

# Pulsar Prediction

Andrew Bergman  
16 September 2019

# Agenda

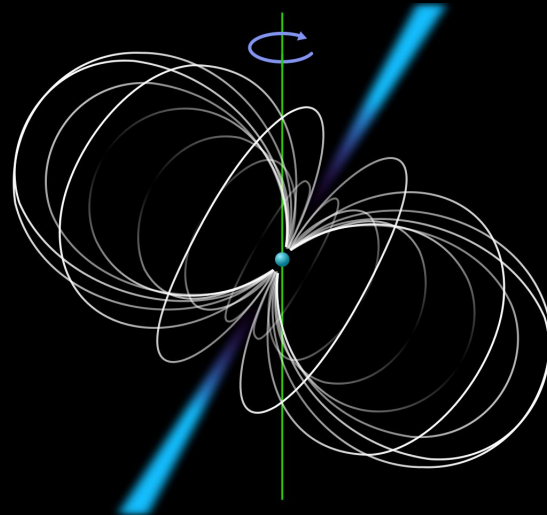
- ❖ Problem statement
- ❖ Pulsars
- ❖ Preprocessing
- ❖ Modeling
- ❖ Evaluation
- ❖ Conclusions & Recommendations

# Problem Statement

- ❖ Pulsars are very important to modern physics and astrophysics: they have been used to study nuclear physics, General Relativity, and have even been instrumental in the discovery of gravitational waves. Because of the vastness of space and how distant they are, pulsars are uncommon and very difficult to identify.

# Pulsars

- ❖ Stellar remnants formed from the cores of giant stars
- ❖ Emit a jet of electromagnetic radiation which is observed as a pulse
- ❖ Have been used to test General Relativity and probe the interstellar space



# Pulsars



# Preprocessing

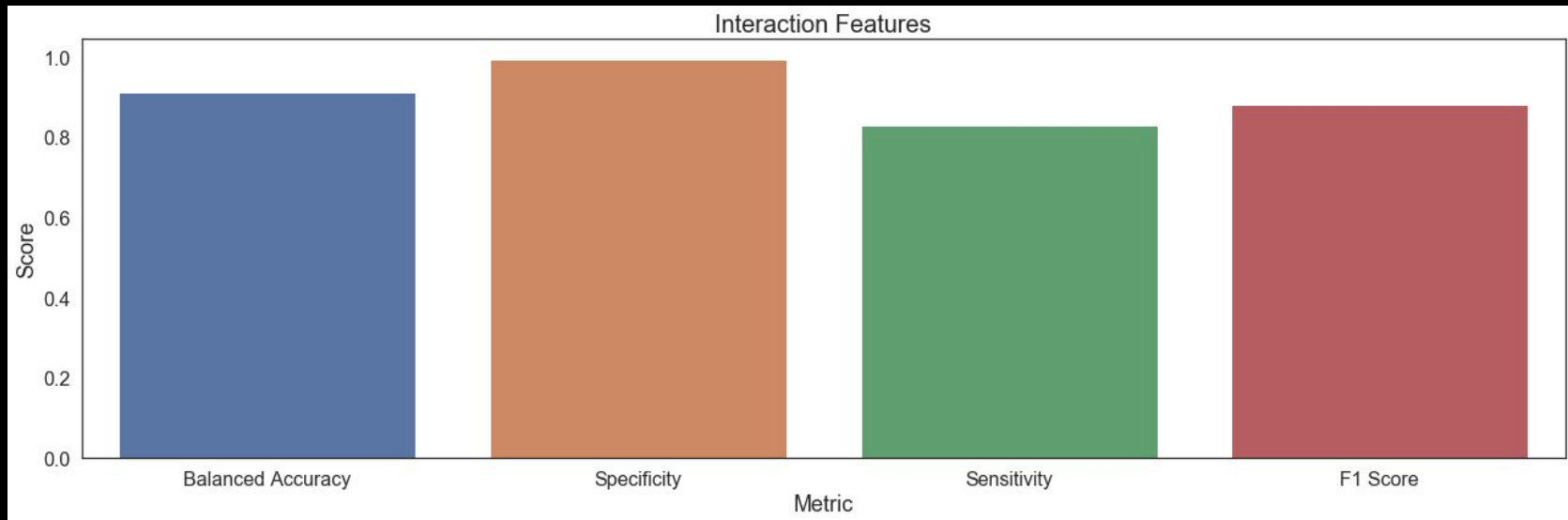
- ❖ The data was *very* clean
- ❖ We only had to shorten column names
- ❖ There were interesting things in the visuals we made
- ❖ Two columns were close to being normally distributed
- ❖ There were very strong correlations between some columns
- ❖ Our classes are very imbalanced: the target is 9.1% of the data

# Modeling

- ❖ We used a neural network because of the severe imbalance
- ❖ The model was run on 3 subsets of the data
- ❖ 2 of the subsets had features we engineered
- ❖ The engineering was based on what we saw in our visuals
- ❖ We kept the models simple with only 1 hidden layer
- ❖ L2 regularization was incorporated to prevent overfitting

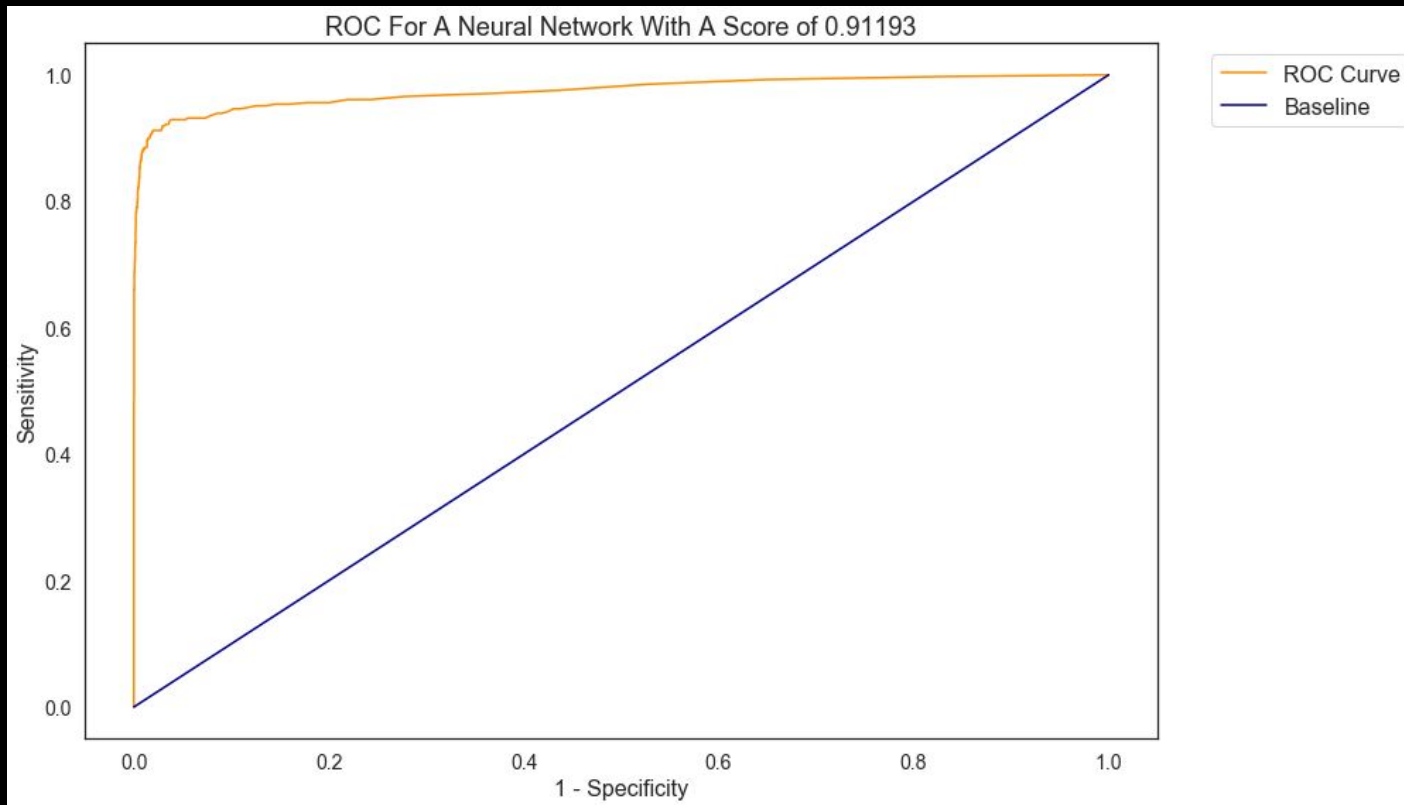
# Evaluation

- ❖ The best model had the interaction features





# Evaluation



# Conclusions & Recommendations

- ❖ We can confidently predict pulsars
- ❖ There were minimal false negatives
- ❖ There was negligible overfitting
- ❖ Going forward we would like to try different regularization methods