Application of mixed data types customer profile clustering for personalized online product ranking

Giovane Silva[1], Samuel Tito Infantas[1], Rute Souza de Abreu[1], Isabelle Cunha Valim[1], Aline Cristina Jacomassi[1], Laion Lima Boaventura[1]
[1] Lendico Brazil

Abbreviated abstract: Over the last few years, data has become extremely important for accurate decision-making. In the financial market, customers' behavior has changed, and there is an increasing demand for personalized products, which leads to a faster and better experience for the user. This project proposes a methodology using artificial intelligence, for ranking the number of installments suggestions in loan requests, according to a subjective relevance score and the user profile.

Related publications:

1- Ren, K., & Malik, A. (2019). Recommendation Engine for Lower Interest Borrowing on Peer to Peer Lending (P2PL) Platform. *IEEE/WIC/ACM International Conference on Web Intelligence*.

2 - Patel, S. B., Bhattacharya, P., Tanwar, S., & Kumar, N. (2021). KiRTi: A Blockchain-Based Credit Recommender System for Financial Institutions. *IEEE Transactions on Network Science and Engineering*, *8*(2), 1044–1054.

3 - Gavalas, D., & Syriopoulos, T. (2014). An integrated credit rating and loan quality model: application to bank shipping finance. *Maritime Policy & Management*, *42*(6), 533–554

4 - Crouhy, M., Galai, D., & Mark, R. (2001). Prototype risk rating system. *Journal of Banking & Finance*, *25*(1), 47–95.

5 - Wu, J., & Xu, Y. (2011). A Decision Support System for Borrower's Loan in P2P Lending. *Journal of Computers*, *6*(6). https://doi.org/10.4304/jcp.6.6.1183-1190

3rd Conference on
Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Problem, Data, Previous Works

01

Financial market recommendation systems use client information to offer loan recommendations **based only** on a computed **risk score** [1-4].

02

Considering exclusively the client's risk and not their **profile adherence** to the **loan requested is not always the best way** to perform grant loan decisions
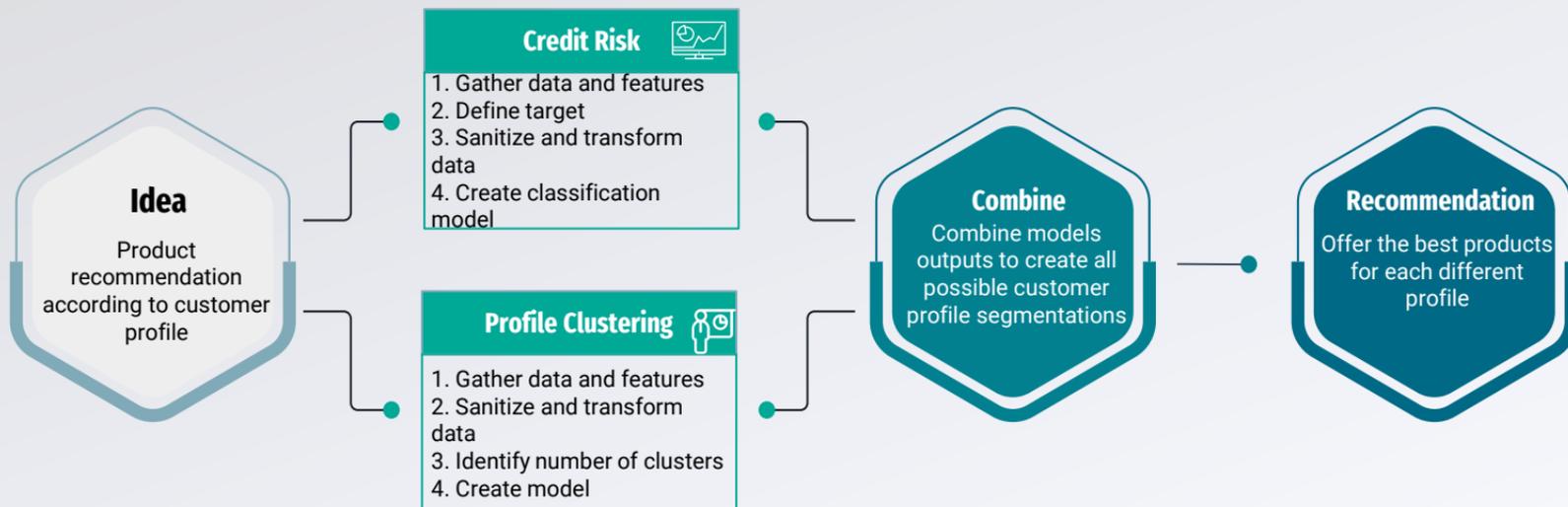
03

In [5], Wu, J., & Xu, Y. proposed a **decision support system** to support borrower's loan, used ontologies for embedding **knowledge** and its **reasoning for making the credit decision.**

04

This paper proposes a methodology that uses a concept similar to [5] but leveraging **artificial intelligence models** to effectively **map the exact customer profile**, such as **clustering** and **regression** and **offer the best loan recommendation.**

3rd Conference on
Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Methods

Idea

Product recommendation according to customer profile

Credit Risk

1. Gather data and features
2. Define target
3. Sanitize and transform data
4. Create classification model

Profile Clustering

1. Gather data and features
2. Sanitize and transform data
3. Identify number of clusters
4. Create model

Combine

Combine models outputs to create all possible customer profile segmentations

Recommendation

Offer the best products for each different profile

3rd Conference on
Statistics and Data Science
Salvador, Brazil (online)
October 28-30, 2021

Results and Conclusions

Leverage **categorical features** during credit risk and profile clustering steps.

At this point, we moved forward with product ordering and offering by using the **mean relevance score**, due to amount of data

Even with a basic initial approach, **good results** were attained, specifically in **group A risk**

It is important to acknowledge that **business decisions** may have had a **great impact on results**, which may lead to different performance in the future

Therefore, the **project presented good initial results and has room for improvements**, leading to the following planned next steps:

Apply the methodologies on customers, measure performance in the proper setup;
Gather more data for next months;
More robust product ordering method, **reinforcement learning algorithms**, automate decision making
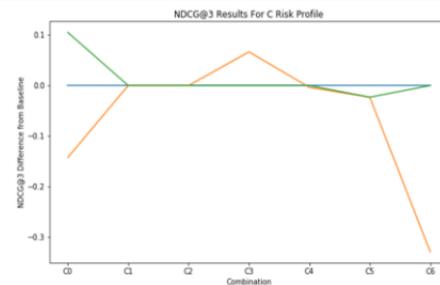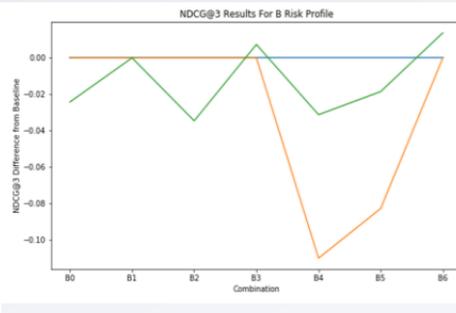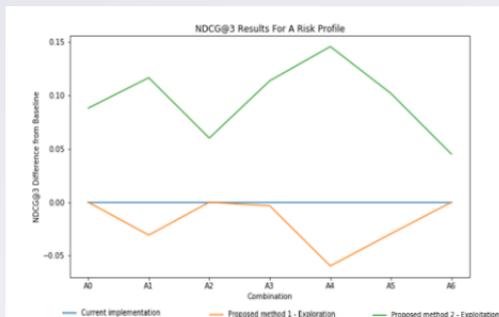
*NDCG: Normalized discounted cumulative gain