



Aalto-yliopisto
Perustieteiden
korkeakoulu

Luottoriskien arviointi logistisella regressiolla (aihe-esittely)

Olivia Antikainen

02.12.2021

Ohjaaja: Prof. *Ahti Salo*

Valvoja: Prof. *Ahti Salo*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

Tausta

- Luottoriski tarkoittaa riskiä, joka liitetään lainanottajan maksukyvyttömyyteen
 - Mallit ennustavat maksukyvyttömyyttä todennäköisyyksinä
- Luottoriskien arviointi on yksi lainanantajien tärkeimmistä työkaluista
 - Haettujen lainahakemusten riskien arviointi → parantaa tehokkuutta ja tuottavuutta
- Luottoriskien arvioiminen on kehittynyt
 - Enemmän dataa ja uusia menetelmiä → edut ja haasteet
- Luottokorttien luottoriskin arvioiminen
 - Luottokorttien käyttöä seurataan jatkuvasti, luottokorttiyhtiöt pystyvät tarvittaessa puuttumaan → tuottavuus kasvaa

Tavoitteet

- Arvioidaan luottokorttiasiakkaiden luottoriskiä
 - Maksukyvyttömyyden todennäköisyys
- Kehitetään malli hyödyntäen dataa luottokorttiyhtiön asiakkaiden luotto- ja muista lähtötiedoista
- Selvitetään mitkä tekijät vaikuttavat riskin todennäköisyyden kasvuun ja miten

Rajaukset

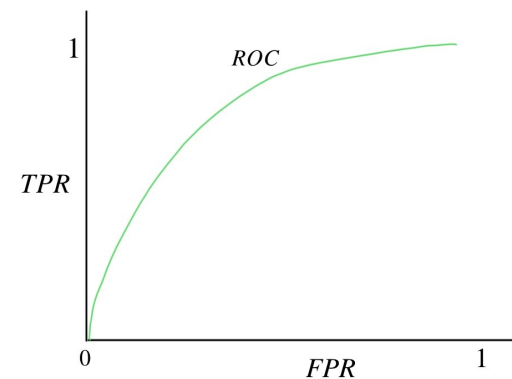
- Toteutetaan logistinen regressiomalli
- Hyödynnetään yhtä data-aineistoa mallin tekemisessä
- Työssä ei tarkastella luottoriskin ehkäisemistä, eikä muuttujien valintaa tai mahdollista puolueellisuutta

Menetelmät ja työkalut

- Logistinen regressio
 - Luottoriski binäärisenä luokitteluongelmana
 - Selitettävänä muuttujana maksukyvyttömyys
 - Selittävinä muuttujina maksuhistoria (määrä ja aika), luottokortin käyttötiedot, luoton määrä sekä kortin haltijan ikä, sukupuoli, koulutus ja siviilisäät

	Ennustettu 0	Ennustettu 1
Todellinen 0	TN	FP
Todellinen 1	FN	TP

- Mallin arviointi:
 - Luokittelutaulukko (Accuracy, Precision, Sensitivity)
 - ROC-käyrä (receiver operating characteristic) → AUC (area under the curve)
- Toteutus RStudion avulla



Aikataulu

- Aiheeseen tutustuminen 11/2021
- Aiheen esittely 12/2021
- Työn kirjoittaminen 12/2021-01/2022
- Mallin rakentaminen 12/2021
- Valmis työ 01/2022

Tietolähteet ja aineisto

- Data-aineisto: University of California, Irvine (UCI) / default of credit card clients
 - Tietolähteet:
 - World Bank (2019) 'Credit Scoring Approaches Guidelines', The World Bank Group.
 - Witzany, J. (2017) '*Credit Risk Management Pricing, Measurement, and Modeling*', 1st ed. 2017.
 - Hilbe, J. M. (2009) '*Logistic Regression Models*', CRC Press.
 - Fawcett, T. (2005) 'An Introduction to ROC Analysis', Pattern Recognition Letters, vol. 27, s. 861-874
-