Benchmark Creation for Aspect-Based Sentiment Analysis in Low-Resource Odia Language and Evaluation through Fine-Tuning of Multilingual Models

Lipika Dewangan[1], Zoyah Afsheen Sayeed[2], Chandresh Kumar Maurya[1]
[1]Indian Institute of Technology, Indore, Madhya Pradesh, India 453552
[2]Kalinga Institute of Industrial Technology, Bhubaneswar, Odisha, India 751024
{phd2101201004, chandresh}@iiti.ac.in, 21052990@kiit.ac.in

Abstract

The rapid growth of online product reviews spurs significant interest in Aspect-Based Sentiment Analysis (ABSA), which involves identifying aspect terms and their associated sentiment polarity. While ABSA is widely studied in resource-rich languages like English, Chinese, and Spanish, it remains underexplored in low-resource languages such as Odia. To address this gap, we create a reliable resource for aspect-based sentiment analysis in Odia. The dataset is annotated for two specific tasks: Aspect Term Extraction (ATE) and Aspect Polarity Classification (APC), spanning seven domains and aligned with the SemEval-2014 benchmark. Furthermore, we employ an ensemble data augmentation approach combining back-translation with a fine-tuned T5 paraphrase generation model to enhance the dataset and apply a semantic similarity filter using a Universal Sentence Encoder (USE) to remove low-quality data and ensure a balanced distribution of sample difficulty in the newly augmented dataset. Finally, we validate our dataset by fine-tuning multilingual pre-trained models, XLM-R and IndicBERT, on ATE and APC tasks. Additionally, we use three classical baseline models to evaluate the quality of the proposed dataset for these tasks. We hope the Odia dataset will spur more work for the ABSA task.

1 Introduction

Aspect-Based Sentiment Analysis is a fine-grained approach that extracts aspect terms from text and predicts their sentiment polarity, providing more detailed insights than traditional sentiment analysis. It aims to automatically identify all aspect terms in a text and determine their associated sentiments. For example, in the sentence *"The khaja prasad at Jagannath Temple is excellent, but the service was slow."*, the aspect terms *"khaja prasad"* and *"service"* carry positive and negative sentiments, respectively. Aspect terms may be single words or multi-word phrases or sometimes may not be explicitly mentioned in the text.

Odia is an Indo-Aryan language[1] recognized by the Indian government as one of the six classical languages alongside Sanskrit, Tamil, Telugu, Kannada, and Malayalam. Odia is the mother tongue of approximately 32 million people in Odisha[2] and is also spoken in neighboring states like West Bengal, Chhattisgarh, and Jharkhand. Despite its rich cultural heritage and a large online community active on social media and in mainstream media, there is a shortage of large-scale, publicly accessible resources for sentiment classification in Odia.

The remaining sections of the paper are organized as follows: section 2 covers related work on sentiment analysis in Indian languages, followed by sentiment analysis in Odia language. Section 3 describes the data collection process, dataset preprocessing, data annotation, and challenges faced during annotation and provides the dataset statistics. In section 4, we present the data augmentation methods employed to expand the training data, while section 5 outlines the experiments conducted to fine-tune state-of-the-art models on the benchmark dataset and section 6 reports the results and introduces baseline models. Finally, section 7 summarizes the findings and concludes the research.

2 Related Work

The ABSA task is initially introduced by SemEval in 2014 for the English language (Pontiki et al., 2014) followed by (Pontiki et al., 2016). Since then, advancements in deep learning significantly improve ABSA, with studies such as (Tang et al., 2016), (Cheng et al., 2017), and (Xue and Li, 2018), and (Wang et al., 2020). An auxiliary sentence was constructed (Sun et al., 2019) from the aspect to convert ABSA from a single-sentence classifica-

[1]https://en.wikipedia.org/wiki/Odia_language
[2]https://culture.odisha.gov.in/

Proceedings of the 31st International Conference on Computational Linguistics, pages 5863–5869
January 19–24, 2025.

tion task into a sentence-pair classification task, such as question answering (QA) and natural language inference (NLI). With the growth of digital technology, Indian regional languages are increasingly prevalent on social media and online platforms, making fine-grained sentiment analysis essential. However, research is hindered by limited datasets in these languages. To address this, ABSA datasets are created for Hindi (Akhtar et al., 2016a) and Telugu (Regatte et al., 2020). The ABSA task is explored in Hindi (Akhtar et al., 2016b), Urdu (Rani and Anwar, 2020), Malayalam (Sunitha et al., 2023), and Bengali (Shimada, 2023). However, native languages like Odia remain under-resourced. Early sentiment analysis in Odia focuses on movie reviews (Sahu et al., 2016) using machine learning techniques but involves a small, non-public dataset of 1,000 sentences. Later, an annotated corpus for Odia sentiment analysis is created in the news domain with 2,045 sentences (Mohanty et al., 2020). Despite these efforts, there is still a lack of large-scale, publicly accessible datasets for ABSA in Odia.

3 Dataset Construction and Annotation

3.1 Data Collection

The dataset is carefully curated by scraping reviews from various sources[3] including regional e-commerce platforms, online forums, and social media, where Odia is predominantly used. The dataset covers seven domains: *Odia Ethnic food, Handloom sarees, Handicrafts, Hotels, Books, Odia movies, and Electronic products*. The electronic product reviews cover smartphones, laptops, DVDs, and water purifiers. The entire dataset is in the Odia script, but for better readability, we also provide transliterated reviews and their English translations. A few samples of dataset are shown in Figure 1.

3.2 Dataset Pre-processing

After data collection, we pre-process it for training by removing irrelevant reviews, excluding predominantly English ones, and discarding those without annotations. Spelling mistakes in Odia words, including in augmented samples, are manually corrected based on context. We ensure consistency by addressing spelling and grammar variations and removing unprintable characters and emoticons. The Odia sentence end marker (Purna Chhed) is removed, but special characters, including diacritical marks (Matras) and unique script elements, are retained to preserve the accuracy of the Odia script and its phonetics.

[3]https://ritikart.com
thrillophilia.com/states/orissa/
https://myodishaproducts.com/
https://odishahandicrafts.com
https://odishaforum.com/
https://www.ameodia.com

Domain	Sentence	Aspect Term	Polarity
Ethnic Food	Odia: ଗାଠିଆ ବହୁତ ଛୋଟ ଛୋଟ ଖଣ୍ଡରେ ଭାଙ୍ଗିଗଲା | Transliterated: Gathiya chhota chhota khandare bhangigala. Translated: Gathiya broken into very small pieces.	ଗାଠିଆ Gathiya Gathiya	ନକାରାତ୍ମକ Nakaratmaka Negative
Handloom Saree	Odia: ଶାଢ଼ୀ ଗୁଣ ବହୁତ ଭଲ, ବହୁତ ସୁନ୍ଦର ଦେଖାଯାଏ | Transliterated: Saadhi guna bahuta bhala, bahuta sundara dekhajaye. Translated: Saree quality is very good, looks very beautiful.	ଶାଢ଼ୀ Saadhi Saree	ସକାରାତ୍ମକ Sakaratmaka Positive
Handicraft	Odia: ମୁଁ ଦୀର୍ଘ ସମୟ ଧରି ଏହି ମୁକ୍ତାପୁରୀ ବାଟିକୁ ଖୋଜୁଥିଲି ।| Transliterated: Mun dirgha samaya dhari ehi Muktapuri batiku khojuthili. Translated: I have been looking for this Muktapuri bowl for a long time,	ମୁକ୍ତାପୁରୀ ବାଟି Muktapuri bati Muktapuri bowl	ନିରପେକ୍ଷ Nirapekhya Neutral

Figure 1: Few samples of Odia dataset

Table 1: Annotation Scores for ATE and APC Tasks

Task	Cohen's Kappa	Fleiss' Kappa	F1 Score
ATE	0.797	0.807	0.850
APC	0.789	0.798	0.830

3.3 Data Annotation

The annotation process is performed by two native Odia-speaking undergraduate students (aged 20-24). They undergo three rounds of training on 60-sample subsets, with doubt clarification and guidance from a linguistic expert after each round. We assess annotation quality by computing inter-annotator agreement using Cohen's Kappa (Cohen, 1960), Fleiss' Kappa (Fleiss, 1971) and the F1 Score. Cohen's Kappa scores are 0.797 for ATE and 0.789 for APC, while Fleiss' Kappa scores are 0.807 and 0.798, indicating strong inter-annotator agreement. High F1 Scores of 0.850 for ATE and 0.830 for APC confirm the accuracy and effectiveness of the Odia ABSA dataset annotation.

3.4 Challenges and Solutions in Odia Data Annotation for ABSA

During the Odia data annotation for ABSA, we encounter several challenges and address them as follows:

1. **Translation Ambiguities:** Accurate translations from English to Odia are challenging, as words like "tasty" and "flavor" both trans-

late to the same word "Swaada" in Odia, despite having distinct meanings in English. We employ bilingual experts and create domain-specific glossaries for accurate context-based translations.

2. **Incorrect Sentiment Translations:** Mistranslations (e.g., "crispy") alter sentiment polarity. A quality control loop with native speakers and an Odia sentiment lexicon ensures accurate polarity.

3. **Homographs:** It is challenging to categorize homographs, words with the same spelling but different meanings. For example, “Paan” can mean either “betel leaf” or “water.” The correct interpretation of such words depends on the context in which they are used. We resolve contextual ambiguity by training annotators to flag unclear cases for expert review.

4. **Missing Aspect Terms:** Sentences without explicit aspects are labeled "No aspect," with sentence-level polarity applied.

5. **Sarcasm:** Sarcastic sentences are labeled negative based on expert guidance for sarcasm detection.

6. **Code-Mixed Text:** English words in Odia reviews are fully translated to ensure uniform annotation.

7. **Toxic Content:** Toxic content is removed from the dataset.

3.5 Dataset Statistics

After pre-processing and data augmentation, our dataset consists of 5,780 sentences with 9,468 tokens across seven domains: Ethnic food, Handloom sarees, Handicrafts, Odisha hotels, Electronic products, Books, and Odia movies. A total of 5,758 aspect terms have been annotated, indicating that sentences may contain zero, one, or more aspect terms. The polarity distribution includes 3,323 positive, 1,604 negative, and 507 neutral aspect terms. This dataset reflects a moderate-sized corpus for Odia text-based sentiment analysis, serving as the foundation for training and evaluating ABSA models. Table 2 presents a summary of the domain-wise statistics for our dataset. The dataset can be requested via email.

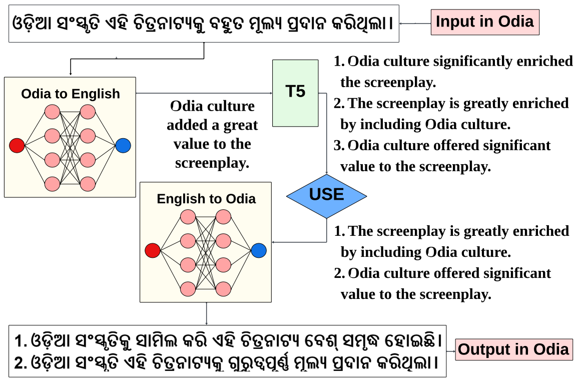

Figure 2: Data Augmentation with ensemble approach

4 Data Augmentation

Data augmentation refers to the automatic generation of additional textual data to expand the training dataset while maintaining class labels. We use an ensemble method combining back-translation (Liu et al., 2021) to preserve the original sentiment with the paraphrasing capabilities of the pre-trained T5 model (Raffel et al., 2020) to generate synthetic samples. A semantic similarity filter, utilizing the Universal Sentence Encoder (Cer, 2018), is applied to eliminate duplicates and retain high-quality sentences. Figure 2 shows an instance of the data augmentation process. The dataset translations are handled using the IndicTrans model from the indicNLP library (Kakwani et al., 2020). The process starts by translating the input text from Odia to English, followed by paraphrasing with the T5 model. Paraphrases are generated using a combination of top_k sampling and top_p nucleus sampling. To filter out these undesirable outcomes, we set a threshold value of 0.6. If the semantic similarity score is below 0.6, we discard such samples. Again, we annotate 1,300 newly augmented samples and achieve a substantial inter-annotator agreement score.

5 Experimental Setup

We conducted our experiments using the PyTorch-based library. To obtain domain-specific performance analysis, we fine-tune the models separately and avoid combining the datasets into a single training file. Running the models individually ensures that each dataset is treated equally, preventing larger datasets from dominating the training process and leading to more balanced performance across tasks. To assess the quality of our proposed dataset benchmark, we conduct experiments on two tasks: ATE and APC. We split the training data into

Table 2: Domain-wise Dataset Statistics

Domains	#Sentences	#Tokens	#Aspect Terms	#Pos	#Neg	#Neu
Ethnic food	850	1519	820	523	142	155
Handloom sarees	600	1461	664	409	234	21
Handicrafts	940	1673	952	523	328	29
Odisha Hotels	1170	1015	1182	508	442	45
Electronic products	1280	1540	1245	792	298	90
Books	540	1214	520	360	93	67
Odia movies	400	1346	375	208	67	100
Total	**5,780**	**9,468**	**5,758**	**3,323**	**1,604**	**507**

90% for training and 10% for validation in each domain. Table 3 shows the details of the hyperparameters used in our experiment. Training is conducted

Table 3: Details of Hyperparameters

Parameters	Details
Learning Rate	2×10^{-5}
Batch Size	16
Training Epochs	10
Weight Decay	0.01
Padding Max Length	128
Hidden Size	768
Hidden Activation	ReLU (Dense Layers)
Output Activation	Softmax
Loss	Categorical CrossEntropy
Optimizer	Adam

for 10 epochs for each model. We use accuracy and F1-score as evaluation metrics. We fine-tune XLM-R (Conneau et al., 2020), a multilingual version of RoBERTa, pre-trained on 2.5TB of Common-Crawl data which supports 100 languages including Odia. In our experiment, we specifically fine-tune the xlm-roberta-base variant[4], which has 279 million parameters. IndicBERT (Kakwani et al., 2020) is a multilingual ALBERT model pre-trained on 12 major Indian languages, including Odia. It is trained on a monolingual corpus of approximately 9 billion tokens and has been evaluated on a wide range of tasks. For our fine-tuning, we use the ai4bharat/indic-bert model[5], which contains 104 million parameters. We additionally employ the following classical baseline models to assess the quality of the proposed dataset for ATE and APC tasks.

BiLSTM-CRF (Alzaidy et al., 2019): Combines BiLSTM for capturing contextual information and CRF for label prediction, considering neighboring words, and is used exclusively for the ATE task.

Table 4: Results of ATE and APC using XLM-R and IndicBERT across seven datasets

Model	Domains	Acc_{ATE}	F1_{ATE}	Acc_{APC}	F1_{APC}
XLM-R	Ethnic Food	97.36	96.99	71.56	65.89
	Odisha Hotels	**99.60**	**98.97**	69.50	65.10
	Books	97.59	96.75	**81.57**	**74.97**
	Odia Movies	98.26	97.61	76.30	68.84
	Electronic Products	96.58	95.45	77.62	73.37
	Handloom Sarees	99.29	97.99	72.28	61.70
	Handicrafts	92.45	90.93	77.15	68.30
IndicBERT	Ethnic Food	**97.98**	**97.95**	70.94	65.11
	Odisha Hotels	96.30	95.04	59.40	55.96
	Books	94.68	92.57	66.76	59.79
	Odia Movies	95.32	94.22	**77.20**	**69.82**
	Electronic Products	94.68	92.57	66.76	60.73
	Handloom Sarees	97.15	96.56	70.72	61.19
	Handicrafts	97.57	95.73	64.90	53.72

ATAE-LSTM (Wang et al., 2016): Enhances LSTM with an attention mechanism to generate aspect-specific attention embeddings for predicting aspect term polarity.

ABSA-BERT (Sentence Level) (Hoang et al., 2019): This model predicts aspects related to a given text for both in-domain and out-of-domain scenarios by utilizing the pre-trained BERT language model and fine-tuning it as a sentence-pair classification model for the ABSA task.

6 Results and Analysis

Table 4 presents the empirical results, comparing the performance of XLM-R and IndicBERT on the ATE and APC tasks across seven datasets. Table 5 evaluates BiLSTM-CRF, ATAE-LSTM, and ABSA-BERT on the same datasets. The results demonstrate that state-of-the-art models outperform classical baselines across ATE and APC tasks. Among the SOTA models, XLM-R achieves the highest ATE accuracy of 99.60% and F1 score of 98.97% on the Odisha Hotels dataset, while IndicBERT excels in Ethnic Food with an ATE accuracy and F1 score of 97.98% and 97.95%, respectively. For APC tasks, XLM-R leads with

[4] https://huggingface.co/xlm-roberta-base
[5] https://huggingface.co/ai4bharat/indic-bert

Table 5: Results of ATE and APC using classical baselines across seven datasets

Model	Domains	Acc_{ATE}	F1_{ATE}	Acc_{APC}	F1_{APC}
BiLSTM-CRF (**Alzaidy et al., 2019**)	Ethnic Food	74.18	68.12	–	–
	Odisha Hotels	74.46	67.73	–	–
	Books	74.68	69.35	–	–
	Odia Movies	73.34	67.59	–	–
	Electronic Products	**78.74**	**72.86**	–	–
	Handloom Sarees	72.44	67.59	–	–
	Handicrafts	73.25	66.08	–	–
ATAE-LSTM (Wang et al., 2016)	Ethnic Food	–	–	71.64	68.76
	Odisha Hotels	–	–	73.92	70.84
	Books	–	–	71.36	69.35
	Odia Movies	–	–	72.34	70.15
	Electronic Products	–	–	70.47	68.54
	Handloom Sarees	–	–	**74.56**	**71.92**
	Handicrafts	–	–	73.04	70.63
ABSA-BERT (Sentence Level) (Hoang et al., 2019)	Ethnic Food	86.34	83.28	72.24	70.17
	Odisha Hotels	87.68	84.55	72.86	68.45
	Books	76.71	74.54	75.74	69.69
	Odia Movies	88.76	86.71	75.71	71.38
	Electronic Products	**89.28**	**87.42**	74.22	69.37
	Handloom Sarees	87.86	83.65	74.54	70.38
	Handicrafts	85.57	81.82	**77.86**	**73.65**

81.57% accuracy and 74.97% F1 on Books, while IndicBERT achieves the best APC performance for Odia Movies (77.20% accuracy, 69.82% F1). Comparatively, classical models like BiLSTM-CRF perform moderately well, especially in Electronic Products and Books, while ATAE-LSTM excels in APC for Handloom Sarees. ABSA-BERT is the most effective classical baseline for both tasks particularly in Electronic Products with an ATE accuracy of 89.28% and F1 score of 87.42%, and in Handicrafts, with an APC accuracy of 77.86% and F1 score of 73.65%. Overall, XLM-R demonstrates the greatest robustness across both ATE and APC tasks.

7 Conclusion and future works

We present a dataset benchmark for aspect-based sentiment analysis in the low-resource Odia language. The dataset, compiled from various online sources, spans seven domains and is annotated with aspect terms and sentiment polarity labels. To expand the training data, we used ensembled data augmentation using back translation and paraphrase generation, applying a semantic filter to reduce noise. Our Odia dataset will serve as a baseline for ABSA system evaluation and support further research. Future work will include multilingual and code-mixed Odia-English ABSA tasks, along with domain adaptation for more robust ABSA models.

8 Limitations of the work

As Odia is an agglutinative language, the models face challenges in properly handling its dialects and matras. Although we set guidelines while training annotators, the correct sentiment polarity outcomes are still lacking. Addressing issues such as implicit aspects, code-mixing, sarcasm and faulty translations remains challenging, highlighting the need for further research in low-resource Odia language. Both XLM-R and IndicBERT struggle to identify the exact aspect term when it contains more than one word, often incorrectly extracting only part of the term. Since pre-trained models require large datasets, we need to further increase our dataset size to improve the performance of the APC task. Although data augmentation helps to some extent, it still lacks versatility and does not provide sufficiently diversified samples.

9 Acknowledgments

We would like to express our gratitude to the reviewers for their valuable comments and suggestions.

References

Md Shad Akhtar, Asif Ekbal, and Pushpak Bhattacharyya. 2016a. Aspect based sentiment analysis in hindi: resource creation and evaluation. In

Proceedings of the tenth international conference on language resources and evaluation (LREC'16), pages 2703–2709.

Md Shad Akhtar, Ayush Kumar, Asif Ekbal, and Pushpak Bhattacharyya. 2016b. A hybrid deep learning architecture for sentiment analysis. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 482–493.

Rabah Alzaidy, Cornelia Caragea, and C Lee Giles. 2019. Bi-lstm-crf sequence labeling for keyphrase extraction from scholarly documents. In *The world wide web conference*, pages 2551–2557.

D Cer. 2018. Universal sentence encoder. *arXiv preprint arXiv:1803.11175*.

Jiajun Cheng, Shenglin Zhao, Jiani Zhang, Irwin King, Xin Zhang, and Hui Wang. 2017. Aspect-level sentiment classification with heat (hierarchical attention) network. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 97–106.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Édouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451.

Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378.

Mickel Hoang, Oskar Alija Bihorac, and Jacobo Rouces. 2019. Aspect-based sentiment analysis using bert. In *Proceedings of the 22nd nordic conference on computational linguistics*, pages 187–196.

Divyanshu Kakwani, Anoop Kunchukuttan, Satish Golla, NC Gokul, Avik Bhattacharyya, Mitesh M Khapra, and Pratyush Kumar. 2020. Indicnlpsuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for indian languages. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 4948–4961.

Xuebo Liu, Longyue Wang, Derek F Wong, Liang Ding, Lidia S Chao, Shuming Shi, and Zhaopeng Tu. 2021. On the complementarity between pre-training and back-translation for neural machine translation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 2900–2907.

Gaurav Mohanty, Pruthwik Mishra, and Radhika Mamidi. 2020. Annotated corpus for sentiment analysis in odia language. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2788–2795.

Maria Pontiki, Dimitrios Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad Al-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, et al. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *International workshop on semantic evaluation*, pages 19–30.

Maria Pontiki, Haris Papageorgiou, Dimitrios Galanis, Ion Androutsopoulos, John Pavlopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. *SemEval 2014*, page 27.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, 21(140):1–67.

Sadaf Rani and Waqas Anwar. 2020. Resource creation and evaluation of aspect based sentiment analysis in urdu. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: Student Research Workshop*, pages 79–84.

Yashwanth Reddy Regatte, Rama Rohit Reddy Gangula, and Radhika Mamidi. 2020. Dataset creation and evaluation of aspect based sentiment analysis in telugu, a low resource language. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 5017–5024.

Sanjib Kumar Sahu, Priyanka Behera, DP Mohapatra, and Rakesh Chandra Balabantaray. 2016. Sentiment analysis for odia language using supervised classifier: an information retrieval in indian language initiative. *CSI transactions on ICT*, 4:111–115.

Kazutaka Shimada. 2023. Dataset construction and evaluation for aspect-opinion extraction in bangla fine-grained sentiment analysis. In *International Conference on Data Science and Applications*, pages 437–449. Springer.

Chi Sun, Luyao Huang, and Xipeng Qiu. 2019. Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence. 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)*, pages 380–385.

R Sunitha et al. 2023. Mabsa: A curated malayalam aspect based sentiment analysis dataset on movie reviews. *Data in Brief*, page 109452.

Duyu Tang, Bing Qin, and Ting Liu. 2016. Aspect level sentiment classification with deep memory network. *arXiv preprint arXiv:1605.08900*.

Kai Wang, Weizhou Shen, Yunyi Yang, Xiaojun Quan, and Rui Wang. 2020. Relational graph attention network for aspect-based sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3229–3238.

Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. 2016. Attention-based lstm for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 606–615.

Wei Xue and Tao Li. 2018. Aspect based sentiment analysis with gated convolutional networks. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2514–2523.

A Appendix

https://ritikart.com
https://thrillophilia.com/states/orissa/
https://myodishaproducts.com/
https://odishahandicrafts.com
https://odishaforum.com/
https://www.ameodia.com
https://www.boyanika.com/
https://utkalamrita.com/
https://odialive.com/

Table 6: Data sources