JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Survey on Diverse Techniques for Sentiment Analysis

[1]Mr.M.YUVARAJA
[1]Ph.d Research Scholar & Assistant Professor
[1]Department of Computer Science
[1]Dr.N.G.P Arts and Science College
[1]Coimbatore-641048
[1]Tamilnadu, India.
[1]yuvaraja0105@gmail.com

[2]DR.R.NITHYA.,M.Sc(CS).,Ph.D., M.B.A
[2]Professor, Department of Computer Technology
[2]Dr.N.G.P Arts and Science College
[2]Coimbatore-641048
[2]Tamilnadu, India.
[2]rnithya@drngpasc.ac.in

Abstract: Recently, Sentiment Analysis (SA) is gaining much importance wherein, sentiments are gathered, examined and combined from text. This has led to the inclusion of various subdomains, each capable of dealing with varied level of examination. In this paper, Aspect-Level SA (ALSA) is focussed on which the main goal is to find and collect sentiments as well as entities in documents. An in-depth summary of the prevailing contemporary schemes which portray great development in assessing the opinion or target related to an entity and the linked sentiment are presented. Diverse AL analysis schemes based on frequency, syntax, supervised as well as unsupervised Machine Learning (ML), Deep Learning (DL) and hybrid schemes are presented. The future possibility of research in AL analysis is also presented.

1. Introduction

The quantity of data upsurges owing to the increased use of web-based applications and sharing of data. As more people use internet, the amount of blogs based on opinion and review show a radical rise. Data sources are built by amassing the blog information based on opinion and review. Organisations are determined in gaining knowledge on customer opinions on products. People deliberately post their opinions and ideas on the web. In contradiction of the past, social media acts as a podium for rapidly sharing the criticism by taking item reviews. Reviews have an impact on individuals to a better degree as customers verify the item scores before buying a product. Depending on customer reviews, the producer can enhance his trade by looking for comments that includes the likes as well as dislikes of the product. The product may be endorsed depending on reviews, thus improving the sales. Through SA, financial market can develop by improving the marketing. The survey offers an outline of descriptions by Pang & Lee (2008) on ALSA. A combination of Natural Language Processing (NLP), Information Retrieval (IR) and Artificial Intelligence (AI) incline to drive SA. 'Opinion mining' in data mining and IR community is employed for identifying opinions of public on a

specific subject. SA from NLP has an objective to bring sentiments from the text. 'Subjective analysis' is a combination of opinion mining and SA. Opinion Mining, SA and subjective analysis examine the existence of opinion, evaluation, sentiment, review, emotion and attitude. For better understanding, they are often stated as sentiment or opinion, although they are different.

Classification focusses on exactly predicting the objective class of unlabelled data by obtaining information from instances well-defined by a group of features along with a label (Sreeja & Sankar 2015). As emotions exist, Polarity is involved in characterizing sentiments into Positive, Negative as well as Neutral. To label positive as well as negative sentiments noticeably, several binary classification schemes are used.

Sentimental Analysis (SA) includes the following steps namely, Detection, Categorisation and Aggregation. It may not be vital to use the order of operations.

1.2 Motivation of the Survey

Sentiment Analysis (SA) is highly concentrated in this study. Document, sentence and aspect are the three levels of SA. The sentiment depends on a single issue at the document level. The statement should only express feelings about a topic at the second level, and it can be untrue in some situations. The entity is considered and sentiment is easily recognised in ALSA. The fundamental goal of ALSA is to locate sentimental-target pairs, and determine the sentiment by looking at the attributes of every entity. A comprehensive examination of text is performed using this method. The focus of this survey is mostly on ALSA, as it pertains to several research fields.

Pang & Lee (2008) have addressed ALSA in depth, covering tactics, methods, ethics, practicality, and theory. Tsytsarau & Palpanas (2012) have dealt with Document Level Sentimental Analysis (DLSA) that identifies sentiments and categorises them as Machine Learning (ML), statistical, semantic or dictionary based. In his survey, Liu (2012) have offered an improved form of SA. It includes ALSA, which handles implicit and explicit statements and incorporates aspect identification. Tang et al. (2015) have conducted a brief study on DLSA, which included and analysed consumer reviews. To increase classification accuracy, many surveys are conducted.

In this survey, initially, ALSA evaluation and the schemes of aspect identification and Sentiment Analysis (SA) are detailed. Following that, relevant concerns with these approaches and their remedies are examined, along with sentiment methods and aspect sentiment scores. This work is ended such that ALSA is thoroughly examined, as well as the research directions.

1.3 Reviews on Opinion Mechanisms

To get a better view of identification of aspect, several schemes are involved. Several schemes in aspect identification are detailed. The survey is detailed with appropriate examples. The assessment schemes and

the scores are detailed. Aspect identification schemes are classified based on frequency, syntax, supervised as well as unsupervised ML, DL and hybrid schemes.

1.3.1 Frequency based Models

Fundamentally, only some words called are used often when compared to other words. This method is effective as several schemes involved in the identification of aspects depend on this. In case of consumer reviews, only a few nouns are considered. A product may be considered along with the reviews of the product. There are chances for missing some features as they may unnoticed. To deal with this issue, frequency based schemes may be extended by using rules.

Hu & Liu (2004) have propounded a scheme, wherein customer views on a product are considered and précised. The reviews focussing on only particular features are considered and categorized into positive or negative. In case of traditional summarization scheme, the reviews are précised and rephrased. Several schemes are proposed to extract opinions. Initially, single as well as compound nouns are taken as aspects and noun frequencies are totalled. The authors have used techniques based on data mining and NLP for better performance. Parts-of-Speech (PoS) tagging plays a dominant role in finding the noun phrases. Feature extraction comprises of PoS tagging, common feature production, pruning, opinion word mining and identification of uncommon features. NLProcessor is involved in examining the complete sentence and finding the type of word. Association Rule Mining (ARM) is employed in identifying the common features using Apriori algorithm. The frequent item sets are found using minimum support and rules. Feature pruning finds the unemployed association rules and removes them depending on compactness and redundancy pruning. ARM focusses on establishing rules, wherein some may not be beneficial. An adjective that is present near the frequent feature is an 'Opinion'. Authors have focussed on only obvious aspects.

Liu et al (2005) have focussed on reviews and an architecture named 'Opinion Observer' is propounded to associate the reviews and examine them. Pros and cons of an item can be identified. The format is of 3 types. The review is examined as the scheme depends on NLP as well as supervised pattern detection. The reviews from websites are considered, followed by analysis to identify features and opinions. The language pattern is identified, a supervised pattern matching scheme is propounded. NLProcessor linguistic parser is employed for producing PoS. In the replacement process, the feature is substituted with the feature. Classification Based on Associations (CBA) is involved in mining rules, and minimum support and confidence are computed. Considering nouns as well as compound nouns as aspects is a shortcoming. Scaffidi et al (2007) have proposed a search system for identifying aspects depending on precision and efficiency of feature mining, precision of product scores and assessed time. The CBA scheme is involved for feature mining, and WordNet for product scoring.

ARM is employed to detect implicit aspects (Hai et al 2011). Depending on sentiment words, ARs are used to identify the aspects. Common item collections are found from transactional databases. Rules are produced

from co-occurrence matrix of sentiment related words as well as explicit aspects. Explicit features followed by implicit features are to be found. To reduce false positives, nouns with close combinations are found. Single and compound nouns containing similar meaning but diverse formats are found. The number of likes and dislikes are not found. Co-occurrence ARM is propounded to deal with this problem. Co-occurrence matrix is built depending on the frequency of opinions and features, and rules are extracted. Depending on the support and confidence, weak rules are pruned. Secondly, to produce solid rules for every opinion, the advanced rules are grouped into feature groups.

Duric & Song (2012) have designed a collection of feature selection mechanisms that involve content and syntax model to study a collection of attributes by splitting the items that are studied from particular expressions that designate them based on polarities. By concentrating on expressions and disregarding items, outstanding features are chosen for DLSA. The outcomes from attributes in a Maximum Entropy (ME) classifier are inexpensive with contemporary ML schemes.Wang & Li (2015) have dealt with the issue of understanding sentiments from more number of images depending on features and background information. Notwithstanding the advances in examining sentiment depending on the text, sentiments after an image content is overlooked. The momentous developments in text-dependent sentiment forecast tasks to improve the process of forecasting the fundamental sentiments in images. Only visual or textual features are capable of precise sentiment labelling. An optimization scheme for determining local-optima is proposed.

Akhtar et al (2017) have designed a technique for uniting DL and standard feature dependent models using a Multi-Layer Perceptron (MLP) based network for analysing financial sentiments. DL models are trained over pre-trained, auto encoder-dependent embeddings of financial words and lexicon features. Ensembles are built by merging the DL and a supervised model is built on Support Vector Regression (SVR). The propounded scheme is assessed on SemEval-2017 shared task. Ding et al (2018) have performed entity-based SA. A dataset with several issue comments from open source projects is constructed. SentiSW is an entity-level SA tool including classification of sentiments and identification of entities is designed. Ahuja et al (2019) have analysed the influence of Term Frequency-Inverse Document Frequency (TF-IDF) word level as well as N-Gram on the SS-Tweet dataset. It is seen that, TF-IDF is better in contrast to N-gram. Dahooie et al (2021) have focussed on narrowing the gaps by designing a framework based on SA and Multi-Criteria Decision-Making (MCDM) schemes using Intuitionistic Fuzzy Sets (IFS). The designed model is employed to rank mobile phones for showing the accessibility and usefulness of the scheme. Sensitivity is analysed to find the best scheme and effective characteristics.

1.3.2 Syntax based Models

In syntax based schemes, the word frequency is not considered for finding the aspects. Syntax based associations are employed for aspect recognition and the adjective modifier association is determined. Less frequent words can be found. Several syntactical relations prevail.

Jakob & Gurevych (2010) have focussed on opinion mining using Conditional Random Field (CRF), wherein the domain likelihood issue is addressed. Evaluation is performed in 3 areas and performance is assessed. Movie database is analysed using supervised scheme. The main demerit of the scheme is that it demands more amount of training to obtain improved results, and the results are related for all databases. It finds whether word and sentiment include direct dependences. It finds the nouns that are closer to sentiment expression. It verifies and finds the word with sentiment. Zhao et al (2010) have developed a model that includes features like intuitive heuristics and syntactic structure likeness. The nouns are taken as characteristics, and convolutional tree kernel is employed. The authors have compared several techniques. The noise ratio rises with domain changes. Zhang et al. (2010) have propounded a scheme named double propagation, wherein noun attributes are mined. It is appropriate for Corpa with average size. For those that are smaller or larger, it involves reduced precision and recall. To improve recall, improvements depending on part-whole and 'no' patterns are necessary. Features are graded to improve the precision of the best candidates. Features are graded based on the feature importance using the implication and frequency of features.

Qiu et al. (2011) have propounded a scheme wherein some sentiments as well as features are mined by involving seed sentiment lexicon. From this, sentiments and features are mined. This is repeated till no more sentimental words are freshly added. Polarity is assigned and rules are built depending on relations. Di Caro & Grella (2013) have developed an algorithm that is dependent on a collection of propagation rules at syntactic level in a parse tree showing dependencies. A context-dependent model with sentiments tuned based on a context is built. SentiVis, a system is designed to implement the ideas using an orthogonal scheme for data visualization. The sentiments are graphically represented in a 2D space. Zhou & Song (2015) have used a widespread topic and syntax based model named PoS LDA (POSLDA). They have proposed many feature assortment schemes for separating entities from modifiers that label entities. Along with ME classifier, the chosen features are employed at document as well as aspect levels. The semantic and syntactic classes are automatically separated, and can be applied to ALSA by relating topics with aspects. Noun-based classes considered as semantic classes must be eliminated to the maximum extent possible so as to lessen their influence on SA.

Word embeddings proficient in finding semantic and syntactic related info from contexts are widely employed for diverse NLP related tasks. Nevertheless, the prevailing schemes for learning context-dependent embeddings are classically incapable of taking adequate sentiment related information. This may lead to words with comparable vector illustrations have a conflicting polarity, thus reducing performance. Yu et al (2017) have propounded a word vector based refinement model which can be implemented on pre-trained vectors. The model adjusts the word vectors, such that they are nearer to semantically and sentimentally alike words and away from dissimilar ones. Naseem et al (2019) have focussed on hybrid word illustration and Bi-directional Long Short Term Memory (BiLSTM) by considering the noise in the text. Syntax, Polysemy,

semantics, Out Of Vocabulary (OOV) and word sentiments are considered. Dai et al (2021) have designed Aspect-Based SA (ABSA) to predict the aspect polarities. Pre-Trained Models (PTMs) should include adequate syntax based information. The trees from PTMs and parsing trees showing dependency applied on several models show that the tree from RoBERTa (FT-RoBERTa) is better as it is sentiment-word-oriented.

1.3.3 Supervised ML based Models

They are limited ML based schemes that are involved in aspect recognition. The power of ML based schemes is dependent on features that are used. The frequency based features are capable of producing several exceptional ones that offer good outcomes when compared to the PoS features. Zhuang et al (2006) have focussed on movie review based datasets along with the domain knowledge that include statistics, movie and words in analysis. Opinion based words, features and review polarity are dealt by involving certain rules and relations. The dependence relation permits obtaining effective outcomes. Depending on the list of keywords, the features are classified into frequent and infrequent items. Infrequent features are not identified. Kessler & Nicolov (2009) have focussed on the data from blogs that include details of a car as well as digital camera. The sentences are tokenized and split, followed by PoS verification using Support Vector Machine (SVM) as well as parsing. The propounded scheme determines the semantic associations amid the expressions and targets. Proximity, heuristic syntax, bloom and SVM are detailed. Though they are not language dependent, sentiment based expressions may not be found thus reducing performance.

Toprak et al. (2010) have designed an annotation mechanism that includes 2 levels i.e., sentence and expression levels. The properties of the sentences are verified at the expression level. This mechanism can be applied to any kind. The authors have considered the reviews from only a few websites. Liu et al (2012) have dealt with noisy labels. It is essential to use physically labelled and noisy labelled data for the purpose of training. To impeccably incorporate these data into a framework is a challenge. The authors have presented Emoticon Smoothed Language Model (ESLAM) for training a language model using physically labelled data. Noisy emoticon data is used for smoothing. Anjaria & Guddeti (2014) have focussed on sentiment forecast over Twitter using ML based schemes by considering social network structure like retweet. Both direct and enhanced terms connected with the event and the effects are considered. Supervised ML based schemes like SVM, Naive Bayes (NB), ME and Artificial Neural Networks (ANNs) to categorize data by using Unigram, Bigram and hybrid model for US Presidential Elections in 2012.

Ghiassi et al (2016) have propounded a scheme to focus on the challenges related to the unique features of Twitter language and brand-dependent class distribution. The efficiency related to 2 unique brands are discussed. Tweet feature representations with 7 dimensions and increased feature density are presented. The reduced dimensionality decreases the classification complexity and feature sparsity. Schouten et al (2017) have designed a text processing architecture that is capable of gathering reviews. The common classes of aspects in review sentences are to be found. An unsupervised scheme that uses ARM on co-occurrence

frequency data from corpus is used to determine the categories. Many ML classifiers are assessed using the propounded feature set (Rustam et al 2021). The set is built by combining BoW and TF-IDF. LSTM architecture is also applied. It is seen that the Trees Classifiers perform well by employing the concatenated feature set.

1.3.4 Unsupervised ML based Models

Hofmann (2000) has dealt with the scheme that finds the likeness amid the text documents from corpus without domain related information. Architecture is designed to determine the likenesses amid documents by involving some values that help in constructing relations amid information reclamation, text learning as well as statistics. A model is built by taking the word frequency and analysed. Posterior likelihoods of terms are computed and Fisher Kernel is employed to determine the log likelihoods. The association amid supervised and unsupervised scheme is produced by employing productive models that supports unlabelled categorization. To support segmentation of unstructured text, Blei & Moreno (2001) have propounded a model. It is constructed using Hidden Markov Model (HMM), wherein aspect model is used. The values are roughly found deprived of any computation that may disturb the model's performance. As the model does not involve any syntactic information, it is easy for segmenting the output. The transition points of the topic are presumed. Contrary to HMM, improved results are not obtained for huge windows.

Titov & McDonald (2008) have designed topic modelling schemes like Latent Dirichlet assignment and Probabilistic Latent SA for mining outputs with multiple grains that are quantitively and qualitatively assessed to designate the enhanced performance when compared to typical models. The reviews are considered and classified. This model is proposed for classifying sentiment, text and opinion. This model performs sentiment scoring for every aspect. It incorporates 2 stages. In the first stage, aspects are mined and sentiments are categorized. In the propounded scheme, the initial stage is enhanced using supervised models. Long et al (2010) have designed Supervised Joint Aspect and Sentiment Model (SJASM), an aspect and sentiment based model from reviews to aid customers in taking decisions. The review features are assessed by associating the performance of SLDA, Joint sentiment topic (JST) along with the Linear Regression (LR) model. A supervised approach is propounded by Lu et al (2011) using the least momentous knowledge that determines the link amid topics and aspects. SA with multiple aspects includes aspects in a review of a customer. The models include multi-aspect labelling of sentences and forecasting of rating. The models are associated and variances are forecast. The weak supervised topic models are improved on tasks depending on multi-aspect based labelling of sentences.

An automatic productive model propounded by Wang et al. (2011) is an improved version of Latent Aspect Rating Analysis (LARA). The user finds the aspect physically and it is challenging for more number of reviews. In LARA, ratings and aspect finding are performed concurrently. Bag of Words (BoW) algorithm is employed. Segmentation is not effectively performed forming imprecise segments.As more amount of

emotions are present in the social media, Hu et al (2012) have considered unsupervised SA using emotional signals. They have focussed on determining whether the signals are capable of supporting SA by offering a combined method to model emotion sign and correlation. The signals are included into an unsupervised learning architecture. Pavlopoulos & Androutsopoulos (2014) have dealt with Aspect Term Extraction (ATE), the processing phases of ABSA for extracting term naming aspects. They have made 3 ATE datasets available. Evaluation measures are introduced. Unsupervised ATE can be enhanced by using incessant space vector illustrations made up of words along with phrases.Suresh (2016) has propounded a fuzzy based clustering model for analysing twitter feeds related to sentiments of a specific brand. A comparative study is done using the prevailing partitioning clustering schemes. García-Pablos et al (2018) have designed W2VLDA, an unsupervised system that depends on topic modelling with few unsupervised mechanisms. The configuration step performs classification of aspects, separation of aspect-term and opinion-word along with the classification of sentiment polarity pertaining to any area and language. Yadav & Chakraborty (2020) have introduced schemes that involve diverse types of multi and cross-lingual based embeddings for proficiently transfering knowledge from monolingual to code-mixed text.

1.3.5 Hybrid Methods

Popescu & Etzioni (2007) have propounded Opine, a scheme defined to extract features from reviews. Relaxation labelling is used to identify robust outcomes and polarity. The problem is broken into sub-tasks. To assess the model, PMI++ and HU++ are used. Opine offers better outcomes using web corpus for finding product features.

Summarization of aspect dependent sentiment is challenging (Blair-Goldensohn et al 2008). Reviews are considered, text is mined and sentiments are classified. It incorporates lexicon and ML algorithms. Aspects are extracted, followed by accumulation and summarization. During classification, the raw-score along with purity are determined. Raju et al. (2009) have propounded an unsupervised and domain free scheme to inevitably determine the features from reviews. The noun phrases are clustered and attributes are found in several phases namely, Pre-processing, clustering and attribute mining. During pre-processing, the noun phrases are mined and pruning is performed. During clustering, the similarity measure is found, followed by clustering and pruning. During attribute extraction, N-grams are obtained from the attribute. Manual data is created and reduced, thus reducing the Recall value.

Yu et al. (2011) have propounded an aspect ranking scheme to find the principal features by using a dependency parser. They are classified using a sentiment based classifier. The problem is determined, followed by recognizing and categorizing features and assigning ranks. The scheme is analysed using F-Measure and is related to frequency, association and hybrid schemes. Ghiassi et al (2013) have a scheme for reducing supervised features involving n-grams and statistical study to design a Twitter-based lexicon. The smaller lexicon set involves lesser modeling complexity, thus maintaining an increased level of coverage

over Twitter corpus, offering enhanced classification. DAN2 offers better classification than SVM for the same Twitter-specific lexicon. Brun et al (2014) have a system for SemEval2014 Task 4 for ABSA. The system involves a parser that offers information to offer linguistic features to diverse classifiers devoted to aspect classes and polarity grouping. Appel et al (2016) have propounded a hybrid scheme to SA at sentence level. The scheme involves NLP based schemes, a lexicon improved using SentiWordNet and fuzzy sets to assess semantic positioning polarity and intensity that offers the basis for finding sentiments. The propounded hybrid scheme is applied to diverse datasets and outcomes attained are likened to the ones got using NB and ME based schemes. It is seen that the proposed scheme is more precise. Ma et al (2018) have propounded a knowledge-based solution to ABSA focusing on leveraging the understanding in deep neural sequential model. To openly model the implication of reliant sentiment, LSTM is augmented with a stacked scheme. To incorporate the open with implied knowledge, Sentic LSTM is proposed that comprises of an output gate that incorporates token based memory and concept based input. LSTM and recurrent additive network are linked to simulate sentic patterns. Iqbal et al (2019) have propounded a unified framework that links the gap amid lexicon and ML based schemes to attain improved accuracy. To deal with scalability, a Genetic Algorithm (GA)-based feature lessening scheme is propounded. Kumar et al (2020) have designed a hybrid DL model for predicting sentiments in multi-modal data. It strengthens DL nets with ML to handle textual and visual data, along with their combination. The propounded circumstantial ConvNet-SVMBoVW model involves discretization, analysis of text and image along with decision module. SVM classifier trained with Bag-of-visual-words (BoVW) is employed for forecasting the sentiment in visual content. A Boolean decision module is added to validate and categorize the output.

1.3.5 Deep Learning based Models

Chen et al (2014) have focussed on visual sentiment classification using deep CNNs. These concepts are Adjective Noun Pairs (ANPs) found from tags of photos, and can be used as efficient statistical signs for identifying emotions in images. A model of deep CNNs that shows better performance enhancement on categorizing largescale image dataset is proposed. The model is trained using Caffe, a DL architecture. To focus on the biased training that includes images with better sentiment and stop overfitting, a model with weights trained from ImageNet is employed. Poria et al (2015) have presented feature mining from small texts depending on activation of inner layer of deep CNN. The mined features in multimodal SA of small video clips signify a sentence each. The shared vectors including textual, visual as well as audio modalities for training a classifier depending on several kernel learning is said to be good on assorted data. A deep CNN is applied on it. The values from hidden layer for a more innovative classifier offer better accuracy.

Chen et al (2017) have propounded a framework using NNs to find the opinion sentiments in a comment. The framework approves multiple attention schemes to find features separated by an extended distance, it is robust against inappropriate data. The weighted scheme overcomes labour exhaustive feature handling, also

offers a perfect memory for diverse opinion targets. Jangid et al (2018) have performed ABSA on microblogs and captions of financial area. A multi-channel CNN are proposed for SA and RNN with bi-directional LSTM to mine aspect from a headline or microblog.

Arora & Kansal (2019) have propounded text standardization with deep convolutional character based embedding NN model for SA of unstructured data. Noisy sentences are treated for sentiment detection and irrelevant memory in word based learning. Precise SA of unstructured data is done. The initial processing phase for carrying out normalization comprises of subsequent steps namely, tokenization, Out Of Vocabulary (OOV) recognition and replacement, lemmatization accompanied by stemming. Character-dependent embedding in CNN is an effectual scheme for SA which involves fewer learnable factors in representation. The propounded scheme deals with sentiment standardization and categorisation of formless sentences. Dashtipour et al (2020) have propounded a hybrid architecture for concept based SA in Persian linguistic which incorporates rules and DL to enhance polarity discovery. Once a pattern is activated, the architecture permits sentiments to move from words to concepts depending on relation of symbolic dependency. When patterns are not activated, the architecture switches to sub-symbolic counterpart and influences DNN for classification. Dang et al (2020) Public opinion provides the valued data. Sentiments are analysed on social networks and is a dominant source of learning about users' opinions. Nevertheless, the competence and accuracy of SA is stalled by challenges met in NLP. It is established that DL models are likely solutions to challenges of NLP. TF-IDF and word embedding are applied on a sequence of datasets.

1.4 Summary

This paper has presented the complete review on diverse types of ALSA in the literature over the past. This confirms that aspect-level study is an optimistic change from word-dependent mechanisms to semantically concept based mechanisms. Recent and substantial analysis schemes are based on Frequency, syntax, Supervised and Unsupervised ML, Deep Learning (DL) and hybrid. The merit of aspect-level study improves the likelihood of dealing with the issue intrinsic in compound language structures as it presents improved leverage in the prevailing human framed knowledge bases. Aspect-level analysis provides a means for enormous quantity of application areas that include the details from the information learned from word-dependent mechanisms and semantically effective concept centric-schemes.

References

1. Sreeja, N. K., & Sankar, A. (2015). Pattern matching based classification using ant colony optimization based feature selection. Applied Soft Computing, 31, 91-102.
2. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis (foundations and trends (R) in Information Retrieval).

3. Tsytsarau, M., & Palpanas, T. (2012). Survey on mining subjective data on the web. Data Mining and Knowledge Discovery, 24(3), 478-514.
4. Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.
5. Tang, D., Qin, B., & Liu, T. (2015, July). Learning semantic representations of users and products for document level sentiment classification. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers) (pp. 1014-1023).
6. Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 168-177).
7. Liu, B., Hu, M., & Cheng, J. (2005, May). Opinion observer: analyzing and comparing opinions on the web. In Proceedings of the 14th international conference on World Wide Web (pp. 342-351).
8. Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H., & Jin, C. (2007, June). Red Opal: product-feature scoring from reviews. In Proceedings of the 8th ACM conference on Electronic commerce (pp. 182-191).
9. Hai, Z., Chang, K., & Kim, J. J. (2011, February). Implicit feature identification via co-occurrence association rule mining. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 393-404). Springer, Berlin, Heidelberg.
10. Duric, A., & Song, F. (2012). Feature selection for sentiment analysis based on content and syntax models. Decision support systems, 53(4), 704-711.
11. Wang, Y., & Li, B. (2015, November). Sentiment analysis for social media images. In 2015 IEEE international conference on data mining workshop (ICDMW) (pp. 1584-1591). IEEE.
12. Akhtar, M. S., Kumar, A., Ghosal, D., Ekbal, A., & Bhattacharyya, P. (2017, September). A multilayer perceptron based ensemble technique for fine-grained financial sentiment analysis. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 540-546).
13. Ding, J., Sun, H., Wang, X., & Liu, X. (2018, June). Entity-level sentiment analysis of issue comments. In Proceedings of the 3rd International Workshop on Emotion Awareness in Software Engineering (pp. 7-13).
14. Ahuja, R., Chug, A., Kohli, S., Gupta, S., & Ahuja, P. (2019). The impact of features extraction on the sentiment analysis. Procedia Computer Science, 152, 341-348.
15. Dahooie, J. H., Raafat, R., Qorbani, A. R., & Daim, T. (2021). An intuitionistic fuzzy data-driven product ranking model using sentiment analysis and multi-criteria decision-making. Technological Forecasting and Social Change, 173, 121158.

16. Jakob, N., & Gurevych, I. (2010, October). Extracting opinion targets in a single and cross-domain setting with conditional random fields. In Proceedings of the 2010 conference on empirical methods in natural language processing (pp. 1035-1045).
17. Zhao, Y., Qin, B., Hu, S., & Liu, T. (2010, June). Generalizing syntactic structures for product attribute candidate extraction. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics (pp. 377-380).
18. Zhang, L., Liu, B., Lim, S. H., & O'Brien-Strain, E. (2010, August). Extracting and ranking product features in opinion documents. In Coling 2010: Posters (pp. 1462-1470).
19. Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion word expansion and target extraction through double propagation. Computational linguistics, 37(1), 9-27.
20. Di Caro, L., & Grella, M. (2013). Sentiment analysis via dependency parsing. Computer Standards & Interfaces, 35(5), 442-453.
21. Zhou, H., & Song, F. (2015, April). Aspect-level sentiment analysis based on a generalized probabilistic topic and syntax model. In The Twenty-Eighth International Flairs Conference.
22. Yu, L. C., Wang, J., Lai, K. R., & Zhang, X. (2017, September). Refining word embeddings for sentiment analysis. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 534-539).
23. Naseem, U., Khan, S. K., Razzak, I., & Hameed, I. A. (2019, December). Hybrid words representation for airlines sentiment analysis. In Australasian Joint Conference on Artificial Intelligence (pp. 381-392). Springer, Cham.
24. Dai, J., Yan, H., Sun, T., Liu, P., & Qiu, X. (2021). Does syntax matter? A strong baseline for Aspect-based Sentiment Analysis with RoBERTa. arXiv preprint arXiv:2104.04986.
25. Zhuang, L., Jing, F., & Zhu, X. Y. (2006, November). Movie review mining and summarization. In Proceedings of the 15th ACM international conference on Information and knowledge management (pp. 43-50).
26. Kessler, J. S., & Nicolov, N. (2009, March). Targeting sentiment expressions through supervised ranking of linguistic configurations. In Third international AAAI conference on weblogs and social media.
27. Toprak, C., Jakob, N., & Gurevych, I. (2010, July). Sentence and expression level annotation of opinions in user-generated discourse. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (pp. 575-584).
28. Liu, K. L., Li, W. J., & Guo, M. (2012, July). Emoticon smoothed language models for twitter sentiment analysis. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 26, No. 1).
29. Anjaria, M., & Guddeti, R. M. R. (2014). A novel sentiment analysis of social networks using supervised learning. Social Network Analysis and Mining, 4(1), 181.

30. Ghiassi, M., Zimbra, D., & Lee, S. (2016). Targeted twitter sentiment analysis for brands using supervised feature engineering and the dynamic architecture for artificial neural networks. Journal of Management Information Systems, 33(4), 1034-1058.
31. Schouten, K., Van Der Weijde, O., Frasincar, F., & Dekker, R. (2017). Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. IEEE transactions on cybernetics, 48(4), 1263-1275.
32. Rustam, F., Khalid, M., Aslam, W., Rupapara, V., Mehmood, A., & Choi, G. S. (2021). A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis. Plos one, 16(2), e0245909.
33. Hofmann, T. (2000). Learning the similarity of documents: An information-geometric approach to document retrieval and categorization. In Advances in neural information processing systems (pp. 914-920).
34. Blei, D. M., & Moreno, P. J. (2001, September). Topic segmentation with an aspect hidden Markov model. In Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 343-348).
35. Titov, I., & McDonald, R. (2008, April). Modeling online reviews with multi-grain topic models. In Proceedings of the 17th international conference on World Wide Web (pp. 111-120).
36. Long, C., Zhang, J., & Zhu, X. (2010, August). A review selection approach for accurate feature rating estimation. In Coling 2010: Posters (pp. 766-774).
37. Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011, December). Multi-aspect sentiment analysis with topic models. In 2011 IEEE 11th international conference on data mining workshops (pp. 81-88). IEEE.
38. Wang, H., Lu, Y., & Zhai, C. (2011, August). Latent aspect rating analysis without aspect keyword supervision. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 618-626).
39. Hu, X., Tang, J., Gao, H., & Liu, H. (2013, May). Unsupervised sentiment analysis with emotional signals. In Proceedings of the 22nd international conference on World Wide Web (pp. 607-618).
40. Pavlopoulos, J., & Androutsopoulos, I. (2014, April). Aspect term extraction for sentiment analysis: New datasets, new evaluation measures and an improved unsupervised method. In Proceedings of the 5th Workshop on Language Analysis for Social Media (LASM) (pp. 44-52).
41. Suresh, H. (2016, October). An unsupervised fuzzy clustering method for twitter sentiment analysis. In 2016 International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS) (pp. 80-85). IEEE.
42. García-Pablos, A., Cuadros, M., & Rigau, G. (2018). W2VLDA: almost unsupervised system for aspect based sentiment analysis. Expert Systems with Applications, 91, 127-137.

43. Yadav, S., & Chakraborty, T. (2020). Unsupervised sentiment analysis for code-mixed data. arXiv preprint arXiv:2001.11384.
44. Popescu, A. M., & Etzioni, O. (2007). Extracting product features and opinions from reviews. In Natural language processing and text mining (pp. 9-28). Springer, London.
45. Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G., & Reynar, J. (2008). Building a sentiment summarizer for local service reviews.
46. Raju, S., Pingali, P., & Varma, V. (2009, April). An unsupervised approach to product attribute extraction. In European Conference on Information Retrieval (pp. 796-800). Springer, Berlin, Heidelberg.
47. Yu, J., Zha, Z. J., Wang, M., & Chua, T. S. (2011, June). Aspect ranking: identifying important product aspects from online consumer reviews. In Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies (pp. 1496-1505).
48. Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. Expert Systems with applications, 40(16), 6266-6282.
49. Brun, C., Popa, D. N., & Roux, C. (2014, August). XRCE: Hybrid Classification for Aspect-based Sentiment Analysis. In SemEval@ COLING (pp. 838-842).
50. Appel, O., Chiclana, F., Carter, J., & Fujita, H. (2016). A hybrid approach to the sentiment analysis problem at the sentence level. Knowledge-Based Systems, 108, 110-124.
51. Ma, Y., Peng, H., Khan, T., Cambria, E., & Hussain, A. (2018). Sentic LSTM: a hybrid network for targeted aspect-based sentiment analysis. Cognitive Computation, 10(4), 639-650.
52. Iqbal, F., Hashmi, J. M., Fung, B. C., Batool, R., Khattak, A. M., Aleem, S., & Hung, P. C. (2019). A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. IEEE Access, 7, 14637-14652.
53. Kumar, A., Srinivasan, K., Cheng, W. H., & Zomaya, A. Y. (2020). Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data. Information Processing & Management, 57(1), 102141.
54. Chen, T., Borth, D., Darrell, T., & Chang, S. F. (2014). Deepsentibank: Visual sentiment concept classification with deep convolutional neural networks. arXiv preprint arXiv:1410.8586.
55. Poria, S., Cambria, E., & Gelbukh, A. (2015, September). Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis. In Proceedings of the 2015 conference on empirical methods in natural language processing (pp. 2539-2544).
56. Chen, P., Sun, Z., Bing, L., & Yang, W. (2017, September). Recurrent attention network on memory for aspect sentiment analysis. In Proceedings of the 2017 conference on empirical methods in natural language processing (pp. 452-461).

57. Jangid, H., Singhal, S., Shah, R. R., & Zimmermann, R. (2018, April). Aspect-based financial sentiment analysis using deep learning. In Companion Proceedings of the The Web Conference 2018 (pp. 1961-1966).
58. Arora, M., & Kansal, V. (2019). Character level embedding with deep convolutional neural network for text normalization of unstructured data for Twitter sentiment analysis. Social Network Analysis and Mining, 9(1), 1-14.
59. Dashtipour, K., Gogate, M., Li, J., Jiang, F., Kong, B., & Hussain, A. (2020). A hybrid Persian sentiment analysis framework: Integrating dependency grammar based rules and deep neural networks. Neurocomputing, 380, 1-10.
60. Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. Electronics, 9(3), 483.