

Unveiling Diverse Vaccine Sentiments: Multi-Label Text Classification

K Shanmukha Naveen^{1,0†}, S Sharon Roshini^{2,0†} and S Karthika^{3,0†}

¹Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chennai

²Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chennai

³Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam, Chennai

Abstract

This paper presents a comprehensive analysis of multi-label text classification for antivaccination tweets, utilizing a dataset of around 10,000 tweets across 12 predefined classes. The study's primary goal was to categorize the various sentiments, concerns in tweets expressed in the realm of antivaccination on social media. To achieve this, three cutting-edge transformer-based models (BERT, XLNet, and RoBERTa) were employed and fine-tuned for tweet classification. The results of our experiments revealed that the BERT model achieved notably high accuracy of 0.88 with F1 macro score being 0.65 in its classification tasks. This research significantly contributes to the field of natural language processing, highlighting the effectiveness of transformer models XLNet, RoBERTa, particularly BERT, in handling multi-label text classification for antivaccination tweets. XLNet and RoBERTa models yielded comparatively lower accuracies of 0.87 and 0.83, respectively. The insights gained from this study offer valuable implementations of these transformer models for better understanding on public concerns related to vaccination efforts and public health initiatives.

Keywords

BERT, XLNET, NLTK, RoBERTa, NLP

1. Introduction

In recent years, the proliferation of social media platforms has given rise to a surge in the dissemination of information and opinions related to various societal issues, including vaccination. The topic of vaccination, has become a subject of heated debate, with a growing presence of antivaccination sentiments on these platforms. Understanding and categorizing the sentiments, concerns, and viewpoints expressed in antivaccination tweets is essential for gaining insights into public perceptions and potentially mitigating the adverse effects of vaccine misinformation. This research analyses the power of advanced transformer models, including BERT, XLNet, and RoBERTa, similar to that of [1] and [2] for the purpose of conducting multi-label text classification. Our primary objective is to explore the impact of data preprocessing on the accuracy of classifying these tweets. While transformer-based models have demonstrated remarkable success in various natural language processing tasks, their performance

Forum for Information Retrieval Evaluation, December 15-18, 2023, India

✉ shanmukhanaveen2010809@ssn.edu.in (K. S. Naveen); sharonroshini2010942@ssn.edu.in (S. S. Roshini); skarthika@ssn.edu.in (S. Karthika)

🌐 <https://github.com/shanboii/AISOME23> (K. S. Naveen)

🆔 0000-0001-8919-5841 (S. Karthika)

© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

can be affected by the quality and characteristics of the input data. Antivaccination tweets are challenging to classify due to their unique language and varying emotions. This paper [3] assisted in understanding the role of data preprocessing in enhancing the models' predictive capabilities. In addition to investigating data preprocessing, this research systematically explores a range of hyperparameters, including batch size, decay rates, learning rates, and epochs, to fine-tune the transformer models. By optimizing these hyperparameters, we aim to push the boundaries of classification accuracy, enabling more precise categorization of tweets across multiple predefined classes. By shedding light on the interplay between data preprocessing, hyperparameter tuning, and transformer-based models, we contribute to advancing the state of the art models in text classification.

The key contributions of this paper are as follows: Firstly, data preprocessing was applied to precisely 9,921 tweets within the provided dataset. Secondly, we trained three advanced models, namely BERT, XLNet, and RoBERTa, on the preprocessed dataset, utilizing their transformative capabilities for effective text classification. Lastly, an optimization process was conducted on hyperparameters to boost classification accuracy, ensuring the models performed at their best.

2. Related work

This paper [4] focuses on creating a multi-labeled Arabic dataset from COVID-19 tweets, exploring both traditional machine learning and deep-learning approaches to achieve higher accuracy and stable performance in sentiment analysis and topic classification on Twitter. Another paper [5] introduces LIAR which improves biomedical document classification using BioBERT and an adaptive loss function, surpassing state-of-the-art methods by 1 percent. [6] conducted sentiment analysis on Twitter discussions about COVID-19 vaccines. [7] analyzes COVID-19-related emotions in Twitter data and finds BERT to be outperforming others in sentiment analysis and classification. This [8] investigated public perceptions of COVID-19 vaccine adverse effects through social media data, with LSTM achieving the highest accuracy. This study [9] uses machine learning and NLP, particularly the BERT model, to analyze sentiments in COVID-19 vaccination tweets, achieving the highest accuracy among classifiers. Here [10] Support Vector Machine (SVM) performed best in analyzing Covid-19 vaccine-related tweets with accuracy of 84.32. [11] utilizes sentiment analysis and machine learning techniques to analyze public attitudes towards COVID-19 vaccination. This paper [12] analyzes COVID-19 vaccine perception, using Twitter data and deep learning models, with LSTM achieving the highest accuracy at 85.7 percent. This study [13] addresses automated topic annotation challenges in COVID-19 literature, presenting the BioCreative LitCovid dataset and achieving remarkable F1 scores of 0.8875 (macro), 0.9181 (micro), and 0.9394 (instance-based) with transformer-based hybrid systems. This paper [14] and [15] introduces CAVES, a substantial dataset of COVID-19 anti-vaccine tweets categorized by specific concerns and classification of vaccine hesitancy on social media.

3. Methodology

The methodology employed for the multi-label text classification of antivaccination tweets involved a systematic approach encompassing data preprocessing, model implementation, and

evaluation. The study was conducted using two distinct strategies: one utilizing preprocessed data, and the other utilizing unprocessed data. The following sections outline the key steps undertaken in this research.

3.1. Data preprocessing

The initial phase of the research involved data preprocessing to ensure the dataset's suitability for subsequent analysis. The Natural Language Toolkit (NLTK) was utilized for text preprocessing tasks such as tokenization, stopword removal, and stemming. The tweets in the dataset were complex, containing symbols like '@', brackets, HTML tags and other emojis. These were removed, and all tweets were changed to lowercase. Additionally, each tweet was categorized by assigning binary labels (0 or 1) based on its relevance to predefined classes. These twelve classes, namely mandatory, country, conspiracy, unnecessary, political, ingredients, side-effect, pharma, none, ineffective, rushed, and religious, were used to classify the tweets in the context of antivaccination discussions. This preprocessing step aimed to enhance the quality of the dataset and establish a basis for subsequent model training.

3.2. Model Implementation

In this study, three prominent transformer-based models, namely BERT, XLNet, and RoBERTa, were employed due to their exceptional ability to capture intricate linguistic relationships and patterns in text data. Each model was explored using two distinct approaches:

a. Preprocessed Data: The preprocessed dataset was utilized for model training. The tokenized and labeled tweets, a result of our preprocessing steps, were fed into the transformer models for fine-tuning. This approach enabled effective classification of tweets into specified classes based on the learned features and patterns.

b. Unprocessed Data: Additionally, the unprocessed dataset containing raw text tweets was employed for model training. This approach aimed to assess the models' capacity to handle noisy and unstructured data directly from the source. Training on unprocessed data provided insights into the models' ability to effectively process and classify tweets without preprocessing.

These two approaches were implemented to determine which method resulted in a higher classification accuracy.

3.3. Evaluation

A comparative analysis was conducted to assess the performance of models trained on two dataset variants: preprocessed and unprocessed. The primary objective was to understand how data preprocessing influenced model performance and classification accuracy. Model evaluation involved employing an 80-20 train-test split approach, where 80 percent of the transformed data from both preprocessed and unprocessed datasets was designated for training, and the remaining 20 percent was allocated for testing. Here, test dataset containing 486 records was utilized for prediction by the models. Subsequently, trained models were applied to predict tweet labels in the test sets. Evaluation metrics, including accuracy and F1 macro validation score, were employed to gauge the models' classification accuracy. After careful comparison of accuracy and F1 macro validation score, the model showing the best performance in classifying

antivaccination tweets was chosen. This was chosen based on its higher accuracy and F1 macro score reports. Moreover, this choice is based on concrete data, ensuring accurate classification of various sentiments and concerns expressed in antivaccination discussions on social media. The chosen model highlights the progress in natural language processing and its vital role in understanding public opinions about vaccination efforts, especially during the global pandemic.

4. Implementation

The three transformer models (BERT, XLNet, and RoBERTa) were trained on the antivaccination tweet dataset, an iterative process was initiated to enhance the classification performance. This iterative approach involved systematically experimenting with various combinations of hyperparameters, including batch sizes, decay rates, learning rates, and epochs.

4.1. BERT

BERT model, utilizes a multi-layer bidirectional transformer encoder to represent text which enables it to grasp the full context of each word in a sentence, significantly enhancing its understanding of text meaning. One of the most interesting things about BERT is that it's a pre-trained model. In our dataset, BERT quickly understood tweets and gave us better accuracy. BERT is good at figuring out what tweets mean, even though tweets are short and informal. This pre-training equips BERT with a deep understanding of language structure and meaning.

4.2. XLNET

XLNet, an extension of BERT, employs a permutation-based training approach, enabling it to understand bidirectional context. Unlike BERT, which reads text in fixed bidirectional sequences, XLNet considers all possible permutations of the input words. To process the tweets in our dataset for classification using XLNet, the tweets undergo initial cleaning to remove irrelevant information. The cleaned text is then broken down into smaller units called tokens through tokenization. These tokens are then numerically encoded using XLNet's vocabulary, assigning a unique ID to each token. These numerical representations are fed into the XLNet model for classification.

4.3. RoBERTa

RoBERTa, an extension of the BERT architecture, stems from the BERT revolution and stands for Robustly Optimized BERT Pretraining Approach. Recognizing that BERT was undertrained despite its remarkable performance, the authors proposed crucial modifications. These included more extensive training with larger batches and data, eliminating the next sentence prediction objective, and incorporating dynamic masking during pretraining. Notably, RoBERTa employs a distinct tokenizer, byte-level BPE, and a larger vocabulary (50k vs. 30k) compared to BERT. Despite the resulting increase in model complexity due to a larger vocabulary, RoBERTa justifies this enhancement with significant performance gains across various tasks.

After training these three transformer models with both preprocessed and unprocessed data, hyperparameters, including batch size, decay rates, learning rates, and epochs, were fine-tuned to achieve higher accuracy and an optimal F1 macro score. The process aimed at refining the models' predictive capabilities for multi-label text classification, ensuring a more precise categorization of the diverse sentiments, concerns, and viewpoints expressed within antivaccination discussions on social media. In essence, this thorough optimization of hyperparameters served as a critical step in maximizing the models' performance and enhancing their ability to accurately classify antivaccination tweets.

5. Observations and Results

From the implementations, our exploration of preprocessed and unprocessed data for multilabel text classification of antivaccination tweets revealed that the models trained on the unprocessed dataset exhibited higher accuracies compared to their preprocessed counterparts. Notably, optimizing hyperparameters, specifically setting the learning rate to 0.001 or 0.0001, decay rate to 0.01 or 0.001, using either 5 or 10 epochs, and a batch size of 32, significantly contributed to higher accuracies and F1 score and Jaccard score.

The evaluation metric used here to evaluate model's accuracy is F1 macro score and Jaccard score.

Macro F1-Score is useful in multiclass classification problems where the classes are imbalanced. It gives equal weight to each class, making it a good metric to use when the data is imbalanced, given by the formula:

$$F1 \text{ score} = 2 \times \left(\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \right) \quad (1)$$

The Jaccard Score is used for evaluating models dealing with binary or multiclass classification focusing solely on the overlap between predicted and actual positive instances, given by the formula :

$$Jaccard(\mathcal{U}, \mathcal{V}) = \frac{|\mathcal{U} \cap \mathcal{V}|}{|\mathcal{U} \cup \mathcal{V}|} \quad (2)$$

Amongst all, the learning rate, set at values like 0.001 or 0.0001, was a critical factor for weight adjustments during training. This parameter played a significant role in achieving convergence, where an apt learning rate ensured the model learned effectively from the data without surpassing optimal parameter values. The "epochs" parameter dictates the total number of full passes through the training dataset, enabling models to iteratively learn and adjust their parameters as they progress through the data. Thus, fine-tuning and optimising these hyperparameters ensured the models robust performance yielding excellent accuracies and F1 scores.

This experimental results are based on the following observations of hyperparameter optimisations:

As a result, the models trained on unprocessed data contributed to yield better results. The preprocessed data did not result in higher accuracy. The unprocessed approach retains authentic language and nuances found in tweets, aiding the model in better understanding and classification. Table 2 specifies the results of models F1 scores and Jaccard scores. Here,

Table 1

Results based on optimizing hyperparameters

Model	Learning Rate	Batch Size	Decay Rate	Epochs	Accuracy
BERT	2e-5	32	0.01	3	0.88
RoBERTa	2e-5	32	0.01	3	0.87
XLNET	2e-5	32	0.01	3	0.83

model-1 , model-2 and model-3 mentioned in the table refers to BERT , RoBERTa and XLNET respectively.

Table 2

Run results

Run Model	MODEL	F1 Macro score	Jaccard score
model1	BERT	0.55	0.57
model2	RoBERTa	0.65	0.63
model3	XLNET	0.64	0.65

6. Conclusions

To sum up, the study aimed to accurately sort tweets about antivaccination. It was found that using BERT, especially with unprocessed data, gave the most accurate results. Different setups were tested to get the best outcomes. The main focus was to get the highest possible score (F1 macro) for properly labeling tweets. The careful tweaking of the models resulted in better scores that confirmed the models' effectiveness in understanding and sorting tweets on the topic of antivaccination, even when the conversation is tricky or complicated. In conclusion, the research emphasizes the value of careful adjustments to the models, leading to strong models that can grasp the complexities of discussions about antivaccination on social media. The excellent performance of the BERT model with unprocessed data highlights its potential in understanding and categorizing complex written content, especially in public health conversations.

References

- [1] L. Cai, Y. Song, T. Liu, K. Zhang, A hybrid bert model that incorporates label semantics via adjustive attention for multi-label text classification, *IEEE Access* 8 (2020) 152183–152192. doi:10.1109/ACCESS.2020.3017382.
- [2] W.-C. Chang, H.-F. Yu, K. Zhong, Y. Yang, I. S. Dhillon, Taming pretrained transformers for extreme multi-label text classification, in: *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, 2020*, pp. 3163–3171.
- [3] S. Vijayarani, M. J. Ilamathi, M. Nithya, et al., Preprocessing techniques for text mining-an

overview, *International Journal of Computer Science & Communication Networks* 5 (2015) 7–16.

- [4] F. M. Alderazi, A. A. Algosaihi, M. A. Alabdullatif, Multi-labeled dataset of arabic covid-19 tweets for topic-based sentiment classifications, in: *2022 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, 2022, pp. 1–8. doi:10.1109/EAIS51927.2022.9787700.
- [5] Z. Chen, J. Peng, Learning label independence and relevance for multi-label biomedical text classification, in: *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2022, pp. 2776–2781. doi:10.1109/SMC53654.2022.9945404.
- [6] Z. Xu, L. Shi, Y. Wang, J. Zhang, L. Huang, C. Zhang, S. Liu, P. Zhao, H. Liu, L. Zhu, Y. Tai, C. Bai, T. Gao, J. Song, P. Xia, J. Dong, J. Zhao, F.-S. Wang, Pathological findings of covid-19 associated with acute respiratory distress syndrome, *The Lancet Respiratory Medicine* 8 (2020) 420–422. URL: <https://www.sciencedirect.com/science/article/pii/S221326002030076X>. doi:[https://doi.org/10.1016/S2213-2600\(20\)30076-X](https://doi.org/10.1016/S2213-2600(20)30076-X).
- [7] V. Battula, S. G. Goli, J. Nasigari, Identification of optimal model for multi-class classification of covid tweets, in: *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2022, pp. 495–499. doi:10.23919/INDIACom54597.2022.9763291.
- [8] K. Kariyapperuma, K. Banujan, P. Wijeratna, B. Kumara, Classification of covid19 vaccine-related tweets using deep learning, in: *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 2022, pp. 1–5. doi:10.1109/ICDABI56818.2022.10041615.
- [9] S. Ningombam, A. Roy, P. Debnath, *An Empirical Analysis of Different Classifiers on COVID-19 Vaccination Data*, Springer Nature Singapore, Singapore, 2023, pp. 285–295. URL: https://doi.org/10.1007/978-981-19-9304-6_28. doi:10.1007/978-981-19-9304-6_28.
- [10] S. K. Akpatsa, X. Li, H. Lei, V.-H. K. S. Obeng, Evaluating public sentiment of covid-19 vaccine tweets using machine learning techniques, *Informatica* 46 (2022).
- [11] N. Gao, *Text Analysis of Twitter Data for COVID-19 Vaccines*, Ph.D. thesis, Instytut Informatyki, 2023.
- [12] K. T. Shahriar, M. N. Islam, M. M. Anwar, I. H. Sarker, Covid-19 analytics: Towards the effect of vaccine brands through analyzing public sentiment of tweets, *Informatics in medicine unlocked* 31 (2022) 100969.
- [13] Q. Chen, A. Allot, R. Leaman, R. Islamaj, J. Du, L. Fang, K. Wang, S. Xu, Y. Zhang, P. Bagherzadeh, S. Bergler, A. Bhatnagar, N. Bhavsar, Y.-C. Chang, S.-J. Lin, W. Tang, H. Zhang, I. Tavchioski, S. Pollak, S. Tian, J. Zhang, Y. Otmakhova, A. J. Yepes, H. Dong, H. Wu, R. Dufour, Y. Labrak, N. Chatterjee, K. Tandon, F. A. A. Laleye, L. Rakotoson, E. Chersoni, J. Gu, A. Friedrich, S. C. Pujari, M. Chizhikova, N. Sivadasan, S. VG, Z. Lu, Multi-label classification for biomedical literature: an overview of the BioCreative VII LitCovid Track for COVID-19 literature topic annotations, *Database* 2022 (2022) baac069. URL: <https://doi.org/10.1093/database/baac069>. doi:10.1093/database/baac069. arXiv:<https://academic.oup.com/database/article-pdf/doi/10.1093/database/baac069/45629681/baac069.pdf>.
- [14] S. Poddar, A. M. Samad, R. Mukherjee, N. Ganguly, S. Ghosh, Caves: A dataset to facilitate explainable classification and summarization of concerns towards covid vaccines, in: *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 3154–3164.

- [15] S. Poddar, M. Basu, K. Ghosh, S. Ghosh, Overview of the fire 2023 track:artificial intelligence on social media (aisome), in: Proceedings of the 15th Annual Meeting of the Forum for Information Retrieval Evaluation, 2023.