International Journal of Computer Science Trends and Technology (IJCST) – Volume 14 Issue 2, Mar-Apr 2026

RESEARCH ARTICLE OPEN ACCESS

An Intelligent Resume Screening and Candidate Ranking Framework Using AI with Edge–Cloud Hybrid Architecture

Kushal B N[1], Madhu G V[2], Rajashekara E[3], Rushitha R[4], Sirisha R[5]
[1,2,3,4,5] Department of Computer Science and Engineering, Constituent College of VTU
Chintamani, Karnataka

ABSTRACT

The rapid expansion of digital recruitment platforms has significantly increased the number of resume submissions, making manual screening a time-consuming process that is often affected by errors and human bias. Conventional Applicant Tracking Systems (ATS) mainly depend on fixed keyword matching techniques and fail to capture the contextual meaning of candidate profiles. This paper presents an AI-driven resume screening and candidate ranking framework based on an Edge–Cloud hybrid architecture that combines machine learning algorithms, Natural Language Processing (NLP), and cloud-based data management. In this approach, edge devices handle preliminary preprocessing tasks to minimize response time and reduce system latency, while cloud services perform advanced operations such as skill extraction, experience evaluation, resume classification, and candidate ranking using intelligent models. Firebase Authentication is used to ensure secure user access, and Firestore supports real-time storage and management of candidate data. Experimental analysis conducted using a publicly available Kaggle resume dataset shows that the proposed system delivers higher accuracy, reduced processing time, and more reliable candidate ranking when compared to traditional rule-based ATS methods. The hybrid model also improves scalability, operational efficiency, and practical usability for enterprises, educational institutions, and employment support organizations.

***Keywords* :—** Resume Screening, Artificial Intelligence, Edge Computing, Cloud Computing, Natural Language Processing, Firebase, Machine Learning, Hiring Automation.

I. INTRODUCTION

The primary contributions of this research work are as follows:

- A novel Edge–Cloud hybrid architecture designed specifically for resume screening to improve processing efficiency and reduce cloud dependency.
- Integration of AI-based NLP and machine learning models for contextual skill extraction, semantic interpretation and candidate scoring.
- Development of a weighted scoring and ranking algorithm that supports real-time candidate evaluation for different job roles.
- Demonstrated accuracy and performance improvements over traditional ATS and cloud-only systems using experimental evaluation on a Kaggle resume dataset.
- Use of a publicly available dataset to ensure transparency, reproducibility and practical applicability for academic and industrial settings.

Recruitment remains one of the most time-intensive aspects of organizational operations. HR teams receive hundreds of resumes per job opening, leading to delays, inconsistencies, fatigue errors and unconscious bias. The extensive use of online job portals has further increased candidate volumes, making conventional manual screening unsustainable.

Existing ATS solutions partially automate this process but primarily rely on keyword matching, causing relevant candi- dates to be overlooked and irrelevant ones to be shortlisted. These systems also struggle with semantic understanding, con- textual skill interpretation and structured candidate ranking. As a result, organizations often miss out on capable candidates while spending significant effort on low-quality applications. Recent advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have enabled automated systems to understand context, extract meaningful patterns and evaluate resumes based on skills and experience rather than mere keyword presence [1], [2]. However, most AI-driven solutions are completely cloud-dependent, resulting in higher latency, increased bandwidth usage and recurring operational costs. Privacy concerns related to sensitive candidate data further complicate cloud-only deployments.

To address these challenges, this paper introduces an AI-powered resume screening and ranking system using a hybrid Edge–Cloud architecture. The edge device performs local preprocessing, thereby reducing cloud load and minimizing latency. The cloud handles intensive AI tasks such as skill extraction, semantic analysis and ranking. This distributed architecture ensures efficient, scalable and secure recruitment automation suitable for enterprises, educational institutions and NGOs involved in placement support.

II. LITERATURE REVIEW

Early resume screening approaches relied on string matching and keyword frequency analysis, resulting in low accuracy and inability to interpret candidate context. Traditional ATS systems focus on rule-based filtering of terms like job titles, degrees and tool names. Such systems often ignore synonyms, variations in phrasing and domain-specific terminology, leading to unstable results [3].

Machine learning (ML) methods introduced classification algorithms such as Na¨ıve Bayes, Support Vector Machines (SVM) and Logistic Regression for resume categorization and suitability prediction [4], [5]. These approaches improved per- formance by learning statistical patterns from labeled resumes. However, they still rely on bag-of-words representations and cannot fully capture deeper semantics or long-range depen- dencies in text.

The emergence of NLP techniques, particularly word em- beddings and contextual models, significantly enhanced semantic understanding. Word2Vec and GloVe embeddings en- abled similarity-based skill matching and clustering of related terms [6]. More recent transformer-based models such as BERT, RoBERTa and their domain adaptations provide contex- tualized embeddings that can interpret skills, responsibilities and achievements with higher fidelity [7], [8]. Several works have applied BERT for job recommendation, skill extraction and professional profile matching.

Cloud-based resume parsing solutions, including commer- cial APIs built on Google Cloud NLP and AWS Comprehend, offer scalable document understanding services [9]. These systems provide language detection, entity recognition and sentiment analysis. However, their dependence on remote infrastructure introduces latency and bandwidth overheads, and subscription pricing may not be feasible for academic institutions or small organizations.

Edge computing has been explored to offload computa- tion closer to data sources, thereby reducing latency and load on centralized servers [10]. Research on Edge–Cloud collaboration shows performance benefits for IoT analytics, video surveillance and mobile applications [11]. Very limited work, however, addresses resume screening or recruitment au- tomation from an Edge–Cloud perspective. Most recruitment-related research continues to assume cloud-only processing and does not consider constrained or intermittent network environments.

Existing literature also highlights the exploration of ex- plainable AI (XAI) for recruitment fairness, fairness-aware ranking and bias mitigation [12], [13]. While these topics are important, many works stop at algorithm design and do not integrate them into deployable systems with secure authentication and data storage.

Considering the above gaps, this paper contributes a com- prehensive Edge–Cloud hybrid pipeline for automated resume screening. It combines edge-side preprocessing, cloud-based contextual NLP, ML scoring and ranking, and secure storage using Firebase. The system is evaluated on a Kaggle resume dataset, demonstrating practical applicability and improved effectiveness over rule-based ATS baselines [14], [15].

III. Proposed System Architecture

The proposed system uses a two-layer architecture: edge and cloud. The edge layer performs file validation and lightweight preprocessing, while the cloud layer handles heavy NLP and machine learning tasks. Fig. 1 shows the high-level architec- ture.

A. Edge Layer

The edge layer is typically the client device (web browser or mobile app) that accepts resume files (PDF/DOCX). Tasks at the edge include:

- Format and size validation of uploaded files.
- Local text extraction using PyMuPDF or Apache Tika libraries.
- Basic cleaning: removal of special characters, numbers and non-ASCII symbols.

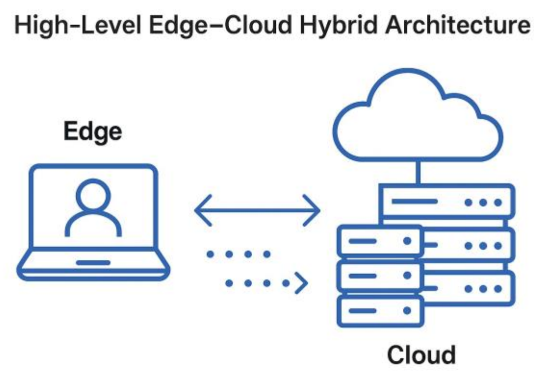

Fig. 1. High-level Edge–Cloud Hybrid Architecture.

- Stopword removal based on a predefined English stop- word list.
- Tokenization and lightweight feature extraction.
- Secure transmission of cleaned text and metadata to cloud APIs.

Advantages of edge processing include reduced bandwidth usage, decreased cloud load and improved responsiveness, particularly in low-connectivity or high-latency networks.

B. Cloud Layer

The cloud layer contains the heavy processing components:

- NLP pipelines for skill extraction, experience detection, education parsing and named entity recognition (NER).
- ML models (Logistic Regression, Random Forest and a BERT-embedding classifier) for job-fit prediction.
- Ranking engine implementing a configurable weighted scoring function.
- Firestore database for storing resume features, scores and ranking results.
- Firebase Authentication for user and role-based access control.
- HR dashboard for visualization and export of ranked candidate lists.

The Edge–Cloud communication is implemented using RESTful APIs. Only processed features and essential metadata are transmitted, reducing bandwidth while preserving privacy.

IV. Methodology

The methodology of the proposed system is divided into several stages involving both edge and cloud operations. Fig. 2 illustrates the end-to-end workflow, beginning from resume upload to the final ranked output.

A. Data Preprocessing at Edge

At the client side, preprocessing reduces computation on the cloud and enhances responsiveness. Steps include:

- **Format Validation:** Identification of acceptable file types (PDF/DOCX) and size limitations.

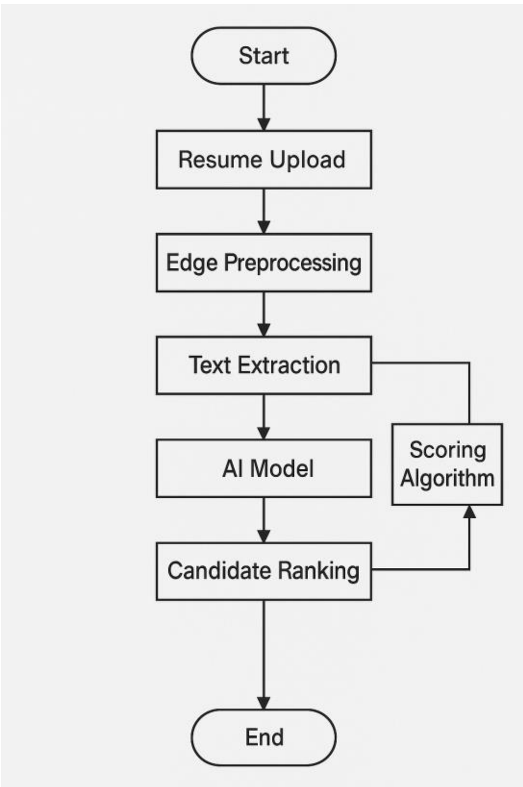

Fig. 2. Overall workflow of the proposed resume screening system.

- **Text Extraction:** Parsing of resumes using PyMuPDF or Apache Tika. For scanned PDFs, OCR (Tesseract) can be integrated.
- **Text Cleaning:** Removal of punctuation, HTML tags and non-informative symbols.
- **Stopword Removal and Tokenization:** Eliminating common words and splitting text into tokens for later processing.
- **Lightweight Vectorization:** TF or hash-based sparse vectors are created for quick preliminary inspection if needed.

B. Cloud NLP Pipeline

Cloud-side NLP is responsible for understanding resume content at a deeper semantic level.

- **NER:** Identification of entities such as technical skills, tools, programming languages, job titles, organizations and locations [6].
- **Semantic Embeddings:** Generation of contextual embeddings for sentences and skill phrases using a pre-trained BERT model fine-tuned on professional text [7], [8].
- **Experience Extraction:** Detection of numerical years of experience or inferred experience range from description segments.
- **Education Parsing:** Extraction of highest degree, institution and graduation year.
- **Soft-Skill Detection:** Identification of soft skills such as communication, leadership and teamwork using a phrase-based classifier.

C. Machine Learning Model

The extracted features are mapped to numerical vectors suitable for classification and scoring:

- One-hot or multi-hot encodings for skill presence.
- Normalized numerical values for experience years.
- Ordinal encoding for education level (e.g., diploma, bachelor, master).
- Binary indicators for presence of certifications and internships.

Several models are evaluated, including Logistic Regression, Random Forest, SVM and a shallow neural network using BERT embeddings as input. Hyper-parameters are tuned using grid search and 5-fold cross-validation. The Random Forest model offered a good trade-off between performance and interpretability, while the BERT-based classifier achieved the highest accuracy.

D. Scoring Algorithm

A weighted scoring model evaluates candidates based on four components:

- S_s: skills match percentage.
- S_e: normalized years of relevant experience.
- S_{edu}: education level score.
- S_c: certifications and internship relevance.

The total score is computed as:

$$Score = W_s S_s + W_e S_e + W_{edu} S_{edu} + W_c S_c, \qquad (1)$$

where W_s, W_e, W_{edu} and W_c are non-negative weights that sum to 1. HR administrators can adjust the weights according to job requirements, allowing role-specific ranking.

E. Candidate Ranking

For each job posting, all screened candidates receive a score computed by the trained model and scoring function. The ranking engine sorts candidates in descending order of score. The dashboard displays:

- overall job-fit percentage,
- contribution of each component to the score,
- top skills and missing skills for each candidate.

A feedback mechanism allows HR users to mark candidates as “suitable” or “not suitable”. These labels can be fed back into the training dataset for incremental model improvement.

V. Dataset and Experimental Setup

A. Dataset Overview

The experimental evaluation uses a publicly available resume dataset hosted on Kaggle [14]. The dataset contains text versions of resumes collected from multiple job domains. For this study, a subset focused on IT-related profiles was used.

Table I summarizes the dataset characteristics.

TABLE I
Summary of Kaggle Resume Dataset Used in This Study

Category	Count	Format	Avg. Skills
Developer	90	PDF/DOCX	18
Analyst	60	PDF/DOCX	15
Tester	50	PDF/DOCX	12
Others	40	PDF/DOCX	10
Total	240	–	14.2

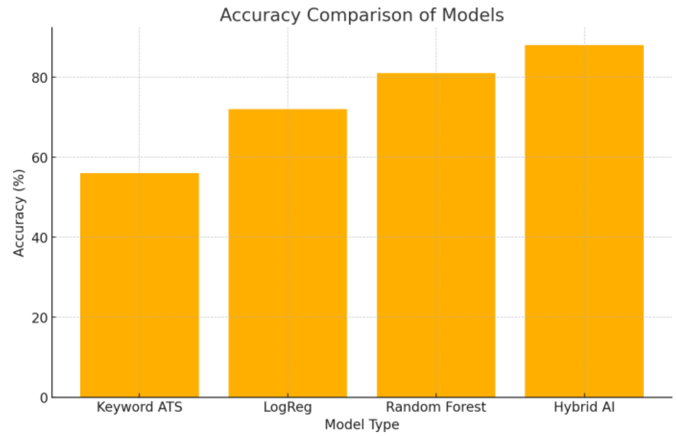

Fig. 3. Accuracy comparison of different models.

TABLE III
Processing Time Comparison

Resumes were converted to plain text using automated parsers. Labels corresponding to job categories were retained from the dataset and used for supervised training.

B. Experimental Setup

The dataset was split into 70% training and 30% testing. Model training was performed on a workstation with Intel Core i5 CPU and 8 GB RAM. For BERT-based experiments, Google Colab GPU was used. The cloud back-end was implemented using FastAPI with Firebase Firestore as the database service and Firebase Authentication for secure login. Evaluation metrics include accuracy, precision, recall and F1- score, along with average processing time per resume.

VI. Results and Data Visualization

A. Model Performance

Table II summarizes the accuracy of different approaches.

TABLE II
MODEL ACCURACY COMPARISON

Model	Accuracy (%)
Keyword-based ATS	56
Logistic Regression	72
Random Forest	81
Hybrid AI System (BERT + RF)	**88**

Fig. 3 provides a visual comparison of accuracies.

B. Processing Time

Average processing time per resume was measured for cloud-only and Edge–Cloud deployments. Results are shown in Table III.

Fig. 4 visualizes the reduction in processing time.

System Type	Avg. Time (s)
Cloud-only	4.8
Edge–Cloud Hybrid	**2.1**

C. Discussion

The results indicate that the proposed hybrid system clearly outperforms a keyword-based ATS baseline in terms of ac- curacy and ranking consistency. The integration of contextual embeddings and ML models enables better understanding of skill descriptions and job responsibilities.

The Edge–Cloud deployment reduces average processing time by more than half compared to a pure cloud implemen- tation. This improvement is particularly beneficial in campus recruitment or job-fair scenarios where large volumes of resumes must be screened in a short time.

Limitations include the moderate size of the dataset and reliance on English resumes. Future work will explore multi- lingual datasets and robustness to noisy or incomplete resumes.

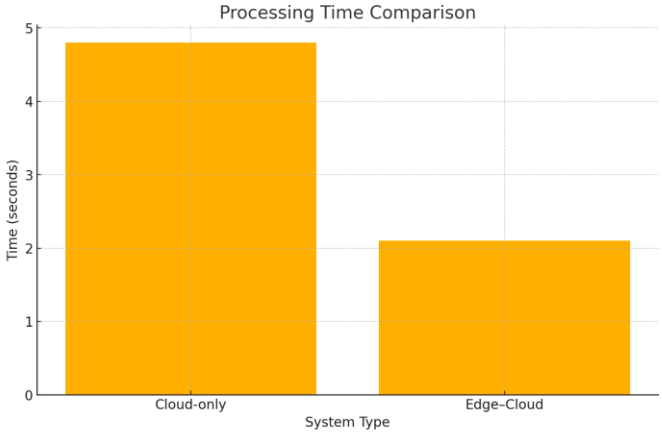

Fig. 4. Average processing time per resume for cloud-only and Edge–Cloud hybrid systems.

Integration of XAI techniques such as LIME or SHAP can also provide interpretable explanations for HR decision makers [12], [13].

VII. Conclusion and Future Work

This paper proposed an AI-powered resume screening and ranking system using an Edge–Cloud hybrid architecture. By combining edge-side preprocessing with cloud-based contex- tual NLP and machine learning models, the system improves accuracy, reduces processing time and offers scalable de- ployment for organizations of different sizes. Experiments on a Kaggle resume dataset show that the proposed approach achieves 88% accuracy and significantly lowers latency com- pared to cloud-only solutions.

Future work includes expanding the dataset with more domains and languages, incorporating fairness and bias-mitigation modules, adding explainable AI components for transparent decision support and integrating the system into full recruitment life-cycle platforms with interview scheduling and feedback analytics.

References

[1] J. Smith, “AI in recruitment: A review,” *IEEE Access*, vol. 11, pp. 1–15, 2023.

[2] A. Kumar and P. Sharma, “Intelligent screening of job applications using NLP,” in *Proc. 2022 Int. Conf.*

Computational Intelligence, pp. 210–215.

[3] L. Brown, “Limitations of keyword-based applicant tracking systems,” *J. Human Resource Technol.*, vol. 8, no. 2, pp. 33–42, 2020.

[4] M. Patel, “Machine learning approaches for automated recruitment systems,” *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–28, 2021.

[5] R. Kumari and S. Gupta, “Resume classification using NLP and deep learning,” *Int. J. Comput. Appl.*, vol. 176, no. 42, pp. 12–18, 2020.

[6] A. Das, S. Roy and P. Saha, “Semantic skill extraction from professional documents using word embeddings,” in *Proc. 2021 IEEE Int. Conf. Data Mining Workshops*, pp. 450–457.

[7] A. Das *et al.*, “Semantic skill extraction from professional documents using BERT,” *IEEE Trans. Computational Social Systems*, vol. 9, no. 4, pp. 955–968, 2022.

[8] Y. Li and H. Chen, “BERT-based job-resume matching for online recruitment,” *Expert Systems with Applications*, vol. 207, 2022.

[9] Google Cloud, “Cloud natural language: Documentation,” 2021. [On- line]. Available: https://cloud.google.com/natural-language

[10] Y. Chen, “Edge computing for intelligent decision-making systems,” *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1854–1865, 2021.

[11] T. Li and J. Wang, “Optimizing cloud-based analytics with edge collab- oration,” *J. Inf. Sci.*, vol. 48, no. 6, pp. 923–937, 2022.

[12] K. Singh and R. Mehta, “Explainable AI for fair recruitment,” in *Proc. 2023 IEEE Conf. Trustworthy AI*, pp. 89–96.

[13] S. George and V. Nair, “Bias-aware ranking models for hiring,” *Pattern Recognition Letters*, vol. 167, pp. 63–70, 2023.

[14] Kaggle, “Resume dataset,” 2020. [Online]. Available: https://www.kaggle.com

[15] M. Rao and D. Kulkarni, “Automated screening of IT resumes using machine learning,” in *Proc. 2021 Int. Conf. Smart Systems*, pp. 300– 305.

[16] S. Zhang *et al.*, “Deep representation learning for job candidate match- ing,” *Knowl.-Based Syst.*, vol. 250, 2022.

[17] N. Khan and F. Hussain, “NLP pipelines for recruitment analytics,” *Applied Soft Computing*, vol. 118, 2022.

[18] P. Roy, “Hybrid edge–cloud systems for NLP,” in *Proc. 2022 Int. Conf. Cloud Engineering*, pp. 101–108. D. Lopez and A. Garcia, “Real-time data processing with edge com- puting: A survey,” *Future Generation Computer Systems*, vol. 135, pp. 45–60, 2022.

[19] S. Hegde and M. Bhat, “Design of a campus recruitment analytics platform,” in *Proc. 2023 Int. Conf. Emerging Computation*, pp. 55–60.

[20] IEEE, “IEEE referencing guidelines for conference papers,” 2020. [Online]. Available: https://ieee.org

[21] N. Verma and A. Rao, “Performance evaluation of hybrid cloud archi- tectures,” *Int. J. Cloud Computing*, vol. 11, no. 1, pp. 1–16, 2023.

Author Profiles:

Mr. Kushal B N is currently pursuing the 4th semester of the Bachelor of Engineering (B.E.) program in Computer Science and Engineering (Artificial Intelligence and Machine Learning) at the Constituent College of VTU, Chintamani. He completed his Secondary School Leaving Certificate (SSLC) from the Karnataka Secondary Education Examination Board in 2024. He has actively participated in the National Conference (NCRIE 2025) held at KS Institute of Technology, No. 14, Raghuvanahalli, Kanakapura Road, Bengaluru – 560109. He has a strong interest in emerging technologies such as Artificial Intelligence, Machine Learning, Big Data, Deep Learning, Cloud Computing, the Internet of Things (IoT), and other related domains.

Mr. Madhu G V is currently pursuing the 4th semester of the Bachelor of Engineering (B.E.) program in Computer Science and Engineering (Artificial Intelligence and Machine Learning) at the Constituent College of VTU, Chintamani. He is passionate about learning and exploring advanced technologies in the field of computer science. He has actively participated in technical events, workshops, and academic conferences that enhance his practical knowledge and professional skills. His areas of interest include Artificial Intelligence, Machine Learning, Data Science, Cloud Computing, Deep Learning, Cyber Security, and the Internet of Things (IoT). He is dedicated to building innovative solutions and continuously improving his technical expertise in emerging technologies.

Mr. Rajashekara Edigara is currently pursuing the 4th semester of the Bachelor of Engineering (B.E.) program in Computer Science and Engineering (Artificial Intelligence and Machine Learning) at the Constituent College of VTU, Chintamani. He completed his Secondary School Leaving Certificate (SSLC) under the Karnataka Secondary Education Examination Board in 2024 and pursued his Pre-University education at Morarji Desai Residential PU College, Arooru, Brahmavara (Tq), Udupi (Dist). He has a strong passion for learning emerging technologies and continuously works to enhance his technical knowledge and practical skills. His areas of interest include Artificial Intelligence, Machine Learning, Big Data, Deep Learning, Cloud Computing, the Internet of Things (IoT), Application Development, Game Development, Full Stack Development, and real-time project implementation. He is highly enthusiastic about building innovative software solutions and exploring modern technologies that solve real-world problems.

Ms. Rushitha R is currently pursuing her 4th semester of the Bachelor of Engineering (B.E.) program in Computer Science and Engineering at the Constituent College of VTU, Chintamani. She completed her 12th standard in PCMB at Royal Composite PU College, Chintamani, Karnataka, affiliated with the Karnataka School Examination and Assessment Board, Bengaluru. She obtained her Secondary School Leaving Certificate (SSLC) from the Karnataka Secondary Education Examination Board in 2022 and joined Royal Composite PU College in the same year for her Pre-University education. She has a strong interest in emerging technologies such as Artificial Intelligence, Machine Learning, Cloud Computing, Cyber Security, Data Science, and related fields. She is passionate about learning new technologies, improving her technical skills, and exploring innovative solutions in the field of computer science and engineering.

Ms. Sirisha S.N is currently pursuing her 4th semester of the Bachelor of Engineering (B.E.) program in Computer Science and Engineering (Artificial Intelligence and Machine Learning) at the Constituent College of VTU, Chintamani. She completed her Secondary School Leaving Certificate (SSLC) from the Karnataka Secondary Education Board in 2022 and joined the Pre-University Course (PUC) in PCMB at Sri Vijaya PUC College, Chintamani, in the same year. She has a strong interest in emerging technologies such as Artificial Intelligence, Machine Learning, Cloud Computing, Internet of Things (IoT), Deep Learning, Quantum Computing, and other related fields. She is passionate about learning new technologies, improving her technical skills, and exploring innovative solutions in the field of computer science and engineering.