**Predicting Criminal Recidivism Using Machine Learning Algorithms**

**Milestone Report for Capstone Project 2**

**Springboard Data Science Career Track**

**By**

**Shishay Bisrat**

**February 2018**

**Predicting Criminal Recidivism Using Machine Learning**

**1. Background Information**

Criminal recidivism refers to the act of a person repeating a criminal offense after being released from prison for a similar offense. Statistics show that recidivism is a major issue in the US: The Bureau of Justice Statistics (BJS) reported 67 percent and 76.6 percent of state prisoners released in 2005 were re-arrested with three and five years of release respectively (Durose et al, 2014). The justice system in the US predicts an individual’s recidivism potential when making key decisions such as parole, probation, and sentencing. The court systems rely on a software package called Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) to predict the likelihood of recidivism. The recidivism prediction made by an individual could be biased depending on one’s life experiences and perspectives. Software packages could also yield biased results if the weight of features is inexplicably fixed and provided by biased individuals. Machine Learning algorithms can be useful in limiting biases and improving recidivism predictions (Teich, 2018). Some justice system institutions such as the Pennsylvania Board of Probation and Parole have begun using machine-learning predictions to help in parole release decisions. It is expected that more states will use the promising field of machine learning to limit biases and make useful contribution in accurately forecasting recidivism (Berk, 2016).

**2. Motivation**

Recidivism is a major societal issue that costs lives and resources. An accurate recidivism prediction would have a major impact to individuals as well as the society. Criminal institutions would target their stretched resources on dangerous offenders increasing the safety of citizens. Non-dangerous, one-time offenders, with little likelihood of recidivism would be treated fairly, encouraging them to integrate to society. This would also ensure sensible use of resources strictly on dangerous offenders.

**3. Data and Data Wrangling**

The original data set consisted of 21,646 rows and 17 features.

|  |  |
| --- | --- |
| Features | Possible Values |
| Year Released | 2010, 2011, 2012, 2013, 2014 |
| Recidivism Year | 2013, 2014, 2015, 2016, 2017 |
| Race/ Ethnicity | White (Non-Hispanic), Black (Non-Hispanic), White, American Indian/ Alaska native, Asian or Pacific Islander, Black (Hispanic) |
| Sex | Male, Female |
| Age at release | Under 25, 25-34, 35-44, 45-54, over 55 |
| Convicting Offense Classification/ New Convicting Offense Classification | D Felony, C Felony, Aggravated Misdemeanor, B Felony, Felony – Enhancement to original penalty, Felony – enhanced, Serious Misdemeanor, Special Sentence 2005, Felony – Mandatory Minimum, Other Felony, A Felony, Simple Misdemeanor, Sexual Predator Community Supervision, Other Misdemeanor |
| Convicting Offense Type/ New Convicting Offense Type | Drug, Property, Violent, Public Order, Other |
| Convicting Offense Subtype/ New Convicting Offense Subtype | Trafficking, Assault, Burglary, Theft, OWI, Sex, Forgery/Fraud, Drug Possession, Other Criminal, Other Violent, Traffic, Murder/ Manslaughter, Weapons, Alcohol, Vandalism, Robbery, Other Drug, Other Public Order, Arson, Sex Offender Registry/ Residency, Flight/Escape, Special Sentence Revocation, Kidnap, Prostitution/Pimping, Stolen Property, Animals |
| Release Type | Parole, Discharged-End of Sentence, Parole Granted, Discharged – Expiration of Sentence, Special Sentence, Paroled with Immediate Discharge, Released to Special Sentence, Paroled to Detainer (INS, Out of State, Iowa, U.S. Marshall), Interstate Compact Parole |
| Main Supervising District | 1JD, 2JD, 3JD, 4JD, 5JD, 6JD, 7JD, 8JD, ISC |
| Recidivism Return | Yes, No |
| Recidivism Type | New Recidivism, Tech, New |
| Days to Recidivism | 0 to 3 years |
| Target Population | Yes, No |

The cleaned dataset consists of 21,611 rows and 17 columns. Each row represents a prisoner released from prison in the five-year period between 2010 and 2014. One of the columns, whether an individual returned to prison within 3 years of release, is used as the label [target feature]. Five of the attributes [*'Recidivism Type'*, *'Days to Recidivism'*, *'New Conviction* *Offense Classification'*, *'New Conviction Offense Type'*, *'New Conviction Offense Sub* T*ype*'] represent information collected for recidivist prisoners and hence are not considered with the predictor features.

The data shows 7,027 individuals returned to prison; therefore the positive to negative class ratio is ~1: 2. Although not as extreme as in problems in areas such as fraud detection and infant mortality, there is imbalance in the data that requires balancing the data before applying machine-learning algorithms.

Figure 1: 3-year period recidivism rate for state of Iowa for prisoners released from 2010 to 2014.

**4. Exploratory Data Analysis (EDA)**

The prisoners in the study were released in the five-year period from 2010 to 2014. The recidivism rate by year ranged from 29.7 (for prisoners released in 2011) to 35.4% (for prisoners released in 2014). The average recidivism for the five-year period is 32.5%.

****

Figure 2: Recidivism rate (by year of release) for state of Iowa for prisoners released from 2010 to 2014.

The recidivism data consists of a column that indicates the race/ethnicity of the prisoners. The data shows that the majority of the prisoners are White (non-Hispanic) followed by Black (non-Hispanic), White (Hispanic) and American Indian/ Alaska native. This roughly mirrors the racial composition of Iowa, which reported a population of White (91.4%), Black (3.7%), American Indian/Alaska Native (0.5%), Asia (2.5%), and Hispanic/ Latino (5.8%) for the year 2017. The recidivism rate for whites (non-Hispanic), blacks, whites (Hispanic), and American Indians/Alaska natives is 33.4, 32.8, 20.9, and 39.4% respectively.

****

Figure 3: Recidivism rate (by race/ethnicity) for state of Iowa for prisoners released from 2010 to 2014.

The distribution of prisoners by age shows that the 36.6% of the total released prisoners belong to that age group of 25-34. The representation of the other age groups in the released prisoners is 35-44 (23.9%), under 25 (17.8%), 45-54 (16.8%), and over 55 (4.9%). The recidivism rate for the age groups under 25, 25-34, 35-44, 45-54, and over 55 is 37.1, 34.6, 31.8, 28.0, and 19.1% respectively. This suggests recidivism is highest for the age group under 25 and the lowest for age group over 55 years. This agrees with studies that showed age is a predictor for recidivism and the likelihood of recidivism diminishes as the prisoner gets older (Langan and Levin, 2002).



Figure 4: Recidivism rate (by age) for state of Iowa for prisoners released from 2010 to 2014.

The gender distribution of the released prisoners shows males and females make up 87.4 and 12.6% respectively. The recidivism rate for males and females is 33.2 and 27.7% respectively.



Figure 5: Recidivism rate (by gender) for state of Iowa for prisoners released from 2010 to 2014.

Based on the offense class of the prison system, the prisoners can be grouped to D Felony (40.5%), C Felony (26.2%), aggravated misdemeanor (18.8%), B Felony (6.9%), Felony – enhanced to original penalty (5.6%), and Felony – enhanced (1%). Other classifications, such as serious misdemeanor, special sentence, A Felony, Other Felony, simple misdemeanor made up less than 1% of the total. The recidivism for each offense class are D Felony (30.2%), C Felony (40.0%), aggravated misdemeanor (33.1%), B Felony (32.0%), Felony – enhanced to original penalty (40.4%), and Felony – enhanced (35.0%).



Figure 6: Recidivism rate (by offense classification) for state of Iowa for prisoners released from 2010 to 2014.

**5. Machine Learning Methods**

**5.1. Resampling**

The data shows 7,027 individuals returned to prison; therefore the positive to negative class ratio is ~1:2. Although the class imbalance is not as extreme as in problems in areas such as fraud detection and infant mortality, there is imbalance in the data that requires balancing the data before applying machine-learning algorithms. I have used under-sampling, over-sampling and combined (under-sampling followed by over-sampling) to balance the data. The resampling step can be applied before or after splitting the dataset into train and test datasets.

**5.1.1 Under-sampling methods**

I used Random Under-Sampling (RUS), RUS/TomekLinks, and Edited Nearest Neighbor (ENN) to sample the subset of the majority class in order to reduce the size of the majority class and balance the dataset.

In random under-sampling (RUS), a subset of the majority class are randomly included creating a smaller, balanced dataset. The loss of information with the removal of data from the majority class is the main disadvantage of this resampling method.

TomekLinks provides a guided undersampling/oversampling of a dataset by identifying and removing “noisy” data (TomekLinks) at the boundary of classes. Elhassan et al. (2016) suggest removing noise observation from majority class followed by RUS improve the performance of classification by reducing the chance of information loss.

In Edited Nearest Neighbor (ENN), the majority class is under-sampled by removing points whose class label differs from a majority of its *k* nearest neighbors.

**5.1.2. Over-sampling methods**

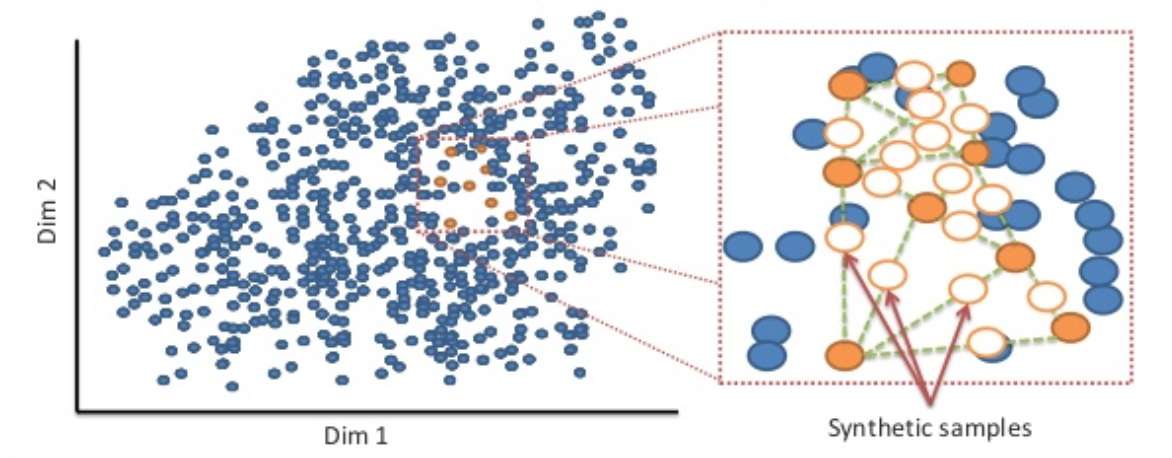
Oversampling methods balance the dataset by increasing the size of the minority class. The two major disadvantages of oversampling are the increased possibility of over-fitting and increased learning time due to increase in learning examples. I use Synthetic Minority Oversampling Technique (SMOTE, Chawla et. al., 2002) and Adaptive Synthetic Sampling Approach (ADAYSN, He et.al, 2008) to oversample the minority class.

Figure 7: Synthetic Minority Oversampling Technique (Bahnsen et. al.).

SMOTE oversamples the minority class by creating synthetic examples, rather than the traditional way of oversampling with replacement.

**5.1.3. Combined over- and under-sampling methods**

SMOTE can generate noisy samples by interpolating new points between marginal data points. TomekLinks and Edited Nearest Neighborhood (ENN) methods can be used to clean the data.

**5.2. Algorithms**

I used logistic regression and random forests for machine learning, algorithms that are commonly used for classification.

**5.2.1 Logistic Regression**

Logistic regression is the common machine-learning tool in classification problems. For binary classification, logistic regression computes the probability of an outcome [between 0 and 1] by computing fitted probabilities of linear combination of predictors. Sigmoid function is used to constrain the probability between 0 and 1.



Figure 8: Sigmoid function.

The procedures below were followed to build the model.

1. Separate the attributes into categorical (nominal), numerical (noncategorical), and binary. Encode the categorical attributes
2. Specify predictor features and target features. The target feature in this project is recidivism of a prisoner..
3. Split the dataset into training and test datasets (80% train and 20% test data set).
4. Use Grid Search and Cross-Validation (using k-fold cross validation with a k= 5) and determine the optimum regularization parameter.
5. Train the best algorithm on the train set and test the algorithm on the test set.
6. Evaluate the classifier using performance metrics. Use classification report (from package imblearn; Lemaire et. al., 2017) to evaluate a classifier trained using the original data.

**5.2.1.1. Results**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | PP | NP | SS | Accuracy | Precision | Recall | TNR |
| LR\_NoResample | Train | 0.32 | 0.68 | 17288 | 0.684 | 0.546 | 0.141 | 0.944 |
| Test | 0.33 | 0.67 | 4323 | 0.675 | 0.539 | 0.140 | 0.940 |
| LR\_ENN\_Split | Train | 0.605 | 0.395 | 9300 | 0.725 | 0.734 | 0.856 | 0.525 |
| Test | 0.601 | 0.399 | 2326 | 0.718 | 0.726 | 0.850 | 0.518 |
| LR\_Split\_ENN | Train | 0.558 | 0.442 | 10085 | 0.721 | 0.731 | 0.790 | 0.633 |
| Test | 0.325 | 0.675 | 4323 | 0.559 | 0.407 | 0.790 | 0.447 |
| LR\_ENN\_selMODE\_Split | Train | 0.432 | 0.568 | 12979 | 0.667 | 0.625 | 0.571 | 0.740 |
| Test | 0.439 | 0.561 | 3245 | 0.667 | 0.635 | 0.570 | 0.744 |
| LR\_Split\_ENN\_selMODE | Train | 0.380 | 0.620 | 14876 | 0.68 | 0.608 | 0.444 | 0.824 |
|  | Test | 0.318 | 0.682 | 4323 | 0.677 | 0.490 | 0.446 | 0.784 |
| LR\_RUS\_Split | Train | 0.499 | 0.501 | 12648 | 0.626 | 0.617 | 0.656 | 0.595 |
| Test | 0.512 | 0.488 | 1406 | 0.617 | 0.619 | 0.656 | 0.577 |
| LR\_Split\_RUS | Train | 0.5 | 0.5 | 12584 | 0.625 | 0.619 | 0.649 | 0.601 |
| Test | 0.34 | 0.66 | 2162 | 0.621 | 0.46 | 0.654 | 0.604 |
| LR\_RUS\_Tomek\_Split | Train | 0.502 | 0.498 | 11243 | 0.635 | 0.628 | 0.666 | 0.603 |
| Test | 0.493 | 0.507 | 2811 | 0.604 | 0.589 | 0.650 | 0.559 |
| LR\_Split\_RUS\_Tomek | Train | 0.500 | 0.500 | 11276 | 0.627 | 0.619 | 0.663 | 0.591 |
| Test | 0.321 | 0.679 | 4323 | 0.598 | 0.421 | 0.664 | 0.567 |
| LR\_CNN | Train | 0.528 | 0.472 | 10702 | 0.616 | 0.611 | 0.75 | 0.466 |
| Test | 0.514 | 0.486 | 2676 | 0.604 | 0.593 | 0.727 | 0.473 |
| LR\_Split\_SMOTE | Train | 0. 5 | 0. 5 | 23334 | 0.628 | 0.619 | 0.666 | 0.589 |
| Test | 0.325 | 0.675 | 4323 | 0.615 | 0.439 | 0.668 | 0.589 |
| LR\_SMOTE\_Split | Train | 0.503 | 0.497 | 23334 | 0.627 | 0.619 | 0.673 | 0.581 |
| Test | 0.487 | 0.513 | 5834 | 0.636 | 0.612 | 0.691 | 0.583 |
| LR\_Split\_ADASYN | Train | 0.509 | 0.491 | 23768 | 0.594 | 0.59 | 0.661 | 0.524 |
| Test | 0.325 | 0.675 | 4323 | 0.581 | 0.417 | 0.723 | 0.513 |
| LR\_ADASYN\_Split | Train | 0.511 | 0.489 | 23809 | 0.601 | 0.596 | 0.684 | 0.515 |
| Test | 0.505 | 0.495 | 5953 | 0.614 | 0.602 | 0.694 | 0.532 |
| LR\_Split\_SMOTEENN | Train | 0.554 | 0.446 | 8196 | 0.798 | 0.794 | 0.857 | 0.724 |
| Test | 0.325 | 0.675 | 4323 | 0.558 | 0.4 | 0.718 | 0.481 |
| LR\_SMOTEENN\_Split | Train | 0.574 | 0.426 | 10217 | 0.771 | 0.782 | 0.833 | 0.688 |
| Test | 0.582 | 0.418 | 2555 | 0.757 | 0.775 | 0.82 | 0.669 |

**5.2.1.2. Performance evaluation and the best model**

I used performance metrics values (accuracy, precision, recall, TNR, AUC) and curves (confusion matrix, precision-recall curve, receiver operating characteristic) to evaluate and visualize the performance of the classifiers.

True Positive (TP, also called hits) refers to a result that correctly identifies a condition that is present. A recidivism prediction result is a true positive when a prisoner is identified as recidivistic and he/she commits an offense after release from prison.

True Negative (TN) refers to a result that correctly identifies the absence of a condition. A recidivism prediction result is a true negative when a prisoner is identified as *not* recidivistic and he/she *does not* commit an offense after release from prison.

False Positive (FP) refers to a result that incorrectly identifies an absent condition as present. A recidivism prediction result is a false positive when a prisoner is identified as recidivistic even though he/she does not commit an offense after release from prison.

False Negative (FN) refers to a result that fails to correctly identify a present condition. A recidivism prediction result is a false negative when a prisoner is identified is not identified as recidivistic and he/she commits an offense after release from prison.

The TP, TN, FP, and FN are normally given in a table format called confusion matrix.

|  |  |  |
| --- | --- | --- |
| **Predicted Condition**  **True Condition** | **Condition absent** | **Condition present** |
| **Condition absent** | True Negative (TN) | False Positive  (FP) |
| **Condition present** | False Negative (FN) | True Positive  (TP) |

The major indicators of performance can easily be derived from the confusion matrix.

The performance indicators show the best model is the model built by using SMOTEENN resampling technique, followed by splitting the dataset into train and test sets. The accuracy, precision, recall, and TNR for the model are 0.76, 0.78, 0.82, and 0.67 respectively.

|  |  |  |
| --- | --- | --- |
|  | **Train Set** | **Test Set** |
| **Confusion Matrix** |  |  |
| **Precision-Recall Curve** |  |  |
| **ROC** |  |  |

Figure 9: confusion matrix, precision-recall curve, and ROC for the logistic regression based model trained on data resampled using SMOTEENN.

**6. Future Work**

**I will continue working with other classification algorithms such as random forests.**

**References**

Berk, Richard (2016): An impact assessment of machine learning forecasts on parole board decision and recidivism, working paper 2016-4.0, Penn Criminology Working Papers Collection.

Breiman, L. (2001): Statisical modeling: the two cultures, Statistical Science, 16(3), 199-215.

Chawla, N. V., Bowyer K. W., Hall L. O., and W. P. Kegelmeyer (2002): SMOTE: Synthetic Minority Over-sampling Technique, Journal of Artificial Intelligence Research, 321-357.

Durose, M. R., Cooper, A. D., and Snyder, H. N. (2014): Recidivism of prisoners released in 30 states in 2005: Patterns from 2005 to 2010. Washington, DC: Bureau of Justice Statistics.

Elhassan T, Aljurf M, Al-Mohanna F, and Shoukri M (2016): Classification of Imbalance Data Using Tomek Link (T-Link) Combined with Random Under-Sampling (RUS) as a Data Reduction Method, Journal of Informatics and Data Mining.

He, H., Bai, Y., Garcia, E. A., and Li, S. (2008): ADASYN: Adaptive synthetic sampling approach for imbalanced learning, IEEE international Joint Conference on Neural Networks, pp. 1322-1328.

Langan, P., and Levin, D. (2002): Recidivism of prisoners released in 1994, Special Report, Bureau of Justice Statistics, Washington, DC.

Teich, David A. (2018): Management AI: bias, criminal recidivism, and the promise of machine learning.