Recidivism of criminal offenders

A machine-learning approach to justice

Estienne Granet & Victor Lei

Harvard University

Abstract

Recidivism of criminal offenders is a difficult topic in criminal justice. If we view the justice system as serving an important rehabilitatory purpose, then recidivism represents indicators for potential improvement in the correctional system. Using a public data-set on recidivism, we build off of prior work that used both traditional techniques, and machine learning techniques to find the probability of a criminal offender re-offending in the near future given his or her background, as well as the amount of time served. From this probability, we derive the optimal sentence given an individual. Our approach employs two classification methods in an attempt to improve accuracy:

- Random Forests
- Neural Networks

Introduction

Our goal is to develop a model that accounts for the offender's background and both predicts the probability of recidivism, and computes the optimal sentence. We define the optimal sentence as the sentence that minimizes the probability of recidivism given a set of constraints like making sure life imprisonment does not become the default answer. In particular, we would like to:

1. derive a predictor P for the probability of recidivism. The predictor would be a function that takes an individual's background, the nature of their offense and the sentence as inputs, and returns the probability of recidivism as an output; and
2. find the optimal sentence for a given individual with a set of constraints.

$$\text{argmin}_{\text{sentence}} P(\text{sentence}, \text{others})$$

under constraints sentence $\leq$ threshold(k) where k is the offense committed

We are using an ICPSR dataset which consists of the profiles of 9,327 prisoners released in 1978 and the profiles of 9,549 prisoners released in 1980. Each entry has 19 variables such as gender, age, sentence, and whether recidivism occurred within the follow-up period (around 4 years.)

Neural network

Neural networks were discovered some time ago, but lately, they have been experiencing a resurgence in usage. There is likely significant complexity in the interactions between the different variables, so simple models like logistic regression are not going to be as capable of modeling them compared to neural networks. Prior work on this data set has considered the use of neural networks and found them to have good predictive power for recidivism compared to traditional modeling techniques like logistic regression models. We seek to harness the improvements in neural networks that have been made over the last few years in an attempt to improve predictive accuracy. We find that we can achieve a small, but notable improvement in accuracy on unseen data.

As this is a plain binary classifications problem, we use a multi-layer perceptron feedforward neural network with a single hidden layer, bias terms, and the logistic function as the activation function (f) for both the hidden and output layers. There is a single output neuron ($\hat{y} \in (0, 1)$) used for classification with a 0.5 threshold. For some input vector x, linear weights W, and bias terms b:

$$\hat{y} = f(W_2 f(W_1 x + b_1) + b_2)$$

$$E = -y \ln(\hat{y}) - (1 - y)\ln(1 - \hat{y})$$

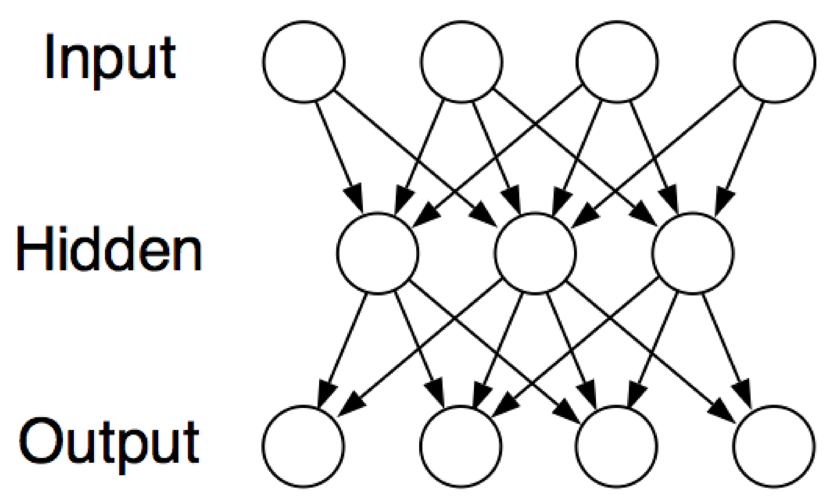

Figure 1: Feedforward Neural Network[1]

Using cross-entropy error (E), we can apply backpropagation and optimize the weights and bias terms by stochastic gradient descent using $\frac{\partial E}{\partial W}$, and $\frac{\partial E}{\partial b}$. In order to counteract overfitting, we use an early-stopping heuristic to stop training the model when the validation set performance does not improve for a certain number of epochs.

Dataset	Palocsay et al.	Our results
1978	69.20% (39)	70.69% (9)
1980	66.98% (26)	68.22% (5)

Table 1: Validation set performance

After splitting the 1978 dataset and the 1980 dataset into training and validation sets (7:3 ratio) and training the neural network, validation set performance is improved compared to the results from Palocsay et al. The results are also achieved with far fewer nodes in the hidden layer, making it faster to train and likely more generalizable. As the figures below reveal, the neural network starts to quickly overfit as the number of nodes in the hidden layer increases.

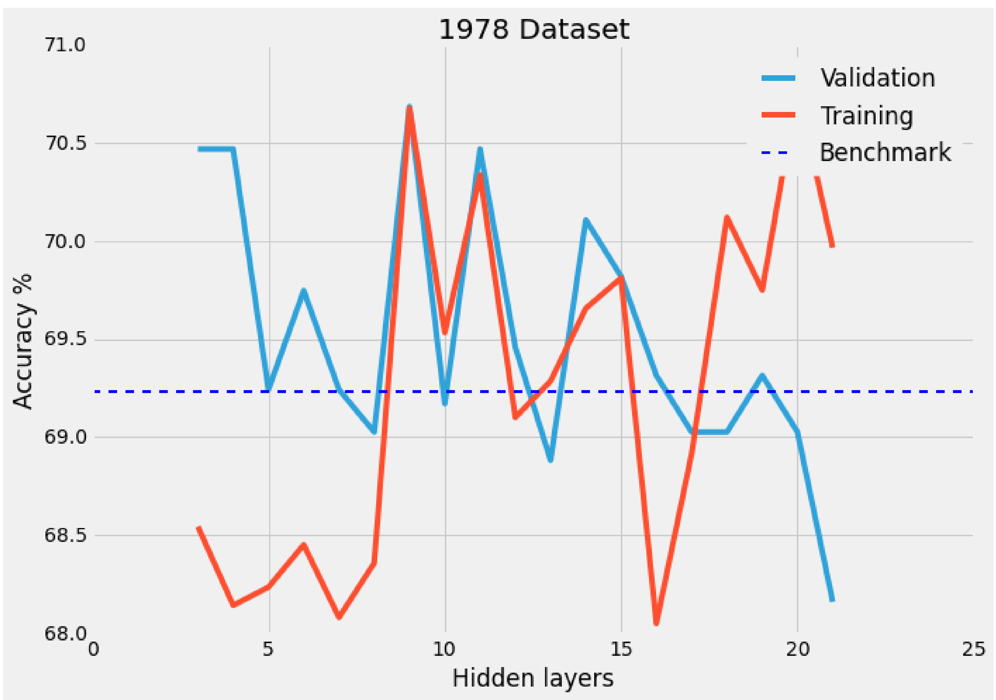

Figure 2: 1978 Dataset

Random Forest

Phasellus imperdiet, tortor vitae congue bibendum, felis enim sagittis lorem, et volutpat ante orci sagittis mi. Morbi rutrum laoreet semper. Morbi accumsan enim nec tortor consectetur non commodo nisi sollicitudin. Proin sollicitudin. Pellentesque eget orci eros. Fusce ultricies, tellus et pellentesque fringilla, ante massa luctus libero, quis tristique purus urna nec nibh.

Nulla ut porttitor enim. Suspendisse venenatis dui eget eros gravida tempor. Mauris feugiat elit et augue placerat ultrices. Morbi accumsan enim nec tortor consectetur non commodo. Pellentesque condimentum dui. Etiam sagittis purus non tellus tempor volutpat. Donec et dui non massa tristique adipiscing. Quisque vestibulum eros eu. Phasellus imperdiet, tortor vitae congue bibendum, felis enim sagittis lorem, et volutpat ante orci sagittis mi. Morbi rutrum laoreet semper. Morbi accumsan enim nec tortor consectetur non commodo nisi sollicitudin.

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 2: Table caption

Vivamus sed nibh ac metus tristique tristique a vitae ante. Sed lobortis mi ut arcu fringilla et adipiscing ligula rutrum. Aenean turpis velit, placerat eget tincidunt nec, ornare in nisl. In placerat.

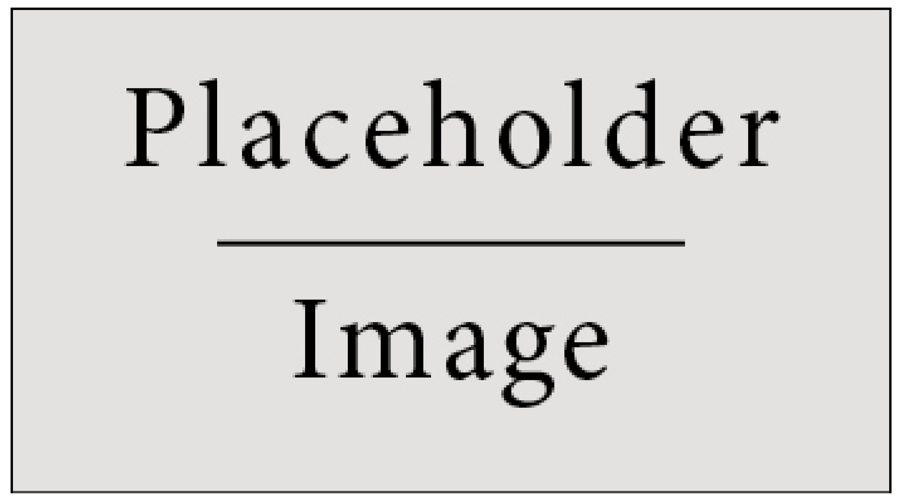

Figure 3: Figure caption

Conclusions

- We improve upon the existing work done by Palocsay et al. on this data set by achieving a higher validation set accuracy using fewer nodes in the hidden layer.

Forthcoming Research

We are still exploring other ways of optimizing the neural network and more advanced regularization strategies like Dropout.

References

[1] Michael I Jordan and Chris Bishop. An introduction to graphical models.