“Loom – The Next-Gen Smart Recruitment Portal”

A PROJECT WORK REPORT SUBMITTED TO

THE NATIONAL INSTITUTE OF ENGINEERING
(An Autonomous Institution Under VTU)

In fulfillment of the requirements for major project work, Seventh semester
Bachelor of Engineering in Computer Science & Engineering
Submitted By

Aayush M D (4NI22CS004)

Under The Guidance Of

Mrs. Chaitra M
Assistant Professor
Department of CS&E,
NIE Mysuru - 570008

DEPARTMENT OF COMPUTER SCIENCE ANDENGINEERING
THE NATIONAL INSTITUTE OF ENGINEERING
(An Autonomous Institution Under VTU)
Mananthavadi Road, Mysuru- 570008

2025-2026

THE NATIONAL INSTITUTE OF ENGINEERING
(An Autonomous institution, affiliated to VTU)
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Mysore -570008

CERTIFICATE

This is to Certify that the project work entitled **"Loom - The Next-Gen Smart Recruitment Portal"** is a work carried out by **Aayush M D** (4NI22CS004) in fulfillment for major project work, seventh semester, Computer Science and Engineering**,** The National Institute of Engineering (Autonomous under VTU) during the academic year 2025–2026. It is certified that all corrections and suggestions indicated for the Internal Assessment have been incorporated in the report deposited in the department library. The major project work report has been approved in fulfillment as per academic regulations of The National Institute of Engineering, Mysuru**.**

Signature of Internal Guides

Mrs. Chaitra M
Assistant Professor
Dept. of CS&E
NIE, Mysuru

Signature of HOD

Dr. Anitha R
Professor and Head
Dept. of CS&E
NIE, Mysuru

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible whose constant guidance and encouragement crown all efforts. First and fore most we would like to thank our beloved principal **Dr. Rohini Nagapadma** for being the patron and the beacon light for this project.

We would like to express our sincere gratitude to our HOD **Dr. Anitha R,** Department of CSE, NIE for her relentless support and encouragement. It gives us immense pleasure to thank our guide **Mrs. Chaitra M,** Assistant Professor, Department of CSE, NIE for her valuable suggestions and guidance during the process of the project and for having permitted us to pursue work on the subject.

-Aayush M D

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Overview

The project titled *"Loom - The Next-Gen Smart Recruitment Portal"* is designed to modernize and streamline the hiring process using advanced Artificial Intelligence and Machine Learning techniques. Unlike traditional recruitment systems, which largely depend on keyword-based searches, Loom leverages semantic similarity models to capture the actual context between resumes and job descriptions. This approach reduces the gap between candidate expectations and recruiter requirements, ensuring more accurate matches. The system integrates resume parsing, explainable AI models, and fairness auditing to maintain transparency and avoid bias. Additionally, it incorporates an interactive chatbot for preliminary screening, thereby saving time and effort for recruiters. By implementing skill gap prediction models, the portal not only shortlists candidates but also guides them on improving their employability. The entire platform is designed as a scalable, cloud-ready system using FastAPI for backend services, Next.js for frontend, and Docker for deployment consistency. The project represents a holistic and future-ready recruitment ecosystem that benefits both recruiters and applicants.

1.2 Objective

The primary objective of this project is to develop a smart recruitment portal that bridges the inefficiencies of traditional hiring systems. It aims to implement semantic embeddings for resume-job matching, thereby overcoming the limitations of keyword filtering. The project also seeks to provide candidates with structured feedback and insights into their skill gaps. Fairness and transparency are core goals, with the system integrating explainable metrics to reduce bias in hiring decisions. Another objective is to build an ATS-friendly resume parser that efficiently extracts structured data. Furthermore, the project intends to deploy a screening chatbot to handle initial candidate assessments. By incorporating predictive models for career guidance, it helps candidates enhance their profiles. The system is also designed with scalability in mind, ensuring deployment across cloud environments. Ultimately, the objective is to create a recruitment solution that is intelligent, fair, and user-centric.

1.3 Scope of the Project

The scope of Loom extends beyond simple job-candidate matching and addresses broader hiring challenges faced by organizations. It covers semantic analysis of resumes and job postings to improve shortlisting accuracy. The portal provides an Applicant Tracking System-compatible environment for recruiters to manage applications efficiently. Candidates benefit from automated feedback and career development insights through skill prediction models. The chatbot interface expands the system's scope by offering conversational preliminary

interviews. From a technical standpoint, the project incorporates cloud deployment, enabling organizations of different scales to adopt it seamlessly. It also integrates explainability in AI models to foster trust among recruiters and applicants. The scope further extends to fairness auditing, ensuring unbiased and inclusive hiring practices. As the platform evolves, it can integrate video interview simulations and advanced analytics, making it a comprehensive hiring ecosystem.

1.4 Existing System

Most existing recruitment portals are built around keyword-based search and filtering mechanisms. These systems often rank resumes higher if they are filled with matching keywords, regardless of actual relevance to the job. As a result, recruiters are overwhelmed with false positives while suitable candidates may get rejected due to insufficient keyword overlap. There is minimal consideration of fairness, and demographic bias often creeps into the process. Moreover, candidates rarely receive feedback, leaving them unaware of areas for improvement. Recruiters also spend excessive time manually screening irrelevant applications. The existing systems are therefore inefficient, opaque, and limited in scope.

1.4.1 Drawbacks of Existing System

The current systems suffer from several critical drawbacks that hinder effective recruitment. Keyword dependency leads to inaccurate candidate-job matching, causing both false positives and false negatives. Candidates often face rejection without explanation, which reduces trust and discourages improvement. There is no mechanism to address fairness or prevent bias, making shortlists potentially discriminatory. Recruiters face significant manual workload while navigating through irrelevant resumes. The lack of semantic understanding of resumes and job descriptions weakens the reliability of the process. Moreover, existing systems fail to predict candidate skill gaps or suggest improvement pathways. This creates inefficiency and dissatisfaction on both sides of the hiring pipeline.

1.5 Proposed System

The proposed system introduces a modern, AI-driven recruitment platform that resolves the limitations of traditional methods. By leveraging transformer-based semantic embeddings, it ensures accurate, context-aware candidate-job matching. It integrates resume parsing tools to structure data efficiently for recruiters. The system features a chatbot that automates the preliminary screening process, saving recruiter time. Fairness metrics and explainability tools are embedded to ensure unbiased and transparent hiring. Additionally, the platform predicts skill gaps, providing candidates with actionable improvement suggestions. Built on scalable cloud-ready architecture, it can serve organizations of varied sizes. Overall, the proposed system represents a fair, efficient, and intelligent approach to recruitment.

1.5.1 Advantages of Proposed System

The new system provides multiple advantages that directly address recruiter and candidate concerns. Semantic matching ensures precise alignment between job requirements and candidate qualifications. Resume parsing simplifies data extraction and reduces manual errors. The screening chatbot accelerates the hiring pipeline by automating early assessments. Fairness monitoring enhances trust by minimizing bias in decision-making. Candidates receive structured feedback, which helps them improve career readiness. Recruiters save valuable time and resources through automated tools. The scalable architecture ensures adaptability for diverse organizations. Collectively, these advantages create a more efficient, transparent, and future-ready recruitment process.

Chapter 2

LITERATURE SURVEY

[1] "Information Extraction from Free-Form CV Documents in Multiple Languages", Davor Vukadin, Adrian Satja Kurdija, Goran Delač, Marin Šilić, IEEE Access, 2021.

Description:

This paper develops a multilingual pipeline for extracting structured information from highly varied CV layouts, combining document segmentation, field classification and sequence labeling to capture personal data, education, work experience and skills. The authors emphasize robustness to template diversity and noisy, scanned PDFs, and describe a hybrid approach that mixes learned models with post-processing rules to normalize outputs. Trained on real-world CVs supplied by industry, the system shows strong generalization across languages and formats, demonstrating practical viability for applicant-tracking workflows. The article also discusses error patterns (e.g., overlapping entities and sparse labels) and design choices to mitigate them, making it a useful reference for production-grade resume parsing.

[2] "A Method for Resume Information Extraction Using BERT-BiLSTM-CRF", Xiaowei Li, Hui Shu, Yi Zhai, Zhiqiang Lin, 2021 IEEE 21st International Conference on Communication Technology(ICCT).

Description:

The authors propose a named-entity extraction model tailored to resumes that fuses contextual embeddings from BERT with a BiLSTM-CRF decoding layer. By leveraging BERT to capture long-range context and a CRF to enforce valid label sequences, the approach improves field tagging for names, schools, companies, positions and dates. The paper reports competitive gains over traditional CRF and LSTM baselines on an annotated resume corpus, particularly in sections with dense, domain-specific terminology. It also details preprocessing for tabular/irregular layouts and post-processing rules for date and organization normalization, yielding outputs that are immediately usable in ranking and matching pipelines.

[3] "A Résumé Evaluation System Based on Text Mining", Chih-Han Chou, Chih-Hsuan Yu, 2019 IEEE International Conference on Artificial Intelligence in Information and Communication(ICAIIC).

Description:

This study builds an end-to-end framework that ingests resumes and job descriptions, applies text mining (tokenization, TF-IDF/keyword weighting, similarity computation) and produces

quantitative suitability scores to support shortlisting. The system emphasizes explainability by surfacing the terms driving a candidate's score, enabling recruiters to audit mismatches in skills and experience. Prototype experiments demonstrate consistent ranking behavior across roles and practical latency for interactive use. The paper contributes a modular architecture that can be extended with topic models or neural encoders, and it highlights operational considerations like handling missing fields and standardizing section headers common in real resumes.

[4] "The Resume Corpus: A Large Dataset for Research in Information Extraction Systems", Fan Su, Wenyuan Zhang, Changlu Lu, Jialei Li, Xujian Xu, 2019 15th International Conference on Computational Intelligence and Security (CIS).

Description:

Addressing the scarcity of public data for resume IE, this work curates a large-scale corpus with standardized annotations over common fields (personal info, education, work history, skills) and documents the collection, labeling protocol and splits. Baseline experiments with CRF and neural taggers establish reference performance and reveal challenging phenomena such as heterogeneous layouts, multi-line entities and mixed languages. By releasing detailed schema definitions and evaluation guidelines, the paper enables reproducible comparisons among extraction models. The dataset and benchmarks aim to catalyze progress on robust parsing that transfers across industries and resume templates.

[5] "Ontologies to Model User Profiles in Personalized Job Recommendation", M. Rimitha, P. Prashanthi, Udaya Lakshmi Ravi, 2018 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER).

Description:

This paper presents an ontology-driven user profile model that encodes skills, roles, qualifications and experience to support personalized job recommendations. Using semantic relationships between competencies and occupations, the system can infer related skills and alleviate cold-start scenarios where explicit data are sparse. The authors detail the ontology construction and reasoning workflow, and show how semantic similarity over the knowledge graph enhances matching beyond bag-of-words methods. The approach also improves transparency by enabling rule- and relation-level explanations for recommended jobs—an aspect valuable for recruiter and candidate trust.

[6] "Automated Analysis and Prediction of Job Interview Performance", Iftekhar Naim, Md Iftekhar Tanveer, Daniel Gildea, Mohammed Ehsan Hoque, IEEE Transactions on Affective Computing, 2016.

Description:

Focusing on the interview stage of hiring, this study builds a multimodal system that quantifies verbal (lexical/prosodic) and non-verbal (facial, head motion) cues from recorded interviews to predict human ratings of traits and overall performance. Trained on real interview data with multiple independent judges, the models achieve strong correlations with expert assessments and identify which features—such as speaking rate, disfluencies and smiles—drive perceived effectiveness. The authors discuss ethical use and the importance of feedback, positioning the technology as a tool to support coaching and screening rather than to replace human judgment. The paper remains a foundational reference for fair, explainable interview analytics.

Chapter 3

SYSTEM REQUIREMENTS AND SPECIFICATION

3.1 Software Requirements

- Operating System: Windows, Linux, macOS
- Drivers: Not Applicable
- Programming Language: JavaScript (React.js), Python (FastAPI)
- Tools: MongoDB

3.2 Hardware Requirements

- Memory: 4GB or higher
- Processor: Dual-core, 64-bit processor

3.3 Functional Requirements

1. Recruiter registration, login, and job posting creation.
2. Candidate registration, profile creation, and resume upload.
3. ATS-friendly resume parser to extract skills, education, and work experience.
4. Semantic matching of resumes with job descriptions using transformer embeddings.
5. Automated scoring and ranking of candidates based on similarity.
6. Preliminary screening through an AI-driven chatbot.
7. Dashboard for recruiters to view ranked lists, fairness metrics, and explainable outputs.
8. Candidate feedback generation with reasons for rejection and improvement suggestions.
9. Skill gap prediction and training recommendation for applicants.
10. Multi-user access with secure cloud deployment and centralized data management.

3.4 Non-Functional Requirements

1. **Scalability** – handle thousands of users and job postings simultaneously.
2. **Performance** – fast response time for parsing, scoring, and chatbot interaction.
3. **Reliability** – minimal downtime, backup and recovery support.
4. **Security** – encrypted data storage, secure login, and compliance with privacy laws.
5. **Usability** – intuitive, responsive UI accessible on web and mobile devices.
6. **Transparency** – explainable AI to justify candidate shortlisting and fairness checks.

7. **Portability** – deployable across operating systems and cloud providers using Docker.
8. **Maintainability** – modular, well-documented codebase for easy upgrades.
9. **Availability** – 24/7 accessible recruitment platform with high uptime.
10. **Extensibility** – ability to integrate future features like video interviews or analytics.

Chapter 4

SYSTEM DESIGN AND ANALYSIS

4.1 High Level Diagram

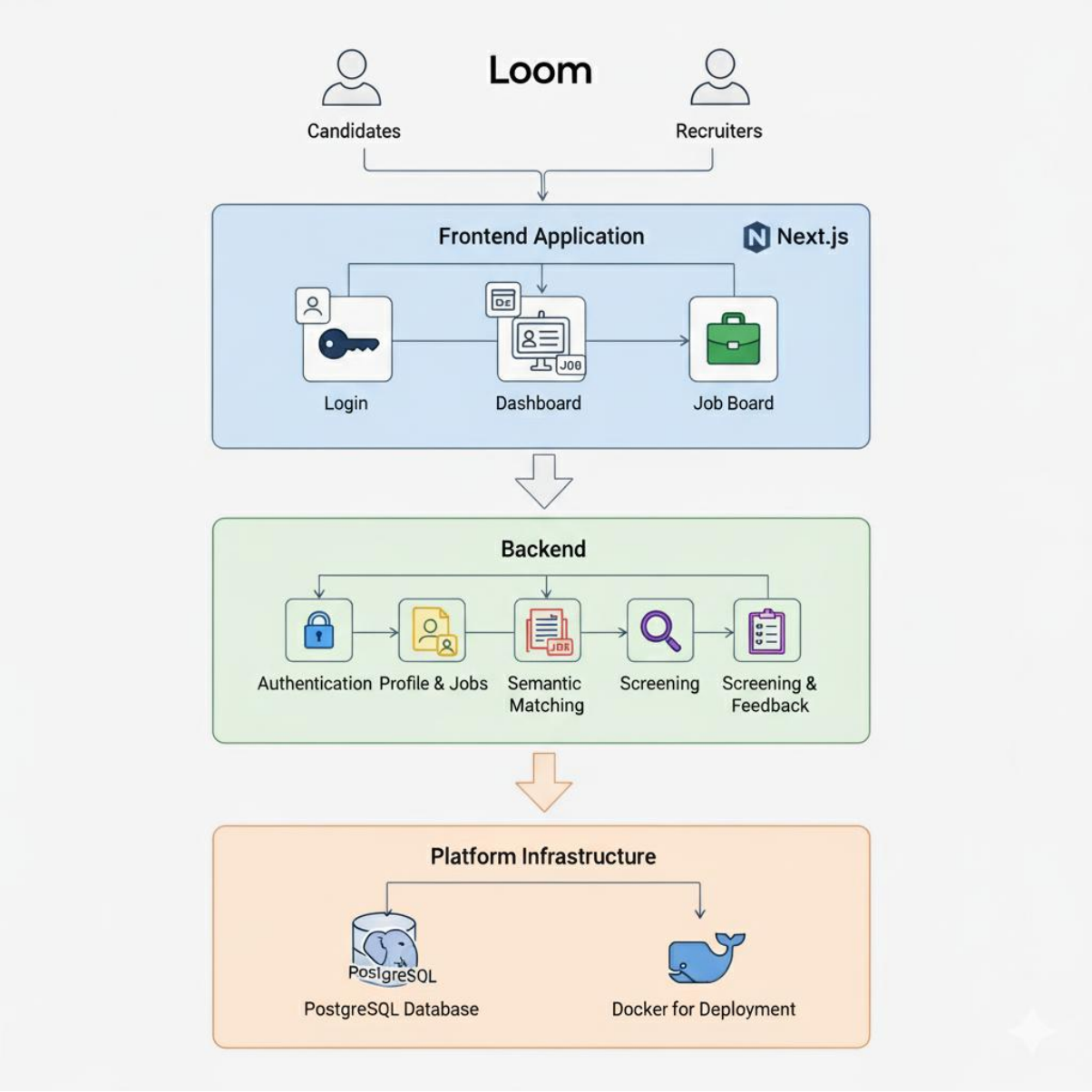

Fig 4.1: System architecture

4.2 Low Level Diagram

Use Case Diagram

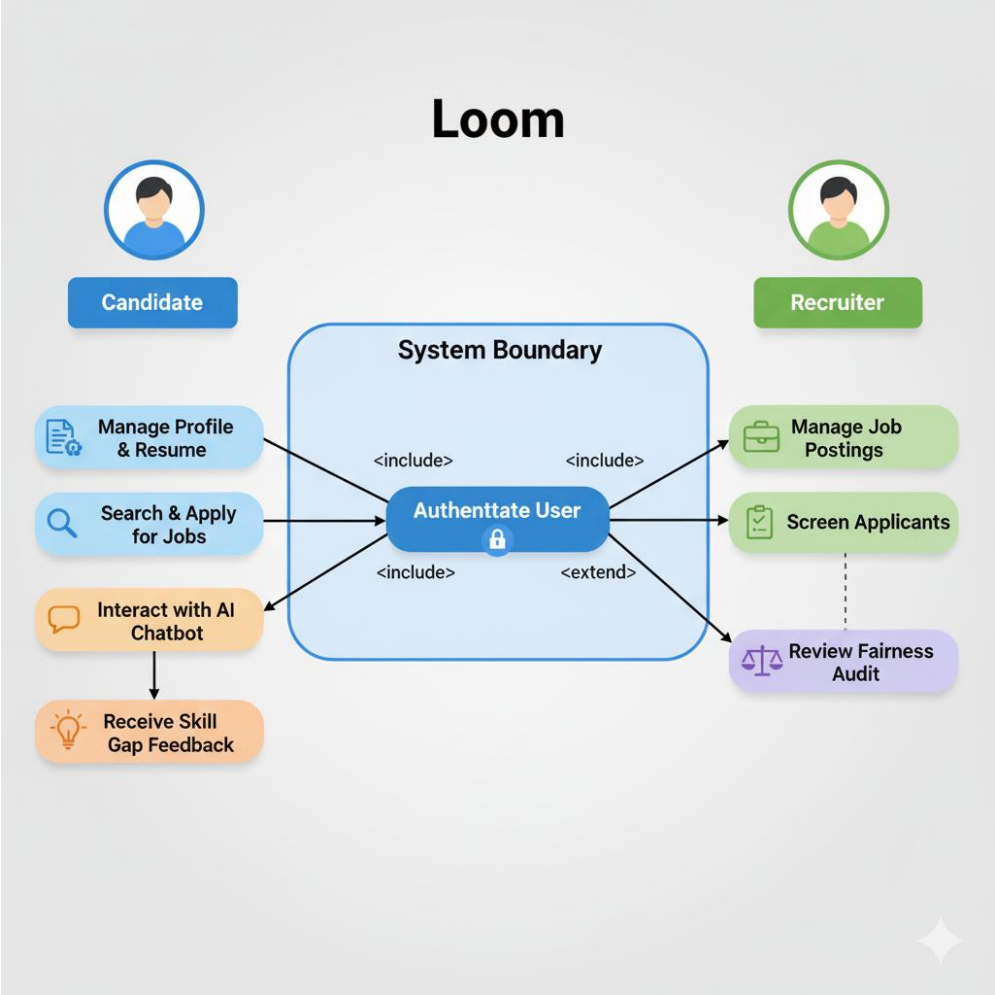

Fig. 4.2.1 Use Case Diagram

Data Flow Diagram

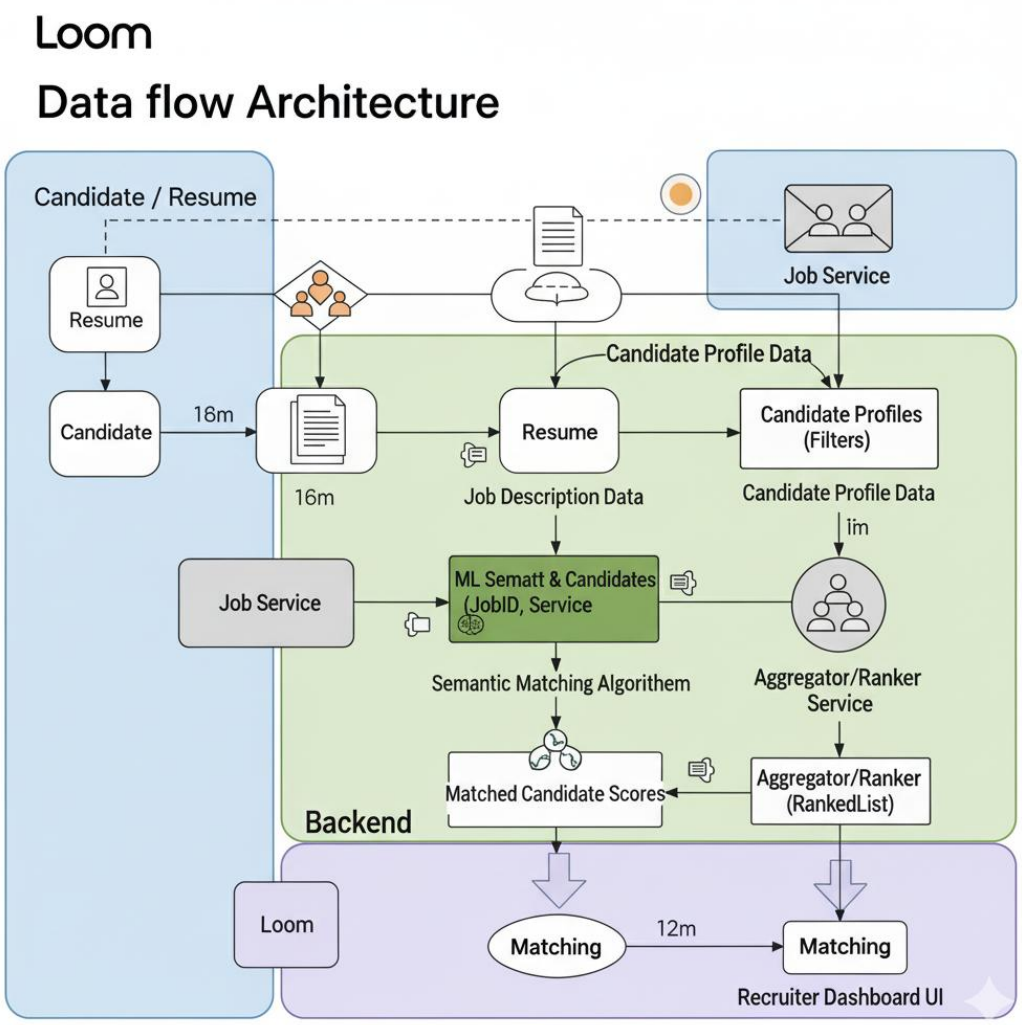

Fig. 4.2.2 Data Flow Diagram

Sequence Diagram

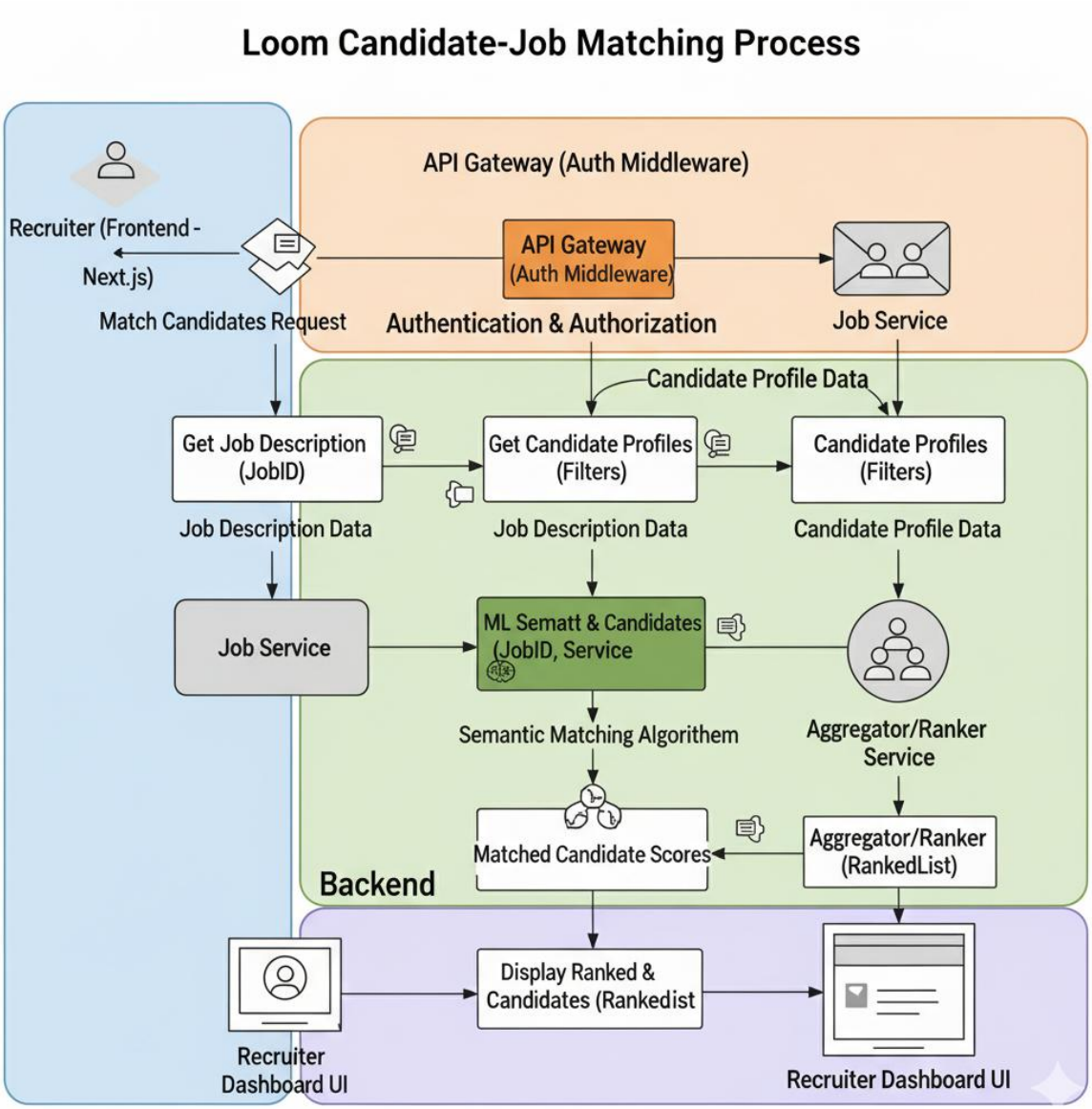

Fig. 4.2.3 Sequence Diagram

Architecture Overview

- **Frontend Layer**

  - Built using **Next.js, React, and Tailwind CSS** for a responsive and modern UI.
  - Provides dashboards for **candidates** (profile creation, resume upload, chatbot interaction) and **recruiters** (job posting, candidate review, fairness analysis).
  - Communicates with backend services through **RESTful APIs**.

- **Backend Layer**

  - Implemented using **FastAPI**, which acts as the core service hub.
  - Manages **authentication, request routing, and API integration** with ML models and databases.
  - Provides endpoints for resume parsing, semantic matching, scoring, chatbot processing, and skill-gap analysis.

- **Machine Learning & NLP Layer**

  - Uses **Sentence-BERT embeddings** for semantic resume-job matching.
  - Employs **spaCy / HuggingFace models** for resume parsing (skills, education, work experience).
  - Integrates **regression models (scikit-learn)** for scoring and ranking candidates.
  - Uses **SHAP/LIME** for explainability of recruitment decisions.
  - Includes a **screening chatbot** for preliminary assessments.
  - Predicts **skill gaps** and provides personalized career recommendations.

- **Database Layer**

  - **SQLite** for development and **PostgreSQL** for production-level deployment.
  - Stores candidate profiles, parsed resume data, recruiter job postings, scoring results, and feedback reports.
  - Ensures relational structure for consistency and advanced queries.

- **Deployment & Infrastructure Layer**

  - Uses **Docker** to containerize services for portability and consistency.
  - Scalable deployment on **cloud platforms (AWS, Azure, GCP)**.
  - Supports orchestration with **Kubernetes or Docker Compose** for load balancing and scaling.

- **Security & Compliance Layer**

- All communication secured via **HTTPS and data encryption**.
- **JWT-based authentication** and **Role-Based Access Control (RBAC)** to separate recruiter and candidate privileges.
- Compliance with **GDPR and data privacy standards**, including user consent and controlled retention policies.

REFERENCES

1. D. Vukadin, A. S. Kurdija, G. Delač, and M. Šilić, "Information Extraction from Free-Form CV Documents in Multiple Languages," *IEEE Access*, vol. 9, pp. 162746–162758, 2021.

2. X. Li, H. Shu, Y. Zhai, and Z. Lin, "A Method for Resume Information Extraction Using BERT-BiLSTM-CRF," in *Proc. IEEE 21st Int. Conf. Communication Technology (ICCT)*, Tianjin, China, Dec. 2021, pp. 1430–1435.

3. C. H. Chou and C. H. Yu, "A Résumé Evaluation System Based on Text Mining," in *Proc. 2019 IEEE Int. Conf. Artificial Intelligence in Information and Communication (ICAIIC)*, Okinawa, Japan, Feb. 2019, pp. 237–241.

4. F. Su, W. Zhang, C. Lu, J. Li, and X. Xu, "The Resume Corpus: A Large Dataset for Research in Information Extraction Systems," in *Proc. 2019 15th Int. Conf. Computational Intelligence and Security (CIS)*, Macao, China, Dec. 2019, pp. 115–119.

5. M. Rimitha, P. Prashanthi, and U. L. Ravi, "Ontologies to Model User Profiles in Personalized Job Recommendation," in *Proc. 2018 IEEE Int. Conf. Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, Mangalore, India, Aug. 2018, pp. 104–108.

6. I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated Analysis and Prediction of Job Interview Performance," *IEEE Trans. Affective Computing*, vol. 9, no. 2, pp. 191–204, Apr.–Jun. 2018.

7. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proc. 2019 Conf. Empirical Methods in Natural Language Processing (EMNLP)*, Hong Kong, China, Nov. 2019, pp. 3982–3992.

8. M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1135–1144.

9. S. Lundberg and S. Lee, "A Unified Approach to Interpreting Model Predictions," in *Proc. 31st Int. Conf. Neural Information Processing Systems (NeurIPS)*, Long Beach, CA, USA, Dec. 2017, pp. 4765–4774.

10. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. 2019 Conf. North American Chapter of the Association for Computational Linguistics (NAACL)*, Minneapolis, MN, USA, Jun. 2019, pp. 4171–4186.