

# Determining How Different Factors Affect Police-Allegation's Sustainability in Chicago using Decision-Tree

Linxin Yang\*

*Korhal Internet, Columbus, Ohio, U.S.A.*

*\*Independent Researcher*

Keywords: Decision Tree, Criminal Justice, Data Analysis.

Abstract: The Citizen Police Data Project (CPDP) is a database of allegations made against the Chicago Police Department. Reports made against officers are rarely sustained, which results in the perception of little officer accountability and contributes to widespread distrust of law enforcement. Using a decision tree model on the CPDP database, this work explores how the following factors: officer years of employment, complainant type, investigation agency, and allegation severity level, affects the outcome of an allegation work together to increase or decrease the sustainability of allegations made against CPD between 2008 to 2018. The results found that when a CPD employee reports an allegation, it has higher chances to be sustained. However, for allegations reported by civilians, a third-party agency increases the likelihood of allegation sustainability.

## 1 INTRODUCTION

The Citizen Police Data Project (CPDP) is a significant source of information related to allegations made against Chicago police officers. The database holds the records of information including allegations made against Chicago police officers. The database also stores the allegations' sustainability, complainants' information, officers' working history and officers' salary. Data that could be buried internally if not for CPDP publishing it and making it available to the public. This database plays a role in serving as a national model for police transparency and a resource for the Chicago citizens to increase the Chicago Police Department's accountability. According to the database, 247,150 allegations were recorded from 1988 to 2020, and only 7% are sustained (CPDP, n.d.). This leads to distrust from the citizens, which negatively affects the police department (Goldsmith, 2005). This results in the public being less likely to make complaints due to the cases not likely being sustained, weakening the department's ability to improve.

Research with the CPDP database has been done to improve the current accountability problem. They were showing such things like how race and ethnicity affect allegation's outcome (Headley et al., 2017), how cases are influenced by the perceptions of citizen (Dowler & Zawilski, 2007), police (Long et al., 2013)

and court (Gottschalk, 2017), and whether outcomes of given cases are socially ecologically correct or not (Kane, 2002). From work listed above, we know that race and ethnicity of the complainant have significant influence over the decision-making process. Though the work mention discovered valuable results, most of the research only focuses on a single factor (e.g. race) and does not consider multiple various factors and how those factors work together when determining the result of an allegation.

Allegation cases usually contain essential and vital factors that could significantly influence the outcome of the case. For example, the complaint is a good factor because a citizen likelihood of their complaint being sustained contributes to the possibility of them filing a complaint in the future. According to Terrill and Ingram (2015), only few of the civilian complaints are sustained, especially those with excessive use of force. The investigation agency as a potential influencer for the case outcome could also be a good factor; the internal investigation bureau could allow the decision being made solely by some chosen officers, where a civilian investigating agency could be more public and transparent (Raymond W. Patterson, 2006). The Bureau of Internal Affairs (BIA) and Independent Police Review Authority (IPRA) are two agencies for the Chicago police department to investigate alleged cases, according to the CPDP database. The BIA is an

internal agency from the police department, where IPRA is a civilian agency that does not belong to the police department. As these agencies are from different perceptions, they surely can be an interesting factor that could influence the outcomes. Littlejohn, states the Civilian Review Board was established to satisfy the long-term dissatisfaction with internal complaint procedures (1981), so this factor is likely to make a difference in the allegation being sustained when a non-bias investigator is working the case. Another factor that can be observed as an influencing factor is the officer's year of employment. Young officers might be given a relatively slight punishment to give them another chance to improve themselves. Another outcome influence is the person filing the complaint (e.g. officer, citizen). Research has shown that less experienced police are more likely to receive complaints (Terrill and Ingram, 2015) which influenced the next factor choice of officer employment history. Besides, the level of severity is also an important factor when reviewing the cases. According to the Police Misconduct Complaint Investigation Manual, low-level allegations should be processed differently from those involving forces or racial bias cases (Attard & Olson, 2020). We can also find support from previous works; server allegations are separated from slight allegations when examining the application of prediction methods (Kyle Rozema & Max Schanzenbach., 2019).

As introduced by Luna et al., machine learning algorithms are widely adopted to the high-stakes areas such as medication and criminal justice; among these algorithms, decision trees are one of the most well-explored algorithms as they can produce rational decision-making processes with even large size of datasets (2019). There are several pieces of research deployed to discuss the feasibility of applying such algorithms when convicting criminals. Corbett-Davis et al. have proposed a machine-learning algorithm to reduce the inequality between races when judging criminals' risk (Corbett-Davies, Pierson, Feller, Goel, & Huq, 2017). Gutierrez and Leroy utilized decision trees to make predictions of whether a crime is reported or not, and eventually improve crime reports' accuracy (2007). We can also see the increasing interest of inventing better algorithms for crime investigation; research was held to enhance criminal recidivism prediction using Machine Learning algorithms (Wang et al., 2010). Since criminal convictions can be operated with reliable machine learning algorithms, it is possible and feasible to apply decision trees into the field of police allegations.

This paper makes use of a decision tree model using data from the CPDP database. By extracting 10,799 allegations from 2008 to 2018, this work looks at allegations to determine the following factors: officer years of employment, complainant type, investigation agency, and allegation severity level affects the outcome of an allegation. This work explores how the chosen factors work together to increase or decrease the sustainability of allegations made against the Chicago police officers. From the result, we have found that misconduct with allegation severity level 2 or higher is more likely to be sustained. When a CPD employee reports an allegation, it has higher chances to be sustained, but for allegations reported by civilians, third-party agencies sustain more cases than the internal agency does. Besides, both agencies sustain experienced officers more than those younger officers.

## 2 METHODS

### 2.1 Decision Tree

The Decision Tree is a model that automatically splits possible consequences by different attributes and ranks those features in order. The structure of Decision Tree is like a flow-chart; an input is given at the top level and passed to lower levels based on analyzed rules, and a prediction will be given as the information reaches the bottom level. A sample Decision Tree is shown in Figure 1.

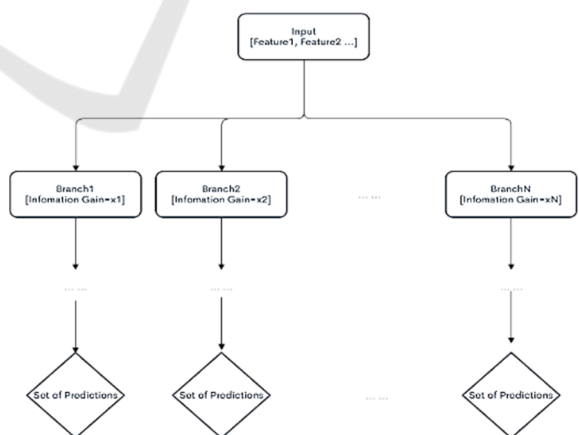


Figure 1: Sample Decision Tree.

The model is trained by a given data set and ranks features based on how important they are to the entire model. For this model, the Decision Tree algorithm from the Scikit-learn package is used. Scikit-learn is an open-source python package that provides various

tools for predictive data analysis. It is one of the most efficient packages available to the public. In this project, the Decision Tree method is used. The Decision Tree construction in Scikit-learn uses Classification and Regression Tree (CART), which uses the Gini index rather than traditional Chi-square. By using CART, the tree will be constructed in binary form, which is much more efficient than using Chi-square with n-ary trees. The error is also controllable and ignorable compared to the other methods. For a dataset containing more than 10,000 samples, CART is perfectly suitable. While having a better performance among other common machine learning analysis algorithms, according to Wibowo and Oesman, a Decision tree can achieve similar accuracy in criminal investigation comparing to those algorithms. (2020)

The importance of each feature is defined by their Gini index, which is defined as follow:

$$Gini(p) = \sum_{k=1}^K p_k(1 - p_k) = 1 - \sum_{k=1}^K p_k^2$$

where  $p_k$  is the possibility of having  $k$ th consequence. This index represents the impurity of the model, and the lower the index is, the better the result will be. When constructing the tree, each node of the level will be computed based on the previous level's possibility. The feature that contributes the most to the information completeness will be placed on the top of the tree. Thus, the nodes on a higher level will be more critical for the decision making.

## 2.2 Data Collection and Preprocessing

The data was collected from the CPDP database from 2000 to 2018, which includes 83,098 complaints and 32,445 officers' serving history. The database consists of reports that provide details on the type of allegation made, the complaints' that made the allegations, and if the allegation is sustained. There's also information about the officer named in the allegation such as their name, age, race, time on the force, and awards they accumulated while being an officer. Some of the data were removed due to the lack of officers' information needed for this analysis. There are also duplicated records as some of the allegations involve multiple charges. After removing the data, we obtained a dataset containing 10,799 cases. We structured the dataset as the following: allegation type, officer start date, officer end date, who made the allegation, was the allegation sustained, if the allegation is sustained what the disciplinary action was and if the allegation is sustained who was the investigating agency.

## 2.3 Outcome Influencers

We evaluated the different types of factors that could determine the outcome of an allegation. The factors complainant and investigation agency are included in the allegation information. Officer employment history is not present and has to be extracted through officer ID matching from another dataset within the CPDP. The complainant information could not be passed into the decision tree because of its format, so it is digitized; we assign 0 to the CIVILIAN and 1 to CPD\_EMPLOYEE. We classify them according to their severity for results and complaint reasons, as shown in Table 1 and Table 2. The allegation types are classified into three classes according to their severity. The basic standard of classifying allegation severity is splitting them according to their outcomes, which is used in another police misconduct research by Kyle Rozema and Max Schanzenbach (2019). Level 1 allegations include authorized weapon discharge and mild allegations like verbal conflicts with civilians. Such allegation usually does not cause any damage to property or human health. Level 2 allegation includes misconduct that usually causes damage either mentally or physically to the complainant. Level 3 allegation is when the incident causes severe casualties such as excessive use of force or related to on-duty felonies such as drug abuse and DUI.

Table 1: Complaint reason classified by severity.

Severity	Complaint Reasons	Label
Level1	Verbal Insults/Violation of Traffic Regulation/Authorized Weapon discharge such as: "SEAT BELTS", "PARKING COMPLAINTS", "ALTERCATION / DISTURBANCE - TRAFFIC"	0
Level2	Religious/Racial Insulting/Unauthorized Discharge of Weapon/Inappropriate Physical Contact such as: "EXCESSIVE FORCE - USE OF FIREARM / ON DUTY - INJURY", "CLOSED HAND STRIKE (PUNCH)", "RACIAL PROFILING"	1
Level3	Excessive Force with or without Injuries/DUI on duty or off duty/Drug Related Behavior such as "EXCESSIVE FORCE - USE OF FIREARM / OFF DUTY - INJURY", "FIREARM DISCHARGE WITH HITS / OFF DUTY", "USE / ABUSE DRUGS / CONTROLLED SUBSTANCE - ON DUTY"	2

Allegations that attract public attention are also classified into this group as they are always prioritized during the investigation according to the Police Misconduct Complaint Investigations Manual (Attard & Olson, 2020).

The classification of the result is referencing the standard of Vancouver Discipline Matrix (Darrel W. Stephens, 2011) and the actual data distribution in the given dataset; 15 days of suspension is common maximum penalty before dismiss, but the most common punishment in the given dataset is around 30 days, so we are using 30 days as the decision basis.

Table 2: Case Results classified by severity.

Severity	Case Result	Label
Level1	"No Action Taken" "Violation Noted" "Unknown" "Penalty Not Served"	0
Level2	Suspended between 1 to 30 Days	1
Level3	Suspended between 30 to 180 Days	2
Level4	Dismiss	3

Level 1 results only include those cases that are not given a penalty. As a suspension under 30 days is always considered a mild penalty, they are classified as a Level 2 penalty, while suspension between 30 days and 180 days is classified as a Level 3 penalty. All other penalties such as dismiss and suspension over 180 days are classified as Level 3 penalties.

Besides, the officer's working experience, as a not-often-mentioned factor, also influences the decision-making a lot. McElvain and Kposowa mentioned that the more experienced the officer is, the less likely they will make discipline mistakes. (2008) In other word, officers with longer experience are less likely to have an unintentional allegation, and they should be punished if they do.

### 3 RESULTS

We first analyzed some normal data patterns from the dataset to better understand the dataset and prove the validation of features selected. The first examined attribute is the complainant type. As shown in figure 1, complaints from the CPD employee are more likely to be sustained, which indicate that this attribute can affect the result.

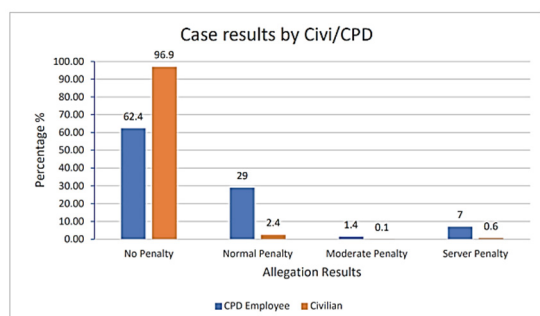


Figure 1: Case results by Civil/CPD.

Then we analyzed the relation between the outcomes and officers' working history. The result shown in Figure 2 shows that officers with more working experience are more likely to receive a severe penalty.

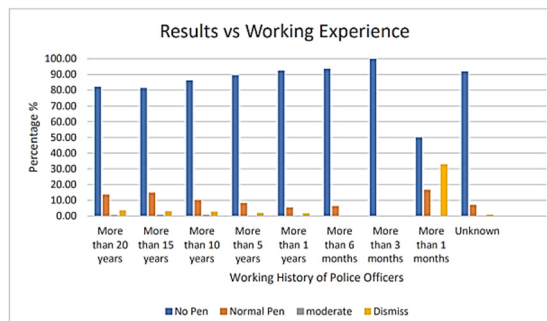


Figure 2: Results vs Working Experience.

We also compared officers' working experiences and allegation levels as shown in figure 3. The results showed a contradiction to the previous one; younger officers are more likely to have severer allegations while not being punished with the same severity. This indicates that both the allegation level and working experience could have an influence on the outcomes.

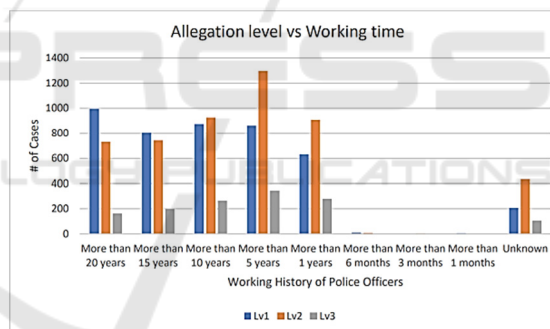


Figure 3: Allegation level vs Working Experience.

As the chosen factors are validated, we can move to the decision tree part. The decision tree is trained by a randomly chosen dataset that is 90% of the original dataset; features are passed into the model in the order of:

- [Severity of Allegation, Complainant type, Employee History, Agency Type]

In the first level, the tree is branched by the complainant type as shown in Figure 4 and Figure 5, which means that the Complainant type is the most essential factor in deciding the penalty; 7,674 complaints are from the civilian, and 3,125 are from the CPD employee. 96.9% of the Civilian cases were acquitted, where 62.4% of the CPD employee cases

are acquitted. As the data and figure show, we can see that the complainant type can effectively influence a case's result.

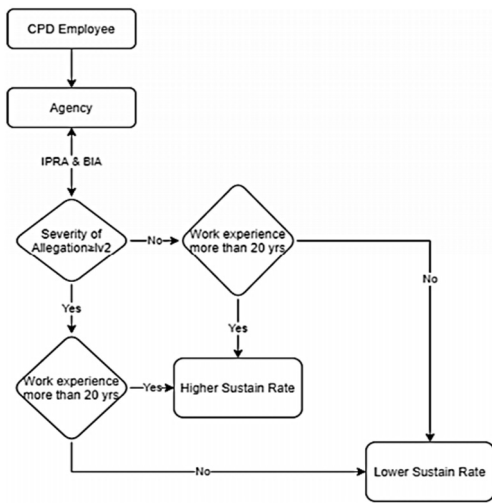


Figure 4: Decision Tree of Cases from CPD.

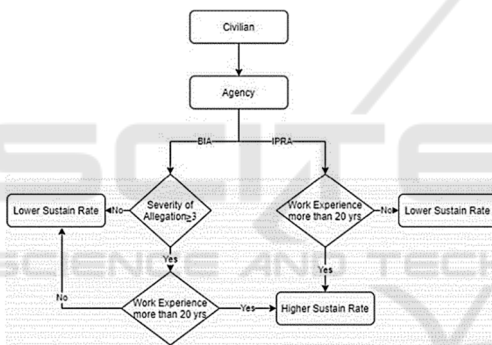


Figure 5: Decision Tree of Cases from Civilian.

Then, both branches take the type of agency as the next most crucial factor. For 7,674 civilian cases, 2,872(37.4%) of them are investigated by BIA, and 4,802(62.6%) of them are investigated by IPRA. For 3,125 CPD cases, 2,647(84.7%) of them are investigated by BIA, and 478(15.2%) cases are investigated by IPRA. The data here shows that, as an internal bureau, the BIA takes more control of the CPD cases, while the third-party organization IPRA takes most of the civilian cases.

The conviction rate also differs between different agencies. Though the innocent percentage of Civilian cases from BIA and IPRA are close (93.4% vs 97%), the innocent percentage of CPD cases differs a lot (48.8% vs 70.0%). It is clear that a CPD case is more likely to be convicted if they are investigated by the internal bureau, which makes the total conviction rate of BIA and IPRA to 3,981 as of 72.1% vs 4995 as of 94.6%.

The next level of the tree shows what different agencies consider as the most important factor. BIA takes the severity of allegation as the first consideration for civilians, while IPRA checks officers' employee history first. According to the tree, civilian cases from BIA are 20% more likely to be convicted if the severity of allegation level is more than lv2, and civilian cases from IPRA are only 1% more likely to be convicted if the officer has worked for more than ten years.

For CPD employees, both agencies examine the severity of the allegation first, but with a higher allegation tolerance (worse than lv12). The conviction rates for BIA are 47.44% and 56.18% for severity less than lv2 and greater than or equal to lv2. 35.08% and 26.48% are the rates for IPRA. The conviction rate of a more severe allegation is more likely to be acquitted under IPRA's investigation. However, this could be caused by the lack of cases given to the IPRA (2,647 cases vs 478 cases).

Apart from these factors, for both agencies, working experience is also a factor that could lead to different sustain rates. If an officer has worked in the department for more than 20 years, they are facing a higher chance of suspension or dismissal.

## 4 DISCUSSION

### 4.1 Data Pattern

Besides the results from the decision tree, we also conclude data patterns from the dataset. These patterns, which are shown in figure 1, 2, 3, also provides insights into the current situation of Police allegation investigation. In figure 1, we can see the inequality between civilians and police in reporting an allegation; apparently, civilians won less trust in the investigation. Such inequality apparently should not be influencing the investigation as instructed in the Police Misconduct Complaint Investigation Manual (Attard & Olson, 2020). The Civil Office of Police Accountability (COPA), which is the new IPRA, as reported by Leven, focuses more on violation of civil rights (Leven et al. 2017). As COPA holds their duty, civilians could expect a more equalized environment.

As we discussed in the previous part, the officer's working experience influences the investigation as well. According to the Investigation Manual, the experiences from parties related to the involving officer should be included when evaluating the officer (Attard & Olson, 2020). A longer working history means more detailed description from their

colleagues, which results in a more reliable decision. This could also be seen in figure 2; experienced officers always receive more severe penalties compared to those new officers. However, new officers are more likely to commit with more severe allegations as shown in figure 3. It is not reasonable that officers with sever allegation receive fewer penalties. This could be caused by the lack of understanding of officers' personality.

## 4.2 Decision Tree

This study has examined the importance of a group of factors during the investigation of police allegations. Four critical factors, including the severity of allegation, complainant type, employee history and agency type, were chosen according to previous research. These factors are then studied by the decision tree with cases retrieved from the CPDP database. The result is reliable as the Gini index of each node in the tree indicates a trustable result. Complaint type was examined to be the most important factor, which matches the data pattern that most reports from the civilian will not be sustained. Such results show how civilians are usually not considered as a reliable report source; agencies would always doubt the Authenticity of these reports. The decision tree also shows that the second most significant factor is the investigating agency, which implies that there might be a huge gap between how these agencies handle cases. This result is reasonable as the BIA represents the internal power of the police department, while the IPRA represents the civilian.

The third most common and important factor is the severity of the allegation. This is also rational as common sense that punishment should be conducted by the crime. However, different complainant types would result in different tolerance of the severity. For cases coming from the police department, the cases are more likely to be sustained when the allegation is level 2 or worse, but, for cases from the civilian, BIA are more likely to sustain the case when the allegation is level 3 or worse. This difference suggests that the standard of sentencing is much different between cases from civilian and the department; police officers who are reported by the civilian have a lower risk of being sustained than those who are reported through the internal system.

While cases are commonly determined by these three factors, IPRA, as a third-party civilian institution, examines the working experience of the police officer in cases reported by civilians. In such cases, experienced officers are expected not to have allegations and have a higher chance of being

sustained when they do. It is hard to tell whether the criteria of investigating allegations are reasonable or not as it is related to other factors such as the rate of fake reporting and officers' personality. However, from the results shown by the study, we have learnt that civilians are not given the same treatment when reporting an allegation. Such inequity could result in trust issues according to Goldsmith (2015), and it could lead to further opposition between civilians and the police department. A proper explanation from the police department, including how different complainant types are considered differently could help relieve such distrust.

## 5 CONCLUSION & FUTURE WORKS

In this research, we have examined how different factors will affect the sustained rate of police allegation investigation, and we have found that there is excessive inequality from the data. Cases from the police department are more likely to get sustained than those from civilians. For cases from the civilian, the civilian agency values the employee history of the officers while the internal bureau focuses on the severity of cases. Only cases that are more severe than level 2 show a higher sustain rate than other cases; for civilian cases, the criteria rise to level 3. According to the result, we found a disturbing truth that, even with improving police supervision activities, civilians are still experiencing a hard time getting along with the police department; excessive force the civilians are experiencing are only judged to be sustained when they cause injury or casualties, and a higher standard might be applied when the case is judged by the internal agency. Even if they have become the victim of police misconduct, they are not very capable of retaking their justice as both agencies do not sustain cases from civilians very often. It is urgent for the police department to figure out a way to aid their accountability with the civilian by improving their allegation investigating policy. Civilians should have equal treatment while reporting police allegations, and allegations that rises the confrontation between civilians and the police department, such as racial slur and excessive forces that do not cause severe injury, should be given a heavier punishment.

As this study only focused on the dataset provided by the CPDP, it has a limitation that only the Chicago Police Department is considered; the detailed information of cases and officers are also not included in the current dataset. A better result could be

produced when more data sets are publicly available for research uses. Besides, more machine learning algorithms could be applied to the data set to explore the inequality the civilians are facing when dealing with police allegations.

## REFERENCES

- CPDP. (n.d.). Retrieved August 09, 2020, from <https://cpdp.co/>
- Goldsmith, A. (2005). Police reform and the problem of trust. *Theoretical Criminology*, 9(4), 443-470. doi:10.1177/1362480605057727
- Headley, A. M., D'Alessio, S. J., & Stolzenberg, L. (2017). The Effect of a Complainant's Race and Ethnicity on Dispositional Outcome in Police Misconduct Cases in Chicago. *Race and Justice*, 10(1), 43-61. doi:10.1177/2153368717726829
- Dowler, K., & Zawilski, V. (2007). Public perceptions of police misconduct and discrimination: Examining the impact of media consumption. *Journal of Criminal Justice*, 35, 193-203.9
- Long, M. A., Cross, J. E., Shelley, T. O., & Ivkovic, S. K. (2013). The normative order of reporting police misconduct: Examining the roles of offense, seriousness, legitimacy, and fairness. *Social Psychology Quarterly*, 76, 242-267
- Gottschalk, P. (2011). Police misconduct behavior: An empirical study of court cases. *Policing*, 5, 172-179
- Kane, R. J. (2002). The social ecology of police misconduct. *Criminology*, 40, 867-896
- Attard, B. O. (2020). *Police misconduct complaint investigations manual*. Place of publication not identified: Routledge.
- Terrill, W., & Ingram, J. R. (2015). Citizen Complaints Against the Police. *Police Quarterly*, 19(2), 150-179. doi:10.1177/1098611115613320
- Littlejohn, E. J. (1981). *The civilian police commission: deterrent of police misconduct*. University of Detroit Journal of Urban Law, 59(1), 5-62.
- Leven, Rachel L., et al. "What's Really New About Chicago's Newest Police Oversight Office?" Better Government Association, (3 Aug. 2020), [www.bettergov.org/news/whats-really-new-about-chicagos-newest-police-oversight-office](http://www.bettergov.org/news/whats-really-new-about-chicagos-newest-police-oversight-office).
- Luna, J. M., Gennatas, E. D., Ungar, L. H., Eaton, E., Diffenderfer, E. S., Jensen, S. T., Valdes, G. (2019). Building more accurate decision trees with the additive tree. *Proceedings of the National Academy of Sciences*, 116(40), 19887-19893. doi:10.1073/pnas.1816748116
- Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic Decision Making and the Cost of Fairness. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi:10.1145/3097983.3098095
- Juliette Gutierrez, Gondy Leroy (Dec 2007). Predicting Crime Reporting with Decision Trees and the National Crime Victimization Survey. *AMCIS 2007 Proceedings*.
- A. H. Wibowo, T. I. Oesman (2020). The comparative analysis on the accuracy of k-NN, Naive Bayes, and Decision Tree Algorithms in predicting crimes and criminal actions in Sleman Regency. *iCAST-ES 2019*.
- James P. Mcelvain, Augustine J. Kposowa (2008). Police Officer Characteristics and the Likelihood of Using Deadly Force. *Criminal justice and behavior*, 2008-04, Vol.35 (4), p.505-521.
- Caroline W., Bin H., Bhrij P., Feroze M., Cynthia R. (May 2020), In Pursuit of Interpretable, Fair and Accurate Machine Learning for Criminal Recidivism Prediction.
- Rozema, Kyle, Schanzenbach, Max. (May 2019), Good Cop, Bad Cop: Using Civilian Allegations to Predict Police Misconduct. *American economic journal. Economic policy*, 2019-05, Vol.11 (2), p.225-268.
- Raymond W. Patterson, (Dec 2006), Resolving Civilian-Police Complaints in New York City: Reflections on Mediation in the Real World. *Ohio State Journal on Dispute Resolution*
- Darrel W. Stephens, (Jun 2011), Police Discipline: A case for Change.
- Wang, P., Mathieu, R., Ke, J., & Cai, H. J. (2010). Predicting Criminal Recidivism with Support Vector Machine. 2010 International Conference on Management and Service Science. <https://doi.org/10.1109/icmss.2010.5575352>