

INVESTIGATING THE IMPORTANCE OF EMOJIS IN SENTIMENT ANALYSIS

Team 5

Praneeth Reddy Mukthapuram

Emmanuel Adebayo

Rohit katta

Vineeth Gangavarapu

INTRODUCTION

- Emojis have become an integral part digital communication
- Their role in natural language processing tasks have been understudied

OBJECTIVE

- Research systematically the impact of emojis in sentiment analysis of twitter data.
- Previous work have explored the subject matter. We Explored modern architecture in context of sentiment analysis
- We also specifically focus on twitter data.

METHODOLOGY

- Pre – processing
- Models
- Performance evaluation
- Result

PRE-PROCESSING

- HTML tags, @ mentions, website links
- Emoji dictionary
 - https://en.wikipedia.org/wiki/List_of_emoticons
- Handle miscellaneous terms (false positives)
 - 2:30
 - < tags in html

```
5
6 emoji_dict = {}
7    ":-)": {"emoji": "😄", "description": "happy face"},
8    ":-)": {"emoji": "😄", "description": "happy face"},
9    ":-)": {"emoji": "😄", "description": "happy face"},
10    ":)": {"emoji": "😄", "description": "happy face"},
11    ":->": {"emoji": "😄", "description": "happy face"},
12    ":->": {"emoji": "😄", "description": "happy face"},
13    "8)": {"emoji": "😄", "description": "happy face"},
14    "8)": {"emoji": "😄", "description": "happy face"},
15    ":-)": {"emoji": "😄", "description": "happy face"},
16    ":-)": {"emoji": "😄", "description": "happy face"},
17    ":-)": {"emoji": "😄", "description": "happy face"},
18    "=)": {"emoji": "😄", "description": "happy face"},
19    "=)": {"emoji": "😄", "description": "happy face"},
20    "☺": {"emoji": "☺", "description": "happy face"},
21    "☺": {"emoji": "☺", "description": "happy face"},
22    "☺": {"emoji": "☺", "description": "happy face"}
```

MODELS

- Distil Bert
 - With emojis
 - No Emojis
 - Unicode representations
- Twitter Roberta(cardiffnlp/twitter-roberta-base)
 - 2 inputs - Text and emojis as different features
 - Adds [sep] token between them

PERFORMANCE EVALUAITON

	Accuracy	Precision	Recall	F1-Score
Bert - Emojis	0.764	0.761	0.761	0.758
Bert – No Emojis	0.732	0.748	0.749	0.748
Bert - unicode	0.765	0.763	0.764	0.764
Twitter Roberta	0.78	0.78	0.76	0.76

CONCLUSION

- Yes, emojis does matter. It should be considered.
- We found out that Twitter Roberta generalizes better. This makes sense because during the training of the twitter Roberta model, it has likely seen all kinds of example.
- We also found out that idea of using Bert Unicode also gives a better result compared to processing without.

THANK YOU