# A Comparative Study of ParsBERT and mBERT in Emotion Recognition for Dari-Farsi Text with Explainable AI


Malika Muradi
Brac University
Dhaka, Bangladesh
malika.muradi@g.bracu.ac.bd

Basit Hussain
Brac University
Dhaka, Bangladesh
basit.hussain@g.bracu.ac.bd

Ehsanur Rahman Rhythm
Brac University
Dhaka, Bangladesh
ehsanur.rahman.rhythm@g.bracu.ac.bd

Annajiat Alim Rasel
Brac University
Dhaka, Bangladesh
annajiat@gmail.com



## ABSTRACT

Even though emotion analysis is a popular research field, most of the studies have been conducted in English, and the number of those considering Dari is very constrained. Dari and Farsi are two different dialect of Persian language. It has limited resources for natural language processing (NLP), posing a major challenge for Dari NLP research. To overcome these challenges, this study explores the application of Bidirectional Encoder Representations from Transformers (BERT) models for emotion detection in the Dari text. Taking advantage of the power of the BERT model, we analyzed emotion classification in Dari. The study used two pre-trained BERT models: Multilingual BERT, a general-purpose model, and ParsBERT, a model specifically designed for the Dari language. In this study, we utilized the ArmanEmo dataset, and 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 utilized two most renowned Explainable AI (XAI) methods: Local Interpretable Model-agnostic Explanations (LIME), and Shapley Additive exPlanations (SHAP). These methods reveal the specific words and phrases that ParsBERT relies on to classify emotions, highlighting its focus on key emotion-related terms and expanding relevant expressions.



## CCS CONCEPTS

• **Computing methodologies** → **Natural language processing**.

## KEYWORDS

Emotion Detection, Natural Language Processing, Multilingual BERT Model, Pars BERT Model, Dari Text, Explainable AI




*Woodstock '18, June 03–05, 2018, Woodstock, NY*



ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

**ACM Reference Format:**
Malika Muradi, Basit Hussain, Ehsanur Rahman Rhythm, and Annajiat Alim Rasel. 2018. A Comparative Study of ParsBERT and mBERT in Emotion Recognition for Dari-Farsi Text with Explainable AI. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/1122445.1122456


## 1 INTRODUCTION

Emotion classification, an important aspect of NLP, has recently received significant attention for its diverse applications, including social media analysis, product reviews, and customer feedback analysis. The ability to understand the emotional tone of text can empower businesses to make data-driven decisions. Although much study has been made in emotion recognition, particularly in English, there remains a gap in research focusing on languages such as Dari. Dari and Farsi are two dialects of the Persian language. While they share a common linguistic root. Dari is predominantly spoken in Afghanistan, whereas Farsi is the primary dialect in Iran [16] native speaker. Understanding and accommodating these differences is crucial for effective emotion recognition in these languages. Expanding research in this area could significantly improve NLP applications for a substantial global population. The real-world applications of emotion detection in Dari are extensive and impactful as many other languages. For instance, businesses can enhance customer service by analyzing feedback to highlight areas needing improvement. Additionally, detecting the emotions allows the companies to tailor their campaigns more effectively through more optimized marketing strategies. In public health, monitoring emotional well-being through text communications can provide valuable insights into community mental health trends. These all require an advanced emotion detection model that ensures adaptability and effectiveness across different sectors.

In recent years, NLP has seen remarkable advancements, including in languages beyond English. Researchers have developed and refined word embedding, sentiment classification, and Named Entity Recognition (NER) models for languages like Dari. However, there is a notable absence of research utilizing the Multilingual BERT model for emotion detection in Dari texts. The Multilingual BERT model is particularly intriguing as it has the potential to handle multiple languages within a single framework, which could significantly benefit languages like Dari with its unique scripts and linguistic nuances.

Adapting an emotion detection system for Dari involves overcoming several challenges, including differences in script, cultural nuances, and local emotion expressions. These challenges play a key role in ensuring the accuracy and cultural sensitivity of sentiment classification. By leveraging the multilingual BERT model, researchers can more effectively explore and understand the emotional content of Dari texts, opening the door to a deeper understanding of language's emotional nuances and expressions.

The proposed study aims to fill this gap by exploring the application of the multilingual BERT model for emotion detection in Dari texts. This study will not only contribute to the field of Dari NLP but also provide valuable insights into the emotional dynamics of the Dari language, paving the way for more culturally sensitive and accurate emotional analysis in diverse linguistic contexts.

## 2 LITERATURE REVIEW

In this section, a comprehensive overview is presented of the predominant methods and relevant research in emotion recognition and sentiment classification in Dari text. These approaches cover a wide range of applications, including market analysis, personality analysis, healthcare, human-computer interaction and many more.

In this study [21], the author explores sentiment classification as a method to extract meaningful information from social media comments in the Persian language. The study focuses on using the ParsBERT model to classify sentiments in informal comments across different domains. By fine-tuning ParsBERT on a corpus of 28,710 Instagram comments labeled as positive or negative, the model achieves an accuracy of 68% when tested on different domains. This outperforms other methodologies, highlighting ParsBERT's effectiveness in analyzing sentiments in diverse social media comments.

In another study [2], the author discusses the difficulties of analyzing Persian text due to the prevalence of the internet and social media. They propose a method to enhance sentiment classification in Persian by using a specific model and lexicon. Their study focuses on analyzing user opinions from the 'Digikala' website, showing that their method performs well, with 88.2% accuracy and a 61.7% F1 score. Improving language models is crucial for understanding subtle sentiments in user content, leading to better efficiency and accuracy in sentiment classification for Persian text mining.

In this study [27], the authors focus on semantic similarity, a crucial aspect of natural language processing, which assesses the meaning-based similarity between two texts. They leverage the ParsBERT language model to convert tokenized sentences into 768-dimensional vectors, extracting optimized vectors through pooling layers. Using cosine similarity, they calculate the similarity between these vectors. Results show the proposed model outperforms previous models, achieving a Pearson correlation coefficient of approximately 0.82, compared to the best previous result of around 0.77. This demonstrates the superior performance of the ParsBERT-based model in semantic similarity assessment. Similarly, in another study [26] the author employ an optimized BERT-BiLSTM model, which demonstrates superior performance compared to traditional models like FastText-BiLSTM and CNN-BERT. The optimized model achieves an accuracy of 88% in binary classification tasks, showcasing its effectiveness in understanding nuanced emotional expressions in Persian literature.

In another study [24], researchers focus on the valuable nature of patent documents in research and development. They emphasize the crucial role of patent classification, including the assignment of International Patent Classification (IPC) codes, which is often done manually. The Iran Patent Office, for instance, uses IPC codes for classification. The researchers introduce a public dataset for Persian patents and examine how pre-trained Transformer models, like ParsBERT, can automate the assignment of IPC codes. They also note the growing trend of using artificial intelligence models for this task in major patent offices such as USPTO and WIPO.

likewise, in another study [1], researchers are addressing challenges in recognizing Persian named entities (NER) across different fields by proposing a new method. Their approach involves a main model with several sets of domain-specific parameters, utilizing techniques like prompt tuning and adapters. This method allows the model to perform as well as individual models for each domain, outperforming existing models. The new model needs only one instance for training and storage but achieves excellent results across all domains, sometimes even surpassing the current best models. Furthermore, the researchers introduce a document-based domain detection pipeline, which improves the adaptability of their method in practical scenarios.

Similarly, in [5] the authors discuss the rapid growth of social networks, particularly Twitter, as a platform for users to express their opinions. They propose a sentiment classification approach for Persian political tweets, aiming to help Iranian politicians. They examine two datasets of Persian political tweets with 3 and 7 classes, investigating various encoding methods such as Bag-of-Words, Word Embeddings, and neural techniques like Word2Vec, FastText, and ParsBERT Embeddings. The research applies sentiment classification in a specific domain of political tweets using ML methods like Random Forests, SVM, and Neural Networks. Their comparisons indicate that CNN+BiLSTM with ParsBERT embeddings demonstrates greater robustness compared to other networks, achieving a score of 89% on Dataset 1 and 71% on Dataset 2.

Furthermore, the study [4] emphasize the importance of social media as a platform for expressing opinions, especially for businesses looking to enhance their products and services. They observe a rising trend in the use of deep learning methods for sentiment classification due to their effectiveness. They introduce a novel deep architecture named DeepSentiParsBERT, which merges ParsBERT and Bidirectional LSTM models, for sentiment classification of Persian texts. Comparative results with state-of-the-art models indicate DeepSentiParsBERT's superior performance on the Digikala corpus, achieving a 91.57% F1-Score.

Furthermore, the paper [7] discusses the impact of pre-trained language models on NLP and introduces ParsBERT, a BERT model tailored for the Persian language. ParsBERT outperforms multilingual models and prior works in tasks like sentiment classification and NER, showcasing its state-of-the-art performance. The paper also presents a sizable dataset for Persian NLP tasks, addressing the challenge of limited data availability in Persian.

In another study [22], researchers gathered more than 12,000 Persian tweets discussing the stock market, which were then manually categorized into positive, neutral, and negative sentiments. They proceeded to fine-tune a pre-trained ParsBERT model using this dataset. The model's performance was evaluated on a separate test dataset and compared against a lexicon-based method using Polyglot. The ParsBERT model demonstrated superior performance, achieving an accuracy of 82%, surpassing the lexicon-based approach.

In this study [25], the researchers employed a method where they added the stem forms of input words and then used a pre-trained language model, ParsBERT, for classifying Persian text. This approach aimed to improve the classifier's ability to work with new, unseen data. They compared this method's performance with traditional ML algorithms and found that the proposed model achieved an accuracy of 0.91, outperforming traditional machine learning by at least +0.4 on both accuracy and F1 score.

The authors [10] investigate how external perceptions impact individual and organizational decision-making. They address a key challenge in sentiment classification for Persian, noting that existing algorithms for languages like English, French, and Arabic are not effective. Their study uses the BERT algorithm to analyze sentiment in Persian text, showing improved accuracy over previous methods. The best results were with ParsBERT on the original dataset, achieving a high F1 score of 96.62 and 94.38% accuracy.

Similarly, the paper [20] examines the impact of the Internet on trading and the value of transaction data for improving marketing strategies, using Iran's Divar online marketplace as a case study. It conducts a competition to predict the percentage of a car sales ad that will be posted on Divar. The authors utilize the Hazm library and two language models, mBERT and ParsBERT, to analyze the dataset and compare the models' performance. They discuss data mining, the Persian language, and the models' setup, presenting findings on model strengths and weaknesses. The paper highlights the challenges and potentials of working with low-resource languages and advanced language models like BERT for text analysis, also touching on data mining processes and machine learning problems. Overall, it provides insights into analyzing text data in low-resource languages using the Divar dataset.

Similarly, the paper [12] addresses the pressing need for more precise emotion classification models tailored to the Farsi and Dari dialects. Here, The authors introduce "LearnArmanEmo," a novel dataset that combines two previously established datasets—LetHerLearn and ARMANEMO—each focusing on different aspects of emotional expression in Persian. The authors propose a hybrid model that integrates the strengths of XLM-RoBERTa-large, a state-of-the-art transformer, with a Bidirectional Gated Recurrent Unit (BiGRU). This innovative approach aims to enhance the accuracy of emotion detection in Persian text, which is often complicated by the nuanced expressions found in social media and other informal contexts. The results demonstrate the model's effectiveness, achieving F1 scores of 72.9% on LetHerLearn, 77.1% on ARMANEMO, and 78.8% on the newly created LearnArmanEmo dataset.

In another work [26], the authors have improved sentiment classification for Central Kurdish using BERT. It includes pre-training BERT with a Kurdish-specific tokenizer and generating models (LSTM, MLP, and fine-tuned BERT) for sentiment classification . The fine-tuned BERT model achieved the highest accuracy at 75.37%, with 86.31% for two types of sentiment classification . This study demonstrates the superiority of BERT over traditional models such as Word2Vec, highlighting the need for language-specific models and enhancing sentiment classification in low-resource languages.

Similarly, another work [13], focusing on Persian and Dari dialects, introduces a new dataset and an advanced ensemble method for emotion classification in Persian text. The proposed model combines XLM-RoBERTa-large and BiGRU and is tested on three datasets: LetHerLearn (Dari), ARMANEMO (Persian), and LearnArmanEmo (both dialects). This model performs highly with F1 scores of 72.9% on LetHerLearn, 77.1% on ARMANEMO, and 78.8% on LearnArmanEmo, highlighting its effectiveness in Persian sentiment classification.

Although all previous approaches focus on emotion recognition from high-resource languages, work on Dari is significantly less. This is especially difficult because it is usually difficult to get the right sentiment from low-resource languages. Several early works have successfully developed a method for recognizing emotions from text in Persian, which is close to Dari, but there is still a huge research gap for Dari.

An important observation is that, among all existing works, we have rarely found XAI implementations for low-resource language emotions recognition. Therefore, perhaps our work has a unique contribution to Dari NLP.

## 3 METHODOLOGY

To start our approach, as illustrated in Fig. 1, we begin with data collection. This is followed by data augmentation and preprocessing. Subsequently, the data is split for training. We then train the BERT model, identify the best-performing model, and implement XAI.

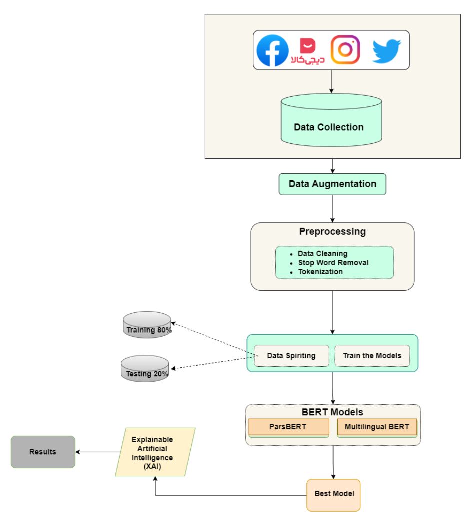


**Figure 1: Workflow of the proposed methodology for emotion recognition.**

## 3.1 Data Collection

In this paper, we use "ArmanEmo", a comprehensive manually labeled emotion dataset consisting of more than 7,000 sentences classified into 7 emotion classes. Data collection includes a variety of sources, including feedback from social media platforms such as Twitter. Instagram, and Digikala, an online marketplace based in Iran. Emotions are primarily categorized according to Ekmans six fundamental groups, namely anger, fear, happiness, hatred, sadness, and wonder. However, the inclusion of an "Other" class is implemented to account for emotions outside Ekman's model, ensuring a comprehensive representation of the complete spectrum of human experiences.

## 3.2 Data Augmentation

To balance the class distribution in our dataset, we used a random oversampling technique [14] for data sampling. This technique involves strategically creating artificial samples for minority groups, effectively increasing their representation and balancing the overall distribution. This approach helps reduce potential biases due to unbalanced data and increases the model's ability to generalize across diverse emotion categories. As the Fig. 2 show dataset distribution before and after Augmentation.

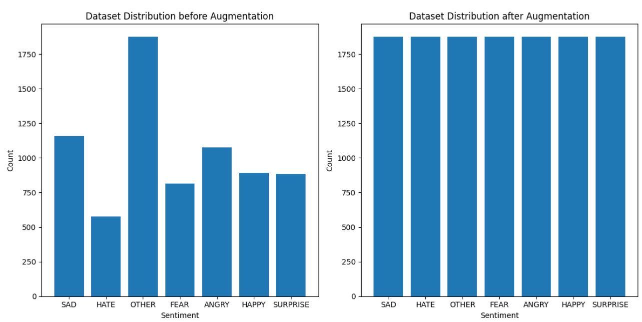


**Figure 2: Dataset distribution before (left) and after (right) augmentation.**

## 3.3 Data Preprocessing

Depending on the dataset and model requirements, we applied data cleaning and preprocessing techniques such as missing value imputation, stop word removal, and noise reduction to clean the dataset. These comprehensive techniques ensure a well-structured and clean dataset that is now suitable for emotion detection model training and evaluation. To preprocess our dataset, we employed two tokenization techniques; ParsBERT tokenizer and the Multilingual BERT tokenizer, as show in Fig. 3. ParsBERT is optimized for Persian, while mBERT is designed to handle multiple languages. Additionally, ParsBERT's vocabulary is designed for Persian, while mBERT's vocabulary is common to many languages. These techniques 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 split the data into two sets, allocating 85% for training and 15% for testing. This split ensures that a significant portion of the data is allotted for training the model to generalize successfully, while also leaving enough data for an unbiased evaluation of its performance.

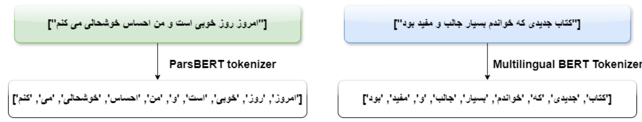


**Figure 3: Tokenization results using the ParsBERT tokenizer (left) and the Multilingual BERT tokenizer (right).**

## 3.4 BERT Model

BERT is a pre-trained language representation model developed by Google. BERT is designed to improve performance on a wide range of NLP tasks by leveraging a deep, two-way understanding of language. BERT's original model is primarily trained on English text, but its architecture has been adapted for other languages through various extensions and modifications. Below, table 1 shows a comparison table of different BERTs and their Utility.

**Table 1: Comparative table of BERT variants and their utility**

| Model | Utility |
|---|---|
| BERT [17] | General-purpose NLP tasks like classification, Q&A. |
| DistilBERT [23] | Scenarios requiring faster inference and less memory. |
| RoBERTa [19] | Improved performance on many NLP benchmarks. |
| ALBERT [18] | Applications needing high efficiency and low resource usage. |
| BanglaBERT [3] | Sentiment classification, named entity recognition, and natural language estimation for Bengali. |
| SpanBERT [15] | Tasks involving span prediction, such as coreference resolution. |
| DeBERTa [11] | Enhanced performance on multiple benchmarks. |

*3.4.1 ParsBERT Model.* In this research, we utilize the pre-trained Pars-BERT model [8] in the challenging domain of emotion detection within Dari text. ParsBERT is a language model built on the architecture of Google's BERT, designed specifically for monolingual purpose. 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 2 [9] present the statistics of ParsBERT Model comprehensive corpus:

**Table 2: Analysis of statistical data and categorization of each source within 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 |

The architecture of the model includes an embedding layer responsible for mapping input tokens to vectors, followed by 12 bidirectional Transformer blocks. Each block in the Transformer utilizes a self-attention mechanism with multiple heads and a connected feed-forward network to capture complex associations and distant dependencies 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 evaluated the model and implemented the LIME and SHAP XAI methods to bolster the interpretability of the model's classifications. The LIME and SHAP 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.

*3.4.2 Multilingual BERT Model.* For comparative analysis, we utilize the Multilingual BERT model [6], 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.

## 4 RESULTS AND DISCUSSION

Table 3 presents the performance evaluation of the Parse BERT and Multilingual BERT for the of emotion detection in Dari text. The results demonstrate that the Pras-BERT performs well as compared to the Multilingual BERT across all the metrics, achieving a significant performance. This superior performance suggests that the Pars BERT model possesses a strong capacity to correctly classify emotions in Dari text. The higher performance of Pars BERT is attributed to several factors. Firstly, it is 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. The Fig. 4 show the Confusion Matrix for Pars BERT and Multilingual BERT Model respectively.

The performance comparisons between different baseline models and BERT-based methods are summarized in table 4. The results are obtained by testing each model on the same dataset under similar conditions. The ParsBERT achieves an accuracy of 86.3%, which is higher than logistic regression 84% and random forest

**Table 3: The model's performance results**

| 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 |

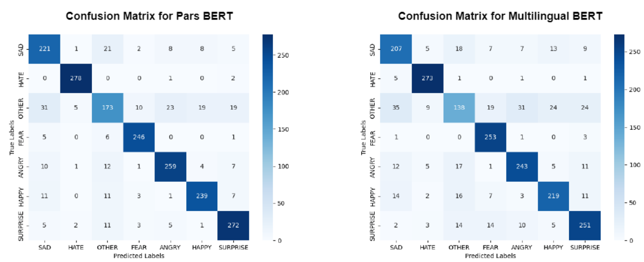


**Figure 4: Confusion matrix for ParsBERT (left) and multilingual BERT (right) model.**

81%. Multilingual BERT outperforms both SVM and LSTM, with an accuracy of 81.2%. Additionally, ParsBERT's precision, F1 score, and recall are all 0.863, which is better than the baseline models. These results confirm the superiority of BERT-based methods over traditional machine learning models, highlighting their superior performance.

To validate the improvement of the BERT-based models, we performed paired t-tests comparing them with the baseline models. A significant difference in precision, accuracy, recall, and F1 score was found for the BERT-based model compared to ParsBERT ($p < 0.05$) as well as other models such as logistic regression, SVM, and Multinomial NB ($p < 0.05$). These results confirm the superior performance of the BERT-based approach in emotion recognition tasks.

**Table 4: Performance comparison of baseline models and BERT-Based approaches**

| Model | Accuracy (%) | Precision | F1 Score | Recall |
|---|---|---|---|---|
| BERT-Based | 81.0 | 0.80 | 0.82 | 0.83 |
| Logistic Regression | 84.0 | 0.83 | 0.82 | 0.84 |
| Multinomial NB | 79.0 | 0.78 | 0.77 | 0.79 |
| SVM | 80.0 | 0.80 | 0.79 | 0.80 |
| Random Forest | 81.0 | 0.81 | 0.80 | 0.81 |
| Gradient Boosting | 65.0 | 0.64 | 0.63 | 0.65 |
| RNN | 80.0 | 0.79 | 0.78 | 0.80 |
| CNN | 82.0 | 0.81 | 0.81 | 0.82 |
| LSTM | 78.0 | 0.77 | 0.76 | 0.78 |
| ParsBERT | 86.3 | 0.86 | 0.86 | 0.87 |
| Multilingual BERT | 81.2 | 0.81 | 0.81 | 0.81 |

## 5 EXPLAINABLE ARTIFICIAL INTELLIGENCE

### 5.1 LIME 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. 5 "لعنت تحریم بائیش چی هست" (Damn, what is the reason for boycotting?) The model accurately detects the emotion expressed by the word "لعنت" (damn). Similarly, in another illustration, for example, "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 the other. The model achieves this by discerning each word's emotional content, excluding those that convey no emotion.

In a different context, as illustrated in the example provided here. "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 "درود", (greetings) "سلام" (salutations), and "عزیزی" (dear) in attributing the emotion of happiness to the sentence.


Figure 5: LIME XAI results for BERT-Based emotion recognition.

## 5.2 SHAP XAI

We used Shap XAI in addition to LIME XAI to illuminate the decision-making process of our model. To enhance clarity, we randomly selected instances of this process, which involved understanding emotion in individual words within a sentence. In Fig. 6 and 7 We can clearly see how the ParsBERT model classified emotions.

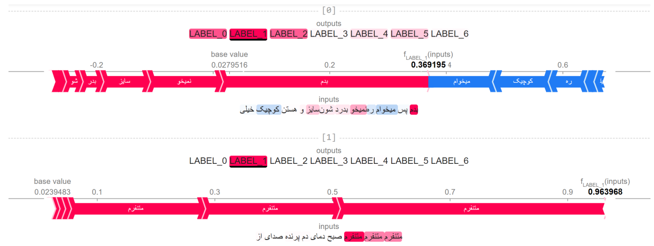


Figure 6: SHAP XAI to visualize The impact on all the output classes.

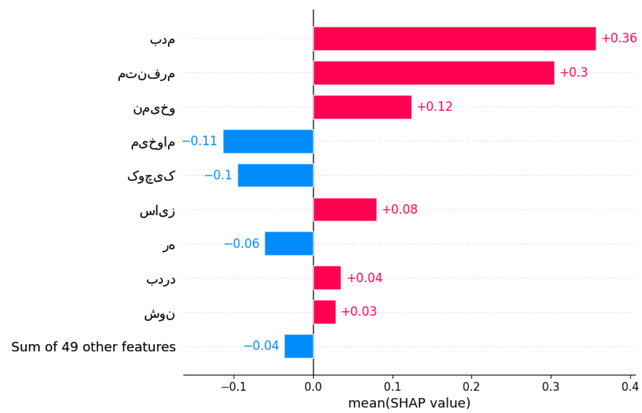


Figure 7: SHAP XAI to plotting the top words impacting a specific class.

## 6 ETHICAL IMPLICATION

Emotion detection models introduce significant ethical concerns, such as privacy, informed consent, and potential misuse. Ensuring participants are fully informed about how their data will be used is essential. Explainable AI (XAI) plays a crucial role in addressing these issues by making the decision-making processes of emotion detection models more transparent. In this study, we used two XAI techniques to help in understanding and interpreting how emotions are detected and classified, thus enhancing trust and accountability.

## 7 CONCLUSION AND FUTURE WORK

In conclusion, this study investigated the application of the BERT model for emotion classifiction in Dari text, highlighting the superior performance achieved by ParsBERT, a model specifically designed for the Dari language. Finally, we implement the LIME and SHAP XAI method for a better understanding of the model classification of emotions. This capability can significantly aid businesses, public health sectors, and government agencies in making data-driven decisions for their development and enhancement. Additionally, it can provide small entrepreneurs with a clear vision of market demand. Future directions 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 the combined ParsBERT-Multilingual BERT model, and exploring the application of the ParsBERT to other NLP tasks relevant to Dari.

## 8 ACKNOWLEDGMENTS

ChatGPT was utilized to generate sections of this Work, including text, and citations.

## REFERENCES


[1] Parham Abed Azad and Hamid Beigy. 2024. Multi-BERT: Leveraging Adapters and Prompt Tuning for Low-Resource Multi-Domain Adaptation. *arXiv e-prints* NA, NA (2024), arXiv–2404.

[2] Farid Ariai, Maryam Tayefeh Mahmoudi, and Ali Moeini. 2024. Enhancing Aspect-based Sentiment Analysis with ParsBERT in Persian Language. *Journal of AI and Data Mining* 12, 1 (2024), 1–14.

[3] Abhik Bhattacharjee, Tahmid Hasan, Wasi Uddin Ahmad, Kazi Samin, Md Saiful Islam, Anindya Iqbal, M Sohel Rahman, and Rifat Shahriyar. 2021. BanglaBERT: Language model pretraining and benchmarks for low-resource language understanding evaluation in Bangla. *arXiv preprint arXiv:2101.00204* 1 (2021), 1.

[4] Omid Davar, Gholamreza Dar, and Fahimeh Ghasemian. 2023. DeepSentiParsBERT: a deep learning model for Persian sentiment analysis using ParsBERT. In *2023 28th International Computer Conference, Computer Society of Iran (CSICC)*. IEEE, 2023 28th International Computer Conference, Computer Society of Iran (CSICC), 1–5.

[5] Mohammad Dehghani and Zahra Yazdanparast. 2023. Sentiment analysis of Persian political tweets using ParsBERT embedding model with convolutional neural network. In *2023 9th International Conference on Web Research (ICWR)*. IEEE, 2023 9th International, Conference on Web Research (ICWR), 20–25.

[6] Jacob Devlin. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* 1, 1 (2018).

[7] Mehrdad Farahani, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. 2021. Parsbert: Transformer-based model for persian language understanding. *Neural Processing Letters* 53 (2021), 3831–3847.

[8] Mehrdad Farahani, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. 2021. Parsbert: Transformer-based model for persian language understanding. *Neural Processing Letters* 53 (2021), 3831–3847.

[9] Mehrdad Farahani, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. 2021. Parsbert: Transformer-based model for persian language understanding. *Neural Processing Letters* 53 (2021), 3831–3847.

[10] Mehrdad Farahani, Mohammad Gharachorloo, and Mohammad Manthouri. 2021. Leveraging ParsBERT and pretrained mT5 for Persian abstractive text summarization. In *2021 26th International computer conference, computer society of Iran (CSICC)*. IEEE, 2021 26th International computer conference, computer society of Iran (CSICC), 1–6.

[11] Pengcheng He, Xiaodong Liu, Jianfeng Gao, and Weizhu Chen. 2020. Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv:2006.03654* 1 (2020), 1.

[12] Mohammad Ali Hussiny, Mohammad Arif Payenda, and Lilja Øvrelid. 2024. PersianEmo: Enhancing Farsi-Dari Emotion Analysis with a Hybrid Transformer and Recurrent Neural Network Model. In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages@ LREC-COLING 2024*. Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages@ LREC-COLING 2024, Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages@ LREC-COLING 2024, 257–263.

[13] Mohammad Ali Hussiny, Mohammad Arif Payenda, and Lilja Øvrelid. 2024. PersianEmo: Enhancing Farsi-Dari Emotion Analysis with a Hybrid Transformer and Recurrent Neural Network Model. In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages@ LREC-COLING 2024*. ELRA and ICCL, Torino, Italia, 257–263.

[14] imbalanced-learn. 2023. Over-sampling — Version 0.9.1. https://imbalanced-learn.org/stable/over_sampling.html Accessed: 2024-08-01.

[15] Mandar Joshi, Danqi Chen, Yinhan Liu, Daniel S Weld, Luke Zettlemoyer, and Omer Levy. 2020. Spanbert: Improving pre-training by representing and predicting spans. *Transactions of the association for computational linguistics* 8 (2020), 64–77.

[16] Sadullah Karimi. 2022. *Survey of Afghan (Dari) Language NLP for Building Afghan NLIDB System*. Ph.D. Dissertation. Brac University.

[17] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacL-HLT*, Vol. 1. Minneapolis, Minnesota, Proceedings of naacL-HLT, 2.

[18] Z Lan. 2019. Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942* 1, 1 (2019).

[19] Yinhan Liu. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* 1, 1 (2019).

[20] Ali Mehrban and Pegah Ahadian. 2023. evaluating bert and parsbert for analyzing persian advertisement data. *arXiv preprint arXiv:2305.02426* 12, 2 (2023), 2.

[21] Mahnaz Panahandeh Nigjeh and Shirin Ghanbari. 2024. Leveraging ParsBERT for cross-domain polarity sentiment classification of Persian social media comments. *Multimedia Tools and Applications* 83, 4 (2024), 10677–10694.

[22] Mohammadjalal Pouromid, Arman Yekkehkhani, Mohammadreza Asghari Oskoei, and Amin Aminimehr. 2021. ParsBERT post-training for sentiment analysis of tweets concerning stock market. In *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*. IEEE, 2021 26th International Computer Conference, Computer Society of Iran (CSICC), 1–4.

[23] V Sanh. 2019. DistilBERT, A Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter. *arXiv preprint arXiv:1910.01108* 1, 1 (2019), NA.

[24] Mohammadreza Shaghouzi, Babak Teimourpour, and Mohammad Ali Soltanshahi. 2024. Fine-tuning BERT for Persian Patent Classification: A Dataset and Model Exploration. In *2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*. IEEE, 2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), 1–3.

[25] Mohammadreza Varasteh and Arefeh Kazemi. 2021. Using parsbert on augmented data for persian news classification. In *2021 7th International Conference on Web Research (ICWR)*. IEEE, 2021 7th International Conference on Web Research (ICWR), 78–81.

[26] Hadi Veisi, Kozhin muhealddin Awlla, and Abdulhady Abas Abdullah. 2024. KuBERT: Central Kurdish BERT Model and Its Application for Sentiment Analysis. *Research Square* NA, NA (2024), NA.

[27] Ali Zareshahi, MohammadAli Javadzade, and Esmaeel Bastami. 2024. Measuring Semantic Similarity of Persian Sentences Using ParsBERT Model. In *2024 10th International Conference on Artificial Intelligence and Robotics (QICAR)*. IEEE, 2024 10th International Conference on Artificial Intelligence and Robotics (QICAR), 316–321.