Using Machine Learning Algorithms for Predicting Electoral Violence

Maya A. Dalton

PLSC 597 - Machine Learning

December 2023

Abstract

A majority of the research on electoral violence emphasizes the estimation of causal parameters – i.e., figuring out what *causes* electoral violence. Even though elections are regular and cyclical in nature, making the timing of electoral violence more predictable than other violent events, scholars have not yet approached the task of *predicting* electoral violence. Now that the state of the literature has a multitude of theoretical explanations for electoral violence, scholars should ensure their theories make correct predictions. In this article, I offer a comparison of the performances of three machine learning approaches to predicting electoral violence: L-2 logistic regression, supervised vector machines, and random forests. Furthermore, I replicate a well-cited causal model of electoral violence using the best-performing algorithm. I find that the driving predictors of a causal model differ from those in a machine learning algorithm. Accurately predicting electoral violence is critical to scholars, policymakers, and the electorate when anticipating hostile election cycles.

Introduction

In recent years, political violence has increased approximately 27% – or by nearly 27,000 events – diverging from several years of declining violence levels. Many of these recorded events also include instances of election-related violence, especially in Kenya, Nigeria, Palestine, India, and Latin America (ACLED, 2022). Electoral violence occurs within a quarter of all national elections worldwide (Hafner-Burton, Hyde and Jablonski, 2014). In 2020, 54% of national elections included some form of violence (Besaw, 2021). With the recent influx of contentious elections, political scientists have worked to figure out what exactly we know about why electoral violence occurs and what can be done to prevent it. The research on electoral violence has especially become more prevalent at the turn of the 21st century, as scholars diverged from previous work on election misconduct and elections as a trigger for civil wars (Birch, Daxecker and Höglund, 2020).

Traditionally, statistical research in the field has been concerned with identifying causal effects rather than prediction (Beck, King and Zeng, 2000). A majority of the research on electoral violence emphasizes the estimation of causal parameters – i.e., figuring out what *causes* electoral violence. This has resulted in a laundry list of predictors of electoral violence, each of which is argued to have "causal effects" on the risk of electoral violence. Even though elections are regular and cyclical in nature, making the timing of electoral violence more predictable than other violent events, scholars have not yet approached the task of *predicting* electoral violence (Birch, Daxecker and Höglund, 2020). Now that the state of the literature has a multitude of theoretical explanations for electoral violence, each of which claims to affect the risk of such events occurring, we should make sure whether the theories make correct predictions.

The international conflict literature was once in a similar state. There was an emphasis on the estimation of causal parameters, resulting in a long list of features that could cause civil war, interstate disputes, or general political instability. Scholars began to dispute over which model of conflict was the best at estimating causal effects, with no regard for predictive power. Now, some work has begun importance of using out-of-sample data to evaluate and compare models (Goldstone et al., 2010; Goldsmith et al., 2013; Muchlinski et al., 2016; Halterman and Radford, 2023), allowing scholars to know if and when their theories are useful in practice. In the electoral violence literature, it now comes time to shift the approach from the causal effects of electoral violence to predictive applications.

Some recent work has intersected machine learning techniques and electoral violence, but only in the sense of classification and measurement (Muchlinski et al., 2021). Predicting electoral violence via machine learning algorithms is beneficial for two reasons. Firstly, the predictive performance of current statistical models of electoral violence has not been examined. Therefore, we do not know if the "true" models of electoral violence are actually useful or have high predictive power. Secondly, theoretically significant indicators of electoral violence remain dependent on context, leading to disputes over which models and variables are even important. For example, specific institutional designs may lead to a higher risk of electoral violence in authoritarian regimes, but not within democracies. Using these algorithms for prediction tasks allows for a benchmark of which predictors are important to predicting electoral violence under specific contexts. It is important to note that a main challenge in examining the causes of electoral violence is that it often takes place in contexts where other forms of violence are already pervasive. Thus, uncovering additional predictors of electoral violence becomes more difficult when a given country is already riddled with a violent past.

Nonetheless, I approach electoral violence with the goal of prediction, comparison of model fit, and assessment of variable importance. In the remainder of this paper, I compare the ability of three machine learning techniques to predict true occurrences of electoral violence. I then examine the predictive power of a well-cited logistic regression model from

Fjelde and Höglund (2016). Taking the machine learning technique with the highest predictive power, I examine variable permutation importance between the algorithm and the logistic regression model to understand which variables indeed influence the likelihood of electoral violence, and which may be theoretically strong, but empirically unimportant. Because electoral violence is not only destructive but undermines democratic ideals of free and fair elections, accurately predicting their occurrence is critical to scholars of electoral and political violence, as well as to policymakers and the electorate when anticipating hostile election cycles.

Understanding Electoral Violence

Electoral violence is commonly defined as "any random or organized act or threat to intimidate, physically harm, blackmail or abuse a political stakeholder in seeking to determine, delay or to otherwise influence an electoral process (Fischer, 2001)". Electoral violence has influential consequences on the societies in which it occurs. Violent electoral strategies can be employed by either the electorate or political elites. In the case of elites, violent acts are often used as a strategy to exclude certain groups or actors from the political sphere (Birch, Daxecker and Höglund, 2020). Violence can also be used by elites to reduce uncertainty about the electoral outcome via disenfranchising voters or repressing opposition (Staniland, 2014; Birch, Daxecker and Höglund, 2020). From the perspective of the electorate, violence can be used to signal dissatisfaction with the current regime and institutions. Electoral violence can not only shape immediate electoral outcomes and elite representation but also undermine the possibility of democratization (Staniland, 2014). For example, Nigeria's most recent general election held in February 2023 was riddled with instances of disorganization and electoral violence. As this was the most contested election in the 21st century, the country saw groups of armed individuals target polling stations, specifically those in areas with larger support for the opposition (Akinwotu, 2023). Similarly, Papua New Guinea's national elections last summer resulted in a larger-scale conflict which led to a death count of 50 civilians, the extensive burning of schools and public buildings, and an estimated 90,000 people displaced (Kuku, 2022).

Why do such acts of violence occur during election time? Existing literature has focused on international, institutional, and societal features that moderate electoral violence. This research has focused primarily on the effect of **institutions**, such as electoral systems (Fjelde and Höglund, 2016; Hafner-Burton, Hyde and Jablonski, 2014), party strength (Fjelde, 2020; Siddiqui, 2022), party competitiveness (Collier and Vicente, 2012; Seeberg, Wahman and Skaaning, 2018), or incumbency effects (Hafner-Burton, Hyde and Jablonski, 2014, 2018). On a larger scale, some research has found that **international** election monitoring has mixed effects on electoral violence (Daxecker, 2012; Smidt, 2016; Birch and Muchlinski, 2018; von Borzyskowski, 2019). Less attention, however, has been paid to the **societal** determinants of electoral violence (Birch, Daxecker and Höglund, 2020). Societal indicators can be further understood as attitudes towards gender norms in the government and society, religious beliefs, political ideology, or access to the political environment by minority groups. For example, research has found that ethnic polarization and the exclusion of ethnic groups from power are linked to greater incentives for electoral violence (Nellis and Siddiqui, 2018; Nellis, Weaver and Rosenzweig, 2016; Fjelde and Höglund, 2016; Wilkinson, 2004). Additionally, there have been a handful of studies looking at gender-based effects of political violence, where women are disproportionally affected by such acts (Agbalajobi, 2016; Kishi, 2021; Krook and Sanín, 2020; Bardall, 2016). Within each of these studies, authors argue their predictors have strong causal effects on electoral violence. This implies that we could "predict" whether a given country will have electoral violence in a given year based on at least some of these attributes. Despite the plethora of research proposing theoretical models of the causes of electoral violence, we have not examined their predictive powers, nor attempted to predict electoral violence using machine learning techniques. The question remains: Do theoretically and causally sound models of electoral violence have high predictive power when using machine learning for prediction?

Research Design

Part I - Comparing Machine Learning Models

My research design is two-fold. I first compare the ability of three popular statistical learning models to correctly predict true instances of electoral violence in out-of-sample data: (1) Regularized (L-2) Logistic Regression, (2) Supervised Vector Machines, and (3) Random Forests. The regularized L-2 logistic regression model (hereafter referred to as L2 logit) handles any potential issues with collinearity or overfitting in my data. This model serves as the "baseline" model to which I compare the other machine learning approaches. The supervised vector machine (SVM) is a supervised learning model often used to analyze data for classification. The goal of the SVM is to design a hyperplane that classifies all the training vectors in two classes: violence or non-violent elections. Finally, random forest algorithms are beneficial because they decrease the model's variance, are less sensitive to outliers in the dataset, and don't require much parameter tuning. This model will recover predicted electoral violence as the average prediction across all decision trees.

My dependent variable is derived from the Electoral Contention and Violence (ECAV) dataset contains information on nonviolent and violent contention related 1,208 to national in 136 countries from 1990-2012 (Daxecker, Amicarelli and Jung, 2019). Y_{ij} is a binary measure of whether electoral violence occurred for a given country, i, in a given year j. $\mathbf{X}$ is a matrix of predictor variables derived from the Varieties of Democracy dataset (Varieties of Democracy, 2023). These variables include previous electoral violence, electoral system, political competition, regime type, international election monitoring, candidate restrictions, social group power opportunities, and gross domestic product (GDP). These predictor variables encompass institutional, international, and societal indicators of electoral violence set forth by previous literature and summarized by Birch, Daxecker and Höglund (2020). While they are not comprehensive, they allow a good overview of variables found to have causal effects on electoral violence.

For analyzing predictive performance, I utilize ten-fold cross-validation (CV) to examine how well each of the five approaches performs out-of-sample. Cross-validation will take the dataset of combined variables from ECAV and VDEM into different "folds". Ten folds will be used to train the model, while a separate fold is held out to test the predictions made by the model in the training data. The model is first trained on the training CV folds and then examines out-of-sample data. Additionally, I visualize the predictive performance of each binary classifier via a receiver operating characteristic area under the curve (ROC-AUC). I also compute the precision-recall area-under-the-curve (PR-AUC) curve to examine the performance of the algorithms in case of class imbalance. As random forests have been shown to generate substantially more accurate predictions than traditional parametric methods (Montgomery, Hollenbach and Ward, 2012), I expect the random forest to perform the best at predicting electoral violence compared to the L2 logit and SVM models.

Part II - Replication Illustration

In the second part of my research design, I analyze the predictive performance of a well-cited logistic regression model developed by Fjelde and Höglund (2016). They test a previous theory that links the use of violent electoral tactics to the high stakes put in place by majoritarian electoral institutions. In their analysis of Sub-Saharan Africa, the authors find that majoritarian systems are more likely to cause electoral violence where large ethnic groups are excluded from power. Specifically, the predicted probability of electoral violence is 2.1% in a majoritarian system compared to 0.8% in a PR system. The authors argue that based on theoretical understandings of institutional design and their findings, scholars can anticipate electoral violence in African majoritarian systems with ethnic exclusions from power (Fjelde and Höglund, 2016).

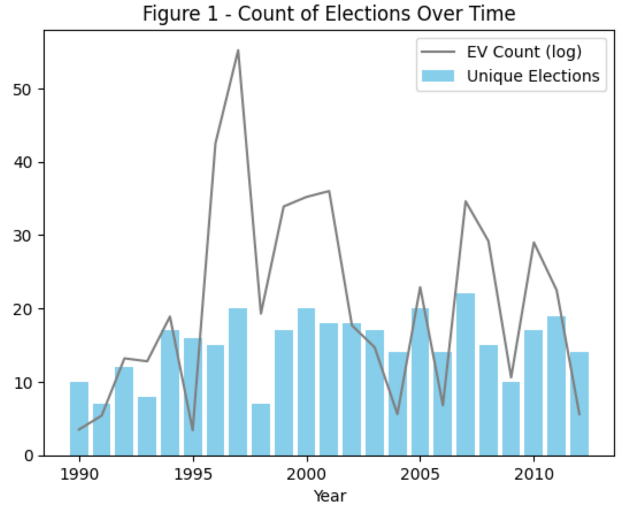
The data is combined from multiple data sources to uncover the predicted probability of electoral violence under certain institutional contexts in Sub-Saharan Africa from 1990-2010. Their dependent variable, derived from the Social Conflict in Africa Database (SCAD), is measured as instances of social disturbances in Africa – such as demonstrations, strikes, riots, and government harassment – focused on those violent events that are coded as explicitly related to elections (Idean Salehyan and Williams, 2012). Their main independent variable – Majoritarian Rules – is a dummy variable coded from the Database of Political Institutions (Regan, Frank and Clark, 2009), which captures whether the country employs either plurality (first-past-the-post) or majoritarian formulas in the election of legislators. Their model further includes control variables of institutional and international indicators for electoral violence, such as mean district magnitude, an indicator of a mixed system, democracy index, population, and GDP. They also include two lagged variables for previous armed conflict and previous electoral violence.

Using the best-performing algorithm from Part I, I replicate Fjelde and Höglund (2016)'s analysis to test if their conclusions hold under a true predictive task. Because their findings are well-cited and empirically strong under a causal approach, it is important to analyze if their theories and results make correct predictions of electoral violence. As with Part I, I analyze predictive performance using ten-fold CV to examine how well this model performs out-of-sample, Additionally, I visualize the predictive performance via the ROC-AUC and PR-AUC. Finally, to compare the machine learning algorithm to Fjelde and Höglund (2016)'s main findings, I examine the permutation importance of the predictor variables from each model. By analyzing which variables are most influential in predicting electoral violence, one can understand if variables that have a causal effect are actually significant in a prediction task. I expect the machine learning algorithm to attribute importance to different variables than those of Fjelde and Höglund (2016)'s conclusion.

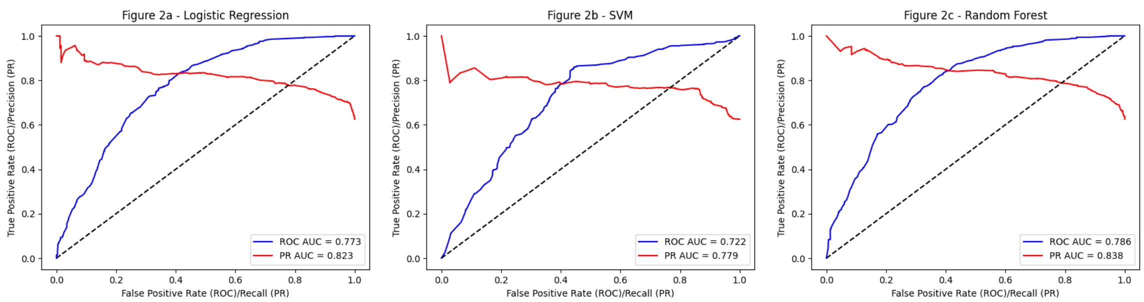
Analysis

The trends in electoral violence from 1990 to 2012 are shown in Figure 1. The figure shows the count of electoral violence per election year alongside the number of national elections that took place each year in the time period studied. The only clear trend in the data is that electoral violence seems to have become more frequent around 1995, which could be coincidental with the dissolution of the former Soviet Union. The peaks in the prevalence of violence do not seem to correspond to the number of elections taking place at that particular time.

Figure 1 - Count of Elections Over Time



I evaluate the performance of three machine learning algorithms in predicting electoral violence. Figures 2 show the receiver operating characteristic (ROC) and precision-recall (PR) curves for each model. The closer the ROC curve is to the upper left corner of the graph, the higher the accuracy of the test. The PR curve, on the other hand, is used for evaluating the performance of binary classification algorithms in situations where classes are heavily imbalanced. For both curves, the AUC provides an aggregate measure of performance across all possible classification thresholds. In other words, the closer the AUC to 1, the better a model performs for predicting electoral violence.



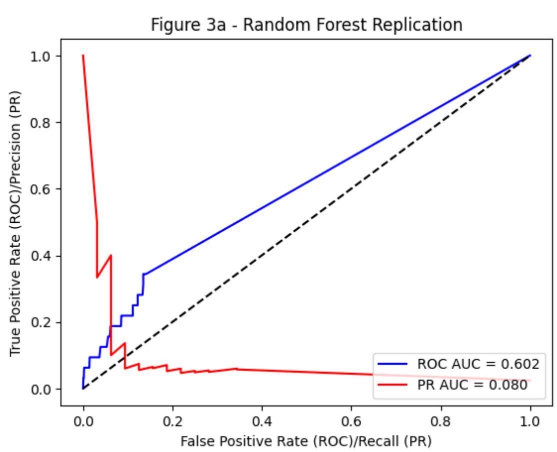
As can be seen in Figure 2a, the logistic regression model performs well at predicting electoral violence, with an ROC AUC of 0.773 and a PR AUC of 0.823. This is significantly better than randomly guessing which cases are positive or negative. Furthermore, as the recall rate increases, the precision only slightly decreases. This means as the model recovers more positive cases, a significant amount of positive predictions actually belong to the positive class. Figure 2b depicts the performance of the SVM model, which is slightly worse than the logit model. This model has an ROC AUC of 0.722 and a PR AUC of 0.779. However, the precision-recall trade-off is more stable in this model, meaning that even as recall increases, the precision score remains relatively the same. Finally, as shown in Figure 2c, the random forest model performs the best of the three algorithms. While this is only slightly better than the logistic regression model, it has an ROC AUC of 0.786 and a PR AUC of 0.838.

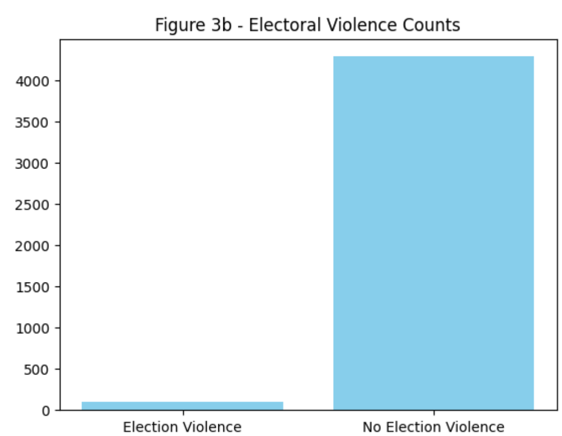
To further compare the three models, Table 1 predicts the mean-squared errors (MSE) and accuracy scores for each. Interestingly, the logit and SVM models have the same MSE and test accuracy, with the logit performing slightly better in training accuracy. In other words, the logit and SVM models perform the same in out-of-sample prediction. The random forest, however, has the highest test accuracy and lowest MSE, meaning it has the smallest distance between actual and predicted values and performs better in out-of-sample prediction.

Table 1: Accuracy Scores

Model	MSE	Training Acc.	Test Acc.
L2 Logit	0.25698	0.75131	0.74302
SVM	0.25698	0.75112	0.74302
Random Forest	0.24651	0.74944	0.75349

Because the random forest performs the best in out-of-sample prediction, I use it to analyze whether theoretical explanations for electoral violence are sufficient for prediction. Using Fjelde and Höglund (2016)'s data, I fit a random forest model to predict electoral violence. I then examine variable permutation importance between the algorithm and the logistic regression model to understand which variables indeed influence the likelihood of electoral violence, and which may be theoretically strong, but empirically unimportant. Figure 3a depicts the ROC and PR curves for this random forest model. Surprisingly, this model does not perform as well as the previous one. The ROC AUC is 0.602, which is only slightly better than random guessing. Furthermore, the PR AUC curve is 0.08. This is because Fjelde and Höglund (2016)'s dataset is heavily imbalanced, as shown in Figure 3b.





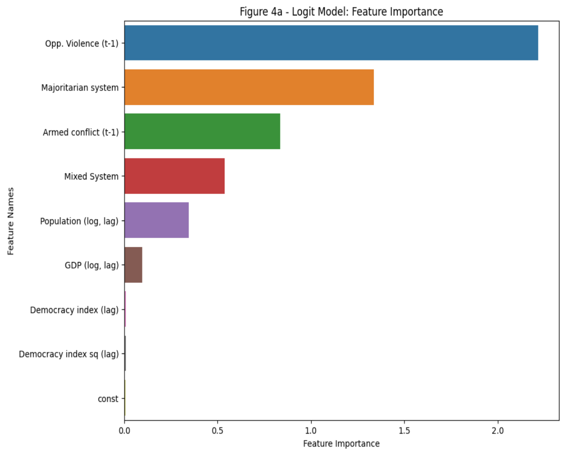
Because of the strong class imbalance, the model produces insufficient accuracy scores.

Table 2: Accuracy Scores

Model	**MSE**	**Training Acc.**	**Test Acc.**
Random Forest (Imbalanced)	0.025	0.97428	0.97494
Random Forest (Class Weights)	0.134	0.90755	0.8656

As shown in Table 2 (Row 1), the random forest model has an almost perfect training and test accuracy score of 0.974. One may be led to believe that because the out-of-sample accuracy score is so high, the model performs extremely well on this data. In this context, however, electoral violence is a rare event, meaning the model is more likely to predict cases in which electoral violence *does not occur* over violent events. Therefore, I refit the random forest model using class weight optimization to handle the imbalanced dataset. Table 2, Row 2 depicts the MSE and accuracy scores of the class-weighted model. While the performance has decreased slightly, the model still performs relatively well in out-of-sample prediction, with a test accuracy of 0.865.

After addressing the class-imbalance issue, I can now turn to assessing of variable importance between the original model and the random forest model. Figures 4 depict the variable permutation importance between the random forest model and the logistic regression model from Fjelde and Höglund (2016), which can help to understand which variables indeed influence the likelihood of electoral violence, and which may be theoretically strong, but empirically unimportant.



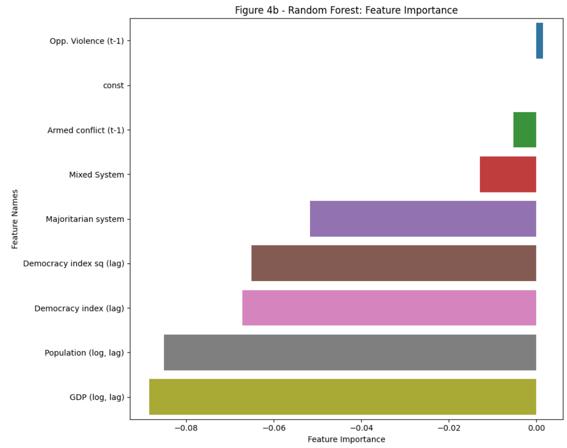


Figure 4a depicts the important features of the logistic regression model from Fjelde and Höglund (2016). As previously mentioned, they find that majoritarian systems are more likely to cause electoral violence where large ethnic groups are excluded from power. Therefore, the driving predictors in their model are previous *opposition violence*, majoritarian systems, and previous *armed conflict*. Figure 4b depicts the feature importance of the class-weighted random forest model. For predicting electoral violence, previous *opposition violence* is the driving feature of this model, and it is a lot less important than in the logit model. All other variables, including the main independent variable from Fjelde and Höglund (2016), have negative permutation importance, meaning they are unimportant in predicting electoral violence. This finding is in alignment with previous theoretical understandings of "conflict traps", in which countries that have previously experienced conflict often struggle to escape the recurring cycles of violence (Malone, 2022; Collier et al., 2003). While both models are driven by previous electoral violence in some way, the main predictor from Fjelde and Höglund (2016) is not empirically important when predicting electoral violence using a random forest model.

Conclusion

This study represents a first step in predicting electoral violence and benchmarking which of the many predictors of electoral violence actually matter. Using cross-national electoral violence data, with various institutional, international, and societal predictors, I find that the random forest model performs best at out-of-sample predicting, closely followed by L-2 regularized logistic regression. To further explore predictions of electoral violence, I replicate a well-cited causal model investigating the relationship between the electoral system and election violence (Fjelde and Höglund, 2016). In this replicated study, I find two important conclusions. Firstly, the authors' data is highly imbalanced, meaning that in their sample of African electoral districts from 1990-2010, electoral violence is extremely rare. Therefore, when using such data for prediction, I use class weight optimization to control for misleading results. Using the class-weighted random forest model for this prediction task, I find that the model does fairly well at out-of-sample prediction of electoral violence. Secondly, I find that the permutation importance of predictors in each model differs significantly. When using their data for prediction, Fjelde and Höglund (2016)'s causal conclusion does not hold. The sole driver of predicting electoral violence in this case is previous electoral violence.

Elections are a highly sourced topic for prediction and forecasting due to their cyclical timing, meaning electoral violence should be more predictable than other violent events. Even so, no other literature has yet taken the steps toward predicting electoral violence with machine learning approaches. Because of this gap in the literature, scholars are left relying on causal effect models to indicate if certain predictors can be used to anticipate violence in election cycles. This is not to say that causal effect models are unimportant. Instead, machine learning tactics for prediction should be used as a cross-reference to ensure that theoretically (and even empirically) strong indicators are strong in prediction tasks as well. As seen in this study, some "true" causal models of electoral violence are not necessarily useful, nor have high predictive power.

This study has focused on only a subset of approaches to prediction with machine learning, as well as, a replication study focused on electoral violence in Sub-Saharan Africa. Firstly, future works could utilize other machine learning algorithms to assess whether these findings are applicable in other contexts. Secondly, there is now the task of testing if other well-cited causal models are up to par in prediction settings. Subsequent works could assess if previous theoretical and empirical findings of the causes and effects of electoral violence can be reproduced with machine learning. Additionally, authors could potentially use machine learning approaches as robustness checks to support their causal findings. Electoral violence is becoming more prevalent in national and sub-national elections, therefore accurately predicting their occurrence is critical to scholars of electoral and political violence, as well as to policymakers and the electorate when anticipating hostile election cycles. My findings emphasize that predicting electoral violence with machine learning should be considered alongside causal effect models, as well as, considering issues of strong class imbalance in conducting such analyses.

References

ACLED. 2022. "ACLED Year in Review: Global Disorder in 2022.".
URL: *https://acleddata.com/2023/01/31/global-disorder-2022-the-year-in-review/*

Agbalajobi, Damilola Taiye. 2016. "Electoral violence and women's participation in electoral processes: challenges and prospects for future elections in Nigeria." *Unilag Sociological Review (USR)* XII:73–94.

Akinwotu, Emmanuel. 2023. "Nigeria's general election saw disorganization and violence." *NPR* .
URL: *https://www.npr.org/2023/02/27/1159822850/nigerias-general-election-saw-disorganization-and-violence*

Bardall, Gabrielle. 2016. Towards a More Complete Understanding of Election Violence: Introducing a Gender Lens to Electoral Conflict Research.

Beck, Nathaniel, Gary King and Langche Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture." *The American Political Science Review* 94(1):21–35. Publisher: [American Political Science Association, Cambridge University Press].
URL: *https://www.jstor.org/stable/2586378*

Besaw, Clayton. 2021. "Election violence spiked worldwide in 2020 – will this year be better?".
URL: *http://theconversation.com/election-violence-spiked-worldwide-in-2020-will-this-year-be-better-153975*

Birch, Sarah and David Muchlinski. 2018. "Electoral violence prevention: what works?" *Democratization* 25(3):385–403. Publisher: Routledge _eprint: https://doi.org/10.1080/13510347.2017.1365841.
URL: *https://doi.org/10.1080/13510347.2017.1365841*

Birch, Sarah, Ursula Daxecker and Kristine Höglund. 2020. "Electoral violence: An introduction." *Journal of Peace Research* 57(1):3–14. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343319889657*

Collier, Paul and Pedro C. Vicente. 2012. "Violence, bribery, and fraud: the political economy of elections in Sub-Saharan Africa." *Public Choice* 153(1):117–147.
URL: *https://doi.org/10.1007/s11127-011-9777-z*

Collier, Paul, V. L. Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol and Nicholas Sambanis. 2003. *Breaking the Conflict Trap : Civil War and Development Policy.* Washington, DC: World Bank and Oxford University Press.
URL: *http://hdl.handle.net/10986/13938*

Daxecker, Ursula E. 2012. "The cost of exposing cheating: International election monitoring, fraud, and post-election violence in Africa." *Journal of Peace Research* 49(4):503–516. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343312445649*

Daxecker, Ursula, Elio Amicarelli and Alexander Jung. 2019. "Electoral contention and violence (ECAV): A new dataset." *Journal of Peace Research* 56(5):714–723. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343318823870*

Fischer, Jeff. 2001. "Electoral Conflict And Violence.".
URL: *https://ciaotest.cc.columbia.edu/olj/et/et$_v$12n1/et$_v$12n1b.pdf*

Fjelde, Hanne. 2020. "Political party strength and electoral violence." *Journal of Peace Research* 57(1):140–155. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343319885177*

Fjelde, Hanne and Kristine Höglund. 2016. "Electoral Institutions and Electoral Violence in Sub-Saharan Africa." *British Journal of Political Science* 46(2):297–320. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/british-journal-of-political-science/article/electoral-institutions-and-electoral-violence-in-subsaharan-africa/3EAE320E957EEFFCC2D23C0253C8CAD0*

Goldsmith, Benjamin E, Charles R Butcher, Dimitri Semenovich and Arcot Sowmya. 2013. "Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003." *Journal of Peace Research* 50(4):437–452. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0022343313484167*

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder and Mark Woodward. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science* 54(1):190–208. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2009.00426.x.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5907.2009.00426.x*

Hafner-Burton, Emilie M., Susan D. Hyde and Ryan S. Jablonski. 2014. "When Do Governments Resort to Election Violence?" *British Journal of Political Science* 44(1):149–179. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/british-journal-of-political-science/article/abs/when-do-governments-resort-to-election-violence/2BE13A1FC0696CC41757F2E733A59B32*

Hafner-Burton, Emilie M., Susan D. Hyde and Ryan S. Jablonski. 2018. "Surviving Elections: Election Violence, Incumbent Victory and Post-Election Repercussions."

British Journal of Political Science 48(2):459–488. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/british-journal-of-political-science/article/abs/surviving-elections-election-violence-incumbent-victory-and-postelection-repercussions/25C94A3E06679735045C309A5CB38DF7*

Halterman, Andrew and Benjamin J. Radford. 2023. "Latent Civil War: Improving Inference and Forecasting with a Civil War Measurement Model." *PolMeth Conference Paper* .

Idean Salehyan, Cullen S. Hendrix, Jesse Hamner Christina Case Christopher Linebarger Emily Stull and Jennifer Williams. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38(4):503–511. Publisher: Routledge _eprint: https://doi.org/10.1080/03050629.2012.697426.
URL: *https://doi.org/10.1080/03050629.2012.697426*

Kishi, Roudabeh. 2021. "Violence Targeting Women in Politics: Trends in Targets, Types, and Perpetrators of Political Violence.".
URL: *https://acleddata.com/2021/12/08/violence-targeting-women-in-politics-trends-in-targets-types-and-perpetrators-of-political-violence/*

Krook, Mona Lena and Juliana Restrepo Sanín. 2020. "The Cost of Doing Politics? Analyzing Violence and Harassment against Female Politicians." *Perspectives on Politics* 18(3):740–755. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/perspectives-on-politics/article/cost-of-doingpolitics-analyzing-violence-and-harassment-against-femalepoliticians/997569433135FA170B2789C88A48DD17*

Kuku, Rebecca. 2022. "PNG election violence: 90,000 displaced since May, 25,000 children unable to attend school." *The Guardian* .
URL: *https://www.theguardian.com/world/2022/sep/28/png-election-violence-90000-displaced-since-may-25000-children-unable-to-attend-school*

Malone, Iris. 2022. "Recurrent neural networks for conflict forecasting." *International Interactions* 48(4):614–632. Publisher: Routledge _eprint: https://doi.org/10.1080/03050629.2022.2016736.
URL: *https://doi.org/10.1080/03050629.2022.2016736*

Montgomery, Jacob M., Florian M. Hollenbach and Michael D. Ward. 2012. "Improving Predictions using Ensemble Bayesian Model Averaging." *Political Analysis* 20(3):271–291.

Muchlinski, David, David Siroky, Jingrui He and Matthew Kocher. 2016. "Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data." *Political Analysis* 24(1):87–103. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/political-analysis/article/abs/comparing-random-forest-with-logistic-regression-for-predicting-classimbalanced-civil-war-onset-data/109E1511378A38BB4B41F721E6017FB1*

Muchlinski, David, Xiao Yang, Sarah Birch, Craig Macdonald and Iadh Ounis. 2021. "We need to go deeper: measuring electoral violence using convolutional neural networks and social media." *Political Science Research and Methods* 9(1):122–139. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/political-science-research-and-methods/article/we-need-to-go-deeper-measuring-electoral-violence-using-convolutional-neural-networks-and-social-media/808197D6EDA72699670325B6320EA5B9*

Nellis, Gareth, Michael Weaver and Steven Rosenzweig. 2016. "Do Parties Matter for Ethnic Violence? Evidence From India." *Quarterly Journal of Political Science* 11:249–277.

Nellis, Gareth and Niloufer Siddiqui. 2018. "Secular Party Rule and Religious Violence in Pakistan." *American Political Science Review* 112(1):49–67. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/american-political-science-review/article/abs/secular-party-rule-and-religious-violence-in-pakistan/3BC44B82306A9C1FFD08998F758BDCE2*

Regan, Patrick M., Richard W. Frank and David H. Clark. 2009. "New Datasets on Political Institutions and Elections, 1972—2005." *Conflict Management and Peace Science* 26(3):286–304. Publisher: SAGE Publications Ltd.
URL: *https://doi.org/10.1177/0738894209104556*

Seeberg, Merete Bech, Michael Wahman and Svend-Erik Skaaning. 2018. "Candidate nomination, intra-party democracy, and election violence in Africa." *Democratization* 25(6):959–977. Publisher: Routledge _eprint: https://doi.org/10.1080/13510347.2017.1420057.
URL: *https://doi.org/10.1080/13510347.2017.1420057*

Siddiqui, Niloufer A. 2022. *Under the Gun: Political Parties and Violence in Pakistan.* Cambridge: Cambridge University Press.
URL: *https://www.cambridge.org/core/books/under-the-gun/406F2F67DB2468BD425B9DF9D8CE60C*

Smidt, Hannah. 2016. "From a perpetrator's perspective: International election observers and post-electoral violence." *Journal of Peace Research* 53(2):226–241. Publisher: Sage Publications, Ltd.
URL: *https://www.jstor.org/stable/43920011*

Staniland, Paul. 2014. "Violence and Democracy." *Comparative Politics* 47(1):99–118.

Varieties of Democracy (V-Dem) Project. 2023.

von Borzyskowski, Inken. 2019. "The Risks of Election Observation: International Condemnation and Post-Election Violence." *International Studies Quarterly* 63(3):654–667.
URL: *https://doi.org/10.1093/isq/sqz024*

Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge Studies in Comparative Politics Cambridge: Cambridge University Press.
URL: *https://www.cambridge.org/core/books/votes-and-violence/DED9BA09D07210B27297837295115*