COMP6576
MACHINE
LEARNING

Text Emotion Detector

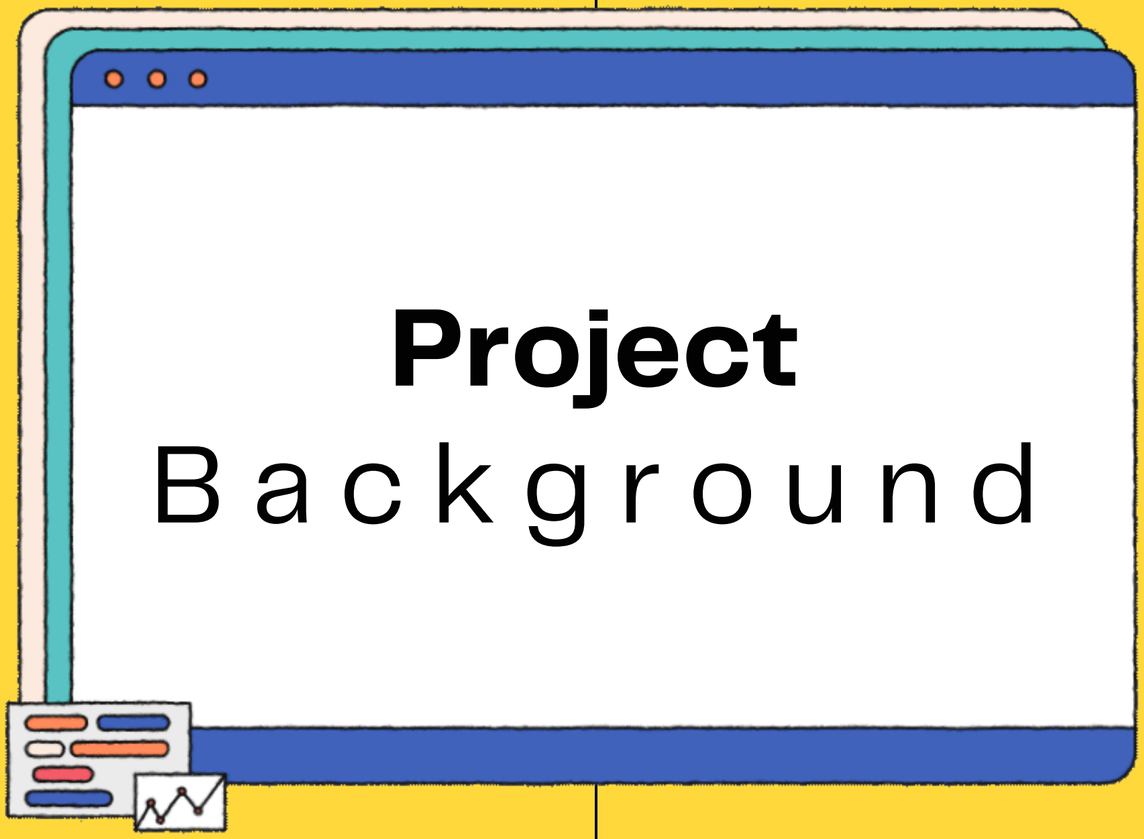
LC01

Adrian Rianto 2301859335
Kelly Cornelya Kesuma 2301920941
Reinaldy Sukamto 2301853136
William Gazali 2301960440

What is the **topic about?**

C h a t t i n g , a common-used technology of this era, has become a substitute for communication media.

C h a t – based application that we commonly use such as **whatsapp**, **line**, **kakaotalk**, etc. Moreover, **non-chat-based** social media also has a **c h a t** feature in it.

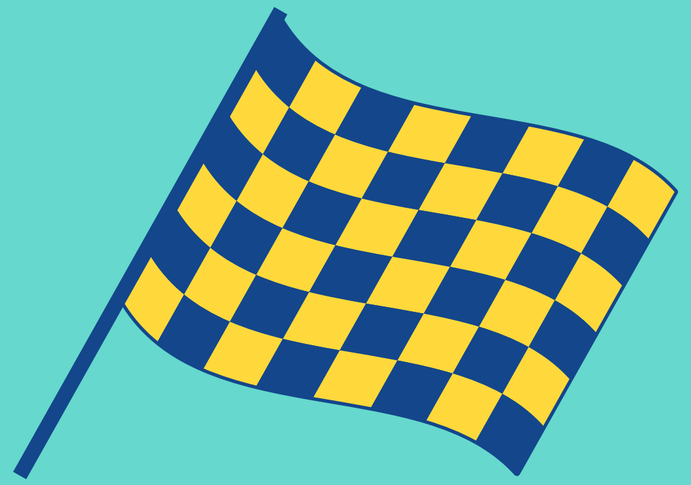


Project
Background

Of the many features and benefits of chatting itself, it is undeniable that chatting also still has various shortcomings. With communication using chat, many things are lost when compared to communicating directly, such as the loss of one's expression and emotions in communicating. This often leads to misunderstandings.

Because of this problem, we have an idea how to bring emotion and expression in chat. We have an idea by classifying a group of chats to conduct training to form a model that can determine a person's emotions and expressions, then present them in the form of emojis or emoticons.

Goals



1. High Accuration
2. Low MSE
3. Accurately Determine Text's Emotion and Expression in a Chat

DATASET

EMOTION DATASET FOR NLP

www.kaggle.com

LITERATURE REVIEW

Rohman, A. N., Utami, E., & Raharjo, S. (2019). Deteksi Emosi Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing. eksplora. stikom-bali. ac. id, doi, 10.

The Methods

PREPROCESSING

1. Data Reading

The first step to developing a new AI model is to read a new data for training purpose

2. Feature Selection

Selecting feature that would be use in classifying text emotions

3. Lemmetizing

Reverting a words back to it's original form within certain context

4. Data Splitting

Splitting data to use for training and testing purpose

5. Data Padding and encoding

Padding and encoding data to make sure all sentences have the same length

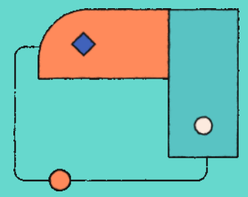
TRAINING

1. Recurent Neural Network

RNN is an algorithm or model that works by using the output of a neuron as input to the neuron itself

2. Multinomial Naive Bayes

Multinomial Naive Bayes algorithm is a probabilistic learning method that is mostly used in Natural Language Processing (NLP).



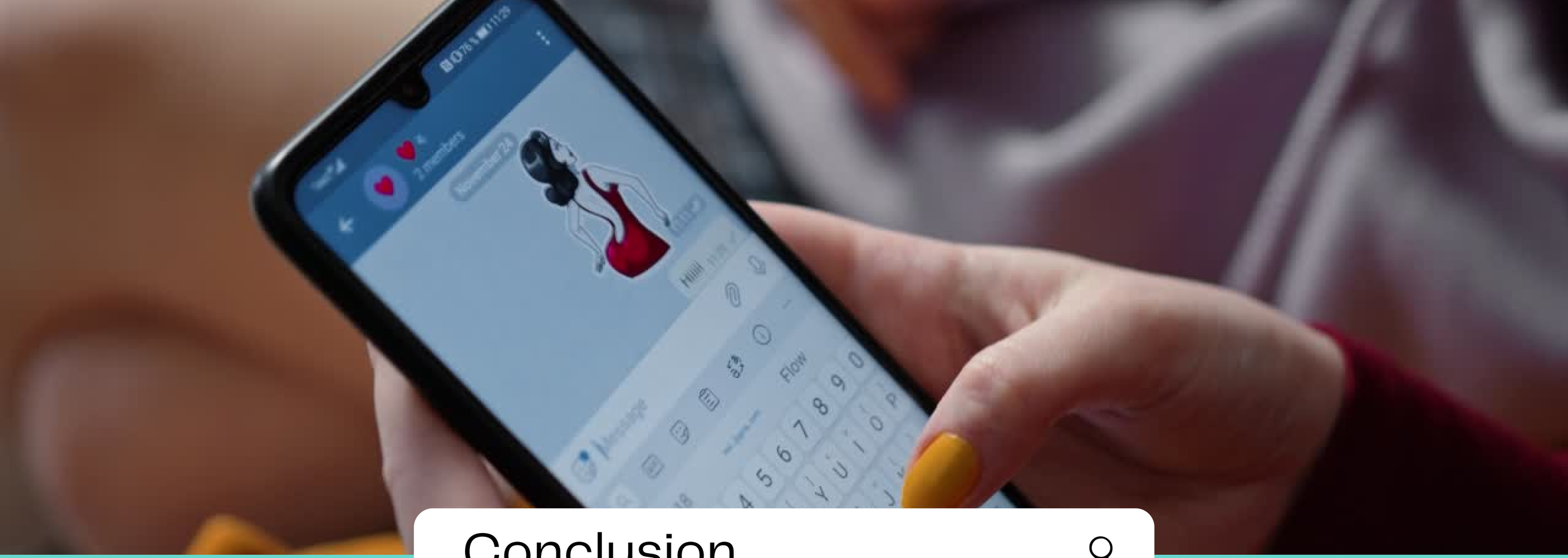
Experiment Result and Analysis

RNN Method

	precision	recall	f1-score	support
0	0.88	0.88	0.88	275
1	0.88	0.85	0.86	224
2	0.93	0.92	0.92	695
3	0.74	0.83	0.78	159
4	0.94	0.94	0.94	581
5	0.65	0.62	0.64	66
accuracy			0.90	2000
macro avg	0.84	0.84	0.84	2000
weighted avg	0.90	0.90	0.90	2000

Multinomial Naive Bayes

	precision	recall	f1-score	support
anger	0.90	0.68	0.77	275
fear	0.83	0.65	0.73	224
joy	0.78	0.95	0.86	695
love	0.87	0.37	0.52	159
sadness	0.78	0.93	0.85	581
surprise	0.64	0.11	0.18	66
accuracy			0.80	2000
macro avg	0.80	0.61	0.65	2000
weighted avg	0.81	0.80	0.78	2000



Conclusion

We tested 2 models against the dataset which is RNN and Multinomial bias and found a fairly general model. But we found that RNN generalise more than the Multinomial bias.

THANKYOU!