

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/385496138>

Aspect Based Sentiment Analysis Datasets for Bangla Text

Article in Data in Brief · November 2024

DOI: 10.1016/j.dib.2024.111107

CITATIONS

5

READS

66

3 authors, including:



Mahmudul Hasan

Khulna University of Engineering and Technology

6 PUBLICATIONS 23 CITATIONS

[SEE PROFILE](#)



K. M. Azharul Hasan

Khulna University of Engineering and Technology

129 PUBLICATIONS 923 CITATIONS

[SEE PROFILE](#)



Data Article

Aspect based sentiment analysis datasets for Bangla text

Mahmudul Hasan^{a,b}, Md. Rashedul Ghani^a, K.M. Azharul Hasan^{a,*}^a Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh^b Department of Computer Science and Engineering, IUBAT— International University of Business Agriculture and Technology, 4 Embankment Drive Road, Sector-10, Uttara Model Town, Dhaka 1230, Bangladesh

ARTICLE INFO

Article history:

Received 31 July 2024

Revised 1 October 2024

Accepted 28 October 2024

Available online 2 November 2024

Dataset link: [Aspect Based Sentiment Analysis Datasets for Bangla Text \(Original data\)](#)

Keywords:

Sentiment analysis

Aspect based sentiment analysis

Bangla sentiment analysis

Opinion mining

Natural language processing

ABSTRACT

Sentiment analysis is becoming rapidly important for exploring social media Bangla text. The lack of sufficient resources is considered to be an important challenge for Aspect Based Sentiment Analysis (ABSA) of the Bangla language. The ABSA is a technique that divides the text and defines its sentiment based on its aspects. In this paper, we developed a high-quality Bangla ABSA annotated dataset namely BANGLA_ABSA. The datasets are labelled with aspects category and their respective sentiment polarity to do the ABSA task in Bangla. Four open domains namely Restaurant, Movie, Mobile phone, and Car are considered to make the dataset. The datasets are called *Restaurant_ABSA*, *Movie_ABSA*, *Mobile_phone_ABSA*, and *Car_ABSA* respectively that contain 801, 800, 975, and 1149 comments. All the comments are either complex or compound sentences. We created the dataset manually and annotated the same by exerting opinions. We organized the dataset as three tuples in Excel format namely $\langle Id, Comment, \{Aspect\ category, Sentiment\ Polarity\} \rangle$. These data are very important that facilitate

* Corresponding author.

E-mail address: az@cse.kuet.ac.bd (K.M.A. Hasan).Social media: [@azhasanlab](#) (K.M.A. Hasan)

the efficient handling of sentiment for any machine learning and deep learning research, especially for Bangla text.

© 2024 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Artificial Intelligence.
Specific subject area	Bangla Natural Language Processing, Sentiment Analysis in Bangla
Type of data	Table (Text/String).
Data collection	The dataset was initially created manually by a large number of people who are undergraduate students. The domains were explained clearly to them and requested to create complex or compound sentences/comments with two aspects and sentiment polarity of the aspect. These types of complex and compound sentences/comments are not commonly found in the web. The initial dataset was divided into three and each part was assigned to three annotators to label the comments into (aspect, polarity) pair. In case of disagreement of a comment, the majority voting was considered.
Data source location	Institution: Khulna University of Engineering & Technology City/Town/Region: Khulna Country: Bangladesh
Data accessibility	Repository name: Mendeley Data Data identification number: 10.17632/998m4jy3m9.3 Direct URL to data: https://data.mendeley.com/datasets/998m4jy3m9/3 Instructions for accessing these data: By using the URL, anyone can view the datasets publicly.
Related research article	None

1. Value of the Data

- Aspect-based Sentiment Analysis (ABSA) is a specific sentiment analysis task that aims to extract the most important aspects of an entity and predict the polarity of each aspect from the text. Bangla is the sixth most spoken language in the world. It is called resource scarce language because very few resources are found especially for ABSA. The datasets will enrich the field of Bangla ABSA as well as the Bangla Natural Language Processing (NLP) task.
- One important property of the ABSA dataset is that all the sentences are complex or compound. All the comments of the datasets are labelled into two pairs of aspect category and sentiment polarity which is different from other existing datasets. However, generally only one labelled pair of aspect category and sentiment polarity for a comment or review is found in the other existing sentiment datasets [1,2].
- The datasets can be utilized for developing machine learning and deep learning models to do ABSA tasks for the Bangla text for the applications of sentiment analysis or opinion mining such as business analysis, product review, market research social media monitoring etc.
- In E-commerce, aspects of a product such as price and size are analyzed separately. Therefore, ABSA provides a complete opinion summary at the aspect level, providing detailed sentiment information and offering more precise sentiment insights. Therefore, the proposed dataset can be used for fine-grained sentiment analysis and recommender systems.
- There is a potential for cross-domain use of the dataset in Bangla ABSA tasks. For example, the Mobile_phone_ABSA dataset can be applied particularly for laptop, personal computer, and digital camera related text analysis. These electronic items have similar aspects and sentiment patterns of text such as performance, design, camera, battery, and storage. In the same way, the Car_ABSA dataset can be used for many aspects that are significantly

common to cars namely such as motorbikes, motorboats, and other vehicles to develop ABSA tasks for performance, exterior, interior, and accessories. Hence with some domain-specific modifications, the proposed dataset can be used in a variety of domains.

2. Background

Natural language processing (NLP) includes sentiment analysis as a key component that becomes essential for interpreting the enormous quantity of content created by users on the internet [3]. It aims to extract sentiments from text, providing insights into people's opinions and attitudes across various topics. With the widespread use of social media, where users spend significant time daily, sentiment analysis has become crucial for understanding public sentiment. Conventional sentiment analysis works focus on overall sentiment at the sentence or document level [4]. This method uses the assumption of a single sentiment towards a single topic, which is often not the case to apply in the actual situation. As a result, recognizing aspect-level opinions and sentiments has gained attention in the past decade [5]. Aspect Based Sentiment Analysis (ABSA) models shift the focus from entire sentences or documents to specific entities or aspects of entities. ABSA task is two types' namely single and compound where the single ABSA is a part of compound ABSA. While ABSA has seen significant development in languages such as English, European languages, Arabic, and Hindi, but the Bangla language has been relatively underrepresented in sentiment ABSA research. Despite being one of the most spoken languages globally, the scarcity of datasets and resources poses a significant barrier to conducting ABSA research in Bangla [6]. Datasets are created in [1,2] for Bangla ABSA tasks. These data are mostly simple sentence with a single aspect. Although there are few sentences for multiple aspects are present there. Therefore, compound and complex ABSA computation is difficult with the dataset. A dataset for sentiment and emotion classification is developed in [7] which provides a comprehensive insight into consumer behavior and preferences in the context of Bangla text. Once again, most of the sentences are simple sentences with a single aspect.

Recognizing this gap, the objective of this work is to contribute to the enhancement of Bangla language resources by developing datasets specifically tailored for ABSA. By addressing the shortage of resources and datasets, this endeavor seeks to pave the way for more extensive research and development in Bangla sentiment analysis, thereby fostering a deeper understanding of public sentiment and consumer behavior within the Bangla-speaking community.

3. Data Description

We have made our datasets openly accessible on a data repository [8]. The dataset is arranged in a folder called "BANGLA_ABSA dataset", which contains *Car_ABSA*, *Mobile_phone_ABSA*, *Movie_ABSA*, and *Restaurant_ABSA* datasets respectively. The *Restaurant_ABSA* contains the comments on restaurant items like food, service, price etc. that shows the sentiment about a restaurant of its customers. The comments are complex or compound sentences with two aspects. For example, the sentence "খাবারটা সুস্বাদু ছিল কিন্তু দাম বেশি ছিল (The food was good but price was high)" has two aspects food (খাবার) and price (দাম). The polarity of the food is positive and the polarity of the price is negative. The dataset is organized as three tuples namely $\langle Id, Comment, \{Aspect\ category, Sentiment\ Polarity\} \rangle$.

The other three datasets namely *Car_ABSA*, *Mobile_phone_ABSA*, and *Movie_ABSA* contain user comments about Car, Mobile phone, and Movie respectively. Table 1 presents an overview of the dataset for four different domains. Snapshots of the dataset are shown in Fig. 1. All the comments are either complex or compound sentences with two aspects and two (same or opposite) polarities. The average length (Average number of words) of *Car_ABSA*, *Mobile_phone_ABSA*, *Movie_ABSA*, and *Restaurant_ABSA* datasets is 10.26, 10.14, 11.78, and 11.22, respectively. The maximum number of words in the dataset of four domains is 19, 20, 35, and 23 respectively. Table 2 shows the statistics of the dataset.

Table 1
Overview of the proposed datasets.

Domain name	Aspect category	Polarity	No. of comments
Car_ABSA	Performance, Exterior, Interior, Accessories, Comfort, Safety, Practicality, Others	Positive, Negative	1149
Mobile_phone_ABSA	Performance, Design, Display, Camera, Battery, Storage, Value_For_Money, Others	Positive, Negative	975
Movie_ABSA	Story, Performance, Music, Miscellaneous	Positive, Negative	800
Restaurant_ABSA	Food, Price, Service, Ambiance, Miscellaneous	Positive, Negative	801

Table 2
Statistics of proposed datasets.

Dataset	No. of comment	Maximum length (in words)	Minimum length (in words)	Average length (in words)	Maximum length (in characters)	Minimum length (in characters)	Average length (in characters)
Car_ABSA	1149	19	5	10.26	128	30	62.13
Mobile_phone_ABSA	975	20	4	10.14	109	23	59.42
Movie_ABSA	800	35	4	11.78	195	23	69.63
Restaurant_ABSA	801	23	4	11.22	144	23	65.43

4. Experimental Design, Materials and Methods

The dataset has one important property that each of the text/sentence is a complex or compound sentence. Each of the text has two aspects with two sentiment polarities. It is unusual to get such text in the web or bog. Initially the dataset was created manually by a large number of people who are mostly undergraduate (BSc. engineering) students. The domains of the dataset are described clearly and requested to create complex or compound sentences within the domain with two aspects and two sentiment polarities (positive or negative). For example the text “the food was good but price was high” has two aspects food and price and two sentiment polarities namely {food, positive} and {price, negative}. All the comments did not meet the ABSA criteria such as complex or compound sentences. These data were pre-processed manually to meet the criteria of ABSA. The initial dataset was divided into three and each part was assigned to three annotators to annotate the text into {aspect, polarity} pair manually. These annotators are also undergraduate students. The majority voting was considered to adopt a {aspect, polarity} text annotating in the dataset.

Table 3 shows an example of annotation. After a majority vote, the comment was finally annotated with a food aspect and positive polarity, and a price aspect and negative polarity. After annotation of the data, the dataset has the scope of improvement. After annotation of

Table 3
An example of annotation.

Comment: যদিও খাবার ভাল ছিল কিন্তু পরিবেশনা ছিল বাজে। (Although the food was good but service was bad)	
Annotator	{Aspect category, Sentiment Polarity}
A1	{food, positive}, {service, negative}
A2	{food, positive}, {service, negative}
A3	{food, positive}, {service, negative}

Id	Comment	(Aspect category, Sentiment Polarity)
1	হ্যাভাক্রেকটা নষ্ট হয়ে গেছে কিন্তু নরমাল ব্রেক গুলা ঠিকই আছে।	(safety, negative), (safety, positive)
2	লেদার সিটগুলো খুব একটা কমফোর্টেবল না হলেও গাড়িটার কেবিনটা সুন্দর।	(comfort, negative), (interior, positive)
3	গাড়িটার একদম নিউ কালার করা হয়েছে কিন্তু এই কালার ক্যান্ডিন টিকবে তার কোনো গ্যারান্টি মাই।	(exterior, positive), (exterior, negative)
4	ইঞ্জিনে পারফরমেন্স হেভার লাগানো আছে কিন্তু ইন্টারকুলারটা একটু বেশি ছোটো।	(performance, positive), (performance, negative)
5	সামনের সিটে কম হেডরুম কিন্তু পিছনে আরো বেটার হেডরুম আছে।	(comfort, negative), (comfort, positive)
6	পিছনের মইন ব্রেক লাইটটা কাজ করে কিন্তু থার্ড ব্রেক লাইটে সমস্যা আছে।	(exterior, positive), (exterior, negative)
7	পিছনের ব্রেক লাইটটা কাজ করে কিন্তু গাড়ির ব্রেক সমস্যা আছে।	(exterior, positive), (safety, negative)
8	বেথডেই শুধু স্পোর্টস কার না, টপস্পীডও অনেক বেশি।	(exterior, positive), (performance, positive)
9	লেদার সিটগুলো একদম নতুনের মতো সাইন মারছে আর কালো রঙে গাড়িটা দেখে মন ভরে যাচ্ছে।	(interior, positive), (exterior, positive)
10	মজান পারফরমেন্স কিন্তু ওন্দকাশন ফুয়েল ইকনমি।	(performance, positive), (practicality, negative)

a) Car_ABSA

Id	Comment	(Aspect category, Sentiment Polarity)
1	বর্তমানে ৩০ হাজার টাকার বাজেটে এই ফোনটি অনেক ভালো কিন্তু ফোনটির জর্নপ্রিয়তা অনেক কম।	(value_for_money, positive), (others, negative)
2	বর্তমানে ৫০ হাজার টাকার বাজেটে এই ফোনটি খুব একটা ভালো না তাও ফোনটির জর্নপ্রিয়তা অনেক বেশি।	(value_for_money, negative), (others, positive)
3	ফোনটির ডিসপ্লে অনেক ভালো এবং ফোনের স্পিড আমার কাছে অনেক বেশি লেগেছে।	(display, positive), (performance, positive)
4	ফোনটির ডিসপ্লে একদম বাজে এবং ফোনের স্পিড আমার কাছে একদম কম লেগেছে।	(display, negative), (performance, negative)
5	ফোনের দাম এতো বেশি না কিন্তু ফোনের প্রিটি ফিচার কাস্টমারদের পছন্দ হচ্ছে।	(value_for_money, positive), (others, positive)
6	ফোনের দাম বেশি তাও ফোনের ফিচারগুলো কাস্টমার পছন্দ করছে না।	(value_for_money, negative), (others, negative)
7	ফোনের দাম একটু বেশি হয়েছে কিন্তু আমার কাছে ফোনের লুক্স জোস লাগছে।	(value_for_money, negative), (design, positive)
8	ফোনের দাম তুলনামূলক কম কিন্তু আমার কাছে ফোনের লুক্স জোস লাগছে।	(value_for_money, positive), (design, positive)
9	ক্যামেরার কোয়ালিটি বাজে এবং দামও অনেক বেশি ফোনটির।	(camera, negative), (value_for_money, negative)
10	ক্যামেরার কোয়ালিটি একদম সেরা, তাই দামও একটু বেশি।	(camera, positive), (value_for_money, negative)

b) Mobile_phone_ABSA

Id	Comment	(Aspect category, Sentiment Polarity)
1	গরমার বস্ত্রের গল্প সুন্দর কিন্তু অভিনয় ভাল ছিলো না।	(story, positive), (performance, negative)
2	যেমন গল্প সুন্দর তেমন অভিনয় চমৎকার।	(story, positive), (performance, positive)
3	গানগুলো সুন্দর বাট একটা বেশি মারামারি ভাল লাগে নি।	(music, positive), (performance, negative)
4	তবে বাচ্চা কিয়ার খুবই অসাধারণ অভিনয় করেছে, এছাড়া বলতে গেলে সিনেমাতোয়াকি অনেক ভালো।	(performance, positive), (miscellaneous, positive)
5	ভাই মুভিটা একটা অবেগ, পরে পুরাই চমক	(miscellaneous, positive), (story, positive)
6	সিনেমার লোকেশন গুলো খুব সুন্দর ছিল কিন্তু পুরো সিনেমা জুড়ে কোন রহস্য নেই।	(miscellaneous, positive), (story, negative)
7	ক্যামেরার কাজও বেশ ভালো হয়েছে তবে গানের কথা একটু এলোমেলো লাগতে পারে।	(miscellaneous, positive), (music, negative)
8	নায়ক তার বিগত সিনেমা গুলির তুলনায় বেশ চেষ্টা করেছেন এছাড়া বিদেশি অভিনয় শিল্পীরাও ভালো অভিনয় করেছেন।	(performance, positive), (performance, positive)
9	অ্যাকশন থ্রিলার গল্প হলেও কোথাও কোন থিল খুঁজে পাবেন না।	(story, positive), (performance, negative)
10	বিশেষ করে তুলতাল ইরেজি যদিও সাবটাইটেল এর ব্যবস্থা করা হয়েছে।	(miscellaneous, negative), (miscellaneous, positive)

c) Movie_ABSA

Id	Comment	(Aspect category, Sentiment Polarity)
1	খাবারটি অত্যন্ত সুস্বাদু ছিল কিন্তু দাম বেশি ছিল	(food, positive), (price, negative)
2	যদিও খাদ্য ভাল ছিল কিন্তু পরিবেশনা ছিল বাজে।।	(food, positive), (service, negative)
3	খাদ্যে সবসময় তাজা স্বাদ এবং তাৎক্ষণিকভাবে পরিবেশিত।	(food, positive), (service, positive)
4	খুব ভাল দাম এবং খুব ভাল সেবা।	(price, positive), (service, positive)
5	এটি একটি চমৎকার শিথিল জায়গা, এখনও খাদ্য ভাল ।	(ambiance, positive), (food, positive)
6	মাংসের অংশ বিশাল ছিল, দাম বিশেষ কিছু না হলেও।	(food, positive), (price, positive)
7	সেবা ছিল অনুপযুক্ত এবং খাদ্যটি ছিল গড়ের নিচে।	(service, negative), (food, negative)
8	খাটি খাদ্য ছিল এবং আমি এক মাস দুইবার যাই।	(food, positive), (miscellaneous, neutral)
9	সেখানে কোন পরিবেশ ছিল না, সেবা ছিল অনুপযুক্ত	(ambiance, negative), (service, negative)
10	সেখানে কোন পরিবেশ ছিল না এবং খাদ্যটি ছিল গড়ের নিচে।	(ambiance, negative), (food, negative)

d) Restaurant_ABSA

Fig. 1. Snapshot of the datasets.

the comments the data is still noisy. The following preprocessing was applied to the proposed datasets after annotation.

- Some of the punctuation characters such as (':', ';') are removed as these punctuations are not very meaningful for the ABSA task. On the other hand some punctuations such as ', ', '!', '?', '|' are used to make meaningful sentences.

Table 4
Aspect category classification performance on the proposed datasets.

Dataset	Model	Precision (%)	Recall (%)	F1-score (%)
Car_ABSA	NB	82.13	73.38	76.00
	LR	81.50	78.50	79.63
	SVM	83.75	82.38	83.13
Mobile_Phone_ABSA	NB	90.63	77.00	81.13
	LR	93.63	86.38	89.13
	SVM	95.63	92.63	93.63
Movie_ABSA	NB	55.20	55.60	55.20
	LR	98.40	87.00	91.80
	SVM	99.00	93.00	95.60
Restaurant_ABSA	NB	99.00	80.17	86.00
	LR	98.67	85.00	90.67
	SVM	99.33	95.00	97.00

Table 5
Sentiment polarity classification performance on the proposed datasets.

Dataset	Model	Precision (%)	Recall (%)	F1-score (%)
Car_ABSA	NB	78.50	78.50	78.50
	LR	81.50	80.50	81.00
	SVM	84.50	84.00	84.00
Mobile_Phone_ABSA	NB	79.50	78.00	79.00
	LR	80.00	79.00	79.50
	SVM	82.50	82.00	82.50
Movie_ABSA	NB	58.33	60.00	59.33
	LR	98.33	74.33	81.33
	SVM	94.33	78.33	83.67
Restaurant_ABSA	NB	92.33	76.00	81.00
	LR	95.67	86.67	90.00
	SVM	99.33	92.33	95.67

- Stop words are inconsequential words with little value in processing specific tasks. Stop words that are used frequently in a text but are not important for the ABSA task. For example, ছিল, অতএব, এই, করবে, তবু, হোক etc. are the stop word for Bangla ABSA task and are removed.
- For example, the punctuation of the text এই গাড়িটা অপূর্ণ একটা প্যাকেজ: আরামদায়ক হলেও গতি একদমই নাই” can be removed to generate “এই গাড়িটা অপূর্ণ একটা প্যাকেজ আরামদায়ক হলেও গতি একদমই নাই.” The stop words can also be removed to process the text “গাড়িটা অপূর্ণ প্যাকেজ আরামদায়ক হলেও গতি একদমই নাই”. We used the stop words listed in [9].

To evaluate the proposed datasets, some supervised machine-learning models namely Naive Bayes (NB) , Logistic Regression (LR) , and Support Vector Machine (SVM) were applied using the proposed datasets separately and achieved satisfactory performance for ABSA. Among the three machine learning models, SVM gained the highest F1 score. The experimental results for the four different datasets are presented for aspect category classification and sentiment polarity classification in [Tables 4 and 5](#) respectively. For aspect category classification, the F1 scores were 83.13 %, 93.63 %, 95.60 %, and 97.00 % respectively. For sentiment polarity classification, the scores were 84.00 %, 82.50 %, 83.67 %, and 95.67 % respectively. From the experimental results with important machine learning models, it is evident that they are producing good and stable results both for aspect detection and sentiment polarity tasks. Hence, we can conclude that the dataset is well balanced for ABSA tasks for Bangla text.

Although the dataset is developed for specific domains but the dataset, in a broad sense, represents consumer behavior and urban experiences, ignoring any rural or other socioeconomic viewpoints and can be applied to other domains with some little modification that can assist the sentiment expressions in a wider context. A comparative assessment of the proposed dataset

Table 6
Comparison of the dataset for SVM model.

Dataset	Precision	Recall	F1-score
Cricket [1]	0.71	0.22	0.34
Restaurant [1]	0.77	0.30	0.38
BAN-ABSA [2]	0.74	0.66	0.69
Car_ABSA	0.84	0.82	0.83
Mobile_Phone_ABSA	0.96	0.93	0.94
Movie_ABSA	0.99	0.93	0.96
Restaurant_ABSA	0.99	0.95	0.97

with [1,2] is presented in Table 6 applying SVM model. The result shows superior precision, recall and F1-score for the proposed dataset.

Limitations

The limitations of the datasets are as follows:

- The dataset is domain dependent namely Car, Mobile phone, Movie, and Restaurant.
- The length of the comments (in words) of the proposed datasets is not long. The maximum length is 35 words. Hence the long dependency may not be captured.

Ethics Statement

The authors do not have any ethical issues for developing the dataset. The datasets are fully anonymous and the data redistribution policies. The identity of the commenter has been obscured to protect their privacy.

Data Availability

Aspect Based Sentiment Analysis Datasets for Bangla Text (Original data) (Mendeley Data).

CRedit Author Statement

Mahmudul Hasan: Conceptualization, Methodology, Data curation, Validation, Writing – original draft; **Md. Rashedul Ghani:** Data curation, Validation, Writing – original draft; **K.M. Azharul Hasan:** Supervision, Writing – review & editing.

Acknowledgements

We express our heartfelt appreciation to the undergraduate students of CSE department KUET for their invaluable support throughout who helped with this study to create and process the sentences in the correct complex or compound form. The authors are thankful to them. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] M.A. Rahman, E.K. Dey, Datasets for aspect-based sentiment analysis in bangla and its baseline evaluation, *Data* 3 (2) (2018) 15.
- [2] M. Ahmed Masum, S. Junayed Ahmed, A. Tasnim, M. Saiful Islam, Ban-absa: an aspect-based sentiment analysis dataset for bengali and its baseline evaluation, in: *In Proceedings of International Joint Conference on Advances in Computational Intelligence: IJCACI 2020*, Springer, 2021, pp. 385–395.
- [3] W. Medhat, A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: a survey, *Ain Shams Eng. J.* 5 (4) (2014) 1093–1113.
- [4] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? Sentiment classification using machine learning techniques, in: *In Proceedings of the Conference Empirical Methods Natural Language Processing*, 2002, pp. 79–86.
- [5] W. Zhang, X. Li, Y. Deng, L. Bing, W. Lam, A survey on aspect-based sentiment analysis: tasks, methods, and challenges, *IEEE Trans. Knowl. Data Eng.* 35 (11) (2022) 11019–11038.
- [6] S.A. Shammi, S. Das, N.R. Chakraborty, S.K. Banshal, N. Nath, A comprehensive roadmap on bangla text-based sentiment analysis, *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* 22 (4) (2023) 1–29.
- [7] M.R.A. Rashid, K.F. Hasan, R. Hasan, A. Das, M. Sultana, M. Hasan, A comprehensive dataset for sentiment and emotion classification from Bangladesh e-commerce reviews, *Data Br.* 53 (2024) 110052.
- [8] M. Hasan, Md.R. Ghani, K.M.A. Hasan, BANGLA-ABSA: unique aspect based sentiment analysis datasets in Bangla Language, *Mendeley Data* V3 (2024), doi:10.17632/998m4jy3m9.3.
- [9] https://github.com/Munna015/Bangla-Stopwords/blob/main/stopwords_bangla.txt.