Leveraging Multi-Task Learning for Detecting Aggression, Emotion, Violence, and Sentiment in Bengali Texts

Shamima Afroz
u1904106@student.cuet.ac.bd

Kawsar Ahmed
kawsarcse18@gmail.com

Mohammed Moshiul Hoque
moshiul_240@cuet.ac.bd
Department of Computer Science and Engineering,
Chittagong University of Engineering and Technology

Abstract

Despite remarkable advances in text classification (TC) for high-resource languages, progress in resource-constrained languages such as Bengali remains limited by the scarcity of standardized corpora, domain adaptation protocols, and robust pre-trained models. We introduce **MTL-MuRIL**, a transformer-based Multi-Task Learning (MTL) framework that jointly learns four interrelated classification tasks—aggression detection, emotion classification, violence detection, and sentiment analysis—within Bengali texts. Our approach leverages shared linguistic representations across tasks to improve generalization and mitigate overfitting in low-resource settings. Comprehensive experiments show that MTL-MuRIL consistently outperforms single-task baselines, achieving F1-scores of 0.893 ($\pm$0.005) for aggression detection, 0.743 ($\pm$0.030) for sentiment analysis, 0.717 ($\pm$0.015) for violence detection, and 0.570 ($\pm$0.020) for emotion classification. These results underscore the effectiveness of multi-task learning for enhancing Bengali text understanding and point toward a scalable paradigm for multilingual low-resource NLP.

1 Introduction

With increasing internet use, people now share their thoughts and emotions on social media platforms and online forums. However, these platforms also face growing abusive content, such as hate speech, cyberbullying, aggression, and obscenity. Understanding human affect in text involves analyzing sentiment and emotion. Sentiment analysis is utilized in customer feedback, product reviews, and the tracking of political opinions. Emotion classification is valuable for mental health monitoring, detecting social media abuse, and improving human–computer interaction. Together, these approaches provide a comprehensive picture of textual affect. Beyond sentiment and emotion analysis, aggression and violence detection address severe, harmful content. Aggression detection targets hostile expressions, while violence detection focuses on content that depicts or promotes harm. Both are vital for automated content moderation, supplementing insights from sentiment and emotion analysis.

The automatic detection of harmful Bengali text is critical. However, this is challenging due to the lack of annotated corpora, limited resources, and underdeveloped language-processing tools. Most studies in Bengali focus on a single domain or task, such as sentiment analysis.

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

However, no work has addressed multiple domains or tasks simultaneously. To address this gap, this work proposes a multi-task learning (MTL) approach: **MTL-MuRIL (Multi-Tasking MuRIL)**. This model classifies Bengali texts into four domains: emotion, aggression, sentiment, and violence. Research shows that MTL enhances performance by enabling knowledge sharing. Using a shared encoder, shared attention, and task-specific outputs, the model learns common patterns across different tasks. This enhances generalization and allows the detection of harmful and emotionally charged content. This setup makes the proposed MTL method more data-efficient and contextually aware compared to single-task models.

2 Related Work

Several studies in high-resource languages (HRLs) have explored MTL and related methods for sentiment analysis, aggression detection, hate speech identification, and emotion recognition. Ghosh et al. [1] proposed a transformer-based MTL framework using XLM-R to jointly detect aggression and hate across four languages with datasets such as TRAC-2020 and HASOC-2019, achieving 77.76% accuracy for Hindi aggression detection and 68.62% for English. A recent study [2] presented MTLsent+emo using BETO with sentiment and emotion as auxiliary tasks, reporting 77.03% F1 on HatEval and 86.58% macro-F1 on MEX-A3T. Xingyi et al. [3] applied a BERT-based MTL model with BiSRU++ and attention for cyberbullying detection on the Kaggle toxic comment dataset (159k samples), achieving an F1 score of 0.86. Tan et al. [4] proposed a BiLSTM-based MTL model for sentiment and sarcasm detection, obtaining a 0.94 F1 score. In low-resource languages (LRLs), few researchers have also explored these four tasks. Kancharla et al. [5] proposed MTTL-Hing-RoBERTa for aggression and offensive text detection in Hindi-English code-mixed tweets, achieving 69.10% accuracy. Hider et al. [6] conducted emotion classification on 5,753 Bangla-English code-mixed texts, reporting 81.75 F1 with BanglaBERT-1. Das et al. [7] studied Bangla emotion classification using an ensemble model, reaching 80.24 F1. These studies highlight the increasing application of multilingual and code-mixed MTL approaches for classification tasks in under-resourced settings.

Recent works in Bangla NLP have specifically addressed sentiment analysis, aggression, violence, and emotion classification. For aggression detection, Hossain et al. [8] developed a dataset of 13,728 texts across five classes, with BanglaBERT achieving an F1 score of 0.89. In contrast, Sharif et al. [9] introduced the 15,650-text M-BAD dataset for multilabel aggression detection, reporting an F1 score of 0.92 with BanglaBERT. For violence detection, Alamgir et al. [10] built a 4,011-text dataset, achieving 71.68 F1 with BanglaBERT, and Hossain et al. [11] achieved 74.59 F1 using GAN+Bangla-ELECTRA on 6,046 texts. Emotion classification has been explored by Hider et al. [6] and Avishek et al. [7], while sentiment analysis was studied by Mukit et al. [12], who used SVM on 4,000 Bangla texts and achieved 82% accuracy. Pal et al. [13] further demonstrate the effectiveness of a cross-task bilingual framework for sentiment and emotion classification on 32.5K Bangla-English texts, achieving 38.92 (Bangla emotion), 71.14 (Bangla sentiment), 83.48 (English emotion), and 67.09 (English sentiment). Despite these advances, Bangla NLP research still relies primarily on single-task approaches. Most studies focus separately on sentiment, aggression, violence, or emotion classification. Only a few studies, such as Hider et al. [14] and Datta et al. [15], have explored MTL in Bangla. However, these remain restricted to narrow domains, such as aspect-based sentiment. To the best of our knowledge, no comprehensive MTL framework has been developed that jointly addresses all four tasks in Bangla. To fill this gap, we propose a transformer-based MTL framework that integrates sentiment analysis, aggression detection, emotion classification, and violence detection. This enables cross-task knowledge sharing and improved performance in low-resource Bangla settings.

3 Datasets

To train and evaluate the proposed MTL model, we constructed a dataset accumulating data from publicly available sources for each task: aggression [16], emotion [17], violence [18], and sentiment classifications [19]) by sampling from publicly available resources. We

preprocessed the accumulated data to remove substantial noise, including extraneous punctuation, numerals, special characters, and non-Bengali words. We eliminated this noise using Python libraries for text normalization, such as NLTK and BanglaNLP for tokenization and stop-word removal, and pandas for efficient dataset handling. After preprocessing, we created balanced data splits for each task. For all tasks, we used 10,800 samples for training, 2496 for validation, and 2500 for testing, ensuring uniform splits across all datasets. To maintain fairness during batch processing, we standardized dataset sizes by identifying the task (sentiment, violence, emotion, aggression) with the fewest samples. This smallest dataset determined split sizes and enabled direct and fair comparisons across all four tasks. This strategy ensured methodological consistency throughout the experiments. Table 1 summarizes the distribution of task datasets in terms of classes, train, development, and test sets. The portion of the datasets used in our experiments is available on GitHub[1].

Table 1: Dataset splits for training, evaluation, and testing

Dataset	Class	Train	Dev	Test	Total
Emotion	Anger	450	67	71	588
	Disgust	450	155	165	770
	Fear	450	89	83	622
	Joy	450	120	114	684
	Sadness	450	129	119	698
	Surprise	450	64	73	587
Subtotal		2700	624	625	3949
Aggression	Aggressive	1350	312	312	1974
	Non-aggressive	1350	312	313	1975
Subtotal		2700	624	625	3949
Sentiment	Positive	1350	312	313	1975
	Negative	1350	312	312	1974
Subtotal		2700	624	625	3949
Violence	Direct	389	196	201	786
	Non-violence	1389	214	212	1815
	Passive	922	214	212	1348
Subtotal		2700	624	625	3949
Total		10800	2496	2500	15796

4 Baseline Models

To establish strong baselines, we experimented with several deep learning and pre-trained language models (PLMs) that support Bangla and multilingual text. Specifically, we evaluated 10 baselines: Bi-GRU, Bi-LSTM, Bi-GRU+CNN, LSTM+Bi-LSTM, LSTM+Bi-LSTM+CNN, MuRIL, XLM-RoBERTa, IndicBERT, mBERT, and BanglaBERT, all of which are available via Hugging Face[2]. These transformer models were selected because they have been trained on large-scale multilingual or Bangla-specific corpora, making them well-suited for downstream tasks in Bangla [20]. To assess their effectiveness, we extended these PLMs with single-task and multi-task architectures for systematic comparison.

4.1 MTL-MuRIL Architecture

Figure 1 (shown in Appendix A) illustrates the proposed transformer-based MTL model. Among the evaluated PLMs, **MuRIL-BERT** consistently delivered superior accuracy and F1-score, making it the backbone of the architecture. Each sentence is first tokenized into input IDs and attention masks, ensuring padding tokens are excluded from computation. Following tokenization, the inputs are processed by the MuRIL transformer, yielding contextual embeddings that act as the shared feature space for subsequent tasks. These shared embeddings are then directed to four task-specific classification heads: Emotion, Violence,

[1] `https://github.com/shamima-afroz/Multi-Task-Learning`

[2] `https://huggingface.co/`

Aggression, and Sentiment. Each branch first applies Multi-Head Attention with residual connections, followed by a lightweight feed-forward network and attention-based pooling implemented via a custom Keras AttentionLayer, which computes attention scores using the tanh function, normalizes them with softmax, and aggregates the input vectors into a summary vector that reflects key information. Finally, the model produces outputs for each task: aggression and sentiment tasks are binary, violence has three classes, and emotion is classified into six categories. Table 2 provides the tuned hyperparameters of the transformer-based MTL model.

Table 2: Hyperparameter values of the MTL transformer model

Parameter	**Value**
Learning Rate	1×10^{-6}
Batch Size	4
Max Sequence Length	100
Epochs	15, 8
Class Weights	Balanced
Activation Function	ReLU, Softmax
Optimizer	Adam
Embedding Dimension	768, 1024
Normalization	LayerNorm ($\epsilon = 1 \times 10^{-6}$)
Loss Function	Sparse CCE, Weighted Sparse CCE

5 Result Analysis

Table 3 shows the evaluation results of the baseline models and the proposed MTL models. The results indicate that MTL-MuRIL performed better than all other models overall. For the single-task models, Bi-GRU + CNN achieved the best Weighted F1 (WF) scores for aggression detection (0.853 ± 0.013), emotion classification (0.473 ± 0.013), and violence detection (0.627 ± 0.034). In contrast, MuRIL achieved the highest sentiment analysis WF score (0.723 ± 0.007). In the multi-task setup, MTL-MuRIL outperformed both single-task models across all four tasks: aggression detection (0.893 ± 0.005), emotion classification (0.570 ± 0.020), sentiment analysis (0.743 ± 0.030), and violence detection (0.717 ± 0.015). Compared to the best single-task models (Bi-GRU + CNN for aggression, emotion, and violence tasks; MuRIL for sentiment), MTL-MuRIL showed relative improvements of 4.69% in aggression detection, 20.51% in emotion classification, 2.77% in sentiment analysis, and 14.36% in violence detection. The smaller improvement in sentiment analysis may be due to negative transfer in multi-task learning. Aggression, violence, and emotion tasks often rely on clear, strong linguistic cues (e.g., hostile or emotional words), whereas sentiment analysis relies on more subtle, context-dependent expressions. When the model shares representations across tasks, it may focus more on the stronger signals that help the other three tasks, paying less attention to the fine details needed for sentiment polarity classification. Overall, these findings demonstrate the benefits of training a single model on multiple datasets. The shared attention layer in MTL-MuRIL helps integrate knowledge across tasks, leading to substantial improvements across three tasks with only a minor compromise in sentiment analysis.

5.1 Statistical Significance Analysis

Table 4 presents the results for all tasks in terms of Precision (P), Recall (R), Weighted F1-score (WF), and Accuracy (A).

Tables 5 and 6 present t-test results comparing MTL-MuRIL with Bi-GRU+CNN and single-task MuRIL across all tasks and metrics. Positive t-statistics and p-values below 0.05 indicate that MTL-MuRIL significantly outperforms both baselines, with notable gains in Aggression, Emotion, and Violence detection, and minor improvements in Sentiment classification.

Table 3: Performance of single-task and multi-task models on four classification tasks

Classifier	Single Task							
	Aggressive		Emotion		Sentiment		Violence	
	WF	Acc	WF	Acc	WF	Acc	WF	Acc
Bi-LSTM	0.842 ± 0.008	0.842 ± 0.008	0.412 ± 0.004	0.416 ± 0.013	0.663 ± 0.005	0.660 ± 0.011	0.579 ± 0.007	0.585 ± 0.021
Bi-GRU	0.841 ± 0.003	0.841 ± 0.003	0.452 ± 0.006	0.444 ± 0.012	0.621 ± 0.012	0.620 ± 0.014	0.592 ± 0.004	0.592 ± 0.004
Bi-GRU+CNN	**0.853 ± 0.013**	0.853 ± 0.013	**0.473 ± 0.013**	0.470 ± 0.008	0.680 ± 0.008	0.680 ± 0.008	**0.627 ± 0.034**	0.627 ± 0.034
LSTM+Bi-LSTM	0.801 ± 0.013	0.801 ± 0.013	0.392 ± 0.005	0.411 ± 0.014	0.669 ± 0.007	0.667 ± 0.011	0.616 ± 0.017	0.615 ± 0.020
LSTM+Bi-LSTM+CNN	0.842 ± 0.005	0.842 ± 0.005	0.433 ± 0.006	0.434 ± 0.012	0.669 ± 0.009	0.668 ± 0.014	0.617 ± 0.008	0.617 ± 0.008
MuRIL	0.827 ± 0.005	0.830 ± 0.000	0.383 ± 0.075	0.433 ± 0.060	**0.723 ± 0.007**	0.723 ± 0.007	0.567 ± 0.036	0.580 ± 0.028
mBERT	0.829 ± 0.003	0.833 ± 0.015	0.332 ± 0.004	0.351 ± 0.018	0.700 ± 0.005	0.699 ± 0.002	0.556 ± 0.008	0.559 ± 0.001
IndicBERT	0.751 ± 0.009	0.754 ± 0.004	0.282 ± 0.006	0.312 ± 0.012	0.621 ± 0.011	0.622 ± 0.007	0.479 ± 0.029	0.486 ± 0.015
BanglaBERT	0.783 ± 0.007	0.781 ± 0.014	0.282 ± 0.011	0.299 ± 0.013	0.655 ± 0.012	0.653 ± 0.007	0.522 ± 0.015	0.532 ± 0.011
XLM-Roberta	0.801 ± 0.010	0.801 ± 0.010	0.374 ± 0.009	0.380 ± 0.013	0.668 ± 0.007	0.675 ± 0.011	0.539 ± 0.004	0.544 ± 0.014
Classifier	**Multitask**							
	Aggressive		**Emotion**		**Sentiment**		**Violence**	
	WF	**Acc**	**WF**	**Acc**	**WF**	**Acc**	**WF**	**Acc**
MTL-MuRIL (Proposed)	**0.893 ± 0.005**	0.893 ± 0.005	**0.570 ± 0.020**	0.567 ± 0.020	**0.743 ± 0.030**	0.750 ± 0.025	**0.717 ± 0.015**	0.717 ± 0.015
MTL (mBERT)	0.829 ± 0.012	0.834 ± 0.014	0.374 ± 0.011	0.380 ± 0.013	0.654 ± 0.007	0.651 ± 0.013	0.573 ± 0.004	0.579 ± 0.015
MTL (IndicBERT)	0.729 ± 0.006	0.732 ± 0.014	0.423 ± 0.010	0.435 ± 0.013	0.532 ± 0.005	0.569 ± 0.011	0.556 ± 0.004	0.559 ± 0.015
MTL (BanglaBERT)	0.824 ± 0.010	0.822 ± 0.014	0.312 ± 0.003	0.309 ± 0.007	0.665 ± 0.002	0.669 ± 0.004	0.481 ± 0.002	0.498 ± 0.003
MTL (XLM-Roberta)	0.847 ± 0.022	0.847 ± 0.022	0.541 ± 0.009	0.534 ± 0.002	0.684 ± 0.012	0.692 ± 0.014	0.669 ± 0.018	0.669 ± 0.018
MTL (LSTM+Bi-LSTM+CNN)	0.840 ± 0.012	0.840 ± 0.012	0.473 ± 0.012	0.470 ± 0.011	0.683 ± 0.013	0.680 ± 0.008	0.680 ± 0.008	0.680 ± 0.008
MTL (Bi-GRU+CNN)	0.832 ± 0.011	0.834 ± 0.013	0.493 ± 0.011	0.497 ± 0.013	0.625 ± 0.014	0.622 ± 0.011	0.594 ± 0.012	0.593 ± 0.014
MTL (Bi-GRU)	0.844 ± 0.010	0.844 ± 0.010	0.458 ± 0.009	0.454 ± 0.011	0.611 ± 0.013	0.611 ± 0.010	0.591 ± 0.012	0.606 ± 0.011
MTL (LSTM+Bi-LSTM)	0.851 ± 0.001	0.852 ± 0.001	0.470 ± 0.013	0.495 ± 0.012	0.634 ± 0.002	0.634 ± 0.002	0.571 ± 0.004	0.571 ± 0.004
MTL (Bi-LSTM)	0.793 ± 0.012	0.793 ± 0.012	0.381 ± 0.012	0.393 ± 0.023	0.598 ± 0.023	0.596 ± 0.031	0.448 ± 0.019	0.486 ± 0.037

Table 4: Performance of single-task and multi-task models on four classification tasks

Classifier	Single Task															
	Aggressive				Emotion				Sentiment				Violence			
	P	R	WF	Acc	P	R	WF	Acc	P	R	WF	Acc	P	R	WF	Acc
Bi-LSTM	0.843 ± 0.007	0.842 ± 0.008	0.842 ± 0.008	0.842 ± 0.008	0.421 ± 0.003	0.418 ± 0.006	0.412 ± 0.004	0.416 ± 0.013	0.661 ± 0.009	0.658 ± 0.014	0.663 ± 0.005	0.660 ± 0.011	0.611 ± 0.004	0.582 ± 0.008	0.579 ± 0.007	0.585 ± 0.021
Bi-GRU	0.841 ± 0.003	0.841 ± 0.003	0.841 ± 0.003	0.841 ± 0.003	0.472 ± 0.005	0.443 ± 0.007	0.452 ± 0.006	0.444 ± 0.012	0.623 ± 0.011	0.619 ± 0.008	0.621 ± 0.012	0.620 ± 0.014	0.603 ± 0.007	0.592 ± 0.004	0.592 ± 0.004	0.592 ± 0.004
Bi-GRU+CNN	0.857 ± 0.013	0.853 ± 0.013	**0.853 ± 0.013**	0.853 ± 0.013	0.50 ± 0.008	0.48 ± 0.014	**0.473 ± 0.013**	0.470 ± 0.008	0.683 ± 0.013	0.680 ± 0.008	0.680 ± 0.008	0.680 ± 0.008	0.67 ± 0.014	0.627 ± 0.034	**0.627 ± 0.034**	0.627 ± 0.034
LSTM+Bi-LSTM	0.822 ± 0.007	0.801 ± 0.013	0.801 ± 0.013	0.801 ± 0.013	0.401 ± 0.003	0.398 ± 0.006	0.392 ± 0.005	0.411 ± 0.014	0.671 ± 0.009	0.668 ± 0.012	0.669 ± 0.007	0.667 ± 0.011	0.642 ± 0.006	0.621 ± 0.008	0.616 ± 0.017	0.615 ± 0.020
LSTM+Bi-LSTM+CNN	0.842 ± 0.005	0.842 ± 0.005	0.842 ± 0.005	0.842 ± 0.005	0.442 ± 0.005	0.431 ± 0.007	0.433 ± 0.006	0.434 ± 0.012	0.671 ± 0.011	0.667 ± 0.008	0.669 ± 0.009	0.668 ± 0.014	0.652 ± 0.007	0.611 ± 0.004	0.617 ± 0.008	0.617 ± 0.008
MuRIL	0.833 ± 0.005	0.830 ± 0.000	0.827 ± 0.005	0.830 ± 0.000	0.507 ± 0.035	0.433 ± 0.060	0.383 ± 0.075	0.433 ± 0.060	0.727 ± 0.012	0.723 ± 0.007	**0.723 ± 0.007**	0.723 ± 0.007	0.60 ± 0.030	0.577 ± 0.032	0.567 ± 0.036	0.58 ± 0.028
mBERT	0.831 ± 0.007	0.834 ± 0.012	0.829 ± 0.003	0.833 ± 0.015	0.341 ± 0.002	0.328 ± 0.006	0.332 ± 0.004	0.351 ± 0.018	0.702 ± 0.009	0.698 ± 0.014	0.700 ± 0.005	0.699 ± 0.002	0.561 ± 0.001	0.558 ± 0.007	0.556 ± 0.008	0.559 ± 0.001
IndicBERT	0.752 ± 0.003	0.748 ± 0.007	0.751 ± 0.009	0.754 ± 0.004	0.291 ± 0.002	0.288 ± 0.004	0.282 ± 0.006	0.312 ± 0.012	0.623 ± 0.005	0.619 ± 0.008	0.621 ± 0.011	0.622 ± 0.007	0.491 ± 0.001	0.482 ± 0.003	0.479 ± 0.029	0.486 ± 0.015
BanglaBERT	0.782 ± 0.012	0.780 ± 0.009	0.783 ± 0.007	0.781 ± 0.014	0.291 ± 0.010	0.288 ± 0.006	0.282 ± 0.011	0.299 ± 0.013	0.652 ± 0.008	0.649 ± 0.014	0.655 ± 0.012	0.653 ± 0.007	0.551 ± 0.004	0.534 ± 0.009	0.522 ± 0.015	0.532 ± 0.011
XLM-Roberta	0.801 ± 0.010	0.801 ± 0.010	0.801 ± 0.010	0.801 ± 0.010	0.291 ± 0.008	0.382 ± 0.011	0.374 ± 0.009	0.380 ± 0.013	0.682 ± 0.012	0.671 ± 0.010	0.668 ± 0.007	0.675 ± 0.011	0.541 ± 0.009	0.545 ± 0.006	0.539 ± 0.004	0.544 ± 0.014
Classifier	**Multitask**															
	Aggressive				**Emotion**				**Sentiment**				**Violence**			
	P	**R**	**WF**	**Acc**	**P**	**R**	**WF**	**Acc**	**P**	**R**	**WF**	**Acc**	**P**	**R**	**WF**	**Acc**
MTL-MuRIL (Proposed)	0.893 ± 0.005	0.893 ± 0.005	**0.893 ± 0.005**	0.893 ± 0.005	0.613 ± 0.010	0.567 ± 0.025	**0.570 ± 0.020**	0.567 ± 0.020	0.760 ± 0.017	0.750 ± 0.025	**0.743 ± 0.030**	0.750 ± 0.025	0.723 ± 0.010	0.717 ± 0.015	**0.717 ± 0.015**	0.717 ± 0.015
MTL (mBERT)	0.831 ± 0.010	0.833 ± 0.007	0.829 ± 0.012	0.834 ± 0.014	0.381 ± 0.009	0.372 ± 0.008	0.374 ± 0.011	0.380 ± 0.013	0.652 ± 0.012	0.649 ± 0.010	0.654 ± 0.007	0.651 ± 0.013	0.591 ± 0.011	0.582 ± 0.006	0.573 ± 0.004	0.579 ± 0.015
MTL (IndicBERT)	0.731 ± 0.008	0.733 ± 0.012	0.729 ± 0.006	0.732 ± 0.014	0.452 ± 0.009	0.411 ± 0.007	0.423 ± 0.010	0.435 ± 0.013	0.603 ± 0.012	0.572 ± 0.008	0.532 ± 0.005	0.569 ± 0.011	0.561 ± 0.007	0.558 ± 0.009	0.556 ± 0.004	0.559 ± 0.015
MTL (BanglaBERT)	0.823 ± 0.019	0.821 ± 0.005	0.824 ± 0.010	0.822 ± 0.014	0.361 ± 0.013	0.358 ± 0.015	0.312 ± 0.003	0.309 ± 0.007	0.692 ± 0.006	0.671 ± 0.005	0.665 ± 0.002	0.669 ± 0.004	0.543 ± 0.011	0.502 ± 0.002	0.481 ± 0.002	0.498 ± 0.003
MTL (XLM-Roberta)	0.847 ± 0.022	0.847 ± 0.022	0.847 ± 0.022	0.847 ± 0.022	0.552 ± 0.010	0.556 ± 0.014	0.541 ± 0.009	0.534 ± 0.002	0.718 ± 0.015	0.691 ± 0.015	0.684 ± 0.012	0.692 ± 0.014	0.667 ± 0.011	0.664 ± 0.013	0.669 ± 0.018	0.669 ± 0.018
MTL (LSTM+Bi-LSTM+CNN)	0.840 ± 0.012	0.840 ± 0.012	0.840 ± 0.012	0.840 ± 0.012	0.472 ± 0.010	0.471 ± 0.011	0.473 ± 0.012	0.470 ± 0.011	0.661 ± 0.013	0.662 ± 0.012	0.660 ± 0.014	0.663 ± 0.013	0.620 ± 0.012	0.581 ± 0.011	0.572 ± 0.012	0.583 ± 0.013
MTL (Bi-GRU+CNN)	0.830 ± 0.012	0.831 ± 0.014	0.832 ± 0.011	0.834 ± 0.013	0.491 ± 0.010	0.512 ± 0.012	0.493 ± 0.011	0.497 ± 0.013	0.621 ± 0.013	0.626 ± 0.012	0.625 ± 0.014	0.622 ± 0.011	0.620 ± 0.012	0.591 ± 0.013	0.594 ± 0.012	0.593 ± 0.014
MTL (Bi-GRU)	0.842 ± 0.012	0.847 ± 0.015	0.844 ± 0.010	0.844 ± 0.010	0.471 ± 0.008	0.453 ± 0.012	0.458 ± 0.009	0.454 ± 0.011	0.613 ± 0.010	0.614 ± 0.012	0.611 ± 0.013	0.611 ± 0.010	0.620 ± 0.011	0.605 ± 0.010	0.591 ± 0.012	0.606 ± 0.011
MTL (LSTM+Bi-LSTM)	0.851 ± 0.001	0.852 ± 0.001	0.851 ± 0.001	0.852 ± 0.001	0.541 ± 0.004	0.456 ± 0.003	0.470 ± 0.013	0.495 ± 0.012	0.645 ± 0.017	0.634 ± 0.002	0.634 ± 0.002	0.634 ± 0.002	0.600 ± 0.005	0.571 ± 0.004	0.571 ± 0.004	0.571 ± 0.004
MTL (Bi-LSTM)	0.793 ± 0.012	0.793 ± 0.012	0.793 ± 0.012	0.793 ± 0.012	0.390 ± 0.010	0.391 ± 0.031	0.381 ± 0.012	0.393 ± 0.023	0.591 ± 0.015	0.593 ± 0.013	0.598 ± 0.023	0.596 ± 0.031	0.512 ± 0.031	0.481 ± 0.022	0.448 ± 0.019	0.486 ± 0.037

Table 5: MTL-MuRIL vs Bi-GRU+CNN: t-test results for each task and metric

Task	**Precision (P)**		**Recall (R)**		**Weighted F1 (WF)**		**Accuracy (Acc)**	
	t-statistic	**p-value**	**t-statistic**	**p-value**	**t-statistic**	**p-value**	**t-statistic**	**p-value**
Aggressive	4.48	0.01102	4.97	0.00763	4.97	0.00763	4.97	0.00763
Emotion	15.28	0.00011	5.26	0.00626	7.04	0.00214	7.80	0.00146
Sentiment	6.23	0.00338	4.62	0.00989	3.51	0.02457	4.62	0.00989
Violence	5.34	0.00594	4.19	0.01375	4.19	0.01375	4.19	0.01375

Table 6: MTL-MuRIL vs MuRIL: t-test results for each task

Task	**Precision (P)**		**Recall (R)**		**Weighted F1 (WF)**		**Accuracy (Acc)**	
	t-statistic	**p-value**	**t-statistic**	**p-value**	**t-statistic**	**p-value**	**t-statistic**	**p-value**
Aggressive	14.70	0.00012	21.82	0.00003	16.17	0.00009	21.82	0.00003
Emotion	5.04	0.00726	3.57	0.02336	4.17	0.01400	3.67	0.02140
Sentiment	2.75	0.05154	1.80	0.14601	1.12	0.32372	1.80	0.14601
Violence	6.74	0.00253	6.86	0.00236	6.66	0.00264	7.47	0.00172

6 Conclusion

This work presents a multi-task learning framework (**MTL-MuRIL**) that leverages transformer-based models to perform multiple tasks concurrently, including emotion classification, aggression detection, violence identification, and sentiment classification. The results demonstrate that the proposed approach is practical in handling multiple classification tasks simultaneously, compared to training separate models for each task. This work lays a strong foundation for robust, scalable multi-task classification models for the Bengali language. Although the model performs well, certain limitations, such as limited dataset size, lack of explainability and generalizability, and high computational requirements, remain. Future work will address these issues and explore LLM-based approaches to achieve improved performance in multi-task learning.

Acknowledgement

This work was supported by the Directorate of Research and Extension and the NLP Lab, Chittagong University of Engineering & Technology (CUET), Chattogram, Bangladesh.

References

[1] Soumitra Ghosh, Amit Priyankar, Asif Ekbal, and Pushpak Bhattacharyya. A transformer-based multi-task framework for joint detection of aggression and hate on social media data. *Natural Language Engineering*, 29(6):1495–1515, 2023. doi: 10.1017/S1351324923000104.

[2] Flor Miriam Plaza-Del-Arco, M. Dolores Molina-González, L. Alfonso Ureña-López, and María Teresa Martín-Valdivia. A multi-task learning approach to hate speech detection leveraging sentiment analysis. *IEEE Access*, 9:112478–112489, 2021. doi: 10.1109/ACCESS.2021.3103697.

[3] Guo Xingyi and H. Adnan. Potential cyberbullying detection in social media platforms based on a multi-task learning framework. *International Journal of Data and Network Science*, 8(1):25–34, 2024.

[4] Yik Yang Tan, Chee-Onn Chow, Jeevan Kanesan, Joon Huang Chuah, and YongLiang Lim. Sentiment analysis and sarcasm detection using deep multi-task learning. *Wireless Personal Communications*, 129(3):2213–2237, 2023. doi: 10.1007/s11277-023-10235-4. URL `https://doi.org/10.1007/s11277-023-10235-4`.

[5] Bharath Kancharla, Prabhjot Singh, Lohith Bhagavan Kancharla, Yashita Chama, and Raksha Sharma. Identifying aggression and offensive language in code-mixed tweets: A multi-task transfer learning approach. In Ruvan Weerasinghe, Isuri Anuradha, and Deshan Sumanathilaka, editors, *Proceedings of the First Workshop on Natural Language Processing for Indo-Aryan and Dravidian Languages*, Abu Dhabi, January 2025. Association for Computational Linguistics. URL `https://aclanthology.org/2025.indonlp-1.14/`.

[6] Md. Ali Hider, Shawly Ahsan, Jawad Hossain, and Mohammed Moshiul Hoque. Emotion classification in bengali-english code-mixed data using transformers. In *2024 27th International Conference on Computer and Information Technology (ICCIT)*, pages 3529–3535, 2024. doi: 10.1109/ICCIT64611.2024.11022096.

[7] Avishek Das, Mohammed Moshiul Hoque, Omar Sharif, M. Ali Akber Dewan, and Nazmul Siddique. Temox: Classification of textual emotion using ensemble of transformers. *IEEE Access*, 11:109803–109818, 2023.

[8] Jawad Hossain, Avishek Das, Mohammed Moshiul Hoque, and Nazmul Siddique. Multilabel aggressive text classification from social media using transformer-based approaches. In *2023 26th International Conference on Computer and Information Technology (ICCIT)*, pages 1–6, 2023.

[9] Omar Sharif, Eftekhar Hossain, and Mohammed Moshiul Hoque. M-BAD: A multilabel dataset for detecting aggressive texts and their targets. In Tanmoy Chakraborty, Md. Shad Akhtar, Kai Shu, H. Russell Bernard, Maria Liakata, Preslav Nakov, and Aseem Srivastava, editors, *Proceedings of the Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situations*, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.constraint-1.9. URL `https://aclanthology.org/2022.constraint-1.9/`.

[10] Refaat Mohammad Alamgir and Amira Haque. Team CentreBack at BLP-2023 task 1: Analyzing performance of different machine-learning based methods for detecting violence-inciting texts in Bangla. In Firoj Alam, Sudipta Kar, Shammur Absar Chowdhury, Farig Sadeque, and Ruhul Amin, editors, *Proceedings of the First Workshop on Bangla Language Processing (BLP-2023)*, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.banglalp-1.18. URL `https://aclanthology.org/2023.banglalp-1.18/`.

[11] Jawad Hossain, Hasan Mesbaul Ali Taher, Avishek Das, and Mohammed Moshiul Hoque. NLP_CUET at BLP-2023 task 1: Fine-grained categorization of violence inciting text using transformer-based approach. In Firoj Alam, Sudipta Kar, Shammur Absar Chowdhury, Farig Sadeque, and Ruhul Amin, editors, *Proceedings of the First Workshop on Bangla Language Processing (BLP-2023)*, pages 241–246, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.banglalp-1.31. URL https://aclanthology.org/2023.banglalp-1.31/.

[12] Mohammad Mukit, Md Moshiul Huq Rabbi, Mohammad Masud Ahmed, Subhenur Latif, and Sajib Saha. Sentiment analysis on bangla and phonetic bangla reviews: A product rating procedure using nlp and machine learning. In *2023 IEEE 9th International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, pages 433–437, 2023.

[13] Pritam Pal, Dipankar Das, and Anup Kolya. Bilingual sentiment and emotion analysis: A multi-task learning framework for bengali and english. In *Proceedings of [Conference Name]*.

[14] Sourav Saha, Shawly Ahsan, and Mohammed Moshiul Hoque. Multsea: Aspect-based sentiment analysis using multitask learning from bengali texts. In *2024 27th International Conference on Computer and Information Technology (ICCIT)*, pages 25–31, 2024. doi: 10.1109/ICCIT64611.2024.11021922.

[15] Nirjhor Datta, Md Hasanur Rashid, Samiur Rahman, Naima Tahsin Nodi, and Moin Uddin. *Deep Learning-Based Hybrid Multi-Task Model for Adrenocortical Carcinoma Segmentation and Classification*. PhD thesis, BRAC University, 2024.

[16] Omar Sharif and Moshiul Hoque. *Identification and Classification of Textual Aggression in Social Media: Resource Creation and Evaluation*, pages 9–20. April 2021. ISBN 978-3-030-73695-8. doi: 10.1007/978-3-030-73696-5_2.

[17] Avishek Das, Omar Sharif, Mohammed Moshiul Hoque, and Iqbal H. Sarker. Emotion classification in a resource constrained language using transformer-based approach. In Esin Durmus, Vivek Gupta, Nelson Liu, Nanyun Peng, and Yu Su, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, pages 150–158, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-srw.19. URL https://aclanthology.org/2021.naacl-srw.19/.

[18] Sourav Saha, Jahedul Alam Junaed, Maryam Saleki, Arnab Sen Sharma, Mohammad Rashidujjaman Rifat, Mohamed Rahouti, Syed Ishtiaque Ahmed, Nabeel Mohammed, and Mohammad Ruhul Amin. Vio-lens: A novel dataset of annotated social network posts leading to different forms of communal violence and its evaluation. In Firoj Alam, Sudipta Kar, Shammur Absar Chowdhury, Farig Sadeque, and Ruhul Amin, editors, *Proceedings of the First Workshop on Bangla Language Processing (BLP-2023)*, pages 72–84, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.banglalp-1.9. URL https://aclanthology.org/2023.banglalp-1.9/.

[19] Md. Arid Hasan, Shudipta Das, Afiyat Anjum, Firoj Alam, Anika Anjum, Avijit Sarker, and Sheak Rashed Haider Noori. Zero- and few-shot prompting with LLMs: A comparative study with fine-tuned models for Bangla sentiment analysis. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, Torino, Italia, May 2024. ELRA and ICCL. URL https://aclanthology.org/2024.lrec-main.1549/.

[20] Md Rajib Hossain, Sadia Afroze, Asif Ekbal, Mohammed Moshiul Hoque, and Nazmul Siddique. Multimodfusenet: Advancing multimodal text classification for low-resource languages through textual-visual feature fusion. *Knowledge-Based Systems*, 328:114085, 2025.

A Dataset Details

Table 7 presents the statistics of the original datasets from which we used a selected portion for our experiments. In total, we used 10,800 samples for training, 2,496 samples for evaluation, and 2,500 samples for testing across the four datasets. From each dataset, 2,700 samples were used for training, 624 for evaluation, and 625 for testing, for a total of 15,796 samples across all splits. Table 1 summarizes the dataset portions used in our experiments, while Table 8 shows the class-wise distribution of both the original datasets and the portions selected for our experiments. In Table 8, "total" means the total samples present in the actual dataset, and "used" means the portion we used for our experiment. For each dataset and class, the table reports the total number of samples in the original dataset and the number of samples used for training, evaluation, and testing. Since the original datasets were not perfectly balanced, we selected samples from the smallest class to maintain as much balance as possible. To achieve this, we randomly selected an equal number of samples from each class for training, evaluation, and testing. For example, in the Emotion dataset, we chose 450 training samples per class, while all available samples were used for assessment and testing when class sizes were smaller. In the Aggression and Sentiment datasets, larger classes were downsampled to 1,350 samples for training and 312 for evaluation. For the Violence dataset, the most common classes, such as Direct, were fully used, while other classes were sampled to keep the splits balanced.

Table 7: Original dataset distribution across training, evaluation, and test splits

Split	**Aggression[16]**	**Emotion[17]**	**Violence[18]**	**Sentiment[19]**
Train	11,326	4,994	2,700	19,153
Dev	1,416	624	1,330	2,826
Test	1,416	625	2,016	5,430

Table 8: Class-wise sample counts and experimentally selected portions

Dataset	**Class**	**Train**		**Dev**		**Test**	
		Total	Used	Total	Used	Total	Used
Emotion[17]	Anger	621	450	67	67	71	71
	Disgust	1233	450	155	155	165	165
	Fear	700	450	89	89	83	83
	Joy	908	450	120	120	114	114
	Sadness	942	450	129	129	119	119
	Surprise	590	450	64	64	73	73
Aggression[16]	Aggressive	5845	1350	769	312	737	312
	Non-aggressive	5481	1350	647	312	679	313
Sentiment[19]	Positive	7342	1350	1126	312	2092	313
	Negative	11811	1350	1700	312	3338	312
Violence[18]	Direct	196	196	196	196	201	201
	Non-violence	1389	1389	717	214	1069	212
	Passive	922	922	417	214	719	212

B Proposed MTL Architecture

The proposed architecture allows the model to exploit shared syntactic and semantic cues across Bangla text while tailoring predictions to task-specific label spaces. The proposed multi-task model employs a pre-trained **MuRIL transformer** as a shared encoder. It takes token IDs and attention masks as input and outputs token-level contextual embeddings, which serve as a shared feature space for four tasks: **Emotion, Violence, Aggression, and Sentiment**. For an input sequence X with attention mask A, the output of the encoder is:

$$H = \text{MuRIL}_\theta(X, A), \tag{1}$$

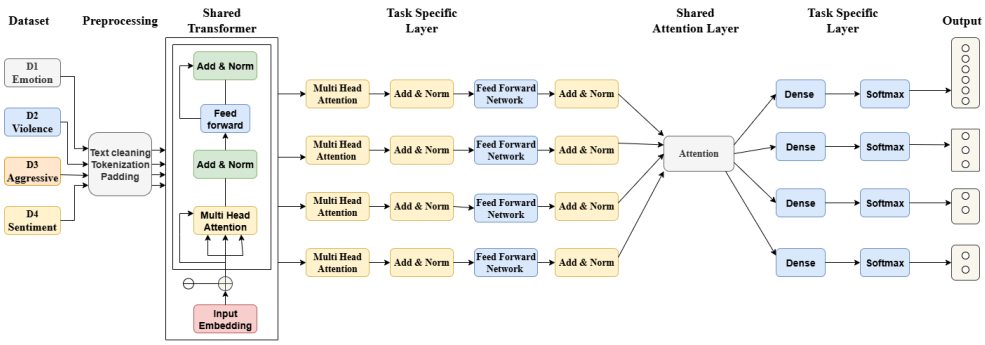


Figure 1: MTL-MuRIL: A Multi-Task Learning Framework

where θ denotes the pre-trained weights, $X \in \mathbb{R}^c$ represents the input token IDs of length c, and $A \in \mathbb{R}^c$ is the corresponding attention mask. The encoder outputs $H \in \mathbb{R}^{c \times d}$, containing contextual embeddings for all tokens, where the hidden dimension d. Each task receives its own embedding $H_{\text{emotion}}, H_{\text{violence}}, H_{\text{aggression}}, H_{\text{sentiment}}$. Here, $H_t \in \mathbb{R}^{c \times d}$ denotes the input hidden states for task t.

Then, each task branch applies Multi-Head Attention (MHA) with residual connections and layer normalization to capture diverse contextual dependencies, followed by a feed-forward network to model task-specific patterns. We obtain Z_t, the combined output of the multi-head attention mechanism.

$$Z_t = \text{LayerNorm}(\text{MHA}(H_t, H_t) + H_t) \tag{2}$$

Following MHA, each branch includes a feed-forward network (FFN) with two dense layers and ReLU activation, along with residual connections and layer normalization. Here, F_t indicates feature representation.

$$F_t' = \text{ReLU}(Z_t W_1 + b_1) W_2 + b_2 \tag{3}$$
$$F_t = \text{LayerNorm}(F_t' + Z_t) \tag{4}$$

To highlight critical input features, a custom attention layer performs attention-based pooling. This layer shares parameters W and b across tasks but computes attention independently for each task. Here, the intermediate score vector is computed as u_t:

$$u_t = \tanh(F_t W_u + b_u) \tag{5}$$
$$\alpha_t = \text{softmax}(u_t) \tag{6}$$
$$s_t = \sum_{i=1}^{c} \alpha_{t,i} F_{t,i} \tag{7}$$

The context vectors s_t are then passed through task-specific Dense layers with softmax activation to produce predictions. Aggression and Sentiment are binary tasks, Violence has three classes, and Emotion has six classes. y_t produces the probability distribution over the target classes.

$$o_t = W_t s_t + b_t \tag{8}$$
$$y_t = \text{softmax}(o_t) \tag{9}$$

Here K_t denotes the number of classes for task t. where N denotes the batch size, $y_i \in \{0, \ldots, K_t - 1\}$ is the true class label of the i-th sample, and $\hat{y}_{t,y_i}$ is the predicted probability

for the true class. For tasks $t \in \{\text{Emotion, Aggression, Sentiment}\}$, sparse categorical cross-entropy is used, and weighted sparse categorical cross-entropy is used for the Violence task to address class imbalance.

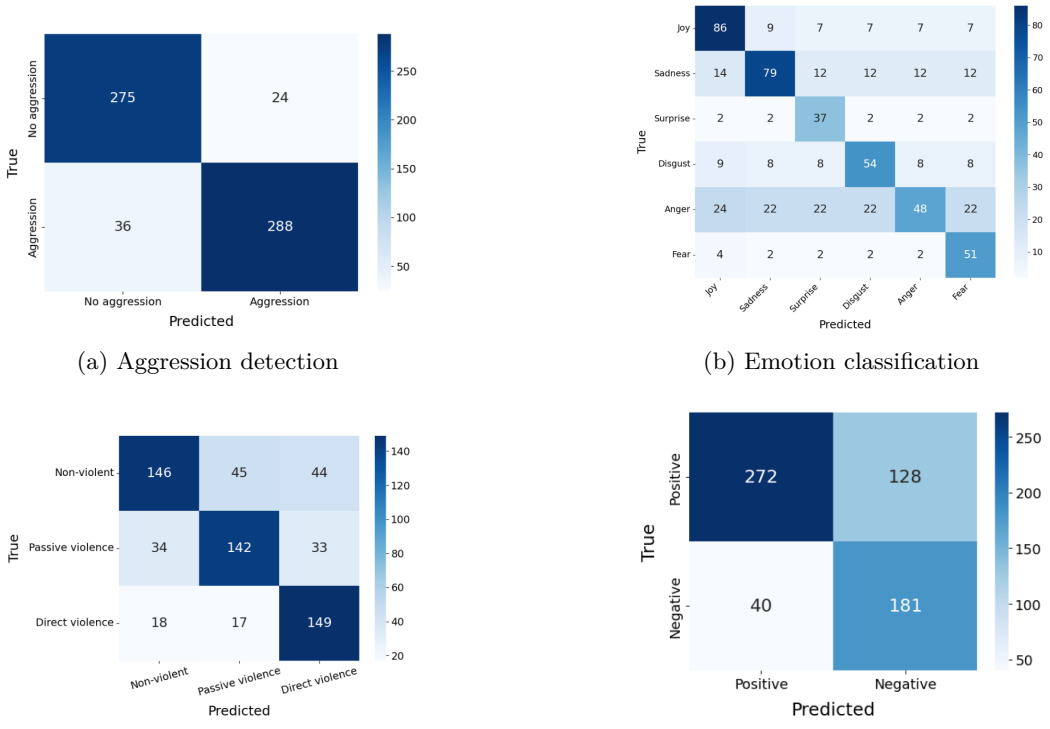
$$L_t = \frac{1}{N} \sum_{i=1}^{N} -\log \hat{y}_{t,y_i} \tag{10}$$

$$L_{\text{violence}} = \frac{1}{N} \sum_{i=1}^{N} -w_{y_i} \cdot \log \hat{y}_{\text{violence},y_i} \tag{11}$$

Where w_{y_i} is the class weight corresponding to the actual class of sample i.

C Error Analysis

Figure 2 shows the confusion matrix of each task for the proposed MTL-MuRIL model. In the emotion classification task, the model correctly predicted 354 samples and incorrectly predicted 271. For the violence detection, it correctly predicted 437 samples and incorrectly predicted 188. The model produced 563 correct predictions and 62 incorrect predictions for the aggression classification task. In terms of sentiment, the model received 456 correct and 169 incorrect predictions. Among these tasks, the highest misclassification occurs in emotion classification. This is because it is a multi-class problem with six categories, whereas aggression and sentiment are binary tasks, and violence detection has only three classes. Emotions also share semantic overlap across classes—for example, anger vs. disgust, fear vs. sadness—making them harder to distinguish. Class imbalance in the dataset further contributes to misclassification, as minority emotion classes are more difficult for the model to predict accurately.



(a) Aggression detection

(b) Emotion classification

(c) Violence detection

(d) Sentiment Analysis

Figure 2: Confusion matrix of the MTL-MuRIL method across four tasks.

Now, we move on to the qualitative analysis of the results. Table 9 presents several sample predictions from the Aggression, Sentiment, Emotion, and Violence datasets. Each example includes the original text, its actual label, and the corresponding predicted label generated by the proposed MTL-MuRIL model.

The model uses shared transformer layers for all tasks, with separate output heads for each task. Learning patterns for one task, such as aggression, helps the model perform better on related tasks, such as emotion and violence classification. This shared structure enables the model to leverage knowledge across tasks, thereby improving overall performance.

Table 9: Sample predictions on Aggression, Sentiment, Emotion, and Violence datasets

Dataset	**Text**	**Actual label**	**Predicted label**
Aggression	বিএনপি যদি, খালেদা ও হাসিনার উভয় পক্ষ যদি প্রকাশ্যে বিরোধ দেখাতো, তবে যাদের কাছে ছিল দেহ, তারা সেটি উন্মুক্তভাবে জানতো। (If BNP, and both Khaleda and Hasina had openly shown their opposition, then those who had the bodies would have known it publicly)	Non aggressive	Non aggressive
	আমি কি বারবার এই কথা বলবো(Do I have to say this again and again?)	Aggressive	Aggressive
	সরকারের চামচামি করে শত ধিক্কার জানাই, যেমন জি মামুনের মত সকল শয়তান সাংবাদিকদের, যাদের নজর উদ্ধ মুখী হয়ে গেছে (I strongly condemn the act of flattering the government, just like journalists such as G. Mamun, whose intentions have become corrupt and self-serving.)	Aggressive	Non-aggressive
	আপনি আসলেই একটা নাস্তিক (You really are an atheist.)	Aggressive	Non-aggressive
	পুলিশে দেওয়ার কথা উল্লেখ পাইলাম না, মাইরের উপর ডিমের মাইর কি চলবে না তাহলে?(I didn't find any mention of reporting to the police; so, is it not allowed to throw eggs at the scoundrel?)	Non-aggressive	Aggressive
Sentiment	বুড়ো হয়ে বুদ্ধি লোপ পেয়েছে (With old age, the wisdom has faded.)	Negative	Negative
	স্কুল-কলেজে সপ্তাহে একদিন ক্লাস নেওয়া হবে শিক্ষা উপমন্ত্রী (Classes will be held one day a week in schools and colleges Deputy Minister of Education.)	Positive	Positive
	আল্লাহ আপনাকে মাফ করুন, এমন মিথ্যা কথা বলার জন্য (May Allah forgive you for saying such lies.)	Positive	Negative
	সাবেক স্ত্রীর পরকীয়া নিয়ে মুখ খুললেন অপূর্ব (Apuvo spoke out about his ex-wife's affair.)	Negative	Positive
	গভীর রাতে গোপন সংবাদের ভিত্তিতে বিস্তারিত নিউজ (A detailed news report based on secret information late at night.)	Positive	Negative
Emotion	কোন মুভির নাম জানো না, আবার কথা বলো সব ভিডিও বানাও, কোথাকার বলদ! (You don't even know the name of the movie, yet you talk nonsense and make videos where are you from, idiot?)	Disgust	Disgust
	মনটি বিভিন্ন মানুষের মধ্যে নানা সমস্যার সৃষ্টি করে, নানা দিক থেকে চিন্তা জন্ম দেয়। (The mind gives rise to different problems in people and creates thoughts from various perspectives.)	Fear	Sad
	এদিকে বেসরকারি বিশ্ববিদ্যালয়গুলোর ক্লাস ও পরীক্ষা বন্ধ করা হলেও, কিছু বিশ্ববিদ্যালয়ে শিক্ষকদের অফিসে যেতে বলা হয়েছে। (Although private universities have suspended classes and exams, some have asked teachers to attend the office.)	Anger	Joy
	সিরিয়াসলি! এটা জলিল ভাইয়ের মুভি? কতটা পরিবর্তন! পুরাই অবাক! (Seriously! This is Jalil Bhai's movie? What a change! Totally surprised!)	Surprise	Joy
	ভদ্র মানুষ দেখতে চান? তাহসানকে দেখুন বুঝবেন মানুষ এত ভালো কীভাবে হতে পারে! লভ ইউ তাহসান (Want to see a gentleman? Look at Tahsan you'll realize how kind a person can be! Love you Tahsan)	Joy	Joy
	মেয়েটা অতিরিক্ত আত্মপরিচয় সংকটে ভুগছে সাবিলা স্যাবলাই রয়ে গেলো, বিউটি উইথআউট ব্রেইন। (The girl is suffering from an identity crisis Sabila remained the same, beauty without brain.)	Disgust	Sad
Violence	আইসিসির প্রধান কর্মকর্তা হচ্ছেন তিনি এবার কি হয়েছে, বিজয়? (He is becoming the chief officer of the ICC. So, what happened this time, Bijoy?)	Passive	Non-violence
	আল্লাহ এইসব অন্যায়ের বিচার করবেন। (Allah will judge all these matters.)	Non-violence	Non-violence
	মুজাহিদ ভাই এখন কোথায়? (Brother Mujahid, where are you now?)	Non-violence	Non-violence
	আমাদের দেশের ইসলাম এত সস্তা হয়ে গেছে যে, যার যেমন ইচ্ছে ব্যবহার করছে। (In our country, Islam has become so cheap that anyone uses it as they wish.)	Non-violence	Passive
	কাওকে মরলে খুশি হতাম এখনও কেউ কেন মরছে না! (I'd be happy if someone died why hasn't anyone died yet!)	Passive	Direct
	পুলিশ বাহিনী, আল্লাহ তোদের বিচার করবেন! (Police force, Allah will judge you!)	Non-violence	Passive