Fall 2023 CSE 297A Final Report: Classification Fairness on improving the prediction of Criminals' recidivate

Introduction

Classification fairness improves machine learning's capability of predicting in various domains in our society. However, there are concerns of decision tree bias against a certain gender or race of committing a crime again, which is known as high false positive divergence. The contribution of decision tree bias or classification unfairness may contribute to a huge dataset. Furthermore, a huge dataset has unintended possibilities for high false positive divergence and decision tree bias. To investigate the causes of high false positive divergence, feature selection method on an input attribute, such as race, may reduce false positive divergence and improve classification fairness in criminal justice domain.

The objective of this research was to study modifications of classification trees for reducing false positive divergence. DivExplorer on a COMPAS dataset encouraged a question of classification fairness for predicting criminals' recidivation. There were concerns of having high false positive divergence for criminals' recidivation based on two attributes of COMPAS dataset: number of prior incidents(#prior) and race. To reduce false positive divergence and solve classification unfairness, there were two decision trees implemented for my study of reducing false positive divergence: pruning to avoid two attributes from appearing in same branches and non-pruning without the two attributes. DivExplorer's divergence method and accuracy were to analyze the results from these two modified decision trees along with the original decision tree. Thus, a comparison of false positive divergences between two modified decision trees and original decision tree had interesting findings to reduce false positive divergence.

Related work

Since my research address the difficulties of lowering the max divergence of subgroups within a given threshold for a given decision tree and a classifier, there are algorithms being introduced from previous research works to address these difficulties along with an attempt to minimize the accuracy loss. After conducting literature review to address the main topic of my research, CART, which is classification and regression tree, had being utilized on feature selection and imbalanced classification problems. Furthermore, it has "identification and construction of a binary decision tree using a sample of training data for which the correct classification is known" (Bittencourt, Clake) for reducing the max divergence in a decision tree with a classifier. Clake's research is related to my research because feature selection may provide efficient results on reducing the false positive divergence. A feature selection in my classification tree had two attributes, which are race and #prior, contribute to most false positive divergence.

The boosting method is very effective for improving accuracy in my decision tree implementations to reduce false positive divergences. From Decision Tree Application to Classification Problems with Boosting Algorithm research article, boosting algorithms improves their performance of the model in accuracy and precision for the CART algorithm applied (Zhao, Lee, etc.)[7]. It utilized the pruning method for their loss matrix and confidence interval in the business domains. Furthermore, these researchers discussed their algorithm had overfitting and decision tree bias as their main concern for their research in business domain. This is related to my research because my classification tree was reducing false positive divergence due to decision tree bias on criminals' recidivation. Even though their research discussed about their classification problems to improve their accuracy for efficient predictions in a business domain, pruning method of the two attributes, which are #prior and race, contributes to a boosting method for improving recidivity prediction accuracy.

Methods

After an analysis of the false positive divergence from a COMPAS dataset, an individual, who had committed a certain number of prior crimes(#prior), and their race contributed to an unfair classification of recidivating. A scikit python library allowed my discovery of creating a decision tree. Initially, a training set must be created to train a classifier, which is crucial for the implementation of two classification trees, from a COMPAS dataset. The randomized training for a testing set prior to creation of a decision tree. After training and testing set had been established, a decision tree classifier would be utilized for prediction of prisoners' recidivating. It established a blueprint for implementation of two classification trees to solve unfair classification issues of the criminal recidivating a crime.

Since #prior and race were the attributes for having false positive divergences of predicting recidivate, two methods of classification trees must be created for examining on having the most classification fairness. The first method was to prune the #prior and race, which are classified as top_attributes, from the same branches of the decision tree with the recursion search technique. The recursion search technique had a conditional statement of identification for the top_features, which are #prior and race attributes, to be removed. The feature selection of removing both #prior and race for accessing high accuracy and lower a false positive divergence in the second classification tree. A second decision tree was to remove #prior and race attributes to examine the differences from a primary method. After the implementation of both decision trees to improve classification fairness, recidivity prediction accuracy was utilized for accuracy calculation to determine which decision tree technique was more efficient in my study. From using a DivExplorer's library for calculating false positive divergence, both methods of classification trees and original classification tree were accessed for classification fairness.

Results

After implementing these two classification trees with respective techniques on improving the classification fairness, the first classification tree had an accuracy of 67.66% and the second classification tree's accuracy was 60.60%. Furthermore, it had an efficient optimization of predicting recidivates of the prisoners' dataset from DivExplorer. An optimization for an efficient accuracy than the second classification tree is the recursion search method in the pruning method. The divergence calculations for both classification trees had interesting results. The first classification tree had high occurrences for #prior =>3 and their age range of 25-45 likely to commit a crime again. The second classification had high possibility of recidivation for individuals, who are male and a young adult. The false positive divergence from the first classification tree for ten items was a consistent 0.78085. In contrast to the first classification tree, a second classification tree had 0.768064 consistently from 0-3 items initially.

Then there is a decline of false positive divergence from item 4 through item 9 with 0.720064 (item 4) to 0.529071 (item 9). The original classification tree had a high occurrence for individuals, who are African American and had most past criminal crimes, may commit a crime again. The lowest false positive divergence of the original decision tree is 0.696 in item 9.

Below are the calculations of the divergence for both classification trees and original decision tree:

```
Tree 1 Accuracy: 67.66%
Divergence for pruning #prior and race on different branches(1st decision tree):
    support                                        itemset   fp   fp_div  \
0  0.103694   (#prior=>3, charge=F, stay=<week, age=25-45)  1.0  0.78085
1  0.122975   (#prior=>3, sex=Male, stay=<week, age=25-45)  1.0  0.78085
2  0.130752   (#prior=>3, charge=F, sex=Male, race=Afr-Am)  1.0  0.78085
3  0.133668     (#prior=>3, charge=F, sex=Male, age=25-45)  1.0  0.78085
4  0.207226                         (#prior=>3, age=25-45)  1.0  0.78085
5  0.121355    (#prior=>3, charge=F, sex=Male, stay=<week)  1.0  0.78085
6  0.142417             (#prior=>3, stay=<week, age=25-45)  1.0  0.78085
7  0.145010            (#prior=>3, race=Afr-Am, age=25-45)  1.0  0.78085
8  0.109527  (#prior=>3, charge=F, race=Afr-Am, age=25-45)  1.0  0.78085
9  0.190376                (#prior=>3, charge=F, sex=Male)  1.0  0.78085

         fp_t  length  support_count
0   90.611825       4          640.0
1   94.792406       4          759.0
2   92.940559       4          807.0
3   94.308537       4          825.0
4  102.966923       2         1279.0
5   95.072317       4          749.0
6   98.431668       3          879.0
7   96.444433       3          895.0
8   88.442089       4          676.0
9  101.747273       3         1175.0
Prediction without top shapley attributes: ['no']
```

2nd classification tree:

```
Prediction without top shapley attributes: ['no']
Tree 2 Accuracy: 60.60%
Second Divergences without pruning. #prior and race never exist in decision tree:
    support                              itemset        fp    fp_div        fp_t  \
0  0.134964   (sex=Male, age=<25, stay=<week)  1.000000  0.768064   98.374244
1  0.107583  (sex=Male, race=Afr-Am, age=<25)  1.000000  0.768064   91.080643
2  0.133020     (charge=F, sex=Male, age=<25)  1.000000  0.768064   96.570877
3  0.178386               (sex=Male, age=<25)  1.000000  0.768064  100.495038
4  0.101426           (stay=1w-3M, age=25-45)  0.952000  0.720064   45.639882
5  0.150518            (stay=1w-3M, sex=Male)  0.813008  0.581072   26.876762
6  0.156837               (charge=F, age=<25)  0.786600  0.554665   25.555923
7  0.177090                      (stay=1w-3M)  0.776786  0.544850   25.940559
8  0.120058   (charge=F, age=<25, stay=<week)  0.762048  0.530112   21.631068
9  0.128808            (charge=F, stay=1w-3M)  0.761006  0.529071   21.138859

   length  support_count
0       3          833.0
1       3          664.0
2       3          821.0
3       2         1101.0
4       2          626.0
5       2          929.0
6       2          968.0
7       1         1093.0
8       3          741.0
9       2          795.0
```

Original classification tree:

```
Tree 3 Accuracy: 68.34%
Original Tree Divergences(3rd tree):
    support                                          itemset        fp  \
0  0.130752   (#prior=>3, charge=F, sex=Male, race=Afr-Am)  1.000000
1  0.128645  (#prior=>3, sex=Male, race=Afr-Am, age=25-45)  1.000000
2  0.145010             (#prior=>3, race=Afr-Am, age=25-45)  0.992424
3  0.109527  (#prior=>3, charge=F, race=Afr-Am, age=25-45)  0.989637
4  0.148412              (#prior=>3, charge=F, race=Afr-Am)  0.956204
5  0.133668     (#prior=>3, charge=F, sex=Male, age=25-45)  0.941667
6  0.190376                 (#prior=>3, charge=F, sex=Male)  0.938889
7  0.103694    (#prior=>3, charge=F, stay=<week, age=25-45)  0.937500
8  0.121355     (#prior=>3, charge=F, sex=Male, stay=<week)  0.931452
9  0.142417              (#prior=>3, stay=<week, age=25-45)  0.928082

     fp_div       fp_t  length  support_count
0  0.768064  90.076191       4          807.0
1  0.768064  89.755774       4          794.0
2  0.760488  77.745696       3          895.0
3  0.757702  65.940939       4          676.0
4  0.724269  49.160522       3          916.0
5  0.709731  41.295748       4          825.0
6  0.706953  47.831274       3         1175.0
7  0.705564  37.598893       4          640.0
8  0.699516  38.954135       4          749.0
9  0.696146  40.866196       3          879.0
```

Discussion

The first classification tree had #prior and race attributes appear in different branches with a pruning method. Furthermore, an implementation of the first decision tree had a recursion search method for the pruning of #prior and race to appear in different branches than second decision tree. An implementation for accuracy for comparison of the two classification trees discussed in my research. Then an utilization of DivExplorer for calculating divergence revealed an amusing result for both classification trees' implementations. Since the first classification tree had #prior and race appear in different branches, age and #prior had high occurrences(fp_t) of criminals' recidivity. In contrast to the first classification tree, second classification tree had high occurrences(fp_t) on sex and young adult age attributes. The second classification tree had lower false positive divergence and 60% accuracy. The pruning method from the first decision tree had better accuracy than the second classification. In contrast to the first classification tree, a feature method was utilized to remove both race and #prior attributes for a low false positive divergence. Thus, a second classification tree may contribute to having a classification fairness of predicting criminals' recidivity due to a low false positive divergence.

References

1. Bittencourt, H. R., & Clarke, R. T. (2004). Feature selection by using classification and regression trees (CART). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.
2. Y. Park and J. Ghosh, "Ensembles of (α)-Trees for Imbalanced Classification Problems," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 131-143, Jan. 2014, doi: 10.1109/TKDE.2012.255.
3. M. Czajkowski and M. Kretowski, "Decision tree underfitting in mining of gene expression data. An evolutionary multi-test tree approach," *Expert Systems with Applications*, vol. 137, pp. 392–404, 2019, doi: 10.1016/j.eswa.2019.07.019.
4. Zhao, L., Lee, S., & Jeong, S. P. (2021). Decision tree application to classification problems with boosting algorithm. *Electronics*, *10*(16), 1903.
5. Machová, K., Puszta, M., Barčák, F., & Bednár, P. (2006). A comparison of the bagging and the boosting methods using the decision trees classifiers. *Computer Science and Information Systems*, *3*(2), 57-72.
6. S. Mitrofanov and E. Semenkin, "An Approach to Training Decision Trees with the Relearning of Nodes," *2021 International Conference on Information Technologies (InfoTech)*, Varna, Bulgaria, 2021, pp. 1-5, doi: 10.1109/InfoTech52438.2021.9548520.
7. Zhao, L.; Lee, S.; Jeong, S.-P. Decision Tree Application to Classification Problems with Boosting Algorithm. Electronics 2021, 10, 1903. https://doi.org/10.3390/electronics10161903