ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 14 (2024) No. 9, pp. 225-240
DOI:10.33168/JSMS.2024.0914

Subjectivity-Aware Aspect-Based Sentiment Analysis using BERT: Improving Performance through Objective Sentence Filtering

Wing Kin Chong[1], Hu Ng[1*], Timothy Tzen Vun Yap[2], Vik Tor Goh[3], Lai Kuan Wong[1], Ian Kim Teck Tan[2], Dong Theng Cher[4]

[1] Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia
[2] School of Mathematical & Computer Sciences, Heriot-Watt University, Putrajaya, Malaysia
[3] Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia
[4] SIRIM Berhad, Shah Alam, Malaysia
nghu@mmu.edu.my

Abstract. This study investigates the impact of incorporating subjectivity detection as a preprocessing step in aspect-based sentiment analysis (ABSA) using the BERT model. The proposed approach, named Sub-BERT, filters out objective sentences from the dataset before performing ABSA, allowing the model to focus on classifying the sentiment polarity of subjective sentences. Experiments are conducted on two datasets, SemEval14 and MEMD, covering various domains such as laptops, restaurants, books, clothing, and hotels. The results demonstrate that Sub-BERT outperforms the baseline BERT model, achieving improvements of up to 22% in accuracy and 28% in F1 score for the SemEval14 dataset, and up to 5% in accuracy and 33% in F1 score for the MEMD dataset. The findings highlight the importance of subjectivity detection in enhancing the performance of ABSA, particularly in scenarios with limited training data. However, the study's limitations, such as the reliance on manual subjectivity filtering and the impact of dataset imbalance, should be considered when interpreting the results. Future research directions include exploring techniques for addressing dataset imbalance and developing a unified framework for end-to-end ABSA with subjectivity detection.

Keywords: Aspect-based Sentiment Analysis, ABSA, objectivity and subjectivity

1. Introduction

The growth of internet commerce has enabled many merchants to reach a larger client base. In the last decade, the rise of microblogs and e-commerce websites has allowed people to voice their thoughts on a variety of topics, including services, products, and political candidates. Using customer feedback to evaluate the quality of products and services is one way for businesses to gain a competitive edge. Merchants will be able to quickly respond to their customers' level of satisfaction and recommend appropriate steps, such as utilizing recommendation systems (Chow et al., 2023; Lim et al., 2023), based on opinion mining or sentiment analysis from these evaluations in order to address any identified issues effectively.

On the other hand, conventional sentiment analysis or opinion mining typically focuses on predicting sentiment at the sentence or document level, aiming to identify overall sentiments expressed in the entire sentence or document. This approach assumes that each sentence contains only one sentiment toward the entire content, which may not accurately reflect real-world scenarios. As a result, merchants may misinterpret true consumer sentiments about their products or services when using this method to make predictions based on reviews. In response to this challenge, researchers have been actively working on aspect-based sentiment analysis for many years.

Aspect-based sentiment analysis expands the scope of traditional sentiment analysis by considering specific aspects or product features as the focal point for determining polarity, rather than treating the entire text as a single unit (Ma et al., 2018). ABSA operates at a more detailed level, where it identifies targets or aspects and then predicts their associated sentiments. It encompasses several subtasks to detect aspect-level sentiment elements such as aspect terms, aspect categories, opinion terms, and sentiment polarities. For instance, in the sentence “The pizza is delicious but the service is bad.”, this sentence contains two aspects which are “pizza” and “service” two contrasting aspects – "pizza" and "service" - with positive and negative sentiments respectively can be identified separately in ABSA instead of being classified neutrally due to opposing sentiments cancelling each other out. This can provide merchants with a more accurate opinion toward their product or service.

Traditional sentiment analysis aims to understand overall sentiments expressed in a piece of text, while Aspect-Based Sentiment Analysis focuses on capturing fine-grained sentiments toward specific aspects or product features. ABSA provides the advantage of enabling merchants to better comprehend how customers perceive individual aspects of their products. Conventional sentiment analysis often overlooks the nuanced nature of customer opinions by assuming that each sentence conveys a single sentiment towards the entire content. This oversimplified approach may lead to misinterpretations and inaccurate assessments of customer sentiments, especially in scenarios where multiple conflicting opinions are present within a single sentence.

Deep learning techniques have been utilized with success in ABSA tasks, for example, Song et.al. (Song et al., 2020) investigated the transformer-based model known as Bidirectional Encoder from Transformer (BERT) and achieved positive outcomes. Kaur & Sharma (Kaur & Sharma, 2023) adopt hybrid feature extraction to obtain review-related features and aspect-related features. While many current approaches focus on using machine learning or deep learning models to discern the relationship between aspect and context, the main challenge in ABSA lies in the limited training data availability. Hence, it is crucial to effectively utilize the available dataset for generating high-quality training data.

In this research, the notion of subjectivity and objectivity are of interest. Subjectivity detection serves as a fundamental preprocessing step in sentiment analysis, especially in ABSA tasks. It involves identifying and distinguishing subjective and objective expressions within textual data. This distinction is essential as it enables sentiment analysis models to accurately capture and interpret subjective opinions and emotions expressed by users, thus providing a more nuanced understanding of customer sentiments towards specific aspects or product features. The proposition is that by performing subjectivity classification before ABSA, a dataset that is only labelled as positive and negative can be obtained which can help to leverage the result of ABSA. The objectives of this work include:

- Exploring the effects of applying subjectivity detection before ABSA.
- Conducting experiment on ABSA dataset with different domains using BERT model to investigate the effect of subjectivity detection before ABSA on different domains.
- Comparing the performance of ABSA when coupled with subjectivity classification versus existing results where ABSA is used alone.

2. Literature Review

Natural language is how people communicate with one another. According to Sirbu (2015), “language is a general, abstract aspect and a sum of organization skills and principles; it is the system that governs any concrete act of communication”. In this era of big data, language processing is crucial for performing a wide range of activities such as sentiment analysis and chatbot. Natural Language Processing (NLP) is one of the research interests in the field of machine learning and artificial intelligence. With the use of NLP, computers and other devices could be able to comprehend language on par with humans. However, teaching machines or computers to understand human language is a difficult endeavor due to several complexities which need to be resolved.

2.1. Aspect-Based Sentiment Analysis

Aspect-Based Sentiment Analysis (ABSA) is defined as the problem to identify sentiment elements of interest for a concerned text item, either a single sentiment element, or multiple elements with the dependency relation between them according to the survey paper by Zhang et al. (2022). Conventional sentiment analysis problem has two key components which are the target and sentiment. While for ABSA, the aspect term α or aspect category c are used to represent the target and use opinion term o and sentiment polarity p to describe the sentiment in a more detailed opinion expression.

Since there are various sentiment element in ABSA, this contributes to various ABSA task according to the elements involved in a task. ABSA tasks can be divided into two categories, namely single and compound ABSA tasks. There are four tasks in single ABSA tasks which are aspect term extraction (ATE), aspect category detection (ACD), opinion term extraction (OTE) and aspect sentiment classification (ASC). Compound ABSA tasks, on the other hand, are the combination of single ABSA tasks that can be further categorized into different tasks which consists of pair extraction, triplet extraction and quad extraction.

Some of the recent works on ABSA shows promising results using the deep learning techniques and attention mechanism such as Kaur & Sharma (2023) perform ABSA with deep learning-based model using hybrid feature extraction approach. Tian et al. (2023) introduced the attention-based multi-level feature aggregation (AMFA) network by applying attention mechanism. Both works uses accuracy and F1 score for evaluating the results since ABSA is a classification problem.

2.2. Sentence Embedding

The concept of sentence embedding involves converting a block of words or sentences into a single numerical vector. This process aims to encapsulate the semantic information of a sentence, enabling machines and computers to comprehend context, similarity, contradiction, and other intricacies within the corpus. In contrast to word embedding, sentence embedding can address the limitation where only limited information is derived through word embedding alone. A basic approach for generating sentence embeddings is by utilizing the word embedding model to encode a given sentence and then obtaining the average of these vectors. While this method serves as an initial benchmark, it does not effectively capture details related to word order and overall semantics of sentences. Sentence embedding holds significant importance in various NLP applications as demonstrated by Kiros et al. (2015); Reimers & Gurevych (2019); Subramanian et al. (2018).

2.3. Transformer

Recently, the use of transformers has been rapidly accelerating research progress in natural language processing (NLP). Many research in NLP has achieved good results with the use of transformer based model. The concept of the original transformer model is introduced in 2017 (Vaswani et al., 2017). It uses an encoder-decoder architecture which is based on attention mechanism. The original transformer model contains a stack of 6 layers encoder and decoder. Figure 1 shows the architecture of a transformer model.

The transformer consists of a 6-layer encoder stack on the right and a 6-layer decoder stack on the left. Input data is processed through an attention sub-layer and FeedForward Network sub-layer in the encoder, while target outputs pass through two attention sub-layers and a FFN layer in the decoder. Transformers employ self-attention, which is a special type of attention mechanism that calculates relationships between words within a sequence using query, key, and value vectors obtained from weight matrix multiplications during training.

The transformer architecture has significantly improved natural language processing tasks by using parallelization to enhance training and inference. Unlike traditional deep learning models, transformers process entire sequences of words at once, which accelerates training and helps capture complex word relationships more effectively. Additionally, transformers use self-attention mechanisms to weigh the importance of different words in a sequence, enabling them to handle long-range dependencies and nuanced relationships between words. Due to these advantages, transformers have shown superior performance in NLP tasks like sentiment analysis and opinion mining, making them the preferred architecture for many NLP applications.

2.4. BERT

Bidirectional encoder representation from transformer (BERT) which is proposed by Devlin et al (2019) make use of the transformer model to achieve the new state of the art and is the first breakout in large language model (LLM). It is a contextual embedding model that can capture semantic information, providing greater results in many NLP tasks such as text generation, sentence classification etc. compared to traditional embedding methods such as Word2Vec.

BERT uses only the encoder from transformer as the name suggests and is designed to pretrain representations from unlabelled text in bidirectional by jointly conditioning on both left and right context in all layers. The encoder can be stacked up to N number and the representation size of each will be the size of the encoder layer. It uses the multi-head attention mechanism to relate each word in the sentence to all other words in the sentence in order to learn the relationship and contextual meaning of the words.

BERT can return the representation for each word of an input sentence as well as returning a special token called [CLS] which is the representation of the whole sentence. This allows BERT to be fine-tunable for different downstream tasks using transfer learning by as simply as adding one additional output layer for a wide range of tasks. Furthermore, BERT can learn the context of the word embedding which results in good performance in many tasks. In this research, BERT is used as a word embedding to capture the semantics of the words in sentences.

2.5. Subjectivity and Objectivity

User reviews have components of subjectivity and objectivity, and quantifying the degree of subjectivity and objectivity in a corpus is difficult. Text can be characterized in two different ways: objective and subjective. Subjective writing is writ-ten with the intention of expressing opinions and personal viewpoints, whereas objective text is created with the intention of communicating facts and information. Furthermore, as virtually all contents are written by humans, an absolute objective corpus may be relative as reviews.

Objectivity and subjectivity are frequent categories used in text classification. Subjectivity classification aims to identify and distinguish objective and subjective information in text. With the exponential growth of opinion-rich resources like social media, there is a need for advanced subjectivity analysis techniques that can separate factual statements from personal viewpoints. Accurate subjectivity classification has become critical for applications like sentiment analysis, where only subjective statements indicate polarity.

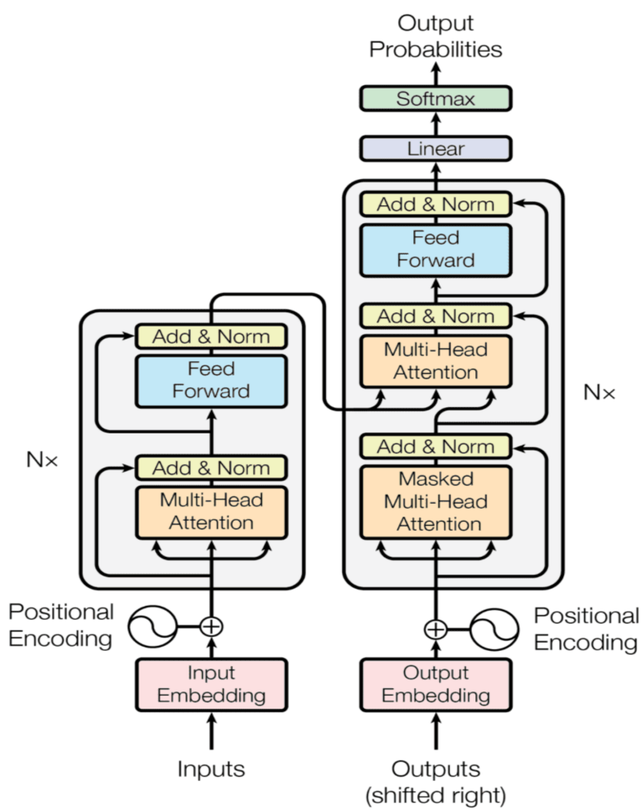


Fig. 1: The architecture of a transformer model (Vaswani et al., 2017)

Some of the works involving subjectivity classification are such as Chiha and Ben Ayed (2020) where they used machine learning approach to do subjectivity classification. Zhao et al. (2022) uses subjectivity text characteristics as the sub-technique to solve bigger NLP problems. Satapathy et al. (2022) combine subjectivity and sentiment detection into a multitask learning framework.

Tkachenko et al. (2018) has shown their interest in the notion of subjectivity. They generated two-word embeddings using objective corpus and subjective corpus using Word2Vec to perform comparative analyses between subjective and objective corpora. Furthermore, Pecar & Simkon (2021) explore the possibility of knowledge transfer about subjectivity between sub-tasks in end-to-end ABSA. By proposing subjectivity-aware learning as a novel auxiliary task, the experiment shows that model performance improve in overall F1 score for ABSA using the SemEval 2015 task 12 dataset.

3. Methodology

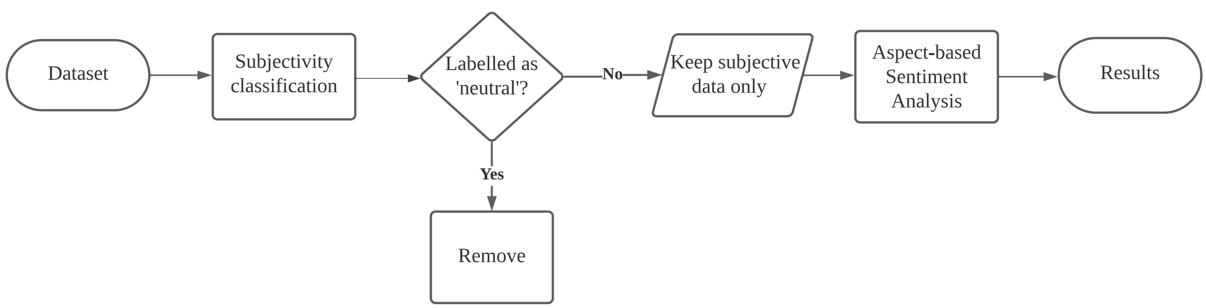
This paper presents a method to enhance ABSA performance by implementing subjectivity detection. In this section, detailed explanation of the steps involved in this experiment including ABSA task involved, dataset and model used will be presented.

3.1. Task Description

In the previous section, ABSA was mentioned to include several subtasks. This experiment focuses on aspect-sentiment classification (ASC), which is a subtask of ABSA. The goal of ASC is to classify the sentiment polarity with respect to each aspect in a sentence. When given a sentence-aspect pair, ASC will predict the sentiment polarity (positive, negative, or neutral) of the sentence over the aspect. For example, if we take the sentence 'The sweet potato were crunchy', and and focus on the aspect word 'sweet potato', the ASC model need to determine the polarity of the aspect word as either positive, negative and neutral.

In this work, subjectivity detection was employed as a technique to filter out sentences that are labelled as neutral before conducting ABSA. where the model only needs to determine either positive or negative given each sentence and aspect word. This approach enabled the model to focus solely on the task of determining whether each sentence and aspect word pair is positive or negative.

The dataset first undergoes a subjectivity classification step to filter out the objective statements and keep only the subjective statements. Then, the remaining dataset goes through a preprocessing step including tokenization and adding special tokens before passing into a classification model to carry out



aspect sentiment classification. Figure 2 provides an overall flowchart for the experiment.

Fig. 2: Flowchart of ABSA with subjectivity classification.

3.2. Dataset

This section introduces the two ABSA datasets used in this research. The first dataset is, referred to as the SemEval14 dataset throughout this paper, is derived from the SemEval2014 Task 4 and is one of the popular datasets among ABSA research. SemEval dataset is a benchmarking dataset in the field of ABSA. It includes details about the aspect word for each sentence and assigns a sentiment polarity (positive, negative, or neutral) to each aspect. By default, it is divided into training and test sets. Table 1 provides the number of sentences in the dataset.

Table 1: Statistic of SemEval14 dataset

Domains	Laptops	Restaurants	Total
Train	2358	3241	5599
Test	654	1120	1774
Total	3012	4361	7373

The second dataset is the Multi-Element Multi Domain ABSA datasets (MEMD) (Cai et al., 2023). This is a dataset containing five different domains with labelling on each element of ABSA. MEMD dataset is a relatively new dataset compared to SemEval dataset where it covers more domains than the SemEval dataset. The domains included in this dataset are books, clothing, hotel, restaurant, and laptop. All domains are labelled with the four elements which are aspect, category, opinion word and sentiment and are suitable for different ABSA tasks. Table 2 shows the statistics of the MEMD dataset. Some of the examples in each domain of the MEMD dataset are presented in Table 2.

Table 2: Statistic of MEMD dataset

Domain	**Books**	**Clothing**	**Hotel**	**Restaurant**	**Laptop**	**Total**
Train	2092	1674	2481	3622	2863	12732
Dev	583	466	696	1020	818	3583
Test	292	233	349	510	405	1789
Total	2967	2373	3526	5152	4076	18094

The SemEval14 dataset includes attributes such as the sentence, aspect, and its corresponding sentiment. MEMD dataset includes sentence and the four elements of ABSA which are aspect, corresponding category, opinion word, and sentiment. Examples for each domain in both datasets can be found in Table 3 and Table 4. In the MEMD dataset, ‘POS’ denotes positive sentiments while 'NEG' represents negative ones; 'NEU' signifies neutral sentiments. In the case of labelling as ‘NULL’ in aspect attribute, this means that the aspect word is not explicitly stated in the sentence. This is one of the reasons where the dataset is domain specified so that it is possible to distinguish the aspects under different subject matter when it is not explicitly stated. One example is shown in Table 4 where in the hotel domain, the sentence does not explicitly state the aspect, but it is known that the review is referring to a hotel room since it is in the hotel domain.

Table 3: Example for SemEval14 dataset

Domain	**Sentence**	**Aspect**	**Sentiment**
Laptops	nightly my computer defrags itself and runs a virus scan.	*virus scan*	neutral
Laptops	it's a great product for a great price!	*price*	positive
Restaurants	the price is reasonable although the service is poor.	*service*	negative

Table 4: Example for MEMD dataset

Domain	**Sentence**	**Aspect**	**Category**	**Opinion**	**Sentiment**
Books	I have read a lot of Stuart McBride books and loved them	Stuart McBride books	Book#General	loved	POS
Clothing	They are also super warm !	They	Bottom#Quality	super warm	POS
Hotel	Clean , comfortable , and friendly near just where we wanted to be .	NULL	Hotel#Comfort	Comfortable	POS
Restaurant	Nice style , nice music , good food and service .	style	Ambience#General	Nice	POS
Laptop	one of the two usb ports is defective , which is an enormous pain .	usb ports	Ports#Operation _Performance	pain	NEG

The datasets are labelled with three types of polarity which are positive, negative and neutral. Figures 3 and 4 show the distribution of the polarity of the SemEval14 and MEMD dataset respectively. It can be observed that the 'neutral' label represents the smallest proportion of sentences in both datasets.

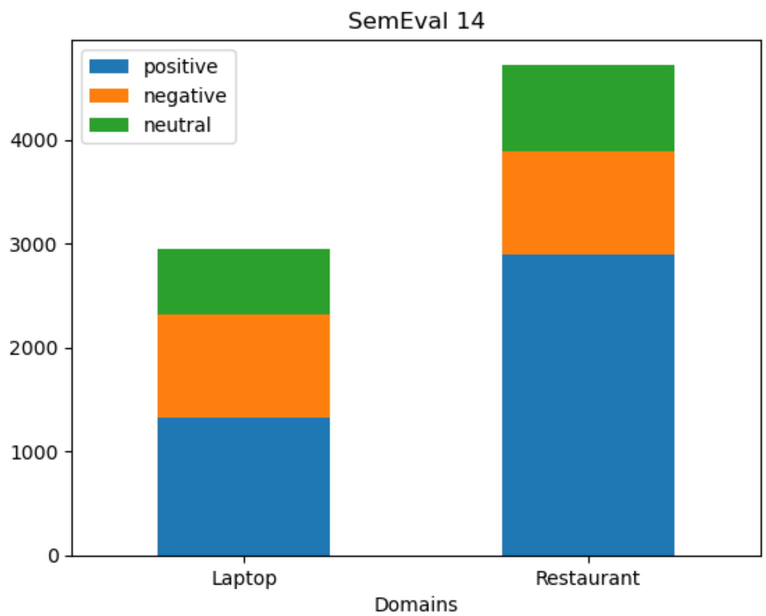


Fig. 3: Polarity distribution per domain of SemEval14 dataset

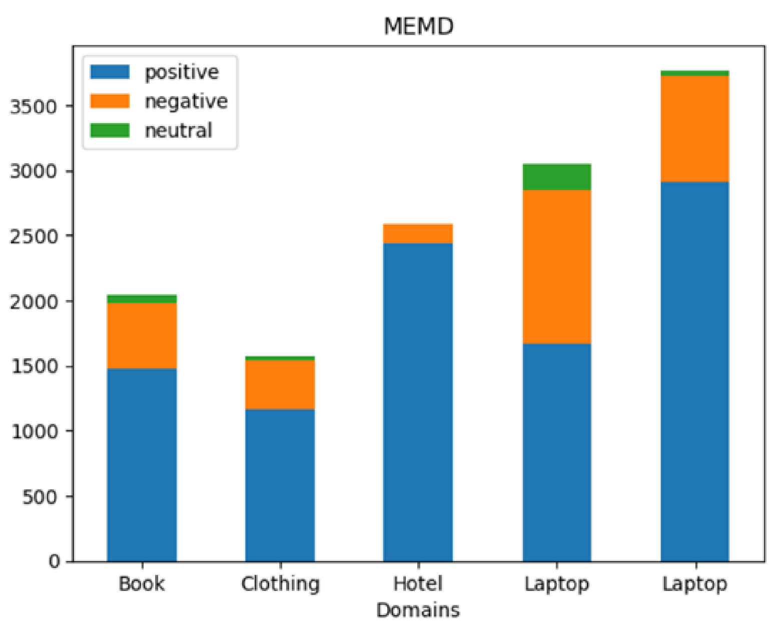


Fig. 4: Polarity distribution per domain of MEMD dataset

3.3. Subjectivity Filtering

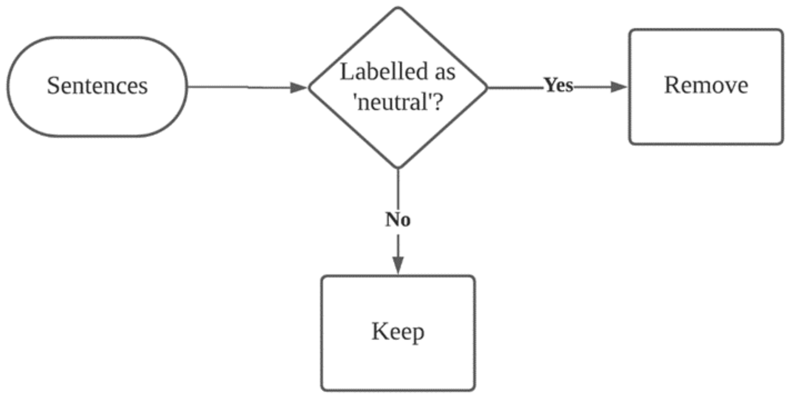


Fig. 5: Flowchart for subjectivity detection

Since the dataset contains sentences with labelling of ‘neutral’, those sentences were considered objective and subsequently removed from the dataset. This filtering process resulted in a dataset containing only subjective sentences labelled as 'positive' or 'negative' for training purposes. Figure 5 shows the flowchart for the subjectivity detection process.

A sentence with a neutral polarity is categorized as an objective statement, whereas a sentence with either positive or negative polarity is labelled as subjective. Examples of both objective and subjective statements from the datasets are presented in Table 5.

Table 5: Examples of objective and subjective statements

Objective	**Subjective**
just made the move from pc to macbook !	They are also super warm !
my computer was used on average a couple hours a day .	i just wonder how you can have such a delicious meal for such little money.
As one might expect of a book that covers so much ground , its not thin .	i charge it at night and skip taking the cord with me because of the good battery life.

After removing the objective statements, the datasets contain only subjective statements. Table 6 and Table 7 display the remaining sentence counts for SemEval14 and MEMD dataset, respectively. The total number of sentences in the SemEval14 dataset has decreased from 7373 to 6215. Similarly, for the MEMD dataset, the number of sentences has been reduced from 18094 to 12675.

Table 6: Number of sentences after subjectivity filtering for SemEval14 dataset

Domains	**Laptops**	**Restaurants**	**Total**
Train	1853	2969	4822
Test	469	924	1393
Total	2322	3893	6215

Table 7: Number of sentences after subjectivity filtering for MEMD dataset

Domain	Books	Clothing	Hotel	Restaurant	Laptop	Total
Train	1397	1084	1814	2612	2051	8958
Dev	196	152	255	372	285	1260
Test	390	302	515	738	512	2457
Total	1983	1538	2584	3722	2848	12675

3.4. Sub-BERT

To conduct the experiment for aspect-sentiment classification (ASC) task, the BERT-base-uncased model is chosen based on the work by Cai et al. (2023) and Song et al. (2020). The BERT-based model is applied since BERT outperforms previous non-BERT-based models on ABSA task. As a result, non BERT-based models are not considered for comparison. BERT model was originally pretrained using masked language modeling and next sentence prediction techniques.

Since BERT is trained on next-sentence prediction task, the second sentence input is utilized to serve as input for the aspect word. Figure 6 shows the architecture of BERT for ABSA in this work. A sentence is passed into 12 layers of encoder to produce the hidden output of each token. The last layer of hidden output serves as the embedding of each word. The embedding of the [CLS] token which is a special token added at the start of a sentence represents the whole sentence and is fed into the classifier. The sentence and aspect word are separated by a special token named [SEP] added at the end of a sentence. The classifier will determine the sentiment based on the embedding of the [CLS] token.

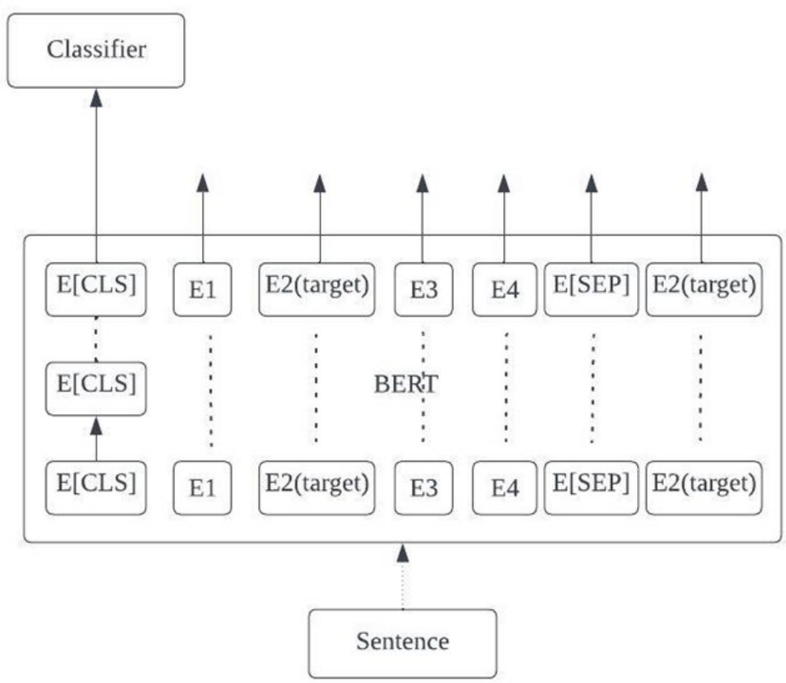


Fig. 6: BERT architecture for ABSA

The default configuration of the BERT model is used following Song et al. (Song et al., 2020) where Adam optimizer with learning rate of 2e-5 is applied to update the parameters and the dropout rate is set to 0.1 to ensure a fair comparison. The number of epochs is set to 10. For evaluation metrics, accuracy and macro-F1 score are used so that results are comparable. Table 8 shows the details configuration and parameters of the BERT model.

Table 8: Configuration of BERT model

Configuration	Value
hidden_size	768
num_hidden_layers	12
learning_rate	2e-5
optimizer	Adam
number of epochs	10
dropout_rate	0.1

4. Performance Evaluation

The performances are presented in this section and the comparison of the results in this investigation with the ones from Cai et al. (2023) and Song et al. (2020) are discussed. The effect of removing objective sentences from ABSA dataset are also discussed in the following Sub-Sections. The approach in this experiment is mentioned as Sub-BERT in this paper.

4.1. SemEval14 Dataset

The performance of aspect-based sentiment analysis on SemEval14 dataset is shown in Table 9. Compared to the work, $BERT_{BASE}$ in Table 9, by Song et al. (2020), the performance of ASC task in laptop domain is improved by 22% for both accuracy and F1 score. While in restaurant domain, the accuracy and F1 score are im-proved by 11% and 17% respectively.

Table 9: Performance of ABSA for each domain in SemEVal14 dataset

Domains	**Laptop**		**Restaurant**	
Metrics	Acc.(%)	F1(%)	Acc. (%)	F1(%)
$BERT_{BASE}$ (Song et al. 2020)	74.66	68.64	82.91	73.20
Sub-BERT	92.32	90.51	93.88	90.60

The confusion matrix in Figure 7 shows that the percentage of false positive and false negative in both laptop and restaurant domain are low which lead to a higher F1 score.

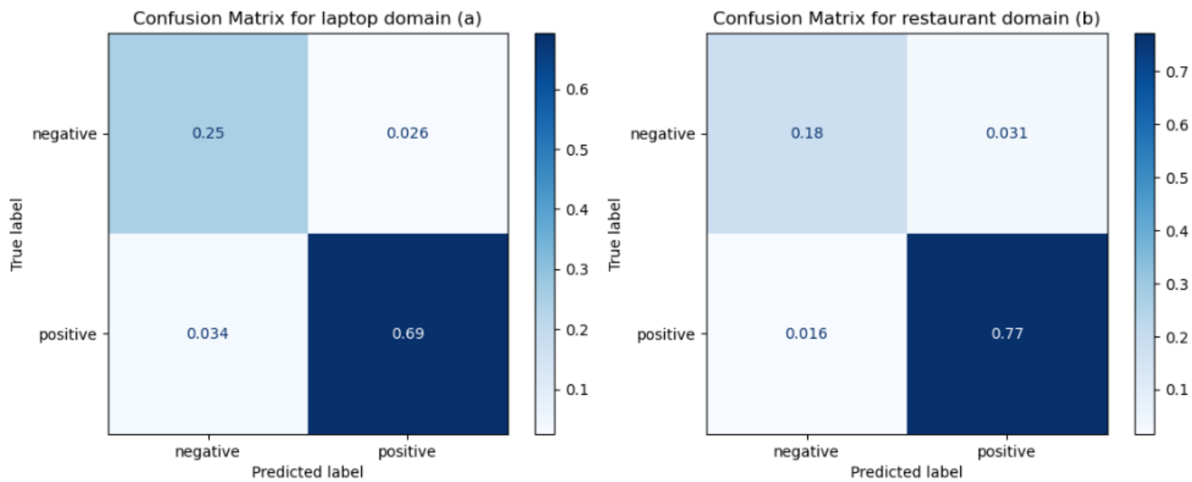


Fig 7: Confusion matrix of different domains in SemEval14 dataset (a) laptop; (b) restaurant

4.2. MEMD Dataset

The results of ABSA using BERT model on MEMD dataset are shown in Table 10. When compared to the work by Cai et al. (2023), labelled as BERT in Table 10, the accuracy for book domain increased by 5% followed by clothing, laptop and restaurant which their accuracy improved by 2-3%. The F1 score for books, clothing, laptop and restaurant increased by 28%, 33%, 18% and 28% respectively. While for hotel domain, both the accuracy and F1 score dropped by 1% and 12% respectively.

Table 10: Performance of ABSA for each domain in MEMD dataset

Domains	**Books**		**Clothing**		**Hotel**		**Laptop**		**Restaurant**	
Metrics	Acc. (%)	F1(%)	Acc. (%)	F1(%)	Acc. (%)	F1(%)	Acc. (%)	F1(%)	Acc. (%)	F1(%)
$BERT_{BASE}$ (Song et al. 2020)	82.39	54.97	90.60	58.70	96.95	82.38	88.11	71.67	88.57	56.74
Sub-BERT	87.79	82.29	93.44	91.18	95.30	70.84	90.00	89.82	90.11	84.90

The confusion matrices for MEMD dataset are shown in Figure 8. For the domains: books, clothing, laptop and restaurant, that the number of true negative and true positive are quite high which is evidence of high F1 score. For hotel domain, since the dataset are highly imbalance, the number of true negatives is almost the same as false positive and false negative which result in a lower F1 score. Interestingly, more type II error than type I error are observed in all of the domains in MEMD dataset

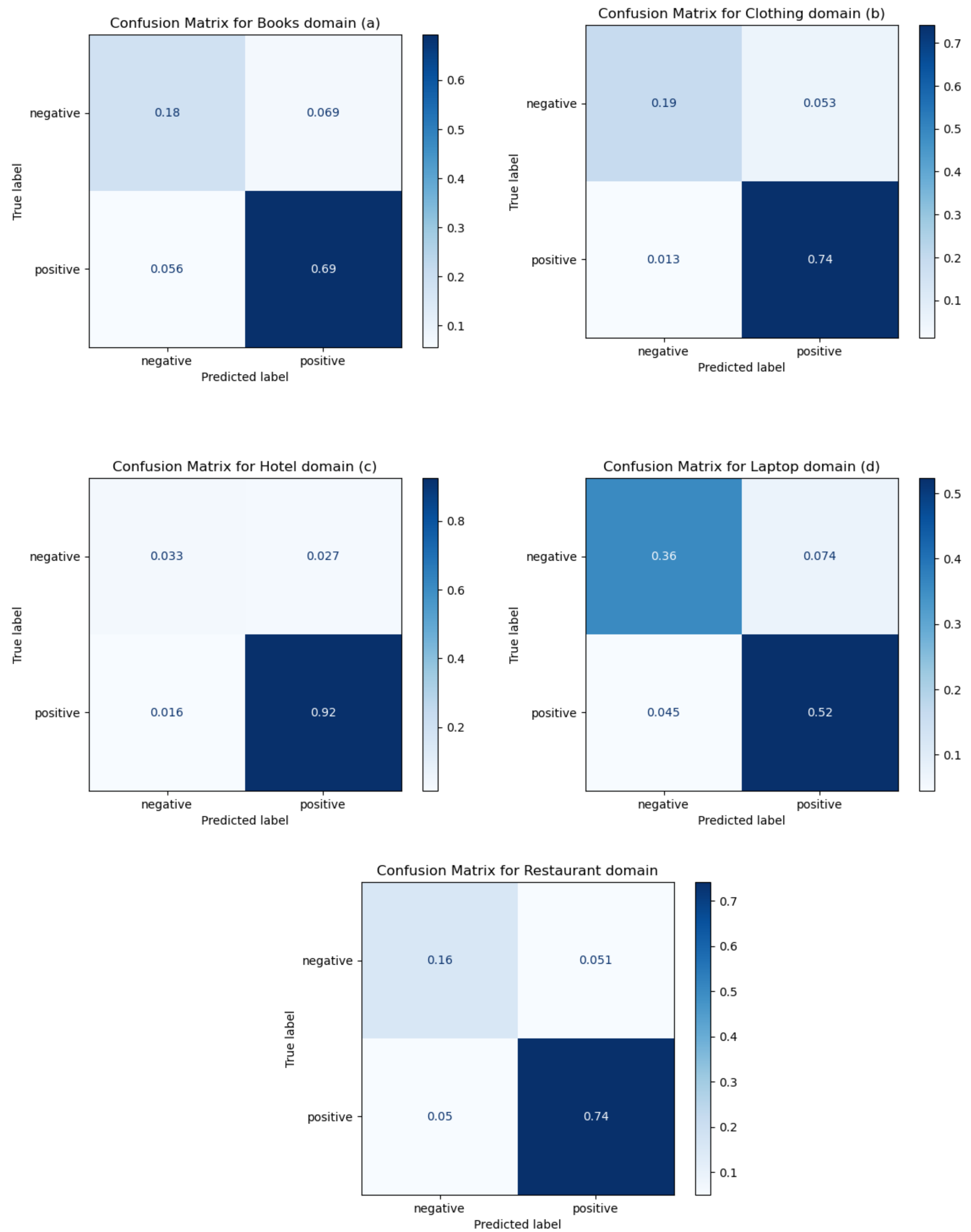


Fig. 8: Confusion matrix for different domains in MEMD dataset (a) Books, (b) Clothing, (c) Hotel, (d) Laptop, (e) Restaurant

4.3. Discussion

The Sub-BERT achieved an overall improvement as compared to the previous work. There is an overall increase in accuracy and F1 score for both datasets.

The accuracy and F1 scores in the SemEval14 laptop domain increased by 22% compared to the results from Song et al (2020). In the restaurant domain, there was a 10% increase in accuracy and a 20% increase in F1 score when compared to the work of Cai et al. (2023). The MEMD dataset showed an average increase of 3% in accuracy across all domains except for hotels. However, for F1 score, books, clothing, and restaurants displayed an average increase of 30%, while laptops increased by 18%. In contrast, there was little change in result performance for hotel domain's accuracy but a decrease in its F1 score. This difference might be attributed to the highly imbalanced nature of the dataset as mentioned in the work of Cai et al. (2023).

In the research conducted by Ma et al. and Song et al., the datasets used in the study comprise both objective and subjective statements. The challenge lies in training the model to categorize statements into one of three classes, despite having a scarcity of training samples for neutrally labeled sentences, which consequently leads to decreased performance. In contrast, this investigation focuses on excluding the objective statements from consideration.

While the proposed approach may contain potential limitations due to simply removal of neutral sentences without considering other factors. Despite the extensive use of factual sentences with neutral sentiment, there is a notable distinction between genuine neutral sentiment in opinionated sentences and the sentiment found in factual statements. Nevertheless, the detection of subjectivity is vital for enhancing aspect-based sentiment analysis, as it offers important insights into the objectivity or subjectivity of expressed opinions or sentiments. These findings propose that rather than conducting ABSA directly on datasets containing both objective and subjective statements, integrating a step for subjectivity classification in advance can be beneficial. This is particularly useful in scenarios with objective statements, allowing for filtering out such statements before performing ABSA on two distinct classes - positive and negative. This approach aims to enhance the model's performance significantly.

Pecar & Simko (2021) have also demonstrated this, showing improved performance in their experiment with an overall increase in F1 score. In addition to the SemEval 2015 dataset, this work utilized the MEMD dataset, which includes new domains such as books, clothing and hotels for ABSA experiments. This demonstrates that subjectivity-aware learning can be generally applied across various domains of the ABSA dataset.

5. Conclusion & Future Work

This study demonstrates the effectiveness of incorporating subjectivity detection as a preprocessing step in aspect-based sentiment analysis using the BERT model. The proposed Sub-BERT approach achieves significant improvements in accuracy and F1 score compared to the baseline BERT model on both the SemEval14 and MEMD datasets, highlighting the importance of filtering out objective sentences before performing ABSA. The findings suggest that subjectivity detection can help mitigate the challenges posed by limited training data in ABSA and improve the model's ability to capture fine-grained sentiments towards specific aspects. However, the study's limitations, such as the reliance on manual subjectivity filtering and the impact of dataset imbalance, should be acknowledged. Future research should explore more sophisticated techniques for subjectivity detection, such as using pre-trained classifiers or unsupervised methods, and investigate the generalizability of the proposed approach to other sentiment analysis tasks and domains. Additionally, the development of a unified framework for end-to-end ABSA with subjectivity detection could further enhance the practicality and efficiency of the proposed approach. Despite these limitations, this study provides valuable insights into the potential of subjectivity detection for improving ABSA performance and lays the foundation for future research in this direction.

Acknowledgements

This research was funded by the Malaysia Ministry of Higher Education (MOHE), under grant No. FRGS/1/2022/ICT06/MMU/03/3.

References

Cai, H., Song, N., Wang, Z., Xie, Q., Zhao, Q., Li, K., Wu, S., Liu, S., Yu, J., & Xia, R. (2023). MEMD-ABSA: A Multi-Element Multi-Domain Dataset for Aspect-Based Sentiment Analysis.

Chiha, R., & Ben Ayed, M. (2020). Supervised Machine Learning Approach for Subjectivity/Objectivity Classification of Social Data (pp. 193–205). https://doi.org/10.1007/978-3-030-44322-1_15

Chow, Y. Y., Haw, S. C., Naveen, P., Anaam, E. A., & Mahdin, H. Bin. (2023). Food Recommender System: A Review on Techniques, Datasets and Evaluation Metrics. Journal of System and Management Sciences, 13(5), 153–168.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. Journal of Big Data, 10(1), 5. https://doi.org/10.1186/s40537-022-00680-6

Kiros, R., Zhu, Y., Salakhutdinov, R., Zemel, R. S., Torralba, A., Urtasun, R., & Fidler, S. (2015). Skip-Thought Vectors.

Lim, Y. F., Haw, S. C., Ng, K. W., & Anaam, E. A. (2023). Hybrid-based Recommender System for Online Shopping: A Review. Journal of Engineering Technology and Applied Physics, 5(1), 12–34. https://doi.org/10.33093/jetap.2023.5.1.3

Ma, Y., Peng, H., & Cambria, E. (2018). Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). https://doi.org/10.1609/aaai.v32i1.12048

Pecar, S., & Simko, M. (2021). Exploiting Subjectivity Knowledge Transfer for End-to-End Aspect-Based Sentiment Analysis (pp. 269–280). https://doi.org/10.1007/978-3-030-83527-9_23

Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androut-sopoulos, I., & Manandhar, S. (2014). SemEval-2014 Task 4: Aspect Based Sentiment Analysis. Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 27–35. https://doi.org/10.3115/v1/S14-2004

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.

Satapathy, R., Pardeshi, S., & Cambria, E. (2022). Polarity and Subjectivity Detection with Multitask Learning and BERT Embedding.

Sirbu, A. (2015, December). THE SIGNIFICANCE OF LANGUAGE AS A TOOL OF COMMUNICATION. https://doi.org/10.21279/1454-864X

Song, Y., Wang, J., Liang, Z., Liu, Z., & Jiang, T. (2020). Utilizing BERT Intermediate Layers for Aspect Based Sentiment Analysis and Natural Language Inference.

Subramanian, S., Trischler, A., Bengio, Y., & Pal, C. J. (2018). Learning General Purpose Distributed Sentence Representations via Large Scale Multi-task Learning.

Tian, D., Shi, J., & Feng, J. (2023). A Self-Attention-Based Multi-Level Fusion Network for Aspect Category Sentiment Analysis. Cognitive Computation, 15(4), 1372–1390. https://doi.org/10.1007/s12559-023-10160-5.

Tkachenko, M., Chia, C. C., & Lauw, H. (2018). Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 1212–1221. https://doi.org/10.18653/v1/P18-1112

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ukasz, & Polosukhin, I. (2017). Attention is All you Need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in Neural Information Processing Systems. (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Zhang, W., Li, X., Deng, Y., Bing, L., & Lam, W. (2022). A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges. IEEE Transactions on Knowledge and Data Engineering, 1–20. https://doi.org/10.1109/TKDE.2022.3230975

Zhao, Z., Zhang, Z., & Hopfgartner, F. (2022). Utilizing subjectivity level to mitigate identity term bias in toxic comments classification. Online Social Networks and Media, 29, 100205. https://doi.org/10.1016/j.osnem.2022.100205