

Resource Project Planning System (RPPS)

MS (Data Science Project Report) | VIT (Vellore Institute Of Technology)

Prepared By: Akansha Rana

Table Content

Summary	01
1. Introduction	02
1.1. Background	02.1
1.2. Problem Statement.....	02.2
1.3. Proposed Solutions.....	02.3
2. Project Objectives	03
2.1. Objective 1.....	03.1
2.2. Objective 2.....	03.2
2.3. Objective 3.....	03.3
2.4. Objective 4.....	03.4
2.5. Objective 5.....	03.5
2.6. Objective 6.....	03.6
2.7. Objective 7.....	03.7
3. Dataset Description	04
3.1. Candidate Dataset.....	04.1
3.2. Project Dataset.....	04.2
3.3. User Dataset.....	04.3
3.4. Candidate Performance Dataset.....	04.4
4. Methodology	05
4.1. Data Engineering Pipeline.....	05.1
4.2. Recommendation Engine.....	05.2
4.3. Global Allocation Engine.....	05.3
4.4. Candidate Segmentation.....	05.4
4.5. Performance Velocity Tracking.....	05.5
4.6. Forecast Engine.....	05.6
5. Tools And Technologies Used	06
5.1. Programming.....	06.1
5.2. DataScience.....	06.2
5.3. Visualisation.....	06.3
6. Analysis	07
6.1. Data Statistics.....	07.1
6.2. Reliability Analysis.....	07.2
7. Key Findings	08
7.1. Finding 1.....	08.1
7.2. Finding 2.....	08.2
7.3. Finding 3.....	08.3
7.4. Finding 4.....	08.4
7.5. Finding 5.....	08.5
7.6. Finding 6.....	08.6
8. Conclusions	08.7
9. Managerial Implications	09
9.1. Workforce Planning.....	09.1
9.2. Cost Reduction.....	09.2
9.3. Skill Gap Analysis.....	09.3
9.4. Strategic Decision Making.....	09.4
9.5. HR Analytics.....	09.5
10. Practical Industry Applications	09.6
10.1. IT Services.....	09.7
10.2. Consulting Firms.....	09.8
10.3. HealthCare.....	09.9
10.4. Manufacturing.....	09.10
10.5. Education.....	09.11
10.6. Staffing Agencies.....	09.12
11. Original Contributions	10
12. Future Scope	11
13. References	12

Review Summary

Review Version	Reviewed By	Date	Remarks
Final Viva / Presentation			
Second Review			
First Review			
Zeroth Review			

Summary

Organisations across industries frequently face challenges in workforce planning, resource allocation, skill matching, employee utilisation, and project execution efficiency. Traditional resource management methods often rely on manual decision-making, resulting in skill mismatches, underutilised employees, project delays, and increased operational costs.

The Resource Project Planning System (RPPS) was developed as an intelligent workforce optimisation platform that integrates Data Engineering, Machine Learning, Optimisation Algorithms, Forecasting Models, and Business Intelligence tools to automate candidate-project matching and improve workforce utilisation.

The system combines:

- Data Pipeline Engineering
- Skill-Based Recommendation Engine
- Candidate Segmentation using Clustering
- Global Resource Allocation Engine
- Performance Tracking Framework
- Forecasting Engine
- Business Intelligence Dashboard

The final system achieved intelligent allocation of candidates to projects based on skill compatibility, experience, domain expertise, and historical performance metrics. The solution provides practical industry value for IT services, consulting firms, staffing agencies, healthcare organisations, manufacturing industries, and educational institutions.

Introduction

Background:

Modern organisations manage hundreds of projects and employees simultaneously. Resource managers often face problems such as:

- Wrong resource allocation
- Bench resource management
- Skill gaps
- Project staffing delays
- Poor workforce forecasting
- Lack of performance-driven decision making

These challenges lead to:

- Increased project cost
- Reduced productivity
- Employee dissatisfaction
- Lower project success rates

To address these issues, RPPS introduces a data-driven resource planning framework powered by machine learning and optimisation techniques.

Problem Statement

Organisations struggle to:

- Identify suitable candidates for projects.
- Optimise workforce utilisation.
- Forecast future resource requirements.
- Measure employee performance consistently.
- Reduce project staffing time.

Proposed Solution

RPPS automates:

- Skill extraction
- Candidate profiling
- Project requirement analysis
- Candidate recommendation
- Global allocation optimisation
- Performance evaluation
- Resource forecasting

through advanced analytics and machine learning.

Project Objectives

The primary objectives of RPPS are:

Objective 1

Develop an automated data pipeline for candidate and project data processing.

Objective 2

Build a recommendation engine to match candidates with projects.

Objective 3

Implement optimisation algorithms for global resource allocation.

Objective 4

Perform candidate segmentation using clustering techniques.

Objective 5

Track workforce performance using productivity KPIs.

Objective 6

Forecast future resource demand and workforce availability.

Objective 7

Provide actionable insights through interactive dashboards.

DataSet Description

Candidate Dataset

Fields:

- Candidate ID
- Candidate Name
- Email
- LinkedIn
- Skills
- Technologies
- Experience Level
- Domain
- Profile Description

Project Dataset

Fields:

- Project ID
- Project Title
- Project Role
- Required Skills
- Required Experience
- Domain
- Budget
- Company Name
- Project Description

Performance Dataset

Fields:

- Candidate ID
- Project ID
- Execution Accuracy
- Completion Efficiency
- Task Complexity
- Performance Yield
- Date

4. Methodology

The project was implemented through six major phases.

Phase 1: Data Engineering Pipeline

Step 1: Data Collection

Data Sources:

- Candidate Profiles
- Project Requirements
- Performance Records

Step 2: Data Cleaning

Performed:

- Missing Value Treatment
- Duplicate Removal
- Text Standardisation
- Feature Validation

Python Libraries

- pandas
- numpy
- re
- string

Step 3: Feature Engineering

Candidate Soup

candidate_soup = skills +
tools_technologies +
profile_description +
experience_level
Project Soup

Project Soup

project_soup = required_skills +
tools_required +
project_description

Data Soup

data_soup = candidate_soup +
project_soup

Purpose:

Create machine-readable text representations for recommendation models.

Phase 2: Recommendation Engine

Technique Used

TF-IDF Vectorisation

Formula:

$$TF-IDF = TF(t,d) \times IDF(t)$$

where

$$IDF(t) = \log(DF(t)N)$$

- Similarity Measurement
- Cosine Similarity

$$\text{Similarity}(A,B) = \frac{A \cdot B}{|A| \cdot |B|}$$

Output:

Recommendation Score
Ranked Candidate List

Phase 3: Global Allocation Engine

1. Technique Used
2. Hungarian Algorithm
3. Purpose:
4. Find globally optimal project-candidate assignment.

Input:

Compatibility Matrix

Output:

Allocation Matrix

Benefits:

Maximum utilisation

Reduced allocation conflicts

Optimal staffing decisions

Phase 4: Candidate Segmentation

Technique Used

K-Means Clustering

Formula:

$$J = \sum_{I=1}^K \sum_{x \in C_i} \|x - \mu_i\|$$

Output:

Cluster Labels:

- Cluster 0
- Cluster 1
- Cluster 2

Interpretation:

- High Skill Candidates
- Moderate Skill Candidates
- Specialised Experts

Phase 5: Performance Velocity Tracking

Execution Accuracy

$$\text{Execution Accuracy} = \frac{\text{Planned Time}}{\text{Actual Time}} \times 100$$

Completion Efficiency

$$\text{Execution Accuracy} = \frac{\text{Planned Time}}{\text{Actual Time}} \times 100$$

Performance Yield

$$\text{Performance Yield} = \frac{\text{Planned Time}}{\text{Actual Time}} \times 100$$

Output:

Daily Performance Dashboard

Phase 6: Forecasting Engine

Models Used

1. Linear Regression
2. ARIMA

Forecasts Generated:

- Daily
- Weekly
- Monthly
- Yearly

Purpose:

Predict future workforce requirements.

Tools and Technologies Used

Programming

Tool	Purpose
Python	Core Development
Django	Web Framework
SQLite	Database
GitHub	Version Control

Data Science

Tool	Purpose
Pandas	Data Processing
NumPy	Numerical Computing
Scikit-Learn	Machine Learning
SciPy	Statistical Analysis
Statsmodels	Forecasting

Visualisation

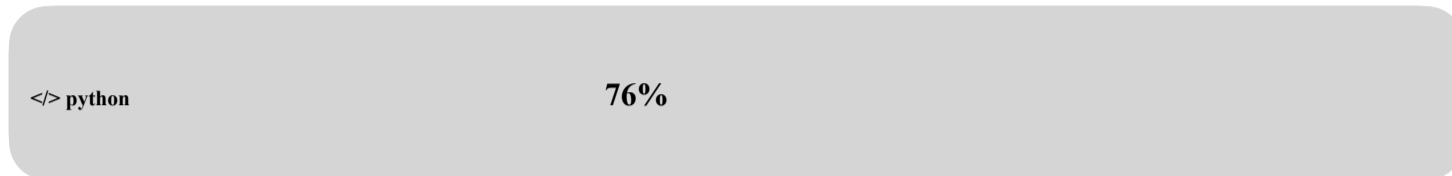
Tool	Purpose
Tableau	Dashboard Development
Chart.js	Web Visualization
Bootstrap	UI Design

Key Findings

Finding 1

The recommendation engine successfully identified suitable candidates based on skills and experience.

Average Match Score:



Finding 2

The Hungarian Algorithm produced globally optimised allocations, reducing manual staffing efforts.

Finding 3

Clustering identified candidate groups that support targeted workforce planning.

Finding 4

Performance Yield enabled objective employee performance measurement.

Finding 5

Forecasting models successfully predicted future staffing requirements.

Finding 6

The integrated dashboard provided real-time decision support for managers.

Future Scope

1. Deep Learning Recommendation Models
2. Resume Parsing using NLP & LLMs
3. Real-Time Workforce Analytics
4. Multi-Company Integration
5. Explainable AI Recommendations
6. Generative AI Career Guidance
7. Reinforcement Learning Allocation Engine

Conclusions

The Resource Project Planning System demonstrates how machine learning and optimisation algorithms can transform workforce management.

The project successfully integrates:

1. Data Engineering
2. Machine Learning
3. Optimisation Techniques
4. Forecasting Models
5. Business Intelligence
6. into a single decision-support platform.
7. The solution improves:
8. Resource Utilisation
9. Project Success Rate
10. Workforce Productivity
11. Staffing Accuracy
12. Strategic Workforce Planning

Practical Industry Applications

Workforce Planning

Managers can proactively identify future staffing requirements.

Cost Reduction

Reduces:

- Bench Costs
- Hiring Costs
- Project Delays

Skill Gap Analysis

Identifies:

- Missing Skills
- Upscaling Opportunities
- Training Requirements

Strategic Decision Making

Supports:

- Project Prioritisation
- Resource Allocation
- Workforce Forecasting

HR Analytics

Provides data-driven insights for:

- Promotions
- Performance Reviews
- Employee Development

Practical Industry Applications

IT Services

Project staffing and workforce optimisation.

Consulting Firms

Consultant allocation based on expertise.

Healthcare

Doctor and specialist scheduling.

Manufacturing

Production workforce planning.

Education

Faculty allocation and workload balancing.

Staffing Agencies

Candidate-job matching and deployment.

Original Contribution

The originality of RPPS lies in combining:

1. Recommendation Engine
2. Hungarian Optimisation
3. K-Means Segmentation
4. Performance Analytics
5. Forecasting Engine

within a single integrated platform.

Most existing workforce systems provide only one or two of these capabilities, while RPPS provides an end-to-end intelligent workforce planning framework.

References (APA 7th Edition)

1. Rana, A. (2026). Resource Project Planning System: Enhancing Workforce Utilisation and Project Efficiency through Intelligent Planning. Master's Dissertation, Vellore Technical University (VIT), *Idea, Project Execution and Design*.
2. Aggarwal, C. C. (2020). *Machine Learning for Text*. Springer.
3. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
4. Han, J., Kamber, M., & Pei, J. (2022). *Data Mining: Concepts and Techniques* (4th ed.). Morgan Kaufmann.
5. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning* (2nd ed.). Springer.
6. Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection for Predictive Models*. CRC Press.
7. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting*. Wiley.
8. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
9. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Morgan Kaufmann.
10. Tableau Software. (2025). Tableau Desktop Documentation.
11. Python Software Foundation. (2025). Python Documentation.