**Employee Performance Analysis**

**INX Future Inc.**

* Candidate Name : Raja Prabu Manivel
* Candidate E-Mail : rpthalaprabhu@gmail.com
* REP Name : DataMites™ Solutions Pvt Ltd
* Assesment ID : E10901-PR2-V18
* Module : Certified Data Scientist - Project
* Exam Format : Open Project- IABAC™ Project Submission
* Project Assessment : IABAC™
* Registered Trainer : Ashok Kumar A
* Submission Deadline Date : 24-MAY-2025

## ****Functional Requirements****

These are the business and analytical goals the project is expected to meet:

**Predict Employee Performance**

Using classification models to predict performance ratings based on historical data.

**Identify Key Performance Drivers**

Use feature importance and correlation analysis to identify what factors influence employee performance most (e.g., job involvement, overtime, satisfaction levels).

**Generate Business Insights**

Understand HR-related trends: which departments perform well, how tenure affects performance, overtime impact, etc.

**Support HR Decision-Making**

Enable smarter decisions on promotions, training needs, and employee retention strategies.

## ****Technical Requirements****

These include tools, frameworks, libraries, and techniques used:

**Programming Language**: Python

**Jupyter Notebooks**: For code execution, visualization, and reporting.

**Libraries Used**:

**pandas** – for data manipulation

**numpy** – for numerical operations

**matplotlib, seaborn** – for visualization

**scikit-learn** – for ML models, preprocessing, evaluation

**warnings, os, joblib** – utility packages

**Machine Learning Algorithms Implemented**:

Logistic Regression

Random Forest Classifier

Support Vector Classifier (SVC)

**Data Preprocessing Steps**:

Handling missing values

Encoding categorical variables

Feature scaling (StandardScaler)

Data splitting (train/test)

**Model Evaluation Metrics**:

Accuracy

Confusion Matrix

Classification Report (precision, recall, f1-score)

## ****Data Requirements****

**Input Data**: Employee-related HR dataset with features like:

Age, Gender, Department, BusinessTravel, Education, OverTime, JobRole, JobInvolvement, etc.

**Target Variable**:

PerformanceRating or a derived binary label indicating high/low performance.

**Data Source**:

Internal HR dataset (you can note it as an assumption if the source isn’t explicitly defined in the notebooks).