**Mass Mobilization Problem Statements:**

As academics who are conducting studies, how can we as data scientists help inform what data to collect and how to collect it, so that it is most useful for predictive power in the future?

* Relationships in the data
* Accuracy and specificity of data collection
* Challenges in the cleaning process and potential impact on model quality/accuracy/usefulness
* Success of a given protest – were protesters and their tactics ‘successful’ to some extent? Not a yes/no – need more nuanced options.

Audience: Mr. Clark, Mr. Regan and members of the Political Instability Task Force. We are here today to discuss the future developments and improvements of the Mass Mobilization project. As you know, the CIA originally funded this study to inform foreign policy and understand the impact of mass mobilizations in foreign states.

However, in light of the mass mobilization events that occurred in the US in 2020 and 2021, the US government now wants to better understand and prepare for protests on US soil. Although there are no US protests or demonstrations in the current data, the government recognizes the need to study US mass mobilizations in order to improve their preparedness for future events, and respond in ways that address protester concerns and encourage peaceful resolutions for all parties.

As data scientists, we have evaluated the data for its quality and usefulness in predicting state responses given a set of attributes about a protest. We have spent time cleaning and modeling the data, and have a number of recommendations for the next stage of the study.

David H. Clark (Binghamton University) and Patrick M. Regan (University of Notre Dame). The Mass Mobilization project is sponsored by the Political Instability Task Force (PITF). The PITF is funded by the Central Intelligence Agency. The views expressed herein are the Principal Investigators' alone and do not represent the views of the US Government.

**Important Metrics for Our Problem:**

Classification metrics to consider:

|  |  |  |  |
| --- | --- | --- | --- |
| Metric | How it’s calculated | What that means in English | When useful for our project |
| Precision | Correct Positive Preds (TP)  All Pos Preds (TP + FP) | Of all of our positive predictions, how many did we get right? | High for NV  Lower for V |
| Specificity | Correct Negative Preds (TN)  All Neg Preds (TN + FP) | Of all our negative predictions, how many were correct? | High for V  Lower for NV |
| Recall/  Sensitivity | Correct Positives Preds (TP)  All Actual Positives (TP + FN) | Of all actual positives, how many did we get right? | High for all, esp imbal class |
| ROC-AUC | No idea | How quickly does our model “front load” correct predictions? |  |

Violent Responses: We would rather predict more violent responses and there to be less violent response in reality, than predict too few. Optimize for FP. Minimize FN. Therefore, specificity which optimizes for the a negative, or non-violent response, is our strongest measure of success.

Non-Violent Response: We would rather predict too few and be wrong, than predict too many. Optimize for FN. Minimize FP. It’s ok to get less TN (which could be violent responses). Therefore, Precision which measures how many of our predictions we got accurate, is the best measure.