

UNIVERSITY OF TARTU
Institute of Computer Science
Conversion Master in IT

Liisa Kallas

Mining Gold and Silver Using Machine Learning: Predicting
Cancelled Orders at the Online Store of Tavex Norway
Master's Thesis (15 ECTS)

Supervisor:
Rajesh Sharma, PhD

Tartu 2019

Mining Gold and Silver Using Machine Learning: Predicting Cancelled Orders at the Online Store of Tavex Norway

Abstract:

Widespread usage of Internet has given rise to huge amounts of online platforms, among which are e-commerce sites. Many traditionally brick-and-mortar stores, like precious metals dealers, have also opened e-commerce platforms to sell their products. Moving business online has benefits of reaching a wider audience and can be more convenient for clients. On the other hand, it has also brought along new business challenges like how to get more website users to purchase and the attempt to predict which website visitors are going to place an order during their visit. This has been a popular topic of many research papers. A less covered area in previous studies is the problem of online orders that had been already placed, but then cancelled. The cancellation problem that happens due to customers not paying for their order has been also noted in the e-commerce site of a precious metals dealer. In this work, a data analytics approach is taken to predict, whether an online order will be cancelled or not after its placement at Tavex Gull og Sølv AS. Three supervised machine algorithms, namely logistic regression, random forest and support vector machine, were used. The best achieved model had the ability to identify 68% of all unpaid orders. These results indicate that data analytics can be very promising in predicting these types of orders.

Keywords:

data mining, machine learning, e-commerce, cancelled orders, precious metals

CERCS: P160 Statistics, operation research, programming, actuarial mathematics

Masinõppe abil kulda ja hõbedat kaevandades: tühistatud tellimuste ennustamine Tavexi e-poes Norras

Lühikokkuvõte:

Laialdase internetikasutuse tõttu on tekkinud palju erinevaid veebikeskkondi. Nende hulka kuuluvad ka e-poes. Paljud traditsioonilised ärid, nagu väärismetalle müüvad ettevõtted, on lisaks olemasolevatele esindustele loonud enda toodete müümiseks ka veebipoed. Ühest küljest võimaldavad need klientidel mugavamalt ostelda, teisalt pakuvad ka ärile võimaluse jõuda suurema hulga potentsiaalsete ostjateni. Ometi on see toonud firmadele ka uusi väljakutseid, näites kuidas panna üha rohkem veebilehe külastajaid ostu tegema. Paljude varasemate uurimuste eesmärk ongi olnud ennustada, kes veebilehe kasutajatest enda külastuse käigus tellimuse teevad ja mis faktorid seda otsust mõjutavad. Varasemates uurimustes on aga vähe käsitletud probleemi, mis tekib juba esitatud veebitellimuste tühistamisest. Tühistatud tellimused, mis tulenevad nende maksmata jätmisest, on probleem muuhulgas väärismetalle müüva ettevõtte Tavexi e-poes Norras. Käesolevas töös kasutatakse kolme masinõppe algoritmi – logistilist regressiooni, juhumetsa ja tugivektormasinat –, et ennustada millised Tavexi veebipoes tehtud tellimused makstakse ja millised tühistatakse. Parim saavutatud mudel suutis tuvastada 68%

kõikidest maksmata jäetud tellimustest. Töötulemused näitavad, et masinõppel on selles valdkonnas ja seda tüüpi tellimuste tuvastamiseks suur potentsiaal.

Võtmesõnad:

Andmekaeve, masinõpe, e-kaubandus, tühistatud tellimused, väärismetallid

CERCS: P160 Statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika

Non-exclusive licence to reproduce thesis

I, Liisa Kallas

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for the purpose of preservation the DSpace digital archives until the expiry of the term of copyright,

Mining Gold and Silver Using Machine Learning: Predicting Cancelled Orders at the Online Store of Tavex Norway

supervised by Rajesh Sharma, PhD.

Publication of the thesis is not allowed.

2. I am aware of the fact that the author retains the right specified in p. 1.
3. This is to certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

14/08/2019