Week 1 Report - Problem Statement and References

Group 2

Members
Garima Singh
Nishkarsh Singhal
Shaivee Sharma

Date of Submission: 17th June, 2025

Index

Problem Statement

Utilization of Machine Learning Techniques to Analyse IBM HR Dataset and Predict Employee Attrition

Research Papers

Paper 1: Optimising HRM practices in call centres: predicting and explaining employee turnover intention using classification and regression trees.
DOI: https://doi.org/10.1108/IJOA-12-2024-5117
Date of Publication: 13th May, 2025

Research Gap

- Lack of focus on turnover drivers: Most prior studies prioritize model accuracy over identifying real-world factors influencing employee attrition.
- Neglect of interpretable models: Models like CART are underutilized, despite their transparency and practical value for HR professionals.
- No consensus on predictors: Previous research using the same datasets often yields inconsistent predictor importance without explaining the discrepancies.
- Lack of novel composite features: Existing studies do not explore derived metrics like workload-to-project ratio to capture hidden dimensions of job stress.

Models Evaluated

- **Primary Model of Focus:** Classification and Regression Trees (CART) – chosen for its high interpretability (white-box) nature, enabling HR professionals to understand decision paths.
- **Comparative Models Evaluated:**
  - XGBoost (Black-box)
  - Random Forest (Black-box)
  - Support Vector Machine (Black-box)
  - Logistic Regression (Partially interpretable)
  - Naïve Bayes (Simple probabilistic)

Key Findings

- **CART's Insightfulness**:
  - Achieved balance between predictive accuracy and interpretability.
  - Clear decision rules helped uncover key drivers like job satisfaction, last evaluation, tenure, and workload stress.
- **Predictor Importance (CART)**:
  - Job Satisfaction
  - Last Evaluation
  - Time Spent at Company
  - Low Salary
  - Department Type
  - Workload Ratio
- **Workload Ratio Discovery**: Demonstrated how even subtle workload mismatches influence attrition—often missed by black-box models.
- **Clarified Prior Conflicts**: By analyzing feature importance and modeling differences, the study explains discrepancies in findings across earlier research using similar datasets.

Limitations

- Synthetic dataset: Based on a structured dataset (IBM HR), limiting real-world complexity and emotional/psychological variables.
- No real-time signals: Does not incorporate live engagement data or stress indicators for dynamic prediction.
- Advanced AI methods excluded: Deep learning models (e.g., LSTM, Autoencoders) and Reinforcement Learning were not explored.
- Domain-specific generalizability limited: The study focuses on general patterns, and findings may differ across industries or cultures.

Paper 2: Predicting employee attrition using machine learning approaches.
DOI: https://doi.org/10.61591/jslhu.20.717
Date of Publication: 15th March, 2025

Research Gap

- Prior studies on attrition relied heavily on traditional statistical methods or single ML models, lacking comprehensive model comparison.
- Most existing models failed to handle data imbalance, which reduced prediction accuracy for minority classes (i.e., employees who left).
- Limited exploration of ensemble methods like Extra Trees, despite their strong potential in handling structured HR data.
- Scarcity of research integrating exploratory data analysis (EEDA) with machine learning pipelines for actionable HR insights.
- Need for systematic model benchmarking and validation (e.g., k-fold cross-validation) to ensure generalization.

Models Evaluate

- Four supervised machine learning classifiers were used:
  - Extra Trees Classifier (ETC) – Proposed primary model
  - Support Vector Machine (SVM)
  - Logistic Regression (LR)
  - Decision Tree Classifier (DTC)
- Data balancing technique: SMOTE (Synthetic Minority Over-sampling Technique)
- Model Evaluation Metrics:
  - Accuracy
  - Precision
  - Recall
  - F1-Score
  - 10-Fold Cross-Validation

Key Findings

- ETC achieved highest performance: 93% accuracy, precision, recall, F1-score, and ROC-AUC
- ETC outperformed SVM (87%), DTC (83%), and LR (72%) in accuracy.
- SMOTE improved class balance, boosting minority class prediction accuracy.
- Top factors influencing attrition (from EEDA):
  Monthly income, hourly rate, job level, age, years with manager
- ETC also outperformed state-of-the-art models from previous studies (DT, LR, RF, SVM).

- 10-fold cross-validation confirmed ETC's robust generalization.
- Visual analysis (age, income, job level, tenure) aligned with model findings, improving interpretability.

Limitation

- Although ETC achieved the highest accuracy (93%), its decision-making process is not easily interpretable due to its black box nature compared to models like Decision Trees or Logistic Regression.
- The study used SMOTE to balance the classes, which can lead to synthetic data bias or overfitting if not carefully tuned.
- Even though ETC provides good feature importance values, it doesn't give rule-based reasoning or conditional logic, which are easier for HR teams to understand.
- It is computationally heavier than Logistic Regression or a single Decision Tree due to multiple randomized trees and random feature splits.
- The model was only tested on the IBM HR dataset. No cross-organizational or cross-domain validation was conducted.

Paper 3: Machine Learning Applications in Human Resource Management: Predicting Employee Turnover and Performance.

DOI: https://doi.org/10.53032/tvcr/2025.v7n2.37

Date of Publication: 30th April, 2025

Research Gap

- Traditional methods (e.g., surveys, regression) lack accuracy and scalability in predicting employee turnover.
- Deep learning methods exist but are often non-transparent and not practical for HR professionals.
- There is limited work on comparing interpretability vs. performance between classical ML models like DT and RF in the HR domain.
- Need for explainable and fair AI in HR, especially considering privacy and bias mitigation.

- Lack of studies focusing on real-world HR factors like salary, job satisfaction, and work-life balance with interpretability tools (e.g., SHAP).

Models Evaluated

- **Decision Tree (DT):**
  - Interpretable, rule-based model
  - Uses criteria like Gini impurity or entropy
  - Prone to overfitting
- **Random Forest (RF)**:
  - Ensemble of Decision Trees
  - Reduces variance and overfitting
  - Higher generalization and robustness
  - SHAP values used for interpretability

Key Findings

- Random Forest outperformed Decision Tree
- Accuracy: RF = 85%, DT = 78%
- Fewer classification errors (False Positives/Negatives)
- Better performance across precision, recall, F1-score
- Important Predictive Features: Annual Salary, Monthly Salary, Job Satisfaction, Work-Life Balance, Experience, etc.
- Interpretability Tools: SHAP used with RF to improve transparency of predictions
- Practical Implications:
  - Helps HR teams identify at-risk employees
  - Enables early intervention and personalized retention plans
  - Supports cost savings, better engagement, and workforce stability

Limitations

- Decision Tree:
  - Overfits to training data.
  - Lower generalization to new dataset.
  - Biased towards dominant classes in imbalanced datasets.
- Random Forest:

- Less interpretable than a single decision tree.
    - Computationally more intensive.
    - May still inherit bias from training data if not addressed.

- Proposed Model:
    - Data bias remains a risk, potentially leading to unfair HR decisions.
    - Static dataset: no real-time data integration.
    - Does not include advanced models like deep learning or hybrid ensembles.
    - Privacy and ethical concerns around employee data use are not fully resolved.

Paper 4: Predicting Employee Attrition Using Machine Learning Approaches
DOI: https://doi.org/10.3390/app12136424
Date of Publication: 24th June, 2022

Research Gap

- Previous studies lacked a comprehensive analysis of factors influencing employee attrition.
- Many studies focused on single models, leading to inconsistent feature importance across research.
- Limited interpretability in prior black-box models made it difficult for HR professionals to act on predictions.
- Overlooked composite features, such as the interaction of salary, job level, and tenure in influencing attrition.

Models Evaluated

- **Primary Model:** Optimized Extra Trees Classifier (ETC) (black-box).
- **Comparative Models:**
    - Support Vector Machine (SVM)
    - Logistic Regression (LR)
    - Decision Tree Classifier (DTC).

- ETC was optimized for performance but lacked inherent interpretability compared to LR and DTC (white-box).

Key Findings

- ETC achieved 93% accuracy, outperforming state-of-the-art approaches in employee attrition prediction.
- Feature Importance Analysis: Key predictors included monthly income, hourly rate, job level, and age.
- Imbalanced Dataset Handling: Applied SMOTE resampling for balanced class distribution.
- Cross-validation with K-Fold analysis validated the model's robustness.
- Exploratory Data Analysis revealed salary and tenure as primary attrition drivers.

Limitations

- The major limitation was it was relying on the accuracy scores for testing of goodness of the model, even though the data is very imbalance and we are generating artificial data points to balance the data, we should rather focus on getting better recall and accuracy for the minority class

---

Paper 5: Machine Learning for Predicting Employee Attrition
DOI: http://dx.doi.org/10.14569/IJACSA.2021.0121149
Date of Publication: 30th November, 2021

Research Gap

- Prior studies lacked a detailed evaluation of machine learning models for employee attrition prediction.
- Limited exploration of feature selection techniques that enhance model performance.
- No comparative analysis of DT, SVM, and ANN models using standardized preprocessing methods.

Models Evaluated

- **Primary Model:** Support Vector Machines (SVM), identified as the optimal choice.
- **Comparative Models:**
    - Decision Tree (DT)
    - Artificial Neural Network (ANN).
- SVM was preferred for high accuracy and generalization capability, whereas DT provided greater interpretability**.**

Key Findings

- SVM achieved 88.87% accuracy, outperforming DT and ANN.
- Feature Importance Analysis: Top predictors included job satisfaction, income, job level, and overtime status.
- SMOTE applied to balance class distribution, improving model fairness.
- Regularization techniques optimized models, reducing overfitting in ANN.

Limitations

- Excluded advanced AI methods such as ensemble learning (e.g., XGBoost, Random Forest), which could potentially yield better results.
- Limited interpretability of SVM and ANN, making HR interventions based on model predictions more challenging.

Paper 6: Employee Attrition: Analysis of Data Drive Models
DOI: https://doi.org/10.4108/eetiot.4762
Date of Publication: 3rd January, 2024

Research Gap

- Prior studies on employee attrition have largely focused on traditional classification techniques without considering deep learning approaches.
- A lack of model interpretability in black-box models limits actionable insights for HR professionals.

- Existing research shows inconsistent predictor importance, with attributes like job satisfaction, overtime, and monthly income yielding varying impacts across studies.
- Composite feature interactions, such as job level combined with work-life balance, remain unexplored.

Models Evaluated

- **Primary model:** Feedforward Neural Network (FNN) (black-box model).
- **Comparative models:** Logistic Regression (white-box), K-Nearest Neighbor (KNN), Decision Tree, Naïve Bayes, Gradient Boosting, AdaBoost, Random Forest, Stacking, XGBoost, and Convolutional Neural Network (CNN) (mostly black-box models).

Key Findings

- They did not balance the dataset
- The FNN model demonstrated superior performance, achieving 97.5% accuracy, 83.93% recall, and 91.26% F1-score, outperforming both ensemble and machine learning classifiers.
- Stacking and XG Boost provided competitive results but fell short of deep learning methods in predictive power.
- Key predictors of attrition include overtime, job satisfaction, job level, monthly income, age, and job involvement, confirming prior studies but reinforcing the need for interaction modeling.
- Logistic Regression showed strong interpretability, making it suitable for HR analytics but less effective in prediction compared to deep learning models.

Limitations

- They have not mentioned that the scores they were achieving were in the majority class or were in the minority class, also they did not balance the dataset, it's probable that they are reporting the scores of the majority class.
- Also a neural network getting high scores as these, on a small dataset, it may suggest that there is some level of overfitting on the data
- Feature Selection Constraints: While attribute importance was assessed, the study did not explore advanced feature engineering techniques such as

composite attribute generation or interaction modeling, which could enhance predictive power.

- Model Interpretability Trade-Off: The deep learning models, particularly the Feedforward Neural Network (FNN), offer strong predictive performance but lack transparent decision-making. This limits direct interpretability for HR professionals compared to white-box models.

Paper 7: Predictive Model of Employee Attrition Based on Stacking Ensemble **Learning**
DOI: https://doi.org/10.1016/j.eswa.2022.119364
Date of Publication: 7th December, 2022

Gaps Identified in Previous Research

- Lack of comparative evaluation across diverse machine learning models.
- Limited exploration of hybrid models that integrate multiple algorithms for better accuracy.
- Inadequate attention to class imbalance and the use of resampling methods like SMOTE.
- Feature engineering and selection were often overlooked or poorly implemented.
- Previous models offered limited interpretability and actionable insights for HR decision-making.

Models Evaluated

- **Primary Model**: Hybrid Model (combines SVM, Decision Tree, Random Forest, and Gradient Boosting).
- **Comparative Models**:
  - Support Vector Machine (SVM)
  - Decision Tree Classifier (DTC)
  - Random Forest (RF)
  - Gradient Boosting (GB)

Key Findings

- The Hybrid Model achieved the highest accuracy at **95.0%**, outperforming all individual models.
- SVM achieved **88.6% accuracy**, leveraging its strength in managing complex, high-dimensional data.
- Random Forest and Gradient Boosting both achieved **87.3%** accuracy.
- Decision Tree recorded the lowest accuracy at **80.5%**, indicating overfitting issues.
- **Feature engineering** (correlation analysis and one-hot encoding) significantly improved model performance.
- **SMOTE** was used to address class imbalance.
- Exploratory analysis identified **monthly income** and **working years** as key attrition predictors.

Limitations

- Overemphasis on accuracy despite class imbalance; metrics like recall or F1-score were not reported.
- The hybrid model, while accurate, lacks interpretability for HR application.
- High computational cost may limit deployment in real-time environments.
- No external validation or industry-specific customization was conducted.

Paper 8: Developing a Hybrid Machine Learning Model for Employee Turnover Prediction: Integrating LightGBM and Genetic Algorithms

DOI: https://doi.org/10.1016/j.joitmc.2025.100557
Date of Publication: 31st May, 2025

Gaps Identified in Previous Research

- Prior studies often relied on either high-accuracy models or interpretable models, rarely achieving both.
- Most works did not effectively address feature selection in high-dimensional HR datasets.
- Few models combined advanced ensemble techniques (like LightGBM) with metaheuristic optimization (such as Genetic Algorithms).
- Limited generalizability in earlier studies due to use of only public datasets.

Models Evaluated

- **Primary Model:** Hybrid model combining Genetic Algorithm (GA) for feature selection and LightGBM for classification.
- **Comparative Models:**
  - Random Forest
  - Logistic Regression
  - Standalone LightGBM
  - K-Nearest Neighbors (KNN)
  - Decision Tree
  - Support Vector Machine (SVM)
  - XGBoost
  - Naive Bayes

Key Findings

- The GA + LightGBM hybrid model achieved AUC of 0.94 and F1-score of 0.87 on the public HR dataset; AUC of 0.78 and F1-score of 0.73 on the organizational dataset.
- Outperformed all baseline models across multiple evaluation metrics (accuracy, precision, recall, F1-score, AUC).
- GA reduced the feature set from 30 to 14 (HR dataset) and 25 to 11 (Organization dataset), improving model interpretability and reducing overfitting.
- LightGBM was selected for its computational efficiency and superior performance on tabular data.
- Evaluation used 10-fold cross-validation for robust performance measurement.
- Exploratory analysis showed that salary, job satisfaction, total working years, and work-life balance were among the strongest predictors of turnover.

Limitations

- Although the model performs well, the hybrid framework increases computational complexity, making it less suitable for real-time deployment without optimization.
- While Genetic Algorithms improve feature selection, the interpretability of LightGBM remains limited compared to white-box models.

- The use of accuracy as a primary metric may be insufficient given the likely class imbalance; additional focus on recall for the minority class is warranted.
- Generalizability still requires validation across more diverse industries and job roles beyond the two datasets used.

Paper 9: Predicting Employee Attrition: A Comparative Analysis of Machine Learning Models Using the IBM Human Resource Analytics Dataset
DOI: https://doi.org/10.1016/j.procs.2025.04.659
Date of Publication: 10th May, 2025

Gaps Identified in Previous Research

- Prior studies primarily relied on single machine learning models, limiting predictive performance.
- Existing models did not fully capture the complexity of employee attrition causes, which are often nonlinear and multifactorial.
- Limited exploration of **stacking ensemble methods** in the employee attrition domain.
- Many models lacked comprehensive feature contribution analysis to support HR decision-making.

Models Evaluated

- **Primary Model**: Stacking Ensemble Model using Random Forest, ANN, and Logistic Regression as meta-learner.
- **Comparative Models**:
  - Logistic Regression
  - Random Forest
  - XGBoost
  - Support Vector Machine (SVM)
  - Artificial Neural Network (ANN)

Three ensemble variants (ESM1, ESM2, ESM3) were tested with different base learner combinations.

Key Findings

- The best-performing model (ESM1: RF + ANN + Logistic Regression) achieved:
  - Accuracy: 97.6%
  - Precision: 99.8%
  - Recall: 95.1%
  - F1-score: 97.5
  - AUC: 97.5%
- All stacking ensemble models outperform individual base models in every metric.
- Feature importance analysis identified **Environmental Satisfaction (4.2%)**, **Overtime (4.2%)**, and **Relationship Satisfaction (3.9%)** as the most influential predictors.
- Other significant features included **Performance Rating** and **Monthly Income**.
- SMOTE was applied to address class imbalance; categorical features were encoded using one-hot and label encoding.

Limitations

- The model does not explain specific relationships between variables and attrition—interpretability remains limited.
- Psychological and subjective factors influencing attrition may be underrepresented in the dataset.
- Further qualitative analysis is needed to identify and incorporate missing predictors.
- Additional techniques like **Partial Dependence Plots** are recommended for deeper interpretability in future work.
