# Nikita McClure Midterm

## 10/25/2024

*Pursuing Open-Source Development of Predictive Algorithms: The Case of Criminal Sentencing Algorithms* argues that open-source penalized regression algorithms are in some ways superior to "black-box" algorithms and should be used, especially in the case of criminal sentencing. Specifically, they argue that specific open-source algorithms: ridge regression, LASSO regression, and elastic net regression are superior to the commonly used black-box algorithm COMPAS for criminal sentencing. They argue that this is the case due to a few factors: increased accuracy, decreased racial bias, improved affordability, the possibility of collaboration, and the ability for the public to test the algorithm. Throughout this essay, the methods that the researchers utilized to prove the aforementioned features will be elaborated upon as we discuss their importance.

When a person is convicted of a crime, the judge must decide upon a fair punishment for their crime. Judges must attempt to account for a few factors in this sentencing, one of which is if or when the person is released, whether they will commit another crime, or "recidivate". This is a very difficult thing for a human to predict, so the opportunity to utilize algorithms that have been designed to assist in this analysis is invaluable. The severity of punishment has a large impact on the course of their future. It impacts all areas of their life, including employment, family planning, and their future lifestyle. Due to the incredible importance this decision has on a person, it should be as accurate to the individual as possible. The possibility of the judge having assistance in making a fair, accurate decision could be incredibly impactful. Having the opportunity to utilize an algorithm that can help the judge predict the likelihood of recidivism can take a huge burden off the judge and can make it more fair for the individual. However, this is only the case if the algorithm itself is a fair and valid tool for the judge to utilize.

An algorithm's fairness can be improved by being an open-source algorithm, rather than a black-box due to its transparency and collaboration potential. As open-source algorithms are public, anyone who desires can view, test, and assist in contributing to it. This means that the algorithm has much potential to be continuously and significantly improved upon because many minds from various disciplines can contribute to it – likely resulting in many more iterations than would occur if it were dependent upon private hires. Also, due to its transparency, defendants have the opportunity to see why they are being sentenced how they are, securing their right to due process.

Additionally, and possibly most significantly, this study demonstrates that each of these specific open-source algorithms can already be shown to be both more accurate and less biased than COMPAS. Findings have shown that COMPAS has a lower accuracy when predicting for Black individuals than for white individuals. Additionally, while inaccurate predictions for White people tend to predict lower recidivism rates, Black people are more often incorrectly predicted to re-offend (Dressel citation). This means that the COMPAS algorithm is more likely to promote more and longer sentences for Black individuals than for White.

One of the most important factors is that fairness of application can be improved with open source. COMPAS, along with other closed algorithms, are privately owned and therefore proprietary information. Because of this, their use and improvements can be very expensive. While maintenance of an open-source algorithm does accrue expenses, its use overall is very affordable, especially in comparison. This can be valuable as it provides the opportunity for under-resourced court systems to benefit from its use, where the more costly option would be more difficult to access. The necessary training for its use is also less time-intensive and costly than possible training in decision-making for humans. This means that the algorithms can be used widely and normalize how punishments are decided. Because every case would adhere to the same parameters, this could reduce the risk of inconsistent and unreliable application, making it more fair overall.

Waggoner and Macmillen ran a series of tests to support their hypothesis that the open-source algorithms ridge regression, LASSO regression, and elastic net regression are more accurate overall and less biased than the COMPAS algorithm. The study used 7214 real criminal records from Broward County Florida. These records included information such as gender, age, home zip code, prior offenses, and information on if the individual reoffended within 2 years. Waggoner and Macmillen replicated how COMPAS would predict recidivism for each offender using the COMPAS prediction feature. They compared the replicated predictions to the observed outcomes – if recidivism actually occurred. This was used to analyze the accuracy, the baseline accuracy was found to be 65.4%. This was used as the baseline to compare each open-source algorithm. Additionally, by addressing the inaccuracies in prediction trends for various racial groups, bias was assessed.

Each open-source model was fit to a penalized regression algorithm (shown in Figure 1).

Notably, each algorithm has a penalty parameter, $\lambda$ to adjust or shrink coefficient estimates to improve model performance. For each algorithm, to find $\lambda$, a K-fold cross-validation, with K of 10, was conducted. A partition of 80% training data and 20% testing data was used. The training subset was fit to a base model, each testing subset was used to predict the outcomes with various $\lambda$ values. The $\lambda$ value that was on standard deviation greater than the minimum $\lambda$ was selected for each algorithm.

Ridge Regression (equation 1) uses $\lambda$ to tune the $\ell_2$-norm penalty. This penalty shrinks the coefficient estimates to very small but does not drop them completely.

Lasso regression (equation 2) uses $\ell_1$-norm penalty, tuned by $\lambda$. $\ell_1$, in contrast to $\ell_2$, penalizes the absolute value of the coefficient estimates. This results in redundant features within the specification being dropped.

Elastic net regression (equation 3) uses both $\ell_1$ (Lasso) and $\ell_2$ (Ridge) penalties, in this case blending the two using a tuning parameter, $\alpha$. Selecting for this parameter is done following the same 10-fold cross-validation methods used for $\lambda$, this time predicting for various $\lambda$ and $\alpha$. The $\lambda$ and $\alpha$ selected were those that jointly minimized the mean squared error".

The mean predictive accuracy across 1,000 iterations was then measured for each model. Each of the open-source algorithms performed slightly better than COMPAS. While COMPAS had an accuracy of _____, LASSO was 67.2%, Ridge was 67.1%, and Elastic Net was 67.4%.

Confusion matrices were also generated for each model in order to observe true positive, true negative, false positive, and false negative rates. Stronger performance was indicated by higher true positive and true negative results.

ROC curves were plotted for each model to visualize the trade-off between true and false positive rates. For all of the open-source models, the area under the curve was between 0.73-0.74, meaning each algorithm had reasonable predictive accuracy.

Lastly, each model was examined to investigate which features contributed the most to predicting recidivism rates. While ridge regression retained all features, LASSO did not include certain features like charge degree and juvenile misdemeanor counts. Additionally, elastic net " " "combined the L1 and L2 penalties to balance feature selection to retain only the most predictive features while minimizing model complexity.

In conclusion, Pursuing Open-Source Development of Predictive Algorithms: The Case of Criminal Sentencing Algorithms demonstrates that open-source penalized regression models, specifically ridge, LASSO, and elastic net regression, have many advantages over proprietary "black-box" algorithms like COMPAS. This is especially the case in high-stakes scenarios such as criminal sentencing. Through improved transparency, lower costs, increased accuracy, and reduced bias, open-source algorithms allow for more fair and reliable decision-making. By comparing each algorithm and its predictions on real-life data the study confirms that open-source methods are often as good as or better than "black box" algorithms in accuracy and ethical transparency. These findings support adopting open-source algorithms for predictive purposes especially in contexts where unbiased results are essential such as in our justice system.

Equation 1, Ridge Regression:

$$\hat{\beta}_\lambda^{ridge} = argmin\{\sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j X_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2\}.$$

Equation 2, LASSO Regression:

$$\hat{\beta}_\lambda^{LASSO} = argmin\{\sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j X_{ij} \right)^2 + \lambda \sum_{i=1}^{p} |\beta_j|\}.$$

Equation 3, Elastic Net Regression:

$$\hat{\beta}_\lambda^{EN} = argmin\{\sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=1}^{p} \beta_j^2\}.$$

Figure 1.