Prof. dr hab. Andrzej Bąk

Department of Econometrics and Computer Science

Wroclaw University of Economics and Business

Jelenia Góra, September 6, 2023

Review of the doctoral dissertation

of Dwijendra Nath Dwivedi, M. Phil. in Development Research

Title: *Analysis of the employer's expectations towards competencies of candidates for employment in the Banking Sector in The UAE using Text Mining Approach*

Scientific supervisor: prof. dr hab. Paweł Lula

Department of Computational Systems, Cracow University of Economics

Formal grounds for the review: resolution of the Discipline Council of Management and Quality Sciences of the Cracow University of Economics of July 13, 2023.

Purpose of doctoral dissertation review: assessment of compliance with the requirements specified in the Act "Prawo o szkolnictwie wyższym i nauce z dnia 20.07.2018 (Dz. U. z 2018 r., poz. 1668 z późn. zm.)" act, and position on allowing Dwijendra Nath Dwivedi, M. Phil. to the public defense of a doctoral thesis.

I.   Characteristics of the doctoral dissertation

The doctoral proceedings of Dwijendra Nath Dwivedi, M. Phil. were initiated on June 29, 2021, in the discipline of management and quality sciences. Dwijendra Nath Dwivedi, M. Phil. wrote his doctoral dissertation entitled *Analysis of the employer's expectations towards competencies of candidates for employment in the Banking Sector in The UAE using Text Mining Approach*

Wroclaw University of Economics and Business
Komandorska 118/120 Street, 53-345 Wrocław Poland
tel.: +48 71 36 80 100
e-mail: kontakt@ue.wroc.pl
www.ue.wroc.pl/en/

at scientific supervision of prof. dr hab. Paweł Lula at the Doctoral School of the Cracow University of Economics.

The doctoral dissertation has 198 pages and consists of 7 chapters, an introduction, a list of references, acknowledgments and abstract.

The structure of the dissertation is very detailed (reaches to the five-digit level – e.g., 5.2.3.2.1 on p. 119) and, in addition to the text, contains tables, drawings and mathematical formulas.

The introduction identified 3 main research gaps justifying the research, the results of which were presented in the doctoral dissertation. The main goal of the research, 5 intermediate goals and 3 hypotheses were also formulated.

Chapter 1 contains an analysis of the labor market, including the characteristics of the banking sector (public and private), in the United Arab Emirates.

Chapter 2 presents definitions of competencies expected by employers and competency research in various industries, with particular emphasis on the banking sector in the UAE.

Chapter 3 concerns methodological aspects of text exploration (text mining), such as the classification of problems solved using text mining and selected methods and models used in text mining. The representation of text documents using a frequency matrix and statistical methods for describing the content of documents using phrases and key words were discussed in more detail, such as: LSI (Latent Semantic Indexing) with SVD (Singular Value Decomposition) of the frequency matrix and a model built based on the Bayesian LDA network. (Latent Dirichlet Allocation model) for latent data.

Chapter 4 presents the results of an empirical analysis of the demand for competences in the banking sector in the United Arab Emirates, which was carried out using text mining methods (mainly using the LDA model discussed in the previous chapter). Text data for analysis was collected from selected portals posting job offers using web scraping scripts (internet data extraction). Visualizations of the analysis results are provided.

Chapter 5 proposes a network model of selected aspects of the labor market. The concept of a competency model expected by employers in the banking sector in the United Arab Emirates was presented. The competency diagram is presented using a weighted graph, where nodes represent competencies and edges represent relationships between two competencies. Weights are assigned to nodes (representing the importance of competencies) and edges (representing the importance of relationships between competencies).

Chapter 6 presents a model of employers' expectations towards candidates for work in the banking sector in the United Arab Emirates. This is the application of a competency model built

on a weighted graph (described in the previous chapter) to the analysis of text data collected based on job offers in the banking sector of the United Arab Emirates. Text data was collected based on an ontological approach - phrases and keywords that matched patterns defined within the ontology (described in the previous chapter) were selected.

Chapter 7 presents conclusions and recommendations (for practitioners and policymakers) and indicates directions for further research in competence modeling using text mining tools and proposed procedures.

II.   Subject and scope of research, scientific problem

The scope of research outlined in the doctoral dissertation covers competencies in the banking sector in the United Arab Emirates (UAE). In particular, the subject of research is the competence gap between employers' expectations and the qualifications of potential employees.

The research problem, formulated based on the analysis of the literature on the subject and observations of the labor market, consists in developing tools (models, software) enabling the identification of the competence gap in the UAE banking sector. To solve the research problem, achieve the set goals and verify hypotheses, the PhD Student used text mining methods and models. The justification for this is understandable because the description of competencies (both on the side of employers' expectations and on the side of job candidates) is represented by text.

The empirical research used selected text mining methods based on the corpus (probabilistic models) and methods referring to domain knowledge (ontological models). The text data needed to conduct the analysis was obtained from public workplaces in the banking sector and from job descriptions posted on websites with job offers in the UAE (Gulf Talent, Bayt, Naukri). The data collection technique used Internet data extraction technique (web scraping).

The research procedure includes stages leading to achieving the goals of the dissertation:
– characteristics of the UAE economy, in particular analysis of the banking sector,
– identification of key competencies expected in the UAE banking sector,
– developing a methodology for analyzing the competencies expected by employers,
– building a network model based on bipartite graphs that allows to describe the relationships between competences and various elements of the labor market,
– development of software tools to analyze the supply and demand side of the labor market.

III.   Evaluation of the doctoral dissertation

Dwijendra Nath Dwivedi's doctoral dissertation concerns the current and important issue of competences in the labor market, both in the social and economic dimensions. Competencies are perceived and assessed from two perspectives: the employer's and the employees. Competencies understood as knowledge and skills are constantly changing under the influence of many factors - nowadays, especially due to the unprecedented pace of development of information technologies, information overload and the digitization of the reality in which humans' function. In this context, employers' expectations regarding the competences of potential employees are changing. Job candidates, however, must quickly adapt their competences to these expectations.

The PhD Student undertook the task of developing tools (model and software) that enable, in a relatively short time, the comparison of competencies perceived from both perspectives and the assessment of discrepancies (competence gap).

I evaluate the choice of the research subject and the formulation of the research problem positively.

To solve the research problem, text mining methods and models were used, which fall more broadly into the area of data analysis (data science, data mining, big data, data engineering, machine learning, artificial intelligence). The PhD Student chose text mining methods and models that are well documented in the literature, available in computer programs and used in empirical research.

I evaluate the choice of text mining methods for analyzing the competence gap represented in the form of text documents positively.

The first part of the empirical research presents the results of the analysis of the demand for competences of the banking sector in the United Arab Emirates. Topic modeling was used using a probabilistic model based on the Bayesian network LDA (Latent Dirichlet Allocation). The results of the analysis indicate a greater importance of front-office competencies (client, sales, customer), and a lower importance for back-office and middle-office staff (report, product, company, financial). This is even more visible when the data is divided into pre- and post-covid periods. The results of this part of the analysis were presented in the form of clear visualizations. They also indicate the expected changes in the competences of potential employees caused by the development of information technologies and forms of office work.

The second part of the empirical research presents the results of competence analysis using network models built on weighted graphs. Text data was collected based on an ontological

approach. The PhD Student identified 84 specific competencies for the banking sector in the UAE. On this basis, a hierarchical model of competences was created. Detailed competences have been aggregated into five higher-level categories (academic, industry, personal, technical, workplace). This part of the research provides information about the importance of individual competencies and the relationships between them. The results of the dependency analysis were illustrated using competency co-occurrence graphs. In assessing the importance of specific competencies, the Gini index was also used, indicating the uneven distribution of the importance of individual competencies. This part of the research ends with an analysis of the distribution of competences in the three largest cities of the UAE. This part of the analysis shows that the most important are industry-wide competencies, followed by professional competencies (workplace competencies), personal competencies (personal effectiveness), technical competencies and academic competencies.

I evaluate the method of conducting empirical research as well as the obtained and interpreted results in both parts positively. At the same time, I believe that the omission of references to own calculation scripts and computer programs that enabled the analysis of text data, and the presentation of results negatively affects the perception of the work as a whole.

IV.   Comments on the content of the doctoral dissertation

1.   The main goal of the dissertation is not precisely defined. The abstract on page 3 reads: "The main objective for the thesis is to perform the analysis of competency gap in the UAE banking sector." In the introduction on page 10, the main objective reads slightly differently: "The main objective of this dissertation is to develop a software to perform the competency gap assessment in the UAE market mainly focusing on the banking sector".

2.   Have the hypotheses formulated on page 10 been formally verified and confirmed? Did the PhD Student assume that these hypotheses could be falsified? If not, these are rather theses.

3.   The 3 research gaps announced in the introduction are not specified.

4.   It is not possible to clearly determine whether the algorithms included in chapter 3 are original, modified by the PhD Student, or taken from literature sources.

5.   There is no presentation of own computer scripts (programs), which were certainly used during the analysis of text documents. From the content of the thesis, it can be learned

(although it is not easy) that the PhD Student used R and Python scripts and program packages. There are no references to this programming, nor a list of packages used and useful in text mining. There is also no information about other computer software that the PhD Student used in his research. References to R and Python programs should be included in the reference list.

6. Other formal and typographical comments:
    - the structure of the dissertation is excessively detailed – some subsections consist of single sentences (e.g., 3.6.7.4, 3.6.7.5, 3.6.7.6, 4.2.2.3, 4.2.2.4, 5.2.1.5, 5.2.1.6),
    - lack of a summary in Polish, which is required (Article 187, section 4 of the Law on Higher Education and Science),
    - the titles of chapter 7 and point 7.2 are the same,
    - no list of tables and list of drawings,
    - mathematical formulas are not numbered,
    - numbering of tables and figures within chapters,
    - some tables and figures are difficult to read (e.g., on p. 141),
    - table 4 on pp. 121-122 has the wrong title,
    - the references section should not be numbered,
    - the title of the section "List of bibliography" is not commonly used in scientific works,
    - references is not composed in accordance with any known standards, e.g. available in the LaTeX system,
    - ambiguous reference to the item in references (Lula et al. 2019),
    - twice included in references (Lula, P., Oczkowska, R., Wiśniewska, S., & Wójcik, K. 2018),
    - the reference on p. 113 is missing from the references (Lula et al. 2017).

Discussion questions:

1. Has own software been developed to enable analysis of the competence gap in the UAE banking sector (main objective in the second version)?
2. What computer software was used in the work, what R and Python program packages?
3. Are the algorithms in Chapter 3 original, modified, or taken from sources?

4. Is it possible to measure the competency gap between employers' expectations and the current competencies of job seekers in the banking sector?

5. Why is the importance of academic competencies relatively small on the labor market under study?

V. Conclusion

Dwijendra Nath Dwivedi's doctoral dissertation concerns the issue of competences perceived from two perspectives – the employer's and the employees. This is a key issue in the modern, dynamically changing labor market and very important for the development of management and quality sciences. The work is correctly constructed, although it also contains the shortcomings indicated above.

The main achievement of Dwijendra Nath Dwivedi, M. Phil. presented in his doctoral dissertation is the development of a procedure enabling the assessment of the competence gap in the banking sector on the UAE market using text mining methods and models. This is an original proposal for solving a scientific problem with significant application values. An important result of the work is the development of proprietary computer software enabling the practical implementation of such a task. These achievements demonstrate the PhD Student's well-established IT and linguistic knowledge, as well as the ability to independently conduct scientific work.

In my opinion, the doctoral dissertation Dwijendra Nath Dwivedi, M. Phil. entitled *Analysis of the employer's expectations towards competencies of candidates for employment in the Banking Sector in The UAE using Text Mining Approach*, submitted for review, meets the requirements set out in the Act "Prawo o szkolnictwie wyższym i nauce z dnia 20.07.2018 (Dz. U. z 2018 r., poz. 1668 z późn. zm.)", and I recommend the PhD Student to be admitted to the public defense.

Andrzej Bąk