ISS PRACTICAL LANGUAGE PROCESSING PRACTICE MODULE COMPETITOR ANALYSIS FOR E-COMMERCE

Cheng Kok Cheong, Daniel Tan Hoong Xiang, Onn Wei Cheng, Yang Jieshen, Yee Zhi Quan Darrel

Institute of Systems Science, National University of Singapore

1 Introduction

E-commerce sales in Singapore is estimated to hit US$8 billion in 2021 [1], in large part driven by a decline in physical retail sales due to COVID-19. With most major brands migrating their businesses onto e-commerce platforms such as Shopee, Lazada and Qoo10, businesses seek market intelligence tools which can give them an edge over competing products.

SELLinALL (SiA), the sponsor for this project, is an e-commerce marketplace integrator. They are developing a suite of market intelligence tools which will include **competitor identification**, **rank boosting**, and **product sentiment analysis** capabilities. While there are existing intelligence tools with such capabilities, SELLinALL would like to make use of NLP techniques to push the envelope and create a more effective product.

Competitor identification is a method which identifies the competition space of a product; products within it compete for marketplace visibility with each other. The vast majority of e-commerce sales start with a buyer making a product search on the marketplace. A product on the first page of search results thus has a vast advantage in visibility (and therefore chance of a sale) over a product on the second or later pages, making it important to identify other products in the competition space. Rank boosting techniques are then used to analyze the competition space for ways to boost one's product's listing position.

Another capability provided is product sentiment analysis. About 93% of consumers indicated that their purchase decisions are influenced by online reviews [2]. This makes reviews an important source of customer sentiments and feedback. Mining product reviews can not only identify the strengths and weaknesses of one's own products, but those of one's competitors as well.

Existing solutions to these problems generally do not take advantage of advanced NLP techniques. A common but naive approach to competitor identification is to search up products in the same category as the target product. Another approach requires the seller to provide the search keywords they are targeting for their products, which are then searched on the marketplace to form the competition space. The subsequent rank boosting is usually serviced by a person in the loop with deep industry knowledge. For sentiment analysis, intelligence tools generally focus on summarizing review scores across the seller's products, instead of performing any deep dive into the content of the reviews.

The team intends to leverage on NLP techniques to improve the level of automation, effectiveness and therefore business value of these capabilities.

2 Dataset

We made use of two datasets in our methodology: 1) an annotated dataset of product reviews from the Amazon marketplace, and 2) raw product data scraped from the Shopee marketplace. To narrow the scope of the project, only a subset of both datasets relating to fashion and lifestyle products will be used.

The full Amazon dataset was obtained from a team in University of California, San Diego [3] which consists of 233.1 million Amazon reviews dating from May 1996 - Oct 2018. Only the "Amazon Fashion" segment of the data was used, which consists of 883 thousand reviews across 186 thousand products.

The raw Shopee data was scraped using a publicly accessible product search API and includes product titles, prices, data on reviews, and other information pertaining to the product. Manual annotations were performed on the product titles and reviews. All annotations were performed using Prodigy, an annotation workflow tool closely integrated with the Spacy framework. Only products with a "Men's Fashion" & "Women's Fashion" were scraped.

3 Methodology

3.1 Competitor Identification

As mentioned in the introduction, one existing technique for competitor identification is for the seller to provide search keywords they think their product will appear in. This requires manual input by the seller, and he may also not know the ideal search space for his product. It is observed that competing products usually share some similar words in their titles which relate to the type of item being sold. For example, in Figure 1, the words "leather" and "shoe" are common words across competing products in the "men's leather shoe" search space, and they also well describe the type of item they are. If item type information can be extracted from the product title, it can be used to automatically generate the competition space instead of requiring manual input.

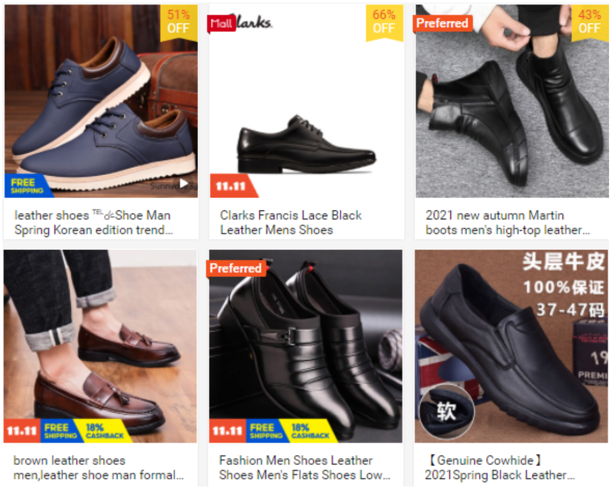

Fig. 1: Shopee products within the same competition space.

For this task, we trained a Named Entity Recognition (NER) model using Spacy to pick out key terms in the product title that can be used to search for similar products. Spacy's NER model uses a transition-based architecture which predicts a sequence of 'actions' required to correctly tag a text sequence; the probability of each action in the sequence is then modelled using a stacked LSTM [4]. The model was able to achieve 86.4% accuracy on NER tasks against the OntoNote 5 dataset while being reasonably fast, making it suitable for our application. We annotated 1,130 product titles from the Shopee marketplace using the tags defined in Table 1. While TYPE information of a product is most relevant to finding its competitor, we included other annotated tags such as BRAND and DEMOGRAPHIC to allow the user to refine their competitor search based on their desired parameters. The additional tags will also allow SiA to automatically catalogue the products of sellers managed by them in the future.

After this information extraction step, the desired tags are joined into a single string and passed to the Shopee product search API to extract a competition space. As a default, TYPE and DEMOGRAPHIC tags are included in the search.

3.2 Rank Boosting

The relevance of the product title to the search term has been observed to be one of the most significant factors determining

Tag	Description
TYPE	The base type of item. E.g. "leather shoe", "dehumidifier"
DEMOGRAPHIC	Relating to demographic information of the product. E.g. "men's", "Ladies's", "kids"
BRAND	Relating to product brand.
MODEL	Relating to product models or brand-specific categories.
VAR_SIZE	Relating to the size of the product. E.g. "11 inch", "M"
VAR_QTY	Relating to the quantity of units in a product. E.g. "pack of 5", "10 bottles"
VAR_COLOR	Relating to details on the design and other variation within a product.

Table 1: Tags used and their description for the competitor identification NER model.

a product's listing rank. While the exact ranking mechanism used on each marketplace is unknown, we can infer that some terms in the title of top-ranking products are contributing to their popularity. By analyzing such terms, we can potentially emulate their success and increase our own products' ranking.

From the competition space, we can generate a word cloud which contains the most frequently used n-gram tokens in the product titles. Based on consultation with several sellers, it was found that using bigrams and single tokens for generating the word cloud provides the most interesting and useful insights for rank boosting.

3.3 Aspect based Sentiment Analysis (ABSA)

It is common for a single review to have different sentiments on multiple aspects. For example, a review on a dress might contain a positive sentiment on the quality of the dress, but a negative sentiment on the size. Traditional sentiment analysis aggregates the full review into a single sentiment and might hide certain sentiments specific to some aspects. To provide the user a deeper level of analytical granularity, we can perform ABSA on the product reviews instead.

The ABSA task is divided into the following sub-tasks:

- Definition of different aspect topics
- Detection of the identified aspect topics
- Analysis of sentiments of the identified topics

For each of the sub-task, one or more models are developed to find the best method for each sub-task. Topic modelling using Latent Dirichlet Allocation (LDA) is used to determine the different aspect topics. A rule-based model and an NER model were developed for the aspect detection subtask, while several machine learning methods were used to determine the sentiments of the aspect detected.

We can add value to the system by performing ABSA on the competition space identified in the competitor identification module mentioned earlier. We can then compare the sentiments broken down by aspects for our own product, against those of our competitors, providing insight into our product's strengths and weaknesses.

3.3.1 Aspect Topics Modelling with Latent Dirichlet Allocation (LDA)

Before developing the ABSA system, we need to determine the aspects we want to detect. LDA is an unsupervised clustering technique that categorizes documents in a text corpus into a predescribed number of topics. The algorithm learns the list of keywords which define the characteristics of each topic based on the assumption that similar documents often contain certain words in common. Being a statistical method, LDA describes each document as a probability distribution over the latent topics, and each latent topic as a probability distribution over all words in the corpus.

LDA can be used to discover aspects of fashion products which are of concern to the users when providing product reviews on e-commercial platforms. This approach is largely suitable because users often use common words to describe specific quality of a product. For example, words such as *big*, *small*, *medium*, *large* and etc frequently appear in reviews that discuss about the *size* property of fashion item.

An LDA model is taught using reviews from the Amazon dataset. The reviews are first cleaned by removing unnecessary *html* elements, hyperlinks, English stop-words and punctuation, followed by lemmatizing and converting each word to lowercase. Finally, the reviews are tokenized and the bag of all tokens are ready for training an LDA model.

We start with learning three topics, and progressively increment the number of topics until the keywords in each latent topic appear conceptually coherent with each other. By observation, the conceptual boundary between topics become clearer as the number of topics grow. Eventually, we decide to stop at 10 topics, at which point the topics could be reasonably distinguished according to either the types or the aspects related to fashion products.

We identified 6 aspects of fashion items that are commonly reviewed - *size*, *comfort*, *washability*, *price*, *quality* and *appearance*. For the subsequent aspect detection models, we will absorb *washability* into *quality* as it is considered a specific property of the material of fashion items. An additional aspect - *delivery*, which is not directly related to the property of fashion items, will also be included to further evaluate the quality of service provided by product vendors.

Topic	Important words	Interpretation
0	ring, gift, watch	Related to fashion items often purchased as gift.
1	size, small, fit, large, big, medium	Related to *size* aspect of fashion items.
2	fit, comfortable	Related to *comfort* aspect of fashion items.
3	wash, make, well	Related to *washability* aspect of fashion items.
4	shoe, foot, sock, boot	Related to footwear.
5	honest, discount, purchase	Related to *price* aspect of fashion items.
6	material, fabric	Related to material *quality* aspect of fashion items.
7	bag, pocket, wallet, purse, zipper, hold, card	Related to fashion items which serve a purpose.
8	belt, money, cheap	Related to belt and *price* aspect of fashion items.
9	nice, look, color, beautiful	Related to *appearance* aspect of fashion items.

Table 2: LDA with 10 topics and topic interpretation

3.3.2 Aspect Detection - Rule-Based Approach

Adjectives are words that describes nouns and often forms an opinion or contains a sentiment of the noun that it is describing. The extraction of such descriptors, a combination of adverb and adjectives, and the noun phrases that it is describing can be used as the aspect terms for sentiment analysis. The method propose the use of a rule-based method to extract the aspect terms.

To ensure that the sentiment of the different aspect topics are identified correctly, it will be important to extract out the noun phrases and descriptors that are connected semantically. To do so, each sentence of the review is passed through a dependency parser to determine the semantics connection.

There are two general ways adjectives can be used to describe a noun; either as an adjectival modifier as illustrated in Figure 2 or as an adjectival complement as illustrated in figure 3. In the case of an adjectival modifier the noun or noun phrase will be the head of the adjective, while in the case of an adjectival compliment both the noun and the adjective will have a verb or auxiliary as the head. Adverbial modifiers such as ”very”, are included as part of the descriptors to ensure that the intensity of the opinion is being considered [5]. Noun phrases, such as ”the maxi dress”, are also considered in entirety instead of just the noun, ”dress”, to capture the full concept of the noun phrase.

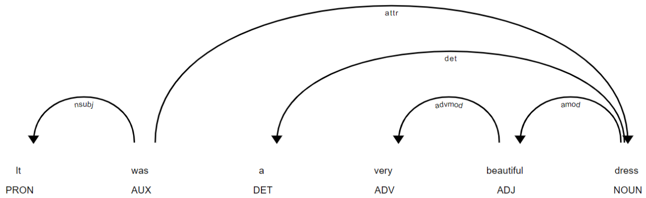

Fig. 2: Adjectival Modifier

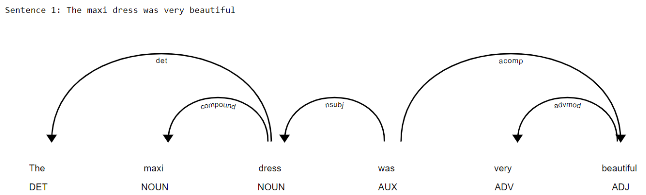

Fig. 3: Adjectival Complement

Using the above described rule-based approach, the noun phrases and descriptors of a particular aspect can be extracted. To determine the topic of the aspect terms, the extracted noun phrases and the descriptors are then matched with pre-defined noun and adjective keywords for each topics to determine the aspect category. This allows the aspect topic to be extracted either in explicit cases where the aspect term is mentioned as a noun, or in implicit cases where it is implied by the adjective.

The explicit case is illustrated in the Figure 4. The adjective detected, ”good”, is a general adjective and does not refer to a specific aspect, but the noun detected, ”quality” is relates to the quality aspect. In the implicit case illustrated in Figure 2, the noun detected, ”dress”, refers to the item purchased and does not match any specific keywords of any aspect. However, the adjective ”beautiful”, matches a keyword relating to the aspect topic appearance. In other words, the specific aspect is implied by the adjective in this case. The same implicit matching will also be used when pronouns are detected as they usually do not contain any aspect information.

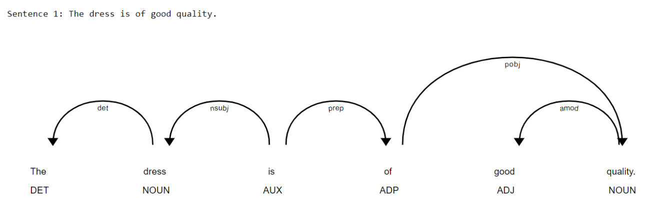

Fig. 4: General Adjective with Matching Noun

The noun phrases and descriptors that are identified are concatenated and put through a sentiment analysis model to determine the sentiment of the aspect term/topic. Negation terms connected to the descriptors, such as ”not” in the phrase ”not beautiful”, will be detected and passed to the sentiment analysis module to negate the sentiment analysed. If multiple aspects are found for a particular topic, the aspect sentiments are aggregated and reported as a sentiment for the topic.

In cases where the noun phrase and descriptors pairs cannot be found by the above mentioned rules either due to incorrect grammar or incomplete sentence, the model will rely on its fallback methods. It will first try to associate the adjective to the noun subject of the sentence as suggested by [5]. If the noun subject and descriptor are found, the aspect topic and sentiment will be identified following the same methods described above.

If no noun subjects are found, the adjective will be matched with the keywords to determined if it is part of any aspect topics. If a match is found, the adjective and the topic will be concatenated and passed into the sentimental analysis model to determine the sentiment of the aspect topic. This is illustrated in Figure 5. The phrase ”Very beautiful” only contains an adjective and does not have any connected noun or a noun subject in the sentence. However, the adjective, ”beautiful”, implies that it is referring to the topic appearance.

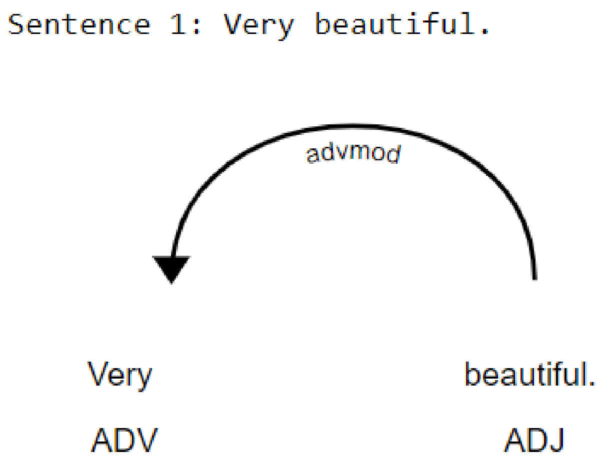

Fig. 5: Sentence with no noun

3.3.3 Aspect Detection - NER Approach

In addition to the rule-based approach, we also experimented with an end-to-end deep learning method of detecting aspect spans. We hypothesize that an NER model should be able identify spans within a review text relating to each of the aspects. These spans can then be passed to the sentiment analysis module to determine its polarity.

The same Spacy NER architecture as the one used for competitor identification was used for this aspect detection approach. A total of 1778 Shopee product reviews with annotated aspect spans were used to train the model.

3.3.4 Sentiment Analysis

The aspect text spans extracted will be fed into the sentiment analysis model, which determines the polarity of the span (either POSITIVE or NEGATIVE). For this purpose we train a text classification model using data from both the Amazon Fashion datasets, and the Shopee review dataset. We experimented with 3 model types: Convolutional Neural Networks (CNN), Logistic Regression and Support Vector Machines (SVM).

The CNN model is provided by Spacy, and we utilize the en_core_web_sm pipeline with pre-trained tok2vec feature extractor. For the logistic regression and SVM, unigrams and bigrams are extracted from the data and passed through a tf-idf vectorizer to create a document term matrix using the methods described by Rahman et al. [6]. This feature matrix is then input to the models.

Annotation of the Amazon dataset utilizes customer product ratings of 1 to 5 to label if the sentiment of a review is POSITIVE or NEGATIVE. Reviews with scores less than 3 will be labelled as negative sentiment while more than 3 will be labelled positive. For the Shopee review data, we extract the annotated aspect spans and re-annotate them as POSITIVE or NEGATIVE. Both Amazon and Shopee datasets will then be subjected to a 80%-20% train-test split to train and evaluate the models. We will train all three models using both datasets and evaluate their performance.

4 Experimental Results & Discussion

In this section, we will discuss the results of our competitor identification, aspect detection and sentiment analysis model.

4.1 Competitor Identification

The NER model achieved an F1-score of 0.74 (0.77 P, 0.71 R) against a holdout set of 282 examples. The per-type score breakdown is shown in Table 3.

Tag	P	R	F1
TYPE	0.70	0.60	0.64
DEMOGRAPHIC	0.93	0.90	0.91
BRAND	0.89	0.90	0.90
MODEL	0.58	0.51	0.54
VAR_SIZE	0.82	0.78	0.80
VAR_QTY	0.75	0.60	0.67
VAR_COLOR	0.78	0.74	0.75

Table 3: Competition Identification NER model per-type score.

DEMOGRAPHIC & BRAND tags performed well, likely due to the limited scope of text that can appear under them. The TYPE tag, probably most important for identifying competitors, performed relatively poorly with only 0.63 F1-score. In spite of this, TYPE tags extracted during testing were found to be still acceptable for searching up competitor products. The loss in accuracy may be due to the model missing words in multi-word TYPE tags or splitting a single multi-word tag into several. In these cases, the TYPE tag result will still be usable for searching up competitor products.

Our NER model was fined-tuned on top of the Spacy en_core_web_md pipeline. This was found to increase overall F1-score by 0.03 points when compared with starting from a blank model, likely due to the additional use of pre-trained GLOVE vectors in the pipeline.

4.2 Aspect-Based Sentiment Analysis

4.2.1 Aspect Detection - Rule-Based Approach

To evaluate the rule-based aspect detection model, we tested if aspect topics are identified correctly against two self-annotated datasets from Amazon and Shopee. The results per label are shown in Table 4 and Table 5 for the Amazon and Shopee data set respectively. The rule-based model achieved an F1-score of 0.68 (0.72 P, 0.64 R) on the Amazon dataset, and an F1-score of 0.62 (0.65 P, 0.58 R) on the Shopee dataset.

Label	P	R	F1
SIZE	0.71	0.80	0.75
COMFORT	0.94	0.67	0.78
APPEARANCE	0.76	0.67	0.71
QUALITY	0.68	0.41	0.51
PRICE	0.47	0.53	0.50
DELIVERY	0.60	0.44	0.51

Table 4: Aspect Detection on Amazon Dataset per label score

Label	P	R	F1
SIZE	0.61	0.63	0.62
COMFORT	0.75	0.71	0.73
APPEARANCE	0.39	0.70	0.50
QUALITY	0.76	0.58	0.65
PRICE	0.50	0.64	0.56
DELIVERY	0.81	0.46	0.58

Table 5: Aspect Detection on Shopee Dataset per label score

It was found that the model performed relatively well when the nouns and adjectives are clearly present in the review. However, it is not able to detect sentiments related to verbs. This is illustrated in Figure 6. The sentence, ”Not well made”, indicated a sentiment on the quality topic. However, as it is not detected as an adjective, it was not recognised by the rule-based model.

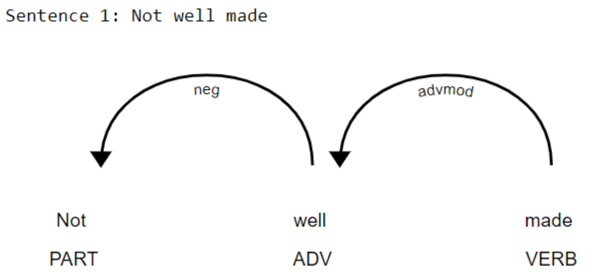

Fig. 6: Unable to detecting sentiments relating to verbs

Comparing the results between the two datasets, it was found that the algorithm works better on the Amazon dataset as compared to the Shopee dataset. Upon inspection of the dataset, the reviews from Shopee seems to have shorter sentence, sometimes with improper grammar and colloquial terms, which is typical in local speech. However, the use of such language style, does not fare well with the dependency parser and part-of-speech tagger provided by Spacy following the English language model. Therefore, the algorithm either makes mistakes or relies on fallback methods, which is of lower accuracy. A sample review from Amazon and Shopee is shown in Figure 7 and Figure 8 respectively.

```
The dress is very pretty, but it runs very small in the waist
and very large in the bust. If you're not an extreme hour-glass,
be prepared to alter this before wearing it. I'm short, so I
plan to take some fabric either from the hem or the oversized
bust to redistribute as needed, personally.
```

Fig. 7: Sample Review from Amazon

```
My lovely OREN T shirt.Comfort and cooling in Singapore weather.
This is my 6th Oren T shirt (5 for working ) and this new one
for exercise. Thank you seller for the fast delivery. Cheers
```

Fig. 8: Sample Review from Shopee

4.2.2 Aspect Detection - NER Approach

The NER model achieved an F1-score of 0.70 (0.72 P, 0.69 R) against a holdout set of 199 examples. The per-type score breakdown is shown in Table 6.

Tag	P	R	F1
DELIVERY	0.81	0.84	0.82
COMFORT	0.59	0.52	0.55
QUALITY	0.66	0.70	0.68
PRICE	0.82	0.82	0.82
SIZE	0.60	0.49	0.54
APPEARANCE	0.68	0.50	0.57

Table 6: Competition Identification NER model per-type score.

SIZE, COMFORT and APPEARANCE tags are somewhat difficult to predict for due to the wide range of expressions that can be used to describe them. SIZE may deal with length, fitting, cut, and a myriad of descriptions for the sizing of the product. COMFORT and APPEARANCE are similar, with different ways to describe them in a review. In contrast, DELIVERY is solely focused on either quality of packaging or speed, while PRICE deals only in a few descriptions of value, making them easier to predict for and thus have higher accuracy. Subjectively, the text spans extracted during testing were quite appropriate, even if the span boundaries may vary from the ideal by 1-2 tokens. See Table 7 for examples of extracted aspects.

The annotation strategy for this aspect detection model was to capture any text span that expressed an opinion belonging to one of the six aspects. As far as possible, aspect annotations were kept to 5 tokens or less to make its easier for the NER model to learn. However, there were examples that required significantly longer text spans to extract a coherent aspect. For example, ”*have to order 2 - 3 sizes larger than suggested*” needs to be annotated in full to highlight the fact that the sizing was inaccurate. Not unexpectedly, the model has trouble extracting such longer spans, contributing to a loss in accuracy. It is expected that with more annotated data, such examples will be better represented during training and thus accuracy will improve.

Similar to the NER model used in competitor identification, the F1-score was improved by 0.02 points by using en_core_web_md as a base model for training.

4.2.3 Sentiment Analysis

The CNN model generally performed better than logistic regression and SVM models, likely due to the pre-trained feature extractor. All models trained on the Amazon dataset suffered a loss in performance when evaluated against the Shopee dataset. This is not surprising due to the differences in grammar and syntax between the datasets, as highlighted in Figure 7 and Figure 8. A comparison of model result is shown in Fig.9.

This performance loss suggests that the Amazon dataset cannot be universally applied to other applications. We conclude that the best approach is to use a CNN model trained on Shopee review data.

Tag	Examples
DELIVERY	”delivery was fast” ”delivery well received” ”delivery took awhile”
COMFORT	”thin comfortable fabric” ”material smooth and cooling” ”inside is very scratchy”
QUALITY	”quality is great” ”very good quality” ”quality was unexpectedly bad”
PRICE	”price was cheap” ”good deal” ”good value”
SIZE	”L fit ok” ”size 2xl seem small” ”cutting is bad”
APPEARANCE	”love the colours” ”this top is so cute” ”nice patterns”

Table 7: NER-extracted aspect examples.

	Testing Datasets	Amazon Fashion			Shopee Aspect		
		Macro Average			Macro Average		
Training datasets	Model	Precision	Recall	F1 score	Precision	Recall	F1 score
Amazon Fashion	Logistic Regression	90%	89%	90%	76%	74%	75%
	SVM	85%	85%	85%	72%	72%	72%
	CNN	99%	99%	99%	82%	72%	74%
Shopee Aspect	Logistic Regression	71%	66%	68%	81%	72%	73%
	SVM	65%	71%	65%	79%	78%	78%
	CNN	62%	62%	62%	92%	92%	92%

Fig. 9: Sentiment Models

5 Conclusion and Future Work

In this report we have proposed a market intelligence system with competitor identification, text analytics, and sentiment analysis capabilities. We have demonstrated that advanced NLP techniques can be employed to improve existing market intelligence systems, and derive greater business value.

Our experiments suggest that our NER-based aspect detection outperforms the rule-based approach. The NER-based approach has other advantages as well, such as being scalable and simpler in implementation.

Future work will involve expanding the scope of prediction to other product categories. This will require exploration into other possible aspects for each product category through LDA. Of interest is whether a single model can service all possible types of products and aspects, or whether a separate model is needed for each product category to maintain a reasonable accuracy.

Another area of improvement is increasing the size of our training dataset with more annotations. During training, it was observed that as more annotations were collected, evaluation accuracy appeared to drop slightly for both NER models. We reason that due to the large variation in possible product titles, greater uncertainty was being injected into the model with every new annotation. With a much larger training dataset, enough examples should have been collected to cover most possible variations, leading to subsequent improvements in accuracy.

6 Contributions

This project was made possible due to the significant efforts made by each member of the team. The individual contributions of each member to the system's design are detailed below.

- Cheng Kok Cheong
    1. Application of the LDA technique to determine the number of aspect topics. Also responsible for the generation of the adjective and noun aspect topic keywords using wordnet.
    2. Generation of the competitor space by integrating Shopee API and the NER model
- Daniel Tan Hoong Xiang
    1. Training and development of multiple sentiment analysis models using Amazon fashion and Shopee Aspect dataset. Also responsible for setting up of the inference API for the various model developed.
    2. Co-contributor to the sentiment annotation for the Shopee Aspect dataset.
- Onn Wei Cheng
    1. Development of the rule-based aspect detection model
    2. Responsible for integrating the rule-based aspect detection model with the sentiment analysis model
- Yang Jieshen
    1. Annotation of Amazon reviews for aspect based model
    2. Integration of overall system and recording of demo video
- Yee Zhi Quan Darrel
    1. Annotated datasets for, and developed NER models for competitor identification and aspect detection
    2. Implemented word cloud technique for rank boosting

7 References

[1] Prisca Ang, "E-commerce sales in singapore forecast to hit $13.4 billion by 2026: Report," 8 2021.

[2] "2017 state of online reviews," http://learn.podium.com/rs/841-BRM-380/images/2017-SOOR-Infographic.jpg, 2017.

[3] Jiacheng Li Jianmo Ni and Julian McAuley, "Justifying recommendations using distantly-labeled reviews and fined-grained aspects," *Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

[4] Jiacheng Li Jianmo Ni and Julian McAuley, "Parsing english in 500 lines of python," 2013.

[5] Aris Pattakos, "Aspect-based sentiment analysis using spacy textblob," https://towardsdatascience.com/aspect-based-sentiment-analysis-using-spacy-textblob-4c8de3e0d2b9/.

[6] Sheikh Shah Mohammad Rahman, Khalid Biplob, Md Rahman, Kaushik Sarker, and Takia Islam, *An Investigation and Evaluation of N-Gram, TF-IDF and Ensemble Methods in Sentiment Classification*, pp. 391–402, 07 2020.