AEQUITAS: A FAIRNESS-DRIVEN REVIEW OF AUTOMATED RESUME PARSING AND CANDIDATE MATCHING TECHNIQUES

[A1]Sneh Lata Singh, [A2]Arpit Upadhyay, [A3]Devansh Thakur
[1,2,3]Department of Computer Science & Engineering
[A]*Veer Madho Singh Bhandari Uttarakhand Technical University*
[A1]Corresponding Author: slsingh003@gmail.com

Abstract: **Background:** As organizations face overwhelming numbers of applicants and the inefficiencies of manual screening, automating the hiring process has turned into a necessity. **Objective:** This review paper assesses the current state of automated hiring systems, with a focus on two major areas: Resume Parsing, which is all about converting unstructured text into structured data, and Candidate Matching, where we rank applicants against job descriptions. **Methodology:** We carried out an in-depth examination of different methodologies, spanning from classic Natural Language Processing (NLP) and rule-based heuristics to the latest Machine Learning (ML) pipelines that incorporate Named Entity Recognition (NER) and Topic Modeling. **Key Findings:** While approaches like TF-IDF and Cosine Similarity are quick and efficient, they do not capture the nuances of meaning. Contemporary systems that utilize spaCy and Latent Dirichlet Allocation (LDA) excel in entity extraction, but come with deployment challenges. **Research Gap:** Upon critical examination, there is a significant shortcoming in how algorithmic bias is handled. Most existing systems prioritize accuracy but neglect fairness across factors like gender, geography, or formatting. **Conclusion:** The Aequitas project addresses these disparities by integrating transformer-based embeddings with strict fairness auditing protocols to ensure hiring automation is fair and just for everyone.

Index Terms - *Resume Parsing, NLP, Candidate Matching, Algorithmic Bias, NER, TFIDF, Aequitas.*

I. INTRODUCTION

Hiring automation has shifted from being a nice-to-have to an absolute must for businesses. Nowadays, companies are grappling with overwhelmed HR teams, a mishmash of resume formats, and a ticking clock, which makes the traditional human screening process slow, inconsistent, and often error-prone [1, 6].

1.1 Industry Context and Rationale

Recruitment data is on the rise like never before. According to recent industry reports, corporate job openings are receiving hundreds of resumes, which can lead to major delays in the "time-to-hire." Additionally, the global HR technology market is set to grow substantially, thanks to the increasing demand for AI-driven efficiency. Still, it is crucial to ensure that this efficiency does not come at the expense of fairness.

1.2 Scope of Review

This article takes a closer look at foundational studies and the latest advancements aimed at achieving two key objectives:

- Effectively breaking down unstructured resumes (PDF/DOCX) into organized fields.
- Evaluating and ranking candidates based on specific job criteria. The review is organized to first delve into parsing techniques, then explore matching algorithms, and ultimately engage in a thoughtful discussion about the ethical concerns and algorithmic biases that the Aequitas project seeks to address [1].

II. REVIEW OF RESUME PARSING TECHNIQUES

Across the literature, two parsing strategies dominate: classical NLP combined with rules, and modern pipelines combining NER with topic modeling [2].

2.1 Classical NLP and Heuristics

Many studies focus on text extraction followed by applying rule-based logic. The standard preprocessing steps typically include cleaning the text, tokenization, removing stop words, and stemming or lemmatization with NLTK [1, 2].

- **Method:** Regex and heuristics are utilized to extract information like email addresses and phone numbers.
- **Critique:** While these pipelines are efficient and straightforward, they can be quite sensitive to different layouts (such as double-column formats or images) and non-standard language.

2.2 NER and Topic Modeling Pipelines

A notable trend is emerging where pre-trained models like spaCy are being used for entity extraction. Recent techniques link spaCy ('en_core_web_lg') with Latent Dirichlet Allocation (LDA) to identify entities like skills and education, as well as to reveal underlying themes that extend past mere keywords.

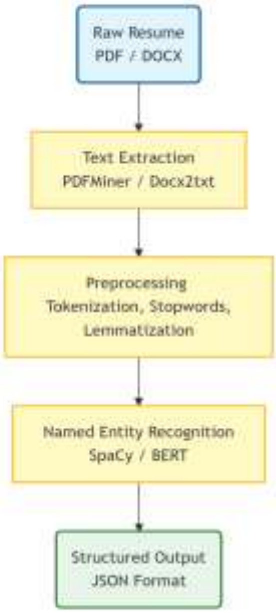

Fig. 1: Workflow of the Resume Parsing and Information Extraction Pipeline.

III. REVIEW OF CANDIDATE MATCHING MODELS

Matching approaches generally fall into three families: lexical vector similarity, classical ML ranking, and contextual embeddings.

3.1 TF-IDF and Cosine Similarity

This is the most common baseline found in literature [4, 5]. Resumes and job descriptions are transformed into TF-IDF vectors and then ranked based on cosine similarity. While this method is computationally inexpensive, it does have a drawback known as the "vocabulary mismatch" problem. For instance, a candidate who uses "ML" might not be recognized for a job that specifies "Machine Learning," unless the system is explicitly programmed to make that connection.

3.2 Classical Supervised Classifiers

When TF-IDF is paired with KNN or SVMs, the reported metrics can be quite high, with accuracy hitting around 98% on synthetic data [0]. However, a thorough analysis points out that these elevated figures are probably a result of limited datasets, which may not translate to effective performance in real-world applications.

IV. COMPARATIVE ANALYSIS OF METHODOLOGIES

To address the technical gaps in existing reviews, Table 1 presents a quantified comparison of the discussed methodologies.

Table 1: Comparative Analysis of Parsing and Matching Methodologies

Methodology	Core Techniques	Performance Metrics	Critical Analysis
Lexical Matching	TF-IDF, Cosine Similarity [5]	High Speed, Low Latency	High False Negatives; fails to capture semantic meaning or synonyms.
Statistical ML	KNN, SVM, Naive Bayes [4]	Acc: ~98% (Synthetic Data)	Prone to overfitting on small datasets; lacks context awareness.
Semantic Topic Modeling	LDA, LSA, spaCy NER [2]	Accuracy: ~82% (Real Data)	Better interpretability; captures latent themes (skills/education).

V. THE CRITICAL RESEARCH GAP: ALGORITHMIC BIAS

The most important insight from this review is that bias and fairness remain the least addressed issues.

- **Data Limitations:** Many studies use small test sets or overly similar datasets, which often miss out on geographic and linguistic diversity [1].
- **Fairness Evaluation:** Most applied research does not provide precision/recall splits by subgroup (like gender or ethnicity). TF-IDF methods can unintentionally highlight historical biases, such as gendered language in job descriptions [3].
- **Automation Risks:** The potential for AI systems to inherit cognitive biases is a significant concern. To tackle this, a collaborative approach between HR managers and developers is crucial [3].

VI. CONCLUSION

This review points out that even though engineering pipelines for resume parsing are evolving thanks to spaCy and NER, the concept of "fairness" is still not fully developed. The field often prioritizes metrics like "does it rank well?" but tends to ignore the important consideration of "does it rank fairly?"

VII. FUTURE PROSPECTS

Future research should shift its focus to:

- **Transformer-based Models:** Utilizing BERT or RoBERTa to achieve a deeper understanding of semantics that goes beyond just matching keywords.
- **Explainable AI (XAI):** Ensuring that the ranking decisions are clear and transparent for applicants.
- **Bias Mitigation:** Applying fairness constraints during both pre-processing and in-processing stages.

The Aequitas project aims to tackle these issues by creating diverse, multi-lingual evaluation sets and incorporating interpretable embedding models, all while carefully measuring disparate impact metrics.

ACKNOWLEDGMENT

The authors would like to thank their institution and mentors for their support and guidance throughout this research.

REFERENCES

[1] A. Fabris et al., "Fairness and Bias in Algorithmic Hiring: A Multidisciplinary Survey," ACM Transactions on Intelligent Systems and Technology, vol. 16, no. 1, art. 16, pp. 1–54, Jan. 2025.

[2] V. Jagwani, S. Meghani, S. Dhage, and K. Pai, "Resume Evaluation through Latent Dirichlet Allocation and Natural Language Processing for Effective Candidate Selection," in 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2023, pp. 1–6.

[3] M. Soleimani, A. Intezari, and D. J. Pauleen, "Mitigating Cognitive Biases in Developing AI-Assisted Recruitment Systems: A Knowledge-Sharing Approach," International Journal of Knowledge Management, vol. 18, no. 1, pp. 1–18, 2022.

[4] R. T. Fareed, V. Rajath, and S. Kaganurmath, "Resume Classification and Ranking using KNN and Cosine Similarity," International Journal of Engineering Research & Technology (IJERT), vol. 10, no. 08, pp. 192–195, Aug. 2021.

[5] A. Agarwal and Dr. Senthilkumar, "Resume Recommendation System Using Cosine Similarity," International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), vol. 04, no. 04, pp. 158–162, Apr. 2022.

[6] P. Pokharel, "Hireability: Information Extraction from Resume," Undergraduate Final Project, Herald College Kathmandu, Kathmandu, Nepal, 2019.

[7] M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices," in Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20), Barcelona, Spain, 2020, pp. 469–481.

[8] S. C. Geyik, S. Ambler, and K. Kenthapadi, "Fairness-Aware Ranking in Search & Recommendation Systems with Application to LinkedIn Talent Search," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19), Anchorage, AK, USA, 2019, pp. 2221–2231.

[9] K. V. Deshpande, S. Pan, and J. R. Foulds, "Mitigating Demographic Bias in AI-based Resume Filtering," in Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20), Genoa, Italy, 2020, pp. 268–275.

[10] C. Qin, H. Zhu, T. Xu, C. Zhu, L. Ma, and E. Chen, "Enhanced Collaborative Filtering with Auxiliary Information for Job Recommendation," in Proceedings of the 2020 SIAM International Conference on Data Mining, Cincinnati, OH, USA, 2020, pp. 604–612.

[11] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A Survey on Bias and Fairness in Machine Learning," ACM Computing Surveys (CSUR), vol. 54, no. 6, pp. 1–35, July 2021.

[12] Y. Bian, Y. Feng, Z. Wang, and J. Huan, "Domain Adaptation for Person-Job Fit with Transferable Deep Global Match Network," in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), Hong Kong, China, 2019, pp. 4810–4820.

[13] A. Pena, I. Serna, A. Morales, and J. Fierrez, "Bias in Multimodal AI: Testbed for Fair Automatic Recruitment," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 28–35.

[14] T. Le, A. G. Baydin, and R. Zinkov, "Towards Explainable Matchmaking in Recruitment with Graph Neural Networks," in 2019 ICML Workshop on Human in the Loop Learning (HILL), Long Beach, CA, USA, 2019.

[15] S. Zhang, L. Wang, and Y. Liu, "Deep Learning Based Job Recommendation: A Survey," IEEE Access, vol. 9, pp. 25687–25700, 2021.

[16] J. Sanchez-Monedero, L. Dencik, and L. Edwards, "What does it mean to 'solve' the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems," Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 2020, pp. 458–468.