DEPARTMENT OF THE ARMY
UNITED STATES MILITARY ACADEMY
Department of Mathematical Sciences
West Point, New York 10996

MADN-MATH 17 JAN 2023

MEMORANDUM THRU MAJ Cruickshank, MA491 Project Advisor, Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996-5000

FOR LTC James Bluman, MA491 Course Director, Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996-5000

SUBJECT: MA491 Project Proposal (Term 23-2)

1. **Title**: Applying Natural Language Understanding to partisan bias in cable news.

2. **Personnel**: CDT Benson, Applied Statistics and Data Science; MAJ Cruickshank, Army Cyber Institute.

3. **Background**. In recent times, much has been written about the trend towards political polarization in the United Staes. This phenomenon describes the increasing prevalence of partisan divides in both the national and local levels of our government. Currently, the literature has heavily documented voting trends that point towards partisan polarization as well as congressional behavior and media consumption patterns. However, one limitation to the current literature is a lack of mathematical-based computational analysis of statements made by political actors. Additionally, while extensive mathematical research has been done discerning differing sentiments towards issues, the application of cross-subject bias models has been limited.

4. **Research/Study Objective**: The goal of the project is to develop Natural Language Understanding techniques to characterize the partisan bias in political statements. These techniques will be applied to a collection of statements made in cable news transcripts to understand the bias of different outlets, speakers, and guests as well as how this has changed over time. Each individual statement made in the extensive collection of cable news transcripts obtained for the project will be separated, with stored metadata on the person who made the statement as well as information about the program they appeared on. From there, statements will go through a pre-processing pipeline and topic modeling in order to become usable data sorted by issue. Finally, writing style bias will be determined through a model examining the word choice and framing of the statements.

MADN-MATH
SUBJECT: MA491 Project Proposal (Term 23-2)

5. Project Timeline & Deliverables:

Timing	Activity and Deliverable	Hours
0-3 weeks	Literature review and data cleaning.	30
3-6 weeks	Initial findings, MORS deadline on February 10	20
6-10 weeks	Model development and mid-semester update brief.	20
10-12 weeks	Model application.	20
12-15 weeks	Sensitivity analysis and further application.	20
Final	Report and presentation.	15
	TOTAL	125

6. **Resources**: This project requires the following: Python (with installed packages), news transcript dataset (already obtained).

a. This topic requires a review or exemption by the Human Research Protection Program: No. A brief description of what the HRP/IRB considerations are.
b. This topic requires data from another organization: No. If yes, describe the plan to obtain a Data Sharing Agreement and state whether the data is on hand or not. State the classification level of the data (data that is classified above CUI is generally not recommended for work at USMA).
c. This topic involves the collection or use of cadet data: No. If yes, describe the plan to obtain approval for the project from the IRB.

7. The point of contact is the undersigned at seth.benson@westpoint.edu.

BENSON.SETH.PHILIP.1253605495
Digitally signed by BENSON.SETH.PHILIP.1253605495
Date: 2023.01.19 10:47:32 -05'00'

SETH P. BENSON
USMA 2023, Company H4
Applied Statistics and Data Science

IAIN CRUICKSHANK
MAJ, MI/49
Research Scientist