

Personalised Recommendations for Students

Author: Ing. Čeněk Žid

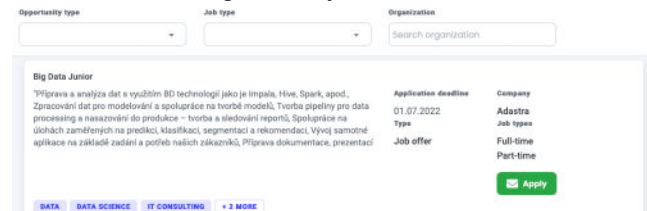
Supervisor: doc. Ing. Pavel Kordík, Ph.D.



FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Abstract

Modern technical universities help students with getting practical experience. They educate thousands of students and it is hard to connect individual students with relevant industrial experts and opportunities. This thesis aims to solve this problem by designing a matchmaking widget powered by a recommendation system (RS). We show that recommendations based on student profile and interaction history improve conversion rate significantly.



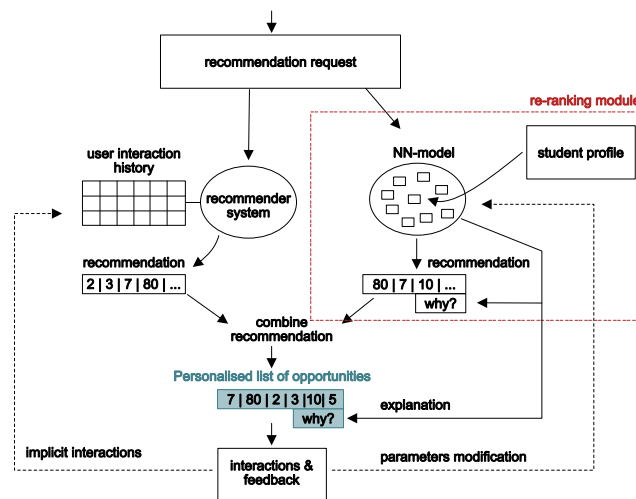
Current State-of-the-art

Standard recommendations based on interaction history are well studied. We wanted to improve cold start performance by employing a student profile. There are some related studies such as Mining Skills from Educational Data for Project Recommendations, or Job recommendation algorithm for graduates based on personalised preference, but none of them uses recommender in combination with a re-ranking module as we do in the thesis.

Data and Proposed Solutions

Overall, we work with 166 opportunities (part-time jobs, internships, etc.) for 1806 potential students. The RS is split into two separate parts:

1. **Interaction-based RS** implemented using pre-prepared solution, the Recombee platform that can work without the other part.
2. **Re-ranking module** that utilises student profiles to enhance recommendation capabilities.



Re-ranking Module

The key part is the re-ranking module implemented using term frequency inverse document frequency algorithm based on a limited set of keywords. The database is created by combining LinkedIn skills, ACM classification skills, and It's your skill dataset. Student profiles are extracted from studied courses into a student keyword embedding that is matched to opportunity embeddings.

$$tfidf(t, d, D) = tf(t, d) \times idf(t, D)$$

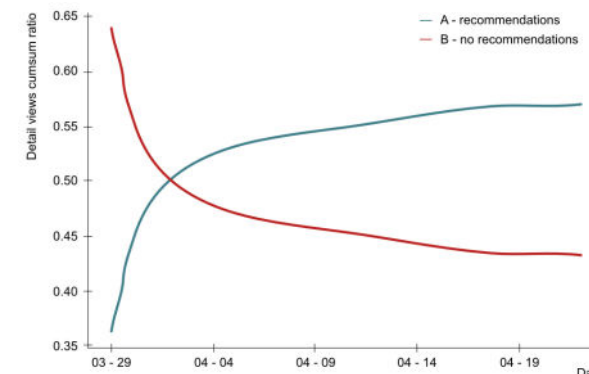
Explainability

The explainability of the recommendations was performed by listing all matched courses and keywords.

```
'Junior Data Scientist'  
# Matching keywords  
'web', 'program analysis', 'parsing', 'code',  
'classification', 'big data', 'databases', ...  
# Matching courses with their keywords  
'bi-big': {'big data', 'databases'},  
'bi-xml': {'programming', 'databases', 'web'},  
...
```

Evaluation

- Two successful A/B tests.
- Positive impact while using interaction-based RS.
- More than 150 engaged students.
- Further improvement with student profiles.



Contributions

- Article submitted to a research journal.
- RS with re-ranking module for student opportunities.
- Functional platform working on the website of the university.

Conclusions

The thesis bring original research contributions as well as real-world results running currently at the university website.

Thanks to the re-ranking module, designed system meets European GDPR norms and works for both subscribed and unsubscribed users simultaneously. We plan to scale the system into a global product.