Review Highlights: Opinion Mining On Reviews

A Hybrid model for Rule Selection in Aspect Extraction

Amit Kushwaha
Zomato
Delhi, India
amit.kushwaha@zomato.com

Shubham Chaudhary
Zomato
Delhi, India
shubham.chaudhary@zomato.com

ABSTRACT

This paper proposes a methodology to extract key insights from user generated reviews. This work is based on Aspect Based Sentiment Analysis (ABSA) which predicts the sentiment of aspects mentioned in the text documents. The extracted aspects are fine-grained for the presentation form known as Review Highlights.

The syntactic approach for extraction process suffers from the overlapping chunking rules which result in noise extraction. We introduce a hybrid technique which combines machine learning and rule based model. A multi-label classifier identifies the effective rules which efficiently parse aspects and opinions from texts. This selection of rules reduce the amount of noise in extraction tasks.

This is a novel attempt to learn syntactic rule fitness from a corpus using machine learning for accurate aspect extraction. As the model learns the syntactic rule prediction from the corpus, it makes the extraction method domain independent. It also allows studying the quality of syntactic rules in a different corpus.

KEYWORDS

rule selection model; hybrid system for aspect extraction; aspect based sentiment analysis; opinion mining; review highlights;

ACM Reference Format:
Amit Kushwaha and Shubham Chaudhary. 2017. Review Highlights: Opinion Mining On Reviews. In *Proceedings of International Conference on Internet of Things and Machine Learning, Liverpool, United Kingdom, October 17–18, 2017 (IML '17),* 6 pages.
https://doi.org/10.1145/3109761.3158385

1 INTRODUCTION

Sentiment analysis has always been an active topic of research in natural language processing (NLP), in both academia and industry. There is an increased concern in the industry to understand user-feedback via reviews or feedback on different product and services. Analysis and understanding of user generated data helps in multiple facets of businesses.

IML '17, October 17–18, 2017, Liverpool, United Kingdom

ACM ISBN 978-1-4503-5243-7/17/10...$15.00
https://doi.org/10.1145/3109761.3158385

Aspect based sentiment analysis (ABSA) is a type of sentiment analysis which predicts aspect-wise sentiment. Consider this sentence for example — *The Chicken Whopper at Burger King was amazing but the service was slow.* In this sentence even though the overall sentiment is mixed, it clearly mentions two different aspects *Chicken Whopper* and *Service* in positive and negative connotation respectively.

There are two types of aspects in ABSA — explicit & implicit. In explicit aspect type, there is a clear mention of aspects which are opinionated. For example in the sentence *They serve the best lasagne*, there is an explicit aspect *lasagne*. On the other hand, in an implicit aspect type, there is an indirect indication of opinionated aspect. In the sentence *The place is quite expensive*, there is no clear mention of an aspect *price*, but *expensive* indirectly indicates towards it. In this work, we only deal with the explicit aspects in the ABSA.

In a syntatic grammer based approach a set of grammar rules are applied to the dataset to extract aspects. A syntactic grammar is defined with a clause and corresponding chunking rule, for example *VBG_DESCRIBING_NN_VV* clause defines the following syntactic pattern:

```
VBG_DESCRIBING_NN_VV:
    <NN|NN.><VB|VB.>+<RB|RB.>*<VB|VB.>
```

This clause chunks the sentence when a verb (VB) describes the opinion on a target. For example in the sentence *The place was awesome*, the verb *awesome* is describing the opinion on target *place*.

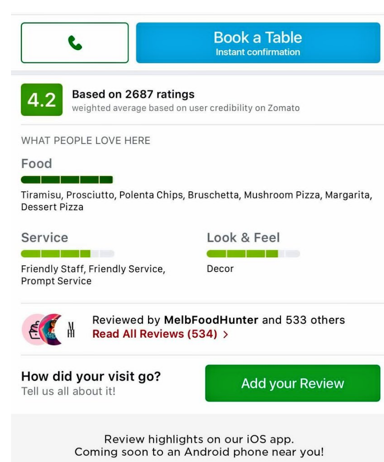

Figure 1: Review Highlight as a Product.

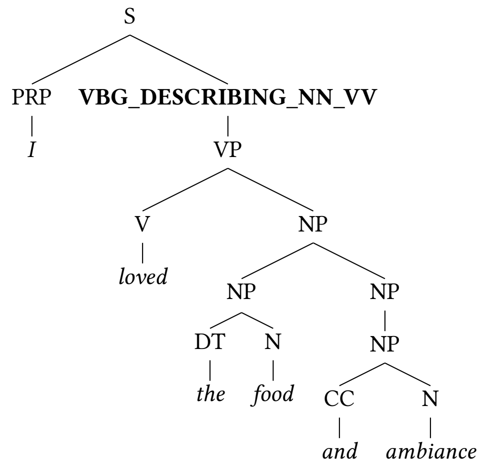

Diagram 1: Expected chunking by the syntactic rule.

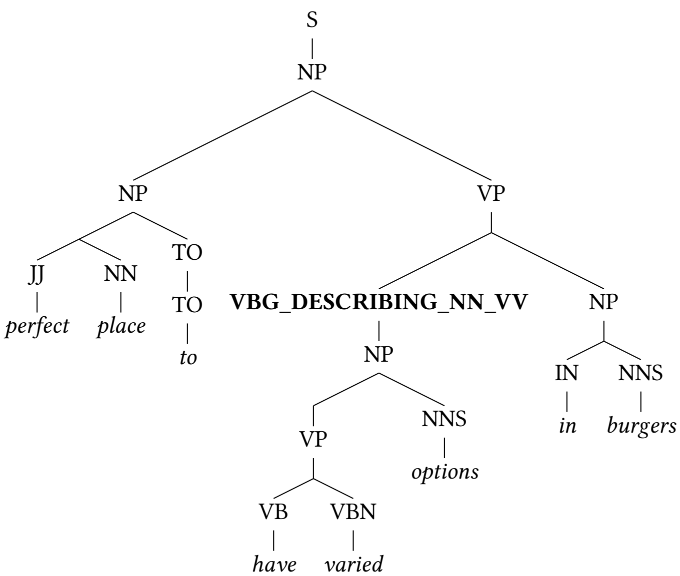

Diagram 2: Unexpected chunking of the sentence. The particular syntactic rule was not expected to parse this sentence. It not only result in incorrect extraction but also blocks the other syntactic rule which have resulted in precise extraction.

Diagram 1 shows chunking by the syntactic grammar clause — *VBG_DESCRIBING_NN_VV*. A relation extractor processes the chunked list of trees for the relationship between entities in the sentence. Though the syntactic approach is effective in parsing, it often suffers from noisy extraction. As the coverage rules are increased, it eventually results in overlapping rules and hence noise extraction. It is clear from diagram 2 that the same rule also interferes with a different sentence. Since the chunked part of diagram 2 has no aspect opinion, it results in a noisy extraction.

To address this problem, we propose a hybrid of rule-based and machine learning model for extracting the aspects and their opinion words. The machine learning model learns the effectiveness of rules to different sentence structures for a given corpus. The model can be trained to adapt for multiple domains, which makes the proposed approach domain independent. For training the model a dataset is prepared with the sentence and aspect-polarity extracted from each rule. A multi-label classifier is trained for syntactic rule prediction followed by relation extraction.

On social media platforms, user generated reviews are opinion rich texts. With the ever increasing collections of these resources, there is a tremendous opportunity to understand user's opinions and views about different aspects of the product. For instance in a restaurant review, users share their experience about several aspects of the visit — food, ambiance, and service, *etc.* As the number of reviews is not limited and each one mentions a different viewpoint, grasping the overall sense of these viewpoints from hundreds of reviews is cumbersome and time-consuming. Can we devise a way to make the decision making process faster?

To answer this, we designed the product — Review Highlight. It provides a way to recapitulate all the reviews in a single glance as shown in figure 1. It is generated by extracting the important entities from a set of user reviews. It has some key challenges for handling a varied set of audience across continents:

- Quality Entity Extraction
- Contextual diversity
- Personalization for maximum utility

Training on the SemEval-2015 restaurant's data[11], our model achieved F1-Score of 0.63 on Task 12 (slot 2) of extracting Opinion Target Extraction (OTE).

The paper is organized as follows: Section 2 presents the related work in aspect extraction tasks. Section 3 describes the main idea of the work including training and extraction. Section 4 mentions challenges in presenting the production ready aspects. Section 5 gives the results of the experimental evaluation. Finally Section 6 concludes the paper with the future extension of this work.

2 RELATED WORK

Sentiment analysis has been widely researched for a long time because of the potential business applications. However, Aspect Based Sentiment Analysis (ABSA) is a recent development in this field of research. The entity extraction can be viewed as a general information retrieval problem. Hidden Markov Model (HMM) and Conditional Random Fields (CRFs) are widely used for the extraction process. In aspect extraction, words or phrases are treated as observations and aspects-opinion expressions as underlying states. Jin et al.[8] utilized lexical HMM to extract product aspects and opinion expressions from reviews. Different from traditional HMM, they integrated linguistic features such as part-of-speech (POS) and lexical patterns into HMM. Jakob et al.[6] used CRF to extract explicit aspects in a custom corpus with data of different domains. Jin et al.[7], S. Huang et al.[5] and Choi et al.[2] also used CRF for extraction of explicit aspects. All of these methods suffer from noisy extractions.

In recent times, Deep Learning models have seen great success in Image Classification and Speech Recognition. Though in the domain of Natural Language Processing (NLP), a major focus has been on the distributed representation of learned word or phrase. Bo Wang and Min Liu[13] introduced Deep Learning framework for ABSA. They proposed parse tree distances for dependency extraction and the lexical sentiment present in the sentence. The Deep Learning, HMM, and CRF are quite suitable for NLP tasks, but these techniques require annotated data.

The work of Poria et al.[12] proposes a rule based parsing method based on dependency structure. The work of Liu et al.[10] proposes a greedy algorithm for selection of syntatic rules.

Our work is an improvement upon these works[10][12], in which we select the suitable syntactic rule for parsing as the first step. This reduces the noise extraction from other ineffective rules. This selection process is fully automated since we use a multi-label classifier.

3 HIGHLIGHT TAGS SELECTION

In ABSA the fundamental task is the extraction of aspects with their opinion words mentioned in a sentence. POS patterns carrying the opinion chunks are frequent and can be traced from simple as well as complex sentences as shown in diagram 3. To capture these patterns, a set of Syntactic Chunking Rules (SCR) is defined which parses the text to a list of chunk tree structures. To extract the opinion targets more efficiently we propose a methodology in which a model is trained to select a syntactic chunking grammar for extraction based on the syntactic features. This step is followed by a Relation Extraction (RE) which extracts the relation between entities mentioned in the sentence.

In the subsequent sub-sections, we will describe the model formulation, training and aspect extraction process.

3.1 Syntactic Grammar Model

Syntactic Grammar Model (SGM) encapsulates the collection of SCR. Each one of these chunking grammar rules are POS tag based chunk parser. A comprehensive list of SCRs is defined to cover exhaustive sentence patterns. Each SCR is typically targeted to parse a specific sentence structure. The SCRs are not necessarily exclusive and often interfere with each other, which is not suitable for the extraction process. This poses an enormous challenge to syntactic approach. Limiting the number of SCR and making them exclusive is not pragmatic when dealing with unstructured user-generated content.

We propose a method in which a multi-label classifier is trained to predict the most effective SCRs. Multi-label classification means that the model can classify more than two classes as true for a single sample. The classifier is used to predict a set of true target labels for each sample[1]. In our case, multiple syntactic rule can be applied to a single sentence predicted by the classifier.

3.2 Dataset Preparation

In algorithm 1 for multi-label classification training, a dataset O is prepared by processing the samples $\mathcal{S}$ and labeled aspect-polarity $\mathcal{AP}$. On line 3, S_i, AP_i denotes the sentence and aspect-polarity of a sample respectively. On line 4, $\mathcal{R}$ denotes a set of syntactic rules. Aspect polarity EAP_{ij} is extracted by applying each rule $r_j \in \mathcal{R}$ to S_i as shown on line 5. On line 7 the labels are marked true when rule $r_j \in \mathcal{R}$ has F1-score above threshold Φ.

Algorithm 1 Preparation of Training Label for Multi-label Classification

1: **procedure** LabelPreperationForClassifier
2: $\mathcal{D} \leftarrow (S, AP)$
3: **for** $S_i, AP_i \in \mathcal{D}$ **do**
4: **for** $syntacticRule, r_j \in \mathcal{R}$ **do**
5: $EAP_i \leftarrow ExtractAspectPolarity(S_i, r_j)$
6: $F1Score \leftarrow CalculateF1Score(AP_i, EAP_i)$
7: $\mathrm{Y}_{ij} = 1_{F1Score>\Phi}$
8: **end for**
9: $\mathrm{X}_i = SyntacticFeature(S_i)$
10: **end for**
11: $O \leftarrow (\mathrm{X}, \mathrm{Y})$
12: **return** O // Extracted Aspect-Polarity for sentence
13: **end procedure**

3.3 Training

For learning the semantic relationship, two types of features namely word features and parse tree features are developed as a feature vector for each sentence. The word features are critical for the baseline predictions while parse tree features capture the semantic structure of a sentence.

3.3.1 ***Word Features****.* Features described on word-level like individual words in a sentence is used to cover the baseline extraction of target words. Similar to the works of He et al.[4] and Giuliano et al.[3], the POS information is appended to each word feature. In our system, we use n-gram features of size up to 3, i.e. unigrams, bigrams, and trigrams. The POS information is appended to each word in a sentence. For instance, *the_DT chicken_whopper_NN at_PP Burger_King_NN are_VB amazing_JJ*, where DT, NN, PP, NN and VB are POS tags.

3.3.2 ***Parse Tree Features****.* Analyzing syntactic patterns of the sentence is used to capture the frequent syntactic patterns which are parsed successfully by one of the SCR. Unlike word features, unigrams and bigrams are not used for parse tree features because these patterns are too short to represent syntactic structures. We use n-gram features of size from 3 to 7 to capture long syntactic patterns. For example, NN -> VBP -> DT -> JJ is one of the most frequent pattern that occurred in our training dataset.

We trained a Support Vector Machine based multi-label classifier model with linear kernel. Our approach for the training the machine learning model is semi-supervised since we require annotated data to learn syntactic rule effectiveness for different sentence syntactic structure.

3.4 Extraction of Aspect-Polarity

Algorithm 2 describes the process of aspect extraction using the multi-label classifier trained on the dataset (O). On line 4, we use the trained classifier $syntacticRuleClassifier$ to predict a set of syntactic rules $\mathcal{PR}$ for given a sentence s. On line 5–7, each predicted

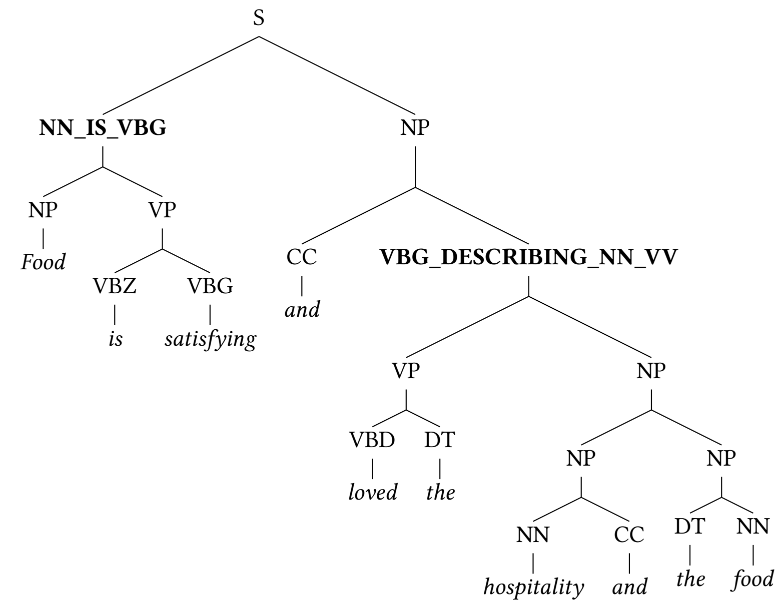

Diagram 3: Multiple syntactic rules capturing chunk trees in a single sentence.

rule $r \in \mathcal{PR}$ is used for the extraction task. The aspect-polarity extracted from each rule is merged for the final extraction results.

Algorithm 2 Extraction of Aspect Polarity Using Trained Classifier

procedure ExtractAspectPolarityUsingClassifier
  $s \leftarrow$ sentence
  $\mathcal{EAP} \leftarrow []$
  $\mathcal{PR} \leftarrow syntacticRuleClassifier.predict(s)$
  for $r \in \mathcal{PR}$ **do**
    $\mathcal{EAP}[s][r] \leftarrow ExtractAspectPolarity(s, r)$
  end for
  return $\mathcal{EAP}$ // Extracted Aspect-Polarity for sentence
end procedure

4 CHALLENGES

So far, we have discussed how an entity and its related sentiments are extracted. Review Highlight is designed to have broader scope for extracting insights from the review content. For an extracted entity to it's final disposal in the product form, it must meet some quality standards. Bringing such model to production comes with some key challenges.

4.1 Frequent words

Some common words are widely used in reviews. Although useful, they cannot be helpful if found on every highlight presented. For example, the sentence *The food is amazing at Bellagio.*, although it describes food, it does not fit into the presentation form of Review Highlights. A preferred entity is more specific like *Kung Pao Soup* in *Kung Pao Soup served was mouth watering*, it has been explicitly described and is a very specific dish which qualifies as an entity for Review Highlight.

4.2 Spelling Consolidation

Colloquial expressions tend to dominate in social media content. Some common words are widely used in reviews. Although useful, they cannot be helpful if found on every highlight presented. For example, the sentence *The food is amazing at Bellagio.*, although it describes the food, it does not fit into the presentation form of Review Highlights. A preferred entity is more specific like *Kung Pao Soup* in *Kung Pao Soup served was mouth watering*, it has been explicitly described and is a very specific dish which qualifies as an entity for Review Highlight.

4.3 Entity Diversity

For review highlights, it is important to maintain diversity. Often similar entity starts dominating. For example, top entities extracted for a restaurant can be a variant of a single type of dish like *Pizzoccheri Pasta, Farfalle Pasta, Bucatini Pasta, Bigoli Pasta, Capellini Pasta.* It mentions only the variant of a single dish, *Pasta.* A context based selection of entity is done to incorporate for multiplicity.

4.4 Time Decay Factor

A business may develop a new hype, trending attributes or a sudden social media coverage. Highlights do not result in historical data. Hence, time factor plays an important role to keep it fresh.

5 EVALUATION AND RESULTS

5.1 Dataset

There are several public annotated datasets available for ABSA[11][9]. We choose SemEval-2015 Task 12 restaurant's data for the evaluation of our work. The annotated data has a total of 2499 tuple of OTE and polarity.

Data	Training	Test	Total
OTE, Polarity	1654	845	2499

Table 1: Number of tuples annotated in Restaurant dataset

5.2 Evaluation Metrics

For dataset $\mathcal{D}$, let AP be the annotated aspect-polarity and EAP be the extracted aspect-polarity, true positive TP be $|AP \cap EAP|$, false negatives FN be $|AP \setminus EAP|$ and false positives FP be $|EAP \setminus AP|$.

We use F1-score[1] for the evaluation of our model. F1-score is the harmonic mean of precision and recall. Precision, Recall and F1-Score are defined as:

$Precision = \frac{TP}{TP+FP}$

$Recall = \frac{TP}{TP+FN}$

$F1Score = \frac{2*Precision*Recall}{Precision+Recall}$

5.3 Results

Our multi-label classifier learned semantic mapping of a sentence to most suitable rule with a rule set of size 17 and training data O of size 820 prepared by algorithm 1.

Phase	Precision	Recall	F1 Score	Support
Training	0.98	0.83	0.89	820
Testing	0.97	0.79	0.86	233

Table 2: Results of SVM Based Multi-Output Regressor

Training results from table 2 show that semantic mapping can be learned with a very high accuracy. This guided us that optimizing the rules will have a direct implication on final extraction results.

Approaches	Precision	Recall	F1 Score
Our Syntactic Grammar Model	**0.72**	0.55	**0.63**
SemEval’15 Highest Precision	0.71	0.55	0.62
SemEval’15 Highest Recall	0.26	0.79	0.50
SemEval’15 Highest F1 Score	0.68	0.71	0.70

Table 3: Aspect Extraction Performances

Table 3 presents the classification report for aspect extraction in comparison with SemEval 2015, task 12 (slot 2). Our model implementation is designed to extract aspect-polarity tuple without using supervised aspect data during prediction.

Our model f1 score is slightly less than the top f1 score acheived in SemEval 2015, task 12 (slot 2). The submissions for this task focused only on aspect extraction. The higher precision implies that our approach was able to reduce the noise in extraction. Our approach paves ways for future improvements in syntactic approach.

Table 4 shows our model’s extraction report for aspect polarity tuple. We are reporting our results without comparative evaluation since there is no established state of the art for aspect polarity tuple extraction.

Approach	Precision	Recall	F1 Score
Our Syntactic Grammar Model	0.74	0.39	0.51

Table 4: Aspect Polarity Tuple Extraction Performances

6 CONCLUSION AND FUTURE WORK

In this paper, we studied the extraction of aspect-polarity and presented in the form of Review Highlights. We showed how a simple rule based model can be more efficient in the extraction process when complemented with machine learning. This is a novel attempt in which a rule based model and machine learning are applied in a hybrid system for the aspect-polarity extraction.

The work opens the door for various relation extraction methods to incorporate our model to minimize the noise. It reduces the noise by solving the challenge of ordering in which the rules are defined in syntactic grammar for chunking process. Currently our system works only with explicit aspects. The immediate extension of this work can be to incorporate the implicit aspects. More efficient syntactic rules and relation extractor can also be the next step.

7 ACKNOWLEDGMENT

The authors would like to acknowledge the support and guidance of Paras Arora for many helpful discussions and insights. This work would not have been possible without his support and motivation. The authors would also like to thank Zomato team members for their support and useful comments on the drafts of this paper.

REFERENCES

[1] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake VanderPlas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. 2013. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 108–122.

[2] 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.

[3] Claudio Giuliano, Alberto Lavelli, and Lorenza Romano. 2006. Exploiting Shallow Linguistic Information for Relation Extraction from Biomedical Literature. In *11th Conference of the European Chapter of the Association for Computational Linguistics*. 401–408. http://www.aclweb.org/anthology/E06-1051

[4] Linna He, Zhihao Yang, Zhehuan Zhao, Hongfei Lin, and Yanpeng Li. 2013. Extracting Drug-Drug Interaction from the Biomedical Literature Using a Stacked Generalization-Based Approach. *PLOS ONE* 8, 6 (06 2013), 1–12. https://doi.org/10.1371/journal.pone.0065814

[5] S. Huang, X. Liu, X. Peng, and Z. Niu. 2012. Fine-grained Product Features Extraction and Categorization in Reviews Opinion Mining. In *2012 IEEE 12th International Conference on Data Mining Workshops*. 680–686. https://doi.org/10.1109/ICDMW.2012.53

[6] Niklas Jakob and Iryna Gurevych. 2010. Using Anaphora Resolution to Improve Opinion Target Identification in Movie Reviews. In *Proceedings of the ACL 2010 Conference Short Papers*. Association for Computational Linguistics, Uppsala, Sweden, 263–268. http://www.aclweb.org/anthology/P10-2049

[7] Feng Jin, Min-Lie Huang, and Xiao-Yan Zhu. 2011. Guided Structure-Aware Review Summarization. *Journal of Computer Science and Technology* 26, 4 (01 Jul 2011), 676–684. https://doi.org/10.1007/s11390-011-1167-y

[8] Wei Jin and Hung Hay Ho. 2009. A Novel Lexicalized HMM-based Learning Framework for Web Opinion Mining. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML ’09)*. ACM, New York, NY, USA, 465–472. https://doi.org/10.1145/1553374.1553435

[9] Bing Liu and Lei Zhang. 2012. A survey of opinion mining and sentiment analysis. In *Mining text data*. Springer, 415–463.

[10] Qian Liu, Zhiqiang Gao, Bing Liu, and Yuanlin Zhang. 2015. Automated Rule Selection for Aspect Extraction in Opinion Mining. In *Proceedings of the 24th*

International Conference on Artificial Intelligence (IJCAI'15). AAAI Press, 1291–1297. http://dl.acm.org/citation.cfm?id=2832415.2832429

[11] Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. *SemEval-2015 Task 12: Aspect Based Sentiment Analysis*. The Association for Computational Linguistics, 486–495.

[12] 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)*. Association for Computational Linguistics and Dublin City University, Dublin, Ireland, 28–37. http://www.aclweb.org/anthology/W14-5905

[13] Bo Wang and Min Liu. 2015. Deep Learning for Aspect-Based Sentiment Analysis.