CS 109a Final Project: Coups D'état Success

Justin Adjasu, Raelene Angle-Graves, Eliel Dushime, Dora Ivkovich, and George Popoola

A. Introduction and Background

Thus far in 2022, there have been nine coup d'état attempts across the world, five of which succeeded, meaning that over 65 million people in total experienced a major governmental change this year [1]. The success of a coup d'état is a major shift in the quality of life and certainty in the future of a country's inhabitants. The literature is divided as to the main determinants of coups, but a comprehensive study by Gassebner, Gutmann, and Voigt (2016) found that slow economic growth rates, previous coup experiences, and political violence to be strongly associated with coup occurrence [2]. However, the literature has not yet explored the determinants of coup success. That is, supposing that a coup or conspiracy is initiated, which factors are most strongly correlated with success?

B. Data and Approach

Our dataset, obtained from the work of the Coup D'état Project hosted by the Cline Center for Social Research, contains information regarding successful, attempted, and conspired coups d'état in 136 countries from 1945 to 2019 [3]. In the dataset, there are 29 columns and 943 observations (rows), where each observation represents a coup event. All of the columns (aside from those indicating date and location) are categorical variables and that indicate the coup status (e.g. realized, conspiracy), coup type (e.g. military, rebel), and deposed executive status (e.g. not harmed, killed, jailed). In order to predict coup success, we chose the "realized" column from the dataset as our response variable, which indicates whether or not the coup led to a change in the ruling body of a country.

From this data, we constructed several additional variables. First, we added numerical variables representing the number of coups in the past 5 years and in the entire history of that country. We also added a variable representing whether the country has undergone major regime changes (e.g., East vs. West Germany into Germany). We also added variables representing the number of previous failed coups and the number of previous successful coups. In our analysis, we plan to disregard the indicators representing deposed executive status, because this status is a consequence of coup success. However, these variables allowed us to construct a measure for the number of times a deposed leader has been hurt or killed, as well as another variable indicating the number of times in the past that a deposed leader had been tried in some way (house arrest, tried, exiled, etc.), which would indicate that a coup government with some power had been established in the past.

To supplement the dataset, we decided to include quantitative information from different data sources we believed would have some predictive power towards a coup's success.

From the Center for Systemic Peace, we merged our data with the Polity5 and Major Episodes of Political Violence datasets [4]. The Polity5 dataset contains measures relating to the type and operation of the current political regime in a country, marking features such as the openness of political office, constraints on the power of the chief executive, etc. The Major Episdoes of Political Violence Dataset contains information on the numbers and types of politically violent events occurring within a country and its neighbors within a given year, including events such as wars between nations, civil wars, assassinations, etc. Together, these two datasets provide information on the nature and stability of the current regime in a country. We believe this information will be useful for training our model as it logically follows that the harshness and/or instability of an existing regime creates greater reason for its deposing through a coup, a deduction that is supported by the work of Gassebner et al.

From Gapminder, we included a number of different columns describing information about the political/economic state (ex: Gini Index, GDP per capita) and standard of living (ex: population growth, gender ratio, life expectancy). As mentioned in the Gassebner study, a lack of economic growth contributes strongly to the likelihood of a coup, and the same can be said for widespread dissatisfaction with the average person's standard of living. While the effects of such on coup likelihood have been documented, it will be interesting to see to what extent and relative strengths they can also be predictive of coup success.

To frame this in Gassebner et al.'s primary predictors, the Major Episodes of Political Violence datasets provide us with information on political violence, the past coup variables give us information on the country's history, and the Gapminder data provides us with economic data. This incorporates the three features Gassebner found to be most predictive, as well as many others that the literature has cited in the past.

After merging all of our data, we obtained a full predictor set of 71 columns. However, some of the Polity-5 and MEPV (Major Episodes of Political Violence) data introduced 90 points of missingness into the dataset. This missingness includes codes for transitional regimes, which could not be scored. To combat this, we have used kNN imputation (countries missing only some observations/transitional regimes) and mean imputation (countries never recorded in the datasets) to fill our dataset. The kNN imputation allows us to approximate what the regime's score would have been had it been measured by comparing it to similar regimes, and the mean imputing fills gaps as if the country is the average in the dataset, which is reasonable giving that we have nothing to which we can compare it. We also built flag variables that would tell us whether a variable was missing an observation. This approach should allow us to determine whether the data were missing at random. If the flag appears as a significant variable, that would indicate that the missingness predicts a different outcome, so the data are not missing at random.

Because of the sheer number of predictors we have, and the potential that some of the data is not missing at random, we have decided to consider two datasets:

1. Including all coup events, dropping any columns with missingness (in other words, all Polity-5 and MEPV data)
2. Including all columns and all coup events, using mean imputation and kNN imputation to the nearest year on missing values for a country, maintaining missingness flags

For these two datasets, we will train 3 models each: a LASSO, a Random Forest, and an AdaBoost. This yields 6 models. We will perform hyperparameter tuning for the number of iterations/trees/the penalization term using a validation split after we split into training/testing data, which ensures we do not overfit to our training data. We feel that using this strategy is appropriate because it gives us the best of both worlds: one dataset allows us to draw conclusions without the potential skew, while using the other allows us to include more influential variables. In other words, they can cross-check one anothers' results.

C. Exploratory Data Analysis

To explore the data we started with a choropleth map that displays coups per country.

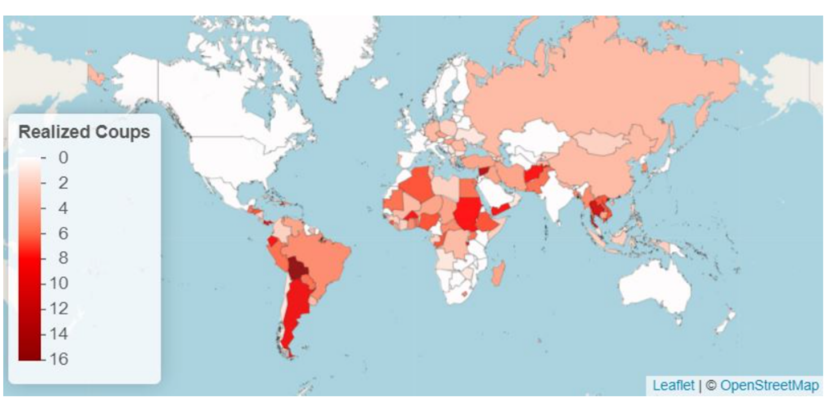

Figure 1a. Choropleth Map of Coups by Country [1]

Data Citation: Peyton, Buddy; Bajjalieh, Joseph; Shalmon, Dan; Martin, Michael; Bonaguro, Jonathan (2021): Cline Center Coup D'état Project Dataset. University of Illinois at Urbana-Champaign. https://doi.org/10.13012/B2IDB-9651987_V3

Next, we created bar charts for all of the relevant predictors, to see if any predictors made a coup more or less likely to succeed. Based on the plots, the only predictor that seems to have a clear visual impact on whether a coup is realized or not is dissident. These plots are provided in the appendix.

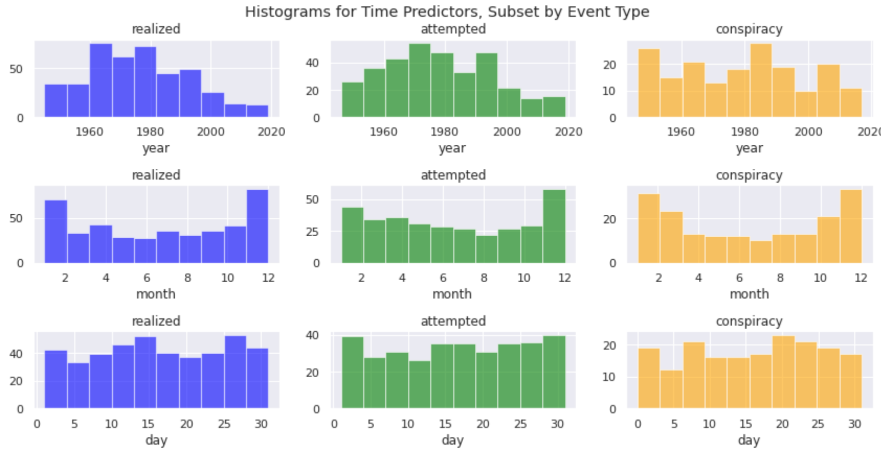

Figure 1b. Histograms for Time Related Predictors, Subset by Event Type

From the above histograms for predictors dealing with time (year, month, day) subset by event type we see that there is roughly a normal distribution in year for all 3 centered around the 1970s (disputable for conspiracy). We also notice what appears to be a seasonality to coup timing, with a preference for winter months (January and December most so).

Finally, when we plot the Polity-5 and the MEPV data, we see relatively Normal

D. Results and Analysis

I. LASSO

As our interest is in the feasibility of predicting a coup's success, we will be fitting classification algorithms to our data to make this prediction. However, we are also interested in what specific factors are most important in determining which coups were successful. Additionally, with our full dataset we run the risk of overfitting due to high dimensionality among our predictor set. In order to solve both of these problems, prior to fitting our classification algorithm we utilized LASSO regression as a method of feature selection.

Due to the high dimensionality of our dataset, we applied multiple rounds of LASSO regularization to limit our target predict set as much as possible. In the first round, we used LASSO with cross validation on the entire predictor set, optimizing the regularization value alpha.

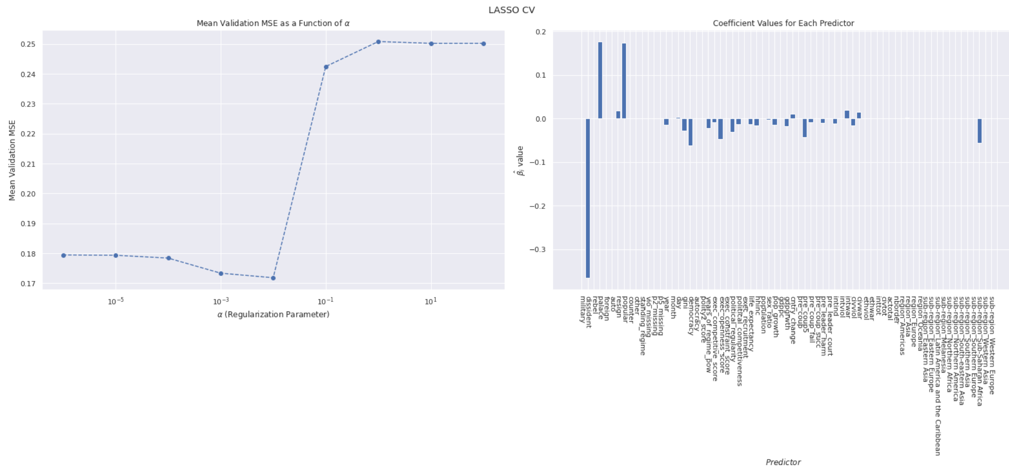

Figure 1. 1st Round LASSO: Alpha Optimization & Predictor Subset

From this figure, we can derive 2 useful pieces of information. The first is the optimal alpha parameter for our LASSO regression, that being .001. The second is the subset of predictors from our first round of LASSO that are worth bringing over to our next round, those being predictors with coefficient magnitudes above .01. Using this new subset of predictors and found alpha value, we performed a second round of LASSO regularization, in which we fit the LASSO regressor to 100 bootstrapped version of our dataset and kept track of how many times each predictor in our subset was found to be significant, which we again defined as having a coefficient magnitude above .01

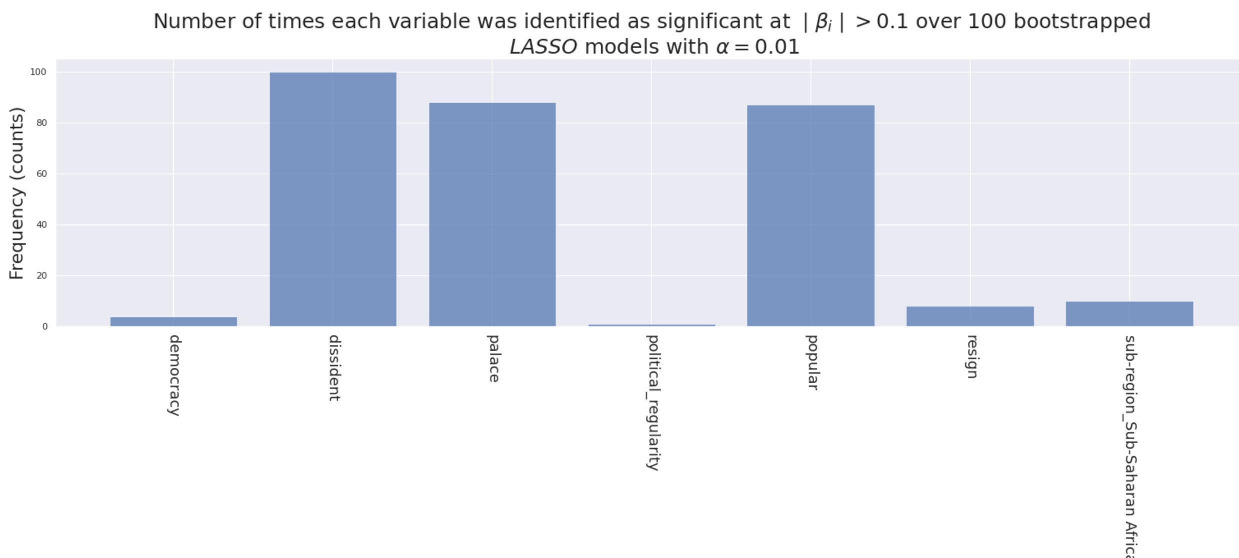

Figure 2. 2nd Round LASSO: Predictor Significance Frequencies

From the predictor significance frequencies, we can see that even among the predictor subset obtained from our first round not all predictors are consistently significant. To resolve this, among the predictors shown above we only included those that were found to be significant at least 20% of the time. Using this second subset of predictors, we decided to perform a 3rd round of LASSO regularization, this time with the inclusion of polynomial features to our design matrix. We chose to do so at this step because, while we knew from the outset that it would be valuable to examine the level of interaction between different predictors in our dataset, doing so on the entire predictor set would expand the number of columns in our design matrix to a very unmanageably large number. For similar reasons of avoiding overly high dimensionality, we chose to limit our polynomial features to a degree of 2.

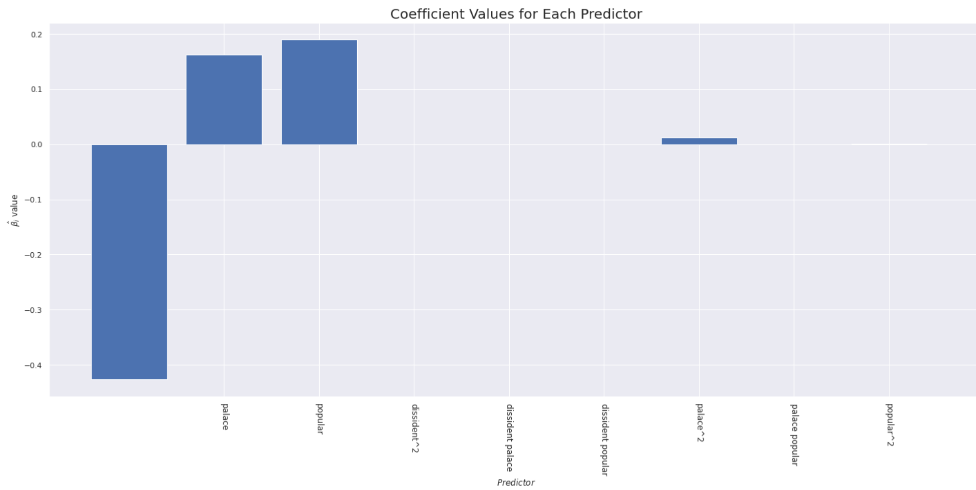

Figure 3. 3rd Round LASSO: Polynomial Features

The above described 3rd round of LASSO regularization led to the following plot of coefficient values, where we can see that there are indeed interaction terms with larger coefficient magnitudes than some of the lone predictors, indicating that these interactions are significant and worthwhile to include in our final models. Repeating the procedure used for the first 2 rounds, we took the columns found to be significant in the prior round and used bootstrapping to determine how frequently each column was found to be significant. Our final predictor subset will be the columns shown below with frequencies of 20% or more.

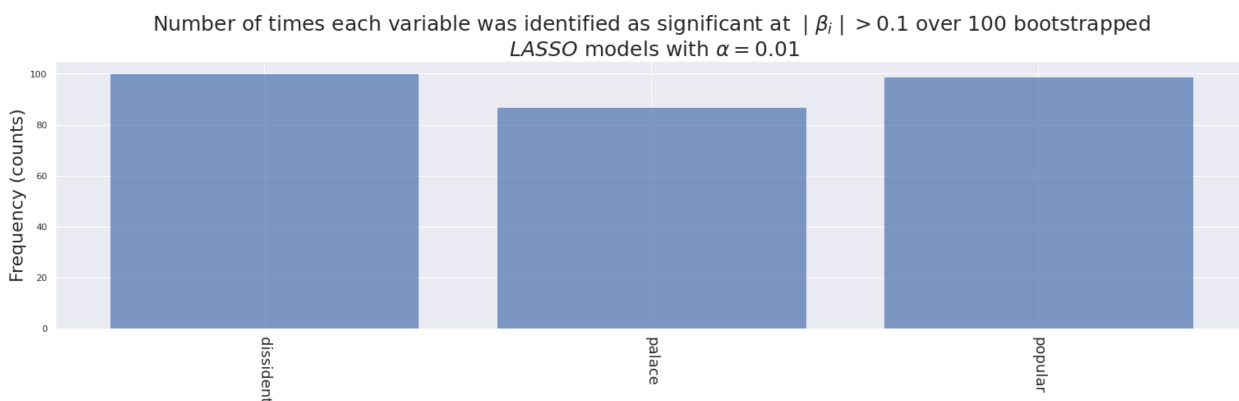

Figure 4. 4th Round LASSO: Significance Frequencies on Polynomial Features

As a quick way of determining whether the chosen predictor subset could accurately be used by a model to distinguish between realized and unrealized coups, we ran a PCA on the dataset including these predictors.

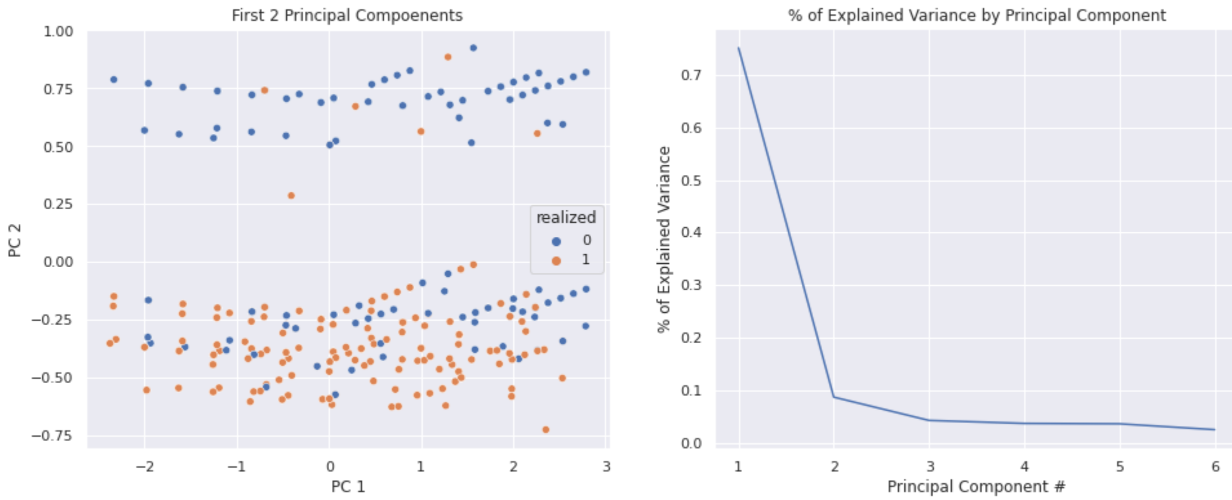

Figure 5. PCA on Final Predictor Subset

The fact that given our chosen predictor subset we were able to see clear separation between the two classes in PC space gives us confidence that the following predictor subset is well tuned for use in our later modeling work.

Our LASSO testing accuracy was 72% with the model trained on data without political indicators and 70.3% with the model trained on the data with political indicators.

II. Random Forest

To fit our classification algorithm, we first opted to use a random forest regression. This is because a random forest model decreases bias by overfitting to the training data. Then, to account for variance, it runs decision classifiers of a specified depth over a predetermined number of bootstraps of the dataset features. We took two approaches to tuning the hyper parameters. To tune the maximum depth, we fit a decision tree with a range of depths and recorded it testing and validation accuracy scores. Then we plotted these scores and determined the max depth. For the set of data not containing political information, we produced the following plot:

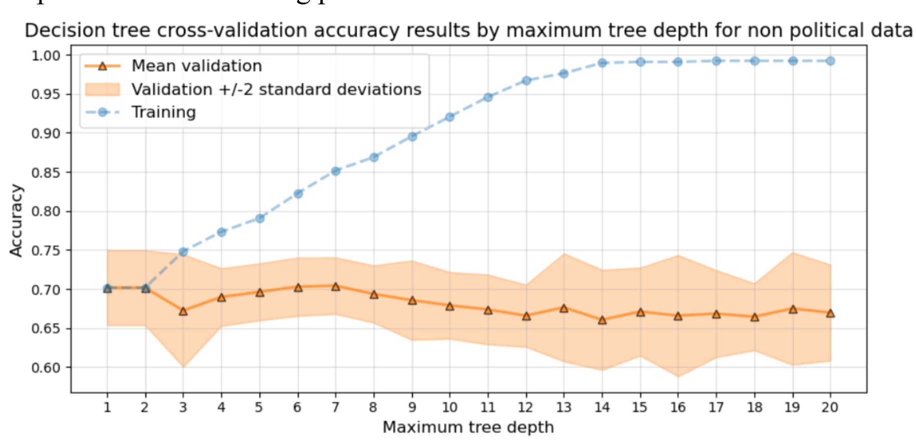

Figure 6. Optimal Max Depth (non-political data)

We selected a max depth of 16 because of the values that overfit to the train data (16-20), 16 has a relatively high validation score while still maintaining a similar train accuracy score. For the set of data not containing political information, we produced the following plot:

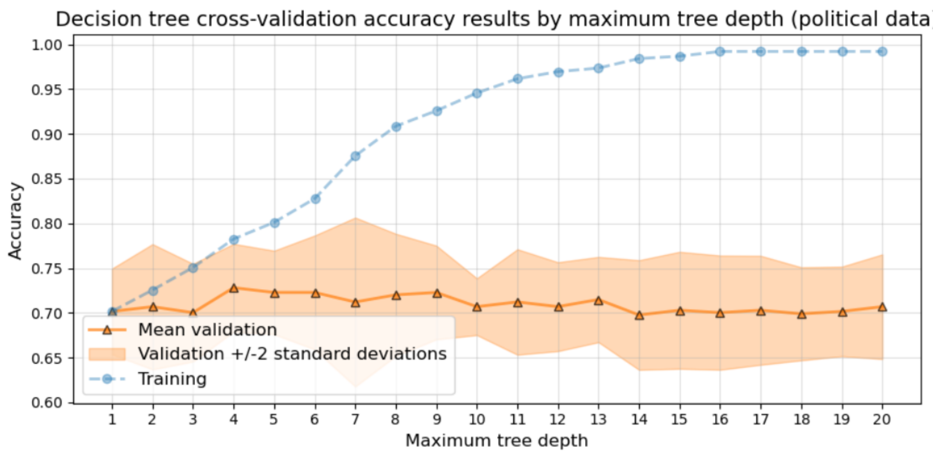

Figure 7. Optimal Max Depth (political data)

We selected a max depth of 20 because of the values that overfit to the train data (16-20), 20 has a relatively high validation score and the highest train score. For the number of bootstraps, we took a trial and error approach to find that 200 bootstraps was optimal, providing an insignificant increase in validation accuracy score.

Our Random Forest testing accuracy was 69.3% with the model trained on data without political indicators and 75.1% with the model trained on the data with political indicators. To determine variable importance, we used Permutation Importance, which we visualized using ELI5, which gave us the following:

Weight	Feature
0.1466 ± 0.0331	dissident
0.0212 ± 0.0125	popular
0.0090 ± 0.0164	palace
0.0074 ± 0.0052	auto
0.0063 ± 0.0132	pre_leader_court
0.0063 ± 0.0123	military
0.0053 ± 0.0047	South-eastern Asia
0.0037 ± 0.0083	pre_coup_fail
0.0032 ± 0.0070	pre_leader_harm
0.0011 ± 0.0114	Latin America and the Caribbean
0.0011 ± 0.0104	standing_regime
0.0011 ± 0.0042	rebel
0.0005 ± 0.0110	resign
0.0005 ± 0.0057	foreign
0.0005 ± 0.0032	Western Asia
0.0005 ± 0.0032	Eastern Asia
0.0000 ± 0.0047	Southern Europe
0 ± 0.0000	Central Asia
0 ± 0.0000	Northern America
0 ± 0.0000	Western Europe
	… 20 more …

Weight	Feature
0.1418 ± 0.0347	dissident
0.0339 ± 0.0307	years_of_regime_pow
0.0053 ± 0.0047	p5_missing
0.0053 ± 0.0047	auto
0.0026 ± 0.0085	exec_openness_score
0.0016 ± 0.0126	sex_ratio
0.0016 ± 0.0083	foreign
0.0011 ± 0.0182	political_regularity
0.0011 ± 0.0132	exec_competitive_score
0.0005 ± 0.0174	popular
0.0005 ± 0.0032	Northern Africa
0.0000 ± 0.0067	civwar
0.0000 ± 0.0047	rebel
0 ± 0.0000	South-eastern Asia
0 ± 0.0000	standing_regime
0 ± 0.0000	intviol
0 ± 0.0000	intwar
0 ± 0.0000	other
0 ± 0.0000	cntry_change
0 ± 0.0000	Northern America
… 44 more …	

Figure 6. Random Forest Feature Weights for limited (left) and full imputed (right) dataset

From the above figure, we can see that the dissident predictor, indicating that the coup was originated by a faction disagreeing with the current government, was the most heavily weighted predictor for both RandomForest and AdaBoost Classifiers, which is in agreement with what was found prior in the LASSO regressions. What is interesting to see is that in the Forest fitted on the full dataset, the weight drop-off past the first 2 most significant predictors is much more drastic than in the limited dataset.

Other variables that represent the actors of the coup are also significant, including popular, palace, military, and auto. The model trained on the non-political dataset finds more terms to be significant than the model trained on the political dataset, which results in the worse fit from the political data. From the political data, years for which the regime has been in power are also significant, indicating that the results from Gassebner et al. 2016 are supported by our results: previous coups, upheavals, and political violence are significant predictors.

III. AdaBoost

Finally, because boosting allows us to iteratively correct for the model's worst predictions, we implement AdaBoost based on a decision tree. For each of the two datasets, using our previous training/testing split, we tune hyperparameters. Using the training data, we randomly select 80% of the data to remain testing data, and 20% to become validation data. Then, we perform AdaBoost regression for multiple maximum tree depths, each for iterations from 0 to 800. This resulted in the following plot for the dataset with no political data:

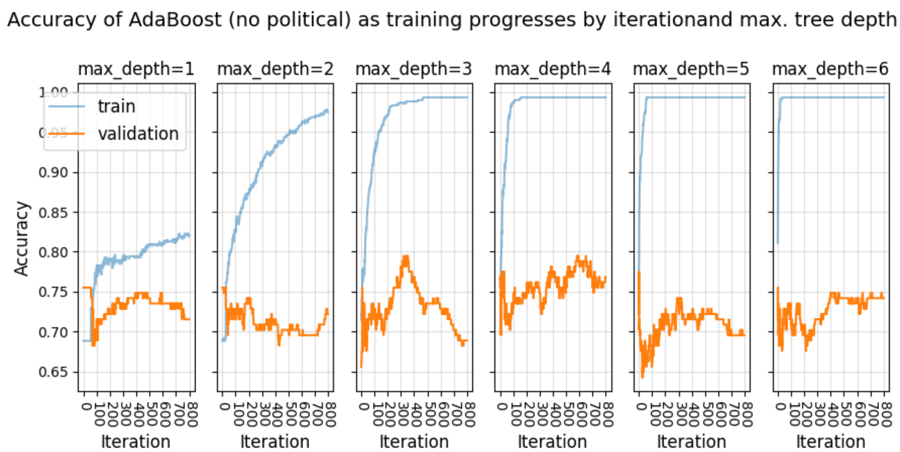

Figure 8: Not Political AdaBoost CV

For the dataset with no political data, we found that a maximum depth of 4 with 322 iterations produced the best validation accuracy.

Then, for the dataset with political data we found the following:

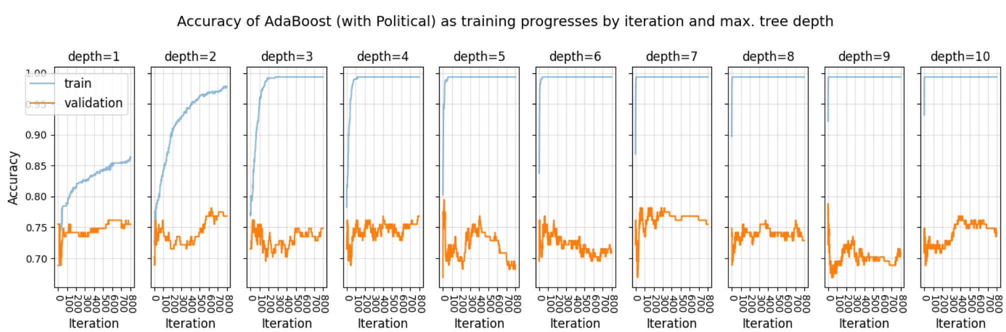

Figure 9: Political AdaBoost CV

From this, we chose a maximum depth of 5 with 15 iterations. Finally, we trained each of these models on their respective full training sets, and then evaluated their accuracy on the testing data.

We found that the model with no political data had a 65.6% accuracy and the model with political data had a 76.2% accuracy, the highest of all of our models. This indicates that fitting with political data improves model accuracy, as we have seen with our other regression approaches. Therefore, we select AdaBoost trained on the political data as our best regression approach. Then, we performed a Permutation Importance test and visualized the results using ELI5.

Weight	Feature
0.1238 ± 0.0432	dissident
0.0111 ± 0.0203	popular
0.0106 ± 0.0195	palace
0.0063 ± 0.0042	rebel
0.0011 ± 0.0104	resign
0.0011 ± 0.0079	pre_coup_succ
0.0005 ± 0.0057	cntry_change

Figure 10: ELI5 results from No-Political Data

Weight	Feature
0.1667 ± 0.0592	dissident
0.0344 ± 0.0405	years_of_regime_pow
0.0323 ± 0.0209	military
0.0175 ± 0.0095	pre_coup_succ
0.0153 ± 0.0198	popular
0.0138 ± 0.0190	standing_regime
0.0127 ± 0.0151	gini
0.0122 ± 0.0095	population
0.0085 ± 0.0233	democracy
0.0079 ± 0.0108	sex_ratio
0.0074 ± 0.0136	gdpgrwth
0.0063 ± 0.0092	palace
0.0058 ± 0.0032	civwar
0.0048 ± 0.0074	polity2_score
0.0048 ± 0.0180	exec_recruitment

Figure 11: ELI5 results from Political Data

Both of the models selected actors as the most important predictors. Dissidents being the actors was most significant for both, with other actors in the top several significant rankings (popular, palace, rebel, and resign for non-political data, and military, popular, and palace for the political data). For both of the training data sets, the Permutation Importance found that the number of previously successful coups was highly significant. Once again, from the political data, years for which the regime has been in power are also significant, indicating that the results from Gassebner et al. 2016 are supported by our results: previous coups, upheavals, and political violence are significant predictors.

E. Conclusions and Further Considerations

Out of all of our analysis techniques, we found that AdaBoost trained on data with political variables performed the best model prediction, with 76% accuracy. Through our multiple analysis techniques as detailed above, there are common features that appeared as being significantly influential in determining whether an attempted coup failed or succeeded. In particular, variables indicating the initiators in the coup are most significant (dissident, military, popular) both when the training set includes the political data and when it does not. From the perspective of the literature, there are several studies which point to military counterbalancing as a significant deterrent in coup activity because the initiators of the coup are particularly determinant of coup violence (Belkin, Shoffer 2003). The theory is that the actors have different resources (e.g., military actors can create a junta, dissidents can create a riot), so certain types of actors having more resources on average is a reasonable result. Although political violence and previous economic downturns was found to play a significant role in inciting coups in previous literature (Gassebner et.al 2016), our results suggest that it does not play a notable role in the success of a coup. In a sense, these might be factors that instigate a coup, but when the coup begins, the actors' resources become more influential.

In the future, this research could be continued by splitting the data by the type of actor causing the coup, and then performing similar analyses on those data. We suspect that, if this were to occur, the results would be most similar to the results of Gassebner et al. 2016.

In summary, we find that, although coup success is difficult to predict, our best model can predict with 76% accuracy. Unlike in predicting coup occurrence, coup success is best predicted by the actors who initiated the coup.

References

1. “List of Coups and Coup Attempts since 2010.” *Wikipedia*, Wikimedia Foundation, 8 Dec. 2022, https://en.wikipedia.org/wiki/List_of_coups_and_coup_attempts_since_2010

2. Gassebner, M., Gutmann, J. & Voigt, S. When to expect a coup d’état? An extreme bounds analysis of coup determinants. Public Choice 169, 293–313 (2016). https://doi.org/10.1007/s11127-016-0365-0

3. Peyton, Buddy, Joseph Bajjalieh, Dan Shalmon, Michael Martin, and Jonathan Bonaguro. 2021. “Cline Center Coup D’état Project Dataset Codebook”. Cline Center Coup D’état Project Dataset. Cline Center for Advanced Social Research. V.2.0.2. February 23. University of Illinois Urbana-Champaign. doi: 10.13012/B2IDB-9651987_V3

4. “Armed Conflict and Intervention (ACI) Datasets.” INSCR Data Page, Center for Systemic Peace, 2022, https://www.systemicpeace.org/inscrdata.html.

Appendix: