Reducing Recidivism in North Carolina

Bhargavi Ganesh, Pete Rodrigue, Vedika Ahuja

06/12/2019

Background and Motivation

North Carolina has one of the highest re-incarceration rates in the country. Nearly half of people released from jail in North Carolina are arrested again within two years, and nearly a third are reincarcerated. Reincarceration is harmful to society, and reducing recidivism is a priority for the State of North Carolina.

But the state's Division of Adult Corrections has limited resources ensure offenders succeed post release and do not return to prison. Currently North Carolina allocates its scarce funding and probation staff using a risk assessment algorithm to classify people into 5 recidivism risk categories. Department staff also conduct a needs assessment for each newly-released individual: a combination of an officer's impressions of the person and the person's self-reported needs. The state currently uses the results of these assessments to decide how frequently a parole officer should visit a person, and how they should respond if the person violates parole.

Our objective is to reduce North Carolina's one-year recidivism rate for people being released for felony convictions by 10%. By recidivism we mean returning to prison for a misdemeanor or felony conviction (again, within one year of release).

Our hypothesis is that by accurately identifying the people at highest risk of returning to prison, the Department will be able to provide these individuals with social workers to help them re-integrate. These interventions will result in fewer people returning to prison, which will lower the one-year recidivism rate.

Our plan is to create a risk assessment tool, similar to those the Department currently uses, which will predict risk of recidivism and allow Department staff to allocate scarce support resources pre- and post- release to the people at highest risk of recidivism.

Related and Previous Work

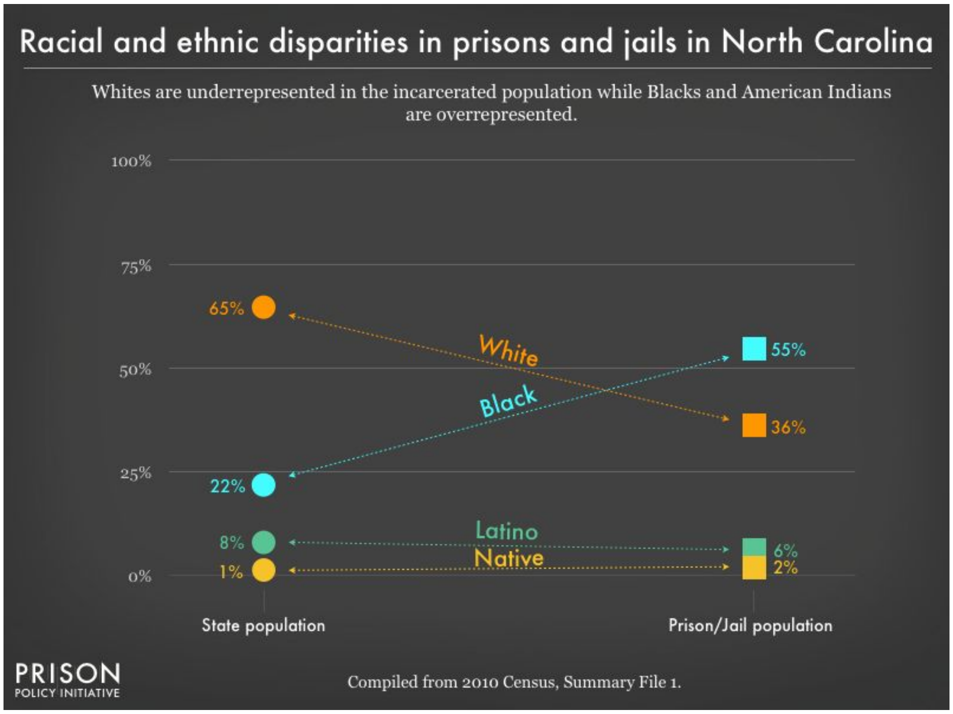
Reducing recidivism reduces human suffering

Extensive research has documented the negative effects of recidivism and imprisonment:

- The Illinois Sentencing Policy Advisory Council estimates that half the total cost associated with a recidivism event, roughly $75,000, is borne by the victims of the crime.
- Employers are averse to hiring formerly incarcerated candidates.
- Incarceration itself increases people's risk of being incarcerated in the future.
- Spouses of inmates suffer from higher rates of depression and financial hardship; children of inmates see their grades slip and experience more behavioral problems.

Recidivism is one factor perpetuating racial inequality in North Carolina.

Black and African American North Carolinians are overrepresented in the state's justice system. One in 40 Black men in North Carolina were imprisoned in 2016.



Source: The Prison Policy Initiative

Reducing recidivism saves government money

A 2011 Pew report found that if the 10 states with the greatest potential cost savings reduced their recidivism rates by 10 percent, they could save more than $470 million in a single year. North Carolina was among those states, and stood to save $23 million annually from reduced recidivism.

It costs up to $42,000 per year to house inmates in North Carolina, several times what the state spends per student on K-12 education.

Our proposed pilot program

Our proposed program is based off of the Michigan Prison Re-entry Initiative, a package of policy changes that led to a reduction in the inmate population in Michigan of 12 percent, and closings of more than 20 correctional facilities.[1] **The program intervenes with prisoners at three points - during a prisoner's sentence, the months prior to a prisoner's release, and the post-release phase involving community supervision.**[2]

North Carolina has recently reformed parole terms and supervision based on a prisoner's likelihood to recidivate, so we propose for the state to now start a pilot program to prepare offenders during their prison term, in the six months prior to their release, and to continue providing social services once released.

The pilot program would serve the people most at risk of recidivating, and would involve various departments coming together to create a Transition Accountability Plan for an inmate when the inmate enters prison, and offer counseling and resources tailored to the inmates needs and strengths. This can include helping the inmate find a job, secure housing, get an education, and build community connections. After the prisoners are released, social workers will continue to check in with inmates about their progress and offer training and resources for the inmate to follow through with the Accountability Plan. For example, in Michigan, certain pilot programs involved the Michigan State Housing Development Authority (MSHDA) and provided residential and case management and federal rent subsidy funds to recently released offenders.[3]

---

[1] https://www.pewtrusts.org/en/research-and-analysis/reports/0001/01/01/state-of-recidivism
[2] https://www.michigan.gov/corrections/0,4551,7-119-1441_1476-103248--,00.html
[3] "Michigan Prisoner Re-entry Initiative is ready to roll," Michigan Department of Corrections. 2004.
https://www.michigan.gov/corrections/0,4551,7-119-1441_1476-103248--,00.html

Policy Goals

1. Reduce the one-year recidivism rate in North Carolina for people being released from prison for a felony conviction by 10%.

2. Allocate pre- and post- release resources through a pilot program in the most efficient way possible to reduce prison incarceration in the state, given the limited resources available in the Division of Adult Corrections.

3. Ensure that the risk assessments used to achieve this reduction in recidivism are not biased against people of color and women.

North Carolina passed the Justice Reinvestment Act in 2011, starting the use of risk-assessment algorithms to determine the level of probation supervision imposed on exiting felons. Other states that have greatly reduced their recidivism rate in the past 15 years, including Oregon and Michigan, have increased resources and training for felons prior to their release and social services once released. For example, "In prison, Oregon inmates receive risk and needs assessments at intake, and targeted case management during incarceration, along with detailed transition planning that begins six months before release."[4] Michigan provides similar programs through the Michigan Prisoner Reentry Initiative.[5] We propose that Community Corrections, part of the Division of Adult Correction, create a pilot program offering comprehensive services prior to release for felons who are most at risk of recidivating according to our model.

[4] "State of Recidivism," Pew Charitable Trusts. April 11, 2011. https://www.pewtrusts.org/en/research-and-analysis/reports/0001/01/01/state-of-recidivism
[5] "Michigan Sees Decline in Prisoner Return Rate," Bridget Bodnar. April 13, 2011. https://www.michiganradio.org/post/michigan-sees-decline-prisoner-return-rate

Machine Learning Problem Formulation

According to the Pew Charitable Trusts, if North Carolina reduced it's recidivism rate by 10%, it could save more than $23 million a year.[6] These savings will only be realized, however, if officials can identify the newly-released individuals who will actually recidivate. A predictive machine learning model is a tool to do that. The model will rank offenders by their probability of recidivating within one year of their release.

We aim to pick a model that most precisely identify newly-released individuals who will return to prison within a year. Such a predictive tool will allow us to allocate our proposed pilot program effectively, and by extension, reduce the recidivism rate by 10% for the lowest cost, assuming resources are most effective when given to those most likely to recidivate. In order to reduce recidivism by 10%, we have to intervene with 5% of the total population, due to less than perfect precision and intervention effectiveness. ***For this reason, we have chosen precision at 5% as our key model performance metric.*** We explain our reasoning in more detail in the evaluation section below.

We also provide the cost-benefit analysis for reducing the recidivism rate by less than 10%, assuming linear savings for each percentage reduction in the one-year recidivism rate.

Data Description, Exploration, and Descriptive Statistics

The data contains all public information on all NC Department of Public Safety offenders convicted from 1972 up to 2019. We subset this dataset to look only at the years 2007-2017 (including 2007 and 2017). This is for two reasons. First, we wanted our models to run faster by looking at fewer years of data, given the time constraints we faced. Given more time, we might try training our models on more years of historical data, to see if their predictive performance improves. Second, we were unsure how informative patterns in recidivism from the 1970s, 80s, and 90s would be when trying to predict recidivism today. This is why we trained our models on more recent data.

We also subset our data to predict recidivism only for people being released after a sentence served for a felony conviction. This is because the majority of the literature on recidivism interventions focuses on reintegrating people being released for felony

[6] "State of Recidivism," Pew Charitable Trusts. April 11, 2011. https://www.pewtrusts.org/en/research-and-analysis/reports/0001/01/01/state-of-recidivism

convictions, and we wanted the setting of our proposed intervention to approximate the setting in which it's been tested.

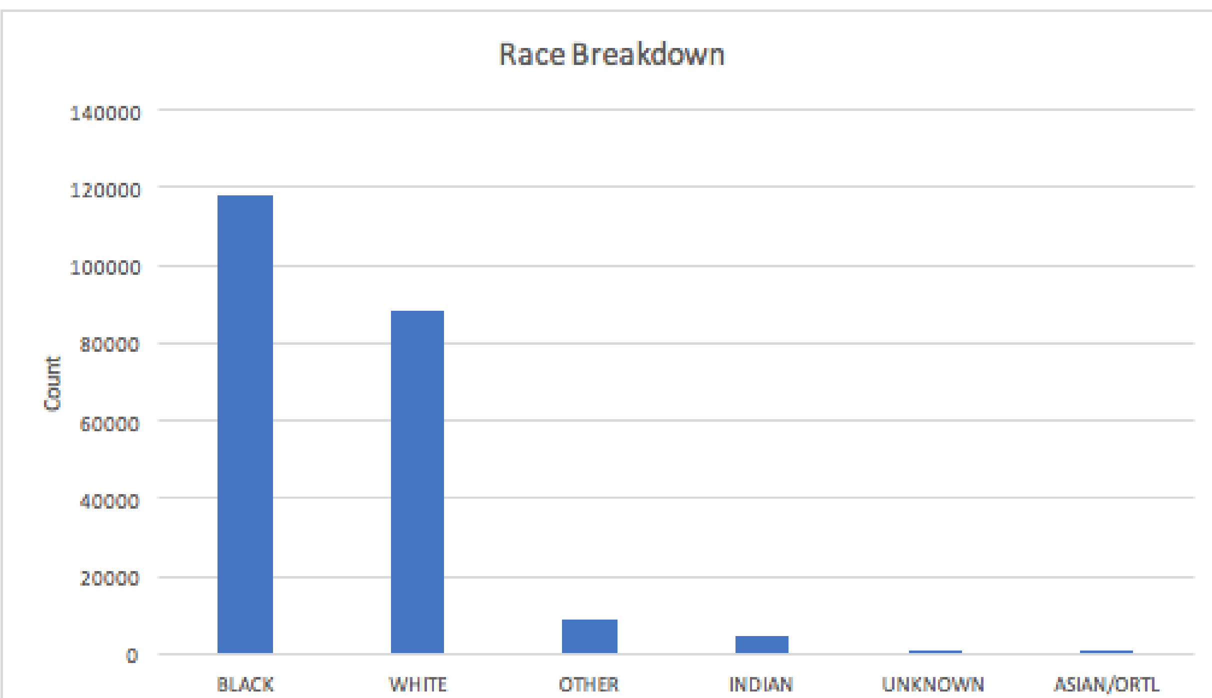
In terms of specific datasets, we merge the following csv files:

- OFNT3AA1: This dataset contains demographic information about people who've been to prison in North Carolina.
- OFNT3CE1: This dataset contains information on the specific counts and crimes that each person is charged with.
- INMT4BB1: This dataset contains information on people's time served in prison.

See appendix for a full list of variables used in our pipeline.

Descriptive Statistics

Eighty-eight percent of our sample was male; the remainder was coded as female. Fifty-three percent of the sample was black, and 40 was white. A small number were coded as "other," Indian, "unknown," or Asian. Ninety-seven percent of people are U.S. citizens.



The table below shows summary statistics for selected variables. Across our whole sample, roughly 10 percent of people returned to prison within a year; the median sentence length for the initial incarceration was 1.1 years. A quarter of the people in our sample were incarcerated before their 23rd birthday.

	count	mean	std	min	25%	50%	75%	max
Total counts for crime	220957.0	1.9	1.5	1.0	1.0	1.0	2.0	86.0
Felony counts for crime	220957.0	1.7	1.3	1.0	1.0	1.0	2.0	86.0
Recidivated?	220957.0	0.1	0.3	0.0	0.0	0.0	0.0	1.0
Person height in inches	220862.0	69.2	3.7	0.0	67.0	69.0	72.0	96.0
Person weight in lbs	220865.0	190.0	40.7	0.0	162.0	184.0	211.0	789.0
Age at time of crime	220957.0	32.6	10.6	14.1	24.1	30.4	39.5	89.8
Sentence length	220957.0	3.2	6.7	-0.9	0.5	1.1	3.0	269.5
Age at first incarceration	220957.0	31.1	10.4	14.1	22.6	28.7	37.8	89.8
Number of previous incarcerations	220957.0	0.5	0.8	0.0	0.0	0.0	1.0	10.0

Finally, we looked at correlations between the variables, which can be found in the appendix.

Overall Setup and Overview of our Machine Learning pipeline

To prepare our data for the machine learning pipeline, we merge the offender file, inmate file, and demographic characteristics file. Each row represents a person being released from prison in North Carolina for a felony conviction at a particular point in time. This means the same person could appear in multiple rows, if they were released from prison for a felony multiple times.

In the raw data, each row represents a count of a crime. Our data cleaning process reshapes the data so that each row is a crime, and then merges the data sources. The process is as follows:

- First this code imports the dataset on offender history, called OFNT3CE1. Each row is a count. There may be multiple counts per crime, and each person may have multiple crimes. We collapse the offender data to the crime level, aggregating details about the counts within each crime. These include the number of counts per crime, the number of felony counts per crime, and counts of various types of charges. See appendix for a full list.
- Then we import the inmate dataset, called INMT4BB1. We clean the data, imputing the release date if the actual sentence release date is missing, using the projected release date.
- Next we import the demographic data, do some cleaning, and drop the clothing size variables, which are mostly missing. After creating some time-based variables, like our outcome label and the age at first incarceration, we merge all three of these datasets together.

After our data is made ready for the pipeline, we make a list of variables that need to be dropped. We dropped ID variables, the start time of next incarceration, and the variable indicating whether a crime was a felony or a misdemeanor. The start time of next incarceration was dropped because it was a variable used to make the recidivate label, and the felony or misdemeanor flag was dropped because we had subsetted the data to include only felonies. We also made a list of variables to be imputed, and made into categorical variables. The imputation, dropping of variables, and categorizing of variables all happens after we do our train, test splits. We loop through a set of parameters for each model, and for each timeframe within the train/test split, and then compare the following evaluation metrics: precision, recall, accuracy, and f1, at 1%, 2%, 5%, 10%, 20%, 30% and 50% of the population.

Evaluation: Train & Test Sets, Metrics

According to the Pew Charitable Trusts, if North Carolina reduced it's recidivism rate by 10%, it could save more than $23 million a year. In order to pick our metric of interest, we defined our goal as a reduction in the recidivism rate of 10%. In order to reduce recidivism by 10%, we need to intervene on 5% of the population of released felons (accounting for the precision of our models at each threshold). We have provided a menu of options if the State of North Carolina and the Division of Adult Corrections wants to adjust the final targeted reduction in recidivism, or the amount spent on the pilot program.

In order to determine that our metric of interest is precision at 5% we calculated that the average number of felons released from prison in our three validation time periods was 16,788, of which 1,324 recidivated. In order to reduce recidivism by 10%, we would have to reduce the number of people who recidivate by about 132 people. If we assume that our proposed policy intervention is 50% effective, and our precision at 5% is about .28 (the precision at 5% of our best models), then targeting 5% of the population would lead to a reduction in about 118 people recidivating, which is close to our goal of 132 people. Our calculations (assuming a 50% effectiveness rate) are shown in the table below.

We have also calculated the expected savings associated with different reductions in recidivism, assuming the savings outlined in the Pew report scales linearly. Finally we calculated the amount the state can spend per person on an intervention to break even, given those projected savings. Policy makers should decide how much they are willing to spend to achieve a given reduction in recidivism. They also need to take into account how much they realistically need to spend on an intervention per person to ensure some effectiveness of the intervention.

Calculations assuming different effectiveness rates are in the appendix. The table below shows potential break-even spending assuming the program effectiveness is 50 percent.

Felons Released	# People who Recidivated	% Targeted by Intervention	Average precision at X% (for % targeted)	# of people targeted	# of People Correctly Identify	Reduction in Recidvism (people)	Total $ Saved	Break Even Spending Per Person	Total
16788	1324	1%	0.36	168	60	30	$ 5,265,223	$ 31,364	5,265,223
16788	1324	2%	0.31	336	104	52	$ 9,067,884	$ 27,008	9,067,884
16788	1324	5%	0.28	839	235	118	$ 20,475,866	$ 24,394	20,475,866
16788	1324	10%	0.24	1679	403	201	$ 35,101,485	$ 20,909	35,101,485

We used temporal splits with training and validation sets of 1 year, with a one year buffer period in between the training and test dates. The reason for this one year buffer is that we would like to leave a year to see the impact of the intervention on the intended population.

The train/test splits are as shown below:

2007 train	--	2009 test	--	2011 train	--	2013 test	--	2015 train	--	2017 test	--

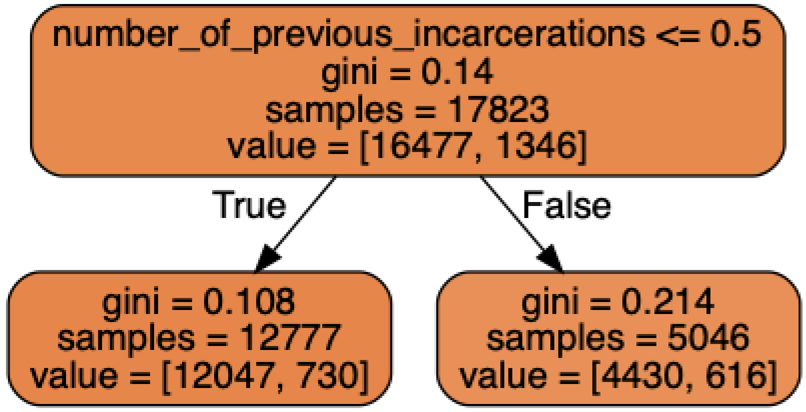
In future work, we would extend the training set, such that the first training set would be from 2007-2008, while 2010 onward would be the test set, and the second training set would be from 2007-2009, while 2011 onward would be the test set, etc. We could also use more historical data, going back to the 1990s, to see if that improved our model's performance.

Results and Feature Importances

Baseline

To benchmark our models, we use a decision tree stump with one variable: number of previous incarcerations. This approximates a rule of thumb that experts use to get a first approximation of the chances that someone will recidivate. This stump sorts based on whether someone has been incarcerated before. Here is the stump for the second split. The baseline recidivism rate is 7.5%. The stump says that someone who's been incarcerated at least once has a 12.2% chance of recidivating, while someone who has no previous incarcerations has only a 5.7% chance. We can think of this 12.2% as the stump model's precision at 5%, if we were to choose randomly from the people who've been incarcerated to meet our 5% threshold. About 12% of them would actually recidivate in a counterfactual.

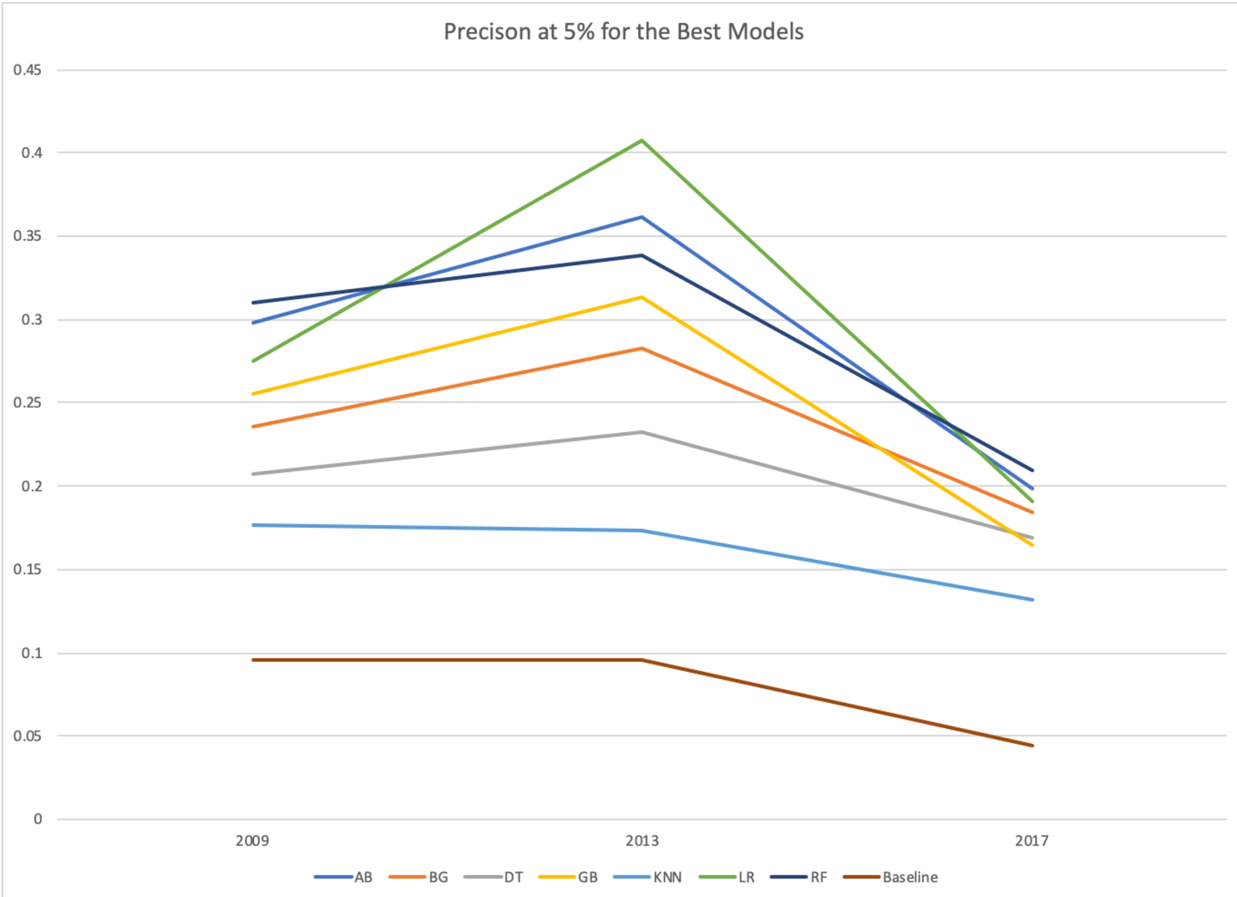
Our preferred model, outlined below, has just over double the precision of the stump model.



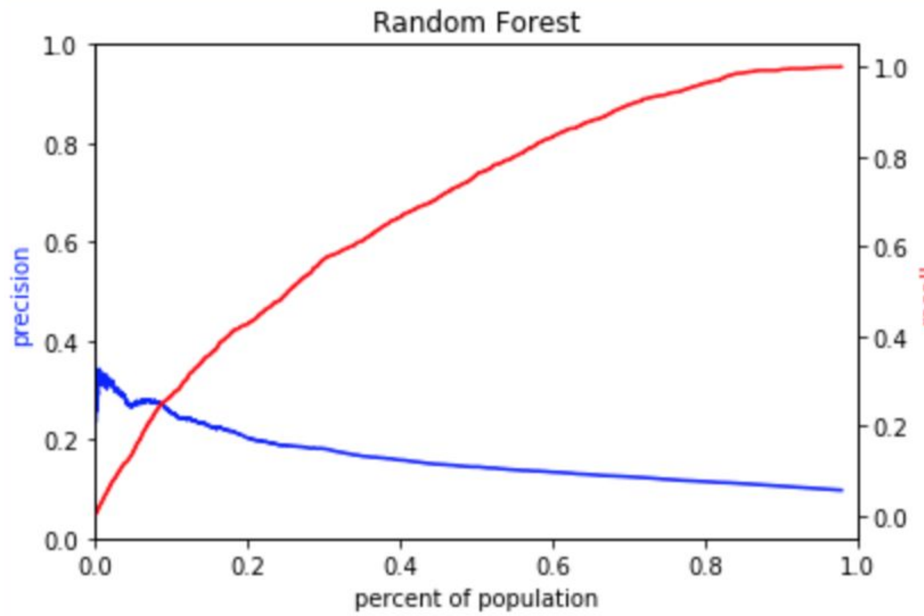
Our preferred model

We recommend the State of North Carolina use the Random Forest model with a max depth of 10, min sample split of 2, and 1000 n-estimators. This model had the best average precision at the 5% level across the time splits of 28.6%. The logistic regression model with an inverse of regularization strength (C=1) of L1 and L2 penalty, as well as the KNN model with a ball tree algorithm, calculated with 25 nearest neighbors, and weight points by the inverse of their distance all had very similar average precision at 5% over the 3 test splits to the best random forest models. The Random Forest model performed best in the first and third time split, and the logistic regression model performed best in the second time split. We recommend the random forest model over the logistic regression model for ease of interpretability, but both could be used for our purposes.

The graph below shows the precision at 5% for the models with the best average 5% precision across the time splits. The best models have a precision about 3 times higher than the baseline across the time splits.



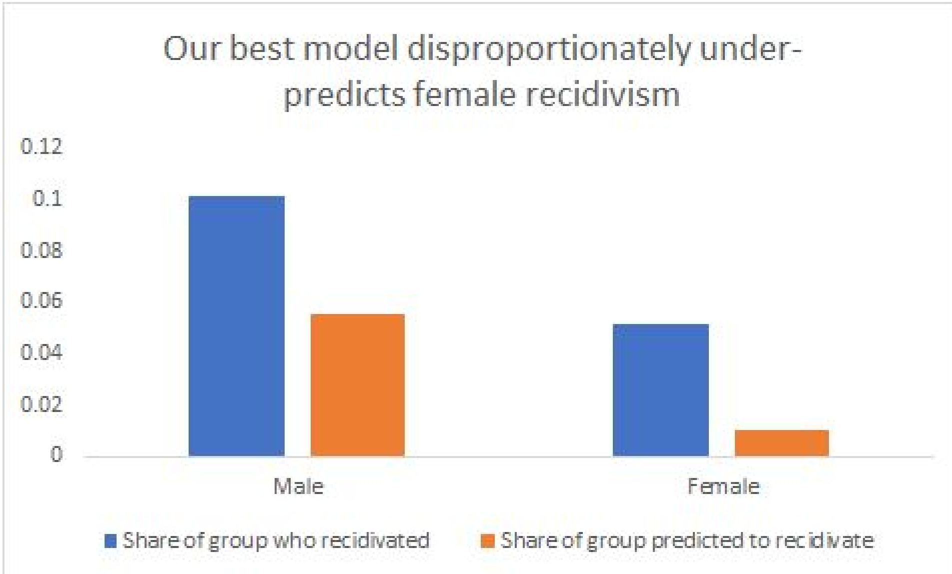
The graph below shows the precision-recall curve of our preferred Random Forest Model in the first time split. See appendix for the precision-recall curves for this model in the second and third time splits.

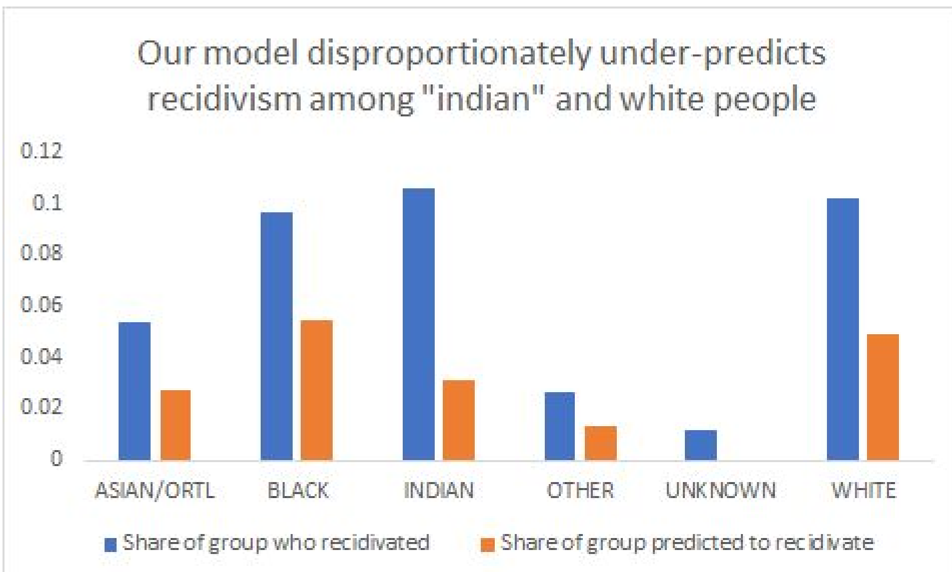


Bias analysis and discussion

We report bias results for our preferred model (the random forest with a maximum tree depth of 100, 1,000 estimators, and a minimum sample split of 2).

This model has some weaknesses with subgroup prediction. It disproportionately underpredicts female recidivism, as well as indian, white, Asian, and “other” recidivism, relative to the prediction rate for black individuals. Only 22 women were predicted to recidivate (out of 2,084 total women); only one Asian person was predicted to recidivate out of 37 total Asian people; only 12 Indian people were predicted to recidivate out of 386, and only 9 people were predicted to recidivate in the “other” category, out of 670 total.

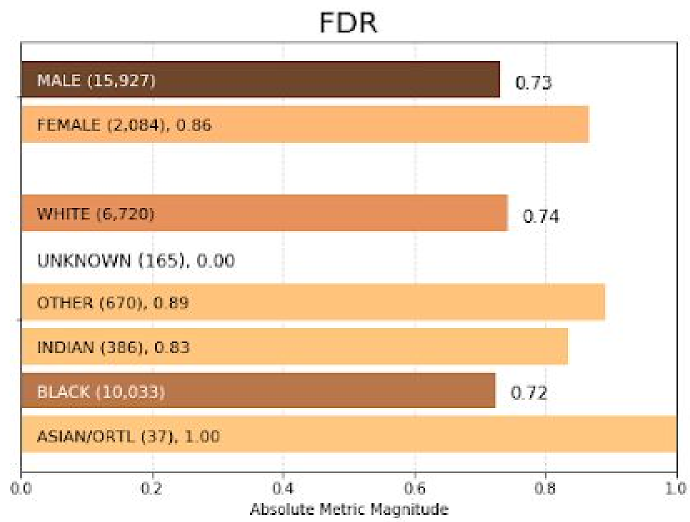




False discovery rate

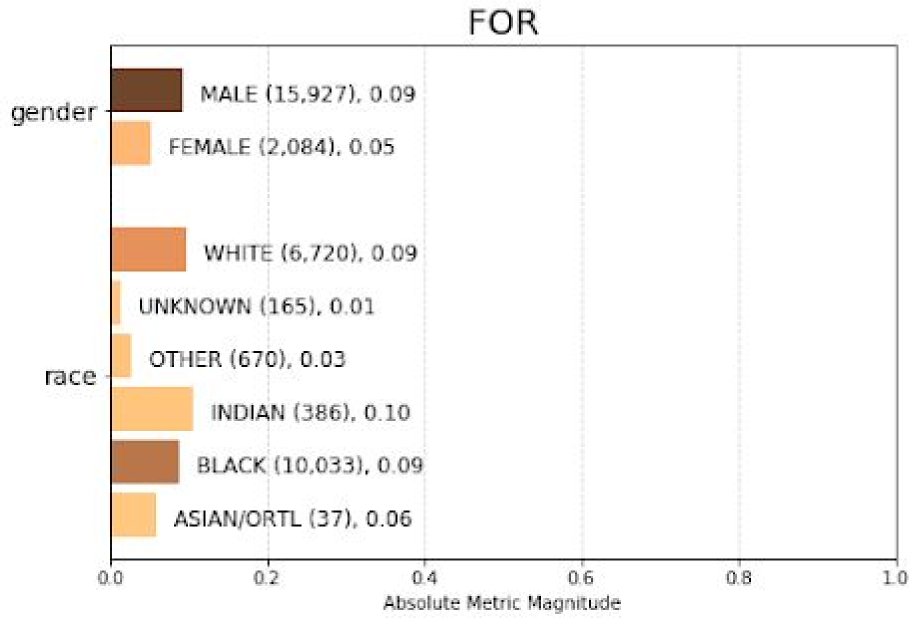
Not only did the model disproportionately under-predict female recidivism, when it *did* predict women would recidivate, it was especially inaccurate. 86 percent of women predicted to recidivate within a year did not actually recidivate. For men that figure was only 73 percent.

Similarly, the single Asian person predicted to recidivate did not; and 89 and 83 percent of "other" and Indian people who were predicted to recidivate did not. Black and white individuals had similar false discovery rates: about three-fourths of people predicted to recidivate did not. This might be concerning because the literature suggests that intervening with lower-risk people might actually backfire, and making them more likely to recidivate. For this reason, we might be especially hesitant to apply this prediction model and intervention to women set to be released from prison.



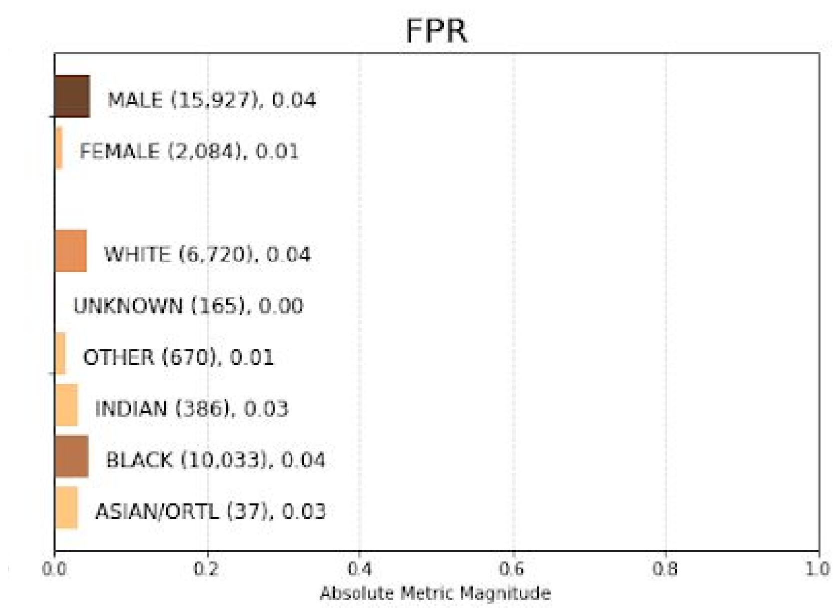
False omission rate

Our preferred model misses more male recidivators that female recidivators; roughly 9 percent of the men we said would not recidivate within a year did. For women that figure was 5%. We missed more white, Indian, and black recidivators than “other” and “unknown” category people, and Asian people. While this is a shame (we’re missing people that could use support services), it’s comforting to know that the false omission rate for our two largest groups, black and white men, is similar.



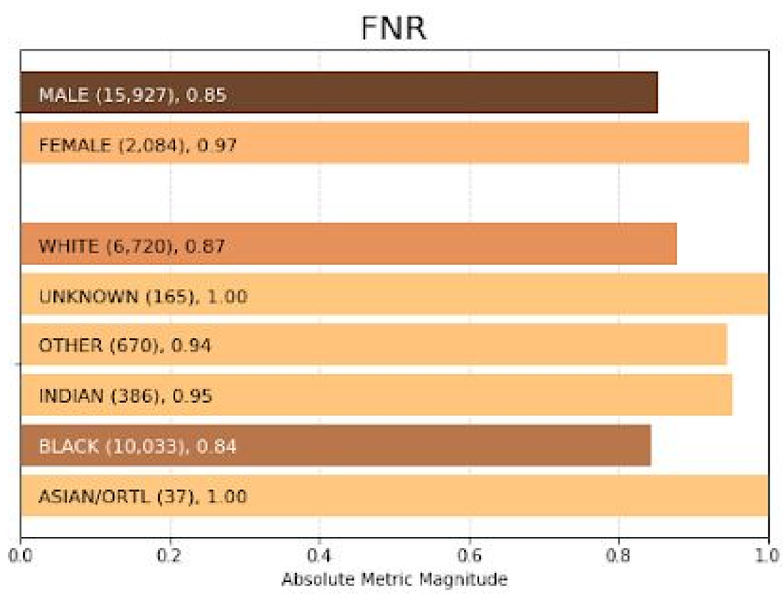
False positive rate

The men we said would recidivate but did not represented about 4 percent of all men who did not recidivate. The women we predicted to recidivate but did not represented only 1 percent of all the women who did not recidivate. The false positive rate for black and white people was similar to that of men: about 4 percent.



False negative rate

As a share of all the people who did actually recidivate, the people we mistakenly said would not recidivate was quite high: between 85 and 100 percent of people in each subgroup. We failed to correctly identify any of the Asian and “unknown” group people who recidivated, so the false negative rate for those groups was 100%. The rate was higher for women than for men; it was similar for black and white people. This is partly a result of our relatively low threshold (5%) and partly a result of our model’s relatively poor precision (around 30%).



When we ran models that omitted the race and gender variables, the results reported in this section were similar. We suspect that is because the other 2,000 variables are correlated with race and gender, and capture some of the same variation.

Caveats

The model has relatively low precision: roughly 70 percent of the people predicted to recidivate will not. The literature suggests that attempting to prevent recidivism among people who are at low-risk of recidivism could backfire. This should be weighed when considering whether to use this model.

The model has especially low precision for women, and Indian prisoners. It is not ready for use with those subpopulations. Before we can recommend use of our model, we would like to create two different models for men and women to see if this reduces bias in our predictions.

We assume the pilot intervention is more effective when given to the people most likely to recidivate within a year. This may not be the case. It could be, for example, that the intervention is most able to prevent recidivism among people who are only moderately at risk of recidivism.

Our preferred pilot intervention is modeled on a program that was developed in Michigan. North Carolina's prison population may differ in important ways from Michigan's, and that should be taken into account when seeking to introduce a similar intervention in North Carolina.

We have not included any data in our pipeline about what happens to people (or what they do) while incarcerated. We could add this to the pipeline given more time, and it would likely add predictive power to our model.

We do not use any data before 2007 to train our model, or seasonal dummy variables. Adding more years of historical data and seasonal dummies might improve our model performance.

We limit our analysis to people being released for felony convictions. This predictive tool should not be used for people being released for misdemeanor convictions without updating the pipeline to include train on data with those individuals.

Policy Recommendations

Due to the disproportionately low precision our model has with women and several race subgroups, **we need to do more work before recommending the Department of Corrections in North Carolina use our model to make predictions for policy purposes**. Specifically, we would want to experiment with using more years of training data, and adding data about people's behavior while incarcerated. Both would likely improve the models' precision.

If the Department of Corrections wants to employ our model immediately for our proposed policy intervention described above, we recommend it creates a pilot specifically focused on Black and White men. Before intervening with individual prisoners, the Department should conduct a screening interview to ensure the offenders are willing to participate in a holistic reintegration program involving pre- and post release counseling and services.

It is important to mention

Although our model outperforms a simple heuristic model (a stump decision tree), the Department should compare the precision of our model with the precision of the model they are already using.

Appendix

Variables in our datasets

OFNT3AA1: This contains demographic information about people who've been to prison in North Carolina.

Variables:

- Age at time of crime (we create this variable based on birth year)
- Gender
- Race
- City, county, state, and country of birth
- Citizenship
- Ethnicity
- Primary language
- Skin complexion
- Height
- Weight
- Body build code
- Hair color
- Eye color

OFNT3CE1: This dataset contains information on the specific counts and crimes that each person is charged with.

Variables:

- County of conviction
- Punishment type code (is the count being punished through the "fair felons," "DWI," "community service," or another process?)
- Component disposition code (was the outcome of a given count guilty, not guilty, a plea, partial revocation, etc)
- Primary offense code (trespass, larceny, assault, etc)
- Court type code (district, superior, magistrate, etc)
- Sentencing penalty class code (how severe is the offense and what is the possible

range of prison times?[7]

- Number of counts per crime (we create this variable)
- Number of felony counts per crime (we create this variable)

INMT4BB1: This dataset contains information on people's time served in prison.

Variables:

- Age at first incarceration (we create this variable)
- Length of sentence (we create this variable)
- Number of previous incarcerations (we create this variable)

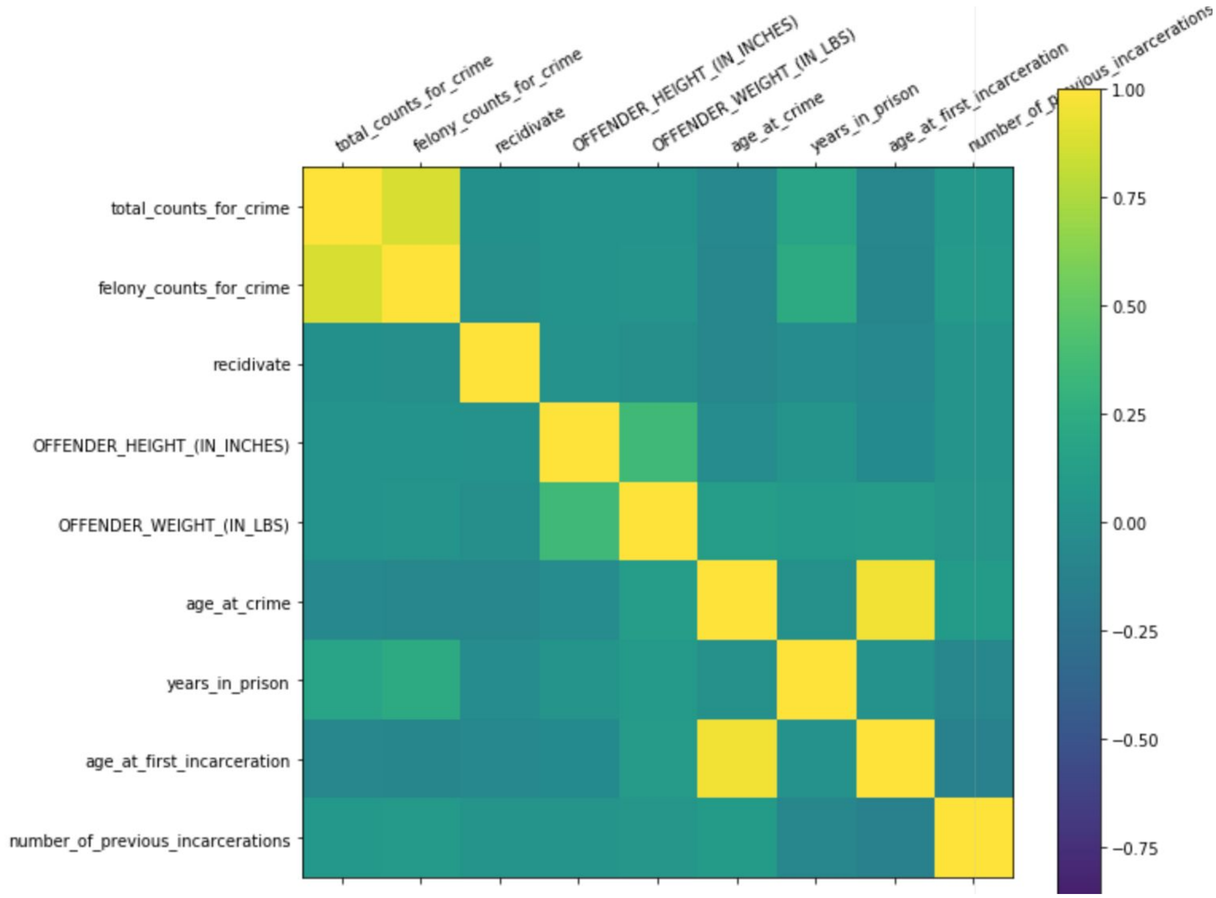
List of offender DOC id numbers predicted to recidivate at the 5% threshold during the first test split using our preferred model. See file "list_of_doc_ids.csv" for the complete list.

rank	predicted probability of recidivating within one year	DOC ID number
1	0.24154	912167
2	0.232872	899361
3	0.230869	1076306
4	0.230851	1096729
5	0.230158	1168060
6	0.228557	907032
7	0.228538	122917
8	0.227056	44379

[7] See here for more: https://www.nccourts.gov/assets/documents/publications/FelonyChart_1013MaxChart.pdf?JOZLdcExFM1TmlzHLiPcH7dUcMjQ8Ls7

9	0.223634	1030837
10	0.223365	837080
11	0.222526	1020170
12	0.221908	1081290
13	0.221564	1038415
14	0.220789	930859
15	0.219731	862754
16	0.2196	1050480
17	0.21833	722922
18	0.217763	1068739
19	0.217741	945586
20	0.217241	47838

Correlations between select variables



People Targeted, reached, and cost assuming different levels of effectiveness of the policy

Assumed Effectiveness: 75%									
Felons Released	# People who Recidivated	% Targeted by Intervention	Average precision at X% (for % targeted)	# of people targeted	# of People Correctly Identify	Reduction in Recidvism (people)	Total $ Saved	Break Even Spending Per Person	Total
16788	1324	1%	0.36	168	60	45	$ 7,897,834	$ 47,045	7,897,834
16788	1324	2%	0.31	336	104	78	$ 13,601,825	$ 40,511	13,601,825
16788	1324	5%	0.28	839	235	176	$ 30,713,799	$ 36,591	30,713,799
16788	1324	10%	0.24	1679	403	302	$ 52,652,227	$ 31,364	52,652,227

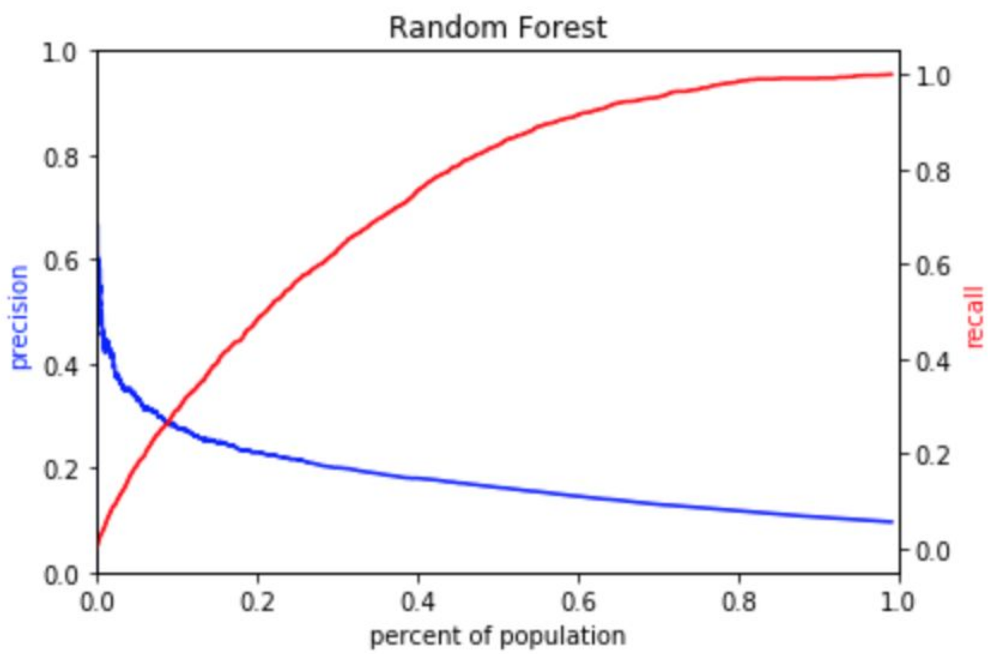
Assumed Effectiveness: 50%									
Felons Released	# People who Recidivated	% Targeted by Intervention	Average precision at X% (for % targeted)	# of people targeted	# of People Correctly Identify	Reduction in Recidvism (people)	Total $ Saved	Break Even Spending Per Person	Total
16788	1324	1%	0.36	168	60	30	$ 5,265,223	$ 31,364	5,265,223
16788	1324	2%	0.31	336	104	52	$ 9,067,884	$ 27,008	9,067,884
16788	1324	5%	0.28	839	235	118	$ 20,475,866	$ 24,394	20,475,866
16788	1324	10%	0.24	1679	403	201	$ 35,101,485	$ 20,909	35,101,485

Assumed Effectiveness: 25%									
Felons Released	# People who Recidivated	% Targeted by Intervention	Average precision at X% (for % targeted)	# of people targeted	# of People Correctly Identify	Reduction in Recidvism (people)	Total $ Saved	Break Even Spending Per Person	Total
16788	1324	1%	0.36	168	60	15	$ 2,632,611	$ 15,682	2,632,611
16788	1324	2%	0.31	336	104	26	$ 4,533,942	$ 13,504	4,533,942
16788	1324	5%	0.28	839	235	59	$ 10,237,933	$ 12,197	10,237,933
16788	1324	10%	0.24	1679	403	101	$ 17,550,742	$ 10,455	17,550,742

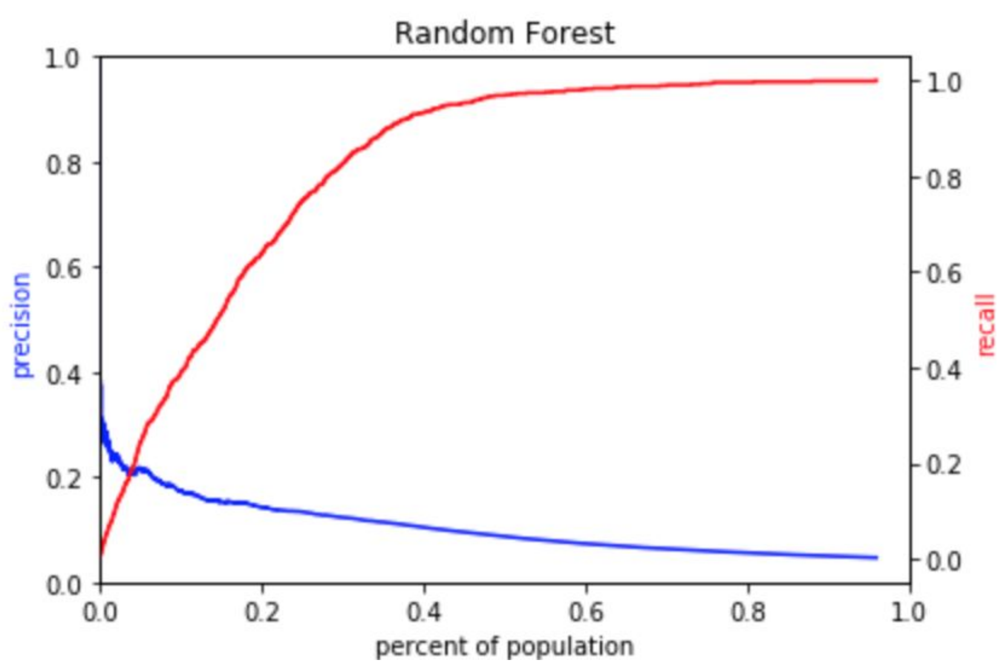
Assumed Effectiveness: 10%									
Felons Released	# People who Recidivated	% Targeted by Intervention	Average precision at X% (for % targeted)	# of people targeted	# of People Correctly Identify	Reduction in Recidvism (people)	Total $ Saved	Break Even Spending Per Person	Total
16788	1324	1%	0.36	168	60	6	$ 1,053,045	$ 6,273	1,053,045
16788	1324	2%	0.31	336	104	10	$ 1,813,577	$ 5,402	1,813,577
16788	1324	5%	0.28	839	235	24	$ 4,095,173	$ 4,879	4,095,173
16788	1324	10%	0.24	1679	403	40	$ 7,020,297	$ 4,182	7,020,297

Precision-Recall Curves

Time 2 Split



Time 3 Split:



Feature Importance

Rank	Variable	Importance
1	age_at_first_incarceration	0.053211409
2	number_of_previous_incarcerations	0.044073264
3	age_at_crime	0.03876745
4	years_in_prison	0.036880498
5	PRIMARY_OFFENSE_CODE_FELONY B&E	0.018838121
6	PUNISHMENT_TYPE_CODE_POST RELEASE	0.01629089
7	SENTENCING_PENALTY_CLASS_CODE_CLASS H	0.016174225
8	OFFENDER_WEIGHT_(IN_LBS)	0.015128403
9	COURT_TYPE_CODE_SUPERIOR	0.014941748
10	PUNISHMENT_TYPE_CODE_ACTIVE SS	0.014156582