Job Recommendation based on Job Seeker Skills: An Empirical Study

Jorge Valverde-Rebaza and **Ricardo Puma** and
Paul Bustios and **Nathalia C. Silva**

Department of Scientific Research

Visibilia

São Paulo, Brazil
2018

Visibilia

FAPESP

Outline

- Motivation and Problem
- Proposal
- Experimental Setup and Results
- Conclusion

Motivation and Problem

Job recommendation sites

Job recommendation task

Job recommendation task

Problem

Improve the recommendation systems to offer job vacancies that fit properly to the job seekers profiles.

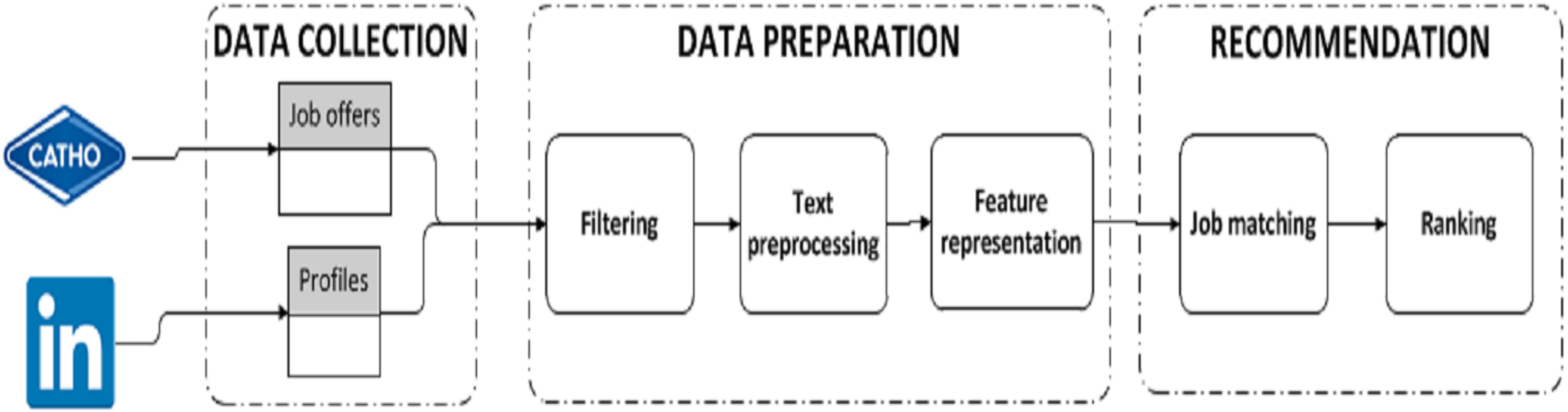
Proposal

Person-job fit premise

- The best **person-job fit** is possible when the personal skills of a job seeker match with the requirements of a job offer

Our Objective

Develop a framework for job recommendation based on **professional skills** of job seekers.



Our Recommendation Framework

DATA COLLECTION

- Job offers collected from Catho website and user profiles/curriculums from LinkedIn
- Job offers related to IT industry
- User profiles of professionals from Brazil
- Datasets collected are publicly available here: http://visibilia.net.br/text2story-job-recomendation/

DATA PREPARATION

- **Filtering** -> data in Portuguese, but parts of some curriculums and/or some expressions are in English
- **Text processing** -> disregard the job offers that are little/nothing related to the IT industry
- **Feature representation** -> representing curriculums and job offers using:
  - *Word Embeddings*
    - Word2Vec: CBOW, Skip-Gram, CBOW + n-grams, Skip-Gram + n-gram
  - *TF-IDF*

Table 1. Word embeddings description

Dataset	# Documents	# Tokens
Profiles	50	111970
Job offers	3877	157576

Description of the corpora used for our embeddings

Table 2. Professional profiles breakdown

Subfield	Profiles
Architect	5
BI consultant	10
Developer	24
Manager	2
Technical Support	9

Distribution of IT subfields within our sample of 50 profiles

RECOMMENDATION

- Given the profile of a job seeker, we search the **job matching** based on:
  - For TF-IDF -> cosine distance
  - For Word Embeddings -> Word Mover's Distance (WMD)

- We offer to the job seeker the **ranking** with the top k-nearest job offers

Experimental setup and results

Experimental setup

- We generate 10 job offer recommendations for 50 diferente profiles
- A group of 5 RH professionals evaluated, manually, these recommendations and allocate a score for them (from 1 to 10)
  - The more accurate the recommendation the greater the RH score
- Based on the scores provided by RH professionals, we calculated the **average score** (accuracy), **precision** and **minimum effectiveness** (ME)

Table 3. Results of job offers recommendation for each technique used.

	Score	**Precision**	**ME**
TF-IDF	0.588	0.775	**0.96**
Word2Vec-CBOW	0.548	0.765	0.92
Word2Vec-SkipGram	**0.590**	**0.814**	**0.96**
Word2Vec-ngrams-SkipGram	0.582	0.784	0.92
Word2vec-ngrams-CBOW	0.580	0.783	**0.96**

Results

Conclusion

Conclusion

- We developed a simple framework for job recommendation process which make possible the use of a variety of text processing and recommendation methods according to the preferences of system designers

- We make publicly available two datasets related to curriculums of job seekers and job offers

References

- Shaha T Al-Otaibi and Mourad Ykhlef. “A survey of job recommender systems". In: International Journal of the Physical Sciences 7.29 (2012), pp. 5127-5142
- N Deniz, A Noyan, and O G Ertosun. “Linking Person-job Fit to Job Stress: The Mediating Effect of Perceived Person-organization Fit". In: Procedia - Social and Behavioral Sciences 207 (2015), pp. 369-376
- M Kusner et al. “From word embeddings to document distances". In: Proc. of the 32nd Int. Conf. on Machine Learning, ICML'15. 2015, pp. 957-966.
- T Mikolov et al. “Efficient estimation of word representations in vector space". In: arXiv preprint arXiv:1301.3781 (2013).

Jorge Valverde-Rebaza and **Ricardo Puma** and **Paul Bustios** and **Nathalia C. Silva**

Thank you

Visibilia

FAPESP