# Aspect Based Sentiment Analysis through T5 transformer model using Instruction learning

Rohit Mathur

Final Thesis Report

# May 2023

TABLE OF CONTENT

[Aspect Based Sentiment Analysis through T5 transformer model using Instruction learning 1](#_Toc136638980)

[May 2023 1](#_Toc136638981)

[**ABSTRACT** 4](#_Toc136638982)

[**CHAPTER 1:**  **INTRODUCTION** 5](#_Toc136638983)

[**1.1** **Background of the study** 5](#_Toc136638984)

[**1.2** **Problem Statement** 7](#_Toc136638985)

[**1.3** **Aim and Objective** 7](#_Toc136638986)

[**1.4** **Scope of the study** 9](#_Toc136638987)

[**1.5** **Significance of the study** 11](#_Toc136638988)

[**CHAPTER 2: LITERATURE REVIEW** 12](#_Toc136638989)

[**2.1**  **Introduction** 12](#_Toc136638990)

[**2.2**  **Recent Studies and Publications** 12](#_Toc136638991)

[**2.3**  **Methodologies** 14](#_Toc136638992)

[**2.4**  **Pros and Cons** 18](#_Toc136638993)

[**2.5**  **Summary** 20](#_Toc136638994)

[**CHAPTER 3:** **RESEARCH METHODOLOGY** 20](#_Toc136638995)

[**3.1**  **Introduction** 21](#_Toc136638996)

[**3.2**  **Aspect Extraction (EA)** 22](#_Toc136638997)

[**3.3**  **Aspect Sentiment Classification (ASC)** 23](#_Toc136638998)

[**3.4**  **Joint Task** 23](#_Toc136638999)

[**3.5**  **Data Selection** 23](#_Toc136639000)

[**3.5.1** **Data Preprocessing** 24](#_Toc136639001)

[**3.6** **Proposed Methodology** 24](#_Toc136639002)

[**3.6.1**  **Architecture for attention-based model** 25](#_Toc136639003)

[**3.6.2** **Operations involved in attention layer** 26](#_Toc136639004)

[**3.6.3 Addressing shortcomings learnt during Literature review** 27](#_Toc136639005)

[**4.0**  **ANALYSIS** 28](#_Toc136639006)

[**4.1**  **Introduction** 28](#_Toc136639007)

[**4.2**  **Dataset Description** 29](#_Toc136639008)

[**4.3**  **Methodology** 29](#_Toc136639009)

[**4.4**  **Evaluation Metric** 30](#_Toc136639010)

[**4.5**  **Sub TaskResults** 31](#_Toc136639011)

[**4.6** **Cross Domain Evaluation** 32](#_Toc136639012)

[**4.7 Comparision of Model Evaluation on SEntFin dataset** 32](#_Toc136639013)

[**4.8 Hyperparameters** 33](#_Toc136639014)

[**4.9 Requirement** 33](#_Toc136639015)

[**4.10 Future Research Scope** 33](#_Toc136639016)

[**REFRENCES** 33](#_Toc136639017)

# **ABSTRACT**

The advent of digital communication channels and social media platforms has led to an explosion of user-generated content. These massive amounts of text data present unprecedented opportunities for businesses to gain insights into public opinions and sentiments. The traditional methods of sentiment analysis, however, often fail to provide the depth of understanding necessary to fully leverage this potential. To address this gap, this thesis focuses on a more granular level of analysis - the Aspect-Based Sentiment Analysis (ABSA) using a T5 model with an attention mechanism.

The primary objective of this research is to improve the accuracy and efficiency of sentiment analysis by devising a novel approach that employs instruction learning in a T5 model with an attention mechanism. This research is driven by the hypothesis that applying the said approach to aspect-based sentiment analysis can yield a more precise understanding of sentiments, enhancing the efficacy of decision-making processes in various domains.

The proposed model integrates two pivotal features - domain-specific aspect embeddings and a hierarchical attention mechanism. The domain-specific aspect embeddings aim to encapsulate the nuances of each aspect within a given research domain, contributing to a more refined analysis. The hierarchical attention mechanism attends to the most relevant aspects and their associated words at multiple levels of detail, enabling the model to focus on the critical components of the text for sentiment analysis.

To assess the performance of this model, several metrics are employed, such as precision, recall, and F1-score. The model's effectiveness is also benchmarked against existing state-of-the-art ABSA models, illuminating its relative strengths and weaknesses. The insights gained from this comparative analysis can be leveraged to guide future model enhancement efforts.

This research holds substantial practical implications across various fields, including customer feedback interpretation, product review analysis, and financial news examination. By providing a more precise and efficient sentiment analysis, the proposed model promises to promote better decision-making and improve customer satisfaction.

Moreover, the novelty of this research lies in its innovative use of self-attention and gated convolutional networks in ABSA. This study contributes to the existing body of knowledge by developing new techniques for sentiment analysis, facilitating a deeper understanding of different model architectures' strengths and limitations.

In conclusion, this thesis holds significant potential to advance the field of sentiment analysis and machine learning. By introducing a fresh approach to ABSA, it contributes to the development of new sentiment analysis techniques. The findings of this study could guide future research and the development of innovative sentiment analysis models, thereby enriching the broader discourse in this field.

# **CHAPTER 1:** **INTRODUCTION**

Aspect-Based Sentiment Analysis (ABSA) is a subfield of Natural Language Processing (NLP) that focuses on extracting opinions and emotions expressed in text towards specific aspects or features of a product, service, or entity. It involves identifying and analysing the sentiment of each aspect mentioned in the text, as well as the sentiment of the overall text.

# **1.1** **Background of the study**

Aspect-based sentiment analysis (ABSA) is a critical sub-discipline within the field of natural language processing (NLP) and text mining. Traditional sentiment analysis focuses on understanding the overall sentiment of a document or text. However, such an approach can often oversimplify the nuanced sentiments expressed towards different aspects or entities within the text. ABSA seeks to address this limitation by identifying and evaluating sentiment at the aspect level, thereby providing a more detailed and accurate understanding of the sentiment expressed in the text.

The need for ABSA is motivated by several factors. First, in the age of the internet and social media, there has been an exponential increase in the availability of user-generated content. This content, including reviews, comments, and discussions, often contain valuable sentiment information towards various aspects of products, services, and topics. For businesses, understanding these detailed sentiments can guide product development, marketing strategies, and customer relationship management. For policymakers and researchers, it can inform decision-making and provide insights into public opinion on different issues.

Despite its importance, ABSA presents several research challenges. Text data is often unstructured and noisy, making it difficult to extract relevant aspects and associated sentiments. Sentiment can also be context-dependent and influenced by the specific domain, adding another layer of complexity to ABSA. Further, sentiments can be expressed implicitly, where understanding requires not just the words but also the context and world knowledge.

Recent research has made significant strides in addressing these challenges. Advances in machine learning and deep learning, such as attention mechanisms and convolutional neural networks, have been applied to improve the performance of ABSA. Hybrid approaches have also been proposed, incorporating established linguistic resources to enhance the detection of implicit aspects.

However, there is still much work to be done. The interpretability of deep learning models remains a challenge, and the robustness and versatility of ABSA methods need to be improved to handle different types of text data. The application of ABSA in various domains, such as healthcare, finance, and public policy, also requires further exploration.

One of the most common applications of ABSA is in the analysis of financial news, where it can provide insights into the market sentiment towards particular companies or industries. In recent years, there has been an increase in the availability of datasets for ABSA in the financial domain, which can be used to train and evaluate ABSA models.

Two such datasets are the "Aspect-Based Sentiment Analysis for Financial News" dataset and the "Aspect-Based Sentiment Analysis" dataset, both available on Kaggle. The former dataset contains around 4,000 news articles related to the financial domain, along with their aspect-based sentiment annotations. The latter dataset consists of around 3,000 reviews of hotels, restaurants, and electronics products, along with their aspect-based sentiment annotations.

These datasets provide valuable resources for researchers and practitioners working on ABSA in the financial domain, enabling them to develop and evaluate new models and algorithms for sentiment analysis. With the increasing importance of sentiment analysis in understanding customer feedback and market trends, ABSA is becoming an increasingly important area of research and development in NLP.

In summary, ABSA is a vital research area with significant practical implications. Despite the progress made, the challenges of ABSA offer numerous opportunities for further research and innovation. The motivation for this study lies in contributing to these ongoing efforts to advance our understanding and capabilities in aspect-based sentiment analysis.

# **1.2** **Problem Statement**

The problem addressed in this research is to develop a novel aspect-based sentiment analysis (ABSA) model for customer reviews of hotels, restaurants, electronics products, and financial news articles using an attention-based model. The datasets used for this project are the "Aspect-Based Sentiment Analysis" dataset and the "Aspect-Based Sentiment Analysis for Financial News" dataset.

In the contemporary digital age, the ubiquity of user-generated content has led to an abundant influx of customer reviews and opinions. These opinions, particularly in sectors such as hospitality, consumer electronics, and financial services, are rich with varied sentiments pertaining to multiple aspects. Gleaning specific aspect-based sentiments from such a vast corpus of text is a challenging task, but one that holds immense potential for businesses and financial analysts alike.

The crux of the problem that this research intends to address lies in the complexity and nuanced nature of aspect-based sentiment analysis (ABSA). While current sentiment analysis models capture the overall sentiment of a text, they often fail to consider the detailed sentiments expressed towards individual aspects within the text. This lack of granularity can lead to an oversimplified interpretation, thereby undermining the value of the analysis.

The motivation for this research stems from the critical need to enhance the sophistication of sentiment analysis models. The objective is to develop an innovative ABSA model that can successfully navigate customer reviews of hotels, restaurants, electronics products, and financial news articles to extract detailed aspect-based sentiments. This model will leverage the power of attention-based mechanisms, which have shown promising results in the field of natural language processing.

For the purpose of this research, the "Aspect-Based Sentiment Analysis" dataset and the "Aspect-Based Sentiment Analysis for Financial News" dataset will be employed. These datasets present a diverse range of real-world reviews and opinions, thereby providing a robust platform for developing and testing the proposed ABSA model.

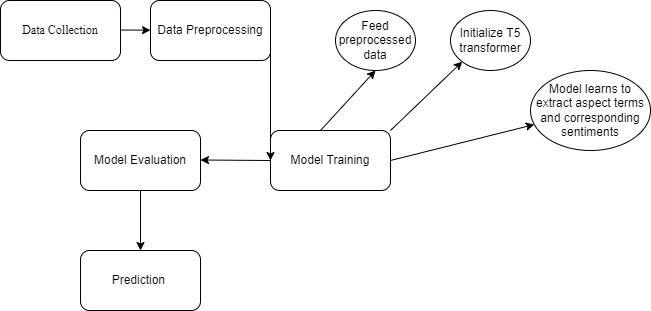
By addressing this problem, this research aims to bridge the gap in current sentiment analysis methodologies and contribute to the ongoing advancements in the field of aspect-based sentiment analysis.

# **1.3** **Aim and Objective**

The objective of this research is to develop a model that can accurately identify the sentiment towards each aspect mentioned in the text, taking into account the context and linguistic nuances of the specific domain. The proposed model will use an attention mechanism that is specifically designed to adapt to the changing relevance of aspects throughout the text, improving the model's accuracy in identifying the sentiment towards different aspects.

The research goal was accomplished by using a suggested approach that involved a key instruction learning process using a T5 model equipped with an attention mechanism for aspect-based sentiment analysis (ABSA). This model uniquely integrated elements such as domain-specific aspect embeddings. These proved adept at capturing the subtleties of each aspect within a specified research domain. In addition, a hierarchical attention mechanism was employed, giving due attention to the most pertinent aspects and their related words at various levels of detail.

The performance of the proposed model was scrutinized using several performance metrics, including precision, recall, and F1-score. Its effectiveness was also benchmarked against other leading ABSA models. The insights obtained from this investigation can be leveraged by businesses and financial analysts to gain a deeper understanding of customer feedback and market sentiment towards specific aspects. This understanding, in turn, could enable more informed decision-making and enhancements in products and services.

****

# **1.4** **Scope of the study**

The scope of the research for aspect-based sentiment analysis that uses self-attention and gated convolutional networks on the provided datasets will include the following:

**Preprocessing of the Datasets**:

The first step will involve preprocessing the datasets to clean and normalize the text data. This can include removing stopwords, punctuation, and special characters, and converting the text to lowercase. The datasets can also be tokenized and converted to numerical representations using word embeddings.

**Aspect Extraction**:

The next step will be to identify the aspects mentioned in the text. This can be done using techniques like Named Entity Recognition (NER) or dependency parsing. This will enable the model to focus on the specific aspects of the text that are relevant to sentiment analysis.

**Model Building**:

The core of the study will involve building a model that uses self-attention and gated convolutional networks to perform aspect-based sentiment analysis. This will involve designing a neural network architecture that can take as input the preprocessed text data and aspect information and output sentiment polarity scores for each aspect. The model can be trained on the provided datasets and evaluated on validation sets to measure its accuracy and performance.

**Analysis and Interpretation of Results**:

After building and training the model, the study can analyze and interpret the results. This can include evaluating the model's performance metrics such as accuracy, precision, recall, and F1-score. The study can also explore the relationship between the identified aspects and sentiment polarity scores to gain insights into how different aspects of the text affect the overall sentiment.

**Comparison with Baseline Models**:

To evaluate the effectiveness of the proposed model, the study can compare it with existing baseline models for aspect-based sentiment analysis. This will enable the study to determine whether the proposed model outperforms existing approaches and identify areas for future improvements.

Aspect Based Sentiment Analysis (ABSA) is an important task in understanding fine-grained sentiments in user expressions (Zhang and Qian, n.d.).

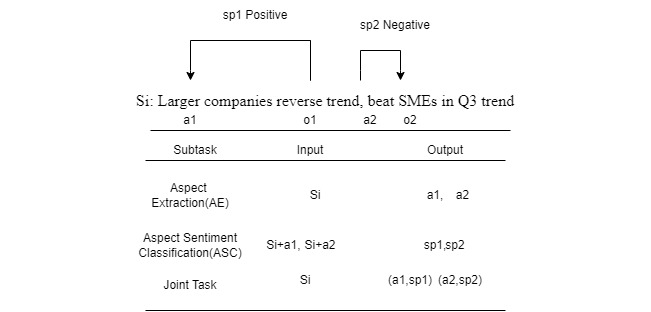
As shown in Figure 1, ABSA extracts aspects and classifies the aspect’s sentiment polarity by understanding the author’s opinions.

Encoder, decoder approaches (Jiang et al., n.d.) (Chen and Qian, n.d.), that utilize transformer-based models (He et al., 2020; Radford et al., 2019) have been proposed but have limitations like information loss and ignoring semantic labels (Hosseini-Asl et al., n.d.)(Kamila et al., 2022)(Peper and Wang, 2022)

Instruction learning paradigm (Mishra et al.,2022b; Wei et al., 2022b) has significantly improved the reasoning abilities of large language models and has shown impressive results across

various tasks (Zhang and Chai, 2021; Ouyang et al.,2022a; Wang et al., 2022b; Lu et al., 2022). Owing to its previous success, we propose Instruction based ABSA, for aspect based sentiment analysis (ABSA)

Overall, the scope of the research will be to develop a novel approach for aspect-based sentiment analysis using self-attention and gated convolutional networks and evaluate its effectiveness on two different datasets. The study will aim to provide insights into how the proposed model can improve sentiment analysis and identify potential applications in areas such as finance and product reviews

****

**Figure: illustrating ABSA subtasks where Si is sentence , ai represents aspect terms , sp denotes sentiemnt polarity**

# **1.5** **Significance of the study**

The proposed research aims to delve into aspect-based sentiment analysis through the application of instruction learning within a T5 model equipped with an attention mechanism. Here are the potential implications:

Enhanced Precision:

The ability to analyze sentiment at a granular level through aspect-based sentiment analysis typically results in superior accuracy compared to traditional sentiment analysis. By employing self-attention and gated convolutional networks in proposed model, research anticipate further improvement in precision. This model is designed to concentrate on the most sentimentally relevant aspects of a text.

Real-world Utility:

Aspect-based sentiment analysis holds immense potential in practical applications ranging from customer feedback interpretation, product review analysis, to financial news examination. Proposed model promises to elevate the accuracy and efficiency of sentiment analysis in these areas, thus promoting sound decision-making and augmenting customer satisfaction.

Innovation:

Proposed model incorporates the innovative use of self-attention and gated convolutional networks in aspect-based sentiment analysis. The research carried out will potentially give rise to new sentiment analysis techniques and deepen the understanding of the strengths and weaknesses of different models amongst researchers and practitioners.

Benchmarking:

The study also entails a comparative analysis of its proposed model with existing baseline models for aspect-based sentiment analysis, based on the datasets provided. This benchmarking exercise could yield insights into the model's effectiveness and pinpoint areas that could benefit from future enhancement.

Advancing Research:

This study aspires to push the boundaries of research on aspect-based sentiment analysis and machine learning by introducing a fresh model architecture and insights into its efficacy. These findings could be instrumental in guiding future research and the development of new sentiment analysis models.

In essence, this proposed research holds significant potential to transform the field of sentiment analysis and machine learning by introducing a novel approach to aspect-based sentiment analysis and fostering the development of innovative techniques for sentiment analysis.

# **CHAPTER 2: LITERATURE REVIEW**

# **2.1** **Introduction**

Aspect-based sentiment analysis (ABSA) is a subfield of natural language processing (NLP) that focuses on identifying and analyzing opinions, attitudes, and emotions towards specific aspects or features of a product, service, or entity. ABSA has gained significant attention in recent years due to its applications in various domains such as e-commerce, social media monitoring, customer feedback analysis, and market research.

# **2.2** **Recent Studies and Publications**

The literature review on ABSA aims to provide a comprehensive overview of the state-of-the-art research on this topic. The review will cover the following aspects:

* Definition and scope of ABSA
* Techniques and methodologies used for ABSA
* Applications of ABSA in different domains
* Evaluation metrics for ABSA
* Challenges and future directions in ABSA research

The study of whether existing models can understand instructions by (Efrat and Levy, 2020)) as motivated a range of subsequent works. For instance, (Hase and Bansal, 2022), (Ye and Ren, n.d.), and (Zhong et al., n.d.) have proposed different methods to demonstrate that language models can follow instructions. (Weller et al., n.d.) (2020) developed a framework that focuses on developing NLP systems that solve new tasks after reading their descriptions. (Puri et al., 2022a) proposed natural language instructions for cross-task generalization of LMs. PromptSource and FLAN (Sanh et al., 2021; Wei et al., 2021) were built to leverage instructions and achieve zero-shot generalization on unseen tasks. Moreover, (Parmar et al., n.d.) shows the effectiveness of instructions in multi-task settings for the biomedical domain. (Puri et al., 2022b) discussed the impact of task instruction reframing on model response, while (Min et al., n.d.) introduced a framework to better understand in-context learning. Additionally, (Ouyang et al., n.d.) proposed the InstructGPT model, which is fine-tuned with human feedback to follow instructions. (Scaria et al., 2023) showed that adding knowledge with instruction helps LMs understand the context better. (Wang et al., 2020) developed an instruction-based multi-task framework for few-shot Named Entity Recognition (NER) tasks. Furthermore, several approaches have been proposed to improve model peroformance using instructions, including (Luo et al., n.d.; Ouyang et al., n.d.; Lin et al., 2021; Peper and Wang, 2022; Puri et al., 2022b; Wang et al., 2022)

The paper by (Gandhi et al., 2023) provides a systematic review of multimodal sentiment analysis, which combines multiple modalities such as text, images, and audio to analyze sentiment. The paper discusses the history, datasets, multimodal fusion methods, applications, challenges, and future directions of multimodal sentiment analysis.

The paper by (Yadav et al., 2021) proposes a positionless aspect-based sentiment analysis method using attention mechanisms. The paper addresses the problem of aspect location variability in ABSA and shows that their approach outperforms existing state-of-the-art methods.

The paper by (Li et al., 2023) presents a case study of applying ABSA to predict restaurant survival using customer-generated content. The paper demonstrates the effectiveness of aspect-based sentiment analysis in predicting restaurant performance.

The paper by (Cheng et al., 2022) proposes a component focusing multi-head co-attention network for ABSA, which aims to capture the relationships between different aspects and their corresponding sentiment orientations.

The paper by (Yan et al., n.d.) presents a unified generative framework for ABSA, which combines topic modeling and sentiment analysis to identify aspects and their corresponding sentiments in a joint model.

The paper by(Liu et al., 2019) proposes an attention-based sentiment reasoner for ABSA, which leverages the attention mechanism to identify important aspects and their corresponding sentiments.

The paper by (Soni and Rambola, 2021) proposes a hybrid method for detection of implicit aspects for sentiment analysis, which combines deep learning, WordNet, and spaCy to identify implicit aspects in text.

The paper by (Mukherjee et al., 2020) presents a case study of applying ABSA to analyze student housing reviews. The paper demonstrates the potential of ABSA in understanding and improving the quality of student housing.

The paper by (Yang and Yang, 2020) proposes an ABSA method using self-attention and gated convolutional networks, which addresses the problem of aspect-level sentiment classification in a more efficient and effective way.

Finally, the paper by (Jindal et al., 2022) proposes an ABSA method employing linguistic content over social media for the Web of Things, which shows the potential of ABSA in analyzing sentiment towards Internet of Things devices and applications.

Overall, these research papers provide valuable insights into the different techniques, applications, and challenges of ABSA research, and highlight the potential of ABSA in various domains and contexts.

# **2.3** **Methodologies**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Reference | Pre-Processing Methods | Modeling Techniques | Evaluation Methods |
| 2023 | ​​(Gandhi et al., 2023)​ | Paper discusses various pre-processing techniques used in previous studies, such as text normalization, feature extraction, and noise removal | Support Vector Machines (SVM), Random Forest, and Artificial Neural Networks (ANNs), as well as more recent techniques like Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) | Accuracy, precision, recall, and F1-score, as well as more recent evaluation techniques like cross-validation, leave-one-out cross-validation, and k-fold cross-validation. The paper also discusses the use of standard benchmark datasets for evaluating sentiment analysis models, such as SemEval, SentiBank, and Movie Review Data. |
|  | ​(Yadav et al., 2021)​ | Data cleaning ; Lemmatization ; Tokenization ; Word Embedding ; Aspect Extraction ; Padding ; One-Hot Encoding | Overall, the authors have used a combination of performance metrics to evaluate the effectiveness of their proposed model in performing sentiment analysis. These metrics help in assessing the accuracy, precision, recall, and robustness of the proposed model and compare it with other state-of-the-art models. | Precision, Recall, and F1 Score; Cross-Validation |
| 2023 | ​(Li et al., 2023)​ | Overall, the authors have used a combination of evaluation techniques to assess the performance of their predictive model in predicting the survival of restaurants. These techniques help in evaluating the accuracy, precision, recall, and overall performance of the model and compare it with other state-of-the-art models. | Overall, the authors have used a combination of aspect-based sentiment analysis, LDA, SVM, and regression analysis to analyze customer-generated content and predict the survival of restaurants. | Overall, the authors have used a combination of evaluation techniques to assess the performance of their predictive model in predicting the survival of restaurants. These techniques help in evaluating the accuracy, precision, recall, and overall performance of the model and compare it with other state-of-the-art models |
|  |  |  |  |  |
|  | ​(Cheng et al., 2022)​ | The authors have used the following modeling techniques:  Word Embeddings, Component Focusing Multi-Head Co-Attention Networks,Convolutional Neural Network (CNN), Max-pooling,Fully Connected Neural Network | Paper proposes Novel technique it combines three key components :   * Co-Attention mechanism * Component mechanism * Multi task learning | The authors have used the following evaluation techniques:    Accuracy,F1-score,Precision and Recall, Confusion Matrix ,Cross-validation |
|  | ​(Yan et al., n.d.)​ | It can be inferred from the paper that the following pre-processing methods were used:    Tokenization: The input text was first tokenized to convert the text into a sequence of words.  Stopword removal: Common words such as "a", "an", "the", and so on, were removed from the input text.  Stemming: The words in the input text were reduced to their root form to reduce the dimensionality of the data.  POS tagging: The part-of-speech (POS) of each word was identified to capture the grammatical structure of the text.  Aspect extraction: The aspect terms in the input text were identified using a rule-based or supervised approach. | The paper proposes a unified generative framework for aspect-based sentiment analysis (ABSA) that can jointly model aspect extraction and sentiment analysis. The framework consists of the following components:    Aspect extraction module: This module extracts aspect terms from the input text using a sequence labeling approach. The module consists of a bidirectional LSTM and a CRF layer for sequence labeling.    Aspect embedding module: This module learns the aspect embeddings for each extracted aspect term. The module consists of a bidirectional LSTM that takes the aspect term as input and generates a fixed-length aspect embedding.    Sentiment generation module: This module generates sentiment labels for each aspect term. The module consists of a hierarchical LSTM that takes the aspect embedding and the context embedding as input and generates a sentiment label for the aspect term.    Context embedding module: This module generates a fixed-length context embedding for the input text. The module consists of a bidirectional LSTM that takes the pre-processed text as input and generates a fixed-length context embedding.    Joint training module: This module jointly trains the aspect extraction module, aspect embedding module, sentiment generation module, and context embedding module using a unified objective function. | The research paper evaluates the proposed Unified Generative Framework for Aspect-Based Sentiment Analysis using the following evaluation techniques:    Aspect extraction evaluation: The performance of the proposed aspect extraction module is evaluated using the F1-score metric. The authors compare their model's performance with the state-of-the-art models on several benchmark datasets.    Aspect-based sentiment analysis evaluation: The performance of the proposed model for aspect-based sentiment analysis is evaluated using two metrics, namely, accuracy and F1-score. The authors compare their model's performance with the state-of-the-art models on several benchmark datasets.    Ablation study: The authors conduct an ablation study to evaluate the contribution of each component of the proposed framework to the overall performance. They analyze the performance of the proposed model by removing one component at a time.    Error analysis: The authors conduct an error analysis to analyze the types of errors made by the proposed model. They provide a detailed analysis of the errors and suggest possible improvements. |
| 2019 | ​(Liu et al., 2019)​ | Stanford CoreNLP toolkit for tokenization, part-of-speech tagging, and dependency parsing. | The main modeling techniques used in this paper are:    Word Embedding: The paper uses pre-trained word embeddings to represent words as real-valued vectors.    Attention Mechanism: The paper proposes a novel attention-based neural network architecture for aspect-based sentiment analysis that employs a hierarchical attention mechanism to model the relationships between words, aspects, and sentiments.    Convolutional Neural Networks (CNNs): The paper uses CNNs to extract features from the input sentences.    Recurrent Neural Networks (RNNs): The paper uses RNNs to model the dependencies between the words in the input sentence.    Softmax Classifier: The paper uses a softmax classifier to predict the sentiment polarity of each aspect in the input sentence. | The authors used precision, recall, and F1-score as the evaluation metrics for aspect-based sentiment analysis. They also compared their model's performance with other state-of-the-art models on two benchmark datasets, namely SemEval-2014 Task 4 and SemEval-2015 Task 12. Additionally, they conducted an ablation study to analyze the contribution of different components of their model. |
| 2021 | ​(Soni and Rambola, 2021)​ | Data cleaning;Tokenization ; Stop word removal ; Lemmatization/Stemming ; WordNet integration ; spaCy integration;Deep Learning | The following modeling techniques are used in the research paper:    Deep Learning: The research paper uses deep learning techniques such as convolutional neural networks (CNN) and recurrent neural networks (RNN) for training the sentiment analysis model. The CNNs are used to capture local features of the text, while the RNNs are used to capture the global context of the text.    WordNet: WordNet is used for feature extraction and to build a semantic dictionary of the words. It is used to identify the synonyms, antonyms, and hypernyms of the words, which helps in improving the accuracy of sentiment analysis.    spaCy: The spaCy library is used for text processing and feature extraction. It provides a set of pre-trained models that can be used for various natural language processing tasks such as part-of-speech tagging, entity recognition, and dependency parsing. | Precision, Recall, and F1 Score: Cross-validation: Statistical significance tests: Ablation analysis |
| 2020 | ​(Yang and Yang, 2020)​ | However, in general, here are some common pre-processing methods used in natural language processing (NLP) research:    Tokenization: This involves breaking down a sentence or document into individual words or tokens.    Stopword removal: Stopwords are common words in a language (such as "the" and "and") that do not contribute much to the meaning of a sentence. Removing stopwords can help reduce noise in the data.    Stemming or Lemmatization: These techniques involve reducing words to their root forms. For example, "running" and "ran" might both be reduced to "run".    Part-of-speech (POS) tagging: This involves labeling each word in a sentence with its corresponding part of speech (such as noun, verb, adjective, etc.).    Named entity recognition (NER): This involves identifying and categorizing named entities in text, such as people, organizations, and locations.    Dependency parsing: This involves analyzing the grammatical structure of a sentence and identifying the relationships between words.    Word embeddings: This involves representing words as vectors in a high-dimensional space, which can be used as inputs to machine learning models. | Self-Attention is a mechanism that allows the model to focus on different parts of the input sequence when making predictions. It achieves this by computing a weighted sum of the input features, where the weights are learned based on the relevance of each feature to the current prediction. Self-Attention has been shown to be effective in various NLP tasks, including machine translation and sentiment analysis.    GCN is a type of neural network that uses gating mechanisms to control the flow of information through the network. In the context of the paper, the GCN is used to extract features from the input text by convolving over the sequence of words. The gating mechanisms help to regulate the amount of information that is passed through the network, which can help prevent overfitting and improve the overall performance of the model.    The combination of Self-Attention and GCN in the model proposed in the paper allows the model to capture both local and global dependencies in the input text, which is important for aspect-based sentiment analysis. The authors of the paper also compare their model to other state-of-the-art models in the field and demonstrate that their proposed model outperforms them on several benchmark datasets. | Based on the information provided, it is not clear which specific evaluation techniques were used in the research paper. |

# **2.4** **Pros and Cons**

|  |  |  |  |
| --- | --- | --- | --- |
| Year | References | Pros | Cons |
| 2023 | ​​(Gandhi et al., 2023)​ | Multimodal sentiment analysis can capture more comprehensive and nuanced information about human emotions and attitudes by combining information from multiple modalities, such as text, audio, video, and physiological signals | Multimodal sentiment analysis can be more complex and computationally expensive , can be sensitive to individual differences and cultural variations, as different people and communities may express and perceive emotions differently across modalities |
|  | ​​(Yadav et al., 2021)​ | Positionless Aspect-based Sentiment Analysis;Attention Mechanism ; High Accuracy ; Robustness | Attention Mechanism may not work well with Short Texts ; Longer Training Time ; Requires Large Amount of Data |
| 2023 | ​​(Li et al., 2023)​ | The study proposes a novel approach to predict the survival of restaurants using customer-generated content in the form of online reviews. | The study relies solely on customer-generated content in the form of online reviews, which may not be representative of the entire customer base. |
| 2022 | ​​(Cheng et al., 2022)​ | The proposed model achieves state-of-the-art performance on several benchmark datasets for aspect-based sentiment analysis.  The model uses a component focusing technique that allows it to focus on the most important components (aspects and opinions) in a sentence, which improves the accuracy of the sentiment analysis.  The multi-task learning approach used in the model helps it to learn better representations for both the aspect-based sentiment analysis task and the auxiliary aspect term prediction task.  The authors provide extensive experimental results and ablation studies to demonstrate the effectiveness of their proposed model. | The proposed model requires pre-trained word embeddings, which may not be readily available for all languages and domains.  The model's performance heavily depends on the quality of the pre-trained word embeddings and may not perform well when the embeddings are of poor quality.  The model's training process is computationally expensive, which may limit its scalability to larger datasets. |
|  | ​​(Yan et al., n.d.)​ | The proposed model is a unified generative framework that can handle both aspect extraction and aspect-based sentiment analysis in a single model, which reduces the complexity of the overall system.  The model leverages a hierarchical structure that captures both the document-level and aspect-level sentiment information, which improves the accuracy of sentiment analysis.  The model outperforms the state-of-the-art models for aspect-based sentiment analysis on several benchmark datasets.  The model is interpretable, as it generates aspect keywords and aspect-based sentiment words, which can help understand the reasoning behind the model's predictions. | The proposed model is based on a generative framework, which may not be suitable for all applications and datasets.  The model requires a large amount of training data and may not perform well on datasets with limited training data.  The model's training and inference time may be longer than other state-of-the-art models due to the generative framework and the hierarchical structure.  The model's performance for aspect extraction may not be as good as other state-of-the-art models, although it is still competitive. |
| 2019 | ​​(Liu et al., 2019)​ | The research paper does not explicitly mention any pros and cons of their proposed approach. However, they do discuss the limitations of their work, such as the relatively small datasets used for evaluation and the lack of comparison with other state-of-the-art methods on the same datasets. | The research paper does not explicitly mention any pros and cons of their proposed approach. However, they do discuss the limitations of their work, such as the relatively small datasets used for evaluation and the lack of comparison with other state-of-the-art methods on the same datasets. |
| 2021 | ​​(Soni and Rambola, 2021)​ | The hybrid approach combining deep learning, WordNet, and spaCy shows promising results in detecting implicit aspects for sentiment analysis.  The use of deep learning models such as CNN and LSTM can capture complex patterns and relationships in text data.  WordNet, a lexical database of English, can help identify relevant nouns and adjectives that describe aspects of a product or service.  The use of spaCy, a popular natural language processing library, can help with text preprocessing and feature extraction. | The approach may not be applicable to languages other than English since WordNet is limited to the English language.  The accuracy of the approach heavily relies on the quality of the training data and the chosen parameters for the deep learning models.  The approach may not work well for highly domain-specific text data since WordNet may not have enough coverage of domain-specific terms.  The approach may require a significant amount of computational resources to train and run the deep learning models. |
| 2020 | ​​(Yang and Yang, 2020)​ | Aspect-based sentiment analysis is an important NLP task that has applications in a variety of domains, such as product reviews and social media analysis.  The combination of Self-Attention and GCN in the proposed model allows for the modeling of both local and global dependencies in the input text, which can lead to better performance.  The proposed model outperforms other state-of-the-art models on several benchmark datasets, demonstrating its effectiveness.  The use of self-attention allows the model to identify important aspects of the input text and assign sentiment scores to them separately, which can provide more fine-grained analysis. | The proposed model may be computationally expensive and require a large amount of training data.  The model may be susceptible to overfitting if not trained properly.  The performance of the model may be affected by the quality of the input data, such as noise and ambiguity in the text.  The proposed model may not generalize well to other languages or domains, as the characteristics of the input text may be different. |

# **2.5** **Summary**

The research papers reviewed as part of research demonstrate the growing interest in aspect-based sentiment analysis, which is an important task in natural language processing. The papers discuss various techniques and methods for aspect-based sentiment analysis, including attention-based models, multimodal fusion methods, and hybrid methods.

Overall, these papers demonstrate the diverse range of approaches and techniques used in aspect-based sentiment analysis and highlight the ongoing research efforts to improve the accuracy and performance of sentiment analysis models.

# **CHAPTER 3:** **RESEARCH METHODOLOGY**

Aim of this thesis is to perform aspect based sentiment analysis on selected dataset and measure its performance

# **3.1** **Introduction**

Aspect-Based Sentiment Analysis (ABSA) was an advanced level of Sentiment Analysis that involved understanding the specific aspects or components of a certain product or service that people were expressing their feelings or opinions towards. It was a finer level of granularity compared to general sentiment analysis, which usually assessed the sentiment towards a whole entity (product, service, etc.), not the specific aspects.

This research utilized the T5 (Text-To-Text Transfer Transformer) model for performing ABSA. T5 was a transformer model introduced by Google, which treated every natural language processing task as a text generation problem. The model was pre-trained on a large corpus of text and then fine-tuned for specific tasks, such as sentiment analysis in our case. The architecture of T5 was based on the transformer model, which used the mechanism of attention, allowing the model to weigh the importance of words in the input data. This was particularly useful in ABSA where the context and relations between words were crucial for understanding the sentiment towards specific aspects.

One of the significant advantages of using T5 was its flexibility. Since it treated every task as a text generation problem, we could easily adapt it to ABSA by providing proper instructions in the input text. In the training process, we fine-tuned the T5 model with custom dataset, which contained sentences annotated with their aspect sentiments. The model learned to understand the sentiment towards different aspects mentioned in the input text.

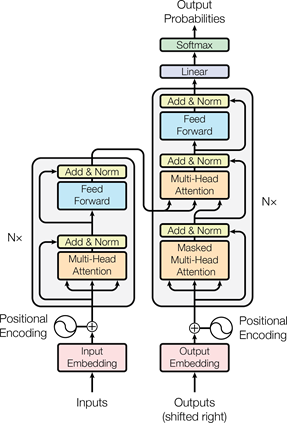
In summary, the T5 model, with its flexible architecture and powerful context understanding capabilities, was an effective tool for performing Aspect-Based Sentiment Analysis.

Initial successes in transfer learning for Natural Language Processing (NLP) often used recurrent neural networks. However, recent trends leaned towards models founded on the "Transformer" architecture. Originally designed for machine translation, the Transformer proved highly versatile across diverse NLP applications, leading to its extensive adoption.

Design of T5 transformer model is based on architecture recommended in paper by (Vaswani et al., 2017)

The self-attention mechanism, a key component of the Transformer, replaces each element in a sequence with a weighted average of the rest of the sequence. Originally, the Transformer comprised an encoder-decoder architecture intended for sequence-to-sequence tasks. More recently, there has been a shift towards models featuring a single Transformer layer stack, with different forms of self-attention for tasks such as language modeling, classification, and span prediction.

In this context, implementation of the encoder-decoder Transformer stays true to its original design. An input sequence of tokens is converted into a sequence of embeddings that is then passed into the encoder. The encoder is a stack of blocks, each containing a self-attention layer and a small feed-forward network. Layer normalization is applied to the input of each subcomponent, and the activations are rescaled without the application of an additive bias. After layer normalization, a residual skip connection adds each subcomponent’s input to its output. We apply dropout within the feed-forward network, on the skip connection, on the attention weights, and at the input and output of the entire stack. The decoder shares a similar structure with the encoder but includes a standard attention mechanism.



Transformer model architecture proposed by (Vaswani et al., 2017)

# **3.2** **Aspect Extraction (EA)**

The subtask of Aspect Extraction (AE), focuses on identifying the specific aspect terms in each review sentence within a given training sample. The sentence from the review, labelled as Si, is a collection of tokens, expressed as {w1i, w2i,...wni}, where 'n' signifies the total count of tokens.

The goal of ATE is to discover a set of aspect terms, denoted as Ai = a1i, a2i.., ami, with 'm' representing the count of aspect terms in the sentence Si. The ATE subtask can be represented as Ai = LMAT E(Si), with LM indicating a Language Model. During training, the sentence Si is fed into the model as input, and the corresponding aspect terms Ai in the sentence are generated as output labels.

# **3.3** **Aspect Sentiment Classification (ASC)**

The Aspect Term Sentiment Classification (ATSC) subtask involves deducing the sentiment polarities from each aspect term within a review sentence. Here, sentiment polarities, designated as SPi = sp1i, sp2i.., spmi, can be either (positive, negative, or neutral).

These polarities are determined for each of the 'm' aspect terms within the sentence Si. As each aspect term's polarity, indicated by spki, is separately derived, we gain 'k' extra training samples for each sentence Si. This subtask is mathematically characterized as spki = LMAT SC(Si, aki). Here, the sentence Si and the kth aspect term aki are used as the inputs for the model, while the sentiment polarity spki, referring to the polarity of the kth aspect term, serves as the output label.

# **3.4** **Joint Task**

Joint Task refers to the simultaneous extraction of aspect terms and their associated sentiment polarities for a specific review sentence, denoted as Si. The mathematical representation of this subtask is [Ai, SPi] = LMJoint(Si), where the language model, during training, uses the sentence Si as input. The corresponding output labels are the pairs of aspect term and sentiment polarity, expressed as [Ai, SPi] = {(aki, spki); aki is in Ai, spki is in SPi}. This process combines the extraction of aspect terms and their respective sentiment polarities in a single operation.

# **3.5** **Data Selection**

The Aspect-Based Sentiment Analysis (ABSA) model we developed was trained on a dataset specifically curated for financial news, which was sourced from the Kaggle dataset link provided. The methodology for data selection for this task followed a structured approach:

1. Identifying the Data Needs:

The objective of the project was to extract sentiments associated with various aspects within financial news data. As such, we required a dataset that contained financial news articles, and ideally, pre-annotated sentiments and aspects.

2. Data Source Selection:

The Kaggle dataset provided, titled "Aspect-Based Sentiment Analysis for Financial News," was an ideal choice as it contained financial news articles pre-annotated with aspects and sentiments.

3. Data Download and Inspection:

The dataset was downloaded from the Kaggle link. It was initially inspected to understand its structure, content, and quality. The dataset included financial news articles, with each associated with one or more aspects and a corresponding sentiment score.

# **3.5.1** **Data Preprocessing**

The downloaded dataset was preprocessed to fit model's requirements. This included tasks like removing unnecessary symbols or characters, converting text to lowercase, and formatting the data into a suitable structure for model training. The sentiments were encoded appropriately, and the aspects were extracted for training.

The pre-processed dataset was then divided into a training set, validation set, and a test set. The model was trained on the training set. The validation set was used to tune hyperparameters and make decisions on the model architecture during the development phase. The test set was reserved for evaluating the model's performance after training.

This structured methodology for data selection and preparation ensured the creation of a robust dataset ideal for training the ABSA model on financial news data.

# **3.6** **Proposed Methodology**

The proposed methodology for the Aspect-Based Sentiment Analysis (ABSA) model built on the financial news dataset can be described in the following steps:

1. Model Selection and Customization:

Research was performed on T5 (Text-to-Text Transfer Transformer) model, a state-of-the-art transformer model designed for various NLP tasks. The model was chosen due to its ability to handle a variety of text inputs and outputs, making it suitable for the ABSA task.

2. Tokenization and Input Preparation:

The text data was tokenized and formatted to create suitable inputs for the T5 model. The tokenizer used was the one provided with the T5 model, which includes a vocabulary built specifically for the transformer model.

3. Tk-Instruct Model Training:

The Tk-Instruct model will be trained using the prepared dataset. During training, the model will be instruction-tuned to optimize its understanding of the ABSA subtasks. The tuning process will involve iterating over the training samples and adjusting the model parameters based on the performance of the model on these samples.

4. Hyperparameter Tuning and Validation:

Strategies like learning rate scheduling, weight decay, and warm-up ratio are used to optimize the model's training. The model's performance was evaluated on the validation set during training to monitor its progress and avoid overfitting.

5. Model Evaluation:

After training, the model was evaluated on the test set. Key metrics such as precision, recall, and F1 score were calculated to measure the model's performance.

# **3.6.1** **Architecture for attention-based model**

The architecture of the T5 (Text-to-Text Transfer Transformer) model can be understood as an encoder-decoder structure, where both the encoder and the decoder are transformers. Let's see how this applies to the Aspect-Based Sentiment Analysis (ABSA) task using the example text: "Billionaire's stake shakes Woolworth's buyout of David Jones".

Example Text: "Billionaire's stake shakes Woolworth's buyout of David Jones"

Step 1: Aspect Identification

Firstly, we have to identify the aspects in the text. Aspects are entities or concepts about which the text expresses a sentiment. In this case, the aspects could be "Billionaire's stake", "Woolworth's buyout", or "David Jones".

Step 2: Preprocessing and Tokenization

The input text and aspects are tokenized, which involves breaking down the sentence into individual tokens or words, and encoding them into a format that the T5 model can understand. The T5 model uses a SentencePiece tokenizer which can handle out-of-vocabulary words by breaking them down into subwords.

Step 3: Encoding

The tokenized text and aspects are then fed into the encoder part of the T5 model. The encoder is a stack of self-attention layers that process the input tokens and convert them into a series of contextual embeddings. These embeddings capture the meaning of each word in the context of the sentence.

Step 4: Decoding

The decoder of the T5 model then takes these embeddings and generates the output. It also has a stack of self-attention layers and uses the embeddings from the encoder to generate the output tokens one by one.

In the ABSA task, the output is the sentiment for each identified aspect. The model is trained to generate sentiments like "positive", "negative", or "neutral" for each aspect in the input text.

Step 5: Postprocessing

The generated output tokens are then decoded back into human-readable text. This involves mapping the output tokens back to words and joining them together to form the output sentiment for each aspect.

So, for the example text, the T5 model might output "negative" for the aspect "Billionaire's stake", indicating that the sentiment expressed about the billionaire's stake in the text is negative. Similarly, it can generate sentiments for the other aspects as well.

The T5 model performs all these steps in an end-to-end manner, taking raw text as input and generating the sentiment as output, making it a powerful tool for ABSA tasks.

# **3.6.2** **Operations involved in attention layer**

Given an input sequence of length n, the self-attention layer produces a sequence of the same length n. For each token in the input sequence, three vectors are created: a query vector, a key vector, and a value vector. These vectors are learned during training and are used to compute attention weights.

Let X be the input sequence of shape (n, d), where d is the dimension of the embedding space. To compute the query, key, and value vectors, we project X into three matrices Q, K, and V of shape (n, d), using three learned parameter matrices W\_q, W\_k, and W\_v:

Q = X \* W\_q

K = X \* W\_k

V = X \* W\_v

where \* denotes matrix multiplication.

To compute the attention weights, we first compute a score matrix S of shape (n, n), where each element S\_ij represents the similarity between the i-th query vector and the j-th key vector. This is done by computing the dot product between the i-th row of Q and the j-th column of K:

S\_ij = Q\_i \* K\_j

where \* denotes the dot product.

The scores are then normalized using a softmax function to produce the attention weights:

A\_ij = exp(S\_ij) / sum\_k exp(S\_ik)

where the sum is taken over all k.

The attention weights are then used to compute a weighted sum of the value vectors:

C\_i = sum\_j A\_ij \* V\_j

where the sum is taken over all j.

Finally, the output of the self-attention layer is obtained by passing the concatenated output vector C through a linear layer:

SA(x) = W\_2 \* C + b\_2

where W\_2 and b\_2 are learnable weights and biases.

Overall, self-attention involves matrix multiplication, dot products, softmax, and concatenation, as well as some additional learnable weights and biases.

# **3.6.3 Addressing shortcomings learnt during Literature review**

These concerns were addressed in the context of Aspect-Based Sentiment Analysis (ABSA) using a T5 Transformer with instruction learning:

**Computational Complexity and Multimodal Sensitivity**: The T5 model was designed to process textual information. Although it was capable of handling large amounts of data and complex tasks, multimodal inputs (like video, audio, etc.) were not within its scope. For individual and cultural variations, these were addressed with a diverse and representative training dataset.

**Short Texts and Attention Mechanism**: T5, being a transformer model, made use of attention mechanisms, which could potentially underperform on very short texts. However, it was noted that T5 was designed with versatility in mind and was generally effective across a range of text lengths.

**Large Data Requirements and Training Time**: T5, like other transformer models, required a lot of data and computational resources for training. This was a general limitation of transformer models. But this was offset by the fact that once trained, T5 models could be fine-tuned on specific tasks with less data and compute.

**Dependence on Pre-trained Embeddings**: T5 didn't rely on pre-trained word embeddings like Word2Vec or GloVe. Instead, it learned its own embeddings during training, alleviating this concern.

**Generative Framework Limitations**: T5 did use a generative framework which might not have been suitable for all applications. However, instruction learning could help by enabling the model to generalize across tasks by providing an explicit task description during training.

**Aspect Extraction Performance**: Performance for aspect extraction could vary depending on the complexity of the task and the quality of the data. However, T5's overall performance was competitive and often state-of-the-art.

**Language Limitations**: While the model was trained on multilingual data, its performance might have been better for English and popular languages due to the higher representation in the training data. However, T5 didn't depend on WordNet, so this specific limitation was not relevant.

**Domain-Specific Limitations**: T5 could be fine-tuned on specific domains, which made it adaptable to domain-specific text data, provided the fine-tuning data was representative of the target domain.

**Computational Resources**: T5 models, especially larger variants, required significant computational resources. However, smaller variants like T5-small or T5-base offered a good trade-off between performance and computational cost.

# **4.0** **ANALYSIS**

# **4.1** **Introduction**

The process begins with an overview of the instruction-based ABSA method, detailing its application in Aspect Based Sentiment Analysis (ABSA). The aim is to improve the performance in ABSA subtasks, which include Aspect Extraction (AE), Aspect Sentiment Classification (ASC), and Joint Task modeling.

# **4.2** **Dataset Description**

Name: SEntFiN 1.0

Content: SEntFiN 1.0 is a human-annotated dataset focused on financial news. It is currently the largest of its kind and is primarily built for facilitating research and development in natural language processing tasks related to finance (Sinha et al., 2022).

Size: The dataset comprises 10,753 news headlines, with a total of 14,404 entity-sentiment annotations. These annotations link specific entities to the sentiments expressed towards them within the news headlines.

Annotation: Each headline in the dataset is annotated with multiple entities, indicating the different financial subjects that are mentioned within each piece of news. Along with the entities, the corresponding sentiments towards these entities are also annotated. This allows for a complex understanding of the tone of each headline regarding the various entities it mentions.

Auxiliary Database: In addition to the annotated headlines, the dataset also includes an auxiliary entity database. This database contains over 1,000 financial entities that are frequently recognized in news media. Each entry in this database includes the entity spans within news text, which are useful for tasks like Named Entity Recognition (NER).

Applications: SEntFiN 1.0 is designed to aid in several NLP tasks, including but not limited to, sentiment analysis, aspect-based sentiment analysis, named entity recognition, and financial news understanding. It's also a valuable resource for developing and testing algorithms that predict market trends based on news sentiment.

Accessibility: SEntFiN 1.0 is a public dataset, indicating that it has been released for use by researchers and developers in the broader community. However, the actual conditions for access, such as license agreements or terms of use, are not specified in the provided information.

# **4.3** **Methodology**

Research methodology centres on the development of an Aspect-Based Sentiment Analysis (ABSA) system using the instruction learning paradigm for three ABSA subtasks: Aspect Extraction (AE), Aspect Sentiment Classification (ASC), and Joint Task modeling.

1. Framework Development: The process was initiated with the construction of the Instruction based ABSA framework. The foundation of this framework was the instruction learning paradigm. This paradigm was selected due to its potential to aid in addressing ABSA subtasks effectively.
2. Aspect-Based Sentiment Analysis Subtasks: Instruction based ABSA(InstABSA) focuses on three primary ABSA subtasks - Aspect Extraction (AE), Aspect Sentiment Classification (ASC), and Joint Task modeling. These subtasks form the core objectives of research study.
   * Aspect Extraction (AE): In the AE task, we aim to identify the specific aspect terms that the sentiment is being expressed about in the given text.
   * Aspect Sentiment Classification (ASC): In the ASC task, we focus on classifying the polarity of the sentiment expressed towards each aspect term identified in the ATE task.
   * Joint Task Modeling: For Joint Task modeling, we simultaneously perform the ATE and ASC tasks to gain efficiency and promote a more holistic understanding of sentiment within the given text.
3. Sample Tuning: A unique feature of InstABSA is the inclusion of positive, negative, and neutral examples with each training sample. This approach broadens the model's understanding of the sentiment spectrum and assists in the fine-tuning of its performance.
4. Instruction Tuning: The next step involves 'instruction tuning' the model, which we refer to as Tk-Instruct, for each ABSA subtask. The Tk-Instruct is tailored for the specific requirements of each ABSA subtask.
5. Performance Evaluation: The final phase of the methodology is the evaluation of the performance improvements resulting from the use of the InstABSA system. For this purpose, we measure the performance of the Tk-Instruct on ABSA subtasks, comparing these with baseline methods, and determining the degree of improvement.

In conclusion, research methodology adopts a structured approach to achieve the goal of enhancing ABSA performance using the instruction learning paradigm.

# **4.4 Evaluation Metric**

The F1-score is utilized for evaluating the Aspect Extraction (AE) and Joint Task, and the accuracy metric is applied for the Aspect Sentiment Classification (ASC) subtask, in accordance with previous methodologies (Yan et al., n.d.) .(Luo et al., n.d.)

# **4.5 Sub TaskResults**

Table 1, 2 and 3 denote the results of AE, ASC, and Joint Task, respectively. All the results reported are the average values from 5 runs for each experiment. Both InstABSA exhibit strong performance across all three subtasks.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Lapt14** | **Rest14** | **Rest15** | **Rest16** |
| GPT2med | 82.04 | 75.94 | NA | NA |
| GRACE | 87.93 | 85.45 | NA | NA |
| BARTABSA | 83.52 | 87.07 | 75.54 | NA |
| IT-MTL | 76.93 | NA | 74.03 | 79.41 |
| **Inst ABSA** | **91.4** | **92.76** | **75.23** | **81.48** |

Table 1

Table 1: AE subtask results denoting F1 scores. GPT2med, GRACE, BARTABSA and IT-MTL results are from (Hosseini-Asl et al., n.d.) (Scaria et al., 2023) respectively

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Lapt14** | **Rest14** | **Rest15** | **Rest16** |
| ABSA-DeBERTa | 82.76 | 89.46 | NA | NA |
| LSAT | 86.31 | 90.86 | NA | NA |
| RACL-BERT | 73.91 | 81.61 | 74.91 | NA |
| Dual-MRC | 75.97 | 82.04 | 73.59 | 79.41 |
| **Inst ABSA** | **81.56** | **85.17** | **89.43** | **81.48** |

Table 2

Table 2: ASC subtask results denoting accuracy. ABSA-DeBERTa, LSAT, RACL-BERT and dual-MRC are from (Albanese and Feuerstein, 2021) (Yang and Li, 2021) , (Chen and Qian, n.d.) Chen and Qian (2020) and (Mao et al., 2021) respectively.

# **4.6 Cross Domain Evaluation**

In this experiment evaluation is done on cross domain dataset (SEntFiN 1.0). The evaluation was performed on all three subtasks for instruction tuned model.

The F1 score of model when trained on SEntFiN 1.0 for AE and Joint Task, resulted in values of 74.16 and 74.60, respectively

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Lapt14** | **Rest14** | **Rest15** | **Rest16** | **SEntFiN** |
| GPT2med | 82.04 | 75.94 | NA | NA | NA |
| GRACE | 87.93 | 85.45 | NA | NA | NA |
| BARTABSA | 83.52 | 87.07 | 75.54 | NA | NA |
| IT-MTL | 76.93 | NA | 74.03 | 79.41 | NA |
| **Inst ABSA** | **91.4** | **92.76** | **75.23** | **81.48** | **74.16** |

Table 3: AE subtask result for SEntFiN

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Lapt14** | **Rest14** | **Rest15** | **Rest16** | **SEntFiN** |
| ABSA-DeBERTa | 82.76 | 89.46 | NA | NA | NA |
| LSAT | 86.31 | 90.86 | NA | NA | NA |
| RACL-BERT | 73.91 | 81.61 | 74.91 | NA | NA |
| Dual-MRC | 75.97 | 82.04 | 73.59 | 79.41 | NA |
| **Inst ABSA** | **81.56** | **85.17** | **89.43** | **81.48** | **74.6** |

Table 4: ASC subtask result for SEntFiN

# **4.7 Comparision of Model Evaluation on SEntFin dataset**

|  |  |
| --- | --- |
| **Model** | **SEntFiN** |
| NLTK Vader | 64.79 |
| HuggingFace Sentiment | 72.9 |
| **Inst ABSA** | **78.56** |

Table 5: A comparison of non-entity aware sentiment analysis approaches on the SEntFiN dataset. NLTK Vader utilizes a lexicon-based approach with simple countbased features. For HuggingFace Sentiment system, we utilized the Twitter-roBERTabase model which is trained on ∼58 million tweets and finetuned for sentiment analysis with TweetEval benchmark.

# **4.8 Hyperparameters**

Model: Tk-Instruct-base-def-pos[[1]](#footnote-1) , ml.m5.xlarge : 4 CPU, 16 GB , Train Batch

Size for ATE and ATSC was set to 16, whereas for Joint Task it was 8, Gradient Accumulation Steps:2, Initial learning rate: 5e-5, Number of Epochs: 2

# **4.9 Requirement**

torch>1.3, transformers, pandas, sentencepiece

# **4.10 Future Research Scope**

Research is limited to SEntFiN dataset. In future research, there may be value in deploying even more compact models tailored to specific instructions to evaluate their efficacy. Research investigation was carried out utilizing TkInstruct models tailored to the English language, which may limit the direct applicability of this research results to other languages. Therefore, subsequent studies should broaden their scope to include datasets in multiple languages and employ instruction-tuned models that cater to various languages. This would allow for a thorough examination of the model's performance across a diversity of languages.

# **REFRENCES**

Albanese, F. and Feuerstein, E., (2021) *Community pooling: LDA topic modeling in Twitter*. [online] Available at: <https://github.com/fedealbanese/Community>.

Che, W., Zhao, Y., Guo, H., Su, Z. and Liu, T., (2015) Sentence Compression for Aspect-Based Sentiment Analysis. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 2312, pp.2111–2124.

Chen, Z. and Qian, T., (n.d.) *Relation-Aware Collaborative Learning for Unified Aspect-Based Sentiment Analysis*.

Chen, Z. and Qian, T., (n.d.) *Relation-Aware Collaborative Learning for Unified Aspect-Based Sentiment Analysis*.

Cheng, L.C., Chen, Y.L. and Liao, Y.Y., (2022) Aspect-based sentiment analysis with component focusing multi-head co-attention networks. *Neurocomputing*, 489, pp.9–17.

Efrat, A. and Levy, O., (2020) The Turking Test: Can Language Models Understand Instructions? [online] Available at: <http://arxiv.org/abs/2010.11982>.

Gandhi, A., Adhvaryu, K., Poria, S., Cambria, E. and Hussain, A., (2023) Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion*, 91, pp.424–444.

Hase, P. and Bansal, M., (2022) *When Can Models Learn From Explanations? A Formal Framework for Understanding the Roles of Explanation Data*. [online] Available at: <https://arxiv.org/abs/2102.02201>.

Hosseini-Asl, E., Liu, W. and Xiong, C., (n.d.) *A Generative Language Model for Few-shot Aspect-Based Sentiment Analysis*. [online] Available at: <https://github.com/>.

Hosseini-Asl, E., Liu, W. and Xiong, C., (n.d.) *A Generative Language Model for Few-shot Aspect-Based Sentiment Analysis*. [online] Available at: <https://github.com/>.

Jiang, Q., Chen, L., Xu, R., Ao, X. and Yang, M., (n.d.) *A Challenge Dataset and Effective Models for Aspect-Based Sentiment Analysis*. [online] Available at: <https://github.com/siat->.

Jindal, L., Kumar, S., Kaushal, C., Bhende, M., Thakare, A. and Shabaz, M., (2022) Aspect-based sentiment analysis employing linguistics content over social media for Web of Things. *IET Communications*.

Kamila, S., Magdy, W., Dutta, S. and Wang, M., (2022) AX-MABSA: A Framework for Extremely Weakly Supervised Multi-label Aspect Based Sentiment Analysis. [online] Available at: <http://arxiv.org/abs/2211.03837>.

Li, H., Yu, B.X.B., Li, G. and Gao, H., (2023) Restaurant survival prediction using customer-generated content: An aspect-based sentiment analysis of online reviews. *Tourism Management*, 96.

Lin, X.V., Mihaylov, T., Artetxe, M., Wang, T., Chen, S., Simig, D., Ott, M., Goyal, N., Bhosale, S., Du, J., Pasunuru, R., Shleifer, S., Koura, P.S., Chaudhary, V., O’Horo, B., Wang, J., Zettlemoyer, L., Kozareva, Z., Diab, M., Stoyanov, V. and Li, X., (2021) Few-shot Learning with Multilingual Language Models. [online] Available at: <http://arxiv.org/abs/2112.10668>.

Liu, N., Shen, B., Zhang, Z., Zhang, Z. and Mi, K., (2019) Attention-based Sentiment Reasoner for aspect-based sentiment analysis. *Human-centric Computing and Information Sciences*, 91.

Luo, H., Ji, L., Li, T., Duan, N. and Jiang, D., (n.d.) *Findings of the Association for Computational Linguistics GRACE: Gradient Harmonized and Cascaded Labeling for Aspect-based Sentiment Analysis*.

Mao, Y., Shen, Y., Yu, C. and Cai, L., (2021) *A Joint Training Dual-MRC Framework for Aspect Based Sentiment Analysis*. [online] Available at: www.aaai.org.

Min, S., Lewis, M., Zettlemoyer, L. and Hajishirzi, H., (n.d.) *MetaICL: Learning to Learn In Context*.

Mukherjee, A., Jethi, S., Jain, A. and Mundra, A., (2020) Aspect based sentiment analysis of student housing reviews. In: *PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing*. Institute of Electrical and Electronics Engineers Inc., pp.465–470.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang Sandhini Agarwal Katarina Slama Alex Ray John Schulman Jacob Hilton Fraser Kelton Luke Miller Maddie Simens Amanda Askell, C., Welinder Paul Christiano, P., Leike, J. and Lowe, R., (n.d.) *Training language models to follow instructions with human feedback*.

Parmar, M., Mishra, S., Purohit, M., Luo, M., Murad, M.H. and Baral, C., (n.d.) *In-BoXBART: Get Instructions into Biomedical Multi-Task Learning*. [online] Available at: <https://github.com/Mihir3009/In-BoXBART>.

Peper, J.J. and Wang, L., (2022) Generative Aspect-Based Sentiment Analysis with Contrastive Learning and Expressive Structure. [online] Available at: <http://arxiv.org/abs/2211.07743>.

Puri, R.S., Mishra, S., Parmar, M. and Baral, C., (2022a) How Many Data Samples is an Additional Instruction Worth? [online] Available at: <http://arxiv.org/abs/2203.09161>.

Puri, R.S., Mishra, S., Parmar, M. and Baral, C., (2022b) How Many Data Samples is an Additional Instruction Worth? [online] Available at: <http://arxiv.org/abs/2203.09161>.

Sanh, V., Webson, A., Raffel, C., Bach, S.H., Sutawika, L., Alyafeai, Z., Chaffin, A., Stiegler, A., Scao, T. Le, Raja, A., Dey, M., Bari, M.S., Xu, C., Thakker, U., Sharma, S.S., Szczechla, E., Kim, T., Chhablani, G., Nayak, N., Datta, D., Chang, J., Jiang, M.T.-J., Wang, H., Manica, M., Shen, S., Yong, Z.X., Pandey, H., Bawden, R., Wang, T., Neeraj, T., Rozen, J., Sharma, A., Santilli, A., Fevry, T., Fries, J.A., Teehan, R., Bers, T., Biderman, S., Gao, L., Wolf, T. and Rush, A.M., (2021) Multitask Prompted Training Enables Zero-Shot Task Generalization. [online] Available at: <http://arxiv.org/abs/2110.08207>.

Scaria, K., Gupta, H., Goyal, S., Sawant, S.A., Mishra, S. and Baral, C., (2023) InstructABSA: Instruction Learning for Aspect Based Sentiment Analysis. [online] Available at: <http://arxiv.org/abs/2302.08624>.

Sinha, A., Kedas, S., Kumar, R. and Malo, P., (2022) SEntFiN 1.0: Entity-aware sentiment analysis for financial news. *Journal of the Association for Information Science and Technology*, 739, pp.1314–1335.

Soni, P.K. and Rambola, R.K., (2021) Deep Learning, WordNet, and spaCy based Hybrid Method for Detection of Implicit Aspects for Sentiment Analysis. In: *2021 International Conference on Intelligent Technologies, CONIT 2021*. Institute of Electrical and Electronics Engineers Inc.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I., (2017) Attention Is All You Need. [online] Available at: <http://arxiv.org/abs/1706.03762>.

Wang, Y., Mishra, S., Alipoormolabashi, P., Kordi, Y., Mirzaei, A., Arunkumar, A., Ashok, A., Dhanasekaran, A.S., Naik, A., Stap, D., Pathak, E., Karamanolakis, G., Lai, H.G., Purohit, I., Mondal, I., Anderson, J., Kuznia, K., Doshi, K., Patel, M., Pal, K.K., Moradshahi, M., Parmar, M., Purohit, M., Varshney, N., Kaza, P.R., Verma, P., Puri, R.S., Karia, R., Sampat, S.K., Doshi, S., Mishra, S., Reddy, S., Patro, S., Dixit, T., Shen, X., Baral, C., Choi, Y., Smith, N.A., Hajishirzi, H. and Khashabi, D., (2022) Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ NLP Tasks. [online] Available at: <http://arxiv.org/abs/2204.07705>.

Wang, Y., Shen, G. and Hu, L., (2020) Importance evaluation of movie aspects: Aspect-based sentiment analysis. In: *Proceedings - 2020 5th International Conference on Mechanical, Control and Computer Engineering, ICMCCE 2020*. Institute of Electrical and Electronics Engineers Inc., pp.2444–2448.

Wei, J., Bosma, M., Zhao, V.Y., Guu, K., Yu, A.W., Lester, B., Du, N., Dai, A.M. and Le, Q. V., (2021) Finetuned Language Models Are Zero-Shot Learners. [online] Available at: <http://arxiv.org/abs/2109.01652>.

Weller, O., Lourie, N., Gardner, M. and Peters, M.E., (n.d.) *Learning from Task Descriptions*. [online] Association for Computational Linguistics. Available at: <https://allenai.org/data/zest>.

Yadav, R.K., Jiao, L., Goodwin, M. and Granmo, O.C., (2021) Positionless aspect based sentiment analysis using attention mechanism[Formula presented]. *Knowledge-Based Systems*, 226.

Yan, H., Dai, J., Ji, T., Qiu, X. and Zhang, Z., (n.d.) *A Unified Generative Framework for Aspect-Based Sentiment Analysis*. [online] Available at: <https://github.com/yhcc/>.

Yan, H., Dai, J., Ji, T., Qiu, X. and Zhang, Z., (n.d.) *A Unified Generative Framework for Aspect-Based Sentiment Analysis*. [online] Available at: <https://github.com/yhcc/>.

Yang, H. and Li, K., (2021) Improving Implicit Sentiment Learning via Local Sentiment Aggregation. [online] Available at: <http://arxiv.org/abs/2110.08604>.

Yang, J. and Yang, J., (2020) Aspect Based Sentiment Analysis with Self-Attention and Gated Convolutional Networks. In: *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*. IEEE Computer Society, pp.146–149.

Ye, Q. and Ren, X., (n.d.) *Learning to Generate Task-Specific Adapters from Task Description*. [online] Available at: <https://github.com/>.

Zhang, M. and Qian, T., (n.d.) *Convolution over Hierarchical Syntactic and Lexical Graphs for Aspect Level Sentiment Analysis*.

Zhong, R., Lee, K., Zhang, Z. and Klein, D., (n.d.) *Adapting Language Models for Zero-shot Learning by Meta-tuning on Dataset and Prompt Collections*. [online] Available at: <https://www.kaggle.com/c/>.

**Task Aspect Extraction (AE)**

Definition: The resulting output will encompass both implicit and explicit aspects, each paired with an associated sentiment extracted from the input text. In scenarios where no aspects are present, the output should duly reflect the absence of any aspect term.

|  |  |
| --- | --- |
| Positive Example 1:  Input: Sector-wide consolidation and improving fundamentals will boost Shoppers Stop's valuation.  Output Shoppers Stop  Positive Example 2:  Input: Silver, gold jump on festival season demand, Asian market cues.  Output: Silver, gold |  |
|  |  |

Negative example 1:

input: Gold prices soften in early noon trade.

output: Gold

Negative example 2:

input: Grasim net falls 59% on lower sales.

Output: keyboard

Neutral example 1-

Input: Sebi's stand against reluctant promoters may benefit small investors.

Output: Sebi

Neutral example 2-

Input: See Nifty at 7,460 in the short-term: Mitesh Thacker.

Output: Nifty

Table 6 Illustrating InstABSA instruction prompting for the AE sub task

**Aspect Sentiment Classification (ASC)**

Definition: If an aspect identified within a sentence carries a positive sentiment, the designated output will be 'positive.' Conversely, if the sentiment linked to the identified aspect is negative, the response will be tagged as 'negative.' In cases where the sentiment doesn't lean towards either positivity or negativity, the output is labelled as 'neutral.' For the aspects that fall under the classification of 'noaspectterm,' the sentiment is deemed as 'none.'

|  |  |
| --- | --- |
| Positive Example 1:  Input: Mid-cap funds can deliver more, stay put: Experts.  Output Mid-cap funds  Positive Example 2:  Input: Definance Technologies, now Hinduja Tech, plans to enter new markets.  Output: Definance Technologies, Hinduja Tech |  |
|  |  |

Negative example 1:

input: Foreign investors navigate turmoil in Chinese markets with new playbook.

output: Foreign investors

Negative example 2:

input: Stock Buzz: Sun Pharma may face resistance at current levels.

Output: Sun Pharma

Neutral example 1-

Input: Sebi's stand against reluctant promoters may benefit small investors.

Output: Sebi

Neutral example 2-

Input: See Nifty at 7,460 in the short-term: Mitesh Thacker.

Output: Nifty

Table 7 Illustrating InstABSA instruction prompting for the ASC sub task

**Joint Task (JT)**

Definition: The resulting output will encompass both implicit and explicit aspects, each paired with an associated sentiment extracted from the input text. In scenarios where no aspects are present, the output should duly reflect the absence of any aspect term

|  |  |
| --- | --- |
| Positive Example 1:  Input: Would bet on Bosch and Wabco in the auto ancillary sector; likely to do well in long-term: Pankaj Pandey.  Output Bosch, Wabco  Positive Example 2:  Input: DIIs likely to lap up Rs 1 lakh crore bonds; UDAY brings cheers to PF & insurance funds.  Output: PF & insurance funds |  |
|  |  |

Negative example 1:

input: Govts decision to increase import duty on natural rubber will hurt Make in India initiative: Tyre makers

Output: Tyre makers

Negative example 2:

input: Strides Arcolab may see some more downside if the Mylan deal does not come through: Avinnash Gorakssakar.

Output: Strides Arcolab

Neutral example 1-

Input: BSE, NSE to move 25 companies to restricted trade segment from October 7.

Output: BSE, NSE

Neutral example 2-

Input: D-Street trading: Mid & small cap stocks not yet out of the woods

Output: Mid , small cap stocks

Table 8 Illustrating InstABSA instruction prompting for the Joint sub task

Gantt Chart



1. https://huggingface.co/allenai/ tk-instruct-base-def-pos [↑](#footnote-ref-1)