



ONTOLOGY-BASED ANALYSIS OF JOB OFFERS FOR DOCTORS IN POLAND

Pawel Lula¹, Marcela Zembura²

¹ Cracow University of Economics, 27 Rakowicka Street, 32-510 Kraków, Poland
e-mail: pawel.lula@uek.krakow.pl

² Medical University of Silesia, 15 Poniatowskiego Street, 40-055 Katowice, Poland
e-mail: marcela.zembura@gmail.com

Abstract:

In the paper the results of the analysis of the labor market for medical specialists in Poland was presented. The research was focused on the demand side of the market and included the analysis of competencies required for doctors. The data describing required competences were retrieved from job offers published on the Internet. The detailed information about requirements were identified with the use of ontology-based tools. Next, graph models were built to show the importance of individual competencies and sets of them. Finally, the analysis of relationships between required competencies and other attributes of the labor market was conducted. All software tools used during the research were prepared in the *R* language.

Keywords: exploratory analysis of text documents, ontology-based analysis of text documents, labor market models, network analysis, bipartite graph models

1. Introduction

The analysis of the labor market and prediction of its behavior is a crucial aspect of national, regional, or sector economy. Labor market models should take into account the demand and supply side of the market, the flow of time, and the behavior of market participants. Models can be descriptive in nature or support decision-making processes. The issue presented in the paper is a part of the medical labor market model which acquires information on the demand for medical doctors from web portals and realizes the analysis of the acquired data focused on the required competencies and geographical location.

2. Model on the required competencies for doctors

The list of main components of the model includes:

- medical specializations,
- description of required competencies,
- geographical localization.

Information about the desired specializations will be obtained using a web scraping technique from portals that publish job offers. The extracted texts will be analyzed using the ontology proposed by the authors. The designed and implemented module is able to analyze offers prepared in Polish and English. Due to the current situation on the Polish labor market, we plan on expanding the module with the ability to analyze texts in Ukrainian. The results of the ontology-based analysis will be compared to the results obtained using the Dirichlet latent allocation method.

The analysis of medical competencies will include an assessment of their importance and an examination of the links between them. For this purpose, competence diagrams will be used which are based on weighted graphs in which:

- competencies are represented by nodes, and links between competencies by edges,
- weights assigned to nodes determine the importance of competences,
- weights assigned to edges describe the importance of links between competences.

The network of links between competencies constructed in this way will allow us to separate groups of competencies which are strongly linked to each other. This task will be performed with the help of methods used to identify homogeneous parts of the graph (communities).

The analysis of the relationship between medical specialties and desired competencies will be conducted using an exploratory approach relying on bipartite graph-based models used in ecology. The usefulness of the proposed approach will be assessed using a comparison with classical statistical models of interdependence based on chi-square statistics.

An analysis of the relationship between physician specialties and the location of the medical institution seeking to hire a physician will be conducted in a similar manner. However, the analysis of relationships between competency performed for each region will allow us to conduct a comparative analysis between.

The system described in this presentation was entirely designed and implemented by the authors of this paper using:

- *Docker* container and *Beautiful Soap* library to perform the job offer acquisition operation (using a web scraping technique),
- *R* language and *tm* and *topicmodels* packages in order to implement analysis
- ontology implemented in *yaml* language describing medical specialties,
- *igraph* and *bipartite* packages available in *R* language for performing network analysis.

3. Conclusions

The results show that ontology-based approach in exploratory text analysis used together with graph-based methods allow us to obtain valuable information concerning the labor market in medical area in Poland.

Ontology-based methods were used to gather detailed information about the requirements defined for medical specialists. The analysis of the importance of competences and relationships between them was conducted with the use of competency schemas based on undirected graphs. Links between competencies and other features describing the labor market were studied with bipartite graphs.

References

- [1] Lula P., Oczkowska R., Wiśniewska S., Wójcik K., Attempt to Estimate the Competency Gap in the IT Sector, *International Entrepreneurship Review*, vol. 5, no. 3 (2019) , s. 95-112.
- [2] Tyrańska M., (ed.), *Wybrane zagadnienia zarządzania kompetencjami pracowników w organizacji*, Wydawnictwo C.H. Beck, Warszawa, 2021
- [3] Wojciechowski J., Pieńkosz K., *Grafy i sieci*, Wydawnictwo Naukowe PWN, Warszawa, 2020