Are you more interested to into know with certainty and precision if an individual will commit a crime, with the trade-off not accounting for all individuals who will commit a crime but weren't predicted as such by the model. Or are you more interested into identifying the greater number of individual that are likely to commit a crime, with the downside of sometimes labeling innocent people as recidivates? , this last one might contribute in the racial group bias. Since I saw that your dataset is unbalance, posibily labeling black people as reoffenders.

Just to clarify, Does the column called is\_recid has the true (real) value if they have commited a crime after release (labeled as 1) or if they haven’t (labeled as 0), and does the column called **two\_year\_recid** is the real (true) value if they recidivated during the two years after they were release (labeled as 0 if not, and 1 if yes) or does it come from the predicted value from the other model? .

**Regarding the data I have some doubts:**

There is a column called names, are does the actual names of the convicts? Shouldn’t be the names anonymized due to data privacy concerns?

To clarify, is dob, day of birth?

For the race, did the data was obtained by self-identification of the convict or by what the police officer thinks?

C\_case\_number: is it unique, or the case is reopened if the convict is detained and/or goes to jail again? It changes when reincide.

Regarding c\_offence\_date and c\_arrest\_date:

When offense date is missing there is usually a value for arrest date. However some rows are missing both, does that mean that they were never arrested but they commited crime? Or does it mean that they didn’t commit a crime?

What does the values for c\_charge\_degree (F, M…) stand for?

C\_charge\_desc , was the description made by the police or by the victims?

What is juv\_fel\_count and is it 10 really the max of felonies someone can commit ?

Juv misd count , is 13 a fixed roof of misdiminors that can be commited?

Priors\_count , does it mean the times the covicts have been in jail before or being arrested or being denounced by victims?

What is compass\_screeing\_date?

Is type\_of\_assesment if they will commit any criminal activity (like from petty crime to murder)?

In theory what decile\_scores correspond to each score\_text?

What does v\_type\_of\_assessment means?

What in theory what v\_decile\_scores correspond to each v\_score\_text?

Are the columns starting with vr\_ mean violence reincide?