Natural Language Processing for Emoticon Prediction in Live-Stream Chats

Jessica Fu, Lucia Zhang, Joseph Gozon

CAIS++
CENTER FOR AI IN SOCIETY'S STUDENT BRANCH

Introduction

Performing natural language processing on live-stream conversations has yet to be extensively explored. However, with the growing importance of online communities in the lives of people across the globe, understanding and analyzing the language used between individuals to communicate online is essential. We chose Twitch as our source for live-stream conversations because it uniquely documents live, online community conversation.

Data

We built the dataset from chat logs of past stream videos (VODs) from Twitch streamer Joe Bartolozzi. Timestamps, usernames, bot commands, mentions, and links were removed from the chat logs. The data consists of roughly 37,000 lines of chat obtained from 8 VODs.

```
[0:00:10] cookie_ptd: JOOOOOOOOE
[0:00:17] mackycatt: JOE
[0:00:27] suteki90: yes
[0:00:27] lilThwonk: catJAM catJ
[0:00:27] flexruler: Yee its goo
[0:00:28] gio_samuari: Yooo
```

```
joooooooooe
joe
yes
catjam catjam
yee its good
yooo
```

Figure 1: Left, raw chat data. Right, cleaned chat data.

The processed data was then labeled with one of five labels (*joebartbusiness*, *joebartlongneck*, *joebartwebelieve*, *lul*, and *catjam*) based on the modal emoticon present within ten lines of chat. These labels were one-hot encoded. There was severe class imbalance, with *joebartwebelieve* having roughly 10,000 chats, *catjam* having 9,000, *lul* having 6,500, *joebartbusiness* having 6,500, and *joebartlongneck* having 5,000.

joebartwebelieve

forza	0	0	1	0	0
yes	0	0	1	0	0
brokeback	0	0	1	0	0
joey bears back	0	0	1	0	0
is brooke moving in wit u	0	0	1	0	0
joebartlongneck	0	1	0	0	0

catjam

Figure 2: Labeled chat data.

Methods

We used the labeled chat rows to fine-tune a Bidirectional Encoder Representations from Transformers (BERT) base model to encode and generate word embeddings for the emoticons. We froze the first five encoding layers and all the embedding layers to prevent overfitting due to the small size of our dataset and computational limitations. We left the remaining layers unfrozen, allowing the model parameters to be tuned by the language present in the data. A batch size of 32 was used for 4 epochs of training. Adam was chosen as the optimizer with a learning rate of 5e-5. We used a sigmoid activation function with binary cross entropy loss. One fifth of the training steps were used as warm-up steps. A train-test split of 0.15 was used to create the training and validation sets. The trained model was then used for classification on new individual lines of chat data.

A web extension was created to allow users to interact with the model. The web extension took real-time chat data and ran it through the model to produce an associated emoticon that it then displayed to the user through a graphic interface.

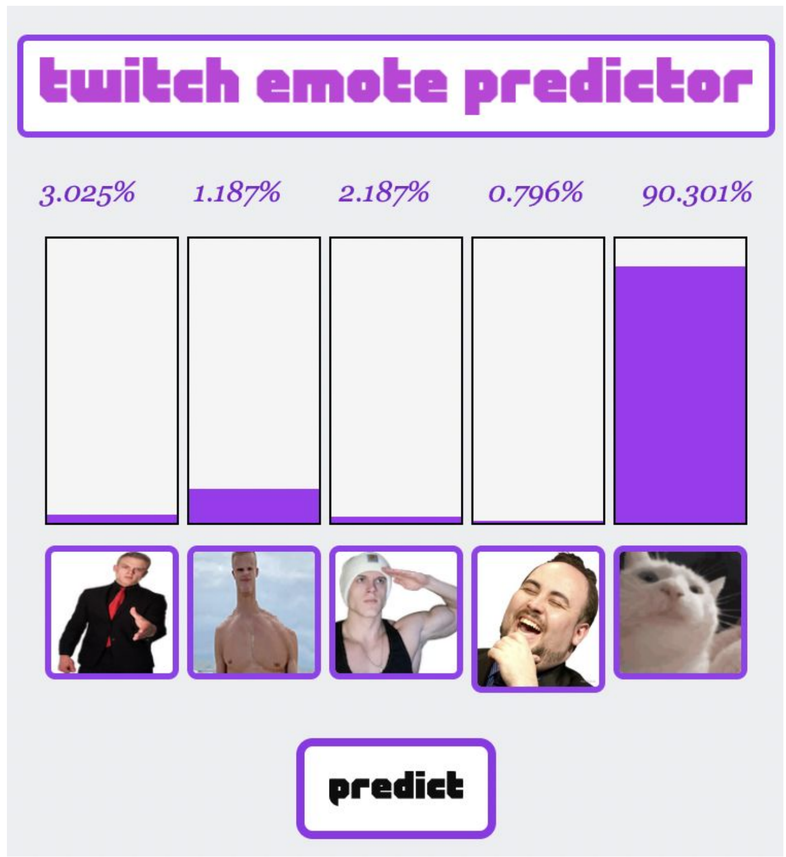

Figure 3: Web extension for emote prediction.

Results

Our model had an accuracy of 0.8418, a precision of 0.75, a recall of 0.27, and a f1-score of 0.37. However, as the classes are imbalanced, AUROC serves as a better representation:

joebartbusiness: 0.73298
joebartlongneck: 0.71129
joebartwebelieve: 0.7687
lul: 0.74796
catjam: 0.85962

The AUROC probabilities reveal that the model was able to obtain fairly accurate results.

```
test comment: everyday is a slay

joebartbusiness: 0.37385496497154236
joebartlongneck: 0.18162694573402405
joebartwebelieve: 0.16215109825134277
lul: 0.2218235284090042
catjam: 0.02542971447110176

final prediction: joebartbusiness
```

Figure 4: Sample test comment and model predictions.

Conclusion

While solely classifying emoticons currently has limited applications, the implications of the model and the logic behind it can be extended to future research regarding online communication, which often differs from in-person communication due to the anonymity afforded by online platforms and presence of different language patterns.

Gaining a clearer understanding of the impact of these online communities on sensitive topics such as racism, sexism, and violence will enable communities to maintain respect and empathy while expressing their own identity. To achieve this goal, future steps would first necessitate expanding the data volume by including a greater number of Twitch communities and covering more diverse genres. Expanding the model beyond classification would reveal patterns in content and speech of various communities, depending on content creator, platform, and the content itself.