Sentencing Algorithms in the Criminal Justice System

Kethayne Trader
kethayne@stanford.edu

Brandon Loong
loong@stanford.edu

CS 229A

Introduction

Goal: Use criminal data & characteristics to predict sentencing lengths and if a life term will be given

Overview of the Criminal Justice System

Timeline of Sentencing Reform [1]

Pre-1900—Retributivism (punish the crime)
- Sentences based on crime committed, not the criminal

1900—Rehabilitation (punish the criminal)
- Shift towards individual sentencing, more consideration for the criminal and discrimination

1984—Sentencing Reform Act (SRA) of Congress
- Established clearer practices and guidelines

2000—Evidence-Based Sentencing (EBS)
- Actuarial approach (scientific & quantitative methods)

Data Processing

- Source: US federal survey of inmates [2]
- Structure: raw data contained 14,499 examples (criminals) and 390 features. Examples that contained only partial feature sets were deleted, leaving 12,330 analyzed. Features that were highly dependent were excluded, leaving 19 features analyzed.
- Challenges: one difficulty was the absence of actual data in this survey. For example, the data says the criminal was 25-34 years old instead of the exact age. It was difficult to interpret trends with this granularity.
- Features: 19 features were included, such as drug usage, incarcerations, race, gender, and age.

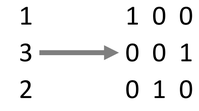
Cited Publications

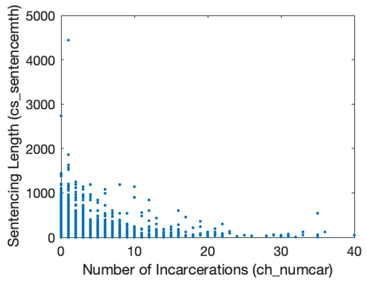
1. D. Kehl, P. Guo, S. Kessler. **Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing.** *Responsive Communities Initiative, Harvard Law School.* (2017)
2. ICPSR. **Survey of Inmates in State and Federal Correctional Facilities [United States]**. *United States Department of Justice, Office of Justice Programs, Bureau of Justice Statistics* (2004)

Linear Regression Model

The linear regression model was used on non-life sentences to predict the lengths based on the 19 features

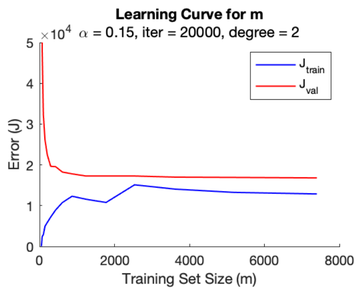
- Since all data was categorical, we converted the features using *one-hot encoding* to properly fit the linear model

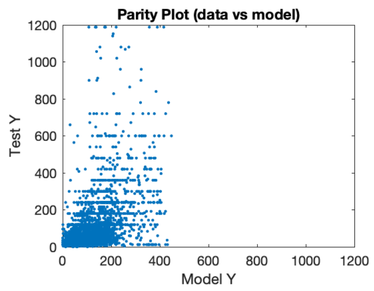




- One feature contained numerical data, so we generated polynomial features from it to improve the complexity

- We excluded significant regularization because our model suffered from high bias, thus, we did not want to underfit any further

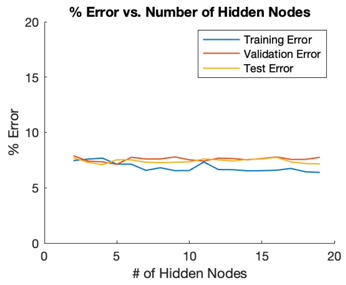




Neural Network Model

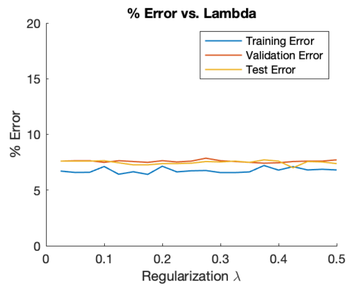
The neural network model was used on life sentences to predict how likely a criminal was to be sentenced for life

- We tried using the NN to predict everything (sentence length & life), so we split the y-data into sentence buckets, but the accuracy was low
- We generated another feature to note the presence or absence of a life sentence.



- The data was again transformed through encoding, and the number of hidden nodes varied to see the impact of model complexity on the accuracy

- Regularization was added, but showed little effect on the final result. Given the complexity of our data, we didn't expect regularization to benefit the model.



Summary

Conclusions

- The data lacked the detail to train a sufficiently complex linear regression model. We expected better results, but underestimated how non-correlated the data was.
- The neural network had better success dealing with the data. It reached accuracies above 90% on smaller subsets.

Future Work

- Obtain better data! Preferably, data that contains actual data rather than categorical data to improve the linear regression model
- Find more features to improve the complexity of the models to enhance the fit
- Incorporate the ideas of precision/recall due to the rare occurrence of life sentences compared to non-life