

SENTIMENT ANALYSIS OF ELECTRIC VEHICLE DISCOURSE IN MALAYSIA USING BERT-BASED LANGUAGE MODEL

Presented By: CHANG ZI YIN

Lecturer: Assoc. Prof. Dr Mohd Shahizan Bin Othman



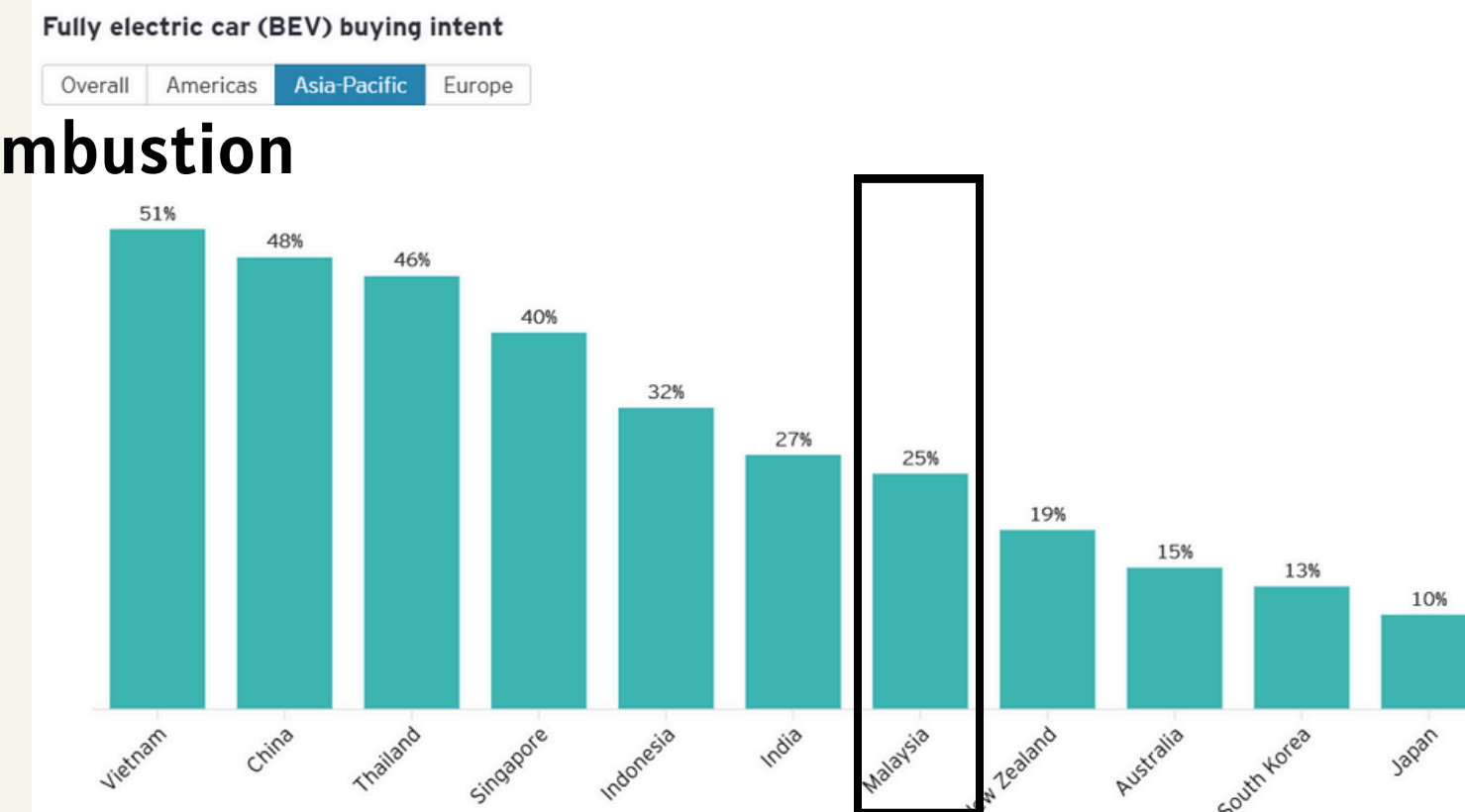
Background of the Study

- The Malaysian automotive market is experiencing a surge in EV adoption
- Malaysia government had incentives to promote the usage of EVs through offering tax exemptions (Asadi et al., 2021)
- Malaysia has aimed for EVs to account for 15% of the total industry volume (TIV) by 2030 and 80% by 2050

Electric Vehicles (EV) continue to show strong growth in Malaysia with **21,789** vehicles registered throughout 2024 according to the latest data shared by JPJ. This marks a **63.81% Year-on-Year (YOY) growth** versus 13,301 units registered in 2023.

According to Ernst & Young's (EY) fifth annual EY Global Mobility Consumer Index

- 25% of Malaysian consumers intent on buying an electric vehicle (EV)
- Nearly 40 % of Malaysians said they would prefer to buy an internal combustion engine, and a hybrid car next (20 per cent)



- Online platform such as social media, forums or video review had become key venue for consumer to share experience, discussing and debating about their concern (Ruan & Qin Lv, 2023)

(EY, 2024)

Problem Background and Statement

- **Lack of Region-Specific Research**
 - Sentiment Research were mostly conducted on global, limited focus on Malaysia's unique market dynamics
- **Limited data diversity,**
 - Reliance on small datasets from single platform, may leading to biased insights
(Sharma, Din, & Ogunleye, 2024)
- **Limitations of Current Sentiment Analysis**
 - Most approaches classify sentiment into simplistic categories (positive, neutral, negative), failing to capture nuanced consumer opinions.
(Wang, You, Ma, Sun, & Wang, 2023)

Research Objective

1

To collect and preprocess electric vehicle related text data from multiple online resources

2

To Implement and compare pre-trained Bert-based model in determining structured sentiment analysis, focusing on aspect-based classification.

3

To analyse the sentiment of Malaysian about the EV

Research Mapping

Research Gap	Problem Statement	Research Question	Research Objective
Most studies focus on global sentiment analyssis of Ev issue, with limited attention to Ev in Malaysia	Lack of localised research hinder on Ev discourse	What are the sentiment analysis of the Ev discourse in Malaysia?	To analyse the sentiment of Malaysian about the EV
Sentiment classification is overly simplistic, often limited to basic categories (positive, neutral, negative).	Simplistic sentiment classification fails to capture the full spectrum of aspect-based sentiment analysis in public discourse about EVs.	How can BERT outperform traditional models?	To Implement and compare pre-trained Bert-based model in determining structured sentiment analysisanalysis, focusing on aspect-based classification.
Limited data diversity due to reliance on small datasets from single platform	The lack of diverse data sources results in potential bias and incomplete insights into EV-related discussions.	How does sentiment vary across platforms?	To collect and preprocess electric vehicle related text data from multiple online resources



Scope of the Project

1

This study will focus exclusively on analysis of discourse within Malaysian social media websites.

2

The web crawler will use to crawl on text published between 2021–2025

3

The analysis only coverage English-language electric vehicle related textual data

4

BERT-based model will be implemented for aspect based analysis, which analysis not just solely on sentiment classification

Literature Review

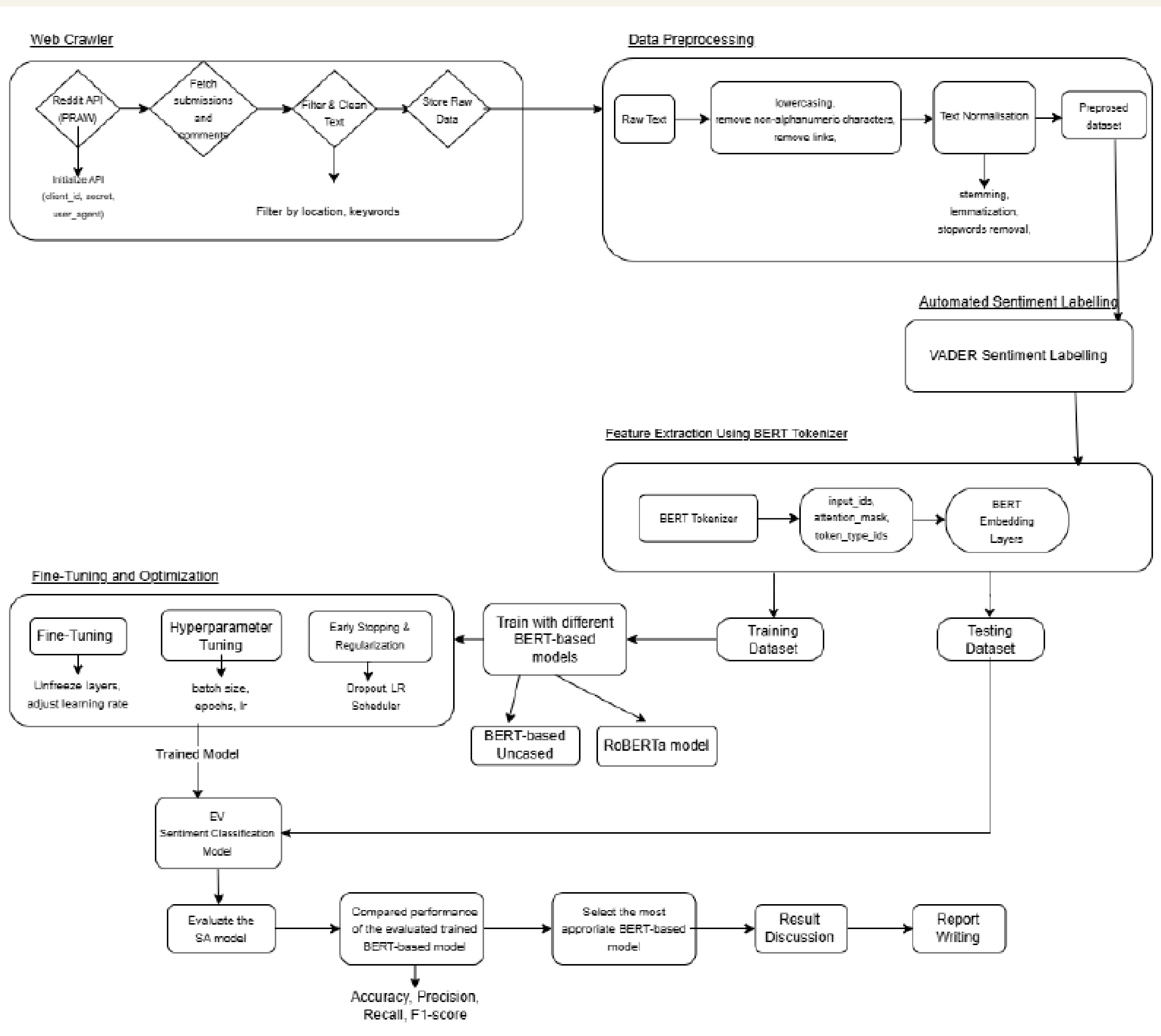
Summary of Model Used

- There are research done on tradisional model such as:
 - Ida topic modeling
 - Sentiment analysis is performed in the NLPIR-Parser platform
 - association rule mining and text mining
 - LSTM model
 - Latent Dirichlet Allocation (LDA)
 - Random Forest
 - combined Enhanced Representation through kNowledge IntEgration (ERNIE)
- and a deep (Convolutional Neural Network) CNN
- There are also adaptation of advance LLM model been used, which is LLM developed by Baidu
- There were also adaptation of deep learning transformer model which is BERT been implemented

Brief summarize:

- Based on the result of the existing research BERT model performance have higher accuracy as compared to the other tradisional machine learning research model
- Eventhough the combined ERNIE and CNN has high accuracy as well but the author suggest to implement advance model fusion method as such simplistic combined classification model cannot capture the full spectrum of aspect-based sentiment analysis in public discourse

Methodology

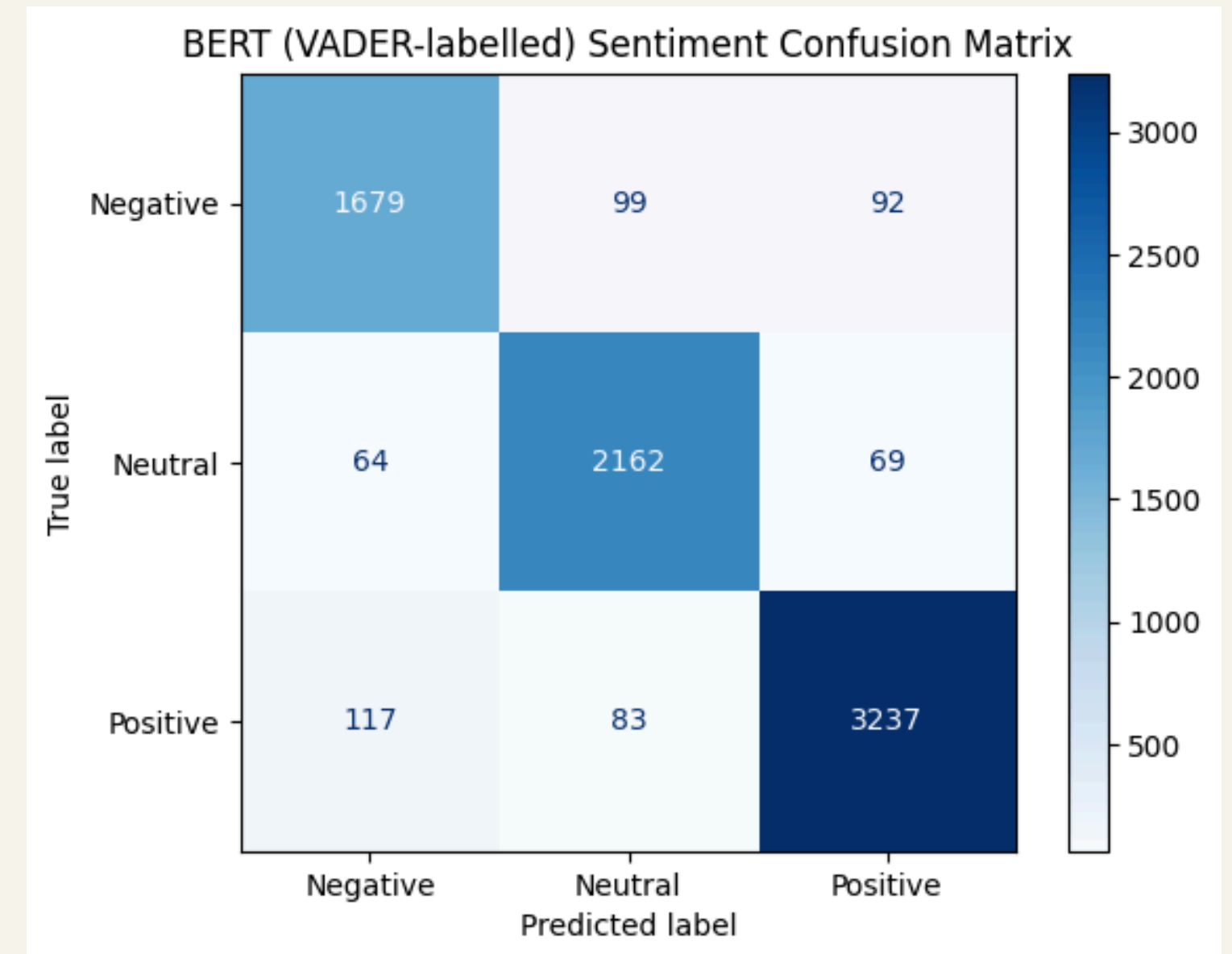


Result

BerT-BASED UNCASED

--- Classification Report ---

	precision	recall	f1-score	support
Negative	0.90	0.90	0.90	1870
Neutral	0.92	0.94	0.93	2295
Positive	0.95	0.94	0.95	3437
accuracy			0.93	7602
macro avg	0.93	0.93	0.93	7602
weighted avg	0.93	0.93	0.93	7602

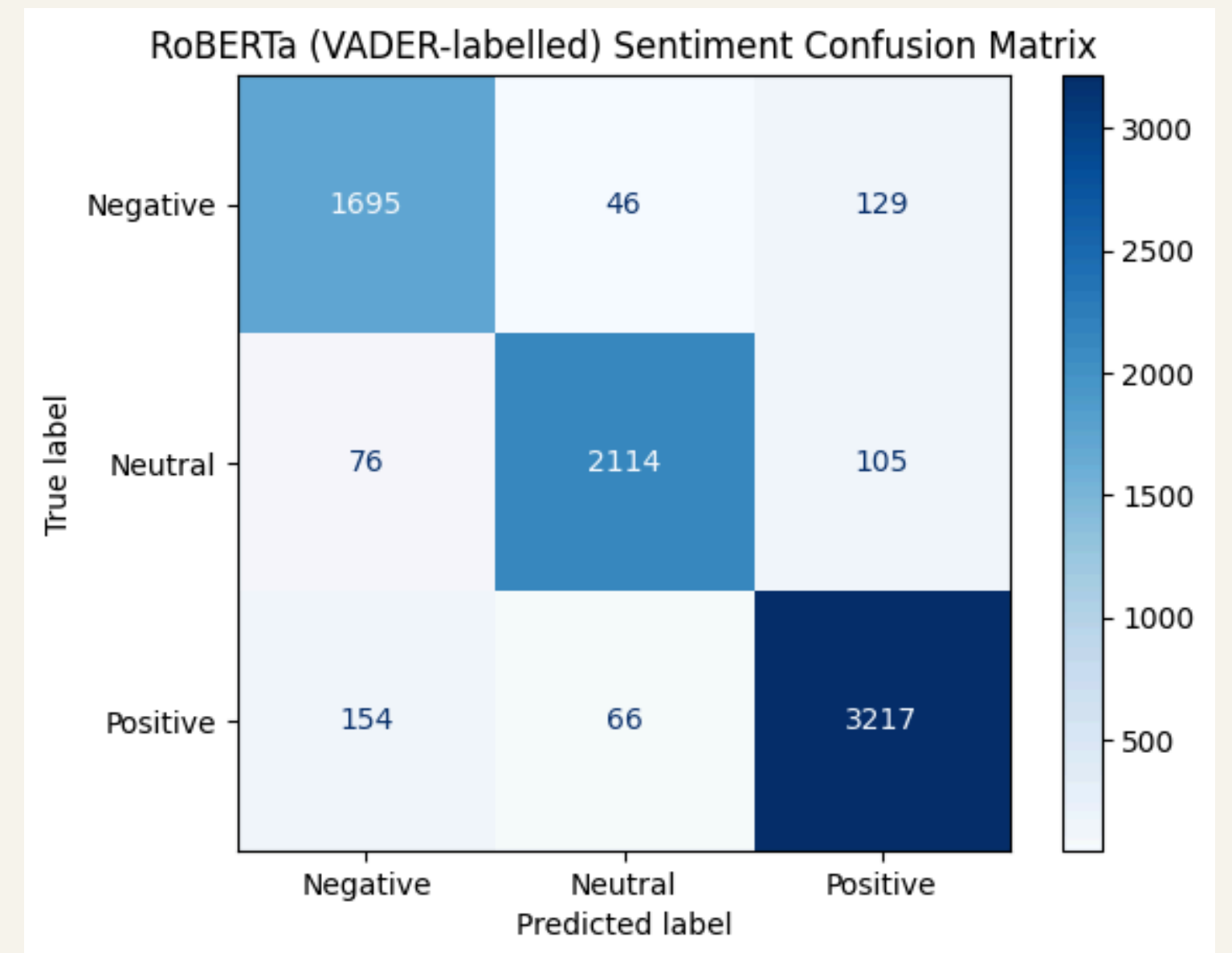


ROBERTA

951/951 [=====] - 65s 59ms/step

--- Classification Report ---

	precision	recall	f1-score	support
Negative	0.88	0.91	0.89	1870
Neutral	0.95	0.92	0.94	2295
Positive	0.93	0.94	0.93	3437
accuracy			0.92	7602
macro avg	0.92	0.92	0.92	7602
weighted avg	0.92	0.92	0.92	7602



Future Work

- 1. data collection on other social media website**
- 2. sentiment labelling using LLM approach**

Thank You

Add a little bit of body text