**Yorùbá Sentiment Analysis Using Transfer Learning**Department of Computer Engineering  
Elizade University,  
Ilara-Mokin Nigeria

Sotonye Adokiye Dagogo  
sotonye.dagogo@elizadeuniversity.edu.ng

Babatunde Mudashiru   
babatunde.midashiru@elizadeuniversity.edu.ng

Isaac Oluwafemi Elesemoyo  
isaac.elesemoyo@elizadeuniversity.edu.ng

Adetomiwa Omotade  
adetomiwa.omotade@elizadeuniversity.edu.ng

Emperor Olowoleni   
emperor.owololeni@elizadeuniversity.edu.ng

**ABSTRACT**

**Sentiment analysis or detection plays an important role in different fields or sectors by identifying the various types of emotions in different languages and make improved decisions based on that. This research developed sentiment classification model across Yorùbá language. The choice of machine learning algorithm or approach was driven by the dataset characteristics and innovation which is Transfer Learning (TL). Useful pre-trained model was collected from Hugging Face data archives and the Yorùbá dataset was collected from Kaggle data archives. The model was trained and evaluated based on accuracy, precision, recall and F1 score. Results show that the final model achieved an accuracy of 70.04%, precision of 69.64%, recall of 70.04% and F1 score of 69.73%. In summary, these findings suggest that transfer learning can significantly enhance sentiment analysis in Yorùbá language.**

***Keywords- sentiment analysis, transfer learning, algorithm***

**INTRODUCTION**

Sentiment analysis can be regarded as a means of determining if a specific data is positive, negative or neutral. Over the years, it has become one of the most active research fields in natural language processing (Hussein, 2018). In contrast, emotion detection is the process of identifying different types of emotions such as happy, sad or angry in various languages (Nandwani & Verma, 2021). In recent times, people are now making use social media sites such as Facebook, X (formally twitter) or Instagram to express their feelings, opinions, reservations and arguments on different range of topics.

The significance of sentiment analysis in social media extends across various fields or sectors such as businesses, healthcare where medical professionals use it to analyze patient feedbacks to enhance their medical services, media and entertainment, politics and public opinions to understand the public opinions on the current state affairs, political candidates and policies, commercial and promotional purposes (Asefon et al., 2024).

There are wide range of applications of sentiment analysis. For instance, the user preference for various types of commodities can be derived by analyzing product reviews to enable companies or organizations adjust their sales strategies and make decisions. Also, sentiment analysis exhibits a relevant role in public opinion control and emergency detection of event comments (Liu et al., 2019). People can hardly afford the luxury of investing resources in data gathering in today’s world since they are rare, inaccessible, often expensive, and difficult to compile. As a result, most people found a better means of data collection, one of the ways is to transfer knowledge between the tasks by means of transfer learning (Torrey & Shavlik, 2010).

Transfer learning aims at improving the performance of target learners on target domains by transferring the knowledge contained in different but related source domains. In this way, the dependence on a large number of target-domain data can be reduced for constructing target learners. Due to the wide application prospects, transfer learning has become a popular and promising area in machine learning (Zhuang et al., 2020).

This research focuses on Nigeria, a country which is made up of many different languages with an enormous cultural diversity that enables a world of social media exchange using many languages in different dialects at the same time. As a result of the constantly changing digital environment, there has become an increasingly relevant necessity to understand and explain the underlying sentiments and opinions that are involved in the multilingual Nigerian social media comments (Asefon et al., 2024).

By utilizing transfer learning approach, this research intends to classify sentiments effectively, providing insights into sentiment trends in Nigeria’s social media space using Yorùbá language as a case study. It employs a machine learning model through transfer learning on Yorùbá language dataset in which the performance of the model will be analyzed.

**RELATED WORKS**

Many studies have been conducted and which have highlighted the effectiveness of monolingual, bilingual or multilingual sentiment analysis. Machine learning approaches or algorithms play a crucial role throughout numerous efficient sentiment analysis. Many strategies were suggested and checked, the existing model’s weakness and strengths were also identified which led to the justification for the proposed model. Some of them will be reviewed in the brief following.

Research was conducted by (Iyanda & Abegunde, 2020) by comparing various machine learning models such as logistic regression, naïve bayes and support vector machine (SVM) to predict sentiment in Yorùbá written texts and identify the best performing model. The Yorùbá texts used were collected from news outlets and social media. The research showed that SVM had the highest accuracy of 81% but the limitation of this work was that it utilized a small dataset of about 1500 texts.

Similarly, another research was proposed by (Asefon et al., 2024) on “Comparative Analysis of Machine Learning Algorithms for Sentiment Analysis of Multilingual Nigerian Social MediaComments”. The authors leveraged machine learning techniques such as long short-term memory (LSTM), random forest, logistic regression and support vector machine to analyze sentiments in Nigerian languages such Yorùbá, Igbo, Hausa and Nigerian-Pidgin. The advantage of this work was that it covered different languages and was able to illustrate how each of the techniques affect the Nigerian languages but the problem associated with this research was it required building multiple machine learning models from scratch for performance analysis.

Shode et al. (2023) proposed their work titled “NollySenti: Leveraging Transfer Learning and Machine Translation for Nigerian Movie Sentiment Classification”. The authors’ purpose was to develop a labelled dataset for popular Nigerian languages such as Yorùbá, Igbo, Hausa, Nigerian-Pidgin and English by making use of Nollywood movie review to evaluate cross-domain adaptation and cross-lingual adaptation using transfer learning. The experiments were carried out using classical machine learning techniques like support vector machine and naïve bayes, and pre-trained language models. The advantages of this research were that it provided an open-source dataset for future research purposes and also represents the first corpus for Nollywood reviews in 5 Nigerian languages. This research required more computational resources and was time consuming.

While these works highlight the potential of machine learning algorithms for sentiment analysis, they utilized the creation of multiple machine learning models for comparative analysis, more computational resources and time. This project aims to fill these gaps by deploying a transfer learning approach.

**METHODOLOGY**

**Research Design**

The main target of this research is to apply transfer learning to fine-tune the AfriBERTa multilingual language model for sentiment analysis of Yorùbá text, providing a robust solution for emotion detection in African language content.

**Research Design Objectives**

1. To create a transfer learning model for sentiment analysis.

2. To see the effectives of the transfer learning model on the Yorùbá dataset using metrics such as F1 score, accuracy, precision and recall.

3. To analyze and observe the performance of the transfer learning model on the Yorùbá dataset.

**Sources and Nature of Data**

For this research, AfriBERTa large which is a pre-trained multilingual language model with around 126 million parameters was utilized. The model has 10 layers, 6 attention heads, 768 hidden units and 3072 feed forward size. The model was pretrained on 11 African languages namely - Afaan Oromoo (also called Oromo), Amharic, Gahuza (a mixed language containing Kinyarwanda and Kirundi), Hausa, Igbo, Nigerian Pidgin, Somali, Swahili, Tigrinya and Yorùbá.

Also, Nigerian Languages Sentiment Dataset which explores the sentiment of Nigerian languages with this dataset sourced from Twitter. It features tweets in Hausa, Yorùbá, Igbo, and Pidgin, categorized into sentiment labels. The dataset contains 5303 entries split into two columns: Tweets and labels. Others have used this dataset for:

1. Learning 1
2. Research 2
3. Application 0
4. LLM Fine-Tuning 2
5. Licensed: Apache 2.0

Source: Kaggle

The pre-trained model and Nigerian Languages Sentiment Dataset were collected from hugging face and Kaggle respectively. The pre-trained model was finetuned for sentiment analysis of the Yorùbá language dataset in the Nigerian Languages Sentiment Dataset.

**Dataset Overview**

The fig. 1 below displays a portion of the Yorùbá text dataset that was utilized for the research and it was provided in tsv format.



Fig. 1. Yorùbá Text Dataset

**Training and Testing Dataset**

The dataset was categorized into three sections for training, testing and validation. A stratified sampling method ensures that each set represents the distributions of sentiments and languages in the entire dataset.

**Partitioning Strategies**

The dataset was partitioned into:

1. Training (70%)
2. Validation (20%)
3. Testing (10%)

The hyperparameters utilized for the project are illustrated below and the fig. 2 shows the initialization process of these hyperparameters.

1. Learning rates:
   1. Classification head: 5e-4
   2. Fine-tuned layers: 1e-5
2. Weight decay: 0.01
3. Batch size: 16 (with gradient accumulation of 4)
4. Number of epochs: 10
5. Max sequence length: 128 (reduced from 512 to speed up training)
6. Warmup steps: 10% of total training steps
7. Gradient clipping: 1.0

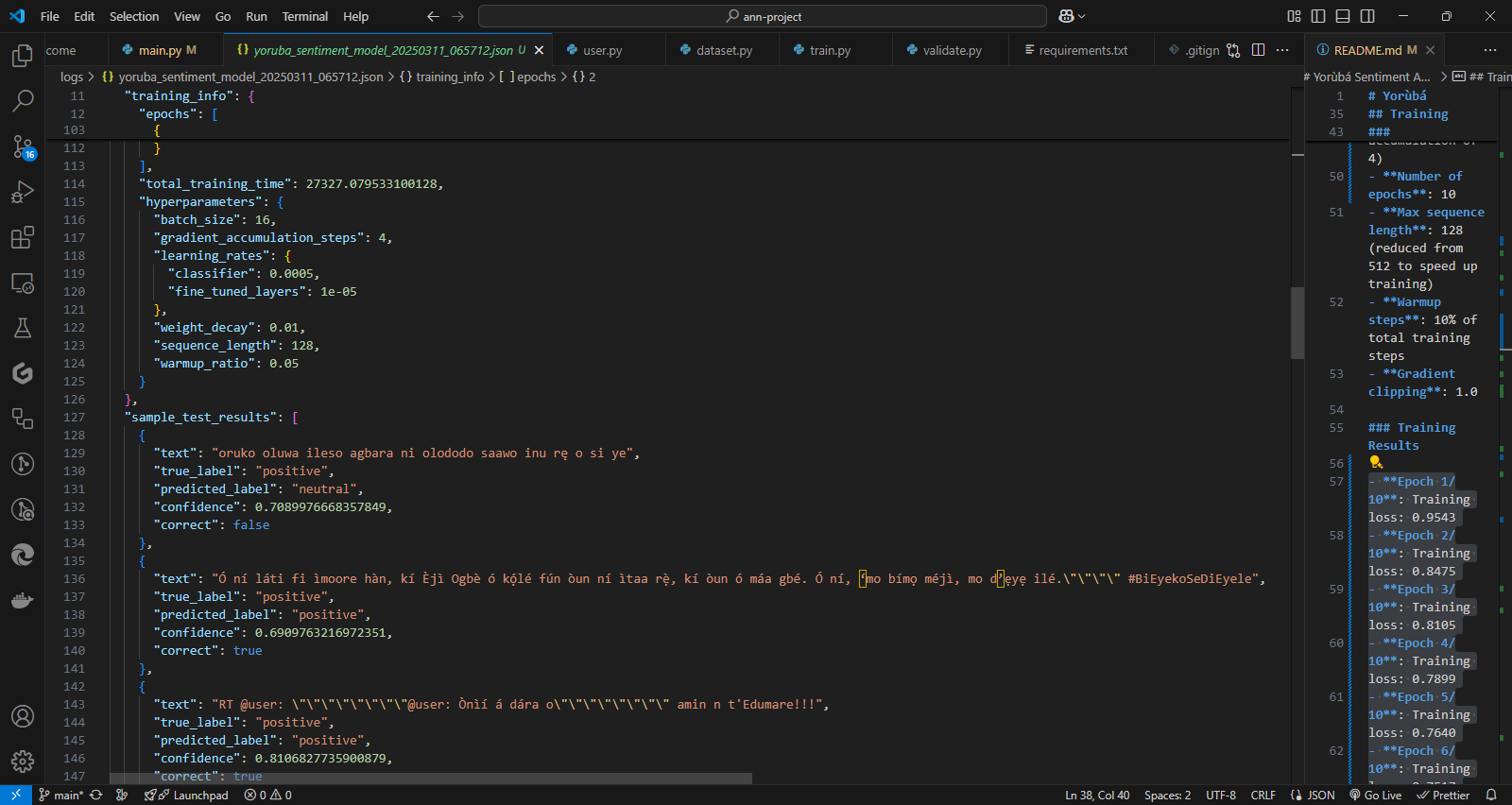


Fig. 2. Hyperparameters Setup

**Performance Metrics**

The performance metric for this project comprises of accuracy, precision, recall and F1 score, the attributes of confusion matrix, used in evaluation of the effectiveness of the model for sentiments classification.

Accuracy can be described as the ability of the model to detect the tone of the statement correctly. This can be calculated as follow:

Accuracy = (1)

The precision metrics can be described as a measure of how the model identifies sentiment or tone.

Precision = (2)

Recall can be described as the proportion of sentiment correctly classified over the total number of sentiment instances.

Recall (3)

F1-score is the harmonic means of both precision and recall.

F1 − score (4)

**Technical Implementation**

1. Programming Languages: Python programming language was utilized because of its comprehensive libraries and frameworks for machine learning and natural language processing.

2. Libraries and Frameworks: Scikit-learn for machine learning, pandas for manipulation and analysis of text dataset, numpy for numerical analysis, tqdm for dataset preprocessing (tokenisation, cleaning), sentencepiece and protobuf for sub word tokenization for neural NLP models and storing large pre-trained model respectively, PyTorch and Transformers for deep learning models.

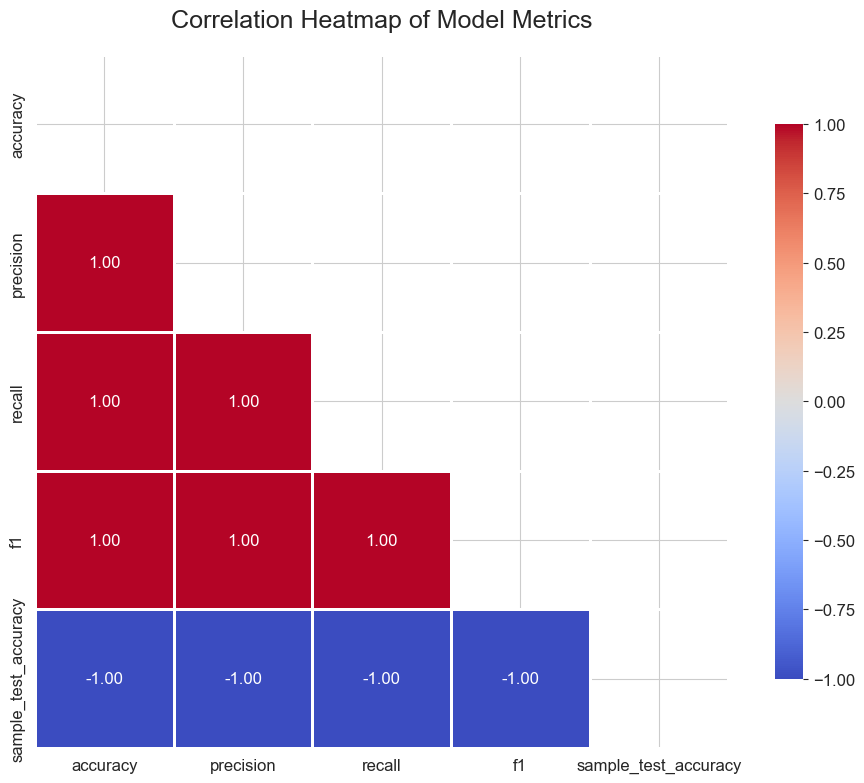
3. Hardware/Software Requirements: To improve the training and evaluation of the transfer learning model, the test was carried out on a system equipped with a powerful GPU.

**RESULTS AND DICUSSIONS**

In this study, two instances of the AfriBERTa-based sentiment classification model were trained to evaluate performance under different training conditions. The first model was trained for 5 epochs using a batch size of 32, while the second was trained for 10 epochs using a batch size of 16. Other than the epochs and batch sizes, the rest of the hyperparameters remained the same. Both instances were evaluated on the same Yorùbá sentiment test dataset.

**Data Correlation and Feature Distribution**

Prior to model training, a correlation heatmap was generated to assess feature relationships in the preprocessed dataset. Although the primary input is textual data tokenized for transformer input, this visualization helped assess label distribution and dataset balance.

****Fig. 3. Correlation Heat-map

**Model Performance Comparison**

After training, both models were evaluated using Accuracy, Precision, and Recall as the main metrics. Table 1 summarizes the performance of both models.

Table 1. Evaluation Metrics for Model 1 and Model 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
| Model 1 | 70.31 | 71.02 | 70.01 | 69.79 |
| Model 2 | 70.04 | 69.64 | 70.04 | 69.73 |

As seen from the table, the first model achieved slightly higher scores across all metrics, indicating that increasing the number of epochs and decreasing the batch size did not lead to performance improvement. Instead, it introduced marginal performance degradation, potentially due to overfitting or batch instability during training.

The following figures further illustrate the comparison of evaluation metrics:

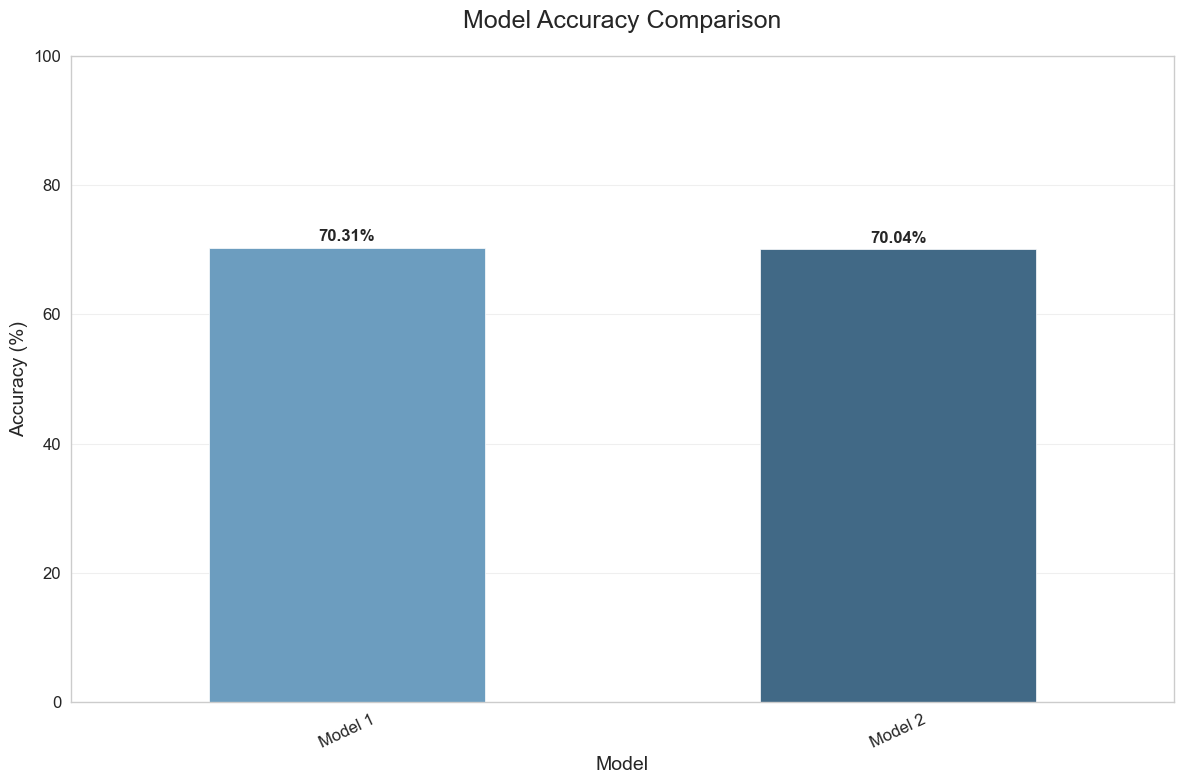
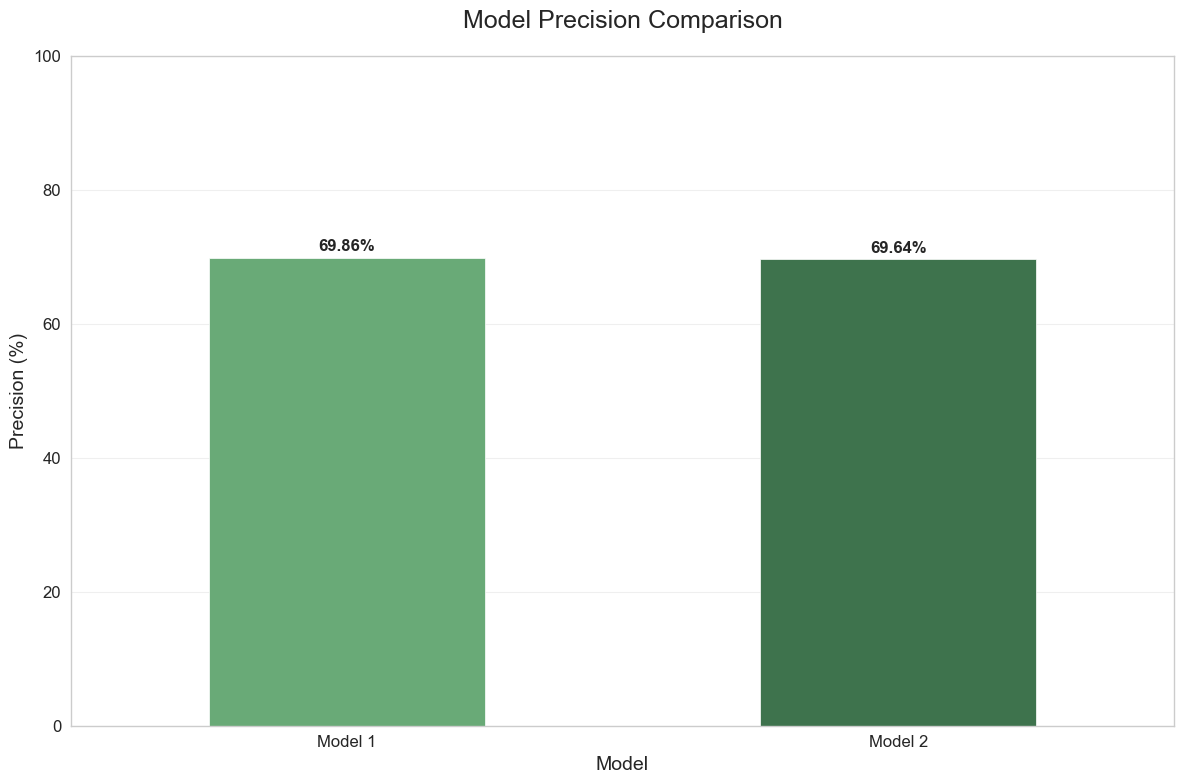
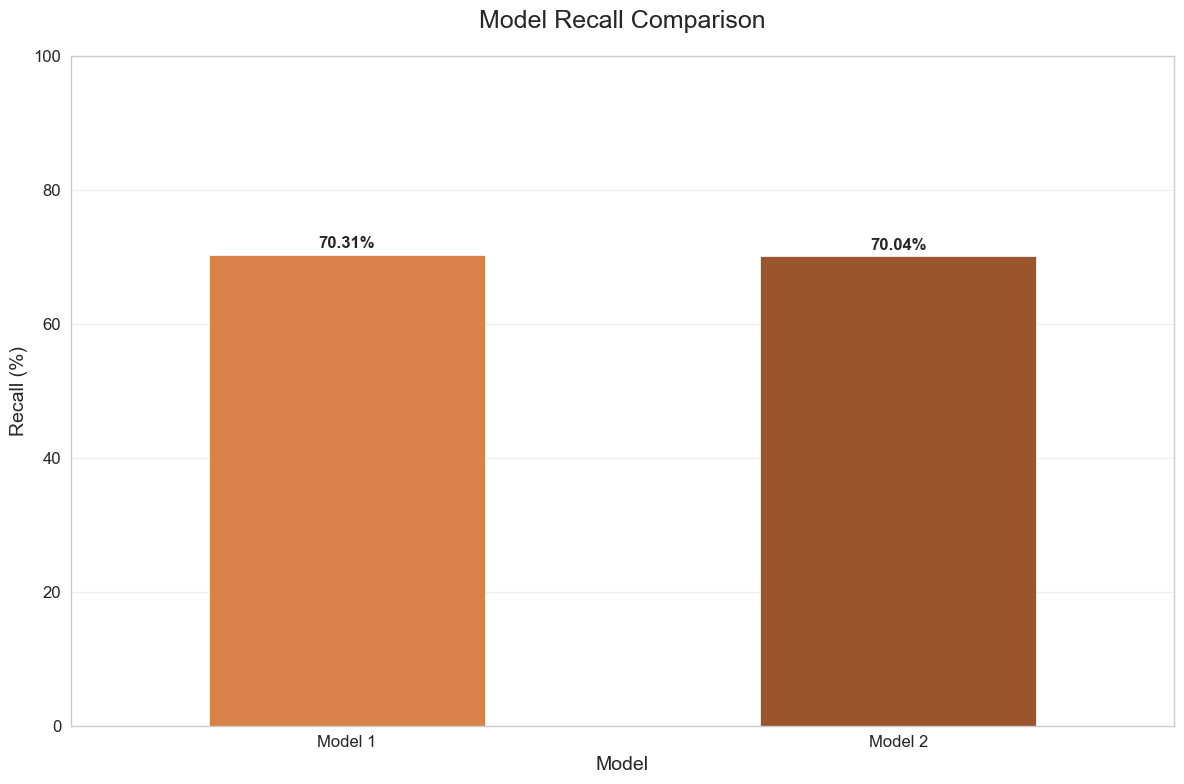


Fig.4. Model Comparison – Accuracy Scores

  
Fig. 5. Model Comparison – Precision Scores

  
Fig. 6. Model Comparison – Recall Scores

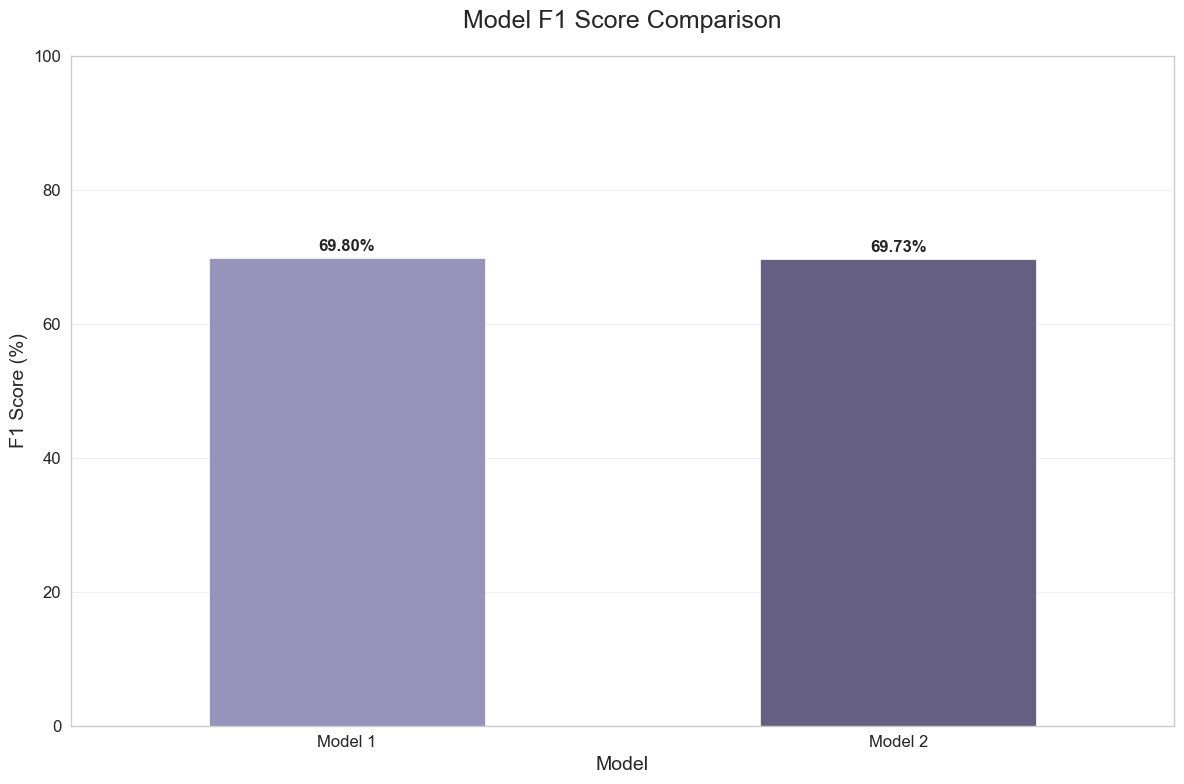
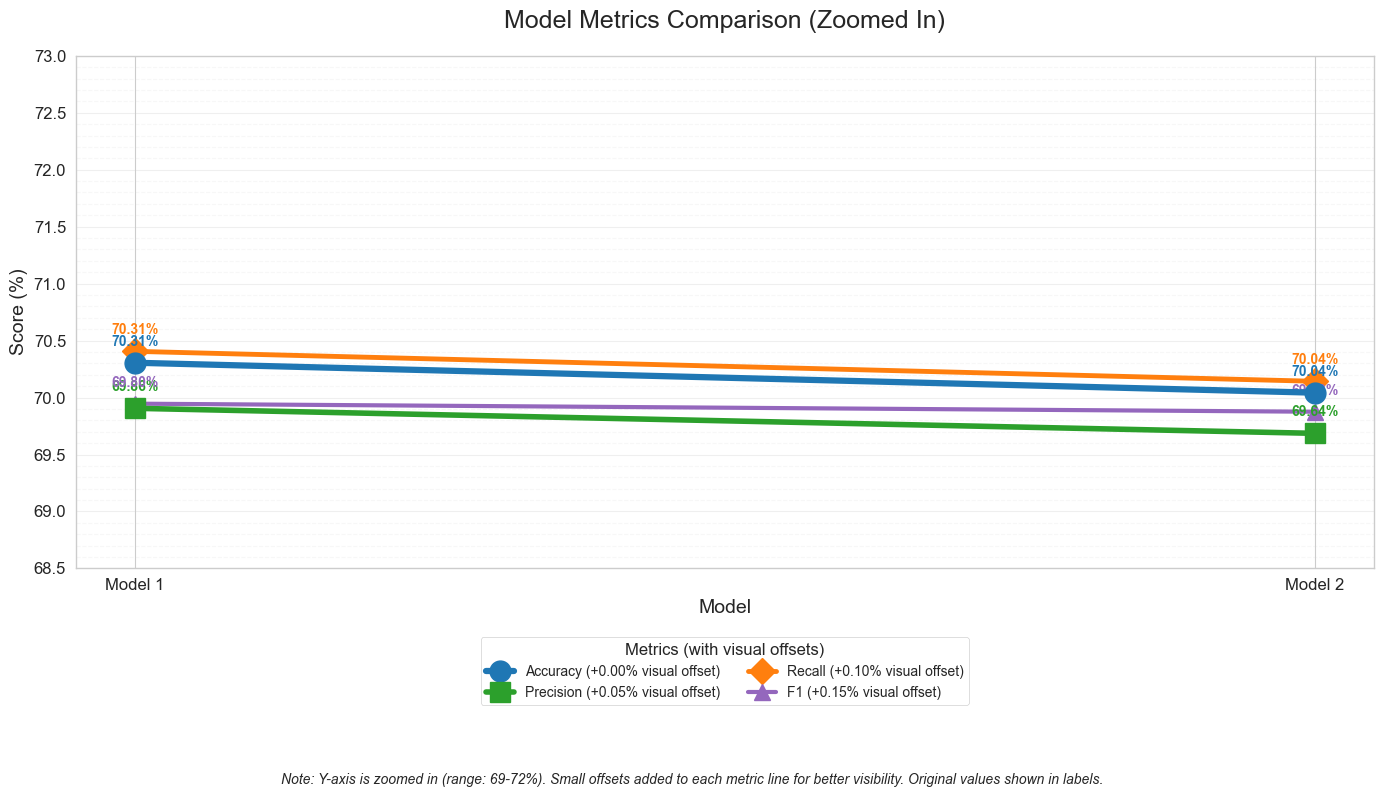


Fig. 7. Model Comparison – F1 Scores

  
Fig. 8. Model Comparison – Accuracy, Precision, Recall and F1

These plots confirm that the differences between the two models are relatively small, but consistently in favour of Model 1. The results suggest that moderate training with an adequate batch size may generalize better for Yorùbá sentiment classification using transfer learning.

**ROC Curve Analysis**

Finally, the ROC curve was plotted to assess the binary classification performance of the models.

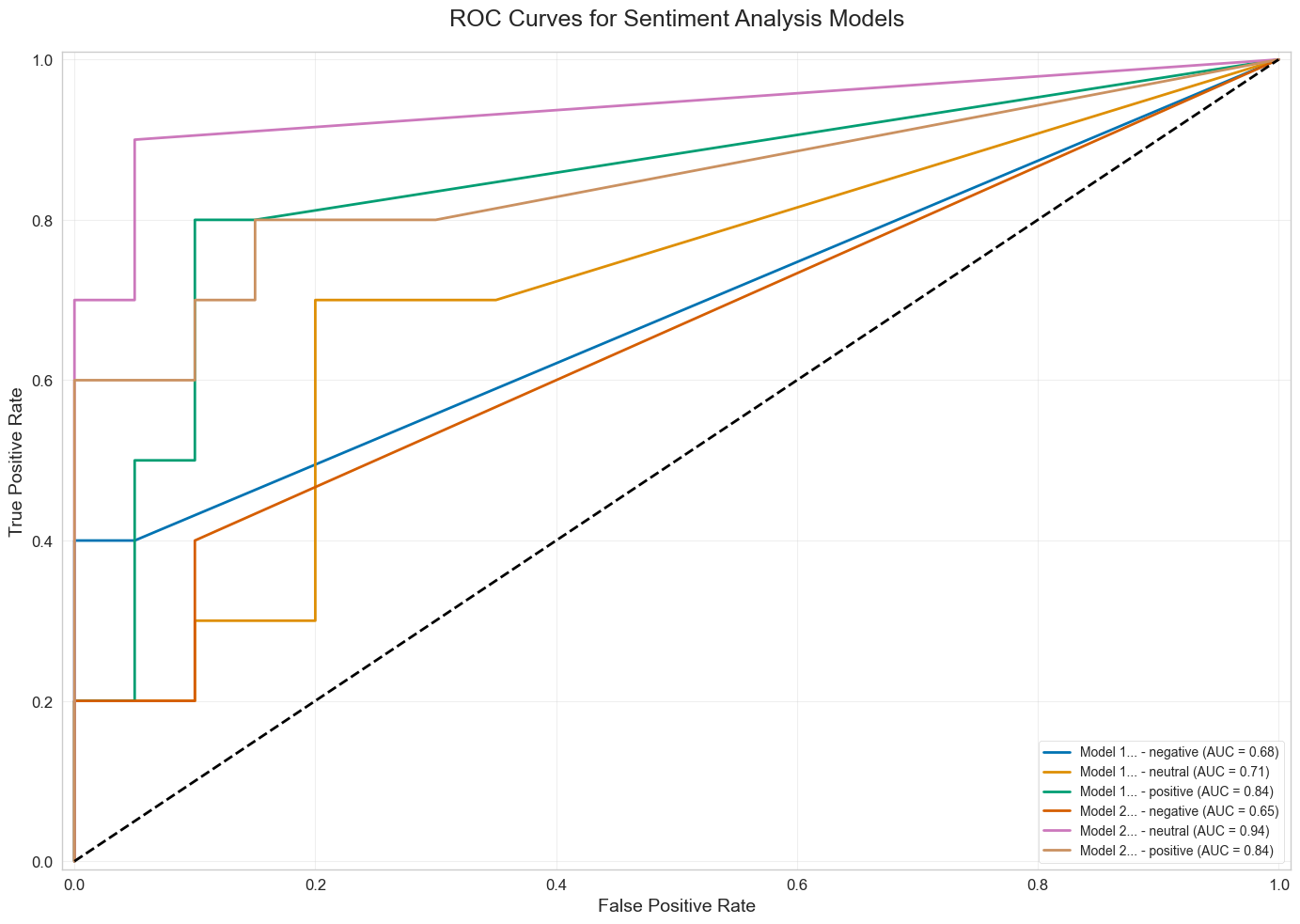


Fig. 9. ROC Intrusion Detection Prediction

The ROC curve shows reasonably strong separation between the classes. This supports the claim that AfriBERTa, even without extensive hyperparameter tuning, can effectively distinguish Yorùbá sentiment tones from textual data.

**CONCLUSION**

This research investigates sentiment analysis in Yorùbá language highlighting the effectiveness of transfer learning approach. Despite, the challenges faced by limited availability of data, the model showed promising result highlighting its potential for real world applications. The results of the model have shown the need for continued research and innovation in this field. The research exhibits valuable insights that can achieved from leveraging computational techniques to discern the huge world of sentiments that are communicated in Nigeria’s activities across different languages on social media platforms.

**REFERENCES**

Iyanda, A. R., & Abegunde, O. (2020). Predicting sentiment in yorùbá written texts: A comparison of machine learning models. In Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys) Volume 2 (pp. 416-431). Springer International Publishing.

Asefon, M. P., Isijola, A. O., Ogude, U. C., & Ntui, S. A. (2024). Comparative Analysis of Machine Learning Algorithms for Sentiment Analysis of Multilingual Nigerian Social Media Comments. FUOYE Journal of Engineering and Technology, 9(4), 615-623.

Shode, I., Adelani, D. I., Peng, J., & Feldman, A. (2023). NollySenti: Leveraging transfer learning and machine translation for Nigerian movie sentiment classification. arXiv preprint arXiv:2305.10971.

Muhammad, S. H., Abdulmumin, I., Ayele, A. A., Ousidhoum, N., Adelani, D. I., Yimam, S. M., ... & Arthur, S. (2023). Afrisenti: A twitter sentiment analysis benchmark for african languages. arXiv preprint arXiv:2302.08956.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43-76.

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. Social network analysis and mining, 11(1), 81.

Liu, R., Shi, Y., Ji, C., & Jia, M. (2019). A survey of sentiment analysis based on transfer learning. IEEE access, 7, 85401-85412.

Hussein, D. M. E.-D. M. (2018). A survey on sentiment analysis challenges. Journal of King Saud University - Engineering Sciences, 30(4), 330–338.

Torrey, L., & Shavlik, J. (2010). Transfer learning. In Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques (pp. 242–264). IGI Global.

Muhammad et al. (2022). NaijaSenti: A Nigerian Twitter Sentiment Corpus.

Ogueji, K. et al. (2021). Exploring Pretrained Multilingual Models for Low-resourced Languages.