## Employee Attrition Prediction and Analysis

**1. Project Planning & Management**

### Overview:

This project aims to analyze employee attrition using data science techniques. By leveraging machine learning and statistical analysis, we seek to identify key factors influencing employee turnover and provide actionable insights to improve retention strategies.

### Objectives:

- Identify patterns and trends in employee attrition.  
- Develop predictive models to estimate the likelihood of employee turnover.  
- Provide data-driven recommendations to HR teams.

### Scope:

- Data preprocessing and exploration.  
- Application of machine learning models.  
- Evaluation of model performance.  
- Interpretation of results and recommendations.

## 2. Project Plan

### Timeline & Milestones:

|  |  |
| --- | --- |
| **Task** | **Duration** |
| Data Collection & Cleaning | 2 Weeks |
| Exploratory Data Analysis (EDA) | 2 Weeks |
| Model Development | 3 Weeks |
| Model Evaluation & Optimization | 2 Weeks |

### Deliverables:

- Cleaned dataset.  
- EDA report.  
- Machine learning models.  
- Final report with insights and recommendations.

## 3. Task Assignment & Roles

|  |  |
| --- | --- |
| **Name** | **Responsibility** |
| Ziad | Data collection, preprocessing, and cleaning |
| Mahmoud | Model development and evaluation |
| Ahmed | Exploratory data analysis and visualization |
| Youssef | Coordination, timeline management, and reporting |

## 4. Risk Assessment & Mitigation Plan

|  |  |
| --- | --- |
| **Risk** | **Mitigation Strategy** |
| Data Quality Issues | Implement data validation and cleaning procedures. |
| Model Overfitting | Use cross-validation and regularization techniques. |
| Lack of Interpretability | Employ explainable AI techniques and feature importance analysis. |
| Computational Limitations | Optimize model efficiency and use cloud computing if needed. |

## 5. Key Performance Indicators (KPIs)

|  |  |
| --- | --- |
| **KPI** | **Measurement** |
| Model Accuracy | Precision, recall, F1-score of predictive models. |
| Data Completeness | Percentage of missing values in the dataset. |
| Retention Rate | Percentage of employees retained post-intervention. |
| Processing Time | Time is taken for data processing and model training. |