A Pipeline of Aspect Detection and Sentiment Analysis for E-Commerce Customer Reviews

Haozheng Tian
The Home Depot, Inc.
Atlanta, Georgia
haozheng_tian@homedepot.com

Morgan White
The Home Depot, Inc.
Atlanta, Georgia
james_m_white@homedepot.com

ABSTRACT

Customers reviews are becoming increasingly important to assist the purchase decision in e-commerce platforms. Reviews from customers usually reflect the aspects of a product or service that are deemed valuable by other customers, which may not be mentioned or emphasized in product descriptions. Accumulating an abundant amount of reviews for products is an efficient approach to build customer trust and often observed with positive correlation of conversion increase. However, at the same time, the enormous amount of reviews becomes an obstacle for a customer to fully grasp the consensus opinion on aspects that he/she truly cares about. Additionally, the vast diversity in vocabulary used in reviews introduces difficulty for quick and accurate comprehension. To address the above-mentioned issues, we present an end-to-end pipeline for product aspects detection from customer reviews and sentiment analysis. In the paper, we describe the aspects detection approach for both explicit and implicit aspects. We further develop a gated-RoBERTa-based sentiment classifier for sentiment analysis, which exhibits outstanding performances in multi-domain corpus.

CCS CONCEPTS

• **Computing methodologies → Supervised learning by classification**; • **Information systems → Language models**.

KEYWORDS

aspect extraction, explicit aspect, implicit aspect, sentiment analysis, RoBERTa

ACM Reference Format:
Haozheng Tian and Morgan White. 2020. A Pipeline of Aspect Detection and Sentiment Analysis for E-Commerce Customer Reviews. In *Proceedings of ACM SIGIR Workshop on eCommerce (SIGIR eCom'20).* ACM, New York, NY, USA, 9 pages.

1 INTRODUCTION

Intuitive and smooth browsing experience of e-commerce websites has been the key to encourage better engagement and drive revenue. While retailers are providing abundant information regarding products specification and features from manufacturer, consensus is reached on the fact that customer reviews play an essential role in purchase decisions. There are several reasons. First, the knowledge gap between manufacturer and customer obstructs a pleasant shopping experience. For example, while the description "This dishwasher is made to perform at only 50 dBA" is straightforward for a professional, a customer may have difficulty in distinguishing whether it is noisy or not. Second, manufacturer may fail to deliver or highlight the specification that a customer deems important. For example, the product specification of a refrigerator may contain no information regarding noise, whereas it is one of the most common topics in reviews. Third, customers frequently refer to reviews before purchase, judging the popularity and quality of a product.

However, reviews of a popular product, which accumulate to the number of several thousand or more, become infeasible to comprehend. Meanwhile, merely reading the top reviews in a webpage is likely to cause bias in the evaluation of a product. As a solution, aspect extraction is the process of summarizing product-relevant information and determining the opinion expressed on it using natural language processing (NLP) techniques. Figure 1 is an example of organizing customer reviews in pros and cons, improving readability and communication.

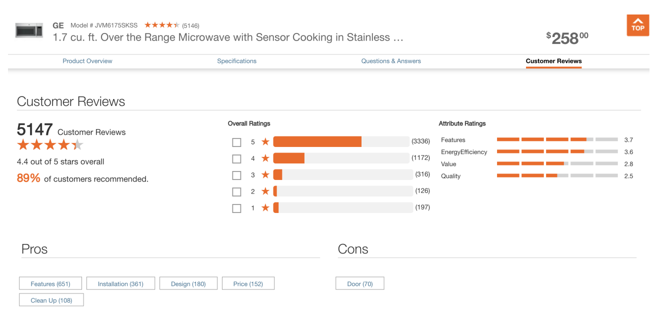

Figure 1: Example of summarized pros and cons based on customer reviews at a product page of e-commerce.

Aspect extraction faces challenges in that mapping colloquial language to technical terms and scaling to dozens of domains without loss of accuracy. WordNet has been leveraged to infer semantic similarity [22], but effectiveness is impaired due to occasional lack of synonymous relationship between words that are actually semantically close. Synonyms based semantic clustering using GloVe [14] and word2vec [26] have also been proposed. However, we found that Google Universal Sentence Encoder (USE) [4] based semantic inference proves to be more accurate. Inferring similarity of phrases using USE outperforms ELMo [30], BERT [9] and XLNET [45] as an encoder, in that USE applies a transformer architecture and is trained to identify the semantic textual similarity (STS) between sentences. Though BERT-based models have been proposed to tackle aspect detection and sentiment analysis as a one stop solution [29], the expense of large scale labeling and difficulty of domain adaption prevent its commercial application.

Sentiment analysis models should be adaptive to the extremely vast difference in vocabulary and language of reviews introduced by millions of products. However, applying one single model for all products reviews could damage performances significantly. It is possible to address the issue by training individual models for subsets of products, but it lacks scalability given dozens or hundreds of categories in a retailer's catalog.

In this work, we present the pipeline of aspect extraction and aspect-based sentiment analysis, deployable at e-commerce platform with multi-domain adaptation. We leverage the capacity of pre-trained transformer architecture, RoBERTa [25], to significantly improve the accuracy in detecting pros and cons of products across domains. The contributions of the study can be summarized as follows:

- Propose the explicit and implicit aspect extraction framework scalable at e-commerce platform.
- Propose semantic-based aspect emerging using USE.
- Propose and benchmark the gated RoBERTa sentiment classifier.

2 RELATED WORK

Aspect extraction is the foundation of aspect-based sentiment analysis. Relevant methods have been proposed by researchers over the past decades. Supervised models generally performs better than unsupervised models, but lack adaption in domain migration [36]. Supervised models include conditional random field (CRF) [6, 13, 20], integration of neural networks and CRF [41, 43], semantic parsing using dependency relations [21, 33], tree-based model [15] and Lexicalized HMM-based model [17, 18]. Recognizing the disadvantage of supervised approach, researcher also proposed novel CRF-based method to adapt the model to new domains [38]. Hybrid approach of rule-based model and neural network have also been developed for domain adaptation [10]. While deep neural network models [2, 31] are gaining popularity with improved performances, recent advance in transformer based language model inspired its application in aspect extraction [46].

Unsupervised models in aspect extraction are widely used and are more robust in diverse kinds of domains. Intuitively, models based on statistical characteristics are proposed, utilizing frequency, association and linguistic features [35, 37]. A rule-based method [14] showed that even though there exists a large disparity in vocabulary between manufacturer and customer, domain-specific information significantly improves aspect discovery. There are also bootstrapping [1, 49], pointwise mutual information (PMI) [34] and word alignment approaches [23, 24].

Semi-supervised model have been proposed to guide clustering of similar aspects by using a few seed words [27, 47]. Seeding aspects extracted from product information are used to guide discovery of related words from reviews by a labeled LDA [42].

While the majority of researches focus on explicit aspect extraction, implicit aspects provide abundant information regarding customers' opinion, which should be fully utilized. Rule-based approach relies on the dependency parsing to mark words that fit in a specific pattern as implicit aspects [32]. CRF is used to recognize implicit aspect indicator which is logically associated with an explicit aspect [3, 8]. Additionally, the idea of association mining is proposed and its applications include utilizing co-occurrence rules [19], calculating PMI to infer semantic association [39] and applying hybrid model with collocation extraction and semi-supervised LDA topic model [16].

Multiple methods have been proposed on sentence level sentiment analysis. Deep learning models have been gaining momentum with extraordinary performances [28, 48]. To further improve deep learning model, a divide-and-conquer approach is developed to group sentences into several types before applying convolutional neural network classifier [5]. The introduction of transformer-base language model [9] has revolutionized natural language processing and significantly improved sentiment analysis compared to recurrent neural network [40]. Transformer models have also been applied in both aspect detection and sentiment analysis [29].

3 EXPLICIT AND IMPLICIT ASPECT EXTRACTION

The length of a product review varies significantly, from one phrase or sentence to a lengthy paragraph with several topics. A long review could be formidable to read, but it is likely to contain some aspects of a product that a customer cares about. Aspect extraction is the process of recognizing and distilling relevant information and assist in better communication. Aspect extraction of review text is performed on sentence level, while one sentence can contain one to two aspects of a product. The methodology is consisted of four parts: 1. pre-processing of raw review text, 2. explicit and implicit aspect extraction, 3. semantic merging of similar aspects, 4. mapping implicit aspects to corresponding explicit aspects. Figure 2 shows the flow of the above mentioned steps.

3.1 Pre-Processing

Pre-processing of raw texts standardizes the language and removes unnecessary information in raw texts. It improves the quality of part-of-speech (POS) parsing and is applied to raw reviews before aspect extraction. The pipeline of pre-processing is consisted of the following steps: 1. a contraction is replaced with its formal form, 2. URLs are removed, 3. information added by retailer/vendors are removed, 4. reviews text are broken down to sentences. The pipeline is implemented using NLP module Spacy v2.2.4 [12].

3.2 Explicit Aspect Extraction

The aspect of a product describe an attribute or feature of a product and is often discussed frequently in reviews. The frequent occurrence of a term is a necessary criteria of it being an aspect. However, customers express opinions on a variety of things, including themselves, family, previous purchase and comparable products. The challenge of identifying a true aspect involves differentiating whether a frequent term is related to the product itself or not. Mostly, the aspect of a product appears in the form of a noun or noun phrase. Meanwhile, certain verb phrases also qualify.

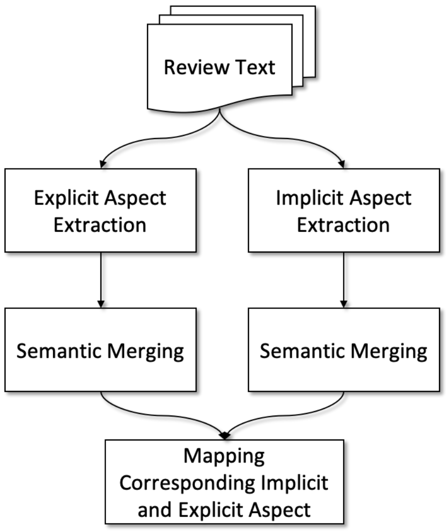

Figure 2: Flowchart of Explicit and Implicit Aspect Extraction.

Explicit aspect are obtained in sentences where the aspect word appears in its lemmatized or a slightly modified form such as in plural. For example, in sentence "Works well, lots of space, loving the side by side design much more than my old french door.", "space" is an explicit aspect of a refrigerator. Extraction of explicit aspect involves three components. First, raw explicit aspects are obtained by recognizing the noun chunks in a sentence. Second, phrases that are consistent with the following POS structures are considered aspect: 1. adjective-to-verb, 2. adjective-preposition-verb, 3. noun-preposition-verb, 4. noun-preposition-noun. Third, raw aspects are filtered with a global and category specific stoplists.

3.3 Implicit Aspect Extraction

Review sentences containing implicit aspect often utilize verb or adjective for conveying information instead of using a noun. Readers could infer the corresponding explicit aspect semantically and logically. For example, in sentence "It is gorgeous and very easy to use!", the adjective "gorgeous" implies the aspect of "design/style". Recognizing implicit aspect is achieved by identifying adjectives that are not in a curated stoplist. The stoplist includes adjectives that are rarely related to objects, or features of products. Additionally, implicit aspects utilizing verbs are conveniently captured in verb phrases described in previous section.

3.4 Semantic Merging of Similar Aspects

An aspect may have several ways of phrasing. For example, aspect "space" of a refrigerator could be expressed as "room", "space" and "space layout". Semantically similar aspects need to be merged before being presented to customers. To perform semantic merging, Universal Sentence Encoder (USE) is utilized to encode aspects to high dimensional vectors. And then pairwise cosine similarity (Equation 1) is calculated for explicit and implicit aspects individually,

$$\cos(\theta) = \frac{\boldsymbol{A} \cdot \boldsymbol{B}}{|\boldsymbol{A}| \cdot |\boldsymbol{B}|} \tag{1}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ represent the embedding vectors of aspects.

Further, merging of raw aspects is guided by hierarchical agglomerative clustering (HAC), where leaf nodes are grouped to represent multiple concepts in aspects. HAC based on the pairwise similarities are created for explicit and implicit aspects respectively. Aspect clusters are obtained by applying cut-off on the hierarchical tree, while the cut-off value is chosen by maximizing the average Silhouette value of all samples. Individual Silhouette value is given by

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{2}$$

where $a(i)$ is the mean distance between i and all other data points in the same cluster and $b(i)$ is the smallest mean distance of i to all points in any other cluster. Customized merging criteria is used as a supplement to the above steps.

For explicit aspects, the most frequent word/phrase in a cluster is chosen to represent the collective meaning of the cluster.

3.5 Merging of Implicit and Explicit Aspects

All the words in a implicit aspect cluster are used to find their corresponding explicit aspect. First, synonyms and antonyms of the adjectives are retrieved from WordNet [11]. Second, the lemmatized noun forms of the synonyms and antonyms are obtained and cross matched with the explicit aspects. Third, the matching explicit aspect with the highest number of occurrence is chosen to be the aspect of the implicit cluster, while clusters with no matches are abandoned.

Figure 3 shows an example of the hierarchical relationship of implicit aspects using USE as the encoder, where implicit clusters are generated for 1) "expensive", "costly", "cheap", "economical", 2) "noisy", "loud", "quiet", 3) "wide", "large", "big", "spacious" given optimized cut-off at 0.59. The three clusters are mapped to explicit aspects 1) price, 2) noise and 3) space, respectively.

3.6 Aspect Ranking

Aspects of the products are grouped into pros and cons based on the common opinion expressed on them. Ranking is assigned to aspects among which the top ones are shown in e-commerce platform. The ranking method should account for both the polarity of the aspect and frequency of mentioning in reviews. For a particular aspect, χ^2 value is calculated using the number of positive, negative and neutral opinions, assuming there is an equal distribution. Aspects with smaller than threshold p-values are ordered and the top ones are selected for display.

4 SENTIMENT ANALYSIS

Sentiment analysis model classifies sentences with aspects in them into three types: positive, neutral and negative. The sentiment labels enable the process of grouping aspects into pros and cons of products, which makes reviews more readable and easier to comprehend. The proposed gated RoBERTa sentiment classifier

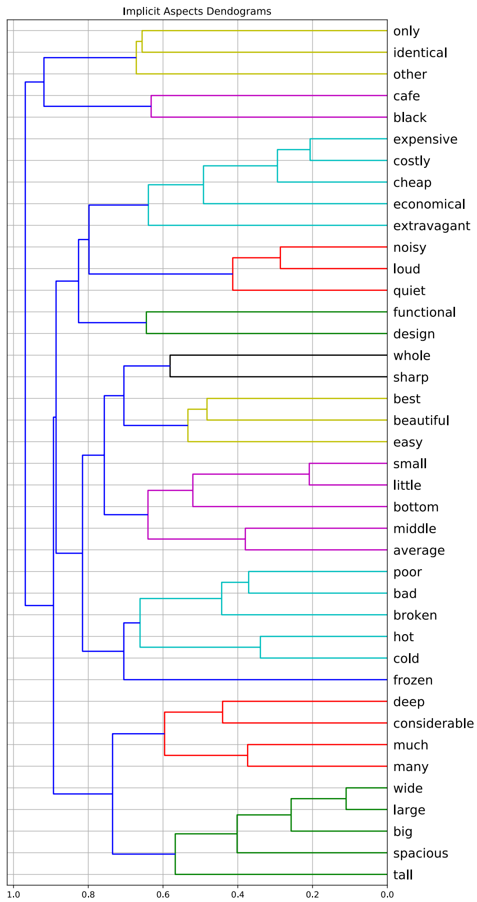

Figure 3: Hierarchical Clustering of Implicit Aspects Based on Cosine Similarity.

utilizes the state-of-the-art transformer architecture and outperforms traditional recurrent neural network models. The model also address the the issue of multi-domain adaptability with the gate mechanism, in which the model decides whether incorporating the categorical information in prediction. The proposed model is trained and benchmarked using labeled customer reviews of Home Depot. Further, performances are compared between four models, 1) model that applies convolutional neural network and LSTM, 2) RoBERTa-based classifier trained separately on two domains, 3) RoBERTa-based classifier trained collectively on two domains, 4) gated RoBERTa classifier trained collectively on two domains.

4.1 Sampling and Labeling

22247 and 15784 review sentences are randomly sampled from two categories, appliances and tools, respectively. Among the two, there are 26 sub-categories and the random sampling is conducted in a way to reflect the distribution of review sentences in the whole dataset (Table 1). Human annotated sentiment labels are assigned to the samples. Positive, negative and neutral sentiments account for 45.01%, 30.65%, 24.34% of all samples, respectively.

Table 1: Labeling of Samples.

Sub-Category	Number
Power Tools	4016
Hand Tools	3012
Power Tool Accessories	2626
Tool Storage	1519
Air Compressors, Tools & Accessories	1450
Ranges	1430
Mini Fridges	1377
Small Kitchen Appliances	1374
Wall Ovens	1373
Microwaves	1371
Dishwashers	1368
Range Hoods	1352
Cooktops	1342
Washers & Dryers	1335
Freezers & Ice Makers	1332
Refrigerators	1326
Appliance Parts	1306
Floor Care	1297
Beverage Coolers	1260
Household Appliances	1252
Garbage Disposals	1236
Wet/Dry Vacuums	969
Trash Compactors	916
Flashlights	875
Safety & Security	730
Welding & Soldering	587

4.2 Gated RoBERTa Classifier

The gated RoBERTa classifier is an integration of pre-trained RoBERTa model and a gate mechanism that regulates the information flow to down stream pipeline. Reviews may contain domain-specific language, such as domain-specific words, or common words with variations of semantic meaning, which can only be accurately interpreted in a context. The gate mechanism consumes both the category information and the text processed with pre-trained RoBERTA model, using a fully connected layer of neural network. It decides whether to pass along the category information to down stream, depending on the criteria that it formulates during training. The gate operates with the logic that it suppresses extra information that may impede the performance when the text is easy to interpret, whereas feeding all information to the model on difficult samples. The objective of the mechanism is to increase domain adaptability of sentiment classifier, given more than dozens of categories in a retailer's catalog.

In Figure 4, category information is appended to the RoBERTa output before going through a fully connected layer, followed by a sigmoid activation. The value from the sigmoid activation decides whether category information should be preserved by multiplying itself with the category information and then feeds the later pipeline together with the output from pre-trained RoBERTa.

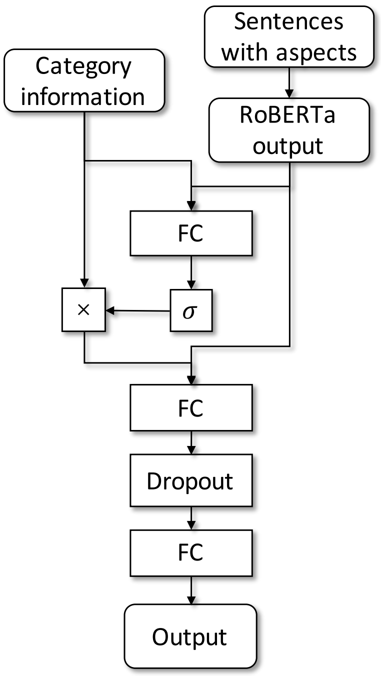

Figure 4: Architecture of Gated RoBERTa Classifier.

5 PERFORMANCE EVALUATION

5.1 Metrics

Metrics for predictive model comparison includes accuracy, Matthews correlation coefficient (MCC) (Equation 3), Cohen's Kappa [7] (Equation 4) and precision.

$$\text{MCC} = \frac{\text{TP} \cdot \text{TN} - \text{FP} \cdot \text{FN}}{\sqrt{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}} \tag{3}$$

$$\text{Kappa} = \frac{p_0 + p_e}{1 - p_e} \tag{4}$$

In the above equations, TP, TN, FP, FN, p_0 and p_e mean true positive, true negative, false positive, false negative, accuracy and hypothetical probability of chance agreement, respectively.

5.2 Models of Sentiment Analysis

The gated RoBERTa model, along with two additional RoBERTa-based schema are fine tuned using the labeled data. First, RoBERTa-based classifier are trained separately on appliances and tools data, which produces two trained models. Second, RoBERTa-based classifier is fine tuned on the dataset without separating the two categories. The RoBERTa-based classifier is consisted of the pretrained RoBERTa model and two layers of fully connected layers, implemented with Huggingface transformer library [44]. To benchmark the performance of transformer based classifier, one recurrent neural network classifier is implemented with one convolutional layer and one LSTM layer. A reserved test set is used to compare the above mentioned models.

Table 2: Performance of Predictive Models

Models	Accuracy	MCC	Kappa
Conv+LSTM	0.666448	0.476716	0.472205
RoBERTa trained seperately	0.8997	0.857224	0.856966
RoBERTa trained together	0.898005	0.841215	0.840983
Gated RoBERTa	**0.916382**	**0.870304**	**0.870078**

Table 2 and Figure 6, 7, 8 show the performance of all four models, where gated RoBERTa exceeds the best of the rest in terms of accuracy, MCC and Kappa. Overall, transformer based classifiers significantly outperform the model utilizing recurrent neural network.

The RoBERTa model collectively trained on the two categories suffers decrease of performance compared to the separately trained one, confirming the significance of the gated mechanism through which contextual information is strategically incorporated. Additionally, precision of negative, neutral and positive cases is compared on 26 sub-categories of appliances and tools. Gated RoBERTa prevails in neutral and positives cases, and achieves a total of 40 winnings, which is more than half of all comparisons (Table 3).

5.3 The Gate Mechanism

To examine the effectiveness of the gate mechanism, its output is grouped into seven bins with equal range. The values of gate output range from 0.58 to 0.99, where small value indicates suppressed category information to down stream processing. The gate output is positively correlated with the average length of sentences, while decreased prediction accuracy is associated with longer sentences (Figure 5). It can be concluded that lengthy sentences tend to trigger the gate mechanism, indicating increased difficulty in interpretation and craving for additional information.

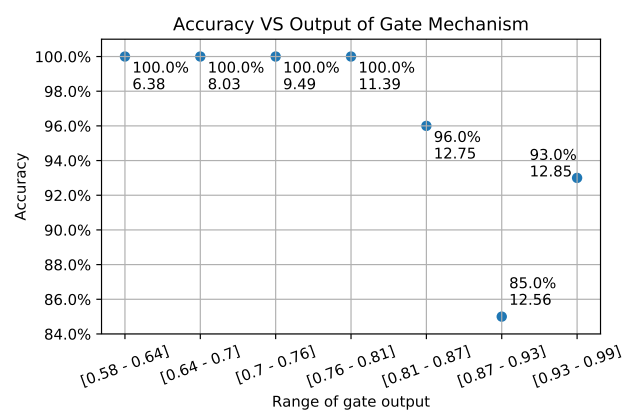

Figure 5: Accuracy of Prediction Given Seven Ranges of Gate Output. The label for each data point means accuracy and average number of words in sentences

Specifically, some example sentences that lightly trigger the gate mechanism include 1) "I adore how quiet this machine is, a must

Table 3: Prevalence of RoBERTa Based Models in Sub-Categories. The values represent the number of times the precision of a particular model outperforms the other two in 26 sub-categories of appliances and tools.

Models	Precision of Negative	Precision of Neutral	Precision of Positive	Total
RoBERTa trained separately	9	8	6	23
RoBERTa trained together	9	5	1	15
Gated RoBERTa	8	13	19	40

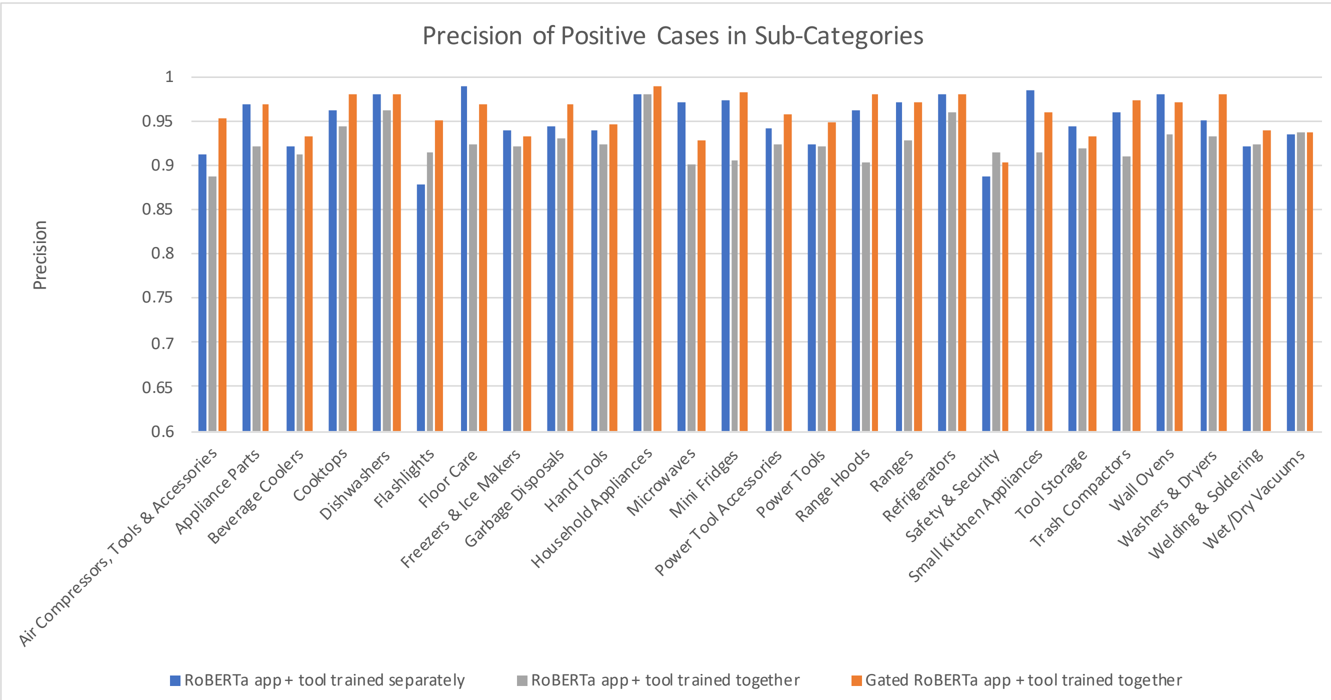

Figure 6: Performance Comparison of Positive Cases of RoBERTa Based Models in Sub-Categories of Appliances and Tools.

for open plan kitchens.", 2) "The dishwasher is easy to use, quiet, and cleans my dishes!", 3) "I am completely satisfied with my new refrigerator." and 4) "This refrigerator has been a wonderful buy for us.". And sentences that excessively rely on category information include 1) "By selecting the Sensor Cycle, the machine automatically selects the right cycle for your load, saving on water and energy.", 2) "And lastly, the chime reminder at the end of the wash cycle is perfect to assist in notifying that things are complete.", 3) "Also the ice dispenser is not capable of doing its job without getting ice on the floor, and the switching between ice and water, is not seamless at all." and 4) "It whines and sounds like whiny dirt bikes reading down my road.".

In general, sentences that are straightforward and contain category-specific keywords usually rely less on the gate output, whereas the contrary for sentences that lack category-specific keywords and express opinions with metaphor or in a intricate manner.

6 CONCLUSION

Customer reviews provide valuable product insights that can be utilized to enhance customer trust and conversion. Novel approaches such as aspect extraction leverage the power of NLP technique to reinforce readability and interpretation of large amount of information in favor of frictionless shopping experience. In this study, we propose the framework to generate product insights based on customer reviews and present them based on the ranking of importance.

To fully exploit the capacity of reviews, we propose the method to extract both explicit and implicit aspects. Additionally, USE embedding is incorporated in hierarchical clustering to infer semantic relationships, guiding the merge of aspects. While the output of HAC is reproducible, it works well regardless of the shape of clusters in hyperspace. Clustering of implicit and explicit aspects is achieved by leveraging synonyms and antonyms of grouped implicit aspects. Our approach requires no prior information of domains, but the performance could be further improved with curated domain knowledge. To address the difficulty of cross domain sentiment analysis, we propose the gated RoBERTa sentiment classifier. We show that the Gated RoBERTa model outperforms not only recurrent neural network models, but also previous transformer-based models as

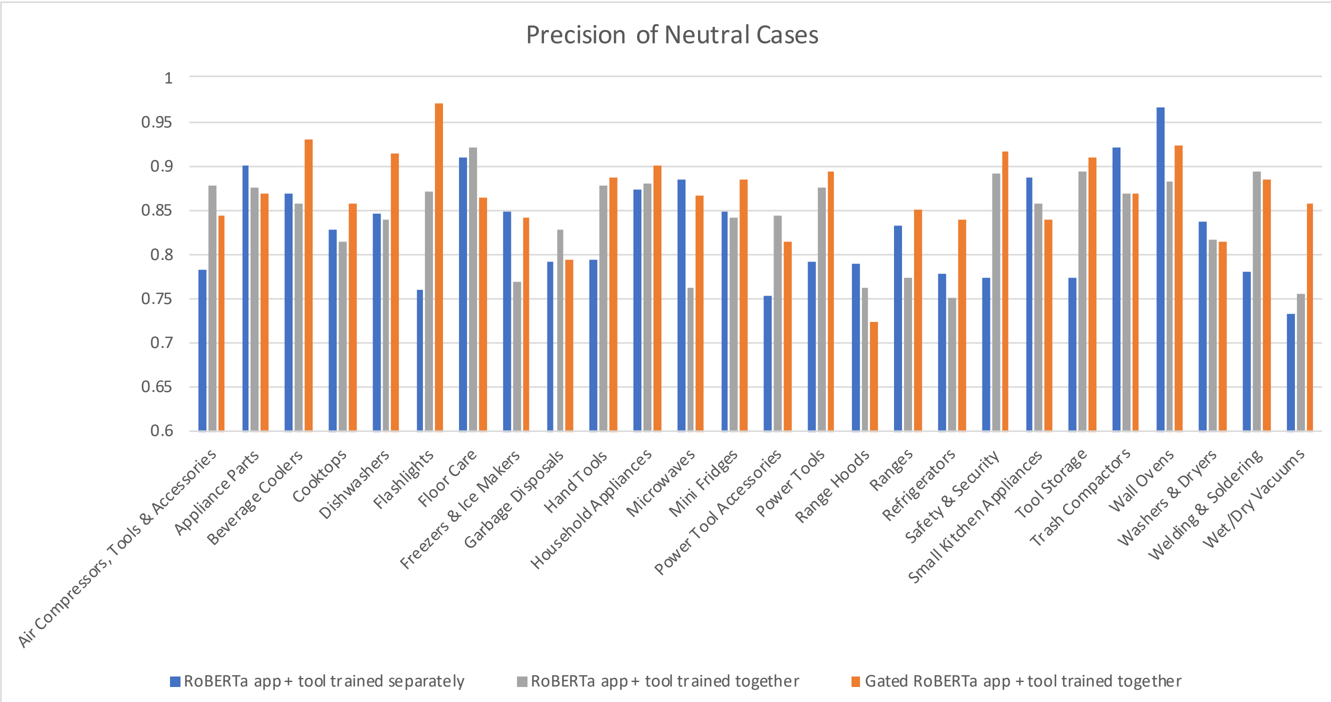

Figure 7: Performance Comparison of Neutral Cases of RoBERTa Based Models in Sub-Categories of Appliances and Tools.

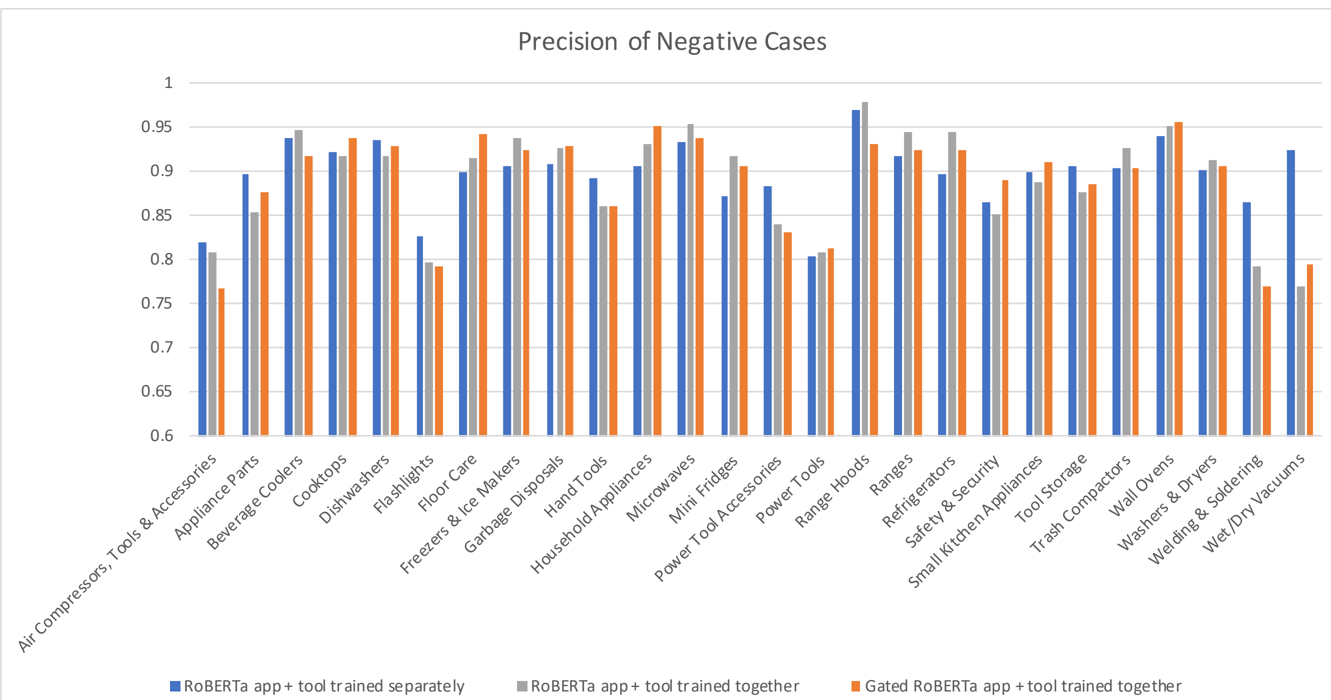

Figure 8: Performance Comparison of Negative Cases of RoBERTa Based Models in Sub-Categories of Appliances and Tools.

well. The deployment of the framework increases customer engagement by 16% from baseline, indicating improvement of browsing experience and potential of revenue increment.

For future steps, aspect extraction could be upgraded to incorporate implicit aspects expressed in nouns. Fine tuning of university sentence encoder using in-house corpus could enhance the inference of semantic similarity in the context of home improvement products. For sentiment analysis, more category-specific information could be explored and leveraged to improve the performance of the gated RoBERTa model.

REFERENCES

[1] Ayoub Bagheri, Mohamad Saraee, and Franciska de Jong. 2013. An unsupervised aspect detection model for sentiment analysis of reviews. In *International conference on application of natural language to information systems*. Springer, 140–151.

[2] Peiman Barnaghi, Georgios Kontonatsios, Nik Bessis, and Yannis Korkontzelos. 2019. Aspect Extraction from Reviews Using Convolutional Neural Networks and Embeddings. In *International Conference on Applications of Natural Language to Information Systems*. Springer, 409–415.

[3] Vishal Bhatnagar, Mahima Goyal, and Md Anayat Hussain. 2016. A Proposed framework for improved identification of implicit aspects in tourism domain using supervised learning technique. In *Proceedings of the International Conference on Advances in Information Communication Technology & Computing*. 1–4.

[4] Daniel Cer, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, et al. 2018. Universal sentence encoder. *arXiv preprint arXiv:1803.11175* (2018).

[5] Tao Chen, Ruifeng Xu, Yulan He, and Xuan Wang. 2017. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications* 72 (2017), 221–230.

[6] Yejin Choi and Claire Cardie. 2010. Hierarchical sequential learning for extracting opinions and their attributes. In *Proceedings of the ACL 2010 conference short papers*. Association for Computational Linguistics, 269–274.

[7] Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20, 1 (1960), 37–46.

[8] Ivan Cruz, Alexander Gelbukh, and Grigori Sidorov. 2014. Implicit Aspect Indicator Extraction for Aspect based Opinion Mining. *Int. J. Comput. Linguistics Appl.* 5 (2014), 135–152.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[10] Ying Ding, Jianfei Yu, and Jing Jiang. 2017. Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction. In *Thirty-First AAAI Conference on Artificial Intelligence*.

[11] Christiane Fellbaum. 1998. A semantic network of English verbs. *WordNet: An electronic lexical database* 3 (1998), 153–178.

[12] Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. (2017). To appear.

[13] Sheng Huang, Xinlan Liu, Xueping Peng, and Zhendong Niu. 2012. Fine-grained product features extraction and categorization in reviews opinion mining. In *2012 IEEE 12th International Conference on Data Mining Workshops*. IEEE, 680–686.

[14] Saratchandra Indrakanti and Gyanit Singh. 2018. A Framework to Discover Significant Product Aspects from e-Commerce Product Reviews.. In *eCOM@SIGIR*.

[15] Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang. 2010. An approach based on tree kernels for opinion mining of online product reviews. In *2010 IEEE International Conference on Data Mining*. IEEE, 256–265.

[16] Wei Jiang, Hao Pan, and Qing Ye. 2014. An improved association rule mining approach to identification of implicit product aspects. *The Open Cybernetics & Systemics Journal* 8, 1 (2014).

[17] Wei Jin, Hung Hay Ho, and Rohini K Srihari. 2009. A novel lexicalized HMM-based learning framework for web opinion mining. In *Proceedings of the 26th annual international conference on machine learning*, Vol. 10. Citeseer.

[18] Wei Jin, Hung Hay Ho, and Rohini K Srihari. 2009. OpinionMiner: a novel machine learning system for web opinion mining and extraction. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1195–1204.

[19] T Karthikeyan and N Ravikumar. 2014. A survey on association rule mining. *International Journal of Advanced Research in Computer and Communication Engineering* 3, 1 (2014), 2278–1021.

[20] Fangtao Li, Chao Han, Minlie Huang, Xiaoyan Zhu, Ying-Ju Xia, Shu Zhang, and Hao Yu. 2010. Structure-aware review mining and summarization. In *Proceedings of the 23rd international conference on computational linguistics*. Association for Computational Linguistics, 653–661.

[21] Shoushan Li, Rongyang Wang, and Guodong Zhou. 2012. Opinion target extraction using a shallow semantic parsing framework. In *Twenty-sixth AAAI conference on artificial intelligence*.

[22] Yan Li, Zhen Qin, Weiran Xu, and Jun Guo. 2015. A holistic model of mining product aspects and associated sentiments from online reviews. *Multimedia Tools and Applications* 74, 23 (2015), 10177–10194.

[23] Kang Liu, Liheng Xu, and Jun Zhao. 2012. Opinion target extraction using word-based translation model. In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*. Association for Computational Linguistics, 1346–1356.

[24] Kang Liu, Liheng Xu, and Jun Zhao. 2014. Co-extracting opinion targets and opinion words from online reviews based on the word alignment model. *IEEE Transactions on knowledge and data engineering* 27, 3 (2014), 636–650.

[25] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).

[26] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. 3111–3119.

[27] Arjun Mukherjee and Bing Liu. 2012. Aspect extraction through semi-supervised modeling. In *Proceedings of the 50th annual meeting of the association for computational linguistics: Long papers-volume 1*. Association for Computational Linguistics, 339–348.

[28] Huy Nguyen and Minh-Le Nguyen. 2017. A deep neural architecture for sentence-level sentiment classification in twitter social networking. In *International Conference of the Pacific Association for Computational Linguistics*. Springer, 15–27.

[29] Songwen Pei, Lulu Wang, Tianma Shen, and Zhong Ning. 2019. DA-BERT: Enhancing Part-of-Speech Tagging of Aspect Sentiment Analysis Using BERT. In *International Symposium on Advanced Parallel Processing Technologies*. Springer, 86–95.

[30] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365* (2018).

[31] Soujanya Poria, Erik Cambria, and Alexander Gelbukh. 2016. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems* 108 (2016), 42–49.

[32] Soujanya Poria, Erik Cambria, Lun-Wei Ku, Chen Gui, and Alexander Gelbukh. 2014. A rule-based approach to aspect extraction from product reviews. In *Proceedings of the second workshop on natural language processing for social media (SocialNLP)*. 28–37.

[33] Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. 2011. Opinion word expansion and target extraction through double propagation. *Computational linguistics* 37, 1 (2011), 9–27.

[34] Changqin Quan and Fuji Ren. 2014. Unsupervised product feature extraction for feature-oriented opinion determination. *Information Sciences* 272 (2014), 16–28.

[35] Santosh Raju, Prasad Pingali, and Vasudeva Varma. 2009. An unsupervised approach to product attribute extraction. In *European Conference on Information Retrieval*. Springer, 796–800.

[36] Toqir A Rana and Yu-N Cheah. 2016. Aspect extraction in sentiment analysis: comparative analysis and survey. *Artificial Intelligence Review* 46, 4 (2016), 459–483.

[37] Kim Schouten, Onne Van Der Weijde, Flavius Frasincar, and Rommert Dekker. 2017. Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. *IEEE transactions on cybernetics* 48, 4 (2017), 1263–1275.

[38] Lei Shu, Hu Xu, and Bing Liu. 2017. Lifelong learning crf for supervised aspect extraction. *arXiv preprint arXiv:1705.00251* (2017).

[39] Qi Su, Xinying Xu, Honglei Guo, Zhili Guo, Xian Wu, Xiaoxun Zhang, Bin Swen, and Zhong Su. 2008. Hidden sentiment association in chinese web opinion mining. In *Proceedings of the 17th international conference on World Wide Web*. 959–968.

[40] Chi Sun, Luyao Huang, and Xipeng Qiu. 2019. Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence. *arXiv preprint arXiv:1903.09588* (2019).

[41] Trang Uyen Tran, Ha Thi-Thanh Hoang, and Hiep Xuan Huynh. 2020. Bidirectional Independently Long Short-Term Memory and Conditional Random Field Integrated Model for Aspect Extraction in Sentiment Analysis. In *Frontiers in Intelligent Computing: Theory and Applications*. Springer, 131–140.

[42] Tao Wang, Yi Cai, Ho-fung Leung, Raymond YK Lau, Qing Li, and Huaqing Min. 2014. Product aspect extraction supervised with online domain knowledge. *Knowledge-Based Systems* 71 (2014), 86–100.

[43] Wenya Wang, Sinno Jialin Pan, Daniel Dahlmeier, and Xiaokui Xiao. 2016. Recursive neural conditional random fields for aspect-based sentiment analysis. *arXiv preprint arXiv:1603.06679* (2016).

[44] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, R'emi Louf, Morgan Funtowicz, and Jamie Brew. 2019. HuggingFace's Transformers: State-of-the-art Natural Language Processing. *ArXiv* abs/1910.03771 (2019).

[45] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R Salakhutdinov, and Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in neural information processing systems*. 5754–5764.

[46] Muhamad Rizky Yanuar and Shun Shiramatsu. 2020. Aspect Extraction for Tourist Spot Review in Indonesian Language using BERT. In *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, 298–302.

[47] Zhongwu Zhai, Bing Liu, Hua Xu, and Peifa Jia. 2010. Grouping product features using semi-supervised learning with soft-constraints. In *Proceedings of the 23rd international conference on computational linguistics*. Association for Computational Linguistics, 1272–1280.

[48] Yuebing Zhang, Zhifei Zhang, Duoqian Miao, and Jiaqi Wang. 2019. Three-way enhanced convolutional neural networks for sentence-level sentiment classification. *Information Sciences* 477 (2019), 55–64.

[49] Jingbo Zhu, Huizhen Wang, Muhua Zhu, Benjamin K Tsou, and Matthew Ma. 2011. Aspect-based opinion polling from customer reviews. *IEEE Transactions on affective computing* 2, 1 (2011), 37–49.