Project Initialization and Planning Phase

Date	18 June 2025
Team ID	SWUID20250177148
Project Title	Machine Learning Approach for Employee Performance Prediction
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) report

The proposal aims to transform employee performance management using machine learning to enable faster and more accurate productivity assessments. It addresses inefficiencies in manual review systems, offering improved operations and decision-making in HR. Key features include a predictive ML model and an interactive web interface for real-time use by managers and HR teams.

Project Overview	
Objective	The primary objective is to enhance employee productivity evaluation by implementing advanced machine learning techniques, ensuring faster, real-time, and more accurate assessments.
Scope	The project covers end-to-end implementation of a machine learning pipeline to predict productivity from structured data and integrates it into a user-friendly web application using Gradio. This enables better team management, resource allocation, and performance support.
Problem Statement	
Description	Addressing inefficiencies in manual productivity assessment systems that cause delays in decisions and limited insights for team leaders and HR.
Impact	Solving this problem results in faster decision-making, early performance alerts, improved employee support, and smarter resource deployment — contributing to organizational productivity and talent retention.

Proposed Solution	
Approach	Employing machine learning techniques (XGBoost, Random Forest, Linear Regression) to predict employee productivity from key performance features. The model is deployed with a real-time Gradio interface.
Key Features	- Predictive ML model using employee productivity dataset - Color-coded productivity insights (Excellent / Moderate / Low) - Input validation and interactive UI via Gradio - Model saved and loaded with .pkl for fast prediction

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	Colab / Local (CPU), optional GPU (T4)
Memory	RAM specifications	8 GB
Storage	Disk space for data, models, and logs	1 GB (code + model + data)
Software		
Frameworks	Python frameworks	Python + Gradio
Libraries	Additional libraries	xgboost, pandas, numpy, scikit-learn, gradio
Development Environment	IDE, version control	Google Colab / VS Code
Data		
Data	Source, size, format	Kaggle Dataset (CSV), ~1200 rows, 14 features