**­­­­MODOO2602 TRI2 FO1CAM- Computing Research Methodologies**

**AI-Powered Resume Screening for Job Recruitment**

**Project Proposal Final Report**

Element 010-1

**─ Module Leader:** **Dr**

Prepared by:

SID:

Submission Date: 11-04-2025

Word count: 3600

## 

[**Table Of Content** 2](#_Toc195247115)

[Abstract 3](#_Toc195247116)

[1. Introduction 4](#_Toc195247117)

[2. Project Summary 4](#_Toc195247118)

[2.1 Aim 4](#_Toc195247119)

[2.2 Objectives 5](#_Toc195247120)

[3. Research Questions 5](#_Toc195247121)

[4. Business Need/Case 6](#_Toc195247122)

[4.1 Problem Background 6](#_Toc195247123)

[4.2 Market Needs 6](#_Toc195247124)

[4.3 Benefits 6](#_Toc195247125)

[4.4 High-Level Risks (Dis-Benefits) 7](#_Toc195247126)

[4.5 Approximate Budget/Time 7](#_Toc195247127)

[5. Work Breakdown Structure (WBS) 8](#_Toc195247128)

[6. Gantt Chart 8](#_Toc195247129)

[7. Literature Review 9](#_Toc195247130)

[7.1 Introduction 9](#_Toc195247131)

[7.2 Critical Review of 10 Related Literature 9](#_Toc195247132)

[7.3 Review Summary 11](#_Toc195247133)

[8. Research Methodology 12](#_Toc195247134)

[8.1 Research Framework 12](#_Toc195247135)

[8.2 Data Collection & Analysis 13](#_Toc195247136)

[8.3 Social, Legal, and Ethical Issues 13](#_Toc195247137)

[9. Findings and Future Directions 14](#_Toc195247138)

[9.1 Future Directions 14](#_Toc195247139)

[10. Self-Reflection 14](#_Toc195247140)

[11. Conclusion 15](#_Toc195247141)

[12. References 16](#_Toc195247142)

# **Abstract**

This study proposes applying natural language processing (NLP) and machine learning (ML) to create and validate an AI-powered resume screening system. The system will automatically screen applicants by extracting and ranking relevant qualifications, experiences, and skills. The system's accuracy, effectiveness, and potential for bias reduction are compared with conventional screening. Legal and ethical issues will be evaluated, academic studies and business practices will be considered, and an AI prototype will be developed. Best practices for AI-based recruitment in a fair, transparent, and lawful way will be informed by the results.

# **1. Introduction**

To find, attract, and hire best talent for the vacant positions aiding in the business success and sustainability in the long run, the recruitment process is critical. In one vacancy, companies receive hundreds or thousands of resumes in the fast-moving and competitive business environment today, and thus manual screening is tedious and time-consuming. This conventional hiring process takes time and increases the chances of implicit bias, decisional prejudice, and suboptimal quality selection of top-level talents.

Increasingly, technology-enabled solutions are deployed to maximize the efficiency of recruiting, equity, and openness. Convergence between artificial intelligence (AI), machine learning (ML), and natural language processing (NLP) in principle can make resume screening automatic, rate candidates on merit, and minimize human mistake. AI in the hiring processes does come with some disadvantages too, including the threat of algorithmic bias, the need for automated decisions are explainable and adherence to privacy laws such as GDPR.

This paper discusses how a resume screening system using AI might be developed in order to attain optimal equality, effectiveness, and adherence to ethics as a response to such challenges.In addition to trying to provide a factual response to recruitment matters currently in progress, it examines the risk and accountability of AI-driven recruitment platforms.

# **2. Project Summary**

## **2.1 Aim**

The objective of the project is to develop and mimic an AI-based resume screening tool that is capable of resume screening of candidates for positions based on machine learning (ML) and natural language processing (NLP). The best candidates will be filtered out by the algorithm, and it will be capable of retrieving the proper resume information, i.e., work experience, education, and skills, and comparing them with the demands of the job. Speed, accuracy, and fairness of the hiring and reducing the HR staff workload is the end objective.

## **2.2 Objectives**

For this, the project will first develop an NLP model for transforming unstructured resume data to structured data and extracting key features. Second, an AI-powered rating system for screening candidates in accordance with predefined hiring parameters will be developed. Concurrently, the project will develop systems to detect and prevent prejudice so that ethical hiring can be guaranteed. Comparative efficiency and fairness of the system will be gauged by assessing its performance against conventional manual screening methods.

In addition to building a natural HR professional user interface, the project shall also determine the means to provide real-time feedback and continuous learning capabilities to the AI model. Seamless integration with current HR management systems shall be achieved by means of guaranteeing legal compliance as well as transparency of data. This project advocates for equal and data-driven recruitment solutions through dealing with central recruitment issues.

# **3. Research Questions**

1. How effective is AI in automating resume screening compared to traditional methods?
2. Which NLP approaches are most powerful in extracting experience and skill from resumes?
3. How can AI mitigate bias during resume screening, and what ethical considerations must be answered?
4. What factors affect the precision and fairness of AI-powered resume screening tools?
5. How can AI-powered resume screening enhance recruitment efficiency and decision-making?
6. What are the challenges that come about in integrating traditional HR management systems with AI-based screening?
7. How can AI resume screening be optimized to ensure transparency and compliance with labor regulations?
8. What is the impact of AI-powered screening on business brand and candidate experience?
9. How can AI-powered recruiting platforms be configured to accommodate diverse industries with dissimilar hiring needs?
10. How can explainability in machine learning promote greater trust in AI-driven employment decisions?

# **4. Business Need/Case**

## **4.1 Problem Background**

Companies are getting too many job applications in the fast-paced digital age of hiring, so it is tedious, error-prone, and inefficient to sort through resumes by hand. With time and capacity constraints, human resource processes tend to make subjective choices, judge candidates unfairly, and prolong the hiring process. Implicit biases based on demographic characteristics like gender, race, and education might also affect traditional hiring processes, leading to discriminatory hiring processes and less diverse organizations. These problems ultimately result in expensive hiring mistakes, talent losses, and suboptimal organizational performance. Even worse, there is no fair, fact-based grading system. By developing an AI-powered resume screening tool that enhances productivity, fairness, and compliance, our project aims to fill these basic gaps.

## **4.2 Market Needs**

The demand for intelligent, automated hiring solutions is increasing in the global HR tech marketplace. With the assurance of faster screening, shorter time-to-hire, and higher-quality leads, AI-based tools are forcing businesses to reconsider how they recruit. Meeting this demand today relies considerably on machine learning (ML) and natural language processing (NLP), particularly when dealing with unstructured resume data. Aside from cost and efficiency savings, company restructurings have to contend with becoming more and more complex legal and regulatory requirements. Compliance with the law for example the US Equal Employment Opportunity Commission (EEOC) and the European Union's General Data Protection Regulation (GDPR) needs equality and transparency in the recruitment process. Additionally, by offering fast response times, open communication, and unbiased consideration, the market demands technologies that may be used to enhance the candidate experience. The business case for the project is further complemented by the alignment of these social, technological, and legal drivers.

## **4.3 Benefits**

There are a number of economic advantages of using an AI-driven resume screening system. Automated activities such as resume processing, first-level screening, and keyword searching have the primary advantage of maximizing the efficiency of recruitment. AI can help to enhance long-term employee alignment and decreasing turnover levels by employing skill-based algorithms that better match the candidate to position. Higher diversity objectives are achievable because of the in-house bias reduction capabilities of the system, which assist human resource departments in ensuring equity and diversity in hiring. Compliance and audit trail company report generation is also facilitated by standardized, compliant AI-based testing, which enhances transparency and accountability in accordance with the law. Predictive intelligence from machine learning algorithms also aids in strategic workforce planning by indicating high-potential candidates. These benefits ultimately save costs, enhance employment procedures, and enhance company image.

## **4.4 High-Level Risks (Dis-Benefits)**

Promising as it is, the use of AI in hiring has several disadvantages that have to be balanced. The most risky thing is to apply biased or unrepresentative recruitment information as the input to the AI system, and it amplifies prejudice. This is bound to lead to discrimination, especially against under-represented groups. Secondly, the "black box" phenomenon of certain machine learning models would render it difficult for HR staff to explain or justify decisions, which would be confidence-eroding and raise moral issues. Another high-risk threat is non-conformity with the law, especially with regards to privacy and equal opportunity law, which would involve financial sanctions and reputational damage. Finally, the resistance of HR units to technology adaptation or lack of digital literacy may limit the application of AI systems, which would subtract from their effectiveness. Dis-benefits have to be addressed with careful design, fair auditing, training, and open communications.

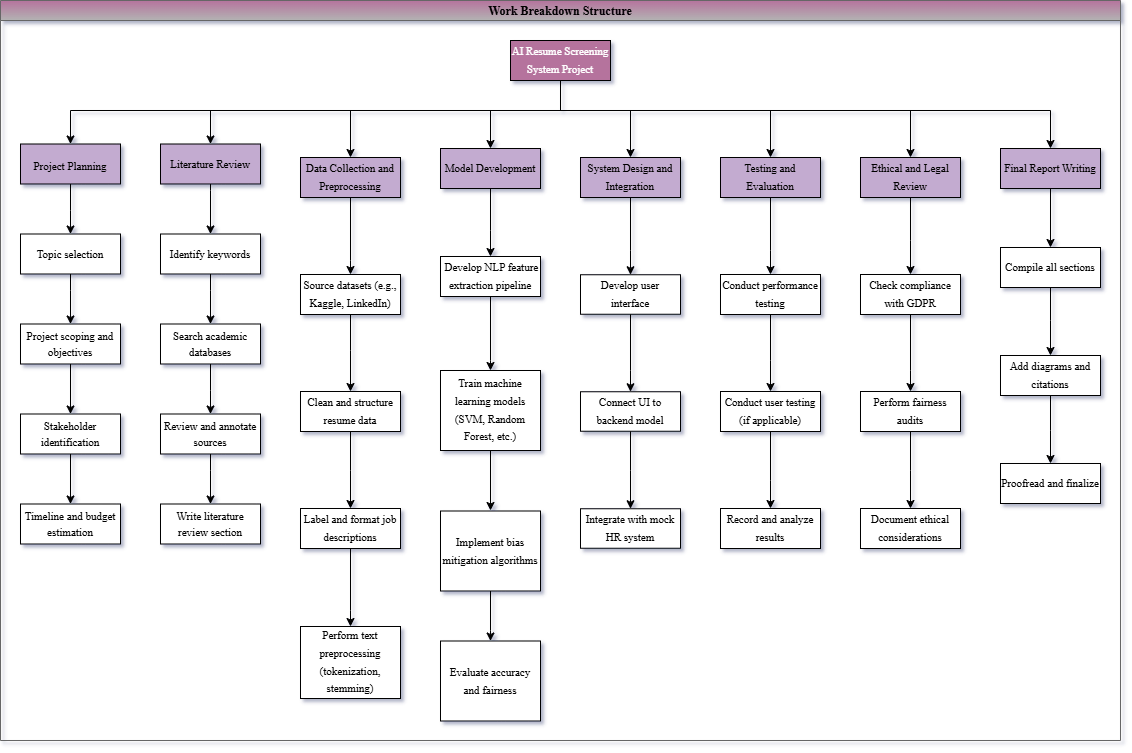
## **4.5 Approximate Budget/Time**

It would take six months to complete and would cost somewhere between £5,000 to £8,000. The following are included in the estimate:

* £1,500 for software development tools and licences
* Preprocessing, cleaning, and data collection: £1,000.
* £2,000 for interfaces integration and building AI models
* £500 for reporting, user testing, and evaluation

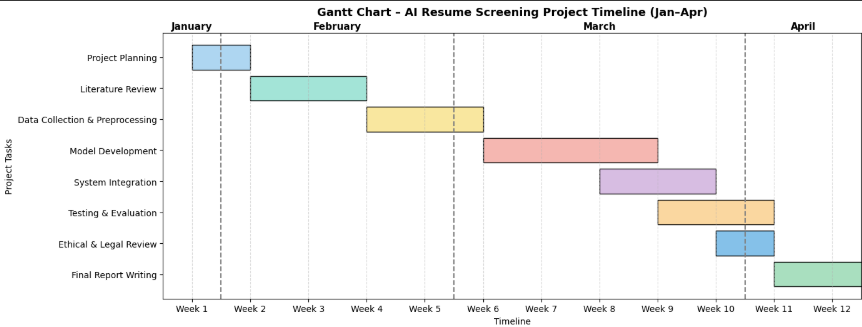
The project remains within budget while trying to offer a high-impact, moral, and innovative recruitment solution based upon open-source materials.

# **5. Work Breakdown Structure (WBS)**



**Figure 1:** The WBS of the AI Resume Screening Project with all the key deliverables and related work packages.

# **6. Gantt Chart**

 **Figure 2:** Gantt chart project timeline of AI Resume Screening System, in line with all the key deliverables, and completed within 12 weeks.

# **7. Literature Review**

## **7.1 Introduction**

The use of Artificial Intelligence (AI) in Human Resource (HR) operations, such as resume scanning, has transformed the workplace. With mounting pressure on businesses to recruit in a timely, accurate, and fair manner, AI-based software has emerged as a top choice for automated candidate shortlisting. The existing literature is discussed here in the context of AI-based recruitment websites, natural language processing (NLP) resume parsing software, and increasing focus on compliance with the law, fairness, and objectivity in automated recruitment software.

Natural language processing (NLP) techniques for extracting structured candidate data from resumes, (1) AI recruitment software and their effects on recruitment efficiency, and (3) ethical, legal, and fairness issues in automated decision-making are the three main themes of this paper. The research goals of the project resulted in these themes. These problems are the theoretical foundations for the design and evaluation of the system developed by this project.

**Keywords:** explainable AI, HR automation, prevention of bias in AI, AI hiring, NLP resume parsing, legal compliance, and machine learning fairness.

## **7.2 Critical Review of 10 Related Literature**

In their study of national origin discrimination in deep learning-based resume screening websites, Li, Li, and Lu (2023) illustrate how algorithmic discrimination can contribute to discriminatory job applicant exclusion. They define their research as setting out the necessity for fairness-sensitive training to reduce unfair outcomes in computer-aided hiring, and that is precisely how our project's fairness audit module is integrated.

Amin et al. (2019) present a rule-based keyword web-based resume screening system. matching. While their method is theoretically efficient in minimizing human effort, it is devoid of semantic information required to analyze resume material. This is evidence that our project demands sophisticated natural language processing models to perform context-sensitive resume parsing and enhance applicant rating accuracy.

To identify and analyze candidate skills, Priyanka and Parveen (2024) use the DeepSkillNER model that incorporates spectral clustering, and Named Entity Recognition (NER). For the NLP module in our project, which is required to extract CV features such as experience and qualifications to match up with job adverts, the model's hybrid framework influenced us.

The text understanding model known as BERT, introduced by Devlin et al. (2019), significantly it improves contextual text comprehension. As it is robust in semantic matching between resumes and job listings, the BERT model is now a de facto industry standard for NLP tasks. This paradigm has the technology substrate to accommodate text embedding to support the technology foundation in our screening software.

The transformer model is introduced to us by Vaswani et al. (2017) in "Attention is All You Need" and forms the foundation on which BERT and other modern NLP models are based on. Their focus mechanism enables them to read long and complex sentences, which are prone to come up on resumes, by recognizing what aspects of the input are content-heavy that they would like to probe.

The process introduced by Gan, Zhang, and Mori (2024) involves adding large language models (LLMs) with resume screening processes. Their research provides insights into how accuracy in matching can be applied to semantically match resumes and job descriptions. The method directly affects our eagerness to adopt transformer-based models into the back end of our project.

In their study on human bias in resume screening, Moore, Livingston, and Susskind (2023) found that recruiters apply cognitive shortcuts in their everyday practice. A trained algorithm is more likely to make merit-based and consistent decisions than human evaluators when faced with cognitive pressure, they established. This is more evidence in support of the use of AI for automated screening.

Aside from recording differential effects in automated decision-making, Feldman et al. (2015) provide ways to identify and eliminate systemic bias in automated systems. While evaluating our system, their standards for fairness and modification techniques, suggested by them, are relevant while verifying training sets and model output.

Chen, Szolovits, and Ghassemi (2020) highlight the importance of explainable AI (XAI) to make significant decisions, e.g., selecting a profession. They recommend that the users employ interpretation tools like SHAP and LIME to become more aware of ML models. Explainability layers are integrated in the screening interface of our solution to offer transparency to HR stakeholders.

Information Commissioner's Office (ICO, 2023) and the United States Equal Employment Opportunity Commission (EEOC, 2023) have developed employment discrimination legislation and personal data processing. GDPR is focused on transparency and rights of data subjects, whereas the EEOC offers legal protection against racism. Subsequent to these legislations, our project design utilizes anonymised data processing, explainability, and audit trails.

| **Title** | **Method** | **Pros** | **Cons** |
| --- | --- | --- | --- |
| Li et al. (2023) | Bias audit in deep learning | Highlights need for fairness-aware training | Focuses mainly on national origin discrimination |
| Amin et al. (2019) | Rule-based keyword matching system | Reduces manual effort in resume screening | Lacks semantic understanding; basic filtering |
| Priyanka & Parveen (2024) | DeepSkillNER – Deep learning + NER + clustering | Effectively extracts and ranks candidate skills | No integration with HR platforms or real-world testing |
| Devlin et al. (2019) | BERT – Contextual NLP model | Strong semantic matching between text inputs | High computational cost for training/inference |
| Vaswani et al. (2017) | Transformer model with attention mechanism | Efficiently processes long and complex texts | Requires large datasets and fine-tuning |
| Gan et al. (2024) | LLMs for resume screening | Improves match precision using semantic alignment | High resource requirements and tuning complexity |
| Moore et al. (2023) | Analysis of human bias in screening | Demonstrates consistency of AI vs. human shortcuts | Observational; no system proposal |
| Feldman et al. (2015) | Fairness auditing and bias detection techniques | Provides tools to detect and reduce bias in automation | Not resume-specific; general bias methods |
| Chen et al. (2020) | Explainable AI using SHAP and LIME | Builds trust by explaining AI decisions to users | Requires additional tools and explanations |
| EEOC (2023) & ICO (2023) | Legal frameworks for data protection & fairness | Ensures GDPR and EEOC compliance in recruitment AI systems | Regulatory in nature; not technical implementation guides |

## **7.3 Review Summary**

In justifying the value and viability of AI-driven resume screening websites, the literature review reveals significant technology, ethical, and legal issues. Particularly, our research uses explainability elements, transformer natural language processing, and fairness audits to counter the absence of semantic context and fairness perceptions in most early systems. Our designed AI-driven recruitment platform is responsibly and practically constructed using the cited models and rules providing guidance and encouragement.

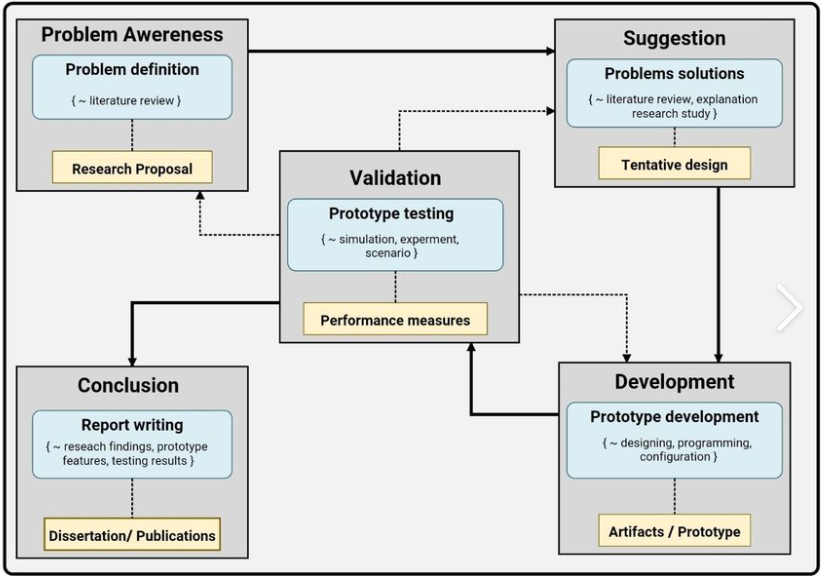
# **8. Research Methodology**

## **8.1 Research Framework**

While creating new IT artifacts with the aim of addressing actual problems, the computer sector greatly favors the Design Science Research (DSR) paradigm, which is utilized in this research. For issue-solving design, conceptualization, evaluation, and iteration, the DSR paradigm is founded on iteration cycles. An AI-based resume screening software, enhancing and streamlining shortlisting of candidates within HR procedures through machine learning and natural language processing will be the project deliverable.

The research begins with the identification of the primary problems with conventional resume screening, including inconsistency, human bias, and inefficiency. System requirements will be identified by carrying out a comprehensive stakeholder analysis and literature review. The AI screening prototype, using machine learning algorithms to grade the candidates on job-specific criteria and natural language processing (NLP) techniques to screen resumes, is created and implemented in the second step. Ongoing feedback from potential end users, repeated testing and evaluation cycles, and simulated testing (using tutors or HR professionals as stakeholders) will all be used to improve the system.

When the DSR process reaches its end, there is also an evaluation phase. Feasibility and performance of the developed artifact will be assessed via quantitative measures (e.g., accuracy classification and AUC score) and qualitative attributes (e.g., fairness, explainability, and usability). The research design ensures that the overall system considers organizational, ethical, and legal considerations in addition to technology performance requirements.



## **8.2 Data Collection & Analysis**

Secondary data from sources like LinkedIn (e.g., publicly published job postings for skill-job matching), GitHub (e.g., job description corpora), and Kaggle (e.g., resume and HR analytics data) will be used in this research. These data are created synthetically or published under public license, thus being morally acceptable as per data usage norms.

Preprocessing of text data will be done using traditional NLP techniques. They are:

* Tokenization of text is the operation of splitting text into tiny portions, i.e., words or phrases.
* Through word-splitting to root words, stemming and lemmatization
* Detection of organized data, i.e., persons, organizations, titles, capabilities, and geopositions, is called Named Entity Recognition (NER).

Following pre-processing, data will be vectorized to store numerical representation of resumes and job postings employing techniques such as TF-IDF or BERT embeddings. Various machine learning algorithms such as Random Forest, Support Vector Machines (SVM), and Logistic Regression will be trained and tested in order to sort resumes on semantic similarity to job requirement.

Performance will be ranked according to the following:

* **Accuracy:** The proportion of data classified appropriately.
* **Tags:** A set of labels to which the input pertains
* **Confusion matrix:** A graphical plot of actual positives vs. false positives and actual negatives vs. false negatives
* Strength of multiple threshold values is expressed in terms of **Area Under Curve (AUC)**
* **Fairness metrics**: Bias detection through metrics that compute equal opportunity, demographic parity, and disparate impact

## **8.3 Social, Legal, and Ethical Issues**

This project addresses the most significant ethical and legal issues of AI-based resume screening in detail. There will be fairness validation at various stages like data preparation, training, and prediction to prevent propagating bias in previous hiring data. Fairness constraints will be incorporated in the model to address gender, ethnicity, and other protected features in a balanced way. Sensitive information will be stripped out in training and testing, avoiding unnecessary data usage, and anonymizing personal data, so the system will be legally compliant under GDPR It will also be consistent with transparency principles by providing clear reasons for its options.

By elucidating the decisions to HR professionals with techniques such as SHAP and LIME, the project uses Explainable AI (XAI) methods to determine the user's confidence. For accountability and post-hoc analysis, audit trails and logging will also be made available. Maintaining the system transparency, observance of law, and moral responsibility, the holistic approach offers businesses with tangible benefit and treats applicants with equity.

# **9. Findings and Future Directions**

The objective of this project is to develop a functional prototype of an AI-based resume screening system based on the use of NLP and ML for auto-scoring and auto-extracting candidate details. It would lower human effort and prejudice in hiring by increasing matching of resumes with job descriptions. Use traditional methods such as Named Entity Recognition and deep architectures presented in recent research, the system's performance will be evaluated based on accuracy, fairness, and explainability measures. Increased hiring equity is one of the significant expected results. The initiative employs open AI techniques like SHAP and LIME, and fairness audits to avoid prejudice from becoming an entrenched characteristic of AI models. Such methods, which are transparent and allow HR professionals to view why AI is making specific decisions, address the problems of trust with earlier black-box systems. Additionally, the system will excel over conventional keyword-based applicant tracking systems using context-aware models such as BERT to analyze resumes more accurately. This allows for a better comparison between the candidate's skills and needs. The software is also GDPR compliant and has fairness testing to fill gaps of legal and ethical requirements that existing AI recruitment technology often overlooks.

## **9.1 Future Directions**

Future real-time hiring can be enabled by integrating the technology with for-profit HR systems such as Workday or SAP SuccessFactors. It can enhance the evaluation of candidates by providing a more holistic view beyond purely technical abilities by searching for soft skills and sentiment in cover letters.

In addition, the program could be modified to scan resumes in foreign languages, which would prove useful for international job postings. The study lays a solid groundwork for ethical AI recruitment by giving precedence to flexibility, fairness, and explainability. In addition, it opens doors for the potential for future research and integration of commercial HR technology.

# **10. Self-Reflection**

I gained excellent technical skills in NLP, machine learning, and data preparation while working on this project, particularly in the application of these techniques in resolving HR problems. In technique learning and design, I also gained improved knowledge in model explainability, bias detection, and ethical AI. Project management-wise, I gained skills in risk assessment and academic documentation and learning the skill of planning, scheduling, and coordinating work using tools like WBS and Gantt charts. Eventually, I hope to complete my degree, gain experience in my line of work through an internship, and pursue a master's degree in AI or in data science. By making equity, openness, and ethical use of smart systems synonymous with my priorities, I hope to be a true specialist in the area of AI for good. I hope to be able to have a career where I can use technical competence and sensitivity to ethics in designing AI systems that positively impact society.

# **11. Conclusion**

This project proposes the development of an AI-based resume screening system using machine learning (ML) and natural language processing (NLP) that can enhance the efficiency, equity, and accuracy of hiring. By offering a data-backed, automated system, it addresses most of the problems in conventional hiring, such as bias, inconsistency, and the time element.

In addition to being technically and morally sound, the system will also be GDPR and EEOC compliant. Literature identifies fairness and transparency barriers and calls for industry AI in recruitment. By combining explainable AI functions with bias identification, the initiative is intended to close such gaps. Future research can delve deeper into soft skill analysis, multilingual resume parsing, and communication with HR platforms—ultimately leading to more intelligent and unbiased recruiting systems—although effectiveness can be data-diversity and test-scope dependent.

# **12. References**

1. Li, S., Li, K. and Lu, H., 2023. *National origin discrimination in deep-learning-powered automated resume screening.* arXiv preprint arXiv:2307.08624.
2. Amin, S., Jayakar, N., Sunny, S., Babu, P., Kiruthika, M. and Gurjar, A., 2019. Web application for screening resume. In: *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC).* Erode, India, 27–29 March 2019. IEEE, pp.135–140.
3. Priyanka, J.H. and Parveen, N., 2024. DeepSkillNER: An automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. *Multimedia Tools and Applications,* 83(1), pp.1–20.
4. Gan, C., Zhang, Q. and Mori, T., 2024. A novel framework for automated resume screening using large language models. *Journal of Information Processing,* 32, pp.881–894.
5. Moore, O.A., Livingston, B. and Susskind, A.M., 2023. Résumé screening heuristic outcomes: An examination of hiring manager evaluation bias. *Equality, Diversity and Inclusion: An International Journal,* 42(1), pp.21–38.
6. Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.* Minneapolis, Minnesota, 2–7 June 2019. ACL, pp.4171–4186.
7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I., 2017. Attention is all you need. In: *Advances in Neural Information Processing Systems (NeurIPS).* Long Beach, California, 4–9 December 2017. pp.5998–6008.
8. Feldman, M., Friedler, S.A., Moeller, J., Scheidegger, C. and Venkatasubramanian, S., 2015. Certifying and removing disparate impact. In: *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* Sydney, Australia, 10–13 August 2015. ACM, pp.259–268.
9. Chen, I., Szolovits, P. and Ghassemi, M., 2020. Can AI help reduce disparities in general medical and hiring decisions? *BMJ,* 368, p.m492.
10. Ribeiro, M.T., Singh, S. and Guestrin, C., 2016. Why should I trust you?: Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* San Francisco, California, 13–17 August 2016. ACM, pp.1135–1144.
11. Information Commissioner’s Office (ICO), 2023. *Guide to the General Data Protection Regulation (GDPR).* [online] Available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/> [Accessed 24 Mar. 2025].
12. U.S. Equal Employment Opportunity Commission (EEOC), 2023. *Employment Laws and AI.* [online] Available at: <https://www.eeoc.gov/> [Accessed 24 Mar. 2025].