IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 5, October 2025, pp. 4271~4278
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i5.pp4271-4278

AI-driven hyper-personalization and transfer learning for precision recruitment

Nour Alqudah, Qusai Q. Abuein, Mohammed Q. Shatnawi
Department of Computer Information Systems, Faculty of Computer and Information Technology,
Jordan University of Science and Technology, Irbid, Jordan

Article Info

Article history:

Received May 1, 2024
Revised Jun 17, 2025
Accepted Jul 10, 2025

Keywords:

Big five personality
Functional areas
Hyper-personalization
Soft skills
Transformer models

ABSTRACT

The research study demonstrates how artificial intelligence (AI)-powered models can transform the hiring process by maximizing the match between candidates and jobs, leading to better hiring options and increased worker productivity. Our research develops highly personalized AI-powered recruitment applications. By using hyper-personalization to tailor job recommendations based on job compatibility and big five personality traits, this study leverages AI to improve job matching. Unlike traditional recruitment models that depend only on complex skill matching, hyper-personalization combines soft skills and personality dimensions to achieve a more precise candidate-job alignment. Transformer-based models, including bidirectional encoder representations from transformers (BERT), RoBERTa, and cross-lingual language model (XLM)-RoBERTa, have shown exceptional performance in natural language processing (NLP) and classification tasks; thus, we apply them. Transfer learning helps us to fine-tune these models to improve the accuracy of personality classification. Compared to conventional models, experimental data achieves up to 80% accuracy in binary classification and 72% in multi-class classification. By demonstrating job-candidate compatibility, this study emphasizes the potential of AI-driven models to transform recruitment, leading to better hiring decisions and workforce productivity. Our outcomes play a crucial role in advancing hyper-personalized AI applications in talent.

Corresponding Author:

Qusai Q. Abuein
Department of Computer and Information Systems, Faculty of Computer and Information Technology
Jordan University of Science and Technology
P. O. Box 3030, Irbid 22110, Jordan
Email: qabuein@just.edu.jo

1. INTRODUCTION

The study aims to enhance the recruitment process by developing artificial intelligence (AI) models based on personal traits and soft skills rather than solely relying on technical qualifications and hard skills as in traditional recruitment. Hyper-personalization refers to personalizing job recommendations based on the candidate's personality traits and soft skills, such as communication, leadership, problem-solving, emotional intelligence, and teamwork [1]. Traditional hiring methods typically emphasize hiring candidates based solely on their technical expertise, but this does not find a proper fit between the applicant and the job or work environment. This focus can lead to employee adaptation issues, resulting in job turnover or poor performance. Today, with digital advancements and AI, there is a shift toward a more personalized hiring experience, which involves matching each person with a position that aligns with their personality and skills. Personalization depends on data companies collect from people's interactions and behaviors, helping them build relationships based on trust and understanding [2]. Moreover, new technologies enable us to analyze vast amounts of information and gain a more precise understanding of individuals.

For example, Gomez *et al.* [3] demonstrated that multilingual natural language processing (NLP) techniques can accurately extract employee opinions from data in different languages, aiding human resources (HR) management decisions. Additionally, Joshi *et al.* [4] showed that using multilingual NLP models increases the accuracy of employee engagement assessments and contributes to developing comprehensive HR strategies. Machine learning (ML) and deep learning (DL) have emerged as powerful tools to tackle many challenges we face in society, particularly in recruitment [5].

Employee and organizational harmony is essential for job satisfaction and retention [6]. As a result, employers seek candidates with the necessary abilities and a personality that complements the position and the company's culture. Personality tests, both online and in-person, have become essential tools in the hiring process [7]. According to Jain *et al.* [8], hyper-personalization gives customers a more individualized experience, which impacts their engagement. This study looks into digital clienteling for hyper-personalization in the fashion industry, with co-creation as a facilitator. Aside from standard sources like cover letters or curriculum vitaes (CVs), social networking sites like Facebook can provide information about a person's personality and predict employment success [9], [10].

Online personality forecasts accurately predict genuine personality traits [11]. These advances in our understanding of behavioral data and personality traits highlight the need for more sophisticated methods for assessing candidates. Recent research has focused on identifying and categorizing relevant skills from text data to improve hiring models, although traditional tools, such as personality tests, are still widely used.

Sayfullina *et al*. [12] explored several skill classification methods by annotating a dataset for hard and soft talents. Few studies have examined how hyper-personalization, which has been effectively used in industries like fashion and marketing, can be used in recruitment to enhance candidate-role alignment despite these developments in skill classification. Tamburri *et al*. [13] classified phrases containing skills by manually labeling job descriptions, whereas the research in [14], [15] approached skill extraction as a multi-label classification problem using bidirectional encoder representations from transformers (BERT). Establishing a relationship between personality characteristics, soft skills, and work fit is vital.

This study builds on our previous research [16], which explored the use of soft skills and big five personality traits to improve candidate-role matching, leveraging them to enhance the efficiency of innovative models to meet the requirements of matching the right employee to the appropriate functional area more accurately and effectively. The current study aims to present the results of applying a pre-trained model built using the same data from our previous study. Additionally, the study focuses on providing personalized recommendations to individuals within the functional area. It extends previous work by deepening our understanding of subtle personality traits and their applications in recruitment. We also present new results demonstrating how pre-trained models can effectively personalize recommendations across different scenarios. In this context, our study relies on modern transformer-based models, DL models designed to understand language and subtle meanings within texts [17]. We employed these models: BERT, robustly optimized BERT RoBERTa, and cross-lingual language model (XLM)-RoBERTa, which support advanced capabilities for precise personalization based on multiple contexts [18], typically based on personality traits.

This study contributes to workplace hyper-personalization by increasing knowledge of how AI, particularly transformer models, can improve model performance in choosing the best candidate for a particular functional area. It aims to match individuals with appropriate jobs based on their personalities and characteristics. Additionally, the study offers a thorough analysis and development of strategies for Transformer models to achieve the best possible candidate-to-job match across various career fields. Accordingly, it represents a significant scientific contribution to our understanding of how AI-powered models can speed up and increase the accuracy of talent acquisition processes. enhance talent acquisition processes directly impacts hiring quality and matching the right person to the right job. The approach introduces a unique method for connecting job seekers with suitable jobs by leveraging personality variables associated with soft skills. This approach supports the hiring process to increase creativity and diversity and match personalities to job requirements, regardless of experience. It aims to reduce favoritism, promote employee retention, and improve company efficiency. The study develops predictive language models with Transformer Models to make it easier to find ideal candidates based on their soft skills for ideal jobs [19].

2. METHOD

This section contains a detailed description of our study's datasets, models, and modeling methodologies. Our research aims to improve knowledge of the complex relationships between soft skills, personality characteristics, and organizational functional roles. Using transformer models, we hope to draw crucial insights and promote informed decision-making in work scenarios. This comprehensive method enables us to identify patterns and correlations that might be critical in matching individuals to the most appropriate employment opportunities based on their unique traits.

2.1. Data set descrabtion

Central to our research is the HYPskill dataset [16], a meticulously curated collection comprising 9,802 distinct sentences sourced from a variety of repositories. These include manually sampled resumes from diverse backgrounds and HR databases, alongside job descriptions extracted from platforms such as Glassdoor and Kaggle's Amazon jobs dataset. Crucially, the dataset is underpinned by the extensive HYPskill Lexicon, encompassing 2,190 categorized terms spanning domains like International Personality Item Pool (IPIP), job postings, resumes, interviews, and URLs.

Our dataset underwent meticulous annotation by three Ph.D. level experts in psychology, utilizing the rigorous IPIP-NEO-120 traits and traits study framework. This process involved binary classification to identify sentences containing soft skills, and multi-class classification to assign sentences to one of the five big five personality trait categories. Moreover, our study introduces a novel approach by mapping these personality traits to specific functional areas within organizations:

- Openness with production
- Conscientiousness with finance
- Extraversion with marketing/sales
- Agreeableness with HR
- Emotional stability with operations

This information sets the foundation of our research, providing insights into the dataset, annotation, and classification tasks conducted in our previous study [16]. These components were crucial in shaping the model design and the mapping of soft skills to personality traits. Building this groundwork allowed us to extend our previous findings into more advanced applications using transformer-based models.

2.2. Data preprocessing

In this part, we will describe some pre-processing procedures that were used in this study. Various pre-processing steps have been tested, and many approaches have been used. Table 1 displays instances of these processes. More examples of what pre-processing processes were employed are as follows.

Table 1. Examples of pre-processing steps

Technique	Examples
Removing punctuations	“!, +, :, ;, ?, @”
Removing identifiers	“the”, “a”, and “an”
Removing stopwords	“He”, “They”, “is”, and “on”
Expanding abbreviations	“I’m”, “can’t”, “into, “I am”, and “Can not”
Lemmatization	“been”, “had”, “into”, “be”, and “has/have”

2.2.1. The oversampling technique

In addressing our study's challenges, we encountered the issue of an imbalanced dataset [20], where certain classes had significantly fewer instances compared to others. This imbalance can severely impact classification model performance, often leading to biased predictions favoring the majority class [21]. To mitigate this issue, we implemented oversampling specifically for our multi-class classification models, which yielded improved results. This technique, a well-established text augmentation method, involves duplicating instances in the minority classes to equalize their representation with the majority classes. We employed a random oversampling approach to achieve a balanced class distribution. Table 2 illustrates the initial and post-resampling counts of training data sentences across each class.

Table 2. The number of data in each class after the resampling technique

Big Five traits	Number from the training data (sentences) before re-sampling	Number from the training data (sentences) after re-sampling
Openness	1,298	-
Conscientiousness	648	1,000
Extraversion	1,040	-
Agreeableness	522	1,000
Emotional stability	537	1,000
Total	4,045	5,080

2.3. Transfer learning models

Building on the success of transfer learning from pre-trained models in NLP tasks, our research includes fine-tuning these models to identify soft skills in a binary classification test and personality characteristics based on the big five in a multi-class classification job. We evaluated the efficacy of multilingual pre-trained models by testing many of them, including BERT, RoBERTa, and XLM-RoBERTa. The purpose of the binary classification exercise was to detect the existence of soft skills in provided texts. The multi-class classification exercise attempted to categorize phrases based on the big five personality qualities of openness, conscientiousness, extraversion, agreeableness, and emotional stability.

The evaluations were performed done in the Google Colab environment, which utilized its computing capabilities for model training and fine-tuning. We used the simple transformers package, which simplifies the design of Transformer-based models and enables rapid fine-tuning of pre-trained language models. The dataset was divided into training and testing sets to carefully analyze the model's performance. The training set was utilized to fine-tune the models, while the testing set was retained for assessing the model's generalization skills. The scikit-learn Python package was used to produce evaluation measures such as accuracy, recall, precision, and F1-score, ensuring a thorough assessment of the models' performance. Accuracy evaluates the model's overall accuracy, whereas precision and recall offer information on the model's performance in certain classes. The F1-score, which is the harmonic means of accuracy and recall, was very important in determining the optimal balance of precision and recall.

We also addressed the issue of class imbalance, particularly in the multi-class classification job, by using oversampling techniques to guarantee that minority classes were appropriately represented in the training set. This method helps to reduce the possibility of the model being biased toward the majority class, boosting the model's robustness and fairness in predictions. In conclusion, our work used transfer learning and fine-tuning of pre-trained models to obtain considerable outcomes in binary and multi-class classification problems. We demonstrated the ability of advanced models such as BERT, RoBERTa, and XLM-RoBERTa to detect soft skills and personality traits, contributing valuable insights to the field of NLP and HR analytics.

2.3.1. BERT (base)

Google's BERT [22], a transformer-based pre-trained model, was used for our tasks. BERT is available in various sizes, such as large and base, differing in hidden layers, attention heads, hidden size, and parameters. For this thesis, the (base) model was fine-tuned for both tasks.

2.3.2. Multilingual BERT model

Multilingual BERT model [23] is pre-trained in the top 104 languages with the most Wikipedia entries. It is available in two versions: old (not recommended) and new (recommended). The new version supports 104 languages, addressing normalization concerns in various languages, making it highly recommended, especially for non-latin alphabets such as Arabic. In this thesis, the new version downloaded from the TensorFlow hub was utilized.

2.3.3. RoBERTa (base)

Facebook's RoBERTa [24], an enhanced version of BERT, was proposed as a robustly optimized model. It differs from BERT in the training process and internal composition. RoBERTa omits the next sentence prediction (NSP) objective, features a dynamic mask that trades tokens between training iterations, and was trained on a more extended sequence of tokens with a larger batch size compared to BERT.

2.3.4. XLM-RoBERTa (base)

The XLM-RoBERTa model introduced by research in[24], [25]. It is based on Facebook's RoBERTa model. XLM-RoBERTa is a large multilingual language model trained using 2.5 TB of filtered CommonCrawl data.

2.4. Pre-trained modeling technique

The pre-trained modeling technique follows a straightforward approach. We first prepare the data, then feed pre-processed sentences into a chosen pre-trained model. Finally, the model delivers its output. This process is illustrated in Figure 1. To ensure consistent training, we employed some standard settings. The training batch size was set to 32, while the evaluation batch size was 16. Additionally, for transfer learning, implementing early stopping with a patience of 3 was crucial. Each model halted training upon reaching a specific epoch, as detailed in Table 3. This table also outlines the individual early stopping and learning rate configurations for each model, along with the corresponding architecture used.

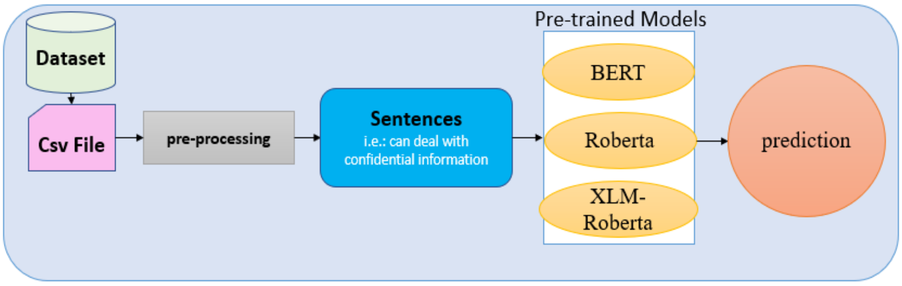


Figure 1. Pre-trained modeling technique

Table 3. Transfer learning models' architecture and IDs

Classification Tasks	Architecture	Model's ID	Number of epochs	Learning rate
Binary classification	BERT	Multilingual	15	3.00E-05
Binary classification	BERT	Base	4	5.00E-06
Binary classification	RoBERTa	Base	8	5.00E-06
Binary classification	XLM- RoBERTa	Base	2	3.00E-05
Multi-class classification	BERT	Multilingual	5	2.00E-06
Multi-class classification	BERT	Base	8	5.00E-06
Multi-class classification	RoBERTa	Base	10	2.00E-05
Multi-class classification	XLM- RoBERTa	Base	5	3.00E-05

3. RESULTS AND DISCUSSION

This section explains the results of research and at the same time gives the comprehensive discussion. Table 4 and Figure 2 show how we profited from the capabilities and performance specifications in transfer learning models. HuggingFace transformers were used to create the models [25].

Table 4. The accuracy, precision, recall, and F1-score for binary classification with transfer learning models

Model	Model ID	F1- score	Recall	Precision	Accuracy
BERT	Multilingual	0.79	0.79	0.79	0.79
BERT	Base	0.79	0.79	0.79	0.79
RoBERTa	Base	0.79	0.80	0.79	0.80
XLM- RoBERTa	Base	0.79	0.80	0.80	0.80

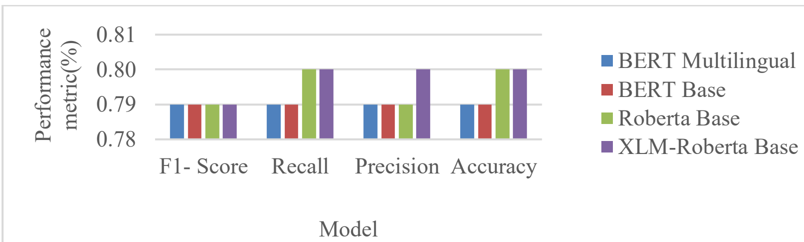


Figure 2. The accuracy, precision, recall, and F1-score for binary classification using transfer learning models

Table 4 and Figure 2 display the results of binary classification, indicating that the XLM-RoBERTA base version model achieved the highest performance. It demonstrated an accuracy of 80%, recall of 80%, F1-score measure of 79%, and precision of 80%. Similarly, RoBERTA also exhibited an accuracy of 80%, signifying that pretrained models surpassed task-trained DL Models. Moreover, Table 5 and Figure 3, focusing on multi-class classification, highlights XLM-RoBERTA's superior performance with an accuracy of 72%, outperforming task-trained DL Models in this specific task. The results showed that XLM-RoBERTA in both tasks (binary and multi-class classification) obtained the highest accuracy and achieved the best performance compared with other models.

These results produced an excellent performance for XLM-RoBERTa in both classification tasks. The superiority is that XLM-R [26], [27] is a cutting-edge multilingual masked language model trained on 2.5 TB of freshly constructed clean Common Crawl data in 100 languages. It outperforms prior multilingual models such as mBERT and XLM in classification, sequence labeling, and question answering, but it reveals the limits of multilingual MLMs, particularly the high-resource versus low-resource trade-off and the relevance of critical hyperparameters. Moreover, the XLM-R multilingual model is the first to surpass typical monolingual baselines that depend on pre-trained models. Multilingual BERT and XLM models have challenges learning usable representations for low-resource languages. XLM-R demonstrates the superiority of multilingual models over monolingual models for low-resource languages [28].

Table 5. The accuracy, precision, recall, and F1-score for multi class classification with transfer learning models

Model	Model ID	F1- Score	Recall	Precision	Accuracy
BERT	Multilingual	0.70	0.70	0.70	0.70
BERT	Base	0.67	0.67	0.67	0.67
Roberta	Base	0.71	0.71	0.71	0.71
XLM-Roberta	Base	0.72	0.72	0.72	0.72

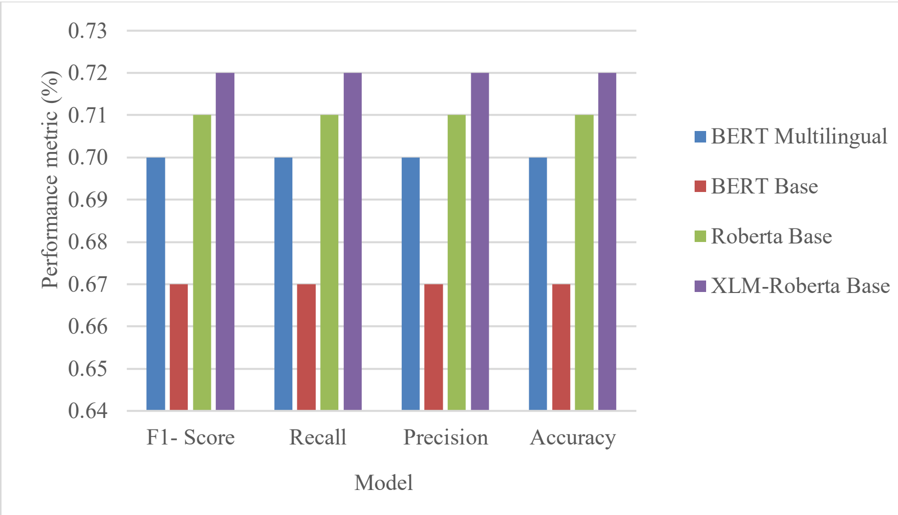


Figure 3. The accuracy, precision, recall and F1-score for multi class classification with transfer learning

4. CONCLUSION

The experimental findings clearly support the objectives outlined in the introduction, particularly in enhancing recruitment precision using transformer-based models. This study addresses critical challenges in talent acquisition by exploring advanced techniques for identifying optimal candidates across various job categories. Our investigation into Transformer models, including BERT, RoBERTa, and XLM-RoBERTa, yielded promising results, with XLM-RoBERTa achieving 80 and 72% accuracy in binary and multi-class classification tasks, respectively. A key innovation lies in our method's ability to automatically extract, categorize, and map novel skills from job descriptions and resumes to career categories based on personality traits. This streamlines the recruitment process and enhances personalized matching between candidates and jobs. However, the study’s reliance on static datasets may affect adaptability across different linguistic or cultural contexts. Future research could focus on utilizing real-time applicant data and reinforcement learning to enhance prediction accuracy and adapt to market fluctuations. By demonstrating the effectiveness of transfer learning in NLP for recruitment, this work highlights the potential of AI in making hiring more precise, fair, and efficient.

FUNDING INFORMATION

Authors state there is no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Nour Alqudah	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓			
Qusai Q. Abuein	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
Mohammed Q. Shatnawi	✓	✓		✓						✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study were created by the authors and are openly available in the HYPskill dataset on Kaggle at https://www.kaggle.com/datasets/nouralqudah/hypskill-dataset.

REFERENCES

[1] G. Koman, P. Boršoš, and M. Kubina, "The possibilities of using artificial intelligence as a key technology in the current employee recruitment process," *Administrative Sciences*, vol. 14, no. 7, 2024, doi: 10.3390/admsci14070157.

[2] M. Tajpour, A. Salamzadeh, and E. Hosseini, "Job satisfaction in IT Department of Mellat Bank: does employer brand matter?," *IPSI BgD Transactions on Internet Research*, vol. 17, no. 1, pp. 15–21, 2021.

[3] L. J. G.-Gomez, S. M. H.-Munoz, A. Borja, J. D. Azofeifa, J. Noguez, and P. Caratozzolo, "Analyzing natural language processing techniques to extract meaningful information on skills acquisition from textual content," *IEEE Access*, vol. 12, pp. 139742–139757, 2024, doi: 10.1109/ACCESS.2024.3465409.

[4] R. Joshi, V. Nair, M. Singh, A. N.-I. J. of AI, and U. 2021, "Leveraging natural language processing and predictive analytics for enhanced AI-driven lead nurturing and engagement," *International Journal of AI Advancements*, vol. 10, no. 1, pp. 1-22, 2021.

[5] R. Mehmood, F. Alam, N. N. Albogami, I. Katib, A. Albeshri, and S. M. Altowaijri, "UTiLearn: a personalised ubiquitous teaching and learning system for smart societies," *IEEE Access*, vol. 5, pp. 2615–2635, 2017, doi: 10.1109/ACCESS.2017.2668840.

[6] M. Tims, D. Derks, and A. B. Bakker, "Job crafting and its relationships with person-job fit and meaningfulness: a three-wave study," *Journal of Vocational Behavior*, vol. 92, pp. 44–53, 2016, doi: 10.1016/j.jvb.2015.11.007.

[7] B. R. Dineen, S. R. Ash, and R. A. Noe, "A web of applicant attraction: person-organization fit in the context of web-based recruitment," *Journal of Applied Psychology*, vol. 87, no. 4, pp. 723–734, 2002, doi: 10.1037/0021-9010.87.4.723.

[8] G. Jain, J. Paul, and A. Shrivastava, "Hyper-personalization, co-creation, digital clienteling and transformation," *Journal of Business Research*, vol. 124, pp. 12–23, 2021, doi: 10.1016/j.jbusres.2020.11.034.

[9] G. N. Burns, N. D. Christiansen, M. B. Morris, D. A. Periard, and J. A. Coaster, "Effects of applicant personality on resume evaluations," *Journal of Business and Psychology*, vol. 29, no. 4, pp. 573–591, 2014, doi: 10.1007/s10869-014-9349-6.

[10] P. L. Roth, P. Bobko, C. H. V. Iddekinge, and J. B. Thatcher, "Social media in employee-selection-related decisions: a research agenda for uncharted territory," *Journal of Management*, vol. 42, no. 1, pp. 269–298, 2016, doi: 10.1177/0149206313503018.

[11] I. Nikolaou, "Social networking web sites in job search and employee recruitment," *International Journal of Selection and Assessment*, vol. 22, no. 2, pp. 179–189, 2014, doi: 10.1111/ijsa.12067.

[12] L. Sayfullina, E. Malmi, and J. Kannala, "Learning representations for soft skill matching," *Analysis of Images, Social Networks and Texts*, Cham, Switzerland: Springer, pp. 141–152, 2018, doi: 10.1007/978-3-030-11027-7_15.

[13] D. A. Tamburri, W. J. V. D. Heuvel, and M. Garriga, "DataOps for societal intelligence: a data pipeline for labor market skills extraction and matching," in *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science, IRI 2020*, 2020, pp. 391–394, doi: 10.1109/IRI49571.2020.00063.

[14] A. Bhola, K. Halder, A. Prasad, and M. Y. Kan, "Retrieving skills from job descriptions: a language model based extreme multi-label classification framework," in *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, 2020, pp. 5832–5842, doi: 10.18653/v1/2020.coling-main.513.

[15] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, pp. 5832–5842, May 2019, doi: 10.18653/v1/2020.coling-main.513.

[16] Q. Q. Abuein, M. Q. Shatnawi, and N. Alqudah, "Improving job matching with deep learning-based hyper-personalization," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 2, pp. 1711–1722, 2024, doi: 10.11591/ijai.v13.i2.pp1711-1722.

[17] A. Rahali and M. A. Akhloufi, "End-to-end transformer-based models in textual-based NLP," *AI*, vol. 4, no. 1, pp. 54–110, 2023, doi: 10.3390/ai4010004.

[18] M. Prytula, "Fine-tuning BERT, DistilBERT, XLM-RoBERTa and Ukr-RoBERTa models for sentiment analysis of Ukrainian language reviews," *Artificial Intelligence*, vol. 29, no. 2, pp. 85–97, 2024, doi: 10.15407/jai2024.02.085.

[19] S. Fareri, N. Melluso, F. Chiarello, and G. Fantoni, "SkillNER: mining and mapping soft skills from any text," *Expert Systems with Applications*, vol. 184, 2021, doi: 10.1016/j.eswa.2021.115544.

[20] X. Li, X. Sun, Y. Meng, J. Liang, F. Wu, and J. Li, "Dice loss for data-imbalanced NLP tasks," *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 465–476, 2020, doi: 10.18653/v1/2020.acl-main.45.

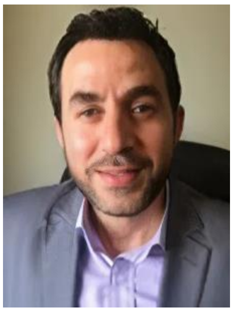
[21] G. Haixiang, L. Yijing, L. Yanan, L. Xiao, and L. Jinling, "BPSO-Adaboost-KNN ensemble learning algorithm for multi-class imbalanced data classification," *Engineering Applications of Artificial Intelligence*, vol. 49, pp. 176–193, 2016, doi: 10.1016/j.engappai.2015.09.011.

[22] M. V. Koroteev, "BERT: a review of applications in natural language processing and understanding," *arXiv-Computer Science*, pp. 1-18, Mar. 2021.
[23] Y. Liu *et al.*, "RoBERTa: A robustly optimized BERT pretraining approach," *arXiv-Computer Science*, pp. 1-13, Jul. 2019.
[24] A. Conneau *et al.*, "Unsupervised cross-lingual representation learning at scale," *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 8440–8451, 2020, doi: 10.18653/v1/2020.acl-main.747.
[25] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.
[26] M. Mars, "From word embeddings to pre-trained language models: a state-of-the-art walkthrough," *Applied Sciences*, vol. 12, no. 17, 2022, doi: 10.3390/app12178805.
[27] O. Mohamed, A. M. Kassem, A. Ashraf, S. Jamal, and E. H. Mohamed, "An ensemble transformer-based model for Arabic sentiment analysis," *Social Network Analysis and Mining*, vol. 13, no. 1, 2023, doi: 10.1007/s13278-022-01009-0.
[28] S. Mutuvi, "Epidemic event extraction in multilingual and low-resource settings," *Ph.D. dissertation*, Laboratoire Informatique, Image, Interaction (L3i), Université de La Rochelle, La Rochelle, France, 2022.

BIOGRAPHIES OF AUTHORS

Nour Alqudah holds master's degree in data science from Jordan University of Science and Technology. She currently works as a digital analytics engineer. Her research interests encompass AI, NLP, big data, deep learning, digital and data analytics. She can be contacted at email: nsalqudah18@cit.just.edu.jo.

Dr. Qusai Q. Abuein an Associate Professor in the Department of Computer Information Systems, Jordan University of Science and Technology. He received his Ph.D. and master's degrees from Ibaraki University, Hitachi, Japan. He teaches web social analysis, web analytics, and data visualization. His research interests include web social analysis, web analytics, data analysis, and information retrieval. He can be contacted at email: qabuein@just.edu.jo.



Dr. Mohammed Q. Shatnawi is a computer science academic and researcher with a strong background in machine learning, data science, and their applications. He earned his B.Sc. in Computer Science from Yarmouk University and went on to complete both his M.Sc. and Ph.D. in Computer Science from The George Washington University, graduating in January 2007. His work bridges theoretical models with real-world applications, contributing significantly to the advancement of intelligent systems and data-driven technologies. He can be contacted at email: mshatnawi@just.edu.jo.