IDSN 542: Introduction to Machine Intelligence

Final Project Part 1 Due: 10/29/2025 @ 11:59PM

Goal

For this part of the final project, you are to find your dataset, make a preliminary decision/hypothesis as to what you are going to be doing, download the dataset and play around with the corr() function to see where the linear correlations might be.

Setup

- You get to create whatever Python scripts you need. You can probably do it in one file for this part of the final project.

```
'''
Name
IDSN 542, Fall 2025
USC email
Final Project Part 1
'''
```

Requirements

A document that contains the following sections:

- **Domain**: What is the type/nature of the problem that you are thinking to solve
- **Dataset**: Where did you get your dataset? How big is it – how many rows and attributes? Why/how does it contain the data you need to solve the problem you are thinking of solving?
- **Problem Type**: Are you creating a predictor or a classifier? Why do you think this is the way to to solve the problem?
- **Attributes**: What are the attributes of the dataset? Which are numeric and which are text? Do you have any missing values for attributes? If you have missing values, how are you thinking of taking care of them?

You are to write some Python code to identify your dataset attributes, their data type, and any missing values. You can also use the info() function to print out the attributes and the number of values each attribute has. Also, use the corr() function to find out if you have any linear correlations on the numeric attributes.

Deliverables

A ZIP file submitted to Brightspace.

Grading

Item	Points
Domain	5
Dataset	5
Problem Type	5
Attributes	5
Python Code	10
Total	**30**