Titanic Kaggle

**Model 1**

Data Input: Pclass, Sex, Age, Fare, Embarked

Inputs:

[is\_first\_class, is\_second\_class, is\_third\_class,   
is\_female,   
age\_normalized,   
fare\_normalised,   
embarked\_s, embarked\_c, embarked\_q]

Architecture:

Linear: in:10, out:6  
Sigmoid  
Linear in:6, out:2

Loss function: CrossEntropy on sigmoid(out), optimiser: Adam

Highest validation\_accuracy: 81.7%

Notes:

I want to change representation of age as float to a one-hot {under\_7, 7\_or\_older, unknown}

Going to start without a hidden layer and no sigmoid, then see if architectural complexity benefits results.

Data is imbalanced, more not\_survived than survived, might balance classes to help learning.

Want to measure precision and recall

Model 2

Data Input: Pclass, Sex, Age, Fare, Embarked

Inputs:

[is\_first\_class, is\_second\_class, is\_third\_class,   
is\_female,   
under\_7, 7\_or\_older, unknown   
fare\_normalised, (filled with mean)  
embarked\_s, embarked\_c, embarked\_q, unknown

Architecture:

Linear: in:10, out:6  
Sigmoid  
Linear in:6, out:2

Loss function: CrossEntropy on sigmoid(out), optimiser: Adam

Highest validation\_accuracy: 82.7%

Notes:

Chart, bar chart, histogram

Description automatically generatedShould add a fourth bin to age, for those aged 64 and over, as suggested by this graph



Going to start without a hidden layer and no sigmoid, then see if architectural complexity benefits results.

Want to measure precision and recall

Model 3

Data Input: Pclass, Sex, Age, Fare, Embarked

Inputs:

[is\_first\_class, is\_second\_class, is\_third\_class,   
is\_female,   
unknown, under\_7, 7\_to\_63, 64\_or\_older   
fare\_normalised, (filled with mean)  
embarked\_s, embarked\_c, embarked\_q, unknown

Architecture:

Linear: in:10, out:6  
Sigmoid  
Linear in:6, out:2

Loss function: CrossEntropy on sigmoid(out), optimiser: Adam

Highest validation\_accuracy: 85.4749%

Notes:  
included checkpointing at maximum validation accuracies.