Multilingual sentiment analysis using connotations

Prashanti Nilayam
nilayam2@illinois.edu

December 30, 2021

Abstract

Sentiment analysis is a challenging task and countless researchers have come up with numerous solutions to capture sentiments for a given corpus from both denotation (literal meaning) and connotation (implied meaning) aspects. Also, both traditional natural language techniques and neural-network based techniques have been applied to effectively capture sentiments over time. In recent times, focus has been to combine the traditional approach and neural methods to capture subtle nuances and hidden sentiments with better performance. This paper presents a few of the related publications which showcases the power of combining both traditional and neural methods for sentiment analysis of multi-lingual corpus.

1 Introduction

Processing appropriate sentiment from a corpus can prove to be a challenging task even for a human being, since the realization and the extent of the sentiment might vary based on person, place, culture or simply the semantics of a particular language. It becomes particularly challenging because the real sentiment may or may not be even stated in the actual corpus. Sometimes, the sentiment needs to be implied based on other factors like the real-world knowledge or the context of the conversation and even from subtle changes in the choice of words. Numerous research work have been published to capture these nuances which uses a combination of both traditional natural language processing techniques and neural methods. This paper reviews 4 of the relevant publications in this area.

The paper presents a brief background followed by the motivation behind selecting the papers. Assessment of all the reviewed papers are presented along with the contributions made by these papers in the area of sentiment analysis, followed by overall conclusion.

2 Topic and Motivation

Extracting sentiment from a given corpus has been one of the heavily researched topics in natural language processing. Understanding sentiments from text is a very convoluted and challenging task due to several reasons including varying perspective of the involved parties, the subjective nature of the sentiments, the possibility of more than one sentiment for the given corpus and implied sentiments, etc.

Despite all these challenges, sentiment analysis has piqued the interest of countless researchers in both academia and industry for several decades due to its manifold possible applications (Feldman, 2013) ranging from product reviews (Haque et al., 2018; Mukherjee and Bhattacharyya, 2012; Wei and Gulla, 2010; Vyas and Uma, 2019) to detecting polarization in social media content (Badami et al., 2017; Mohammad et al., 2015), and from financial domain (Mittal and Goel, 2012) to health care applications (Denecke and Deng, 2015).

The methods (Ahlgren, 2016) of sentiment analysis have evolved over time. Some of the methods follow a more traditional approach including keyword-analysis (van Eck and Waltman, 2007) using a similarity metric, topic modeling using LDA (Blei et al., 2003), etc, while others follow the neural network approach using deep learning (Zhang et al., 2018; Araque et al., 2017).

Both of these approaches have been widely successful and are being used in several applications, which begs the question ”Can a model which combines the techniques of both traditional methods of natural language processing and neural techniques yield even better result in sentiment analysis as compared to using either one of the techniques?”

3 Background

The origin of research on Sentiment Analysis can be traced back to 60's. One of the earliest computer systems developed for sentiment analysis was General Inquirer [Stone and Hunt (1963)] which used punch cards to load data into the system and predict sentiments based on word-reference count and possibly co-reference.

With the onset of statistical machine learning techniques, researchers started applying some of these algorithms to propose more innovative solutions for sentiment analysis. Pang et al. (2002) published one of the earliest paper exploring the possibilities of using some of these techniques (Naive Bayes, maximum entropy classification, and support vector machines (Boser et al., 1992)) for sentiment analysis, which underlines various challenges in sentiment analysis.

Liu (2012) presents a very comprehensive research on sentiment analysis highlighting all the challenges, discussing possible multiple levels of sentiments which includes Document level, Sentence level and Entity and Aspect level, along with the possible approaches to tackle the problems in sentiment analysis.

Once computer systems became powerful enough and the techniques of deep learning came into application, researchers started comparing the performance of using statistical methods to that of the neural network methods. Moraes et al. (2013) present one of such comparisons, successfully showing that the neural method comfortably outperform one of the best statistical methods, SVM in prediction. Also, with increase in training data, the difference becomes even more pronounced.

The performance of neural methods have been quite satisfactory, however, the problem of sentiment analysis cannot be completely solved by just brute power of neural nets. It requires the finesse of the concepts of natural language understanding which often goes beyond the evidence presented in the training data and depends on world knowledge, the context of the paragraph, sometimes entire document. Even subtle choice of words in the sentence add rich connotations (implied meaning) to the sentence.

4 Reviewed papers

4.1 Motivation

Capturing and representing the implied meaning from a given corpus is quintessential for understanding true sentiments presented in the corpus. The papers selected for review, propose several solutions to solve this problem.

In *Connotation Frames: A Data-Driven Investigation* Rashkin et al. (2016) present a novel approach to formally represent connotations in English language based on verb predicates. The paper presents a new method to create embeddings which can then be learned to predict aspect-level or frame-level sentiment.

Rashkin et al. (2017) extend their work in *Multilingual Connotation Frames: A Case Study on Social Media for Targeted Sentiment Analysis and Forecast* from the previous paper and demonstrate that the connotation frames for English can be transferred to other languages as well, to make the solution multi-lingual.

While Rashkin et al. (2016, 2017) present efficient ways to represent connotations based on *verb* predicate, Allaway and McKeown (2021) propose a new embedding method in *A Unified Feature Representation for Lexical Connotations* to capture connotation from other parts of speech like *noun* and *adjectives* as well. The paper also presents a way to combine the techniques presented by Rashkin et al. (2016, 2017) for verb connotations and the connotations for other parts of speech to create a unified embedding model which can be then used to learn attention-based bidirectional model to predict the sentiment.

One of the challenges to develop a truly global and multilingual solution for sentiment analysis is the lack of enough annotated corpus in most of the languages. In fact, there are only a handful of languages with enough annotated lexicon to train an effective model. Buechel et al. (2020) present a very smart solution to translate and generate arbitrarily large annotated emotion lexicon for languages with scarce gold labels.

4.2 Connotation Frames: A Data-Driven Investigation (Rashkin et al., 2016)

4.2.1 Summary

One of the important aspects of sentiment analysis is to capture the unsaid sentiment. It requires world knowledge and the ability to understand the subtle connotations presented in the statements. While several sentiment analysis techniques capture the denotation (direct/literal) of the statement quite successfully, it is a difficult task for a machine to capture connotations presented in the statements by just training at the text/embedding.

This paper presents a novel approach to create a formal representation of the connotations for a given predicate along with a new crowd-sourced dataset. The idea presented in the paper, is to create a connotation frame which captures five sets of information which consists of writer's perspective, entities' perspective, effect value and mental states of the entities. In total, nine relationships are captured based on the five sets of information discussed above for all entities with each of these relationships having the sentiment value as one of (positive, negative or neutral).

Rashkin et al. (2016) propose two different models: 1) Aspect level and 2) Frame level. Aspect level model is trained on a 300-dimensional dependency based embedding vector (Levy and Goldberg, 2014). It is used to predict sentiment for each individual relation.

Frame level model is based on a newly introduced factor-graph algorithm as shown in Figure 1 in which overall sentiment of the predicate is predicted based on three factors: 1) Embedding factors, 2) Interdependency factors and 3) Belief propagation. Embedding factors is used to capture the results of aspect-level model for each of the nine nodes as shown in Figure 1. Interdependency factors are used to capture dynamics between different relation types and belief propagation is used to handle unseen words.

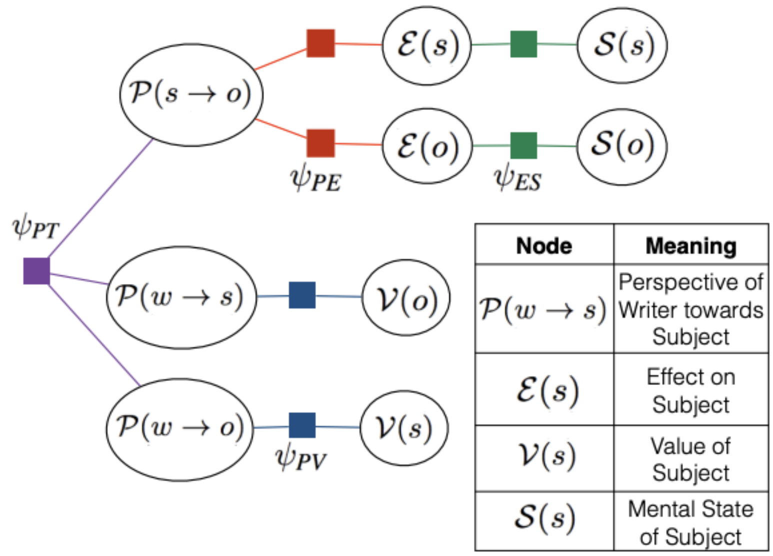

Figure 1: Figure depicting factor-graph model (Rashkin et al., 2016)

4.2.2 Assessment, Interpretation, Analysis

The experimental setup described in the paper is quite elaborate which uses New York Times corpus (Sandhaus, 2008) to create crowd-sourced annotations. Based on the results, (Rashkin et al., 2016) were able to demonstrate that both the proposed models (aspect-level and frame-level) outperform the selected baseline models.

4.2.3 Contributions

One of the major contributions of this paper is the approach to represent connotations as embeddings. There have been several embedding techniques like *word2vec* (Mikolov et al., 2013) and *fastText* (Bojanowski et al., 2017) which can efficiently represent literal meaning as embeddings, however, in case of sentiment analysis, the connotations provide very rich information and capturing that information may greatly improve the overall model performance by learning subtle hidden/implied meanings.

4.3 Multilingual Connotation Frames: A Case Study on Social Media for Targeted Sentiment Analysis and Forecast (Rashkin et al., 2017)

4.3.1 Summary

This paper is an extension of the above discussed paper by (Rashkin et al., 2016). The paper expands on the idea of connotation frames (see 4.2.1) and presents a solution to extend it to 10 different European languages in which the connotation frames from English is projected to all the different languages.

The major challenges in doing this projection is ensuring that the projections are being made in

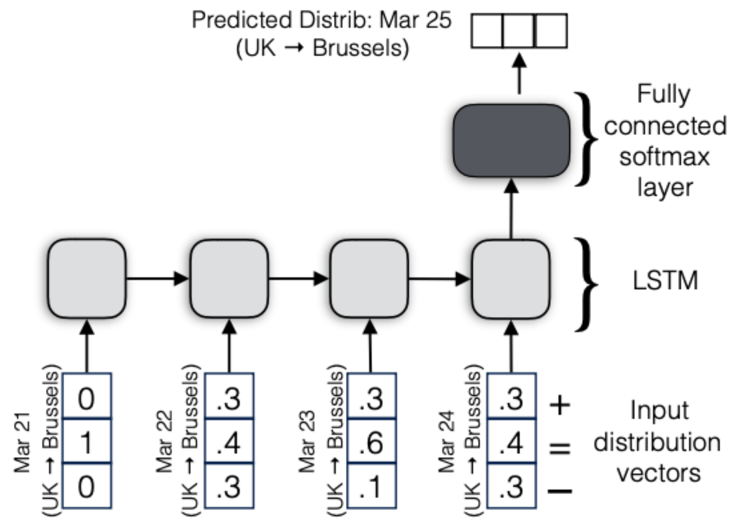

Figure 2: LSTM model architecture to predict location-based perspectives (Rashkin et al., 2017)

similar emotions across all the languages. To ensure this, the authors take a probabilistic approach to determine the most probable English word after translation from a target word. The connotation frames are transferred to the target word with highest translation probability.

The paper uses multi-lingual twitter dataset having certain specific tags like *#news*, *#breaking*, etc, and compute the distribution of the targeted sentiments. These distrubtions are then used in an LSTM (Hochreiter and Schmidhuber, 1997) model to forecast the sentiment dynamics for a given entity at a specific location for a given day (see Figure 2). For example, as shown in Figure 2, the distribution is used in the model to predict the sentiment dynamics of the word ”Brussels” in ”UK” on ”March 24” (Rashkin et al., 2017).

4.3.2 Assessment, Interpretation, Analysis

Rashkin et al. (2017) successfully demonstrate that the connotation frames might be transferred from English to other languages based on probabilistic translation. However, the number of languages explored is very limited (10). Also, the method assumes that there is a one-to-one mapping from target to source words for translation.

4.3.3 Contribution

Building upon the contributions in *Connotation Frames: A Data-Driven Investigation* (Rashkin et al., 2016), this paper presents a solution to project and use connotation frames in multiple languages. This greatly enhances the applicability of using connotation frames, since most of the real-world products using sentiment analysis are multi-lingual and having consistent prediction performance across all the languages is very important for usability.

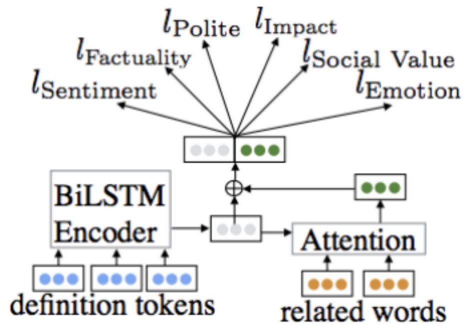

Figure 3: Model architecture to create unified connotation embedding (Allaway and McKeown, 2021)

4.4 A Unified Feature Representation for Lexical Connotations (Allaway and McKeown, 2021)

4.4.1 Summary

This paper takes the work of (Rashkin et al., 2016, 2017) in the papers discussed above, a step further and presents a unified connotation framework for all parts of speech. In (Rashkin et al., 2016, 2017), primary focus is given to *verb* and the connotation frames are formulated based on the verbs in the statements.

This paper employs similar idea to formulate connotation frames for nouns and adjectives. To represent connotations for nouns and adjectives, the paper introduces since new connotation aspects: 1) Social value, 2) Politeness, 3) Impact, 4) Factuality, 5) Sentiment and 6) Emotional aspect. The connotations generated for nouns and adjectives is then combined with the connotations generated for verbs into a single vector space.

The proposed solution uses an attention-based (Vaswani et al., 2017) bidirectional LSTM model (Schuster and Paliwal, 1997) as shown in Figure 3 to learn a dense connotation feature representation. This model jointly predicts all of the connotation labels for a given word and can be used to generated a simplified single connotation feature representation for all parts of speech.

The paper uses distant labeling to build the lexicon dataset and applies the proposed solution to *synonym analysis* and *stance detection*. The authors successfully demonstrate that the trained model based on the unified connotation embedding is well suited for tasks like stance detection.

4.4.2 Assessment, Interpretation, Analysis

The methodologies presented in this paper can truly act as the single solution to generate connotations for all parts of speech, instead of specific models for different types of words. The paper also presents a solution to create a new kind of embedding learned from the combined representation of connotation frames from all the words. This reduces the dimensions of the inputs for models to predict sentiments from the corpus.

4.4.3 Contribution

The contributions of this papers extends the contributions from above discussed papers *Connotation Frames: A Data-Driven Investigation (Rashkin et al., 2016)* and *Multilingual Connotation Frames: A Case Study on Social Media for Targeted Sentiment Analysis and Forecast (Rashkin et al., 2017)* by presenting a connotation representation for nouns and adjectives and providing a unified connotation embedding for all parts of speech.

4.5 Learning and Evaluating Emotion Lexicons for 91 Languages (Buechel et al., 2020)

4.5.1 Summary

While there is plenty of research done to analyze emotional lexicons, there is one major limitation. Most of these research is limited to a handful of languages and the lexicons sets are limited to small sizes in many cases. In this paper, (Bojanowski et al., 2017) propose an innovative technique to generate arbitrary number of emotional lexicons for any language with minimal assumptions.

The proposed method requires three individual and generally available components to be present which includes a source language emotion lexicon as the base for generation, a bilingual translation model which can translate from the given source to a specific target language and a target embedding.

The core idea presented in the paper is to create an architecture which can automatically generate a train/dev/test sets for the target language from any available source language, which can then be used to train a classification model to predict the emotion labels in the target language. As first step, the bilingual translation model is used to translate all the source words to the target language. The actual labels from the source embeddings are copied as-is for the translated target embeddings. These generated embeddings are then used to train a classification model to predict the emotion labels for the embeddings.

Once the model is trained, it is used to predict the labels for the translated embeddings to create a second set of target emotion labels (first set is copied directly from the source as discussed above). This ensures that the model is trained well enough to generalize for the entire translated embedding.

The authors were able to apply the proposed solution to generate emotion lexicons for 91 languages using the pre-trained models (Grave et al., 2018) based on *fastText* embeddings (Bojanowski et al., 2017).

4.5.2 Assessment, Interpretation, Analysis

To be able to generate the emotional lexicons for 91 different languages with minimal gold labels is a truly commendable achievement and will definitely help in advancing further research on truly global sentiment analysis. Also, the paper demonstrates that even without the gold labels, the embeddings generated for the lexicons are comparable to gold labels and generalize well for all the languages.

4.5.3 Contribution

This paper introduces a very simple and yet very effective solution to generate emotion lexicons for any number of languages, specially for the languages which lack gold labels. It can greatly impact the application of sentiment analysis to many of the less documented languages.

5 Conclusion

Clearly, a lot of progress has been made to formally represent both connotations and denotations of a given corpus as embeddings or similar forms which can be learned using neural nets, in an effort to capture both direct and implied sense presented in the data and subsequently understand and extract the true sentiment from the corpus.

In all the papers reviewed above, the methodologies follow similar pattern. In first step, the focus is given to the linguistic aspects to capture and represent the complete sense of the corpus as embedding, and in the second step, these embeddings are used as inputs to different architectures of neural nets including LSTM (Hochreiter and Schmidhuber, 1997), Bidirectional LSTM (Schuster and Paliwal, 1997), attention models (Vaswani et al., 2017), etc.

One of the noteworthy observations from all the reviewed papers is that, all the words cannot be treated equal for an efficient model. Including the linguistic aspects like the part of speech, related words, etc greatly improve the model performance and it is closer to the way a human understands the language. And with the current pace of the research, sooner than later, machine learning models will have comparable or even better performance in understanding the true sentiment from a given corpus as compared to a human being.

References

Oskar Ahlgren. 2016. Research on sentiment analysis: The first decade. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 890–899.

Emily Allaway and Kathleen McKeown. 2021. A unified feature representation for lexical connotations. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2145–2163, Online. Association for Computational Linguistics.

Oscar Araque, Ignacio Corcuera-Platas, J Fernando Sánchez-Rada, and Carlos A Iglesias. 2017. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Applications*, 77:236–246.

Mahsa Badami, Olfa Nasraoui, Welong Sun, and Patrick Shafto. 2017. Detecting polarization in ratings: An automated pipeline and a preliminary quantification on several benchmark data sets. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 2682–2690.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. 1992. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152.

Sven Buechel, Susanna Rücker, and Udo Hahn. 2020. Learning and evaluating emotion lexicons for 91 languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1202–1217, Online. Association for Computational Linguistics.

Kerstin Denecke and Yihan Deng. 2015. Sentiment analysis in medical settings: New opportunities and challenges. *Artificial intelligence in medicine*, 64(1):17–27.

Nees Jan van Eck and Ludo Waltman. 2007. Vos: A new method for visualizing similarities between objects. In *Advances in Data Analysis*, pages 299–306, Berlin, Heidelberg. Springer Berlin Heidelberg.

Ronen Feldman. 2013. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82–89.

Edouard Grave, Piotr Bojanowski, Prakhar Gupta, Armand Joulin, and Tomas Mikolov. 2018. Learning word vectors for 157 languages. *arXiv preprint arXiv:1802.06893*.

Tanjim Ul Haque, Nudrat Nawal Saber, and Faisal Muhammad Shah. 2018. Sentiment analysis on large scale amazon product reviews. In *2018 IEEE international conference on innovative research and development (ICIRD)*, pages 1–6. IEEE.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Omer Levy and Yoav Goldberg. 2014. Dependency-based word embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 302–308.

Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Anshul Mittal and Arpit Goel. 2012. Stock prediction using twitter sentiment analysis. *Standford University, CS229 (2011 http://cs229. stanford. edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis. pdf)*, 15.

Saif M. Mohammad, Xiaodan Zhu, Svetlana Kiritchenko, and Joel Martin. 2015. Sentiment, emotion, purpose, and style in electoral tweets. *Information Processing Management*, 51(4):480–499.

Rodrigo Moraes, João Francisco Valiati, and Wilson P GaviãO Neto. 2013. Document-level sentiment classification: An empirical comparison between svm and ann. *Expert Systems with Applications*, 40(2):621–633.

Subhabrata Mukherjee and Pushpak Bhattacharyya. 2012. Feature specific sentiment analysis for product reviews. In *Computational Linguistics and Intelligent Text Processing*, pages 475–487, Berlin, Heidelberg. Springer Berlin Heidelberg.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of EMNLP*, pages 79–86.

Hannah Rashkin, Eric Bell, Yejin Choi, and Svitlana Volkova. 2017. Multilingual connotation frames: A case study on social media for targeted sentiment analysis and forecast. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 459–464, Vancouver, Canada. Association for Computational Linguistics.

Hannah Rashkin, Sameer Singh, and Yejin Choi. 2016. Connotation frames: A data-driven investigation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 311–321, Berlin, Germany. Association for Computational Linguistics.

Evan Sandhaus. 2008. New york times corpus: Corpus overview. *LDC catalogue entry LDC2008T19*.

M. Schuster and K.K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681.

Philip J Stone and Earl B Hunt. 1963. A computer approach to content analysis: studies using the general inquirer system. In *Proceedings of the May 21-23, 1963, spring joint computer conference*, pages 241–256.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Vishal Vyas and V Uma. 2019. Approaches to sentiment analysis on product reviews. In *Sentiment Analysis and Knowledge Discovery in Contemporary Business*, pages 15–30. IGI Global.

Wei Wei and Jon Atle Gulla. 2010. Sentiment learning on product reviews via sentiment ontology tree. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 404–413.

Lei Zhang, Shuai Wang, and Bing Liu. 2018. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4):e1253.