

ACAD 499: Machine Intelligence

Final Project Part 2

Due: 11/22/2021 @ 11:59PM

Goal

For this second part of the final project, you are to explore and prepare your data for training with the ML algorithm that you've chosen for your analysis. You aren't training your ML algorithm for part 2, you are just preparing the data so that it is ready for ML algorithm training.

Setup

You get to create whatever Python scripts you need. You probably will need some custom transformer classes. They should go into their own Python script file.

```
...
Name
IDSN 599, Fall 2021
USC email
Final Project Part 2
...
```

Requirements

A document that contains the additional sections to your document from part 1. Just add to your document that you submitted for part 1

- **ML Algorithm:** You are to choose three ML algorithms and describe why you have selected those algorithms. Justify your choices of ML algorithms.
- **Correlations:** Show some numbers, or screen shots of plots to show which pairs of your data attribute correlations are most likely to be of value in your analysis. I suggest also showing plots as it really shows the quality of the correlation, instead of just numeric values.
- **Transformers:** Discuss each of your transformers, whether custom or not. Explain why you chose this transformer and what are each of them doing. I suggest using at least the Pipeline class. You may also want to use the ColumnTransformer class if you have categorical (text) data.

Deliverables

A compressed folder containing your document and all Python code you used to prepare your dataset for training of your ML algorithms.

Grading

Item	Points
ML Algorithms	10
Correlations	10
Transformer discussion	10
Transformer Code	10
Python Code	10
Total	50