

Vulgar Word Extraction in Chittagonian Dialect of Bangla

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Tanjim Mahmud , Michal Ptaszynski , Fumito Masui

The proliferation of the internet, especially on social media platforms, has amplified the prevalence of cyberbullying and harassment. Addressing this issue involves harnessing natural language processing (NLP) and machine learning (ML) techniques for the automatic detection of harmful content. However, these methods encounter challenges when applied to low-resource languages like the Chittagonian dialect of Bangla. This entry compares two approaches for identifying offensive language containing vulgar remarks in Chittagonian. The first relies on basic keyword matching, while the second employs machine learning and deep learning techniques. The keyword-matching approach involves scanning the text for vulgar words using a predefined lexicon. Despite its simplicity, this method establishes a strong foundation for more sophisticated ML and deep learning approaches. An issue with this approach is the need for constant updates to the lexicon.

vulgar remark detection

vulgar term extraction

low-resource language

logistic regression

recurrent neural network

1. Introduction

Bangladesh has seen a remarkable increase in its use of the Internet over the past two decades. There were more than 125 million Internet users in Bangladesh as of November 2022, according to the Bangladesh Telecommunication Regulatory Commission (BTRC) ^[1]. Additionally, with the help of the implementation of the Digital Bangladesh initiative ^[2], the vast majority of people in Chittagong ^[3], Bangladesh's second-largest city, now have access to the Internet and actively use social media. According to a survey ^[4], the number of Facebook users in Bangladesh is the highest among social media (see **Figure 1**). Moreover, with the benefit of Unicode being widely used on most communication devices, such as tablets or smartphones, speakers of underrepresented languages, such as Chittagonian, can express their thoughts in their native languages and dialects. Many people in Chittagong now use social media on a regular basis, regularly using platforms like Facebook ^[5], imo ^[6], various blogs, and WhatsApp ^[7]. These platforms offer a venue where people can express themselves freely and informally. However, the pervasiveness of social media has also resulted in unfavorable influences that are difficult to shake. Excessive use of social media has the potential to cause addiction ^[8], which as a result could cause young people to spend more time on these platforms than they spend with their family and friends ^[9]. Their general health and social interactions may suffer as a result of this addiction. Additionally, social media witnesses the growing problem of the increase in online abuse and cyberbullying, which can have a negative impact on a person's self-esteem and even violate their privacy ^[10]. The spread of misinformation and hatred online has also

contributed to an uptick in violent crimes in society [11]. Receiving messages with vulgar language is a startling realization of this unwelcome and damaging phenomenon. The likelihood of encountering such vulgar remarks rises as social media use increases.

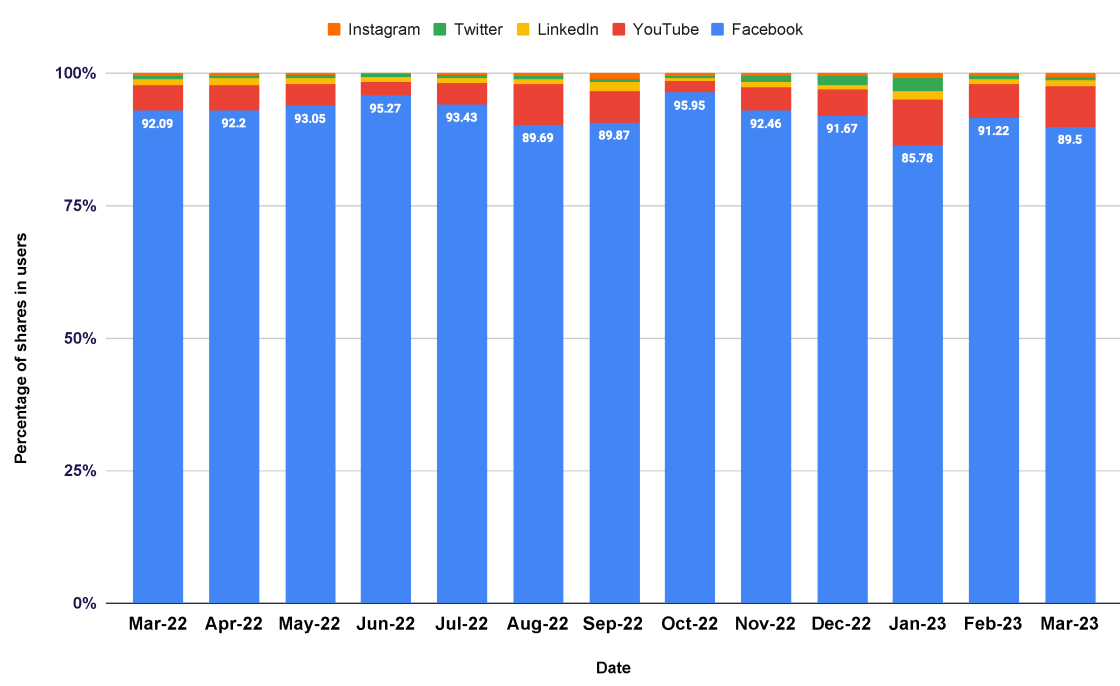


Figure 1. Yearly social media usage statistics in Bangladesh as of March 2023 [4].

Vulgarity or obscenity with regards to language refers to terms used to describe the use of vulgar language, such as swearing, taboo words, or offensive expressions [12][13]. Unfortunately, such language has become more and more common in contemporary culture [14], especially on social media sites like Twitter [15]. The majority of the time, however, vulgar language is used in the context of online harassment and negativity. While there are some instances where vulgar language may be used in a positive context to convey informality, express anger, or establish a sense of belonging with a particular group [16], these are usually rare in comparison to its use in negative contexts or in closed message groups. Therefore, detecting such use of language quickly and effectively is necessary to allow social media platforms to efficiently moderate their contents.

2. Automatic Vulgar Word Extraction Method with Application to Vulgar Remark Detection in Chittagonian Dialect of Bangla

Table 1 shows research in Bengali on topics related to detecting vulgarity.

Traditionally, vulgar expression lexicons have been developed as a means of vulgarity detection [17]. These lexicon-based approaches need to be updated frequently to remain effective, however. In contrast, machine learning (ML) techniques provide a more dynamic approach by classifying new expressions as either vulgar or non-vulgar without

relying on predetermined lexicons. Deep learning has made significant contributions to the field of signal and image processing [18], diagnosis [19], wind forecasting [20] and time series forecasting [21].

Beyond lexicon-based techniques, vulgarity detection has been the subject of several studies. Moreover, numerous linguistic and psychological studies [22] have been carried out to comprehend the pragmatic applications [13] and various vulgar language forms [23].

For machine learning-related studies, for example, Eshan et al. [24] ran an experiment in which they classified data obtained by scraping the Facebook pages of well-known celebrities using the traditional machine learning classifiers multinomial naive Bayes, random forest, and SVM (support vector machine). They gathered unigram, bigram, and trigram features and weighted them using *TF-IDF* vectorizers. On datasets of various sizes, containing 500, 1000, 1500, 2000, and 2500 samples. The results showed that when using unigram features, a sigmoid kernel had the worst accuracy performance, and SVM with a linear kernel had the best accuracy performance. However, MNB demonstrated the highest level of accuracy for bigram and trigram features. In conclusion, TfidfVectorizer features outperformed CountVectorizer features when combined with an SVM linear kernel.

Akhter et al. [25] suggested using user data and machine learning techniques to identify instances of cyberbullying in Bangla. They used a variety of classification algorithms, such as naive Bayes (NB), J48 decision trees, support vector machine (SVM), and k-nearest neighbors (KNN). A 10-fold cross-validation was used to assess how well each method performed. The results showed that SVM performed better than the other algorithms when it came to analyzing Bangla text, displaying the highest accuracy score of 0.9727.

Holgate et al. [16] introduced a dataset of 7800 tweets from users whose demographics were known. Each instance of vulgar language use was assigned to one of six different categories by the researchers. These classifications included instances of aggression, emotion, emphasis, group identity signaling, auxiliary usage, and non-vulgar situations. They sought to investigate the practical implications of vulgarity and its connections to societal problems through a thorough analysis of this dataset. Holgate et al. obtained a macro F1 score of 0.674 across the six different classes by thoroughly analyzing the data that were gathered.

Emon et al. [26] created a tool to find abusive Bengali text. They used various deep learning and machine learning-based algorithms to achieve this. A total of 4700 comments from websites like Facebook, YouTube, and Prothom Alo were collected in a dataset. These comments were carefully labeled into seven different categories. Emon et al. experimented with various algorithms to find the best one. The recurrent neural network (RNN) algorithm demonstrated the highest accuracy among the investigated methods, achieving a satisfying score of 0.82.

Awal et al. [27] demonstrated a naive Bayes system made to look for abusive comments. They gathered a dataset of 2665 English comments from YouTube in order to evaluate their system. They then translated these English remarks into Bengali utilizing two techniques: (i) Bengali translation directly; (ii) Bengali translation using dictionaries. Awal et al. evaluated the performance of their system after the translations. Their system impressively

achieved the highest accuracy of 0.8057, demonstrating its potency in identifying abusive content in the context of the Bengali language.

Hussain et al. [28] suggested a method that makes use of a root-level algorithm and unigram string features to identify abusive Bangla comments. They gathered 300 comments for their dataset from a variety of websites, including Facebook pages, news websites, and YouTube. The dataset was split into three subsets, each of which contained 100, 200, and 300 comments. These subsets were used to test their system, which resulted in an average accuracy score of 0.689.

Das et al. [29] carried out a study on detecting hate speech in Bengali and Romanized Bengali. They extracted samples from Twitter in order to gather the necessary information, producing a dataset with 5071 samples in Bengali and Romanized Bengali. They used a variety of training models in their study, including XML-RoBERTa, MuRIL, m-BERT, and IndicBERT. Following testing, they discovered that XML-RoBERTa had the highest accuracy, at 0.796.

Sazzed [30] collected 7245 YouTube reviews manually and divided them into two categories: vulgar and non-vulgar. The purpose of this process was to produce two benchmark corpora for assessing vulgarity detection algorithms. Following the testing of several methods, the bidirectional long short-term memory (BiLSTM) model showed the most promising results, achieving the highest recall scores for identifying vulgar content in both datasets.

Jahan et al. [31] created a dataset by using online comment scraping tools to collect comments from public Facebook pages, such as news and celebrity pages. SVM, random Forest, and AdaBoost were the three machine learning techniques used to categorize the comments for the detection of abusive content. Their approach, which was based on the random forest classifier, outperformed other methods in terms of accuracy and precision, scoring 0.7214 and 0.8007, respectively. AdaBoost, on the other hand, demonstrated the best recall performance, earning a score of 0.8131.

Ishmam et al. [32] collected a dataset sourced from Facebook, categorized into six distinct classes. The dataset was enriched with linguistic and quantitative features, and the researchers employed a range of text preprocessing techniques, including punctuation removal, elimination of bad characters, handling hashtags, URLs, and mentions, as well as tokenization and stemming. They utilized neural networks, specifically GRUs (gated recurrent units), alongside other machine learning classifiers, to conduct classification tasks based on the historical, religious, cultural, social, and political contexts of the data.

Karim et al. [33] used a combination of machine learning classifiers and deep neural networks to detect hate speech in Bengali. They analyzed datasets containing comments from Facebook, YouTube, and newspaper websites using a variety of models, including logistic regression, SVM, CNN, and Bi-LSTM. The researchers divided hate speech into four distinct categories: political, religious, personal, and geopolitical. With F1 scores of 0.78 for political hate speech, 0.91 for personal hate speech, 0.89 for geopolitical hate speech, and 0.84 for religious hate speech detection in the Bengali language, their results showed satisfying performance.

Sazzed [34] created a transliterated corpus of 3000 comments from Bengali, 1500 of which were abusive and 1500 of which were not. As a starting point, they used a variety of supervised machine learning methods, such as deep learning-based bidirectional long short-term memory networks (BiLSTM), support vector machines (SVM), logistic regression (LR), and random forest (RF). The SVM classifier displayed the most encouraging results (with an F1 score of 0.827 ± 0.010) in accurately detecting abusive content.

User comments from publicly viewable Facebook posts made by athletes, officials, and celebrities were analyzed in a study by Ahmed et al. [35]. The researchers distinguished between Bengali-only comments and those written in English or a mix of English and other languages. Their research showed that 14,051 initial comments in total, or approximately 31.9% of them, were directed at male victims. However, a significant number of the 29,950 comments, or 68.1% of the total, were directed at female victims. The study also highlighted how comments were distributed according to the different types of victims. A total of 9375 comments were directed at individuals who are social influencers. Among these, 5.98% (equivalent to 2633 comments) were aimed at politicians, while 4.68% (or 2061 comments) were focused on athletes. Additionally, 6.78% (about 2981 comments) of the comments were centered around singers, and the majority, which is 61.25% (totaling 26,951 comments), were directed at actors.

For the classification of hate speech in the Bengali language, Romim et al. [36] used neural networks, including LSTM (long short-term memory) and BiLSTM (bidirectional LSTM). They used word embeddings that had already been trained using well-known algorithms such as FastText, Word2Vec, and Glove. The largest dataset of its kind to date, the extensive Bengali dataset they introduced for the research includes 30,000 user comments. The researchers thoroughly compared different deep learning models and word embedding combinations. The outcomes were encouraging as all of the deep learning models performed well in the classification of hate speech. However, the support vector machine (SVM) outperformed the others with an accuracy of 0.875.

Islam et. al. [37] used large amounts of data gathered from Facebook and YouTube to identify abusive comments. To produce the best results, they used a variety of machine learning algorithms, such as multinomial naive Bayes (MNB), multilayer perceptron (MLP), support vector machines (SVM), decision tree, random forest, and SVM with stochastic gradient descent-based optimization (SGD), ridge classifier, and k-nearest neighbors (k-NN). They used a Bengali stemmer for preprocessing and random undersampling of the dominant class before processing the dataset. The outcomes demonstrated that, when applied to the entire dataset, SVM had the highest accuracy of 0.88.

In their study, Aurpa et al. [38] used transformer-based deep neural network models, like BERT [39] and ELECTRA [40], to categorize abusive comments on Facebook. For testing and training, they used a dataset with 44,001 Facebook comments. The test accuracy for their models, which was 0.85 for the BERT classifier and 0.8492 for the ELECTRA classifier, showed that they were successful in identifying offensive content on the social media platform.

Table 1. Research on vulgarity detection or related topics in Bengali (Facebook (F), YouTube (Y)).

Paper	Classifier	Highest Score	Language	Sample Size	Class and Ratio	Data Sources
[24]	Multinomial Naive Bayes, Random Forest, Support Vector Machines,	80% (Accuracy)	Bengali	2.5K	-	F
[25]	Support Vector Machines, Naive Bayes, Decision Tree, K-Nearest Neighbors	97% (Accuracy)	Bengali	2.4 K	Non-Bullying Bullying (10%)	F, T
[26]	Linear Support Vector Classification, Logistic Regression, Multinomial Naive Bayes, Random Forest Artificial Neural Network, RNN + LSTM	82.2% (Accuracy)	Bengali	4.7 K	Slang (19.57%), Religious, Politically, Positive, Neutral, violated (13.28%), Anti-feminism (0.87%), Hatred (13.15%), Personal attack (12.36%)	F, Y, News portal
[27]	Naive Bayes	80.57% (Accuracy)	Bengali	2.665 K	Non-Abusive, Abusive (45.55%)	Y
[28]	Root-Level approach	68.9% (Accuracy)	Bengali	300	Not Bullying, Bullying	F, Y News portal
[30]	Logistic Regression, Support Vector Machines, Stochastic Gradient Descent, Bidirectional LSTM	89.3% (F1 Score) 82.4% (F1 Score)	Bengali	7.245 K	Non Vulgar, Vulgar	Y
[31]	Support Vector Machines, RF, Adaboost	72.14% (Accuracy) 80% (Precision)	Bengali	2 K	Non Abusive, Abusive (78.41%)	F
[32]	Gated Recurrent Units, Support Vector Classification, LinearSVC, Random Forest, Naive Bayes	70.1% (Accuracy)	Bengali	5.126 K	Religious comment (14.9%), Hate speech (19.2%), Inciteful (10.77%), Communal	F

Paper	Classifier	Highest Score	Language	Sample Size	Class and Ratio	Data Sources
					hatred (15.67%), Religious hatred (15.68%), Political comment (23.43%)	
[33]	Logistic Regression, Support Vector Machines, Convolutional Neural Network, Bidirectional LSTM, BERT, LSTM	78% 91% 89% 84% (F1 Score)	Bengali	8.087 K	Personal (43.44%), Religious (14.97%), Geopolitical (29.23%), Political (12.35%)	F
[34]	Logistic Regression, Support Vector Machines, Random Forest, Bidirectional LSTM	82.7% (F1 Score)	Bengali	3 K	Non abusive, Abusive (10%)	Y
[36]	Long Short-term Memory, Bidirectional LSTM	87.5% (Accuracy)	Bengali	30 K	Not Hate speech, Hate speech (33.33%)	F, Y
[37]	Multinomial Naive Bayes, Multilayer Perceptron, Support Vector Machines, Decision Tree, Random Fores, Stochastic Gradient Descent, K-Nearest Neighbors	88% (Accuracy)	Bengali	9.76 K	Non Abusive, abusive (50%)	F, Y
[38]	ELECTRA, Deep Neural Network, BERT	85% (Accuracy) (BERT), 84.92% (Accuracy) (ELECTRA)	Bengali	44.001 K	Troll (23.78%), Religious (17.22%), Sexual (20.29%), Not Bully (34.86%), Threat (3.85%)	F

usive and
rch gaps
of small-
sized datasets without a well-defined annotation process, and the lack of benchmarking efforts to assess dataset quality. Additionally, class imbalance in datasets remains an issue, and limited attention has been given to vulgarity detection in low-resource language Bengali, with only a single work [30] addressing this area. Many papers fail to specify the source of their datasets and conduct limited experiments. Field surveys are often superficial or nonexistent. Furthermore, none of the papers considered ethical considerations in data collection, such as preserving user privacy through dataset anonymization. Addressing these research gaps is essential for advancing

the field of vulgarity detection and related areas, ensuring the development of more robust, ethical, and well-defined detection systems.

References

1. Bangladesh Telecommunication Regulatory Commission. Available online: <http://www.btrc.gov.bd/site/page/347df7fe-409f-451e-a415-65b109a207f5/> (accessed on 15 January 2023).
2. United Nations Development Programme. Available online: <https://www.undp.org/bangladesh/blog/digital-bangladesh-innovative-bangladesh-road-2041> (accessed on 20 January 2023).
3. Chittagong City in Bangladesh. Available online: <https://en.wikipedia.org/wiki/Chittagong> (accessed on 1 April 2023).
4. StatCounter Global Stats. Available online: <https://gs.statcounter.com/social-media-stats/all/bangladesh/#monthly-202203-202303> (accessed on 24 April 2023).
5. Facebook. Available online: <https://www.facebook.com/> (accessed on 28 January 2023).
6. imo. Available online: <https://imo.im> (accessed on 28 January 2023).
7. WhatsApp. Available online: <https://www.whatsapp.com> (accessed on 28 January 2023).
8. Addiction Center. Available online: <https://www.addictioncenter.com/drugs/social-media-addiction/> (accessed on 28 January 2023).
9. Prothom Alo. Available online: <https://en.prothomalo.com/bangladesh/Youth-spend-80-mins-a-day-in-Internet-adda> (accessed on 28 January 2023).
10. United Nations. Available online: <https://www.un.org/en/chronicle/article/cyberbullying-and-its-implications-human-rights> (accessed on 28 January 2023).
11. ACCORD—African Centre for the Constructive Resolution of Disputes. Available online: <https://www.accord.org.za/conflict-trends/social-media/> (accessed on 28 January 2023).
12. Cachola, I.; Holgate, E.; Preotjuc-Pietro, D.; Li, J.J. Expressively vulgar: The socio-dynamics of vulgarity and its effects on sentiment analysis in social media. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 2927–2938.
13. Wang, N. An analysis of the pragmatic functions of “swearing” in interpersonal talk. Griffith Work. Pap. Pragmat. Intercult. Commun. 2013, 6, 71–79.

14. Mehl, M.R.; Vazire, S.; Ramírez-Esparza, N.; Slatcher, R.B.; Pennebaker, J.W. Are women really more talkative than men? *Science* 2007, 317, 82.
15. Wang, W.; Chen, L.; Thirunarayan, K.; Sheth, A.P. Cursing in English on twitter. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Baltimore, MD, USA, 15–19 February 2014; pp. 415–425.
16. Holgate, E.; Cachola, I.; Preoțiuc-Pietro, D.; Li, J.J. Why swear? Analyzing and inferring the intentions of vulgar expressions. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, 31 October–4 November 2018; pp. 4405–4414.
17. Sazzed, S. A lexicon for profane and obscene text identification in Bengali. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, Online, 1–3 September 2021; pp. 1289–1296.
18. Das, S.; Mahmud, T.; Islam, D.; Begum, M.; Barua, A.; Tarek Aziz, M.; Nur Showan, E.; Dey, L.; Chakma, E. Deep Transfer Learning-Based Foot No-Ball Detection in Live Cricket Match. *Comput. Intell. Neurosci.* 2023, 2023, 2398121.
19. Mahmud, T.; Barua, K.; Barua, A.; Das, S.; Basnin, N.; Hossain, M.S.; Andersson, K.; Kaiser, M.S.; Sharmen, N. Exploring Deep Transfer Learning Ensemble for Improved Diagnosis and Classification of Alzheimer’s Disease. In *Proceedings of the 2023 International Conference on Brain Informatics*, Hoboken, NJ, USA, 1–3 August 2023; Springer: Cham, Switzerland, 2023; pp. 1–12.
20. Wu, Z.; Luo, G.; Yang, Z.; Guo, Y.; Li, K.; Xue, Y. A comprehensive review on deep learning approaches in wind forecasting applications. *CAAI Trans. Intell. Technol.* 2022, 7, 129–143.
21. Gasparin, A.; Lukovic, S.; Alippi, C. Deep learning for time series forecasting: The electric load case. *CAAI Trans. Intell. Technol.* 2022, 7, 1–25.
22. Pinker, S. *The Stuff of Thought: Language as a Window into Human Nature*; Penguin: London, UK, 2007.
23. Andersson, L.G.; Trudgill, P. *Bad Language*; Blackwell/Penguin Books: London, UK, 1990.
24. Eshan, S.C.; Hasan, M.S. An application of machine learning to detect abusive bengali text. In *Proceedings of the 2017 20th International Conference of Computer and Information Technology (ICCIT)*, Dhaka, Bangladesh, 22–24 December 2017; pp. 1–6.
25. Akhter, S.; Abdhullah-Al-Mamun. Social media bullying detection using machine learning on Bangla text. In *Proceedings of the 2018 10th International Conference on Electrical and Computer Engineering (ICECE)*, Dhaka, Bangladesh, 20–22 December 2018; pp. 385–388.
26. Emon, E.A.; Rahman, S.; Banarjee, J.; Das, A.K.; Mitra, T. A deep learning approach to detect abusive bengali text. In *Proceedings of the 2019 7th International Conference on Smart*

- Computing & Communications (ICSCC), Sarawak, Malaysia, 28–30 June 2019; pp. 1–5.
27. Awal, M.A.; Rahman, M.S.; Rabbi, J. Detecting abusive comments in discussion threads using naïve bayes. In Proceedings of the 2018 International Conference on Innovations in Science, Engineering and Technology (ICISSET), Chittagong, Bangladesh, 27–28 October 2018; pp. 163–167.
 28. Hussain, M.G.; Al Mahmud, T. A technique for perceiving abusive bangla comments. *Green Univ. Bangladesh J. Sci. Eng.* 2019, 4, 11–18.
 29. Das, M.; Banerjee, S.; Saha, P.; Mukherjee, A. Hate Speech and Offensive Language Detection in Bengali. *arXiv* 2022, arXiv:2210.03479.
 30. Sazzed, S. Identifying vulgarity in Bengali social media textual content. *PeerJ Comput. Sci.* 2021, 7, e665.
 31. Jahan, M.; Ahamed, I.; Bishwas, M.R.; Shatabda, S. Abusive comments detection in Bangla-English code-mixed and transliterated text. In Proceedings of the 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET), Dhaka, Bangladesh, 23–24 December 2019; pp. 1–6.
 32. Ishmam, A.M.; Sharmin, S. Hateful speech detection in public facebook pages for the bengali language. In Proceedings of the 2019 18th IEEE International Conference on Machine Learning and Applications (ICMLA), Boca Raton, FL, USA, 16–19 December 2019; pp. 555–560.
 33. Karim, M.R.; Dey, S.K.; Islam, T.; Sarker, S.; Menon, M.H.; Hossain, K.; Hossain, M.A.; Decker, S. Deephateexplainer: Explainable hate speech detection in under-resourced bengali language. In Proceedings of the 2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA), Porto, Portugal, 6–9 October 2021; pp. 1–10.
 34. Sazzed, S. Abusive content detection in transliterated Bengali-English social media corpus. In Proceedings of the Fifth Workshop on Computational Approaches to Linguistic Code-Switching, Online, 11 June 2021; pp. 125–130.
 35. Faisal Ahmed, M.; Mahmud, Z.; Biash, Z.T.; Ryen, A.A.N.; Hossain, A.; Ashraf, F.B. Bangla Text Dataset and Exploratory Analysis for Online Harassment Detection. *arXiv* 2021, arXiv:2102.02478.
 36. Romim, N.; Ahmed, M.; Talukder, H.; Islam, S. Hate speech detection in the bengali language: A dataset and its baseline evaluation. In Proceedings of the International Joint Conference on Advances in Computational Intelligence, Dhaka, Bangladesh, 20–21 November 2020; Springer: Singapore, 2021; pp. 457–468.
 37. Islam, T.; Ahmed, N.; Latif, S. An evolutionary approach to comparative analysis of detecting Bangla abusive text. *Bull. Electr. Eng. Inform.* 2021, 10, 2163–2169.

38. Aurpa, T.T.; Sadik, R.; Ahmed, M.S. Abusive Bangla comments detection on Facebook using transformer-based deep learning models. *Soc. Netw. Anal. Min.* 2022, 12, 24.
39. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, MI, USA, 2–7 June 2019; Association for Computational Linguistics: Minneapolis, MI, USA, 2019; pp. 4171–4186.
40. Clark, K.; Luong, M.T.; Le, Q.V.; Manning, C.D. Electra: Pre-training text encoders as discriminators rather than generators. In *Proceedings of the International Conference on Learning Representations, ICLR 2020, Virtual, 26 April–1 May 2020*.

Retrieved from <https://encyclopedia.pub/entry/history/show/116892>