

[← Go Back to Data Mining](#)[:≡ Course Content](#)

Project - Data Mining

Submission type	:	File Upload
Due Date	:	Dec 09, 2020, 11:59 PM
Total Score	:	60
Available from	:	Nov 20, 2020, 8:00 AM
Your Score	:	43/60

Description



Scoring guide (Rubric) - Project - Data Mining

Evaluated

Criteria	Ratings	Points
1.1 Read the data and do exploratory data analysis (3 pts). Describe the data briefly. Interpret the inferences for each (2 pts).	Inference missing. EDA done perfectly.	3/5
1.2 Do you think scaling is necessary for clustering in this case? Justify	What type of scaling would suit ? StandardScaler() function	4/5
1.3 Apply hierarchical clustering to scaled data (3 pts). Identify the number of optimum clusters using Dendrogram and briefly describe them (4.5).	Dendrogram not there. Inferences not clear. To generate some more insights, segmentation with more than 2 clusters is preferred	4/7.5


Criteria	Ratings	Points
1.4 Apply K-Means clustering on scaled data and determine optimum clusters (3 pts). Apply elbow curve and silhouette score (3 pts). Interpret the inferences from the model (1.5 pts).	No plots available in pdf.	3/7.5
1.5 Describe cluster profiles for the clusters defined (2.5 pts). Recommend different promotional strategies for different clusters in context to the business problem in-hand (2.5 pts).	Good	5/5
2.1 Data Ingestion: Read the dataset. Do the descriptive statistics and do null value condition check (3 pts), Interpret the inferences from the descriptive statistics in a detailed manner (2 pts).		5/5
2.2 Data Split: Split the data into test and train (1.5 pts), build classification model CART (2 pts), Random Forest (2 pts), Artificial Neural Network(2 pts).	Test - Train ratio not shown in report	6/7.5
2.3 Performance Metrics: Check the performance of Predictions on Train and Test sets using Accuracy (1.5 pts), Confusion Matrix (2 pts), Plot ROC curve and get ROC_AUC score for each model (2 pts), Write inferences on each model (2 pts).	Confusion matrix and interpretation missing	4/7.5

Criteria	Ratings	Points
2.4 Final Model - Compare all models on the basis of the performance metrics in a structured tabular manner (3 pts). Describe on which model is best/optimized (2 pts).		5/5
2.5 Based on your analysis and working on the business problem, detail out appropriate insights and recommendations to help the management solve the business objective.	Recommendation need to be elaborate	4/5
Please reflect on all that you learnt and fill this reflection report: https://docs.google.com/forms/d/e/1FAIpQLSd7e_bJVCiFpZAYbBTMtKrr9TLRnl8kuvZT7IsZ5MSjRtfjcQ/viewform?usp=sf_link		0
Points		43/60

Submitted Assignment

 Shreyansh_M4_Data_Mining_W4_Business_Report.doc 

 Problem 1 Clustering.ipynb 

 Problem 2 CART-RF-ANN 2.ipynb 

Submitted at Sun, Dec 06, 2020

Comments:[+ Add comments](#)[← Previous](#)[Next →](#)

Proprietary content. ©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

© 2021 All rights reserved

[Privacy](#) [Terms of service](#) [Help](#)