**



Fairness in Recidivism

Prediction:

A Data-Driven Approach

A Comprehensive Analysis and Model Development

Approach for Fairer Criminal Justice Decision-Making

**Prepared for:**

Florida Department of Corrections

**Prepared** **by:**

João Pinto

Data Scientist

Awkward Problems Solutions™

***23rd of June 2024***

# Table Of Contents

[**Table Of Contents 2**](#_h7iu9kcwg2hi)

[**1. Client requirements 2**](#_a1urbwwtw60r)

[1.1 Summary 2](#_cuenmm8f7du8)

[**2. Dataset analysis 2**](#_u7bqwphaydaa)

[2.1 General analysis 2](#_ksyxaqbeg2va)

[2.2 Business questions analysis 4](#_ehucbpozyeok)

[2.3 Recommendations 4](#_dbvykaypk5t2)

[**3. Modelling 4**](#_qdc15q24lfzg)

[3.1 Model specifications 4](#_6kpnymbaau8s)

[3.2 Model performance and expected outcomes 5](#_wxzo0eyfmz9x)

[3.3 Alternatives considered 5](#_fd6abcv7t03x)

[**4. Model Deployment 5**](#_crxjsqjiqobd)

[4.1 Deployment specifications 5](#_6qxoj3yo3xvf)

[4.2 Known issues and risks 5](#_fjg8ieulds8f)

[**5. Annexes 6**](#_e1ja9be29svj)

# Client requirements

## 1.1 Summary

This project focuses on enhancing our client's ability to predict recidivism risk among criminal defendants through a data-driven model and a responsive API. The client, Florida Department of Corrections, seeks an analysis of their dataset and the development of a model that can provide real-time predictions to improve judicial decision-making.

The initial task involves a detailed examination of the provided dataset, which will aim to uncover trends and patterns, particularly looking at variables like gender, age, and race. Addressing potential biases in the recidivism scores and suggesting metrics to evaluate and reduce these biases is a crucial part of this task. The next step is to create a predictive model that determines the likelihood of recidivism, providing a True/False outcome for each defendant. This model needs to strike a balance between accuracy and fairness, ensuring that predictive performance does not come at the cost of unjust bias.

In addition to the model, an API will be developed with two main functions: "will\_recidivate/" and "recidivism\_result/." The "will\_recidivate/" endpoint will take defendant data and return a prediction about recidivism risk, while the "recidivism\_result/" endpoint will accept true recidivism outcomes to help validate model predictions. The API must manage data efficiently and respond accurately to these requests.

The project also includes comprehensive reporting, which involves submitting a proposal outlining the planned approach for analysis and modeling, as well as a final report. The final report will detail the dataset analysis, the development process of the model, and the results, providing clear justifications for the chosen methodologies. The report should be accessible to both technical and business readers.

# Dataset analysis

## 2.1 General analysis

This section presents a comprehensive overview of the dataset utilized for developing a predictive model to assess recidivism risk. The dataset includes information on criminal defendants, encompassing socio-demographic details, criminal history, and COMPAS assessment scores. The analysis focuses on synthesizing this information to provide insights that are useful for business decision-makers.

**Dataset Overview**

The dataset contains several key variables for each defendant, which are crucial for understanding recidivism patterns and developing predictive models. These variables include:

* Socio-Demographic Information: Includes details such as the defendant's age, gender, and race.
* Criminal History: Includes data on previous juvenile and adult criminal activity.
* Criminal History: Includes the number of juvenile felonies, misdemeanors, and other convictions, as well as the number of prior crimes committed as adults.
* COMPAS Scores: These scores range from 1 to 10, with higher scores indicating a higher risk of recidivism. The scores are divided into categories: "Low" (1-4), "Medium" (5-7), and "High" (8-10).
* Recidivism cases: Includes information related to a convict’s criminal activity within two years after the current jail period, including general recidivism and violent recidivism.

**Visualizations and Key Insights**

Recidivism Distribution

A graph with blue rectangular bars

Description automatically generated

The recidivism distribution (our target for this project) chart shows that around half of the convicts have committed recidivism,

Gender Breakdown

A blue rectangles and a black rectangle

Description automatically generated

The gender distribution indicates a higher proportion of male defendants. This disparity is crucial for understanding and addressing potential biases in recidivism predictions.

Charge Description Breakdown

A graph with blue and white lines

Description automatically generated

The distribution of charge descriptions associated with higher recidivism occurrences provides insights into the types of offenses that commonly lead to repeated criminal behavior. Offenses such as "Driving License Suspended," "Possess Cannabis/20 Grams Or Less," and "Resist/Obstruct W/O Violence" appear frequently among individuals who reoffend. These charges often involve non-violent or lower-level offenses that may reflect underlying socio-economic challenges or substance-related issues contributing to recidivism, where offenses related to driving violations, drug possession, and minor theft offenses are prevalent.

**Racial Composition**

A screenshot of a graph

Description automatically generated

The analysis of recidivism rates across different racial groups reveals notable disparities that are essential to consider in the context of criminal justice reform. African-American defendants, for instance, exhibit the highest recidivism rate at 54.73%, indicating that over half of this demographic group reoffends within the specified timeframe. In contrast, Hispanic defendants have a recidivism rate of 38.64% and Caucasian defendants at 41.96%. Asian and Native American defendants have a recidivism rate of 42.31% and 53.85%, respectively. However, these are not representative at they originate from a very small sample, as seen in a graph. Finally, the "Other" racial category shows the lowest recidivism rate at 36.92%.

These disparities are crucial because they highlight potential inequities within the criminal justice system. Higher recidivism rates among African-American and Native American defendants suggest underlying social and systemic factors that may contribute to their increased likelihood of reoffending. Such insights emphasize the importance of developing policies and interventions that address these disparities, ensuring fairness and equity in the treatment of all individuals within the criminal justice system.

Understanding these racial disparities is not only essential for improving predictive models but also for shaping policies that aim to reduce recidivism rates overall. By acknowledging and addressing these differences, policymakers and stakeholders can work towards creating a more just and equitable criminal justice system that supports rehabilitation and reduces the cycle of incarceration for all racial and ethnic groups.

## 2.2 Business questions analysis

The primary focus of our investigation revolves around examining the potential biases inherent in the recidivism prediction algorithm previously employed by the Florida Department of Corrections. This algorithm was designed to assess the likelihood of individuals reoffending based on various demographic and criminal history factors. There have been concerns raised regarding the fairness of this algorithm, particularly with respect to its impact on different ethnic and socio-demographic groups. Our task is to conduct a thorough analysis to ascertain whether these concerns are substantiated by empirical evidence.

One of our key objectives is to evaluate whether there are disparities in the recidivism predictions made by the algorithm across different ethnicities and socio-demographic backgrounds. This involves examining the data to determine if certain groups, such as African-Americans, Hispanics, Caucasians, Asians, Native Americans, and others, are disproportionately classified as high-risk or low-risk for recidivism. Our analysis will not only focus on the overall distribution of recidivism scores among these groups but also delve deeper into specific cases where disparities may be evident. By scrutinizing these patterns, we aim to identify any systematic biases that may exist and propose adjustments to mitigate such disparities.

In addition to ethnicity, we will explore how socio-demographic factors such as age, gender, and possibly socioeconomic status influence the recidivism predictions. This analysis seeks to uncover whether the algorithm exhibits differential treatment based on these variables. For instance, we will investigate whether younger defendants are more likely to be assigned higher recidivism scores compared to older individuals with similar criminal histories. Similarly, we will examine gender disparities in recidivism predictions to determine if there are systematic biases that need addressing.

Another critical aspect of our investigation involves analyzing the algorithm's predictions in relation to different categories of criminal charges. We will categorize offenses based on their severity and type, ranging from minor misdemeanors to serious felonies. By examining recidivism predictions across these categories, we aim to identify any trends or biases that may exist. For example, we will assess whether individuals charged with non-violent offenses are consistently classified differently from those charged with violent crimes, and whether such categorizations align with actual recidivism rates observed in the data.

To effectively evaluate the fairness and accuracy of the algorithm, we will propose and apply specific metrics and methodologies. These metrics will help us quantify disparities in predictions across different groups and identify areas where adjustments may be necessary. Metrics such as disparate impact ratio, predictive parity, and equal opportunity difference will be employed to assess whether the algorithm's outcomes are equitable across ethnicities and socio-demographic groups. Additionally, we will leverage statistical techniques such as regression analysis and machine learning models to validate the predictive performance of the algorithm while ensuring fairness.

Based on our findings, we will provide actionable recommendations for refining the recidivism prediction algorithm to enhance fairness and accuracy. These recommendations may include recalibration of predictive models, adjustment of thresholds for risk classifications, or incorporating additional variables that mitigate biases. Our goal is to propose modifications that align with ethical standards and regulatory requirements, ensuring that the algorithm contributes positively to decision-making processes within the criminal justice system.

In conclusion, our investigation into the recidivism prediction algorithm used by the Florida Department of Corrections is guided by a commitment to fairness, transparency, and accountability. By addressing concerns of bias and disparities across ethnicities, socio-demographic factors, and criminal charge categories, we aim to deliver a comprehensive analysis that informs responsible algorithmic decision-making. Through rigorous evaluation and informed recommendations, we endeavor to support the Department in its mission to uphold justice and equity in criminal justice practices.

## 2.3 Recommendations

{**Audience:** Business

**Description:** In this section, you should give your professional opinion about more subjective topics, such as what specialities may merit further investigation or interpretations of your findings. Any recommendation to your client is expected to be done here.

**Requirement:** Mandatory  
**Page limit:** 0.5 pages}

# Modelling

## 3.1 Model specifications

{**Audience:** Technical

**Description:** This is the section the IT department will read to understand your proposed model implementation and the know-how to replicate it. Describe your proposal and briefly justify it, with the expectation that your reader is someone who knows data science, but not necessarily Python. Since they might want to redo your setup in another language, aim for a more generic and explicit description of your setup and avoid sticking too much to Python specifics (without having to worry about the natural slight differences in implementation between the various languages). Discussion about treatments to address biases (if there are any) in your data should be mentioned here.

**Requirement:** Mandatory

**Page limit:** 1 page}

## 3.2 Model performance and expected outcomes

{**Audience:** Technical

**Description:** Present here your model prototype and preliminary results. Based on this, describe what outcomes you expect from the production run and set justified expectations on model performance (and other topics you consider relevant). Consider all the client requirements at this point and mention some of the critical thinking that leads to your implementation choice. Discussion about model choices or tradeoffs should be mentioned here.

**Requirement:** Mandatory  
**Page limit:** 1 page}

## 3.3 Alternatives considered

{**Audience:** Technical

**Description:** What alternatives did you consider prior to deciding on a model? Why were those alternatives discarded?

**Requirement:** Mandatory  
**Page limit:** 0.5 page}

# Model Deployment

## 4.1 Deployment specifications

{**Audience:** Technical

**Description:** Explain to a technical audience how to replicate the model deployment. Specifications on your application structure, and the structure of the data it receives, should be clarified here.

**Requirement:** Mandatory  
**Page limit:** 1 page}

## 4.2 Known issues and risks

{**Audience:** Technical

**Description:** Regarding your application and your implemented model currently in production, explain what issues and risks you expect. What would potentially break your application? Be critical about your model and the expectations that you set. Be mindful that a “no risk model” does not exist.

**Requirement:** Mandatory

**Page limit:** 0.5 page}

# Annexes

{**Audience:** Technical

**Description:** All additional information used to support your conclusions that could not be included in the main report body. Large figures and tables can go here to keep within the report page limitations

**Requirement:** Optional

**Page limit:** 5 pages}