Project Milestones – Ensemble Techniques

Criteria	Pts	Week	Topic
1 a. Read 'TelcomCustomer-Churn_1.csv' as a DataFrame and assign it to a variable	1.0 pts	Common part	Basic Data prep operations
1 b.Read 'TelcomCustomer-Churn_2.csv' as a DataFrame and assign it to a variable	1.0 pts	Common part	Basic Data prep operations
1 c.Merge both the DataFrames on key 'customerID' to form a single DataFrame	2.0 pts	Common part	Basic Data prep operations
1 d.Verify if all the columns are incorporated in the merged DataFrame by using simple comparisonOperator in Python	1.0 pts	Common part	Basic Data prep operations
2 a Impute missing/unexpected values in the DataFrame	2.0 pts	Common part	Basic Data prep operations
2 b Make sure all the variables with continuous values are of 'Float' type	2.0 pts	Common part	Basic Data prep operations
2 c Create a function that will accept a DataFrame as input and return pie-charts for all theappropriate Categorical features. Clearly show percentage distribution in the pie-chart.	4.0 pts	Common part	Python functions
2 d.Share insights for Q2.c	2.0 pts	Common part	Basic Data prep operations
2 e.Encode all the appropriate Categorical features with the best suitable approach	2.0 pts	Common part	Supervised Learning
2 f.Split the data into 80% train and 20% test.	1.0 pts	Common part	Supervised Learning
2 g.Normalize/Standardize the data with the best suitable approach	2.0 pts	Common part	Basic Data prep operations
3.a Train a model using Decision tree and check the performance of the model on train and test data	4.0 pts	week1	Ensemble techniques
3.b Use grid search and improve the performance of the Decision tree model , check the performance of the model on train and test data , provide the differences observed in performance in Q3.a and Q3.b	5.0 pts	week2	
3.c Train a model using Random forest and check the performance of the model on train and test data	4.0 pts	week2	Ensemble techniques
3.d Use grid search and improve the performance of the Random tree model , check the performance of the model on train and test data , provide the differences observed in performance in Q3.c and Q3.d	5.0 pts	week2	Ensemble techniques
3.e Train a model using Adaboost and check the performance of the model on train and test data	4.0 pts	week2	Ensemble techniques
3.f Use grid search and improve the performance of the Adaboost model , check the performance of the model on train and test data , provide the differences observed in performance in Q3.e and Q3.f	5.0 pts	week2	Ensemble techniques
3.g Train a model using GradientBoost and check the performance of the model on train and test data	4.0 pts	week2	Ensemble techniques

3.h Use grid search and improve the performance of the GradientBoost model , check the performance of the model on train and test data , provide the differences observed in performance in Q3.g and Q3.h	5.0 pts	week2	Ensemble techniques
3.i Provide detailed analysis of the below steps i Compare the performance of each model in train stage and test stage II) Provide your observation on which model performed the best III)Provide your reasoning on why the model performed best iv)Provide your final conclusion on your observation	4.0 pts	observation on previous steps	Ensemble techniques