

First Workshop on Bangla Language Processing (BLP 2023)

Singapore
7 December 2023

ISBN: 978-1-7138-9045-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2023) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Offensive Language Identification in Transliterated and Code-Mixed Bangla</i> Md Nishat Raihan, Umma Hani Tanmoy, Anika Binte Islam, Kai North, Tharindu Ranasinghe, Antonios Anastasopoulos and Marcos Zampieri	1
<i>BSpell: A CNN-Blended BERT Based Bangla Spell Checker</i> Chowdhury Rafeed Rahman, MD.Hasibur Rahman, Samiha Zakir, Mohammad Rafsan and Mo- hammed Eunus Ali	7
<i>Advancing Bangla Punctuation Restoration by a Monolingual Transformer-Based Method and a Large- Scale Corpus</i> Mehedi Hasan Bijoy, Mir Fatema Afroz Faria, Mahbub E Sobhani, Tanzid Ferdoush and Swakkhar Shatabda	18
<i>Pipeline Enabling Zero-shot Classification for Bangla Handwritten Grapheme</i> Linsheng Guo, Md Habibur Rahman Sifat and Tashin Ahmed	26
<i>Low-Resource Text Style Transfer for Bangla: Data & Models</i> Sourabrata Mukherjee, Akanksha Bansal, Pritha Majumdar, Atul Kr Ojha and Ondrej Dusek	34
<i>Intent Detection and Slot Filling for Home Assistants: Dataset and Analysis for Bangla and Sylheti</i> Fardin Ahsan Sakib, A H M Rezaul Karim, Saadat Hasan Khan and Md Mushfiqur Rahman	48
<i>BEmoLexBERT: A Hybrid Model for Multilabel Textual Emotion Classification in Bangla by Combining Transformers with Lexicon Features</i> Ahasan Kabir, Animesh Chandra Roy and Zaima Sartaj Taheri	56
<i>Assessing Political Inclination of Bangla Language Models</i> Surendrabikram Thapa, Ashwarya Maratha, Khan Md Hasib, Mehwish Nasim and Usman Naseem 62	
<i>Vio-Lens: A Novel Dataset of Annotated Social Network Posts Leading to Different Forms of Communal Violence and its Evaluation</i> Sourav Saha, Jahedul Alam Junaed, Maryam Saleki, Arnab Sen Sharma, Mohammad Rashidu- jjaman Rifat, Mohamed Rahouti, Syed Ishtiaque Ahmed, Nabeel Mohammed and Mohammad Ruhul Amin	72
<i>BanglaCHQ-Summ: An Abstractive Summarization Dataset for Medical Queries in Bangla Conversa- tional Speech</i> Alvi Aveen Khan, Fida Kamal, Mohammad Abrar Chowdhury, Tasnim Ahmed, Md Tahmid Rah- man Laskar and Sabbir Ahmed	85
<i>Contextual Bangla Neural Stemmer: Finding Contextualized Root Word Representations for Bangla Words</i> Md Fahim, Amin Ahsan Ali, M Ashraful Amin and Akmmahbubur Rahman	94
<i>Investigating the Effectiveness of Graph-based Algorithm for Bangla Text Classification</i> Farhan Noor Dehan, Md Fahim, Amin Ahsan Ali, M Ashraful Amin and Akmmahbubur Rahman 104	
<i>SynthNID: Synthetic Data to Improve End-to-end Bangla Document Key Information Extraction</i> Syed Mostofa Monsur, Shariar Kabir and Sakib Chowdhury	117

<i>BaTEClacor: A Novel Dataset for Bangla Text Error Classification and Correction</i>	
Nabilah Tabassum Oshin, Syed Mohaiminul Hoque, Md Fahim, Amin Ahsan Ali, M Ashraful Amin and Akmmahbur Rahman	124
<i>Crosslingual Retrieval Augmented In-context Learning for Bangla</i>	
Xiaoqian Li, Ercong Nie and Sheng Liang	136
<i>Pseudo-Labeling for Domain-Agnostic Bangla Automatic Speech Recognition</i>	
Rabindra Nath Nandi, Mehadi Hasan Menon, Tareq Al Muntasir, Sagor Sarker, Quazi Sarwar Muhtaseem, Md. Tariqul Islam, Shammur Absar Chowdhury and Firoj Alam	152
<i>BanglaNLP at BLP-2023 Task 1: Benchmarking different Transformer Models for Violence Inciting Text Detection in Bangla</i>	
Saumajit Saha and Albert Aristotle Nanda	163
<i>Team CentreBack at BLP-2023 Task 1: Analyzing performance of different machine-learning based methods for detecting violence-inciting texts in Bangla</i>	
Refaat Mohammad Alamgir and Amira Haque	168
<i>EmptyMind at BLP-2023 Task 1: A Transformer-based Hierarchical-BERT Model for Bangla Violence-Inciting Text Detection</i>	
Udoy Das, Karnis Fatema, Md Ayon Mia, Mahshar Yahan, Md Sajidul Mowla, Md Fayeze Ullah, Arpita Sarker and Hasan Murad	174
<i>nlpBDpatriots at BLP-2023 Task 1: Two-Step Classification for Violence Inciting Text Detection in Bangla - Leveraging Back-Translation and Multilinguality</i>	
Md Nishat Raihan, Dhiman Goswami, Sadiya Sayara Chowdhury Puspo and Marcos Zampieri	179
<i>Score_IsAll_You_Need at BLP-2023 Task 1: A Hierarchical Classification Approach to Detect Violence Inciting Text using Transformers</i>	
Kawsar Ahmed, Md Osama, Md. Sirajul Islam, Md Taosiful Islam, Avishek Das and Mohammed Moshiul Hoque	185
<i>Mavericks at BLP-2023 Task 1: Ensemble-based Approach Using Language Models for Violence Inciting Text Detection</i>	
Saurabh Page, Sudeep Mangalvedhekar, Kshitij Deshpande, Tanmay Chavan and Sheetal S. Sonawane	190
<i>VacLM at BLP-2023 Task 1: Leveraging BERT models for Violence detection in Bangla</i>	
Shilpa Chatterjee, P J Leo Evenss and Pramit Bhattacharyya	196
<i>Aambela at BLP-2023 Task 1: Focus on UNK tokens: Analyzing Violence Inciting Bangla Text with Adding Dataset Specific New Word Tokens</i>	
Md Fahim	201
<i>SUST_Black Box at BLP-2023 Task 1: Detecting Communal Violence in Texts: An Exploration of MLM and Weighted Ensemble Techniques</i>	
Hrithik Majumdar Shibu, Shrestha Datta, Zhalok Rahman, Shahrab Khan Sami, Md. Sumon Miah, Raisa Fairouz and Md Adith Mollah	208
<i>the_linguists at BLP-2023 Task 1: A Novel Informal Bangla Fasttext Embedding for Violence Inciting Text Detection</i>	
Md. Tariquzzaman, Md Wasif Kader, Audwit Nafi Anam, Naimul Haque, Mohsinul Kabir, Hasan Mahmud and Md Kamrul Hasan	214

<i>UFAL-ULD at BLP-2023 Task 1: Violence Detection in Bangla Text</i> Sourabrata Mukherjee, Atul Kr Ojha and Ondrej Dusek	220
<i>Semantics Squad at BLP-2023 Task 1: Violence Inciting Bangla Text Detection with Fine-Tuned Transformer-Based Models</i> Krishno Dey, Prerona Tarannum, Md. Arid Hasan and Francis Palma	225
<i>LowResourceNLU at BLP-2023 Task 1 & 2: Enhancing Sentiment Classification and Violence Incitement Detection in Bangla Through Aggregated Language Models</i> Hariram Veeramani, Surendrabikram Thapa and Usman Naseem	230
<i>Team Error Point at BLP-2023 Task 1: A Comprehensive Approach for Violence Inciting Text Detection using Deep Learning and Traditional Machine Learning Algorithm</i> Rajesh Kumar Das, Jannatul Maowa, Moshfiqur Rahman Ajmain, Kabid Yeiad, Mirajul Islam and Sharun Akter Khushbu	236
<i>NLP_CUET at BLP-2023 Task 1: Fine-grained Categorization of Violence Inciting Text using Transformer-based Approach</i> Jawad Hossain, Hasan Mesbaul Ali Taher, Avishek Das and Mohammed Moshuiul Hoque	241
<i>Team_Syrax at BLP-2023 Task 1: Data Augmentation and Ensemble Based Approach for Violence Inciting Text Detection in Bangla</i> Omar Faruqe Riyad, Trina Chakraborty and Abhishek Dey	247
<i>BLP-2023 Task 1: Violence Inciting Text Detection (VITD)</i> Sourav Saha, Jahedul Alam Junaed, Maryam Saleki, Mohamed Rahouti, Nabeel Mohammed and Mohammad Ruhul Amin	255
<i>BanglaNLP at BLP-2023 Task 2: Benchmarking different Transformer Models for Sentiment Analysis of Bangla Social Media Posts</i> Saumajit Saha and Albert Aristotle Nanda	266
<i>Knowdee at BLP-2023 Task 2: Improving Bangla Sentiment Analysis Using Ensembled Models with Pseudo-Labeling</i> Xiaoyi Liu, Mao Teng, SHuangtao Yang and Bo Fu	273
<i>M1437 at BLP-2023 Task 2: Harnessing Bangla Text for Sentiment Analysis: A Transformer-based Approach</i> Majidur Rahman and Ozlem Uzuner	279
<i>nlpBDpatriots at BLP-2023 Task 2: A Transfer Learning Approach towards Bangla Sentiment Analysis</i> Dhiman Goswami, Md Nishat Raihan, Sadiya Sayara Chowdhury Puspo and Marcos Zampieri	286
<i>Ushoshi2023 at BLP-2023 Task 2: A Comparison of Traditional to Advanced Linguistic Models to Analyze Sentiment in Bangla Texts</i> Sharun Akter Khushbu, Nasheen Nur, Mohiuddin Ahmed and Nashtarin Nur	293
<i>EmptyMind at BLP-2023 Task 2: Sentiment Analysis of Bangla Social Media Posts using Transformer-Based Models</i> Karnis Fatema, Udoy Das, Md Ayon Mia, Md Sajidul Mowla, Mahshar Yahan, Md Fayeze Ullah, Arpita Sarker and Hasan Murad	300
<i>RSM-NLP at BLP-2023 Task 2: Bangla Sentiment Analysis using Weighted and Majority Voted Fine-Tuned Transformers</i> Pratinav Seth, Rashi Goel, Komal Mathur and Swetha Vemulapalli	305

<i>Semantics Squad at BLP-2023 Task 2: Sentiment Analysis of Bangla Text with Fine Tuned Transformer Based Models</i>	
Krishno Dey, Md. Arid Hasan, Prerona Tarannum and Francis Palma	312
<i>Aambela at BLP-2023 Task 2: Enhancing BanglaBERT Performance for Bangla Sentiment Analysis Task with In Task Pretraining and Adversarial Weight Perturbation</i>	
Md Fahim	317
<i>Z-Index at BLP-2023 Task 2: A Comparative Study on Sentiment Analysis</i>	
Prerona Tarannum, Md. Arid Hasan and Krishno Dey	324
<i>Team Error Point at BLP-2023 Task 2: A Comparative Exploration of Hybrid Deep Learning and Machine Learning Approach for Advanced Sentiment Analysis Techniques.</i>	
Rajesh Kumar Das, Kabid Yeiad, Moshfiqur Rahman Ajmain, Jannatul Maowa, Mirajul Islam and Sharun Akter Khushbu	331
<i>UFAL-ULD at BLP-2023 Task 2 Sentiment Classification in Bangla Text</i>	
Sourabrata Mukherjee, Atul Kr Ojha and Ondrej Dusek	336
<i>Embeddings at BLP-2023 Task 2: Optimizing Fine-Tuned Transformers with Cost-Sensitive Learning for Multiclass Sentiment Analysis</i>	
S.m Towhidul Islam Tonmoy	340
<i>LowResource at BLP-2023 Task 2: Leveraging BanglaBert for Low Resource Sentiment Analysis of Bangla Language</i>	
Aunabil Chakma and Masum Hasan	347
<i>BLP-2023 Task 2: Sentiment Analysis</i>	
Md. Arid Hasan, Firoj Alam, Anika Anjum, Shudipta Das and Afiyat Anjum	354
<i>BLP-2023 Task 1: Violence Inciting Text Detection (VITD)</i>	
Sourav Saha, Jahedul Alam Junaed, Maryam Saleki, Mohamed Rahouti, Nabeel Mohammed and Mohammad Ruhul Amin	365