ETDT

Emerging Trends in Digital Transformation

Volume-2, Issue-2, April-June 2026 ISSN: **3107- 4308** *Paper ID:* ***ETDT-V2-I2-38***

EMPLOYEE PERFORMANCE ANALYSIS USING MACHINE LEARNING

#1**Vilasagar Tharuni**, *Department of MCA,*

#2**Dr. T. Ravikumar**, *Professor, Department of CSE,*

Vaageswari College of Engineering (Autonomous), Karimnagar, TG.

ABSTRACT: Evaluating employee performance is a central responsibility of human resource management, yet conventional appraisal methods depend heavily on subjective manager opinions and are slow, inconsistent and prone to bias. This paper presents a machine learning based system that analyses historical employee records and predicts an employee's performance category (Low, Medium or High) from measurable attributes such as experience, training hours, projects completed, attendance, average working hours, peer review score and job satisfaction. The workflow consists of data collection, cleaning, exploratory analysis, feature scaling, stratified train/test splitting, training of six classifiers (Logistic Regression, Decision Tree, K-Nearest Neighbours, Support Vector Machine, Gradient Boosting and Random Forest), cross-validation, hyper-parameter tuning and comparison using accuracy, precision, recall, F1-score and ROC-AUC. Random Forest gave the best overall balance on the held-out test set with an accuracy of 87.0% and a weighted F1-score of 0.871, with Gradient Boosting performing almost identically. Feature importance analysis shows which factors influence performance most, making the predictions explainable to HR managers. The trained model is exposed through a Flask web application with administrator and user modules.

Keywords: *Employee performance, machine learning, Random Forest, classification, feature importance, HR analytics, Flask.*

1. INTRODUCTION

Organisations make decisions on promotions, training, rewards and retention based on how well employees perform. Traditionally this is judged through periodic appraisals that rely heavily on a manager's personal opinion, which makes the process slow, inconsistent and open to bias. At the same time, organisations already hold large amounts of routine HR data, such as attendance, working hours, projects completed, training and review scores, that is rarely used in a systematic way.

Machine learning can learn the relationship between such measurable attributes and past performance outcomes, and then predict the performance category of an employee in a consistent and repeatable manner. The objective of this project is to build a classification system that predicts whether an employee's performance is Low, Medium or High, to compare several algorithms fairly, to explain which factors matter most, and to make the final model usable by non-technical HR staff through a web application.

DOI: https://doi.org/10.5281/zenodo.23254095

2. LITERATURE REVIEW

Research on workforce analytics has moved from simple statistical appraisal scoring towards supervised learning on HR records. Studies in this area commonly treat performance prediction as a classification problem and compare tree-based ensembles with linear and distance-based models. Ensemble methods such as Random Forest, introduced by Breiman [1], and gradient boosting, introduced by Friedman [2], are repeatedly reported to handle mixed numeric attributes, non-linear relationships and noisy data better than single models, and they provide built-in measures of feature importance.

Support Vector Machines [3] remain a strong baseline on scaled numeric data, while Logistic Regression, Decision Trees and K-Nearest Neighbours are widely used for their simplicity and interpretability. Practical toolkits such as scikit-learn [4] have made it straightforward to build, cross-validate and tune such pipelines. A recurring concern in the literature is explainability: HR decisions affect people, so managers need to understand why a model produced a prediction. Methods such as SHAP [5] attribute a prediction to individual input features and are increasingly proposed for HR analytics. Fairness is the second recurring concern, since models trained on historical data can reproduce historical bias unless sensitive attributes are audited or removed.

From this survey, the gaps addressed by this project are a like-for-like comparison of several classifiers under the same preprocessing and validation, an interpretable view of the drivers of performance, and delivery of the result as a simple application rather than only a notebook.

3. SYSTEM ANALYSIS

EXISTING SYSTEM

In the existing system, employee performance is assessed through manual appraisal forms and manager discussions held at fixed intervals. Scores depend on the evaluator's judgement, different managers apply different standards, and historical data is seldom analysed to find patterns.

Disadvantages

- Evaluation is subjective and prone to bias.
- The process is slow and must be repeated manually for every cycle.
- Available HR data is not used to find which factors actually drive performance.

PROPOSED SYSTEM

The proposed system trains machine learning classifiers on historical employee records and predicts the performance category for a new employee from attributes such as experience, training hours, projects completed, attendance, average working hours, peer review score and job satisfaction. Six algorithms are trained and compared, the best model is selected, and feature importance is reported so that predictions can be explained. A Flask web application provides administrator and user modules.

DOI: https://doi.org/10.5281/zenodo.23254095

Advantages

- Consistent, data-driven predictions that reduce subjectivity.
- Several algorithms are compared fairly using the same split, scaling and cross-validation.
- Feature importance makes the result explainable to HR managers.
- A simple web interface lets non-technical staff use the model through forms and charts.
- Early identification of employees who may need training or support.

4. METHODOLOGY

The dataset of employee records is first cleaned and explored to understand distributions and relationships between attributes. Numeric features are scaled, and the data is divided into training and test sets using a stratified split so that the Low, Medium and High classes keep the same proportions in both sets. Six classifiers are trained: Logistic Regression, Decision Tree, K-Nearest Neighbours, Support Vector Machine, Gradient Boosting and Random Forest. Five-fold cross-validation and hyper-parameter tuning are used to select settings, and the models are compared on the held-out test set using accuracy, precision, recall, F1-score and ROC-AUC. The selected model is saved and loaded by the Flask application.

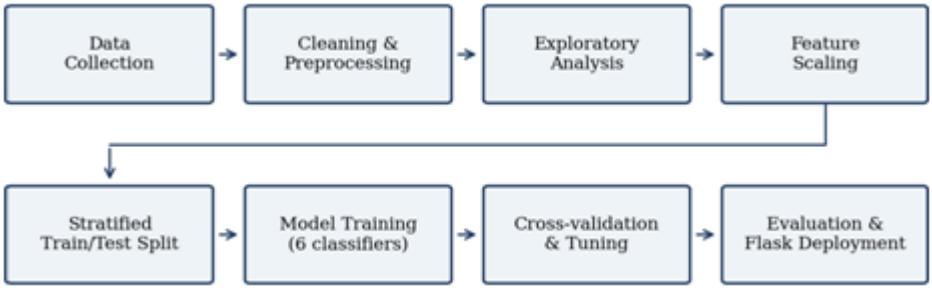

Fig 1: Workflow of the Employee Performance Analysis System

5. RESULTS

All models were evaluated on the same held-out test set. Random Forest gave the best overall balance and was selected as the final model, with Gradient Boosting performing almost identically.

Model	Test Accuracy	Weighted F1-score
Random Forest (selected)	87.0%	0.871

Table 1: Performance of the selected model on the test set

Feature importance analysis from the final model shows which attributes contribute most to the predicted performance category, which allows HR managers to see why a prediction was made and where training or support may help. In the web application, the administrator manages data and the model, while users enter employee attributes through a form and receive the predicted category together with charts.

DOI: https://doi.org/10.5281/zenodo.23254095

6. CONCLUSION

This project demonstrated that machine learning can predict employee performance categories from routine HR data. After exploratory analysis, preprocessing and scaling, six classifiers were compared using a hold-out test set, five-fold cross-validation and ROC-AUC. Random Forest was selected as the final model with 87.0% test accuracy and a weighted F1-score of 0.871, closely matched by Gradient Boosting. Feature importance analysis made the results interpretable, and a Flask web application made the model usable by non-technical HR staff. The system reduces subjectivity, saves effort and helps managers act earlier on training and retention, while keeping the final decision with people.

FUTURE ENHANCEMENT

- Train and validate the model on real organisational data from multiple departments.
- Add deep learning and time-series models to track performance trends across appraisal cycles.
- Integrate SHAP-based per-employee explanations into the web interface.
- Perform fairness audits and remove sensitive attributes to avoid bias.
- Connect to HRMS or ERP systems for automatic data collection and dashboards.
- Add attrition prediction and personalised training recommendations.
- Include text analysis of manager feedback and survey comments using natural language processing.
- Deploy the application on a cloud platform with role-based access control and audit logs.

REFERENCES

1. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
2. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232
3. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
4. M. Shahimoon, H. Rasheed, D. R. Rao, K. R. Krishnaiah, K. K. Gajula, et al., “Evolving Predictive Models for Customer Churn in Telecommunications,” in Proc. 2nd International Conference on New Frontiers in Communication Technologies, 2025.
5. Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
6. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
7. K. K. Gajula, “Deep Neural Forecasting Models for Climate Change Analysis Using Multivariate Time-Series Data,” Journal of Computational Analysis and Applications, vol. 33, no. 8, pp. 6356–6368, 2024.

DOI: https://doi.org/10.5281/zenodo.23254095