Unit 4: Assignment
Due: Sun Apr 13, 2025 11:59pm



Unlimited Attempts Allowed

∨ Details



ASSIGNMENT



Take the dataset from last week and re-run the analysis using tree-based models. Include the following:

- A bagging model, such as Random Forest or Extra Trees.
- A boosting model, such as XGBoost, LightGBM, or Catboost.
- A SHAP plot of one bagging model and one boosting model. Discuss the differences in the SHAP values.
- A partial dependence plot of the same feature (ideally, one that is important per SHAP) for both a bagging model and a boosting model. Discuss the differences between the two plots.
- Please show all your code.



Criteria for Success

• You can reference the attached rubric that I will use to grade this assignment. As always, let me know if you have questions!

• Due Sunday, 11:59 p.m., CT.

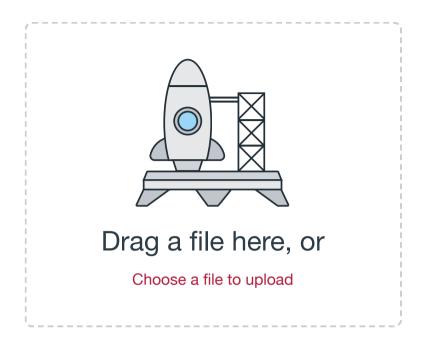
∨ View Rubric

Assignment Rubric				
Criteria				Points
Content organization and relevance	Excellent Content is well organized and relevant to the assignment.	Average Content organization has minor issues; or/and minor parts of the assignment are missing.	Insufficient Content is poorly organized, or irrelevant to the assignment.	/30 pts
	25.1 to 30 pts	15.1 to 25 pts	15 pts	
Writing and APA	Excellent	Average	Insufficient	/10 pts
style	The content is well-written, sources are properly referred in APA style when applicable.	The content has minor writing issues, or there are minor issues with reference formatting.	The writing is poor. APA format is ignored when applicable.	'
	7.1 to 10 pts	5.1 to 7 pts	5 pts	
Methodology	Excellent	Average	Insufficient	/40 pts
	Methodology appropriate for the task has been fully followed	Minor issues with the methodology	Major issues with the methodology	
	30.1 to 40 pts	20.1 to 30 pts	20 pts	

Choose a submission type







or Canvas Files

⟨ Previous |

Submit Assignment

Next >

(https://canvas.park.edu/courses/85581/modules/items/6163552)

(https://canvas.park.edu/courses/85581/modules/items/616356