

David Teran

CS4375

March 4, 2023

## ML Algorithms from Scratch

```
Opening file: titanic.csv for data exploration.
Headings: "", "pclass", "survived", "sex", "age"
--- Model Coefficients and Data Output ---

Weight 1: 1.20469
Weight 2: -0.758992

Accuracy: 0.662602

Training time after 1k iterations: 364 ms

--- Confusion Matrix Result ---

Treating values 1 'sex' as the positive class

pred  0   1
0 112  64
1  19  51

Sensitivity: 0.443478
Specificity: 0.854962
```

Using the titanic.csv data, the logistic regression using survived and sex data columns, with sex as the predictor, came out more accurate, but with a lower sensitivity and a higher specificity/

```
c2r.cpp:8: NaiveBayes_Part2 @ /Users/david/Documents/CS4375/CS4375_Project3/NaiveBayes_Part2
Opening file titanic.csv for data exploration.
Headings: "", "pclass", "survived", "sex", "age"
Coefficient Output:

Prior Probabilities:
0: 0.61 1: 0.39

Pclass Probability:
0.172131 0.22541 0.602459
0.416667 0.262821 0.320513

Sex Probability:
0.159836 0.840164
0.679487 0.320513

Age Statistics:
Mean: 30.4182, Variance: 205.153
Mean: 28.8261, Variance: 209.155

Training Algorithm Duration (Milliseconds): 0
Accuracy: 0.000666667
Sensitivity: nan
Specificity: 0.000666667
```

With Naïve Bayes, the results became less accurate when using sex, age, and pclass as predictors for predicting survivability.