CS447 Literature Review: Sentiment analysis in NLP

Prashanth Giriyappa Shankarappa,
pg14@illinois.edu

December 6, 2021

Abstract

People often express their option and stance through their choice of words. Words chosen often have *connotation* of implied sentiment beyond literal meaning towards entities in a context. Sentiment analysis tries to extract this information from text and classify writer's views by analyzing polarity of words in a particular context. In this literature review we look at few approaches to Sentiment Analysis for English and other languages. We look at using *connotation frames* for verb predicates for English and non-English and a unified feature representation that adds additional relationships for nouns and adjectives.

1 Introduction

Sentiment can be defined as a view or orientation towards an object from a writer's perspective. People express these sentiments through social media, writers express their view through articles about events. A writer can influence a readers view by their choice of words. The implied sentiment can also convey writer's sentiment towards parties involved in an event. Sentiment analysis tries to extract this information from text and classify writer's views by analyzing polarity of words in a particular context. Sentiment Analysis is used to extract and analyze views from content. We have observed news media reporting the same information with words that convey subtle biases or sentiments towards a subject or object. How can natural language processing(NLP) be used for Sentiment analysis to help classify opinion or view expressed by the author? What are challenges when dealing with Sentiment classification?

2 Background

Understanding connotation is key to deciphering the subtle sentiment or bias expressed by an author. Words may have similar meanings but can represent different connotations - positive, neutral or negative. **Connotation** of a word refers to implied sentiment that goes beyond literal meaning. For instance, unique*(positive)* and peculiar*(negative)*, innocent*(positive)* and immature*(negative)* even though have similar meaning can impart different opinions or views. Positive or negative analysis is referred to as **sentiment**.

Emotion can be considered multi-dimensional and can be considered to have 8 dimensions - {anger, joy, fear, trust, anticipation, sadness, disgust, surprise}.

Analysis of sentiment as intended by the author, can express a positive, neutral or negative notion about an object, event or situation. Sentiment analysis can thus be considered as a **classification** task. **Generative classifier** like *Naive Bayes* can be used to model given a set to labeled data.

However, this model is not ideal in cases where there is insufficient training data or when there are negation words in the corpus. Neural models have been used for sentiment analysis which has resulted in considerable improvements. Few neural models are mentioned in approaches below:

- BiLSTM Graves et al. (2005) Long Short Term Menmory(LSTM) is a type of Recurrent Neural networ(RNN) which have gates that can be used to remember or forget past hidden states. With Bidirectional LSTM propagates input in forward and backward directions.
- BERT Devlin et al. (2019) BERT stand for Bidirectional Encoder Representation for Transformers. BERT architecure is a stacked directional transformer encoder model Vaswani et al. (2017).

3 Connotation Frames

Description Writers or Authors express their option or biases through their choice of words. Predicate chosen usually have implied sentiment attached to them which may not be explicitly mentioned. The writer with his opinion tries to convey a message and influence readers perspective. Connotation frames Rashkin et al. (2016) is a representation that uses multiple dimensions of connotation. A verb predicate can be used to express sentiment about entities involved. Below are two example sentences:

- A violated B's airspace
- Family is reunited with their dog

In the first sentence, some sentiment is implied by the use to verb *violated* and toward involved entities:

1. B is being portrayed as victim
2. A as the perpetrator
3. Violated had an effect on B
4. B may be disturbed by the event

In the second example,

1. Family is being portrayed as victim
2. Sympathetic towards the dog
3. Reunited had a positive effect on Family
4. Family and dog are both happy that they were united

In this review, different relationships that form connotation frame is explained with a representative example. These relationships can be assigned different polarities - positive, negative or neutral. Connotation frames are modeled as 9 relationship aspects. Some of these relationships have inter dependencies which can be inferred. We can then model connotation frame using aspect-level and frame-level. Aspect-level predicts polarity for each relationship aspect independently. For a predicate, aspect-level polarities are used to calculate frame-level assigns polarity which takes into account the interaction between different aspects.

3.1 Connotation Frame Representation

Authors represent connotation frame using 4 different relations:

1. perspective - P(x→y) perspective of x towards y. P(w→agent),P(w→theme), P(agent→theme)
2. value - V(x) value of x. V(agent), V(theme)
3. effect - E(x) effect of event on x either good or bad. E(agent), E(theme)
4. mental state - S(x) Mental state of x as a result of event

Each of these typed relations can have polarities of positive, negative or neutral {+, -, =}. A representation is presented in Figure 3.

Example of typed relations

Verb	Perspective	Value,Effect and State	Sentence
Reunited	P(w→family)=+ P(w→dog)=+ P(family→dog)=+	E(family)=+ V(family)=+ S(family)=+ E(dog)=+ V(dog)=+ S(dog)=+	Family is reunited with their dog

Table 1: Showing specific example.

This example shows relation of verb *Reunited* to entities *family* and *dog*. It shows the perspective P(x → y) from writers perspective to both subject(agent/family) and object(theme/dog) and to each other. Connotation for Value, Effect and Mental state are also captured for agent and theme.

3.2 Using data to analyze

In any language, meaning depends on context in which it is used, so exact connotation may also vary depending on the context. There could be differences in connotation from different sources or across languages. However, there are still common understanding of most connotation for many predicates. Authors used Subjectivity Lexicon (Wilson et al., 2005) to analyze verbs and their effect on Agent(Subject) and theme(Object). They found consistency in what was observed for words including suffer, guard and uphold. For suffer, 64% of words in agent position had polarity of positive, while 94% of words in theme position had negative. This indicates for word suffer, objects are most likely to have a negative perspective because they would have suffered from an event.

3.3 Inter dependencies

Polarity assignments are inter dependent. Consider the following inter dependencies

- Perspective - writer, agent and theme: If writer feels positively towards agent and negatively towards theme, then polarity of agent towards them and vice versa may not be positive.

- Perspective - Effect: If predicate has positive effect on subject, perspective between Subject and Object was positive
- Perspective - Value: If A is valuable, then writer view A positively. If A is not valuable, writer views A negatively.
- Effect - Mental state: If predicate has positive effect on A, then A will experience positive mental state from the event. Likewise for negative effect, mental state is also negative.

3.4 Modeling

Model should be able to predict relations. Authors use two levels for modeling relations - *Aspect-Level* and *Frame-Level*. A predicate can have 9 relationship aspects:

- Perspective: P(w→s),P(w→o), P(s→o)
- Effect: E(s), E(s)
- Value: V(s), V(o)
- Mental state: S(s), S(o)

Aspect-Level model makes prediction about the relations independently based on context in which the predicate appears. It uses 300 dimensional word embedding and Maximum Entropy*MaxEnt* classifier to maximize average F1 score.
Frame-Level model makes prediction based on connotation frames collectively that are associated with a predicate. It used a factor graph as shown below in Figure 2. Graph contains 9 nodes each containing relation and each can have different polarities {+,-,=}

Y_i is the relational aspect of i^{th} verb predicate. $P(Y_i)$:= Probability of assignment of polarity to node in Y_i is the product of individual function probabilities given below

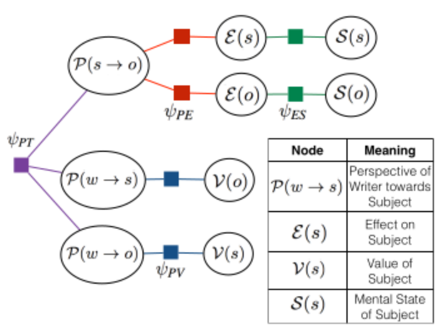

Figure 1: Probability assignment $P(Y_i)$ from (Rashkin et al., 2016, figure 2)

$$
\begin{aligned}
P(\mathbf{Y}_i) \propto\ & \psi_{\mathrm{PV}}(\mathcal{P}_{ws}, \mathcal{V}_s)\, \psi_{\mathrm{PV}}(\mathcal{P}_{wo}, \mathcal{V}_o) \\
& \psi_{\mathrm{PE}}(\mathcal{P}_{so}, \mathcal{E}_s)\, \psi_{\mathrm{PE}}(\mathcal{P}_{so}, \mathcal{E}_o) \\
& \psi_{\mathrm{ES}}(\mathcal{E}_s, \mathcal{S}_s)\, \psi_{\mathrm{ES}}(\mathcal{E}_o, \mathcal{S}_o) \\
& \psi_{\mathrm{PT}}(\mathcal{P}_{wo}, \mathcal{P}_{ws}, \mathcal{P}_{so}) \prod_{y \in \mathbf{Y}_i} \psi_{emb}(y)
\end{aligned}
$$

Figure 2: P(Yi) from (Rashkin et al., 2016, sec. 3.2)

Figures, Tables, References
Paper - Rashkin et al. (2016)
Sections - (Rashkin et al., 2016, sec. 1-3)

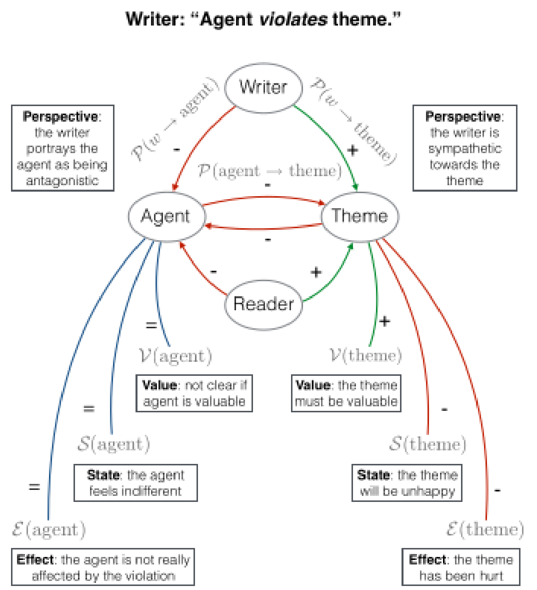

Figure 1: An example connotation frame of "violate" as a set of typed relations: perspective $\mathcal{P}(x \rightarrow y)$, effect $\mathcal{E}(x)$, value $\mathcal{V}(x)$, and mental state $\mathcal{S}(x)$.

Figure 3: Connotation frame from (Rashkin et al., 2016, figure 1)

4 Multilingual Connotation Frames

Description We have already seen how *Connotation frames* Rashkin et al. (2016) can be used to facilitate Sentiment analysis for English verb predicates, *multilingual connotation frames* Rashkin et al. (2017) aims to extend it to 10 additional European languages. With social media readily accessible across geographies and languages people express their views and reflections through social media. An event like *Brussels bombing* can invoke reaction on social media that could last for days across geographies and languages. In this paper authors propose creating a multilingual lexicon for non-English languages by location(country of origin of tweet) while trying to forecast sentiment on events for the next few days after the event first occurred. By using concepts of connotation frame and targeted sentiment, a multilingual lexicon is created. While this is a specific case study, this literature review does not include prediction and sentiment dynamics from the paper. This approach of using Connotation frames is a good foray that would enable further analysis into multilingual sentiment analysis.

The below review defines connotation frames and targeted sentiment. It also shows the approach to modeling multilingual connotation frame based on connotation frame of corresponding English verb. It uses twitter dataset to extract tweets from reliable sources to build tuples of (agent,verb,theme) in multiple languages. Context based projection of English connotation frames is performed on 10 additional European languages to assign most probable English connotation frame to non-English predicate. Targeted sentiment can then be assigned to verb predicate.

Definition 1 - Connotation frames For a given verb predicate, connotation frame expresses rela-

tionships between predicate towards its arguments as given in Figure 4 for verb survive. Connotation frames are relationships between predicate and entities which will aid in assigning targeted sentiment.
Definition 2 - Targeted Sentiment Sentiment of a source towards a target is a targeted sentiment. In the example in Figure 4, few targeted sentiments are:

1. sentiment(writer→teenager) as *victim(positive)*
2. sentiment(writer→bombing) as *effect(negative)*
3. sentiment(teenager→bombing) as *disturbed(negative)*

4.1 Multilingual twitter dataset

Multilingual twitter dataset was used for 15 days duration after Brussels attack. This dataset contains location information about the origin of tweets. Tweets from trusted news sources and certain hashtags were chosen. By using SyntaxNet depedency parser Andor et al. (2016) 10 non-english languages were trained resulting in agent-verb-them tuples for these languages depicted . About 1.2 million tuples were extracted for all these languages as in Figure 5.

4.2 Building Multilingual connotation frame

To build connotation frame for 10 European languages languages parallel corpora is used - Opus Corpus Tiedemann and Nygaard (2004), Multi-UN parallel data Eisele and Chen (2010) and EuroParl parallel data Koehn (2005).
For each non-English verb *v'* probability of it being translated to a verb *v* in English is calculated. Connotation frame for a Non-English verb is the connotation frame of English verb *F(v*)* with the highest translation probability *F(v')*.

$$v^* = \mathrm{argmax}_v\, p(v \mid v')$$
$$F(v') = F(v^*)$$

For example, connotation frame for dutch word *buit* may come from connotation frame for English word *loot*.

4.3 Extracting targeted sentiments

By using connotation frame lexicon generated for different languages, polarity is calculated for most frequently used named entities by aggregating sentiments of writers location in each country, e.g, distribution of positive, neutral and negative perspectives towards *Obama* in *dutch* tweets. Aggregated polarities are represented as 3 dimensional probability vector p=[$p^+ p^= p^-$]. This vector is used to predict sentiment forecast for the next few days.

Figures, Tables, References
Paper - Rashkin et al. (2017)
Sections - (Rashkin et al., 2017, sec. 1-3)

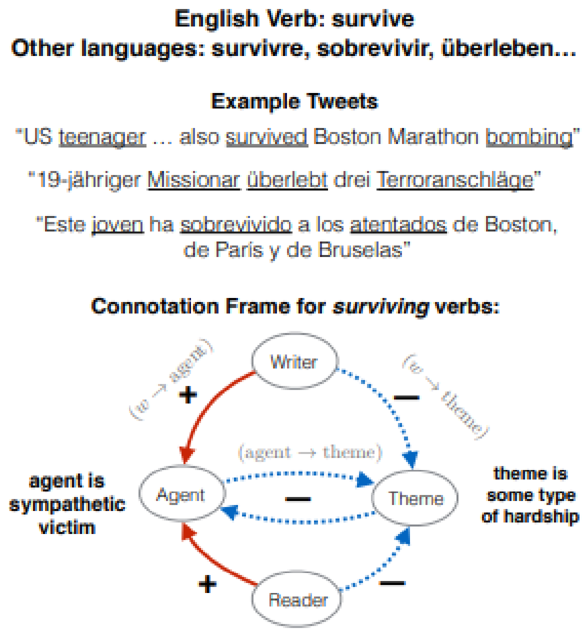

Figure 4: Connotation frame of *"survive"* from (Rashkin et al., 2017, figure 1)

Lang	# Tuples	Examples
EN	643,004	(korea, fires, missile)
ES	305,310	(acuerdo, vulnera, derecho)
FR	85,286	(obama, quitte, cuba)
PT	76,849	(valentino, renova, contrato)
RU	28,511	(путин, обсуди, мостa)
DE	23,197	(seehofer, lösen, flüchtlingskrise)
NL	14,091	(artiesten, bevestigd, komst)
IT	13,586	(conte, lascia, nazionale)
FI	2,859	(hans, tekemään, valintansa)
SV	2,229	(rubio, avslutar, kampanj)
PL	2,226	(papież, wyrazi, zgodę)

Figure 5: Number of language tuples(agent,verb,theme) from (Rashkin et al., 2017, table 1)

5 Unified Feature Representation for Lexical Connotations

Description Unified Feature Representation for Lexical Connotations Allaway and McKeown (2021) can be perceived as an extension of *Connotation frames* Rashkin et al. (2016). Connotation frames defined aspects like perspective, effect, value, mental state revolving around verb predicates. While Connotation Frames paper focused on Verb predicate, this paper defines connotation aspects to include nouns and adjectives. It also introduces finer details for polarity. The result is a labeled lexicon for English that includes words from all parts of speech.

In this review, meaning of 6 new aspects for noun and adjective is stated. With the help of available lexicons, new aspects for words are are labeled with connotation to generate a new lexicon. The resulting lexicon has over 7000 words that are fully labeled for all aspects and over 93k words labeled with some aspects. New lexicon and verb lexicon are both used to train an BiLSTM model for all parts of speech. This approach fills the gap that was left behind in the Connotation frame modeling which focused on verb predicates.

5.1 Definition

Six new aspects are defined to include nouns and adjectives.

- Social value: Similar to *"value"* relation in Connotation frame, Social value represents both objects and concepts. They are objects and concepts that people value. Value also now includes 'power' for verbs and subjects that are in context.
- Politeness: These are words that are pleasant towards the addressee and are also formal. Words like 'admire' can be considered polite. This aspect is extended to noun and adjectives as they were part of verb connotation frames.
- Impact: Impact is the effect word have. Verbs already have arguments, this aspect is extended to nouns and adjectives though they don't have arguments. For instance, 'charity' though doesn't take any arguments has an impact on society.
- Factuality: It captures if words are real world concepts, attributes or objects.
- Sentiment: Sentiment conveys overall connotation
- Emotional associations: Different types of emotions which conveys overall connotation. Ex. anger, anticipation, disgust, fear, joy, sadness, surprise, trust

5.2 Creating a lexicon

Lexica was created by combining different lexicons each of which were chosen to represent different connotation aspects.

- Social Value, Politeness, and Impact: Harvard General Inquirer (Stone and Hunt, 1963)
- Factuality: 'Imagery' dimension from Dictionary of Affect in Language (Whissell, 2009)
- Polarity: Connotation WordNet (Kang et al., 2014)
- Emotional Association: 8 emotions from the NRC Emotion Lexicon (Mohammad and Turney, 2013) in Language (Whissell, 2009)

Result was a lexicon with 7,578 fully labeled words for all aspects and about 93K words labeled for some aspects. Total resulting words were 100,176. Emotion was the only multi-class aspect with each emotion labeled l ε {0,1}. All other aspects had label l ε {-1,0,1}.
It was noted that aspects exhibit uneven distributions. For instance, 10.5% words were considered polite while 1% impolite and 32% were considered to have social value while 15.5% were not.
Word-sense level labeling was not used since lexias mentioned above are word-sense independent.

5.3 Human evaluation of Lexica

To evaluate the lexia, labels generated by distant supervision were compared against labels generated by human annotators. It was observed that labels match 64.2% of the time. For non-conflicting value pairs - (+, neutral) and (-,neutral), the math was about 90%.

5.4 Modeling

A combined lexica with lexica produced for nouns and adjectives and 2 other verb lexica were used to create a single vector space.
A bi-directional LSTM model is used to predict 6 aspects - Social value, Politeness, Impact, Factuality, Sentiment and Emotion, for each word. For each word w, a set of definition words D_{w^t} and related words R_{w^t} are passed to the encoder of the model. The encoder produces a connotation feature v_{w^t} from D_{w^t} and R_{w^t} of 300 dimension. v_{w^t} is used to predict label l_a for each aspect a. Below is a representation of the model used Figure 6

Figures, Tables, References
Paper - Allaway and McKeown (2021)
Sections - (Allaway and McKeown, 2021, sec. 1-4)

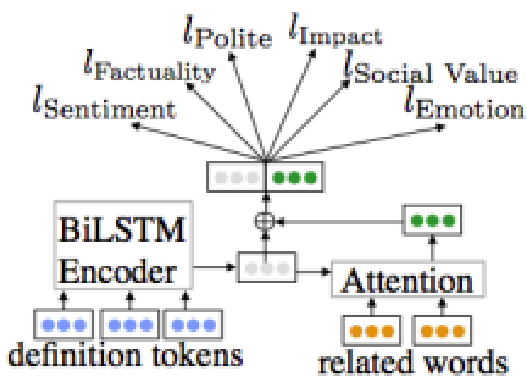

Figure 6: Connotation embedding modeling from (Allaway and McKeown, 2021, figure 2)

6 Assessing and probing sentiment classification

Description While previous papers dealt with Connotation towards entities, this paper uses with Sentence-level sentiment analysis using Neural networks. Neural models have been used for Sentiment Analysis which have led to improvements over previous approaches. Neural models have achieved over 90% accuracy for sentence-level sentiment analysis. Even with these improvements there are still challenges when dealing with linguistic phenomena.
Negation, Polarity shifters and Modality have been studied for Sentiment analysis on n-gram based models but they haven't been studied on Neural networks.
Negation which has been studied for sentiment analysis, can transform sentiment from one to the other. For example, **'not'** in sentence *'He did **not** pass the test'* changes polarity from + to -.
Verbal polarity shifter is similar to Negation since they also change the polarity. They can act on

positive and negative statements. For example, **'failed'** in *'He failed$_{shifter}$ the test'*.
Modality is used to express possibility. Modal verbs like *can, could, must, shall, should* when accounted for in Sentiment classification can improve results.
This paper attempts to discover and classify errors in Sentiment analysis for English. This review is organized into different sections:

- Dataset creation
- Models used for Sentiment analysis
- Dataset analysis - Error annotation for 18 linguistic and paralinguistic phenomena

A subset of sentences which are considered as gold label from different datasets are collected and used on different models. The setup uses three Neural models - BERT Devlin et al. (2019), ELMo Peters et al. (2018) and BiLSTM Barnes et al. (2017). A non-neural bag-of-words classifier is also part of the setup. The sentences are annotated into 19 linguistic and paralinguistice labels. Sentences whose polarities are misclassified by each model forms the error dataset. The resulting errors in classification is then manually annotated.

6.1 Dataset creation

Subset of sentences from 6 English datasets which are annotated with sentence level polarity are collected. Sentences are annotated with 5 polarities [++,+,0,–,-] for *Strongly positive, Positive, Neutral, Negative and Strongly negative*. Refer to Figure 7 for statistic about sentence-level annotations.

6.2 Models

Three neural models listed below were trained on all datasets to analyze errors that are generated from them. A bag-of-words classifier is also trained which forms a good reference for text classification.

- **BERT** Devlin et al. (2019) model is a birectional transformer that is trained on each sentiment dataset.
- **ELMo** Peters et al. (2018)
- **BiLSTM** Barnes et al. (2017) A single layer bidirectional LSTM is pretrained with skipgram embeddings which creates a sentence by concatenating the final hidden layer from left and right directions. Softmax layer is used for classification.
- **Bag-of-words classifier** A support vector machines(SVM) classifier is trained on bag-of-words representation of training sentences. SVM is a supervised machine learning algoritm that can be used for classification.

Out of the four models, BERT and ELMo were the best performing with great accuracy of results.

6.3 Dataset analysis

Resulting error sentences from these models are categorized into 19 categories. Categories are listed in descending order of number of errors encountered after classification task was completed.

- **Mixed polarity** These were the largest set of errors. Sentences with errors had two different polarities expressed in the same sentence that can be associated with two different entities or same entity. For example, *'The plot was good, but a little slow'*.
- **Non-standard spelling** Many of sentences had hash-tags. Example sentence - *'It was veeery gooood.'*
- **Idioms** Sentences that contained idioms were also misclassified. Example sentence - *'It was raining cats and dogs yesterday.'*
- **Strong labels** Strong positive or negative are usually harder to classify since they can expressed with different vocabulary. Consider the sentence *'These items should be put in a box and buried in a deep trench.'*, which indicates strong negative polarity is difficult to classify.
- **Negation** It inverses the polarity of a sentence and is thus difficult to classify. Some negation was correctly classified while many were not. An analysis of incorrect classification indicated that when negator occurs far away from negated element, it becomes difficult to classify. Example, *'I don't think that this is a particularly interesting idea.'*
- **World knowledge** Sentences that require world knowledge when doing a comparison. *'I would give it 10 out of 10.'*
- **Amplifiers** These are sentences that are strongly positive or negative polarities. *'It was a really good movie.'*
- **Comparative** When sentence contains multiple entities and we have to determine entity that for which a certain polarity applies. *'Between the two, I think the first drink was better than the second.'*
- **Sarcasm/Irony** It is required to correctly capture Sarcasm to classify negative and strongly negative sentences.
- **Shifters** When certain words are mentioned in a sentence they shift polarity. Words like 'fail','reject' are not common in the dataset but move polarity from positive to negative.
- **Emoji** Chosen models mostly classify sentences with emojis correctly but they fail on negative examples.
- **Modality** Dataset contains sentences with modal verbs that express different sentiment than the same sentence without modal verbs, e.g., "I would have liked the room, if it had been bigger."
- **Morphology** Concatenated words could lead to strong positive or negative polarity, e.g., *"It was fan-freakin-tastic"*
- **Reducers** Reducers like 'less', 'kind of' occur with both positive and negative examples, e.g., *"It was a lot less of a hassle"*

Figures, Tables, References
Paper - Barnes et al. (2017)
Sections - (Barnes et al., 2017, sec. 1-5)

Label	MPQA	OP.	Sem.	SST	Ta.	Th.
++	–	379	–	1,852	–	–
+	193	879	3,499	3,111	923	2,727
0	527	–	4,478	2,242	1,419	1,779
–	413	399	1,310	3,140	1,320	1,828
– –	–	74	–	1,510	–	–
Total	1,133	1,731	9,287	11,855	3,662	6,334

Figure 7: Statistics for the sentence-level annotations in each dataset. (Barnes et al., 2017, Table 1)

7 Discussion

Connotation Frames Rashkin et al. (2016) Sentiment analysis using Connotation frame representation is a good approach to organizing connotation towards entities involved. With Perspective interaction between writer, subject and object is captured. Effect and Value captures polarity with reference to Subject. Mental state captures subjects Mental state. We may have observed biases and choice of words, for instance, when left or right leaning news media reports the same information. With this representation, inferred sentiment can be assessed that the writer is trying to convey towards entities involved. The paper reveals that taking *"Obama"* as example, right leaning media portrays as someone who *attacks or criticizes* while left leaning media portrays him as the victim of those *attacks or criticism*. For the experiment, authors used human annotators to annotate polarity for 9 aspects and compared that against results produced from the model. Overall, percentage agreement between human annotators and what is predicted by the model is encouraging. Connotation frames can thus be used to understand writer's view or opinions expressed towards relevant entities.

Multilingual Connotation Frames Rashkin et al. (2017) Concept of Connotation frames was extended to include 10 additional European languages. Through the use of parallel corpa for multiple languages, connotation frame for non-English language predicate was mapped from English word by maximizing probability distribution.

Unified Feature Representation for Lexical Connotations Allaway and McKeown (2021) Sentiment analysis using Connotation frames focused on *verb* predicates while Unified feature representation added additional aspects for *nouns and adjectives*. The 6 new aspects or relations addresses gaps in Connotation frame representation. It also uses emotion which is a multi-dimensional feature. The new lexicon which is created adds these new dimensions to further enhance human opinion assessment. While connotation frames and Unified feature representation both provide means to assess human judgements for languages, further analysis is required to explore interaction between these aspects and other linguistic aspects like context and word sense.

Assessing and probing sentiment classification Barnes et al. (2017) While sentiment analysis has been studied there still are challenges with some lingusitic phenomena as mentioned in section 6.3. Building a dataset to identify and categorizing are the first steps towards finding a solution that could address these phenomena. Some features of languages like Idioms, Emoji, Morphology may challenging because in English you can generate words that don't normally exist in the vocabulary. Languages do tend to differ in topological structure and syntax. While many of these categories may exist in other languages they may also include additional linguistic phenomena.

8 Conclusion

In this review, four papers were presented - Connotation Frames sec. 3, Multilingual Connotation Frames sec. 4, Unified Feature Representation for Lexical Connotations sec. 5 and Assessing and probing sentiment classification sec. 6.
This review explores how natural language processing(NLP) can be used for Sentiment analysis to help classify opinion or view expressed by the author and Challenges with Sentiment classification. Connotation frames are a novel approach that can be used to classify implied sentiments by way of assessing relationships between events and entities. Connotation frames can also be extended to include additional non-English languages, however, they are limited to verb predicates. With Unified Feature representation, nouns and adjectives can also be used to classify connotation. They also add new relations include multi-dimensional feature representation for emotion aspect. More research is however, needed to understand interaction between these aspects and other linguistic aspects like context and word sense. While sentiment analysis can be performed at sentence-level, there are multiple challenges which have been categorized. Sentence-level sentiment analysis may include additional challenges when for other languages because of linguistic phenomena that are associated with those languages.

References

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.

Daniel Andor, Chris Alberti, David Weiss, Aliaksei Severyn, Alessandro Presta, Kuzman Ganchev, Slav Petrov, and Michael Collins. 2016. Globally normalized transition-based neural networks. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2442–2452, Berlin, Germany. Association for Computational Linguistics.

Jeremy Barnes, Roman Klinger, and Sabine Schulte im Walde. 2017. Assessing state-of-the-art sentiment models on state-of-the-art sentiment datasets. In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 2–12, Copenhagen, Denmark. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human*

Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Andreas Eisele and Yu Chen. 2010. Multiun: A multilingual corpus from united nation documents.

Alex Graves, Santiago Fernández, and Jürgen Schmidhuber. 2005. Bidirectional lstm networks for improved phoneme classification and recognition. In *Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005*, pages 799–804, Berlin, Heidelberg. Springer Berlin Heidelberg.

Philipp Koehn. 2005. Europarl: A parallel corpus for statistical machine translation. In *Proceedings of Machine Translation Summit X: Papers*, pages 79–86, Phuket, Thailand.

Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, New Orleans, Louisiana. Association for Computational Linguistics.

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.

Jörg Tiedemann and Lars Nygaard. 2004. The OPUS corpus - parallel and free: `http://logos.uio.no/opus`. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal. European Language Resources Association (ELRA).

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need.