Smart Lender -Human Resource Management:Predicting Employee Promotion Using Ml

Milestone 1: Project Initialization and Planning Phase

The project begins with defining the problem statement and gathering relevant data on employee characteristics, performance, and promotion history. The data is then preprocessed to ensure quality and consistency, followed by feature engineering to extract relevant predictors. The data is split into training, validation, and testing sets, and appropriate ML algorithms are selected for the prediction task.

Activity 1: Define Problem Statement

Problem Statement: The goal is to develop a predictive model that accurately forecasts which employees are likely to receive promotions within the next 12-18 months, based on their historical performance, skills, experience, and other relevant factors. The model should identify top performers and high-potential employees, enabling the organization to proactively develop and retain talent, improve succession planning, and enhance overall business performance.

Ref. template: Click Here
SmartLender Problem Statement Report: Click Here

Activity 2: Project Proposal (Proposed Solution)

This project proposes the development of a machine learning model to predict employee promotions, enabling proactive talent development and retention. The project will collect and analyze HR data, including performance evaluations, skill assessments, and career progression history. The goal is to build a predictive framework that identifies high-potential employees and forecasts promotion likelihood within 12-18 months.

Ref. template: Click Here
SmartLender Project Proposal Report: Click Here

Activity 3: Initial Project Planning

The project begins with defining the scope, goals, and timeline, and assembling a team consisting of a project lead, data scientist, HR analyst, and IT support. Next, the team collects and identifies relevant HR data, historical promotion data, and business metrics, ensuring data quality and preprocessing.

Ref. template: Click Here
SmartLender Project Planning Report: Click Here

Milestone 2: Data Collection and Preprocessing Phase

The data collection phase gathers relevant information from various HR systems, including employee demographics, job descriptions, performance evaluations, skill assessments, promotion history, and tenure.

Additional data points may include training participation, certifications, and business metrics like sales performance or productivity.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

The dataset for SmartLender - The data collection plan involves gathering data from various HR systems, including HRIS (Workday), performance management software (SuccessFactors), and learning management systems (Coursera). Raw data sources identified include employee master data, performance evaluation records, skill assessment data, promotion history, and training participation records

Ref. template: Click Here
SmartLender Data Collection Report: Click Here

Activity 2: Data Quality Report

The dataset for SmartLender - The data quality report reveals a completeness rate of 90%, with some missing values in skill assessment data (10%) and performance evaluation records (5%). While accuracy and consistency are high at 95% and 92%, respectively, there are minor errors in job titles and skill categorization. Timeliness is perfect, with all data up-to-date.

Ref. template: Click Here
SmartLender Data Quality Report: Click Here

Activity 3: Data Exploration and Preprocessing

The data exploration phase revealed interesting patterns and correlations, including a strong correlation between tenure and promotion likelihood, and a significant difference in average performance ratings between promoted and non-promoted employees. Data preprocessing involved handling missing values through mean imputation, encoding categorical variables, and normalizing numerical features.

Ref. template: Click Here
DataSmartLender Exploration and Preprocessing Report : Click Here

Milestone 3: Model Development Phase

The model development phase involved training and evaluating several machine learning algorithms, including logistic regression, decision trees, random forests, and neural networks. Hyperparameter tuning was performed using grid search and cross-validation to optimize model performance.

Activity 1: Feature Selection Report

The feature selection report identifies the most relevant predictors of employee promotions, based on feature importance analysis and recursive feature elimination. The top-ranked features include: tenure (feature importance score: 0.35), average performance rating (0.23), skill density (0.17), career progression velocity (0.12), and job satisfaction (0.08). These features collectively explain 85% of the variance in promotion outcomes.

Ref. template: Click Here
SmartLender Feature Selection Report: Click Here

Activity 2: Model Selection Report

The model selection report compares the performance of various machine learning algorithms for employee promotions prediction, including logistic regression, decision trees, random forests, and neural networks. The random forest model outperforms others, achieving an accuracy of 95%, precision of 0.93, recall of 0.96, and F1-score of 0.95.

Ref. template: Click Here
SmartLender Model Selection Report: Click Here

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code utilizes Python's scikit-learn library to implement a random forest classifier. The model is trained on a dataset of employee features, including tenure, performance ratings, and skill density. Model validation and evaluation report reveals an accuracy of 93%, precision of 0.92, recall of 0.95, and F1-score of 0.93 on the validation set.

Ref. template: Click Here
SmartLender Model Development Phase Template: Click Here

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

The Gradient Boosting model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Gradient Boosting model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Gradient Boosting as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring employee promotion predictions.

Ref. template: Click Here
SmartLender Model Optimization and Tuning Phase Report: Click Here

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, Kindly click the link and refer to the flow. Click Here

For the documentation, Kindly refer to the link. [Click Here](#)

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.