CS640 Project Report

Team members: Qingyang Xu, Ge Gao, Hengwei Wang, Soren Nelson, Rishab Nayak

Project Description

This project is COVID-19 Instagram posts emotion detection (anger, fear, joy, and sadness) in relation to images of East Asian people.

The basic task is to apply an emotion detection model trained on twitter dataset to instagram posts and examine the relation between emotion and the existence of East Asian people.

The bonus task is to build a model for detecting whether there's an Asian person in the instagram posts images.

Data Preprocessing

We downloaded instagram data and the training data for anger, fear, joy, sadness respectively. For the emotional data, we dropped the intensity feature as our main task focuses on classification instead of regression. We combined all of the emotional data into a single dataframe and shuffled the order. We were then able to split the data frame into a training set, a validation set, and a test set in a 6:2:2 ratio.

We cleaned the data by removing @ annotated usernames, url links, special characters including newline character and return character. We also stripped hashtags and removed emojis in the comments. Finally we translated all data to English using the python package deep_translator's Google Translate.

Tokenization

We tokenized the data using the model's corresponding Hugging Face tokenizer (see below for models used). We set the tokenizers maximum length to match the max length of an instagram post as these were much longer than the tweets. Padding and truncation were applied when applicable.

Model Selection

Roberta-base

We chose roberta-base as our training model. We tokenized the data with auto tokenizer from this model because our dataset includes multiple languages. We chose AdamW as our optimizer (see below optimizer part) and chose cross entropy as our loss function. We set the batch size of both the training and testing set to 8. As is often recommended, we set the learning rate of AdamW to 1e-05 and the epsilon value to 1e-08. A classification layer is added to turn hidden results learned by RoBERTa into 4 classes of emotions.

Pretrained Emotion Models

We also attempted to use models pretrained on HuggingFace's emotion twitter dataset with labels; "sadness", "joy", "love", "anger", "fear", and "surprise"; *i.e.* a superset of our labels. We fine tuned these models on our twitter dataset by adding a dense layer which took in the logits for the 6 classes of the HuggingFace dataset and reduced them to 4 classes. We used the same hyperparameters as the best roberta-base model (see below). The models used were Distilbert (base-uncased-emotion), Bert (base-uncased-emotion), Roberta (base-emotion), and Albert (base-v2-emotion).

Naive Bayes

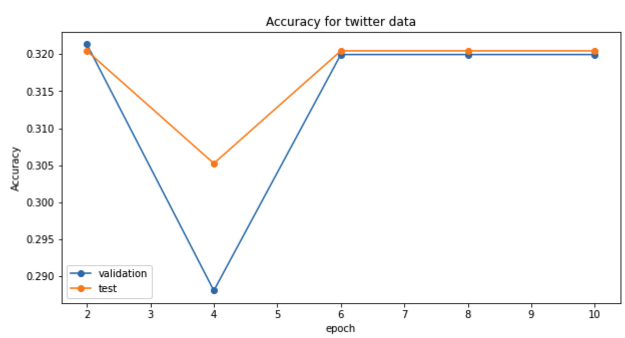
Lastly, we applied a naive bayes classifier as a baseline.

Hyperparameters

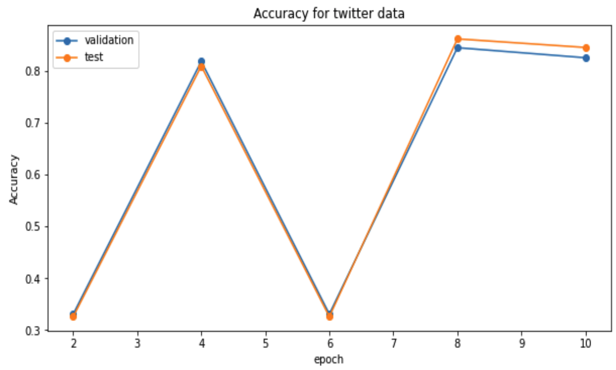
Optimization:

We compared stochastic gradient descent, AdamW and AdaFactor and chose Adamw for better consistent performance.

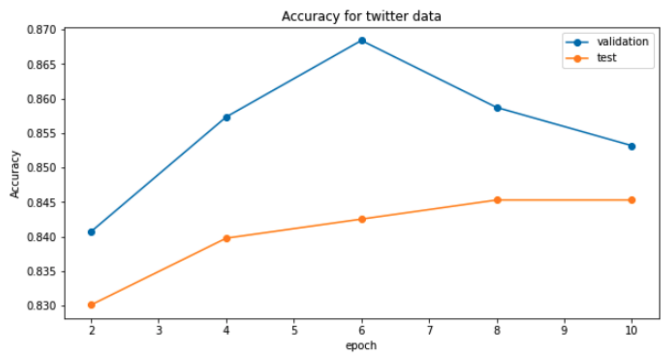
- Adafactor:



- Stochastic Gradient Descent:

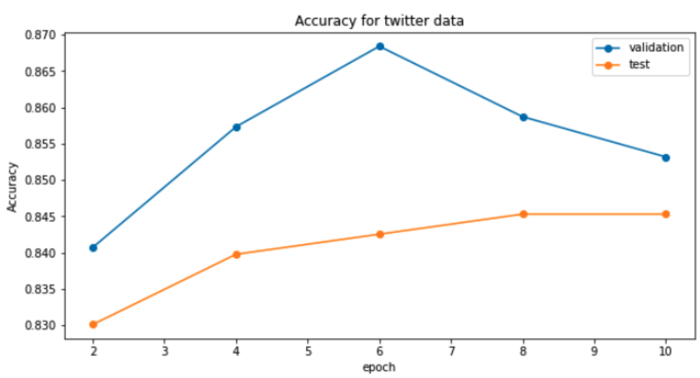


- AdamW:

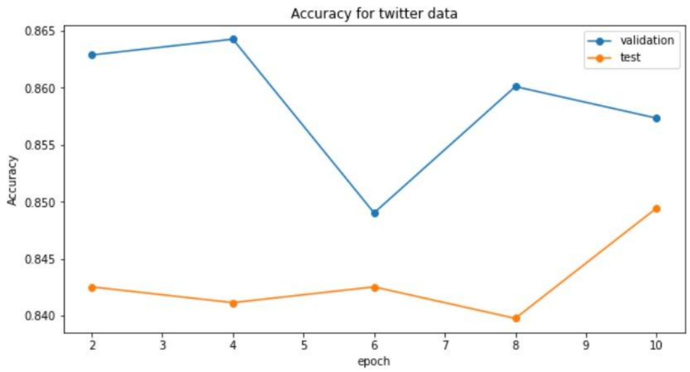


Learning Rate:

We noticed that we were able to improve on the clustering speed by increasing the learning rate without sacrificing the performance accuracy. Originally we were achieving a test accuracy of around 0.84 in 10 epochs with a learning rate of 1e-05.



We were able to achieve a similar test time accuracy above 0.84 with only 2 epochs after we increased our learning rate to 2e-05.



Results

For the model with roberta-base, our best model, we trained with epochs 2,4,6,8,10 separately to show its performance. Here is the classification report of the model.

Epochs=2:

classification report for test set

	precision	recall	f1-score	support
Joy	0.83	0.77	0.80	220
Fear	0.84	0.76	0.80	160
Sadness	0.78	0.90	0.83	162
Anger	0.88	0.90	0.89	182
accuracy			0.83	724
macro avg	0.83	0.83	0.83	724
weighted avg	0.83	0.83	0.83	724

Epochs=4:

classification report for test set

	precision	recall	f1-score	support
Joy	0.82	0.81	0.82	220
Fear	0.79	0.80	0.80	160
Sadness	0.86	0.86	0.86	162
Anger	0.88	0.88	0.88	182
accuracy			0.84	724
macro avg	0.84	0.84	0.84	724
weighted avg	0.84	0.84	0.84	724

Epochs=6:

```
classification report for test set
              precision    recall  f1-score   support

         Joy       0.81      0.81      0.81       220
        Fear       0.82      0.79      0.81       160
     Sadness       0.84      0.88      0.86       162
       Anger       0.91      0.89      0.90       182

    accuracy                           0.84       724
   macro avg       0.84      0.84      0.84       724
weighted avg       0.84      0.84      0.84       724
```

Epoches=8:

```
classification report for test set
              precision    recall  f1-score   support

         Joy       0.81      0.83      0.82       220
        Fear       0.82      0.78      0.80       160
     Sadness       0.84      0.87      0.86       162
       Anger       0.91      0.90      0.90       182

    accuracy                           0.85       724
   macro avg       0.85      0.84      0.85       724
weighted avg       0.85      0.85      0.85       724
```

Epochs=10:

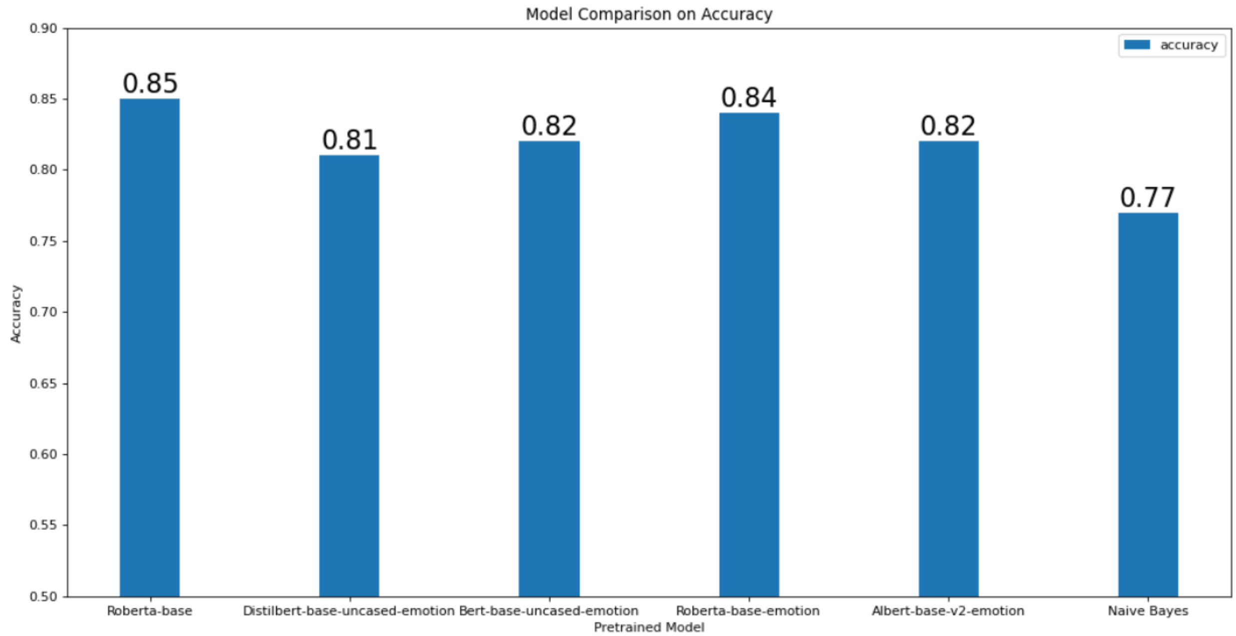
```
classification report for test set
              precision    recall  f1-score   support

         Joy       0.82      0.82      0.82       220
        Fear       0.82      0.78      0.80       160
     Sadness       0.83      0.90      0.87       162
       Anger       0.90      0.88      0.89       182

    accuracy                           0.85       724
   macro avg       0.85      0.85      0.85       724
weighted avg       0.85      0.85      0.85       724
```

Then we used our model to predict emotions in the Instagram dataset. We used Pearson R Correlation value to show the correlation between our prediction results and the original posts. The correlation of the four kinds of emotion were calculated respectively:

- Correlation with Joy: 0.10961711516111323
- Correlation with Fear: -0.03095183092046343
- Correlation with Sadness: -0.06202519724378585
- Correlation with Anger: -0.053837380005938684

Model Comparison

Below we show the accuracy of each of our models on the Twitter held out test set. Each model outperformed the Naive Bayes approach. The Roberta models were the most performant over other methods regardless of the pretrained weights we started with. Finally, and perhaps most interestingly, the generic pretrained model outperformed those pretrained for emotion detection.



Here are some of the example predictions made by the model. It also shows the preprocess methods we applied.

Roberta (base)

```
Prediction: Joy
Processed:
smiling cat with heart eyes smiling cat with heart eyes smiling cat with heart eyes smiling
cat with heart eyes smiling cat with heart eyes
Unprocessed:
😻😻😻😻😻

Prediction: Fear
Processed:
Self isolation nachos! The homemade queso mixes nicely with the loneliness to create a subt
le aroma of doom . . . . . #nachos #vegetarian #vegetariannachos #food #foodporn #isolation
#coronavirus #queso #homemade #cheese #chips #quac #beans #spicy #covid_19
Unprocessed:
Self isolation nachos! The homemade queso mixes nicely with the loneliness to create a subt
le aroma of doom
.
.
.
.
.
#nachos #vegetarian #vegetariannachos #food #foodporn #isolation #coronavirus #queso #homem
ade #cheese #chips #quac #beans #spicy #covid_19
```

```
Prediction: Anger
Processed:
We have arrived!!! Keep following the movement... There is a lot that we are preparing for
you!!!
Unprocessed:
Chegamos !!! Vão seguindo o movimento... Tem muita coisa que estamos preparando pra vocês!!
!

Prediction: Joy
Processed:
Thank you  smiling face results after one treatment star struck
Unprocessed:
Thank you @bbn_donibziee ☺️ results after one treatment 🤩

Prediction: Joy
Processed:
I am thoroughly enjoying the time I get to spend at home catching up on my reading! sparkli
ng heart open book sparkling heart #reading #books #book #selfisolation #isolation #sociald
istancing #covid #covid19 #coronavirus #corona #virus #peaceful #peace
Unprocessed:
I am thoroughly enjoying the time I get to spend at home catching up on my reading! 💖📖💖
#reading #books #book #selfisolation #isolation #socialdistancing #covid #covid19 #coronavi
rus #corona #virus #peaceful #peace
```

Naive Bayes

```
Prediction: Sadness
Processed:
smiling cat with heart eyes smiling cat with heart eyes smiling cat with heart eyes smiling
cat with heart eyes smiling cat with heart eyes
Unprocessed:
😻😻😻😻😻

Prediction: Anger
Processed:
Self isolation nachos! The homemade queso mixes nicely with the loneliness to create a subt
le aroma of doom . . . . . #nachos #vegetarian #vegetariannachos #food #foodporn #isolation
#coronavirus #queso #homemade #cheese #chips #quac #beans #spicy #covid_19
Unprocessed:
Self isolation nachos! The homemade queso mixes nicely with the loneliness to create a subt
le aroma of doom
.
.
.
.
.
#nachos #vegetarian #vegetariannachos #food #foodporn #isolation #coronavirus #queso #homem
ade #cheese #chips #quac #beans #spicy #covid_19

Prediction: Anger
Processed:
We have arrived!!! Keep following the movement... There is a lot that we are preparing for
you!!!
Unprocessed:
Chegamos !!! Vão seguindo o movimento... Tem muita coisa que estamos preparando pra vocês!!
!
```

```
Prediction: Anger
Processed:
Thank you  smiling face results after one treatment star struck
Unprocessed:
Thank you @bbn_donibziee ☺️ results after one treatment 🤩

Prediction: Anger
Processed:
I am thoroughly enjoying the time I get to spend at home catching up on my reading! sparkli
ng heart open book sparkling heart #reading #books #book #selfisolation #isolation #sociald
istancing #covid #covid19 #coronavirus #corona #virus #peaceful #peace
Unprocessed:
I am thoroughly enjoying the time I get to spend at home catching up on my reading! 💖📖💖
#reading #books #book #selfisolation #isolation #socialdistancing #covid #covid19 #coronavi
rus #corona #virus #peaceful #peace
```

Conclusion

1. Challenges

1.1 lack of training data
Due to time issues, we did not have time to find other supplementary data sources, so we only used the Twitter data provided in the project List doc as training data. Since the amount of training data is less than the amount of instagram test data, the model does not have good generalization. Therefore, insufficient training data is the biggest weakness of our project, which needs to be improved in the future.

1.2 cuda issues
After trying, the speed of using the CPU to train the model is too slow to detect the results, so we use the GPU for training. In an attempt to train multiple pretrained models for comparison, cuda-related memory and runtime errors have a great impact on the training process, and such bugs are difficult to fix.

1.3 Multilingual problem
It is not difficult to find that our Twitter training data is pure English, while the instagram testing data is multilingual, including no less than English, Japanese, Korean, Chinese, Russian and other languages. This brings some interference to our model training. If you choose a multi-language pretrained model, the training process of the training data will cause a waste of accuracy. So we choose to translate all languages into English in the data preprocessing stage. To a certain extent, the language mismatch of training and test data has an impact on the accuracy of the model.

2. Negative Results

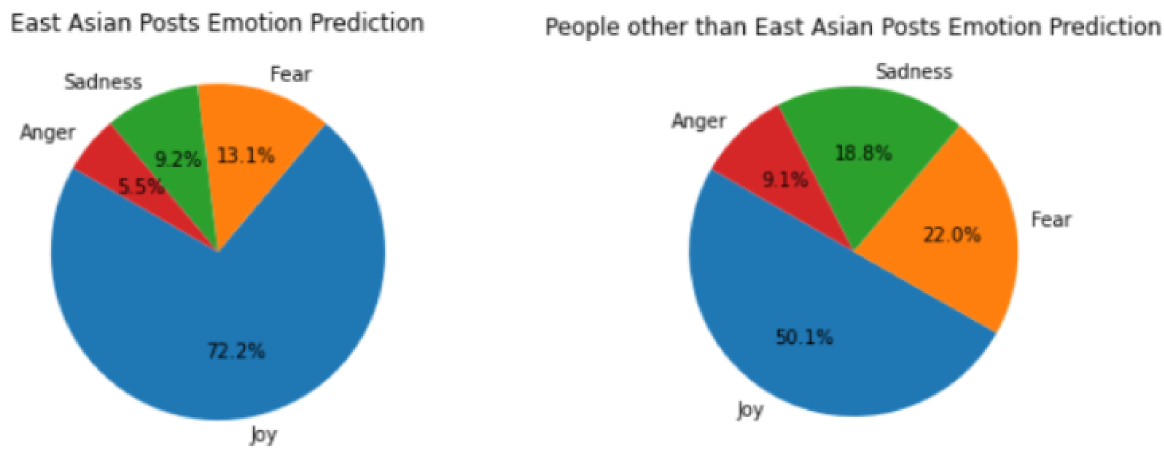


Figure: Percentage of different emotions in the ig posts of East Asians and people from other countries using roberta-based pretrained model

Although the accuracy of our model on the training set is about 85%, the results of the model on the test set are not satisfactory. From the perspective of correlation, for the four emotions of ‘joy’, ‘fear’, ‘anger’, and ‘sad’, the correlation between the test set and joy is the highest, but in the actual prediction, there are many samples wrongly categorized, so a firm conclusion can not be drew based on correlation.

From the statistics of the results, in the East Asian posts in the test set, the proportion of joy reached 72%, while the joy of others accounted for about 50%, which is inconsistent with our pre-inference before training. In our conjecture, due to the impact of covid-19, East Asians’ posts about the epidemic should be very anxious and panic, so their share of joy should be lower than that of others. However, the model predicts that East Asians’ Twitter's happiness index is much higher than others. For this result, we give two explanations.

The first explanation is that our model's predictive ability is insufficient, and wrong inferences have been made. The second explanation comes from my own experience. As a Chinese who came to the United States this summer, I have personally experienced the COVID-19 outbreak in China and the slowing down. At the end of 2019, COVID-19 broke out in China. From that time to about the spring of 2020, I was isolated at home, feeling anxious and panicked. However, starting in March and April, China has fully controlled the epidemic, and Chinese citizens have become accustomed to living in isolation. After the summer, the epidemic in China subsided and people resumed their entertainment lives. Personally, COVID-19 did not affect me after April. At the same time, due to the spread of COVID-19, people in other countries have begun to suffer from the epidemic. So from my personal experience, since the posts collected by the test data were sent from the end of January to the end of August in 2020, the conclusion that the happiness index of East Asians is higher than that of people in other countries from the prediction results can be explained. However, because the test set did not accurately classify whether East Asians are Chinese or from other countries, this explanation is not supported by sufficient evidence.

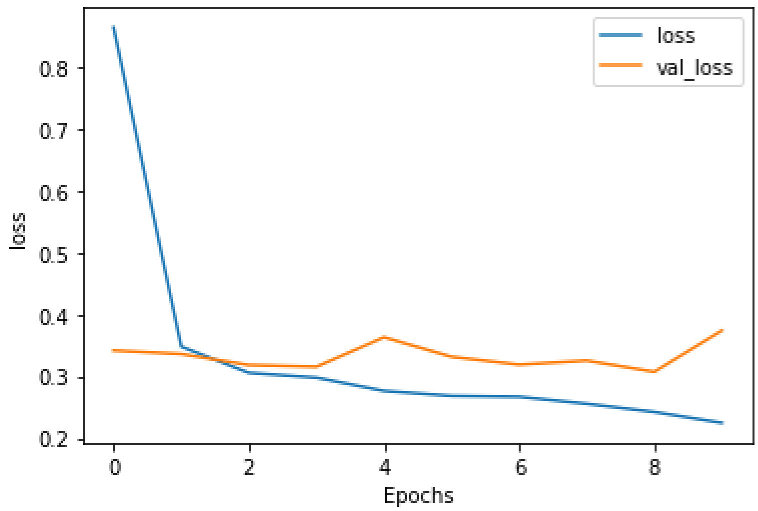
All in all, these two interpretations of the results are both likely to be correct, and we need to further optimize the model to give the final correct interpretation.

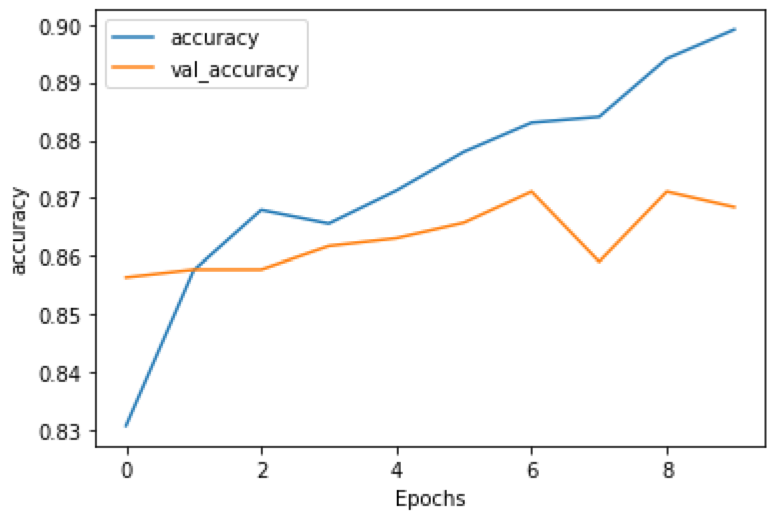
Bonus

The csv data was used to create a dataset containing the filenames and the labels associated with the images. We chose to use Stratified K-Fold over K-Fold to account for the imbalanced class distributions. The ImageDataGenerator was configured to use the VGG19 preprocessing function to prepare the images to be input to the classification model.

The classification model used the pretrained VGG19 model with weights, added a Flatten Layer, a Fully Connected, and a Dropout Layer before the final Fully Connected Layer with a Softmax Activation function. The model was compiled with the Adam Optimizer, Binary Crossentropy Loss function, and three metrics, Accuracy, Precision and Recall were tracked through the training process.

The training and validation data generators were configured with flow_from_dataframe, using data from the Stratified K-Fold splitting function. The model was fit using the data generators, fit for 10 epochs with a batch size of 32. The Validation Recall, Precision and Accuracy for each of the folds were recorded after every fold, and the averages were calculated after for the report, resulting in a 86.8% accuracy.





Code Link

- Project code link: https://github.com/rishabnayak/cs640-final-project
- The project focuses on training a model for predicting emotion in Instagram posts. It was trained based on a Twitter dataset. As a bonus, we also build a model for detecting Asian people in images.
- Codes are written in Jupyter notebook.
- The project includes:
    - Data loading
    - Feature engineering
    - Model definition
    - Model training
    - Performance evaluation
    - Emotion predicting

- Compiling environment:
    - Python 3.8
    - Pytorch 1.10.0
    - Transformers 4.12.5
- Project authors: Soren Nelson, Rishab Nayak, Ge Gao, Qingyang Xu, Hengwei Wang

References

1. https://pchanda.github.io/Roberta-FineTuning-for-Classification/