Homework 2 Question 2

50 Points Possible

01/11/2024

Attempt 1

In Progress

NEXT UP: Submit Assignment

Add Comment

Unlimited Attempts Allowed

04/10/2024

Details

Introduction

This portion of Homework 2 will be done individually, not in a group. Same as in Question 1, your assignment should be submitted by uploading your code (in the form of a **Jupyter Notebook (.ipynb) AND pdf copy of the files** – so we can make comments directly on the file) to Canvas. **Be sure to run the file before committing so that we can directly see your results.** Please mention all the resources that were used to solve the problem (e.g., websites, books, research papers, other people, etc.). To complete the assignment, you can use any Python (or R) package that you want, but we recommend using Scikit-Learn.

Question

To gain a better understanding of the differences across datasets, perform the same tasks as in Question 1, but on a dataset of your choice (if you worked on a team for Question 1, please do not select the same dataset as your team members). The dataset should contain multiple features (attributes) and you can perform binary or multi-class classification. Make sure to create a train/test/validation split as you find appropriate. Note that part 6 of Q1 is specific to the rocks dataset so you can ignore that part when answering this question.

View Rubric

Assignment 2

Criteria	Ratings		Pts
Question -1	**2 to >0 pts** **Full Marks** Statistical descriptions and Visualizations :1.5 If any special treatment required :0.5	**0 pts** **No Marks**	/ 2 pts
Question-2	**3 to >0 pts** **Full Marks** Computing the PCC:1.5 Scatter Plots :1.5	**0 pts** **No Marks**	/ 3 pts
Question-3	**5 to >0 pts** **Full Marks** Splitting the data in testing , validation and training sets correctly 2.5 Verification of splitting 2.5	**0 pts** **No Marks**	/ 5 pts
Question 4 a Multinomial Logistic Regression	**10 to >0 pts** **Full Marks** Model is implemented correctly :2 Different hyperparameters (C, solver,max number of iterations) have been tried:3 Training, Validation and Testing Performance have been reported :3 Discussion on the impact of different hyper parameters has been done :2	**0 pts** **No Marks**	/ 10 pts
Question 4 b view longer description	**10 to >0 pts** **Full Marks**	**0 pts** **No Marks**	/ 10 pts

Assignment 2

Criteria	Ratings		Pts
	Model is implemented correctly :2 Different hyperparameters (C, Kernel, Gamma, degree) have been tried:3 Training, Validation and Testing Performance have been reported :3 Discussion on the impact of different hyper parameters has been done :2		
Question 4 c view longer description	**10 to >0 pts** **Full Marks** Model is implemented correctly :2 Different hyperparameters(no. of trees, max depth ,the minimum number of samples required to split an internal node, the minimum number of samples required to be at a leaf node) have been tried:3 Training, Validation and Testing Performance have been reported :3 Discussion on the impact of different hyper parameters has been done :2	**0 pts** **No Marks**	/ 10 pts
Question 5 view longer description	**10 to >0 pts** **Full Marks** Ensemble classifier has been implemented via all the models with the best hyperparameters :4 Accuracy of the ensemble is greater than all the individual classifiers :2 Test set Accuracy :1 Discussion on Findings	**0 pts** **No Marks**	/ 10 pts
			Total Points: 0

Choose a submission type

Upload More

Choose a file to upload

or

Webcam Photo