Semantic-Aware Sentiment Analysis: A Survey

Urmi Saha[1] , **Aditya Joshi**[1,2] , **Pushpak Bhattacharyya**[1]
[1]Indian Institute of Technology, Bombay
[2]CSIRO, Sydney
{urmisaha, pb}@cse.iitb.ac.in, aditya.joshi@csiro.au

Abstract

Ontologies are known to serve well as a knowledge base for several NLP problems, including sentiment analysis. This report talks about the importance of semantics in the task of sentiment analysis. In this paper, we present a concise survey of various approaches of 1) aspect-based sentiment analysis, 2) ontology-based sentiment analysis and 3) semantics in sentiment analysis, which provide a base for our research. We observe that, although ontologies have proved to be important for various NLP tasks, their benefits in neural sentiment analysis is a much less-explored area of research.

1 Introduction

In today's world, sentiment analysis and opinion mining is of utmost importance for e-commerce giants like Flipkart, Amazon, etc. to analyze their product reviews, which in turn would help them to improve on various aspects and deal with customers. In addition to these sites, even social networking websites like Facebook, Twitter, etc have become sources of huge amount of review data as people have taken to these platforms for venting out complaints and express dissatisfaction about some product or service. Often, people find social media platforms more reliable to spread awareness among public, as e-commerce sites can get rid of negative reviews, or can pump up fake positive reviews due to biasness. Thus there are loads of data getting accumulated by internet users waiting for getting processed and to be put to useful analysis. Semantics is a very crucial aspect in this important task of sentiment analysis. Consider the following examples,

1. *The weight of the camera is too heavy* is a negative review for camera.

2. *The lens of the camera is quite good* is a positive review for camera.

In general sentiment analysis task, the two sentences put together would generate a neutral sentiment for the entire review, as there is one positive and one negative polarity assigned. But if we keep the domain in mind, we know for a camera *lens* is more important than *weight* and thus the overall polarity should have been positive. But classification in this level is difficult to achieve. More information is required about the domain to assign scores (which can be considered as a importance measure) to different aspects to determine the degree of positivity and negativity of an expression. Semantics of a sentence along with domain knowledge from a domain ontology graph is thus the main concern of this research work.

To attend to this problem we first perform some background research on three fields which can be considered part of our final research goal: 1) Aspect-based Sentiment Analysis, 2) Ontology-based Sentiment Analysis and 3) Semantics Incorporation in Sentiment Analysis. Some interesting research works are mentioned briefly in the following sections to give an overview of these research problems which are significant parts of our research work.

The rest of the paper is organized as follows. We first explain semantic analysis in Section 2. In Section 3, we talk about ontology and its contributions in Natural Language Processing tasks. We describe various approaches in Section 4. Finally, we conclude the paper in Section 5.

2 Semantic Analysis

According to The Free Dictionary[1], *semantics is the linguistic and philosophical study of meaning*. Semantic analysis is the study of semantics or meanings of a word, phrase, or document. Construction of meaning representations of linguistic input and extracting out common-sense knowledge from it is the main work of semantic analysis. Thus semantic analysis can be approximately considered as understanding knowledge.

Importance of semantic analysis is quite significant in various Natural Language Processing tasks, namely text summarization, information retrieval, machine translation, document classification, sentiment analysis, human-computer interaction, etc.

2.1 Sentiment Analysis

Sentiment analysis in the simplest form is the method of automatically determining whether sentiment expressed about

[1]https://www.thefreedictionary.com/

a given subject is positive, negative, or neutral. Other sentiments include anger, surprise, and many others.

2.2 Semantics in Sentiment Analysis

Sentiment Analysis is one of the fundamental tasks in Natural Language Processing. The goal is to assign polarities to product reviews. Basic sentiment analysis uses a bag of words representations where sentiment is decided by calculating the number of positive and negative tokens in the text. This does not always yield proper results due to phenomena like negations, sarcasm, thwarting, discourse analysis, etc. Moreover, when multiple aspects of an object are mentioned with different sentiments, it is important to understand which aspect is more important for the concerned object and accordingly weights should be assigned to each aspect, which in turn will decide the overall polarity of the object. Here, semantics has got a key role to play. Semantic-aware sentiment analysis finds the semantics or meanings of the aspects and their importance for the object or topic under consideration and helps in sentiment prediction more accurately. Few interesting phenomena in sentiment analysis where semantics play a good role and which when taken care generates better results are *negations*, *sarcasm*, *thwarting* and *discourse analysis*.

3 Ontology

As mentioned in [Guarino *et al.*, 2009], the word ‘ontology’ has different senses in two different communities, namely the branch of philosophy and the branch of computer science. We deal with the latter where ontology is a mean to model structure of a system with entities and relations created out of observations and help in representing useful information for various important computational tasks. The backbone of an ontology is a hierarchy of concepts with generalization or specialization relation.

To understand the hierarchy, we can consider the following example: ontology for human resources will have *Person*, *Manager*, and *Researcher* as relevant concepts. *Person* is a superconcept of the other two. *Cooperates* can be considered a relevant relation between persons.

[Gruber, 1991] originally defined the notion of an ontology as an “explicit specification of a conceptualization” which is considered to be the most prevalent definition of ontology among other definitions. [Borst and Borst, 1997] defined an ontology as a “formal specification of a shared conceptualization”. [Studer *et al.*, 1998] merged these two definitions stating that: *“An ontology is a formal, explicit specification of a shared conceptualization.”*

Conceptualization can be considered as an abstract, simplified view of the world of knowledge, represented for some purpose. The role of ontology is to explicitly specify a conceptualization. An ontology can thus be considered as just a set of such axioms, i.e., a logical theory designed in order to capture the intended models corresponding to a certain conceptualization and to exclude the unintended ones, resulting in an approximate specification of a conceptualization.

Some popular ontologies and knowledge bases are SUMO (The Suggested Upper Merged Ontology) [Pease *et al.*, 2002], WordNet, Knowledge Graph by Google, CYC, UMLS, Freebase, etc.

3.1 Ontology-driven NLP

As mentioned in [Polpinij and Ghose, 2008], for sentiment analysis tasks, text classification techniques are not good enough because domain knowledge of text classification comes from the defined task and it is hard to transfer that knowledge to a variety of domains of interest. A solution to this problem is to use ontology which can bring about improvement as ontology represents understanding of domains. Hence ontology is believed to enhance performance of information processing systems. In addition to this, domain information is also beneficial to understand the semantic orientation of words which contribute to the NLP tasks significantly.

[Sharma and Bhattacharyya, 2015] in their research has shown that there are some Domain Dedicated Polar Words (DDPW) (eg. *blockbuster* is a positive polar word while *miscast* is a negartive polar word specific to movie domain), which when used as features results in high accuracy. This paper also mentions *chameleon words* which have fixed polarity in a particular domain but fluctuating polarity across different domains. E.g. *unpredictable driving* expresses negative sentiment while *unpredictable story line* expresses positive sentiment. The method mentioned here first identifies DDPWs, finds out the class (positive/negative) of the encountered word by performing the Chi-Square Test. These are then used as features for three machine learning based classification algorithms, that are, Neural Network, Logistic Regression and SVM. Figure 1 is the proposed methodology of this research paper.

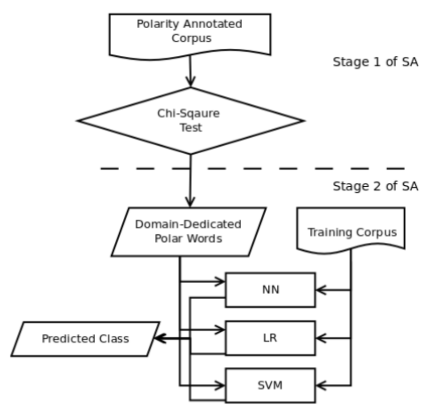

Figure 1: Two-stage approach Sentiment Analysis[Sharma and Bhattacharyya, 2015]

Another very interesting NLP task is detection of thwarting in a sentence containing sentiment expressions. **Thwarting**, as mentioned in [Ramteke *et al.*, 2013] is looked upon as the phenomenon of polarity reversal at a higher level of ontology compared to the polarity expressed at the lower level. More simply, if polarity of majority of a document contradicts with the overall polarity of the document, then the document is considered to be thwarted. An example of a thwarted document is:

> I love the sleek design. The lens is impressive. The pictures look good but, somehow this camera dis-

appoints me. I do not recommend it.

A machine learning method with added annotation of thwarted/non-thwarted results in more accuracy than rule-based system. So, authors of this paper have built a model to detect turnaround of sentiment from entity level to overall document level, with the help of domain ontology of camera. The first step described in the paper is to build domain ontology by first a) identifying features and entities, and then b) connecting them in a form of hierarchy. For this step, Latent Dirichlet Allocation (LDA) [Blei *et al.*, 2003] have been used on a corpus containing camera reviews. The paper then mentions both rule-based and machine learning-based approach. In the rule-based method, adjective-noun dependencies are identified by dependency parsing and polarities against the nouns are updated in the nodes of the ontology graph. Finally, document is considered thwarted if there is contradiction of sentiment between different levels of the ontology graph. The machine learning method includes two major steps: 1) learning polarity weights for each word in a sentence which belongs to the ontology graph, 2) using these weighted polarities as feature vector in a SVM classifier.

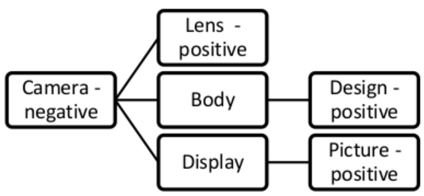

Figure 2: Ontology with polarity marking [Ramteke *et al.*, 2013]

4 Approaches

In this section, we describe past approaches in sentiment analysis research. We classify them into three categories: aspect-based sentiment analysis, ontology-based sentiment analysis and semantics in sentiment analysis approaches.

4.1 Approaches for Aspect-based Sentiment Analysis

Sentiment analysis of an item provides a general opinion of an item whether it is good or bad. But a positive opinion of an item does not necessarily mean that all the aspects of the item has been given a positive opinion. This calls for deeper analysis of sentiment i.e. sentiment expressed in aspect level.

The approach presented in [Tang *et al.*, 2016] consists of multiple computational layers with shared parameters. Each layer is a content and location based attention model, which first learns the importance/weight of each context word and then utilizes this information to calculate continuous text representation. The text representation in the last layer is regarded as the feature for sentiment classification. As every component is differentiable, the entire model could be efficiently trained end-to-end with gradient descent, where the loss function is the cross-entropy error of sentiment classification.

Given a sentence s = $w_1, w_2, ..., w_i, ...w_n$ and the aspect word w_i ,each word is mapped into its embedding vector. These word vectors are separated into two parts, aspect representation and context representation. If aspect is a single word like *"food"* or *"service"*, aspect representation is the embedding of aspect word. For the case where aspect is multiple word expression like *"battery life"*, aspect representation is an average of its constituting word vectors. To simplify the interpretation, we consider aspect as a single word w_i. Context word vectors $\{e_1, e_2...e_{i-1}, e_{i+1}...e_n\}$ are stacked and regarded as the external memory m $\in R^{dX(n-1)}$, where *n* is the sentence length. An illustration of this approach is given in Figure 3[2].

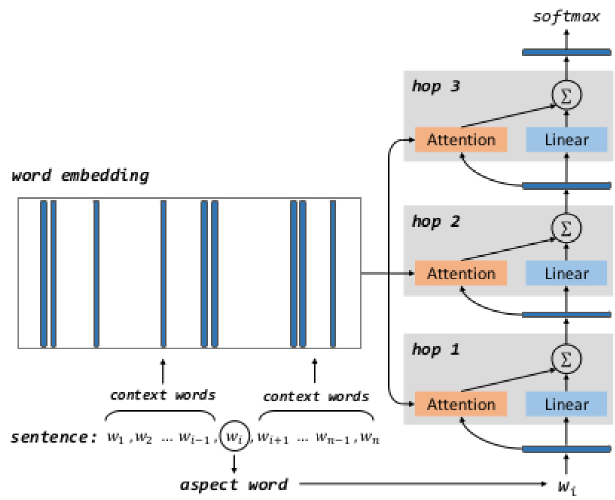

Figure 3: An illustration of deep memory network with three computational layers (hops) for aspect level sentiment classification. [Tang *et al.*, 2016]

This approach consists of multiple computational layers (hops), each of which contains an attention layer and a linear layer. In the first computational layer (hop 1), aspect vector is regarded as the input to adaptively select important evidences from memory *m* through attention layer. The output of attention layer and the linear transformation of aspect vector are summed and the result is considered as the input of next layer (hop 2). The output vector in last hop is considered as the representation of sentence with regard to the aspect, and is further used as the feature for aspect level sentiment classification.

Another recent approach[Wang *et al.*, 2018] jointly addresses aspect extraction and sentiment prediction by introducing the dual-label scheme to integrate sentiment labels with general BIO labels. Aspect-based Sentiment Analysis is considered here as a sequential tagging problem following a novel label schema where sentiment labels are integrated with general BIO labels: a dual-label scheme "BIO-sen" indicating both location of aspect and its sentiment orientation.

- B-sen(beginning of an aspect term with "sen" emotion), I-sen(inside of an aspect term with "sen" emotion),

[2] [Tang *et al.*, 2016]

O(others)

- "sen" represents four emotions, i.e., pos(positive), neg(negative), neu(neutral) and con(conflict)

MNNs output a sequence of dual-labels for a review sentence S, and the dual-label corresponds to a word in S.

Aspect-based Sentiment Analysis, a joint tagging problem has a novel multi-task neural learning framework:

- Two types of high-level sentence representations are generated via Convolutional Neural Network (CNN) [Kim, 2014] and Long Short-Term Memory Network (LSTM) [Gers *et al.*, 1999] separately.
- Interactive attention on these two representations to capture important interactive information of different representations and self-attention on LSTM representation to learn the inter-relationship between aspect and sentiment.
- Conditional Random Fields (CRF) is used to decode the sequence of labels.

Previous works on aspect-based sentiment analysis generally divide the task into several subtasks and address them in a pipeline solution. But in this method leads to error propagation along with intensive labour and external resource dependency for each subtask. To avoid these problems, multi-task neural network is proposed by the authors in their work where a novel multi-task neural learning framework jointly tackles aspect extraction and sentiment prediction subtasks at the same time, and leverage attention mechanisms to learn the joint representation of aspect-sentiment relationship. Unlike other multi-task models, here both subtasks are granted as main tasks because both subtasks are equally important for the whole ABSA.

The two multi-task neural networks in this paper are described:

- MNN-1 model consists of an **embedding layer**(character-level and word-level embeddings combined), a **single BiLSTM encoder with self-attention** to capture the inter-relationship between aspect and sentiment, and a **CRF decoding layer**
- a modified model MNN-2 is proposed by adding a **CNN component** to learn a different sentence representation and **interactive-attention** to capture the interactive information of CNN and LSTM sentence representations
- output is a sequence of dual-labels for a given review sentence S, for each word w_i S

The experimental results on benchmark corpora show effectiveness of proposed multi-task models. The performance of aspect extraction outperforms the state-of-the-art systems and the performance of sentiment prediction is quite competitive compared with most existing works.

4.2 Approaches for Ontology-based Sentiment Analysis

Ontology represents knowledge of a particular domain in a hierarchical form through concepts and relationships between those concepts. Ontology graph makes the task of sentiment

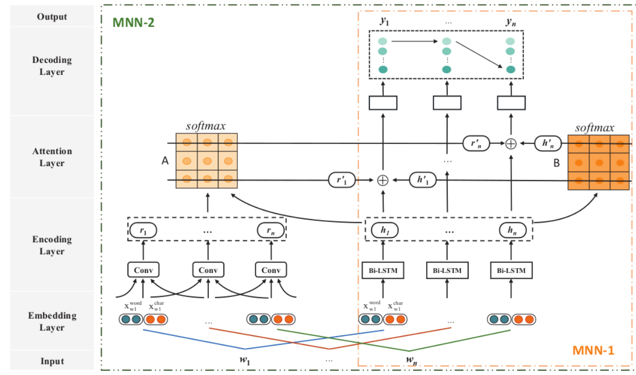

Figure 4: Proposed model of [Wang *et al.*, 2018]

classification easier. It is useful for considering unseen words too which are not found in the labelled corpus, by calculating nearest mean.

[Tamilselvam *et al.*, 2017] in their work have addressed the problem of sentiment aggregation by assigning weights to aspects appearing in the text content. These weights are considered to be aspect importance in the concerned domain. The weights of the aspects are calculated from domain-specific ontologies extracted from ConceptNet, using graph centrality measures. These weights are then incorporated while calculating the sentiment aggregation. This system outperforms the state of the art by a F-score of about 20%. The requirement of ontology graph in such case is the diversity of aspect. Simple summation of aspect-level sentiment polarity values for calculating overall aggregated sentiment is not enough. This paper suggests a novel approach that extracts ontology graph from ConceptNet and assign weights using a measure of their centrality. These weights are combined with the feature sentiments and a weighted aggregation is carried out to obtain overall sentiment for each review. Various centrality measures are used to measure the importance of nodes and the amount of sentiment information that can be shared across nodes during sentiment aggregation. Thus interrelationships amongst aspects are maintained while calculating sentiment aggregation.

The authors of this paper has described their approach of sentiment aggregation using Graph centrality in the following steps:

- **Ontology graph construction:** For each concept or domain, a seed word is identified and then for each vertex in the graph, if there is a concept in ConceptNet which has atleast one relationship with the vertex, the concept is added into the graph. The relationships can be 1) Hierarchical, representing parent-child relationship, 2) Synonymous, identifying related concepts and 3) Functional, identifying purpose of a concept.
- **Graph centrality computation:** The different graph centrality measures described are 1) Closeness Centrality, 2) Betweenness Centrality and 3) PageRank [Brin and Page, 1998]. Relationship type is not considered while computing this measure. Closeness centrality measure performs best.

- **Feature/Aspect-Specific Sentiment Computation:** Dependency parsing of each review determines the sentiment polarity expressed towards an aspect(feature). Each feature acts as a cluster head and each word is assigned to the cluster whose cluster head is the closest i.e. path contains least number of edges as compared to that to other cluster heads. Majority voting of sentiment values of each word in a cluster determines opinion about that feature.
- **Sentiment Polarity Aggregation:** Overall polarity of review is considered a weighted sum of a) sentiment polarity towards each aspect and b) importance of aspect in given domain.

  If c_{m_i} represents the centrality score for an aspect m_i computed using the ontology graph for the domain which P belongs to, polarity value p(sum) for a review text R can be calculated as:

$$p(sum) = \sum_{i=1}^{M} m_i^p \times c_{m_i} \tag{1}$$

Finally aggregated sentiment polarity is assigned as:

S = Positive $\widehat{\text{if}}$ p(sum) >0
S = Negative $\widehat{\text{if}}$ p(sum) <0
S = Neutral $\widehat{\text{if}}$ p(sum) = 0

This research thus shows that in a particular domain, incorporating significance of entities along with their sentiment weights observe good performance outperforming state of the art methods.

4.3 Approaches for Semantics in Sentiment Analysis

Semantics play an important role in various Natural Language Processing tasks. Semantics analysis deals with understanding the data in a deeper level. Here we focus on the task of sentiment classification. Simple sentiment classification of a document will be to label it with positive, negative or neutral class. But with the help of semantics, more fine-grained information can be extracted about the various aspects mentioned in the document and the relevance or importance of different aspects to the domain under consideration.

[Verma and Bhattacharyya, 2009] proposes a methodology where SentiWordNet (English wordnet[3] with polarity scores) is used to incorporate **word level sentiment** to feature vector of a document. This methodology is deviced to address the following challenges of sentiment analysis task:

- **Subjectivity detection**, i.e., selecting opinion containing sentences.
  Example:*Singapore's economy is heavily dependent on tourism and IT industry. It is an excellent place to live in.*
  The first sentence is s sentimentless objective or factual sentence and should be filtered out.
  SVM classifier is trained with different values of cost parameter C. Best results of F-Measure 88.53% was produced for C=1. Classifier then classifies preprocessed sentences of corpus as subjective or objective.
- **Word Sense Disambiguation**:
  Example: *an unpredictable plot in the movie* is positive, *an unpredictable steering wheel* is negative
- **Thwarting** i.e., sudden deviation from positive to negative polarity:
  Example: *The movie has a great cast, superb storyline and spectacular photography; the director has managed to make a mess of the whole thing*
- **Negations:**
  Example: *not good* is replaced with *not_good* and sentiment score of *not_good* is the negative of the sentiment score of *good*
- **Keeping the target in focus:**
  Example: *my camera compares nothing to John's camera which is sleek and light, produces life like pictures and is inexpensive*.

SentiWordNet[Esuli and Sebastiani, 2006] is considered a very important resource for the proposed methodology. SentiWordNet is based on English wordnet along with sentiment information incorporated into each synset, which have three numerical scores *Pos(S), Neg(S) and Obj(S)*, describing *positivity, negativity and objectivity* of the synset. Thus SentiWordNet is considered an important resource for subjectivity detection too along with polarity identification.

The methodology described in this paper includes mainly two steps:

1. **Preprocessing of Data:** This includes tokenizing, stemming and stop word removal on the documents
2. **Feature pruning:** Feature pruning consists of three sub tasks performed one after the other:
   (a) **Sentiment score based pruning on full review:** This process removes all the non-opinion words from the text. Sentiment scores of all words are calculated using SentiWordNet and words with score higher than a particular threshold only are accepted. Also, to address the Word Sense Disambiguation (WSD) problem, sentiment score of a word is calculated from sentiment scores of all words in its synset.
   After this process is completed, accuracy obtained was 71.80% (movie review) and to 81.80% (product review)
   (b) **Information gain based pruning on full review:** This process removes domain specific stop words and noisy words (eg. cast, hero, comedy, etc).
   After this step, accuracy improved to 78.9% (movie review) and to 83.91% (product review)
   (c) **Subjective review:** This process considers only sentences containing opinion.
   After this final step, accuracy improved to 85.61% (movie review).

[3] http://princeton.wordnet.edu

Figure 5: Closeness centrality measure for Camera Ontology [Tamilselvam *et al.*, 2017]

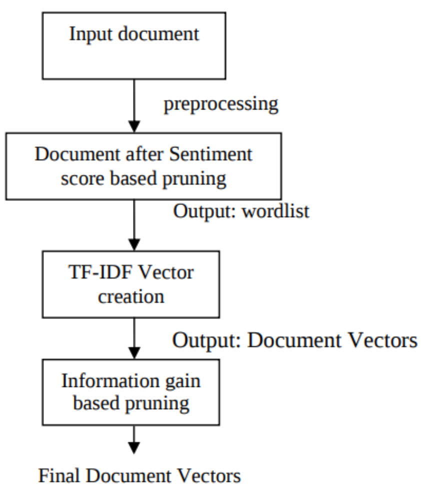

Figure 6: Document vector creation [Verma and Bhattacharyya, 2009]

This method proves to be quite better in comparison with state of the art - 85.61% as compared to

- 82.9% (method used by [Pang *et al.*, 2002], which applies machine learning techniques on unigram and bigram features) on the same corpus
- 86.4% (method used by [Pang and Lee, 2004], which applies min cut algorithm with SVM classifier)

[Socher *et al.*, 2013] has claimed that althrough semantic words have been quite useful, they still fail to express meanings of longer phrases. Authors of this paper have provided solutions to this problem by introducing a **Sentiment Treebank** which includes sentiment labels for 215,154 phrases in the parse trees of 11,855 sentences. This resulted in new challenges for sentiment composionality, which has been addressed by a new model called **Recursive Neural Tensor Network**. This model, trained on the new treebank outperforms previous methods by significant amount. This model also captures accurately effects of negation at various levels of the tree for both positive and negative phrases.

The **Stanford Sentiment Treebank** is the first corpus with fully labeled parse trees. The corpus is based on movie review excerpts from the rottentomatoes.com website originally collected and published by [Pang and Lee, 2005]. The dataset consisted of $10,662$ sentences, half of which were considered positive and the rest negative.

The powerful **Recursive Neural Tensor Network(RNTN)** predicts compositional semantic effects. It represents a phrase through word vectors and a parse tree and then compute vectors for higher nodes in the tree using the same tensor-based composition function. Several supervised, compositional models get significant boost when trained with the new dataset, but RNTN obtains the highest performance with 80.7% accuracy.

The paper describes three recursive neural models which compute compositional vector representations for phrases of variable length and syntactic type, which are used as features for classification. Such compositional model parses an n-gram into a binary tree, where each leaf node corresponds to a word and is represented as a vector. Recursive neural models, as the name suggests compute parent vectors in a bottom up fashion using different compositionality functions. Fig 7 depicts the how parent node vectors are calculated from children node vectors.

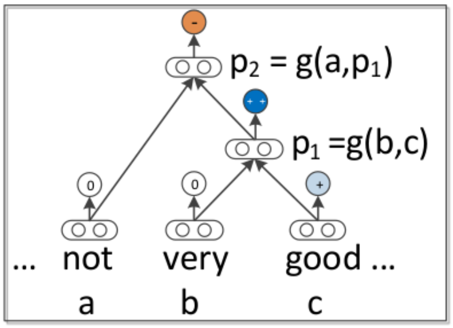

Figure 7: Approach of Recursive Neural Network models for sentiment: Compute parent vectors in a bottom up fashion using a compositionality function g and use node vectors as features for a classifier at that node. This function varies for the different models.[Socher *et al.*, 2013]

Various recursive neural models are described and compared here: Recursive Neural Network(RNN), Matrix-Vector RNN(MV-RNN) and Recursive Neural Tensor Network(RNTN). Common operations in these models are: 1) word vector representations and classification, 2) each word is represented as a d-dimensional vector.

RNN:

This is the simplest member of the neural network family, where parent vectors are calculated all of whose children are already computed. With reference to Fig.7,

$$p_1 = f(W \begin{vmatrix} b \\ c \end{vmatrix}), p_2 = f(W \begin{vmatrix} a \\ p_1 \end{vmatrix}) \tag{2}$$

$W \epsilon R^d \times 2^d$ is the main parameter to learn.
$f = \tanh$

Parent vectors must be of the same dimensionality to be recursively compatible and must be given the same softmax classifier.

MV-RNN:

This method represents every word and longer phrase in a parse tree as both a vector and a matrix. While combining two constituents, matrix of one is multiplied with the vector of the other and vice versa. Each n-gram is represented as a list of (vector,matrix) pairs, together with the parse tree.

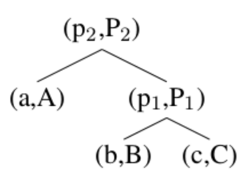

Figure 8: Vector representations of nodes in MV-RNN[Socher *et al.*, 2013]

$$p_1 = f(W \begin{vmatrix} Cb \\ Bc \end{vmatrix}), P_2 = f(W_M \begin{vmatrix} B \\ C \end{vmatrix}) \tag{3}$$

RNTN:

The main idea of this model is to use the same, tensor-based composition function for all nodes. Output of a tensor product $h \epsilon R^d$:

$$h = \begin{vmatrix} b \\ c \end{vmatrix}^T V^{[1:d]} \begin{vmatrix} b \\ c \end{vmatrix} \tag{4}$$

$$h_i = \begin{vmatrix} b \\ c \end{vmatrix}^T V^i \begin{vmatrix} b \\ c \end{vmatrix} \tag{5}$$

$$p_1 = f(\begin{vmatrix} b \\ c \end{vmatrix}^T V^{[1:d]} \begin{vmatrix} b \\ c \end{vmatrix} + W \begin{vmatrix} b \\ c \end{vmatrix}) \tag{6}$$

$$p_2 = f(\begin{vmatrix} a \\ p_1 \end{vmatrix}^T V^{[1:d]} \begin{vmatrix} a \\ p_1 \end{vmatrix} + W \begin{vmatrix} a \\ p_1 \end{vmatrix}) \tag{7}$$

We can interpret each slice of tensor as capturing a specific type of composition. 9

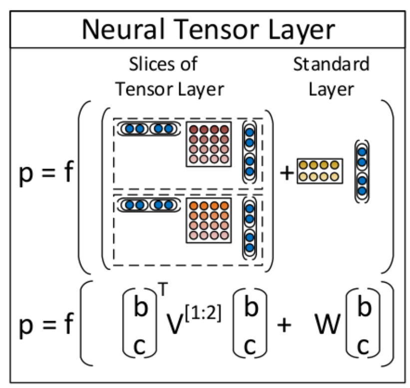

Figure 9: A single layer of the Recursive Neural Ten- sor Network. Each dashed box represents one of d-many slices and can capture a type of influence a child can have on its parent.[Socher *et al.*, 2013]

Evaluation metric analyzes the accuracy of fine-grained sentiment classification for all phrases, where RNTN gets the highest performance, followed by the MV-RNN and RNN models.

Figure 10 shows the sentiment prediction of sentences.

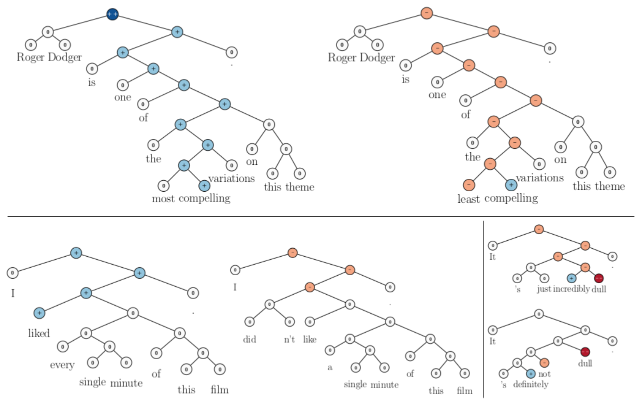

Figure 10: RNTN prediction of positive and negative (bottom right) sentences and their negation.[Socher *et al.*, 2013]

5 Conclusion

This paper presented definitions of semantics and ontology and there relevance and usefulness in NLP tasks, specifically sentiment analysis. We discussed recent trends as reported in the past work in sentiment analysis research. We focused mainly on three areas: 1) aspect-based sentiment analysis, 2) ontology-based sentiment analysis and 3) semantics in sentiment analysis. We observed that significant research has been performed on these domains. However, incorporating semantics into a neural architecture through a domain ontology is a relatively new area of research and is less explored. We conclude that, exploring the benefits of incorporating semantics, extracted from an ontology into a neural network in order to enhance aspect-based sentiment signals is a promising line of work in sentiment analysis research.

References

[Blei *et al.*, 2003] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[Borst and Borst, 1997] Willem Nico Borst and WN Borst. Construction of engineering ontologies for knowledge sharing and reuse. 1997.

[Brin and Page, 1998] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

[Esuli and Sebastiani, 2006] Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In *LREC*, volume 6, pages 417–422. Citeseer, 2006.

[Gers *et al.*, 1999] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. 1999.

[Gruber, 1991] TR Gruber. A translation approach to portable ontologies knowledge acquisition 5 (2) 199220. *Neches, R, Fikes, R, Finin, T, Gruber, T, Patil, R, Senator, T and Swartout, W*, 1991.

[Guarino *et al.*, 2009] Nicola Guarino, Daniel Oberle, and Steffen Staab. What is an ontology? In *Handbook on ontologies*, pages 1–17. Springer, 2009.

[Kim, 2014] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[Pang and Lee, 2004] Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics*, page 271. Association for Computational Linguistics, 2004.

[Pang and Lee, 2005] Bo Pang and Lillian Lee. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd annual meeting on association for computational linguistics*, pages 115–124. Association for Computational Linguistics, 2005.

[Pang *et al.*, 2002] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics, 2002.

[Pease *et al.*, 2002] Adam Pease, Ian Niles, and John Li. The suggested upper merged ontology: A large ontology for the semantic web and its applications. In *Working notes of the AAAI-2002 workshop on ontologies and the semantic web*, volume 28, pages 7–10, 2002.

[Polpinij and Ghose, 2008] Jantima Polpinij and Aditya K Ghose. An ontology-based sentiment classification methodology for online consumer reviews. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 1, pages 518–524. IEEE, 2008.

[Ramteke *et al.*, 2013] Ankit Ramteke, Akshat Malu, Pushpak Bhattacharyya, and J Saketha Nath. Detecting turnarounds in sentiment analysis: Thwarting. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, volume 2, pages 860–865, 2013.

[Sharma and Bhattacharyya, 2015] Raksha Sharma and Pushpak Bhattacharyya. Domain sentiment matters: A two stage sentiment analyzer. In *Proceedings of the 12th International Conference on Natural Language Processing*, pages 130–137, 2015.

[Socher *et al.*, 2013] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642, 2013.

[Studer *et al.*, 1998] Rudi Studer, V Richard Benjamins, and Dieter Fensel. Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1-2):161–197, 1998.

[Tamilselvam *et al.*, 2017] Srikanth Tamilselvam, Seema Nagar, Abhijit Mishra, and Kuntal Dey. Graph based sentiment aggregation using conceptnet ontology. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 525–535, 2017.

[Tang *et al.*, 2016] Duyu Tang, Bing Qin, and Ting Liu. Aspect level sentiment classification with deep memory network. *CoRR*, abs/1605.08900, 2016.

[Verma and Bhattacharyya, 2009] Shitanshu Verma and Pushpak Bhattacharyya. Incorporating semantic knowledge for sentiment analysis. *Proceedings of ICON*, 2009.

[Wang *et al.*, 2018] Feixiang Wang, Man Lan, and Wenting Wang. Towards a one-stop solution to both aspect extraction and sentiment analysis tasks with neural multi-task learning. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.