

Resume Bias Detection and Mitigation using AI/ML

Anushka Adak, Disha Barapatre, Aarya Nalawade, Gargi Pawar (Information Technology)
Marathwada Mitra Mandal's College of Engineering, Pune
Affiliated by Savitribai Phule Pune University

Guide: Sayali Joshi

Abstract - Automatic resume screening is done using AI and ML techniques that enable rapid evaluation of a large number of applications received by employers for specific jobs. This process saves time and money. The drawback of these technologies is that they may have some prejudices ingrained in them from the past data used in developing them. This leads to biased hiring practices that can discriminate against certain individuals on the grounds of gender, age, and geographic origin.

Keywords - AI Bias, Recruitment Algorithms, Disparate Impact, Algorithmic Fairness, Machine Learning Ethics, HR Technology

1. INTRODUCTION

Organizations receive stacks of resumes in a competitive labour market. Manual screening becomes inefficient, time-consuming, and biased. AI combined with NLP can extract key details like education, experience, and skills from resumes and rank candidates based on job requirements.

The proposed system is an Artificial Intelligence driven resume screening system, which not only processes resumes in multiple formats but also supports worldly wise NLP for contextual candidate evaluations beyond simple keyword matching. The system minimizes bias, reduces manual effort, and improves efficiency while supporting fair, data driven hiring decisions.

1.1 Research Objectives

The primary objectives of this research are:

1. To quantify algorithmic bias in AI recruitment tools using standardized fairness metrics
2. To identify patterns of discrimination across different job categories and candidate profiles
3. To develop a reproducible methodology for bias testing that organizations can implement
4. To provide recommendations for bias mitigation in AI powered recruitment systems

1.2 Scope and Objectives

It is a process that includes the identification and reduction of unjustified discriminatory actions in resume screening through

manual and AI based recruitment systems. These are discrimination based on factors like gender, age, race, ethnicity, educational background, and many more. All these are considered at all stages of early recruitment, particularly at the resume shortlisting phase. The proposed research has the following objectives, ensuring non discriminatory and merit based candidate screening, promoting diversity and inclusion, enhancing the accuracy and ethics of candidate screening, ensuring compliance with anti discrimination laws, and ensuring white box based and unbiased candidate screening systems that can recruit candidates purely based on their qualifications, skills, and experience.

2. LITERATURE REVIEW

The automation of resume screening has witnessed tremendous growth in recent years owing to the ability to increase the speed of the recruitment process while minimizing biases in the selection process. Recent breakthroughs in the development of natural language processing and machine learning have authorized the selection of the most suitable candidates by understanding the information presented in the resume more effectively. Research has proven the effective utilization of different language models in the development of vector embeddings for achieving the best possible similarity between the resume and the job requirements, as well as the accuracy of the ranking system. Learning algorithms such as Support Vector Machines and Decision Trees have also witnessed tremendous success in the classification of the resume by utilizing the most effective feature extraction mechanisms. Furthermore, the development of stacked models and the utilization of real time recommendation systems through the utilization of NLP, graph neural networks, and LSTM models have also witnessed tremendous growth in the accuracy of the resume screening system. Despite the development of the most effective resume screening system, there are still a number of challenges that remain to be addressed in the future by utilizing the most advanced NLP techniques.

Approach	Table Column Head		
	Method	Application	Limitation
LLM-based screening	LLMs, embeddings	Better semantic matching	Format diversity
ML Resume Ranking	Feature-based ML	Accurate ranking	Real-time load
Auto ML Screening	NLP, ML models	High precision filtering	Scalability
ML for HR Classification	NLP, ML	Improved shortlisting	Improved shortlisting
Model Stability Testing	Cross-format testing	Stability analysis	Data diversity

3. METHODOLOGY

The methodology for resume bias detection and mitigation using AI/ML begins with collecting resume data that includes demographic and job-related features and identifying sensitive attributes such as gender or country.

3.1. Data Collection and Preprocessing:

The resumes are collected in various formats such as PDF and/or DOCX and processed using Natural Language Processing (NLP) techniques. The data in each resume will then be scraped and then cleansed, which includes removal of stop words, tokenization, and text normalizations. This assists in minimize irrelevant data as much as possible, and important details are highlighted.

3.2. Feature Extraction:

Applications such as Named Entity Recognition (NER) in NLP are used for identifying important pieces of information, including the education, experience, skills, and certifications. Massive language models, help to produce vector representations for each resume, incorporating semantic meaning relevant for comparisons with job requirements. Skill comparisons are made according to specific job description search terms linked to the target position.

3.3: Machine Learning Model for Ranking:

The processed data will then be fed into machine learning algorithms such as SVMs for ranking is either appropriate or inappropriate as needed. Additionally, a rank model (KNN model or a stack model) is used to determine how relevant each candidate is to the job position. on a per-extracted skill set basis relative to job requirements, to give a top-ranked list of applicants.

3.4. Testing and Evaluation:

The system performance measurement is done through measures such as accuracy, precision, remember, and F1-score from a labeled resume. The data set is a basis. Cross-validation will be done to optimize model parameters that will ensure ranking robustness and classification.

3.5. Deployment and user Interface:

An implementation of the model on a server with a web-based interface, which will enable uploading resumes and also inputting of job demands and ranked candidates. Being provided with real-time updates, this interface is meant for

showing rankings: matched skills, and job fit scores of the candidate.

4. ANALYSIS AND DISCUSSION

4.1 Effectiveness of NLP on Resume Parsing:

The skills/experience extraction on resumes performed using various NLP tasks like NER and vector embeddings with LLM (BERT/SBERT) is accurate regardless of resume formats and language use.

4.2 Ranking of Candidates and Accuracy of ML Models:

Models such as SVM, XGBoost, KNN, and ensemble models perform very well in ranking the candidates based on relevance to the job and achieve accuracy of more than 90% after validating and adjusting the models.

4.3 Comparison With the Traditional Method of results:

Automated screening is much faster than the manual method of screening, and it is bias-free, ensuring equal opportunities to all candidates qualified for the position.

4.4 Future Enhancements:

System functionality can be optimized with the help of adopting better transformer models and multi-lingual capabilities to adapt for the future trends in the labor market.

5. EXPERIMENTAL RESULTS SUMMARY

5.1 Model Performance:

Random Forest is gaining 88.15% accuracy but problem showed as high bias (DPD = 1.0, EOD = 1.0), favoring candidates from certain countries. Reweighting and Disparate Impact Remover failed to reduce bias effectively. Adversarial Debiasing eliminated bias (DPD = 0, EOD = 0) while maintaining accuracy at 88.15%, making it the most effective technique.

5.2 SHAP Analysis:

Before debiasing, organization was a dominant feature, overshadowing valid predictors like education and previous salary. After debiasing, the influence of Country decreased, and legitimate features gained importance, promoting fairer decision-making.

5.3 Confusion Matrix Analysis:

Original model had imbalanced errors, disadvantaging candidates from underrepresented countries. Debaised model showed balanced errors across groups, ensuring equitable evaluation.

6. LIMITATIONS

There are few limitations in model designing which is not much concerning, but this can be addressed if we have more data to train this model. Current limitations of this model are as

follows: i) Model is designed to take input in CSV format. But in reality, resumes are either in .doc format, .pdf format, etc. But due to the limitations of this dataset, this model could not be enhanced to take input either in .doc format or .pdf format. But this can be done using a library “text extract.” This library can read input in any format and can convert it into a single format that can be used to this model. This is because this library can read input in any format and convert it into a single format that can be used to this model. ii) Generation of summary using “genism” library. This might cause a loss of information because of implicit compression of information due to summarization. This can be tuned so that information is not lost. For example, important features of information, like experience and skill of a candidate, are not lost.

7. CONCLUSION

The AI-based Resume Screening System is a futuristic tool that speeds up hiring by efficiently extracting relevant skills and ranking candidates using NLP and ML techniques like SVMs, KNN, vector embeddings, and few models. It minimizes the scope of bias, effort, and fairness in a variety of resume formats. Although the presence of bias and standard data in the past has been a challenge, future improvements in the tool may include the use of advanced “Natural Language Processing” and “Multilingual Support”. Overall, the system offers a reliable, efficient, and fair approach to recruitment, transforming the hiring process.

8. REFERENCES

- [1] AlOtaibi, S.T., Ykhlef, M., 2012. A survey of job recommender systems. *International Journal of Physical Sciences* 7, 5127–5142.
- [2] Breaugh, J.A., 2009. The use of bio data for employee selection: Past research and future directions. *Human Resource Management Review* 19, 219–231.
- [3] Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32.
- [4] Carrer-Neto, W., Hernandez-Alcaraz, M.L., Valencia-García, R., García-Sánchez, F., 2012. Social knowledge-based recommender system. application to the movies domain. *Expert Systems with Applications* 39, 10990–11000.
- [5] Celma, O., 2010. Music recommendation, in: *Music recommendation and discovery*. Springer, pp. 43–85.
- [6] Das, A.S., Datar, M., Garg, A., Rajaram, S., 2007. Google news personalization: scalable online collaborative filtering, in: *Proceedings of the 16th international conference on World Wide Web*, ACM, pp. 271–280.
- [7] Diao, Q., Qiu, M., Wu, C.Y., Smola, A.J., Jiang, J., Wang, C., 2014. Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars), in: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 193–202.
- [8] Farber, F., Weitzel, T., Keim, T., 2003. An automated recommendation approach to selection in personnel recruitment. *AMCIS 2003 proceedings*, 302.
- [9] Golec, A., Kahya, E., 2007. A fuzzy model for competency-based employee evaluation and selection. *Computers & Industrial Engineering* 52, 143–161.
- [10] Howard, J.L., Ferris, G.R., 1996. The employment interview context: Social and situational influences on interviewer decisions. *Journal of applied social psychology* 26, 112–