

# Predicting Recidivism with Machine Learning: An Analysis of Risk Factors and Proposal of Preventions

Vienna Li<sup>1</sup>, Srinitha Sridharan<sup>2</sup>, Sandeep Sethuraman<sup>3</sup> and Georgios Avdis<sup>#</sup>

<sup>1</sup>Ridge High School, USA

<sup>2</sup>Quarry Lane High School, USA

<sup>3</sup>BASIS Chandler

<sup>#</sup>Advisor

## ABSTRACT

Despite efforts to support the re-entry of prisoners into society, a significant proportion of released offenders eventually return to crime. To identify areas for improvement within correctional facilities, researchers are directing their focus towards recidivism, the tendency of an offender to recommit a crime. Although past studies have identified factors that are correlated to recidivism, there is still uncertainty about the most significant combinations of factors that drive it. Due to the complexity of this issue, the goal of our project is to create a machine-learning model to predict whether an individual will relapse into crime. Such a model will help experts study the effectiveness of specific forms of punishment and develop personalized correctional programs to target individuals based on their recidivism risk factors. We applied the Decision Tree, Random Forest, and Gradient Boosted Decision Tree algorithms to classify a prisoner as likely to recidivate or not, tuning the hyperparameters to optimize accuracy. To evaluate our hypotheses, we analyzed the top nodes of our trees and confirmed several of our initial predictions. Furthermore, we found that an individual's relative placement in their community, such as the percentage of individuals in their community with lower or equal education levels, was a significant predictor of recidivism. The results of this research may help law enforcement officers make more informed decisions about how to allocate their resources based on predictions of recidivism.

## Introduction

With the U.S. housing the largest incarcerated population globally (Scommegna, 2012), it has become increasingly important to develop effective intervention methods to address the growth in crime. Recidivism, the tendency of a prisoner to return to crime after release, has posed a significant threat to correctional facilities that are unable to accommodate the influx of prisoners. As 68% of prisoners are rearrested within three years of release (National Institute of Justice, 2008), researchers have recognized the need to develop a new framework for parole policies that will directly target the causes of recidivism.

Although previous studies identified major risk factors for recidivism (Grieger & Hosser, 2014) and (Håkansson & Berglund, 2012), the lack of consensus on the most significant factors has created a knowledge gap, preventing policymakers from effectively reforming the parole system. However, taking a data-driven approach to criminology has the potential to fill this gap and facilitate the development of innovative approaches to minimize recidivism rates.

To prevent the depletion of prison resources, law enforcement officers must be able to allocate services effectively so that individuals at a high risk of recidivating receive sufficient support upon reentry into society. The ability to predict whether a prisoner will recidivate can provide crucial insights into the resource distribution

process. However, it is not an easily obtainable ability: a lack of publicly available data due to privacy laws, coupled with factors affecting different demographic groups to different extents, makes it very difficult to make meaningful predictions. As the volume of prison data grows, researchers have begun to explore the utility of machine learning in addressing this issue. Notably, recent findings have pointed to the conclusion that AI is less susceptible to bias and more accurate than humans in predicting recidivism (Nader, 2021). However, bias can be introduced into AI models if the training dataset is not representative of the prison population.

The purpose of this project is to develop and test different machine learning models to predict recidivism and determine its most significant causes. In addition to analyzing the known causes, we explored how a prisoner's similarity to their community affects their likelihood to recidivate. We compared the performance of three supervised machine learning models in predicting recidivism within three years, the target variable in our dataset. Our models use classification to make a binary prediction of whether or not an individual is likely to recidivate, and we specifically chose models that operate entirely on boolean logic to make the predictions easily interpretable. Translationally, the results of this research are intended to help law enforcement officers design parole services that eliminate the incentive for high-risk prisoners to recidivate.

## Literature Review

Recidivism is a major problem in modern society, as it has contributed to an increased public safety threat due to the reintroduction of criminals, who go on to commit more crimes. By no means is this a new problem. Even in the mid-to-late 20th century, significant proportions of released inmates returned to crime, with nearly 70% being arrested after their release, and almost 50% seeing more jail time (Visher & Travis, 2003). These large figures contribute to a huge societal and financial cost, especially in cases involving serious crimes such as murder or sex offenses (Quinsey et al., 1993). On average, it costs \$36,000 to house one inmate for a single year, and when this figure is extrapolated to the hundreds of thousands of inmates in the United States, the financial burden of recidivism becomes astronomical (Bureau of Prisons, Justice, 2018).

While being a clear detriment to society as a whole, recidivism also ensnares released inmates in a cycle of crime that they do not have the resources to escape. Inmates who feel abandoned by the society around them begin to believe that there are no opportunities for them in the world outside of crime. Not being visited by friends and family who alienate them, among other forms of social stigma, can contribute greatly to this phenomenon (Bales & Mears, 2008). Isolation can contribute to mental health issues among prisoners which in turn can contribute to a loss of hope for redemption from crime, perpetuating a cycle of prison time and isolation (Solbakken & Wynn, 2022). This effect of rejection by society can be particularly detrimental for female prisoners when it combines with a weak support system and already existing feelings of judgment (Korzh, 2022).

Every year, more than 600,000 individuals transition from correctional institutions to communities (Pogorzelski & Wolff 2005). These individuals face numerous disadvantages upon their arrival, including drug and alcohol addiction, mental illness, outdated job skills, and limited education. Their criminal records hinder their ability to secure housing, treatment services, and employment. Successful reintegration into society can greatly improve the lives of inmates outside of prison, as shown by Koschmann and Peterson (2013) who advocated for mentorship programs to support released prisoners as they adjust to society. In their case study, they found that such a policy would foster a conversational environment for parolees and provide them with valuable skills. However, because these policies are not yet widespread, recidivism remains prevalent in the majority of prisons across the United States.

To address this issue, the federal government established programs such as The Second Chance Act, which advocates for greater support for people after release. From public assistance, housing, and mental health services to education and job training, the programs aim to help prior offenders effectively integrate into society. With the support of the criminal justice, public health, and social service systems, the legislation draws attention to the necessity of such programs for prisoners, which would help the prisoners foster healthy relationships with

their families and children. However, the many policy restrictions and their non-explicit caveats have limited the government's attempts to address the issue. Hence, adjustments to existing policies must be made to facilitate community reintegration.

As awareness of the negative effects of recidivism has spread, researchers have investigated which intervention methods have the greatest potential to decrease recidivism. Policy measures include three primary classification groups: general deterrence, incapacitation, and rehabilitation. General deterrence is a method that employs media and attention to dissuade the public from criminal conduct. Unfortunately, research has shown that deterrence is a weak form of punishment because perpetrators under the influence are unable to think rationally and weigh the risks versus the benefits. Incapacitation is a crime reduction method in which criminals are physically locked up, making it impossible for the perpetrator to commit further crimes. Although incapacitation is effective in lowering the crime rate, it is very costly to operate and maintain prisons. Increasing probation time and austere regulations in prison are counterproductive as violent environments result in lasting mental effects on the inmates, which may result in recidivism once they are released (Tobón, 2022). The final and arguably the most effective method in decreasing recidivism rates is rehabilitation. Providing facilities to help people recover and immerse back into society, although more successful than the other two methods, is not a panacea in addressing recidivism. A large number of absences, lack of funding, and efficient assessment of effectiveness are all aspects of rehabilitation programs that the government should focus on.

Government officials and private companies are also looking for algorithms to predict and prevent recidivism. This process of quantification in the realm of recidivism and social issues as a whole has become possible, largely because of the introduction of big data (Mattiuzzo, 2019). An existing software known as COMPAS, capable of determining the risk of recidivism on a scale from 1-10, has been applied in many prisons across the United States. While algorithms such as COMPAS have in some part facilitated large-scale classification of recidivism likelihoods, racial bias has been observed in the algorithm, as the false positive rates for black prisoners are higher than those of white prisoners (Dressel & Farid, 2018). In another case, an investigation by Angwin et al. (2016) showed that the widely used COMPAS model was twice as likely to misclassify Black inmates as likely to recidivate. Although the validity of this argument has been disputed, it highlights the ethical considerations involved with the use of machine learning in this field. It is important to use datasets that have a diverse representation of the incarcerated population to make accurate and fair predictions.

Researchers have also explored the possibilities associated with using Machine Learning models to enhance the effectiveness of recidivism predictions to improve fairness in the algorithmic approach to recidivism. This change in methodology is supported by Nader (2021), who suggested that AI models are less prone to bias in predicting recidivism than human-based models are. The extent to which machine learning algorithms have a higher performance than conventional risk assessment tools, such as the logistic regression tool, has been evaluated in past literature. Duwe and Kim (2015) found that although newer models produced more accurate results, the difference between older and newer algorithms was not significant. Their research suggests that several machine learning models should be taken into consideration when developing risk assessment measures. This finding was corroborated in a systematic review of research on different Machine Learning models and their applicability across different datasets and circumstances (Travaini et al., 2022).

There is an abundance of literature centered on developing recidivism prediction models using machine learning, and much of this research consists of an examination of the applicability of a single or couple of models in recidivism prediction. While there has been published research comparing different models, these studies have generally compared Machine Learning models to more traditional prediction models (Tollenaar & van der Heijden, 2013). However, there is a gap in this literature: little to no investigation has been made into the effect of an individual's status in their community on their likelihood of recidivism. With factors such as familial rejection and isolation being shown to significantly increase the risk of recidivism, it follows that an individual's characteristics relative to the characteristics of those in their immediate community would also

have a significant effect on recidivism. In our research, we intend to bridge this gap by defining quantitative metrics, based on which, we can test this theory.

## Hypotheses

To turn our theory on the effect of an individual's community status into a testable hypothesis, we decided to compare the accuracies of a model including features about the individual's community status, and a model without such features. Each one of these features was the numerical form of a percentage of persons in an individual's Public Use Micro Area (PUMA), a subsection of Georgia analogous to a community, with either a lower or an equal characteristic to the individual. We propose H1: Individuals who have a higher social status than larger percentages of people within their PUMA Residences will recidivate at lower rates than other individuals.

Because we included multiple individual-level characteristics in all of our models, we were able to test the effect of an individual's age on their recidivism prediction. Prior research has shown that older individuals tend to have lower sexual recidivism rates (Fazel et al., 2006), but there is a gap when it comes to generalizing this discovery to recidivism as a whole. To fill this, we propose H2: An individual's age will negatively correlate with their likelihood of recidivating for any offense.

Since financial stability and employment opportunities can contribute to a lower likelihood of re-offending (Kassem, 2017), we also hypothesized that individuals who are employed for a higher number of days would have success reintegrating into society. In support of this hypothesis, a research article by Stephen et al. (2009) titled "Is Employment Associated With Reduced Recidivism?: The Complex Relationship Between Employment and Crime," found evidence suggesting that employment significantly reduces the risk of recidivism among offenders. The study analyzed data from multiple studies and concluded that individuals who secure and maintain employment following their release from correctional facilities have a decreased probability of returning to criminal activities. Although this research article supports our belief that longer periods of employment can contribute to reducing recidivism, we were curious whether this was also the case in a smaller time frame (within 3 years). H3: A high number of days employed is an indicator of greater job stability, which may reduce the likelihood of reoffending.

## Methodology

### Data Collection

To evaluate each of our models on equal grounds, we used a public dataset from the National Institute of Justice, containing information on inmates in Georgia in 2013. This dataset had 54 features, including the 3-year recidivism outcomes, criminal history, offense severity, supervision risk score, educational and employment statuses, and PUMA Residence of the individual, with 25,835 rows of data. We chose this dataset because it contained features we had hypothesized would affect recidivism rates, allowing us enough data to test our hypotheses.

For one version of the models, we combined this NIJ data with a dataset from the US Census Bureau's American Community Service (ACS). This dataset was a Public-Use Microdata Sample (PUMS), consisting of both personal characteristics and societal statuses of residents in all PUMA Residences of Georgia. This larger dataset with 283 features and 486,232 instances included many of the same features that the NIJ dataset did but wasn't limited to prisoners. One unique feature of this dataset was the "PWGTP" column which measured the relative "weight" of an individual in the dataset, or the amount of times we needed to duplicate each individual's entries to construct an accurate representation of the population of Georgia. We opted to add this dataset because

it allowed us to test our hypothesis on the effect of an individual's "community status" on their recidivism outcome.

## Data Preprocessing

To convert our data into a format comprehensible by the machine learning model, we performed several data preprocessing steps including data cleaning, data transformation, and feature selection. In both datasets, all columns were converted to integers or floats, and categorical columns were assigned numerical values for each level. For example, in the education level column, 'less than high school diploma' would correspond to 0, 'high school diploma' would correspond to 1, and 'At least some college' would correspond to 2. Columns with numerical ranges were converted to the lower bound or the average of the lower and upper bound in the range. For columns with less than 7% missing values, the null values were imputed with the median. However, for *Gang\_Affiliated* and *DrugTests\_THC\_Positive*, which had over 13% missing values, the null values were imputed with the distribution of the non-null values. For example, if 80% of the non-null values for gang affiliation were 1, then 80% of the null values would be assigned 1.

For our combined models, we constructed a new category of features to be used in predictions of recidivism. This category, which we chose to call "community status," consists of features that represent an individual's relative standing in their community, or in our case, their PUMA Residence. For example, the community status feature for education would be the percentage of people in an individual's PUMA Residence that had a lower or equal level of education to them. The creation of this category is significant because it can provide more context into features such as an individual's income, by comparing it with other individuals in the same community.

To include these additional features in our model, we used the Pandas library in Python to transform our PUMS dataset. After we had duplicated all of the entries in our dataset using the PWGTP feature, we grouped the dataset by PUMA Residence and combined all of the entries in each PUMA Residence into one that represented the distribution of all entries in the same manner. For numerical variables, the mean or median values were used, and for categorical or boolean variables, we created a feature containing the percentage of individuals in each category. This resulted in a dataset with 25 rows (representing each of the 25 PUMA Residences within Georgia) and multiple extra columns.

To combine this new dataset with the NIJ dataset, we added the features to each entry based on the PUMA Residence, and then implemented a Python code to combine the different percentages one variable for each feature that contained the percentage of individuals with a lower or equal value for each variable when compared to the prisoner. This allowed us to refine our method of predicting based on an individual's societal status, as this societal status was now based on concrete data rather than common belief.

## Data Analysis

To better understand the dataset, we employed multiple statistical methods to gain a more cohesive picture of the topic. One trend in the data showed that the percentage of males returning to recidivism within 3 years was about 4 times as much as women, which could not be explained by highly correlated factors such as Supervision Risk Scores, Drug Use, or Amount of Dependents. To understand why males were more likely to become recidivists, we looked at the distribution of the data and noticed that it was highly skewed towards men. We also found a positive correlation between the Supervision Risk Score, the Percent Days Employed as well as Jobs per Year. As Supervision Risk Score increased, Percent of Days Employed initially increased until it hit a maximum at a risk score of around 3.0 and then steadily decreased, while Jobs Per Year increased at a decreasing rate until it hit an asymptote of around 0.8 jobs per year. This trend suggests one of two things: either the

supervision risk score is a fairly effective method at predicting the likelihood of recidivism, or that the supervision risk score might cause some prisoners to become recidivists. Ultimately, these findings helped us determine the most significant factors of recidivism to include in our final dataset.

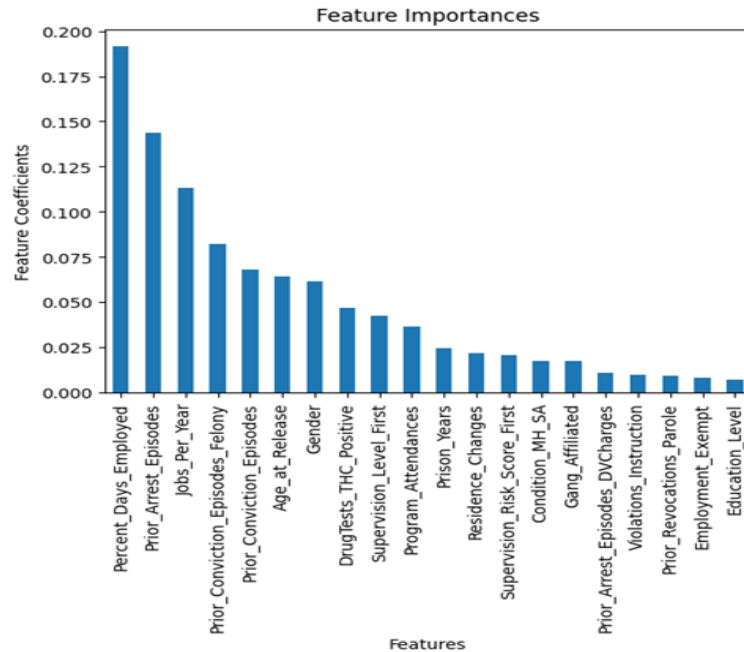
Next, we used a factor analysis technique to identify underlying structures in our dataset. The technique involves analyzing the relationships among a set of variables to determine which factors are most important in identifying the patterns and models within the data. In the context of predicting recidivism, factor analysis helps researchers identify and understand the independent factors or variables that contribute to the likelihood of reoffending. By splitting the data into two categories - one with the target variable (recidivism within three years) and one without - factor analysis uncovers new independent factors that describe the relationships and patterns among the original dependent variables. This method allows for a comprehensive exploration of the complex topic of recidivism and provides valuable insights into the factors that may contribute to the likelihood of reoffending.

To further explore our dataset, we proceeded to use several graphs, charts, and other visual models to portray complex patterns within the dataset. By doing this, we were able to discover which factors were most closely correlated with each other and which factors had clear effects on the recidivism rate in the dataset. This, in turn, allowed us to develop more accurate predictive models and make informed decisions on which factors to include in our model.

## **Machine Learning Model**

### **Feature Selection**

Because we intended to compare the effect of including different features in our prediction, we created three model versions with different feature selections. To create our first two model categories (Model A and Model B), we only used data from the NIJ dataset. To avoid overfitting, we eliminated features that were unrelated to our prediction (e.g. the individuals' Serial Number) and consolidated features that were separated into several different categories into a single column (e.g. combining certain prior arrest columns into one column with the total number of prior arrests). Then, to further increase the robustness of our model, we determined the correlations of each feature to the prediction variable and eliminated the features with correlations close to 0. Figure 1 was particularly useful in visualizing the disparities in the correlations across different features, allowing us to narrow down the columns.



**Figure 1.** Feature Importance. This graph displays the correlations of various dataset features to the target variable, recidivism within three years.

In our third model category (Model C), we augmented our data with the PUMS dataset, so our feature selection process was more complex. We included the most important features from the NIJ dataset, combining them with the community status variables with the highest correlation to recidivism rates. To minimize the bias in our data, we decided not to create community status variables for immutable characteristics, such as race. The inclusion of such a feature could have contributed to bias, as we would have been required to designate certain races as higher status than others, which is ethically problematic. However, we did decide to include the immutable characteristic of an individual’s disability status, as it is known that having a disability makes it more difficult to gain employment and other such resources. In this feature, for individuals with disabilities, the percentage of people within their PUMA Residence who were also classified as disabled was recorded. By including this factor, we were able to determine whether individuals classified as disabled were more susceptible to recidivism in communities where they felt alienation, which was an important insight to make.

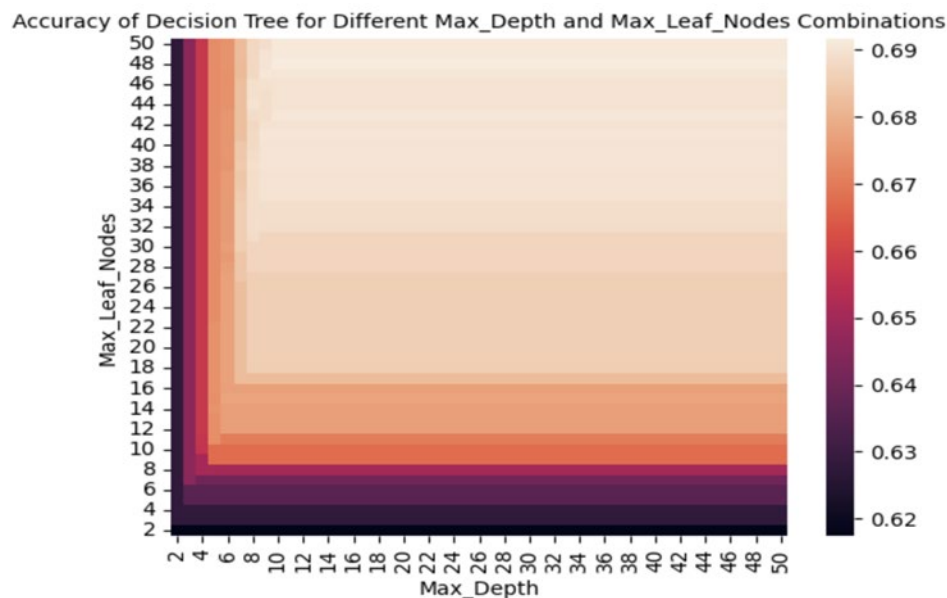
The mutable characteristics that we included were the ones that had clear counterparts in the NIJ dataset, which was essential; without an individual’s personal trait, we wouldn’t be able to make a comparison between the individual and the community centering on the trait. As a result, we were able to make comparisons on features including employment status, educational level, disability status, government aid status, and marital status. Including these features in our model was vital to the significance of our findings, as this metric of community status has not been included in any prior research to predict recidivism.

### Decision Tree Model

The first machine learning model employed was the Decision Tree algorithm, which uses the features from the data to create a classification model. The Decision Tree classifier from the Python scikit-learn library recursively divides the data into nodes based on patterns in the dataset. At each decision node, the tuple of data points will encounter a condition and follow the corresponding branch to the next decision node. The leaf node is used to classify a data point, allowing the model to make predictions.

Scikit-learn uses the Classification and Regression Tree (CART) algorithm, which constructs a binary tree that optimizes the information gain at the nodes. The CART algorithm essentially determines the best-split point for each input. Based on the split points, the model identifies new “best” split points. This process of splitting continues until the data point reaches the leaf nodes. Because our target variable, recidivism within three years, is a binary value, we employed a classification tree and the Gini index metric, which concludes the probability of a variable being incorrectly classified and a Gini coefficient. A Gini index of 0 indicates that all the elements are skewed towards a certain class, a Gini index of 1 depicts that all elements are randomly sorted into classes, and a Gini index of 0.5 signifies that all the elements are equally distributed across the classes.

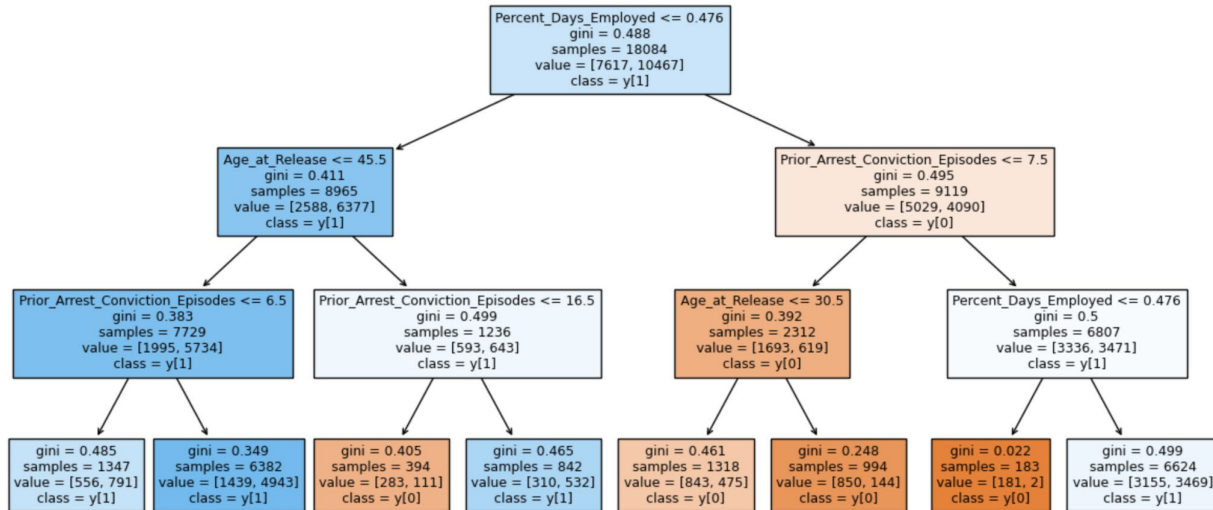
To create the model, we first split our data into a training, validation, and testing set. Using different visualization techniques, we concluded that the optimal training, validation, and testing split was 0.70-0.15-0.15. To optimize model accuracy, we used Randomized Search algorithm to determine the best hyperparameters for the tree. This algorithm selects random combinations of hyperparameters within a defined space and outputs the values producing a model with the optimal score. To further tune the hyperparameters, we graphed a heatmap (Figure 2) to determine the optimal combination of max\_depth and max\_leaf\_nodes.



**Figure 2.** Accuracy by max\_depth and max\_leaf\_nodes. This heatmap shows the Decision Tree accuracy of Model B for various combinations of the two hyperparameters.

The decision tree (Figure 3) shows the top decision nodes of a model constructed with a maximum depth of 14 and maximum leaf nodes of 45, producing an accuracy of 70.3%. The root node of the decision tree is *Percent Days Employed*, which indicates that employment was a significant condition in splitting the data. Using the correlation matrix, we noticed that the features in the top nodes had the highest correlations to the target variable.





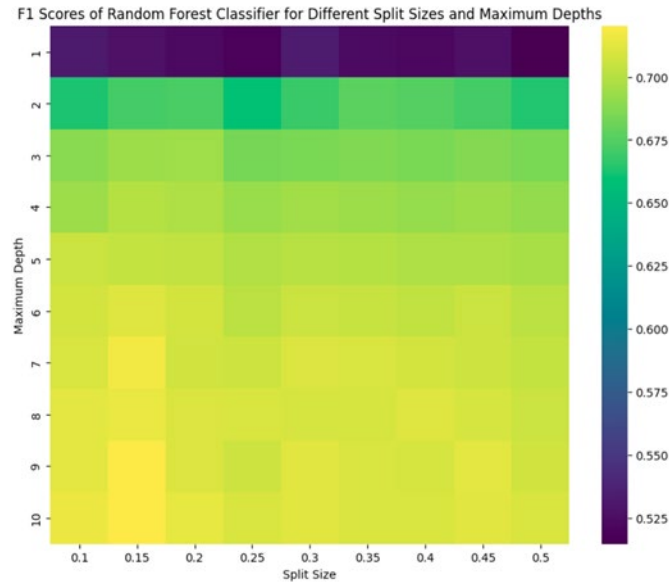
**Figure 3.** Decision Tree. This visualization depicts a decision tree (Model B) with a max depth of 3 and a root node of *Percent Days Employed*.

Although the decision tree performed fairly well, there were several assumptions that we made about the dataset. We first assumed that the features selected were appropriately correlated with the target variable, recidivism within three years, and that the data itself was collected ethically and accurately. We are also assuming that the dataset is a good representation of the incarcerated population in Georgia and that the outcomes the model predicts are unbiased. However, since decision trees are very sensitive to the data, we also constructed a Random Forest model to attain the optimal results.

### Random Forest Model

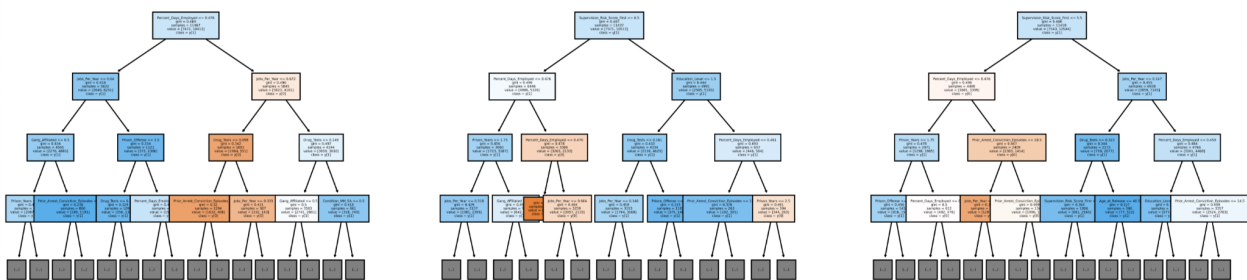
The random forest classifier is an ensemble model that makes a prediction based on the most common output from multiple decision trees. As individual decision trees have high variance and sensitivity to the data, random forest can substantially improve the generalizability of the algorithm by combining several uncorrelated decision trees. Random forest uses bootstrap aggregation to create samples of the dataset with subsets of the original features and randomly selected rows. Each decision tree is trained independently on these samples, and their outcomes are aggregated to make a final prediction. We chose to apply the random forest model to our prediction problem due to its ability to mitigate overfitting and minimize bias.

After building random forest classifiers with the scikit-learn library using a similar approach to the decision tree, we evaluated model performance using two key metrics: accuracy and F1 score. Since the F1 score gives equal weightage to precision and recall, we found it useful in understanding how the model predictions compared to reality. To optimize the accuracy of the model, we tuned the hyperparameters using the Randomized Search, Grid Search, and Bayesian Search algorithms. While randomized search attempts random combinations of hyperparameters within a parameter space, grid search attempts every possible combination and can be exhaustive, but it is more likely to identify the most favorable hyperparameters. Bayesian search improves the efficiency of randomized search by further exploring regions of the parameter space that maximize model performance. Among the three searches, grid search produced the highest accuracy while randomized search resulted in a slightly lower accuracy with the fastest runtime. Another method we used to identify the optimal combination of Maximum Depth and Split Size was to graph a heatmap of the F1 scores (Figure 4).



**Figure 4.** F1 Scores by Maximum Depth and Split Size. This heatmap shows the effect of Maximum Depth and Split Size on the F1 Scores of our Random Forest Classifier for Model C. A maximum depth of 9 and Test/Validation Split Size of 0.15 were determined as the optimal hyperparameters. A similar visualization was created for accuracy.

Similarly to the decision tree, we plotted the first three trees in the random forest model to analyze the topmost nodes and determine which features the model considered important (Figure 5). Although *Percent Days Employed* was one of the root nodes and had a high correlation to the target variable, *Supervision Risk Score First* had a comparatively smaller correlation than other features included in the model. This is likely because random forest constructs uncorrelated trees that are based on samples of the data rather than its entirety.



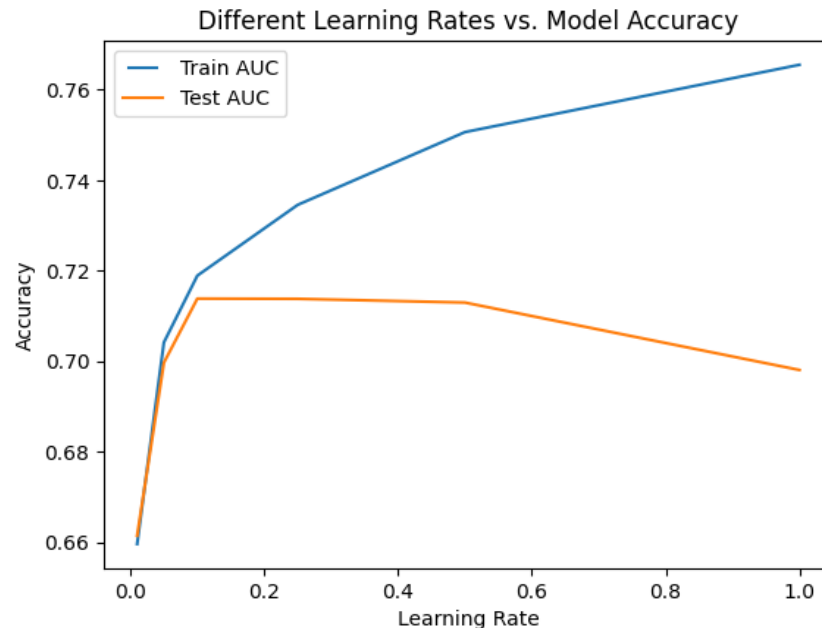
**Figure 5.** Random Forest. This plot shows the first three trees in a random forest model (Model C) with tuned hyperparameters. The root nodes are ‘Percent\_Days\_Employed’, ‘Supervision\_Risk\_Score\_First’, and ‘Supervision\_Risk\_Score\_First’.

### Gradient Boosted Decision Tree Model (GBDT)

Boosting algorithms, such as GBDT, leverage the strengths of sequential weak learners. Weak learners are singular decision trees that attempt to minimize the errors made by previous trees in the boosting process. By combining multiple decision trees in series, where each subsequent tree focuses on the errors made by the

previous trees, the boosting algorithm creates a strong learner that is highly efficient and accurate in predicting recidivism.

Unlike random forests, GBDT does not employ bootstrap sampling but instead, fits on a modified version of the dataset every time a new decision tree is added. The sequential nature of boosting algorithms allows for a slow learning rate, which can be advantageous in improving the performance accuracy of the model, continually adjusting and improving predictions throughout the course of its runtime. This is especially important in recidivism prediction, in which sophisticated patterns and factors influence an individual's likelihood of reoffending. To improve the accuracy of the GBDT, we tuned the hyperparameters using randomized search. We also employed the Area Under the Curve (AUC) metric to determine a better model.



**Figure 6.** Gradient Boosted Decision Tree. This plot shows the learning rate vs. the accuracy of the GBDT for Model A. The optimal learning rate is 0.15.

## Results

Ultimately, we created three machine learning models for both the Decision Tree and Random Forest using slightly different variations of features for the input data. Two versions of the model were trained solely on the NIJ dataset (Model A and Model B) while the third version was fitted to an augmented dataset using the PUMS and NIJ datasets (Model C). We compared the performances between the Decision Tree, Random Forest, and Gradient Boosted classifiers, and evaluated the extent to which varying the features affected these models.

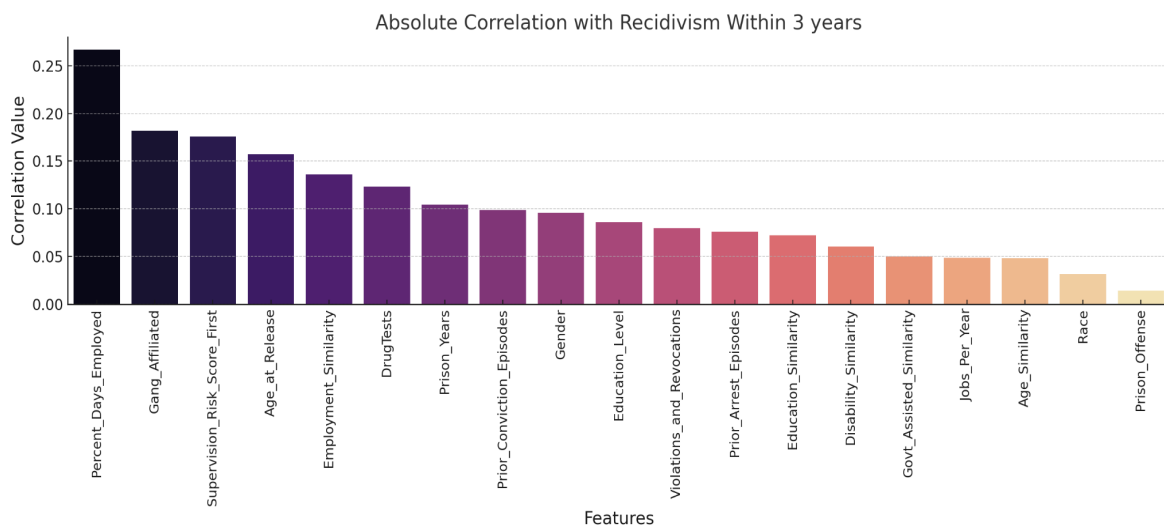
Model A used a total of 20 features whereas Model B was trained on a subset of 10 features out of these 20. Although having a larger number of features can uncover more patterns within the dataset, a model with fewer features may be less susceptible to noise in the data and have greater generalizability. The Decision Tree for Model A had a validation accuracy of 67.4% and a testing accuracy of 67.5%. The Random Forest produced a validation accuracy of 72.0% and a testing accuracy of 73.0%. Comparatively, the Decision Tree for Model B produced a validation accuracy of 71.5% and a testing accuracy of 70.6% after hyperparameter tuning. Random Forest produced a validation accuracy of 73.0% and a testing accuracy of 71.6%, which was a slight improvement from the Decision Tree.

Model C differed from Model A and Model B because it included the “community status” characteristics, and with this model, we used the Decision Tree and Random forest to determine whether the inclusion of community status as a feature would improve the accuracy of our model. We found that with this model, the Decision Tree accuracy had no significant change with a validation accuracy of 70.0% and a testing accuracy of 71.3%. However, with the Random Forest, when we used hyper-tuned parameters, Model C produced a validation accuracy of 81.3% and a testing accuracy of 80.6%, which was a fairly significant increase over our previous models, indicating that factoring in community status features can provide useful context to make more informed predictions of recidivism.

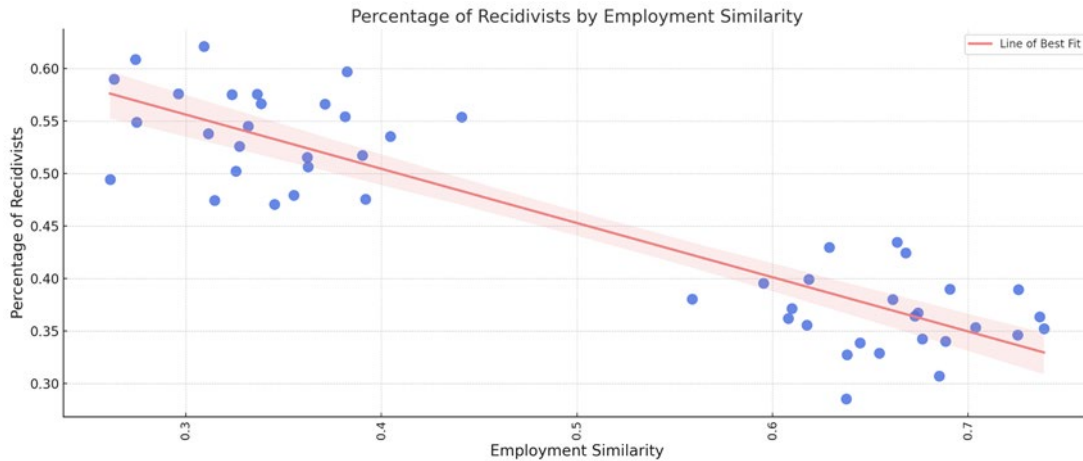
For Model A, we also employed the Gradient Boosted Tree algorithm which had a validation accuracy of 83.0% and a testing accuracy of 84.2%. Because the boosted tree used sequential learning and adjusted its predictions throughout the process, we concluded that the Gradient Boosted Tree algorithm had the highest accuracy rate, and thus was the best out of the three models.

## Conclusion

Due to the complexity of factors associated with recidivism, the purpose of our research was to develop a machine learning approach to predict recidivism and compare the performances of different algorithms to determine the most accurate and unbiased solution. Based on our results from all three models, we were able to conclusively test all four of our hypotheses. To test H1, we analyzed the recidivism rate based on each of the community status features in the dataset used in Model C. We did this in two ways: firstly, by plotting all of the feature importances and evaluating the implications for the utility of the community status as a whole (Figure 7), and secondly by plotting the recidivism rates based on the percentages for each community status feature (Figure 8).



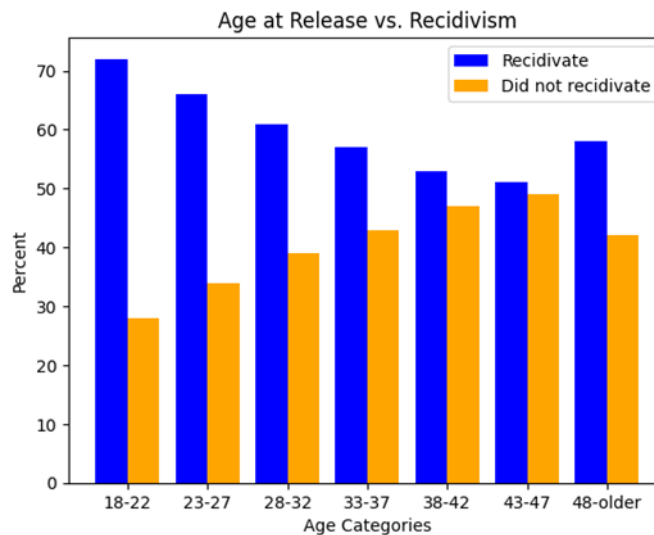
**Figure 7.** Absolute Correlation with Recidivism Within Three Years. This graph includes both negative and positive correlations of features to the target variable.



**Figure 8.** Percentage of Recidivists by Employment Similarity. This scatterplot plots the percentage of recidivists in groups of individuals with certain Employment Similarities and generates a line of best fit for the data, showing a clear negative correlation.

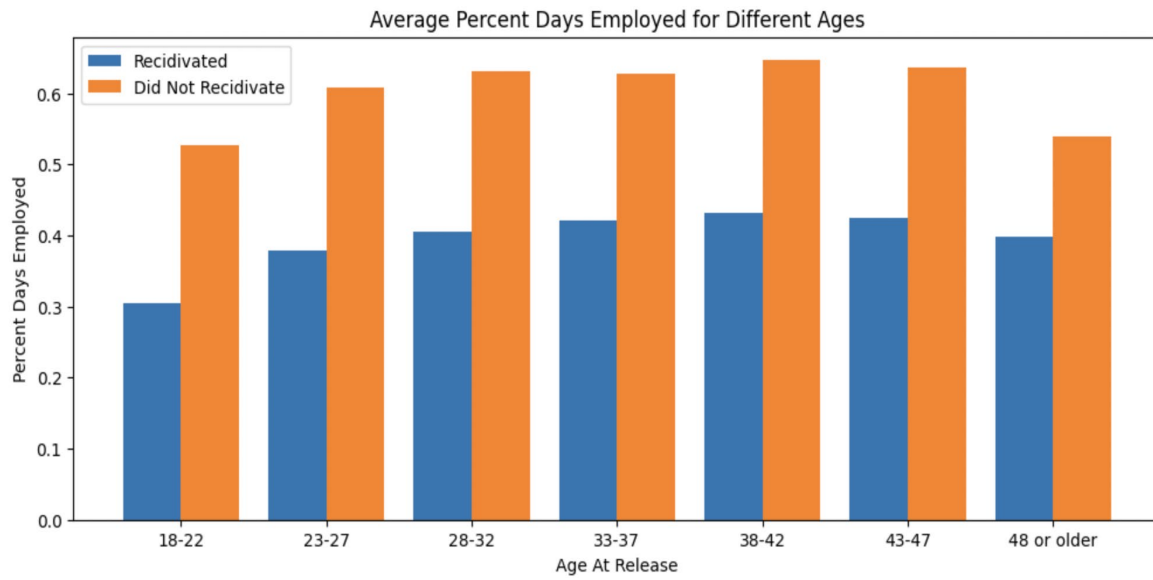
Figure 8 confirms H1 because the line of best fit shows a statistically significant negative correlation between the community status feature for employment, which was the community status feature with the most importance, and the recidivism rate. The first graph supports our theory that including the recidivism rate can better our predictive models, because the community status features for employment and education similarity have significant importance, indicating that they should be included in future research into Machine Learning predictions on recidivism.

With regards to H2, we found significant negative correlations between the age at release for an inmate and the recidivism rate within 3 years in each of our Models. Additionally, in each Model, the age at release feature was consistently either the root node or near the top of the decision trees. This applied to the Decision Tree, Random Forest, and Gradient-Boosted Decision Tree.



**Figure 9.** Age at Release vs. Recidivism. This grouped bar chart shows the percentage of prisoners who recidivated for various age categories.

Figure 9 indicates that the recidivism rate decreases as the age increases, validating H2. However, an interesting phenomenon occurs at ages above 48, where the recidivism rate increases. Our proposed explanation for this is that this category encompasses all inmates above the age of 48, including retired inmates, which could contribute to their financial instability. This increased financial instability could contribute to motivation for them to turn to crime. To support our theory, we created the visualization below.



**Figure 10.** Average Percent Days Employed for Different Ages. This graph displays the mean percentage of days prisoners were employed in a year across different age groups. The bar colors correspond to whether or not the prisoner recidivated.

Because the 48 or older category in Figure 10 breaks the trend of increasing employment percentages, this supports our theory that the sudden spike in recidivism rates could be caused by a lack of usable income. However, there are clear limitations to our theory, as we didn't account for the reception of government assistance and several other nuances. Because of this, further research is required to determine whether this phenomenon was unique to our dataset, or if it is a widespread effect. Furthermore, Figure 9 shows that prisoners who did not recidivate had a higher average percentage of days employed across all age groups, supporting H3.

The results of this project have broader implications for the development of future data-driven solutions to address recidivism and the enhancement of parole policies. By analyzing the top nodes of the trees and feature correlations, we identified several major risk factors for recidivism, one of the most significant being the percentage of days employed. We propose that parole services be targeted toward securing jobs for newly released prisoners to promote long-term financial stability for parolees. Additionally, we recommend prioritizing younger parolees, as our visualizations indicate a trend of younger prisoners recidivating more frequently. Furthermore, the GBDT appears to be the most useful for creating a model with high accuracy. The significantly higher performance of Model C than the other model versions points to the conclusion that education and employment similarity are important predictors of recidivism and should be further evaluated in future studies.

## Limitations

With the significant implications of model predictions in mind, it is important to acknowledge the limitations within the datasets they were trained on. Since both datasets were confined to the state of Georgia, the model may not be generalizable to other states and countries. Additionally, certain characteristics of the incarcerated population were not represented in the NIJ dataset, as there were no races other than White and Black, and there were no females who were gang-affiliated. Finally, the data used to train our model was taken from 2013 to 2015, which is a long period in which the causes of recidivism could have evolved. Especially critical is to evaluate whether these same findings hold for prisoners released after the COVID-19 pandemic. Despite these limitations, the identification of new risk factors relating to community similarity and our comparisons of various model performances lay the groundwork for future recidivism research.

## Directions for Future Research

To expand upon our current research, we could use a regression model to predict the percentage likelihood of a prisoner to recidivate rather than a binary prediction of whether or not they will recidivate. Using this method, we could compare our model to risk assessment tools based on scales, such as COMPAS, and leverage techniques to improve accuracy and minimize bias. As an extension to our project, we could select another target variable within the Georgia dataset for prediction, such as supervision risk score, which is related to recidivism in that prisoners considered more likely to recidivate are assigned higher supervision risk scores. Features that may have been overlooked due to their low correlation with recidivism may prove to be more significant in predicting supervision risk scores, allowing for a more robust analysis of the factors that determine parole supervision policies. To account for the limitations we identified within our dataset, we could use a more representative dataset encompassing prisoner information within the U.S. and further explore methods to reduce algorithmic bias.

## Acknowledgments

We would like to thank our Prof. Avdis for guiding us through the research process and making this project possible. We would also like to thank our data science research program for ensuring the quality of our paper and providing valuable feedback.

## References

- Bales, W. D., & Mears, D. P. (2008). Inmate Social Ties and the Transition to Society. *Journal of Research in Crime and Delinquency*. <https://doi.org/10.1177/0022427808317574>
- Bureau of Prisons. (2018). Annual determination of average cost of incarceration. *Federal Register*. <https://www.federalregister.gov/documents/2018/04/30/2018-09062/annual-determination-of-average-cost-of-incarceration>
- Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1). <https://doi.org/10.1126/sciadv.aao5580>

Duwe, G., & Kim, K. (2015). Out with the old and in with the new? an empirical comparison of supervised learning algorithms to predict recidivism. *Criminal Justice Policy Review*, 28(6), 570–600. <https://doi.org/10.1177/0887403415604899>

Fazel, S., Sjöstedt, G., Långström, N., & Grann, M. (2006). Risk Factors for Criminal Recidivism in Older Sexual Offenders. *Sexual Abuse: A Journal of Research and Treatment*, 18(2). <https://doi.org/10.1007/s11194-006-9009-0>.

Grieger, L., & Hosser, D. (2014). Which Risk Factors are Really Predictive?: An Analysis of Andrews and Bonta’s “Central Eight” Risk Factors for Recidivism in German Youth Correctional Facility Inmates. *Criminal Justice and Behavior*, 41(5), 613-634. doi:10.1177/0093854813511432

Håkansson, A., & Berglund, M. (2012). Risk factors for criminal recidivism - a prospective follow-up study in prisoners with substance abuse. *BMC psychiatry*, 12, 111. <https://doi.org/10.1186/1471-244X-12-111>

Kassem, G. L. (2017). *The Effects of Employment on Recidivism Among Delinquent Juveniles* (Master's thesis, East Tennessee State University). East Tennessee State University.

Korzh, A. (2022). ‘You have been punished in prison. And then when you are released, you are punished for life’: Post-incarceration barriers for women in Ukraine. *International Sociology*, 37(3), 373–390. <https://doi.org/10.1177/02685809221084447>

Koschmann, M. A., & Peterson, B. L. (2013a). Rethinking recidivism. *Journal of Applied Social Science*, 7(2), 188–207. <https://doi.org/10.1177/1936724412467021>

Mattiuzzo, M., (2019) Algorithms and Big Data: Considerations on Algorithmic Governance and its Consequences for Antitrust Analysis. *Revista de Economia Contemporanea*, 23 (02). <https://doi.org/10.1590/198055272328>

Moore, K. E., Stuewig, J. B., & Tangney, J. P. (2015). The effect of stigma on criminal offenders’ functioning: A longitudinal mediational model. *Deviant Behavior*, 37(2), 196–218. <https://doi.org/10.1080/01639625.2014.1004035>

Nader, J. (2021, May 19). Serving data justice when predicting recidivism. Data Society. <https://datasociety.com/serving-data-justice-when-predicting-recidivism/>

National Institute of Justice. Measuring recidivism. (2008, February 28). <https://nij.ojp.gov/topics/articles/measuring-recidivism>

Pogorzelski W, Wolff N, Pan KY, Blitz CL. Behavioral health problems, ex-offender reentry policies, and the "Second Chance Act". *Am J Public Health*. 2005 Oct;95(10):1718-24. doi: 10.2105/AJPH.2005.065805. Epub 2005 Aug 30. PMID: 16131635; PMCID: PMC1449426.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Lalumière, M. L. (1993). Assessing Treatment Efficacy in Outcome Studies of Sex Offenders. *Journal of Interpersonal Violence*. <https://doi.org/10.1177/088626093008004006>



- Scommegna P., (2012). U.S. Has World's Highest Population Rate. *Population Reference Bureau*.  
<https://www.prb.org/resources/u-s-has-worlds-highest-incarceration-rate/>
- Solbakken, L. E., & Wynn, R. (2022). Barriers and opportunities to accessing social support in the transition from community to prison: a qualitative interview study with incarcerated individuals in Northern Norway. *BMC psychology*, *10*(1), 185. <https://doi.org/10.1186/s40359-022-00895-5>
- Tobón, S. (2022). Do Better Prisons Reduce Recidivism? Evidence from a Prison Construction Program. *The Review of Economics and Statistics*, *104*(6), 1256–1272. [https://doi.org/10.1162/rest\\_a\\_01007](https://doi.org/10.1162/rest_a_01007)
- Tollenaar, N., & van der Heijden, P. G. M. (2013). Which method predicts recidivism best?: a comparison of statistical, machine learning and data mining predictive models. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, *176*(2), 565–584. <http://www.jstor.org/stable/23355205>
- Travaini, G. V., Pacchioni, F., Bellumore, S., Bosia, M., & De Micco, F. (2022). Machine Learning and Criminal Justice: A Systematic Review of Advanced Methodology for Recidivism Risk Prediction. *International journal of environmental research and public health*, *19*(17), 10594.  
<https://doi.org/10.3390/ijerph191710594>
- Tripodi, S. J., Kim, J. S., & Bender, K. (2010). Is Employment Associated With Reduced Recidivism?: The Complex Relationship Between Employment and Crime. *International Journal of Offender Therapy and Comparative Criminology*, *54*(5), 706–720. <https://doi.org/10.1177/0306624X09342980>
- Visher, C. A., & Travis, J. (2003). Transitions from Prison to Community: Understanding Individual Pathways. *Annual Review of Sociology*. <https://doi.org/10.1146/annurev.soc.29.010202.095931>