

AI HR Attrition Prediction System

Final Report – Artificial Intelligence Term Project

Group Members:

Muhammad Usman – F2023105003

Noman Aslam – F2023105043

Project Repository: [<https://github.com/muhammadasmancode/AI-HR-Attrition-Prediction>]

1. Introduction

Employee attrition is a critical challenge that can lead to financial loss, decreased productivity, and operational disruption. Early identification of employees who may leave enables HR departments to implement effective retention strategies.

The **AI HR Attrition Prediction System** employs **advanced machine learning algorithms** to predict employee attrition based on historical HR data and individual employee profiles. The system is designed to provide **actionable insights** and support HR teams in making **data-driven decisions**.

Objectives: - Accurately predict employee attrition (Leave or Stay) using AI. - Offer insights to improve HR retention strategies. - Provide an interactive, user-friendly interface for real-time predictions.

2. Problem Statement

Employee turnover can negatively impact organizations in several ways: - High recruitment and training costs for replacing employees. - Loss of experienced and skilled workforce. - Operational delays and reduced productivity.

Traditional HR assessment methods are often subjective, slow, and prone to errors. This project applies **machine learning models** to predict attrition, enabling HR departments to take proactive measures to retain talent.

3. Methodology

3.1 Dataset

- **File:** `dataset.csv`
- Includes features such as age, department, salary, job satisfaction, years at company, work-life balance, and other relevant employee attributes.
- **Data preprocessing steps:**
 - Handling missing values and inconsistent data.
 - Encoding categorical variables (e.g., department, gender).
 - Normalizing numerical features to ensure consistency.
 - Splitting dataset into training (80%) and testing (20%) sets for model evaluation.

3.2 Tools and Technologies

- **Python:** Programming language for data analysis and machine learning.
- **Pandas & NumPy:** Data manipulation, cleaning, and numerical computations.
- **Scikit-learn:** Machine learning library for model training, testing, and evaluation.
- **Joblib:** Efficiently save and load trained ML models.
- **Streamlit:** Framework to develop interactive web applications for predictions.

3.3 Model Training

- The ML model was trained using classification algorithms such as **Logistic Regression** and **Random Forest Classifier**.
- **Training Steps:**
 1. Load dataset with Pandas.
 2. Preprocess features: encode categorical variables, scale numerical data.
 3. Split dataset into training and testing sets.
 4. Train the classifier on the training set.
 5. Evaluate model performance using accuracy, precision, recall, and F1-score.
 6. Save trained model with Joblib (`model.pkl`).

Training Command:

```
python train_model.py
```

3.4 Web Application

- Streamlit app provides a clean, interactive interface.
- **Workflow:**
 1. User inputs employee details via a form.
 2. Click **Predict** to obtain attrition prediction (Leave/Stay).

3. Real-time predictions are generated using the saved ML model.

Run Web App:

```
python -m streamlit run app.py
```

- Fully compatible with modern browsers such as Chrome, Firefox, and Edge.
-

4. Features

- Real-time employee attrition prediction.
 - Intuitive and interactive web interface.
 - Based on real HR dataset for accurate modeling.
 - Actionable insights for HR teams.
 - Scalable and modular design for future expansion.
-

5. Challenges and Solutions

- **Data Preprocessing:** Ensuring data quality by cleaning and normalizing inconsistent values.
 - **Feature Selection:** Identifying features most influential in predicting attrition.
 - **Model Accuracy:** Iteratively tuning hyperparameters for improved performance.
 - **UI Integration:** Connecting the trained model with Streamlit for smooth user experience.
 - **Resource Optimization:** Ensuring efficient app performance even on low-end systems.
-

6. Results

- The trained model achieved strong predictive performance on test data.
- Streamlit interface provides seamless real-time predictions.
- HR teams can leverage insights to make proactive retention strategies.

Example Output: - Employee: Age 28, Department: Sales, Years at Company: 3 -
Prediction: **Stay**

Screenshots Placeholder:

AI HR Attrition Prediction System

Enter employee details so we can check if he is willing to leave or stay

Age

30

-

+

Monthly Income

5000

-

+

Job Satisfaction (1 = Low, 4 = High)

3

Years at Company

5

-

+

OverTime

No

▼

Job Satisfaction (1 = Low, 4 = High)

3

1

4

Years at Company

5

-

+

OverTime

No

▼

Predict

Low Risk: Employee wants to stay

7. AI Usage Declaration

- AI tools assisted in learning programming concepts, debugging, and generating boilerplate code.
 - All model design, customization, and implementation logic were fully understood and executed by the students.
-

8. Project Files

File	Description
dataset.csv	HR dataset used for training
train_model.py	Script to train the ML model
model.pkl	Trained ML model file
app.py	Streamlit application for predictions
README.md	Project documentation
requirements.txt	Python dependencies

Python Requirements:

```
pandas
numpy
scikit-learn
streamlit
joblib
```

9. Conclusion

The **AI HR Attrition Prediction System** offers a professional, AI-driven solution for predicting employee attrition. Integrating machine learning with an interactive web application empowers HR departments to take proactive measures, improve retention strategies, minimize costs, and maintain a stable workforce.
