Document Title:	Team MattMarifelSora: NIJ Recidivism Forecasting Challenge Report
Author(s):	Matthew Motoki, Marifel Barbasa, Sorapong Khongnawang
Document Number:	305044
Date Received:	July 2022
Award Number:	NIJ Recidivism Forecasting Challenge Winning Paper

Matthew Motoki
Marifel Barbasa
Sorapong Khongnawang

Team MattMarifelSora
NIJ Recidivism Forecasting Challenge Report

Table of Contents

I. Introduction

The National Institute of Justice (NIJ) defines *recidivism* within the Recidivism Forecasting Challenge as "an arrest for a new crime." The objective of the challenge was to predict

- Stage 1: How likely it was that individuals would recidivate in the first year after their release from prison.

- Stage 2: How likely it was that individuals would recidivate in the second year given they did not recidivate in the first year.
- Stage 3: How likely it was that individuals would recidivate in the third year given they did not recidivate in the first two years.

In each stage, teams were evaluated in five categories: how well they could predict male recidivism, female recidivism, average recidivism, male recidivism with a racial bias penalty, and female recidivism with a racial bias penalty. Teams were partitioned into three types: Student Team, Small Team, and Large Team.

We applied data processing and machine learning techniques to predict how likely it was that individuals would recidivate. We applied hierarchical Bayesian target encoding and trained models that are known to perform well on binary classification and multiclass classification problems involving tabular data. Following the industry standard in machine learning competitions, we combined predictions from many models into an ensemble to boost our score. Our high-level pipeline is shown in Figure 1.

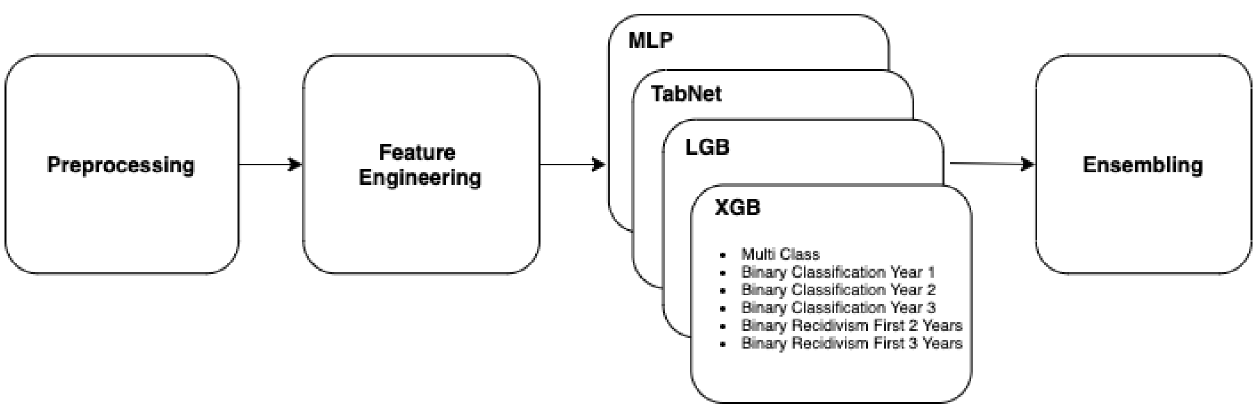

Figure 1. Solution Pipeline

Our team's results are summarized in Table 1.

Table 1. Final competition results for team MattMarifelSora.

Place	Stage	Team Type	Category
1st	1	Large Team	Male Recidivism with the Racial Bias Penalty
2nd	1	Large Team	Male Recidivism
2nd	1	Large Team	Female Recidivism
2nd	1	Large Team	Average Recidivism

II. Relevant Literature

Predictive tooling in criminology as a whole is not a recent initiative. Prediction concerns in criminology and the design of the first risk assessment tool started in the 1920s. Ozkan [1] groups the historical methodology into four generations: highly subjective and unstructured professional judgments (generation 1), empirical-based actuarial assessments (generation 2), theoretically informed tools (generation 3), and treatment matching with a focus on the risk-need-responsivity or RNR model (generation 4).

Ozkan mentions demographics and criminal history as recidivism factors, along with factors such as antisocial tendencies, peer delinquency, and substance abuse. He cited that criminal history is the most important factor, though components of it, such as “frequency, seriousness, and recency,” should also be taken into account. Other individual-level factors include "dominance, entitlement, self-justification, displacing blame, optimistic perceptions of realities, and blaming society" [1].

In an ideal world, we want to maximize the forecasting accuracy and worry about nothing else, because it’s always easier to focus on one metric. But it’s important to also determine the costs associated with false positives and false negatives; these costs are often not equivalent in the real world. In regards to recidivism, predicting more false positives than false negatives leads to classifying more individuals into the high-risk category, which requires an increased number of high-risk prison accommodations and community supervision officers (we assume the author implies that this can be costlier than false negatives for certain non-violent offense types) [1]. On the other hand, in regards to domestic violence, classifying more false negatives can be costlier; the Los Angeles Sheriff's Department found that not responding to a domestic violence call (that actually needs police) is more costly than responding to one (that happens to be a false alarm) [1]. In weighing costs, cost ratios can be considered, but there are no rules of thumb, and it can change depending on jurisdiction. If used, ratios should be adjusted accordingly and over time as policies and needs change [1].

In Ozkan's literature review, he cited that the models used in forecasting recidivism were logistic regression, discriminant analysis, decision trees, random forests, the classification and regression tree (CART) model, stochastic gradient boosting, support vector machines (SVMs), and neural networks, including multi-layer perceptrons (MLPs). There was no substantial conclusion on which models performed best in the literature, as different studies claimed different results [1].

In our work, we utilized gradient boosted decision trees via the XGBoost and LightGBM libraries and created a custom MLP with skip connections using the PyTorch library. Additionally, we used the dreamquark implementation of a modern neural network architecture known as TabNet, which takes advantage of attention mechanisms to selectively focus on input features. We also tried NODE and SVM models, but their performances were notably worse and not included in our pipeline.

On the topic of racial bias, one might think that excluding race alone as a factor can mitigate that bias. As Rudin [2] of the Zeng et al. study [3] claims, when excluding race, accuracy in their machine learning models did not differ significantly as compared to when they included race. However, another study by Dressel and Farid found that, despite isolating the factors to just two variables, age and number of prior crimes, the false positive rate (as likely to recidivate when they didn't) of black men was still much higher compared to white men. This study was compared to ProPublica's 2016 study of the commercially available proprietary COMPAS risk assessment software by Northpointe, which also had a much higher false positive rate for black men but using 137 variables instead of two. Race was not a variable used in COMPAS and the study; the results suggest that the number of prior crimes is a proxy for race [4]. In addition, Ozkan cited in his thesis that "zip codes can serve as a proxy for race" as well when it comes to risk assessment tools [1].

Therefore, as much as we strive to reduce racial bias by searching and optimizing for the right metric, the underlying data, collected throughout an individual's journey through the criminal justice system, has racial biases embedded within. To help combat racial injustice, Farid makes two suggestions in his TED talk on "The danger of predictive algorithms in criminal justice": Let's follow the example set by the General Data Protection Regulation (GDPR), in which European citizens have the right to audit data and algorithms used against or toward them, and let's establish a national or international cyberethics panel of experts akin to bioethics panels established in the 1980s and beyond, when advances in biology and medicine caused struggles with ethics, morality, and religion [5].

Thus, as long as we have bias in initial arrests, those will persist in data that informs recidivism, reflected in the model predictions that utilize such data. Making decisions that can impact the rest of a previously incarcerated person's life carries huge weight. Machine learning methods are known to reproduce biases in the data. Predictive models for recidivism should be used with caution.

III. Variables

NIJ provided challenge participants with a dataset of previously incarcerated individuals from the State of Georgia, under parole supervision between January 1, 2013 through December 31, 2015. The dataset contains 53 variables in total that, aside from the identifier variable, fall within these seven categories [6]:

- **Supervision case information**, including demographics and whether or not the individual is part of a gang
- **Prison case information**, including education level and number of dependents
- **Prior Georgia criminal history**, including number of prior arrests by offense type
- **Prior Georgia community supervision history**, including whether or not the individual has parole and probation revocations
- **Georgia Board of Pardons and Paroles conditions of supervision**, in which certain individuals had to abide by specific conditions set by the Board before they were released for parole
- **Supervision activities**, including violations, drug tests, employment, and other data reflecting the supervision period
- **Recidivism measures**, including target variables: whether or not the individual recidivated per year after their release from prison

We didn't utilize any external datasets, but we did construct variables based on the challenge dataset.

Step 1: Data Preprocessing

We performed common data preprocessing techniques on the raw challenge dataset, which involved replacing missing values, converting string values to integers, and discretizing continuous variables. For more details about the preprocessing that we applied to the variables, see Appendix A.

Step 2: Feature Engineering

Feature engineering is the step where preprocessed data is transformed into a form that is more readily usable by the machine learning model. This can involve constructing new variables (feature columns) out of existing ones that help further inform the model.

We employed hierarchical Bayesian target encoding (HBTE) to construct features. Target encoding converts categorical variables into numerical features by mapping each category to the mean target value within that category. In HBTE, we created additional categories by further partitioning a category according to the hierarchy, as shown in Figure 2. We further made use of this hierarchy to construct Bayesian estimates of the mean target value. See Appendix B for more details.

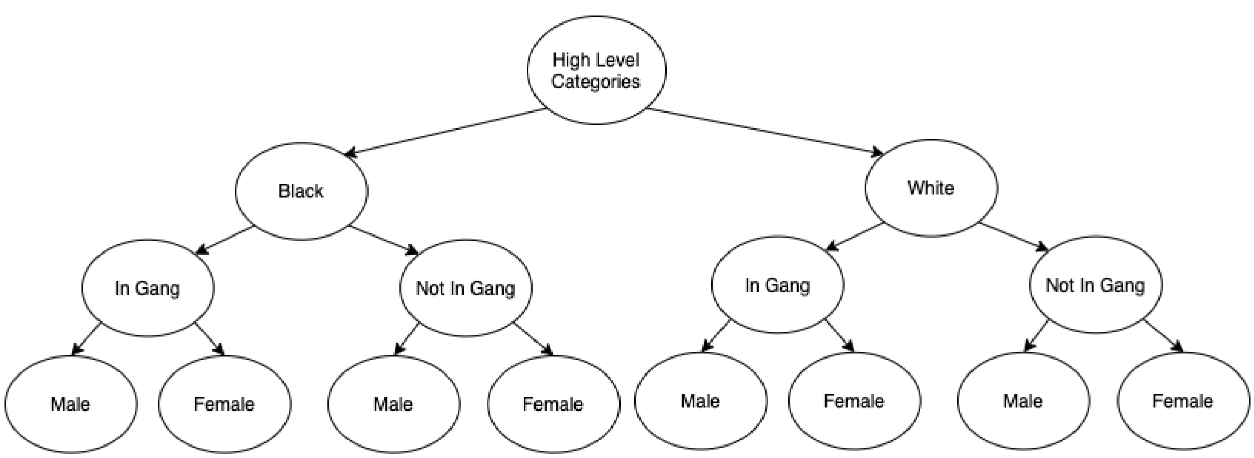

Figure 2. Hierarchical Baysian Target Encoding

Feature Importance

In tree-based models, we measure feature importance using the number of times a variable is used to split our data. This is a common metric used to evaluate feature importance. For brevity, we show only the top 10 features of our best single model (LightGBM); however, the feature importance from our other models were similar.

Table 2. The top 10 most important features from our best single model.

rank	name	importance
0	te1_Age_at_Release	263.44
1	te1_Residence_PUMA	225.95
2	te123_Age_at_Release	209.78
3	te3_Residence_PUMA	182.05
4	te1_Prior_Arrest_Episodes_Felony	170.06
5	te1_Supervision_Risk_Score_First	169.88
6	te123_Residence_PUMA	166.37
7	te1_Prison_Years	153.28
8	te2_Residence_PUMA	152.92
9	te123_Supervision_Risk_Score_First	144.74

Note that the importance includes features that have been target encoded. Features with the “te” prefix indicate that the feature is engineered; the number within the prefix indicates the target variable, according to the following mapping:

- 1: Recidivism_Arrest_Year1
- 2: Recidivism_Arrest_Year2
- 3: Recidivism_Arrest_Year3
- 123: Recidivism_Within_3years

For example, `te1_Age_at_Release` means that the feature is constructed using HBTE with `Age_at_Release` as the base feature and `Recidivism_Arrest_Year1` as the target.

For simplicity, we did not drop any variables from our models. Unlike unregularized linear models, gradient boosted trees and regularized neural networks are relatively robust to redundant features. That said, proper feature selection could possibly further improve our models.

IV. Models

Although we were evaluated on male, female, and average metrics, we decided to optimize specifically for the accuracy metric calculated using the entire dataset. Accordingly, we trained a single model that we used for all three categories. Similarly, we did attempt to optimize for the racial bias penalty. Our thinking was that a single submission cannot simultaneously minimize the penalized and unpenalized metrics; therefore, we chose to focus on the unpenalized metrics that made up the majority of the competition categories.

Stage 1: Predicting Recidivism in Year 1

In the first stage of the challenge, we were tasked with predicting recidivism within the first year. Our final model was a stacked ensemble where the base models were XGBoost, LightGBM, TabNet, and MLP.

We trained 5 versions of each base model with the following binary target variables (dependent variables):

- No recidivism
- Recidivated in year 1
- Recidivated in year 2
- Recidivated in year 3
- Recidivated in year 1 or year 2

We also trained another version of each base model with a multiclass target variable with the following 4 classes:

1. No recidivism
2. Recidivated in year 1
3. Recidivated in year 2
4. Recidivated in year 3

Thus, we had 6 versions of each base model type with 9 predictions (5 binary, 4 multiclass). So in total, we had 24 models and 36 predictions.

We repeated each version of the base models with 100 different seeds and averaged their predictions. We combined the predictions of the base models using a simple two-step procedure:

1. Train an elastic net regression model using the base model predictions as input (36 inputs total). We kept only the predictions associated with the non-zero coefficients. This step dropped all predictions except those for “No recidivism” and “Recidivated in year 1.”
2. Train a ridge regression model using only the base predictions selected in the previous stage (4 predictions for each model type, 2 of which were from binary classification models and 2 were from the multiclass model; thus, 16 predictions total).

We evaluated the performance of our models using repeated stratified cross validation with 10 folds and 10 repeats and tuned hyperparameters using the Optuna library.

Table 3. The model errors for the “Recidivated in year 1” target.

Model	Target	Male Brier Score	Female Brier Score	Average Brier Score
LightGBM	Binary	0.19115	0.15497	0.17306
LightGBM	Multiclass	0.19109	0.15488	0.17298
XGBoost	Binary	0.19098	0.15544	0.17321
XGBoost	Multiclass	0.19107	0.15524	0.17315
MLP	Binary	0.19093	0.1548	0.17286
MLP	Multiclass	0.19089	0.1542	0.17254
TabNet	Binary	0.19129	0.1544	0.17284
TabNet	Multiclass	0.1911	0.1534	0.17225
Final Ridge	Binary	0.19075	0.15429	0.17252

Stage 2: Predicting Recidivism in Year 2

The second stage differed from the first in that we were tasked with predicting recidivism in year 2 given that the parolees did not recidivate within the first year. Our modeling procedure in the second stage followed the modeling procedure in the first stage very closely except that we dropped the parolees who recidivated in year 1 from our dataset.

We trained 3 versions of each base model with the following binary target variables:

- No recidivism
- Recidivated in year 2
- Recidivated in year 3

We also trained another version of each base model with a multiclass target variable with the following 3 classes:

1. No recidivism

2. Recidivated in year 2
3. Recidivated in year 3

Thus, we had 4 versions of each base model with 6 predictions (3 binary, 3 multiclass). So in total, we had 16 models and 24 predictions.

We used the same two-step procedure as stage 1 to combine the base model predictions.

Table 4. The model errors for the "Recidivated in year 2" target.

Model	Target	Male Brier Score	Female Brier Score	Average Brier Score
LightGBM	Binary	0.17241	0.12876	0.15058
LightGBM	Multiclass	0.17272	0.12924	0.15098
XGBoost	Binary	0.17204	0.12976	0.1509
XGBoost	Multiclass	0.172	0.13021	0.1511
MLP	Binary	0.17183	0.12927	0.15055
MLP	Multiclass	0.17127	0.12839	0.14983
TabNet	Binary	0.17287	0.12948	0.15117
TabNet	Multiclass	0.17102	0.12791	0.14947
Final Ridge	Binary	0.17072	0.12854	0.14963

Stage 3: Predicting Recidivism in Year 3

In the final stage, we were tasked with predicting recidivism in year 3 given that the parolees did not recidivate within the first two years. Our modeling procedure was very similar to the procedure in the second stage except that we dropped the parolees who recidivated in year 2 from our dataset. In this stage, the multiclass classification reduces to binary classification, and so only a single binary classification model that predicted "Recidivated in year 3" could be trained. Thus, we had 4 models and 4 predictions in total. We used the same two-step procedure as stage 1 to combine the base model predictions.

Table 5. The model errors for the “Recidivated in year 3” target.

Model	Target	Male Brier Score	Female Brier Score	Average Brier Score
LightGBM	Binary	0.1542	0.11872	0.13646
XGBoost	Binary	0.15368	0.11857	0.13613
MLP	Binary	0.15402	0.11827	0.13614
TabNet	Binary	0.15337	0.11901	0.13619
Final Ridge	Binary	0.15342	0.11854	0.13598

V. Conclusion

We’ve shown that these techniques are capable of producing relatively accurate recidivism predictions. Our models predicted “male recidivism with the racial bias penalty” in year 1 relatively well (achieving 1st place in this category). This is surprising since we did not tune our models for this metric.

The most important variables were not very surprising:

- Age_at_Release
- Residence_PUMA
- Prior_Arrest_Episodes_Felony

Our models performed well in stage 1 but not so well in stages 2 and 3. One reason for the decline may be because we didn’t make use of the individuals who recidivated in years 1 and 2 in stages 2 and 3. One way to possibly improve our models would be to use our predictions from stages 1 and 2 to inform our stages 2 and 3 models.

Although we've shown that these techniques are capable of producing relatively accurate recidivism predictions, and we do believe that the models are intended to be used in good faith, we don't recommend using our model in practice because it will replicate any biases already present within the data. Furthermore, since we took a black-box approach to forecasting recidivism, it's difficult to extract any practical or applied findings that could help the field.

VI. Future Considerations

Although we placed in the year 1 racial bias category, we only trained the model on the full dataset, optimizing for the Brier score. Thus, the fact that the fairness penalty only considered false positives did not affect our submission, to our knowledge. Given more time, we would have performed ablation studies. Lastly, since we did not get to experiment with changing the threshold and stuck with 0.5, our team has no recommendations for using a different threshold at this time.

Our team chose not to optimize for the metric with the racial bias penalty. We believe that a single submission cannot simultaneously minimize the penalized and unpenalized metrics. Therefore, in future competitions, it may be beneficial to allow teams to submit two submissions—one for the penalized metric and one for the unpenalized metric.

For future challenges, NIJ should consider improving them by changing the following: gathering more data from other states and gathering more information, apart from Condition_Cog_Ed, about whether or not parolees received educational programming during their time in prison. A variable highlighting whether or not they signed up for educational programs and were put on the waitlist could be a useful indicator of the parolee wanting to change their lives for the better.

For symmetry, NIJ should also consider false negatives in the penalty function. It would be interesting to consider the penalties based on offense type and gravity of offense. For example, non-violent offense types can probably prioritize the false positive rate, while violent offense types can try prioritizing the false negative rate. One metric that has often been used in the literature is to measure predictive performance using the area under the curve (AUC) [1]. Another metric worth considering is binary cross entropy because whether or not someone recidivates is a classification task. NIJ can also try categorical cross entropy to predict all years at once.

Our complex model is not very interpretable. Researchers opt for simpler, more transparent models that are also capable of achieving the same level of accuracy as more complex state-of-the-art models [2], as researchers have shown in [3]. This example of an interpretable model by Zeng et al. proposes using simpler machine learning models to determine factor coefficients, and then if used in the real world, the model is clear for decision-makers because it’s a simple checklist, where points are applied to certain factors [3]. In this way, these models that make life-altering decisions for parolees can be scrutinized, challenged, and understood by all parties involved, including the parolees themselves. Otherwise, as it stands, most risk assessment tools are still proprietary, making transparency and accountability difficult or impossible [7].

We have to caution using machine learning algorithms, especially when used to help make life-altering decisions. Perhaps we need to shift our focus: Apart from aiming for more accurate models and trying out different measures of fairness and accuracy, we first need to ensure fairness throughout all levels of the criminal justice system, from where we send our police to parole supervision and beyond [8]. This is because at all levels, data is collected, and that is the very data that we end up feeding into our recidivism forecasting models.

Outside of predictive models, what are ways we can focus on reducing recidivism systematically? [8] outlines short-term, medium-term, and long-term reforms with an emphasis on rehabilitation and reintegration into society. Education, employment, housing, and behavioral health rehabilitation is also the focus of [9], in which they emphasize that we need to reframe reentry “from a piecemeal approach to an integrative, ecosystem-based approach.” In addition, the mitigation of criminal record stigma can lead to better outcomes in terms of education, employment, and housing, as reducing stigma “is one of the most important and well-documented barriers to successful reentry and reintegration” [8]. [8] goes on to make a number of stigma reduction suggestions, which includes ending “restrictions on living in publicly subsidized housing for those with criminal records.” Another concern that [8] raises is in regards to the parole and probation system, for which violations “account for large shares of prison admissions in many states,” due to the system’s emphasis on surveillance and punishment.

Lastly, on the dangers of predictive algorithms, Farid reminds us that “technology can be a force for tremendous progress and tremendous good, but as we have seen over the last few years, left unchecked, it can just as well plunge us into a digital dystopia” [5]. And in the subject of her TED talk on “Reducing Recidivism,” Hawes leaves us with this powerful thought: “What was your greatest mistake ever? What if you were reduced to that one moment in time? If that were to happen, your skills, your talents, your gifts, your dreams will not be realized. If you were minimalized to that moment only and defined by that bad decision, would that be fair? Look, we all deserve a second chance, and our community deserves one as well” [10].

Appendix A: Preprocessing Details

Variable Name	Description
Gender	Map “Gender” feature to integer values. Female labels (“F”) are mapped to 0. Male labels (“M”) are mapped to 1.
Race	Map “Race” feature to integer values. White race labels (“WHITE”) are mapped to 0. Black race labels (“BLACK”) are mapped to 1.
Age_at_Release	Map “Age_at_Release” feature to integer values. Age “18-22” labels are mapped to 0. Age “23-27” labels are mapped to 1. Age “28-32” labels are mapped to 2. Age “33-37” labels are mapped to 3. Age “38-42” labels are mapped to 4. Age “43-47” labels are mapped to 5. Age “48 or older” labels are mapped to 6.
Residence_PUMA	Label encode “Residence_PUMA” feature with LabelEncoder from the sklearn library.
Gang_Affiliated	Map “Gang_Affiliated” feature to integer values. Map “missing” labels to 0. Map “False” labels to 1. Map “True” labels to 2.
Supervision_Risk_Score_First	Fill missing values in “Supervision_Risk_Score_First” feature with 0.
Supervision_Level_First	Map “Supervision_Level_First” feature to integer values. Map “missing” labels to 0. Map “Standard” labels to 1. Map “High” labels to 2. Map “Specialized” labels to 3.
Education_Level	Map “Education_Level” feature to integer values. Map “Less than HS diploma” labels to 0. Map “High School Diploma” labels to 1. Map “At least some college” labels to 2.
Prison_Offense	Map “Prison_Offense” feature to integer values. Map empty labels to 0. Map “Other” labels to 1. Map “Property” labels to 2. Map “Drug” labels to 3. Map

	“Violent/Non-Sex” labels to 4. Map “Violent/Sex” labels to 5.
Prison_Years	Map “Prison_Years” feature to integer values. Map “Less than 1 year” to 0. Map “1-2 years” labels to 1. Map “Greater than 2 to 3 years” to 2. Map “More than 3 years” labels to 3.
Avg_Days_per_DrugTest	Discretize a normalization of a log transform of “Avg_Days_per_DrugTest” feature to be in 5 groups, from group 0 to group 4.
Jobs_Per_Year	Transform and group “Jobs_Per_Year” feature to discrete numbers between 0 to 4 jobs per year.
DrugTests_THC_Positive DrugTests_Cocaine_Positive DrugTests_Meth_Positive DrugTests_Other_Positive Percent_Days_Employed	Map empty and 0 to 1. Map 1 to 0.
Dependents Prior_Arrest_Episodes_Felony Prior_Arrest_Episodes_Misd Prior_Arrest_Episodes_Violent Prior_Arrest_Episodes_Property Prior_Arrest_Episodes_Drug Prior_Arrest_Episodes_PPViolationCharges Prior_Conviction_Episodes_Felony Prior_Conviction_Episodes_Misd Prior_Conviction_Episodes_Prop Prior_Conviction_Episodes_Drug Delinquency_Reports Program_Attendances Program_UnexcusedAbsences Residence_Changes	Map string labels to integer values. To perform conversion, first remove the suffix string “or more” then perform integer conversion on the remaining string.
Prior_Arrest_Episodes_DVCharges, Prior_Arrest_Episodes_GunCharges, Prior_Conviction_Episodes_Viol, Prior_Conviction_Episodes_PPViolationCharges, Prior_Conviction_Episodes_DomesticViolenceCharges, Prior_Conviction_Episodes_GunCharges, Prior_Revocations_Parole, Prior_Revocations_Probation, Condition_MH_SA, Condition_Cog_Ed, Condition_Other, Violations_ElectronicMonitoring,	Map boolean values to integer values. “True” labels are converted to 1. “False” labels are converted to 0.

Violations_Instruction, Violations_FailToReport, Violations_MoveWithoutPermission, Employment_Exempt, Recidivism_Within_3years, Recidivism_Arrest_Year1, Recidivism_Arrest_Year2, Recidivism_Arrest_Year3	

Appendix B: Hierarchical Bayesian Target Encoding

In hierarchical Bayesian target encoding (HBTE), we model the target variable as a Bernoulli random variable and we are interested in estimating its parameter p. We use a Beta distribution to model the prior and posterior distributions for p. We parameterize the prior distribution in terms of

- alpha = pseudo count of the number of times the target is a 1
- beta = pseudo count of the number of times the target is a 0

We set the mean of the prior distribution to be the mean of the posterior from one level up in the hierarchy. This gives us the following equation for the prior mean: `mu = alpha / (alpha + beta)`. Next we treat the total pseudo count tau as a hyperparameter that we optimize using cross-validation. This leads to the following formulas for the prior parameters:

- alpha = tau * mu
- beta = tau * (1 - mu)

The posterior parameters are simply

- alpha_prime = alpha + N1
- beta_prime = beta + N0

where

- N1 is the number times the target is a 1 within the category
- N0 is the number times the target is a 0 within the category

The posterior mean is `mu_prime = alpha_prime / (alpha_prime + beta_prime).`

HBTE can be viewed as a very simple model. The model learns the posterior mean for p within a category and uses that value to predict the target variable.

References

[1] Ozkan, T. (2017, May). (dissertation). Predicting Recidivism Through Machine Learning. The University of Texas at Dallas: Eugene McDermott Library. Retrieved from https://utd-ir.tdl.org/bitstream/handle/10735.1/5405/ETD-5608-7434.54.pdf.

[2] Rudin, C. (2015, September 9). New models to predict recidivism could provide better way to deter repeat crime. The Conversation. https://theconversation.com/new-models-to-predict-recidivism-could-provide-better-way-to-deter-repeat-crime-44165.

[3] Zeng, J., Ustun, B., & Rudin, C. (2016). Interpretable classification models for recidivism prediction. Journal of the Royal Statistical Society: Series A (Statistics in Society), 180(3), 689–722. https://arxiv.org/pdf/1503.07810v2.pdf.

[4] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. Science Advances, 4(1). https://advances.sciencemag.org/content/4/1/eaao5580.full.

[5] TEDxTalks. (2018, October 2). The danger of predictive algorithms in criminal justice | Hany Farid | TEDxAmoskeagMillyard. https://www.youtube.com/watch?v=p-82YeUPQh0.

[6] Recidivism Forecasting Challenge. National Institute of Justice. (n.d.). https://nij.ojp.gov/funding/recidivism-forecasting-challenge.

[7] AI is sending people to jail—and getting it wrong. MIT Technology Review. (2019, January 21). https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/.

[8] A better path forward for criminal justice: Prisoner reentry. Brookings. (2021, April). https://www.brookings.edu/research/a-better-path-forward-for-criminal-justice-prisoner-reentry/.

[9] Reducing recidivism. Deloitte Insights. (2018, December 17). https://www2.deloitte.com/us/en/insights/industry/public-sector/reducing-recidivism-ecosystem-approach-for-successful-reentry.html.

[10] TEDxTalks. (2020, December 10). Reducing recidivism | Valerie Hawes | TEDxKingLincolnBronzeville. YouTube. https://www.youtube.com/watch?v=Xpyc9c31XyE.