International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancing Aspect Term Extraction from Customer Reviews with Sparse Gated Recurrent Units (SGRUs) in the Context of BERT and NER

Mohammed Ziaulla*[1], Dr. Arun Biradar[2]

Submitted: 28/01/2024 **Revised**: 06/03/2024 **Accepted**: 14/03/2024

Abstract: A product evaluation system is a potent instrument that analyzes many online product reviews and recommends suitable products to consumers. A large amount of unlabeled data can be found on social networking sites. Fine-grained data annotation, on the other hand, is both costly and time-consuming. Aspect-based sentiment analysis (ABSA) aims to identify aspect terms in online reviews and predict their polarity. Sentiment analysis is a difficult and complex operation. Common subtasks include Aspect Term Extraction (ATE) and Aspect Polarity Classification (APC). When these two subtasks are trained independently, the relationship between ATE and APC should be addressed. Extracting aspect phrases from customer evaluations is crucial for sentiment analysis and opinion mining. This work proposes a novel technique for improving aspect phrase extraction in the context of BERT (Bidirectional Encoder Representations from Transformers) and Named Entity Recognition (NER) by leveraging Sparse Gated Recurrent Units (SGRUs). In order to address the challenges of aspect word extraction, we provide a synergistic combination of cutting-edge methods such as Sparse GRUs, BERT, and NER. Sparse GRUs benefit from efficient computing and improved generalization by including sparsity requirements in the GRU architecture. This novel approach seeks to gather local and contextual information to improve the precision and relevance of derived aspect phrases. Our experimental results on a benchmark data set demonstrate the effectiveness of the proposed technique. By combining Sparse GRUs with BERT and NER, we can significantly improve the accuracy of aspect term extraction over earlier methods. The tests' results suggest that Sparse GRUs can increase the identification of aspect phrases inside customer reviews, resulting in more accurate sentiment analysis and a better understanding of customer ratings.

Keywords: *Aspect Term Extraction, Customer Reviews, aspect-based sentiment analysis (ABSA), Sparse Gated Recurrent Units (SGRUs), Named Entity Recognition (NER), Aspect Term Extraction (ATE), Aspect Polarity Classification (APC)*

1. Introduction

Aspect-based sentiment analysis (ABSA) is a method for gathering detailed sentiment information from text [1-2]. Its primary objective is to recognize aspect phrases in comments and determine their sentiment polarity. "Service" and "environment" are used in the example. The two aspects produce positive and negative emotional polarities [3]. Doing a more in-depth investigation than a sentiment analysis of the full sentence is preferable because the attitudes connected to these two features are incompatible. Two subtasks—ATE and APC—are the core topics of ABSA's primary research goal. Aspect phrases and their polarity must be manually marked before the APC process is completed. However, most of the models suggested for aspect-based sentiment analysis tasks have mainly concentrated on increasing the precision of aspect polarity categorization. Unfortunately, these models have ignored the investigation of the Chinese Aspect Term Extraction (ATE) subtask. This lack of research focus has made it difficult to complete some tasks without a clear aspect extraction approach, especially when applying transfer learning to aspect-based sentiment analysis.

[1]Research Scholar, School of Computer Science and Engineering REVA University, Bangalore, India
[2]Professor, School of Computer Science and Engineering REVA University, Bangalore, India
drmdziaulla@gmail.com[1] and arun.biradar@reva.edu.in[2]
Corresponding Author Email: drmdziaulla@gmail.com[1]

ABSA's subtasks are aspect term extraction (ATE) and aspect polarity classification (APC). The APC is a text classification task. Several deep learning models based on the Transformer network technique [4-5] have been developed to solve the APC subtasks. The APC aims to forecast the precise sentiment polarity connected with each unique aspect phrase rather than ambiguously measuring the overall sentiment of phrases or documents. The APC task frequently categorizes positive, negative, and neutral polarities into three groups. Decision-makers have a more exact frame of reference when categorizing sentiment polarity based on particular factors, which enables them to extract complex emotional tendencies from reviews or tweets more accurately. The Aspect Term Extraction (ATE) uses named entity recognition (NER) to extract aspect phrases from reviews or tweets. Most methods [6] focus on ATE as a distinct activity from the APC work. Tokens from a review are separated by the ATE models, which decide which aspects each token relates to. Although numerous approaches exist to label tokens in various studies, most papers employ the IOB2 labels to annotate tokens [7].

Two new statistical models, the Maximum-Entropy Seeded Aspect Sentiment model (ME-SAS) and the Seeded Aspect Sentiment model (SAS), were put forth in [8] to address the problem in different situations where the user provides multiple headwords for numerous aspect categories and the model simultaneously extracts and groups the aspect terms into categories. Despite having a close relationship, aspect term extraction (ATE) and aspect category categorization (ACC) are frequently the focus of different studies. Knowledge obtained from one learning exercise would be transferable to another. Researchers developed a multi-task learning model (MTNA) based on neural networks to address both objectives [9]. A sequential labeling job known as ATE requires that word tokens about the assigned qualities be identified using a specific tagging method, such as IOB (Inside, Outside, Beginning). A subset of predetermined aspect labels must be used to label the phrase in the supervised classification task known as ACC. They incorporate CNN for ACC and Bi-LSTM for ATE inside a multi-task framework. To develop mixed classifiers for aspect extraction, the upgraded CNN is combined with an SVM that employs cutting-edge, manually generated features [10].

The study's use of cutting-edge natural language processing (NLP) technologies enhanced the science of aspect term extraction greatly. The Bidirectional Encoder Representations from Transformers (BERT) paradigm and Named Entity Recognition (NER) techniques have emerged as significant players in information extraction from textual data. While NER techniques concentrate on finding named entities within the text, BERT, noted for its contextualized word embeddings, has revolutionized several NLP tasks by capturing intricate semantic relationships.

This study explores a novel method for improving the precision and accuracy of aspect phrase extraction from customer ratings. This is achieved by looking into incorporating Sparse Gated Recurrent Units (GRUs) into the BERT-NER architecture. Sparse GRUs are a feasible choice for improving the extraction process because they have a reputation for forecasting long-term dependencies while effectively managing computer resources and combining the skills of BERT's contextual understanding and NER's entity recognition expertise with the special abilities of the sparse GRU results in a comprehensive and effective technique. Showcase the significant improvements in aspect term extraction efficiency and accuracy through arduous testing and careful analysis. Concerning sentiment analysis, customer feedback research, and product creation, our findings show that Sparse GRUs can be successfully incorporated into the BERT-NER framework and highlight this novel method's potential.

The contribution of this paper:

- To develop a synergistic combination of cutting-edge techniques, including Sparse GRUs, BERT, and NER, to address the challenging problem of aspect term extraction.
- By incorporating sparsity constraints into the GRU architecture, sparse GRUs offer the advantages of efficient computation and increased generalization.
- This novel approach attempts to capture both local and contextual information, thereby increasing the precision and relevance of derived aspect phrases.
- By combining Sparse GRUs with BERT and NER, we have significantly improved the accuracy of aspect word extraction compared to prior techniques.
- The experimental findings show that Sparse GRUs may enhance the recognition of aspect phrases in customer reviews, leading to more precise sentiment analysis and a better comprehension of customer reviews.

The following is the order in which the paper is organized: The existing works on aspect term extraction are illustrated in Section 2 of this article. Section 3 provides a detailed explanation of the suggested methodology. The experimental setup and datasets used for the studies are shown in Section 4 and the parameter values. Lastly, Section 5 presents the study's conclusion and insights into its future directions.

2. Literature Survey

Hu et al. [11] developed a framework based on span boundaries to identify aspect concepts for the ATE challenge. Their principle backbone network was BERT. Our model is distinct from the span boundaries-based technique and the Seq2Seq4 ATE. The suggested IANN encoder shows an improved ability to collect essential insights for solving the ATE problem by utilizing the MCRN approach to combine information from nearby words towards a single word inside a phrase. Our suggested model's contextualized embedding layer might be able to capture words' changing meanings. Additionally, the proposed IANN uses the creative AO tags as its tagging scheme.

Liu et al. [12] presented movie reviews and a framework to mine aspect-based opinions for decision assistance. Based on sentence features, we separate movie reviews from Douban, China's most famous movie community, into short and long reviews. First, we create methods to extract global and local fine-grained characteristics from short and long reviews. Second, a lexical updating method is suggested to recognize opinion terms on various topics. The new approach used in this study looks into several elements and views of a film through rigorous inquiry, diverging from most research initiatives that primarily concentrate on the polarity of sentiment. Experimental results using Douban data confirm the approaches' effectiveness and accuracy.

Liu et al. [13] present AO tags, a less difficult tagging method consisting of "Aspect" and "Outside" labels. The contextualized embedding layer's goal is to record words' changing definitions. This layer creates context-sensitive word embeddings by utilizing BERT. The MCRN model, meanwhile, is calculated to take in particularly illuminating data. The MCRN excels in identifying extensive connections and sequencing patterns beyond recognizing a phrase's nearby components and contextual cues. The MCRN also gains new features that provide deeper insights because of its multi-layer architecture. Extensive testing on three well-known datasets is used to assess the efficacy and usability of the proposed Integrated Artificial Neural Network (IANN).

Zhao et al. [14] used graph-based techniques to provide Sentiment Dependencies Using Graph Convolutional Networks (SDGCN). Each node in the graph structure receives an updated feature representation vector via GCN that includes pertinent information about its nearby nodes. Edges represent the sentiment-dependent relationship between two nodes, and aspects are handled as nodes. The model uses the graph to learn how the sentiments of the qualities relate to one another. Increasing the depth of Graph Convolutional Networks (GCNs) does not result in additional improvements because of excessive smoothing. However, a breakthrough was made when Hou et al. [15] created the Selective Attention-based GCN block (SA-GCN), which allowed researchers to overcome this restriction. This breakthrough made it possible to pinpoint the most significant contextual words. This useful information was then effortlessly incorporated into the aspect word representation.

When used for ASBA jobs, RNN and CNN perform less than ideal. They need to accurately capture the complex relationships between words and the overall textual framework in a given manuscript. Chen et al. [16] suggested GCNSA, a novel neural network strategy, to overcome these difficulties. By treating text as a graph, this method can identify some aspects inside specific areas of the graph structure. They put forth an enhanced structural attention model that built a full-text hidden state using a convolutional operation on the text graph and an LSTM to collect specific information.

Al-Ghuribi et al. [17] proposed an overall review sentiment score by assessing the weight and rating of each retrieved review element. An improved TF-IDF weighting method and a lexicon relevant to the topic are used to determine the importance of each extracted aspect and its related rating. The discovered elements' efficiency is next evaluated compared to recognized and removed characteristics listed in previous publications. General and domain-specific lexicons were used for evaluation. The suggested technique with the domain-specific lexicon outperformed Amazon and Yelp datasets in F-measure and accuracy.

Luo et al.'s [18] main goal is to isolate a product's core components while reducing superfluous semantic content. They start by identifying and compiling a list of distinctive characteristics. They then assess how much these attributes' semantic similarity stacks up. The final output is a presentation of the product's most salient qualities. An analysis of their approach's performance compared to six different frameworks shows its superiority.

Wang et al. [19] proposed aspect amount and aspect boundary errors in extracted aspect terms, and these errors had a substantial impact on the model's performance on the ATE test. Developed an aspect term error repair post-processing technique comprised of an aspect boundary adjusting module and an aspect number determining module. The aspect amount determination module can change the amount of extracted aspect words to match the actual aspect terms. To address the issue of the extracted aspect term's boundary not matching the boundary of the ground-truth aspect term, we use the aspect boundary modifying module to change the extracted aspect term's border to meet the ground-truth aspect term. Using four SemEval datasets, we display that our model outperforms current SOTA models and that the post-processing technique can correct errors in extracted aspect terms.

Luo et al. [20] created a one-of-a-kind bidirectional dependency tree network to extract information about dependency structure from phrases. The core idea is explicitly combining top-down and bottom-up propagation representations within a designated dependency syntax tree. The difficulty of aspect term extraction is then successfully met by combining these embedded illustrations with Bi-LSTM and CRF components to create a comprehensive system capable of training sequential and tree-structured features. The novel model exhibits higher performance across four benchmark SemEval datasets when associated with the current state-of-the-art baseline models.

Michelle et al. [21] presented a method for automatically creating aspect and opinion lists. We expanded the word list's coverage as well. As a result, the subsequent word list is known as a domain-specific lexicon. This study used the word embedding method to build 1000-word domain-specific lexicons. The information for the domain-specific lexicon was gathered from review websites using a specialized web crawler. The revised Sequential Covering approach made use of the obtained domain-specific lexicons. The F1 scores from the test results showed that this strategy performed better than the conventional Sequential Covering technique.

The study by Abdelminaam et al. [22] examined the opinions of Twitter users who spoke Arabic. They used three different datasets of Arabic tweets to accomplish this, grouping user sentiments into positive, negative, and neutral clusters. The study used various classification algorithms, including Naive Bayes, K-Nearest Neighbor, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, and Deep Learning approaches like LSTM, GRU, and ASTM.

Sharbatian et al. [23] suggested an aspect recognition approach for opinion mining. Giving to the consequences, the presented model outperforms previously proposed methods such as SVM and NB. Furthermore, deep learning models provide conditions that prevent the need for a single language. The consequences reveal that CNN and LSTM are quite good at detecting product features at the document level. Because of the fast technique of labeling and annotating, there was a large number of unlabeled data in this study.

Dutta et al.[24] developed the Element-Based Opinion Analysis technique to forecast sentiment polarity within multiple Twitter attributes during the lockdown and subsequent relaxation stages. This inquiry aims to ascertain Indians' opinions about the government's efforts to impose a lockdown to prevent the coronavirus's spread. They proposed a deep learning model to categorize the Twitter dataset. It succeeded with noteworthy accuracy of 82.35% for the Lockdown dataset and 83.33% for the Unlock dataset.

Chouikhi et al. [25] employ a Transfer Learning (TL) technique to tackle Arabic ATE and APD. This strategy bridges the gap by utilizing the strengths of language models that have already undergone training. The Arabic BERT model serves as the foundation for the models presented. Furthermore, the paper investigates several BERT implementations. The study conducts experimental assessments using an ABSA dataset as the standard. The experiment consequences demonstration that our models outperform both the baseline model and previous techniques.

3. Proposed System

This section discusses a synergistic combination of cutting-edge techniques, including Sparse GRUs, BERT, and NER, to address the challenging problem of aspect term extraction. By incorporating sparsity constraints into the GRU architecture, sparse GRUs offer the advantages of efficient computation and increased generalization. This novel approach attempts to capture both local and contextual information, thereby increasing the precision and relevance of derived aspect phrases.

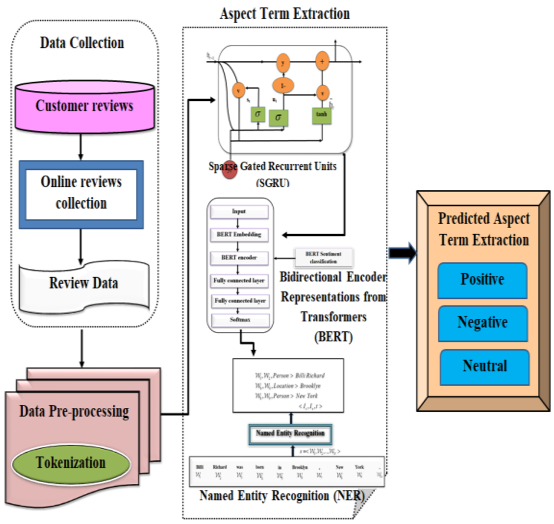

Fig 1: Proposed method of SGRU-BERT

Fig.1 depicts the SGRU-BERT method's block diagram. The free source R application was used to examine the collected data. Topic modeling processes are broken into three sections. Collect the data that will be used to apply topic modeling first. Sparse GRUs, BERT, and NER, are used to tackle the difficult challenge of aspect term extraction. Sparse GRUs provide the benefits of efficient computing and better generalization by introducing sparsity constraints into the GRU architecture. This unique approach aims to collect local and contextual information, boosting the precision and relevance of derived aspect phrases. We can significantly increase the accuracy of aspect term extraction over previous methods by integrating Sparse GRUs with BERT and NER. The Aspect Word Extraction (ATE) label establishes the relationship between the token and an aspect word. In contrast, the Aspect Polarity Classification (APC) label reveals the sentiment polarity of a specific aspect term.

3.1. Datasets

Two unique datasets were used in the trials. Online customer reviews from numerous online marketplaces were gathered in the first category. These reviews were then compared to reference models using a specially designed web crawler [26]. The second category, on the other hand, included benchmark datasets widely used by academics to compare their performance to reference models. The following datasets were utilized in the experiments:

3.1.1. Collected Datasets

- Amazon: The dataset includes 28,125 online customer reviews from 45 laptop models.

- BestBuy: The dataset comprises 33 different laptop models and 47,905 consumer reviews from online sources.
- Dell: The collection consists of 44,175 online reviews from customers who purchased 19 other laptop models.
- Lenovo: The dataset includes evaluations of 21 laptop models from 85,711 online consumer reviews.

3.1.2. Benchmark Datasets

In aspect-based sentiment analysis, the fourth task is our main focus, concentrating on the first subtask known as aspect term extraction. This work was carried out in Dublin, Ireland, as a part of the SemEval-14 competition, a platform for assessing semantic knowledge. These statistics include consumer reviews for two product categories: laptops and restaurants. The main facts about the dataset are briefly summarized in Table 1.

Table 1: The main characteristics of the SemEval-14 datasets

Dataset		Laptops			Restaurants		
		Trial	Test	Total	Trial	Test	Total
Sentence Count		3048	800	3848	3044	800	3844
Aspect Term	With Repetition	2373	654	3027	3699	1134	4833
	Unique Items	1048	419	1315	1259	555	1657

3.2. Data pre-processing

3.2.1. Tokenization

The process of dividing a continuous stream of text-based information into separate entities known as tokens is referred to as tokenization. Characters, numbers, words, and even punctuation are all included in these tokens. To increase the accuracy of the subsequent analysis, some symbols and punctuation are usually deleted from the input data during tokenization. These tokens are then used as the primary input for all stages of building a classification model [27–28].

3.3. BERT

The BERT architecture is useful for tasks like natural language processing, sentiment analysis, and question-answering. The BERT model [29], developed in 2018, significantly changed from earlier machine learning methods. BERT, which Google researchers created, is distinguished by its exceptional bidirectional capacity, which allows it to understand textual contexts in both left-to-right and right-to-left orientations. Pre-training on a sizable amount of unlabelled text from many sources, including Wikipedia and a library of literature, makes this accomplishment possible. Consider the word "matches" in the phrases "The ring matches the bracelet perfectly" and "Today's matches are both interesting" as examples. Traditional Context techniques frequently produce a single-word representation for each lexicon entry without considering the context. However, BERT [30] acts in an incredibly bidirectional way by considering the immediate prior and subsequent situations. There are two main ways that BERT can be used. The initial method entails masking a subset of the input text's words. The BERT model then processes this disguised input and forecasts the hidden words. The model uses the visible words before and after the masked term to predict the concealed phrase. Its goal is to ascertain if phrase Y follows sentence X or if the next few words are chosen randomly.

The BERT Model completes two unique jobs in two stages: first, during pre-training, it becomes familiar with the input textual data and its contextual nuances; second, during the fine-tuning phase, it takes in the input and precisely selects the proper replies. To obtain the most recent results, fine-tune the pre-trained BERT model by adding one layer. The BERT Model has two subcategories: BERTBASE and BERTLARGE. Both cases employ four feed-forward layers.

BERTBASE:

Blocks of the BERTBASE transformer are layered. =12.

Measurements of the output = 768.

Multiple-Headed Attention Count = 12.

Parameters total = 110 M. BERTLARGE.

The transformer blocks' total number of layers = 24.

Dimensions of the output = 1024.

Multiple-Headed Attention Count = 16.

Total Parameters = The Hugging Face Transformers library's uncased BERT Base model categorizes text reviews. The Amazon dataset is then uploaded to Google Colab and set up with the necessary GPU runtime parameters.

3.3.1. BERT model structure.

To comply with the specifications of the BERT model, the order of the input tokens should be changed. According to the requirements of the model, each input sequence must begin with the [CLS] (Classification token) and end with the [SEP] (Separation token). The entire input text sequence is categorized using the inserting associated with the [CLS] token in the output.

Preprocessing: Finishing the basic actions listed in the next section [31] is crucial before entering the review text data into the model.

a) Canonicalization: Initialization ignores all numerals, punctuation, distinguishing special characters, and lowercase letters are applied to any uppercase letters.

b) Tokenization: The BERT Base Uncased Transformer Model's processing vocabulary comprises 30,522 terms. On the other hand, tokenization entails splitting input review text data into predetermined tokens. A structure derived from the specified lexicon is used to accomplish this. When a term is absent in this glossary, the WordPiece tokenization method is used. By removing prefixes and suffixes, the WordPiece Tokenizer divides a word's whole content into subwords or root words. For instance, "looking" becomes "look ##ing."

- ip_ids' – A list of integers represents the input sentences. The special tokens to be inserted are represented by the integer values 101 and 102, while the padding token is represented by the value 0.
- 'attention_mask' - representing by 1s and 0s. The mask values with '0' can be disregarded, and the model is instructed to focus on masked tokens with the value '10'.

BERT Layer: Dropout regularization and a softmax classifier layer are combined to create a simple pre-trained BERT Base Uncased Model. The four essential techniques include the required preprocessing and embedding phases. Dropout regularization is used in the third step with a probability factor 0.1 to prevent overfitting. A classification approach called the softmax activation function is used. Softmax converts input scores into resulting probabilities that match output class labels via an exponential process. This guarantees that the final result keeps a cumulative probability of 1. The equation below can be used to represent the Softmax function mathematically.

$$P(\vec{y_i}) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \quad (1)$$

With "n" stands for the number of output classes, $P(\vec{y_i})$ for the softmax function, and e^{x_i} for in this case, the standard exponential function is applied to both the input and output vectors.

In an example utilizing the category cross-entropy loss function, the following sums are summed to determine the loss:

$$Loss = -\sum_{i=1}^{n} y_i . \log x_i \quad (2)$$

In this case, y_i stands for the value associated with the goal, x_i stands for the model's output with a scalar value, and 'n' signifies the entire amount of scalar values in the model's output. Figure 2 displays the visual representation of the BERT Model's procedure.

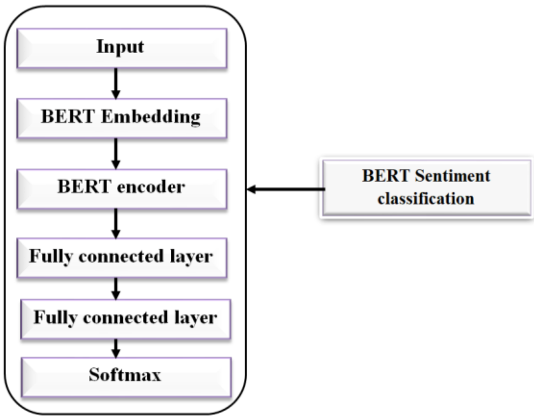

Fig 2: The flow chart of the BERT model for sentiment classification

The Adam optimizer is used to optimize the BERT-Base-Uncased model. Throughout five epochs, various learning rates are tested to confirm its efficacy. For loss reduction, the categorical cross-entropy loss function is used. Throughout the training phase, the parameters of the BERT-Base-Uncased model are adjusted. By tuning hyperparameters, the ultimate objective is to determine the best model, emphasizing minimizing validation loss. Changing the hyperparameters has a substantial impression on the model's performance. Regarding results, the learning rate value 2e-05 outperforms other learning rate values.

3.4. Named Entity Recognition (NER)

A named entity is a word or phrase that explicitly separates a single element from a collection of related components [32]. Organizations, people, and places are examples of named entities in typical contexts. A technique for locating and classifying these named entities inside text and assigning them to specified entity categories is known as named entity recognition (NER).

Named Entity Recognition (NER) is anticipated to provide a $< I_s, I_e, t >$ collection of tuples when given a set of tokens. A named entity listed in the provided sequence, s, is represented by each tuple in this collection.

$$s =< \mathrm{W}_1, \mathrm{W}_2, .., \mathrm{W}_N > \quad (3)$$

$I_s \in [1, N]$ and $I_e \in [1, N]$ stand for the beginning and end indices of a named entity mention, respectively, whereas "t" here stands for the entity type taken from a specified category list. Figure 3 illustrates a NER system identifying three named items in a text. The initial MUC-6 task for NER was identifying names of people, companies, places, and expressions of time, money, and percentage in text. It should be highlighted that only a few coarse entity types—one for each named object—are the focus of the study. Coarse-grained NER is the name given to this kind of NER task. Recent fine-grained NER research has focused on a considerably broader range of entity types, with various fine-grained types mentioned.

NER is a crucial pre-processing step for various downstream applications, including information retrieval, question answering, machine translation, and others. We emphasize the importance of Named Entity Recognition (NER) in supporting various applications by using semantic search as an example. Semantic search refers to a wide range of techniques that enable search engines to decipher the ideas, intentions, and meanings buried inside user searches. Seventy-one percent of search requests include at least one recognizable entity. We can give users more relevant search results by improving our understanding of user intents through the capacity to recognize named things inside these search queries. We introduced entity-based language models for including named entities in the search process. These models consider groups of terms that have been classified as entities in documents and queries, in addition to single phrases. To improve the user experience, research has also looked at the use of named entities, such as entity cards, query completion, and suggestions.

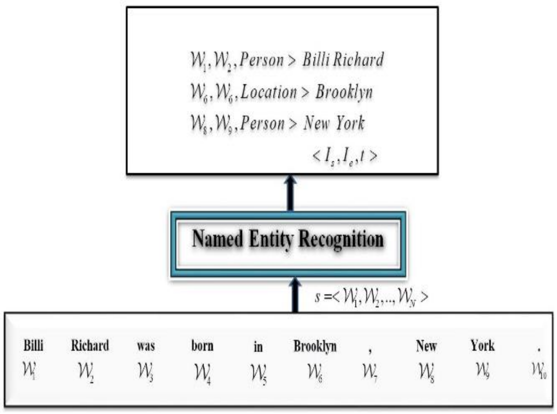

Fig 3: An illustration of the named entity recognition task

3.5. Sparse Gated Recurrent Units (GRUs)

To successfully capture the temporal dependencies and contextual information included in customer evaluations, we use Sparse Gated Recurrent Units (GRUs), a popular recurrent neural network design variation. Introducing sparsity into the GRU architecture can increase the model's capacity to focus on critical features while reducing the noise inherent in lengthy and frequently verbose evaluations.

Our study's primary assistances can be summarized as surveys:

- Sparse Gated Recurrent Units (GRUs): To improve the aspect term extraction process, we propose the concept of sparsity inside the GRU architecture. This advancement enables the model to selectively attend to crucial temporal relationships while filtering out less relevant information, resulting in higher extraction accuracy.
- Contextual Understanding: Our proposed solution uses the sequential nature of customer assessments to capture the inherent context and semantics that support precise aspect phrase detection. The surrounding words and phrases help the system understand the reviewer's sentiment and viewpoint.
- Domain Adaptability: We assess the efficacy of our strategy across various domains and product categories. Our method may be applied to multiple review datasets due to its robustness and adaptability, eliminating the need for intensive domain-specific customization.

This publication gives a thorough overview of relevant research in sentiment analysis, aspect term extraction, and neural network designs. The architecture and implementation of the Sparse GRU model for aspect term extraction are then covered in great detail. We thoroughly assess our method's performance on standard benchmark datasets and compare our findings to those of cutting-edge approaches to confirm its efficacy. The results show that Sparse GRUs can improve sentiment analysis jobs by significantly enhancing aspect phrase extraction accuracy. We address a critical component of sentiment analysis by introducing a new aspect phrase extraction technique using Sparse Gated Recurrent Units. By effectively capturing contextual information and emphasizing essential features, our method advances the accuracy and reliability of sentiment analysis, ultimately benefiting businesses and consumers alike in making informed decisions based on customer reviews.

An LSTM version called Sparse Gated Recurrent Units (SGRUs) was introduced in [33]. Since SGRUs are built with more robust memory, they are perfect for storing long-term dependencies between sequence segments. An individual update gate comprises an input gate and a forget gate that, among other things, interact with the hidden and cell states. The subsequent model, which is less sophisticated than normal LSTM models, has develop increasingly popular in various fields. In SGRU, there are reset (s_t) and update (u_t) gates. The second element determines how much the previous hidden state h_{t-1} should be reduced if it is deemed irrelevant in calculating the current state. The first aspect dictates how much the previous hidden state h_{t-1} should impact the following state h_t. The input y_t and the prior state h_t determine the output h_t of an SGRU, which is calculated as follows:

$$s_t = \sigma(V_s.[h_{t-1}; y_t] + b_s \qquad (4)$$

$$u_t = \sigma(V_u.[h_{t-1}; y_t] + b_u \qquad (5)$$

$$\tilde{h}_t = \tanh(V_h.[(r_t \odot ; h_{t-1}); y_t] + b_h) \qquad (6)$$

$$h_t = (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \qquad (7)$$

The reset and update gates are denoted by the s_t and u_t, respectively. The symbol represents vector concatenation, sigmoid functions, and element-wise multiplications. In this case, "dh" stands for the dimension of the concealed state, and $V_s, V_u, V_h \in \mathbb{R}^{d_h \times (d + d_h)}$ stands for the appropriate values for the reset and update gates.

Because the SGRU model is more straightforward than the LSTM model, numerous academics have begun to build ABSA problems using this model. Most ABSA tasks accept the target or aspect is known before defining sentiment polarity. Figure 4 depicts the SGRU architecture.

4. Result and Discussion

This section defines the Experimental setup and consequences in terms of baseline metrics for comparison and evaluation. Finally, the experimental findings are presented and discussed.

4.1. Experimental results

In the first section, we compare the effectiveness of the suggested method to that of recognized models. The performance of the suggested technique across all datasets is then presented. The next subsections elaborate on the specifics of each comparison.

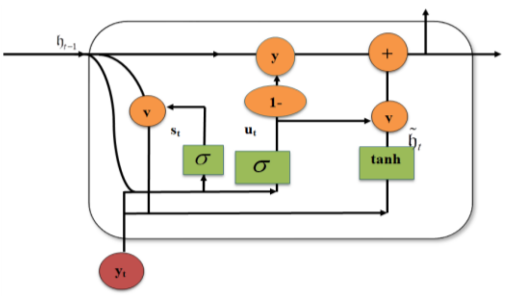

Fig 4: Architecture of SGRU

4.1.1. Baselines

DE-CNN [34]: In the context of the ATE challenge, the paper "DE-CNN: Dual embeddings CNN " investigated several embeddings. Both the general domain embedding and the specialized domain embedding are considered inside the DE-CNN architecture.

AGCN [35]: To improve the target nodes' ability to represent data, a combined aggregated graph convolutional network (AGCN) is used. Two aggregator functions are shown to take advantage of the information about the intrinsic node features. These routines routinely update the node representations based on information about their immediate local environment.

MGAN [36]: To understand the association among aspects and contexts, a multi-granularity attention mechanism is used with a bidirectional LSTM neural network to gather contextual information.

Co-LSTM [37]: "Collaborative Long Short-Term Memory" is the expanded form of "Co-LSTM." This neural network architecture integrates numerous Long Short-Term Memory networks to interpret and make sense of sequential input.

GA-GAT [38]: The target-dependent graph attention network (TD-GAT) is an example of a neural network like this that is used in machine learning and natural language processing. Its main goal is to manipulate data represented as graphs, where connections join nodes (or objects) together.

4.2. Performance metrics

In a multi classification situation, aspect sentiment classification is the goal. This study evaluates the effectiveness of triple classification using aspect-based sentiment classification (APC). The accuracy and the macro average F1 score are used to evaluate the effectiveness of the suggested model. The larger average F1 values and increased accuracy are indicators of improved model performance. The most accessible performance metric for the general public is accuracy, which shows the percentage of accurately predicted outcomes among all forecasts. Accuracy gauges the predictions made by the entire sample. The calculation for accuracy is as follows:

$$Accuracy = \frac{C_{pre}}{T_{pre}} \qquad (8)$$

Where C_{pre} represents the number of correct predictions and T_{pre} represents the total number of forecasts.

First, macro averaging is used to determine the overall mean of the arithmetic values for all classes by tabulating the index values for each class. In this procedure, we first compute the F1 score for each class, and then we average these values to get the overall average F1 score. The F1 score, which includes the weighted averages of both recall and precision, is the primary statistic. In multiclass classification, the class being considered affects precision, recall, and F1 measures. Class A is mainly the positive class in our circumstance, while all other classes are negative. The classification has three categories: positive, neutral, and negative. For each of these categories, precision, recall, and F1 score are calculated using the subsequent method:

$$Precision_A = \frac{C_A}{T_{A1}} \tag{9}$$

$$Recall_A = \frac{C_A}{T_{A2}} \tag{10}$$

$$F1_A = \frac{2 \times Precision_A \times Recall_A}{Precision_A + Recall_A} \tag{11}$$

Where A is "Positive, Negative, Neutral," CA is the amount of Class A examples that were accurately predicted, TA1 is the entire amount of samples projected to be Class A, and TA2 is the whole amount of Class A examples.

It is determined how to calculate the macro average F1 score:

$$Macro - F1 = \frac{1}{n}\sum_{i=1}^{n} F1_i \tag{12}$$

4.2.1 Comparative Analysis of first group of dataset

A custom web crawler collects consumer reviews from online marketplaces in the first category.

Accuracy Analysis

Table 2: Accuracy Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	93.598	96.766	95.876	94.887	97.765	99.728.
BestBuy	92.454	95.343	94.344	93.276	96.356	98.672
Dell	92.518	95.675	94.565	93.543	96.459	98.524
Lenovo	93.632	96.843	95.998	94.954	97.954	99.952

In Fig. 5 and Tab. 2, the accuracy of the SGRU-BERT strategy is contrasted to other existing techniques. This graph shows how the deep learning method increases efficiency while maintaining accuracy. The accuracy values of DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models are 93.598%, 96.766%, 95.876%, 94.887%, and 97.765% respectively, while the SGRU-BERT model has an accuracy of 99.728.% for Amazon dataset. The SGRU-BERT model has an accuracy of 98.672% under BestBuy dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have accuracy values of 92.454%, 95.343%, 94.344%, 93.276%, and 96.356%, respectively. Similarly, the SGRU-BERT model has an accuracy of 98.524% under the Dell dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have accuracy values of 92.518%, 95.675%, 94.565%, 93.543%, and 96.459%, respectively. Likewise, the SGRU-BERT model has an accuracy of 99.952% under the Lenovo dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have accuracy values of 93.632%, 96.843%, 95.998%, 94.954%, and 97.954%, correspondingly.

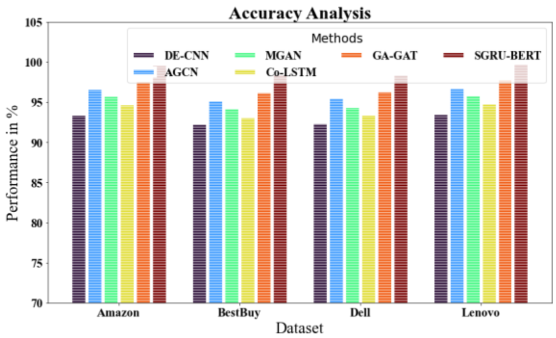

Fig 5: Accuracy Analysis for SGRU-BERT method with Existing systems

4.2.2. Precision Analysis

Table 3: Precision Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	93.827	92.926	91.627	94.928	95.127	96.826
BestBuy	92.311	91.726	90.425	93.726	94.526	95.524
Dell	92.416	91.614	90.627	93.826	94.627	95.624
Lenovo	93.624	92.727	91.728	94.826	95.225	96.926

In Fig.6 and Tab.3, the SGRU-BERT strategy's precision is contrasted to other methods currently in use. This graph shows how the deep learning method increases efficiency while maintaining precision. The precision values of DE-CNN, AGCN, MGAN, Co-LSTM and GA-GAT models are 93.827%, 92.926%, 91.627%, 94.928%, and 95.127% respectively, while the SGRU-BERT model has a precision of 96.826.% for Amazon dataset. The SGRU-

BERT model has a precision of 95.524% under BestBuy dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have precision values of 92.311%, 91.726%, 90.425%, 93.726%, and 94.526%, respectively. Similarly, the SGRU-BERT model has a precision of 95.624% under the Dell dataset; compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have precision values of 92.416%, 91.614%, 90.627%, 93.826%, and 94.627%, respectively. Likewise, the SGRU-BERT model has a precision of 96.926% under the Lenovo dataset; compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have precision values of 93.624%, 92.727%, 91.728%, 94.826%, and 95.225%, respectively.

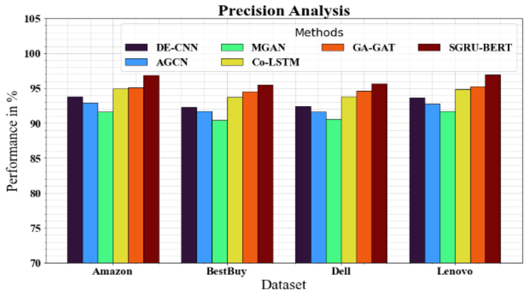

Fig 6: Precision Analysis for SGRU-BERT method with Existing systems

4.2.3. Recall Analysis

Table 4: Recall Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	89.826	92.826	90.627	91.938	93.027	94.928
BestBuy	88.422	91.726	89.344	90.627	92.526	93.627
Dell	88.726	91.624	89.542	90.726	92.726	93.728
Lenovo	89.927	92.926	90.826	91.826	93.125	94.827

In Fig. 7 and Tab. 4, a comparison of the recall of the SGRU-BERT strategy to several existing techniques is presented. The graph shows how the deep learning method increases efficiency while maintaining recall. The recall values of DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models are 89.826%, 92.826%, 90.627%, 91.938%, and 93.027% respectively, while the SGRU-BERT model has a recall of 94.928.% for Amazon dataset. The SGRU-BERT model has a recall of 93.627% under BestBuy dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have recall values of 88.422%, 91.726%, 89.344%, 90.627%, and 92.526%, respectively. Similarly, the SGRU-BERT model has a recall of 93.728% under the Dell dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have recall values of 88.726%, 91.624%, 89.542%, 90.726%, and 92.726%, respectively. Likewise, the SGRU-BERT model has a recall of 94.827% under the Lenovo dataset; compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have recall values of 89.927%, 92.926%, 90.826%, 91.826%, and 93.125%, correspondingly.

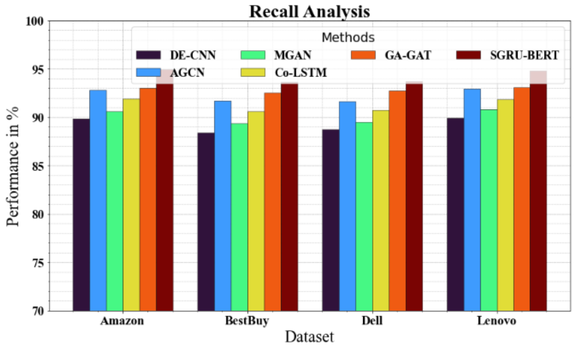

Fig 7: Recall Analysis for SGRU-BERT method with Existing systems

4.2.4. F1-Score Analysis

Table 5: F1-Score Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	85.892	88.937	87.637	86.827	89.026	90.827
BestBuy	84.356	87.526	86.536	85.435	88.324	89.627
Dell	84.536	87.674	86.435	85.536	88.452	89.725
Lenovo	85.723	88.726	87.738	86.927	89.382	90.924

Fig. 8 and Tab. 5 compare the f1-score of the SGRU-BERT strategy to several existing techniques. The graph demonstrates how deep learning increases efficiency while maintaining the f1-score. The f1-score values of DE-CNN, AGCN, MGAN, Co-LSTM and GA-GAT models are 85.892%, 88.937%, 87.637%, 86.827%, and 89.026% respectively, while the SGRU-BERT model has an f1-score of 90.827.% for Amazon dataset. The SGRU-BERT model has an f1-score of 89.627% under BestBuy dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM and GA-GAT models, which have f1-score values of 84.356%, 87.526%, 86.536%, 85.435%, and 88.324%, respectively. Similarly, the SGRU-BERT model has an f1-score of 89.725% under the Dell dataset; compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have f1-score values of 84.536%, 87.674%, 86.435%, 85.536%, and 88.452%, respectively. Likewise, the SGRU-BERT model has an f1-score of 90.924% under the Lenovo dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have f1-score values of 85.723%, 88.726%, 87.738%, 86.927%, and 89.382%, respectively.

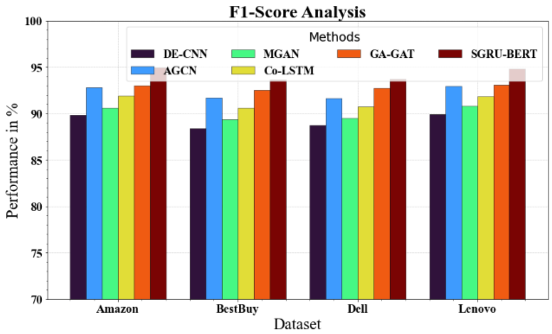

Fig 8: F1-Score Analysis for SGRU-BERT method with Existing systems

4.2.5. Macro average F1 Analysis

Table 6: Macro average F1 Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	93.565	96.839	95.023	94.837	97.124	98.822
BestBuy	92.675	95.526	94.547	93.526	96.425	97.835
Dell	92.985	95.627	94.637	93.625	96.526	97.924
Lenovo	93.675	96.926	95.213	94.927	97.298	98.725

In Fig.9 and Tab.6, the macro average f1 of the SGRU-BERT approach is compared to several existing strategies. The graph shows how the deep learning strategy is more effective while maintaining the macro average f1. The macro average f1 values of DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models are 93.565%, 96.839%, 95.023%, 94.837%, and 97.124% respectively, while the SGRU-BERT model has a macro average f1 of 98.822% for Amazon dataset. The SGRU-BERT model has a macro average f1 of 97.835% under BestBuy dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have macro average f1 values of 92.675%, 95.526%, 94.547%, 93.526%, and 96.425%, respectively. Similarly, the SGRU-BERT model has a macro average f1 of 97.924% under the Dell dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have macro average f1 values of 92.985%, 95.627%, 94.637%, 93.625%, and 96.526%, respectively. Likewise, the SGRU-BERT model has a macro average f1 of 98.725% under the Lenovo dataset, compared to the DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT models, which have macro average f1 values of 93.675%, 96.926%, 95.213%, 94.927%, and 97.298%, respectively.

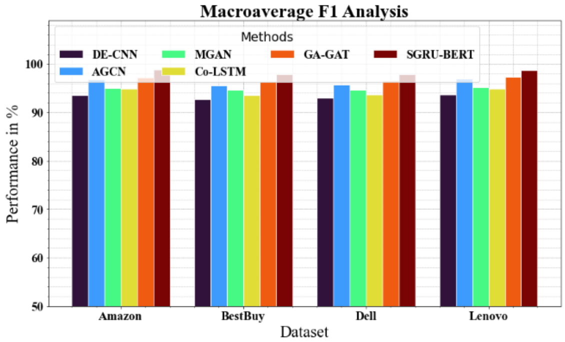

Fig 9 Macro average F1 Analysis for SGRU-BERT method with Existing systems

4.2.6. Execution time Analysis

Table 7: Execution time Analysis for SGRU-BERT method with Existing systems

Dataset	DE-CNN	AGCN	MGAN	Co-LSTM	GA-GAT	SGRU-BERT
Amazon	0.061	0.032	0.044	0.051	0.021	0.012
BestBuy	0.066	0.038	0.047	0.058	0.029	0.018
Dell	0.068	0.039	0.049	0.059	0.026	0.017
Lenovo	0.063	0.033	0.041	0.053	0.023	0.013

In Fig.10 and Tab.7, the execution time of the proposed SGRU-BERT methodology is compared to that of existing techniques. The data clearly shows that the SGRU-BERT technique outperformed all other strategies. The suggested SGRU-BERT approach, for example, took only 0.012ms to execute the Amazon dataset. In contrast, other current methods such as DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT have taken 0.061ms, 0.032ms, 0.044ms, 0.051ms, and 0.021ms, respectively as their execution time. However, the suggested SGRU-BERT approach takes 0.018ms to execute the BestBuy dataset, while existing techniques like DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT have taken 0.066ms, 0.038ms, 0.047ms, 0.058ms, and 0.029ms, respectively to execute. Likewise, the suggested SGRU-BERT approach takes 0.017ms to execute the Dell dataset, while existing techniques like DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT have taken 0.068ms, 0.039ms, 0.049ms, 0.059ms, and 0.026ms, respectively to execute. Similarly, the suggested SGRU-BERT approach takes 0.013ms to execute the Lenovo dataset, while existing techniques like DE-CNN, AGCN, MGAN, Co-LSTM, and GA-GAT have taken 0.063ms, 0.033ms, 0.041ms, 0.053ms, and 0.023ms, respectively as their execution time.

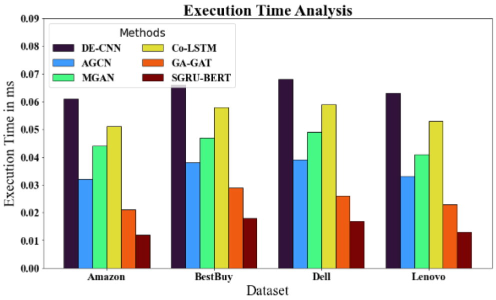

Fig 10: Execution time Analysis for SGRU-BERT method with Existing systems

4.2.7. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) Curve has frequently been used to evaluate the effectiveness of classifiers. Equation (13) represents the mathematical expression for this performance metric.

$$ROC = \frac{CP(i / Positive)}{CP(i / Negative)} \qquad (13)$$

Conditional probabilities are expressed using the notation $CP(i/l)$, where l stands for a class label. The ROC curve shows how categorization results evolve from extremely positive to severely negative.

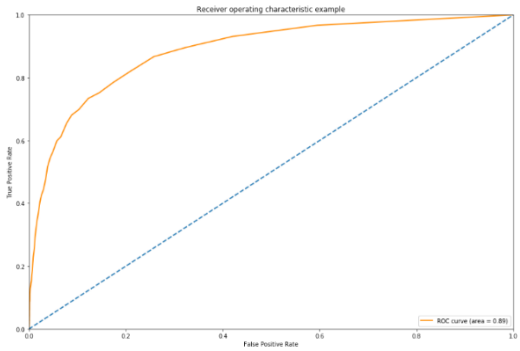

Fig 11: ROC Curve Analysis

4.2.8 Comparative Analysis of second group of dataset

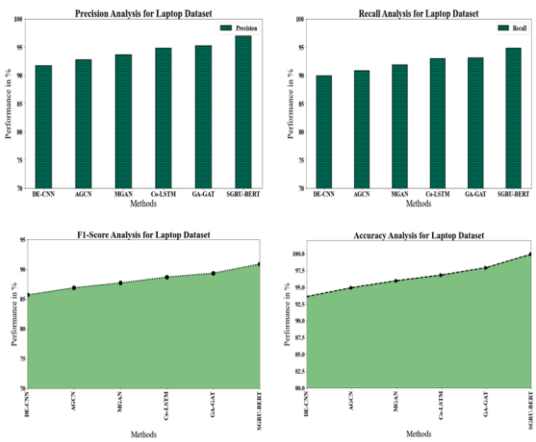

Fig 12: Comparative Analysis for Laptop dataset

A comparison of the proposed SGRU-BERT methodology with currently used methods is shown in Fig.12 and Tab.7 using laptop dataset. Regarding accuracy, precision, recall, and f1-score, the deep learning technique performs better, as seen in the graph. The SGRU-BERT model has a precision value of 96.926%, recall value of 94.827%, the f1-score value of 90.924%, and an accuracy of 99.952%. The proposed method performs better in every aspect.

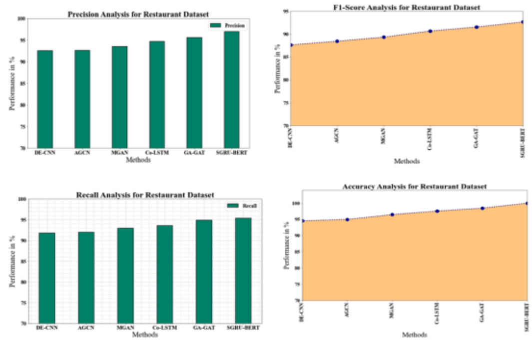

Fig 13: Comparative Analysis for Restaurant dataset

A comparison of the proposed SGRU-BERT methodology with currently used methods is shown in Fig.13 and Tab.8 using Restaurant dataset. Regarding accuracy, precision, recall, and f1-score, the deep learning technique performs better, as seen in the graph. The suggested SGRU-BERT has a precision of 96.965%, the recall value of 95.377%, the f1-score value of 92.675%, and an accuracy of 99.999%, The proposed method performs better in every aspect.

5. Conclusion

In conclusion, our research ventured into aspect term extraction from customer evaluations, utilizing innovative methodologies to optimize the process. The accuracy and effectiveness of aspect phrase extraction are improved using Sparse Gated Recurrent Units (GRUs) in the context of BERT and Named Entity Recognition. The experiment’s findings demonstrate that the suggested model successfully captures the complex relationships between words and aspects, leading to a more precise identification of aspect terms in the context of customer reviews. Contextual embedding from BERT offer a strong foundation for comprehending text semantics and enable the model to recognize minute variations in aspect expressions. The model's overall performance improved with the addition of Sparse GRUs, improving its ability to understand long-term dependencies. This study offers a significant advancement in sentiment and user comment analysis in natural language processing. The statistics show that all proposed models function admirably, but one model stands out regarding precision, recall, f1-score, and accuracy. We evaluated our top-performing model amalgamation using the SemEval-14 datasets and associated the consequences to industry best practices to demonstrate its trustworthiness. The data strongly suggest that our approach outperforms other methodologies regarding F-Score results. To improve the efficacy of aspect term extraction across various domains and languages, future research could focus on enhancing the architectural design, experimenting with other combinations of deep learning components, and identifying new linguistic features.

References

[1] Zhao, G., Luo, Y., Chen, Q., & Qian, X. (2023). Aspect-based sentiment analysis via multitask learning for online reviews. *Knowledge-Based Systems*, *264*, 110326.

[2] Venugopalan, M., & Gupta, D. (2022). A reinforced active learning approach for optimal sampling in aspect term extraction for sentiment analysis. *Expert Systems with Applications*, *209*, 118228.

[3] Venugopalan, M., & Gupta, D. (2022). An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis. *Knowledge-based systems*, *246*, 108668.

[4] Y. Song, J. Wang, T. Jiang, Z. Liu, Y. Rao, Attentional encoder network for targeted sentiment classification, arXiv preprint arXiv:1902.09314 (2019). https://arxiv.org/abs/1902.09314.

[5] B. Zeng, H. Yang, R. Xu, W. Zhou, X. Han, Lcf: A local context focus mechanism for aspect-based sentiment classification, Applied Sciences 9 (16) (2019) 3389, https://doi.org/10.3390/app9163389.

[6] W. Xue, T. Li, Aspect based sentiment analysis with gated convolutional networks, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (vol. 1: Long Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 2514–2523. doi:10.18653/v1/P18- 1234. https://www.aclweb.org/anthology/P18-1234.

[7] Yang, H., Zeng, B., Yang, J., Song, Y., & Xu, R. (2021). A multi-task learning model for chinese-oriented aspect polarity classification and aspect term extraction. *Neurocomputing*, *419*, 344-356.

[8] Mukherjee, A., and B. Liu. 2012. Aspect extraction through semi-supervised modelling. In Proceedings of the 50 Annual Meeting of the Association for Computational Linguistics July, 2012 (Association for Computational Linguistics) Jeju Island, Korea, 339–348.

[9] Xue, W. 2017. MTNA: A neural multi-task model for aspect category classification and aspect term extraction on restaurant reviews. Proceedings of the The 8th International Joint Conference on Natural Language Processing, Taipei, Taiwan, 151–156.

[10] Senarath, Y., N. Jihan, and S. Ranathunga. 2019. A hybrid approach for aspect extraction from customer reviews International Journal on Advances in ICT for Emerging Regions (ICTer) . 12 (September):1–8.

[11] Hu M, Peng Y, Huang Z et al (2019) Open-domain targeted sentiment analysis via span-based extraction and classification. In: Proceedings of the 57th annual meeting of the association for computational linguistics, pp 537–546

[12] Liu, Z., Sun, Q., Yang, Z., Jiang, K., & Yan, J. (2020, December). Fined-grained aspect extraction from online reviews for decision support. In *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)* (pp. 1536-1541). IEEE.

[13] Liu, N., & Shen, B. (2023). Aspect term extraction via information-augmented neural network. *Complex & Intelligent Systems*, *9*(1), 537-563.

[14] A. Zhou.Li, L., and Y. Liu. 2018. Hierarchical attention-based position-aware network for aspect-level sentiment analysis. Conference of Computational Natural Language Learning (CoNLL), Brussels, Belgium, Association for Computational Linguistics, 181–189.

[15] Hou, X., Huang, J., Wang, G., Qi, P., He, X. 2019. Selective attention-based graph convolutional networks for aspect-level sentiment classification Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15) June,2021 (Association for Computational Linguistics) 83–93 Mexico City, Mexico.

[16] Chen, J. 2019. Graph convolutional networks with structural attention model for aspect-based sentiment analysis. 2019 International Joint Conference on Neural Networks (IJCNN). Budapest, Hungaray: IEEE, July, 1–7.

[17] Al-Ghuribi, S. M., Noah, S. A. M., & Tiun, S. (2020). Unsupervised semantic approach of aspect-based sentiment analysis for large-scale user reviews. *IEEE Access*, *8*, 218592-218613.

[18] Luo Z, Huang S, Zhu KQ. Knowledge empowered prominent aspect extraction from product reviews. Inform Process Manag. 2019; 56(3): 408-423.

[19] Wang, R., Liu, C., Zhao, R., Yang, Z., Zhang, P., & Wu, D. (2022). Post-processing method with aspect term error correction for enhancing aspect term extraction. *Applied Intelligence*, *52*(14), 15751-15763.

[20] Luo, Huaishao, Tianrui Li, Bing Liu, Bin Wang, and Herwig Unger. "Improving aspect term extraction with bidirectional dependency tree representation." *IEEE/ACM Transactions on Audio,*

Speech, and Language Processing 27, no. 7 (2019): 1201-1212.

[21] Michelle, P., Ruskanda, F. Z., & Purwarianti, A. (2020, September). Development of domain-specific lexicon for aspect-based sentiment analysis. In *2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA)* (pp. 1-6). IEEE.

[22] Abdelminaam DS, Neggaz N, Gomaa IAE, Ismail FH, Elsawy AA. ArabicDialects: an efficient framework for Arabic dialects opinion mining on twitter using optimized deep. Neural Netw. 2021; 9: 97079-97099.

[23] Sharbatian, K., & Moattar, M. H. (2023). Deep aspect extraction and classification for opinion mining in e-commerce applications using convolutional neural network feature extraction followed by long short term memory attention model. *Applied AI Letters*, e86.

[24] Dutta, R., Das, N., Majumder, M., & Jana, B. (2023). Aspect based sentiment analysis using multi-criteria decision-making and deep learning under COVID-19 pandemic in India. *CAAI Transactions on Intelligence Technology*, *8*(1), 219-234.

[25] Chouikhi, Hasna, Mohammed Alsuhaibani, and Fethi Jarray. "BERT-Based Joint Model for Aspect Term Extraction and Aspect Polarity Detection in Arabic Text." *Electronics* 12, no. 3 (2023): 515.

[26] Ali, Noaman M., Abdullah Alshahrani, Ahmed M. Alghamdi, and Boris Novikov. "Extracting Prominent Aspects of Online Customer Reviews: A Data-Driven Approach to Big Data Analytics." *Electronics* 11, no. 13 (2022): 2042.

[27] X. Feng, Y. Zeng, Neural collaborative embedding from reviews for recommendation, IEEE Access 7 (2019) 103263–103274.

[28] Venugopalan, Manju, and Deepa Gupta. "A reinforced active learning approach for optimal sampling in aspect term extraction for sentiment analysis." *Expert Systems with Applications* 209 (2022): 118228.

[29] N. Shrestha, F. Nasoz, Deep learning sentiment analysis of Amazon.Com Reviews and Ratings, Int. J. Soft Comput. Artif. Intel. Appl. 8 (No. 1) (2019), 01–15.

[30] Sridevi M, Murugeswari, Bheema Lingaiah, "Deep Learning Approaches for Cyberbullying Detection and Classification on Social Media", Computational Intelligence and Neuroscience, vol. 2022, 2022. https://doi.org/10.1155/2022/2163458

[31] S. Yu, J. Su, D. Luo, Improving BERT-based text classification with auxiliary sentence and domain knowledge, IEEE Access 7 (2019) 176600–176612.

[32] Sridevi M, Saravanan Chandrasekaran, Murugeswari, Bheema Lingaiah, "Deep Learning Approaches for Cyberbullying Detection and Classification on Social Media", Computational Intelligence and Neuroscience, vol. 2022, 2022. https://doi.org/10.1155/2022/2163458

[33] Li, J., Sun, A., Han, J., & Li, C. (2020). A survey on deep learning for named entity recognition. IEEE Transactions on Knowledge and Data Engineering, 34(1), 50-70.

[34] Trisna, Komang Wahyu, and Huang Jin Jie. "Deep learning approach for aspect-based sentiment classification: a comparative review." Applied Artificial Intelligence 36, no. 1 (2022): 2014186.

[35] Abbas Mardani, Sudhanshu Maurya & N. Arulkumar (2023) Eagle Strategy Arithmetic Optimisation Algorithm with Optimal Deep Convolutional Forest Based FinTech Application for Hyper-automation, Enterprise Information Systems, DOI: 10.1080/17517575.2023.2188123

[36] Gangathimmappa, M, Subramani, N, Sambath, V, Ramanujam, RAM, Sammeta, M. Deep learning enabled cross-lingual search with metaheuristic web-based query optimization model for multi-document summarization. Concurrency Computation Practice Experience. 2022;e7476. doi:10.1002/cpe.7476

[37] Xu H, Liu B, Shu L et al (2018) Double embeddings and CNN based sequence labeling for aspect extraction. In: Proceedings of the 56th annual meeting of the association for computational linguistics (volume 2: short papers), pp 592–598

[38] M. Zhao, J. Yang, J. Zhang, S. Wang, Aggregated graph convolutional networks for aspect-based sentiment classification, Inform. Sci. 600 (2022) 73–93.

[39] F. Fan, Y. Feng, D. Zhao, Multi-grained attention network for aspect-level sentiment classification, in, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 3433–3442, https://doi.org/10.18653/v1/D18-1380

Table 8: Comparative Analysis of second group of dataset

Methods	Laptop				Restaurant			
	Precision	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**Accuracy**
DE-CNN	91.728	89.927	85.723	93.632	92.555	91.788	87.65	94.566
AGCN	92.727	90.826	86.927	94.954	92.587	91.987	88.455	94.987
MGAN	93.624	91.826	87.738	95.998	93.532	92.965	89.343	96.487
Co-LSTM	94.826	92.926	88.726	96.843	94.665	93.565	90.677	97.565
GA-GAT	95.225	93.125	89.382	97.954	95.564	94.877	91.556	98.455
SGRU-BERT	96.926	94.827	90.924	99.952	96.965	95.377	92.675	99.999