Profiling Social Media Users

Hacettepe University, Computer Science Engineering & AI Engineering
İhsan Çağatay ERASLAN, Umut KALMAN, Adem Baran ORHAN
Supervisor: Asst. Prof. Fuat Akal

Try our Model!

TÜBİTAK
Funded by 2209-A

Introduction

Our project, "Profiling Social Media Users," leverages **natural language processing** techniques to predict demographic traits such as age, gender, sexual orientation, race, and education level, thereby enhancing the analysis of public opinion and societal trends on social media platforms.

Objective

The objective is to **fine-tune** a robust **transformer** model that accurately predicts profiles by analyzing Twitter data. This will enable more precise insights into public opinion and improve the ability of researchers and organizations to understand and engage with diverse audiences effectively.

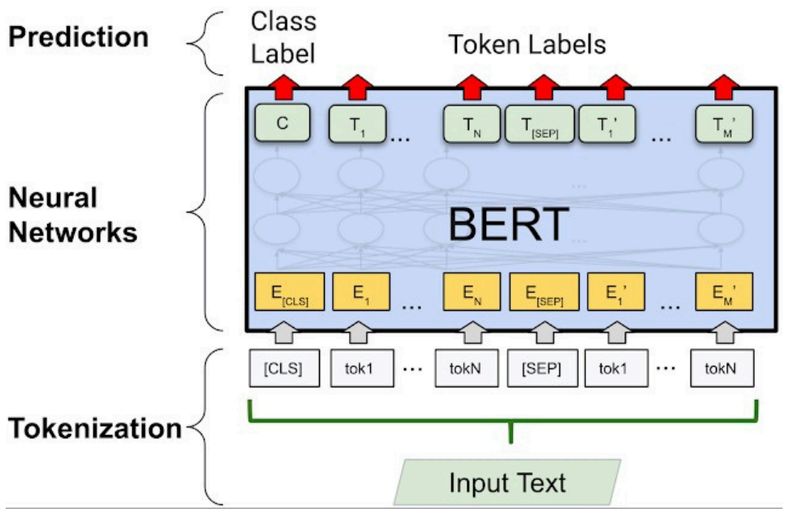
Methodology

Data Acquisition:
- Collected tweets and bio information using **Apify Twitter Scraper.**

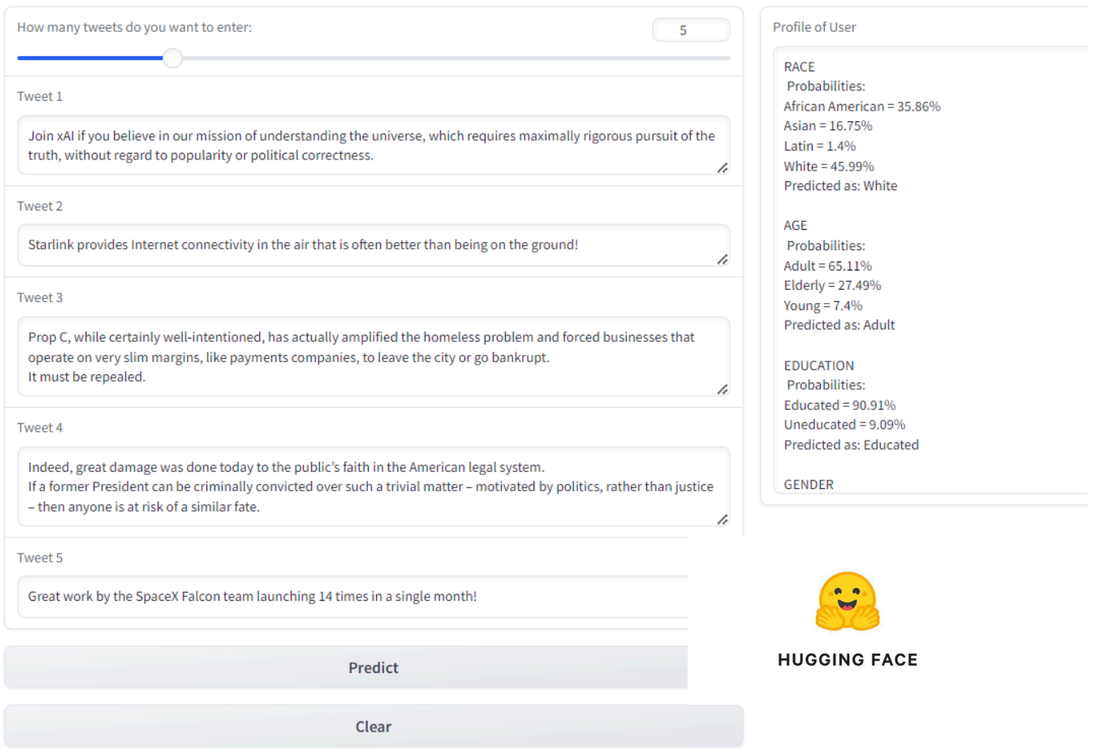
Data Preprocessing:
- Analyzed hashtags and emojis, filtered tweets, and combined data into a labeled DataFrame.

Model Development:
- Utilized BERTweet for tokenization and classification.
- Split data into training, validation, and test sets.

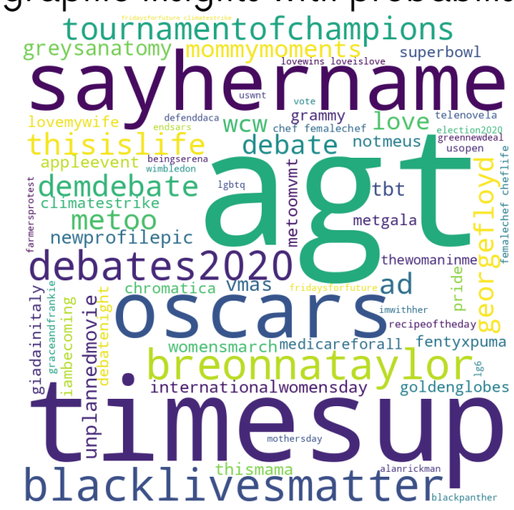
Training and Evaluation:
- Trained with cross-entropy loss and Adam optimizer.
- Applied early stopping, evaluated performance, and predicted on unseen data.



Model Deployment and Insights



We deployed our model on Hugging Face Space and gave to model latest five tweets from **Elon Musk**. The model successfully predicted his demographic profile accurately. You can try it yourself using the QR code on our poster—simply input tweets from a different user, and the model will provide key demographic insights with probabilistic scores.



Example Word Cloud for Female Class

Acknowledgements

- This research is supported under the TÜBİTAK 2209-A University Student Research Projects Support Program.
- This project was completed within the context of BBM479-480 Design Project courses in Hacettepe University, Faculty of Engineering, Department of Computer Engineering.

Application Areas

- Academic Studies: Supporting sociological and psychological research by providing demographic data analysis.
- Public Opinion Analysis: Understanding the demographic distribution of viewpoints on social media.