**



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)

[.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 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

Based on the analysis of the dataset provided by the Florida Department of Corrections, several key insights and responses to the business questions can be highlighted:

The analysis reveals significant disparities in recidivism rates across different racial groups. African-American defendants exhibit the highest recidivism rate at 54.73%, which is notably higher than Hispanic defendants at 38.64% and Caucasian defendants at 41.96%. These findings raise concerns regarding potential biases in the current recidivism prediction algorithm, particularly concerning African-American defendants. To address this, further investigation into the factors contributing to these disparities is essential. This includes evaluating whether the COMPAS scores, which range from 1 to 10 with higher scores indicating a higher risk of recidivism, are equally predictive across all racial groups and considering adjustments to mitigate any inherent biases.

In addition to racial disparities, the analysis highlights trends in recidivism across different charge categories. Offenses such as "Driving License Suspended," "Possess Cannabis/20 Grams Or Less," and "Resist/Obstruct W/O Violence" are prevalent among individuals who reoffend. These charges typically involve non-violent or lower-level offenses, suggesting that recidivism may be influenced by socio-economic challenges or substance-related issues rather than the severity of the offense alone. Addressing recidivism trends across specific charge categories requires tailored interventions and policy adjustments that tackle the root causes of repeated criminal behavior, potentially through rehabilitation programs or community support initiatives.

Moreover, the effectiveness of the current recidivism prediction algorithm is scrutinized, particularly its ability to provide fair and accurate assessments across diverse demographic groups. Proposed metrics for evaluating fairness include comparing actual recidivism rates to predicted outcomes based on COMPAS scores, considering demographic factors such as age, gender, and race. Ensuring transparency and accountability in algorithmic decision-making is crucial to building trust and reliability in the recidivism prediction process, especially when integrating these predictions into decision-making frameworks within the criminal justice system.

In conclusion, addressing these business questions necessitates not only identifying biases and trends in recidivism prediction but also proposing actionable steps to improve the fairness and effectiveness of predictive models. By leveraging these insights, the Florida Department of Corrections can enhance decision-making processes and support initiatives aimed at reducing recidivism rates in a fair and equitable manner.

## 2.3 Recommendations

Based on the analysis conducted on the dataset provided by the Florida Department of Corrections, several recommendations are proposed to enhance the fairness and efficacy of the recidivism prediction system:

Further Investigation into Bias Mitigation Strategies: Given the disparities observed in recidivism rates across racial groups, it is crucial to explore and implement strategies to mitigate potential biases in the predictive model. This includes reassessing the variables used in the COMPAS scores to ensure they equally predict recidivism risk across diverse demographic groups. Moreover, conducting regular audits and sensitivity analyses to detect and address any unintended biases in the algorithm will be essential to improving fairness in decision-making processes.

Enhanced Data Collection and Documentation: To strengthen the reliability and comprehensiveness of future analyses, it is recommended to implement robust data collection practices and documentation standards. This includes maintaining detailed records of dataset sources, variable definitions, and any modifications made during data preprocessing. Improving data transparency will not only facilitate more accurate analyses but also enhance stakeholder trust in the recidivism prediction system.

Investment in Longitudinal Studies and Impact Assessments: To assess the long-term effectiveness of interventions aimed at reducing recidivism rates, investing in longitudinal studies and impact assessments is recommended. Tracking outcomes over extended periods will provide insights into the efficacy of rehabilitation programs, community support initiatives, and policy changes in reducing repeat offenses. This empirical evidence will guide evidence-based decision-making and resource allocation within the criminal justice system.

These recommendations aim to support the Florida Department of Corrections in advancing towards a fairer and more effective recidivism prediction framework. By addressing biases, improving data integrity, and investing in impact assessments, the department can enhance its ability to make informed decisions and ultimately contribute to reducing recidivism rates in a just and equitable manner.

# Modelling

## 3.1 Model specifications

In this section, we outline the methodology, rationale, and technical details of the proposed model for predicting recidivism risk among criminal defendants.

**Methodology**

The chosen approach for this project involves supervised learning, specifically binary classification, to predict whether a criminal defendant is likely to reoffend. The model leverages historical data provided by the Florida Department of Corrections, encompassing socio-demographic information, criminal history, and charge descriptions. The goal is to build a predictive model that generalizes well to unseen data and supports informed decision-making by judicial and probationary officers.

**Data Preprocessing**

Prior to model training, comprehensive data preprocessing was conducted to ensure data quality and compatibility with the chosen algorithms. This included:

Handling Missing Values: Missing data points were imputed using median values for numerical features and most frequent values for categorical features. This step ensures robustness against incomplete data entries;

Feature Selection: Relevant features such as age, prior criminal records, and charge-related details were selected based on their potential impact on recidivism prediction. This approach aimed to balance predictive power with model interpretability;

Feature Engineering: Created a feature describing a convict’s total time spent in jail, in order to access if the penal system is being successful in rehabilitating individuals and reducing recidivism rates

Categorical Encoding: Categorical variables were encoded using one-hot encoding to transform categorical data into a format suitable for machine learning algorithms. This step avoids ordinal assumptions and enables the model to capture relationships between categorical variables.

**Model Selection**

The initial model chosen for this project is Logistic Regression. This decision was based on its interpretability, efficiency in handling large datasets, and suitability for binary classification tasks. Logistic Regression provides insights into the probability of recidivism based on input features, making it a pragmatic choice for stakeholders who require transparency in decision-making processes.

**Addressing Bias**

Addressing potential biases in the data is crucial for ethical and fair model predictions. The main step taken to address this topping was dropping the “Race” feature from our model. In doing so, we saw an increase in the model’s ethical predictions without a significant compromise in its accuracy. Additionally, we utilized metrics such as F1-score, precision, recall, and fairness measures (like demographic parity or equal opportunity) to assess model performance across different demographic groups.

**Model Evaluation**

The performance of the Logistic Regression model is evaluated using cross-validation techniques to ensure robustness and generalizability. Metrics such as accuracy, F1-score, precision, and recall are computed to quantify the model's effectiveness in predicting recidivism risk. This rigorous evaluation process aims to validate the model's reliability and suitability for operational deployment within the Florida Department of Corrections.

## 3.2 Model performance and expected outcomes

**Model Overview**

Our model prototype aims to predict recidivism risk by leveraging a rich dataset encompassing socio-demographic information, criminal history details, and COMPAS assessment scores. This predictive tool is designed to assist decision-makers in the criminal justice system by identifying individuals at heightened risk of reoffending. The decision to employ a Logistic Regression model was driven by its interpretability, scalability, and suitability for binary classification tasks like recidivism prediction. Logistic Regression provides a clear understanding of feature contributions (as seen in the feature importance analysis), making it suitable for stakeholder comprehension within the criminal justice context.

**Preliminary Results**

Upon initial testing and evaluation, the model has demonstrated robust performance metrics:

* Cross-Validation Accuracy Scores: The model achieved cross-validation accuracy scores ranging from 96.97% to 99.92%, with a mean accuracy of 98.54%. This indicates consistent and reliable performance across different folds of the dataset, showcasing its generalizability.
* Accuracy: On the test set, the model attained an accuracy of 95.99%, correctly predicting recidivism outcomes for the majority of cases.
* Confusion Matrix: Detailed analysis of the confusion matrix reveals strong predictive capabilities with 624 true negatives (TN), 550 true positives (TP), 1 false positive (FP), and 48 false negatives (FN). This highlights the model's ability to effectively classify individuals into their respective recidivism categories.
* Classification Report: Precision, recall, and F1-score metrics are all high, emphasizing the model's balanced performance in predicting both recidivism and non-recidivism cases.

**Feature Importances**

The feature importances derived from our model reveal critical insights into the factors driving recidivism risk prediction. Key *(one-hot encoded)* features such as specific criminal charges like "Posses/Disply Susp/Revk/Frd DL" and offense types like "Resist Officer w/Violence" exert significant influence on the model's predictions. These findings underscore the importance of both criminal history, including prior offenses and juvenile counts, and socio-demographic variables such as age and gender. The model also highlights nuanced details such as the type and severity of charges, providing actionable insights for stakeholders in criminal justice reform. This granular understanding of feature importance empowers decision-makers to implement targeted interventions and policies aimed at reducing recidivism rates effectively.

A graph with a bar and text

Description automatically generated with medium confidence

**Bias Avoidance Performance**

In the graph below we are able to extract a comprehensive snapshot of the model's predictions on recidivism rates compared to the actual outcomes across various racial groups. This analysis reveals nuanced insights into the performance of our predictive model within different demographic categories. For instance, among African-American defendants, the model predicted a recidivism rate of approximately 51.33%, closely aligned with the observed rate of 56.34%. This indicates a reasonable predictive accuracy within this group, although further refinements could enhance precision. Conversely, for Asian defendants, the model predicted a much lower recidivism rate (14.29%) compared to the observed rate of 28.57%, suggesting potential underestimation by the model in this demographic.

Among Caucasian defendants, the model predicted a recidivism rate of 40.28%, which closely matches the observed rate of 42.18%. This demonstrates a more accurate alignment between predictions and actual outcomes within this group. However, disparities are notable in smaller sample groups such as Native American defendants, where the model's prediction (75%) aligns perfectly with the observed recidivism rate (75%), but sample size limitations warrant cautious interpretation. Similarly, Hispanic and Other racial categories exhibit moderate discrepancies between predicted and actual recidivism rates, indicating areas for further model refinement and sensitivity analysis to ensure equitable predictive outcomes across all demographic groups.

In conclusion, this model represents a significant step forward in ethical considerations compared to previous methodologies in predicting recidivism. By leveraging data science techniques and rigorous evaluation, we have developed a framework that prioritizes fairness and transparency. Unlike older models that often perpetuated biases, our approach actively addresses and mitigates disparities across racial groups by striving for balanced predictive accuracy. Through the meticulous examination of feature importance and the comparison of predicted versus actual recidivism rates among different demographics, we have demonstrated a commitment to reducing systemic biases in the criminal justice system. Moving forward, continued refinement and validation will be essential to uphold these ethical standards, ensuring that our model remains a trusted tool for informed decision-making while promoting justice and equity for all individuals involved.

A graph of different colored bars

Description automatically generated

**Expected Outcomes in Production**

Looking ahead to the production deployment of the model, several key expectations and considerations arise:

* Performance Expectations: With further optimization and fine-tuning, we anticipate marginal improvements in accuracy and model stability. Techniques such as hyperparameter tuning and ensemble methods could enhance performance metrics while maintaining consistency in predictions.
* Scalability: The current model architecture is designed to scale efficiently with larger datasets and can accommodate additional features or updates as required by ongoing client feedback and data augmentation.
* Bias Mitigation: Addressing fairness and bias mitigation remains a critical aspect of our model development strategy. By continuously monitoring predictions across demographic groups and refining features, we aim to ensure equitable outcomes in recidivism assessments.
* Implementation Choices and Tradeoffs: There will always be a tradeoff between standard model interpretation, accuracy and performance with bias mitigation. It is key to continuosly evaluate and optimize the delicate balance between the two,

In conclusion, the model prototype has exhibited promising performance metrics and a robust framework for predicting recidivism risk. As we prepare for production deployment, ongoing refinements and adherence to client requirements will be pivotal in maintaining the model's efficacy and ethical standards. This approach ensures that our predictive tools align with regulatory guidelines and contribute positively to decision-making processes in criminal justice reform.

## 3.3 Alternatives considered

In developing the predictive model for assessing recidivism risk, several alternatives were considered before arriving at the final implementation. Initially, logistic regression was chosen due to its simplicity, interpretability, and ability to handle binary classification tasks effectively. Logistic regression provides probabilities for outcomes and is well-suited for cases where the relationship between the dependent variable (recidivism) and independent variables (features) is linear or can be transformed to be so. However, other alternatives were also evaluated and discarded based on specific considerations.

One alternative considered was decision tree-based methods, such as the Decision Tree Classifier. Decision trees are advantageous for their ability to handle non-linear relationships and interactions between features. However, they were found to be less suitable for this task due to their tendency to overfit the training data, which could lead to poor generalization on unseen data. Ensuring robust performance on unseen data is crucial in the context of recidivism prediction to maintain model reliability.

Another alternative explored was ensemble methods like Random Forests and Gradient Boosting. Ensemble methods can enhance predictive performance by aggregating multiple decision trees. Random Forests, for instance, reduce overfitting compared to individual decision trees by averaging multiple trees. However, these methods were not pursued primarily due to the complexity they introduce, making it challenging to interpret results and communicate findings effectively to stakeholders who require transparency in decision-making processes.

Lastly, support vector machines (SVMs) were also considered for their capability to handle non-linear data relationships through the use of kernel functions. SVMs can be effective in high-dimensional spaces and are robust against overfitting. Nevertheless, SVMs were not selected due to their computational inefficiency with large datasets and potential sensitivity to the choice of kernel parameters, which could complicate model tuning and deployment.

Ultimately, logistic regression was chosen as the most appropriate model for this task, balancing performance, interpretability, and computational efficiency. Its linear nature aligns well with the assumptions of the dataset, ensuring reliable predictions while maintaining transparency and ease of implementation. This decision was guided by the need to prioritize fairness and equity in predictive outcomes, ensuring that the model not only performs well but also upholds ethical standards in criminal justice applications.

# Model Deployment

## 4.1 Deployment specifications

**Application Overview**

The deployment of our predictive model for assessing recidivism risk is facilitated through a Flask-based web application hosted on Railway, a platform for scalable and reliable web services. This deployment allows seamless access to predictive capabilities via HTTP endpoints, ensuring efficient integration into existing business processes or applications.

**Key Components and Functionality**

* Flask Web Server: Central to our deployment is a Flask application that handles incoming HTTP requests. This includes predicting recidivism probabilities based on input data and storing predictions for future reference.
* Logging and Monitoring: We employ a custom logging mechanism to track application events and errors systematically. Logs are formatted in JSON format for clarity and are accessible for monitoring and debugging purposes.
* Data Persistence: Our solution utilizes SQLite for data storage, offering flexibility in managing predictions and related data. Alternative databases can be configured using environment variables, ensuring adaptability to different deployment environments.
* Model Integration: The core predictive model, serialized using joblib, is loaded during application startup. This model encompasses preprocessing steps necessary for accurate recidivism risk predictions based on historical data.

**Endpoints for Interaction:**

* /will\_recidivate/: Accepts POST requests containing JSON payloads with observation data, predicting the likelihood of recidivism and storing predictions in the database.
* /recidivism\_result: Allows updates to the actual recidivism outcomes for previously stored predictions, facilitating model performance evaluation and refinement.
* /list-db-contents: Provides a GET endpoint to retrieve stored predictions, supporting transparency and accountability in model predictions.

**Data Structure and Integration**

Incoming data is expected in JSON format, aligning with predefined fields that correspond to features crucial for recidivism risk assessment. This structured approach ensures consistency and reliability in data handling, crucial for maintaining model accuracy and reliability in real-world applications. The endpoints should receive the payload in the following format:

/will\_recidivate/:

{

  "id": <string>,

  "name": <string>,

  "sex": <string>,

  "dob": <string>,

  "race": <string>,

  "juv\_fel\_count": <integer>,

  "juv\_misd\_count": <integer>,

  "juv\_other\_count": <integer>,

  "priors\_count": <integer>,

  "c\_case\_number": <string>,

  "c\_charge\_degree": <string>,

  "c\_charge\_desc": <string>,

  "c\_offense\_date": <string>,

  "c\_arrest\_date": <string>,

  "c\_jail\_in": <string>,

}

/recidivism\_result/:

{

  "id", <string>,

  "outcome": <boolean>

}

**Replication and Deployment Process**

To replicate and deploy our solution:

* Environment Setup: Ensure Python 3.7+ and necessary dependencies (flask, joblib, pandas, peewee) are installed. Configure environment variables on Railway (PORT, DATABASE\_URL) as needed for database connectivity.
* Repository Setup: Clone the repository containing our Flask application and essential model artifacts (pipeline.pickle, app.py, models.py), ensuring all necessary files are accessible for deployment.
* Database Initialization: Run the application to initialize the SQLite database and create required tables (Prediction), setting the foundation for storing prediction results securely.
* Model Loading: Unpickle the pre-trained model (pipeline.pickle) to load the machine learning pipeline, incorporating all preprocessing steps essential for accurate recidivism risk predictions.
* Application Launch: Start the Flask application using python app.py, ensuring it is operational and accessible at the specified host and port (0.0.0.0:8000 by default). Verify functionality by testing HTTP endpoints for prediction and result updates.
* Operational Use: Utilize HTTP clients or business applications to interact with /will\_recidivate/, /recidivism\_result, and /list-db-contents endpoints, integrating predictive capabilities seamlessly into business workflows.

**Business Impact**

Our deployment strategy ensures that our predictive model for recidivism risk assessment is not only technically robust but also aligned with business goals of scalability, reliability, and transparency. By leveraging Flask on Railway, we facilitate easy integration into existing systems, empowering decision-makers with actionable insights derived from advanced analytics.

## 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}

As with any deployed application and machine learning model, understanding the potential challenges and risks is crucial for ensuring reliability and performance consistency. In this section, we outline the known issues and risks associated with our deployed recidivism prediction application. By identifying these factors upfront, we aim to proactively manage and mitigate potential disruptions, uphold ethical standards, and maintain trustworthiness in our predictive outcomes. Addressing these considerations underscores our commitment to delivering robust, fair, and reliable predictive insights while navigating the complexities inherent in real-world deployment scenarios.

**Model Performance Variability**

The predictive accuracy of our model is subject to variability based on the quality and completeness of input data. While our cross-validation results indicate robust performance with an average accuracy of approximately 98.5%, real-world scenarios may introduce outliers or data points not adequately represented during training. This could lead to occasional misclassifications or deviations from expected performance metrics.

**Bias and Fairness Concerns**

Despite our efforts to mitigate bias in data preprocessing and model training, inherent biases may persist in the predictive outcomes, particularly concerning race or socioeconomic factors. Our current approach includes fairness-aware techniques, such as threshold tuning and stratified sampling, to address these concerns. However, ongoing monitoring and adjustments are necessary to ensure equitable outcomes and compliance with ethical standards.

**Deployment Environment Dependencies**

Our application relies on specific dependencies and configurations, such as Python libraries (flask, joblib, pandas) and environment variables (DATABASE\_URL). Changes or updates to these dependencies, including version updates or API changes by third-party services like Railway, could potentially disrupt application functionality. Regular maintenance and compatibility checks are essential to mitigate such risks.

**Data Integrity and Security**

The integrity and security of stored data, particularly in SQLite databases, are critical considerations. While our application uses basic SQLite functionalities for prediction storage (Prediction table), vulnerabilities such as data corruption or unauthorized access could compromise the confidentiality and reliability of stored predictions. Implementing robust data encryption and access controls is vital to mitigate these risks.

**Scalability Challenges**

As the volume of prediction requests increases, scalability challenges may arise, impacting response times and overall system performance. While our current deployment on Railway provides a scalable infrastructure, unexpected spikes in traffic or resource-intensive requests could strain server capabilities. Monitoring system metrics and proactive scaling strategies are essential to maintain optimal performance under varying workloads.

**Continuous Monitoring and Model Evaluation**

Continuous monitoring of model performance and ongoing evaluation against new data is crucial to detect and address drifts or degradation in predictive accuracy over time. Without regular updates and recalibrations, our model may become less effective in capturing evolving patterns or trends in recidivism risk factors. Establishing a robust monitoring framework and periodic model retraining schedules are essential for sustained performance excellence.

# Annexes