ΔΕΣΠΟΙΝΑΣ ΒΕΡΓΙΔΟΥ

Επιβλέπων καθηγητής: ΙΩΑΝΝΗΣ ΠΑΠΑΝΑΣΤΑΣΙΟΥ

Οκτώβριος 2022

Περίγραμμα παρουσίασης

- Στόχοι
- Ερευνητικά ερωτήματα
- Βιβλιογραφική ανασκόπηση
- Δεδομένα έρευνας
- Μεθοδολογία έρευνας
- Ανάλυση αποτελεσμάτων
- Συμπεράσματα
- Περιορισμοί
- Προτάσεις για μελλοντική έρευνα
- Βιβλιογραφία

Στόχοι

1. Εύρεση των προσδιοριστικών παραγόντων της αθέτησης δανείων P2P
2. Σύγκριση της προβλεπτικής ικανότητας ενός παραδοσιακού μοντέλου με ένα μοντέλο μηχανικής εκμάθησης

Ερευνητικά ερωτήματα

1. Ποιές μεταβλητές επηρεάζουν την πιθανότητα αθέτησης;
2. Είναι εφικτή η ικανοποιητική πρόβλεψη του κινδύνου αθέτησης με αυτές τις μεταβλητές;
3. Παραδοσιακό μοντέλο έναντι μηχανικής εκμάθησης-ποιό υπερτερεί;

Βιβλιογραφική ανασκόπηση (1/2)

Μακροοικονομικοί παράγοντες

- COVID-19
- επιτόκια δανεισμού
- πληθωρισμός
- επιχειρηματικό περιβάλλον
- οικονομικές συνθήκες της χώρας

Βιβλιογραφική ανασκόπηση(2/2)

Παράγοντες σχετικοί με τον δανειολήπτη και το δάνειο

- φύλο, εμφάνιση
- πιστωτική βαθμολογία
- λόγος DTI, εισόδημα
- τρέχουσα κατάσταση κατοικίας
- σκοπός, διάρκεια, ποσό, επιτόκιο δανείου
- μορφωτικό επίπεδο, εργασιακή εμπειρία
- συχνότητα λήψης δανείων

Δεδομένα έρευνας

- Πλατφόρμα Bondora, Εσθονία
- 152.542 παρατηρήσεις-δάνεια και 16 μεταβλητές
- Χρονική περίοδος: 2009-μέσα 2022
- Εξαρτημένη μεταβλητή (τελικού δείγματος):
Κατάσταση δανείου (Status)
- Επεξηγηματικές μεταβλητές:
 - Μεταβλητές χαρακτηριστικές του δανείου
 - Μεταβλητές χαρακτηριστικές των δανειοληπτών

Μεθοδολογία έρευνας (1/4)

Προσαρμοσμένο μοντέλο λογιστικής παλινδρόμησης

Χρήση προσαρμοσμένου μοντέλου λογιστικής παλινδρόμησης διότι:

- εξηγεί καλύτερα τους παράγοντες που επηρεάζουν την πιθανότητα ενός γεγονότος
- κατάλληλο για την παροχή προβλέψεων της πιθανότητας ενός γεγονότος
- ερευνά τη μη γραμμικότητα
- δημιουργεί έναν κανόνα ταξινόμησης των παρατηρήσεων με βάση την προβλεπόμενη πιθανότητα

Μεθοδολογία έρευνας (2/4)

Εξίσωση λογιστικής παλινδρόμησης

- Έστω διωνυμική ποιοτική μεταβλητή (πχ A ή B),
- Ψευδομεταβλητή Y (πχ 1 αν A και 0 αν B),
- $\pi = P(Y=1)$,
- k παράγοντες πρόβλεψης, τότε η εξίσωση λογιστικής παλινδρόμησης είναι:

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + u \quad (1.1)$$

Μεθοδολογία έρευνας (3/4)

Μέθοδος εκτίμησης

- Μέθοδος εκτίμησης των μοντέλων λογιστικής παλινδρόμησης → μέγιστης πιθανοφάνειας
- Μέθοδος εκτίμησης των μοντέλων κλασικής γραμμικής παλινδρόμησης με συνεχή εξαρτημένη μεταβλητή → ελαχίστων τετραγώνων

Μεθοδολογία έρευνας (4/4)

Μέτρηση της προβλεπτικής ικανότητας

- Σύγκριση με το μοντέλο μηχανικής εκμάθησης BForest
- Αξιολόγηση των προβλέψεων σε 3 στάδια:
 1. Διαχωρισμός συνολικού δείγματος σε δείγμα εκπαίδευσης (75%) και δείγμα ελέγχου (25%)
 2. Εκτίμηση των δύο μοντέλων με χρήση του δείγματος εκπαίδευσης
 3. Υπολογισμός δεικτών μέτρησης της ακρίβειας των προβλέψεων σε σχέση με το δείγμα ελέγχου

Ανάλυση αποτελεσμάτων (1/7)

Τελική συνάρτηση λογιστικού μοντέλου

$$\begin{aligned} Y_i = & \beta_0 + \beta_1(\text{Age})_i + \beta_2(\text{IcomeTotal})_i + \beta_3(\text{Interest})_i + \beta_4(\text{LoanDuration})_i + \\ & \beta_5(\text{NoOfPreviousLoansBeforeLoan})_i + \beta_6(\text{Male})_i + \beta_7(\text{Woman})_i + \beta_8(\text{Primary})_i + \\ & \beta_9(\text{Basic})_i + \beta_{10}(\text{Vocational})_i + \beta_{11}(\text{Secondary})_i + \beta_{12}(\text{Tenant, Other})_i + \beta_{13}(\text{AA or A})_i \\ & + \beta_{14}(\text{B})_i + \beta_{15}(\text{C})_i + \beta_{16}(\text{D})_i + \beta_{17}(\text{E})_i + \beta_{18}(\text{F})_i + \beta_{19}(\text{No})_i + u_i, \end{aligned} \quad (1.2)$$

όπου $Y = \log\left(\frac{\pi}{1-\pi}\right)$ και $\pi = P(\text{Status} = 1)$ πιθανότητα αθέτησης εξόφλησης του δανείου.

Ανάλυση αποτελεσμάτων (2/7)

Ποσοτικές μεταβλητές

- επιτόκιο → θετική σχέση \Leftrightarrow επιτόκιο \uparrow 10%, PD \uparrow 1,8%
- διάρκεια → θετική σχέση \Leftrightarrow διάρκεια \uparrow 10 μήνες, PD \uparrow 2,5%
- ηλικία → θετική σχέση \Leftrightarrow ηλικία \uparrow 10 έτη, PD \uparrow 2%
- εισόδημα → θετική σχέση \Leftrightarrow εισόδημα \uparrow 10%, PD \uparrow 0,01%
- αριθμός προηγούμενων δανείων → θετική σχέση \Leftrightarrow αρ. προηγ. δανείων \uparrow 10 δάνεια, PD \uparrow 2,6%

Ανάλυση αποτελεσμάτων (3/7)

Ποιοτικές μεταβλητές (1/2)

- φύλο:
 - απροσδιόριστο → θετική σχέση, PD ↑ (PD απροσδιόριστο 81% > PD αντρών και 2,3 φορές > PD γυναικών)
 - άνδρες/γυναίκες → αρνητική σχέση, PD ↓ (PD αντρών κατά 1,24 > PD γυναικών)
- εκπαίδευση:
 - βασική ή δευτεροβάθμια → αρνητική σχέση, PD ↓ (PD δευτεροβάθμια 44% > PD βασική)
 - επαγγελματική ή πρωτοβάθμια → θετική σχέση, PD ↑ (PD επαγγελματική 94% > PD βασική, 35% > PD δευτεροβάθμιας και 56% > PD ανώτερης), (PD πρωτοβάθμια 156% > PD βασική, 1,3 φορές > PD επαγγελματική, 2,1 φορές > PD ανώτερης και 1,8 φορές > PD δευτερεύουσας)

Ανάλυση αποτελεσμάτων (4/7)

Ποιοτικές μεταβλητές (2/2)

- κατάσταση στέγασης:
 - ιδιοκτήτες ή με υποθήκη → αρνητική σχέση, PD ↓
 - ενοικιαστές ή άλλο → θετική σχέση, PD ↑ (PD ενοικιαστών 1,59 φορές > PD ιδιοκτητών)
- αξιολόγηση → αρνητική σχέση ⇔ αξιολόγηση ↑, PD ↓ (PD δανείων βαθμού D > PD δανείων βαθμού B)
- μη-αναδιάρθρωση → αρνητική σχέση, PD ↓ (PD αναδιαρθρωμένων 198% > PD μη-αναδιαρθρωμένων)

Ανάλυση αποτελεσμάτων (5/7)

Μέτρα αξιολόγησης των προβλέψεων (1/3)

1. Μέτρα πρόβλεψης των μοντέλων

	RMSE	MAD	MR
Logistic	0,4523	0,4103	0,3238
BForest	0,4462	0,4160	0,3068

- ✓ BForest → μεγαλύτερη ακρίβεια πρόβλεψης, με βάση τους RMSE και MR, διότι $0,4462 < 0,4523$ και $0,3068 < 0,3238$, ενώ
- ✓ Λογιστικό → ακριβέστερο με βάση τον MAD, καθώς $0,4103 < 0,4160$

Ανάλυση αποτελεσμάτων (6/7)

Μέτρα αξιολόγησης των προβλέψεων (2/3)

2. Πίνακας σύγχυσης (Confusion Matrix)

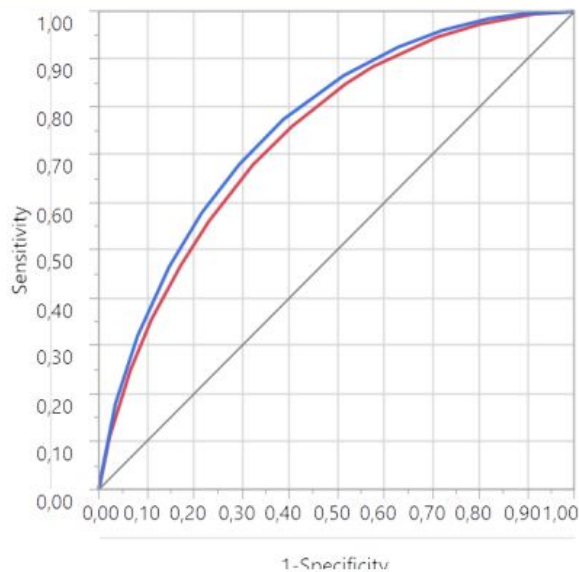
Status	Logistic		BF	
	P-Repaid	P-Default	P-Repaid	P-Default
A-Repaid	0,734	0,266	0,731	0,269
A-Default	0,392	0,608	0,351	0,649

- ✓ BForest → μεγαλύτερη ακρίβεια πρόβλεψης, καθώς σωστή πρόβλεψη αθέτησης BForest (64,9%) > σωστή πρόβλεψη αθέτησης λογιστικού (60,8%)

Ανάλυση αποτελεσμάτων (7/7)

Μέτρα αξιολόγησης των προβλέψεων (3/3)

3. Καμπύλη ROC (Receiver Operating Characteristic)



	Predictor	AUC
—	Logistic	0,7436
—	BForest	0,7638

- ✓ BForest → υπερτερεί στην ακρίβεια πρόβλεψης με ποσοστό 76,38%, έναντι 74,36% του λογιστικού

Συμπεράσματα (1/2)

Για πρώτο στόχο

- Μεταβλητές χαρακτηριστικές του δανείου:
 - επιτόκιο, διάρκεια → θετική σχέση με PD
 - μη-αναδιαρθρωμένο δάνειο, δάνεια με υψηλή βαθμολογία → αρνητική σχέση με PD
- Μεταβλητές χαρακτηριστικές των δανειοληπτών:
 - ηλικία, εισόδημα, αριθμός προηγούμενων δανείων, δανειολήπτες που δηλώνουν φύλο απροσδιόριστο → θετική σχέση με PD
 - δανειολήπτες με βασική ή δευτεροβάθμια εκπαίδευση, ιδιοκτήτες, με υποθήκη → αρνητική σχέση με PD

Συμπεράσματα (2/2)

Για δεύτερο στόχο

- Μέτρα αξιολόγησης των προβλέψεων:
 1. Μέτρα πρόβλεψης των μοντέλων
 - ✓ BForest → μεγαλύτερη ακρίβεια πρόβλεψης, με βάση τους RMSE και MR ενώ,
 - ✓ Λογιστικό → ακριβέστερο με βάση τον MAD
 2. Πίνακας σύγχυσης (Confusion Matrix)
 - ✓ BForest → μεγαλύτερη ακρίβεια πρόβλεψης
 3. Καμπύλη ROC (Receiver Operating Characteristic)
 - ✓ BForest → μεγαλύτερη ακρίβεια πρόβλεψης

Περιορισμοί

- Δεδομένα από μία μόνο πλατφόρμα P2P δανεισμού
- Μη χρήση τρεχόντων (current) δανείων (μόνο εξοφλημένα και αθετημένα) στο τελικό δείγμα
- Χρήση μόνο τριών μέτρων σύγκρισης

Προτάσεις για μελλοντική έρευνα

- Σύγκριση μεταξύ διαφορετικών πλατφορμών P2P, σε διαφορετικές ηπείρους και οικονομικές συνθήκες
- Ένταξη των τρεχόντων δανείων
- Μεγαλύτερη σύγκριση με εναλλακτικούς δείκτες αναφοράς
- Εύρεση του μέγιστου ποσού έκθεσης των loan originators σε μία χώρα, σε σχέση με το συνολικό ποσοστό έκθεσης σε όλες τις άλλες.

Βιβλιογραφία

- Nigmonov, A., & Shams, S. (2021). COVID-19 pandemic risk and probability of loan default: evidence from marketplace lending market. *Financial Innovation*, Vol. 7, 83
- Polena, M., & Regner, T. (2018). Determinants of Borrowers' Default in P2P lending under Consideration of the Loan Risk Class. *MDPI Games*.
- Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of Default in P2P Lending. *Plos One*, pp. 1-22.
- Chen, X., Huang, B., & Ye, D. (2019). Gender gap in peer-to-peer lending: Evidence from China. *Journal of Banking & Finance*, Vol. 112, 105633.

Ευχαριστώ για την προσοχή σας!
Ερωτήσεις-Παρατηρήσεις;