The Effects of Artificial Intelligence in the Process of Recruiting Candidates

Lasha Abuladze[1](✉) and Lumbardha Hasimi[2]

[1] University of Economics in Bratislava, Petržalka, Slovakia
Lasha.abuladze@euba.sk
[2] Comenius University in Bratislava, Bratislava, Slovakia
Lumbardha.hasimi@uniba.sk

Abstract. AI-based solutions have found a great application in filling the gap and enhancing the massive recruiting processes. With the recent developments, the role of gamification in the overall managerial processes, especially in recruiting has proven to be crucial. AI as a powerful tool towards the challenges the hiring process faces, appears as contradictory in a number of issues. In this paper, we have observed and analyzed the advantages and disadvantages of AI in recruiting, followed by a proposed model for resume screening based on keywords and phrases against job description. Furthermore, a case has been presented and assessed regarding results and implications of AI-based tools, namely machine learning models in a simple scenario of hiring process.

1 Introduction

Human resource management (HRM) is constantly looking for innovative technologies to increase efficiency and effectivity in recruitment and selection. Recent technological advancements have enabled it to move from a traditional style of operation to a modern, digitized one in some key areas of their recruitment process [1]. The use of AI technologies is currently described and understood by different researchers and authors as an emerging trend in recruitment and selection activities, which has brought new opportunities and challenges for the enterprise instances. Using artificial intelligence represents one of the biggest evolutions in HRM. With the help of AI, specifically the Machine Learning as a subset of AI, HRM has gained opportunities to automate some basic and important areas of recruitment and selection that were previously done manually by human [2–4].

1.1 Theoretical Background

HRM is constantly looking for innovative technologies to increase efficiency and effectivity in recruitment and selection. Recent technological advancements have enabled HRM to move from a traditional style of operation to a modern, digitized one in some key areas of their recruitment process [3]. The process of finding and recruiting candidates electronically has brought new aspects into the recruiting scene. Most popular methods of e-recruitment are:

M. Papadaki et al. (Eds.): EMCIS 2022, LNBIP 464, pp. 465–473, 2023.
https://doi.org/10.1007/978-3-031-30694-5_34

– Social media sourcing. Social platforms like Facebook, LinkedIn, Glassdoor, Instagram are widely used by companies to attract, search and recruit candidates online. It is also a great way and an effective tool to spread a positive name among candidates and build the brand status of the employer.
– Employer's website. It provides information about new job opportunities in the company, allows applicants to apply for it and collects their data.
– Job boards. Websites used by HRM to advertise and promote job offers while job seekers can use it to find new vacancies and job opportunities.
– Online interviews. Communication technologies (Skype, Teams, Zoom, G-meet) allow HRM to interview candidates online without using any physical space.
– AI powered software. Software, that use humankind intelligence to find, screen candidates and simplify the entire process of recruitment for the HRM.

In today's recruitment use of artificial intelligence became one of the biggest evolution and trend for HRM. With the help of AI, HRM has gained opportunities to automate some basic and key areas of recruitment and selection that were previously done manually by humans. According to [5] approximately 70% of companies will adopt some form of AI by 2030, and countries that can establish themselves as AI leaders could reap up to 20 to 25% more economic benefits than current levels.

1.1.1 Artificial Intelligence - Its Use in the Recruitment Process

AI as a sub field of computer science is focused on training computers to perform initially designed human tasks. While seeking to imitate and therefore improve the human intelligence, it compares natural human intelligence to artificial intelligence. As a model of science, it aims and targets alleviating the human physical and mental labour using computational intelligence behavioral models with the goal of developing reasoning, learning, decision making, and performing complex issues that can be executed by human brain [6, p.]. We have witnessed in the recent years the expansion in the unprecedented impact of AI in recruitment processes. Enormous software in today's recruitment markets find AI- based solutions, has a great hand to helping employers mitigate the overall recruitment process. The AI systems up to date are available across wide range functions. Specifically speaking, AI has found a great application in: Candidate Sourcing, Engagement Candidate Tracking, CV Screening, Pre-Employment Assessments, Video Interviewing etc. Recruitment AI encompasses a wide array of technologies functioning at different points in the recruitment process [7].

AI-based solutions help employers scan a considerable number of apps for the best possible candidates. In fact, this is one of the most widely used forms of AI recruitment solutions today [7]. The use of AI in the recruiting, has completely redefined the relationship between employer and the applicant, not only in the process of selection, but in the overall employer experience. AI tools like Chatbot provide applicants with new features and enhanced employer practices. The entire assessment process, such as interview scheduling, customized candidate profiles, customized offers, reference checking are all tasks easily handled by AI-imbued applications. Despite the wide use of AI, only 10% of companies currently use it in a high context, whereas 36% of organizations expect to make full use of AI in the future [8]. Some of the commonly known AI applications adopted by big companies include:

– ATS and CRM Systems Applicant Tracking Systems (ATS) – platforms that offer the recruiters the chance to track every stage of hiring process, from the initial contact to the end of hiring
– Candidate Relationship Management (CRM)
– CV/Resume Screeners CV screening
– Conversational Agents Recruitment conversational agents, or chatbots
– Pre-Employment Assessments
– AI powered interviewing

In the case of resume screeners, AI powered recruitment innovation addresses the issue of high-volume data processing. Trained models screen and detect the characteristics in the content given, through keywords and main expressions that are crucial for the criteria. Whereas for the chatbots, the models are designed to mimic human conversational skills, using NLP technique to analyze and reply effectively. AI-powered interviewing analyze facial expressions, evaluating on basis of the tone, language, non-verbal measures, while applicating standardized process [9].

AI technologies used in recruitment are seen as highly capable and objective tools. However, the practical use and benefit of AI to support recruitment are opposing seen in different perspectives [7]. Among many concerns regarding the benefits of AI in the recruitment processes researchers have demonstrated records of inaccurate results as well as issues with the disability problem. In the case of the assessment of the disability fairness, the issue has gathered quite little attention. The impact of AI aside, the structural issues affecting people with disabilities in gaining and maintaining employment is a complex and ongoing concern. Therefore layering a complex system of automated assessment of candidates risk complicating the situation and expanding the risk of harm [10]. Through implementation of AI in the business strategies the recruiters can identify candidates and effortlessly get data on persona. Most of the repetitive tasks disappear because of the efficiency and use of a living soul the recruiters free to deal with more strategic issues.

Although the AI is designed to overcome the bias during the selection process, the technique itself can still suffer the bias inherited from their programming and data sources. How much this affects the AI systems of recruitment, it is a prone to many questions and complexities levels within the primary functions of systems. In the case of primary sources of bias such as name, age, gender etc., such data unbiasedly pass through the support of AI [8, 11, 12].

Despite the contradictive attitudes towards benefits of AI, many studies argue that AI recruiting does not inherently conflict with human rights [13]. Although on the issue of conflicting with ethical principles, highly depends on the conditions under AI tools are used and trained. However, it is of a crucial importance to derive possible implications and responsibilities in terms of human rights standards in context of AI recruiting [8]. This alongside with bias issues of AI methods, represents a new gap that requires further research and standardization.

Majority of companies which adopt AI-enabled recruiting are unfamiliar with the implementation process. It is especially important to identify certain categories of talent pool, before applying the AI-based tools. In massive applications scenarios, AI tools are applied as cost-effective solutions [14].

Alongside cost-efficiency, AI in the recruitment helps on evaluation process, processing of data, ranking and qualification screening, reduction of administrative and routine tasks, communication, the speed of the overall process carrying etc. Most of all, at the bias context AI allows the equal opportunity to all candidates, leaving no room for human tendency. As an output, AI tools impact business competitiveness [6], while help to gain better insight into the talent and recruitment process.

2 Proposed Methodology

In the theoretical part of research, we defined e-recruitment, artificial intelligence with their characters based on articles, research papers from different authors and researchers. With help of deduction and analysis methods, we screened, filtered existing knowledge, and explained mutual relations between e-recruitment and artificial intelligence. By using methods of logic and induction, we compared and synthesized existing knowledge and have found limited or no information, knowledge on the implications of AI-based solutions in recruiting processes [15].

While analyzing the current state-of-art in the field, we have summarized research gaps and current issues, trying to answer the research questions of this study:

1. What are the implications of AI-based solutions in recruiting processes?
2. At what extent, the machine learning models can affect the performance and result of the screening process?

Hence, the main goal of the research is to answer the research questions stated, while investigating and exploring the effects of AI-based solutions implemented in recruiting processes. Namely, the importance of time-consuming and number-efficient processing of resume datasets.

In the practical part of the research, to elaborate further what has been discussed above, we have implemented a simple model for *resume* screening using machine learning, based on text classification and keyword similarities found within resumes or candidate profiles. Another element that would enhance the screening process, is the readability of the resume. Hence, a proposed readability score as a metrics of screening and classification would add an asset. The dataset that is used in this paper is "*Resume dataset*" available at Kaggle.com [16] containing over 2000 resumes from various fields and categories (Table 1).

Table 1. Dataset used to train the model.

Label	Description
ID	Unique identifier and file name for the respective pdf.
Resume_html	Contains the resume data in html format as present while web scrapping.
Resume_str	Contains the resume text only in string format.
Category	Category of the job the resume was used to apply.

The dataset is separated into training and test sets by a 75/25 split, respectively. After importing relevant modules, and loading the dataset, a pre-processing of the data was conducted, using stemmer function from nltk library, namely, Porter Stemmer [17] Algorithm as a text normalizing algorithm.

Table 2. The pre-processing of data

Function	Description
Stop words	Tokenizing the input words into individual tokens and stored it in an array. StopWords [8]
Tokenization	Converting the corpus to a vector of token counts. Count Vectorizer (sklearn)
Lemmatization	Transform the corpus of text into a list of words and assign words to lemmas.

Considering the importance of the dataset and the pre-processing phase, throughout the process, the data had undergone the cleaning stage, tokenization, stemming and lemmatization, as seen in the Table 2. To ignore the case of letters, all words were converted to lower case, whereas to avoid words without high semantic load, we used a special blacklist of words from the natural language toolkit (*nltk*) library. To avoid overload of the dictionary with equal words in various forms, and increase the overall accuracy, the stemmer-function from the '*nltk'* library was found to be highly effective.

3 Implementation and Results

The proposed solution uses various techniques with the aim of achieving automated screening of candidate's resume that mainly focuses on the content of the resume, namely focusing on the keywords used throughout the content. For this reason, the feature extraction stage covers keyword extraction followed by the next stage which includes the similarity computation. In the extraction stage we perform the extraction of keywords, phrases, and related parameters to match candidates with the job description of the company.

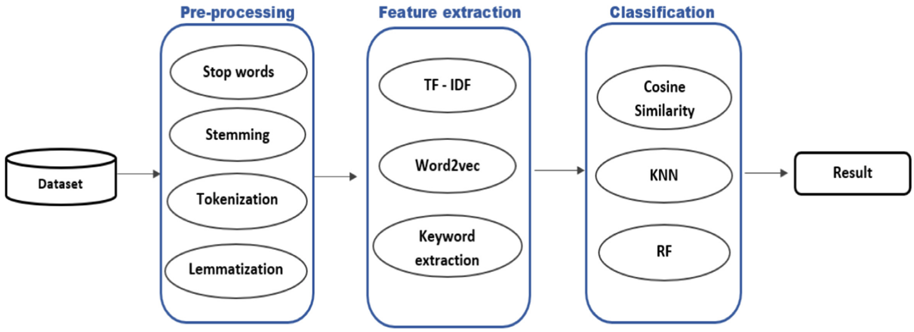

Fig. 1. The architecture and activity overflow of the implemented model

As seen in the Fig. 1, the architecture of the proposed model – the basic concept consists of three phases, namely the pre-processing, data extraction and final classification.

Tf-Idf. In the second phase, Tf-idf, was used to calculate the terms. As a technique it ranks each word in accordance with the number of times it appears within a document [18–20]. This term is then normalized by dividing by the number of words in the entire group. Considering that words with shorter segment tend to hold more weight. Furthermore, the inverse document frequency ranks the words based on the individuality against the specific segment. Meaning that words are separated towards their use in a section of the text, from the global used words (Fig. 2).

```
wordfrequencydist = nltk.FreqDist(total_Words)
mostCommon = wordfrequencydist.most_common(50)
print(mostCommon)
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from scipy.sparse import hstack
required_Text = resumeData['structured_resume'].values
required_Target = resumeData['Category'].values
word_vectorizer = TfidfVectorizer(sublinear_tf=True,stop_words='english',max_featu
res=1500)
word_vectorizer.fit(required_Text)
WordFeatures = word_vectorizer.transform(required_Text)
```

Fig. 2. Feature extraction and vectorizing

Word2Vec. Considering that the relationship between the words is important for our model, the Word2vec model was an effective solution. Word2vec is the most common technique as a word embedding model that processes text. Its effectiveness lies on the ability to group together vectors of similar words [21]. A word2vec model can be generated with two different approaches as a Skip-gram and Common Bag of Words. The skip-gram model takes a certain word and tries to predict its surrounding word or context words, whereas in the case of Common Bag of Words, the context of the words is taken and predicted to the target word (Fig. 3).

```
wordscores = Calculate_Word_Scores(phraseList)
keywordresume = Generate_resume_Keyword_Scores(phraseList, wordscores)
sortedKeywords = sorted(keywordresume.items(), key=operator.itemgetter(1), reverse=True)
totalKeywords = len(sortedKeywords)
rake = Rake("db\\Stop_Word_List.txt")
keywords = rake.run(text)
print (keywords)
```

Fig. 3. Keyword generating and printing after extraction phase

K-Nearest Neighbors Algorithm (k-NN). Nearest neighbors is a supervised learning model which is used for classification and regression analysis [20]. KNN does not have parameters we can change or optimize in order to achieve better performance. This model is used to identify the resumes that are nearest matching the given job description. First, to have similar parameter the open-source library "gensim" was taken into account, generating so the summary of the provided text in the provided word limit.

Final output on running similarity on words and phrases, with the percentage score reached against parameters defined (Fig. 4).

```
# import cosine similarity
from sklearn.metrics.pairwise import cosine_similarity

#similarity score
matchpercentage = cosine_similarity(count_matrix)[0][1]
matchpercentage = round(matchpercentage*100,2)
print('Resume matches {} % score to job description! ',cosine_similarity(count_matrix))
```

```
Resume matches 83.26 % score to job description!
```

Fig. 4. Cosine similarity output

Considering that the core functionality of the model is to find out the the level at which a resume fits within a job description, the final technique engaged in this case for the similarity measure is the cosine similarity. Cosine similarity is applied to calculate the similarity between vectors and measure a cosine of angle between the corresponding vectors. In our case the count matrix checked containing 0 and 1 values, was implemented to check on the similarity and have the final output in the percentage as the final result on the candidates resume profile.

4 Conclusion

In the theoretical part of the research, we used the methods of deduction and analysis to screen, filter existing knowledge and then defined and explained e-recruitment, artificial intelligence with their mutual relationships, characteristics, elements, principles, and potential benefits. To find a research gap, we used logical and inductive methods, through which we compared, synthesized existing information, knowledge, and found the research gap on the implications of AI-based solutions in recruiting activities.

Based on the AI powered recruitment process, we found that using of AI powered software in recruitment process can reduce recruiters' time to screen and shortlist applicants by seeking special keywords in their resumes and comparing, matching specific job requirements with applicants' skills, knowledge, and experience. It can improve quality of hiring process by quickly and efficiently finding the most suitable candidates, while reducing the human bias, preventing similarity attraction effect confirmation bias, halo effect, demographic discrimination, and other factors. However, despite the minimizations on the human bias, the machine learning models can be only as good as human, in terms of the model training and parameter setting. This means that human bias can be inherited indirectly to the AI-based model itself, what leaves huge room for future scientific and research endeavors.

AI-based models can positively impact the overall performance and outcome of the recruitment process, but then again there are limitations and challenges to consider before using it in the recruitment process. Programming limitations, ability to inherit human bias, certain dependencies, lack of human judgement, lack of know-how are all elements that need to be considered and backed up. By answering research questions, we succeeded to fulfil main goal of our research, in investigating and exploring the effects of AI-based solutions implemented in recruiting processes. Namely, the importance of time-consuming and volume-efficient processing of resume datasets.

In addition to all the provided information and knowledge, our research has several barriers and limitations in the application of its results. In our case we focused on the profile of the candidate, taking into account only the skill attribute and the keywords extracted from the resume. Future work could consider more thoroughly the other attributes of the candidates' profile, such as the sentiment score of the overall content using both sentiment analysis and recommendation systems concept.

Acknowledgement. This research was supported by Slovak Academy of Sciences VEGA project No. 1/0623/22 Virtualization in people management – employee life cycle in businesses in the era of digital transformation.

References

1. Kirovska, Z., Josimovski, S., Kiselicki, M.: Modern trends of recruitment - introducing the concept of gamification. J. Sustain. Dev. **10**, 55–65 (2020)
2. Mantello, P., Ho, M.-T., Nguyen, M.-H., Vuong, Q.-H.: Bosses without a heart: socio-demographic and cross-cultural determinants of attitude toward Emotional AI in the workplace, AI & Soc, November 2021. https://doi.org/10.1007/s00146-021-01290-1
3. Pissarides, C., Bughin, J.: Embracing the New Age of Automation | by Christopher Pissarides & Jacques Bughin, Project Syndicate, 16 January 2018. https://www.project-syndicate.org/commentary/automation-jobs-policy-imperatives-by-christopher-pissarides-and-jacques-bughin-2018-01. Accessed 27 Oct 2022
4. Joy, M., Assistant, J.: An investigation into gamification as a tool for enhancing recruitment process. Ideal Res. **3** (2017)

5. Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A., Subramaniam, A.: Skill shift: automation and the future of the workforce | VOCEDplus, the international tertiary education and research database. https://www.voced.edu.au/content/ngv%3A79805. Accessed 27 Oct 2022
6. Al-Alawi, A.I., Naureen, M., AlAlawi, E.I., Naser Al-Hadad, A.A.: The role of artificial intelligence in recruitment process decision-making. In: 2021 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, pp. 197–203, December 2021. https://doi.org/10.1109/DASA53625.2021.9682320
7. FraiJ, J., László, V.: A literature review: artificial intelligence impact on the recruitment process. Int. J. Eng. Manag. Sci. **6**(1), Art. no. 1 (2021). https://doi.org/10.21791/IJEMS.2021.1.10
8. Hunkenschroer, A.L., Kriebitz, A.: Is AI recruiting (un)ethical? A human rights perspective on the use of AI for hiring, AI Ethics, July 2022. https://doi.org/10.1007/s43681-022-00166-4
9. Armstrong, M., Landers, R., Collmus, A.: Gamifying Recruitment, Selection, Training, and Performance Management: Game-Thinking in Human Resource Management, 2016, pp. 140–165 (2016). https://doi.org/10.4018/978-1-4666-8651-9.ch007
10. Nugent, S.E., et al.: Recruitment AI has a disability problem: questions employers should be asking to ensure fairness in recruitment, SocArXiv, preprint, July 2020. https://doi.org/10.31235/osf.io/emwn5
11. Obaid, I., Farooq, M.S., Abid, A.: Gamification for recruitment and job training: model, taxonomy, and challenges. IEEE Access **8**, 65164–65178 (2020). https://doi.org/10.1109/ACCESS.2020.2984178
12. Hunkenschroer, A.: How to improve fairness perceptions of AI in hiring: the crucial role of positioning and sensitization. AIEJ **2**(2) (2021). https://doi.org/10.47289/AIEJ20210716-3
13. Fritts, M., Cabrera, F.: AI recruitment algorithms and the dehumanization problem. Ethics Inf. Technol. **23**(4), 791–801 (2021). https://doi.org/10.1007/s10676-021-09615-w
14. Black, J.S., van Esch, P.: AI-enabled recruiting: what is it and how should a manager use it? Bus. Horiz. **63**(2), 215–226 (2020). https://doi.org/10.1016/j.bushor.2019.12.001
15. Korn, O., Brenner, F., Börsig, J., Lalli, F., Mattmüller, M., Müller, A.: Defining Recrutainment: a Model and a Survey on the Gamification of Recruiting and Human Resources (2017)
16. snehann bhawal, Resume Dataset. https://www.kaggle.com/datasets/snehaanbhawal/resume-dataset. Accessed 07 Oct 2022
17. Chandra, R.: Co-evolutionary Multi-task Learning for Modular Pattern Classification, presented at the Neural Information Processing (ICONIP 2017), PT VI, 2017, vol. 10639, pp. 692–701 (2017). https://doi.org/10.1007/978-3-319-70136-3_73
18. Keita, Z.: Text data representation with one-hot encoding, Tf-Idf, Count Vectors, Co-occurrence Vectors and…, Medium, 06 April 2021. https://towardsdatascience.com/text-data-representation-with-one-hot-encoding-tf-idf-count-vectors-co-occurrence-vectors-and-f1bccbd98bef. Accessed 12 July 2022
19. Paullada, A., Raji, I.D., Bender, E.M., Denton, E., Hanna, A.: Data and its (dis)contents: a survey of dataset development and use in machine learning research. Patterns **2**(11), 100336 (2021). https://doi.org/10.1016/j.patter.2021.100336
20. Wendland, A., Zenere, M., Niemann, J.: Introduction to text classification: impact of stemming and comparing TF-IDF and count vectorization as feature extraction technique. In: Yilmaz, M., Clarke, P., Messnarz, R., Reiner, M. (eds.) EuroSPI 2021. CCIS, vol. 1442, pp. 289–300. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85521-5_19
21. McCormick, C.: Word2Vec Tutorial - The Skip-Gram Model. http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/. Accessed 12 July 2022