

# Sentiment analysis classification system using hybrid BERT models

Amira Samy Talaat (2023)

Group Number: 2

21141064 Basit Hussain

RA: Ehsanur Rahman Rhythm

ST: Mehnaz Ara Fazal



# Introduction

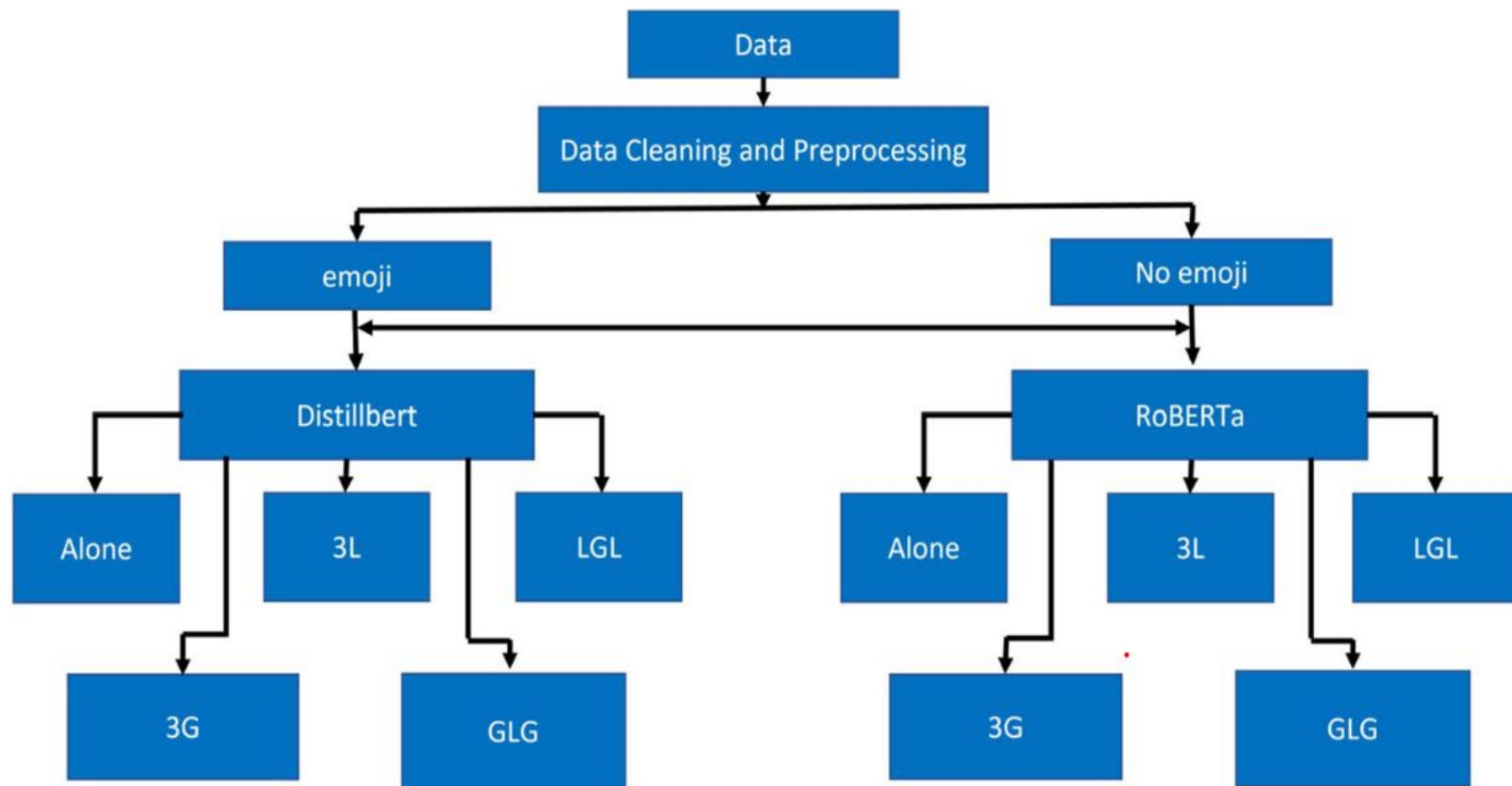
- Highlights the importance of social media in public opinion.
- Addresses the need for more accurate sentiment analysis.
- Proposes four deep learning models that combine BERT with BiLSTM and BiGRU.
- Evaluates model performance on various datasets and compared with pre-trained BERT models
- Propose four innovative deep learning models for emotion classification.
- Compare four RoBERTa and four DistilBERT models to select the best hybrid model.
- Utilize BiGRU and BiLSTM networks to extract context information.
- Evaluate models on emoji datasets to test the impact of emojis on classification.

# Objective

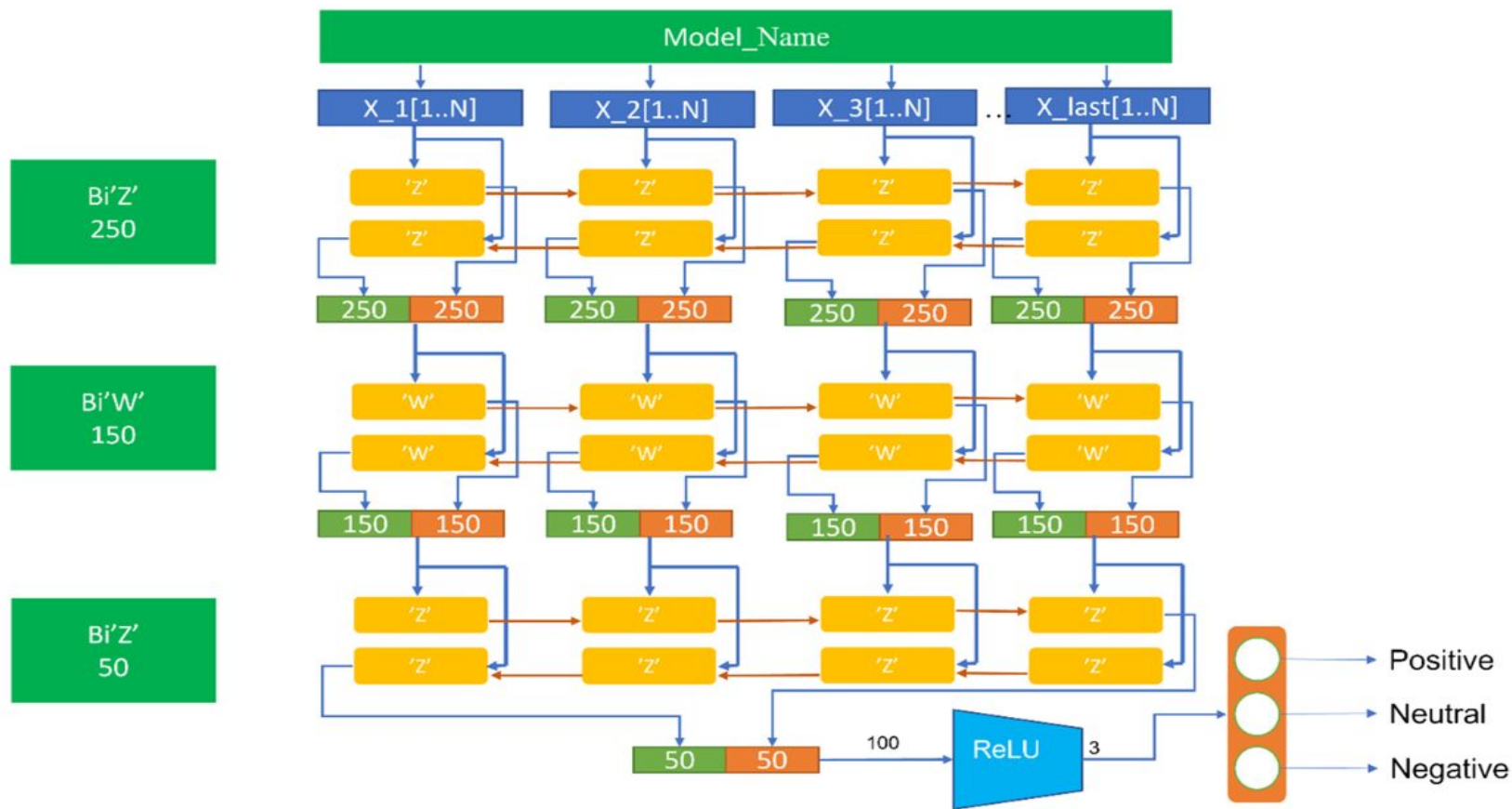
- Emphasize the focus on enhancing sentiment analysis accuracy.
- Highlight the comparison of classical machine learning and hybrid BERT models.

# Methodology

- BERT Models
- Usage of two BERT models (BERT and BERT-mini)
- Datasets
- Transformation of datasets
- Model Development
- Data preprocessing tasks
- Hybrid Models
- Model Training
- Second Round



**Fig. 1** The structure of the proposed framework



**Fig. 2** Hybrid Models Architecture

# Results and Discussion

**Table 2** Testing accuracies of classical methods techniques trained on three datasets

Dataset	Decision Tree	KNN	Logistic Regression	SVM	Naive Bayes	Random Forest	XGBoost
Airlines	66.67	73.16	80.62	80.33	68.56	76.91	76.16
CrowdFlower	63.57	69.18	73.73	73.12	72.15	71.8	71.45
Apple	71.57	78.32	81.6	84.05	79.75	82.41	81.39

**Table 3** Accuracies of airlines, crowdflower, and apple datasets

Model		with emojis			without emojis		
		Accuracy Airlines	Accuracy CrowdFlower	Accuracy Apple	Accuracy Airlines	Accuracy CrowdFlower	Accuracy Apple
DistilBERT	-	83.5	78.71	84.97	83.4	77.92	86.2
	3G	81.8	76.48	83.74	82.17	76.61	87.12
	3L	81.9	74.9	83.13	81.83	76.74	82.82
	GLG	83.74	80.42	85.89	83.47	79.24	88.04
	LGL	82.55	78.71	86.81	83.27	78.98	87.42
RoBERTa	-	85.72	82.39	91.72	85.69	79.63	90.18
	3G	86	79.63	91.1	85.72	79.63	91.72
	3L	85.66	82.26	90.18	85.66	81.34	89.57
	GLG	85.28	81.21	91.41	85.93	80.55	90.18
	LGL	85.52	81.34	89.88	84.97	80.16	90.49

# Conclusion

- The study introduced a framework for predicting user emotions from tweets using BERT models and classical machine learning.
- Three datasets were used
- Hybrid models with BERT and BiGRU layers demonstrated strong performance, especially with large datasets.
- RoBERTa models generally outperformed DistilBERT, achieving higher accuracy.



# Future Work

- Extend this work in the future by combining it with classical text classification algorithms.
- To increase the performance of the present system, the most up-to-date approaches to feature extraction and feature selection will be integrated with traditional methods.

# Limitations

- Dataset Dependency
- Lack of Interpretability

**Thank you!**