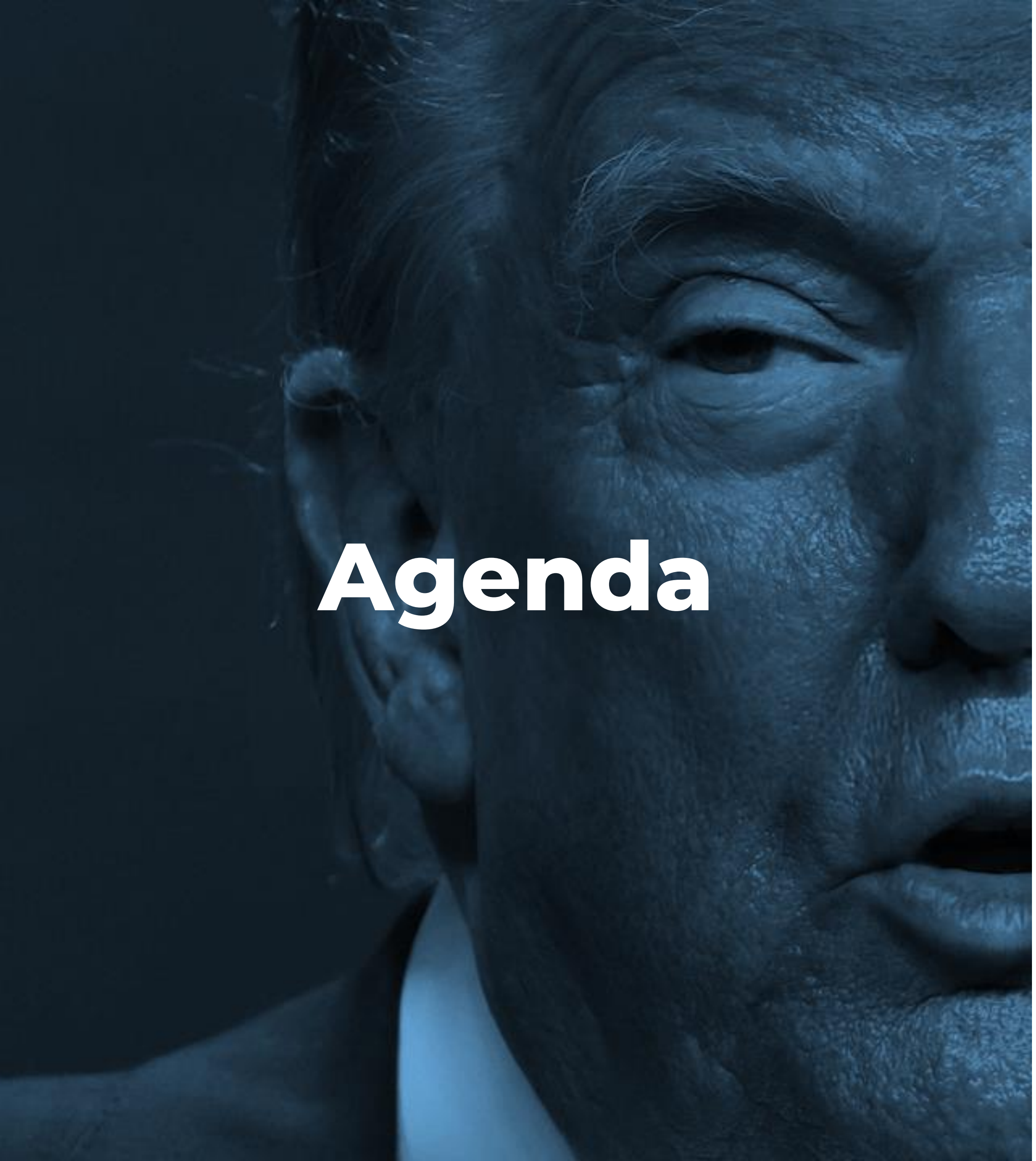
Trump on Twitter: Sentiment Before and After Taking Office



Agenda

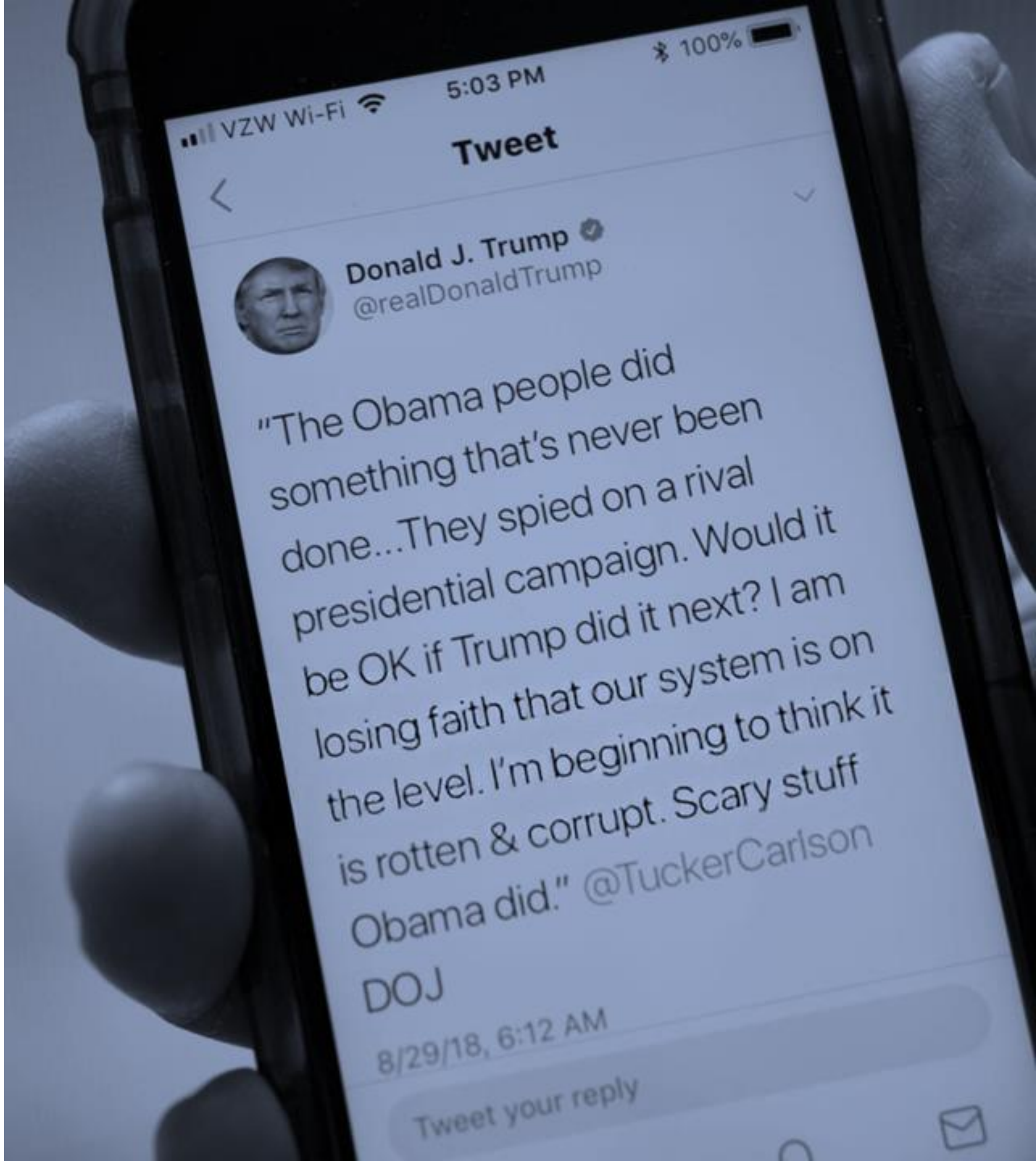
1. Project Overview
2. Methodology
3. Experimental Results
4. Final Remarks

PROJECT OVERVIEW

The core objectives involve constructing a pipeline using **transformer-based models** (e.g., RoBERTa or BERT fine-tuned on emotion datasets) to **classify emotional tones** in texts like speeches or social media posts, analyzing variations across **political groups** or temporal contexts, and generating comparative visualizations to highlight trends.

DEVELOPED WORK

Focusing on **Donald Trump's tweets,** this study applied **sentiment analysis** to track shifts in the emotional tone of his communication **over time**.



Research Question

Is there a **significant difference** in the emotional tone of Donald Trump's tweets **before** and **after** becoming **president**?

November 8, 2016

Donald Trump won the U.S. presidential election, defeating Hillary Clinton to become the 45th President of the United States.

Methodology

DATASET DESCRIPTION

Data overview and EDA

MODEL DEVELOPMENT

Discussion of the modeling pipeline and performance evaluation

KEY FINDINGS

Analysis and insights from emotion classification applied to our dataset to answer the research question

Dataset Description

The dataset used for the analysis comprises 56,571 tweets authored by **Donald Trump** between **2009** and **2021**. During preprocessing, tweets containing the token *[url]* were removed from the dataset. These entries were considered sentimentally uninformative, as the token represents a hyperlink whose content is inaccessible and thus provides no usable context for sentiment analysis. This step reduced the dataset from 58,011 to **40,694 entries**, contributing to a cleaner and more meaningful corpus.

ID

Unique numeric identifier assigned to each tweet.

TEXT

Full textual content of the tweet.

IS_RETWEET

Boolean indicating whether the **tweet is a retweet** (TRUE or FALSE).

IS_DELETED

Boolean indicating whether the **tweet was deleted**.

DEVICE

String specifying the **device or platform used to post the tweet**.

FAVORITES

Integer count of **likes** the tweet received

RETWEETS

Integer count of **retweets** the tweet received.

DATETIME

String representation of the **exact posting date and time**.

IS_FLAGGED

Boolean indicating if the **tweet was flagged.**

DATE

Date-only timestamp in datetime format (**YYYY-MM-DD**).

Exploratory Data Analysis

TWITTER ACTIVITY OVER TIME

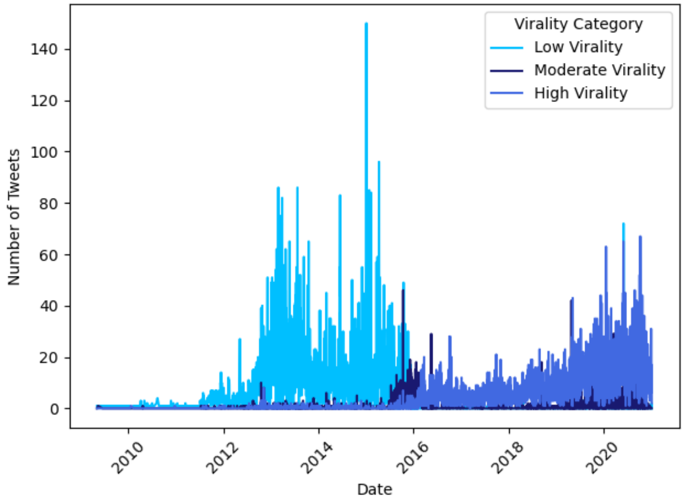
Tweet volume **peaked** during the **2016** election, then **declined** but stayed stable **until 2020**.

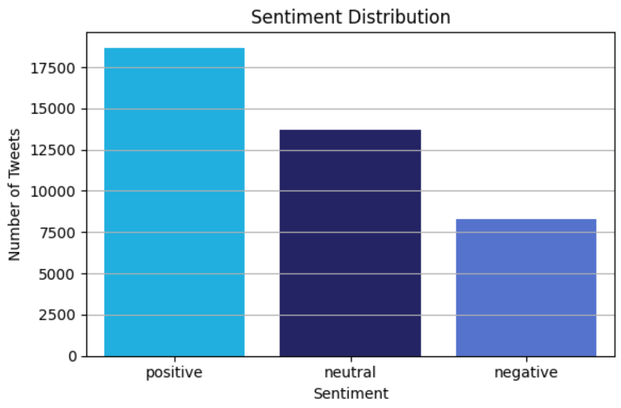
INCREASED VIRALITY

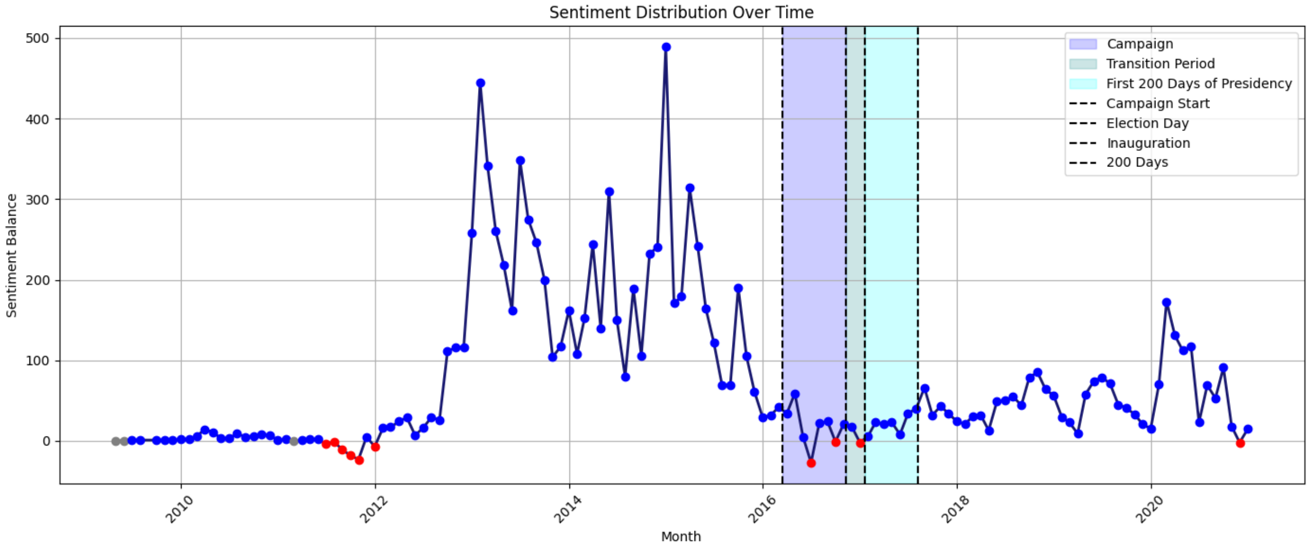
Presidency increased tweet virality, suggesting **higher impact** despite lower frequency.

SENTIMENT DISTRIBUTION

Most tweets were **positive**, but sentiment **shifted** toward neutral/negative during the 2016 campaign.







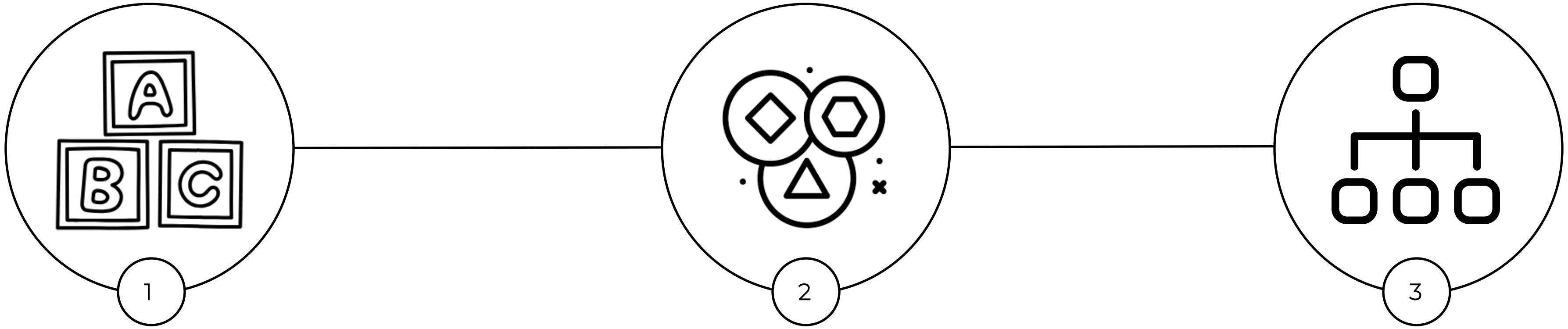
MODEL DEVELOPMENT

The project implements **a multi-stage emotion classification** pipeline designed to capture both broad sentiment and nuanced emotional content in political texts. Leveraging the **GoEmotions** dataset—comprising 58,000 Reddit comments annotated with 27 emotion categories plus a neutral label. .

Four distinct models were trained to support the classification strategy. The **first model** performs sentiment-level classification, distinguishing tweets as **positive, negative, or neutral**. To address class imbalance, this model was trained in two phases: one on the full dataset and a **second** using **undersampling**. The **third model** targets full emotion classification across **all 28 categories** in GoEmotions. The **fourth model** focuses on fine-grained emotional detection **while excluding the neutral class**, offering a more targeted understanding of emotional nuance.

01 **MODEL 1** → 02 **MODEL 2** → 03 **MODEL 3** → 04 **MODEL 4**

Modeling Pipeline



INPUT TEXT CLEANING

Preprocessing included removing special characters, expanding contractions, replacing URLs with “[url]”, and eliminating numeric-only text. Extra whitespace was trimmed, and all text was converted to lowercase for consistency.

TOKENIZATION

Cleaned text was tokenized using a RoBERTa tokenizer, converting input into subword tokens compatible with transformer models. This ensured structured, model-ready input.

TRAIN/VAL/TEST SPLIT

The data was divided into training (81%), validation (9%), and test (10%) sets. The test set was reserved for final model evaluation to ensure unbiased performance assessment.

Performance Evaluation

	ACCURACY	PRECISION	RECALL	F1 SCORE
1	64%	64%	64%	0.64
2	66%	67%	66%	0.66
3	43%	41%	43%	0.41
4	44%	43%	44%	0.43

KEY FINDINGS

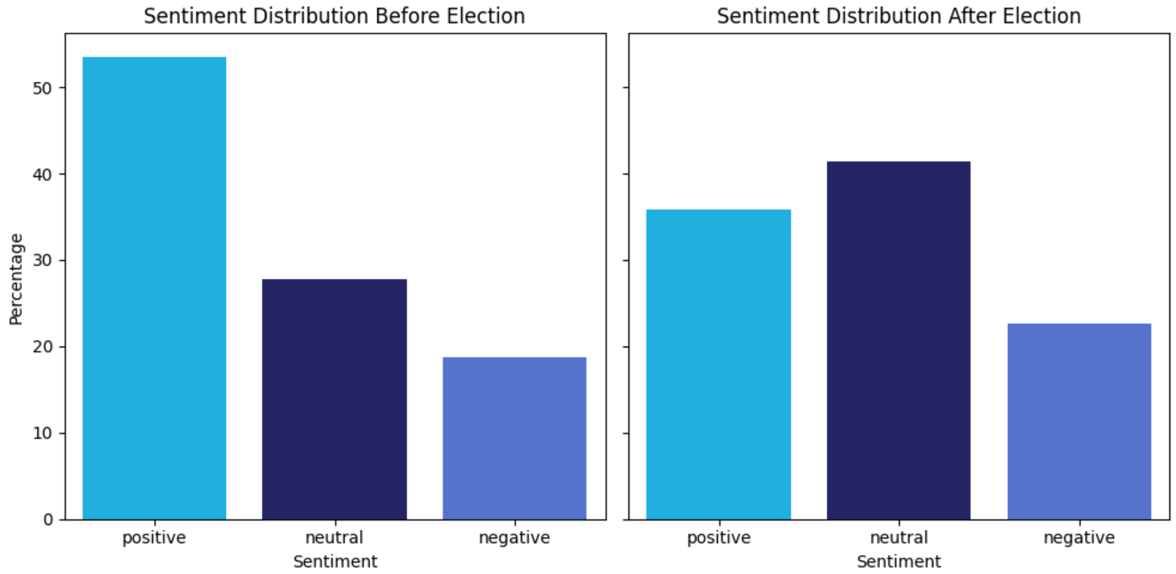
Coarser Classification: Before and After November 8, 2016

CHI-SQUARED TEST

CHI-SQUARED STATISTIC	DEGREES OF FREDOOM	P-VALUE
1313.63	2	5.62×10^{-286}

CHANGE IN DISTRIBUTION

POSITIVE	NEUTRAL	NEGATIVE
53.56% → **35.90%**	27.71% → **41.43%**	18.73% → **22.68%**



Beyond Neutrality

SENTIMENT ANALYSIS CHALLENGES

- Rise in neutral tweets complicates **emotional interpretation**.
- Neutral tone lacks **clear polarity** for sentiment classification

MISLEADING NEUTRALITY

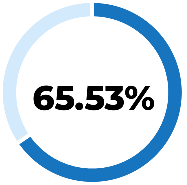
- Frequent positive or neutral words often used **sarcastically** or **politically**

Neutral Sentiment Breakdown

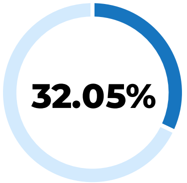
UNMASKING NEURAL TWEETS

- **Admiration** emerges as the dominant emotion among these reclassified tweets, accounting for 28%
- **Approval** follows closely at 20.75%
- Negative and inquisitive tones are also present, with 12.71% of tweets expressing **annoyance** and 6.89% showing **curiosity**.

POSITIVE TWEETS

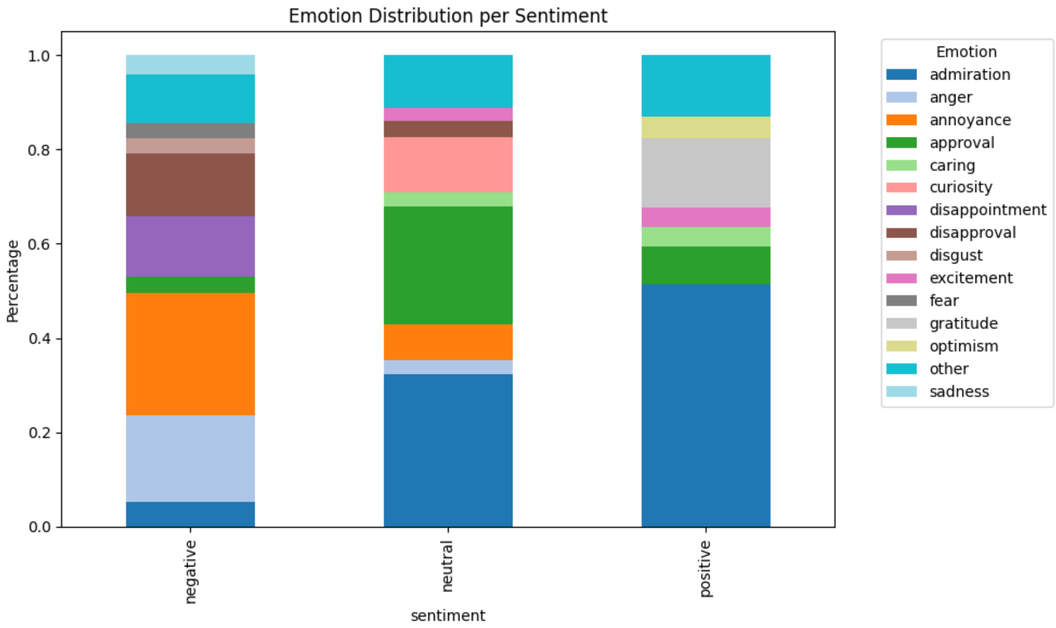


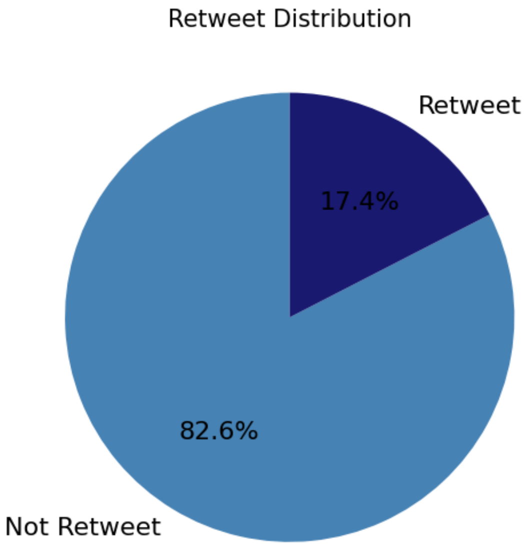
NEGATIVE TWEETS

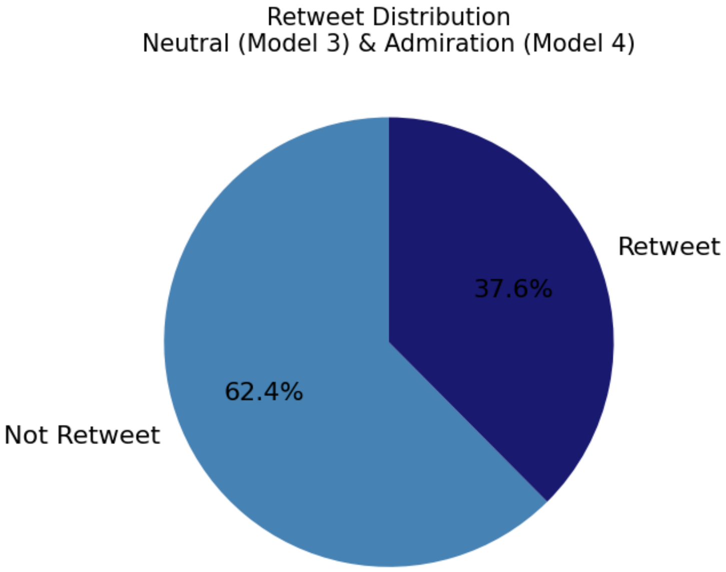


SENTIMENT SKEW IN RETWEETS

- While **retweets** make up only 17% of the total dataset, they represent **37.5%** of the tweets **reclassified as admiration**.







Fine-Grained Classification: Before and After November 8, 2016

Emotion	Percentage Before the Election	Percentage After the Election	Percentage Change
gratitude	9.18%	3.62%	-5.56%
annoyance	6.27%	10.58%	4.31%
approval	11.01%	15.10%	4.09%
admiration	36.84%	33.81%	-3.03%
anger	3.90%	6.55%	2.65%
disapproval	3.01%	5.40%	2.39%
caring	3.99%	2.09%	-1.90%
love	1.64%	0.45%	-1.19%
curiosity	5.83%	4.72%	-1.11%
excitement	3.10%	2.46%	-0.64%

CHI-SQUARED TEST P-VALUE

p = 4e-309

PERCENTAGE CHANGE IN DISTRIBUTION

POSITIVE EMOTIONS

-11.01%

NEGATIVE EMOTIONS

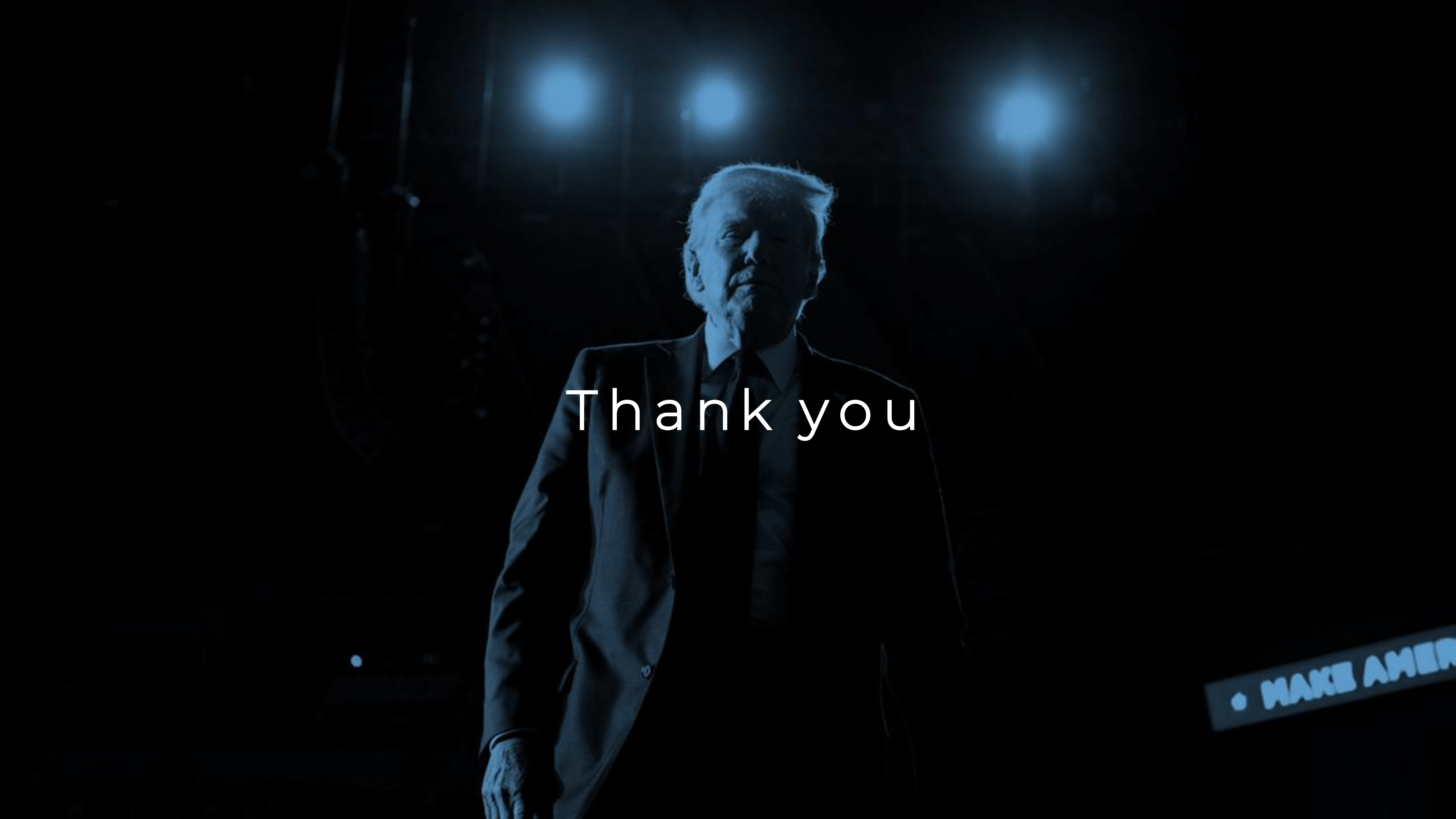
10.12%

Final Remarks

We observed a **clear shift** in the emotional landscape of his Twitter communication. Most notably, while positive emotions consistently represented the majority of tweets across models, there was a **notable increase** in **negative emotions** in the **post-election period**. The fourth model provided deeper insight by breaking down the neutral class into discrete emotions, revealing that **admiration** was the **most prominent category** and that emotions such as **annoyance**, **disapproval**, and **anger** became **more pronounced after the election**.

FUTURE IMPROVEMENTS

- **Enhance emotion classifiers** to match or surpass state-of-the-art transformer accuracy, improving emotional representation in tweets and yielding more precise research insights
- Manual review indicates difficulty detecting contextual nuances such as **sarcasm**; integrating contextualized embeddings (BERT) with static embeddings (GloVe) may improve classifier performance
- Fine-tune emotion classifiers on political tweets to capture **domain-specific nuances** absent in models trained on general content


Thank you