

Portfolio Component: ML From Scratch

```
Survived (Coefficients):  
0.999877    -2.41086  
  
Predicted values:  
Accuracy:    0.784553  
Sensitivity:  0.695652  
Specificity:  0.862595  
Training Time of the algorithm: 2451ms
```

From this output we can see that The Logistic Regression model is showing some hopeful values. Based off the predictors; class, age, and sex from the Titanic dataset we can see that the formula deduced the accuracy for survivability to be at a rounded 78%. The other numbers Sensitivity nearly 70% and Specificity at 86% show that this formula was able to almost accurate calculate and deduce who had survived and died on the Titanic.

The training time for the algorithm is slow in my opinion at 2447ms. I was not able to create the Naïve Bayes model, but I assume that it would take less time than logistic regression. Since, in logistic regression the program must take more time to calculate the weights of each of the predictors.

“Reproducibility is the ability to be recreated or created (towardsdatascience).” An algorithm without reproductivity will be hard to navigate and implement. Reproducibility in Machine Learning means that you can run your algorithm multiple times and obtain the same results with accuracy. You must be able to accurately record changes in data. This is what allows the data to be reproducible and the same output achieved again. Using modeling such as Logistic Regression and Naïve Bayes we can record changes that happen in data clearly.

Citations

<https://bdtechtalks.com/2021/03/01/papers-without-code-machine-learning-reproducibility/>

<https://neptune.ai/blog/how-to-solve-reproducibility-in-ml>

<https://blog.ml.cmu.edu/2020/08/31/5-reproducibility/>