Project Proposals

## Titanic Dataset, Machine Learning from Disaster:

Analyze the Titanic dataset to predict the survival of passengers. This is a learning project on Kaggle which provides a lot of insight into logistic regression.

From Kaggle website: “the shipwreck led to such loss of life was that there were not enough lifeboats for the passengers and crew. Although there was some element of luck involved in surviving the sinking, some groups of people were more likely to survive than others, such as women, children, and the upper-class”

Using the dataset provided on kaggle, we would like to analyze what kind of people are likely to survive.

The dataset for this exercise is available on Kaggle here.

<https://www.kaggle.com/c/titanic/data>

## Sentiment Analysis on UCI dataset:

UCI provides a dataset, with positive and negative sentiment sentences from three sources, Amazon, Imdb and Yelp. As part of this project, these datasets will be analyzed to predict the sentiment of each of these sentences.

The Dataset is available at the below link on UCI:

<http://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences>

## Whats Cooking on Kaggle:

This project has been picked from one of the Kaggle competitions. Even though this might not address a real problem (or may be it does?), it has a lot of scope for learning and similar principles can be applied to several other classification problems.

The primary objective of the project is to predict the type of cuisine based on the list of ingredients used in the dish. A classification model will be developed based on a training dataset and then be used to predict the cuisines in a test dataset.

The Data source for this project is provided by Yummly on Kaggle [here](https://www.kaggle.com/c/whats-cooking/data).