

# IDSN 542: Machine Intelligence

Final Project Part 2

Due: 11/19/2025 @ 11:59PM

## Goal

For this second part of the final project, you are to explore and prepare your data for training with the ML algorithms that you've chosen for your analysis. You are to train 3 ML models, but not test them. That is part 3.

## 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.

## 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. You can use <https://www.geeksforgeeks.org/choosing-a-suitable-machine-learning-algorithm/> to help with this part.
- **Correlations:** Show some numbers, or screen shots of plots to show which pairs of your data attribute linear correlations (if you have any) 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. This only works with numeric data. You are to include ROC/AUC charts, or something equivalent, of your 3 chosen algorithms to determine which performs the best.
- **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.
- Train your prepared data on the 3 algorithms you chose. Use GridSearchCV or RandomSearchCV to explore the hyperparameter space for the 3 ML models you chose.

## Deliverables

A compressed folder containing your document, the dataset(s) you are using and all Python code you used to prepare and train 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