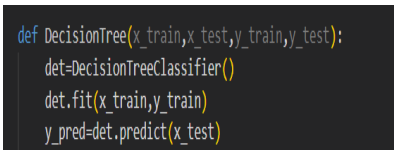
Model Development Phase Template

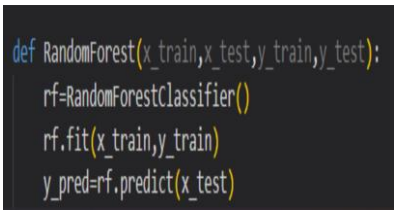
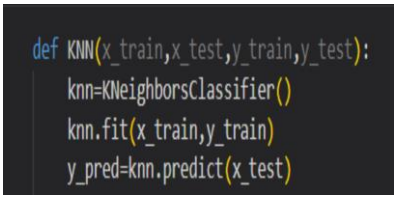
Date	9 July 2024
Team ID	SWTID1720104839
Project Title	Human Resource Management: Predicting Employee Promotion using Machine Learning
Maximum Marks	6 Marks

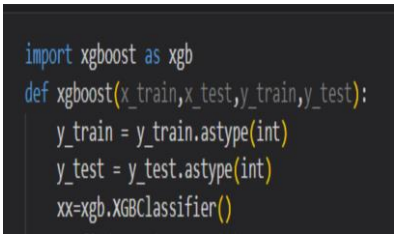
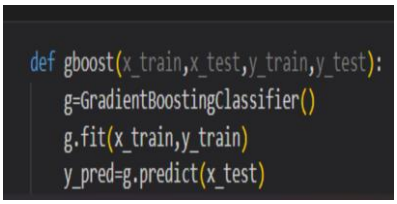
Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
Decision Tree	A decision tree is a simple, tree-based model that splits data into subsets based on features. It's a supervised learning method that predicts the value of a target variable by learning simple decision rules from the data. Decision trees are easy to interpret, but	 Using random state ensures that the randomness involved in initializing and training the model is controlled, leading	Accuracy Score = 93%

	they can be prone to overfitting.	to consistent and reproducible results	
Random Forest	A random forest is an ensemble learning method that combines multiple decision trees to improve the accuracy and robustness of predictions. It works by training multiple decision trees on random subsets of the data and features, and then combining their predictions. Random forests are less prone to overfitting and can handle high-dimensional data.	 Using random state ensures that the randomness involved in initializing and training the model is controlled, leading to consistent and reproducible results	Accuracy Score = 95 %
KNN	KNN is a simple, supervised learning method that predicts the target variable by finding the k most similar instances (nearest neighbors) to a new instance. It's a lazy learner, meaning it doesn't build a model until a new instance is presented. KNN is sensitive to the choice of k and can be computationally expensive.	 **Initializing the model**	Accuracy Score = 90 %

XG Boost	XG Boost (Extreme Gradient Boosting) is a popular, open-source implementation of gradient boosting. It's an ensemble learning method that combines multiple decision trees to improve the accuracy and speed of predictions. XG Boost is known for its high performance, scalability, and ability to handle large datasets.	 Using random state ensures that the randomness involved in initializing and training the model is controlled, leading to consistent and reproducible results	Accuracy Score =96 %
Gradient Boosting	Gradient boosting is an ensemble learning method that combines multiple weak models to create a strong predictive model. It works by iteratively training decision trees on the residuals of the previous tree, with each tree trying to correct the errors of the previous one.	 Using random state ensures that the randomness involved in initializing and training the model is controlled, leading to consistent and reproducible results	Accuracy Score = 82%