

<https://github.com/bbelandrr/TitanicMachineLearning>

1. Provide a justification for your design and experimental decisions.

In both of my models, I removed most columns except for Pclass, Age, SibSp, Parch, and Sex. I believed that these features would be the most important for determining if someone died or not.

Sex was one-hot encoded because it is categorical data.

The rest were normalized to make the log loss model work better with the features.

2. Report validation accuracy on SKLearn and From Scratch models. Is the accuracy the same?

Can you provide any insights on model performance?

SKLearn gave me 75.4% accuracy, and my personal model apparently was 0% accuracy.

I believe that my personal model performed so poorly because of the way that I submitted my data? I'm not entirely sure why it literally has a 0% score. I spent so long trying to get this to work that I'm just going to submit what I have.

I would like to know what I did wrong? The code I have written may not be the best organized code in existence.

3. Submit to the competition your best performing model. What was your final test score and ranking?

My SKLearn model achieved a score of 75.4%. That puts me at 11,385<sup>th</sup> place.

I beg for your mercy in regards to the late policy; I am submitting this around 50 minutes after 12:00 AM after a full day of work