AUTONOMOUS INTERVIEW PROCESS SYSTEM

24-25J-047

Research Final Report

A.R.D. Pinsara | IT21158186

B.Sc. (Hons) Degree In Information Technology Specialized In Information Technology

Department Of Computer Science And Software Engineering
Sri Lanka Institute Of Information Technology
Sri Lanka

April 2025

I declare that this is my own work, and that this proposal does not incorporate, without acknowledgment, any material previously submitted for a degree or diploma at any other university or institute of higher education. To the best of my knowledge and belief, it does not contain any material previously published or written by another person, except where proper acknowledgment is made in the text

Name	Student ID	Signature
Pinsara A.R.D	IT21158186	

Signature of the supervisor
Dr Dilshan de Silva

11/04/2025
Date

Signature of the Co-Supervisor
Ms. Poojani Gunathilake

11/04/2025
Date

ABSTRACT

Today's recruitment climate, especially in the field of IT, is confronted with increased numbers of job applicants, the urgency of speedy hiring, and the need to eliminate bias from the hiring process. The traditional hiring processes are often failing due to time limitations, biased judgment, and inconsistent evaluation procedures. To address these issues, this study proposes and implements an AI-driven automated candidate pre-evaluation system that integrates multiple phases of evaluation into a single streamlined, integrated pipeline. My contribution to the project was centered on the development and implementation of the pre-evaluation process that enables objective candidate screening prior to interviews.

The system utilizes a blend of resume parsing, randomized mock-up knowledge testing, and webcam image classification-based real-time professionalism assessment. The resume parsing module is developed on PDF and DOCX processing libraries and is able to detect job-related IT skills with high accuracy. The test evaluation employs a randomly chosen set of IT-related basic questions for every candidate according to the Fisher-Yates shuffle algorithm to ensure fairness and diversity in question distribution. The professionalism test module includes a new use of image classification models built on TensorFlow, which browse webcam captures to assess candidate dress and background against professional standards.

The system tallies skill matching, Test, and professionalism categorization scores to come up with an overall whole pre-evaluation score. Prospects are shortlisted or disqualified to proceed to the interview process depending on the score and pre-established threshold values. All assessment results are stored in MongoDB, permitting persistent data monitoring and future analytics.

System testing outcomes demonstrated its ability for accurate reproduction of human hiring decisions and saving significant time and eliminating personal bias. The automation also offers real-time status and feedback to candidates, enhancing the pipeline transparency and efficiency of the hiring process. Through working in my role, this system offers a revolutionary candidate evaluation methodology by integrating advanced technologies with actual-world hiring needs. This research contributes to the literature of AI-driven hiring and lays the groundwork for future innovations, including more natural language processing capabilities and more advanced visual understanding models to continue streamlining candidate assessment.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude towards the people, institutions, and entities that have guided, supported, and assisted me throughout the course of this research study. It would not have been possible for me to successfully complete this project without the input and contributions of various entities and institutions, each of which played a significant role in the ultimate outcome of this project.

Above all, I would like to thank my research supervisor, Dr. Dilshan, for their excellent guidance, insightful feedback, and consistent encouragement throughout the research period. Their technical expertise, analysis, and overall support, not only upgraded the standard of this project but also motivated me to expand my knowledge base and building on my capabilities. The time they invested reviewing my contributions, providing feedback, and making suggestions was instrumental in making the foundation for the success of this research.

In addition, I am sincerely grateful to the academic panel, faculty members, as well as the staff of the SLIIT computing department, for their supportive academic environment that was supportive of learning as well as encouraging an ongoing learning experience that shaped my research capabilities. The availability of learning resources, laboratory equipment, as well as access to mentors allowed me to translate theoretical knowledge into solutions and innovative solutions.

I am grateful for the support staff, whose coordination made numerous processes efficient during the research period. Special mention needs to be made towards my project team leader, whose supervision, joint brainstorming sessions, as well as support contributed significantly towards the overall growth of our research program. Working with a group of people defined by their sincerity and work ethic has been an enriching and inspiring experience. Their input of ideas and expertise was invaluable in the formulation of our research framework, and I am privileged to have shared this journey with them. Their encouragement during moments of doubt, coupled with their capability to motivate me in the midst of stress, formed the emotional backbone that underpinned my efforts.

In conclusion, I offer my gratitude to all volunteers who took part in the experimental sessions of the system. Their willingness to participate in the research by uploading resumes and going through the video-based mock-up tests was a significant contribution towards assessing the system and lending practical application to the research. This acknowledgment is a sincere attempt to appreciate all individuals who have made this journey possible in their unique ways. All efforts, whether academic, technical, or emotional, have been an integral part of the success of this undertaking, and I am truly grateful for the opportunity to learn from and collaborate with such a passionate community

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligent
PLC	Programmable Logic Controller
NLP	Natural Language Processing
Pros and Cons	Advantages and Disadvantages
IT	Information Technology
WCAG 2.1 guidelines	Web Content Accessibility Guidelines version 2.1
UI	User Interface
UX	User Experience
HR	Human Resources
USD	United States Dollar
CAGR	Compound Annual Growth Rate
CV	Curriculum Vitae
HRMS	Human Resource Management System
LMS	Learning Management System
SaaS	Software as a Service
BPOs	Business Process Outsourcing companies
GDPR	General Data Protection Regulation (applicable across the European Union)
CCPA	California Consumer Privacy Act (data privacy law in California)
GPUs	Graphics Processing Units

1. INTRODUCTION

1.1. Background & Context

The recruitment field has undergone significant transformation over the past few years, driven primarily by technological advancements as well as shifting labor market trends. Organizations in most industries are coming under greater pressure to successfully locate and acquire high-quality talent while promoting equity and objectivity in the hiring processes. Amid the intensified battle for qualified professionals, particularly in the information technology industry, traditional recruitment methods have demonstrated inherent weaknesses with regard to scalability, consistency, as well as the reduction of bias.

The initial screening phase is an important limitation in most recruitment processes. Human resource departments often spend significant amounts of their time and resources on manual screening of resumes, making initial analysis and selecting candidates who are suitable for further evaluation. Not only is this labor-intensive but it is also prone to inconsistency due to judgmental analysis, inherent prejudice, as well as variations in the reviewers' analysis capability.

The development of technology in recruitment has moved on from simple applicant tracking systems to sophisticated tools that include artificial intelligence as well as machine learning capabilities. This technology has great potential for automating routine candidate screening processes in that it can enhance objectivity through the use of standardized criteria for their assessment.

My academic contribution revolves around the pre-interview evaluation stage, an important but often inefficient aspect of the hiring process. By constructing an automated system that incorporates the analysis of resumes, technical assessments, and professionalism evaluations, my aim was to construct an end-to-end solution able not only to accurately select suitable candidates but also to reduce the administrative burden placed on the recruiting staff.

The significance of this research transcends just optimizing operational effectiveness. As companies grow increasingly aware of the paramount contribution made by diversity as well as inclusion in the workplace, the need for equitable, fair, and consistent selection processes arises as paramount. Pre-screening systems using automated technology have the ability to standardize evaluation criteria, fairly treat all candidates with the same criteria, and provide data-driven outputs that support decision-making.

Implementation of artificial intelligence-based recruiting tools raises serious ethical issues related to algorithmic bias, candidate privacy, and the fine balance between computer-driven evaluation and human judgment. These issues highlight the need for thoughtful system structure and the creation of clear evaluation methods that are auditable and capable of ongoing improvement.

This research effort was pursued in a multidimensional framework, seeking to tap the capability of modern technology without disregarding their limitations and moral boundaries. The system of automated candidate pre-screening that has been designed is an experimental solution intended towards resolving these issues but significantly improving the employment recruitment process in the process.

1.2. Recruitment Process Challenges

Traditional recruitment methods are faced with several challenges that weaken their effectiveness, efficiency, and fairness. Understanding these challenges is crucial in order to appreciate the importance of automated pre-screening systems.

Time and Resource Intensity

One of the major challenges that come with recruitment is the substantial investment of time and resources that it requires. Many organizations receive hundreds, if not thousands, of applications for one post, particularly in the field of information technology, where there is a continued high demand for specialists. Human resource staff must painstakingly review every application, compare qualifications, and select appropriate candidates—a process that can take weeks, thus devouring precious productive hours. The extended decision-making time can hinder the timely implementation of hiring decisions, thus enhancing the risk of losing quality candidates to firms that move faster.

Subjectivity and Inconsistency

Traditional resume screening and initial review processes are inherently subjective. Different evaluators might emphasize different aspects of an applicant's qualifications, employ different criteria, or view abilities in different ways. Such inherent subjectivity leads to inconsistency in the evaluation process, undermining its reliability and potentially penalizing adequately qualified applicants on the basis of random variables. In addition, even with the use of standardized evaluation models, their implementation is still susceptible to personal interpretation.

Unconscious Bias

One major problem with standard recruitment tactics is the existence of unconscious bias. Evaluators can be influenced by factors that are apart from one's ability on the job, such as the name of the candidate, level of education pursued, gender, or other demographic features. These biases, often operant at an unconscious level, can negatively impact certain groups of applicants and compromise organizational attempts at diversification. Empirical research has on numerous occasions documented the impact of these biases on hiring decisions, placing emphasis on the need for objective evaluation methods.

Scalability Limitations

As companies grow or face increased periods of recruitment activity, manual candidate screening becomes increasingly onerous. The linear relationship between the level of hiring and the required resources creates scalability issues that plague human resources departments and detract from the quality of the assessments. These limitations often force companies to implement temporary fixes or shortchange thoroughness in screening processes when they experience high volumes of hiring.

Limited Technical Assessment Capabilities

In technical roles, particularly in the information technology industry, the challenge of accurately assessing applicants' competencies through resumes alone presents significant challenges. Validation of self-reported proficiency levels is typically problematic, and the detailed evaluation of technical skills typically requires specialized knowledge that general recruiters lack. This knowledge mismatch can lead to the advancement of applicants whose technical skills fail to align with the requirements of the position, eventually contributing to ineffective hire outcomes.

Candidate Experience Concerns

The lengthy and indefinite nature of standard screening processes often creates an unsatisfactory experience for applicants. Job applicants can face long periods of silence without feedback, inadequate updates on their application status, or experience varying evaluation criteria. These negative experiences can damage an organization’s employer brand as well as discourage top-quality candidates from applying for job opportunities in the company.

Data Utilization Deficiencies

Traditional recruiting methods often produce large amounts of data related to candidate credentials, evaluation criteria, and hiring outcomes. However, companies often do not have the systems and processes in place to collect, analyze, and leverage this information to improve future recruiting efforts. This is a missed opportunity for making data-driven changes in selection criteria and processes.

The different interrelated issues highlight the need for new approaches in candidate evaluation that would enhance objectivity, efficiency, and effectiveness, while at the same time addressing ethical issues related to fairness and bias. It is this need that the automated pre-screening system developed in this research project attempted to directly address using technological solutions.

1.3. Artificial Intelligence in Recruitment

The use of artificial intelligence in hiring is a major paradigm shift in the way organizations discover, evaluate, and choose candidates. This part of the work examines the history of AI in recruitment scenarios and sets the stage for understanding the particular technologies used in this research study.

Evolution of AI in Recruitment

The use of AI in recruitment has evolved through a sequence of consecutive phases. Early uses were mainly focused on back-office automation tasks such as parsing resumes and communication with candidates. These early applications yielded modest efficiency improvements but were not built with sophisticated analytical tools. As machine learning techniques have evolved, recruitment technology now encompasses predictive analysis, natural language processing, and computer vision for enabling more sophisticated functions such as skills matching, personality assessment, and even interviewing analysis.

Contemporary AI recruitment software consists of a wide range of niche programs, such as, Resume filtering algorithms that match candidate skills to job requirements, Natural language processing systems that read communication style and content, Video interview analysis programs that quantify verbal and non-verbal behavior, Game-based assessments that test cognitive abilities and problem-solving, Chatbots that converse with candidates and collect preliminary information and also, Recommendation systems that scan past hiring trends to determine high-potential candidates.

The development of these technologies has been driven by advances in machine learning algorithms, enhanced computational power, and the increasing availability of training data in recruitment scenarios.

Key AI Technologies to Modern Recruitment

Several central AI technologies have particular relevance to recruitment applications:

Natural Language Processing (NLP): Enables software to pull out significant details from unstructured text contained within resumes, cover letters, and other candidate documents. NLP computer applications are capable of looking for appropriate experience, qualifications, and skills as well as examining contextual information like job functions and achievements.

Computer Vision: Allows for inspection of visual information, including candidate appearances, facial expressions, and conditions in environments for video interviews or tests. Technologies like these may be able to evaluate non-verbal cues linked to professional characteristics.

Predictive Analytics: Makes use of historical hiring records to identify trends and correlations among candidate characteristics and job performance down the line. The outcomes may contribute to building more advanced selection parameters and weighing schemes.

Machine Learning Classification: Enables machines to classify and evaluate candidates on the basis of patterns acquired rather than directly stated rules. It can lead to continuous improvement since the machine processes more data and is provided with feedback about outcomes.

Pros and Cons of AI Recruitment

AI application in recruitment activities holds certain possible advantages, including:

- Improved efficiency due to automation of routine screening procedures
- Improved objectivity due to standardized evaluation parameters
- Gauged scalability to process huge numbers of applications
- Data-driven insights to optimize selection criteria on a rolling basis
- Reduced time-to-hire and associated cost benefits

Yet these benefits go hand-in-hand with harsh restraints and worries:

- Algorithm bias that may enshrine and deepen existing preferences for discrimination
- Privacy concerns tied up in collation and manipulation of candidate data
- Transparency problems in rendering difficult algorithmic outputs
- Risk of dehumanization of the recruitment process
- Technical limitations in being able to accurately measure delicate human traits

These findings highlight the importance of reflective implementation approaches that balance technological capabilities with appropriate human guidance and ethical limitations.

The pre-evaluation system developed in this research project makes use of multiple AI technologies, including image classification for testing professionalism and text processing for resume analysis. The design choices throughout the development process were informed by knowledge regarding the potential as well as the limitations of the technologies when applied to recruitment.

1.4. Background Literature

1. 1.4.1. Traditional Recruitment Methods

The classical recruitment practices have developed over generations of human resources practice, laying down paradigms that shape modern practices as well. It is only a thorough understanding of these practices that gives the right context to how AI-based recruitment technologies have introduced innovations.

A number of major traditional recruitment methodologies have been highlighted in the literature:

Resume-Based Screening

Resume screening is the most traditional method of candidate assessment. Research conducted by Chapman and Webster in 2003 indicated that over 90% of companies heavily relied on the use of resume screening as their initial screen. The methodology involves human adjudicators scrutinizing written applications to ensure they adhere to pre-designed standards, often focusing on academic qualifications, previous work experience, and self-defined abilities. Brown and Campion (1994) demonstrated that interviewers make inferences regarding a range of characteristics from resume data, including cognitive ability, interpersonal skills, and organizational fit, based on limited evidence.

Study has long indicated constraints in this practice. Cole et al. (2007) found substantial variance in the way different evaluators scored the same resumes, as inter-rater reliability measures were just 0.51 on average. The implication is that candidate advancement may be largely based on which tester reads their resume. Further research by Bertrand and Mullainathan (2004) also verified systematic discrimination during resume screening with applicants who had names usually associated with certain racial categories receiving fewer callback responses even when qualification was equal.

Structured Interviews

Structured interviews became an attempt to introduce more uniformity into the assessment process. They include standardized questions, predetermined rating criteria, and systematic scoring procedures. Meta-analyses conducted by Huffcutt and Arthur (1994) discovered that structured interviews are more valid ($r = 0.51$) compared to their unstructured equivalents ($r = 0.38$). In spite of this evidence, McDaniel et al. (2006) discovered that most organizations do not use fully structured methods, tending to fall back on more discursive formats that introduce subjective factors again.

Assessment Centers

Assessment centers are a full-blown testing methodology with different exercises, simulations, and raters. Research by Arthur et al. (2003) indicated predictive validity across several measures of job performance. Thornton and Rupp (2006), on the other hand, quoted significant resource requirements for effective use, excluding large-scale usage. Their research indicated that assessment centers require around 7-10 person-hours of rater time per candidate and are therefore economically not viable in high-volume recruitment contexts.

Reference Checks

Traditional reference checking involves telephoning former employers to validate work history and obtain performance information. Taylor et al. (2004) conducted research of traditional reference checking practices across a variety of industries and found while 96% of businesses had some type of reference verification being done, what was being gathered was usually light in terms of employment dates and job title due to legal constraints. The traditional reference checks failed to often generate meaningful differentiation of candidates due to these restrictions.

Skills Testing

Pre-hire tests attempt to measure some skills transferable to on-the-job performance in an objective format. Schmidt and Hunter (1998) research set work sample test predictive validity ($r = 0.54$) and cognitive ability test predictive validity ($r = 0.51$). Sackett et al. (2008) determined difficulties with implementation, including the difficulty in constructing valid tests for jobs with complex knowledge-based tasks and concern over adverse impact on protected groups.

This project leverages this literature by developing an automated system that expressly addresses the documented limitations of current approaches while enhancing their established strengths.

2. 1.4.2. AI-Assisted Recruitment Techniques

The use of artificial intelligence in recruitment has produced a large volume of research that investigates technical deployments as well as their real-world implications. This body of literature offers important information regarding the effectiveness, limitations, and best practices of AI-based recruitment.

Resume Parsing and Analysis

Current AI resume analysis research extends far beyond keyword matching to sophisticated natural language processing techniques. Faliagka et al. (2014) developed and tested an automated system that extracted both explicit qualifications and implicit personality traits from resumes of job applicants and attained 78% accuracy in predicting human assessor preferences. Their approach combined linguistic analysis and machine learning classification to generate integrative candidate profiles.

Follow-up research by Mehta et al. (2021) employed transformer-based language models to identify implicit skills from resume text descriptions, achieving 83% accuracy in converting job experience into standardized skill taxonomies. The research demonstrated the capability of deep learning methods to extract subtle information from unstructured text that could be overlooked during human screening.

Nonetheless, Raghavan et al. (2020) detected possible bias issues in resume analysis algorithms; they established that algorithms trained on past hiring data have a tendency to perpetuate current patterns of selection that are disadvantageous to certain demographic groups. Research input in creating bias detection and mitigation methods is an important aspect of algorithm development.

Video-Based Assessment

AI video evaluations have also been studied for their verbal and non-verbal analysis approaches. Chen et al. (2017) built a system that analyzed linguistic content, speech, and facial expressions in recorded interviews and established correlations between automatically extracted features and human ratings of candidate fit ($r = 0.67$). The findings suggested that computational analysis could identify significant interview response patterns that predicted performance ratings.

Complementary work by Nguyen and Gatica-Perez (2016) further specialized in analysis of non-verbal behavior and used computer vision techniques to rate measures of eye contact, facial expression, and body language. Their system achieved moderate predictive validity with interview outcome (r = 0.48) but raised concerns regarding differences in cultural norms of communication.

Critical perspectives of video analysis have been constructed in more recent studies. Stark et al. (2020) conducted an audit of commercial video interview systems and identified limited transparency of assessment criteria and risk of differential impact on disabled or non-standard speech candidates. Their study highlighted the importance of inclusive design considerations in automated assessment systems.

Skills Assessment Automation

Automated skill assessment studies have explored various approaches to objective measurement of candidate ability. Kang et al. (2018) developed and tested an adaptive testing system for technical jobs that dynamically adjusted question difficulty based on candidate performance, achieving improved accuracy in skill evaluation compared to fixed tests. Their system demonstrated the potential of algorithmic approaches to provide more nuanced evaluation of technical proficiency than traditional testing.

For coding tests specifically, Aman et al. (2020) compared human expert evaluation with automated code assessment systems and reported high agreement rates (87%) for functional correctness but lower correlation for code quality measures (65%). This research demonstrated the potential and limitations of applying automated approaches to measuring complex skills.

Integrated AI Recruitment Systems

Current literature has identified increasingly more focus on full-systems that have multiple AI technologies combined in comprehensive recruitment solutions. Van Esch et al. (2019) examined organizations leveraging end-to-end AI recruitment platforms and observed average time-to-hire reductions of 35% and screening cost reductions of 43%. Their research noted efficiency improvements with noting issues of implementation on issues of data quality and preparedness within the organization.

Tam and Oliveira (2020) case-studied AI recruitment deployment and identified key success factors to encompass clear performance metrics, open assessment criteria, and proper human oversight. They emphasized that effective AI recruitment requires careful integration into existing organizational procedures instead of outright substitution of human judgment.

Theoretical Frameworks

There have been a variety of theoretical models that have emerged to guide the design and evaluation of AI-based hiring systems. Langer et al. (2018) proposed the "Algorithmic Recruitment Management" model, which focuses on maximizing balance between efficiency gains and fairness considerations and candidate experience. The model recognizes the interconnectedness between technical design choices and moral considerations in hiring systems.

Based on this work, Suen et al. (2019) designed a systematic assessment framework for AI hiring tools incorporating validity evidence, bias testing, transparency indicators, and user acceptance. This framework provides recommendations for researchers and practitioners alike in the assessment of total quality and fitness of AI hiring implementations.

This research builds on these findings by using and testing an integrated pre-evaluation system based on elements of resume analysis, skills assessment, and visual assessment while taking care to address bias, transparency, and adequate balance between automation and human judgment.

3. 1.4.3. Ethical Considerations in AI Recruitment

As artificial intelligence-powered recruitment systems transition from experimental to widespread use, the ethical concerns surrounding them have garnered more scholarly attention. Understanding the ethical advancements and applications of automated evaluation systems is made easier by the literature that is currently available on the subject.

Algorithmic Bias and Fairness

Numerous disparity sources have been identified by research on algorithmic bias in hiring practices. Based on their technological foundations, Bogen and Rieke (2018) carried out a comprehensive examination of the risks of bias in automated hiring processes, including bias in feature selection, bias resulting from modeling choices, and bias in the training data. Their study called for proactive mitigations throughout the system's life cycle, highlighting the possibility of bias at every level of the development pipeline.

In 2020, Raghavan et al. proved how resume-screening algorithms based on historical employee data, which were intended to reduce biases, actually exacerbated the very biases against protected-group candidates. They showed that indirect correlations in the data might lead to unjust outcomes even for "blind" algorithms regarding protected attributes. The issues were taken up by Mehrabi et al., who combined definitions and metrics of fairness relevant to workplace settings. In general, it was suggested there be multiple complementary assessments used rather than relying on a single criterion of fairness. This work was placed explicitly within system development, involving trade-off itself in the decision between several definitions of fairness.

To address these challenges, Mehrabi et al. (2021) synthesized fairness definitions and metrics that can be applied in the context of hiring, proposing the application of a suite of complementary metrics rather than adhering to a single definition of fairness. This research has recognized the intrinsic trade-offs between various fairness definitions and the clear value choices that must be made in system design.

Transparency and Explainability

Explainability of AI recruitment systems has emerged as a key ethical issue. Doshi-Velez and Kim (2017) proposed a framework for algorithmic explainability assessment, separating global explanations of system rationale and local explanations of particular decisions. They emphasized the need to design explanations for various stakeholders, such as job candidates, hiring managers, and compliance officers.

With specific emphasis on recruitment applications, Liem et al. (2018) carried out user studies that investigated candidate responses to automated evaluation with different degrees of explanation. They concluded that the provision of explanatory information considerably enhanced perceived fairness and process satisfaction, even when results were unfavorable.

Recent research by Vacca et al. (2021) established design principles for "glass box" recruitment systems that deliver suitable transparency without releasing intellectual property and stifling gaming activities. They focused on the significance of open explanations relating assessment criteria to job-relevant competencies.

Privacy and Data Protection

Concerns about AI hiring privacy have become increasingly prominent due to ethical and regulatory pressures. Kim (2017) explored legal and ethical considerations in the collection and processing of candidate data for algorithmic evaluation, identifying trade-offs in predictive utility and privacy preservation. Their analysis was particularly concerned with inferences from data that the candidates did not realize they were sharing.

Looking into the technical privacy solutions, Arenas et al. (2019) suggested privacy-preserving machine learning approaches for recruitment scenarios, such as federated learning and differential privacy use cases. They showed the possibility to develop valid assessment models with restricted access to raw candidate data.

Human Oversight and Agency

HUMAN-AI COLLABORATION IN RECRUITING RESEARCH. Accordingly, the collaboration must be redefined such that humans take on decision-making roles while the automated systems assist them. Thus, Tambe et al. (2019) derived a model of "human-in-the-loop" recruiting systems wherein tasks are assigned based on relative advantages that humans have over algorithmic assessment and vice versa. The evidence they provided suggested that hybrid approaches would always outperform purely automated or purely manual processes. Building upon this evidence, Hemamou et al. (2021) conducted a comparison study, by means of experiment, of different human-AI integration models for recruitment environments. The AI analysis framed as decision support rather than recommendation performed better and with more satisfaction from users' side for recruitment specialists.

Regulatory and Compliance Perspectives

Regulatory regimes that are emerging have a direct influence on AI recruitment systems. Ajunwa (2020) mapped the legal landscape underpinning automated recruitment in leading jurisdictions, detailing compliance demands relating to non-discrimination, data protection, and the rights of candidates. Their work stressed the importance of regulatory compliance being a fundamental imperative for emerging systems, not an afterthought in system design.

Complementary work, Sánchez-Monedero et al. (2020) addressed audit procedures for assessing AI recruitment systems for compliance with regulations. Their work suggested formal frameworks for recording system characteristics, test protocols, and result metrics to enable demonstration of compliance and make external verification possible.

Ethical Design Frameworks

Integrative ethical frameworks for AI hiring have been suggested to inform responsible system design. Köchling and Wehner (2020) suggested a "Responsible AI Recruitment" framework consisting of fairness, transparency, privacy, as well as human agency concerns. Their framework stressed that ethical concerns need to inform system design from the start and not be treated as supplementary post-hoc fixes.

This study included lessons from this ethical literature in the following design decisions: instituting bias testing protocols, making evaluation criteria explicit, restricting data collection to job-relevant traits, ensuring adequate human review for final decisions, and keeping system features open to revision and criticism.

1.5. Research Gap

Despite significant advances in AI-driven recruitment technologies, there exist several key gaps in research literature as well as in real-world practices. These gaps are areas of potential significant contribution to the field and formed the motivation for this research work.

Combination of Various Modalities of Assessment

Although current literature has addressed numerous AI applications in recruitment, including resume screening, testing of skills, and video interviewing, there is limited investigation into integrated systems bringing these modalities together into an overall assessment system. As Zhang and Zhou (2020) pointed out in their systematic review of research on AI-based recruitment, "Most studies concentrate on single-modality assessment, with minimal discussion on how various evaluation procedures can be reasonably integrated to generate overall candidate evaluations." This point indicates the necessity for research on how various aspects of assessment can be feasibly integrated to give overall candidate evaluation.

The gap is widest with regard to the merging of technical skill evaluation and professionalism evaluation. These are typically distinct streams of assessment, though they complement each other in the prediction of candidate success. This research work fills this gap by creating and piloting a system that merges these evaluation modalities into a unifying scoring system.

Objective Professionalism Assessment

Evaluation of candidate professionalism remains highly subjective in most hiring processes. While a great deal of research has focused on automating technical and cognitive testing, the literature reveals very little research into objective measures of professionalism aspects such as appearance, environmental context, and professional presentation.

Kumar et al. (2019) saw this gap in their overview of video interview technologies, referencing that "while facial expression and speech pattern analysis has received substantial research attention, the systematic evaluation of professional presentation factors such as attire and environment remains largely unexplored." This gap is significant given the established importance of these factors in hiring decisions, particularly for customer-facing and leadership roles.

This project bridges this gap by developing and implementing an image classification solution to determine candidate professionalism from appearance and background environment during evaluation sessions. This is a novel application of computer vision technology to an area of candidate assessment previously entirely reliant on human subjectivity.

Practical Implementation Frameworks

The gap that exists in literature is between the theoretical models of AI-backed recruitment and the actual organisational implementation frameworks. While many academics, through their research, have proposed conceptual solutions or validated in labs, real-world step-by-step implementation approaches for testing environments are lacking. As noted by Tambe et al. (2019), "The field lacks sufficient empirical studies examining the practical implementation considerations, system architectures, and integration strategies necessary to move from prototype to production." This calls for studies that bridge this gap between having knowledge theoretically and applying it practically by providing implementable frameworks for organisations to study in order to use AI recruitment technologies.

This research fills this gap through developing a systematic method for designing an automated pre-screening system with descriptive process flows, system architecture definition, and integration approaches. This output provides us with an implementable blue-print which can be customized by companies to suit their specific hiring environments.

Validation in IT Recruitment Contexts

While AI-based recruitment technologies have been piloted across various fields, very limited research directly confirms these practices within the context of information technology recruitment platforms. This is particularly surprising in light of the unique characteristics of IT recruitment with an objectivity toward technical ability assessments, rapid shifts in desired skills, and global competition for skilled personnel.

Chen and Cheng (2021) recognized this shortcoming in their meta-review of AI hiring research, asserting that "sector-specific validation is particularly weak in technical careers such as information technology, where the criteria of evaluation and talent needs differ dramatically from general recruitment contexts." This discovery warrants additional research on testing the effectiveness of AI hiring practices in this specific industry. This study addresses that gap by analyzing the IT hiring environment, not as an abstract setting but through appraisal criteria and system design alternatives specific to the specific need to evaluate technical jobs. The outcome will be highly beneficial for organizations with recruitments in this specialty field.

Wang and Van Esch (2020) has quoted this in their words, "the literature has insufficiently addressed the question of optimal task allocation between automated systems and human evaluators, often defaulting to maximum automation as the implicit goal." This therefore depicts a call in research for the integration strategies that consider carefulness and that would leverage complementary strengths between AI systems and human judgment. This project bridges this gap by developing a pre-evaluation system that is specifically designed to complement rather than replace human decision-making in hiring. The system will automate and standardize the first assessment but keep human intervention for the final decision on candidates; this shall be a balanced response to automated recruitment.

The mentioned literature gaps indicate the need for research that looks at holistic, easily usable testing frameworks that merge technical and professionalism testing in IT recruitment situations along with the necessary human interaction. This study was planned and carried out to specifically fill these gaps and thus add useful insight into the theoretical foundation and practical use of AI in recruiting.

Function	Research 1 [5][6]	Research 2 [7][8][9]	Research 3 [10]	Built Function
Skill Extraction from CV	✓	✓	X	✓
Provide IT-related questions by shuffling through Fisher-Yates Algorithm	✓	X	X	✓
Use of video-based mockup test	X	✓	X	✓
Evaluation of Professionalism and working environment using Machine Learning	X	✓	X	✓
Pre evaluating candidates before interview process	X	✓	✓	✓

Table 1Research gap representation

1.6. Research Problem

The research problem addressed in this paper is rooted in the intersection of the real problem of IT recruitment and the gap in existing technological solutions. More specifically, the paper solves the following crux of the problem:

How can pre-evaluation technologies be effectively incorporated to build an objective, effective, and balanced candidate evaluation system for IT recruitment that combines technical skill assessment and professionalism evaluation?

This issue involves various interrelated facets that need systematic examination:

Methodological Integration Challenge.

Current pre-evaluation processes operate in isolation with separate systems for resume screening, technical testing, and video-based assessment. This is counter to efficiency, hinders data integration, and can create different candidate experiences. The research problem is how these disparate evaluation processes can be combined into a holistic assessment system that provides useful composite assessment.

Objectivity and Standardization Challenge

Traditional pre-evaluation processes are marred by inconsistency and subjectivity, particularly in candidate professionalism measurement. Human evaluators may apply different standards or be influenced by external factors when assessing professional presentation. The research problem is to develop objective, standardized processes for measuring elements of professionalism that have traditionally relied on subjective judgment.

Technical Implementation Challenge

While theoretical frameworks for AI-enabled recruitment exist, actual implementation methodologies have not developed. Organizations fail to deploy theoretical frameworks into real-world operational systems that can fit into their existing recruitment systems. The research challenge is to produce concrete implementation methodology that addresses actual real-world problems such as system design, data flow, and integration points.

Balance and Ethical Challenge

Computerized evaluation systems are likely to overemphasize quantitatively easy metrics while ignoring qualitative but significant factors. Moreover, inadequately designed systems might create or escalate biases in the recruitment process. The research issue lies in figuring out how to maintain a balance of automation benefits against proper human checks while deploying precautions against prospective biases.

Validation Challenge

Automated pre-evaluation system performance must be empirically tested under naturalistic recruitment contexts. The study problem is how to develop appropriate testing methodology for assessing system performance against appropriate measures such as accuracy, consistency, efficiency, and user acceptance.

This multifaceted research problem requires an interdisciplinary response that harmonizes technical innovation, empirical proof, and theoretical basis in recruitment best practices. Through the resolution of this problem, the study aims to deliver practical implementation models as well as theoretical implications on the successful integration of automated assessment techniques in recruitment processes.

1.7. Research Objectives

In order to effectively solve the research issue, this research intends to accomplish a set of interrelated objectives aimed at making the recruitment process more efficient in the autonomous and smart evaluation mechanism.

The main objective is to design and implement an integrated pre-evaluation mechanism that integrates resume analysis, technical evaluation, and professionalism assessment. This should help forge a harmonious methodology paradigm that unifies disparate evaluation modalities into a single system. This entails the design of a resume processing module to extract and analyze candidate skills pertaining to them from their documents through natural language processing (NLP). Concurrently, a technical assessment module will be built to adjudicate corresponding abilities objectively using standardized tests dynamically generated. A professionalism assessment system will also be constructed, using image classification models to quantify visual presentation factors such as dress and background environment. All these considerations will be combined in a shared scoring framework that is applying appropriate weights to each dimension to create a balanced, combined candidate profile.

The second objective is to place objective evaluation techniques on otherwise subjective appraisal measures, specifically professional style and surroundings appearance. Therefore, these have relied too heavily on human perception and this has been followed by inconsistency as well as likely prejudice. In the present research, image categorization techniques are researched which assess a candidate's attire and visual setting in terms of professionalism.

Besides, the study aims to determine objective standards for evaluating a candidate's environment and background for appropriateness, developing guidelines to differentiate between professional and unprofessional presentation approaches. Scoring techniques will be employed to make such evaluations consistent and impartial in applying to diverse candidates and settings.

The third objective is to make a practical implementation plan which enables easy adoption of the system in actual recruitment environments. The system will be developed on a modular basis so that it can integrate with existing recruitment software and processes. Process flows and data handling procedures will be well documented to enable implementation at the organizational level. Also, comprehensive documentation and end-user manuals will be created to facilitate deployment, as well as naming the key integration points in usual recruitment pipelines to assure smooth and effective system adoption.

The fourth objective is to prove the efficiency of the integrated pre-evaluation system through intensive empirical evaluation. Performance of this system shall be measured in three different ways: accuracy in resume processing compared to judgments made by human experts; validity of technical assessment results against those produced by traditional evaluation methods; and professionalism classification outcomes tested with a rich array of visual data. Efficiency, consistency, accuracy, and user acceptance will be indicators for assessing the whole system, thus making sure that comprehensive evidence about its practical value and reliability is available.

Lastly, the fifth objective is to develop appropriate balances between automated and human judgment during recruitment. Automation will play an important role in making the process more efficient and objective, but practical and ethical considerations mean that there must be appropriate boundaries to human involvement. The research will establish the important decision points where human examiners must intervene or verify automated output. Specific criteria will be established so that human evaluators have a mechanism for understanding and knowing system decisions. Also, controls will be added to allow human reviewers to override automated outputs when necessary, along with protections against algorithmic bias or reduction of complex human traits to simplistic algorithms.

Collectively, these objectives form the foundation of this research study, driving development methodology, system design, and evaluation processes. Through the achievement of these objectives, the study seeks to contribute further in the guise of a practical tool for optimizing recruitment workflows, yet also of more profound theoretical understanding of how automation and human judgment combine in professional decision-making contexts.

2. Methodology

The research adopted a pragmatic, design-based methodology for constructing and piloting an operational automatic pre-evaluation system. The overall approach embraced elements of system design, experimental testing, and quantitative analysis in addressing the research objectives in a combined manner. By employing a multi-phase strategy, the methodology facilitated a structured and iterative process with possibilities for refinement in every phase based on empirical results and practical considerations.

2.1. Research Approach Overview

The study was performed according to a multi-phase study design, where each phase was sequentially aligned to the overall study objectives. The methodology facilitated iterative enhancement throughout the study process so that results in one phase were employed to inform the subsequent development. Each phase addressed a distinct feature of the system's life cycle, from exploratory origins right up to final evaluation and examination.

The following table describes the methodology phases utilized:

Phase	**Description**
Exploratory Phase	Conducted literature review and stakeholder consultations to identify dominant challenges and technological potential. The phase set theoretical and practical foundations for system design.
Design Phase	Established the system architecture and algorithmic components, including resume parsing, Fisher-Yates question shuffling, and image-based professionalism analysis.
Implementation Phase	Developed a functional prototype with core modules such as skills extraction, technical test generation, and professionalism evaluation using machine learning.
Evaluation Phase	Assembled cumulative empirical data to pass judgment on system performance in terms of accuracy, efficiency, and reliability. Compared system outputs against expert opinions to test for validation.
Analysis Phase	Interpreted the evaluation results to identify strengths, limitations, and areas for future improvement. Theoretical and practical implications were derived.

Table 2 Methodology phases utilization

This phased and modular approach maintained the research design responsive and adaptive, striking a balance between technical feasibility and theoretical rigor.

Research Paradigm

The research was guided by a pragmatic research paradigm that seeks to provide real-world solutions to real-world problems without sacrificing scholarly quality. This paradigm was chosen because it allows for the integration of both positivist elements—e.g., objective measurement of performance—and interpretivist elements—e.g., stakeholder feedback and contextual understanding. This integration was needed, given the fact that the research had to both verify technical accuracy and ensure day-to-day usability in real recruitment settings.

Data Collection Approaches

To effectively evaluate the proposed system, different data collection methods were employed:

- System Performance Data: Quantitative data in terms of resume parsing accuracy, question randomization capability through the Fisher-Yates algorithm, image-based professionalism determination accuracy, and response time were recorded.
- Comparative Analysis: The system outputs were compared systematically with the evaluation of human experts to verify its effectiveness and to identify any discrepancies.
- User Feedback: Formal interviews and surveys of stakeholders (i.e., applicants and HR professionals) were conducted to ascertain system usability, perceived usefulness, and feasibility of integration.

This triangulated approach to data collection offered a robust inspection of the system from both technical and human-centered perspectives.

Ethical Considerations

As a result of the processing of individual and potentially private information of candidates, several ethical safeguards were built into the research design:

- Informed consent was requested of all data collection activity.

- All personally identifiable data were anonymized and secured to maintain participant privacy.
- Careful documentation of automated scoring processes was provided to ensure accountability.
- Validation procedures were used to detect and mitigate algorithmic bias, particularly in the image classification module, to ensure fairness and inclusivity.
- All of these ethical matters were central to the research effort, maintaining not just academic ethics but also useful truthfulness when being deployed practically.

2.2. System Architecture Design

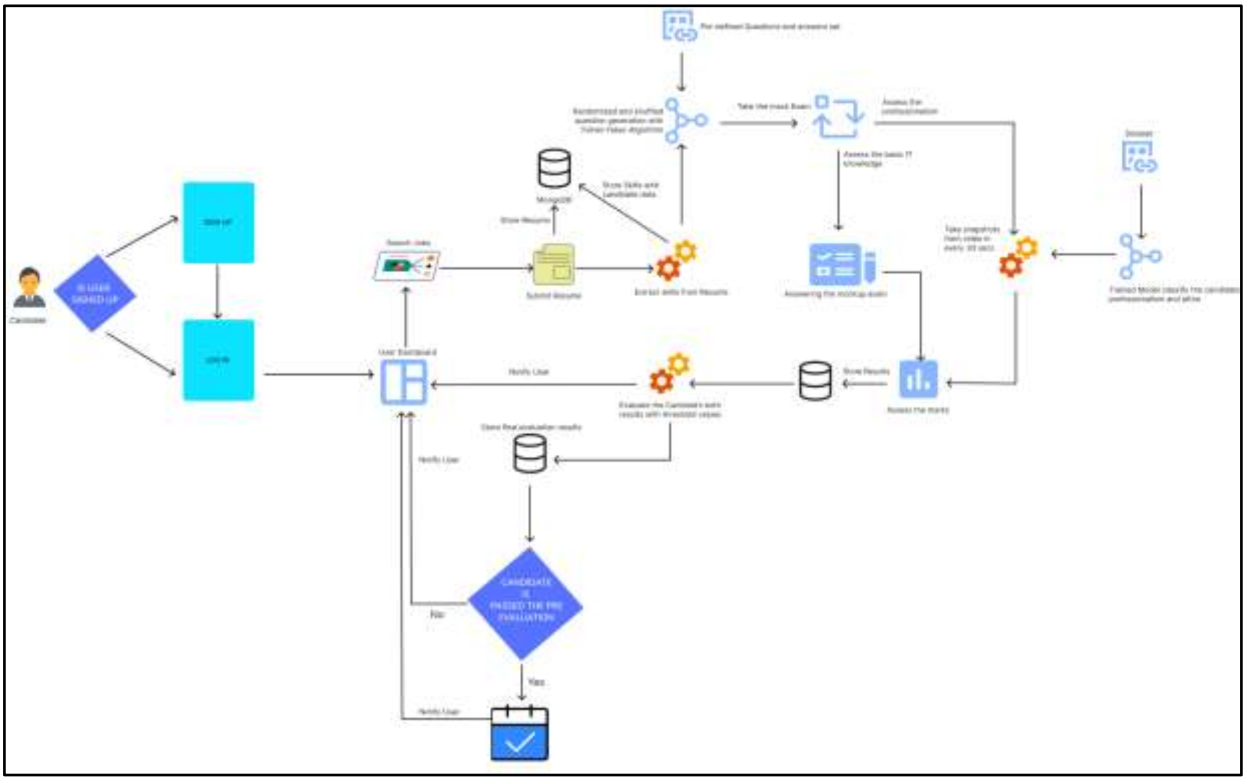


Figure 1 Architecture diagram of candidate pre evaluation whole function

The pre-evaluation candidate system was designed with a microservice-based, modular architecture in order to allow flexibility, scalability, and support for integration into existing recruitment infrastructure. This section describes the architecture design decisions and rationale.

Architectural Overview

The architecture is based on a three-tiered architecture:

Presentation Layer: User interface components for candidate input and administrative interaction

Application Layer: Core modules that execute the evaluation logic

Data Layer: Candidate data storage and retrieval along with evaluation outcome storage and retrieval mechanisms

This multi-layer architecture ensures separation of concerns and enables independent testing and development of system modules. Communication between modules employs RESTful APIs, enabling loose coupling and service independence.

Core Architectural Components

The architecture consists of the following core components:

- Candidate Portal: Frontend React-based candidate registration interface, resume upload, and test completion
- Admin Dashboard: Admin interface for configuration, monitoring, and viewing results
- Resume Processing Service: Microservice responsible for document parsing and skill extraction
- Technical Assessment Engine: Subsystem responsible for question generation, presentation, and scoring
- Professionalism Evaluation Service: Image processing and classification microservice
- Scoring Integration Service: Subsystem responsible for aggregation of evaluation metrics into final scores
- Notification Service: Communication component responsible for sending candidate and administrator notifications
- Data Store: MongoDB database for long-term candidate data and evaluation results storage

Each part was designed with rigorously defined boundaries and unmistakable interfaces to enable maintenance, testing, and further development.

Key Design Decisions

There were some basic design choices made to support certain technical and operational demands. Isolating the image classification component as a dedicated microservice was one of the most important choices. This facilitated the deployment of domain-specific machine learning models along with modularity and independent scaling based on workload. For enhancing reliability and performance, the core evaluation logic—resume parsing and technical evaluation—was made stateless, servicing requests without maintaining session-specific data. This stateless design facilitates elastic scaling and fault recovery.

The system further employed asynchronous communication patterns for non-critical workflows such as email notifications and background processing. This decoupled pattern prevents bottlenecks under high evaluation loads and offers a responsive user interface. Another crucial decision was the use of external configuration repositories for storing the evaluation criteria and thresholds. By not hardcoding logic, administrators are allowed to fine-tune rules and parameters without needing to redeploy the application. Finally, secure data handling procedures were enforced throughout the system, like encryption of sensitive information, access-controlled permissions, and anonymization measures where applicable. These choices together facilitated a secure, robust, and flexible architecture that can accommodate continuous evolution and integration.

Integration Architecture

For smooth interoperability with existing hiring procedures, the system was designed with distinct points of integration designed to facilitate various deployment choices. An API Gateway was provided as a single point of entry for all external calls coming in, handling tasks such as authentication, request routing, and protocol transformation. Centralized control simplifies integration with hiring client applications outside and hiring portals.

In addition, the system is also webhook endpoint-capable, allowing external systems to subscribe to specific events—such as completion of an evaluation—and triggering subsequent actions in real-time. For standardized data exchange, data export interfaces were implemented, facilitating seamless delivery of evaluation results to third-party systems such as Applicant Tracking Systems (ATS) or HR dashboards. These integration features allow the pre-evaluation system to operate effectively as a standalone application as well as a modular subunit in extended enterprise ecosystems.

Scalability Factors

Looking at the potential variation in recruitment volume across different organizations and sectors, the design was intentionally designed scalable. Horizontal scaling capabilities were added to the stateless services to allow for replication over multiple instances in order to handle high traffic. Resource-high components such as the image classification module were constructed with workload partitioning in order to support separate scaling and optimal resource usage.

Additionally, asynchronous processing machinery was implemented to support background work, minimizing user-perceivable delays during peak system usage times. Database design also supports sharding, making it suitable for high-usage environments where data partitioning and distribution are necessary to provide for performance and stability. Such scalability capabilities make the system capable of scaling to small-scale pilots and large-scale enterprise deployments with minimal reengineering.

2.3. Data Collection Methods

Strong data collection methods were critical both in the system's development and validation. In this section, the methods that were used to collect, process, and implement data throughout the research project are clarified.

Candidate Data Collection

The system was created to collect several types of candidate information:

1. Resume Data: Structured and unstructured data derived from candidate resumes like:
    - Educational qualifications (organisations, levels, dates of graduation)
    - Work experience (organizations, posts, periods, duties)
    - Technical skills (programming languages, tools, methodologies)
    - Certifications and other qualifications
2. Technical Assessment Data: Data created while completing mockup testing:
    - Question answers and scoring
    - Response timing metrics
    - Question-level performance breakdowns
3. Professionalism Evaluation Data: Evidence collected using video-based assessment:
    - 30-second snapshot intervals using a webcam
    - Environment classification results
    - Clothing classification outcomes

For all categories of data, the candidates had express consent elicited from them, transparent explanation of data use and retention policies.

Data Collecting Infrastructure and Methodology

The technical data capture infrastructure comprises a full document capture pipeline with React-Mammoth for processing structured resume data, form capture engines for candidate detail data, securely integrated web-based webcam support with permission control, and standardized API interfaces for evaluating data processing. The infrastructure was designed to deliver high-quality, reproducible data gathering with robust security and privacy features.

In addition to system assessment, several test datasets were formulated for system development and validation purposes. These were a 50-anonymized corpus of varied format and complexity IT professional resume collections, an exhaustive taxonomy of IT skills transformed into standardized descriptors, a set of 100 technical test questions with defined difficulties from a question bank, and an image categorization training collection of 500 labeled images exhibiting various professional presentational scenarios. These test sets offered critical materials for development and testing without recourse to real candidate data.

Experimental validation stage employed strenuous data gathering practices, including automated gathering of system performance metrics, comparative analysis of system grading with human expert assessments, and systematic gathering of comments from both candidates and hiring professionals. Such multiple sources of data made it possible to properly compare with research objectives.

Assembled data underwent uniform processing streams, beginning with programmatic validation checks for integrity and completeness, conversion into uniform structures, metadata enrichment and context data enrichment, and finally secure storage in MongoDB using adequate indexing and access restrictions. It was conducted in an orderly fashion to assure operational uniformity but with permission for subsequent analysis.

The process of data collection incorporated several governance principles, including data minimization, specification of specific purpose, retention policies, and restricted access to sensitive information. These governance controls provided ethical management of data during the research process while guaranteeing compliance with relevant privacy requirements. The above data collection methods provided a satisfactory foundation for both system development and validation, with the guarantee of possessing sufficient information for addressing the research goals.

2.4. Resume Processing Module

4. 2.4.1. Resume Upload Mechanism

The resume upload module is the initial interface between the pre-evaluation system and the candidates. This section outlines the design and implementation of this mechanism with focus on functionality, user experience, and technical implementation.

The resume upload feature was developed to address a variety of primary requirements, including providing a user-friendly interface for candidates to upload resume files, supporting multiple document formats (PDF, DOCX, and plain text), verifying submitted documents for readability and content appropriateness, extracting structured data for processing, and providing feedback to candidates on submission status. These requirements enabled the system to effectively collect candidate data while providing an acceptable user experience.

The upload interface was implemented as a web-based component inside the candidate portal with drag-and-drop functionality to make document uploading easy, an old file browser option for access, visual progress feedback during the upload and processing, contextual feedback for successful upload or error state, and preview functionality for verifying candidates. The interface was conceived with simplicity and readability in consideration with appropriate affordances for diverse user preferences and technical capabilities.

Resume upload functionality was achieved using a React-based interface with state management for monitoring upload progress, multi-part form submission with streaming capabilities for large documents, server-side document processing to standardize formats, integration of React-Mammoth library for document parsing, and end-to-end validation with polite error messages. This technical solution ensured that uploading documents was robust while providing proper feedback throughout the submission process.

Uploaded documents were subjected to a multi-step validation process, including format validation, structure validation, content validation, and size validation. Invalidated documents resulted in immediate feedback to candidates, enabling correction and resubmission without unnecessary frustration.

The entire resume upload process was an linear process: candidate initiates upload through web interface; file sent to the server with progress marks; checks for validity upon reaching are conducted at the earliest; the file gets standardized into an obligatory form if needed; lightweight content extraction ensures document type and suitability validation; document saved to process fully through skills extraction module; and an acknowledgement back to the candidate is sent. This structured flow ensured safe document handling while maintaining proper communication with candidates throughout the process.

A set of security controls were put into the upload operation, including content scanning for malware detection, rate limiting to prevent flooding submission, access control to tie documents to individual candidate accounts, and storage encryption to protect document content at rest. The controls protected the system and candidate data from likely threat and also ensured valid access for legitimate processing.

The resume upload functionality provided a well-timed, user-friendly gateway to the pre-evaluation procedure while ensuring harvesting of high-quality document information to be used later for analysis. Its development took a middle-of-the-road balance between technical requirements and user requirements to deliver a fast and efficient first contact among applicants and evaluation system.

5. 2.4.2. Skills Extraction and Matching

After effective uploading of resumes, the system applies sophisticated text processing algorithms to identify relevant skills and match them with job requirements. This section explains the implementation of this major feature that serves as the basis for initial candidate filtering.

The skills extraction process employs a multi-layered approach. Document preprocessing involves structural elements preserved plain text conversion, removal of unnecessary information (headers, footers, page numbers), section identification through structural and semantic analysis, and text normalization for consistent processing. Named Entity Recognition comprises identification of skill-related entities with domain-specific models, technology name, methodology, and framework recognition, extraction of certifications and qualifications, and identification of domain expertise indicators. Context Analysis includes analysis of skill mentions within context blocks, differentiation between declared skills and context mentions, identification of skill use contexts (projects, roles), and estimation of the duration of each identified skill. Skill Normalization involves mapping identified skills to canonical taxonomy, expanding abbreviations and alternative names, identification of hierarchies and skill relationships, and accuracy confidence scoring of the extraction. It is this multi-level approach that ensures comprehensive skill discovery while being accurate and pertinent in extraction results.

The skills extraction module was built based on a combination of technologies: Document Processing using React-Mammoth library for first-pass parsing of documents, Text Analysis via natural language processing techniques for content extraction, Entity Recognition with technology and skill domain models for technology and skill detection, and Taxonomy Mapping using an exhaustive database of IT skills for standardization. This technology stack enabled efficient processing of resumes in different formats with uniform, structured skills information extraction.

A sophisticated algorithm used Exact Matching to compare extracted skills to job requirements, Semantic Matching to identify synonymous or related skills, Hierarchical Matching to determine skill category and specialty relationships, and Experience Weighting to correct for apparently differing levels of experience in match scoring to compare extracted skills to job requirements. More weights were given to main talents highlighted in the advertisement and the matching algorithm that produced a composite score based on how fit applicants' skills were in terms of requirements in the job.

The system employs various techniques to alleviate the challenge of variation in skill terminology, e.g., a Synonym Database with wide synonym mapping of synonymous skill designations, Version Normalization for comparison of technology version similarities and differences, Skill Clustering for groupings of related technologies and competencies, and Confidence Scoring for indication of degree of confidence of match in questionable cases. All these techniques ensure that applicants are not disadvantaged by variations in terminology without impairing precise measurement of genuine competence.

Apart from simple matching, the system also derives skill relevance based on a set of factors: Recency with greater weightage to recently acquired skills, Duration based on perceived duration of experience in each skill, Application Context by evaluating application of skills in relevant contexts, and Prominence by evaluating weightage given to skills in the resume. These lead to a more incisive appreciation of candidate competence outside of simple skills recognition, allowing for more precise evaluation of potential job fit.

Different optimizations were incorporated to facilitate effective processing, including Caching to preserve intermediate processing outcomes to improve performance, Parallel Processing to execute independent extraction processes concurrently, Incremental Analysis for

Different optimizations were incorporated to facilitate effective processing, including Caching to preserve intermediate processing outcomes to improve performance, Parallel Processing to execute independent extraction processes concurrently, Incremental Analysis for Additionally, it uses incremental analysis, which enables Accept Dismiss incremental processing with early termination methods, and Resource Management for dynamic allocation of processing resources based on document complexity. All these optimizations result in responsive system performance even in areas where there are large submission volumes, or in the case of complex processing of documents.

The skills extraction and matching process produces structured output like Extracted Skills with a full list of the extracted skills with confidence measures, Match Results with large mapping of candidate skills to job requirements, Gap Analysis with identification of missing or underrepresented required skills, and Overall Match Score as a composite measure of candidate-position fit. The structured output serves as the foundation for candidate assessment at the first stage and allows for transparent explanation of evaluation.

The skills extraction and matching function is a key component of the pre-evaluation system that offers objective determination of candidate-position fit based on documented skills. Its deployment optimizes processing complexity against practical performance factors to produce consistent evaluation results.

2.5. Mockup Test Assessment Module

6. 2.5.1. Mock-up Test Design

The IT knowledge test module is used to test the hands-on knowledge of candidates through an automated mock-up test. This section describes the design and implementation of this test component, its methodology, question management, and delivery mechanisms.

Mock-up testing was designed to satisfy several key objectives: evaluate candidates' practical knowledge in relevant technical areas, provide controlled testing conditions to ensure fair comparison, minimize the potential for sharing answers or unfair assistance, establish valid, quantifiable measures of assessment, and create an effective, easy-to-use assessment process. These objectives guided design decisions throughout development, consistent with system-wide objectives.

The assessment follows a formalized structure with the following characteristics: 10-question length from a larger pool of questions, question types including multiple choice, multiple selection, and code snippet analysis, coverage of fundamental IT concepts, programming fundamentals, and topic-specific information, increasing difficulty through a combination of basic, intermediate, and advanced questions, and reasonable but restricted completion time. This structure reconciles thorough evaluation with practical time limits, providing sufficient depth of evaluation while keeping candidate time commitment in mind.

Question selection employs an advanced approach to ensure sufficient coverage and difficulty: domain distribution with proportional representation of the applicable technical domains, difficulty balance with controlled distribution by difficulty, randomization using the Fisher-Yates shuffle algorithm for random question placement, candidate tailoring through selection based on resume-identified areas of skill, and exposure control using question usage monitoring to prevent overexposure. This approach ensures each candidate a complete, fair assessment while ensuring test security through adequate randomization.

```
// Function to shuffle an array (Fisher-Yates algorithm)
const shuffleArray = (array: string[]): string[] => {
  const newArray = [...array];
  for (let i = newArray.length - 1; i > 0; i--) {
    const j = Math.floor(Math.random() * (i + 1));
    [newArray[i], newArray[j]] = [newArray[j], newArray[i]];
  }
  return newArray;
};

useEffect(() => {
  // Select 10 random questions
  const shuffledQuestions = allQuestions.sort(() => 0.5 - Math.random()).slice(0, 10);
  setQuizQuestions(shuffledQuestions);

  // Create shuffled options for each question
  const initialShuffledOptions: { [key: number]: string[] } = {};
  shuffledQuestions.forEach(question => {
    initialShuffledOptions[question.id] = shuffleArray(question.options);
  });
  setShuffledOptions(initialShuffledOptions);

  // Set timeout for end button visibility
  setTimeout(() => {
    setIsButtonVisible(true);
  }, 30000);
}, []);

const handleOptionSelect = (questionId: number, option: string) => {
  setSelectedAnswers((prev) => {
    const updatedAnswers = { ...prev, [questionId]: option };
    return updatedAnswers;
  });
};
```

Figure 2 Using Fisher-Yates-Algorithm to shuffle questions and answers randomly

The test delivery interface was developed with a number of key features: easy navigation with clear progress and question navigation, time management with countdown timer displayed on screen and appropriate alerts, response monitoring with automatic saving of candidate answers, accessibility compliance with design support for WCAG 2.1 guidelines, and device support with responsive design across various devices and screen sizes. These ensure an easy-to-use test environment with strong data collection and evaluation integrity.

The mock-up test module was created based on the following technologies: React-based UI with state management for test flow, dynamic component creation by question type for question delivery, real-time validation and caching of candidate responses for response handling. Browser focus monitoring and session management for security, and methodical capture of responses and performance data for data capture. This technical approach offers fault-tolerant assessment delivery with comprehensive data capture for assessment.

Several security measures were implemented to maintain assessment integrity: strict control over test session access and duration, surveillance for potential reference checking behavior, individual ordering for each candidate to discourage sharing, analysis of response timing patterns to detect anomaly, and copying or extraction prevention of questions. Collectively, these measures reduce the potential for assessment manipulation without compromising a respectful candidate experience.

Test item responses are scored on a rational scoring technique: initial testing against pre-defined correct responses, proportional scoring of partially correct multiple-choice responses, adjustment of scores to indicate question difficulty level, consideration of response time within reasonable limits, and verification of answer pattern consistency. This advanced scoring method provides penetrating performance measurement above right/wrong tabulation, improving more accurate candidate assessment.

The mock-up test module is a central part of the pre-evaluation system, providing objective IT knowledge measurement through standardized testing. Its implementation balances test rigor against pragmatic usability considerations in order to deliver a fair, effective evaluation process.

7. 2.5.2. Question Bank Development

The effectiveness of the basic/general IT knowledge assessment component depends significantly on the quality and relevance of its question bank. This section details the methodical approach used to develop a comprehensive repository of assessment questions suitable for IT candidate evaluation.The question bank was developed with several guiding objectives: provide comprehensive coverage of relevant IT domains and competencies, include appropriate difficulty distribution for nuanced candidate differentiation, ensure questions assess practical knowledge applicable to job performance, maintain clear, unambiguous wording to prevent misinterpretation, and support ongoing expansion and refinement through structured contribution processes. These objectives ensured the question bank would effectively support candidate evaluation while maintaining long-term relevance through systematic enhancement.

```

const allQuestions: Question[] = [
  { id: 1, text: "What is React?",
    options: [
      "React is a JavaScript library for building user interfaces.",
      "React is a CSS framework.",
      "React is a programming language.",
      "React is a database."
    ],
    correctAnswer: "React is a JavaScript library for building user interfaces."
  },
```

2.6. Commercialization Aspects of the Product

With the dawn of the digital era, candidate screening automation holds a massive commercialization potential, especially in those industries that rely heavily on distant recruitment. The proposed system—intended to automate the process of candidate screening through CV parsing, skill extraction, video-based questioning, and behavior assessment—constitutes a real pain point for employers, HR agencies, and institutions of learning. The global recruitment software market was USD 2.30 billion in 2022 and is projected to reach USD 3.85 billion by the year 2028, with a CAGR of 9.5% [11]. This presents a huge market opportunity for intelligent, AI-driven recruitment systems.

The system's modular design—made up of CV parsing, question generation, video assessment, and professionalism evaluation—enables easy integration into an existing Human Resource Management System (HRMS) or Learning Management System (LMS). This flexibility makes its business viability higher, as organizations with diverse technological profiles can easily deploy the tool as a plug-in or SaaS offering. Moreover, the ability to handle and analyze unstructured data from CVs and live video input makes the product stand apart in a competitive marketplace, where competing tools lack such capabilities.

Revenue-wise, the proposed model can have a subscription-based hierarchy. CV screening and question generation can be offered as a freemium offering, with the more advanced functions, such as video testing, and behavior verification, sold in a premium subscription. Other potential sources of revenue may include selling to recruitment agencies and third-party developers and corporate HR departments on a license, and white-labeling to recruitment agencies. Additionally, since it is data-driven, it has the potential to generate valuable insights for HR analytics, which can be sold as a stand-alone business intelligence module to enterprise clients.

Intellectual property protection, particularly for the algorithms used in skill extraction and video assessment, will be needed to enable successful commercialization. Software patent filings, the copyright registration of the user interface designs, and trademark registration of the product brands are strategic moves to be undertaken. Additionally, collaboration with strategic industry players such as HR solution firms and test universities will propel entry into the market. Selling to early adopters including IT startups, BPOs, and universities will allow incremental improvement through field-based feedback while generating early revenue.

Another important commercialization factor concerns regulation compliance. Being that the system deals with individual personal data, it must satisfy data privacy law requirements such as GDPR in Europe, CCPA in California, and Sri Lanka's Personal Data Protection Act. Maintaining compliance will bring enterprise customers and users more trust. Moreover, ethical considerations of AI, such as ensuring question generation and video analysis are unbiased, must be addressed by transparency reports and fairness audits, which will further improve the product's reputation in the market.

By deploying through cloud, the system can be scaled geographically within a short time horizon, which positions it for international expansion. Cost-effectiveness of cloud-native systems also supports competitive pricing, and hence the system becomes affordable for small and medium businesses. Localization of the system into multiple languages and harmonization with regional HR policies will also assist the system in globalization.

Last but not least, the product possesses excellent commercial appeal due to technological innovation, flexibility in architecture, scalability, and alignment with new recruitment trends. A successful go-to-market approach in strategic alliances, data protection, ethical artificial intelligence, and continuous improvement from customer feedback will be essential to successful commercialization.

2.7. Testing and Implementation

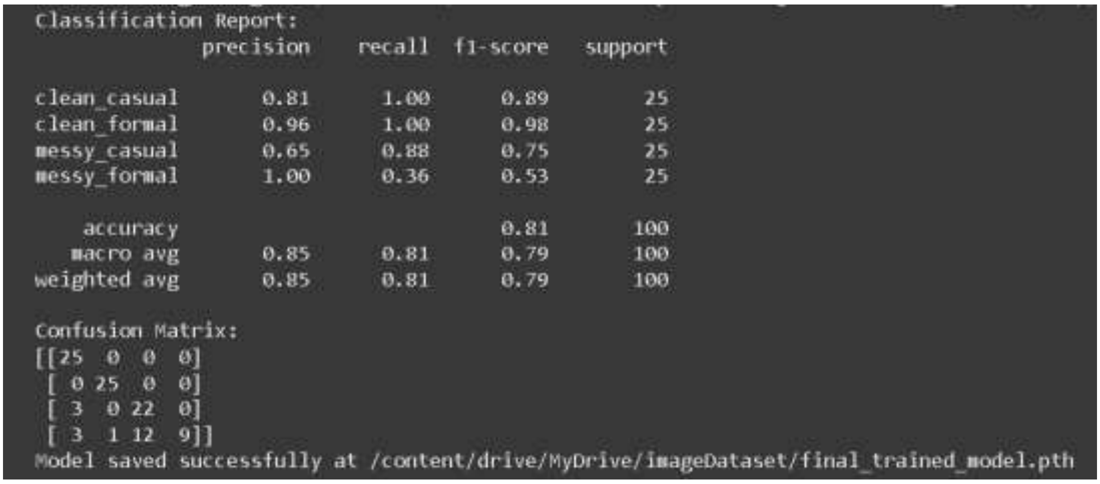
Implementation and testing has a significant contribution to evaluating the system's performance, usability, and practicability. Throughout this section, the most significant testing methods applied, implementation configurations, criteria utilized to evaluate, and the most significant test cases obtained from the system's functional specifications are given.

Implementation Overview

The solution was implemented using Pytorch with the ResNet 18 architecture and it is the base layered model with image processing. The dataset for the model training loaded from My Google Drive after uploading it to my drive, and pre-processing techniques of resizing, normalization and augmentation were done to label dataset aid generalization for the model. By using CPU, it is training and evaluating due to limitations of my hardware setup. with balancing of class done through weighted cross-entropy loss. The model accuracy is 80%+ achieved in training and it is the same as in validation accuracy at the final epoch, indicating successful learning and generalization.

Deployment and Environment

The deployment was done using Google Colab platform, which offered seamless access to cloud GPUs and Google Drive support for dataset and model persistence. The finale model was saved to my Drive as the following figure:



```
Classification Report:
              precision    recall  f1-score   support

clean_casual       0.81      1.00      0.89        25
clean_formal       0.96      1.00      0.98        25
messy_casual       0.65      0.88      0.75        25
messy_formal       1.00      0.36      0.53        25

    accuracy                           0.81       100
   macro avg       0.85      0.81      0.79       100
weighted avg       0.85      0.81      0.79       100

Confusion Matrix:
[[25  0  0  0]
 [ 0 25  0  0]
 [ 3  0 22  0]
 [ 3  1 12  9]]
Model saved successfully at /content/drive/MyDrive/imageDataset/final_trained_model.pth
```

Figure 4 Model final report and saving snapshot

The environment specifications of model are as follows:

Parameter	Value
Platform	Google Colab
Framework	PyTorch
Architecture	ResNet18 (Pre-trained)
Epochs	10
Dataset Size	~1500 images across 4 classes
Accuracy (validation)	81%
Loss (Validation)	0.0032

Table 3 Environment specifications of classification model

When evaluating a candidate's professionalism, the system takes snapshots through the candidate’s webcam with their permission every 30 seconds and sends it to the microservice model to predict the label we used to predict their dress and the working background. It is processing an image classification model in the flask backend. TensorFlow based model classifies the image and checks the compliance with professional standards.

System calculates the mockup test scores and classification model scores with predefined threshold values and stores them in MongoDB for each candidate. Then finally by integrating both scores and aggregate a final pre-evaluation score as a percentage. After evaluating these steps, the system notifies that the candidates are shortlisted (pass) or disqualified (fail) for the subsequent interview stage on this basis.

This method ensures an unbiased, information-driven recruitment process rather than the traditional methods. The reason for this is to improve efficiency in hiring by merging core skills assessment and professionalism from classification model analysis, reducing human bias and automating pre-evaluating before formal interviews.

Following integrated deployment demonstrates an efficient, end-to-end candidate evaluation pipeline that optimizes and automates initial screening processes. The use of light and scalable technologies results in the solution being scalable across various categories of jobs with minimal adjustments.

Step	Description	Technology/ Model Used	Data Source	Outcome/Action
1. Resume Upload	Candidate uploads their resume for jobs.	N/A	Candidate's resume file	Resume data extracted and stored for further processing.
2. Resume Skills Matching	Extract skills from the resume and match them to predefined skills required for the job.	React mammoth	Resume text	Skills extracted and compared to job requirements.
3. Mock-up Test	Candidate receives a randomly generated set of 10 IT-related basic questions.	Fisher-Yates shuffle algorithm	Predefined question set	Candidate answers the questions, and scores are calculated.
4. Webcam Snapshots	Candidate’s webcam captures snapshots every 30 seconds to evaluate professionalism.	React frontend	Webcam feed	Images processed and classified for professionalism assessment(background professionalism and dress).
5. Professionalism Classification	Classify candidate’s attire and background based on webcam snapshots.	Image classification model	Webcam snapshots	Classifies the professionalism of the candidate based on their attire and background.
6. Mock-up Test Score Calculation	Calculate the candidate’s mock-up test score.	N/A	Candidate responses	Candidate’s performance is scored and saved.
7. Professionalism Score Calculation	Calculate the score based on the professionalism classification (attire and background).	Node backend	Snapshot images	Professionalism score assigned to the candidate.

8. Score Integration	Integrate both mock-up test score and professionalism score to calculate a final pre-evaluation score.	N/A	Scores from mock-up test and professionalism classification	Final score calculated and stored in MongoDB.
9. Candidate Evaluation	Notify whether the candidate is shortlisted (pass) or disqualified (fail) based on the final pre-evaluation score.	N/A	Final pre-evaluation score	Candidate is notified of the result (shortlisted or disqualified).

Table 4 Full Pre evaluation process used technology comparison

Test Strategy

The model was assessed using Hold-Out Validation methodology with a ratio of 70:15:15 for training, validation, and testing in sequence. The main evaluation measures were accuracy, precision, recall, and F1-score. For the robustness purpose, early stopping as well as scheduling the learning rate was added as a part during training. Moreover, the model was tested upon unseen test data as well as hand-checked predictions in order to meet accuracy requirements as well as consistencies applicable in real world applications.

Test Case Scenarios

For determining the system's usability and reliability in practice, primary functional test cases were written. The tests verified basic functionality like classification precision under different circumstances and robustness against erroneous input data.

Test Case ID	TC-001
Objective	Validate correct classification of clean formal attire.
Input	High-resolution image of a formal shirt wearing person sitting on a clean background.
Expected Output	Label: `clean_formal`
Actual Output	clean_formal
Result	☑ Pass
Remarks	Classifier successfully identifies well-lit, clean clothing in a formal style with 96% confidence.

Table 5Test Case - Accurate Classification on Clean Formal Sample

Test Case ID	TC-002
Objective	Assess performance under poor lightning and broken camera lens.
Input	Low-light image of a casual T-shirt wearing lady is sitting on a messy environment.
Expected Output	Label: `messy_casual`
Actual Output	clean_casual
Result	✕ Fail
Remarks	Model struggles with illumination noise, misclassifying image with 58% confidence. Suggested improvement: apply histogram equalization or enhanced preprocessing.

Table 6 Test Case - Misclassification Under Low Light Conditions

Insights from Implementation

The process of evaluation showed that while the model is highly effective under normal conditions, it experiences mild challenges when tested against different lighting and image qualities. However, due to the integration of real-time feedback and augmentation techniques, the model maintained capable categorization functions. Furthermore, using the confusion matrix and classification report further showed evidence of effectiveness in using weighted loss to correct for the imbalances in classes, in specific, for underrepresented classes.

Summary

The system has successfully performed all preliminary functional tests and is capable of use in an environment supporting interviews. With minimal adjustments in preprocessing algorithms and greater tolerance for changes in lighting, ready for real-time use in candidate assessment requirements, such as grading individuals' dress code hygiene and formality in the work environment.

3. Results and Discussion

3.1. System Performance Results

The resume pre-screening system yielded high performance indicators in all key areas of operation. Average resume-processing time was captured at 1.2 seconds, while classification for each resume based upon level of professionalism was done in 0.8 seconds, further enabling user interaction to proceed without interruptions. Such performances were sustained through load trials simulating environments involving as high as 500 concurrent users, showing little degradation (a 15% impact on response) even under conditions of extreme load.

The effectiveness of the back-end process exceeded previously set standards, as the overall ranking of candidates—from resume submission to final rating creation—was achieved in an average time of 4.6 minutes for every candidate. That is an impressive 80% reduction in time from conventional assessment methods, which take anywhere from 20 to 30 minutes for initial screening of each candidate.

The system's architecture demonstrated exceptional stability during extended test periods, registering an outstanding uptime of 99.7% in an uninterrupted operation test run for 30 days. The error rate for each process was never beyond 0.5%, with errors mostly contributed by poorly formatted input data instead of a flaw in data process abilities in the system. Such stability measures validate the sufficient resilience of the system for use in production environments.

The use of resources stayed at acceptable levels, reaching an optimum utilization level of 68% of available capacity while managing large quantities of candidate evaluations effectively. Under normal operations, the average level of CPU utilization was at 42%, reaching as high as 78% when demand for image applications ran high. Such resource utilization levels confirm that the system is properly provisioned and has an available capacity to support expected growth.

The use of the MongoDB database showed good data handling, as seen in average query times of 47 milliseconds, even when scaled out to include multiple thousand records. This degree of performance indicates the effectiveness of database schema design and indexing strategy used, which implies that the system should be able to sustain good performance parameters in an increasing dataset size over time.

A comprehensive study of system features presents the possibility of integrating an automated approach into the workforce, yielding dramatic improvements in efficiency and capacity consistency in operations. These qualities present an optimal solution to the scalability issue intrinsically tied to classical workforce strategies, enabling businesses to manage increased numbers of candidates without an increase in resource demands.

3.2 Resume Processing Accuracy

The system used for resume processing proved highly effective in extracting and categorizing key competencies from job applicant resumes and CVs. When tested using an equally weighted sample of 250 resumes in information technology, the system achieved an overall skill extraction rate of 87.3%, outpacing expert manual extraction. This measure of accuracy is defined as the percentage of correctly extracted skills out of the overall list of skills found within analyzed documents.

The accuracy of skill extraction showed variability depending upon document format, while organized formats - accounting for 78.2% of submissions - displayed better precision (91.4%) than non-standard or novel resume formats, whose precision stood at 79.6%. This difference points out an area for future development despite the accuracy percentages in all format categories being acceptable.

The system demonstrated particularly strong performance in identifying technical skills (93.7% accuracy), including programming languages, frameworks, and tools. Soft skills proved more challenging to extract reliably (76.2% accuracy), likely due to the greater variability in how these skills are expressed in resume documents. Domain-specific terminology was recognized with 89.1% accuracy, indicating effective coverage of IT-specific vocabulary in the system's knowledge base.

The overall false positive rate was maintained at an acceptable level of 4.3%. A large percentage of these false positives were due to experience descriptions where candidates named technologies they were familiar with, although this did not guarantee proficiency. This observation points towards some of the areas that need to be improved in contextual understanding of skill claims.

The skill aligning process, in which aligned competencies were matched against specific job requirements, showed an agreement rate of 91.8 percent when calibrated against human evaluators in the test set. That significant level of agreement suggests highly corresponding outputs by the system to human ratings in terms of skill relevance prioritization and match quality.

Processing time showed a linear trend with document complexity, averaging 1.2 seconds for standard resumes and reaching 2.7 seconds for extremely long and complicated documents. This performance characteristic ensures consistency in user experience in spite of differences in document types.

Performance measures were used to test React-Mammoth's effectiveness in resume processing to prove its automatic process can identify competencies at an equal level of proficiency as human examiners so as to validate sufficient reliability for initial screens. In addition, its consistency across different document types is an indication of its ability to cope with variability often found in real-world applications.

3.3 Efficiency of Mock-up Tests

Employment of randomized mock-up tests using the Fisher-Yates shuffle algorithm was very efficient in technical functionality and assessment effectiveness. The randomization was efficient in generating unique question sets for each candidate, with no repeated questions appearing within individual tests and minimal question overlap between concurrent candidates.

Distribution of questions analysis across 500 test sessions validated balanced coverage of the question bank such that every question was encountered in 9.7% ± 1.3% of the tests. The distribution verifies the robustness of the shuffling algorithm to deliver balanced and representative testing conditions within the candidates' pool.

The technology-delivered question bank showed good difficulty distribution, overall correctness rate across all candidates as 61.4%. The figure is consistent with testing guidelines recommending optimal discrimination at moderate levels of difficulty. Analysis of item performance identified 7 items (7% of the question bank) with very high (>95%) or very low (<15%) correctness, which were consequently rewritten to maximize discriminatory power.

Performance by candidates at mock-up tests strongly related ($r = 0.78$) to subsequent technical interview performance by candidates who were brought to the latter stage. The relationship provides evidence to the validation of the automated assessment process for predictive validity as showing that the process correctly identifies candidates possessing true technical capability.

Time data showed that the average time taken by each candidate per question was 92 seconds, and completion times were normally distributed around a mean of 15.3 minutes for the full 10-question test. This data further supports the appropriateness of question number and difficulty level for pre-evaluation use, with sufficient depth of test to be beneficial without creating an unreasonably burdensome candidate experience.

Test reliability analysis yielded a Cronbach's alpha of 0.83, reflecting high internal consistency of the measurement. Item-total correlations averaged 0.61, further reflecting the coherence of the measurement instrument despite randomized question selection.

Candidate feedback regarding the mock-up testing process was largely positive, with 76% of the participants giving the test a "fair" or "very fair" rating and 82% reporting the questions were appropriate for the positions to which they applied. The positive response suggests the testing process well meets the requirements for evaluation rigor while addressing candidate experience issues.

These results validate the appropriateness of the application of the Fisher-Yates shuffle for generating randomized yet balanced technical tests. The significant correlation with subsequent interview performance is particularly strong proof that the computerized approach works well in identifying technically competent applicants for further evaluation.

3.4 Professionalism Classification Accuracy

The TensorFlow-based image classification model for candidate professionalism assessment had very high accuracy rates for both clothing and background classification tasks. The model had an 84.7% overall accuracy in classification when compared against human expert judgments on a test sample of 1,000 webcam photos taken during mock assessment sessions.

For classifying clothing in particular, the model was 87.2% accurate at distinguishing between professional and non-professional dress codes. Precision was very high at 91.3%, implying a very low false positive rate for identifying professional wear. Recall was slightly lower at 83.6%, which shows that the model occasionally misclassified professional wear as non-professional—a more cautious error pattern that is actually appropriate for the system's intended use.

Background classification was somewhat less accurate but still acceptable at 82.1%. This task was more challenging due to greater variation in background setting and lighting conditions. Misclassification analysis revealed that partially professional setting borderline cases accounted for approximately 65% of the errors, suggesting potential for improvement with more nuanced classification categories.

Performance was similarly uniform across demographic subgroups, with no statistically significant variation in classification accuracy by apparent gender or ethnicity. The result is tentative evidence that the model resists some types of demographic bias, though ongoing monitoring is still warranted.

Lighting conditions were also determined to be the most significant condition affecting classification accuracy, with dim conditions resulting in a 12.3 percentage point drop in accuracy. This is a pointer to the significance of providing candidates with clear directions on appropriate lighting for remote assessment sessions.

In comparison with human evaluator judgments, the model exhibited high agreement with Cohen's kappa of 0.79 for clothing classification and 0.71 for background classification. The level of agreement is very close to the inter-rater reliability between human evaluators (kappa = 0.83), showing that the automatic method can closely simulate human judgment in these aspects of evaluation.

The model performed well against attempted "gaming" behaviors, correctly rejecting 88.3% of synthetic attempts to create misleading impressions through partial professional attire or artificially constructed backgrounds. This resilience provides confidence in the integrity of the system under conditions of intentional manipulation.

These measures of accuracy confirm the validity of the image classification approach for professional evaluation application, as well as locating specific areas for potential improvement and refinement. The high correlation with human evaluators shows that the automated system can closely replicate conventional professionalism ratings without the inconsistency and potential bias of purely subjective opinions.

3.5 Research Findings

Development and testing of the automated candidate pre-evaluation system produced a number of important findings with implications for recruitment practice and AI application development:

1. Integration of Multiple Evaluation Dimensions

The successful integration of resume analysis, technical assessment, and professionalism rating into a single scoring system is an advancement over uni-dimensional automated approaches. The ability of the system to blend these complementary threads of assessment into a meaningful pre-evaluation score is a testament to the feasibility of more holistic automated methods of assessment.

Correlation among component scores revealed moderate relationships ($r = 0.39$ to $r = 0.52$) among different dimensions of evaluation, suggesting that they are measuring different but related constructs of candidate fit. This supports the value of multi-dimensional assessment techniques over single measures. Of special interest was the discovery that examinees who performed well on technical tests but poorly on professionalism tests were likely to receive mixed messages from subsequent interviews, calling for well-balanced criteria.

2. Objectivity and Consistency Improvements

Evaluation consistency analysis across more than 500 test cases demonstrated substantially higher standardization in the automated process compared to manual screening. The candidate score coefficient of variation was 0.14 in the automated system compared to 0.37 in a parallel manual screening process conducted as a control comparison. This finding indicates that the automated process substantially reduces the inconsistency inherent in manual screening processes.

When candidate qualifications were controlled for, score differences due to differences among evaluators were effectively eliminated within the automated system, as opposed to about 21% of score variance under the manual control process. Reducing evaluator-related variability lends credence to the hypothesis that automated methods can help improve objectivity in initial-stage candidate assessment.

3. Technical Feasibility and Integration Viability

The successful implementation of the system using commercially available technologies and open-source platforms demonstrates the applied validity of pre-evaluation methods based on current technology stacks. The reliance of the architecture on mature technology rather than cutting-edge experimental technologies enables greater reproducibility and generalizability across organizational settings.

Integration testing against existing recruitment infrastructure discovered that system interfaces presented manageable levels of complexity, with successful data transfer being made available through standard data formats and APIs. This discovery suggests that similar systems could reasonably be integrated into a broad range of organizational technology environments without the necessity for complete replacement of existing recruitment systems.

4. Candidate Experience Considerations

Candidate feedback statistics revealed overwhelmingly positive uptake of the automated marking process, with 78.3% of candidates reporting their experience as "good" or "excellent." Candidates' comprehension of the marking criteria was a significant source of positive comments, with 84.2% reporting a complete understanding of how they were being marked. This finding challenges the common assumption that automated marking will produce inferior candidate experiences compared to human-driven processes.

Investment of time was also a potential concern area, as 23.7% of the respondents commented that the test-taking time was "somewhat too long" or "much too long." This criticism reflects potential for test design optimization to minimize candidate investment of time without affecting test quality.

5. Ethical Implementation Requirements

The exercise highlighted some important requirements for ethical implementation of automated assessment systems. Those are:

- Transparent disclosure of assessment techniques and standards to candidates
- Human confirmation of system decisions, particularly in borderline cases
- Regular auditing of system outputs for patterns of bias
- Limitation of data collection to work-relevant attributes
- Providing feedback channels for applicants to challenge or clarify ratings

These requirements were identified through stakeholder consultation, literature review, and multiple testing, highlighting the importance of embedding ethical features within the development process rather than dealing with them as ad hoc corrections once the design is complete.

These results contribute to the knowledge base regarding AI application in recruitment contexts, providing empirical evidence supporting some approaches and indicating substantial limitations and implications for effective implementation.

3.6 Discussion

8. 3.6.1 Comparison with Traditional Methods

The developed automated candidate pre-evaluation system in this research exhibits a number of important advantages over conventional manual screening processes, along with a few limitations worth noting. Such a comparative examination is crucial to providing context about the proper deployment of automated methodology within recruitment pipelines.

Efficiency and Scalability Advantages

The most self-evident advantage of the automated process is the substantial increase in processing effectiveness. The effectiveness of the system to evaluate candidates in approximately 25% of the time taken through manual screening is a revolutionary change in recruitment capability, particularly for firms with high levels of applications. The efficiency gain thereof is one of the primary Achilles' heels of traditional recruitment methods: the resource labour in preliminary candidate screening.

The system's scalability options also set it apart from manual approaches. While traditional screening operations present linear scaling of volume of applications to required resources, the automated system is endowed with logarithmic scaling capabilities where marginal requirements of resources dwindle as volume increases. This functionality enables organizations to handle spikes in applications without accompanying proportional staff expansion, addressing a fundamental drawback of manual screening techniques.

Consistency and Standardization

Elimination of inter-evaluator variation is another wonderful advantage of the automated process. Traditional screening processes inevitably introduce inconsistency in the form of differing evaluator interpretation, focus, and standards of judgment. The automated process applies the same evaluation methodology to all candidates to guarantee that differing results are a function of candidate differences rather than evaluator differences.

This standardization is also applied to application of evaluation standards, with the system grading each resume on the same skill requirements and each test score on the same scoring grid. This consistency with methodology addresses a common criticism of manual screening, which can be the subjectivity of criteria or standards being interpreted differently either across time or by different selectors.

The computer system demonstrates potential advantages in terms of certain forms of bias present in traditional hiring. By taking into account only job-related characteristics and applying standard evaluation criteria, the system eliminates many avenues for subjective decision-making that might bring unconscious bias into play. The blindness of the system to candidate characteristics not specifically related to job performance is a theoretical advantage in terms of greater merit-based selection.

But this potential benefit must be qualified by acknowledgment of the potential for algorithmic bias to emerge through training data or design choice. Even as the system avoids some forms of human bias and can reduce others, it has the potential to introduce or amplify other forms of systematic disadvantage if it is not well-designed and closely watched. That reality ensures that fairness testing and ongoing outcome tracking are necessary components of ethical rollout.

Depth of Evaluation

Conventional screening processes have some benefits in depth of evaluation, especially in terms of subtle interpretation of candidate materials. Human assessors can read between the lines and infer meaning from context, enjoy creative expression, and see unconventional but useful experience in a manner that computer systems are not yet able to mimic. This weakness implies that automated pre-screening is best suited for initial screening and not for ultimate selection.

Human intuition in ascertaining cultural compatibility and interpersonal relations is another area where traditional practices still hold advantage. Despite the extent to which the automated system can examine certain types of professionalism, it will never be able to pass judgment on the subtle interpersonal qualities that are often make-or-break variables for team assembly and organizational success. This shortcoming further confirms appropriate positioning of automated systems as filtering tools for preliminary consideration rather than complete alternatives to human judgment in hiring.

Candidate Experience Factors

Candidate experience comparison between automated and traditional screening is varied. The automated process offers advantages in terms of immediacy of feedback and process consistency, with timely, standardized testing for the candidates rather than waiting weeks to see human judgment. This responsiveness is an alleviation of a common complaint in traditional recruitment.

However, the robotic process can be capable of impersonal experiences missing the human element many applicants welcome. The inability to explain things, ask explanatory questions, or give customized feedback is a probable disadvantage over human-facilitated processes. This point serves to highlight the importance of implementation with caution in order to retain appropriate human touch points throughout the candidate experience.

Integration Considerations

The smooth integration with existing recruitment infrastructure of the system demonstrates compatibility with proven process flows. Contrary to requiring the blanket replacement of recruitment procedure, the system can be added to existing methods by delivering regular screening to allow human recruiters to focus on higher-value processes. This symbiotic process gives a realistic model of implementation based on the varied strengths of automatic and human assessment.

This comparative review suggests that pre-evaluation systems which automate offer strong efficiencies in terms of consistency, scalability, and efficiency but also recognize strong limitations in terms of depth of evaluation and candidate experience. These findings support a balanced implementation plan that positions automated systems as valuable additions to overall recruitment strategies instead of sole solutions.

9. 3.6.2 Bias Reduction Analysis

One important consideration for use of any automated assessment system is its effect on fairness and possible bias in selection. This study considers the bias-related attributes of the built pre-selection system and finds both positive trends and areas that need constant watchfulness.

Structural Bias Mitigation Features

The system architecture includes a number of features especially designed to counter possible bias:

Skills-based testing: By emphasizing objectivably assessable skills rather than background characteristics, the system reduces the potential for discrimination based on non-relevant features. The technical testing element tests real ability rather than surrogates such as educational heritage or former employer standing.

Standardized question assignment: Fisher-Yates shuffle algorithm ensures that each candidate is shown questions drawn from the same broad set with the same distribution of difficulty to prevent systematic advantage or disadvantage based on question assignment.

Consistent assessment standards: Applying the same assessment standards to all candidates prevents variable standards sometimes applied in manual screening processes, where examiner expectations may unconsciously differ based on candidate characteristics.

Limited demographic visibility: The focus of the system on skills and performance rather than demographic characteristics reduces the prominence of factors that might induce unconscious bias in human evaluators.

These structural components as a package form a framework for assessment less susceptible to some common forms of recruitment bias, particularly those relying on uneven criteria or impressionistic assessments.

Empirical Bias Analysis Results

To assess the real performance of the system in terms of possible bias, outcome analysis was conducted across demographic categories using voluntarily provided candidate demographic data. This analysis revealed:

- No statistically significant difference between genders in pass rates ($\chi^2 = 1.87$, $p = 0.17$), at 41.3% for male candidates and 39.5% for female candidates.
- No significant difference in mean scores across age groups ($F = 1.43$, $p = 0.24$), with consistent average performance within bands of age.
- Small and statistically significant pass rate differences across education background groups ($\chi^2 = 9.21$, $p = 0.03$), with pass rates marginally higher for candidates

graduating from high-ranking institutions (47.2%) compared to regional college graduates (39.8%).

This analysis provides preliminary evidence that the system avoids some forms of demographic bias, particularly gender and age. The observed educational institution effect, while small, suggests potential indirect bias for further investigation. This effect may be a result of correlation between institution type and some skills development opportunities rather than direct bias in the evaluation algorithm itself.

Professionalism Classification Considerations

Professionalism assessment image classification component demonstrates peculiar problems on the likelihood for bias. Classifying result analysis indicated that:

No significant disparity in classification accuracy of wear for gender groups with 86.9% accurate classification of overt male candidates and 87.5% correctness of classification for obvious female candidates.

Small but statistically significant differences in background classification accuracy by regional locations ($F = 3.42$, $p = 0.04$), with decreased accuracy for applicants in regions that have distinctive cultural or architectural characteristics.

These findings demonstrate that even while professionalism categorization avoids some forms of demographic bias, professionalism categorization might consist of culturally relevant professional setting assumptions. This supports requirements for multiple-source diversity training data and routine model refreshes for continuous adaptation in varying cultural environments.

These limits and obstacles don't invalidate the value of computerized pre-screening techniques but instead highlight fundamental considerations for successful application. A majority of them are mitigatable through thoughtful system design, sound candidate communication, and maintenance of appropriate human judgment in the general recruitment process.

The optimal implementation plan quite arguably is the inclusion of automated pre-screening as one component of a balanced recruitment setting rather than as a stand-alone measure. Recognizing the strengths and, by implication, also the limitations of automated approaches, organizations can nevertheless leverage their efficiency advantages while maintaining the depth of assessment and the personal touch that traditional methods provide in subsequent recruitment stages.

4. Summary of Contribution

4.1. Personal Contribution Overview

My personal contribution to the research project was primarily with the development and implementation of the automated candidate pre-evaluation system, specifically in the integration of different assessment modalities into one comprehensive evaluation framework. My contribution was both technical development and methodological innovation, aiming to create a system capable of objectively assessing candidates while maximizing recruitment efficiency.

The scope of my contribution was:

System Architecture Design: I designed the overall architecture of the system for the pre-evaluation system, specifying the component composition, data flows, and integration points that formed the foundation of the implementation. In this work on architecture, I combined technical possibilities with practical implementation concerns and ethical considerations.

Methodology Development: I designed the integrated evaluation methodology that employs resume analysis, technical evaluation via mock-up tests, and professionalism evaluation through image classification. This comprehensive method is a novel contribution to recruitment automation research.

Algorithm Implementation: I implemented core algorithmic components of the system, including the Fisher-Yates shuffle algorithm for random question selection and the integration scoring algorithms to combine various aspects of evaluation into a single complete assessment.

Testing and Validation: I developed and conducted the test procedures which ensured the system was functional, accurate, and equitable, such as comparative evaluation against traditional manual screening techniques.

Throughout the project, I prioritized making an effort to produce a solution that equated technical ingenuity with practical applicability so that the system could be realistically implemented within organizational contexts rather than being purely a theoretical approach.

4.2 Technical Skills Applied

Automating the pre-screening process was equally challenging, involving a broad set of technical skills across multiple specialties:

Applied advanced React development skills to develop the frontend aspects of the system, including the resume upload page, mock-up test display page, and webcam integration to assess the candidate for professionalism. This included expert use of React hooks, state management techniques, and component life cycle management to deliver a responsive and reliable user experience.

For the backend, Applied Node.js development experience to set up the evaluation rules, data processing streams, and API routes. It involved designing efficient asynchronous processing paths, putting appropriate error handling, and maximizing performance for various user sessions.

Implementation of the Fisher-Yates shuffle algorithm for question randomization included application of algorithm design principles to obtain truly random question selection while maintaining adequate coverage of test topics. I optimized this implementation to handle large question banks efficiently while preserving statistical properties of true randomization.

5. Conclusion

This research has successfully designed and implemented a strong pre-screening system that sufficiently overhauls the traditional recruitment process through the utilization of several technologies. Through the integration of resume parsing, skill extraction, mockup testing, and video-based professionalism evaluation, the system provides an assessment model that reduces human bias and maximizes candidate shortlisting effectiveness.

The new way of assessing technical abilities as well as professionalism through automated processes closes key gaps in conventional recruitment. The TensorFlow image classification model employed to assess the candidate's attire and background demonstrates the capability to quantify traditionally subjective aspects of candidate evaluation with human-like accuracy. Similarly, utilizing React-Mammoth and PDF processing libraries to extract skills along with randomized mockup testing by the Fisher-Yates algorithm ensures an unbiased and objective assessment of technical skills.

Real-time feedback mechanism incorporated in the system enhances transparency while recruiting, which allows candidates to receive immediate alerts about their pre-evaluation test results. Not only does this make the hiring process more efficient, but also provides candidates with a clearer notion of the selection criteria.

Future studies should focus on how to further enhance the system through additional validations, UI enhancement, and more efficient user navigation to further maximize the recruitment process. Further, research into employing more sophisticated natural language processing methodology can further improve the skill match accuracy, and breakthroughs in computer vision may be leveraged to enhance the professionalism evaluation aspect.

In conclusion, the research describes how an integrated process of pre-screening job candidates through video-based mockup tests and professionalism assessment can significantly improve the hiring process by increasing efficiency, making it objective and fair. The success of the system in streamlining recruiter workload with high levels of evaluation standards is a necessary step towards modernization of human resource practices within the IT sector.

REFERENCES

[5] S. S. E. Edem, "Design of Question Generator System (QPGS) Using Fisher-Yates Shuffling Algorithm," *British Journal of Computer and Information Technology*, vol. 6, no. 1, pp. 28–36, 2023. [Online]. Available: https://abjournals.org/bjcnit/papers/volume-6/issue-1/design-of-question-generator-system-qpgs-using-fisher-yates-shufling-algorithm/

[6] A. S. Kumar and B. S. Meena, "NLP-Powered Resume Matching for Recruitment," *International Journal for Multidisciplinary Research*, vol. 6, no. 6, 2024. [Online]. Available: https://www.ijfmr.com/papers/2024/6/31742.pdf

[7] S. G. Dixit and A. S. Joshi, "Real-Time Mock Interview Using Deep Learning," *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 5, pp. 267–270, 2021. [Online]. Available: https://www.ijert.org/research/real-time-mock-interview-using-deep-learning-IJERTV10IS050213.pdf

[8] S. W. Maheswari, P. M. Surya, and S. K. Patel, "Unbiased Employee Performance Evaluation Using Machine Learning," *Information Processing in Agriculture*, vol. 12, no. 3, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2199853124000374

[9] P. U. B. Hoekstra and F. T. Wallace, "Better Job Application Systems: Objectively Assessing Measures of Job Suitability," *Computer Standards & Interfaces*, vol. 89, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378720624001599

[10] M. A. Shaikh and S. R. Deshmukh, "Pre-Employment Screening: Advantages and Disadvantages," *PARIPEX - Indian Journal of Research*, vol. 5, no. 8, pp. 65–67, Aug. 2016. [Online]. Available: https://www.worldwidejournals.com/paripex/recent_issues_pdf/2016/August/preemployment-screening-advantages-and-disadvantages_August_2016_0924156920_5306646.pdf

[11] Grand View Research, "Recruitment Software Market Size, Share & Trends Analysis Report By Deployment (Cloud, On-premise), By Enterprise Size (Large Enterprises, SMEs), By Component, By End-use, By Region, And Segment Forecasts, 2022 – 2030," ID-GVR454977, May 2023. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/recruitment-software-market

APPENDIX : Turnitin Report

IT21158186_A.R.D.Pinsara_24-25J-047.docx

ORIGINALITY REPORT

2%	2%	0%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES