Sentiment Analysis Project Proposal

Adam Lehavi
Machine Learning Intern
alehavi@usc.edu

CONTENTS

I. Executive Summary

As Israel has become a larger topic in campus conversations, many college students have felt a fear in expressing their views, and even their Jewish identity. In a survey by Brandeis University, 65% of $AE\Pi$ and $AE\Phi$ members have reported not feeling safe as a Jew on their campus [1]. However, the largest study of the correlation between antisemitism and antizionism only encompasses a 400 student survey on a single college campus [2]. Papers have also set out to measure antisemitism, but focus on broadly tying it to hate on anonymous message boards [3]. Overall, there is a lack of both data and models for measuring and predicting sentiment related to this issue.

The sentiment analysis project would aim to bridge this gap. The project would focus on measuring **antisemitism**, **antizionism**, **hate speech and crimes**, **Jewish student mental health**, and **Jewish student comfort of expression** in regards to both establishing current records and setting up a system to capture and compare trends over time. The project would also cross-reference these scores to attacks, protests, or articles on campus and in surrounding areas to see if there exists correlation between any metrics.

The goal of these measurements would be to inform and make informed conclusions, and the results could be applied to a variety of applications. If the trend established by [2] can be tied to more universities, of antisemitism being correlated to antizionism, action can be taken to change campus curriculum and policies to make a more welcoming environment for Jewish students by means of reducing anti-Israel sentiment in classes, clubs, and discussion. The results of surveys or data scraped could be used to generate more accurate rankings of safest or best colleges for Jewish students to attend. The surplus of data would also allow for the creation of digestible yet powerful info-graphics. Tying a campus's ranking to tangible messages of hate or examples of students feeling out of place would aid in conveying a message or generating responses in media. Furthermore, trends of Jewish student mental health, participation, and expression would allow for Olami as an organization to make more informed decisions about how to best market and connect to campus resources.

This initiative would aim to produce multiple final deliverables. A database of responses, sentiment, or scraped information would be made, centralized, and processed. Machine learning models would be trained, tested, and presented to find accurate rankings, determine the sentiment of text inputs, and predict future trends or attacks based on sentiment and other external measurements over time. Trends found would either be turned into informative graphics themselves or sent to other groups within Olami to neatly process into a digestible format. Lastly, if findings are significant, results would be turned into a presentable paper, either for the sake of a conference or for suggesting changes to colleges, cities, or organizations.

II. Project Description

Background

The debate over the tie between antisemitism and antizionism is one that has persisted for a fair amount of time. However, no study has gone about truly defining this correlation particularly as it applies to multiple localities, whether it be cities or campuses. Beyond this, other metrics related to these could answer other common questions. The tie between other common forms of hate and antisemitism and the tie between antisemitism and antizionism could lead to interesting proposals about the relation between progressive ideologies on campuses and how they may lead to fringe discrimination of Jews because of stereotypes related to control of media and wealth. Metrics related to mental health and student participation, while partially separated, are also important to measure if there is an effect of antisemitism or antizionism on student participation and expression of Judaism.

Objective

The main objectives of the project would be lined up as such, in rough sequential order:

1) Survey common or past methods for measuring hate, antisemitism, and antizionism with a focus on model explainability, accuracy, and training distribution.
2) Survey datasets of both Jewish student metrics and campus messages related or unrelated to hate.
3) Establish a novel approach to gathering, processing, and storing data that best benefits removing bias and obtaining high accuracy results, or communicating with the authors of valid data to see feasibility of reproducing or obtaining the data.
4) Create or gather high-accuracy explainable models for measuring sentiment.
5) Create models for seeking trends between sentiments and attacks, whether simple statistical methods or explainable AI models.
6) Look at scoring metrics for best schools for Jewish or Zionist students, compare and examine scoring metrics, and aim to generate more reflective rankings.
7) Based on correlation between sentiments and attacks generate trends to be used for the generation of informative graphics and action in policy changes.
8) Compile models and data into a clean and organized code repository, and compile information learned into presentations or papers.

III. Project Work Plan

The work plan, timeline, and deliverables are to be discussed based upon the size of team allocated and the level of proficiency of the team.

The plan of weekly action would be as follows:

- See which objective as outlined in the project description is next in line.
- Individual workers would make progress on advancing some portion of the goal, independent of others. The main benefit of measure multiple sentiments is that it would give way for many people to work on advancing individual portions at once.
- The team would then have short weekly meeting to re-divide work an analyze errors, along with potential monthly meetings to present progress.
- Some member would be in charge of managing the status and workload of other members, relative to their skill and availability.

REFERENCES

[1] L. Prior. 1st poll of 'openly jewish' college students reveals 65% felt unsafe, 50% hid jewish identity. [Online]. Available: https://brandeiscenter.com/1st-poll-of-openly-jewish-college-students-reveals-65-felt-unsafe-50-hid-jewish-identity/

[2] R. Shenhav-Goldberg and J. S. Kopstein, "Antisemitism on a california campus: Perceptions and views among students," *Contemporary Jewry*, vol. 40, no. 2, pp. 237–258, 2020.

[3] J. Finkelstein, S. Zannettou, B. Bradlyn, and J. Blackburn, "A quantitative approach to understanding online antisemitism," *arXiv preprint arXiv:1809.01644*, 2018.