

PROFORMA FOR THE APPROVAL PROJECT PROPOSAL

(Note: All entries of the proforma of approval should be filled up with appropriate and complete information. Incomplete proforma of approval in any respect will be summarily rejected.)

PRN No.: **2023016401383291**

Roll no: **22358**

1. Name of the Student: **Raj Chandrapratap Yadav**
2. Title of the Project: **AI-Based Company Fit Index System**
3. Name of the Guide: **Asst.Prof. Khalil Mujawar**
4. Teaching experience of the Guide: **10 Years**
5. Is this your first submission? Yes ☐ No ☐

Signature of the Student

Signature of the Guide

Date:

Date:

Signature of the Co-Ordinator

Date:



1. Project Title:

AI-Based Company Fit Index System

Prepared By: Raj Chandrapratap Yadav

2. Introduction

In the era of digital transformation and competitive global markets, recruitment has become one of the most critical strategic functions for modern organizations. Companies today receive a massive volume of applications for every job opening through online portals, professional networks, and recruitment platforms. Manually reviewing such a large number of resumes is not only time-consuming and resource-intensive but also highly susceptible to inconsistency and human error. Furthermore, traditional hiring processes often emphasize technical qualifications while overlooking equally important factors such as personality traits, communication skills, adaptability, teamwork, and alignment with organizational values. This imbalance frequently leads to mismatched hiring decisions, reduced job satisfaction, and increased employee turnover.

Artificial Intelligence (AI) and Machine Learning (ML) have rapidly evolved as transformative technologies capable of extracting meaningful patterns from large and complex datasets. In the recruitment domain, these technologies enable automated resume parsing, semantic analysis of job descriptions, detection of skill gaps, behavioral inference from textual data, and predictive modeling of candidate success. Natural Language Processing (NLP), a subfield of AI, plays a particularly important role in understanding unstructured text found in resumes, cover letters, interview responses, and organizational documents. By leveraging such techniques, recruitment systems can move beyond keyword-based filtering toward more intelligent, context-aware candidate evaluation.

The proposed AI-Based Company Fit Index System aims to integrate these advanced analytical techniques into a unified, web-based decision-support platform for human resource professionals. The system will collect and analyse candidate information, organizational culture profiles, and job role requirements, and then compute a quantitative Fit Index Score that represents the level of compatibility between a candidate and a specific company or position. This score will be used to rank applicants, highlight strengths and weaknesses, and provide actionable insights through interactive dashboards. By automating early-stage screening and offering data-driven recommendations, the system seeks to modernize the hiring workflow, reduce subjective bias, improve recruitment efficiency, and support organizations in selecting candidates who are more likely to succeed and remain engaged in the long term.

3. Problem Statement

Despite technological advances, most recruitment systems face the following limitations:

- Manual resume screening consumes significant time and resources
- Cultural and personality factors are rarely evaluated systematically
- Interview-based judgments introduce unconscious bias
- Lack of predictive analytics leads to poor long-term hiring decisions
- High employee attrition due to role mismatch
- Difficulty managing large applicant pools

These problems create inefficiencies in hiring processes and increase costs for organizations.

4. Proposed Solution

The proposed system introduces an automated and intelligent recruitment framework with the following features:

- Resume parsing using NLP to extract education, skills, experience, and certifications
- Job description analysis to identify required competencies and responsibilities
- Personality assessment through structured questionnaires and behavioural indicators
- Company culture profiling to define organizational values and work styles
- Machine learning models to compute compatibility
- Generation of a Fit Index Score
- Ranking of candidates for HR teams
- Interactive dashboards and analytics

This integrated approach ensures holistic evaluation rather than resume-only screening.

5. Objectives

The major objectives of the project are:

- To automate initial candidate screening
- To reduce bias through algorithmic evaluation
- To analyse technical and non-technical attributes
- To predict long-term suitability
- To assist HR professionals with decision-making tools
- To improve employee retention
- To provide transparent and explainable scoring mechanisms
- To build a scalable recruitment platform

6. Scope

The scope of the system includes:

- Online candidate registration
- Resume upload and processing
- Company profile creation
- Culture and value mapping
- Personality questionnaires
- Fit score calculation
- Candidate ranking
- HR dashboards
- Data storage and reporting
- Administrative management

The system is designed as a decision-support tool, not a replacement for human recruiters.

7. System Modules

1) Candidate Profile Module

Collects and stores resume data, academic qualifications, work experience, technical skills, and certifications in a structured database.

2) Company Culture & Value Mapping Module

Allows organizations to define cultural parameters such as teamwork, leadership style, innovation, flexibility, and ethics.

3) Personality & Behaviour Analysis Module

Uses psychometric questionnaires and NLP-based text analysis to infer behavioral traits relevant to workplace success.

4) AI Fit Score Generator

Implements machine learning algorithms to compute similarity scores between candidate attributes and organizational expectations.

5) HR Dashboard Module

Provides recruiters with visual analytics, candidate rankings, filters, and comparative reports.

6) Ranking & Recommendation Module

Sorts applicants based on their Fit Index Score and recommends the best matches for each role.

8. Methodology

The project will follow an **Agile-based development methodology**, consisting of iterative stages:

- Requirement gathering from HR use cases
- Dataset collection and cleaning
- Feature extraction from resumes and questionnaires
- Model selection and training
- Validation and evaluation
- Frontend and backend integration
- Testing (unit, integration, system)
- Deployment on cloud platforms
- Documentation and reporting

9. Tools and Technologies

Frontend:

HTML5, CSS3, JavaScript, React.js

Backend:

Python (Django / Flask) or Node.js

Database:

MySQL or MongoDB

AI / ML:

Scikit-Learn, TensorFlow, SpaCy, NLTK

Visualization:

Chart.js / Power BI

Version Control:

GitHub

10. Project Timeline

Phase	Duration
Week 1	Requirement analysis and planning
Week 2	UI design and dataset preparation
Week 3	Resume parsing module
Week 4	Job description analysis
Week 5	Personality assessment system
Week 6	Model training and validation
Week 7	Fit score engine
Week 8	HR dashboard development
Week 9	Testing and optimization
Week 10	Deployment and documentation

11. Expected Outcomes

The system is expected to:

- Speed up recruitment processes
- Improve hiring quality
- Reduce human bias
- Provide transparent evaluation
- Support large-scale hiring
- Enhance employee retention
- Deliver data-driven HR insights
- Demonstrate real-world AI application
- Enable scalable future upgrades

12. References

1. **Russell, S., & Norvig, P.** (2021).
Artificial Intelligence: A Modern Approach (4th ed.). Pearson Education.
2. **Jurafsky, D., & Martin, J. H.** (2023).
Speech and Language Processing (3rd ed.). Pearson.
3. **Goodfellow, I., Bengio, Y., & Courville, A.** (2016).
Deep Learning. MIT Press.
4. **Aggarwal, C. C.** (2018).
Machine Learning for Text. Springer.
5. **Han, J., Kamber, M., & Pei, J.** (2012).
Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann.
6. **Sommerville, I.** (2016).
Software Engineering (10th ed.). Pearson Education.
7. **Pressman, R. S., & Maxim, B. R.** (2020).
Software Engineering: A Practitioner's Approach (9th ed.). McGraw-Hill.
8. **Manning, C. D., Raghavan, P., & Schütze, H.** (2008).
Introduction to Information Retrieval. Cambridge University Press.
9. **Géron, A.** (2022).
Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow (3rd ed.). O'Reilly Media.
10. **Silberschatz, A., Korth, H. F., & Sudarshan, S.** (2020).
Database System Concepts (7th ed.). McGraw-Hill.
11. **Kleppmann, M.** (2017).
Designing Data-Intensive Applications. O'Reilly Media.
12. **Tan, P.-N., Steinbach, M., & Kumar, V.** (2019).
Introduction to Data Mining (2nd ed.). Pearson.