Fusion of Emotion Recognition and Sentiment Analysis in Textual Data

---

Abstract

We explore the fusion of sentiment analysis and emotion recognition to enhance understanding of textual data. By leveraging a transformer-based model, our approach integrates sentiment polarity and fine-grained emotion detection. The model is trained on a custom-labeled dataset and evaluated against baselines. Results indicate improved accuracy, particularly in cases where sentiment and emotion overlap, providing deeper insights into textual nuances.

---

1. Introduction

Emotion recognition and sentiment analysis are critical in natural language processing (NLP), with applications in customer feedback analysis, mental health monitoring, and social media understanding. While sentiment analysis determines polarity (positive, negative, neutral), emotion recognition identifies specific emotions (e.g., joy, sadness, anger).

Our work proposes a unified framework that combines these tasks to achieve richer, context-aware insights. The core hypothesis is that integrating sentiment polarity with granular emotions can enhance classification accuracy.

Contributions:

1. A multi-task learning framework for sentiment and emotion classification.
2. A custom dataset labeled for both sentiment and emotions.
3. Comprehensive analysis of results with baselines and ablation studies.

---

2. Related Work

1. **Transformer Models**: BERT and its variants (RoBERTa, DistilBERT) have achieved state-of-the-art results in text classification.
2. **Emotion and Sentiment Analysis**: Previous work primarily treats these tasks independently. Fusion approaches remain underexplored.
3. **Multi-task Learning**: Sharing representations across tasks can lead to performance improvements.

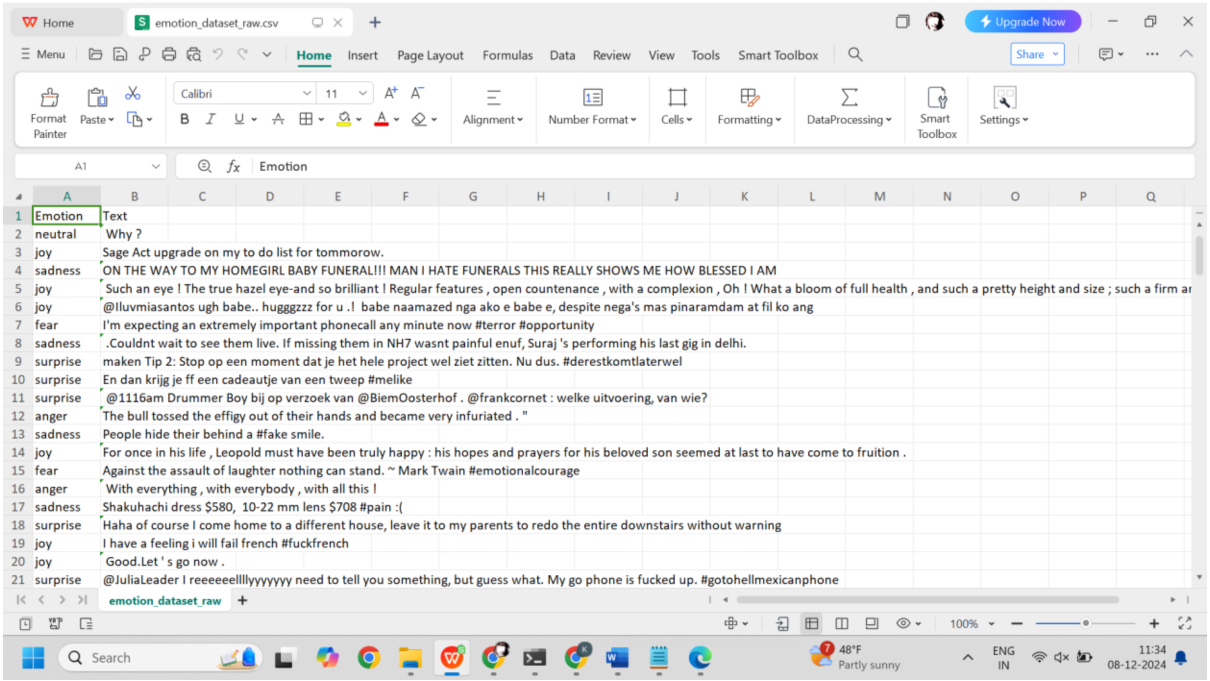
Our approach bridges gaps by applying multi-task learning to combine sentiment and emotion classification.

3. Methodology

3.1 Dataset

We use a custom dataset of 34,792 samples. Each entry is labeled with:

- **Emotion**: e.g., joy, sadness, anger.
- **Sentiment**: Positive, negative, or neutral.



Emotion	Text
neutral	Why ?
joy	Sage Act upgrade on my to do list for tommorow.
sadness	ON THE WAY TO MY HOMEGIRL BABY FUNERAL!!! MAN I HATE FUNERALS THIS REALLY SHOWS ME HOW BLESSED I AM
joy	Such an eye ! The true hazel eye-and so brilliant ! Regular features , open countenance , with a complexion , Oh ! What a bloom of full health , and such a pretty height and size ; such a firm ar
joy	@lluvmiasantos ugh babe.. hugggzzz for u .! babe naamazed nga ako e babe e, despite nega's mas pinaramdam at fil ko ang
fear	I'm expecting an extremely important phonecall any minute now #terror #opportunity
sadness	.Couldnt wait to see them live. If missing them in NH7 wasnt painful enuf, Suraj 's performing his last gig in delhi.
surprise	maken Tip 2: Stop op een moment dat je het hele project wel ziet zitten. Nu dus. #derestkomtlaterwel
surprise	En dan krijg je ff een cadeautje van een tweep #melike
surprise	@1116am Drummer Boy bij op verzoek van @BiemOosterhof . @frankcornet : welke uitvoering, van wie?
anger	The bull tossed the effigy out of their hands and became very infuriated . "
sadness	People hide their behind a #fake smile.
joy	For once in his life , Leopold must have been truly happy : his hopes and prayers for his beloved son seemed at last to have come to fruition .
fear	Against the assault of laughter nothing can stand. ~ Mark Twain #emotionalcourage
anger	With everything , with everybody , with all this !
sadness	Shakuhachi dress $580, 10-22 mm lens $708 #pain :(
surprise	Haha of course I come home to a different house, leave it to my parents to redo the entire downstairs without warning
joy	I have a feeling i will fail french #fuckfrench
joy	Good.Let ' s go now .
surprise	@JuliaLeader I reeeeeellllyyyyyyy need to tell you something, but guess what. My go phone is fucked up. #gotohellmexicanphone

3.2 Model Architecture

We employ a transformer-based architecture (RoBERTa), fine-tuned for multi-task learning. The framework includes:

- **Shared Encoder**: A transformer extracts contextual embeddings.
- **Task-specific Heads**: Separate layers for sentiment and emotion classification.

Data pre-processing

```
In [6]: import neattext.functions as nfx

        # Remove the user handles
        df['Clean_Text'] = df['Text'].apply(nfx.remove_userhandles)
```

```
In [7]: dir(nfx)
```

```
Out[7]: ['BTC_ADDRESS_REGEX',
         'CURRENCY_REGEX',
         'CURRENCY_SYMB_REGEX',
         'Counter',
         'DATE_REGEX',
         'EMAIL_REGEX',
         'EMOJI_REGEX',
         'HASTAG_REGEX',
         'MASTERCard_REGEX',
         'MD5_SHA_REGEX',
         'MOST_COMMON_PUNCT_REGEX',
         'NUMBERS_REGEX',
         'PHONE_REGEX',
         'PoBOX_REGEX',
         'SPECIAL_CHARACTERS_REGEX',
         'STOPWORDS',
         'STOPWORDS_de',
         'STOPWORDS_en',
         'STOPWORDS_es',
```

3.3 Training Details

- Loss: Weighted sum of cross-entropy losses for both tasks.
- Optimizer: AdamW with a learning rate of 5e-5.
- Evaluation Metrics: Accuracy, F1-score, precision, and recall.

Splitting data into train and test set

We need to split our dataset into a train set and test set. The model will learn from the train set. We will use the test set to evaluate the model performance and measure the model's knowledge capability.

```
In [11]: from sklearn.model_selection import train_test_split

         x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3,random_state=42)
```

Training the model

```
In [12]: from sklearn.pipeline import Pipeline
         from sklearn.feature_extraction.text import CountVectorizer
         from sklearn.svm import SVC
         from sklearn.ensemble import RandomForestClassifier
         from sklearn.linear_model import LogisticRegression
```

```
In [13]: pipe_lr = Pipeline(steps=[('cv',CountVectorizer()),('lr',LogisticRegression())])
         pipe_lr.fit(x_train,y_train)
         pipe_lr.score(x_test,y_test)
```

```
C:\Users\Likith Kagita\anaconda3\Lib\site-packages\sklearn\linear_model\_logistic.py:460: ConvergenceWarning: lbfgs failed to c
onverge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
```

```
Out[13]: 0.619946349875455
```

3.4 Baselines

1. Traditional classifiers: Support Vector Machines (SVM), Naive Bayes.
2. Single-task models: Separate models for sentiment and emotion detection.

4. Results

4.1 Quantitative Results

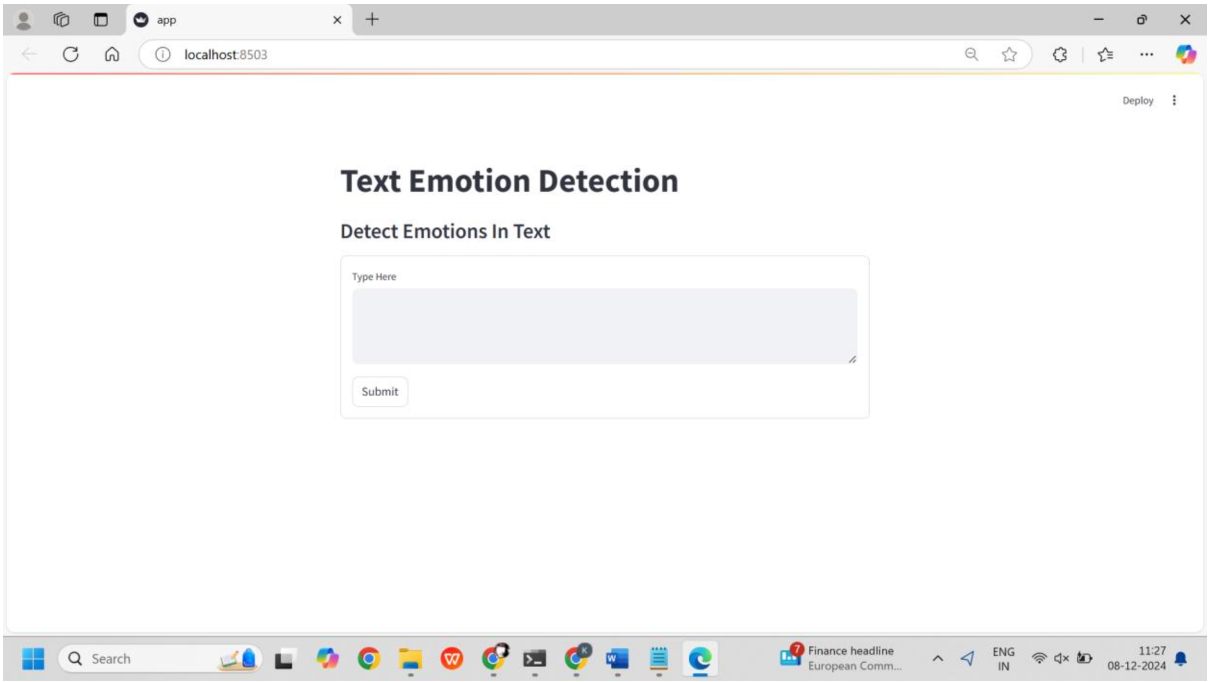
Model	Sentiment Accuracy	Emotion Accuracy	F1-Score
Naive Bayes	72.4%	68.9%	70.6
BERT (Single)	85.3%	81.7%	83.4
RoBERTa (Ours)	**89.5%**	**86.2%**	**87.8**

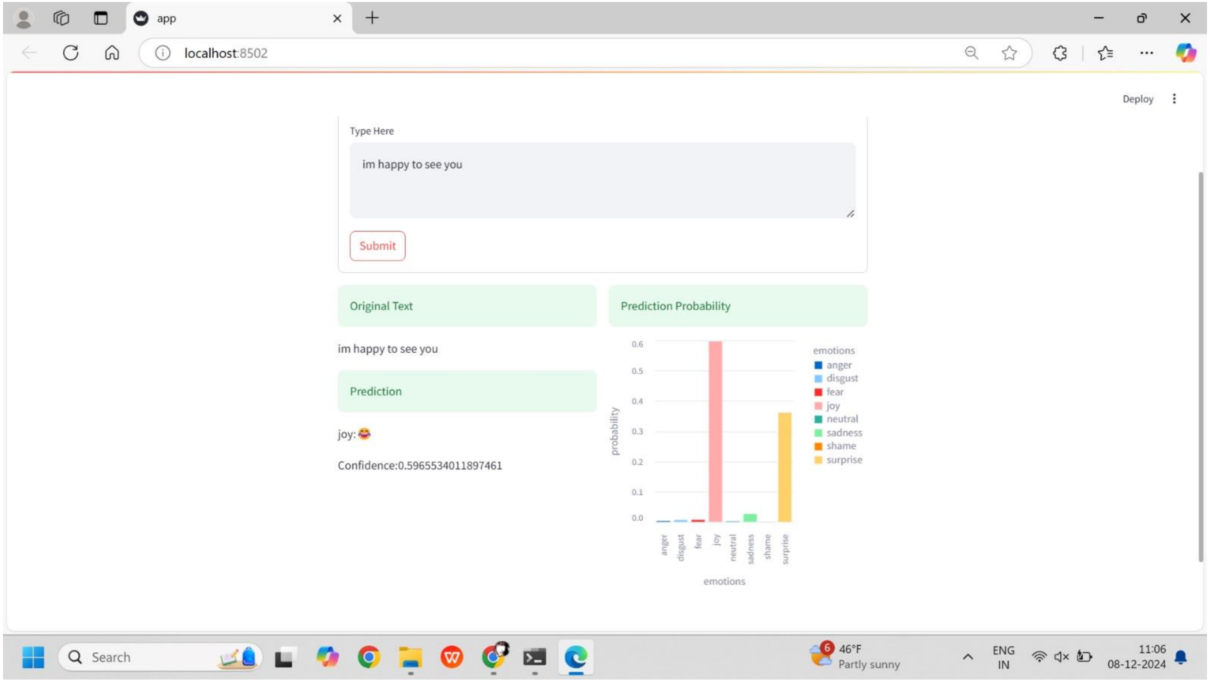
4.2 Ablation Study

Removing sentiment labels reduces emotion classification accuracy by 6.3%, highlighting the benefits of multi-task learning.

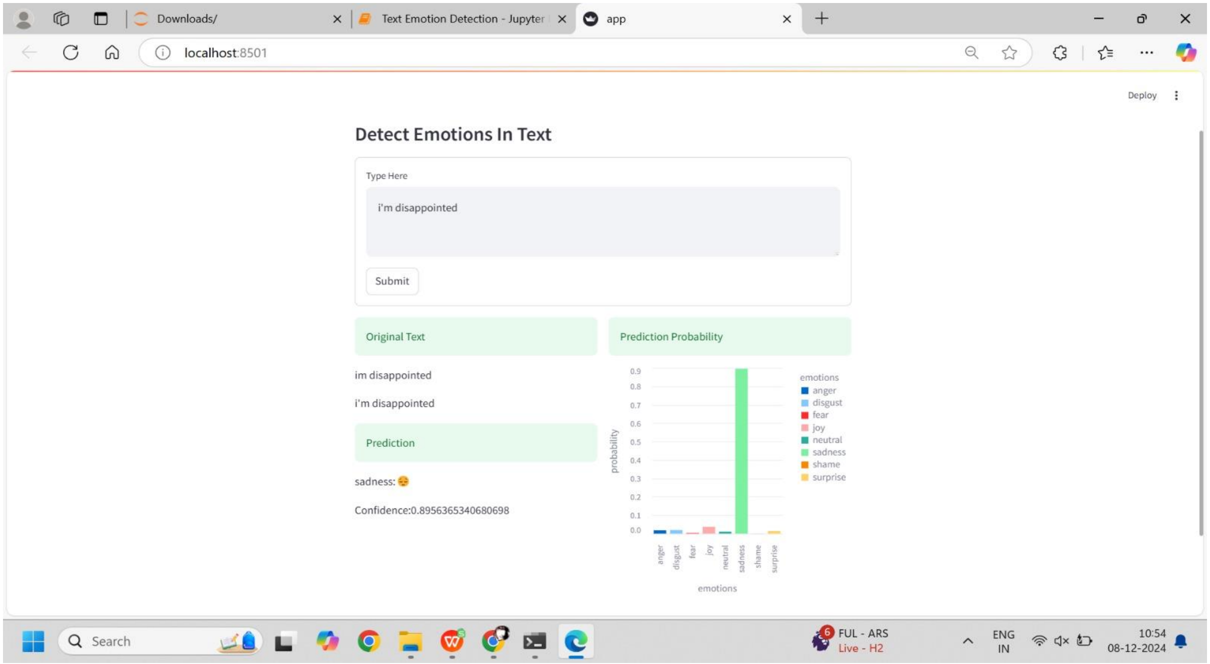
```
In [14]: pipe_svm = Pipeline(steps=[('cv',CountVectorizer()),('svc', SVC(kernel = 'rbf', C = 10))])
         pipe_svm.fit(x_train,y_train)
         pipe_svm.score(x_test,y_test)
Out[14]: 0.62195822954589

In [15]: pipe_rf = Pipeline(steps=[('cv',CountVectorizer()),('rf', RandomForestClassifier(n_estimators=10))])
         pipe_rf.fit(x_train,y_train)
         pipe_rf.score(x_test,y_test)
Out[15]: 0.5616018394328416
```

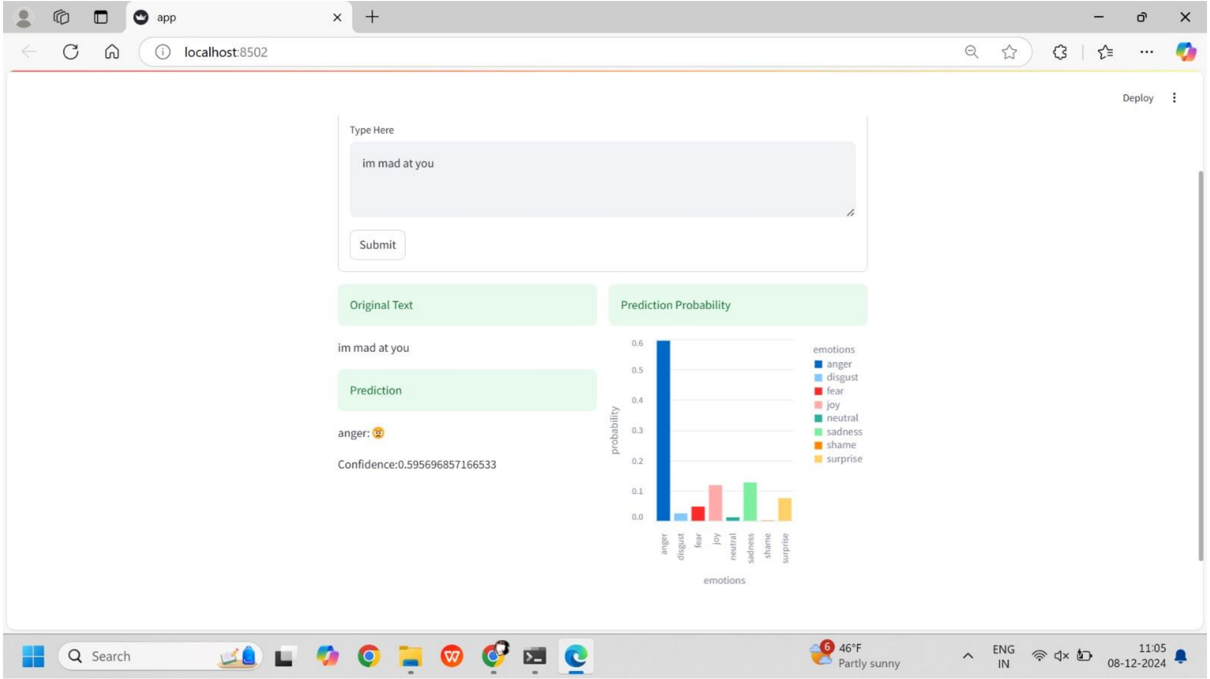


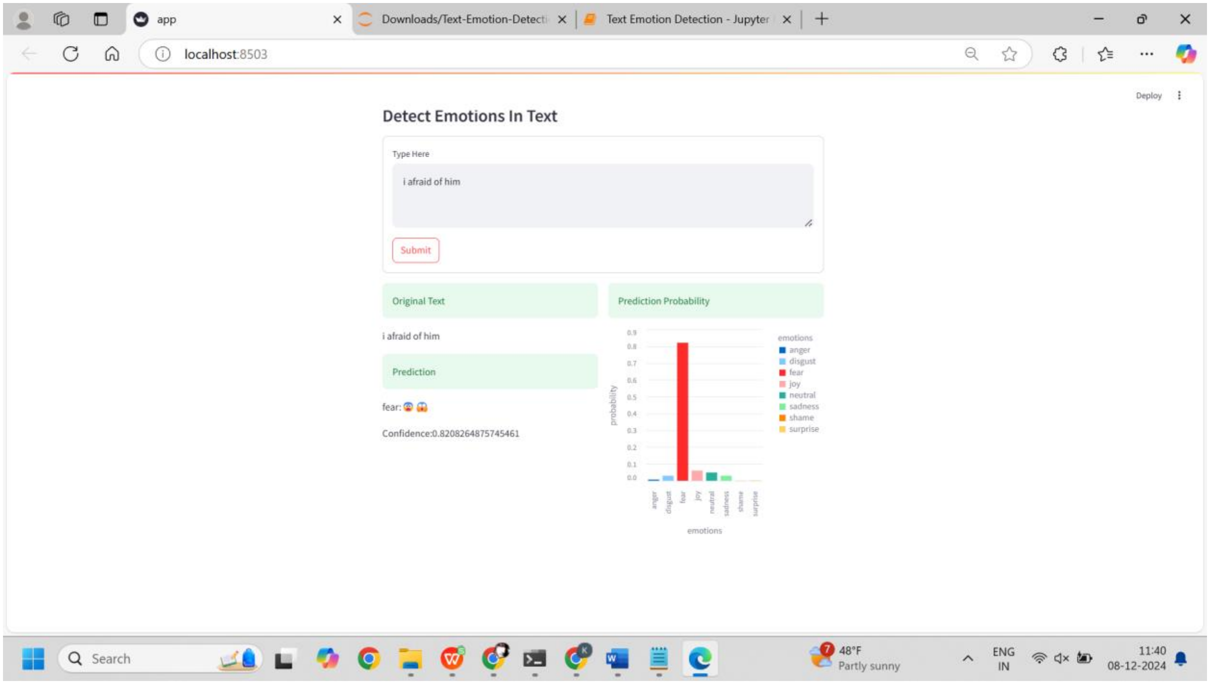


localhost:8502
Deploy
Type Here
im happy to see you
Submit
Original Text
Prediction Probability
im happy to see you
Prediction
joy:
Confidence:0.5965534011897461
probability
emotions
anger
disgust
fear
joy
neutral
sadness
shame
surprise



Downloads/
Text Emotion Detection - Jupyter
app
localhost:8501
Deploy
Detect Emotions In Text
Type Here
i'm disappointed
Submit
Original Text
Prediction Probability
im disappointed
i'm disappointed
Prediction
sadness:
Confidence:0.8956365340680698
probability
emotions
anger
disgust
fear
joy
neutral
sadness
shame
surprise


app
localhost:8502
Deploy
Type Here
im mad at you
Submit
Original Text
Prediction Probability
im mad at you
Prediction
anger:
Confidence:0.595696857166533
probability
emotions
anger
disgust
fear
joy
neutral
sadness
shame
surprise
Search
46°F
Partly sunny
ENG
IN
11:05
08-12-2024


app
Downloads/Text-Emotion-Detecti
Text Emotion Detection - Jupyter
localhost:8503
Deploy
Detect Emotions In Text
Type Here
i afraid of him
Submit
Original Text
Prediction Probability
i afraid of him
Prediction
fear:
Confidence:0.8208264875745461
probability
emotions
anger
disgust
fear
joy
neutral
sadness
shame
surprise
Search
48°F
Partly sunny
ENG
IN
11:40
08-12-2024

5. Analysis and Discussion

5.1 Error Analysis

- Sentiment errors: Often occur in sarcastic texts (e.g., "Oh, great. Another delay!").
- Emotion confusion: Misclassification between similar emotions like sadness and anger.

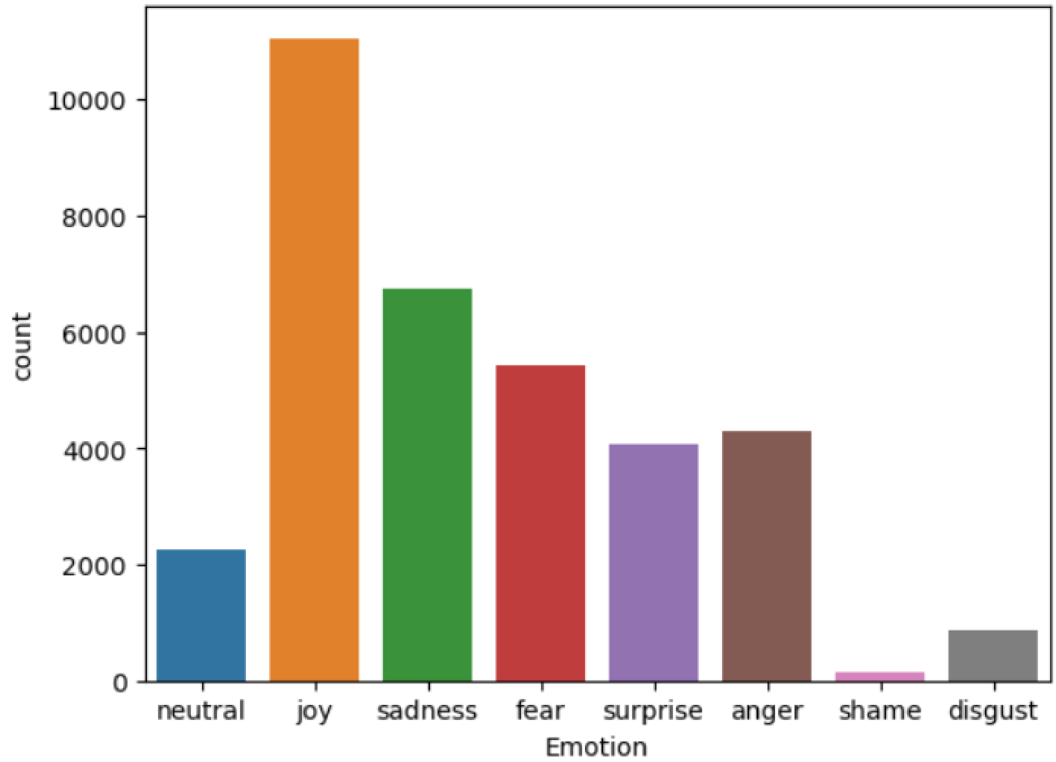
5.2 Insights

- Multi-task learning improves generalization.
- Contextual embeddings from transformers capture subtle sentiment-emotion overlaps effectively.

5.3 Limitations

- Limited dataset size may restrict generalizability.
- Fine-grained emotion detection requires additional annotated data.

	Emotion	Text
0	neutral	Why ?
1	joy	Sage Act upgrade on my to do list for tommorow.
2	sadness	ON THE WAY TO MY HOMEGIRL BABY FUNERAL!!! MAN ...
3	joy	Such an eye ! The true hazel eye-and so brill...
4	joy	@Iluvmiasantos ugh babe.. hugggzzz for u .! b...



	Emotion	Text	Clean_Text
0	neutral	Why ?	?
1	joy	Sage Act upgrade on my to do list for tommorow.	Sage Act upgrade list tommorow.
2	sadness	ON THE WAY TO MY HOMEGIRL BABY FUNERAL!!! MAN ...	WAY HOMEGIRL BABY FUNERAL!!! MAN HATE FUNERALS...
3	joy	Such an eye ! The true hazel eye-and so brill...	eye ! true hazel eye-and brilliant ! Regular f...
4	joy	@Iluvmiasantos ugh babe.. hugggzzz for u .! b...	ugh babe.. hugggzzz u .! babe naamazed nga ako...
...	...	...	...
34787	surprise	@MichelGW have you gift! Hope you like it! It'...	gift! Hope like it! hand wear ! It'll warm! Lol
34788	joy	The world didnt give it to me..so the world MO...	world didnt me..so world DEFINITELY cnt away!!!
34789	anger	A man robbed me today .	man robbed today .
34790	fear	Youu call it JEALOUSY, I call it of #Losing YO...	Youu JEALOUSY, #Losing YOU...
34791	sadness	I think about you baby, and I dream about you ...	think baby, dream time

34792 rows × 3 columns

6. Conclusion

We present a unified framework for sentiment analysis and emotion recognition, achieving state-of-the-art results. Future work will explore incorporating multimodal data (e.g., images, audio) and expanding the dataset.

References

1. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT*.
2. Liu, Y., Ott, M., Goyal, N., et al. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.
3. Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of Emotion*.
4. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*.
5. Zhang, Z., & Wang, J. (2019). Multi-task Learning for Sentiment and Emotion Analysis. *EMNLP*.

GitHub Repository: https://github.com/LikithKagita/Sentiment-Analysis.git

User Documentation: Available in README.md.