# SENTIMENT ANALYSIS FOR ARABIC TWEETS

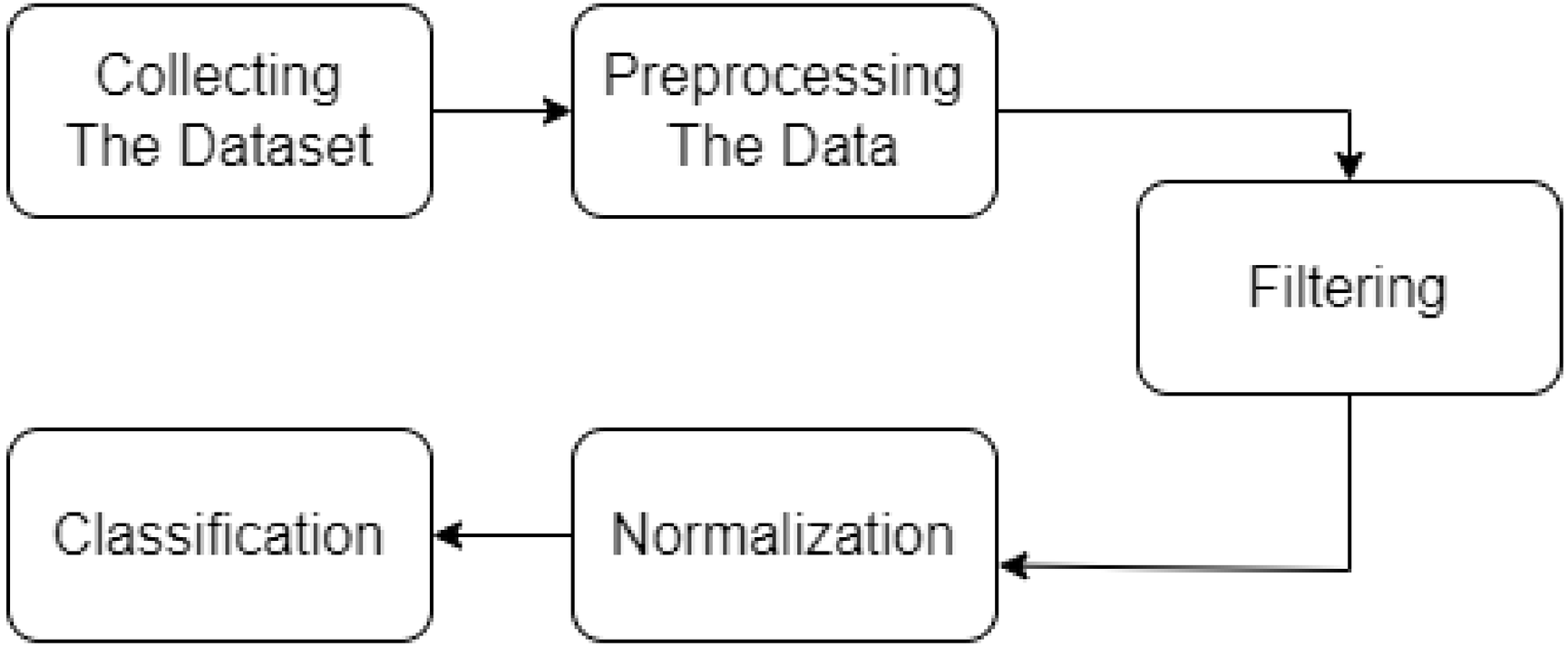
Linah Hossam
Toka Ossama
Nerimane ElRefai

Team 19

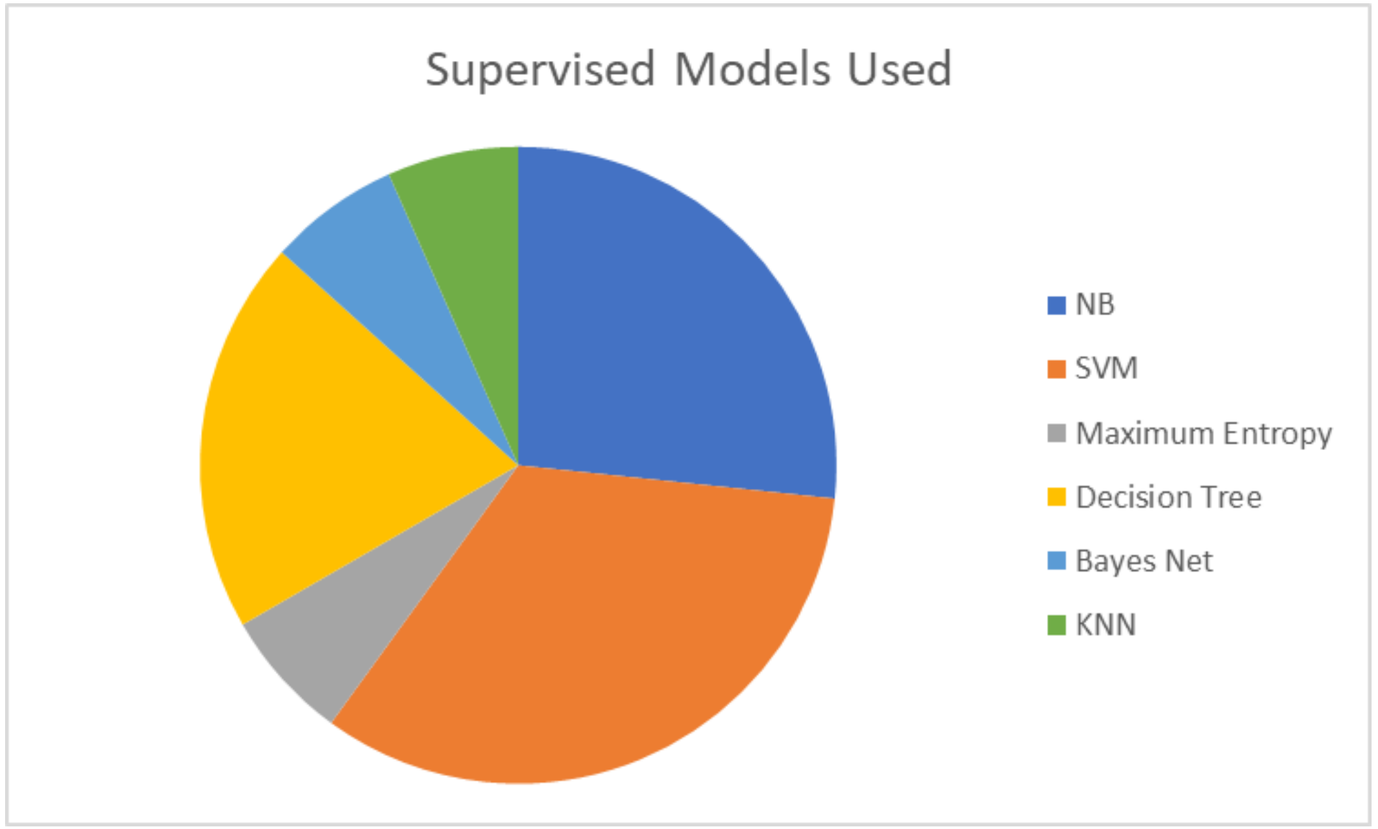
# Outline

- Related Work
- Dataset Overview
- Methodology
  - Logistic Regression
  - Naive Bayes
  - Arabert
- Experiments
- Results
- Dataset Limitations

# Related Work




Sarah O Alhumoud, Mawaheb I Altuwaijri, Tarfa M Albuhairi, and Wejdan M Alohaideb. Survey on arabic sentiment analysis in twitter. International Journal of Computer and Information Engineering, 9(1):364–368, 2015.




[2] Amira Shoukry and Ahmed Rafea. Sentence-level arabic sentiment analysis. In 2012 International Conference on Collaboration Technologies and Systems (CTS), pages 546–550, 2012.
[3] Muhammad Abdul-Mageed, Mona Diab, and Sandra K ̈ubler. Samar: Subjectivity and sentiment analysis for arabic social media. Computer Speech Language, 28:20–37, 01 2014.
[4] Soha Ahmed, M. Pasquier, and Ghassan Qadah. Key issues in conducting sentiment analysis on arabic social media text. pages 72–77, 03 2013.
[5] Janan Ben Salamah and Aymen Elkhlifi. Microblogging opinion mining approach for kuwaiti dialect. international conference of computer technology and information management, 04 2014.
[6] Rehab Duwairi, Raed Marji, Narmeen Sha'ban, and Sally Rushaidat. Sentiment analysis in arabic tweets. pages 1–6, 04 2014.
[7] Salha Al-Osaimi and Khan Badruddin. Role of emotion icons in sentiment classification of ara- bic tweets. MEDES 2014 - 6th International Conference on Management of Emergent Digital EcoSystems, Proceedings, pages 167–171, 09 2014.

None

Joy

Symapthy

Love

8 Labels

Anger

fear

surprise

sadness

# Methodology

## Text Classification

1. Logistic Regression

2. Multinomial Naive Bayes

3. AraBERT

# Logistic Regression

- A common statistical technique for binary classification
- For multi-class classification, scikit-learn’s logistic regression internally uses the ”one  vs-rest” (OvR) strategy.
- This strategy depends on training multiple binary logistic regression models, where each model differentiates between one class and the rest of the classes
- The feature extraction used is TF-IDF

# Multinomial Naive Bayes

- It is a probability algorithm usually used for text classification tasks.
- This classifier is based on Bayes' theorem.
- Multinomial is a specific variant of Naive Bayes .It assumes that the features represent word counts or frequencies.
- The feature extraction used is TF-IDF

# AraBERT

- Arabic pretrained lanaguage model based on Google's BERT architechture

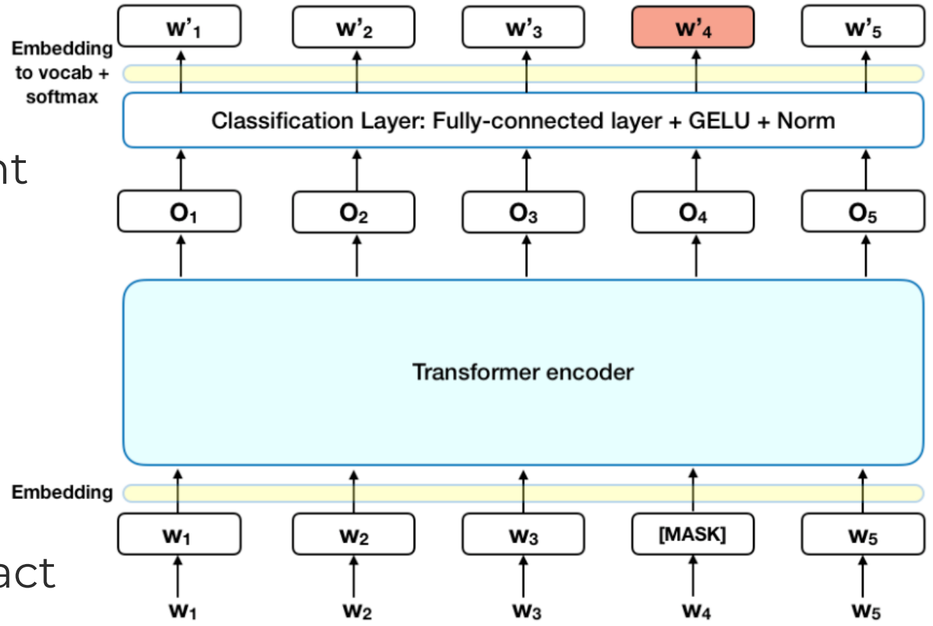
- Used to demonstrate how much of a specific emotion a tweet represents

# Fine Tuning

- **Number of epochs:** 15
- **Learning Rate:** 1e-4
- **batch size:** 32

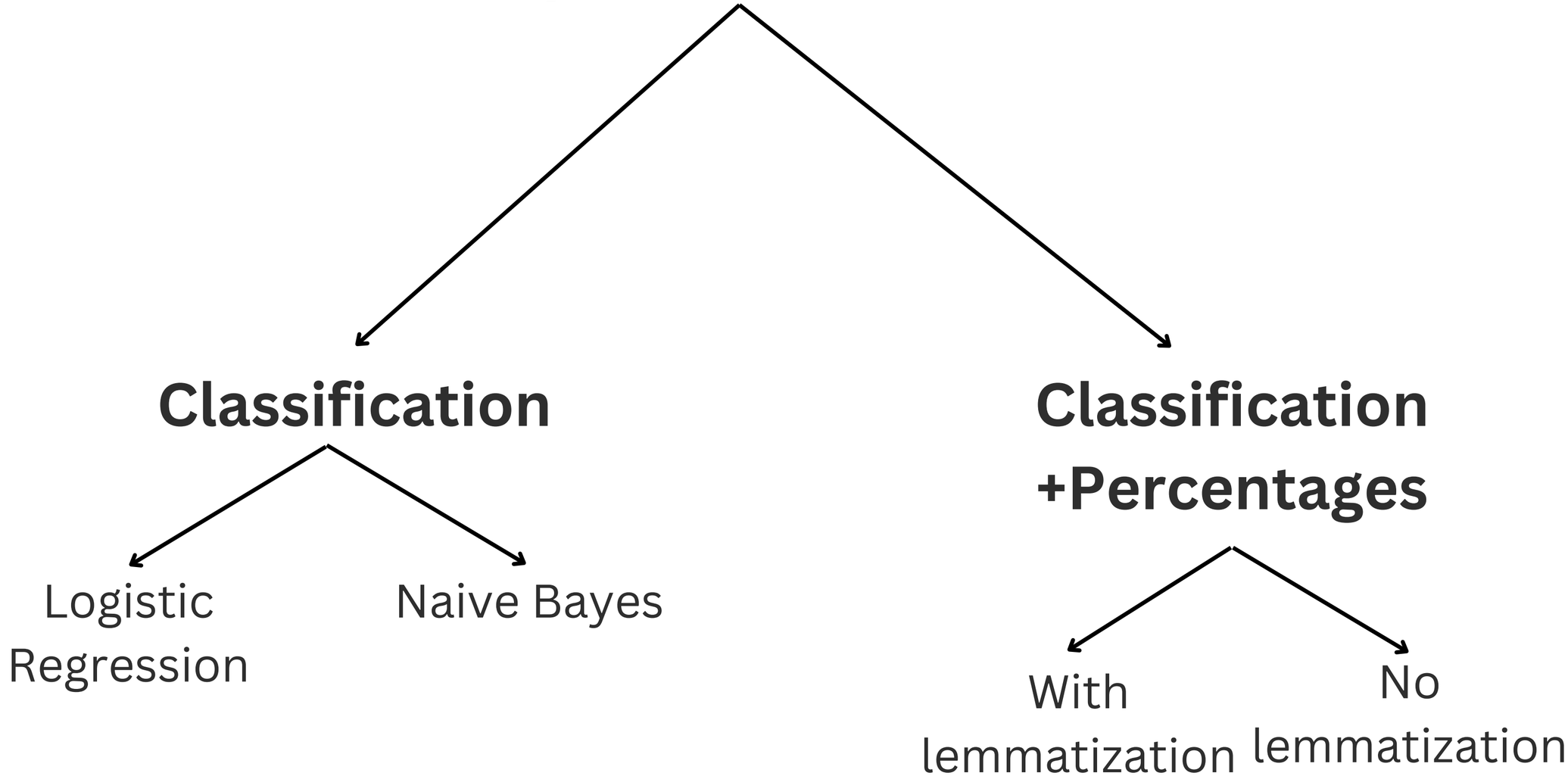
# Architecture

- The tokenized input text is fed to the BERT input layer.
- The transformer encoder is the central component of BERT that captures the contextual representations of the input tokens
- A classification layer is added on top of the BERT encoder to adapt it for this task.
- The linear classification layer transforms the abstract token embeddings to the space of the sentiment vocabulary
- Softmax layer is used to give final probabilities for every class in the label column of the data.




https://www.kaggle.com/code/heyytanay/bert-with-pytorch-lightning-tpu-for-sentiments

# Experiments:



# Results:

Samples of our architeture's output:

```
1  inference(sentiment_module_trained,tokenizer, 'وداعا رافت الهجان وداعا الفنان الاسكندراني الاصيل الله يرحمه' )
```

```
torch.Size([1, 768])
torch.Size([1])

('sadness',
 [('sadness', 0.8231801390647888),
  ('anger', 0.13717156648635864),
  ('surprise', 0.0348108746111393),
  ('none', 0.002486994257196784),
  ('joy', 0.0009141888585872948),
  ('love', 0.0007736752741038799),
  ('fear', 0.0004719136923085898),
  ('sympathy', 0.00019061418424826115)])
```

# Results:

## Logistic Regression

Accuracy score is 0.68

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| anger | 0.62 | 0.67 | 0.64 | 243 |
| fear | 0.98 | 0.86 | 0.92 | 200 |
| joy | 0.58 | 0.51 | 0.55 | 193 |
| love | 0.71 | 0.73 | 0.72 | 193 |
| none | 0.64 | 0.86 | 0.73 | 260 |
| sadness | 0.48 | 0.40 | 0.44 | 196 |
| surprise | 0.58 | 0.47 | 0.52 | 176 |
| sympathy | 0.83 | 0.81 | 0.82 | 192 |
| accuracy | | | 0.68 | 1653 |
| macro avg | 0.68 | 0.67 | 0.67 | 1653 |
| weighted avg | 0.68 | 0.68 | 0.67 | 1653 |

## Naive Bayes

Accuracy score is 0.59

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| anger | 0.65 | 0.59 | 0.62 | 243 |
| fear | 0.84 | 0.83 | 0.84 | 200 |
| joy | 0.57 | 0.34 | 0.42 | 193 |
| love | 0.68 | 0.74 | 0.70 | 193 |
| none | 0.40 | 0.95 | 0.56 | 260 |
| sadness | 0.60 | 0.22 | 0.33 | 196 |
| surprise | 0.74 | 0.18 | 0.29 | 176 |
| sympathy | 0.84 | 0.73 | 0.79 | 192 |
| accuracy | | | 0.59 | 1653 |
| macro avg | 0.67 | 0.57 | 0.57 | 1653 |
| weighted avg | 0.65 | 0.59 | 0.57 | 1653 |

# Results:

| | Lemmatized Data | Non-lemmatized Data |
|---|---|---|
| **Training Accuracy** | 100% | 95% |
| **Validation Accuracy** | 65% | 69% |

# Dataset Limitations

- The Data samples that were labeled as "None".
- Some corrupted data samples
- The emotion faces (emoji)

Thank You