

# Assignments Part 3 - ML Algorithms

You can share your work by uploading jupyter notebooks to a github repo and sharing a pull request for each assignment.

## **Grading Criteria:**

1. Use of libraries/algorithm selection
2. Execution
3. Comments
4. Model Performance - accuracy

## **Bonus Points For:**

1. Deeper understanding of the algorithm - The math behind it OR Various applications of the algorithm.
2. What other algorithms can be used to solve the same problem?
3. Execute a different algorithm for the same problem and compare accuracy.

# Assignment 1 - Supervised Learning

**Objective:** Titanic Prediction - Determine how of the passengers is most likely to survive to sink of the ship? Use two different algorithms for this purpose and compare them.

**Data Set:** <https://www.kaggle.com/c/titanic>

## Tasks:

1. Select an algorithm (Suggestion : **Logistic Regression**)
2. Data cleaning
  - a. Add default value for age column
  - b. Classify the p class column into 1st class, 2nd class and 3rd class
3. Determine features and label
4. Split data into training and testing sets
5. Train the data set
6. Test the data set
7. Find the accuracy

**Expected Output:** Based on the passengers information predict if they will survive or not.

# Assignment 2 - Unsupervised Learning

**Objective:** Investigating what people think about masculinity

**Data Set:** [https://drive.google.com/drive/u/0/folders/1c7fKVQ9RX\\_Ywnr8JKEGIZegQjsS\\_8TPR](https://drive.google.com/drive/u/0/folders/1c7fKVQ9RX_Ywnr8JKEGIZegQjsS_8TPR)

## Tasks:

1. Select an algorithm (Suggestion: Kmeans)
2. Import libraries
3. Explore and Clean Data- You may use visualizations where you deem appropriate
4. Convert the rating("often", "likely" etc) to numeric values
5. Get Clustering! ( experiment with different values of n and how the output differs)
6. Train and test
7. Bonus: Use a visualization to show the best case of 'n' clusters

**Expected Output:** Based on the passengers information predict if they will survive or not.

# Assignment 3 - Deep Learning

**Objective:** Cat vs Dog : Identify if a certain image contains cat or dog

**Data Set:** <https://www.kaggle.com/shaunthesheep/microsoft-catsvsdogs-dataset>

## Tasks:

1. Select an algorithm (Suggestion: CNNs)
2. Import libraries
3. Define image properties
4. Prepare dataset for training model
5. Create the neural net model
6. Analyze the model
7. Define callbacks and learning rate
8. Split data into training and testing
9. Train, Test, Predict!

**Expected Output:** Teach a model how to differentiate between cats and dogs. Upon inputting image, model should be able to predict which animal is in the provided picture