BERT-Based Emotion Recognition: A Comprehensive Study in Dari Texts with Explainable AI

1st Malika Muradi
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
malika.muradi@g.bracu.ac.bd

2nd Basit Hussain
Department of Computer Science and Engineering
Brac University
Dhaka, Bangaladesh
basit.hussain@g.bracu.ac.bd

3rd Ehsanur Rahman Rhythm
Department of Computer Science and Engineering
Brac University
Dhaka, Bangladesh
ehsanur.rahman.rhythm@g.bracu.ac.bd

4th Mehnaz Ara Fazal
Department of Computer Science and Engineering
Brac University
Dhaka, Bnagladesh
mehnaz.ara.fazal@g.bracu.ac.bd

5th Annajiat Alim Rasel
Department of Computer Science and Engineering
Brac University
Dhaka, Bnagladesh
annajiat@gmail.com

***Abstract*—This study explores the application of Bidirectional Encoder Representations from Transformers (BERT) models for emotion detection in Dari text. Emotion detection is a complex task, especially when dealing with diverse linguistic contexts. Leveraging the power of BERT model, we conducted analysis of emotion classification in Dari language. The study utilized two pre-trained BERT model: Multilingual BERT a general purpose model and ParseBERT, a model specifically designed for the Dari/Persian language. In this study we Utilize the ArmanEmo dataset, our analysis demonstrates ParsBERT's superior performance across all evaluation metrics, achieving an accuracy of 86.3% compared to Multilingual BERT's 81.2%. This advantage is attributed to ParsBERT's deeper understanding of Dari intricacies and its domain-specific adaptation. Further analysis through Lime XAI reveals the specific words and phrases that ParsBERT relies on for emotion classification, highlighting its focus on key emotion-related terms and expanding relevant expressions.**

***Index Terms*—Emotion Detection, Natural language processing, Multilingual BERT model, Pars BERT model, Dari text, Explainable AI**

I. Introduction

Emotion detection is a crucial field in Natural Language Processing (NLP). It involves determining the emotional tone or sentiment expressed in text data.It is valuable in various applications, including social media analysis, product reviews, customer feedback analysis and so on which understanding the emotional tone of text can help businesses make data-driven decisions. The field of emotion detection has witnessed significant improvement in recent years, with the advancement of NLP techniques in so many languages, specifically English. Though, we witnessed the good improvement of other languages including Dari language in NLP too where the researchers have built and worked on word embedding, sentiment analysis and NER model which are quite good improvements. However we won't find research on emotion detection using the Multilingual Bert model which can be designed to handle multiple languages within a single modal. Dari is a language that have their own scripts and nuances, and our goal is to adapt our emotion detection system to work effectively in these specific linguistic contexts. This includes considerations for script differences, cultural nuances, and local sentiment expressions to ensure accurate and culturally sensitive emotion analysis for Dari.

II. literature Review

This section provides a thorough description of the most widely used approaches for emotion recognition and sentiment analysis in Dari text. These approaches cover a wide range of applications, including market analysis, personality analysis, healthcare, human-computer interaction and many more [1].

In the study [2], the researcher's emphasis is on the crucial role of emotion detection, particularly in addressing the research gap for low-resource languages such as Urdu. The paper introduces the Urdu Nastalique Emo-

tions Dataset (UNED) and presents a deep learning-based technique for emotion classification. Through an evaluation of the UNED corpus, the research showcases the superior performance of the deep learning model compared to generic machine learning approaches. With an impressive F1 score of 85% on sentence-based and 50% on paragraph-based corpora, the paper underscores the significance of UNED and its potential applications in emotion classification for Urdu text.

In another study [3], the researchers demonstrate the challenge of emotion identification in Roman Urdu (RU), a resource-poor oriental language prevalent in Asia. The study introduces a varied RU text dataset annotated with six basic emotions following Paul Ekman's theory. Utilizing dense word embeddings, specifically BERT, the proposed method surpasses baseline algorithms and existing state-of-the-art techniques, achieving an average accuracy of 91% in empirical assessments.

In the other study [4], the researcher examined two different methods to analyze the emotion in Dari text. The first approach, FastText and the Bi-LSTM methods to classify the emotion and the second approach,a deep bidirectional transformer to extract features from text. resulting in notable improvement of multilingual classification.

In the other study [5], the researchers introduce the Urdu sentiment Analysis Bidirectional Encoder Representation from Transformers (USA-BERT) approach for sentiment analysis in Urdu, a low-resource language. Leveraging the UDSA-23 dataset and BERT embedding, the study demonstrates enhanced sentiment classification through re-sampling techniques and deep learning classifiers.

In on other study [6], the researchers introduce two new human-labeled emotion datasets for the Dari language, EmoPars and ArmanEmo. The study evaluate these dataset and addressing the imbalance problem using data augmentation, resampling, and class-weights with Transformer-based Pretrained Language Models (PLMs).

Additionally, feature selection enhance the model performance by finding the specific text features.The proposed model achieves a F1-score of 0.81 and 0.76 on ArmanEmo and EmoPars, establishing new state-of-the-art results in these benchmarks.

In an other research paper [7], Aspect-based sentiment analysis (ABSA) has emerged as a nuanced approach within sentiment analysis, delving into the identification of opinion polarity directed towards specific aspects in a given text. This method has achieved increasing attention within the researcher, owing to its ability to fit more comprehensive and insightful information.The primary objective of the research is to enhance ABSA performance, focusing on the Persian Pars-ABSA dataset. Leveraging the power of pre-trained BERT models and employing a sentence-pair input approach for ABSA tasks.The findings demonstrate that the incorporation of the Pars-BERT pre-trained model, coupled with a natural language inference auxiliary sentence (NLI-M), significantly enhances ABSA accuracy, reaching up to 91%.

Similarly [8], an other research paper have been done on comparing the performance of ParsBERTwith with the multilingual model.The model's performance is evaluated in various natural language processing (NLP) tasks, including sentiment analysis, text classification, and named entity recognition. ParsBERT demonstrates superior performance compared to multilingual BERT and previous models, showcasing its effectiveness in understanding and processing the Persian language.

In the study [9], the researcher addresses the challenges of analyzing emotions in low resource languages, particularly focusing on Dari literary texts. They introduce the JAMFA dataset, which consists of Dari literary sentences labeled with basic emotions. The study compares the use of pre-trained language models, such as XLM-R, and the Catboost algorithm to improve classification accuracy in emotion detection. The findings highlight the significance of deep learning methods and pre-trained models in addressing the lack of trained data in low resource languages, offering valuable insights for emotion detection and sentiment analysis in diverse linguistic contexts.

III. Dataset

A. Data Collection

In this paper we utilize "ArmanEmo", citation(?) a comprehensive manually labeled emotion dataset consisting of over 7,000 Dari sentences categorized into 7 emotion classes. The data collection encompasses diverse source, including comments from tweeter, Instagram, and Digikala, an Iranian e-commerce platform. Labels are assigned based on Ekman's six basic emotions (Anger, Fear, Happiness, Hatred, Sadness, Wonder), with an additional category ("Other") accounting for emotions not covered in Ekman's model.

B. Data Augmentation

To address the imbalanced class distribution in our dataset, we employed the random oversampling technique for data argumentation. This technique involves strategically generating synthetic samples for the minority classes, effectively increasing thire representation and balanced the overall distribution. This approach helps to mitigate potential biases caused by imbalanced data and enhances the model's ability to generalize across diverse emotion categories.

C. Data Preprocessing

Depending on the dataset and model requirements, we applied data cleaning and preprocessing technique such as missing value imputation, stop word removal, and noise reduction to clean the dataset. these comprehensive techniques ensure well-structured, and clean dataset that are now suitable for emotion detection model training and evaluation. To preprocesses our dataset, we employed two

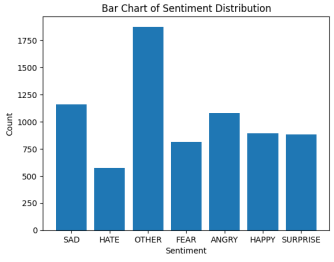

Figure 1. Distribution of dataset befor augmentation

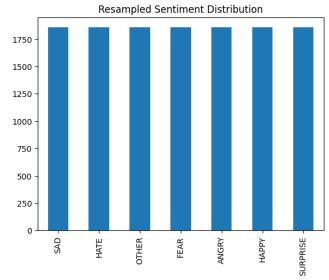

Figure 2. Distribution of dataset after augmentation

tokenization techniques; ParsBERT tokenizer and Multilingual BERT tokenizer. These technique break down text into finely-grained word parts, allowing for a more accurate and nuanced representation of the text compared to generic tokenization methods. All sentences undergo tokenization, and each token is then mapped to its corresponding word ID. To ensure a robust and unbiased dataset for training and evaluating our models, we implemented a representative split of 85% for training data and 15% for testing data. This split ensures that a significant portion of the data is available for training the model to generalize effectively, while reserving a sufficient amount of data for unbiased evaluation of its performance.

IV. Methodology

A. ParsBERT Model

In this research, we utilize the pre-trained Pars-BERT model in the challenging domain of emotion detection within Dari text. ParsBERT is a monolingual language model founded upon the Google BERT architecture. It undergoes pre-training using extensive Persian corpora that encompass diverse writing styles across a multitude of subjects, including scientific, literary, and journalistic domains. This comprehensive training dataset encompasses over 3.9 million documents, comprising 73 million sentences and a voluminous lexical repository of 1.3 billion words. Table I demonstrates the statistics of ParsBERT Model general-domain corpus:

The model architecture comprises an embedding layer that maps input tokens to vectors, followed by 12 bidirectional Transformer blocks. Each Transformer block utilizes a multi-head self-attention mechanism and a fully

Table I
Statistics and types of each source in the corpus

Source	Type	Total Documents
Persian Wikipedia	General (encyclopedia)	1,119,521
BigBang Page	Scientific	135
Chetor	Lifestyle	3,583
Eligasht	Itinerary	9,629
Digikala	Digital magazine	8,645
Ted Talks	General (conversational)	2,475
Books	Novels, storybooks, short stories	13
Miras-Text	News categories	2,835,414

connected feed-forward network to capture long-range dependencies and complex relationships within the text. Notably, the architecture incorporates residual connections and layer normalization to enhance its ability to learn intricate linguistic nuances and navigate the complexities of Dari text.

In the model training process, we utilize a batch size of 16, a learning rate of 2e-5, and a loop for 4 epochs. Throughout training, the model's performance is diligently monitored on a held-out test set, ensuring its effectiveness in discerning emotions within Dari text. After the model training, we evaluate the model and implemented the Lime XAI (Explainable Artificial Intelligence) method to bolster the interpretability of the model's classifications. The LIME XAI method facilitates a more transparent understanding of the decision-making process behind the emotion classifications, enriching the interpretation of the results and providing valuable insights into the model's inner workings.

B. Multilingual BERT Model

For comparative analysis we utilize the Multilingual BERT model, specifically the BERT-base-multilingual-cased variant, to tackle the challenging task of emotion detection in Dari text. This pre-trained model possesses a remarkable ability to comprehend diverse languages. Its training relies on a masked language modeling objective, ensuring its proficiency in processing and understanding the intricacies of textual data across linguistic boundaries. The training loop operates on data batched into sets of 16 and iterates for 4 epochs. During this period, the model undergoes fine-tuning on the provided Dari text dataset, encompassing input IDs, token type IDs, attention masks, and corresponding emotion labels. This iterative process dynamically adjusts the model's internal parameters, progressively refining its ability to discern and classify emotions within the context of Dari text.

V. Results And Discussion

Table II presents the performance evaluation of the Parse BERT and Multilingual BERT for the of emotion detection in Dari text. The evaluation metric includes accuracy, precision, F1 score, and recall.

Table II
STATISTICS AND TYPES OF EACH SOURCE IN THE CORPUS

Model	Accuracy	Precision	F1 Score	Recall
ParsBERT	0.863	0.863	0.863	0.866
Multilingual BERT	0.812	0.806	0.807	0.813

The results demonstrate that the Pras-BERT perform well as compare to Multilingual BERT across all the metrics, achieving a significant performance. This superior performance suggests that Pars BERT model possess a strong capacity to correctly classify emotions in Dari text. The higher performance of Pars BERT attribute to several factors. Firstly, it's pre-training on a large Persian corpus allows it to capture the intricate nuances and specific characteristics of the language, leading to a better understanding of the underlying semantic and emotional cues within Dari text. Secondly, Pars-BERT's architecture is better suited for the task of emotion detection compared to the more general-purpose Multilingual-BERT.

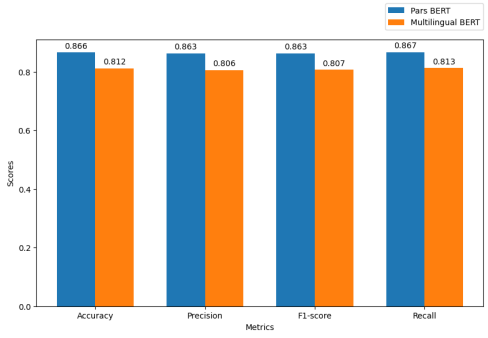

Figure 3. A Comprehensive Performence Comparison between Pars BERT and Multilingual BERT Model

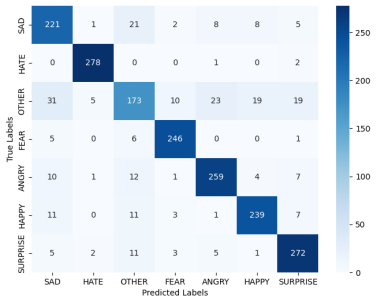

Figure 4. Confusion Matrix for Pars BERT Model

VI. Explainable Artificial Intelligence (XAI)

We employed LIME XAI to illuminate the decision-making process of our model. To enhance clarity, we randomly selected instances of this process, which involves discerning emotions in individual words within a sentence. In the fig 6"لعنت تحریم بانیش چی هست",(Damn, what is the reason for boycotting?), the model accurately detects the

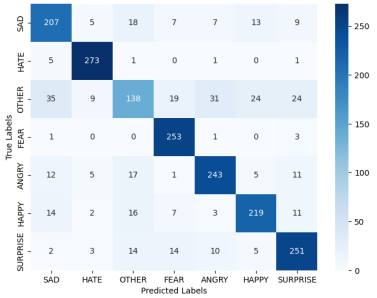

Figure 5. Confusion Matrix for Multilingual BERT Model

emotion expressed by the word "لعنت"(damn). Similarly, in another illustration, Figure 7 "مردم باید آزاد دوست انتخاب کنند حتی نامطلوب سقلمه",(people should be free to choose friends, even when they are undesirable) is correctly identified as other. The model achieves this by discerning each word's emotional content, excluding those that convey no emotion. In a different context, the 8"درود و سلام بر همه هموطنان عزیز ک امروز در راه پیمایی شرکت کردن و چ اون بزرگوارانی ک ب دلایلی نتونستن شرکت کنن خدا پشت وپناه همگی باشه و حلال مشکلات مولا علی یارشون و زیارت کربلا قسمتشون",(Greetings and salutations to all my dear compatriots who participated in the march today and to those honorable people who could not participate for some reason, may God protect them all and solve the problems of Mullah Ali Yaar and their pilgrimage to Karbala.) exemplifies the model's excellent performance in emotion detection. It accurately distinguishes the positive sentiment conveyed by words such as "رودو",(greetings), "سلام" ("salutations"), and"عزیزی" (dear) in attributing the emotion of happiness to the sentence.

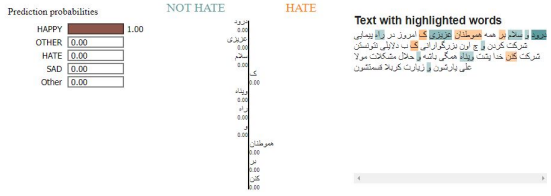

Figure 6.

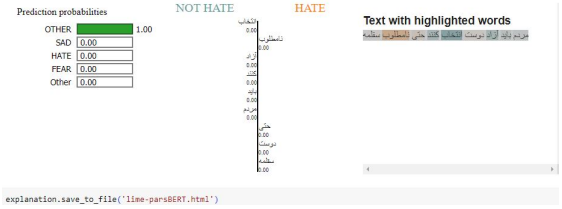

Figure 7.

VII. Conclusion and Future work

In conclusion, this study investigated the application of BERT model for emotion detection in Dari text, highlight-

Figure 8.

ing the superior performance achieved by ParsBERT, a model specifically designed for the Dari/Perian language. Finally, we implement the LIME XAI method for better understanding of the model classification of emotions. Future direction include extending the study to diverse emotion datasets, investigating the impact of alternative data augmentation techniques, implementing additional XAI methods for enhanced interpretability, exploring the potential of combined ParsBERT-Multilingual BERT model, and exploring the application of the ParsBERT to other NLP task relevant to Dari language.

References

[1] S. Kaur and Dr. N. Kulkarni, "Emotion Recognition - A Review," SSRN Electronic Journal, 2020, doi: https://doi.org/10.2139/ssrn.3647958

[2] M. F. Bashir, A. R. Javed, M. U. Arshad, T. R. Gadekallu, W. Shahzad, and M. O. Beg, "Context Aware Emotion Detection from Low Resource Urdu Language using Deep Neural Network," ACM Transactions on Asian and Low-Resource Language Information Processing, Apr. 2022, doi: https://doi.org/10.1145/3528576

[3] N. Ali et al., "Towards Enhanced Identification of Emotion from Resource-Constrained Language through a novel Multilingual BERT Approach," ACM Transactions on Asian and Low-Resource Language Information Processing, Apr. 2023, doi: https://doi.org/10.1145/3592794

[4] Azadeh Khodaei, Azam Bastanfard, Hadi Saboohi et al. Deep Emotion Detection Sentiment Analysis of Persian Literary Text, 05 July 2022, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-1796157/v1]

[5] "BERT-Based Sentiment Analysis for Low-Resourced Languages: A Case Study of Urdu Language | IEEE Journals & Magazine | IEEE Xplore," ieeexplore.ieee.org https://ieeexplore.ieee.org/abstract/document/10271297

[6] Abaskohi, A., Sabri, N., Bahrak, B. (2022). Persian Emotion Detection using ParsBERT and Imbalanced Data Handling Approaches. arXiv preprint arXiv:2211.08029.

[7] H. Jafarian, A. H. Taghavi, A. Javaheri and R. Rawassizadeh, "Exploiting BERT to Improve Aspect-Based Sentiment Analysis Performance on Persian Language," 2021 7th International Conference on Web Research (ICWR), Tehran, Iran, 2021, pp. 5-8, doi: 10.1109/ICWR51868.2021.9443131.

[8] Farahani, M., Gharachorloo, M., Farahani, M. et al. ParsBERT: Transformer-based Model for Persian Language Understanding. Neural Process Lett 53, 3831–3847 (2021). https://doi.org/10.1007/s11063-021-10528-4

[9] Azadeh Khodaei, Azam Bastanfard, Hadi Saboohi et al. Deep Emotion Detection Sentiment Analysis of Persian Literary Text, 05 July 2022, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-1796157/v1]