Aspect Based Sentiment Analysis of Unlabeled Reviews Using Linguistic Rule Based LDA

Nikhlesh Pathik, Institute of Engineering and Technology, DAVV, India

https://orcid.org/0000-0002-3303-8684

Pragya Shukla, Institute of Engineering and Technology, DAVV, India

https://orcid.org/0000-0002-3497-6824

ABSTRACT

In this digital era, people are very keen to share their feedback about any product, service, or current issues on social networks and other platforms. A fine analysis of this feedback can give a clear picture of what people think about a particular topic. This work proposed an almost unsupervised aspect based sentiment analysis approach for textual reviews. Latent Dirichlet allocation, along with linguistic rules, is used for aspect extraction. Aspects are ranked based on their probability distribution values and then clustered into predefined categories using frequent terms with domain knowledge. SentiWordNet lexicon used for sentiment scoring and classification. The experiment with two popular datasets shows the superiority of the strategy as compared to existing methods. It shows 85% average accuracy when tested on manually labeled data.

KEYWORDS

Aspect Extraction, Aspect Ranking, Aspect based Sentiment Analysis, Classification, LDA, Linguistic Rules, NLP, SentiWordNet, Topic Modeling, Unsupervised

INTRODUCTION

Nowadays, people are very expressive on the web. Due to the exponential growth in user feedback data, it becomes necessary for every product and service provider to perform the mining of these feedbacks. People regularly share their views on current activities on Twitter or similar platforms. A fine-grained analysis of these tweets or reviews can give a clear picture of what people think about a particular topic. That is why aspect based sentiment analysis has gained popularity, and a lot of work has been done in this area in the last decade. Still, it is an active research area, especially unsupervised approaches that require improvements (Yue et al. ,2019, Do et al. , 2019).

The primary differentiation in sentiment analysis and aspect specific sentiment analysis is that the former only detect the sentiment of an overall text. Later, investigate each text sentence to find out various aspects and then determine the emotion associated with each of them. We can say, instead of evaluating the overall sentiment of a text, an aspect based approach allows us to associate specific opinions with various aspects or features of a product and service. The aspect based analysis looks more closely at the information behind a text. That is why results are more detailed and accurate.

Suppose we consider the example of COVID public sentiment analysis based on social network data. Then we require to analyze the various issues or aspects related to COVID and public sentiment about that. Here overall polarity may not be a good indicator. We required sentiment about a particular issue. The same analysis is required for every business and service related feedbacks or opinion. Due to regularly generating massive feedback data, unsupervised and semi-supervised approaches are gaining popularity.

Topic modeling is an unsupervised NLP technique representing a group of text documents with several topics that can best explain the underlying information in each document. It seems similar to clustering with a difference. Instead of numerical features, it has a collection of words. These words need to be grouped so that each group represents a topic in a document. Latent Dirichlet Allocation (LDA) is the most well-known method for modeling thematic information, i.e., topics from the document collection. It is an unsupervised learning approach that views documents as a bag of words. LDA is used in an extensive collection of documents to classify topics(Beli et al., 2003). It is helpful for Search Engine Optimization, automation of customer service, and any other instance where knowing the theme of documents is essential. It applies to the role of describing topics that best represent a collection of documents. During the topic modeling method, these topics emerge and are therefore named latent.

The main contributions to this work include the following:

1. An unsupervised aspect extraction approach using optimized LDA configuration and Parts of Speech (POS) rule for unlabeled reviews.
2. Categorization of aspects, using very few domain words.
3. Aspect specific analysis of sentiment using SentiWordNet(SWN).

The remaining structure of the paper is as follows: Section 2 sheads light on the latest work in the field. The background and intuition of LDA and SWN described in section 3. The methodology and proposed algorithms are explained in section 4. Section 5 presented experimental details and results. The paper concluded with summarization and future directions in Section 6. In this paper, the word sentiment and opinion are used interchangeably, similarly word aspect and feature.

RELATED WORK

For this study, topic modeling based approaches are mainly considered for sentiment analysis. Some hybrid models based on deep neural networks and LSTM are also discussed. We focused on very recent work of the last 3-4 years in this field.

The various survey describes the present state of arts in sentiment analysis research, mostly online reviews, and social media data(Yue et al., 2019). Detailed analysis of different Deep learning based approaches discussed along with their performance issues(Do et al., 2019). LDA was presented by Blei et al. (2003), and even after almost two decades, it is still increasing its popularity in unsupervised topic extraction.

Tai et al. (2015) presented LDA based approach for predicting person mental stability by analyzing online dairies. Non-sentiment words removed from data, and LDA extracted concept keywords. SWN is used to calculate the emotion score. An aspect based review summarization using LDA and sentiment lexicon is presented by Akhtar et al. (2017). Manually annotated hotel review data used for their study. Wang et al. (2017) presented a hybrid, supervised text classification method for social media data. Feature selection performed using LDA and optimized SVM used for sentiment classification. The maximum entropy-based joint aspect-sentiment model presented by Tang et al. (2019) for sentiment analysis of online reviews.

Yiran and Shrivastava(2019) presented an Aspect Based Sentiment Analysis(ABSA) approach using ontology-based LDA with a weighted aspect on Amazon mobile reviews dataset. Concept-LDA presented for better aspect extraction and quality topics. It is a combination of words, named entities, and concept terms (Ekinci & İlhan, 2020). Pathik and Shukla(2020) proposed an algorithm using Simulated Annealing for LDA hyperparameter tuning for better coherence and more interpretable output. A topic-model based approach used for aspect-oriented opinion mining. The manual effort requires assigning topics to the extracted aspects (Anoop & Asharaf, 2018).

Sokhin & Butakov (2018) discussed a semi-supervised LDA-based approach with additive regularization for sentiment analysis. Recently Long Short Term Memory (LSTM) based approaches are getting popularity in the ABSA task. LDA combined with various Deep Neural Networks for the same job. A hybrid model using LDA, LSTM, and CRF applied for rare and emerging Name Entities from user-generated social media data(Jansson and Liu,2017). A mortality prediction model from ICU admitted patients' clinical remarks discussed by Jo et al. (2017). The model uses LDA and LSTM with simultaneous training and learning. Pergola et al. (2019) presented a hybrid model based on LDA-GRU for feature extraction and sentiment classification. Topic embeddings and internal attention applied for jointly dealing with topic and sentiment.

Ma et al. (2018) discussed an LSTM attention model for aspect level sentiment analysis using embedding and commonsense knowledge. Dieng et al. (2020) presented an embedded topic model that uses both LDA and embeddings' properties to learn more interpretable topics. An amortized variation inference algorithm was proposed for the same. Neural Network based various sentiment analysis approaches presented for Indian languages. Different configurations of models prepared to demonstrate hyperparameters' effect on the performance(Bhargava et al.,2019).

The main challenge with DNN based methods is the target variable in the data, which is not there in unsupervised data. The majority of approaches considered manually labeled sentiment datasets. Another difficulty is the interpretation and output representation of LDA output. It is not in a favourable format for applying any classification approach directly to it. No well-known representation of LDA output is well suited for DNN models.

Many works have been done in this field, but most of them required human annotation for better results. An efficient ABSA approach requires (i) efficient aspect extraction, (ii) identification of sentiment about that aspects, and (iii) sentiment polarity classification. In the presented work, all these are attempted to make our approach better than the existing one. An optimized LDA model with linguistic rule is used to extract aspects efficiently. SWN, along with some rules, improves the sentiment classification task.

BACKGROUND

In this section, the background and intuitions of the LDA and SWN discussed in detail.

Latent Dirichlet Allocation(LDA)

LDA assumes that every document is a mixture of topics, and every word has a certain probability of falling into a particular topic. Consider an extensive collection of documents, LDA attempts to map these documents to the latent topics so that these topics mostly capture words present in each of the documents.

In LDA, each word in each document comes from a topic. The topic selects from a per-document distribution over topics. The topic distribution Θ for each document is proportional to Dirichlet(α), and the word distribution Φ is proportional to Dirichlet(β). Hyperparameters α and β plays a vital role in the generative process of LDA. α controls the documents-topic concentration. The low value of α represents the documents as a mixture of a few topics, whereas its high value results in more topics per documents. The optimal value of α controls document similarity within a topic. Similarly,

β controls topic-word concentration. A low value of β represents the topics with fewer unique words making it more distinct, and its high value results in more unique words in each topic.

LDA topic generation process depends on these two probability distributions:

1. P(t|d) = The probability distribution of topics in documents Θ_{td}.
2. P(w|t) = The probability distribution of words in topics Φ_{wt}.

The utmost goal of LDA is to estimate the probability of a word given document, i.e., P(w|d) with the help of above two probabilities, which is given by

$$P\left(\frac{w}{d}\right) = \sum t \in T \ P\left(w \mid t\right) P(t \mid d) \tag{1}$$

It is the dot product of Θ_{td} and Φ_{wt} for each topic t.

Gibbs sampling applied for successively sampling conditional distributions of variables. In the long-run, distribution over states converges to the accurate one. The equation for the same is below:

$$p(z_{d,n} = k \mid \vec{z}_{-d,n} , \vec{w} ,\alpha, \beta) = \frac{n_{d,k} + \alpha_k}{\sum_i^k n_{d,i} + \alpha_i} \ \text{X} \ \frac{v_{k,w_{d,n}+\beta_{w_{d,n}}}}{\sum_i v_{k,i} + \beta_i} \tag{2}$$

where

$n_{d,k}$: # document d use topic k
$v_{k,w}$: # topic k uses the given word

α and β: Dirichlet parameter for the document to topic and topic to word distribution, respectively

[3]

SentiWordNet (SWN)

SWN is a sentiment lexicon derived from WordNet. To each synset of WordNet, it assigns three sentiment scores: positive, negative, objective. SWN based approach is unsupervised learning, which means there is no need for training data and training steps. It has its tagger (POS). It tokenizes sentence and token entry match with SWN synset data. If an entry exists, it returns positive, negative, and neutral sentiment scores (Esuli et al.,2006).

SWN computes aggregated sentiment scores for a sentence based on the polar words present in the sentence. It means the sentiment of each synset aggregated to compute the sentiment score of a sentence. The same way sentiment of a complete document is computed (Baccianella et al.,2010).

In most cases, SWN gives higher accuracy in sentiment scoring. In some cases, It does not produce accurate results due to the non-matching of a synset with tokens. It may be possible with a new word or some time lemmatization results in incorrect word form. In the case of negation and qualifiers, scores are not accurate.

METHODOLOGY

The primary focus of this work is on developing an almost ABSA approach for online reviews based on LDA. The considered dataset reviews preprocessed first, then tokenized into sentences and words. The bag of words generated through tokenized words. Tokenized sentences are stored separately for sentiment analysis further.

LDA applied to a bag of words created by tokenized terms and generate topic word distribution. Based on POS rules and probability values, essential aspects identified, and a dictionary designed for them. Based on the probability distribution value in the topic, these aspects are ranked. These aspects are categorized into various clusters using domain knowledge. Some domain-related words also are added as extended aspects terms for improving the aspect categorization.

Now, based on the aspect categories, review sentences are processed for sentiment analysis. Each sentence is related to a particular aspect map with its sentiment values. Sentiment calculation for each aspect category is done. SWN is used as a sentiment lexicon for sentiment classification and evaluation. The sentiment score for each sentence added to get the average sentiment score about a particular aspect. The same process repeated for all aspect categories, and the average sentiment score is calculated for all. Finally, based on the average sentiment score, we can determine a particular aspect's sentiment strength in the given review data. The evaluation of accuracy is performed on manually annotated data. Figure 1 represents the flow of the proposed methodology.

The complete ABSA process is divide into three sub-procedures or Algorithms:

1. Aspect Extraction and Ranking
2. Aspect Expansion and Categorization
3. Aspect Based Sentiment Analysis

Now each of the above Algorithm discusses separately with their explanation.

1. Aspect Extraction and Ranking Algorithm

Input: Review data
Output: Most probable Aspect list with their rank
Step 1: Input review corpus and preprocess it.
Step 2: Tokenize preprocessed corpus into words.
Step 3: Remove outliers from tokenized words.
Step 4: Create a Bag of Words (BOW) representation.
Step 5: Apply LDA on BOW to get topic-word probability distribution.
Step 6: Apply POS rules on LDA output distribution to get the most probable aspects. Create a list of aspects.
Step 7: Rank the topmost aspects based on the probability value.

In this work, two popular datasets from different domains are considered for ABSA. For any text input based method, preprocessing is the first and most crucial step. The same is done in this work also. Standard preprocessing steps are as follows: HTML tags removal, email removal, punctuations, and other special symbols removal. After that, stop word removal and, and the text normalization performed using lemmatization.

Now another important step is the removal of outliers. Outliers are the terms that have either shallow frequency or very high frequency both are not in the favour in the case of unsupervised algorithms. For example, in mobile reviews, the word "mobile" came in almost every review and "hotel" in the hotel domain. Similarly, some words are coming very rarely, so not affecting much. Removal of outliers gives better probability distribution as LDA output.

Figure 1. Flow chart of the proposed ABSA methodology
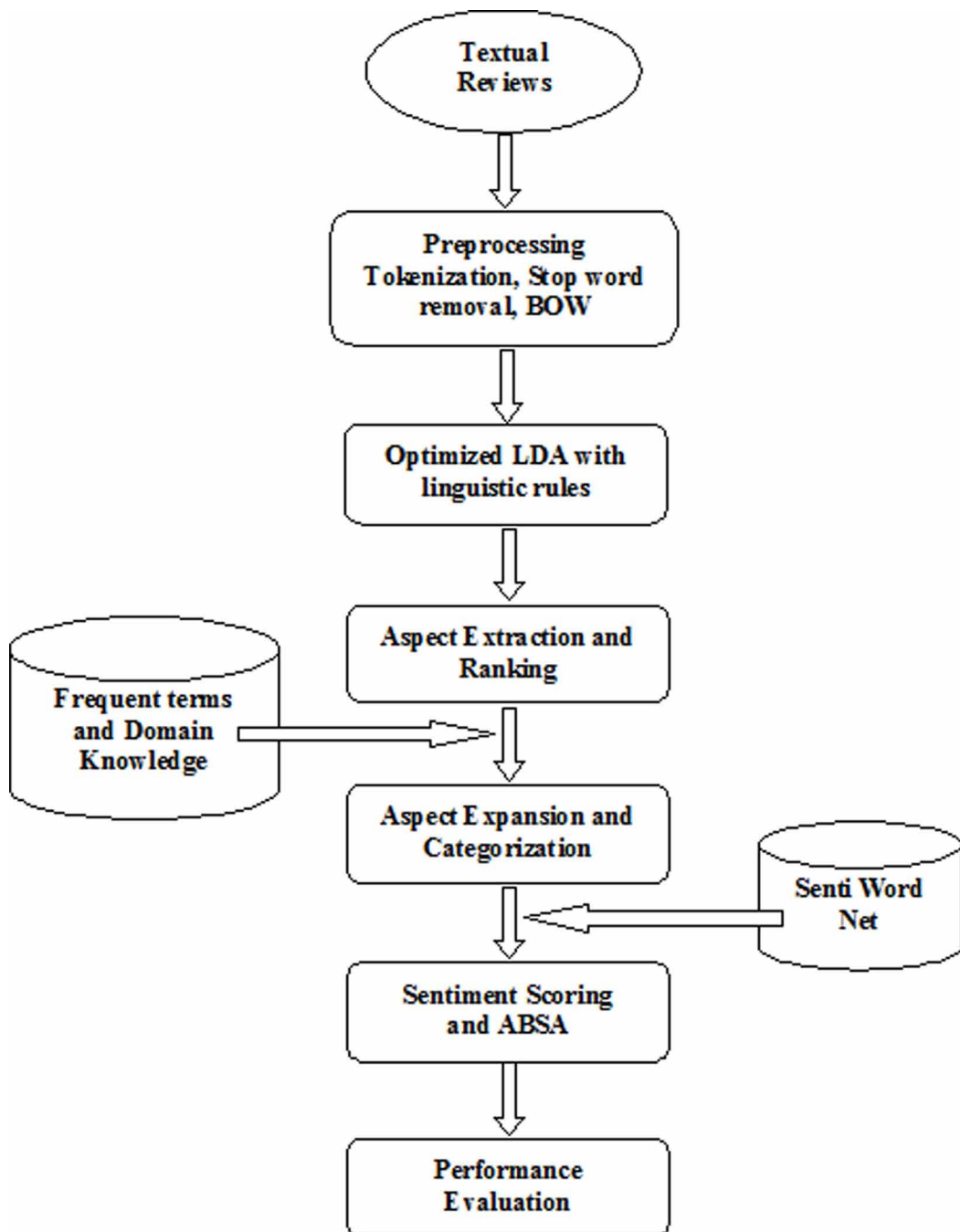
Corpus is tokenized into word and finally converted into the BOW. LDA operates on BOW and generates topic word probability distribution for supplied input. Linguistic rules(POS) applied to LDA output and extracted aspects from their probability distribution value. These aspects are stored for further process.

LDA output quality measures using the Coherence value. LDA hyperparameter tuning produces cohesive output. So we have tuned LDA hyperparameters using SA-LDA(Pathik and Shukla,2020). This optimized LDA configuration generates more coherent aspects as compare to standard LDA. Based on the LDA output, i.e., topic probability distribution values, and POS rules, the highest probability aspect is selected from each topic. Aspect are also ranked based on this probability value.

2. Aspect Expansion and Categorization Algorithm

Input: Aspect list from Algorithm-1 with additional aspect terms taken from frequent words.
Output: Aspect categories with extended terms and separate reviews per aspect-category.
Step 1: Take the aspect list from Algorithm-1 as input.
Step 2: Extend aspect list using frequent words from the corpus with human inception.
Step 3: Apply domain knowledge to categorize aspects into various categories.
Step 4: Scan the review sentence, and separate them as per the aspect category.
Step 5: Create separate files for these review sentences for further process.

The aspect list generated using Algorithm-1 expanded using frequent corpus terms and domain knowledge. In the same way, aspects are categorized into various categories based on their correlations (Akhtar et al., 2017). For example, battery, charger, and backup all are put into the same cluster in the mobile domain. Similarly, in the hotel domain, food, breakfast, and tea should be kept in the same category. By applying the same approach, we can divide the aspects into different logical categories and separate the reviews for further process. The aspect expansion help us to cover more reviews, and aspect categorization gives a better picture of various aspects.

3. Aspect Based Sentiment Analysis Algorithm

Input: Categorized aspects and review data with Sentiment Lexicon
Output: Sentiment scores for each Aspect category
Step 1: Process review sentences containing aspects for each aspect category separately.
Step 2: Compute the sentiment of each sentence about each aspect category using the SWN lexicon.
Step 3: Aggregate the sentiment scores for all sentences containing a particular aspect category.
Step 4: Repeat (3) for all aspect categories and calculate the sentiment score.
Step 5: Summarize the sentiment score for each aspect category.

Aspect categories generated by Algorithm-2 with the separated reviews, we are now ready for ABSA. For this, SWN is used as a sentiment lexicon. Every review sentence is scanned for a particular category aspect to calculate its sentiment. The positive and negative sentiments are stored separately. It repeats for all the reviews. In the end, overall sentiment (positive and negative both) along with its strength for a particular category evaluated. For each category, the same procedure applied to get the sentiment scores for all aspect categories. Sentiment scoring is perform using SWN, which operates on Synsets. The sentiment score is the sum of sentiment values for each synset present in the sentence for a given sentence. For example, the sentiment score of the sentence " This camera of this mobile is good." is evaluated as follows:

```
POS : [('The', 'DT'), ('camera', 'NN'), ('of', 'IN'), ('this',
'DT'), ('mobile', 'NN'), ('is', 'VBZ'), ('good', 'JJ')]
Sentiment Scores: [  [], [0.0, 0.0, 1.0], [], [], [0.0, 0.0, 1.0],
[],[0.75, 0.0, 0.25]]
```

POS represents the parts of speech tag for each word. Scores represent a ternary value for each word. The empty "[]" represent that word is not present in the list. The sentiment score is represented in three values [positive, negative, neutral]. If the positive value is greater than the negative, then the sentiment for that synset is positive. Similarly, if the negative value is greater than the positive, then the sentiment is negative. If negative and positive are equal, then the word is neutral. The sentiment score for a sentence is the sum of the sentiment of all the synset present. The sentiment score for the above sentence is 0.75, and it is positive.

SWN does not give an accurate sentiment score for sentences having negations. So we separately deal with review sentences containing negation. Their sentiment scores evaluate separately, and in the end, their polarity has reversed. For example, the sentiment score of the sentence " This camera of this mobile is not good." is 0.75, and now it is negative. In the same way, the sentiment scores of the sentences containing negation words are computed.

EXPERIMENT SETUP AND RESULTS

Gensim implementation of LDA used with customization on the Anaconda platform using Python. The experiments conducted on a laptop Core I5 CPU @ 2.5 GHz 2.49 GHz with 8GB RAM on Windows 8 OS. Orange Canvas is also used for some visualization purposes.

Two popular datasets from two different domains considered for analysis. For the Mobile domain, Amazon unlocked mobile reviews[1] are chosen, and for Hotel reviews[2] Yelp is taken. Both these datasets are available at Kaggle. In the Mobile dataset, 1836 reviews are taken on HTC mobile, whereas, in Hotel reviews, 1000 reviews consider for processing. Figure 2 shows the word cloud for both datasets, which visualized frequent terms present in both datasets.

Figure 2. Frequent terms visualization through a word cloud for Hotel and Mobile dataset

As mentioned in Algorithm-1, optimized LDA configuration use for improved cohesive output, as shown in table 1.

Table 1. Optimized LDA configuration calculated using SA-LDA[10]

Model	Alpha	Beta	Coherence
LDA	0.1	0.1	0.49
SA-LDA	0.415	0.626	0.567

Aspects are ranked based on the topic probability proportion. The sample distribution for the Mobile domain is as follows:

```
Topic0:(0.025*"battery"+0.014*"one"+0.013*"month"+0.013*"screen
"+0.012*"it"+0.012*"problem"+      0.010*"buy" + 0.010*"get" +
0.010*"work" + 0.009*"htc"')
Topic1:(0.015*"camera"+0.014*"phone"+0.010*"good"+0.008*"price
"+0.008*"it"+0.008*"working"+      0.007*"battery" + 0.007*"well"
+ 0.007*"work" + '  '0.007*"htc"')
Topic2:(0.023*"screen"+0.012*"phone"+ 0.009*"battery" +
0.008*"htc" +0.008*"get" +    0.008*"work"      + 0.007*"it" +
0.007*"new" + 0.007*"problem" + 0.007*"would"')
```

In the above example, we have only shown three topics with ten words in each. In the above distribution, we can see that "battery" aspect is coming in multiple topics. We have considered its highest probability value. The same is applied to each aspect. Table 2 represents the top-ranked aspects generated by applying Algorithm-1 for the Mobile domain.

Table 2. Top Aspects Extracted from LDA output using linguistic rules for Mobile

Aspect	Probability
Battery	0.025
Screen	0.023
Camera	0.015
Charge	0.011
Service	0.011
Volume	0.009
Wifi	0.008
Video	0.007
Network	0.007

Figure 3 represents the proportion of the extracted features from LDA output using POS(NN) rules. Similarly, for the Mobile domain, the probability distribution is as follows:

Topic0:(0.023*"room"+0.023*"hotel"+0.015*"stay" +0.008*"one" + 0.008*"would" +0.006*"service"+ 0.006*"time" + 0.005*"staff" + 0.005*"get" + ' '0.005*"night"')
Topic1:(0.033*"room"+0.018*"hotel"+0.015*"staff" + 0.009*"night" + 0.007*"would" +'0.007*"floor" + 0.006*"one" + 0.006*"stay" + 0.006*"time" + 0.005*"bathroom")

Table 3 represents the top-ranked aspects generated by applying Algorithm-1 for the Hotel domain.

Figure 4 represents the proportion of the extracted features from LDA output using the POS(NN) rule for the Hotel.

It is clear from Tables 2 and Table 3 with their visualization that the small aspect list is not able to extract deep insight about the entire dataset. However, one can get a reasonable idea about the domain and its essential features. Now using Algorithm-2, aspect expansion and categorization are performed on both the domains. Table 4 and Table 5 represent the extended aspect list with their categories for both Mobile and Hotel domains, respectively.

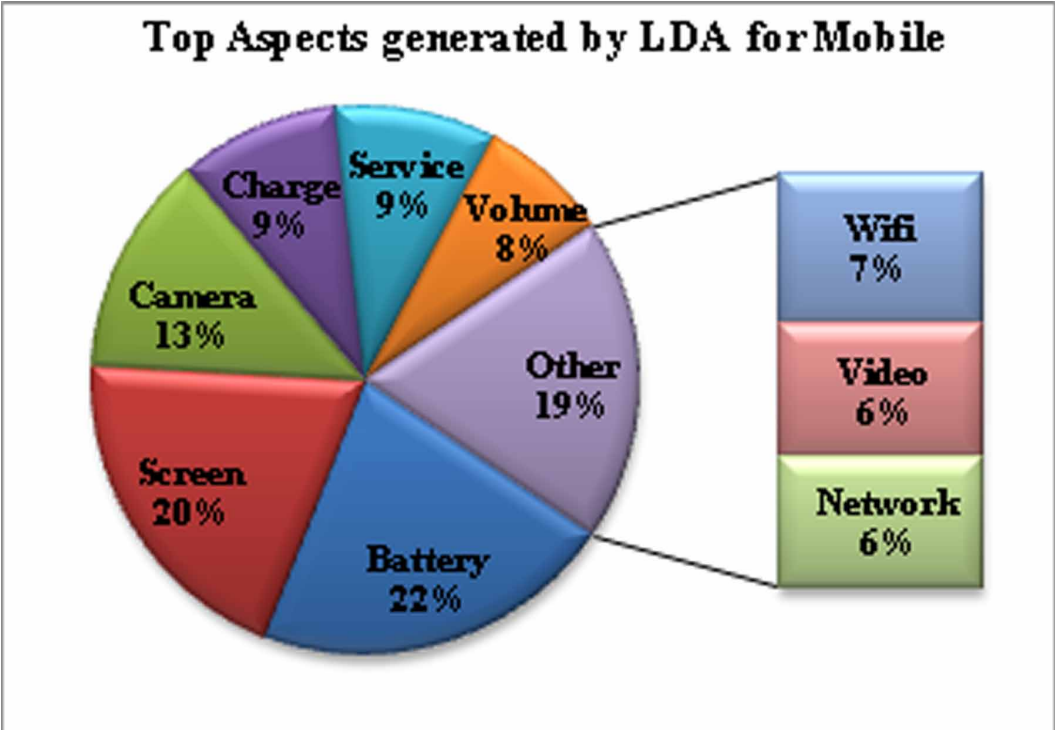**Figure 3. Top aspects generated by LDA for Mobile Domain**

Table 3. Top Aspects Extracted from LDA output using linguistic rules for Hotel

Aspect	Probability
Room	0.033
Staff	0.015
Stay	0.015
Breakfast	0.011
Location	0.011
Clean	0.01
Service	0.006
Bed	0.006
Desk	0.006
Pool	0.005

Now, Algorithm-3 applied to these extended aspect categories represented in Table 5 for ABSA. Table 6 represents the sentiment scores with positive and negative percentage for each category in the Mobile domain.

From Table 6, we get a clear picture of various aspects of the Mobile domain for given reviews. Sometimes the same aspects are represented by different words that can be seen by analyzing Price

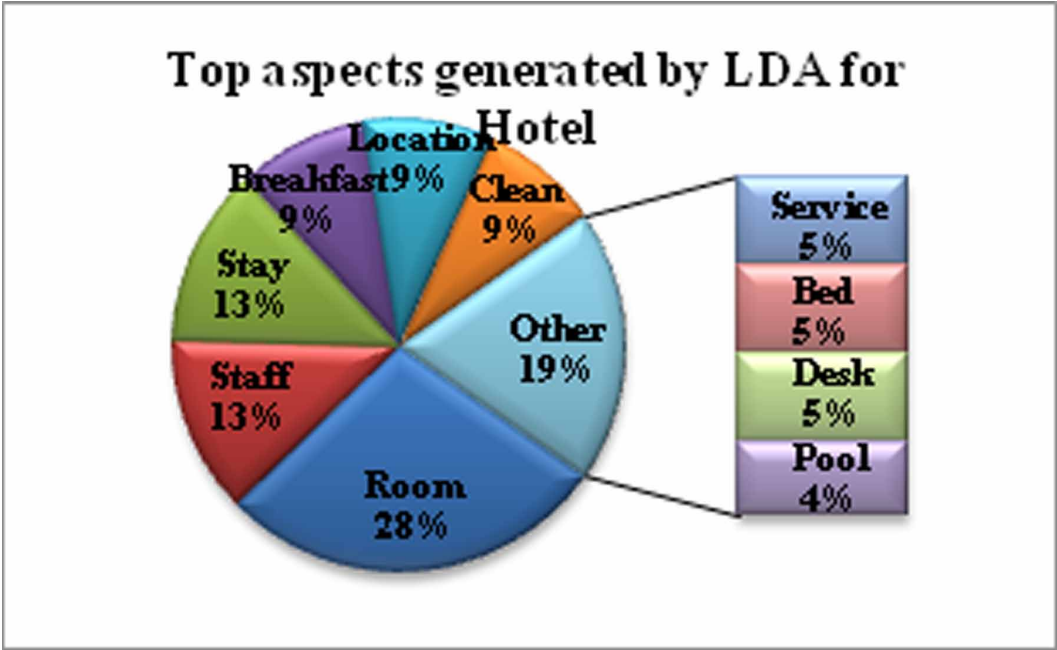**Figure 4. Top aspects generated by LDA for Hotel Domain**

Table 4. Aspects Expansion and Categorization for Mobile

Camera	Picture	Battery	Price	Service
camera	picture	battery	price	wifi
photo	screen	charger	cost	network
lens	video	backup	charge	service
zoom	glass	power	money	volume

Table 5. Aspects Expansion and Categorization for Hotel

Room	Food	Location	Price	Staff
room	food	location	price	staff
bed	breakfast	airport	cost	service
bathroom	restaurant	railway	money	time
water	coffee	city	expensive	desk
clean	meal	pool	charge	luxury

aspects. Price is represented by money and cost. The same is there in other aspect categories also. That is why it is vital to extend the aspect list for more in-depth insight.

We are aggregating the values for each category to represent the average ABSA. Table 7 shows the final sentiment score for the Mobile domain in various aspect categories.

Table 6. Sentiment Scores for each aspects categories with their percentage for Mobile

| Aspect Category	Extended Aspects	Score Positive	Score Negative	% Positive	% Negative
Camera	Camera	135.75	74.375	64.604	35.395
	Photo	15.375	14.375	51.68	48.319
	Lens	0.875	0.25	77.777	22.222
	Zoom	0.125	0.625	16.666	83.333
Screen	Screen	225.125	167.75	57.30	42.70
	Picture	73.5	45.125	61.96	38.04
	Video	20.75	13.25	61.03	38.97
	Glass	7.125	6.625	51.82	48.18
Price	Price	108.25	47.875	69.335	30.664
	Money	42.489	51.261	45.321	54.678
	Cost	11.739	18.636	38.646	61.353
Battery	Battery	232.989	236.011	49.677	50.322
	Charger	62.75	51.375	54.983	45.0164
	Power	34	28.25	54.618	45.381
	Backup	0.375	1.875	16.666	83.333
Service	Charge	122.25	153.375	44.353	55.646
	Service	46.125	49.5	48.235	51.764
	Network	21.5	18.375	53.918	46.081
	Wifi	14.125	12.125	53.8	46.19

Table 7 shows that not only do we get sentiment proportions, but also sentiment scores support its strength or evidence in the reviews. The camera is the most positive aspect of Mobile with 62.93% positive reviews, and service is the most negative with 53.36% negative reviews. The battery is the

Table 7. Sentiment Scores for various aspects categories as a whole for Mobile

Aspect Categories	Score Positive	Score Negative	% Positive	% Negative
Camera	152.125	89.625	62.93	37.07
Screen	326.5	232.75	58.38	41.62
Battery	330.114	317.511	50.97	49.03
Price	162.478	117.772	57.98	42.02
Service	204	233.375	46.64	53.36

most talked-about aspect in reviews with the highest sentiment score. Figure 5 graphically represents the sentiment scores shown in Table 7.

**Figure 5. Aspect Based Sentiment Scores in % for Mobile reviews**
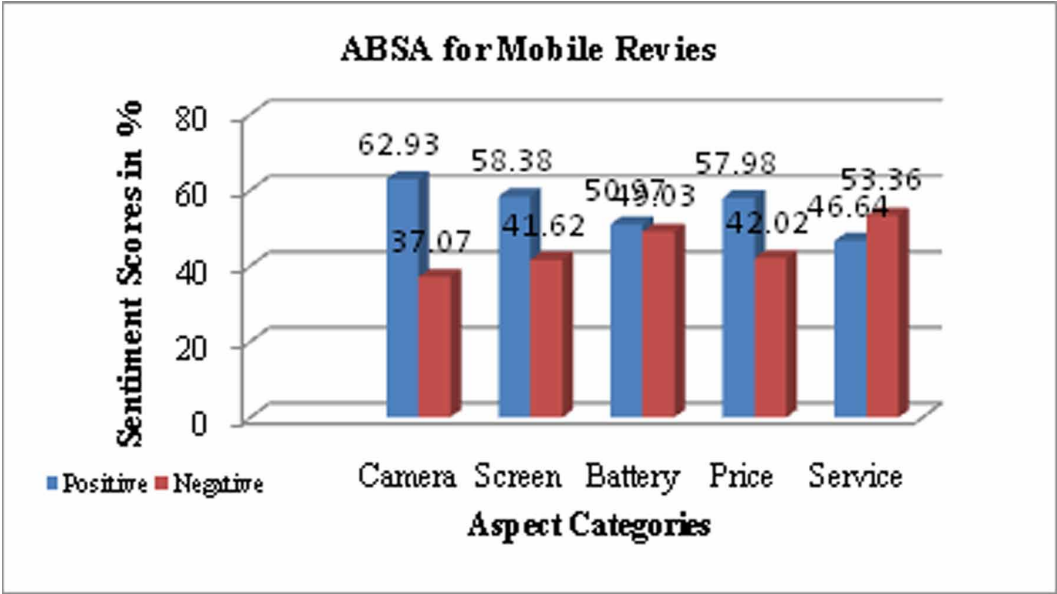
From sentiment scores, we can see how much a particular aspect is discussed in the reviews. Figure 6 represents the positive and negative proportion of the total sentiment scores for each category in the Mobile domain.

Battery 32% and Service 24% in the negative pie chart represent that these are the most discussed aspects of overall negative reviews. Same way, Screen 28%, and Battery 28% are the most discussed aspects in positive reviews. A similar analysis is performed for the Hotel domain.

Now, Algorithm-3 is applied to extended aspect categories presented in Table 5 for ABSA. Table 8 represents the sentiment scores with positive and negative percentage for each category in the Hotel domain.

Figure 6. Aspect Category wise distribution of Sentiment Scores in percentage for Mobile reviews
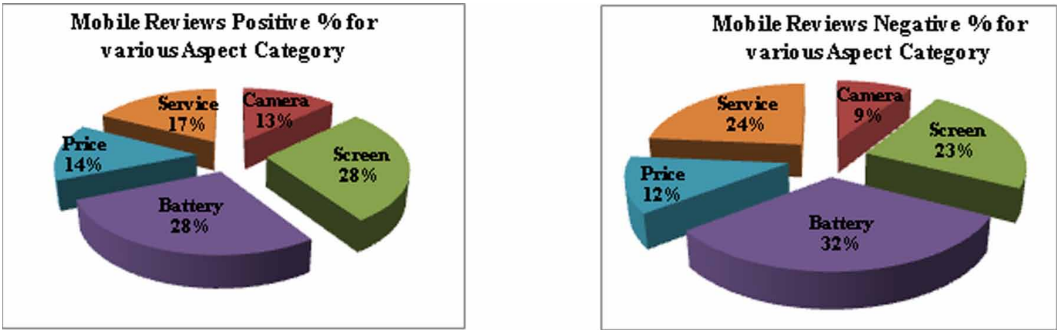
From Table 8, we get a clear picture of various aspects of the Hotel domain for given reviews. Here also the price is represented by money, expensive and cost. We are aggregating the values for each category to represent the average ABSA. Table 9 shows the final sentiment score for the Hotel domain in various aspect categories.

From Table 9, 'Food' is the most positive aspect of the Hotel with 65.78% positive reviews, and 'Room' is the most negative with 48.20% negative reviews. The 'Room' and 'Staff' are the most discussed aspects in reviews with the highest sentiment scores. Figure 7 graphically represents the sentiment scores shown in Table 9.

Figure 8 represents the positive and negative proportion of the total sentiment scores for each category in the Hotel domain.

The pie chart representing the negative sentiment score, 'Room' 44%, and 'Staff' 24% are the most talked aspects in overall negative reviews. In positive reviews, 'Room' and 'Staff' are also the most discussed aspects.

For every Algorithm, it is essential to evaluate the accuracy of the output. Aspect extraction quality evaluated using the coherence value of LDA topic generation that can be seen from table 1. Extended aspects and categorization significantly improve the ABSA, which can visualize in the pie charts, and sentiment scores show the significance of extended aspects.

As labeled reviews are not considered, to evaluate the proposed ABSA algorithm's accuracy, 100 reviews for each domain are selected and manually labeled in terms of sentiment. Twenty reviews per aspect category considered with an equal positive number of positive and negative. Now proposed Algorithm applies to this testing dataset. Table 10 represents the sentiment scores of testing data for each aspect category. The confusion matrix is prepared based on this, and F-score is calculated accordingly.

For test data in the Mobile domain, the F- score lies between 80%-95% for various categories, and the average is 89%. Similarly, for Hotel domain test data, the F- score lies between 77%-84% for various categories, and the average is 80%. If we aggregate both the accuracy results and take the average, we can say the average accuracy or F-score for the presented Algorithm is 84.5%.

Figure 9 shows the graphical representation of the F-score on testing data for the Mobile and Hotel domain.

CONCLUSION AND FUTURE DIRECTION

This work presented a straightforward and efficient ABSA approach. The aspect extraction task performed with a tuned LDA model enriched with POS rule, producing more coherent output. Few domain words used for Aspect extension, and their categorization prepared by domain knowledge. The sentiment classification achieved using the sentiment scores generated by SWN. An evaluation performed on two different datasets shows the effectiveness of the presented approach that is verified with synthetic data. The average accuracy of the proposed approach is 85%. This approach can be used to prepare datasets from unstructured review data, like tweets, and perform sentiment labeling using SWN.

In the future ensemble learning approach can be used for LDA topic generation. Sentiment scoring using SWN can be more accurate by adding a few more rules for dealing with negation, conjunctions, and qualifiers. Conversion of the LDA output into a well-adopted representation by DNN for classification tasks could extend this work.

Table 8. Sentiment Scores for each aspects categories with their percentage for Hotel

Aspect Category	Extended Aspects	Score Positive	Score Negative	% Positive	% Negative
Room	Room	1286.375	1173.125	52.302	47.69
	Clean	548.625	439.5	55.52	44.47
	Bed	294.25	351	45.6	54.39
	Bathroom	163.25	172.25	48.65	51.34
	Water	66.75	59.375	52.92	47.07
Food	Food	134.055	59.069	69.41	30.58
	Breakfast	273.125	125.375	68.53	31.46
	Restaurant	243.29	136.955	63.982	36.01
	Coffee	77.75	55.625	58.29	41.7
	Meal	18.625	11.5	61.82	38.17
Location	Location	420.556	204.569	67.27	32.72
	Pool	114.375	68.125	62.67	37.32
	City	97.875	73.875	56.98	43.01
	Airport	46.875	34.625	57.51	42.48
Price	Price	181.18	108.694	62.503	37.49
	Expensive	78	51.75	60.11	39.88
	Money	51.875	41.5	55.55	44.44
	Cost	16.25	24.25	40.12	59.87
Staff	Staff	543.364	291.88	65.05	34.94
	Service	312.972	206.15	60.28	39.71
	Time	262.5	200.375	56.71	43.289
	Desk	138.5	120.625	53.44	46.55
	Luxury	18.5	18.5	50	50

Table 9. Sentiment Scores for various aspects categories as a whole for Hotel

Aspect Category	Score Positive	Score Negative	% Positive	% Negative
Room	2359.25	2195.25	51.80	48.20
Staff	1275.84	837.53	60.37	39.63
Food	746.85	388.52	65.78	34.22
Location	679.68	381.19	64.07	35.93
Price	327.31	226.19	59.13	40.87

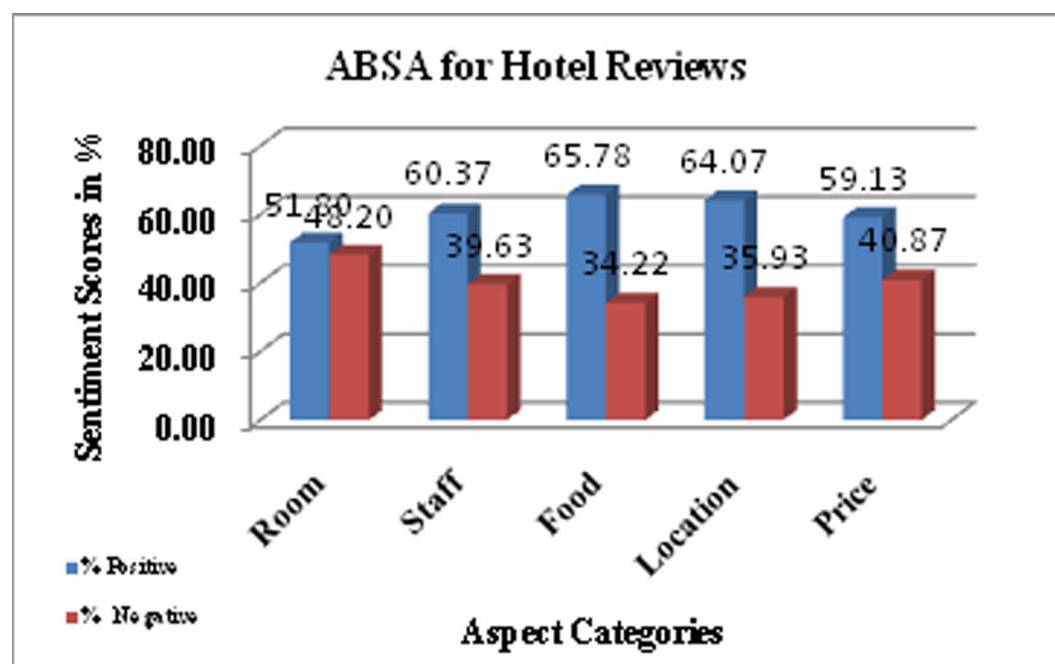**Figure 7. Aspect Based Sentiment Scores in % for Hotel reviews**

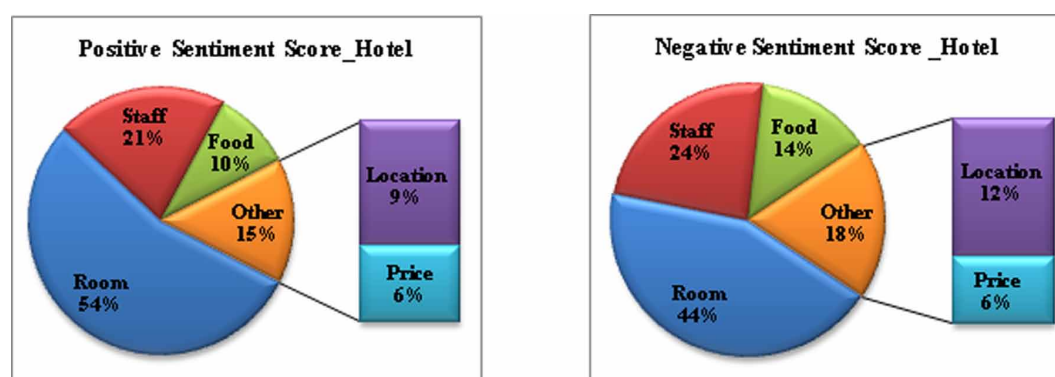**Figure 8. Aspect Category wise distribution of Sentiment Scores in percentage for Hotel reviews**

Table 10. Sentiment Scores for Testing Data for Mobile and Hotel Domains

Domain	Aspect Category	Sentiment Score (+)	Sentiment Score (-)	F-Score
Mobile	Camera	21.5	25.375	0.92
	Picture	40	60	0.80
	Battery	29.25	26.125	0.95
	Price	14.75	17.125	0.93
	Service	18.75	13.5	0.86
	Average	24.85	28.425	0.89
Hotel	Room	5.5	3.125	0.78
	Food	15.125	8.125	0.77
	Location	10.375	6.75	0.83
	Price	6.75	4.125	0.81
	Service	14.875	10.125	0.84
	Average	10.525	6.45	0.80

Figure 9. F-score for testing data of various Aspect Categories Mobile and Hotel reviews
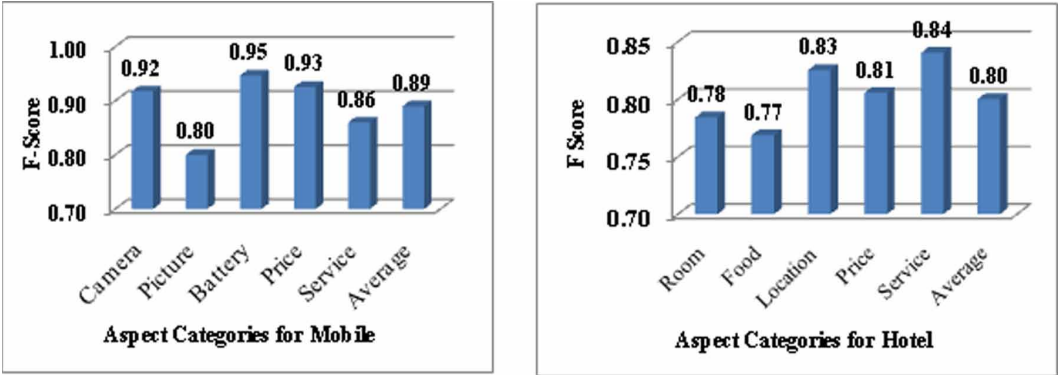
REFERENCES

Akhtar, N., Zubair, N., Kumar, A., & Ahmad, T. (2017). Aspect based sentiment oriented summarization of hotel reviews. *Procedia Computer Science*, *115*, 563–571. doi:10.1016/j.procs.2017.09.115

Anoop, V. S., & Asharaf, S. (2018). Aspect-Oriented Sentiment Analysis: A Topic Modeling-Powered Approach. *Journal of Intelligent Systems*, *29*(1), 1166–1178. doi:10.1515/jisys-2018-0299

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of LREC*, *10*, 2200–2204.

Bhargava, R., Arora, S., & Sharma, Y. (2019). Neural network-based architecture for sentiment analysis in Indian languages. *Journal of Intelligent Systems*, *28*(3), 361–375. doi:10.1515/jisys-2017-0398

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Dieng, A. B., Ruiz, F. J., & Blei, D. M. (2020). Topic modeling in embedding spaces. *Transactions of the Association for Computational Linguistics*, *8*, 439–453. doi:10.1162/tacl_a_00325

Do, H. H., Prasad, P. W. C., Maag, A., & Alsadoon, A. (2019). Deep learning for aspect based sentiment analysis: A comparative review. *Expert Systems with Applications*, *118*, 272–299. doi:10.1016/j.eswa.2018.10.003

Ekinci, E., & İlhan Omurca, S. (2020). Concept-LDA: Incorporating Babelfy into LDA for aspect extraction. *Journal of Information Science*, *46*(3), 406–418. doi:10.1177/0165551519845854

Esuli, A., & Sebastiani, F. (2006). Sentiwordnet: A publicly available lexical resource for opinion mining. *Proceedings of LREC*, *6*, 417–422.

Jansson, P., & Liu, S. (2017). Topic modeling enriched LSTM models for the detection of novel and emerging named entities from social media. In *Proceedings of IEEE International Conference on Big Data* (pp. 4329-4336). IEEE.

Jo, Y., Lee, L., & Palaskar, S. (2017). *Combining LSTM and latent topic modeling for mortality prediction*. arXiv preprint arXiv:1709.02842.

Ma, Y., Peng, H., & Cambria, E. (2018). Targeted Aspect based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM. In *Proceedings of AAAI* (pp. 5876-5883). AAAI.

Pathik, N., & Shukla, P. (2020). Simulated Annealing Based Algorithm for Tuning LDA Hyper Parameters. In *Proceedings of Soft Computing: Theories and Applications* (pp. 515–521). Springer. doi:10.1007/978-981-15-4032-5_47

Pergola, G., Gui, L., & He, Y. (2019). TDAM: A topic-dependent attention model for sentiment analysis. *Information Processing & Management*, *56*(6), 102084. doi:10.1016/j.ipm.2019.102084

Sokhin, T., & Butakov, N. (2018). Semi-automatic sentiment analysis based on topic modeling. *Procedia Computer Science*, *136*, 284–292. doi:10.1016/j.procs.2018.08.286

Tai, C. H., Tan, Z. H., Lin, Y. S., & Chang, Y. S. (2015). Mental disorder detection and measurement using latent Dirichlet allocation and SentiWordNet. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1215-1220). IEEE. doi:10.1109/SMC.2015.217

Tang, F., Fu, L., Yao, B., & Xu, W. (2019). Aspect based fine-grained sentiment analysis for online reviews. *Information Sciences*, *488*, 190–204. doi:10.1016/j.ins.2019.02.064

Wang, X., Wang, J., Yang, Y., & Duan, J. (2017). Labeled LDA-Kernel SVM: A Short Chinese Text Supervised Classification Based on Sina Weibo. In *Proceedings of 4th International Conference on Information Science and Control Engineering (ICISCE)* (pp. 428-432). IEEE. doi:10.1109/ICISCE.2017.96

Yiran, Y., & Srivastava, S. (2019). Aspect based Sentiment Analysis on mobile phone reviews with LDA. In *Proceedings of the 4th International Conference on Machine Learning Technologies* (pp. 101-105). doi:10.1145/3340997.3341012

Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. *Knowledge and Information Systems*, *60*(2), 1–47. doi:10.1007/s10115-018-1236-4

ENDNOTES

1.   Dataset for Mobile is taken from Kaggle from the following link: https://www.kaggle.com/PromptCloudHQ/amazon-reviews-unlocked-mobile-phones
2.   Dataset for Hotel is also taken from Kaggle from the following link: https://www.kaggle.com/yelp-dataset/yelp-dataset