

Predicting Survivors of Titanic Shipwreck

Heath Thompson

Computer Science

CS 477 Spring 2021

04/26/2021

Feasibility

After looking at the datasets provided by this link: <https://www.kaggle.com/c/titanic/data>, I was able to determine that the amount of time necessary for preprocessing the data was minimal. This is important as I would like to focus my time on not only developing the initial model but also put forth ample time in tuning it for better accuracy. The data is not too monstrous or complex for me to handle as it mostly contains pertinent information that is given in the necessary format.

Model development is within my understanding of ML concepts. We have discussed several methods for approaching a classification problem type. I can utilize one of these discussed methods to solve this problem.

Lastly, I am able to tune my model using concepts taught in class.

Approaches

Because the problem's goal is to determine whether a person survived or died in the wreck, I'm dealing with a classification problem. From here, I can utilize many different techniques. Logistic Regression and K nearest neighbors are a couple of approaches I could use for predicting outcomes.

Timeline

04/26 - 04/29: Data processing and understanding, formatting and modifying for use in model

04/30 - 05/02: Initial training and testing of model with some tuning

05/03: Final ideas/changes, write up report for FP M2 discussing project progress

05/04 - 05/09: Further tuning to model, finalize and/or utilize any ideas for project

05/10 – 05/12: Write up results for FP M3 presenting my model, results, and project evolution