Predicting High-Risk Countries for Political Instability and Conflict

Blair Huffman, Emma Marriott, April Yu
Stanford University

Abstract

We present this paper in order to suggest a new tool for political science in respect to identifying unstable countries and their root indicators that could suggest areas for political intervention. With more than 98% accuracy, we were able to predict severely unstable political states based on countries with a Fragile States Index (FSI) score of > 100. In addition, we were able to determine the features that best indicate a country's instability through the use of filter feature selection.

1 Introduction

Everyday, political leaders around the world make decisions about international intervention. By being able to detect trends within a nation and respond with the right political, economic, and developmental sanctions, they can avoid the use of military intervention and prevent conflict or total governmental collapse. Our motivation is to be able to capture and interpret these trends on a grand scale and build a model that can indicate the fragility of a nation, as well as identify the crucial indicators that attribute to its instability. Currently, there are organizations that rank countries based on their political instability, but these systems use targeted quantitative and qualitative data, and are annual reviews rather than instantaneous predictions [1]. We hope that by applying data mining to a quantitative representation of country trends, we can create accurate, real-time predictions of a country's fragility, and eventually scale to include trending news, legislative updates, and other qualitative resources for more accurate inferences.

2 Dataset

2.1 Label - Fragile States Index (FSI)

We obtained our high-risk classifier data from the Fragile States Index (FSI) which is released annually from The Fund for World Peace[2]; each country is given a score based on Social, Economic, Political and Military indicators from a variety of quantitative and qualitative sources where a low score indicates high stability (i.e. Finland is 18) and a high score indicates low stability (i.e. Somalia is 113). In order to conduct tests with binary classification, we categorized an unstable state as that having an FSI score above 100. We chose this particular ranking system because it is a popular standard in political science and the product of two reputable organizations, *Foreign Policy* Magazine and The Fund For Peace[3].

2.2 Data - World Bank

We obtained all our data from the World Bank database, which contains over 1300 indicators for each country with an FSI score [3]. We have **1343** features for our **179** countries with FSI scores over 9 years (2006-2014) [4]. Our feature categories include:

- Environment
- Economic Policy and Debt
- Financial Sector
- Health
- Infrastructure
- Social Protection and Labor
- Poverty
- Private Sector and Trade
- Public Sector

2.3 Features and Preprocessing

Due to the nature of data collection in the social science field and the sheer size of our feature space, our data set was extremely sparse. Knowing this, we built and tested each model with an additional two variations on our data: 1) A version that corrected for the missing values by assuming linear growth in between existing values; 2) A version which contained the *change* in value from the previous year. Surprisingly, we found that the original, unmodified data set with missing values actually had the greatest success in failed state prediction.

3 Results

We used 4 algorithms to model our data: K-means, SVM, SMO, and SMO Regression. A summary of our results are below:

Table 1: Indicator Frequency (by number of years indicator appeared)

Algorithm **Testing Size**	**Training Error**	**Test Error**	**Training Error**	**Training Size**
SVM	0.6%	1.7%	1432	179
SMO	0.9%	2.8%	1432	179
SMO Regression	2.9*	9.3*%	1432	179

*Indicates root-mean squared (RMS) error as opposed to percent error

3.1 K-Means

We initially began our investigation using K-means due to its ability to separate data into k groups with the hope that K-means would be able to group based on stability using all features. We performed K-means with cluster sizes 2,3,4,5,10, and 20, and also used correlation thresholds to improve our clusters. Our best result (shown below) contained 4 clusters and used only the 10 highest correlated features.

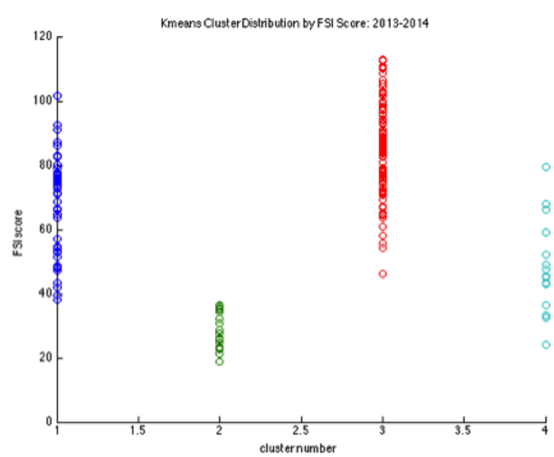

Figure 1: K-means with correlation threshold of 10 and cluster size 4

3.2 SVM

Next, we decided to use SVM because of its advantage in a high-dimensional feature space. In addition to obtaining training/test performance results, we created an ROC curve for our our positive classification (high-risk nations).

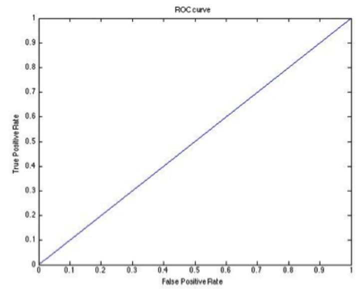

Figure 2: ROC curve for SVM algorithm

3.3 SMO

We then used a variant of SVM, SMO in order to see if we could improve the accuracy specifically for our positive class. With SMO, we also modeled the predictive power by plotting training/test errors over an accumulation of years.

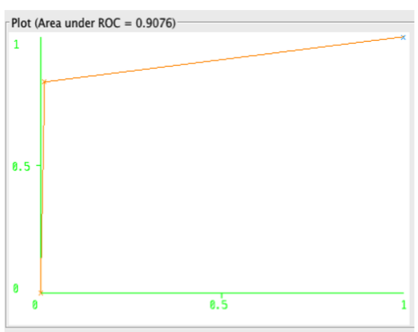

Figure 3: ROC curve for SMO algorithm. Horizontal axis is false positive rate; vertical axis is true positive rate.

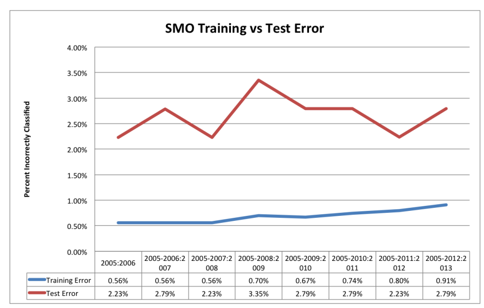

Figure 4: Training and Test error for accumulation of different years *(e.g. 2005-2012:2013 means model trained on the accumulation of data from 2005 to 2012 and tested on data from 2013).*

3.4 SMO Regression

Because of our success with SMO (as discussed later), we decided to try SMO Regression to model a non-binary prediction of the FSI scores. Below is the classifier error of our SMO Regression, with the X-axis as the actual FSI score and the Y-axis as the predicted FSI score. For this example, we trained on all data from 2005-2012 and tested on data from 2013.

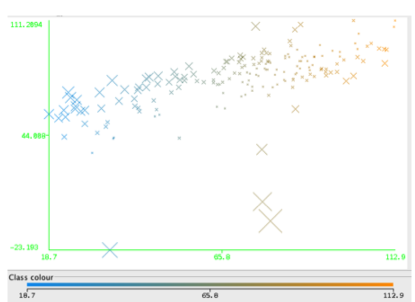

Figure 5: Visualization of classifier errors from SMO regression (larger error margins correspond to larger data points).

3.5 Feature Filter Selection

Because we have over 1300 features, we decided to run filter feature selection. We used a correlation-based feature selector with forward search in order to eliminate repetitive features on each year and determined those that have the strongest predictive power [5]. We display the results of the highest correlated indicators in the table below with their respective frequencies (number of years they were selected).

Table 2: Indicator Frequency (by number of years indicator appeared[1])

FREQUENCY	INDICATOR(S)
10	Under five mortality
8	Net bilateral aid flows from DAC donors[1], Neonatal mortality rate
	Refugee population by country of origin, Domestic credit to private sector
7	Time to import (in days), Female secondary education
6	Out of pocket health expenditure
5	Immunization of measles for infants, Public health expenditure
	Female primary education, Immunization of DPT for infants
4	Adjusted savings in education expenditure, Newborns protected against tetanus
	Urban population with access to improved water source
3	Public health expenditure, Maternal mortality ratio
	Tuberculosis case detection, Principal repayments on external debt (public)

4 Discussion

4.1 K-Means

We found that although K-means is effective in separating countries that are stable and sustainable (FSI < 40), it cannot make the necessary distinctions for countries that are classified as warning and above (FSI > 60). While countries with FSI < 40 cluster nicely together (as seen in figure 1), the other clusters contain countries whose FSI scores range from 60 to 120. This is expected, as many extremely stable nations are very similar in the indicators provided by the World Bank, while the more unstable countries differ from each other only in a few indicators. As we are primarily interested in classifying extremely unstable states as opposed to extremely stable ones, our K-means result is not very helpful. We clearly need an algorithm that puts more weight on the subset of indicators that are more commonly linked to unstable states.

4.2 SVM

SVM was relatively efficient in categorizing countries as either stable or highly unstable, with high testing accuracy ($> 97.8\%$) regardless of the size of the training set. However, there is one particular phenomenon that is troubling about the SVM algorithm, despite the high accuracy. As seen above, the ROC curve for SVM is linear. From this, we know that the SVM algorithm has a high false negative rate, in which some of the most unstable countries were labeled stable. In all trials, Nigeria and Syria were labeled stable despite having some of the most unstable (highest) FSI scores. This is particularly concerning because they represent potential threats that would be overlooked.

4.3 SMO

SMO also had a high test accuracy ($> 97.2\%$) and could reasonably classify countries as stable or unstable. Although the accuracy was not quite as good as SVM, the predictive power of instability was better for highly unstable states. Additionally, the ROC curve for SMO indicated that this algorithm had a lower false negative rate, which allows it to more accurately label the highly unstable states that we are targeting. With a comparable testing accuracy and a better ROC curve, SMO was a more appropriate algorithm to run with our data set than SVM.

[1]DAC donors are an sub-organization of the OECD aiding development programs. The currently have 29 members including US, Australia, Canada, UK, EU, etc.

4.4 SMO Regression

We looked further in to the 3 outliers (bottom 3 points with negative FSI scores) and found they corresponded to the US, India, and China (from right to left). Although the US would have been correctly labeled as a low-risk country, China and India may have biased indicators that account for the high error margin. As we know from filter feature selection, all of the most correlated indicators are social or economic. India and China are outliers in these sectors, having extremely high economic statistics and limited social indicator data.

4.5 Feature Filter Selection

The indicators selected by our feature filter selection could be used to suggest political actions that could improve a countrys stability. For example, public health expenditure, immunization, and mortality rates are shown to be significant indicators. It is possible that addressing these concerns directly through a focus in health programmes and related funding could have some impact in alleviating dissatisfaction leading to instability and conflict.

5 Conclusions

When utilizing SVM and SMO to simply recognize extremely unstable countries for a given year, we realized that while SVM had better accuracy, it also resulted in a high false negative rate. Given the application of the results, the categorization of an extremely unstable state as stable results in potentially dire consequences. As a result, we used the SMO algorithm, which resulted in similar accuracy and a lower false negative rate.

We then ran SMO regression to see if we could predict country stability with a higher granularity than binary classification. Through this process, we realized the bias our algorithms and data contain. Particularly, while social factors correlate most closely with a country's stability, many countries lack this data, which allows other factors to have greater influence in the labeling of that country. In particular, China and India lacked a lot of the social information needed to make an accurate prediction of their stability and thus the algorithm depended heavily on their economic statistics, which mirrors those of extremely stable countries.

Lastly, we decided to run filter feature selection to determine the features that were most predictive of a failed state. We discovered that health care related features and female education were strong indicators of a politically fragile country. This finding further confirmed the biases of our data set and algorithm, which we had begun to suspect when running SMO regression.

6 Future Directions

A major drawback with our data is that we are using the results from possibly semi-biased FSI scores in order to train and test our data. So if we had another 6 months, we would work on better predicting the scores based on more opinions than just the FSI. Another source of error is the sparsely filled data matrix from the World Bank. As shown from our SMO Regression, if we had social data for all countries along with the financial data, the data would appear less biased. In addition, we would try to extend our work to continually adjust our predictions of a countrys risk based on major events occurring in each country.

References

[1] Margolis, J. Eli. "Estimating State Instability." Studies in Intelligence 56.1 (2012). [2] Fund for World Peace. Fragile States Index [Online]. Available: http://ffp.statesindex.org

[3] World Population Day and 'Failing States'. Huffington Post [Online]. Available: http://www.huffingtonpost.com/robert-walker/world-population-day-and-_b_3568231.html

[4] World Bank. World Databank [Online]. Available: http://databank.worldbank.org/data/databases.aspx

[5] Hall, Mark A. Correlation-based feature selection for machine learning. Diss. The University of Waikato, 1999.

[6] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.