# Mental Health and Incarceration

Abhilash Biswas, Yu Hou, Pearry Khachichit, Matt Lampl, Sara Maillacheruvu

## Executive Summary

Incarceration and mental illness exist at an uncomfortable intersection. In the United States, jails and prisons have often served as the largest housing facilities for individuals with mental illnesses. This is of particular concern in Johnson County, Kansas. Over 3 in 5 individuals in Johnson County's local jails have documented mental health illness, while over half have some form of substance abuse or dependence. Between November 2016 and June 2019, over one in four individuals who were booked at the local jail were referred to mental health services, based on the Brief Jail Mental Health Screen (BJMHS). Despite the widespread prevalence of mental illness among the incarcerated population, the National Alliance on Mental Illness reports that only 45% of people in local jails receive mental health care while incarcerated.

Working together with the Johnson County Mental Health Center (JCMHC), we aim to break the cycle of reincarceration by providing proactive mental health outreach and treatment to formerly incarcerated people. We will do this by developing a machine learning model to identify 100 individuals each month with the highest risk of recidivating. The JCMHC can then take this list and offer personalized outreach and care to those individuals. By intervening before recidivation, we aim to strengthen communities, preventing families and neighborhoods from undergoing the traumas associated with incarceration.

Using a metric of precision at the top 100 people identified as at-risk for reincarceration, we found that the two best machine learning models were a Scaled Logistic Regression model and a Random Forest Classifier model. These models had a precision of 0.86 and 0.84 respectively. (Precision here is the number of people in the top 100 who actually recidivated within one year divided by 100.) Both models outperformed our three baselines (ranking individuals based on number of bookings, number of visits to JCMHC, and average days out of jail between an individual's current release and next booking) in terms of precision. We found that the most important features tended to stem from inmate data, with the number of bookings being important for both. Notably, mental health features and simple demographic features such as age did not show up in the top 15 important features for either model. Additionally, we found that the models only identified those who are either Black or White as at-risk, and those identified tended to be males and in younger age groups.

With these results in mind, we first propose that the JCMHC further develop the BJMHS. This will allow the machine learning team to further hone our work by including mental illness in whom we consider in our treatment population. More importantly, it will allow the JCMHC to provide more tailored outreach and care to individuals, potentially increasing the efficacy of its work. Further, continued mental health screening while individuals are in jail will enable the JCMHC to identify individuals who have developed mental illnesses while in jail. We finish this report by proposing a menu of field trial options, including a randomized controlled trial, an instrumental variables analysis, and a regression discontinuity design. Each of these methods has strengths and weaknesses, and we invite the JCMHC to consider them in selecting a design that aligns with its values, ethics, and policy goals.

# Table of Contents

# 1 Background and Introduction

## 1.1 Problem Statement

There is a known link between mental health and reincarceration.[1] In fact, American jails and prisons have often served as the largest housing facilities for individuals with mental illnesses, a phenomenon that reflects the criminalization of mental illness.[2] This is of particular concern in Johnson County, Kansas, where 64% of individuals in local jails have documented mental health illnesses and 55% of individuals have some form of substance abuse or dependence.[3] While the prevalence of mental illness is high — both in Johnson County and across the US — the National Alliance on Mental Illness notes that only 45% of individuals in local jails receive mental health care while incarcerated.[4] Thus, local need is immense, but national trends indicate that the current provision of services is insufficient.

Untreated mental health conditions further a negative spiral, in which the criminal justice system ensnares people in a cycle of reincarceration. This leads to harmful social and community outcomes, taking parents away from children, community members away from neighborhoods, and weakening the social ties that knit places and people together. These impacts are also economic: the Brennan Center reports that being convicted of a misdemeanor is associated with a 16% reduction in annual earnings, whereas spending time in prison is associated with a sizable 52% reduction in annual earnings.[5] Moreover, being incarcerated may harm individuals' physical and mental health.[6] In fact, 1,200 people died in local jails in 2019, marking a 5.44% increase in local jail morbidity rates from 2018.[7] Tellingly, nearly 30% of local jail deaths occurred by suicide, making it the single-most leading cause of death and reinforcing the importance of adequate mental health care for current and formerly incarcerated people.[8]

There are important equity implications to highlight here. In a nationwide analysis, the Vera Institute determined that Black people are 2.17 times more likely to be arrested than white people across the US.[9] The Bureau of Jail Statistics determined that, between 1999 and 2014, incarceration rates for American Indians and Alaska Natives increased at three times the rate of people of other races.[10] Based on current incarceration rates, one in six Latino men will go to jail in their lifetime, compared to one in seventeen white men — the latter of which, even as a referent, is still too high.[11]

The criminalization of race can interact with mental health, placing a double burden on people of color — in particular, on Black people. While documented rates of mental illnesses tend to be lower among people of color, it is possible that armed with this prior knowledge on "base rates" and paired with poor physician-patient communication, medical providers may racially discriminate in how they treat patients, leading to lower treatment rates of mental health conditions for people of color.[12] Another study identified that, between 2004 and 2012, racial disparities with regard to access to mental health care[13] stagnated — and, in some cases, increased — with Black, Hispanic, and Asian people having reduced access to mental healthcare as compared to white people.[14] Further, the medical establishment's reaction to addiction can be racialized, e.g., in the medicalization of the opioid epidemic (associated with middle- and upper class white people) as compared to the criminalization of the crack-cocaine epidemic (associated with low-income Black and white people).[15]

In summary, mental health and incarceration are intertwined; an effective approach to reducing recidivism must, in some way, address the mental health needs of formerly incarcerated people. Moreover, research has shown that early interventions, ongoing care throughout individuals' incarceration periods, and post-incarceration care is crucial in reducing recidivism among individuals with mental illnesses.[16] However, at present, the JCMHC implements a reactive solution to address the mental health needs of formerly incarcerated people. At the same time, the organization has limited resources and must optimize the time and money spent on proactive outreach.

Despite these constraints, the JCMHC can have widespread, community-level impacts if it can begin to reduce recidivism via proactive mental healthcare. As discussed above, formerly incarcerated people face huge burdens upon reentry. The unemployment rate among formerly incarcerated people is 27%.[17] Further, the homelessness rate is nearly fourteen times higher for people who have formerly been *incarcerated more than one time* as compared to the general population (2.79% as compared to 0.21%).[18] In addition, incarceration impacts quality of life for families with incarcerated members.[19] One study noted that incarcerated fathers were likely to be less active in their families' lives and that divorce or relationship disunion rates are higher among incarcerated men.[20] In an interview, one researcher emphasized how detrimental incarceration was to children whose parent(s) have been incarcerated:

> [T]he incarceration of a parent sets off a downward spiral in which negative responses from teachers, correctional officials and even peers due to the stigma of parental incarceration interacts with negative behavioral responses to the trauma of that event to lead children down a difficult path that has dire consequences for their transition to adulthood.[21]

This is to say that while formerly incarcerated individuals may be ensnared in a pernicious cycle, so, too, are their family members — in particular, their children — with the potential for longitudinal impacts.

Thus, by breaking the cycle of incarceration, the JCMHC can begin to address the bevy of issues commonly associated with incarceration and reincarceration: reduced employment rates, impaired family relations, elevated trauma levels for children, increased homelessness rates, and reduced earnings, among others. Of course, these efforts must be long-sighted: change will require years to have deep and lasting impacts. Incarcerated and formerly incarcerated people belong to families and communities; like all other people, they are imbricated in a web of social relations. Through proactive mental health outreach, the JCMHC can begin to strengthen and restore these ties.

## 1.2 Potential Impact

The JCMHC aims to identify individuals at the highest risk of reincarceration to proactively provide mental health care in order to break the cycle of reincarceration by reducing the rate of reincarceration. Because the JCMHC has limited resources for this new proactive outreach effort, the organization estimates that its current level of budget and staffing will allow them to attempt to contact 100 individuals each month. Hence, our goal is to identify the 100 formerly incarcerated individuals with the highest risk of recidivating. As the resource constraint is at a monthly level, we will run the model every month to assess the target population.

The JCMHC currently offers a range of services to residents of Johnson County, including counseling, mobile response teams, etc. For this project, the JCMHC will reach out to identified individuals through calls and/or home visits. Based on the JCMHC's budget and individuals' desires, the organization may then choose to assign individuals to continued mental health services provided by JCMHC, based on their assessed needs.

## 1.3 Policy Goals

We aim to improve the quality of life and positively impact community health by reducing the rate of reincarceration stemming from untreated mental health conditions among the formerly incarcerated population in Johnson County, Kansas. By identifying at-risk populations and sharing our findings with JCMHC, we hope to address the well-documented association between mental illness and incarceration. To be more specific, we break down the goals into three types: effectiveness, efficiency, and equity.

### Effectiveness

Our effectiveness goal is to reduce the reincarceration rate among former inmate populations who are at risk of recidivism due to untreated mental health conditions. While it is our goal to break the negative spiral for as many people as possible, we are under the constraint that in every month, the JCMHC can conduct 100 interventions.

The intervention (proactive mental health services) would be considered effective if the probability of recidivating decreases among individuals who were given the intervention.

### Efficiency

Due to time and resources constraints, we want to ensure that the highest risk individuals receive the 100 interventions. Ideally, of the people that the model identifies to be at risk for reincarceration, the people who receive treatment should be those who are most likely to be reincarcerated. To do this, the model will assign a risk score to each individual in the model. In order to optimize JCMHC's resources, we want to choose those individuals with the highest risk scores. Once an individual has been selected by the model and receives JCMHC services, they will be removed from the pool.

### Equity

We understand that policing has disproportionate impacts on communities of color — in particular, on Black communities. As such, we want the model to not produce more false discoveries or less true discoveries on protected groups – Black people and women. This would mean that the model, in its top 100 predicted people who are at risk of recidivism, does not make more mistakes for Black people as compared to White people and also women as compared to men. We also want to ensure that of all the Black people that go back to jail, the model is able to detect similar proportions of White people as compared to Black people, and similarly for women as compared to men.

### Trade Offs

To achieve the effectiveness goal, we could simply target people with low chances of reincarceration, thereby reducing overall reincarceration. However, this will negatively impact the efficiency goal because people with the lowest chances of reincarceration are not those with the greatest need. Additionally, people with the lowest chances of reincarceration may come from communities of higher privilege, leading to limited equity effects.

## 1.4 Related Work

Currently, the JCMHC offers a reactive solution in breaking the cycle of incarceration. Once booked, each incoming inmate completes a health screening assessment (BJMHS). Based on responses to this assessment, individuals are flagged for referral to the JCMHC for mental health services. However, these mental health services are only provided to current inmates: once released, no mental health services are readily available to formerly incarcerated people.

In 2010, Johnson County received a Justice and Mental Health Collaboration Program grant from the U.S. Department of Justice to implement a mental health co-responder program with the Olathe Police Department (OPD).[22] Through this program, social workers ("co-responders") accompanied officers from the Olathe Police Department[23] and hosted training sessions to teach them about mental health, as well as equip them with best practices on responding to calls where individuals with mental illnesses were involved.[24] The program was successful in that it was associated with a decrease in police calls that resulted in jail arrests; a 59% decrease in police calls to the same person; and a substantial increase in referral rates, from 1.2% to 38.9%.[25]

The Mecklenburg County Jail in North Carolina has also implemented simple programs to address the intersection of mental health and incarceration. Since 2008, the Charlotte-Mecklenburg Crisis Intervention Team (CIT) has taken a trauma-informed treatment approach in training 800 officers in de-escalation skills.[26] The CIT has also addressed the intersectionality of housing, health,[27] and incarceration through its FUSE initiative, an inter-agency collaboration that provides 45 supporting housing units to individuals who have "been frequent users of Mecklenburg's jail, street camps, shelters and hospitals."[28] The FUSE program has been associated with decreased health costs per participant, fewer self-reported arrests/jail stays, and better self-reported health.[29]

While both of these programs show promise, the OPD and CIT approaches are limited in that they react to instances in which mental illness and the criminal justice system intersect. The OPD's new training regime demonstrated strong results, but much is left to officers' discretion in incorporating the training into their approach. Moreover, police-based approaches allow untreated mental illnesses to apex to the point of law enforcement involvement. In other words, they fail to recognize individuals' mental health needs outside of a carceral context; in a way, the criminal justice system, then, is a legitimizing force for mental illness — and the care that people likely needed long before interacting with this system.

Our approach marks an improvement to these existing strategies in that it offers a proactive solution to address the intersection of mental illness and the carceral system. Through machine learning, we will identify 100 individuals each month (a number that is limited by the JCMHC's budget) who have the highest risk of recidivating. Here, we acknowledge that agencies have implemented machine learning models in the criminal justice system in the past, with damaging effects that entrench racial bias into a carceral system which already criminalizes race.[30] However, our approach centers not on assisting the carceral system, but on building pathways to keep people out of it. Moreover, our findings will only be shared with the JCMHC; we will have strict data sharing agreements with the JCMHC to prevent our findings from being leaked to law enforcement agencies to be used in predictive policing.

To break the cycle of repeated incarceration, the JCMHC must proactively identify formerly incarcerated people who have a high risk of recidivating, in order to provide the with mental health services. The intention here is that this intervention can improve individuals mental health and prevent crises from developing that may have otherwise resulted in reincarceration. Our machine learning models will enable the JCMHC to quickly identify individuals to whom they should provide outreach, enhancing the efficiency of the agency's outreach processes.

## 1.5 Analytical Formulation

Due to the JCMHC's time, budget, and resource constraints, our analytical decisions are on a monthly basis, i.e., on the first date of every month, so that the JCMHC can prepare for the intervention. The targeted population included in the cohort is all individuals who were released from Johnson County jail in the last two years. The output of the model is a list of 100 individuals of formerly incarcerated people who we have identified as having the highest risk of recidivating in the next one year (365 days). Here, we consider one year as the relevant timespan because it may offer the JCMHC enough time to intervene to prevent incarceration. Thus, our analytical formulation is as follows:

> *On the first of every month, of all the individuals who were released from Johnson County Jail in the last 2 years and are not currently in jail, we aim to identify 100 individuals who are most likely to return to jail in the next year (365 days) to prioritize proactive mental health interventions to prevent re-incarceration.*

We acknowledge that our present formulation does not include demonstrated mental illness or mental health need as a condition for being included in the cohort. While we have access to the "Brief Jail Mental Health Screen" — which includes information on whether someone should be referred to JCMHC — we only had this data for recent years (2016-2019). As an equity concern, we did not want to restrict our cohort to individuals who only had recent interactions with the criminal justice system. Moreover, it is possible that there is bias in the BJMHS due to erroneous or missing data. Individuals may inaccurately complete the survey out of fear of repercussion or misunderstanding, meaning that the current BJMHS tool may not be an effective proxy of mental illness or mental health in general.

While we could leverage the BJMHS along with data from the JCMHC (i.e., the DLA and CAFAS score assessment), data missingness in the more distant past is again a concern. Further, mental illness and mental health in general are complex phenomena; simple, checkbox-based screening tools may fail to account for nuance. As such, we opt for an analytical formulation that enables all formerly incarcerated individuals released within the past 2 years to be considered for treatment. As documented in our policy recommendations, we will add further complexity to the cohort definition vis-à-vis mental illness and mental health in future iterations of the project.

## 1.6 Solution Overview

In order to identify the 100 individuals who are most likely to recidivate among the formerly incarcerated population, we performed both descriptive and predictive analysis using supervised machine learning techniques. Descriptive analysis allowed us to see the demographic distribution of our target population over time and the relationship between potential features, especially for features relating to mental health and criminal justice. In turn, this helped us to target our efforts in generating relevant features, which the machine learning model used to generate predictions. Further, exploring the distribution of inmate booking information enabled us to identify the population that should be included in the model (e.g., by determining that data are sparse for more temporally distant bookings). We discuss these details further in the data section.

In our predictive analysis, our goal is to build the most efficient, effective, and equitable model that, each month, identifies the 100 individuals with the highest risk of going back to jail in the next year. Hence, our labels are whether or not a person will recidivate within the next year (365 days) of the model prediction date. We will then use the features identified from the descriptive analysis and set the temporal validation for each training and testing set, described at length in the temporal validation section of this report.

We used the Decision Tree Classifier and Random Forest as initial “simple-run” models. Our model selection process drew upon the logic that out of the 100 people who have the highest risk scores we want to select individuals who, without the outreach/treatment from JCMHC, would have otherwise recidivated. In other words, we analyze the precision rate of the 100 people with the highest risk scores. Since there is a cost associated with false negatives, we also consider recall rate in assessing model performance. In addition, we will measure our models against the base rate, which is defined as the total number of people who went back to jail within one year, divided by the total number of formerly incarcerated people in the cohort. We also compare our model to three simple baselines: 1) A list of the 100 people by top number of previous bookings, 2) A list of 100 people with the highest number of JCMHC visits, and 3) A list of 100 people with the highest average time between bookings.

# 2 Data

## 2.1 Data Description

Johnson County provided administrative data from its mental health center, jail system, police arrests, and ambulance runs. The data summary of each database are as follows:

The Justice Information Management System Department (JIMS) provides the records on inmate demographic and background information related to mental health associated during the incarceration such as pretrial risk assessment and the mental health referral to JCMHC for mental health care services based on the BJMHS results. The database also provides inmates' incarceration history records, e.g., booking and release information as well as charges and bail associated. In summary, the data covers the information from early 2000s to 2019 with approximately 180,000 inmates and 558,000 cases in total.

The JCMHC database is recorded at the patient level (unique patient ID). The data includes the demographic and socioeconomic background of each patient, including information on race, sex, and income. It also includes mental health assessment data: the Child and Adolescent Functional Assessment Scale (CAFAS) and Daily Living Activities– 20 (DLA-20) score. JCMHC also provide crisis and emergency call information, including risk associated, e.g., family conflicts, depression, and suicide risk.

Local police departments mainly provide the arrest and charge information, including date, location, status, and type of arrest and charges. Lastly, the Medical Emergencies Department (MEDACT) dataset contains the information about patients' encounters with emergency ambulance services, such as emergency type and detail of emergency, i.e., date and time of emergency occurrence, medication provided, and associated location.

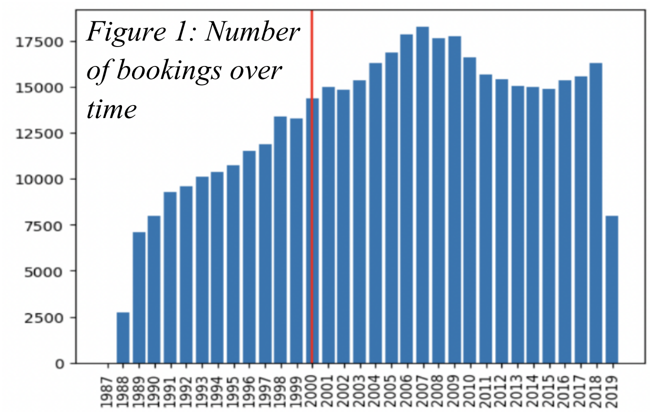


*Figure 1: Number of bookings over time*

## 2.2 Data Exploration

Here, we first look into the distribution of the number of bookings each year over time from 1987 to 2019. Prior to 2000, denoted by the red vertical line, the data are quite sparse and incomplete. Thus, we included in target population individuals who were formerly incarcerated from 2000 onward for our machine learning models.

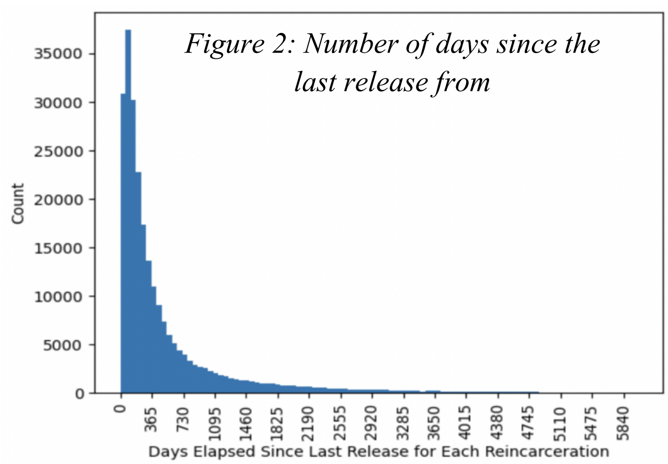


*Figure 2: Number of days since the last release from*

We investigated the distribution of days elapsed since release for all reincarceration bookings. We found that most reincarceration occurs within 2 years of release, accounting for roughly 78.52% of all bookings of people who have previously been incarcerated. Hence, we targeted the individuals who were released from Johnson County jail in the last two years in our analytical formulation.

In addition, we visualized how many times a person was booked into jail against how many times that person received mental health services to see the relationship between reincarceration and mental health history. We did not see a significant pattern between mental health history and reincarceration. For this reason and others detailed above, we did not include mental health in our problem formulation.

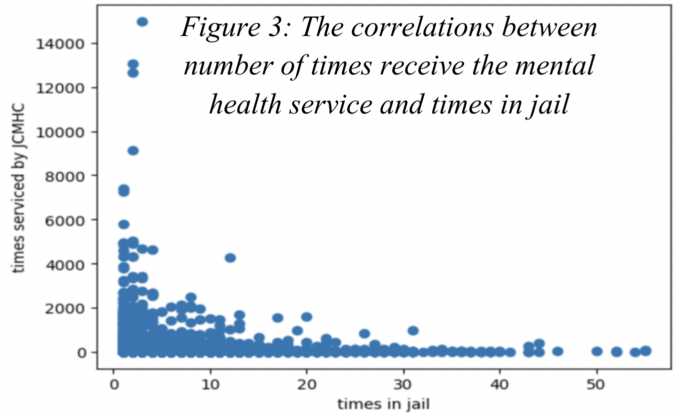


*Figure 3: The correlations between number of times receive the mental health service and times in jail*

# 3 Analysis

## 3.1 Methods

The methods that were used to build the analysis were employed in conjunction with the analytical formulation. Since the problem was formulated as identifying those who were released from the Johnson County Jail in the last two years (of the as-of-date), and who were not currently in jail, the cohort definition followed that logic. Therefore, each row in the final dataset represented a former inmate who had previously been incarcerated, but was released within the last two years. Similarly, the outcome variable was formulated as a categorical binary variable where it was equal to 1 if the former inmate was reincarcerated within one year of the as of date, and 0 if the former inmate was not reincarcerated within one year. All trained machine learning models were compared against both the base rate and three baselines (explained below). This is done to make sure that the ML models perform better than a simplistic, and cheaper method of identifying at-risk individuals.

## 3.2 Temporal Validation

Because of the temporal nature of this problem, we cannot use a traditional k-fold cross validation technique. The problem attempts to predict if a person will return to jail in the future. Therefore, to evaluate the model, the validation set must be comprised of data from a future period, while the training set consists of historical data. Each training set contains two years worth of data. Further, the training set is built with a one year rolling window such that there are multiple subsets of data occurring every 6 months in one training set. The validation set is a one year duration that starts one year after the latest date from its corresponding training set. The choice of one year for each training subset as well as the validation set was made under the assumption that any intervention to a person's mental health would take one year to make an impact on their likelihood of recidivism. This training validation set method is repeated in a two year rolling window with each set being built one year apart. The training data starts from January 1st, 2009 and ends January 1st, 2017 while the testing data starts January 1st, 2010 and ends January 1st, 2018. Although data were available prior to 2000, the data become sparse and unreliable. Therefore, any feature data must come from January 1st, 2000 or later.

## 3.3 Model Types and Hyperparameters

Because the outcome variable is a binary indicator, all of the machine learning models that were trained and tested were classifiers. The types of models trained were Random Forest Classifiers, Decision Tree Classifiers, Scaled Logistic Regression, Gradient Boosting Classifier, and a Neural Network.

## 3.4 Features

The features we used can be categorized into demographic information, criminal justice history, and mental health history of our cohorts. Demographic information includes age on the date of prediction, race, sex, marital status, income, and employment status. Criminal justice history includes the count of bookings in the past, average time elapsed between one release and the next incarceration, time elapsed since last release, age at the time of the first booking, number of times referred to receive mental health center while in jail, count of different types of charges, count of different flags raised in pretrial assessment (e.g., drug/substance abuse flag), count of different arrest types. Finally, mental health history includes count of admissions into the mental health facility, DLA and CAFAS scores, count of services received that are related to crisis events, count of different call types (e.g., crisis call, referral call), count of times participated in medication clinics with JCMHC, count of times engaged with emergency services, count of different triage codes (e.g., yellow code, green code, red code), count of different primary impressions for calls (e.g., anxiety, drug, alcohol), and whether the entity has been evaluated to be at risk of conducting violence towards themselves/others.

For each feature, we used different multiple past timeframes to generate. The detailed configuration can be found in the appendices.

## 3.5 Evaluation Metrics and Baselines

The problem formulation called for 100 interventions to be made each time the model is run. Therefore, the evaluation metric that was produced each time a model was trained was the precision at the top 100 people that the model had identified. Precision is defined as the number of people that the model had correctly predicted would return to jail, divided by the total number of former inmates that the model predicted would return to jail. In this case, the precision at 100 metrics answers the question "Out of the 100 people that were predicted to return to jail, how many of them actually went back to jail?". We used precision at top 100 as the selection metric because of the resource constraint faced by JCMHC. Since they have the budget and resources to only intervene on 100 people every month, we wanted to make sure that those 100 people are actually at high risk of recidivism and need the intervention.

To have an effective metric to measure against, there were four main baseline methods. The first was the *Base Rate* which is defined as the total number of people who went back to jail within a year, divided by the total number of former inmates in a given set. Depending on the evaluation start time of the validation set, the base rate hovered approximately between 0.19 and 0.21.

Next, three common-sense baselines were created. All three of these ranked individuals based on a metric and labeled the top 100 people as at risk for recidivism (the ranking algorithms are ascending). The first common-sense baseline ranked the former inmates by the number of bookings each individual had. The next baseline ranked people by the number of times they had visited the JCMHC. The last baseline ranked people on the average number of days between an individual's most recent release and their next booking.

Below is a table showing the precision at the top 100 people from the three baseline ranking methods, followed by the top performing logistic regression model, and top performing random forest model.

| Baseline | Lowest Precision@100 | Highest Precision@100 |
|---|---|---|
| *Base Rate* | *0.19* | *0.21* |
| **Number of bookings** | **0.42** | **0.57** |
| **Number of visits to JCMHC** | **0.31** | **0.41** |
| **Days between release and next booking** | **0.15** | **0.29** |
| *Scaled Logistic Regression Model* | *0.58* | *0.86* |
| *Random Forest Model* | *0.63* | *0.84* |

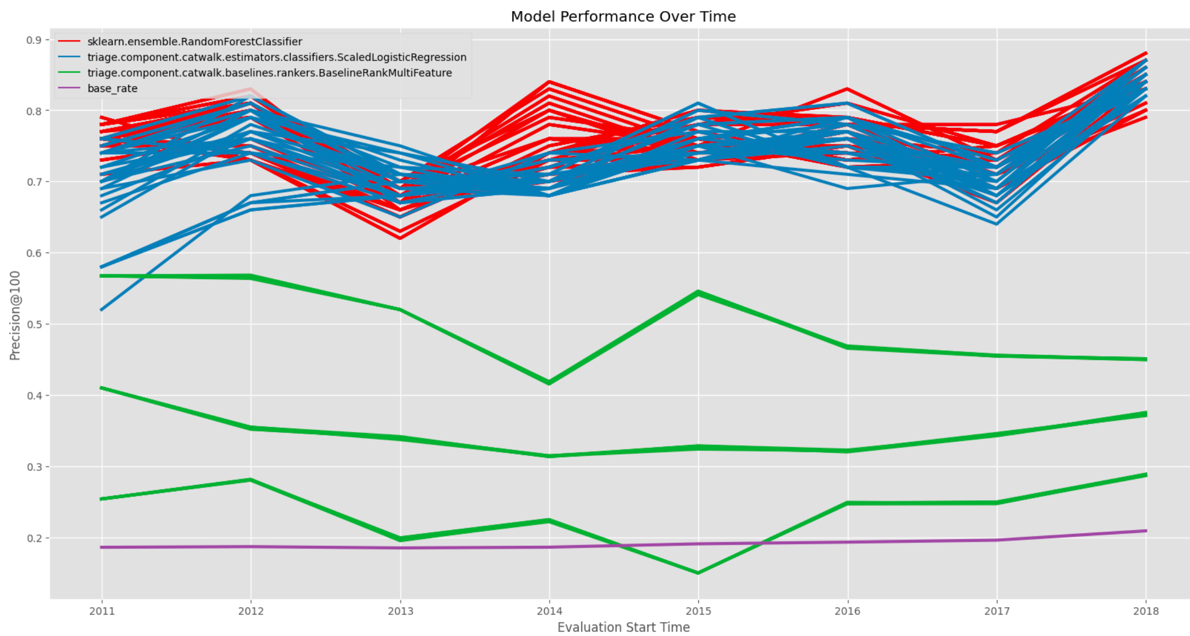


*Figure 4: Model performance over time graph showing the base rate (purple), the three baselines (green), and the logistic regression and random forest models (blue and red)*

## 3.6 Results

Based on our selection criteria of precision at the top 100 people, the top two performing models were a Scaled Logistic Regression model, and a Random Forest Model. The logistic regression was tuned with a *C* (inverse of regularization strength) of 0.01 and an L1 penalty term. The random forest model has 100 estimators, a max depth of 100 and needs a minimum of 1,000 samples to split. As shown in both the table and performance graph above, the logistic regression model had a top precision of 0.86 while the random forest model had a top precision of 0.84. The two models traded as the top performer depending on the evaluation start time.

The Precision and Recall at k (PR-k) curves below show the two models' precision (true positives divided by predicted positives) in blue and recall (true positives divided by actual positives) in red. As more of the population is being predicted, it becomes harder for both models to maintain precision. The precision will eventually converge with the base rate if we predict 100% of the population to be true; that is, we will only be correct for 21% of the population. The opposite is true for recall: if we predict 100% of the population to be true, then all 100% of the actual at-risk population will be caught by the model. Since we are only focusing on the top 100 people, we are only looking at the far right portion of the graphs. If the resource constraint was higher, it is possible that the best model would also be different.

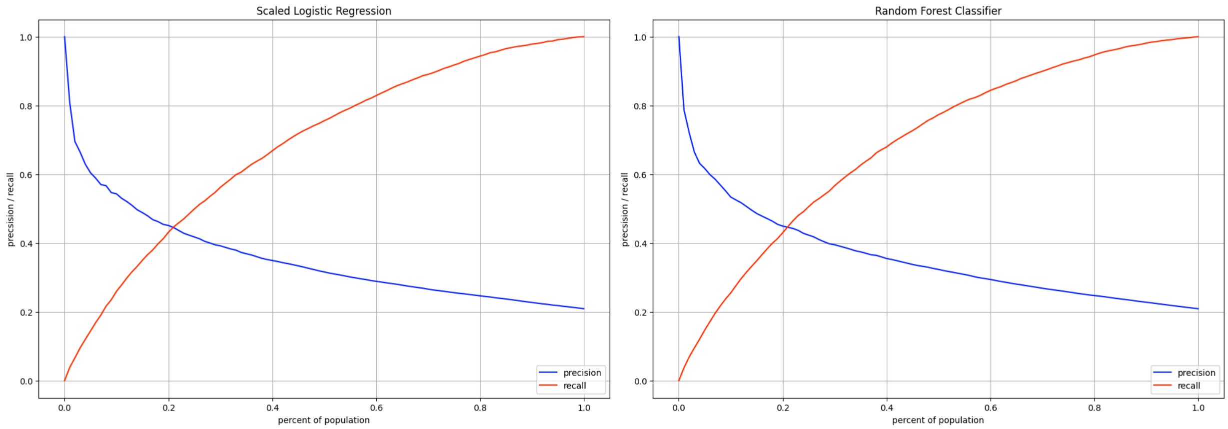


*Figure 5: PR-k curves for the top two performing models*

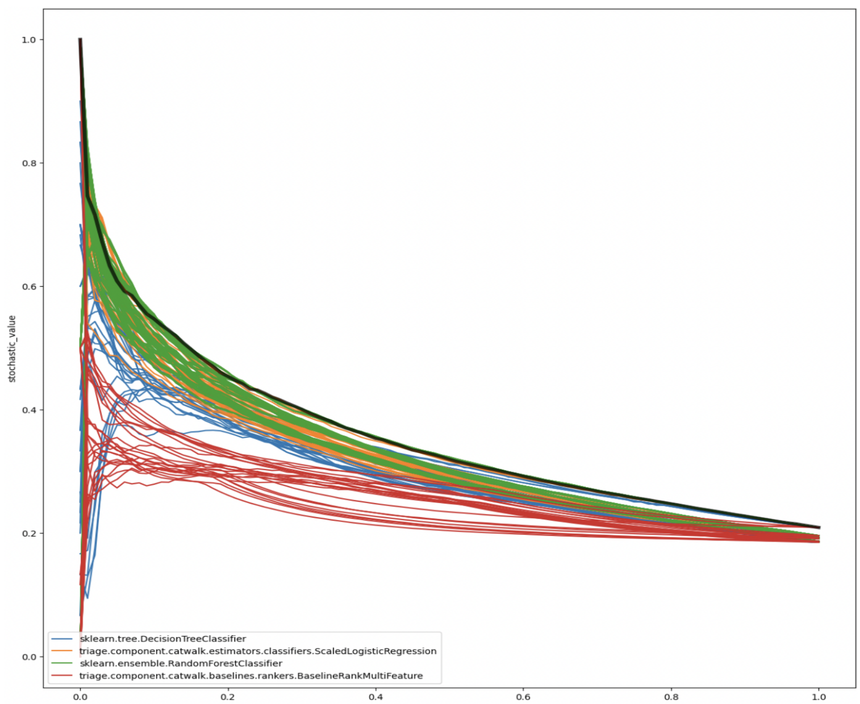


*Figure 6: Precision curves for all models in last 2 runs (top logistic regression model in black)*

The above graph plots the precision curves for all the models in our last 2 runs. The most recent model of our best model group is highlighted in black. We can see that this model performs well in terms of the precision at top 100 people. However, as our threshold increases the performance of other models and other model groups will be higher than this model, which indicates that if more resources are allocated to Johnson County to address the mental health issue of the reincarcerated people, the best model may be different.

Both models had similarities in terms of important features. Looking at the top 15 most important features for both models, the number of bookings over 6 months, 1 year, 2 years, and all time shows up for both models. The models have different top features, however, with the logistic regression model flagging the number of charges in a 2 year period and the random forest using the number of days since an individual's release as their respective top features. Of note, both models do not have any mental health features, or simple demographics, like age, in their top 15 features.

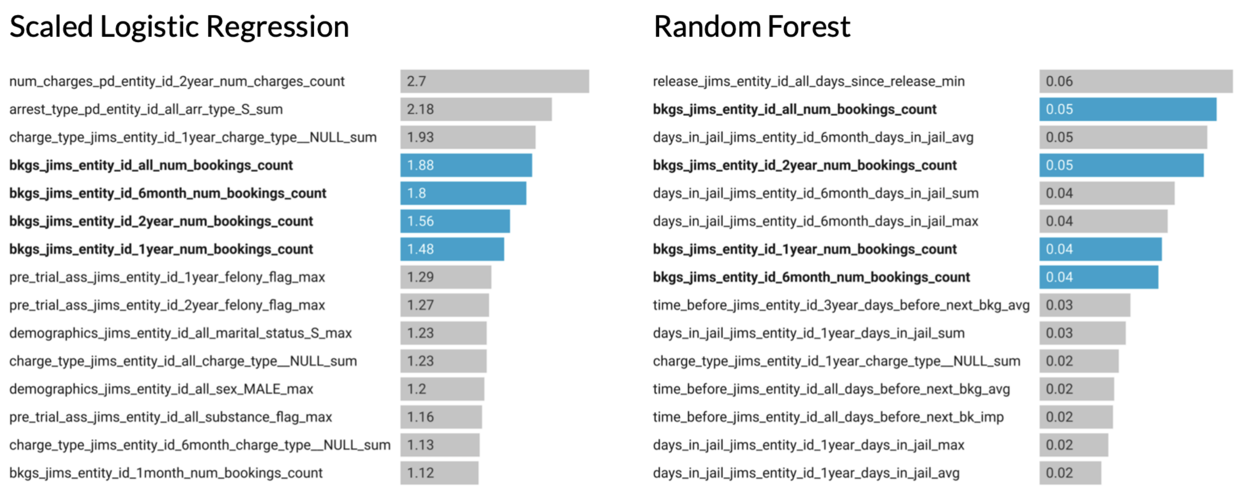


*Figure 7: Feature Importances for the top two performing models*

When looking at the 100 people that the best performing logistic regression model identified as at-risk, the people skew towards being male and in younger age ranges. Additionally, only people who are Black or people who are white are identified as being at risk for recidivism by the model.

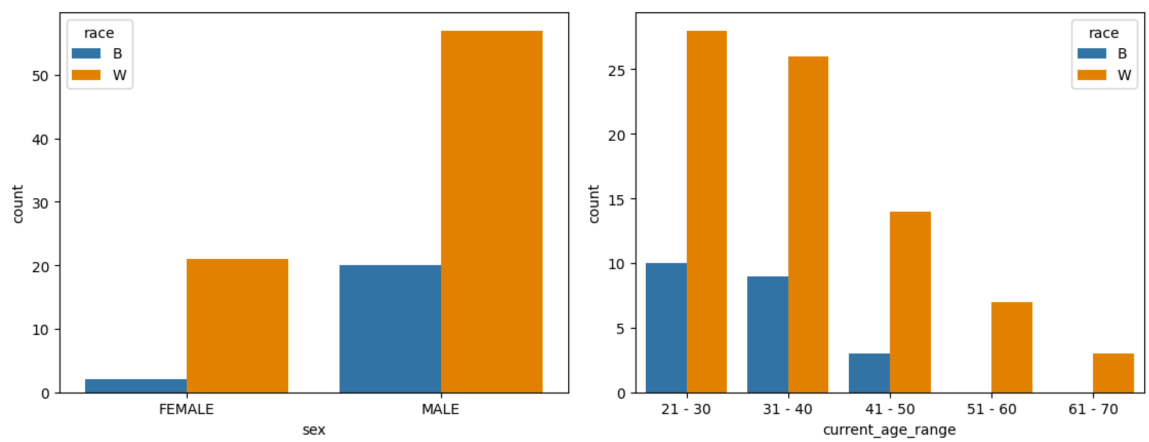


*Figure 8: Breakdown of sex and race (right) and breakdown of current age and race (left)*

# 4 Model Limitations and Fairness Analysis

## 4.1 Limitations and Caveats

JCMHC aims to break the repeat cycles of incarceration due to untreated mental health needs for incarcerated and formerly incarcerated people. The causal path of the intervention assumes that people who the JCMHC reaches out to proactively to provide mental health services are in fact those who are at risk of going back to jail if they are not provided mental health services. To that end, the biggest limitation we face is that mental health needs are not explicitly baked into the analytical structure of the cohort or labels. The labels simply indicate whether someone goes back to jail or not within a year of the prediction. The cohort contains only those people who have been released from jail in the last 2 years and are not in jail currently. This limitation may affect the effectiveness of the intervention, irrespective of model precision, if those that the model flags at being high risk are at risk of going back to jail, but not necessarily because of untreated mental health conditions.

We analyze the difference in mental health outcomes in the cohort between people that go back to jail within a year and those who don't. To assess the mental health outcomes, we consider two features — the number of times someone has been admitted in JCMHC and the number of times someone has been referred from jail through the screening form. Between people who go back to jail (outcome label = 1) and those who don't (outcome label = 0), the difference in average number of times someone has been referred from jail is negligible. However, those who go back to jail are roughly twice as likely to have been admitted to JCMHC as compared to those who don't. Further, the best models predict 100 people who have the highest risk of going back to jail. Between those for whom the predictions are correct vs. those for whom it's incorrect, the correctly predicted people (i.e., those who actually go back to jail) are more likely to have mental health needs. Finally, comparing the top 100 risk scores vs. everyone else, we see that the mental health needs for the top 100 are higher than others.

In spite of these differences, mental health features do not appear as the most important features in almost all models. This could be due to the following reasons:

- Criminal justice features are correlated with mental health features, so those features, to some extent, already capture unmet mental health needs
- The features don't appear to be important because the analytical structure doesn't explicitly account for mental health needs
- Mental health features are not good predictors of who would go back to jail.

In future variations of the model, we would like to incorporate mental health into the structure to see if these results change. Additionally, during the field trial, it would be useful to gather information on the actual mental health needs of people who are predicted as high risk. That would inform how big of a limitation not incorporating mental health into the analytical structure is.

## 4.2 Bias and Fairness Audit

Since the eventual goal of this machine learning activity is to help people experiencing the cycle of reincarceration, it is important to evaluate who are the people the model is predicting to be at high risk and which groups would get the intervention based on the model's results.

The composition of different groups of people experiencing reincarceration can be different from their distribution in the population. Hence, this project's aim is not to ensure that the top 100 high risk people predicted by the model are similar in composition to their occurrence in the population. However, as the equity goals mention, the project aims to ensure that certain protected groups who may be at disadvantage in the society, especially from a criminal justice perspective, are not unfairly treated by the model.

To that end, we care about two specific metrics: False discovery rate and True positive rate. False discovery rate means that of all the people who were predicted to be at high risk of going back to jail, what proportion were incorrectly predicted. True positive rate means that of all the people who actually go back to jail, what proportion was our model able to detect. In this project, we identify three main protected groups to analyze: Black people (race), women (sex), and people who were flagged as needing mental health support (from the in-jail screening form). The false discovery rates and true positive rates for these groups were compared against a base group in each category: White people (race), men (sex), and people who were flagged as not needing mental health support.

For the best performing models (Random Forest and Scaled Logistic regression based on average precision at top 100), we find that the model makes lower false discoveries and similar true discoveries for Black people as compared to White people. These models also make lower false discoveries and higher true discoveries for people flagged as needing mental health support as compared to those who do not. However, the models make higher false discoveries for women as compared to men.

Agreeing on what extent of disparity in false discoveries and true detection between the protected groups and base groups are we comfortable with is a complex value judgment. However, it is encouraging to see that the best models don't discriminate against Black people and people needing mental health support. We need additional investigation to understand why the models are unfair towards women as compared to men. The suggested next steps would include diving deeper into how the model predicts risk scores for women, and more importantly, gaining an understanding from the field trials on whether we are missing women who are at high risk and should have been prioritized for the interventions. The appendix contains additional information on comparative numbers between the protected and base groups for all the models.

# 4 Policy Recommendations and Future Work

While our models effectively predict whether or not an individual will recidivate within 1 year, they may not identify individuals who also have mental health conditions. That is, our treatment may not be effective for those to whom it is proffered. As discussed before, there were many reasons that informed the modeling choice to exclude mental illness from our cohort definition. However, in the future, we would like to actively incorporate mental health in our cohort definition, to ensure that the JCMHC's outreach can be as effective as possible.

To this end, we suggest that the JCMHC expand and assess the BJMHS, the screening tool to identify if individuals in jail should be referred to its services. The Government Accountability Office noted that individuals should be screened for mental illness early in the booking process, with continued surveying throughout their time in the carceral system.[31] The BJMHS was introduced within the last decade, and the first recorded referral via the BJMHS to the JCMHC occurred on November 7, 2016. Between then and the end of June 2019, there have been 10,000 referrals from the tool. In fact, over one in four unique inmates were flagged for referral over this timeframe. Thus, the screening has the potential to be an effective tool both in redefining the machine learning cohort, as well as in providing more nuanced, targeted treatment and outreach approaches.

First, the JCMHC should assess the BJMHS for bias. Individuals may be predisposed to incorrectly report information out of fear of perceived repercussions. To address this, the JCMHC should partner with a surveying firm or statisticians from the nearby University of Kansas to assess the survey tool. Next, the JCMHC should work with the firm to expand the BJMHS questionnaire, which will allow the JCMHC to collect more data. Because mental health is nuanced and complex, the JCMHC should offer personalized outreach and care plans to the 100 individuals identified each month. The expanded survey will enable the JCMHC to offer more nuanced treatment plans to patients, ensuring that its initial outreach to clients will meet them where they are. This is of particular importance, as the quality of the initial outreach very well could determine whether or not an individual chooses to uptake long-term care with the JCMHC. Lastly, the JCMHC should complete follow-up surveying throughout individuals' time in jail, to identify individuals whose mental health may suffer during jail, as well as to offer more personalized care and outreach, as discussed above.

In addition to the above recommendations, we also propose a field trial using the results of our top-performing machine learning model. Here, we present the JCMHC with three options: a randomized control trial measuring the causal effect of outreach on recidivism rates, an instrumental variables approach measuring the causal effect of longer-term mental healthcare on recidivism rates, and a regression discontinuity design measuring the causal effect of either outreach or longer-term mental healthcare on recidivism rates. We detail the strengths and weaknesses of each approach below and invite the JCMHC to weigh them, selecting the field trial design that best aligns with their goals and ethics.

First, we propose a randomized control trial completed over the course of one year (12 months), which is the timeframe proposed for all field trial designs. Each month, our model will assign reincarceration risk scores to formerly incarcerated people, which will be used to rank individuals from highest to lowest risk. The JCMHC has the capacity to complete outreach to 100 individuals each month. Thus, we will randomly select 100 individuals from the top 400 highest-ranked individuals. We restrict the sample to the top 400 individuals because the treatment, once validated, will be applied to individuals with the highest recidivism probability scores — that is, including all individuals in the sample would contradict how it will be operationalized in practice. Among the people with the highest recidivism risk scores for the top-performing logistic regression model, recidivism risk scores ranged from 0.92 to 0.62. The JCMHC will then complete initial, personalized outreach to these individuals. In this approach, the JCMHC would stop here, aggregating outcomes over one year to see whether individuals who received outreach are less likely to recidivate than those who did not receive outreach.

While the randomized control trial may be the simplest way to establish a causal pathway of the intervention on recidivism rates, it is not without serious considerations. We are primarily concerned around the ethics of withholding outreach to individuals who, given the output of the model, would have otherwise been prioritized for care. This is a concern not to be taken lightly: it could be the difference between a person's going back to jail or their staying out of jail. In addition, a randomized control trial may prove costly, requiring the JCMHC to maintain contact with 4,800 individuals over the course of 24 months. This may also pose a burden on staff time and/or force the organization to hire additional caseworkers. For these reasons, the JCMHC will likely want to contract with a research firm or partner with a local university to complete data collection and analysis on behalf of the organization.

Second, we propose an instrumental variables analysis that measures the impact of longer-term mental health care — defined here as 4 continuous months of meeting with a mental health practitioner (therapist or psychiatrist), participating in a mental health support group, or taking medication for a mental illness — on recidivism rates. This design is similar to the randomized control trial, but takes it one step further, as we suspect that a single instance of outreach may not provide enough support to alter someone's likelihood of recidivating. Thus, we will randomly assign treatment to 100 individuals from the top 400 highest-ranked individuals. Our rationale for sample restriction holds from above. However, some individuals will likely not complete 4 full months of treatment. Thus, an instrumental variables approach is required. The instrument is being selected for services, under the assumption that being selected for services impacts recidivism only through the provision of services. Again, we suggest that the JCMHC contract with a firm in carrying out this analysis.

Our concerns for this method, particularly around ethics, echo those of the randomized control trial. This method will prevent individuals who have a higher risk of recidivating, as assessed by our model, from receiving treatment. In addition, the instrumental variables approach may have limited external validity if individuals who comply with treatment are systematically different from other individuals who would receive treatment. However, we should note that the JCMHC's "lever" to break the cycle of incarceration requires that patients elect to participate in its services; that the findings of the instrumental variables analysis are limited may not be particularly concerning. Moreover, this is why we emphasize the importance of personalized care and outreach plans: to ensure that the organization's approach is sensitive to individuals' personal, cultural, and social needs.

The third method we propose is a regression discontinuity design. This approach differs from the previous two in that the JCMHC would proffer services to the 100 individuals who have the highest recidivism scores. The JCMHC could complete either a single instance of outreach, or a longer-term treatment plan, based on its capacity. (If the JCMHC chooses longer-term treatment plans, we will implement a regression discontinuity design with instrumental variables.) We will then compare recidivism outcomes around the 100 cutoff (those just above, who receive treatment; and those just below, who do not) one year after treatment. As is standard econometric procedure, the bandwidth of individuals to choose around the cutoff can be determined using Mean Squared Error optimal selection methodology.[32]

The regression discontinuity design perhaps poses a less ethically dubious method than the randomized control trial and instrumental variables method. That is, services will be offered to the individuals with the highest risk scores; we will not withhold services to people who would have otherwise received them, in the absence of an experimental trial. At the same time, the findings from this design only apply to individuals with risk scores right around the 100 cutoff. In other words, the treatment effect is local to people with risk scores around the cutoff.

If the JCMHC intends to complete a single instance of outreach as its intervention, the randomized control trial is the most relevant method. However, if the JCMHC wants to work with patients over a longer time horizon, then the instrumental variable method with cohort randomization will prove more effective. While we defer to the JCMHC's expertise in this matter, we favor the instrumental variables method, under the assumption that mental health evolves and cannot be treated as a point-in-time phenomenon. The regression discontinuity design is a quasi-experimental method that attenuates some of the earlier ethical concerns and can be adapted either for outreach or longer-term care. However, its generalizability may be limited, as compared to the previous two methods.

We invite the JCMHC to consider these policy recommendations and proposed field trial designs. We are grateful for the opportunity to collaborate with the JCMHC and consider it our privilege to have worked on this project. We admire the mission of this project and look forward to further discussion.

# 5 Appendices

## 5.1 GitHub Links: Configuration and Run Files

Please reference our label configuration YAML file here. Please reference our Python run file here. Feel free to explore our repository, linked here.

## 5.1 Temporal Configuration

The parameters for temporal configuration are as follows:

Feature start time: 2015-01-01 and feature end time: 2019-01-01
Label start time: 2009-01-01 and label end time: 2019-01-02
Test durations: 1day
Max training histories: 2year
Training and testing as of date frequencies: 6month
Training and testing label timespans: 1year

The output is shown below:

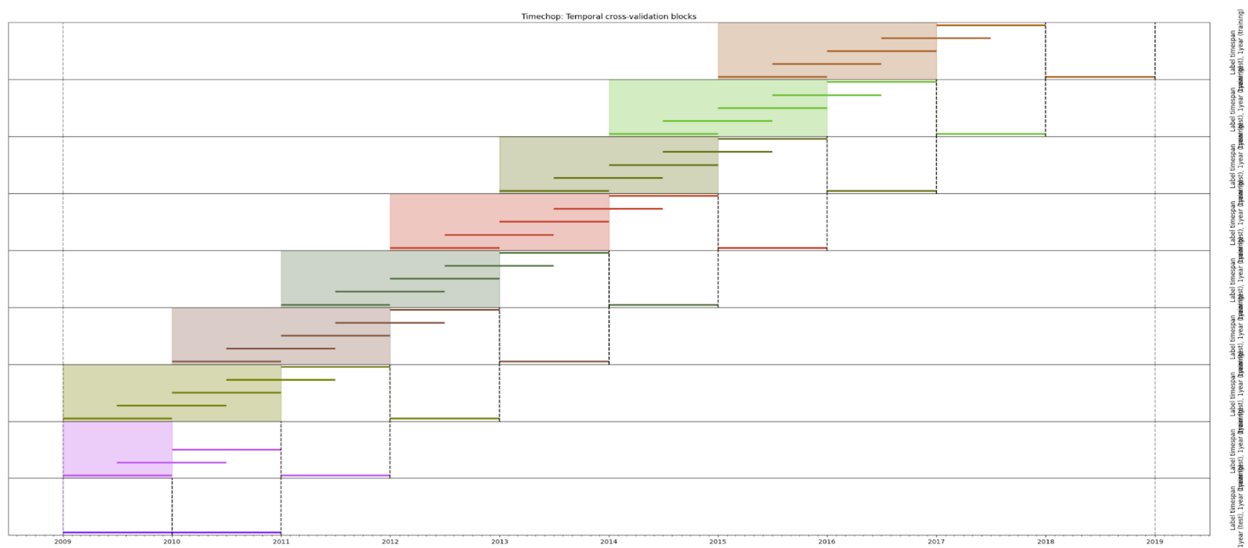


*Figure 9: temporal configuration*

Starting from the most recent train-validation pair, the sets are as follows:

| Train-Validation Pair ID | Train Set | | | | Validation Set | | | |
|---|---|---|---|---|---|---|---|---|
| | Earliest row as of date | Latest row as of date | Start date for labels | End date for labels | Earliest row as of date | Latest row as of date | Start date for labels | End date for labels |
| 1 | 2015-01-01 | 2016-12-31 | 2015-01-01 | 2017-12-31 | 2018-01-01 | 2018-12-31 | 2018-01-01 | 2018-12-31 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | 2014-01-01 | 2015-12-31 | 2014-01-01 | 2016-12-31 | 2017-01-01 | 2017-12-31 | 2017-01-01 | 2017-12-31 |
| 3 | 2013-01-01 | 2014-12-31 | 2013-01-01 | 2015-12-31 | 2016-01-01 | 2016-12-31 | 2016-01-01 | 2016-12-31 |
| 4 | 2012-01-01 | 2013-12-31 | 2012-01-01 | 2014-12-31 | 2015-01-01 | 2015-12-31 | 2015-01-01 | 2015-12-31 |
| 5 | 2011-01-01 | 2012-12-31 | 2011-01-01 | 2013-12-31 | 2014-01-01 | 2014-12-31 | 2014-01-01 | 2014-12-31 |
| 6 | 2010-01-01 | 2011-12-31 | 2010-01-01 | 2012-12-31 | 2013-01-01 | 2013-12-31 | 2013-01-01 | 2013-12-31 |
| 7 | 2009-01-01 | 2010-12-31 | 2009-01-01 | 2011-12-31 | 2012-01-01 | 2012-12-31 | 2012-01-01 | 2012-12-31 |
| 8 | 2009-01-01 | 2009-12-31 | 2009-01-01 | 2010-12-31 | 2011-01-01 | 2011-12-31 | 2011-01-01 | 2011-12-31 |
| 9 | 2009-01-01 | 2009-12-31 | 2009-01-01 | 2009-12-31 | 2010-01-01 | 2010-12-31 | 2010-01-01 | 2010-12-31 |

## 5.1 Model and Hyperparameters Tested

The details of hyperparameters tested in each model are as follows:

| Model Type | Hyperparameters Tested |
|---|---|
| Random Forest Classifier | n_estimators: [100, 1000, 5000, 7000]<br>max_depth: [100, 300, 500]<br>min_samples_split: [100, 1000] |
| Decision Tree Classifier | max_depth: [5, 10]<br>max_features: None<br>min_samples_split: [50, 100] |
| Scaled Logistic Regression | C (inverse of regularization strength): [0.1, 1, 0.01]<br>penalty: [‘l1’, ‘l2’] |
| Gradient Boosting Classifier | n_estimators: [1000, 2500]<br>subsample: [0.5, 0.8]<br>criterion: squared error<br>min_samples_split: 300<br>max_depth: 300<br>max_features: [‘sqrt’, ‘auto’] |

| | |
|---|---|
| Neural Network MLP Classifier | hidden_layer_sizes: [100, 75, 50, 25]<br>activation: [‘logistic’, ‘relu’]<br>learning_rate: [‘adaptive’]<br>max_iter: 500 |

## 5.2 Features

The features used as the attribute input to the models are compute over defined time span as follows:

| Feature Category | Feature Name | Aggregate metrics | Time Span |
|---|---|---|---|
| Demographic Information | Age on the date of prediction | max | all |
| | Race | max | all |
| | Sex | max | all |
| | Marital status | max | all |
| | Income | max, avg | 1year, 2year, all |
| | Employment status | sum | 1year, 2year, all |
| Criminal Justice History | Count of bookings in the past | count | 1month, 6month, 1year, 2year, all |
| | Average time elapsed between one release and the next incarceration | avg | all, 3year |
| | Time elapsed since last release | min | all |
| | Age at the time of the first booking | min | all |
| | Number of times referred to receive mental health center while in jail | sum | 6month, 1year, 2year, all |
| | Count of different types of charges | sum | 6month, 1year, 2year, all |
| | Count of different flags raised in pretrial assessment | max | 1year, 2year, all |
| | Count of different arrest types | sum | 1year, 2year, all |

| Feature Category | Feature Name | Aggregate metrics | Time Span |
|---|---|---|---|
| | Count of charges | count | 2year, all |
| | Number of days spent in jail | Max, avg, sum | 6month, 1year, 2year, all |
| Mental Health History | Count of admissions into the mental health facility | count | 6month, 1year, 2year, all |
| | DLA and CAFAS scores | min, avg, max | 1year, 2year, all |
| | Number of services received that are related to crisis events | sum | 6month, 1year, 2year, all |
| | Count of different call types (e.g., crisis call, referral call) | sum | 6month, 1year, 2year, all |
| | Count of times participated in medication clinics with JCMHC | sum | 6month, 1year, 2year, all |
| | Count of times engaged with emergency services | count | 6month, 1year, 2year, all |
| | Crosstab of different emergency code types (e.g., yellow code, gray code) | sum | 6month, 1year, 2year, all |
| | Crosstab of different emergency code types (e.g., patient being anxious, patient on drug) | sum | 6month, 1year, 2year, all |
| | Whether the entity has been evaluated to be at risk by the mental health center | sum | 2year, all |

## 5.3 Limitations and Caveats

The two graphs below show the difference in mental health needs in the cohort, grouped by the true outcome labels. We see that there is a difference in mental health needs, but not as large as a criminal justice feature such as number of bookings

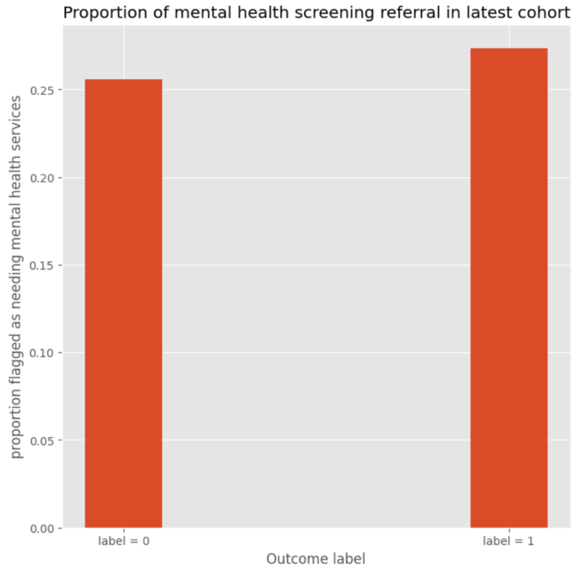


*Figure 10: proportion of mental health screening (BJMHS) between true outcome label*

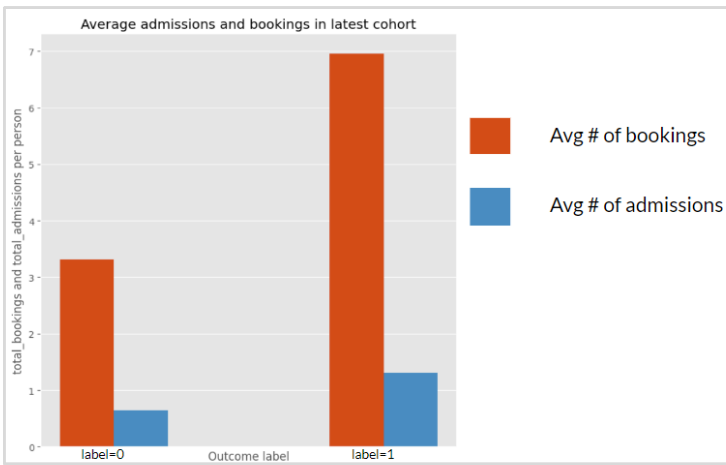


*Figure 11: proportion of average number of bookings and admission between true outcome label*

The graphs below show the difference in mental health outcomes between the correct and incorrect predictions in top 100 high risk scores. Again, we see that those who actually went back to jail have a slightly higher mental health need (higher for average number of referrals as compared average number of admissions in JCMHC).

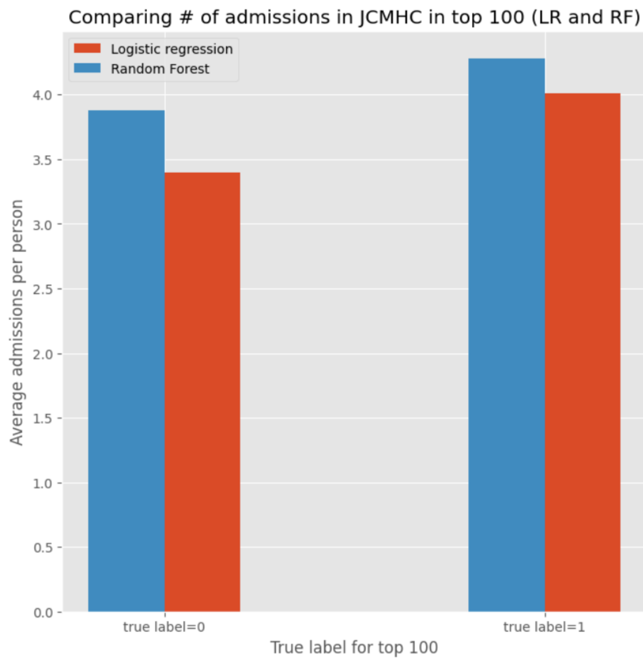


*Figure 12: mental health outcomes between correct and incorrect predictions at top 100 at risk of recidivism*

Finally, we see from the graphs below that there is a starker difference in mental health needs between the top 100 high risk people vs the rest (although the difference is lower as compared to criminal justice features)

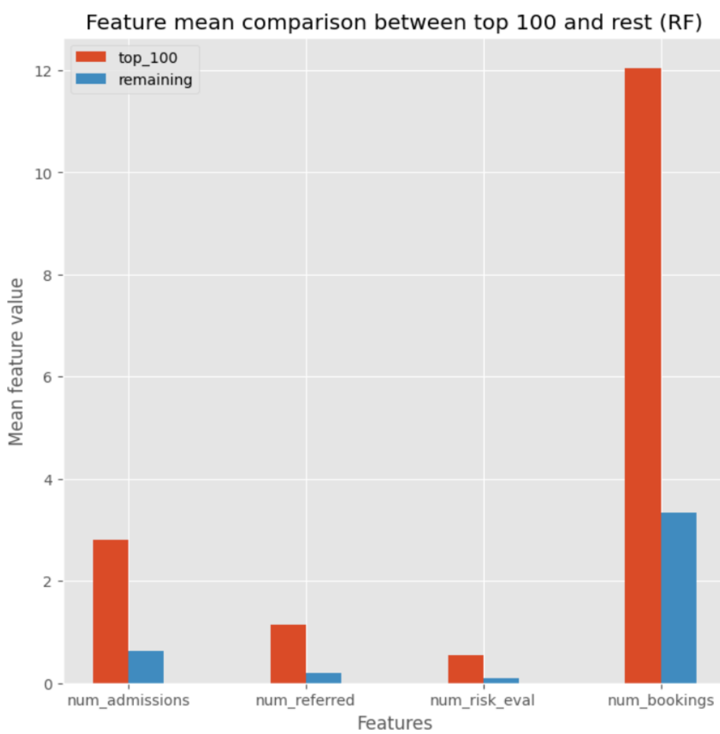


*Figure 13: Comparative mean value of features between top 100 most at risk of recidivism and the rest for random forest model*

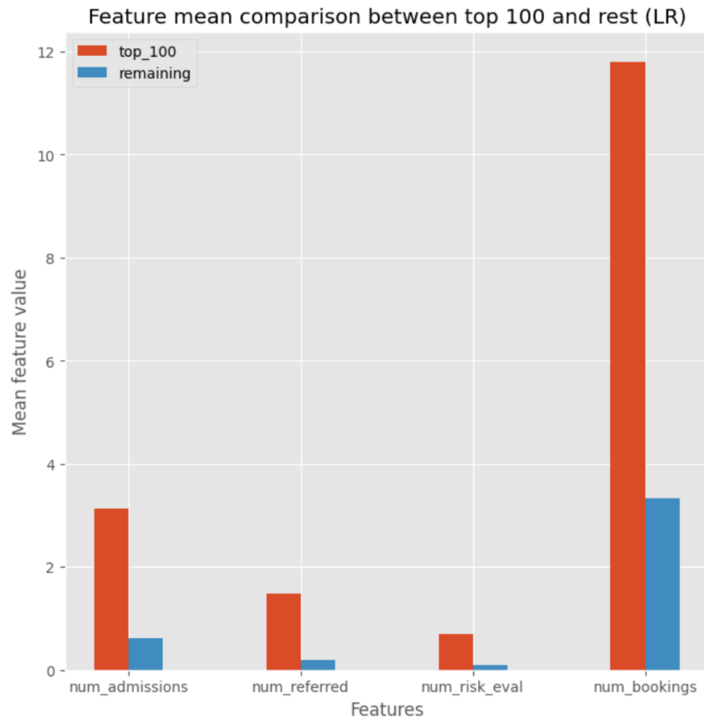


*Figure 14: Comparative mean value of features between top 100 most at risk of recidivism and the rest for logistic regression model*

## 5.4 Bias and Fairness Analysis

The graphs below show false discovery rates and true positive rates against the precision of all models, for the protected group as compared to the base group. Each dot represents a unique model. The false discovery rate (fdr) ratio indicates the fdr ratio for the protected group as compared to the base group. An fdr ratio of less than 1 indicates that the model makes lower false discoveries for the protected group. A similar inference can be drawn for the true positive rate ratio.

### 5.4.1 Comparison of Black to White people

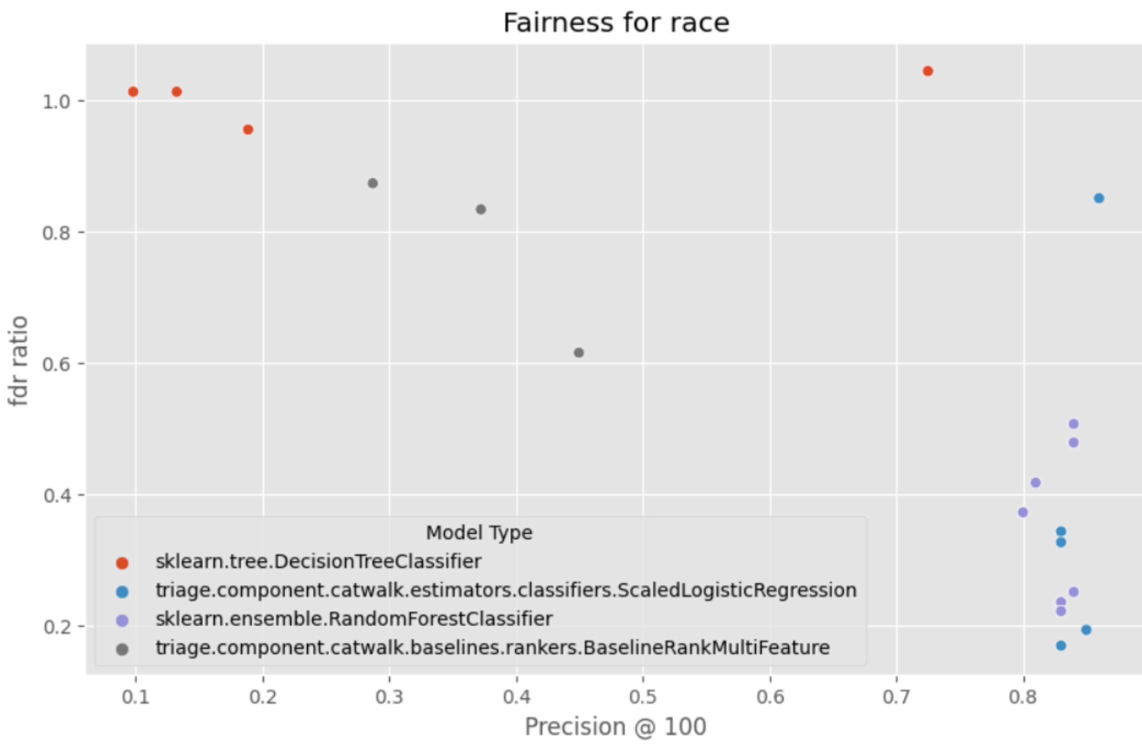


*Figure 15: False discovery rate of precision at 100 for race comparison*

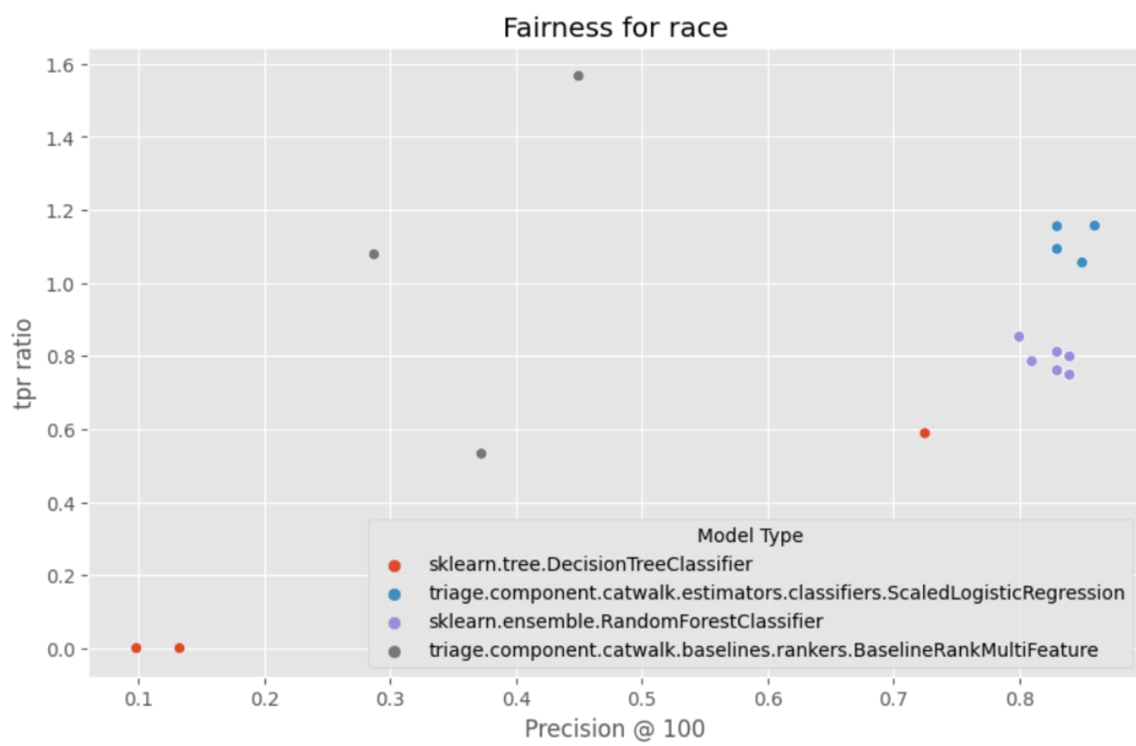


*Figure 16: True positive rate of precision at 100 for race comparison*

5.4.2 Comparison of female to male

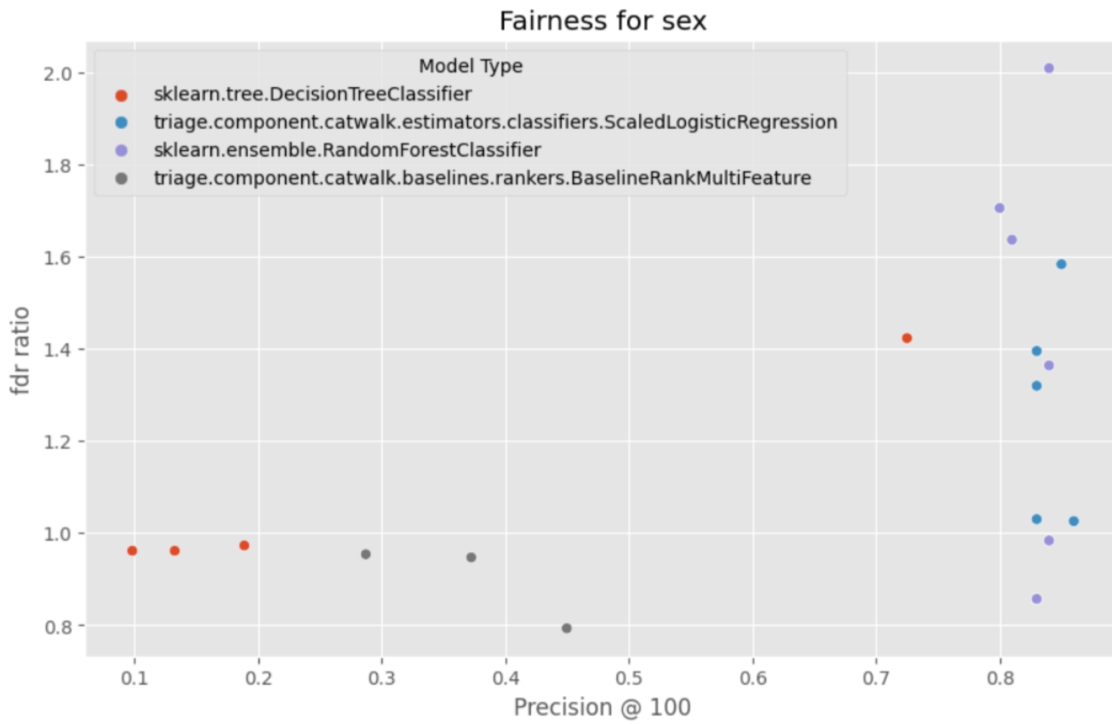


*Figure 17: False discovery rate of precision at 100 for sex comparison*

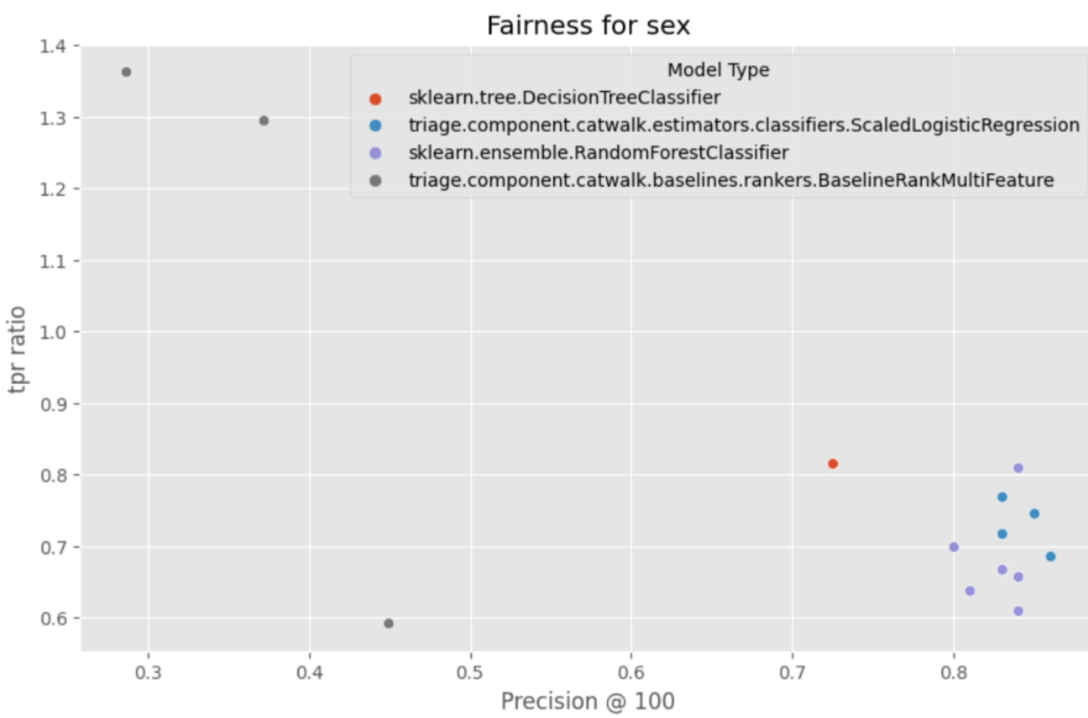


*Figure 18: True positive rate of precision at 100 for sex comparison*

5.4.3 Comparative between those needing mental health support and those who don’t (as per the screening form)

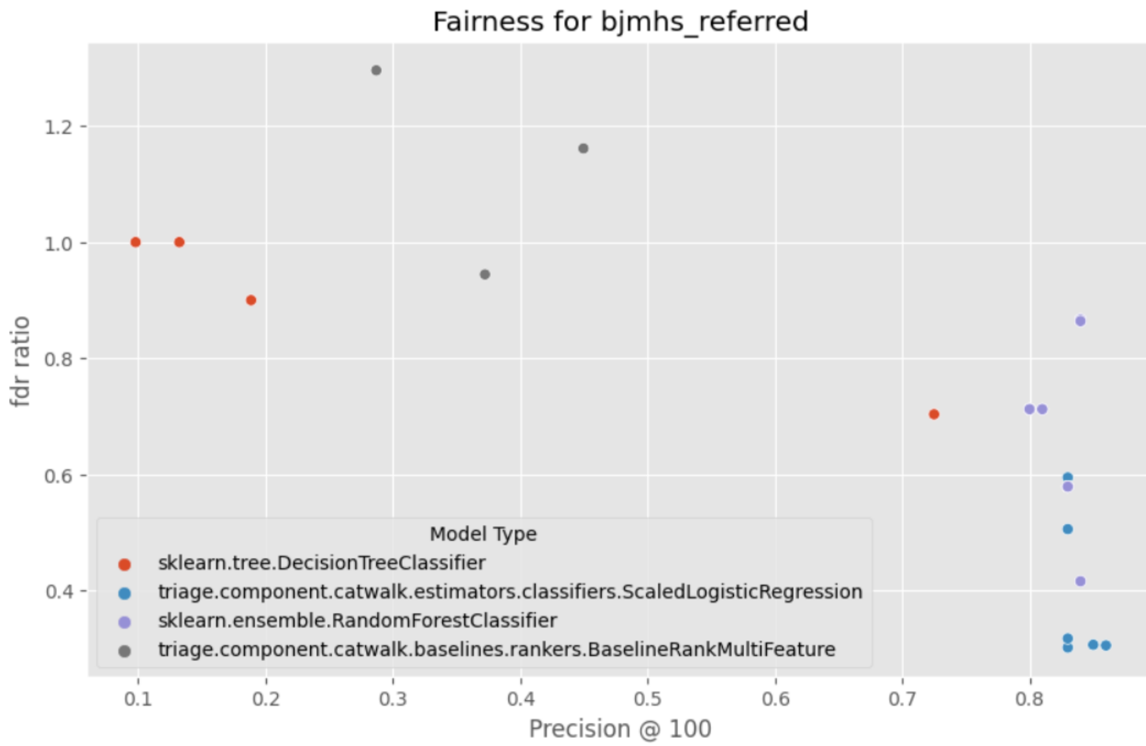


*Figure 19: False discovery rate of precision at 100 for BJMHS referral comparison*

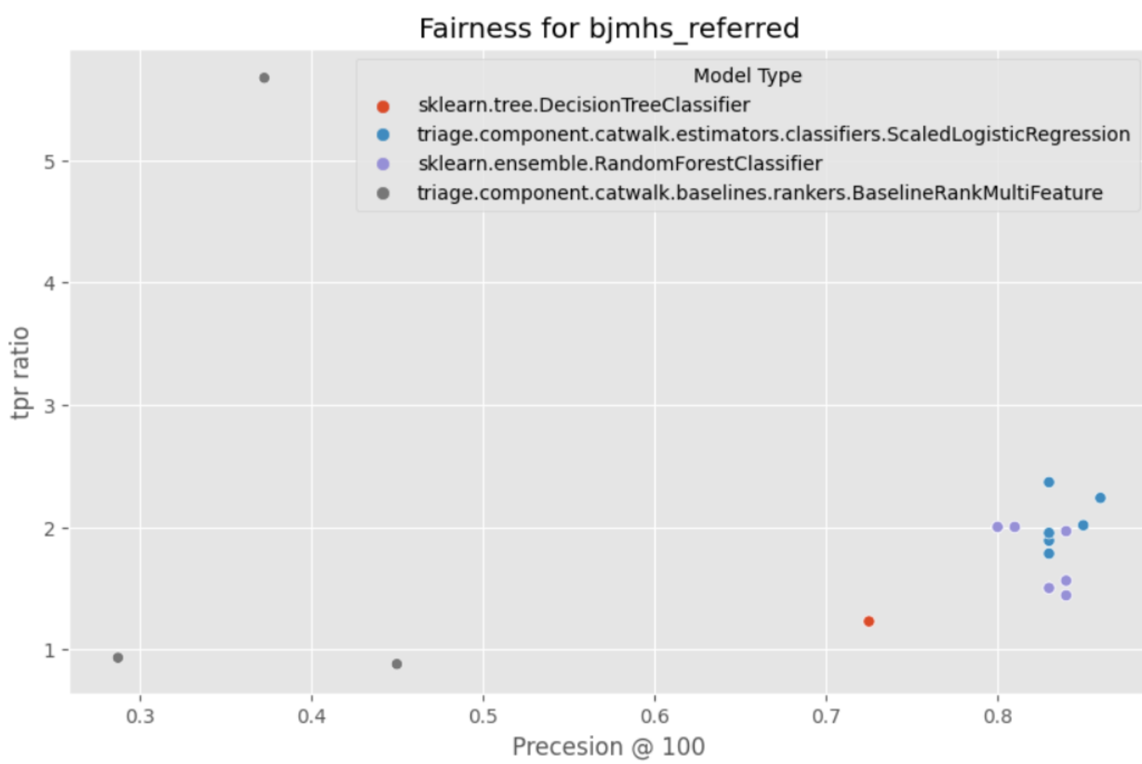


*Figure 20: True positive rate of precision at 100 for BJMHS referral comparison*

5.4.4 Comparative between people aged 41-50 at the time of their first booking and those in the age range of 21-30 at the time of their first booking

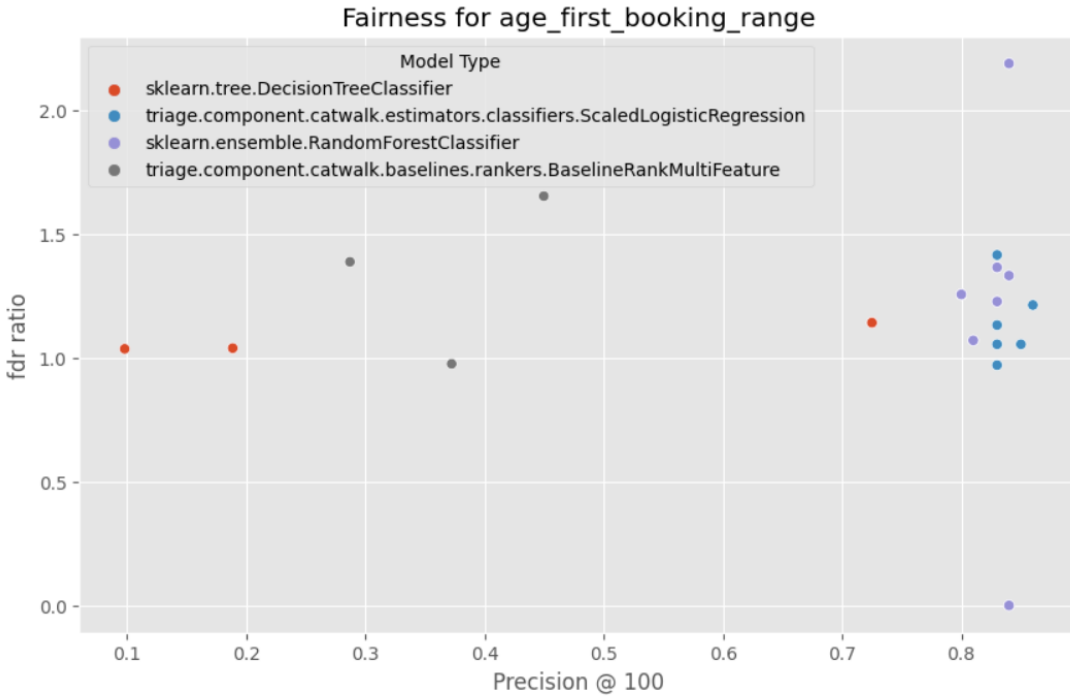


*Figure 21: False discovery rate of precision at 100 for first age booking comparison*

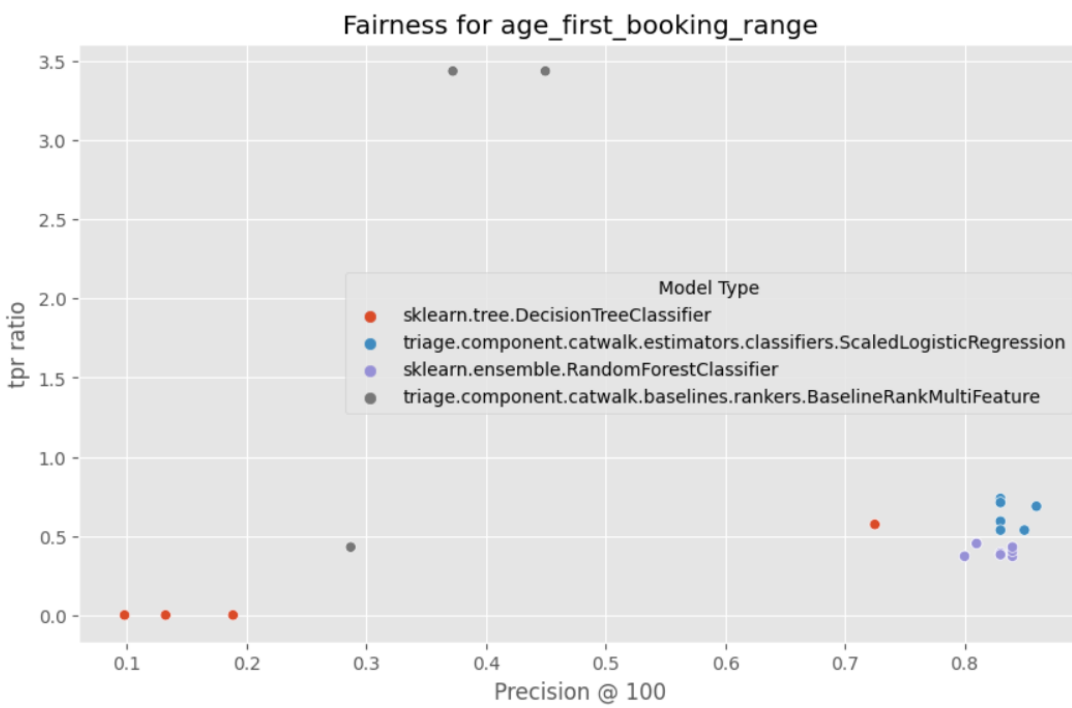


*Figure 22: True positive rate of precision at 100 first age booking comparison*

## 5.5 PR-k Curves for Top 5 Models and Baselines

| Model Number | Model Type | Hyperparameters | Precision@100 |
|---|---|---|---|
| 1. | Scaled Logistic Regression | {'C': 0.01, 'penalty': 'l1'} | 0.86 |
| 2. | Scaled Logistic Regression | {'C': 0.01, 'penalty': 'l2'} | 0.85 |
| 3. | Random Forest Classifier | {'max_depth': 500, 'n_estimators': 100, 'min_samples_split': 1000} | 0.84 |
| 4. | Random Forest Classifier | {'max_depth': 500, 'n_estimators': 1000, 'min_samples_split': 100} | 0.84 |
| 5. | Random Forest Classifier | {'max_depth': 500, 'n_estimators': 100, 'min_samples_split': 100} | 0.84 |
| 6. | Baseline (Number of Bookings) | {'rules': [{'feature': 'bkgs_jims_entity_id_all_num_bookings_count', 'low_value_high_score': False}]} | 0.45 |
| 7. | Baseline (JCMHC Admissions) | {'rules': [{'feature': 'num_adms_jcmhc_entity_id_all_num_admissions_count', 'low_value_high_score': False}]} | 0.37 |
| 8. | Baseline (Avg Days Before Next Booking) | {'rules': [{'feature': 'time_before_jims_entity_id_all_days_before_next_bkg_avg', 'low_value_high_score': True}]} | 0.29 |

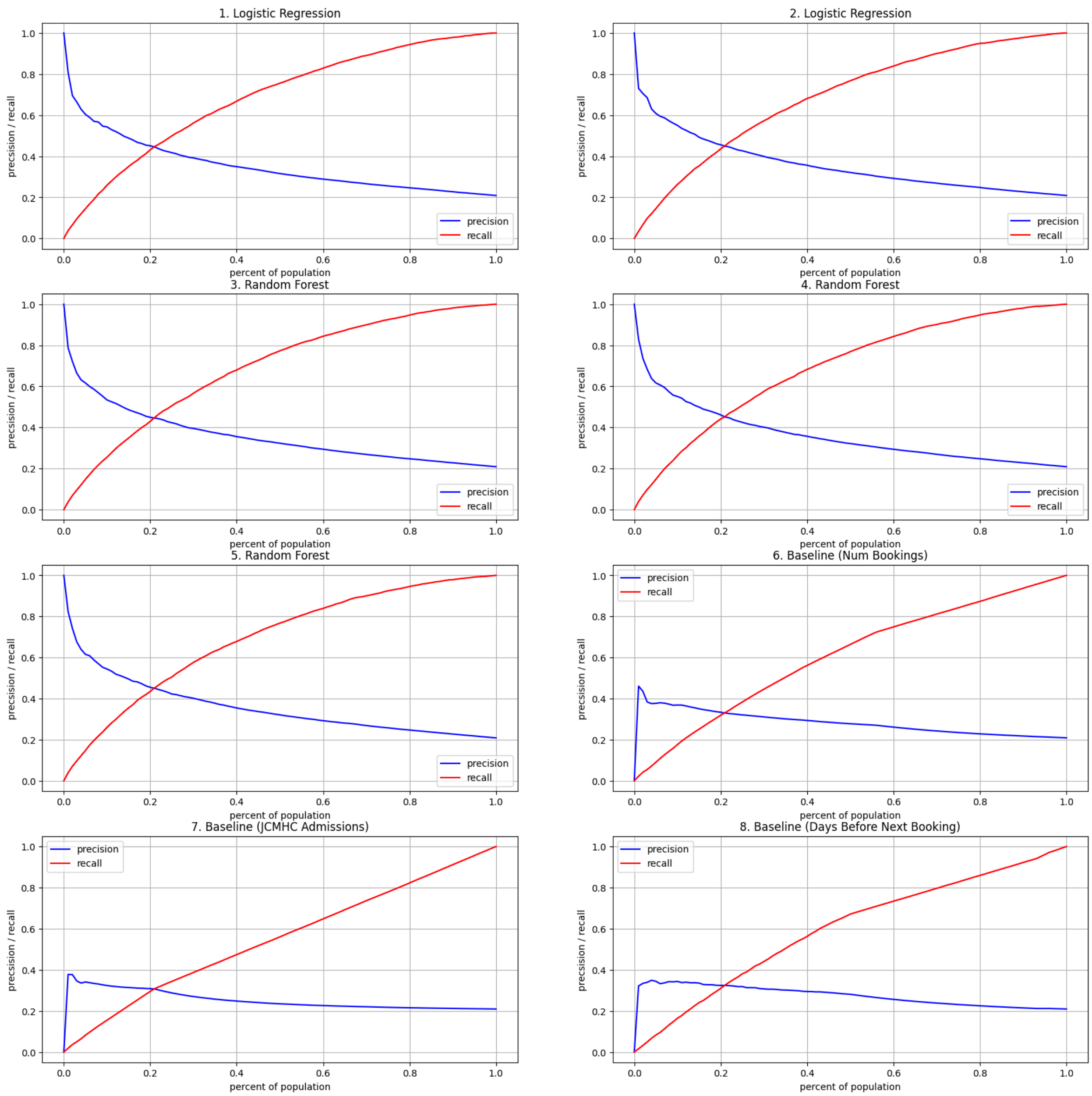


*Figure 23: PR-k Curves for top 5 models and 3 baselines*

## 5.6 Feature Importances for Top 5 Models

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| num_charges_pd_entity_id_2year_num_charges_count | 2.703968048 | 8.320539475 | 0.014891642 | 0.016029107 | 0.015539434 |
| arrest_type_pd_entity_id_all_arr_type_S_sum | 2.176589727 | 1.848301888 | 0.006964435 | 0.006794098 | 0.007038302 |
| charge_type_jims_entity_id_1year_charge_type__NULL_sum | 1.92667222 | 4.064843178 | 0.023779932 | 0.010714355 | 0.01351548 |
| bkgs_jims_entity_id_all_num_bookings_count | 1.883547306 | 2.100995302 | 0.053739882 | 0.024597903 | 0.032991104 |
| num_bookings_jims_entity_id_all_num_bookings_fixed_count | 1.883547306 | 2.100995302 | 0.034746845 | 0.022164355 | 0.022142755 |
| bkgs_jims_entity_id_6month_num_bookings_count | 1.796027422 | 2.010220528 | 0.036072198 | 0.021474578 | 0.020120185 |
| num_bookings_jims_entity_id_6month_num_bookings_fixed_count | 1.796027422 | 2.010220528 | 0.035018036 | 0.025427214 | 0.024454178 |
| bkgs_jims_entity_id_2year_num_bookings_count | 1.56002152 | 1.320848823 | 0.049807975 | 0.027575978 | 0.028688328 |
| bkgs_jims_entity_id_1year_num_bookings_count | 1.479387999 | 1.692485333 | 0.036977861 | 0.031659745 | 0.023240827 |
| pre_trial_ass_jims_entity_id_1year_felony_flag_max | 1.292210102 | 1.28353858 | 0.007645827 | 0.005337598 | 0.005727571 |
| pre_trial_ass_jims_entity_id_2year_felony_flag_max | 1.271722198 | 1.411779761 | 0.00764847 | 0.005199815 | 0.005954044 |
| demographics_jims_entity_id_all_marital_status_S_max | 1.232668877 | 1.052756429 | 0.001645248 | 0.002494427 | 0.00246412 |
| charge_type_jims_entity_id_all_charge_type__NULL_sum | 1.228088141 | 1.86233139 | 0.012360605 | 0.013137716 | 0.011408651 |
| demographics_jims_entity_id_all_sex_MALE_max | 1.204458475 | 1.647773027 | 0.000691323 | 0.00222276 | 0.002204934 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| pre_trial_ass_jims_entity_id_all_substance_flag_max | 1.156162381 | 1.169273853 | 0.002467631 | 0.003567288 | 0.003585197 |
| charge_type_jims_entity_id_6month_charge_type__NULL_sum | 1.128252149 | 4.697709084 | 0.005434371 | 0.005416742 | 0.005569475 |
| bkgs_jims_entity_id_1month_num_bookings_count | 1.116710186 | 1.486956239 | 0.001562149 | 0.002771242 | 0.002295131 |
| demographics_jims_entity_id_all_marital_status_OTHER_max | 1.096473575 | 1.069653511 | 0.000732728 | 0.002044451 | 0.001884098 |
| demographics_jims_entity_id_all_race_B_max | 1.093962669 | 0.992893696 | 0.000937274 | 0.002201556 | 0.002121691 |
| pre_trial_ass_jims_entity_id_all_prior_jail_time_max | 1.086017966 | 1.196864605 | 0.003941335 | 0.003075199 | 0.003293525 |
| pre_trial_ass_jims_entity_id_1year_substance_flag_max | 1.070348382 | 1.130299449 | 0.005010976 | 0.003281096 | 0.004276891 |
| pre_trial_ass_jims_entity_id_2year_drug_flag_max | 1.06009078 | 1.102051973 | 0.001025519 | 0.001612352 | 0.001538534 |
| pre_trial_ass_jims_entity_id_all_drug_flag_max | 1.00062716 | 0.973963499 | 0.001509144 | 0.001851156 | 0.001871559 |
| charge_type_jims_entity_id_1year_charge_type_F_sum | 1 | 1.479854703 | 0.008333919 | 0.004980382 | 0.005519555 |
| charge_type_jims_entity_id_1year_charge_type_FU_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_1year_charge_type_M_sum | 1 | 1.058657289 | 0.000639055 | 0.002364188 | 0.00268708 |
| charge_type_jims_entity_id_1year_charge_type_ON_sum | 1 | 0.811284959 | 0.001323199 | 0.003138926 | 0.003382655 |
| charge_type_jims_entity_id_1year_charge_type_OY_sum | 1 | 0.885604441 | 0.000278047 | 0.001009479 | 0.001035052 |
| charge_type_jims_entity_id_1year_charge_type_T_sum | 1 | 1 | 0 | 0 | 0 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| charge_type_jims_entity_id_1year_charge_type_X_sum | 1 | 0.519582987 | 0.00037573 | 0.000986794 | 0.000960275 |
| charge_type_jims_entity_id_2year_charge_type_C_sum | 1 | 5.051925182 | 0.00079998 | 0.002781725 | 0.002771482 |
| charge_type_jims_entity_id_2year_charge_type_F_sum | 1 | 2.10017848 | 0.002636779 | 0.004952703 | 0.004914084 |
| charge_type_jims_entity_id_2year_charge_type_FU_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_2year_charge_type_M_sum | 1 | 1.648352861 | 0.001213297 | 0.004107094 | 0.004223506 |
| charge_type_jims_entity_id_2year_charge_type__NULL_sum | 1 | 1.452066183 | 0.011472264 | 0.010325661 | 0.01003004 |
| charge_type_jims_entity_id_2year_charge_type_ON_sum | 1 | 2.614035845 | 0.002012398 | 0.005858617 | 0.005748165 |
| charge_type_jims_entity_id_2year_charge_type_OY_sum | 1 | 1.800143123 | 0.000777392 | 0.002114334 | 0.002019359 |
| charge_type_jims_entity_id_2year_charge_type_T_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_2year_charge_type_X_sum | 1 | 0.40165481 | 0.000641501 | 0.001826025 | 0.00184675 |
| charge_type_jims_entity_id_6month_charge_type_C_sum | 1 | 1.0381037 | 0.000132686 | 0.000643911 | 0.000625759 |
| charge_type_jims_entity_id_6month_charge_type_F_sum | 1 | 2.163896561 | 0.004102617 | 0.002971889 | 0.003189613 |
| charge_type_jims_entity_id_6month_charge_type_FU_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_6month_charge_type_M_sum | 1 | 1.147766113 | 0.000718414 | 0.001306652 | 0.001125827 |
| charge_type_jims_entity_id_6month_charge_type_ON_sum | 1 | 0.705198884 | 0.001305224 | 0.001789848 | 0.002073834 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| charge_type_jims_entity_id_6month_charge_type_OY_sum | 1 | 1.094228983 | 0.000171442 | 0.000551875 | 0.000564122 |
| charge_type_jims_entity_id_6month_charge_type_T_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_6month_charge_type_X_sum | 1 | 0.786840558 | 0.000250377 | 0.000516832 | 0.000590646 |
| charge_type_jims_entity_id_all_charge_type_C_sum | 1 | 0.659873903 | 0.001201159 | 0.004073986 | 0.004131275 |
| charge_type_jims_entity_id_all_charge_type_F_sum | 1 | 1.945049286 | 0.004443108 | 0.009586805 | 0.008878575 |
| charge_type_jims_entity_id_all_charge_type_FU_sum | 1 | 1 | 0 | 0 | 0 |
| charge_type_jims_entity_id_all_charge_type_M_sum | 1 | 1.114827275 | 0.003691014 | 0.007449579 | 0.007359634 |
| charge_type_jims_entity_id_all_charge_type_ON_sum | 1 | 1.352997184 | 0.009024168 | 0.010072952 | 0.010872791 |
| charge_type_jims_entity_id_all_charge_type_OY_sum | 1 | 0.849305809 | 0.000926454 | 0.003261378 | 0.003180499 |
| charge_type_jims_entity_id_all_charge_type_T_sum | 1 | 1.153010845 | 4.72345E-05 | 0.000195762 | 0.000185718 |
| charge_type_jims_entity_id_all_charge_type_X_sum | 1 | 1.620463133 | 0.000672197 | 0.002669065 | 0.002654804 |
| days_in_jail_jims_entity_id_1year_days_in_jail_avg | 1 | 2.091010094 | 0.018663033 | 0.020579857 | 0.015188169 |
| days_in_jail_jims_entity_id_1year_days_in_jail_max | 1 | 0.465720028 | 0.020714257 | 0.018112732 | 0.014500843 |
| days_in_jail_jims_entity_id_1year_days_in_jail_sum | 1 | 0.581657708 | 0.025992756 | 0.018688836 | 0.019322297 |
| days_in_jail_jims_entity_id_2year_days_in_jail_avg | 1 | 0.531978726 | 0.002435857 | 0.010329634 | 0.010114906 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| days_in_jail_jims_entity_id_2year_days_in_jail_sum | 1 | 0.563806713 | 0.006645371 | 0.012369404 | 0.011714997 |
| days_in_jail_jims_entity_id_6month_days_in_jail_avg | 1 | 1.822249413 | 0.050707807 | 0.023644208 | 0.032176655 |
| days_in_jail_jims_entity_id_6month_days_in_jail_max | 1 | 1.000913739 | 0.038628202 | 0.022392489 | 0.026583537 |
| days_in_jail_jims_entity_id_6month_days_in_jail_sum | 1 | 0.40354228 | 0.041026472 | 0.021531079 | 0.023397208 |
| days_in_jail_jims_entity_id_all_days_in_jail_avg | 1 | 1.647384524 | 0.004796319 | 0.01398566 | 0.01337649 |
| days_in_jail_jims_entity_id_all_days_in_jail_max | 1 | 1.57632041 | 0.011937099 | 0.01586876 | 0.015138617 |
| days_in_jail_jims_entity_id_all_days_in_jail_sum | 1 | 0.501878798 | 0.015693057 | 0.020348036 | 0.018149414 |
| demographics_jims_entity_id_all_marital_status__NULL_max | 1 | 1 | 0 | 0 | 0 |
| time_unemployed_jcmhc_entity_id_all_time_unemployed_sum | 1 | 0.6070081 | 0.00018941 | 0.00057606 | 0.00050594 |
| age_jims_entity_id_all_age_on_asofdate_max | 1 | 0.000517056 | 0.01193108 | 0.024174722 | 0.024893486 |
| arrest_type_pd_entity_id_1year_arr_type_O_sum | 1 | 1.196696997 | 0.001046164 | 0.002341826 | 0.002405078 |
| arrest_type_pd_entity_id_1year_arr_type_R_sum | 1 | 1.292584658 | 4.5956E-06 | 1.9992E-06 | 0 |
| arrest_type_pd_entity_id_1year_arr_type_S_sum | 1 | 3.336013556 | 0.002898079 | 0.003881568 | 0.003331433 |
| arrest_type_pd_entity_id_1year_arr_type_T_sum | 1 | 1.103129387 | 0.006519271 | 0.005950115 | 0.005844013 |
| arrest_type_pd_entity_id_2year_arr_type__NULL_sum | 1 | 1.074918866 | 0.002192032 | 0.00209147 | 0.00213083 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| arrest_type_pd_entity_id_2year_arr_type_O_sum | 1 | 0.669612467 | 0.000954239 | 0.0026519 | 0.003009334 |
| arrest_type_pd_entity_id_2year_arr_type_R_sum | 1 | 6.141424179 | 7.66786E-05 | 0.000117808 | 0.000121536 |
| arrest_type_pd_entity_id_2year_arr_type_S_sum | 1 | 0.941028476 | 0.004146328 | 0.004197794 | 0.004470237 |
| arrest_type_pd_entity_id_2year_arr_type_T_sum | 1 | 0.455572575 | 0.007841332 | 0.006488652 | 0.006133376 |
| arrest_type_pd_entity_id_all_arr_type_O_sum | 1 | 1.165222645 | 0.001068007 | 0.003794538 | 0.003772292 |
| arrest_type_pd_entity_id_all_arr_type_R_sum | 1 | 0.75201577 | 8.57438E-05 | 0.000347171 | 0.000368545 |
| arrest_type_pd_entity_id_all_arr_type_T_sum | 1 | 2.654327393 | 0.011996568 | 0.010832767 | 0.011262487 |
| call_type_jcmhc_entity_id_1year_num_crisis_call_sum | 1 | 0.777889967 | 0.000156446 | 0.000417155 | 0.000455208 |
| call_type_jcmhc_entity_id_1year_num_referral_call_sum | 1 | 1.239898086 | 0.000626429 | 0.001396316 | 0.00135259 |
| call_type_jcmhc_entity_id_2year_num_crisis_call_sum | 1 | 0.635787487 | 0.000196154 | 0.000766073 | 0.000699904 |
| call_type_jcmhc_entity_id_2year_num_referral_call_sum | 1 | 0.826757789 | 0.000769435 | 0.002027779 | 0.00222609 |
| call_type_jcmhc_entity_id_6month_num_crisis_call_sum | 1 | 1.132942915 | 5.46514E-05 | 0.000274226 | 0.00030231 |
| call_type_jcmhc_entity_id_6month_num_referral_call_sum | 1 | 0.618748367 | 0.000379308 | 0.000855055 | 0.000899084 |
| call_type_jcmhc_entity_id_all_num_crisis_call_sum | 1 | 1.500422478 | 0.000270649 | 0.001442071 | 0.001587616 |
| call_type_jcmhc_entity_id_all_num_referral_call_sum | 1 | 1.222845316 | 0.00373442 | 0.005796888 | 0.006174131 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| charge_type_jims_entity_id_1year_charge_type_C_sum | 1 | 0.997112453 | 0.000410715 | 0.001360339 | 0.001250957 |
| medact_imptypes_entity_id_2year_drug_count_sum | 1 | 1.28501761 | 0.00038738 | 0.000814984 | 0.000792191 |
| medact_imptypes_entity_id_2year_pysch_count_imp | 1 | 1.008613348 | 0.000318483 | 0.000653236 | 0.000681987 |
| medact_imptypes_entity_id_2year_pysch_count_sum | 1 | 0.883956075 | 0.000249013 | 0.00066423 | 0.000665151 |
| medact_imptypes_entity_id_6month_alc_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_6month_anx_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_6month_drug_count_sum | 1 | 1.283497214 | 0.000125271 | 0.000286106 | 0.000244751 |
| medact_imptypes_entity_id_6month_pysch_count_sum | 1 | 1.605869174 | 0.000174886 | 0.000239756 | 0.000243196 |
| medact_imptypes_entity_id_all_alc_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_all_anx_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_all_drug_count_sum | 1 | 1.291442394 | 0.000626089 | 0.001348171 | 0.001312045 |
| medact_imptypes_entity_id_all_pysch_count_sum | 1 | 1.283555269 | 0.000306371 | 0.000981509 | 0.000970136 |
| num_adms_jcmhc_entity_id_1year_num_admissions_count | 1 | 0.979016602 | 0.001352015 | 0.001622614 | 0.001554989 |
| num_adms_jcmhc_entity_id_2year_num_admissions_count | 1 | 0.814469457 | 0.000911816 | 0.001819892 | 0.00195693 |
| num_adms_jcmhc_entity_id_6month_num_admissions_count | 1 | 1.936677575 | 0.000928245 | 0.001499665 | 0.001557309 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| num_adms_jcmhc_entity_id_all_num_admissions_count | 1 | 0.919063032 | 0.003849037 | 0.005696797 | 0.004924146 |
| num_charges_pd_entity_id_all_num_charges_count | 1 | 0.847034335 | 0.014227995 | 0.019333076 | 0.017198586 |
| num_crisiseventsvcs_jcmhc_entity_id_1year_num_referred_sum | 1 | 0.921776176 | 0.00015398 | 0.000736996 | 0.000639885 |
| num_crisiseventsvcs_jcmhc_entity_id_2year_num_referred_sum | 1 | 0.938561499 | 0.000227865 | 0.001129872 | 0.001204118 |
| num_crisiseventsvcs_jcmhc_entity_id_6month_num_referred_sum | 1 | 0.642319024 | 9.99611E-05 | 0.000564682 | 0.00059728 |
| num_crisiseventsvcs_jcmhc_entity_id_all_num_referred_sum | 1 | 0.607027829 | 0.001267805 | 0.002486939 | 0.002700598 |
| num_medact_entity_id_1year_num_medact_count | 1 | 0.6264413 | 0.000514002 | 0.001240445 | 0.00120198 |
| num_medact_entity_id_2year_num_medact_count | 1 | 0.869449914 | 0.000649026 | 0.001675654 | 0.001693969 |
| num_medact_entity_id_6month_num_medact_count | 1 | 1.260637999 | 0.000433582 | 0.000880208 | 0.000805789 |
| num_medact_entity_id_all_num_medact_count | 1 | 0.368488342 | 0.000862959 | 0.0026552 | 0.002791373 |
| num_med_clinics_jcmhc_entity_id_1year_num_med_clinics_sum | 1 | 0.545424163 | 6.83934E-05 | 0.000339486 | 0.000334819 |
| num_med_clinics_jcmhc_entity_id_2year_num_med_clinics_sum | 1 | 2.176107407 | 0.000246522 | 0.000707966 | 0.000716642 |
| num_med_clinics_jcmhc_entity_id_6month_num_med_clinics_sum | 1 | 1.178309679 | 3.73161E-05 | 0.000220386 | 0.000234477 |
| num_med_clinics_jcmhc_entity_id_all_num_med_clinics_sum | 1 | 0.852544427 | 0.000482583 | 0.001963983 | 0.001946271 |
| num_referred_jims_entity_id_1year_num_referred_sum | 1 | 1.073843956 | 2.46439E-05 | 0.000158917 | 0.000139643 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| num_referred_jims_entity_id_2year_num_referred_sum | 1 | 1.073843956 | 3.04794E-05 | 0.000154597 | 0.000159841 |
| num_referred_jims_entity_id_6month_num_referred_sum | 1 | 1.073843956 | 4.29654E-05 | 0.000145979 | 0.000188061 |
| num_referred_jims_entity_id_all_num_referred_sum | 1 | 1.073843956 | 3.88568E-05 | 0.000163085 | 0.00017428 |
| pre_trial_ass_jims_entity_id_1year_drug_flag_max | 1 | 0.981434524 | 0.001065783 | 0.001204674 | 0.000942452 |
| pre_trial_ass_jims_entity_id_1year_employed_max | 1 | 1.036080957 | 0.000378554 | 0.000978814 | 0.000997454 |
| pre_trial_ass_jims_entity_id_1year_ks_resident_max | 1 | 0.912626565 | 0.000838898 | 0.001586445 | 0.001192883 |
| pre_trial_ass_jims_entity_id_1year_prior_jail_time_max | 1 | 1.035408378 | 0.003798925 | 0.003012812 | 0.003021149 |
| pre_trial_ass_jims_entity_id_2year_prior_jail_time_max | 1 | 0.980885029 | 0.002116507 | 0.003059163 | 0.003096442 |
| pre_trial_ass_jims_entity_id_2year_substance_flag_max | 1 | 1.005371094 | 0.002938136 | 0.003355484 | 0.003688804 |
| pre_trial_ass_jims_entity_id_all_felony_flag_max | 1 | 0.912840545 | 0.006449894 | 0.004760179 | 0.004932355 |
| pre_trial_ass_jims_entity_id_all_ks_resident_max | 1 | 1.011731863 | 0.00152834 | 0.001723712 | 0.001682254 |
| time_unemployed_jcmhc_entity_id_1year_time_unemployed_sum | 1 | 1.072007299 | 7.66437E-05 | 0.00030788 | 0.000291557 |
| time_unemployed_jcmhc_entity_id_2year_time_unemployed_sum | 1 | 1.196664453 | 0.00010435 | 0.00040475 | 0.000403653 |
| time_unemployed_jcmhc_entity_id_6month_time_unemployed_sum | 1 | 0.700910986 | 0.00005718 | 0.000202464 | 0.000221577 |
| age_first_bkg_jims_entity_id_all_age_first_bkg_min | 1 | 0.000848089 | 0.012056946 | 0.021647777 | 0.021563747 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| demographics_jims_entity_id_all_marital_status_W_max | 1 | 1.155031681 | 9.77814E-05 | 0.000773285 | 0.000759107 |
| demographics_jims_entity_id_all_race_A_max | 1 | 0.587690473 | 3.01433E-05 | 0.000313608 | 0.000325802 |
| demographics_jims_entity_id_all_race_I_max | 1 | 0.380477488 | 0 | 1.95491E-05 | 2.27785E-05 |
| demographics_jims_entity_id_all_race__NULL_max | 1 | 1 | 0 | 0 | 0 |
| demographics_jims_entity_id_all_race_O_max | 1 | 1 | 0 | 0 | 0 |
| demographics_jims_entity_id_all_race_W_max | 1 | 0.905145168 | 0.000670967 | 0.002201169 | 0.002295242 |
| demographics_jims_entity_id_all_sex__NULL_max | 1 | 0.8002581 | 0 | 1.32E-08 | 0 |
| dla_cafas_scores_jcmhc_entity_id_1year_cafas_score_avg | 1 | 0.797470808 | 3.44097E-05 | 0.000374554 | 0.000337694 |
| dla_cafas_scores_jcmhc_entity_id_1year_cafas_score_imp | 1 | 0.896995127 | 3.38807E-05 | 0.000177887 | 0.000141664 |
| dla_cafas_scores_jcmhc_entity_id_1year_cafas_score_max | 1 | 0.796578646 | 0.000029592 | 0.000362707 | 0.000406879 |
| dla_cafas_scores_jcmhc_entity_id_1year_cafas_score_min | 1 | 0.879648745 | 6.13403E-05 | 0.000364979 | 0.000381484 |
| dla_cafas_scores_jcmhc_entity_id_1year_dla_score_avg | 1 | 1.124377608 | 0.000365917 | 0.002431111 | 0.002596155 |
| dla_cafas_scores_jcmhc_entity_id_1year_dla_score_imp | 1 | 0.896995127 | 6.85409E-05 | 0.000172803 | 0.000167518 |
| dla_cafas_scores_jcmhc_entity_id_1year_dla_score_max | 1 | 0.781947792 | 0.000373443 | 0.002590252 | 0.00252604 |
| dla_cafas_scores_jcmhc_entity_id_1year_dla_score_min | 1 | 1.381865025 | 0.000397478 | 0.002502853 | 0.00250395 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| dla_cafas_scores_jcmhc_entity_id_2year_cafas_score_avg | 1 | 0.880112171 | 5.26026E-05 | 0.000385659 | 0.000391341 |
| dla_cafas_scores_jcmhc_entity_id_2year_cafas_score_imp | 1 | 1.376170754 | 4.73128E-05 | 0.000165471 | 0.000156407 |
| dla_cafas_scores_jcmhc_entity_id_2year_cafas_score_max | 1 | 0.833204866 | 4.16255E-05 | 0.000389851 | 0.00042107 |
| dla_cafas_scores_jcmhc_entity_id_2year_cafas_score_min | 1 | 0.884264529 | 3.42687E-05 | 0.000380518 | 0.000375826 |
| dla_cafas_scores_jcmhc_entity_id_2year_dla_score_avg | 1 | 0.894749582 | 0.000308737 | 0.002443438 | 0.002365009 |
| dla_cafas_scores_jcmhc_entity_id_2year_dla_score_imp | 1 | 1.376170754 | 3.59882E-05 | 0.000170031 | 0.000143164 |
| dla_cafas_scores_jcmhc_entity_id_2year_dla_score_max | 1 | 0.797926903 | 0.000354085 | 0.002566646 | 0.002691223 |
| dla_cafas_scores_jcmhc_entity_id_2year_dla_score_min | 1 | 0.85804534 | 0.000299025 | 0.00256359 | 0.002629977 |
| dla_cafas_scores_jcmhc_entity_id_all_cafas_score_avg | 1 | 0.879721522 | 3.82467E-05 | 0.000381917 | 0.000336775 |
| dla_cafas_scores_jcmhc_entity_id_all_cafas_score_imp | 1 | 0.869692981 | 3.42683E-05 | 0.000178962 | 0.000225149 |
| dla_cafas_scores_jcmhc_entity_id_all_cafas_score_max | 1 | 0.821482718 | 4.11174E-05 | 0.000375392 | 0.000403364 |
| dla_cafas_scores_jcmhc_entity_id_all_cafas_score_min | 1 | 0.884183466 | 5.36382E-05 | 0.000361844 | 0.000374602 |
| dla_cafas_scores_jcmhc_entity_id_all_dla_score_avg | 1 | 1.132922769 | 0.000374863 | 0.00253762 | 0.002412213 |
| dla_cafas_scores_jcmhc_entity_id_all_dla_score_imp | 1 | 0.869692981 | 1.99947E-05 | 0.000180116 | 0.000193184 |
| dla_cafas_scores_jcmhc_entity_id_all_dla_score_max | 1 | 1.298517585 | 0.000361866 | 0.002606414 | 0.002675045 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| dla_cafas_scores_jcmhc_entity_id_all_dla_score_min | 1 | 0.857229352 | 0.000323502 | 0.002592276 | 0.002590483 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_hom_sum | 1 | 1.603467464 | 0.000038153 | 0.000144824 | 0.000164996 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_sui_sum | 1 | 1.575650215 | 0.00027531 | 0.000746915 | 0.000792152 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_vio_sum | 1 | 0.569176674 | 0.000281128 | 0.000589516 | 0.000579682 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_hom_sum | 1 | 0.71914494 | 5.84594E-05 | 0.000344976 | 0.00037115 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_sui_sum | 1 | 0.619092941 | 0.000424009 | 0.001602563 | 0.001627028 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_vio_sum | 1 | 2.212601423 | 0.000615384 | 0.001682233 | 0.001724224 |
| income_jcmhc_entity_id_1year_income_avg | 1 | 0.884850025 | 0.000694105 | 0.004548599 | 0.004650035 |
| income_jcmhc_entity_id_1year_income_imp | 1 | 1.050230265 | 0.000221049 | 0.000460755 | 0.000516575 |
| income_jcmhc_entity_id_1year_income_max | 1 | 0.96087873 | 0.000826859 | 0.004501732 | 0.004473276 |
| income_jcmhc_entity_id_2year_income_avg | 1 | 0.640667141 | 0.001194846 | 0.005102708 | 0.005059709 |
| income_jcmhc_entity_id_2year_income_imp | 1 | 0.936791539 | 0.000398286 | 0.000562191 | 0.000580136 |
| income_jcmhc_entity_id_2year_income_max | 1 | 1.062606335 | 0.000835306 | 0.004936421 | 0.004950053 |
| income_jcmhc_entity_id_all_income_avg | 1 | 0.801687181 | 0.001879721 | 0.00799809 | 0.007937333 |
| income_jcmhc_entity_id_all_income_imp | 1 | 1.01240325 | 0.00390617 | 0.00267703 | 0.002144278 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| income_jcmhc_entity_id_all_income_max | 1 | 1.6474123 | 0.001992405 | 0.007949244 | 0.007840332 |
| medact_calltypes_entity_id_1year_blue_count_sum | 1 | 1.64740324 | 0.000009889 | 1.22236E-05 | 1.41628E-05 |
| medact_calltypes_entity_id_1year_green_count_sum | 1 | 0.825525224 | 0.000554215 | 0.00121461 | 0.001355803 |
| medact_calltypes_entity_id_1year_grey_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_calltypes_entity_id_1year_red_count_sum | 1 | 0.779115439 | 0.000146327 | 0.000256936 | 0.000219491 |
| medact_calltypes_entity_id_1year_yellow_count_sum | 1 | 0.644843042 | 0.000157465 | 0.000495208 | 0.000533043 |
| medact_calltypes_entity_id_2year_blue_count_sum | 1 | 0.867939651 | 8.4647E-06 | 1.55278E-05 | 0.00001405 |
| medact_calltypes_entity_id_2year_green_count_sum | 1 | 0.866753697 | 0.000574842 | 0.001560567 | 0.001494946 |
| medact_calltypes_entity_id_2year_grey_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_calltypes_entity_id_2year_red_count_sum | 1 | 1.514492512 | 0.000241819 | 0.000498529 | 0.000527909 |
| medact_calltypes_entity_id_2year_yellow_count_sum | 1 | 1.749136686 | 0.000242238 | 0.000897718 | 0.000923584 |
| medact_calltypes_entity_id_6month_blue_count_sum | 1 | 1.263364315 | 0 | 0.000012601 | 1.25091E-05 |
| medact_calltypes_entity_id_6month_green_count_sum | 1 | 1.576318502 | 0.000335216 | 0.000832061 | 0.000814448 |
| medact_calltypes_entity_id_6month_grey_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_calltypes_entity_id_6month_red_count_sum | 1 | 1.347320676 | 0.00010329 | 0.000186599 | 0.000209157 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| medact_calltypes_entity_id_6month_yellow_count_sum | 1 | 0.985566735 | 0.000133078 | 0.000331758 | 0.000365275 |
| medact_calltypes_entity_id_all_blue_count_sum | 1 | 0.627079129 | 3.6194E-06 | 1.56257E-05 | 1.85083E-05 |
| medact_calltypes_entity_id_all_green_count_sum | 1 | 0.392239153 | 0.000870346 | 0.002524721 | 0.002548692 |
| medact_calltypes_entity_id_all_grey_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_calltypes_entity_id_all_red_count_sum | 1 | 0.621717811 | 0.000233075 | 0.000675749 | 0.000647343 |
| medact_calltypes_entity_id_all_yellow_count_sum | 1 | 0.381890148 | 0.000431246 | 0.001468781 | 0.001425769 |
| medact_imptypes_entity_id_1year_alc_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_1year_anx_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_1year_drug_count_sum | 1 | 1.20354259 | 0.000249131 | 0.000461186 | 0.000467591 |
| medact_imptypes_entity_id_1year_pysch_count_sum | 1 | 0.80825001 | 0.000199274 | 0.000433984 | 0.000430879 |
| medact_imptypes_entity_id_2year_alc_count_imp | 1 | 1.008613348 | 0.000225119 | 0.000666769 | 0.000689365 |
| medact_imptypes_entity_id_2year_alc_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_2year_anx_count_imp | 1 | 1.008613348 | 0.000228 | 0.000646263 | 0.000629956 |
| medact_imptypes_entity_id_2year_anx_count_sum | 1 | 1 | 0 | 0 | 0 |
| medact_imptypes_entity_id_2year_drug_count_imp | 1 | 1.008613348 | 0.000275081 | 0.000658678 | 0.000629204 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| demographics_jims_entity_id_all_sex_FEMALE_max | 0.999992251 | 1.324894667 | 0.000695252 | 0.002225229 | 0.00224889 |
| pre_trial_ass_jims_entity_id_2year_employed_max | 0.993653178 | 0.941717565 | 0.000369731 | 0.001256207 | 0.001229695 |
| demographics_jims_entity_id_all_marital_status_D_max | 0.986210346 | 1.049796701 | 0.000253082 | 0.001536608 | 0.001570596 |
| call_type_jcmhc_entity_id_2year_num_referral_call_imp | 0.986143053 | 0.978261471 | 0.000685023 | 0.001022438 | 0.000916718 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_vio_imp | 0.986143053 | 0.978261471 | 0.001243557 | 0.001067649 | 0.001392317 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_sui_imp | 0.986143053 | 0.978261471 | 0.000977549 | 0.001024802 | 0.000854707 |
| evaluated_as_risk_jcmhc_entity_id_2year_num_referred_ho_imp | 0.986143053 | 0.978261471 | 0.000677427 | 0.000953796 | 0.000915995 |
| call_type_jcmhc_entity_id_2year_num_crisis_call_imp | 0.986143053 | 0.978261471 | 0.000738203 | 0.001008833 | 0.000862003 |
| medact_imptypes_entity_id_1year_anx_count_imp | 0.98107332 | 0.962122202 | 0.000599741 | 0.000747613 | 0.000886256 |
| medact_imptypes_entity_id_1year_pysch_count_imp | 0.98107332 | 0.962122202 | 0.000406153 | 0.000683736 | 0.0007149 |
| medact_imptypes_entity_id_1year_drug_count_imp | 0.98107332 | 0.962122202 | 0.00040867 | 0.000684707 | 0.000717656 |
| medact_imptypes_entity_id_1year_alc_count_imp | 0.98107332 | 0.962122202 | 0.00049687 | 0.000754603 | 0.000760872 |
| medact_imptypes_entity_id_all_alc_count_imp | 0.980791628 | 0.96652019 | 0.000468167 | 0.000988032 | 0.001006186 |
| medact_imptypes_entity_id_all_anx_count_imp | 0.980791628 | 0.96652019 | 0.000429726 | 0.001003242 | 0.00112823 |
| medact_imptypes_entity_id_all_pysch_count_imp | 0.980791628 | 0.96652019 | 0.000419038 | 0.000997772 | 0.000974364 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| medact_imptypes_entity_id_all_drug_count_imp | 0.980791628 | 0.96652019 | 0.000322132 | 0.001001142 | 0.000884316 |
| medact_imptypes_entity_id_6month_pysch_count_imp | 0.978600621 | 0.959295929 | 0.000234605 | 0.000594539 | 0.000568536 |
| medact_imptypes_entity_id_6month_anx_count_imp | 0.978600621 | 0.959295929 | 0.000254249 | 0.000620689 | 0.000569623 |
| medact_imptypes_entity_id_6month_drug_count_imp | 0.978600621 | 0.959295929 | 0.000280421 | 0.000626305 | 0.000703842 |
| medact_imptypes_entity_id_6month_alc_count_imp | 0.978600621 | 0.959295929 | 0.000255936 | 0.000632963 | 0.000679004 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_sui_imp | 0.976768315 | 0.986593247 | 0.002636972 | 0.002096022 | 0.002536199 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_vio_imp | 0.976768315 | 0.986593247 | 0.00278326 | 0.002589254 | 0.002639536 |
| call_type_jcmhc_entity_id_all_num_referral_call_imp | 0.976768315 | 0.986593247 | 0.003328307 | 0.002628614 | 0.002328117 |
| call_type_jcmhc_entity_id_all_num_crisis_call_imp | 0.976768315 | 0.986593247 | 0.002025584 | 0.002517969 | 0.001482632 |
| evaluated_as_risk_jcmhc_entity_id_all_num_referred_ho_imp | 0.976768315 | 0.986593247 | 0.004001651 | 0.002322836 | 0.002723765 |
| call_type_jcmhc_entity_id_6month_num_crisis_call_imp | 0.975496054 | 0.939109385 | 0.000955803 | 0.000771943 | 0.000787199 |
| call_type_jcmhc_entity_id_6month_num_referral_call_imp | 0.975496054 | 0.939109385 | 0.000562054 | 0.000852419 | 0.000795395 |
| time_before_jims_entity_id_all_days_before_next_bkg_avg | 0.968196094 | 0.955521226 | 0.022437118 | 0.029727162 | 0.02737825 |
| days_in_jail_jims_entity_id_2year_days_in_jail_max | 0.965951622 | 0.89594996 | 0.004799133 | 0.00966872 | 0.009849892 |
| arrest_type_pd_entity_id_all_arr_type__NULL_sum | 0.961370826 | 0.967794418 | 0.000743981 | 0.001585319 | 0.001679564 |

| Features | 1. Logistic Regression | 2. Logistic Regression | 3. Random Forest | 4. Random Forest | 5. Random Forest |
|---|---|---|---|---|---|
| pre_trial_ass_jims_entity_id_2year_ks_resident_max | 0.956480742 | 0.913592994 | 0.001220747 | 0.001630581 | 0.001546037 |
| pre_trial_ass_jims_entity_id_all_employed_max | 0.950053096 | 0.895309627 | 0.000327542 | 0.001296533 | 0.001309614 |
| num_adms_jcmhc_entity_id_2year_num_admissions_imp | 0.947914541 | 0.950505555 | 0.00680444 | 0.003754423 | 0.003521451 |
| num_adms_jcmhc_entity_id_all_num_admissions_imp | 0.947914541 | 0.950505555 | 0.009490707 | 0.003520265 | 0.004577673 |
| num_adms_jcmhc_entity_id_1year_num_admissions_imp | 0.947914541 | 0.950505555 | 0.006969655 | 0.003646145 | 0.004400775 |
| num_adms_jcmhc_entity_id_6month_num_admissions_imp | 0.947914541 | 0.950505555 | 0.004344099 | 0.003566562 | 0.004036874 |
| call_type_jcmhc_entity_id_1year_num_referral_call_imp | 0.945459008 | 0.928968906 | 0.001233957 | 0.001199641 | 0.001121234 |
| call_type_jcmhc_entity_id_1year_num_crisis_call_imp | 0.945459008 | 0.928968906 | 0.000751462 | 0.001198329 | 0.000819146 |
| time_before_jims_entity_id_all_days_before_next_bk_imp | 0.934132338 | 1.125908375 | 0.022234978 | 0.008431459 | 0.011740566 |
| demographics_jims_entity_id_all_marital_status_M_max | 0.914051473 | 0.909020305 | 0.001161299 | 0.002097008 | 0.002003414 |
| arrest_type_pd_entity_id_1year_arr_type__NULL_sum | 0.874307632 | 0.940301597 | 0.014490697 | 0.009080472 | 0.00820485 |
| time_before_jims_entity_id_3year_days_before_next_bk_imp | 0.708525121 | 0.618996084 | 0.010398468 | 0.010870011 | 0.007634829 |
| time_before_jims_entity_id_3year_days_before_next_bkg_avg | 0.481585085 | 0.568156719 | 0.02745585 | 0.029581727 | 0.030960629 |
| release_jims_entity_id_all_days_since_release_min | 0.232309535 | 0.216455415 | 0.058542433 | 0.055406521 | 0.053084385 |

## 5.7 Crosstab of 10 Most Different Features for Top 5 Models

| **1. Logistic Regression** | | | |
|---|---|---|---|
| **Feature** | **Label=1** | **Label=0** | **Difference** |
| income_entity_id_1year_income_max | 31754.04 | 25871.53 | 5882.51 |
| income_entity_id_1year_income_avg | 31298.82 | 25466.53 | 5832.30 |
| income_entity_id_all_income_avg | 20736.93 | 25964.62 | 5227.69 |
| income_entity_id_2year_income_avg | 53489.10 | 58330.68 | 4841.57 |
| income_entity_id_2year_income_max | 55927.05 | 60448.12 | 4521.08 |
| time_before_entity_id_3year_days_before_next_bkg_avg | 202.21 | 1366.73 | 1164.52 |
| age_entity_id_all_age_on_asofdate_max | 11628.71 | 12746.35 | 1117.64 |
| time_before_entity_id_all_days_before_next_bkg_avg | 193.35 | 1274.91 | 1081.56 |
| income_entity_id_all_income_max | 37197.46 | 38073.83 | 876.37 |
| release_entity_id_all_days_since_release_min | 47.09 | 324.63 | 277.54 |
| **2. Logistic Regression** | | | |
| **Feature** | **Label=1** | **Label=0** | **Difference** |
| income_entity_id_1year_income_max | 33461.56 | 25862.67 | 7598.89 |

| | | | |
|---|---|---|---|
| income_entity_id_1year_income_avg | 32946.73 | 25457.98 | 7488.75 |
| income_entity_id_all_income_avg | 22194.48 | 25957.06 | 3762.58 |
| age_entity_id_all_age_on_asofdate_max | 11299.98 | 12748.06 | 1448.08 |
| income_entity_id_2year_income_avg | 57133.55 | 58311.77 | 1178.22 |
| time_before_entity_id_3year_days_before_next_bkg_avg | 217.52 | 1366.65 | 1149.13 |
| time_before_entity_id_all_days_before_next_bkg_avg | 204.65 | 1274.85 | 1070.20 |
| income_entity_id_2year_income_max | 59546.92 | 60429.34 | 882.42 |
| release_entity_id_all_days_since_release_min | 47.80 | 324.62 | 276.82 |
| income_entity_id_all_income_max | 37829.08 | 38070.55 | 241.47 |
| **3. Random Forest** | | | |
| **Feature** | **Label=1** | **Label=0** | **Difference** |
| income_entity_id_1year_income_max | 31519.65 | 25872.75 | 5646.91 |
| income_entity_id_1year_income_avg | 31092.99 | 25467.59 | 5625.40 |
| income_entity_id_2year_income_max | 55549.34 | 60450.08 | 4900.74 |
| income_entity_id_all_income_avg | 21422.38 | 25961.07 | 4538.68 |
| income_entity_id_2year_income_avg | 53806.53 | 58329.03 | 4522.50 |

| | | | |
|---|---|---|---|
| income_entity_id_all_income_max | 36568.37 | 38077.09 | 1508.72 |
| age_entity_id_all_age_on_asofdate_max | 11519.13 | 12746.92 | 1227.79 |
| time_before_entity_id_3year_days_before_next_bkg_avg | 181.35 | 1366.84 | 1185.49 |
| time_before_entity_id_all_days_before_next_bkg_avg | 192.59 | 1274.91 | 1082.33 |
| release_entity_id_all_days_since_release_min | 44.08 | 324.64 | 280.56 |
| **4. Random Forest** | | | |
| **Feature** | **Label=1** | **Label=0** | **Difference** |
| income_entity_id_1year_income_max | 33369.61 | 25863.15 | 7506.46 |
| income_entity_id_1year_income_avg | 32962.03 | 25457.90 | 7504.13 |
| income_entity_id_all_income_max | 40720.66 | 38055.55 | 2665.11 |
| income_entity_id_all_income_avg | 23987.94 | 25947.76 | 1959.82 |
| age_entity_id_all_age_on_asofdate_max | 11212.32 | 12748.51 | 1536.19 |
| income_entity_id_2year_income_avg | 57005.72 | 58312.43 | 1306.71 |
| time_before_entity_id_3year_days_before_next_bkg_avg | 200.96 | 1366.74 | 1165.78 |
| time_before_entity_id_all_days_before_next_bkg_avg | 186.44 | 1274.94 | 1088.51 |
| income_entity_id_2year_income_max | 59361.92 | 60430.30 | 1068.38 |

| | | | |
|---|---|---|---|
| release_entity_id_all_days_since_release_min | 42.73 | 324.65 | 281.92 |
| **5. Random Forest** | | | |
| **Feature** | **Label=1** | **Label=0** | **Difference** |
| income_entity_id_1year_income_avg | 31704.66 | 25464.42 | 6240.24 |
| income_entity_id_1year_income_max | 32104.13 | 25869.71 | 6234.42 |
| income_entity_id_all_income_avg | 22967.44 | 25953.05 | 2985.61 |
| income_entity_id_2year_income_avg | 56337.33 | 58315.90 | 1978.56 |
| age_entity_id_all_age_on_asofdate_max | 10843.43 | 12750.43 | 1907.00 |
| income_entity_id_2year_income_max | 58714.73 | 60433.66 | 1718.93 |
| time_before_entity_id_3year_days_before_next_bkg_avg | 191.30 | 1366.79 | 1175.49 |
| time_before_entity_id_all_days_before_next_bkg_avg | 180.06 | 1274.98 | 1094.92 |
| income_entity_id_all_income_max | 38533.64 | 38066.90 | 466.75 |
| release_entity_id_all_days_since_release_min | 43.82 | 324.64 | 280.82 |

# *Page intentionally left blank.*

Refer to our source material on the following page.

## 5.8 Source Material

[1] Wallace, Danielle, and Xia Wang. “Does In-Prison Physical and Mental Health Impact Recidivism?” SSM - Population Health 11 (August 1, 2020): 100569. https://doi.org/10.1016/j.ssmph.2020.100569.
[2] Dumont, Dora M., Brad Brockmann, Samuel Dickman, Nicole Alexander, and Josiah D. Rich. “Public Health and the Epidemic of Incarceration.” Annual Review of Public Health 33, no. 1 (April 21, 2012): 325–39. https://doi.org/10.1146/annurev-publhealth-031811-124614.
[3] Rodolfa, Kit. “JoCo Mobile Crisis Response Team.” DSSG. https://github.com/dssg/mlpolicylab_fall22_mcrt3/blob/main/README.md
[4] Pescosolido, Bernice A., Bianca Manago, and John Monahan. “Evolving Public Views On The Likelihood Of Violence From People With Mental Illness: Stigma And Its Consequences.” Health Affairs 38, no. 10 (October 1, 2019): 1735–43. https://doi.org/10.1377/hlthaff.2019.00702.
[5] Craigie, Terry-Ann, Ames Grawert, and Cameron Kimble. “Conviction, Imprisonment, and Lost Earnings: How Involvement with the Criminal Justice System Deepens Inequality.” Brennan Center (September 15, 2020). https://www.brennancenter.org/our-work/research-reports/conviction-imprisonment-and-lost-earnings-how-involvement-criminal
[6] Danielle Wallace and Xia Wang. “Does In-Prison Physical and Mental Health Impact Recidivism.”
[7] Carson, Ann. “Mortality in Local Jails, 2000–2019 – Statistical Tables.” Bureau of Justice Statistics (December 2021). https://bjs.ojp.gov/library/publications/mortality-local-jails-2000-2019-statistical-tables.
[8] Ibid.
[9] “Arrests | Arrests Trends.” Accessed December 12, 2022. https://arresttrends.vera.org/arrests.
[10] Bureau of Justice Statistics. “American Indian and Alaska Natives in Local Jails, 1999-2014.” Accessed December 12, 2022. https://bjs.ojp.gov/library/publications/american-indian-and-alaska-natives-local-jails-1999-2014.
[11] Vera Institute of Justice. “An Unjust Burden.” Accessed December 12, 2022. https://www.vera.org/publications/for-the-record-unjust-burden.
[12] McGuire, Thomas G., and Jeanne Miranda. “Racial and Ethnic Disparities in Mental Health Care: Evidence and Policy Implications.” Health Affairs (Project Hope) 27, no. 2 (2008): 393–403. https://doi.org/10.1377/hlthaff.27.2.393.
[13] As measured by access to outpatient mental health visits, psychotropic medication, and any form of mental healthcare.
[14] Cook, Benjamin Lê, Nhi-Ha Trinh, Zhihui Li, Sherry Shu-Yeu Hou, and Ana Progovac. “Trends in Racial-Ethnic Disparities in Access to Mental Health Care, 2004–2012.” Psychiatric Services (Washington, D.C.) 68, no. 1 (January 1, 2017): 9–16. https://doi.org/10.1176/appi.ps.201500453.
[15] Dollar, Cindy Brooks. “Criminalization and Drug ‘Wars’ or Medicalization and Health ‘Epidemics’: How Race, Class, and Neoliberal Politics Influence Drug Laws.” Critical Criminology 27, no. 2 (June 1, 2019): 305–27. https://doi.org/10.1007/s10612-018-9398-7.
[16] “FEDERAL PRISONS: Information on Inmates with Serious Mental Illness and Strategies to Reduce Recidivism.” Government Accountability Office (February 2018). https://www.gao.gov/assets/gao-18-182.pdf
[17] Prison Policy Initiative. “Recidivism and Reentry.” Accessed December 12, 2022. https://www.prisonpolicy.org/research/recidivism_and_reentry/.
[18] Prison Policy Initiative. “Nowhere to Go: Homelessness among Formerly Incarcerated People.” Accessed December 12, 2022. https://www.prisonpolicy.org/reports/housing.html.
[19] Savat, Sara. “For 50 Years, Mass Incarceration Has Hurt American Families. Here’s How to Change It - The Source - Washington University in St. Louis.” The Source, October 14, 2021. https://source.wustl.edu/2021/10/for-50-years-mass-incarceration-has-hurt-american-families-heres-how-to-change-it/.
[20] Ibid.
[21] Ibid.
[22] Sullivan, Robert. “Mental Health and Criminal Justice - Case Study: Johnson County, Kansas.” Accessed December 12, 2022. https://www.jocogov.org/sites/default/files/files/2021-10/cmo-johnson-county-mental-health-jails-case-study.pdf.
[23] Olathe is the largest city in Johnson County. Its current population is 143,014, per the 2021 5-year ACS. (https://www.census.gov/quickfacts/olathecitykansas)

[24] Ibid.
[25] Ibid.
[26] National Association of Counties. "Mental Health and Criminal Justice - Case Study: Mecklenburg County, NC." Accessed December 13, 2022. https://www.naco.org/sites/default/files/documents/Mecklenberg%20County%20Case%20Study%20FINAL.pdf
[27] Health here is taken broadly, but more than half of the program's participants take medication for mental health conditions.
[28] Ibid.
[29] Ibid.
[30] Mattu, Julia Angwin, Jeff Larson,Lauren Kirchner,Surya. "Machine Bias." ProPublica. Accessed December 13, 2022. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
[31] Government Accountability Office. "FEDERAL PRISONS: Information of Inmates with Serious Mental Illness and Strategies to Reduce Recidivism." February 2018. Accessed December 5, 2022. https://www.gao.gov/assets/gao-18-182.pdf
[32] Gray, Colin, Adam Leive, Elena Prager, Kelsey B. Pukelis, and Mary Zaki. "Employed in a SNAP? The Impact of Work Requirements on Program Participation and Labor Supply." National Bureau of Economic Research, June 2021. https://doi.org/10.3386/w28877.