MACS 33002: Project Proposal

Ipek Cinar, Pete Cuppernull, Seth Harrison, Casey Mallon

1/24/2020

Puzzle and Research Question

In this project we seek to address the puzzle of state voting in the United Nations. More specifically, how do countries vote on human rights issues in the United Nations? We will first train our learner to sort through the existing roll-call voting records via text analysis and use this data to predict how countries will vote in the future. In doing so, we hope to devise a way to predict voting patterns in the United Nations that will prove analytically useful for policymakers and international advocates targeting United Nations officials.

The United Nations is the central node around which modern international cooperation takes place. In seeking to collaborate to formalize international norms into loosely upheld "international laws," states come together through the United Nations to vote on pivotal global issues such as climate change, norms of warfare, and increasingly establish and uphold norms of human rights. Human rights includes issues regarding the political, social, and economic equality and maintenance of rights for all persons regardless of group affiliation.[1] Hence, human rights is becoming more important to international debate as populations are increasingly mobile, bringing into question how the human rights of persons are to be guaranteed outside of their home countries, whether said persons are migrant workers, refugees, or asylum-seekers.ˆ[See Giordani, Paolo E. and Michele Ruta. 2012. "Coordination in immigration policy." *Journal of International Economics* 89: 55-67., Binder, Martin. 2015. "Paths to intervention: What explains the UN's selective response to humanitarian crises?" *Journal of Peace Research* 52(6): 712-726., Dowty, Alan and Gil Loescher. 1996. "Refugee Flows as Grounds for International Action." *International Security* 21(1): 43-71., Noll, Gregor. 2010. "Why Human Rights Fail to Protect Undocumented Migrants." *Journal of Migration Law* 12: 241-272., Bosniak, Linda S. 1991. "Human rights, state sovereignty and the protection of undocumented migrants under the International Migrant Workers Convention." *International Migration Review 25: 737-770.]*

Having this robust portrait of roll-call voting history is important because it clearly shows how states have voted on human rights issues based on if the motion passed or not. Instead of theorizing how a state voted based on their individual preferences, we are simply able to look at the results and focus on

---

[1]See Park, Han S. 1987. "Correlates of human rights: Global tendencies." *Human Rights Quarterly* 9: 405-413

how that action came about and how we can predict similar future actions. Furthermore, it is important to analyze roll-call voting and predict future voting patterns since this is the process through which human rights norms would be formalized into international law. For policymakers and norm entrepreneurs at all levels of the international system, it is important to be able to predict state voting patterns to preventatively advocate for the legalization of specific human rights.

Computation Approach

The computational approach to our project will heavily rely on natural language processing to predict the voting behavior of states. Our theory assumes that the text of the resolutions themselves – in conjunction with a handful of other pieces of available metadata on the proposals and external political and economic data - can explain voting behavior. We identify two main components to our computational approach: first, we need to build a learner which leverages an appropriate text processing technique that is both statistically insightful and computationally efficient. Second, we need to adopt a statistical approach which properly deals with time series data.

On the language processing side, our initial research suggests that we should adopt an N-gram model in processing the resolutions. While our experience here is somewhat limited, we expect to consult Rochelle Terman for guidance on how best to process the text for our project. It seems that it will be necessary for our learner to appropriately identify the topic(s) of the resolution, and then determine a general sentiment or stance on the topic – from here, a learner should be able to learn a state's preferences on given topics to then predict their voting behavior.

Furthermore, it will be necessary to adopt a computational approach to accommodate time series data. While we will be better equipped to select an appropriate method as we progress through the course material, our initial research suggests that we should employ a neural network or Gaussian process.[2] As for assessing the accuracy of the model, we hope to leverage a blocked cross-validation approach as suggested by Bergmeir and Benitez (2012) for time series data.[3]

[2] Ahmed, Nesreen K., Amir F. Atiya, Neamat El Gayar, and Hisham El-Shishiny. 2010. "An Empirical Comparison of Machine Learning Models for Time Series Forecasting." *Econometric Reviews* 29, no. 5(6): 594–621. https://doi.org/10.1080/07474938.2010.481556.

[3] Bergmeir, Christoph, and José M. Benítez. 2012. "On the Use of Cross-Validation for Time Series Predictor Evaluation." *Information Sciences* 191: 192–213. https://doi.org/10.1016/j.ins.2011.12.028.

Data Sources

For the dependent variable, the project will use United Nations General Assembly (UNGA) voting data.[4] The dataset was originally coded to derive ideal point estimates, which describe a country's political orientation towards the United States, and includes every roll-call vote held in the UNGA from 1946-2012. Coding for the dataset is based on existing compilations of UNGA roll call votes.[5] Each observation in the dataset is a country-roll call vote pair and is assigned one of five vote choice values: Yes, Abstain, No, Absent, or Not a member. (See Figure 1 for the visualization of how 6 countries voted on UN Resolutions over the course of 1946-2018.)

UNGA roll–call votes (1946–2018)
Broken down by country and vote choice (Source: UNGA Voting Data)
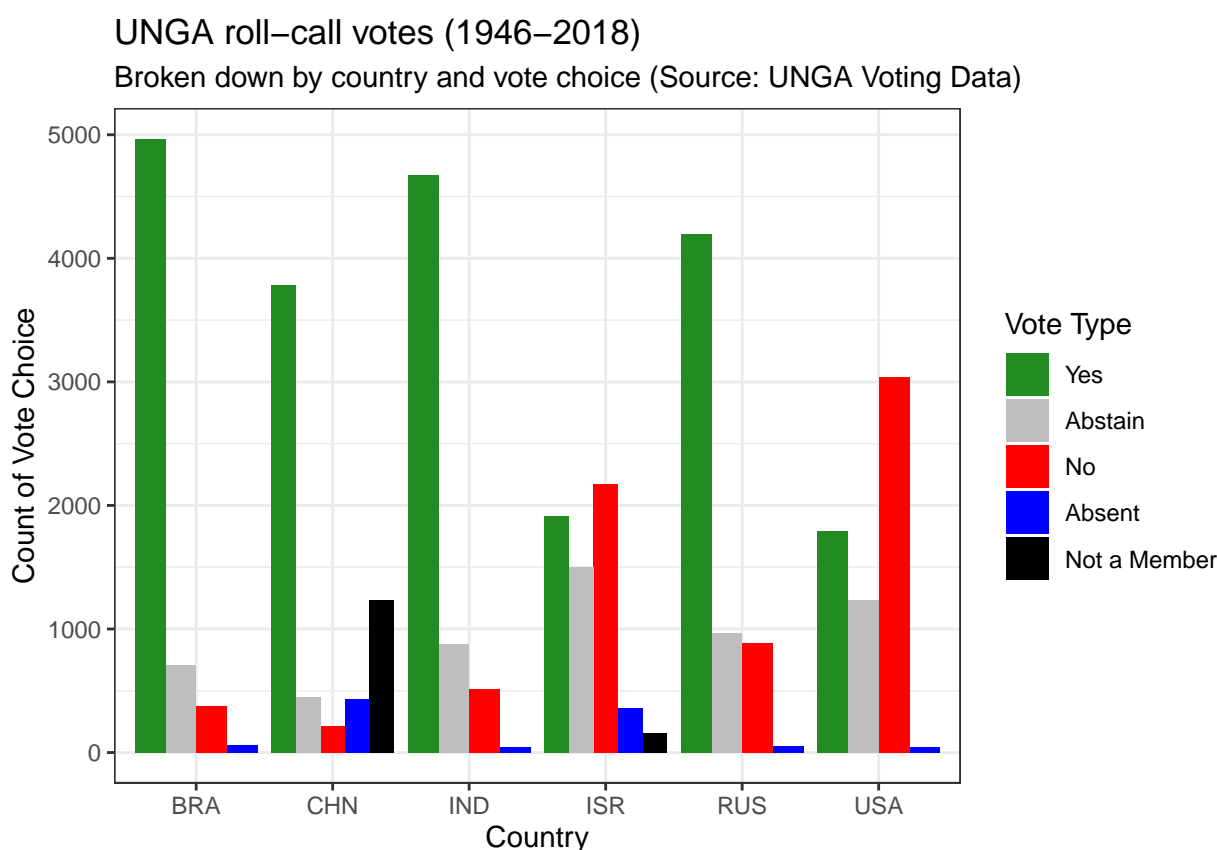
In addition to coding vote choice, the data is coded topically, allowing for easy identification of

[4]Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating dynamic state preferences from United Nations voting data." *Journal of Conflict Resolution* 61(2): 430-456.

[5]Sources identified in the codebook include: Inter-university Consortium for Political and Social Research (ICPSR). *United Nations Roll Call Data, 1946-1985* [Computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1982.; Gartzke, Erik and Dong-Joon Jo. *UN General Assembly Voting* V3.0 January 2002. http://dss.ucsd.edu/~egartzke/;Jo, Dong-Joon. Dundas Dataset. Used in: Jo, Dong-Joon. 2000. "Power Resources and Influence at the UN General Assembly." *Presented at the Annual Meeting of the Northeastern Political Science Association.*; Schopen, Lynn; Newcombe, Hanna; Young, Chris; Wert, James, *Nations on Record: United Nations General Assembly Roll-Call Votes (1946-1973).* Oakville-Dundas, ON: Canadian Peace Research Institute, 1975. (and subsequent supplements).;Kim, Soo Yeon; Russett, Bruce, "The new politics of voting alignments in the United Nations General Assembly." *International Organization.* Aut 1996, 50, (4), 629 - 652.;Various UN Resources including the *Official Records to the Proceedings of the United Nations General Assembly, UNBISNET:* http://unbisnet.un.org/, and UN documentation on-line: http://www.un.org/documents/resga.htm.

human rights-related votes. (Please see Figure 2) Percentage agreement with six countries (United States, Russia, Brazil, China, India, and Israel) is also recorded, along with descriptions of the vote item.

UN Resolution Topics
Broken down by percentage (Source: UNGA Voting Data Codebook)
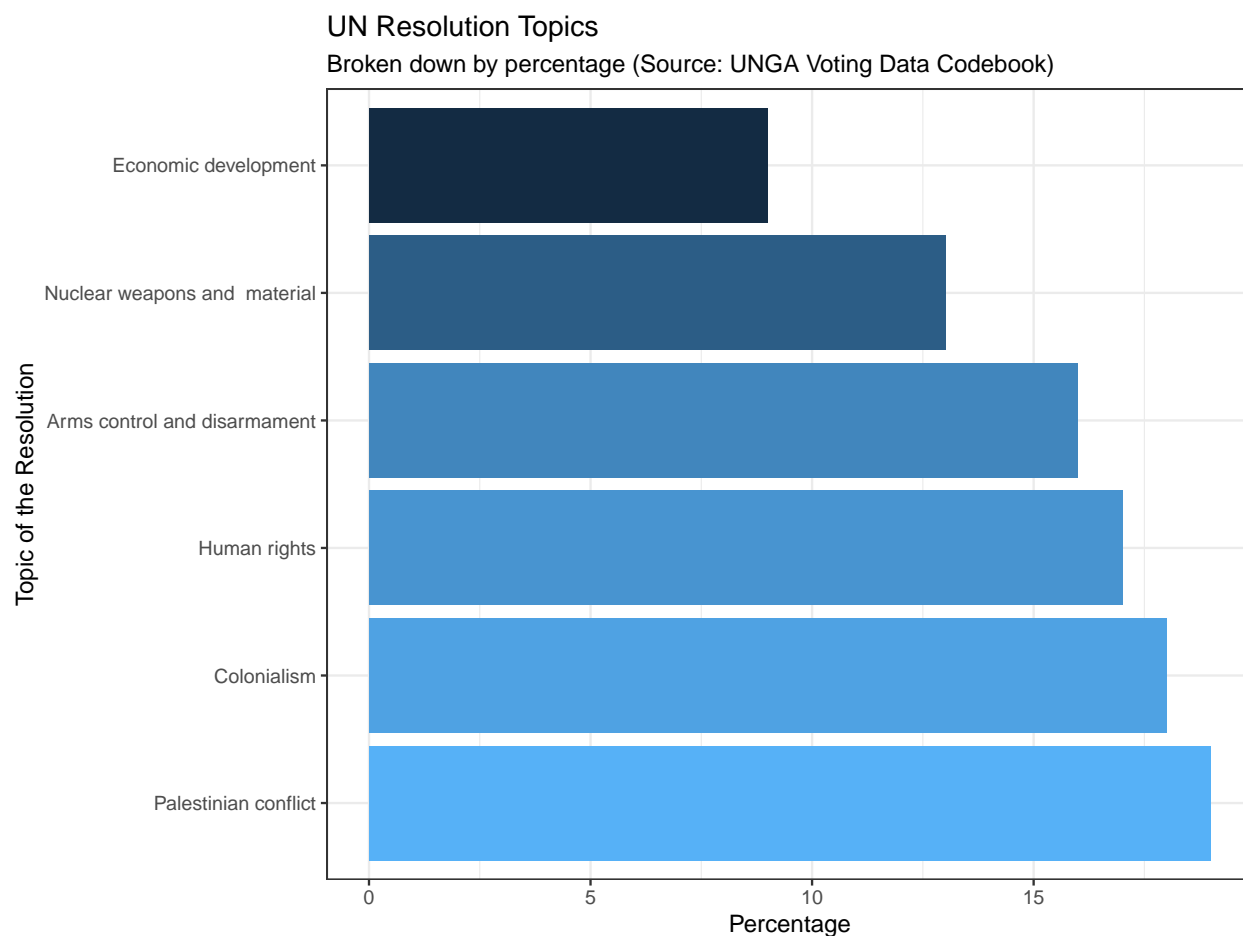
For the text data, the project will use an existing corpus of 2,100 UNGA resolutions.[6] The corpus, published in 2009, contains approximately 3 million tokens for each of the UN's six official languages– though we anticipate to input only the English-language tokens to the learner. The corpus will be obtained from the OPUS project, an open source corpus distribution project.

---

[6]Rafalovitch, Alexandre and Robert Dale. 2009. United Nations General Assembly Resolutions: A Six-Language Parallel Corpus. In Proceedings of the MT Summit XII, pages 292-299, Ottawa, Canada, August