Chapter 2

Background Study

The primary focus of ABSA research is to determine different parts of sentiment at the aspect level. It includes aspect terms, aspect categories, opinion terms and sentiment polarities [124]. Despite sentiment analysis's widespread and effective use, it is still insufficient for fully comprehending aspect-level opinions. This task necessitates identifying and extracting many sentiment parts. For instance, the process of extracting aspect category-sentiment polarity pairs involves identifying the aspect category and its corresponding sentiment polarity in a compound form. In other words, it entails extracting the pair (*Accessory*, *Positive*) from the review shown in table 1.1.

This chapter provides a comprehensive background study of traditional sentiment analysis, aspect term extraction, aspect category detection, aspect sentiment classification and aspect category-sentiment polarity pair in subsequent sections. It also describes Hindi dataset developed and well-accepted existing Hindi datasets. These datasets are used for experiments in this work.

2.1 Sentiment Analysis

Pang and Lee [87] conducted the initial study on determining the sentiment expressed in a text towards a specific subject. The text underwent a thorough examination and analysis of the computer handling of *opinion*, *sentiment* and *subjectivity*.

Aditya Joshi et al. [56] reported first work in Hindi language based sentiment analysis in 2010. They proposed a fall back strategy to do sentiment analysis for Hindi documents by three approaches. In the first approach, they constructed a sentiment annotated Hindi corpora and trained a classifier model. Then a new Hindi document is classified based on the trained classifier. In the second approach, author trained a classifier on English language and translate the given Hindi document into English to classify. In the third approach, they applied a majority based classifier for Hindi lexicon. These approaches achieved accuracy of 78.14% , 65.96% and 60.31% respectively. However the approach was not much effective to handle ambiguous nature of words to extract the actual aspect term polarity. Accuracy is a metric for performance evaluation of classification models. It tells the percentage of accurate predictions over total number of predictions.

Bakliwal et al. [12] did opinion mining in Hindi, however their work examined positive and negative classifications only. The methodology employed fusion of POS tagged n-gram and basic n-gram methods in conjunction with Hindi subjectivity lexicons. The accuracy scores were inferior in comparison to analogous experiments conducted in English.

Sharma et al. [100] discusses sentiment analysis of Hindi movie reviews, focusing on polarity detection. System highlights successful results in identifying positive, negative, and neutral sentiments using natural language processing techniques tailored to Hindi's linguistic structure. However, the system faces challenges such as handling mixed sentiments, context sensitivity, and complex negations in

Hindi sentences, which affect its accuracy. Despite these shortcomings, the study marks an important step toward improving sentiment analysis for Hindi language.

Ghosh and Dutta [41] conducted a real-time analysis of the sentiment expressed in Hindi tweets. They used a resource-based method to classify the messages into positive, negative, or neutral attitudes. The study conducted a comparative analysis of various techniques for removing stopwords [37] and performing POS tagging to determine their respective levels of efficiency. In addition, the study suggested methods to enhance the Hindi SentiWordnet as described by [56].

2.2 Aspect Term Extraction

Md Shad Akhtar et al. [2] were the first to develop a supervised method for extracting aspect terms. They accomplished this by creating a meticulously annotated dataset and constructing a machine learning model for sentiment analysis. Their goal was to illustrate how the dataset may be used in real-world scenarios. The aspect term extraction was performed using conditional random field(CRF)-based probabilistic classification. The F1-score reported was 41.07, with an accuracy of 54.05%.

In their study, Hetal Gandhi and Vahida Attar [36] conducted a comprehensive analysis of the sub-tasks involved in extracting aspect terms in Hindi. They employed CRF and Bi-LSTM classifiers for this purpose. The CRF-based model achieved an F1-score of 44.54, whereas the Bi-LSTM classifier-based model achieved an F1-score of 44.49 on the IIT, Patna [2] Hindi dataset.The stated results did not include any remark of accuracy.

Bhattacharya et al. [16] introduced a new Seq2Seq4ATE framework for ATE tasks. The researchers conducted experiments using the dataset [1] from

IIT, Patna as well as their own datasets. The researchers attained an F1-score of 35.04 on the IIT, Patna dataset and 68.61 on their own dataset. The results did not include information about accuracy, and the F1-score was found to be lower in the IIT, Patna dataset.

Kush Shrivastava and Shishir Kumar [101], enabled a Gated Recurrent Unit (GRU) network [84] to capture the semantic and syntactic relation between Hindi words and to classify them into sentiment classes. Here, authors adopted a genetic algorithm-GRU approach with tuned hyper-parameters. Experimental results showed better results as compared to deep learning based approaches. Experiments showed 65.96% accuracy for CNN-SVM Approach, 72.01% accuracy for RNN Approach and 88.02% accuracy for proposed approach.

Vandana Yadav et al. [120] performed experiments using LSTM neural networks to perform sentiment analysis on Hindi-language e-commerce reviews. The experiment demonstrated an accuracy rate of 87%.

Sujata Rani and Prateek Kumar [90] introduced a sentiment analysis technique based on a dependency parser [82]. The researchers conducted experiments on the Hindi movie dataset and achieved an accuracy of 83.2%.

2.3 Aspect Category Detection

ACD is vital to extracting critical insight from the input text. It can help with content recommendation, information retrieval etc. Xue et al. [119] proposed a multi-tasking neural network that addresses aspect category detection and aspect term extraction for English. The model employed a BiLSTM layer. Experiments were demonstrated on the SemEval dataset [32] and F1-score of 0.7642 F1-score was reported.

Existing literature [24], [70], [64], [65],[49] reports aspect category detection in Arabic [8], Turkish [86] and Vietnamese [109] languages. Machine learning and deep learning - based ACD is reported for the Hindi language, too. All these experiments on Hindi were performed on IIT Patna dataset. Akhtar et al. [5] proposed binary relevance and label power set [48] approaches. They obtained 0.46, 0.56, 0.30 and 0.64 F1-scores for ACD tasks in *Electronics*, *Mobile apps*, *Travels* and *Movies* domains, respectively.

Hetal V et al [51] performed aspect category detection based on the Feedforward Neural Network (FNN) approach. They also performed experiments on IIT Patna dataset [3] and achieved 0.67 F1-score for *Electronics* domain, 0.69 F1-score for *Mobile apps* domain, 0.56 F1-score for *Travels* and 0.72 F1-scores for *Movies* domains, respectively. FNN treats words as independent features and does not explicitly consider the order of words. Hence, FNN has limitations for text processing.

2.4 Aspect Sentiment Classification

Classifying the polarity of a review statement is a multi-class classification. Raksha Sharma and P Bhattacharyya [98] introduced a method that relies on a lexical resource to analyze sentiment in the Hindi language. The approach is formulated to retrieve words with the right polarity orientation, which is an essential undertaking, as a word can possess both positive and negative connotations simultaneously. The technique attained a classification accuracy of 87% on both a manually gathered movie review dataset and a product dataset.

The polarity determination models proposed by Brun et al. [20] explores an ensemble modeling approach combining syntactic and semantic knowledge to enhance Aspect-Based Sentiment Analysis (ABSA) and reported an accuracy of

88% and 79% for English and French reviews, respectively. The system utilizes techniques to generate an aspect-focused representation of features and assigns polarity to a term or sentence with the greatest likelihood using CRF. Additionally, the aspect category that was recognized is also utilized as the feature, along with syntactic features extracted from the parser.

Xenos et al. [118] employed an ensemble of two SVM models and reported an accuracy over 76% for both domains. The initial model was trained via manually designed and lexicon-driven characteristics. The second model was constructed by extracting characteristics from word-embeddings.

In their study, Md Shad Akhtar et al. [4] introduced a new and innovative deep learning framework that combines different techniques to provide effective sentiment analysis in Hindi language text. The sentiment vectors are created by a convolutional neural network and then enhanced by incorporating a set of optimum features picked using a Multi-Objective Optimization(MOO) framework. The SVM classifier was employed for both coarse-grained (sentence-level) and fine-grained (aspect-level) sentiment classification. The experiments yielded a sentence-level sentiment analysis accuracy of 62.52% and an aspect-level sentiment analysis accuracy of 65.96%. The experiment were conducted on four Hindi datasets and two English datasets. The summary is prsented in table 2.1.

Table 2.1: Summary of related work on Aspect Sentiment Classification

Author(s)	**Approach**	**Language(s)**	**Technique/Model**	**Level & Accuracy**
Raksha Sharma & P. Bhattacharyya [98]	Lexicon-based sentiment analysis using polarity orientation	Hindi	Lexical resource-based method	Sentence-level, 87%
Brun et al. [20]	Ensemble of syntactic and semantic features with CRF	English, French	Aspect-focused representation using CRF	Aspect-level, 79%
Xenos et al. [118]	Ensemble of two SVM models: lexical + word embeddings	Likely English	SVM with manual + embedding-based features	Domain-level,76%
Md Shad Akhtar et al. [4]	Deep learning with CNN, MOO, and SVM	Hindi, English	CNN + Multi-Objective Optimization + SVM	Sentence & Aspect, 62.52% & 65.96%

2.5 Aspect Category-Sentiment Polarity Pair Prediction

There has not been much research done on multitask learning for aspect category detection(ACD) and aspect sentiment classification(ASC). If ASC is a multi-class classification problem (where each sentiment polarity is a class), then ACD is essentially a multi-label classification problem (considering each category as a label). There are not many works that are known to identify aspect category-sentiment polarity consistently.

Cai et al. [21] introduced a hierarchical graph-CNN [123] that initially detects aspect categories and subsequently predicts polarity for each recognized category in English dataset. A hierarchical graph is constructed to represent relationships between words and aspect categories. The Graph Convolutional Network analyzes this graph, efficiently combining contextual and sentiment data. The evaluation metrics for the model are as follows: accuracy of 85.3%, F1 Score of 82.1%, precision of 83.5% and recall of 81.2%. This meticulous analysis improves precision by concentrating on precise elements and categories inside the text, resulting in a more intricate sentiment classification in contrast to conventional approaches.

Li et al.[69] address the problem of limited data by using a shared sentiment prediction layer to exchange sentiment knowledge among different aspect categories. The authors suggest the use of a pre-trained generative model. The desired output is represented using natural language sentences. The shared sentiment prediction layer thereby simultaneously learns and predict sentiments for both aspects and categories, hence improving accuracy. The approach guarantees consistent and coherent sentiment predictions by exchanging information across various features and categories. The utilization of this collaborative learning method greatly enhances sentiment analysis by offering more intricate and precise observations in contrast to conventional techniques. The efficacy of the model is evidenced by trials, which highlight its advantage in capturing subtle nuances of sentiment within the English text. The stated outcomes of the collaborative model for aspect-category sentiment analysis consist of an accuracy of 82.5%, an F1 Score of 79.3%, a precision of 80.1%, and a recall of 78.7%. These approaches are summarized in table 2.2.

Table 2.2: Summary of multitask learning approaches for aspect-category detection and aspect sentiment classification

Author(s)	**Approach**	**Language**	**Technique/Model**
Cai et al. [21]	Hierarchical Graph-CNN for detecting aspect categories and predicting sentiment polarity	English	Graph Convolutional Network (GCN) with hierarchical word-category graph
Li et al. [69]	Shared sentiment prediction layer to enable knowledge sharing across aspect categories using a pre-trained generative model	English	Collaborative multitask learning with natural language generation outputs

2.6 ABSA datasets

A number of new datasets have been created in the field of ABSA research in recent years due to its rapid advancement. Table 2.3 provides a summary of key details for Hindi datasets. These datasets used in experiments in this work.

Table 2.3: Summary of Hindi datasets

Dataset	Language	URL
1.	TU-HSA	Gathered in Excel format
2.	IITP-I [2]	https://www.aclweb.org/anthology/L16-1429
3.	IITP-II [3]	https://doi.org/10.1007/978-3-319-75487-1_19

- **TU-HSA(Tezpur University-Hindi Sentiment Analysis)**: We gathered manually the post-web reviews that users had created. These evaluations come from seven distinct disciplines. The three polarity categories that these reviews fall into are *positive*, *negative* and *neutral*. Polarity labels are used to identify reviews: (1) for positive reviews, (-1) for negative reviews, and (0) for neutral reviews. We translated reviews from Hindi to English using Google translate. TU-HSA is in Excel format. Table 2.4 presents the TU-HSA dataset statistics.

Table 2.4: TU-HSA dataset details

S No	Domain	Positive	Negative	Neutral	Total
1	Games	686	252	70	1008
2	Corona	454	481	79	1014
3	Bollywood	655	308	48	1011
4	Mobiles	763	150	102	1015
5	Cars	797	160	44	1001
6	Bikes	875	77	56	1008
7	Laptops	671	189	156	1016
	Total	4901	1617	553	7071

Three labelled Hindi review examples from each domain and translated English review are presented below. Each review have different polarity.

i) **Domain** - *Games*

1. भारतीय टीम दूसरा क्वार्टर आने तक पूरी रंगत में खेलने लगी। 1

   The Indian team started playing in full form till the second quarter came. 1

2. इस हार के बाद भारत की मुश्किलें बढ़ गई हैं । -1

   India's troubles have increased after this defeat. -1

3. जेमिमा रोड्रिग्ज ने नाबाद दो रन बनाए । 0

   Jemima Rodriguez scored two not out. 0

ii) **Domain** - *Corona*

1. गौरतलब है कि अब तक 5 करोड़ 30 लाख टेस्ट करने वाला उत्तर प्रदेश अकेला राज्य है। 1

   Significantly, Uttar Pradesh is the only state to have conducted 5 crore 30 lakh tests so far. 1

2. इससे पहले भारत समेत दुनियाभर में डेल्टा वेरिएंट ने कहर मचाया था। - -1

   Earlier, the Delta variant had wreaked havoc across the world including India. - -1

3. 21 मार्च को भारत में कोविड के 1410 नए मामले सामने आए थे। 0

   On March 21, 1410 new cases of Covid were reported in India. 0

iii) **Domain** - *Bollywood*

1. फिल्म को दर्शकों से बेहिसाब प्यार मिला। 1

   The film received immense love from the audience. 1

2. बाद में फिल्म कई विवादों में घिर गई। -1

   Later the film got embroiled in many controversies. -1

3. लोग केवल चलता-फिरता चित्र देख सकते थे। 0

   people could only see moving pictures. 0

iv) **Domain** - *Mobiles*

1. कुछ चुनिंदा बैंकों के साथ डिस्काउंट और क्रेडिट कार्ड ऑफर्स भी दिए जा रहे हैं। 1

   Discounts and credit card offers are also being offered with select banks. 1

2. हालांकि सैमसंग ने प्रोसेसर की जानकारी नहीं दी है। -1

   Although Samsung has not given details of the processor. -1

3. Infinix InBook X1 Slim की भारत में कीमत 29,990 रुपये से शुरू होती है। 0

   Infinix InBook X1 Slim price in India starts from Rs 29,990. 0

v) **Domain** - *Cars*

1. कुल मिलाकर नई Scorpio N को चलाने में आपको मजा आएगा। 1

   Overall, you will enjoy driving the new Scorpio N. 1

2. AC के लिए यहां टेम्परेचर और ब्लोअर स्पीड कंट्रोल नहीं दिया। -1

   Temperature and blower speed control not given here for AC. -1

3. इलेक्ट्रिक सनरूफ, डुअल जोन क्लाइमेट कंट्रोल, लेदर अपहोलस्ट्री, पुश बटन स्टार्ट, कनेक्टेड कार टेक्नोलॉजी, ऑटोमैटिक हेडलैंप्स, ऑटोमैटिक वाइपर्स, फ्रंट और रियर कैमरा और एक वायरलेस फोन चार्जर आपको इसमें मिल जाता है। 0

   Electric sunroof, dual zone climate control, leather upholstery, push button start, connected car technology, automatic headlamps, automatic wipers, front and rear cameras and a wireless phone charger. 0

vi) **Domain** - *Bikes*

1. भारत में लॉन्चिंग के बाद से ही पल्सर रेंज काफी लोकप्रिय रही है। 1

   The Pulsar range has been quite popular since its launch in India. 1

2. इसमें वह आवाज नहीं थी जिसकी ग्राहक चाह रख रहे थे। -1

   It didn't have the sound customers were looking for. -1

3. नंबर प्लेट माउंट, इंडिकेटर और रिफ्लेक्टर सभी एक यूनिट पर लगे हैं। 0

   Number plate mount, indicator and reflector all on one unit. 0

vii) **Domain** - *Laptops*

1. स्टूडेंट्स और ऑफिस वर्कर्स के लिए यह रोजमर्रा में बहुत पोर्टेबल है। 1

   It is very portable for everyday use for students and office workers. 1

2. अफसोस कि इसपर टाइपिंग करना बहुत सहूलियत भरा नहीं है। -1

   Sorry it's not very comfortable to type on. -1

3. 25W के टाईप-सी चार्जर से यह लैपटॉप करीब तीन घंटे में फुल चार्ज होता है। 0

   This laptop is fully charged in about three hours with a 25W Type-C charger. 0

Manual collection is done for these post-web reviews. Additionally, annotation is done by hand. These evaluations have been gathered from the websites listed below:-
www.bbc.com, www.espncricinfo.com, www.bhaskar.com, www.zeenews.india.com, www.ndtv.com, https://www.bollywoodnews.org/, https://hindi.news18.com/, www.jagran.com, https://www.bollywoodhungama.com/hindi/, https://hindi.gadgets360.com/, https://drivesparls.com.
TU-HSA is employed to perform machine learning classifiers based SA for Hindi-English language. Data description for TU-HSA is presented in table 2.5

Table 2.5: Domain-wise Dataset Details and Sources

S.No.	Domain	Sentences	Sources	Post-web
1	Games	1008	1. www.bbc.com 2. www.espncricinfo.com 3. www.bhaskar.com 4. www.essayinhindi.com	Yes
2	Corona Virus	1014	1. www.bbc.com 2. www.bhaskar.com 3. www.zeenews.india.com 4. https://nibandh.net 5. www.ndtv.com	Yes
3	Bollywood News	1011	1. https://www.bollywoodnews.org 2.https://hindi.news18.com/ 3. www.jagran.com	Yes
4	Mobiles	1015	1.https://hindi.gadgets360.com/	Yes
5	Cars	1001	1. https://hindi.drivespark.com/car-reviews 2. https://www.patrika.com/car-reviews	Yes
6	Bikes	1008	1. https://drivesparls.com	Yes
7	Laptops	1014	1. https://amarujala.com	Yes
Total Sentences		**7071**		

- **IITP-I(Indian Institute of Technology Patna-I):** The dataset in XML [26] format. There are twelve domains in which these Hindi review sentences fall. The following categories are included: (i) laptops; (ii) mobile phones; (iii) tablets; (iv) cameras; (v) headphones; (vi) home appliances; (vii) speakers; (viii) televisions; (ix) smart watches; (x) mobile apps; (xi) travel; and (xii) movies. Table 2.6 presents a few of the dataset's fundamental statistics.

Table 2.6: Some basic statistics

Metrics	**IITP-I**
Sentences	5417
Aspect terms	4509
Tokens	96140
Average of aspect terms per sentence	0.81
Total number of domains	12

Table 2.7 shows the two instances from IITP-I together with their respective structures. The <sentences> node is the XML's root node. It includes <sentence>, or the review's sentences as its offspring. Every <sentence> has a 'id' attached to it in order to uniquely identify it. It displays information about the domain. There are two offspring per <sentence> node: <text> and <aspectTerms>. While <aspectTerms> has <aspectTerm> nodes, the <text> node only has one review phrase. At least one can be <aspectTerm>. Aspect term information is handled by these nodes. The four properties that each <aspectTerm> node possesses are 'term', 'from', 'to', & 'polarity'. The attribute "term" designates the aspect term that the current node represents, and "polarity" points how the user feels about the "term", which may be "*pos*". The options are '*neu*' (neutral), '*neg*' (negative) or '*pos*' (positive).

Table 2.7: IITP-I labelled structure

Labelled structure

```
<sentences>
   <sentence id="app_2">
      <text>
फेसबुक का सिक्योरिटी चैकअप फीचर पॉपअप की तरह यूजर्स को दिखाइ देगा।
      </text>
         <aspectTerms>
<aspectTerm from="10" polarity="neu" term= "सिक्योरिटी चैकअप फीचर" to="31"/>
         </aspectTerms>
</sentence>
   <sentence id="app_94">
      <text>
इसमें बहुत उच्च स्तर के ग्राफ़िक्स नहीं हैं लेकिन यह गेम बहुत आकर्षित करने वाला है।
      </text>
......
......
</sentences>
```

IITP-I dataset is used for aspect term extraction task.

- **IITP-II(Indian Institute of Technology Patna II):** The dataset is in XML format. These 5417 review sentences same as IITP-I belong to fourteen different categories. These categories are budget, design, ease of use, GUI, hardware, misc, music, performance, place, price, reachability, scenery, software and story. Within the dataset, these categories are not dispersed equally. Distribution of different domains in four aspect categories is shown in table 2.8. There are total of 2250 *'positive'*, 635 *'negative'*, 2241 *'neutral'* and 128 *'conflict'* instances of aspect categories.

Table 2.8: Aspect categories corresponding to distinct domains

Domains	Aspect categories
Electronics	Design, Software, Hardware, Ease of use, Price, Misc
Mobile apps	GUI, Ease of use, Price, Misc
Travels	Scenery, Place, Reachability, Misc
Movies	Story, Performance, Music, Misc

Two instances of the dataset with their structure are depicted in table 2.9.

The <sentences> node is the XML root node. It contains sentences of the review as its children i.e. <sentence>. To uniquely identify each <sentence>, an 'id' is associated with it. It indicates an information about domain. Each <sentence> node has two children, namely <text> and <aspectCategories>. The <text> node consists one review sentence, whereas <aspectCategories> contains <aspectCategory> nodes. <aspectCategory> can be 1 or more. These nodes handles aspect category information. Each <aspectCategory> node holds two attributes: 'category' & 'polarity'. Attribute 'category' defines aspect term category represented by current node while 'polarity' stores the sentiment towards the 'term' which is either '*pos*'(positive), '*neg*'(negative) or '*neu*'(neutral). IITP-II is used for ACD, ASC and aspect category-sentiment polarity pair prediction task.

Table 2.9: IITP-II labelled structure

Labelled structure
<sentences>
<sentence id="app_2">
<text>
फेसबुक का सिक्योरिटी चैकअप फीचर पॉपअप की तरह यूजर्स को दिखाइ देगा।
</text>
<aspectCategories>
<aspectCategory category="gui" polarity="neu"/>
<aspectCategories>
</sentence>
<sentence id="app_94">
<text>
इसमें बहुत उच्च स्तर के ग्राफ़िक्स नहीं हैं लेकिन यह गेम बहुत आकर्षित करने वाला है।
</text>
<aspectCategories>
<aspectCategory category="gui" polarity="neg"/>
<aspectCategory category="misc" polarity="pos"/>
</aspectCategories>
</sentence>
……
……
</sentences>

2.7 Classifications in Machine Learning

In machine learning, classification divides data into several categories, classes and groups. Based on the characteristics of the target variable, machine learning classification can be broadly divided into three types: binary, multi-class and multi-label classification. Below is a synopsis of each:-

1. **Binary Classification:** In binary classification, the target variable has only two possible classes or outcomes.
   Examples:
   - Spam detection: classifying emails as "spam" or "not spam".
   - Medical diagnosis: classifying a tumor as "benign" or "malignant".

   Common Algorithms: Logistic regression(LR), Support Vector Machine(SVM), Decision Tree(DT), Random Forest(RF), Naive Bayes(NB).

2. **Multi-class Classification:** In multi-class classification, the target variable has more than two classes and each instance is assigned to one and only one class.
   Examples:
   - Handwriting recognition: Classifying handwritten digits (0-9).
   - Species classification: classifying animals into species such as "cat", "dog", "rabbit" etc.

   Common Algorithms: Multinomial Naive Bayes(MNB), Decision Tree(DT), Random Forest(RF), Neural Networks(NN), k-Nearest Neighbours(KNN) and Naive Bayes(NB).
   Approaches:
   - **One-vs-Rest (OvR)** [108]: Train one classifier per class, each distinguishing that class from all other classes.

- **One-vs-One (OvO)** [35]: Train a classifier for every pair of classes.

3. **Multi-label Classification**[50]: In multi-label classification, each instance can belong to more than one classes simultaneously. *Examples:*

    - Text categorization: tagging a news article with multiple topics such as "politics", "economy" and "sports."
    - Image tagging: assigning multiple labels to an image, such as "beach", "sunset" and "vacation."

    Common Algorithms:

    - Adapted versions of algorithms like decision tree, random field and neural network to handle multi-label outputs.
    - Algorithms specifically designed for multi-label classification like Binary relevance, Classifier chains and Label power-set.

    Approaches:

    - **Binary relevance**[122]: Treat each label as a separate single-label binary classification problem.
    - **Classifier chains**[91]: Chain classifiers in such a way that the prediction of one label is used as a feature for predicting the next label.
    - **Label power-set**[57]: Transform the problem into a multi-class problem with one class for every label combination observed in the training data.

2.8 Summary

In this chapter, we presented a background study of the approaches to traditional sentiment analysis, aspect term extraction, aspect category detection, aspect sentiment classification and aspect category-sentiment polarity pair prediction tasks.

Additionally, we established the gaps present in the existing works. We presented an introduction to TU-HSA, a dataset we developed and well accepted datasets IITP-I and IITP-II.