

UNIVERSITY OF TARTU
Institute of Computer Science
Data Science Curriculum

Joosep Hook, Kea Kohv

Data Analysis and Machine Learning for Financial Insights

Capstone Project (15 ECTS)

Supervisors: Üllar Rannik, PhD
Jaak Vilo, PhD

Data Analysis and Machine Learning for Financial Insights

Abstract:

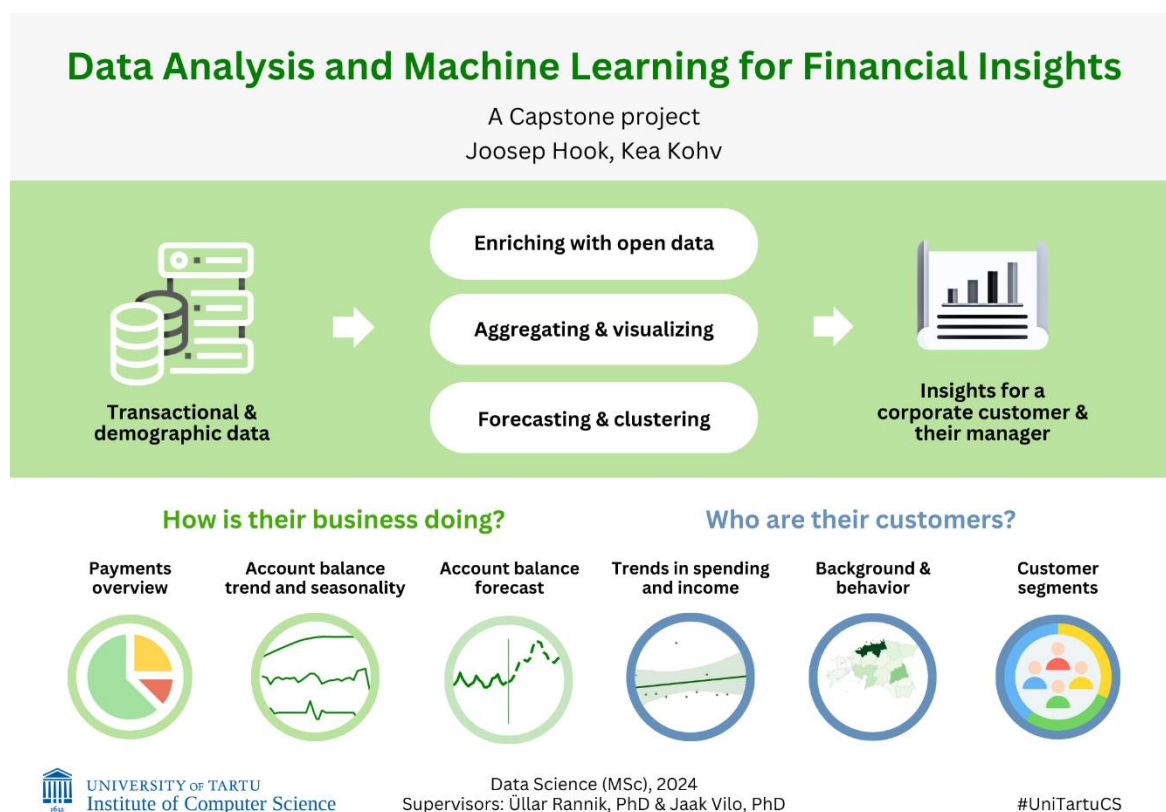
In banking, automated data analysis and machine learning tools can support high-level decision making and reduce the burden of manual work. In addition, automated data analysis can provide up-to-date insights for both the bank and its customers. The goal of this Capstone project was to explore what meaningful data-driven insights about Estonian corporate customers can be derived from transactional and socio-demographic data. The project involved developing relevant queries, enriching the dataset with open data, aggregating and visualizing, decomposing time series, detecting trends and change points, time series forecasting, and segmenting customers. By providing the corporate customer's identification code and data period, the created tool automatically generates a report specific to that corporate customer. The report covers two perspectives: firstly, an analysis of the business's financials; secondly, profile of its customers. Only aggregated statistics are used to describe the household customers of the business. This tool benefits customer managers by saving time and enabling a better understanding of the business and the needs of the corporate customer.

Keywords:

Banking, trend detection, forecasting, segmentation

CERCS: P176 - Artificial intelligence, P160 - Statistics, operation research, programming, actuarial mathematics

Visual abstract:



Ettevõtte ülevaade andmeanalüüsi ja masinõppega

Lühikokkuvõte:

Automaatse andmeanalüüsi ja masinõppe toel saab panganduses teha paremaid otsuseid, vähendada käsitööd ja leida uusi kasulikke teadmisi nii pangale kui ka panga klientidele. Lõpuprojekti eesmärk oli leida, milliseid andmepõhiseid teadmisi on võimalik Eesti ettevõtete tehingute ja nende klientide sotsiaaldemograafiliste andmete põhjal tuletada. Projekt hõlmas andmebaasipäringute kirjutamist, andmete täiendamist avaandmetega, andmete koondamist ja visualiseerimist, aegridade dekomponeerimist, trendide ja muutuspunktide tuvastamist, aegridade prognoosimist ja klientide segmenteerimist. Ettevõtte koodi ja soovitud perioodi sisestamisel genereerib loodud tööriist selle ettevõtte kohta aruande. Aruanne hõlmab kahte perspektiivi: ettevõtte finantsnäitajate analüüsi ja klientide profiili ülevaadet. Eraklientide kirjeldamiseks kasutatakse ainult üldistatud andmeid. Loodud tööriist võimaldab säästa kliendihaldurite aega ning aitab neil paremini mõista ärikliente ja nende vajadusi.

Võtmesõnad:

Pangandus, trendi tuvastamine, aegridade prognoosimine, segmenteerimine

CERCS: P176 - Tehisintellekt, P160 - Statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika

Visuaalne abstrakt:



License

Non-exclusive licence to reproduce the thesis and make the thesis public

We, Joosep Hook and Kea Kohv

1. grant the University of Tartu a free permit (non-exclusive licence) to:

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, our Capstone project report

Data Analysis and Machine Learning for Financial Insights,

supervised by Üllar Rannik, PhD, and Jaak Vilo, PhD,

2. We grant the University of Tartu the permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work from **11/04/2029** until the expiry of the term of copyright,
3. We are aware that the author retains the rights specified in points 1 and 2.
4. We confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Tartu, 15.05.2024