Are we spreading Hate or Reach? Towards Generating Contextually Relevant Reactions for LinkedIn Data

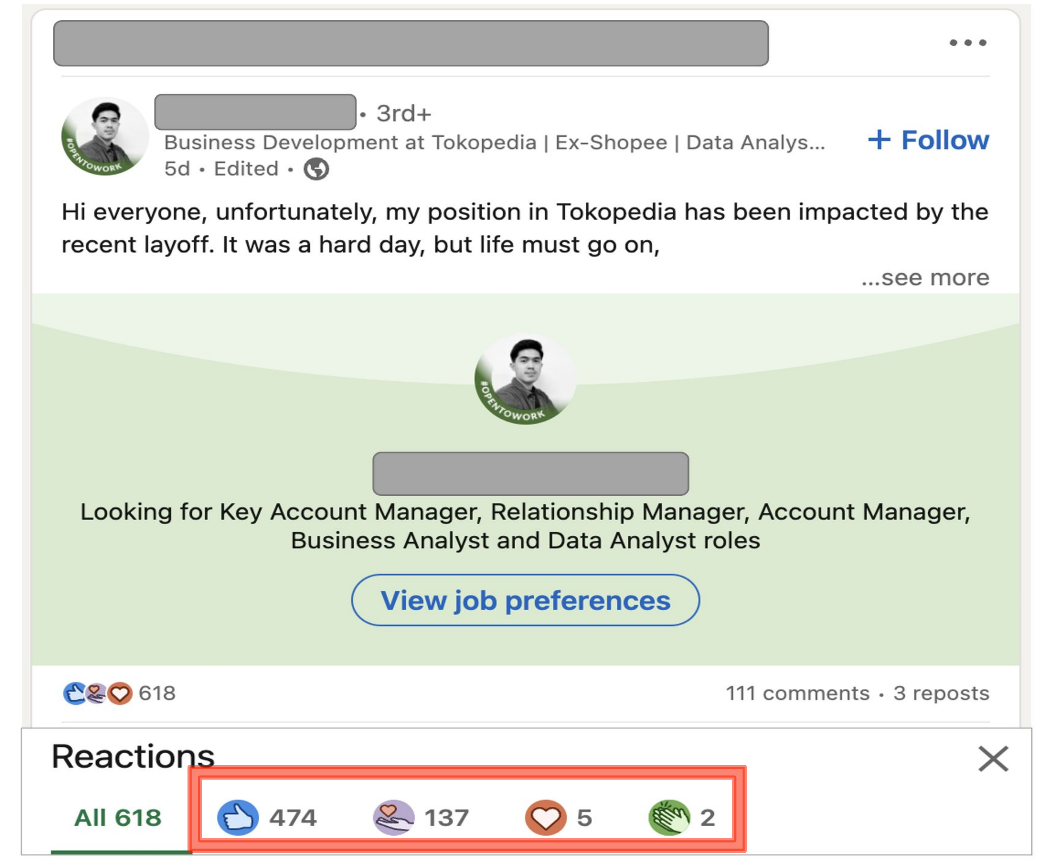
Group 17

Ajinkya Nagare, Amit Singh, Anay Dhombe, Ashwin Chafale, Sanket Bhilare

USC Viterbi School of Engineering

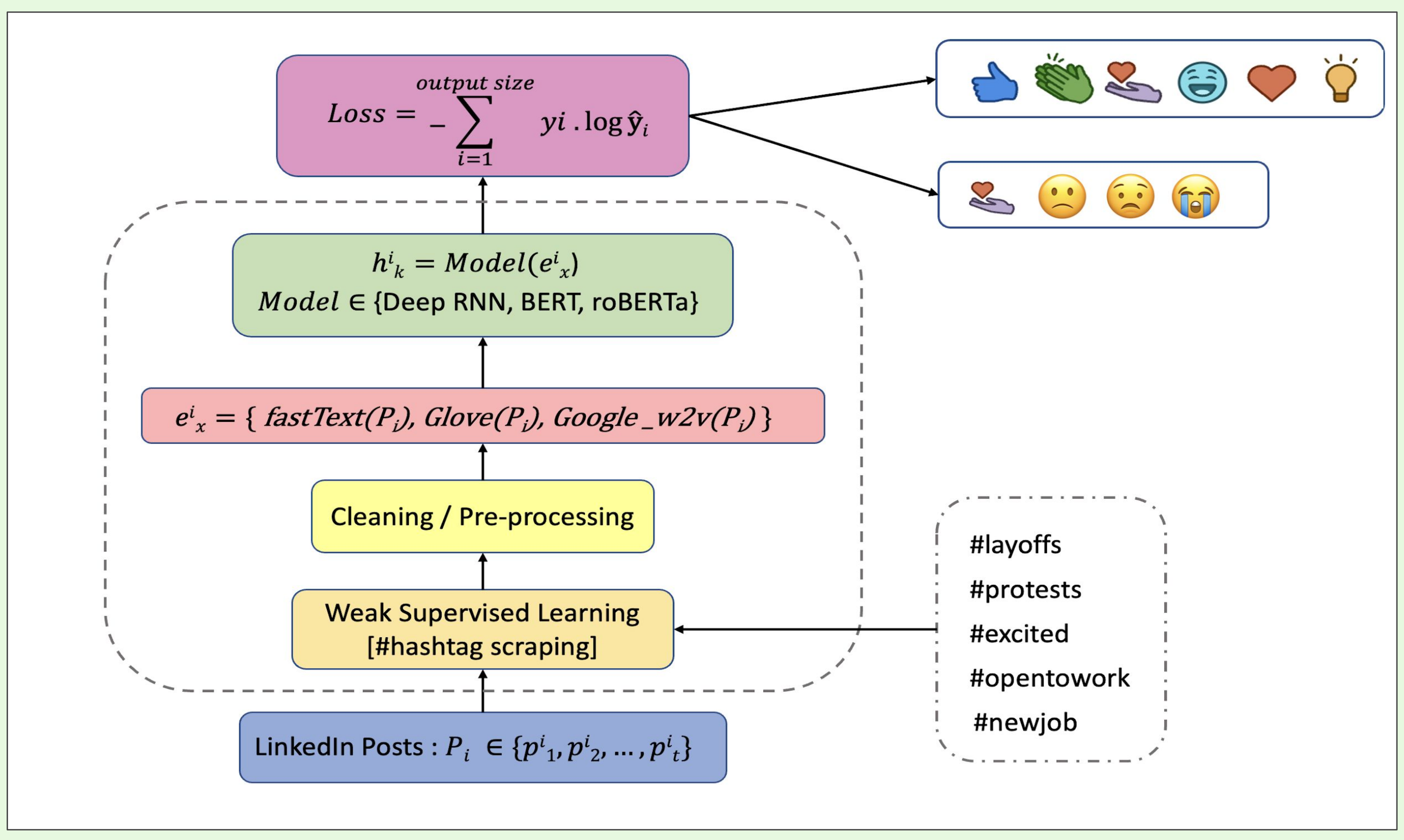
Motivation

- LinkedIn gets flooded with 1000s of **posts**, and there are **multiple reactions** to react to a post
- Most of the time, such **reactions** often tend to be **against the sentiment of the post**
- This **spreads negative sentiments** amongst the users creating a negative psychological impact on the user. (E.g., like/celebrate/funny reaction to a post related to unemployment/protests)
- **Novelty**: We propose a system that generates only contextually relevant reactions based on the sentiment.
- Our system aims to answer the following **research questions**:
  - How many people react with a reaction opposite to the LinkedIn post sentiment?
  - How can we generate contextually relevant reactions to restrict the negative hate spread?



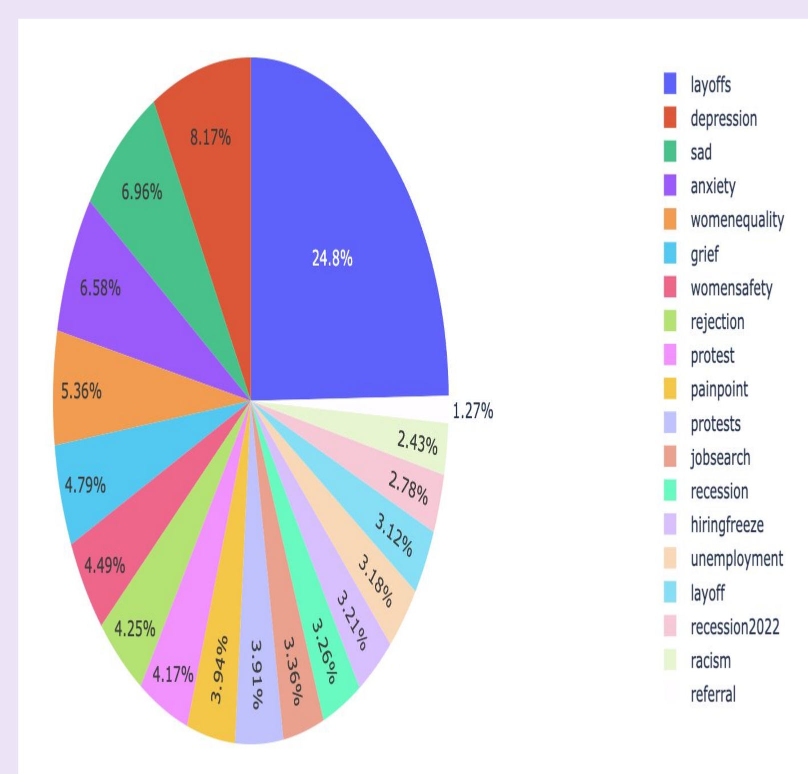
Design

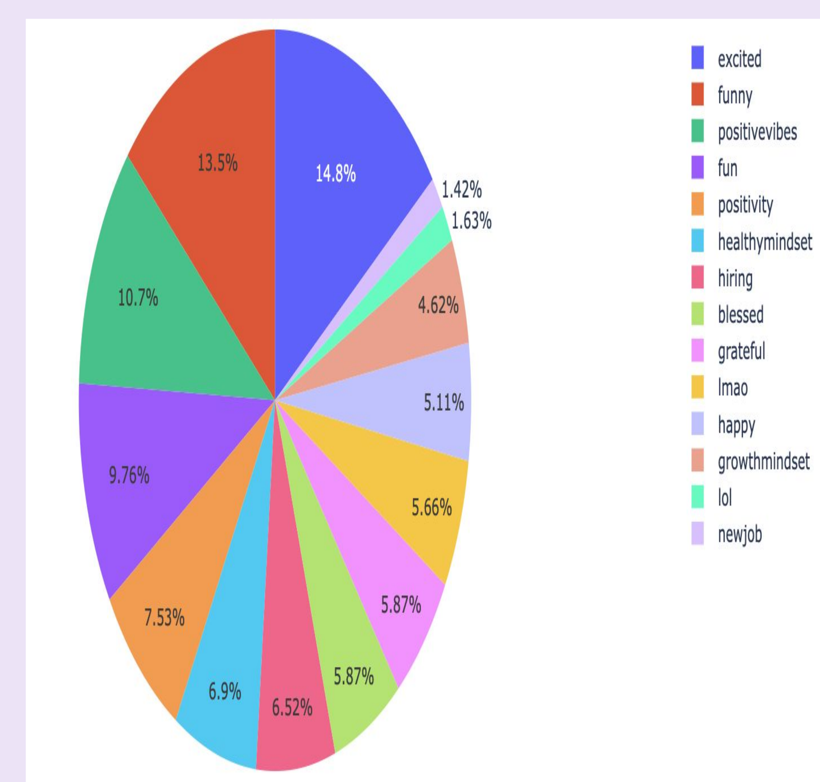
- Custom model architectures such as **Deep Bidirectional- LSTMs, GRUs** and **Stacked LSTM-GRUs** were trained on **20K posts.**
- **FastText, Google pre-trained w2v,** and **GloVe** were used to obtain vector representation for words.
- Pre-trained **BERT$_{BASE}$** and **RoBERTa$_{BASE}$** models were fine-tuned on our training data.



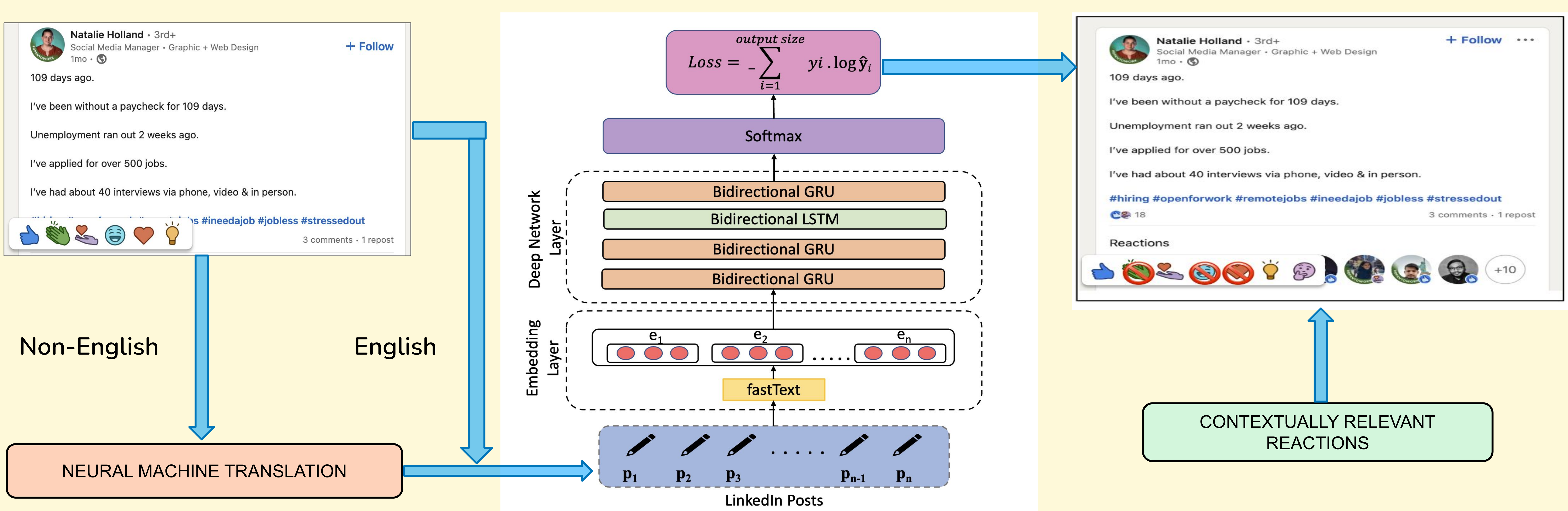
Dataset

- Used **Weak Supervised Learning** approach to generate and annotate data from raw LinkedIn posts
- Around **20K posts** were scraped using a **scraping script** and 30 different hashtags
- Posts were further classified into respective sentiments based on the hashtags (E.g., Neg sentiment for #layoffs, #protests, etc.)





Model Architecture & Pipeline



Analysis and Results

- Models were evaluated on **~3K test data posts.**
- **False positives** are of higher concern to us. Hence, we want **precision to be high.**
- **Stacked GRU-LSTM** performed better as compared to other custom models.
- Fine-tuned BERT$_{BASE}$ and RoBERTa$_{BASE}$ models were the **top performers.**

Model	Embedding	Avg. P	Avg. F1	Avg. Accuracy
Deep Bidirectional-GRUs	Fast Text	0.87	0.87	0.87
	Glove	0.85	0.85	0.85
	Google-W2V	0.86	0.86	0.86
Stacked GRU-LSTM	Fast Text	0.89	0.88	0.89
	Glove	0.86	0.85	0.85
	Google-W2V	0.87	0.87	0.87
BERT$_{BASE}$	Pre-trained	0.96	0.96	0.96
RoBERTa$_{BASE}$	Pre-trained	0.96	0.96	0.96