

## EPPS 6323 Project Proposal

Alden Felix, Wei-chen Huang, Jim Pan

February 22, 2023

### **Area/Topic**

As a team, we were inspired by Dr. Vito D’Orazio’s work at UTD, where he utilized automated machine learning to create a dynamic forecasting model to predict inter-state violence in Africa. After looking into his work, we decided to further generalize the idea and create a model to predict inter-state affinity: whether two states are allies, trade partners, neutral, have conflicting interests, or hostile towards one another. In order to do this, we will utilize political, social, and economic data.

### **Research Statement/Method**

This study intends to employ machine learning methods to predict affinity between states using social and economic variables. These variables will be dyadic, for example trade flow between a pair of states, and variables unique to a state, such as GDP. The dependent variable, the measure of a state’s affinity, could take two forms within this study. Using CAMEO codes and intensity scores found in coded event data, a continuous dependent variable could be derived that represents general affinity between two states. The more complex method would involve using a clustering method to find “affinity communities”, which would result in a categorical dependent variable (Pauls 2017).

Instead of limiting the study to a single machine learning method, it would prove insightful to test multiple methods for predictive power against a traditional OLS model, or a LASSO model in the case of the categorical dependent variable. Some of the possible models that could be employed alongside these are random forests, gradient boosted machines, and

neural networks. In this way the study can test machine learning methods with varying degrees of complexity.

**Data/Variable:**

This study will use data from various sources, including the World Bank, International Monetary Fund, and UTD Event Data which compiles coded event data from multiple sources including Phoenix\_RT, Phoenix NYT, Phoenix FBIS, Phoenix SWB, ICEWS: Integrated Crisis Early Warning System from Harvard Dataverse, and TERRIER. All sources were cited appropriately, and their references can be found in the bibliography section.

Our analysis considers a range of variables, including dyadic variables such as trade flow between a pair of states, and state-specific variables such as GDP, population, military expenditure, education statistics, gender, and other relevant economic and social factors. In addition, we will use CAMEO codes and intensity scores from the event data to calculate a measure of affinity between different countries.

## Works Cited

Pauls, Scott & Cranmer, Skyler. (2017). Affinity communities in United Nations voting: Implications for democracy, cooperation, and conflict. *Physica A: Statistical Mechanics and its Applications*. 484. 10.1016/j.physa.2017.04.177.