Aspect Based Sentiment Analysis for Competitive Market Intelligence

Saurabh Shirodkar and **Arvind Govindarajan** and **Yuchen Cai**
College of Information and Computer Sciences
University of Massachusetts Amherst

Abstract

In this report, we discuss the task of aspect based sentiment analysis applied to a dataset of reviews of products in the SBD portfolio. The task was carried out in an unsupervised manner to extract aspects within reviews followed by the application of sentence level sentiment analysis to analyze the sentiment associated with aspects. We present the findings via interactive visualizations- a spider chart and a qualitative textual view.

1 Introduction

1.1 Task Description

The objective of this project is to be able to provide signals to Stanley Black and Decker based on review data that can be used to make investment decisions on new products or updating old products. The signals are derived by analyzing sentiment on product-level and category-level features (aspects) that can provide the company with a consumer perspective of product features and expectations.

In order to provide the aforementioned signals, the natural questions that arise are- (1) how to train a model to recognize features, and (2) how to classify sentiment on these features? From a research point of view, the questions we are trying to answer are- (1) how to apply unsupervised aspect term extraction to a large, diverse, and non-annotated dataset, and (2) how to classify sentiment polarity on the extracted features?

The task of Aspect Based Sentiment Analysis (ABSA) has been widely researched since 2014 largely due to the SemEval tasks of 2014 (Pontiki et al., 2014), 2015 (Pontiki et al., 2015), and 2016 (Pontiki et al., 2016).

The SemEval tasks have steered the current efforts in ABSA but the problem with existing research is two-fold:

1. The bigger task of ABSA can be subdivided into smaller tasks of (i) Aspect Term Extraction (Opinion Term Extraction), and (ii) Aspect Sentiment Detection/Polarity Classification. Researchers are pursuing each sub-task individually in its own right. Thus, to address the larger task of ABSA, one needs to create a pipeline merging the two sub-tasks which is cumbersome and prone to errors that propagate through the pipeline and accumulate across tasks (Li et al., 2019).

2. Most of the research being conducted is bound by the lack of annotated datasets. This is limiting application of research efforts mostly in a supervised and semi-supervised setting with application of methods to predefined domains. This scenario can be described as a domain lock-in. Practically speaking, many datasets are unlabelled and the task of manually annotating the datasets just to be able to apply the current SOTA models for ABSA isn't a viable solution.

1.2 Technical Challenges

Due to the very nature of our dataset being unlabelled, we faced several challenges in finding an appropriate model to carry out the task of ABSA. We conducted several iterations of literature survey as the problem statement was iteratively refined. Initially, we spent time looking into the current SOTA models for ABSA. We were quick to realize that most of the SOTA models were supervised models and this wasn't applicable to our problem. We then considered tackling the problem in a semi-supervised setting (with the annotation of a subset of the dataset) but were quick to shut down the idea as annotation can be cumbersome.

Research papers that address the task of ABSA in an unsupervised setting are few and far between. Before landing on the unsupervised model that we employed for this project, our efforts were primarily focused on solving the problem by applying domain transfer techniques through Selective Adversarial Learning as described in the paper (Li et al., 2019). We focused our efforts into that model because it is end to end. It performed the tasks of aspect extraction and aspect sentiment analysis with a single model and is the first of its kind. While this model is unsupervised and achieves SOTA results on the SemEval tasks, we decided to pursue alternate solutions due to the following reasons: (i) the model performed domain knowledge transfer from a labelled dataset to an unlabelled dataset. Given that most of the reviews in the dataset belong to the 'Home Improvement' and 'Power Tools' categories, transfer of knowledge from labelled datasets restricted to a few select domains like 'Laptop' and 'Restaurant' to an unrelated domain like 'Home Improvement' would result in poor extraction of aspects; (ii) the proposed code for the model was strongly coupled to the experiments that the authors conducted and applying the model to our dataset became a challenge.

After deciding to tackle ABSA by splitting it into its constituent aspect extraction and aspect sentiment detection tasks, we faced a challenge again in analyzing sentiment once aspects were extracted. The challenge faced was in trying to employ a sentiment analysis model that determines sentiments based around an aspect. In order to overcome this issue, we decided to employ a simple assumption that the sentiment of a sentence is reflective of the sentiment of the aspect that it contains. By making this assumption, we were able to employ a standard sentiment analysis model in the pipeline.

1.3 Contributions of your project

1. An ML pipeline that performs the end to end task of ABSA by extracting aspects followed by the application of sentence level sentiment analysis to determine sentiment of the extracted aspects. The pipeline is completely unsupervised and end-to-end.

2. Novel assumption to associate sentence level sentiment to aspects based on qualitative review of the dataset.

3. Visualization tool to present findings via radar charts and qualitative textual views.

2 Related Work

In their work, (Li et al., 2019) tackle the problem of unsupervised aspect extraction and sentiment classification by capturing latent relations between aspect and opinion terms by leveraging their co-location and syntactical usage. A global memory neatly captures highly correlated aspect or opinion words from the local memories and SAL enables dynamic alignment of the captured relations from the source to the target domain. This model in conjunction with the model described in (Yang et al., 2019), to the best of our knowledge, are the only two models that perform the task of ABSA in a single model thereby being end to end.

(Yang et al., 2019) propose a local context focus based model that performs aspect term extraction and polarity classification simultaneously. The model works in a supervised setting, however, and requires annotated training data.

BERT post-training approach has been proposed to improve ABSA task (Xu et al., 2019). The authors pointed out that BERT doesn't have domain knowledge about the reviews, because it was trained from Wikipedia corpora (no subjective content), which would induce model bias and lower the ABSA performance. The authors addressed this issue by introducing a post-training stage before fine-tuning on the end task. The post-training is conducted on unsupervised data (review data in our case), and composed of two tasks which are Masked Language Model (MLM) and Next Sectence prediction (NSP). The authors argue that MLM is crucial for injecting review domain knowledge and for alleviating the bias of the knowledge from Wikipedia NSP encourages BERT to learn contextual representation beyond word-level.

The paper (Song et al., 2019) Attentional Encoder Network for Targeted Sentiment Classification replaces RNNs with AENs and introduces a label smoothening process. In the paper is also described BERT-SPC and Bert-ADA. BERT-Sentence Pair Classification introduces a new type of input token format for words as “CLS - tokens - SEP - aspects - SEP” instead of “CLS - tokens - SEP” that is usually used. BERT-SPC was specially designed for the task of ABSA. BERT-ADA reinforces the need for fine-tuned domain specific input for better performances.

To overcome the issue of a lack of annotated datasets for the Apsect Term Extraction (ATE) sub-task, (Giannakopoulos et al., 2017) present a B-LSTM and CRF classifier which they use for feature extraction and aspect term detection for both supervised and unsupervised ATE. Moreover, they introduce a new, automated, unsupervised and domain independent method to label tokens of raw opinion texts as aspect terms with high precision. We could use this method for our ATE subtask as our dataset is not annotated.

In recent years, Latent Dirichlet Allocation and its variants became the dominant unsupervised models for aspect extraction. LDA methods model the corpus as a mixture of aspects, and aspects as distributions over word types. Aspects extracted by LDA-based methods tend to describe the corpus well but are prone to being unrelated and loosely-coupled because LDA models do not leverage word co-occurrence to preserve aspect coherence as backed by research conducted in (Mimnoet al., 2011). LDA models assume that each word is generated independently. Furthermore, LDA-based models need to estimate a distribution of aspects for each document. Review documents tend to be short, thus making the estimation of topic distributions more difficult.

Unsupervised aspect extraction by exploiting linguistic phenomenon has been proposed in (Liao et al., 2019). They noted that aspect words generally distinguish themselves from other words in their occurrence patterns within global and local context, and they proposed a Neural model that is capable of coupling globel and local representation to discover aspect words. This model is different from the one we employ based on (He et al., 2017) in that the latter model leverage the word co-occurrences through the use of neural word embeddings word2vec. Word embedding models encourage words that appear in similar contexts to be located close to each other in the embedding space. In addition, attention mechanism is used to de-emphasize irrelevant words during training.

3 Proposed Model

In this section, we explain the ML pipeline that we have employed to perform the task of ABSA. The code can be accessed on `https://www.github.com/saurabhshirodkar/abae-pytorch`.

3.1 Components of Model

The model is composed of 2 models:

1. An unsupervised aspect extraction model based on (He et al., 2017) that uses an attentional neural mechanism to extract aspects.
2. Valence Aware Dictionary and sEntiment Reasoner (VADER) model to analyze the sentence-level sentiment. VADER is a lexicon and rule-based sentiment analysis model that is finetuned to social media sentiment analysis. The model nicely generalizes to product reviews.

3.1.1 Unsupervised Neural Attention Model for Aspect Extraction

The model proposed in (He et al., 2017) disovers coherent aspects by exploiting the distribution of word co-occurrences through the use of neural word embeddings. Instead of assuming independently generated words, word embeddings encourage words that appear in similar contexts to be located close to each other in the embedding space. Further, attention mechanism is applied to remove irrelevant words.

To begin with each word is represented by a word embedding feature vector $_w \in \mathbb{R}^d$. The feature vectors associated with the words correspond to the rows of a word embedding matrix $\mathbb{E} \in \mathbb{R}^{V \times d}$, where V is the vocabulary size. We intend to learn the aspect embeddings $\mathbb{T} \in \mathbb{R}^{k \times d}$, where k is the number of aspect categories to be learned. The aspect embeddings are used to approximate aspect terms in the vocabulary and the attention mechanism is used to filter down the aspect terms.

The input to the model is a list of indexes of words. Two steps are performed in the model: (i) filter out non-aspect words by down-weighting using attention mechanism followed by the construction of sentence embedding (z_s) from

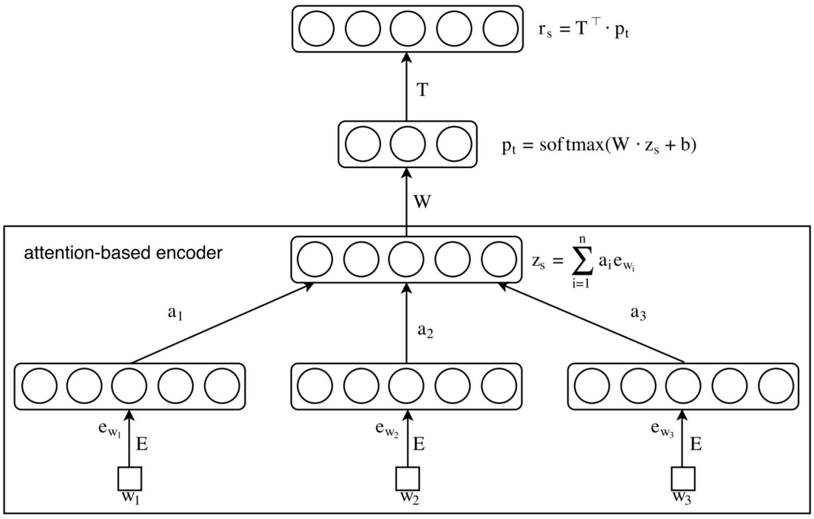

Figure 1: Aspect Extraction Model Architecture

weighted word embeddings, and (ii) reconstruction of the sentence embedding (r_s) from the aspect embeddings $\mathbb{T}$. This process preserves the crucial aspect information within the embeddings.

The process of sentenece embedding construction from weighted word embeddings is motivated by the need to capture the most relevant information w.r.t the aspect of a sentence. The sentence embedding is defined as follows:

$$z_s = \sum_{i=1}^{n} a_{iw_i} \tag{1}$$

For each word we compute a positive attentional weight to check if the word is to be focused on as an aspect term. The attentional weight is conditioned on the word embedding and is defined as follows:

$$a_i = \frac{\exp(d_i)}{\sum_{j=1}^{n} \exp(d_j)} \tag{2}$$

$$d_i = {}_{w_i}^{T} \cdot M \cdot y_s \tag{3}$$

$$y_s = \frac{1}{n} \sum_{i=1}^{n} w_i \tag{4}$$

As a reminder, word embeddings contain within then word co-occurence statistics. This allows us to capture the global context. $M \in \mathbb{R}^{d \times d}$ is a matrix mapping between the global context embedding y_s and word embeddings e_w. y_s is the average of the word embeddings. The word embeddings are learned during training.

The attention mechanism can be visualized as follows:

1. Construction of sentence embedding by averaging the word embeddings y_s
2. Weight of each word is assigned by (i) filtering word through M to capture relevance of word to the k aspect embeddings and (ii) capture relevance by taking inner product with global context y_s.

Once sentence embedding is obtained, the sentene embedding is recomputed using the aspect embeddings $\mathbb{T}$. The reconstruction process mimics that of an autoencoder and is constituted of two transformations:

$$r_s = \mathbb{T}^T \cdot p_t. \tag{5}$$

$$p_t = \textbf{softmax}_(W \cdot z_s + b) \tag{6}$$

where r_s is the reconstructed sentence embedding and p_t is the weight vector over k aspect embeddings. Each weight in the p_t vector represents the probability that the sentence belongs to the related aspect. It is obtained by reducing z_s from d to k dimensions followed by the application of a softmax layer to yield normalized non-negative weights.

3.1.2 VADER

VADER (Hutto and Gilbert, 2014) is a parsimonious rule based model for general sentiment

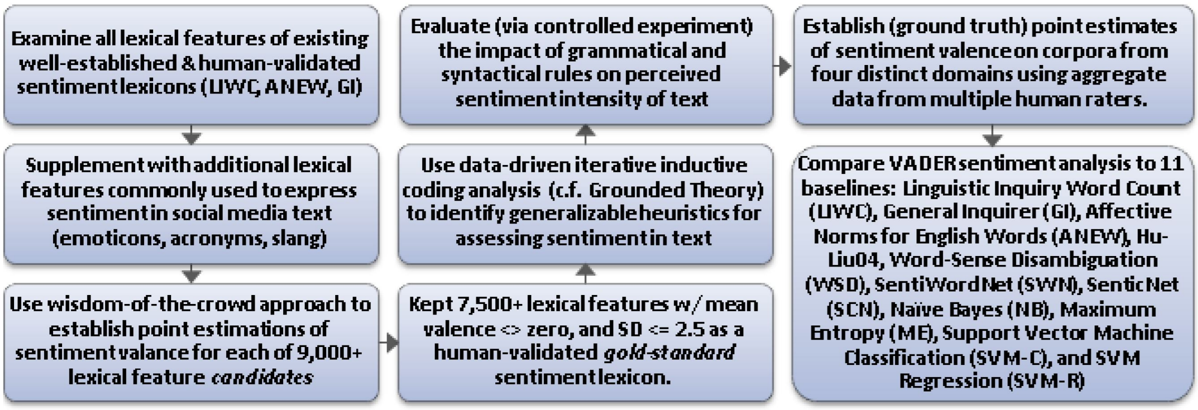

Figure 1: Methods and process approach overview.

Figure 2: Methodology of producing lexicon framework

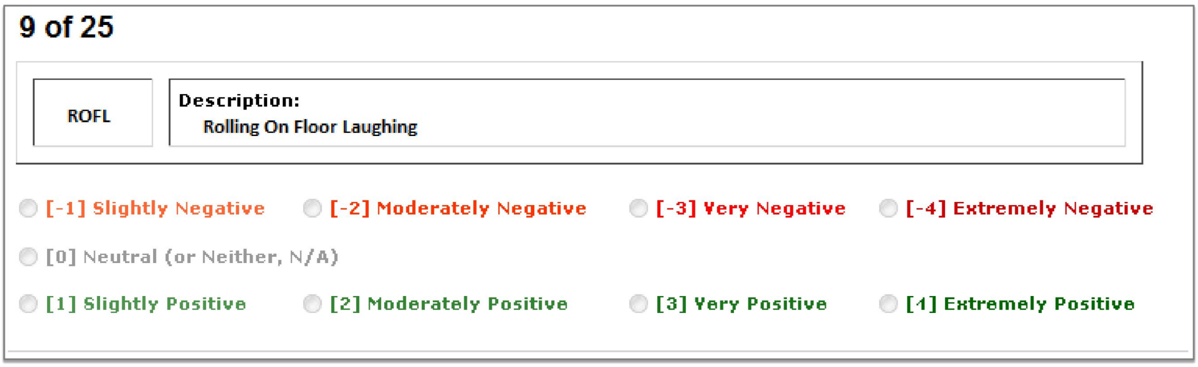

Figure 3: Interface to obtain gold-standard

analysis. It was developed by producing and validating a human-annotated gold-standard sentiment lexicon that is especially attuned to microblog-like contexts. The methodology of generating the lexicon is described in figures 2 and 3. This includes review data. The lexical features were combined into five generalizable rules that embody grammatical and syntactical conventions that humans use when expressing or emphasizing sentiment intensity.

After validation, the authors constructed a set of 5 generalizable rules:

1. Punctuation, namely the exclamation point (!), increases the magnitude of the intensity without modifying the semantic orientation. For example, "*The food here is good!!!*" is more intense than "*The food here is good.*"

2. Capitalization, specifically using ALL-CAPS to emphasize a sentiment-relevant word in the presence of other non-capitalized words, increases the magnitude of the sentiment intensity without affecting the semantic orientation. For example, "*The food here is GREAT!*" conveys more intensity than "*The food here is great!*"

3. Degree modifiers (also called intensifiers, booster words, or degree adverbs) impact sentiment intensity by either increasing or decreasing the intensity. For example, "*The service here is extremely good*" is more intense than "*The service here is good*", whereas "*The service here is marginally good*" reduces the intensity.

4. The contrastive conjunction "*but*" signals a shift in sentiment polarity, with the sentiment of the text following the conjunction being dominant. "*The food here is great, but the service is horrible*" has mixed sentiment, with the latter half dictating the overall rating.

5. By examining the tri-gram preceding a sentiment-laden lexical feature, we catch nearly 90% of cases where negation flips the

polarity of the text. A negated sentence would be "*The food here isn't really all that great*".

3.2 Pipeline Description

1. Aspect Extraction (Unsupervised) – Given a set of reviews, identify the essential aspects or features that are mentioned in the reviews.

   Example: The product was durable. The instructions were unclear. The price was too high.
   Essential aspects: durable, instructions, price

   Similar aspect terms get grouped to form one aspect: price, cost, expensive, pricey, cheap

2. Aspect Refinement – This optional step is used to get rid of irrelevant aspect terms

   (a) Split coarse aspects into multiple fine-grained aspects
   Example: [tightened, tightening, tighten], [loosened, loosening, loosen], [torqueing, torqued, torx]

   (b) Rename the aspects into meaningful names
   Overall Quality (Aspect 2): {reliability, usability, durability, longevity, performance, term, usefulness, efficiency, versatility, exceeds, ratio, sacrifice, workmanship, consumption, exceptional}

   (c) Remove aspect terms that don't make sense
   Overall Quality: {reliability, usability, durability, longevity, performance, usefulness, efficiency, versatility, workmanship, exceptional}

3. Aspect Polarity – Identify a sentiment associated with each of the aspects for some product. This can be done by following an algorithm:

   (a) Split reviews by sentences
   (b) Extract and store sentiment of each sentence
   (c) For each aspect, find sentences in which it occurs
   (d) Assign the average sentiment of those sentences to the aspect

   Example: Ergonomics = {heft, heavier, grippy, ergonomic, hefty, weights, stable, weigh, weight, substantial, slippery, stability, lightweight, beefy, heavy}

   (a) This saw is very easy to set up and very stable once in place. $\rightarrow$ 0.8
   (b) It is a lightweight that packs the punch of a heavyweight. $\rightarrow$ 0.6
   (c) Very light weight and easy to use. $\rightarrow$ 1.0
   (d) a little heavier however than brand X however still light enough to be mobile. $\rightarrow$ 0.4

   Ergonomics = 0.7 (Positive)

3.3 Training

The model is trained on the reconstruction error based on the contrastive max-margin objective function as defined in (Weston et al., 2011). For each input sentence, m sentences are randomly sampled as negative samples n_i. The objective is to ensure similarity between r_s and z_s while keeping r_s different from the negative samples.

To ensure that the aspect embeddings $\mathbb{T}$ that is learned doesn't suffer from redundancy and are unique, the loss function is regularized as follows:

$$U(\theta) = \|\mathbb{T}_n \cdot \mathbb{T}_n^T - \mathbf{I}\| \tag{7}$$

where I is the identity matrix and T_n is the normalized aspect embedding vector. The regularization term promotes orthogonality between the aspect embeddings.

The loss function can thus be defined as the hinge loss that maximizes similarity between r_s and z_s and dissimilarity between r_s and n_i added to the regularization term

$$L(\theta) = \sum_{s \in D} \sum_{i=1}^{m} \max(0, 1 - r_s z_s + r_s n_i) + \|\mathbb{T}_n \cdot \mathbb{T}_n^T - \mathbf{I}\| \tag{8}$$

where θ is the representation of model parameters $\{\mathbb{E}, \mathbb{T}, \mathbb{M}, W, b\}$.

In order to determine the sentiment and associate it to the extracted aspects, we make the simple assumption that the sentiment of a sentence is reflective of the aspects it contains. We arrived at this assumption by performing a qualitative study of the reviews in the dataset. Reviews that contain two aspects within a single sentence are very few in comparison to ones that contain only a single aspect. The qualitative study can be verified via the visualizations provided.

4 Experiments

4.1 Datasets

The review dataset we use is provided by Stanley Black & Decker, which contains 607755 product reviews between January 2019 to August 2019. There are 15 product categories (see Figure 4) from which reviews come from, and besides the review text, in each record, we also have information about product name, product description, brand name, URL to the product page and the rating.

A reviews contains 40 words on average with standard deviation 55. The maximum number of words is slightly above 2400, and the 25 percentile and 75 percentile word count are 11 and 50 respectively. Moreover, Table 1 shows that over 400000 reviews have 5 star ratings.

Ratings	Number of reviews
5	417142
4	82957
3	31482
2	22905
1	53269

Table 1: The Distribution of Rating

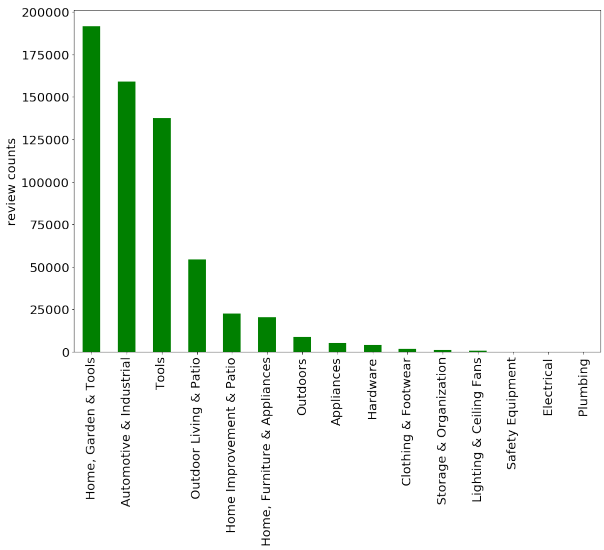

Figure 4: Product Categories

After filtering the dataset to remove redudnacy, we had a little over 632,000 reviews in our dataset. Given the product categories within the dataset, the number of reviews isn't sufficient to conduct effective training. Thus, we augment the dataset with the dataset with a dataset created and maintained by Stanford of 134,476 Amazon reviews of products in the 'Tools and Home Improvement' category.

4.2 Baselines

Due to the sporadic improvements in unsupervised ABSA, it is clear from the study conducted in (He et al., 2017) that the chosen model outperforms LDA based models, BTM and many other rule based methods.

The neural attention model chosen is competitive despite there being newer developments in the domain. We tried implementing (Li et al., 2019) for our setting but due to reasons mentioned previously we had to work with the attention model.

The results obtained as described later in the results section show that the aspects extracted are coherent and sensible. In addition, there is the advantage of categorization of aspect terms. Each aspect term in the aspect embedding vector consists of d dimensions. The d terms are aspect words within the vocabulary. Thus, if we choose to have k=30 aspect terms with d=15, we get a total of 450 aspect terms while keeping the number of parameters relatively low. This enables us to cover a wide array of aspect words in the dataset.

4.3 Tasks

1. Experimentation of (Li et al., 2019) in pursuit of delivering an end-to-end solution which we replicated to obtain results on the Laptop, Restaurants and Service datasets which were presented in the midterm presentation.

2. Implementation of (He et al., 2017) to the augemented dataset described in section 4.1. The aspects extracted are presented in the following section.

3. Application of VADER to the original dataset to obtain sentence-level sentiment applied to the aspects.

4. Building of a radar chart and a qualitative textual view keeping in line with the current trends in visualization as cited in (Kucher et al., 2018).

4.4 Main Results

For our implementation, we set k=30 to extract 30 different aspect categories with d=15. Some of the aspect categories that we obtained include ergonomics, instructions, quality, price, and durability. Some of the aspect words within the aspect categories have been described in figures 7 and **??**

Aspect 1: ["loud", "noise", "hum", "audible", "buzzing", "hear", "sound", "humming", "buzz", "clicking", "beep", "louder", "pitched", "loudly", "vibration"]

Aspect 2: ["heft", "heavier", "grippy", "ergonomic", "hefty", "weighs", "stable", "weigh", "weight", "substantial", "slippery", "stability", "lightweight", "beefy", "werner"]

Aspect 3: ["instruction", "diagram", "english", "manual", "direction", "documentation", "written", "booklet", "youtube", "pdf", "web", "illustration", "website", "printed", "instructional"]

Aspect 4: ["flow", "intake", "cfm", "airflow", "psi", "filtration", "circulate", "gpm", "allergen", "air", "circulating", "stream", "micron", "vent", "clogged"]

Figure 6: Aspects extracted by the unsupervised model

ERGONOMICS => ["heft", "heavier", "grippy", "ergonomic", "hefty", "weighs", "stable", "weigh", "weight", "substantial", "slippery", "stability", "lightweight", "beefy", "heavy"]

Figure 7: Aspect category: Ergonomics

Once the aspects were extracted, we computed the sentiment of individual sentences, matched the aspects to that sentiment followed by the computation of the average of all the sentiments of all the aspect words within an aspect category to calculate the overall sentiment of the aspect category. The method has been neatly visualized in figure 8.

1. This saw is very easy to set up and very stable once in place. => 0.8
2. Its a lightweight that packs the punch of a heavyweight. => 0.6
3. Very light weight and easy to use. => 1.0
4. a little heavier however than brand X however still light enough to be mobile. => 0.4

ERGONOMICS => 0.7
(POS)

Figure 8: Sentiment calculation methodology

In order to enable taking decisions based on the results obtained, we built visualizations coded in D3.js. A radar chart represents the average sentiment across the aspect categories. An interactive visualization is also present which contains boxes of all the aspect categories which when clicked upon provides the user with a descriptive textual view of all the reviews contained within an aspect category filtered by aspect words with each review highlighted in varying saturation levels of green and red to denote intensity of positive and negative sentiment respectively. A snapshot of the visualization is presented in figure 5.

4.5 Analysis

The model that we proposed to extract aspects outperforms several other baselines as backed strongly by research as found in (He et al., 2017). While newer methods exist, our neural model pipeline possesses the advantage of being simple, end to end, and most importantly completely

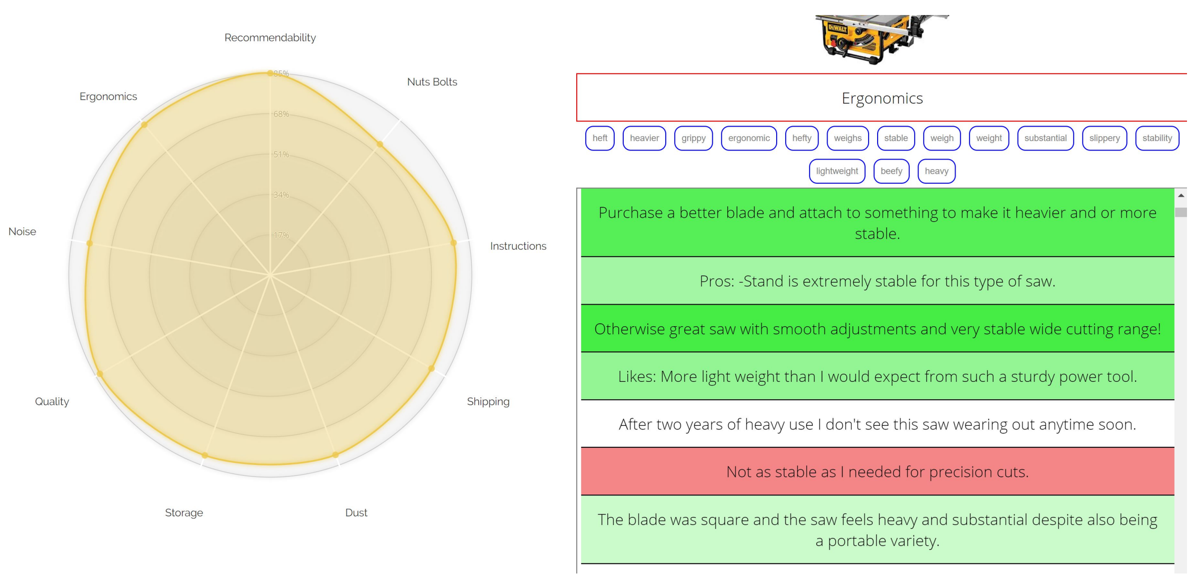

Figure 5: Snapshot of visualization tool

unsupervised consequently being adaptive to any domain.

While we could have validated the performance of our model by comparing results with an annotated dataset, we didn't do so due to a lack of time. Once the task of ABSA was complete, our focus was on building a visualization tool which we justify by relying on the results provided in the papers that proposed the models (He et al., 2017) and (Hutto and Gilbert, 2014) used in the pipeline. In addition, the results obtained were verified qualitatively which further backs the visualization tool.

From the results provide in figures 8, it can be seen that the aspects extracted are of high quality, inter-related and highly coherent. We admit that there can be improvements made in the sentiment analysis component of our pipeline and mention it as a scope for future work.

5 Conclusion and Future Work

To conclude, in this project, we build an end-to-end, fully unsupervised aspect based sentiment analysis pipeline to extract and determine sentiment of product features to drive informed decisions on product modifications/introductions. We use a simple but effective attention based mechanism to leverage word co-occurrence to extract highly coherent aspects. We apply the idea of associating sentiment of a sentence to the aspects it contains opening up the sentiment analysis component of the pipeline to several generic sentiment analysis models. This way SOTA models can be applied to detect sentiment of sentences and thereby aspects. The assumption made is backed by a qualitative review of the dataset.

In order to aid the decision making process, we provide a visualization tool to easily digest the results. The visualization elements are in line with the current trend of visualizing sentiment.

Since the pipeline is unsupervised, there is a lot of scope of improvement. The sentiment analysis model, to begin with, can be swapped for a better performing model. While improvements would be relatively marginal, it would still provide better depiction of sentiment contained within the reviews.

Acknowledgements

We would like to thank SBD Data Scientists Dr. Sunil Nakrani and Dr. Mark Kanner for their invaluable advice and guidance. This project wouldn't have taken its current shape without their input. We would also like to profess our gratitude to our PhD mentor Tu Vu and our TA Rajarshi Das for their constant support. We wouldn't have been able to overcome certain roadblocks if it weren't for their input. Lastly, we would like to thank Dr. Andrew McCallum for playing his part in making an oppurtunity like this viable.

References

Athanasios Giannakopoulos, Claudiu Musat, Andreea Hossmann, and Michael Baeriswyl. 2017. Unsupervised aspect term extraction with b-lstm crf using automatically labelled datasets.

Ruidan He, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2017. An unsupervised neural attention model for aspect extraction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 388–397, Vancouver, Canada. Association for Computational Linguistics.

Clayton J. Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*. The AAAI Press.

Kostiantyn Kucher, Carita Paradis, and Andreas Kerren. 2018. The state of the art in sentiment visualization. *Computer Graphics Forum*, 37(1):71–96.

Zheng Li, Xin Li, Ying Wei, Lidong Bing, Yu Zhang, and Qiang Yang. 2019. Transferable end-to-end aspect-based sentiment analysis with selective adversarial learning.

Ming Liao, Jing Li, Haisong Zhang, Lingzhi Wang, Xixin Wu, and Kam-Fai Wong. 2019. Coupling global and local context for unsupervised aspect extraction. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4579–4589, Hong Kong, China. Association for Computational Linguistics.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeniy Kotelnikov, Nuria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit.

2016. SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 19–30, San Diego, California. Association for Computational Linguistics.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. SemEval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 486–495, Denver, Colorado. Association for Computational Linguistics.

Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.

Youwei Song, Jiahai Wang, Tao Jiang, Zhiyue Liu, and Yanghui Rao. 2019. Attentional encoder network for targeted sentiment classification.

Neil Weston, Iain Greenlees, and Richard Thelwell. 2011. The impact of a performance profiling intervention on athletes' intrinsic motivation. *Research quarterly for exercise and sport*, 82:151–5.

Hu Xu, Bing Liu, Lei Shu, and Philip S. Yu. 2019. Bert post-training for review reading comprehension and aspect-based sentiment analysis.

Heng Yang, Biqing Zeng, JianHao Yang, Youwei Song, and Ruyang Xu. 2019. A multi-task learning model for chinese-oriented aspect polarity classification and aspect term extraction.