Applying Sentiment Product Reviews and Visualization for BI Systems in Vietnamese E-Commerce Website: Focusing on Vietnamese Context

Ngoc-Bao-Van Le 1 and Jun-Ho Huh 2,*

1 Department of Data Informatics, (National) Korea Maritime and Ocean University, Busan 49112, Korea; vanlnb@g.kmou.ac.kr
2 Department of Data Science, (National) Korea Maritime and Ocean University, Busan 49112, Korea
* Correspondence: 72networks@kmou.ac.kr

Abstract: Product reviews become more important in the buying decision-making process of customers. Exploiting and analyzing customer product reviews in sentiments also become an advantage for businesses and researchers in e-commerce platforms. This study proposes a sentiment evaluation model of customer reviews by extracting objects, emotional words for emotional level analysis, using machine learning algorithms. The research object is the Vietnamese language, which has special semantic structures and characteristics. In this research model, emotional dictionaries and sets of extract rules are combined to build a data training data set based on the semantic dependency relationship between words in sentences of the given Vietnamese context. The recurrent neural network model (RNN) solves the emotional analysis issue, specifically, the long short-term memory neural network (LSTMs). This analysis model combines the vector representations of words with a continuous bag-of-words (CBOW) architecture. Our system is designed to crawl realistic data in an e-commerce website and automatically aggregate them. These data will be stored in MongoDB before processing and input into our model on the server. Then, the system can exploit the features in products reviews and classify customer reviews. These features extracted from different feedback on each shopping step and depending on the kinds of products. Finally, there is a web-app to connect to a server and visualize all the research results. Based on the research results, enterprises can follow up their customers in real-time and receive recommendations to understand their customers. From there, they can improve their services and provide sustainable consumer service.

Keywords: Natural Language Processing; e-commerce website; product reviews; Vietnamese; sentiment analysis; text classification; big data; application

1. Introduction

In recent years, the development of information technology (IT) and internet applications has allowed worldwide users to create a massive amount of information. Contents are created daily on various online platforms such as status social media streams, videos or pictures on e-commerce websites, applications, etc. Social media, e-commerce websites, mobile platforms, applications, etc., allow people to share information and represent their attitudes, perspectives on products and services and other social issues. Product reviews become more important in the buying decision-making process of customers [1]. Reviewing customer’s comments is a great way to develop businesses by understanding the strengths and weaknesses of a company’s goods and services in order to capture what are their demands and give them the best choices. Exploiting and analyzing customer product reviews also become an advantage for businesses in many sectors, especially e-commerce platforms.

In the past, it was common that humans are in charge of handling customer opinions. These ways of processing led to unnecessary errors and increase cost of businesses and
organizations. Natural language processing technology, language mining techniques have been presented to help solve this problem, which can be mentioned in Refs. [2–4], etc. In the field of e-commerce, classification and sentiment analysis are often researched on the Amazon platform, which is the largest online supplier in the world by having a wide range of products and numerous reviews. For example, a model of extracting features from user opinion information and evaluated customer opinions about iPhone 5 reviews with three basic levels—good, average and bad—was conducted with data crawled from Amazon [2]. A continuous naïve Bayes learning framework is used to classify sentiment product reviews in research [3]. Currently, emotion analysis usually has three main methods: The primary method bases on the dictionary of emotional word; method deep learning neural network method of combining rule-bases and corpus-bases. Lately, the problem of analyzing user emotions based on reviews has been carried out mainly in the English language, with many different technologies and models applied. However, the analysis of reviews, especially Vietnamese product reviews, has not been applied in terms of popularity. Our problem is detecting and extracting the discussion sentences of each evaluation aspect in the article while predicting the user’s rating and importance of each aspect. This is a form of opinion analysis and language processing problems in Vietnamese.

In this research, to address the problem, our system is designed to crawl realistic data in an e-commerce website and automatically aggregate them. These data will be stored in MongoDB before processing and input into our model on the server. Then, the system will exploit all features of the products that customers were interested in through reviews in the Vietnam e-commerce market and different feedback on shopping processes depending on the kinds of products. Finally, a web-app is built, connected to a server to visualize all the research results so that small and medium enterprises can follow up their products and their customers in real-time. Based on the research results, enterprises can follow up their customers in real-time and receive recommendations to understand their customers. From there, they can improve their services and provide sustainable consumer service. Aiming at the problem of sentiment analysis, this research uses machine learning algorithms to propose a sentiment evaluation model for customer reviews by extracting objects and sentiment words. Combining our method with the sentiment dictionary and the extraction rule set, a data training data set is constructed based on the semantic dependency in the Vietnamese context. The sentiment analysis problem is solved by the recurrent neural network model (RNN), which has a commonly used variant currently, called the long short-term memory neural network (LSTM), which combines the vector representation of words with a continuous bag of words (CBOW) architecture.

The rest of this paper is organized as follows: Chapter 2—Research Background provides a brief overview of E-commerce development in Vietnam, as well as the chosen research objectives; Chapter 3—Literature Review presents previous research or works related to sentiment reviews; Chapter 4—Design and Implementation discusses the methodologies for collecting reviews, research design and process of analyzing and classifying sentiment reviews; the results and discussions are presented in Chapter 5 of paper—Result and Discussion; Chapter 6, the last chapter, presents the conclusions of this research and future work aiming to improve the results.

2. Research Background

The e-commerce market is showing strong growth in Vietnam, with a high internet accessing rate (ranked 17th in the world) and rising smartphone penetration rates (E-commerce Industry in Vietnam, 2018). According to the Vietnam E-commerce Association (VECOM) in the Vietnam Online Business Forum 2019 (VOBF), the retail market started out at 4 billion USD in 2013 and reached 7.8 billion USD in 2018. If the growth rate in 2019 and 2020 continues to be maintained at 30%, by 2020, the scale of the retail e-commerce market will reach about 13 billion USD, which is higher than expected in the overall plan for e-commerce trade development for 2016–2020 [5]. In addition to transacting via traditional e-commerce platforms, consumers are gradually shifting to online transactions via social
network sites, with a tendency to use multi-platform channels because of intelligent searching or convenient shopping experiences. Vietnamese consumers now contribute to the greater reach of the internet trend by using multi-platform channels [6].

The trend of the digital economy is changing the shopping habits of customers in many industries. The e-commerce market is increasingly open, with many models, participants and supply chains also gradually changing in a more modern direction, with digitalization and information technology support. The e-commerce trends and consumer behavior in the world have changed rapidly. Restricted by lockdown, self-isolation and social distance, these driving factors have led to the increase in households purchasing luxury products through e-commerce [7]. COVID-19’s effects on five global e-commerce companies were studied in Ref. [8]. The results indicated that new coronavirus cases, current coronavirus cases and cumulative deaths affected almost all of the companies studied. A study on the impact of COVID-19 on the e-commerce market in Malaysia also shows that shipment processes are affected, resulting in a shortage of some products and kits manufactured in China [9]. In particular, due to the impact of the COVID-19 pandemic, the e-commerce market is becoming more vibrant because of the change in consumers’ buying habits, switching from traditional buying habits to e-commerce purchases. Consumers’ purchasing decisions have changed significantly during this period. Consumers’ awareness and their experiences become important influencing factors when making decisions [10].

There were seven factors affecting consumer’s online shopping buying behavior: ease of use, perceived risk, perceived usefulness, website design, economic factor, availability of products and customer satisfaction [11]. Normative determinants and hedonic motivation are related to German consumers’ intention of online purchasing were also investigated [12].

Especially, the influence of social distancing has made the hedonic motivation factor of gen Z even higher. Another piece of evidence comes from research in Bangladesh, which has been significantly impacted by COVID-19, resulting in changes in online shopping such as products, time-saving, payment and administrative factors [13]. In Ref. [14], five factors influencing the online shopping intentions of consumers in Ho Chi Minh City during the COVID-19 period were arranged in order of importance from high to low: perceived usefulness; reference group; safety and confidentiality; reputation. Nowadays, customers have become increasingly meticulous when shopping. Notably, customers will also measure the value they receive when they pay for a product, not only the price. This means they will consider the quality, features and also fashion. When people become busier, reviews from other customers can convince them and also can affect purchasing decisions. Almost 90% of consumers read reviews before visiting a business, 31% are likely to spend 31% more on a business with excellent reviews and 72% will take action only after reading a positive review [15].

With the development of multi-platform channels in e-commerce, primarily via mobile applications and e-commerce websites through the internet [16], customers are increasingly able to conveniently and freely comment on products after they purchase them or share their shopping experiences.

With such numerous customer comments, effectively exploiting insights from customer comments plays an important role. Collecting and analyzing customer comments can create timely changes in product quality, logistics issues and even the customer relationship management (CRM) system [17] to improve the business efficiency of each business and earn revenue to attract customers more effectively. Currently, there are many research studies worldwide that analyze the sentiments of customer comments in many different languages, as in Refs. [18–21]. Other researchers conducted many studies and works related to the analysis and classification topics of Vietnamese sentences, as in Refs. [20,21], or works related to specific service products, such as hotel services, as in Ref. [22]. However, there has not been much research on sentiment customer reviews in the context of Vietnamese e-commerce transactions, despite the volume of transactions and fertile information in the current e-commerce environment.

In this research, a process is established to collect and polarize opinion and feature classifications of Vietnamese customer reviews on e-commerce websites. Especially, prod-
uct review data, which was used for building data training sets and analysis, were collected from a famous e-commerce website in Vietnam. Tiki is an all-in-one commercial ecosystem consisting of member companies, such as TikiNOW Smart Logistics, which provides end-to-end logistics services; Ticketbox, which offers top-notch event and movie ticket services; and Retailer Tiki Trading and Exchange, which offers 10 million products in 26 categories and serves millions of customers nationwide. Tiki’s unique features include 10%+ discounts on 90% of its products, 24 h or 2 h delivery, customer hotline offering support 12 h a day/7 days a week and return and exchange policies [23]. Tiki had 19.74 million visits in December, 2017. With the diversity of goods as well as the trust of Vietnamese customers, Tiki’s products review data are a reliable source of data reflecting customers’ thoughts and many research angles that are exploitable.

3. Literature Review

3.1. Mining Product Reviews and Sentiment Analysis

In English, there have been several works conducted. Opinion mining was applied to Vietnamese and American online reviews to express hotel service preferences introduced in Ref. [24]. The study collected 37,712 online reviews of Vietnamese hotels by domestic and foreign customers in both English and Vietnamese on hotel booking and review channels Agoda.com (accessed on 17 April 2020). The study was conducted with five steps: data collection, data preprocessing, lexicon-based model design, machine learning model design and hybrid model design. Lexicon-based models were used to classify opinions based on lexicon resources. As a result of the study, the authors showed the Vietnam hotel aspect summary reviewed by Vietnamese customers and American customers. There were differences in opinions between domestic customers and American customers because of the differences in cultures of each nation. In this study, the authors also introduced a hybrid model of lexicon-based and machine learning. The proposed model can be applied and used in English and non-English languages. This model can be modified with specific characteristics and used in our analytics reviews in the present study.

In Vietnamese, a process using a machine learning model was proposed to classify emotion into two polarities: positive and negative [25]. The hotel services and experimental data on Vietnamese reviews collected from booking.com (accessed on 21 April 2020) were used as a case study. The techniques used were overfitting training, 5-FoldCV, confusion matrix and ensemble methods. Because the author only trained the data set in the hotel services context, this data training set can only be effective in this case study and is still not used popularly in other contexts. Vietnamese sentiment classification on one public dataset was studied in research [26]. The authors built five ensemble schemes and proved feature importance methods with five techniques. The results showed that ensemble methods could perform better than the others. In another study, the authors used the Vietnamese lexicon to build a sentiment dictionary with scoring method. The domain in this work is electronic products from e-commerce websites. This dictionary can extract nouns, adjectives and adverbs, which can be highly applied in social analysis. In our research, we also expect to build a dictionary for standardizing acronyms and some punctuation in the context of Vietnamese.

In other languages such as Chinese in research [27], textual factors in online product reviews were identified through both qualitative and quantitative methods. The authors interviewed and surveyed 136 participants and used the product “Mobile phones” as a sample product in this study. This study also followed the decision-making process and emotion of the online product reviews following each step of this process. From this study, the researchers identified six measures related to influence on perception, including accuracy, comparison to other products, information on customer support, overall star rating, technical information and persuasive words. The study also defined the presence of rude and racist words, ambivalence/neutrality, spelling and grammar mistakes, brevity and other reviewers’ ratings and comments as less influential. The study considered “Textual factors in online product reviews: a foundation for a more influential approach
to opinion mining” conducted only in specific regions such as China and with a defined product, “Mobile phones”, so it still cannot be accepted or deemed helpful in other contexts, especially in younger e-commerce such as Vietnamese e-commerce. Moreover, other factors such as products, product details, sales and promotions in e-commerce websites would have some influence on consumer buying decisions with online reviews. Our study researched and presented this idea.

In the field of e-commerce, classification and sentiment analysis are often researched on the Amazon platform because this is the largest online supplier in the world with a wide range of products and a large number of product reviews. In Ref. [2], the author presented a model of extracting features from user opinion information and evaluated customer opinions about iPhone 5 reviews with three basic levels: good, average and bad. A continuous naive Bayes learning framework is used to classify and sentiment product reviews were present in research [3]. There are other research studies about mining and sentiment analysis, such as Refs. [28–30]. In these research studies, aspect extraction is essential and it can affect the result of sentiment classification.

3.2. Overview of Technologies Used in Supporting Sentiment Reviews

The term natural language processing, also known as NLP, is a collection of computational techniques for automatic analysis and representation of human languages, motivated by theory [31]. NPL plays an important role and can be applied in many areas such as machine translation, focusing on automatic translation converting from one language to another; information retrieval (IR), which is used to solve the problem of providing the most relevant information and information with the input a question of the user; and information extraction (IE), which is used to extract the information automatically, question answering, text summarization, topic modeling and, more recently, opinion mining [32]. As the primary task in NLP, tokenization is used to break up a string of words into semantically functional units called tokens and works by defining boundaries and criteria of where each token begins and ends [33]. Sentence tokenization can be used both to split sentences within a text and words within a sentence [31]. Removing stop words is an essential step in our text processing, which involves filtering out high-frequency words that add little or no semantic value to a sentence. In this research, the NLP techniques are used to remove stop words, tokenize and embed entities of a user’s statement.

As a subtopic of NLP in artificial intelligence, natural language understanding (NLU) is understood as machine reading comprehension and considered an AI-hard problem [31]. Most NLU systems share some standard components, which include a lexicon of the language and a parser, as well as grammar rules to break sentences into an internal representation [32], and the construction of a rich lexicon with suitable ontology [33]. Generally, NLU is a field of NLP that deals with the conversion of natural language into a semantic representation that the computer can interpret [32,33].

To effectively support and implement NLP and NLU, recurrent neural networks (RNNs) are often used in many systems and models [34,35]. The main idea of RNNs is to use a series of information to theoretically perform the same task for all sequence elements with the output dependent on previous calculations [36]. The neural network consists of three key sections: input layer, hidden layer and output layer [35].

Recurrent neural network (RNN) models into a flat form with $X_t$, $X_{t+1}$, $X_{t+2}$ as the input data at time $t$, $t + 1$ and $t + 2$ as follows (here, $t$, $t + 1$, $t + 2$, etc., are called the timesteps). RNN uses only a single neural network (usually a layer) to calculate the output value of each timestep. Therefore, outputs, upon becoming input, are multiplied by the same weight matrix (here is $WR$). Figure 1 shows a simple recurrent neural network [30].
Figure 1. A simple recurrent neural network.

Figure 2 presents an unfolding recurrent neural network process. First, $X_t$ is taken from the sequence of input, subsequently outputting $Y_t$, which, together with $X_{t+1}$, is the input for the next step. Thus, $Y_t$ and $X_t$ are the input for the next step. Similarly, $Y_{t+1}$ from the next step is the input with $X_{t+2}$ for the next step and so on. This way, it keeps remembering the context while training.

The formula for the current state is:

$$Y_t = f(Y_{t-1} - X_t)$$

Applying the activation function:

$$Y_t = \tanh (W_{hh} Y_{t-1} - 1 + W_{xh} X_t)$$

$W$ is weight, $Y$ is the single hidden vector, $W_{hh}$ is the weight in the previous hidden state, $W_{xh}$ is the weight in the current input state and $\tanh$ is the activation function implementing a non-linearity that squashes the activations to the range from $-1$ to $1$.

Output:

$$H_t = (W_{hy})(Y_t)$$

The unfolding process of the recurrent neural network can bring two main advantages for using this network and building our machine learning model to classify polarity words in each product review in this paper [37]. First, regardless of the sequence length, the learned model always has the same input size because it is specified in terms of the transition from one state to another, instead of being specified in terms of the variable-length history of states. Second, it is possible to use the same transition function $f$ with
the same parameters at every time step. These two factors make our training model able to learn a single model $f$ that operates on all-time steps and all sequence lengths instead of needing to learn a separate model for all possible time steps. Learning a single shared model allows generalization to sequence lengths that did not appear in the training set and enables the model to be estimated with far fewer training examples than required without parameter sharing.

In this paper, a training model is built, in which the sentiment analysis problem is solved by the recurrent neural network model (RNN), which has a commonly used variant currently called the long short-term memory neural network (LSTM), which combines the vector representation of words with a continuous bag of words (CBOW) architecture.

4. Design and Implementation

4.1. Proposal of Research Model

In this research, a sentiment evaluation model of customer reviews is proposed with functions extracting objects, and emotional words for the problem of emotional level analysis, using machine learning algorithms. We combined emotional dictionaries and a set of extract rules to build a data training data set based on the semantic dependency relationship between words in sentences in the given Vietnamese context. In Figure 3, the research framework is presented. The process starts with crawling review data automatically. Then, crawled data will be sent to data processing, with functions such as standardization, removing stop words, etc., and extract into features. There are five main groups of features that directly affect customer reviews that were found in this research. After the training program is run, we can input a new review to the model. This review can be classified into sentiment reviews. If a new feature is defined, it will be updated and added to the feature library, as shown in Figure 3. With the input of a new product review statement, the model can analyze the sentiment of the review and visualize analysis results on the real-time website. A research framework is proposed as shown in Figure 3 and the details of each primary process in this research framework are explained in the following parts.

![Figure 3. Research Framework.](image-url)
4.2. Crawling Reviews Data

We used Python combined with Splash-Scrpy to crawl data of sales on e-commerce websites. Splash is a lightweight headless browser that works as an HTTP API. With Splash, we can render JavaScript pages, follow URLs and extract data from specific HTML paths on each page. Our collected data were saved in MongoDB of AWS. After collecting data, we cleaned and filtered data in Python. Key conditions expression was used every time crawling data to filter data.

Next, after reviewing the data source of Tiki, the Hot Deal data of the website Tiki was chosen. As a result, there were 56,484 raw reviews from 16 product categories in Hot Deal of the website Tiki as in Table 1.

Table 1. Data collected from Hot Deal of Tiki.

| No. | Category                          | Number of Reviews |
|-----|-----------------------------------|-------------------|
| 1   | Fashion                           | 10,378            |
| 2   | Toy—Mother & Baby                 | 5921              |
| 3   | Beauty—Health                     | 5513              |
| 4   | House—Life                        | 5391              |
| 5   | Tiki Bookstore                    | 5221              |
| 6   | Electrical Appliance              | 5061              |
| 7   | Grocery Online                    | 5023              |
| 8   | Laptop—Computers                  | 4037              |
| 9   | International Commodity           | 3318              |
| 10  | Digital Device—Digital Accessory  | 2840              |
| 11  | Phone—Tablet                      | 1317              |
| 12  | Sport—Picnic                      | 887               |
| 13  | Camera—Camcorder                  | 618               |
| 14  | Electronics—Refrigeration         | 547               |
| 15  | Voucher—Service                   | 241               |
| 16  | Cars—Motorcycles—Bicycles         | 171               |
|     | Total                             | 56,484            |

4.3. Data Preprocessing and Feature Defining

4.3.1. Data Preprocessing

After crawling reviews from Tiki’s Hot Deal, the process proceeded with the data standardizing phase with simple commands in Python based on four missions:

- convert all words into lowercase;
- remove numbers, punctuation, extra spaces, special characters, duplicate characters;
- standardize fonts, eliminate meaningless words, correct some abbreviations, spelling errors and spaces;
- eliminate extra characters (duplicate characters); standardize Vietnamese accents (example: “u’, “ý”, etc.).

Then, we used two libraries of Python, ViTokenizer and ViPostagger, to define tagged files containing word-processing paragraphs (word segments) in Vietnamese and the library Gensim for unsupervised topic modeling and natural language processing using modern statistical machine learning. In this same step, we used TF-IDF to determine the importance of the words to the sentences in the corpus [38]. TF-IDF works by determining the relative frequency of words in a specific document compared to the inverse proportion of such words over the entire document corpus [38]. The TF-IDF weight is composed of two terms: the first computes the normalized term frequency (TF), or the number of times a word appears in a document, divided by the total number of words in that document; the second term is the inverse document frequency (IDF), measures how important a term is [39]. The TF-IDF weight is computed as the logarithm of the number of documents in the corpus divided by the number of documents where the specific term appears [39–41].
In our research, we used TF-IDF for stop word filtering in various subject fields, including text summarization and classification. Figure 4 presents the results of the stop word and punctuation process.

![Figure 4. Building stop word and punctuation.](image)

Figure 5 shows the command built-in Python to remove all the stop words in the data set.

![Figure 5. Command Remove stop word.](image)

In addition, after processing data, we found that customers usually used many acronyms, misspelt words, spoken language, etc., in product reviews out of habit. We also built a separate dictionary for this particular word-based.

4.3.2. Feature Defining

In this step, we conducted a review of the data set and extracted essential features of the product reviews. Then, we also conducted a frequency analysis and extracted terms in customer reviews related to process features and added them to the final feature word list as shown in Table 2.

| Feature Word Vietnamese | Feature Word Translate to English | Opinion Word (Defined by Adjective of Sentence) Vietnamese | Opinion Word (Defined by Adjective of Sentence) Translate to English |
|-------------------------|----------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Giá                     | Price                            | Mắc, rẽ, cao, thấp, không mắc, không rẽ, không cao, không thấp | Expensive, cheap, high, low, not expensive, not cheap, not high, not low |
| Chất lượng               | Quality                          | Chất lượng, bền, nhiều chức năng, ổn định, tốt | Quality, durability, function, performance, worse, not good |
| Văn chuyển              | Shipment                         | Nhanh, chậm, không nhanh, cân thận, nhanh, thời gian dài, lâu | Long, slow, not fast, careful, fast |
| Thiết kế                 | Design                           | Đẹp, tệ, sang trong, phù hợp, trẻ trung, không tệ, không đẹp | Nice, Bad, Awful, luxury, suitable, young, not bad, not good |
| hài lòng                 | Satisfy                          | Cảm thấy, thất vọng, không thất vọng, hài lòng, hạnh phúc, thích, không thích | Feeling, disappointed, not disappointed, satisfy, happy, like, dislike |

After overviewing the opinion feature word dictionary, as in Table 2, the following process was feature extraction and calculation of the weight of each feature in reviews data by the topic modeling process presented in Figure 6.
First, collected product reviews go to standardization sentences and spoken tagger as subjects, objects, verbs and adjectives by comparing with the sentence tagger dictionary. Second, sentences are represented as vectors with vectorization steps. The system defines the entity and feature word by machine learning algorithm with the topic data-training set model. After this step, the process can define the adjectives and verbs corresponding to each feature’s attribute. Finally, the process calculates the weight of features with each entity and summary feature and opinion word in reviews. The main purpose of the topic modeling process is to explore feature words, as well as the main opinion of review sentences of customers in the data set.

Text files are series of words (ordered). We needed to convert the text files into numerical feature vectors to run the machine learning algorithm. We used a bag of words model. Briefly, we segmented each text file into words (for English splitting by space), counted the # of times each word occurs in each document and, finally, assigned each word an integer ID. Each unique word in our dictionary corresponds to a word of the table opinion feature words. The representation of text documents is a challenging task in machine learning. Nowadays, many techniques are used, such as doc2vec method, a concept presented in [40]. In this research, we used the Tfidfvectorizer method.

4.4. Machine Learning Algorithms

The workflow of the building machine learning model is presented in Figure 7. The process starts with the creating reviews sentiment. After that, the process extracts feature from text files into labels. We continued with defining the model, which is a training model. After drawing up a training model, we evaluated this model with a comparison index of the model (accuracy, val_loss, loss). If all indices are suitable, the process ends. If not, the process notifies the situation, reviews the dictionary and adds more or changes the model if needed.
Phase 1: Extracting features from text files

After this process, we store collected review data as the value upon processing into the corresponding characteristic columns. Then, we calculated statistics on the number of comments for each feature, pivoted the data and saved it to the database. We prepared statistics on comments by topic. Because the review details are the objects, they need to be processed inside the loop of objects. We defined the evaluation functions according to each word feature. Then, we exported them into a file used as a training data model.

In the beginning, all collected product review data are stored in MongoDB with the following format of key-value. All the keys are ID, product event ID, rating, reviews count, rating count, review detail (review author, review date post, review title, content and useful), link of review and time updated of this review. Each key contains the corresponding.

The statistics on reviews by category product are prepared with df_overview_topic functions. At this step, all the reviews are counted with their product categories in two main columns: name of category and frequency. There are 14 product categories found in this step, as shown in Figure 8.
Next, data will be moved to process extracts all features from text files into labels. We define the evaluation functions according to the Feature dictionary and continue to polarize this feature into 3 levels: 1 is negative, 2 is neutral and 3 is positive. We analyze product reviews with five features of price, qualify, shipment, design and customer satisfy. The reviews continue to polarize into three levels: 1 is negative, 2 is neutral and 3 is positive. We use command conditions to extract the levels of these features in customer reviews easily. After that, the implemented value is stored in the database. We exported all the implemented values into a file used as a training data model. and after those data are saved in CSV files.

**Phase 2: Building the Data model**

We use the trained model to build the classify function in order to predict the future comment in real-time. In our training model, we separated our data into two parts: 80% data for training and 20% remaining of the data for testing. The training data contains 30,000 records and the test data contains 8,778 records. Training data are processed, such as standardization and POS tagger, features extraction. Each sentence will be divided into basic tokens, punctuations and spaces before training. The input text of the model can be one sentence or a short paragraph of 3 to 5 sentences in the same reviews. Vietnamese is an isolated language, so that word boundaries are not automatically defined by spaces. Determining the semantics of words is greatly influenced by word segmentation so that the model achieves higher accuracy. Therefore, there is also a dictionary for phrases that often appear but do not carry much meaning as well as words that will affect the operation of the model. The continuous bag-of-words (CBOW) model is used to represent the word set into vector space at this stage. Thanks to the vector space of this vocabulary, the more similar the words are, the larger the cosine distance value (closer to 1); conversely, the less similar words are smaller (closer to 0). As a result of the training process, we obtained the set of weights of the neural network LSTMs stored in the file (params), along with the LSTMs network configuration hyperparameters (conf). These two files will be loaded into the network of LSTMs for testing and operation. About sampling procedure was applied to test the model; we chose a random selection to test in order to ensure the accuracy and fairness of the model evaluation.

**4.5. Building of Actual Application Model**

Figure 9 shows the entire UML framework of our system. The user can access MainActivity, which has a login interface. MainActivity connects with the server through the Internet network.
5. Result and Discussion

5.1. Topic Modeling

The results of the topic modeling process show that there are 42 distributions of words structured into five main features (quality, price, design, shipment and satisfaction), as shown in Figure 10. They are manually named quality, price, design, shipment and satisfaction. For each topic, the most frequent words (vietnamese_word and english_word) are listed together with their frequency and probabilities (Prob.).
We designed Word Cloud charts as in Figure 11 to visualize all the topic distributions for users to follow and understand the overview of Product Review Data easily. As shown in Figure 11, the ten most commonly used words in customer’s reviews is “Hàng” (Product), “Không” (No), “Sản phẩm” (Product), “Giao” (Ship), “Tốt” (Good), “Được” (Get), “Nhanh” (Fast), “Đẹp” (Beautiful), “Nhưng” (But) and “Mới” (New).

5.2. Data Training Set

As a result, this research built a data training set that can classify the sentiment of each Vietnamese review with 38,778 different rows and reviews with five features (price, shipment, quality, design, satisfy) and three levels (1—negative, 2—neutral, 3—positive) of sentiment, as shown in Table 3.

In our model, that val_loss starts to decline and val_acc starts to increase; this means that the model is learning and working well. Only the loss function was used to update the model’s parameters and the accuracy was only used to see how well the model worked.

When inputting a new review, this model can classify the sentiment. For example, the review was “Fast shipping, good product quality, prices are quite cheap, so I feel great” (”mình mua hàng giao nhanh_chất lượng sản phẩm đẹp_gói cách chẩn giá khá_rẻ hàng đảm bảo nên cùng khác an_tâm”). The model can correctly classify the sentiment of this review with status (33222), Price_Positive, Shipment_Positive, Quality_Neutral, Design_Neutral and Satisfy_Neutral.
Figure 11. Word cloud for customer’s reviews.
### Table 3. Data training set.

| ID  | Vietnamese Review                                                                 | English Review                                                                 | Sentiment Results                                                                 |
|-----|----------------------------------------------------------------------------------|--------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| train_0 | kẹy_vong vao olay vi day cuc xíng và hệ thống thắt_vong                           | expectation on olay because this is also a reputable company, but using it is disappointing. | (Neutral price, neutral shipment, neutral quality, neutral design, negative satisfy) |
| train_1 | mình là dân nghiêm mà bài viết có nội dung giới thiệu một số với chất lượng     | I am mask addict because of it's convenient, the price with promotion is so great compared to the quality | (Positive price, neutral shipment, positive quality, neutral design, neutral satisfy) |
| train_2 | đập vào khả năng và không có hiệu quả nhiều                                     | applying is quite sunburn and ineffective.                                                                 | (Neutral price, neutral shipment, negative quality, neutral design, neutral satisfy) |
| train_3 | da mình thích loại khỏi và khả nhạy cảm để tim loai phù hợp rất khó khăn đã thử nhiều loại mặt na nhưng chưa tìm được loại nào phù hợp. Đợt này dùng olay thì kết quả rất rất thích hợp. | my skin is dry and quite sensitive to find the right type very difficult, tried many types of masks but haven’t found any suitable. This time using olay, the results are very good. | (Neutral price, neutral shipment, neutral quality, neutral design, positive satisfy) |
| train_4 | diện trong là giao hàng lâu vi order hàng từ nức mediocre                           | minus point is long delivery because order from abroad                                                                            | (Neutral price, neutral shipment, neutral quality, neutral design, neutral satisfy) |
| train_5 | trong những mình có rất nhiều sản phẩm và hệ thống thắt sergeant                   | during the month, I ordered olay mask to give my wife a promotional product, but after using it, my wife really liked this product. I also just received a discount code from tiki for the last purchase of products. | (Neutral price, neutral shipment, neutral quality, neutral design, positive satisfy) |
| train_6 |                         | in general, it is quite pleasant to use because the paper of the mask is quite thick and absorbs the nutrients she puts after washing the face, the skin is smooth and seems a bit brighter | (Neutral price, neutral shipment, positive quality, neutral design, positive satisfy) |
| train_7 |                         | the stretchy mask can cover the entire face after being applied on the face with a white layer of cotton like cotton coming out from the mask. I don’t know if any of you are like this | (Neutral price, neutral shipment, neutral quality, neutral design, neutral satisfy) |
| train_8 |                         | bought at the promotion price, bought many times, packaged goods with plastic wrap outside to avoid getting wet, but the box is still a paper carton, so the wide box is heavy and dented. | (Neutral price, neutral shipment, negative quality, neutral design, neutral satisfy) |
| train_9 |                         | I really like the mask. It is quite thick and the thick mask translation is very easy to use. When applying the mask, I feel comfortable, after applying it, my skin is bright white and very smooth. | (Neutral price, neutral shipment, positive quality, neutral design, positive satisfy) |
| train_10 |                         |                                                                 | (Neutral price, neutral shipment, neutral quality, neutral design, neutral satisfy) |

#### 5.3. Visualization of the Results

Some charts and tables are presented for readers or users can get more information and insight from our research. We also designed a web-app to show the visualization of all the results.

As shown in Figure 12, the product with the highest topic review ratio is fashion (18, 37%), followed by toys—moms and children (10, 48%) and beauty—health (9, 76%). The main customers who put their time into feedback were female customers, so the company should focus on designing and providing products that are more suitable and attractive to female customers. The research results also show the ratio of feedback and polarity of each feedback.
Figure 12. Summarizations of topic review ratio chart.

Table 4 presents an example of a visualization screen of total reviews of each product category in our web-app. This example shows the total number of reviews of each product category. The top three product categories with the most reviews are fashion (10,378 reviews), toy—mom and children (5921 reviews) and beauty—health (5513 reviews).

Table 4. Total reviews of each product category.

| Product Category                      | Customer Reviews |
|---------------------------------------|------------------|
| Fashion                               | 10,378           |
| Toy—Mom and Children                  | 5921             |
| Beauty—Health                         | 5513             |
| Home—Lifestyle                        | 5391             |
| Tiki Bookstores                       | 5221             |
| Electric Appliances                   | 5061             |
| Grocery Online                        | 5023             |
| Laptop—Desktop—Accessories            | 4037             |
| International Product                 | 3318             |
| Digital Devices—Digital Accessories   | 2840             |
| Phone—Tablet                          | 1317             |
| Sport—Picnic                          | 887              |
| Camera—Video Recorder                 | 618              |
| Electronics—Refrigeration             | 547              |

Figure 13 presents the sentiment results of product reviews by product categories according to five features: customer satisfaction, good design, good quality, good shipping and low price.
6. Conclusions

This research aims to build a machine learning model to classify Vietnamese product reviews based on the five main features of the online buying process aspect in an e-commerce website. Specifically, we crawled 56,484 customer reviews from 16 kinds of products. We collected details and descriptions of products and screened the product reviews to find the main feature words. As a result, we found five main features that appear and affect almost all customer reviews include price, shipment, quality, design and satisfaction. By extracting feature words, we built a solitary word dictionary with three levels (negative, neutral and positive) in Vietnamese. After that, we built a machine learning model using an RNN model with a data training set having 38,778 data records. When input a customer review or a file of customer reviews, our model could portray the reviews with five features and three levels. Finally, there is a web-app to connect to a server and visualize all the research results. Based on the research results, enterprises can follow up their customers in real-time and receive recommendations to understand their customers.

In this research, we only used a simple model and polarity reviews with five defined features and three levels; we have still not considered the relationship and ratio of each feature with the products. The testing method of research model and evaluation of the sampling procedure have not yet been established. This current research is necessary to evaluate the utility and cross-validation metrics with different training and test sets. Building cross-validation and evaluation will be researched more in future work. In addition, future work on more different e-commerce websites shall be conducted. In addition, the potential of building a multi-platform product reviews polarity model with a variety of products is essential, as well as deep learning models and natural language processing techniques will be researched in the future to solve these problems. We also plan to build an android app based on cloud computing to enable users follow up realistic sale processing and customer feedback in real-time. This research will be a suitable reference and highly applicable not only in the field of e-commerce. In addition, this methodology can be used to train and sentiment reviews in other specific domains such as education, government, baking, service, healthcare or scientific domain.
Author Contributions: Conceptualization, N.-B.-V.L.; data curation, N.-B.-V.L. and J.-H.H.; formal analysis N.-B.-V.L. and J.-H.H.; funding acquisition, J.-H.H.; investigation, N.-B.-V.L.; project administration, J.-H.H.; resources, N.-B.-V.L. and J.-H.H.; software, N.-B.-V.L. and J.-H.H.; validation, N.-B.-V.L.; visualization, N.-B.-V.L.; writing—original draft, N.-B.-V.L. and J.-H.H.; writing—review and editing, J.-H.H. All authors will be informed about each step of manuscript processing, including submission, revision, revision reminder, etc., via emails from our system or assigned Assistant Editor. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are openly available in researchgate.net at doi:10.13140/RG.2.2.24115.96800.

Computer Code and Software: Some scripts are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sharma, R.; Nigam, S.; Jain, R. Mining of product reviews at aspect level. Int. J. Found. Comput. Sci. Technol. (IJFCST) 2014, 4, 87–95. [CrossRef]
2. Bhatt, A.; Patel, A.; Chheda, H.; Gawande, K. Amazon review classification and sentiment analysis. Int. J. Comput. Sci. Inf. Technol. 2015, 6, 5107–5110.
3. Xu, F.; Pan, Z.; Xia, R. E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework. Inf. Process. Manag. 2020, 57, 10221. [CrossRef]
4. Fang, X.; Zhan, J. Sentiment analysis using product review data. J. Big Data 2015, 2, 1–14. [CrossRef]
5. Vietnam E-commerce Association (VECOM). Vietnam E-Business Index 2019 Report; Vietnam E-commerce Association: Hanoi, Vietnam, 2019.
6. Department Foreign Investment. Report of Foreign Direct Investment in 2019; Ministry of Planning and Investment of Vietnam: Hanoi, Vietnam, 2019.
7. Galhotra, B.; Dewan, A. Impact of COVID-19 on digital platforms and change in E-commerce shopping trends. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020; pp. 861–866. [CrossRef]
8. Abdelrhim, M.; Elsayed, A. The Effect of COVID-19 Spread on the E-Commerce Market: The Case of the 5 Largest E-Commerce Companies in the World. 2020. Available online: https://ssrn.com/abstract=3621166 (accessed on 21 September 2021). [CrossRef]
9. Hasanat, M.W.; Hoque, A.; Shikha, F.A.; Anwar, M.; Hamid, A.B.A.; Tat, H.H. The impact of coronavirus (COVID-19) on e-business in Malaysia. Asian J. Multidiscip. Stud. 2020, 3, 85–90.
10. Gu, S.; ´Slusarczyk, B.; Hajizada, S.; Kovalyova, I.; Sakhbieva, A. Impact of the COVID-19 Pandemic on Online Consumer Purchasing Behavior. J. Theor. Appl. Electron. Commer. Res. 2021, 16, 125. [CrossRef]
11. Pandey, A.; Parmar, J. Factors Affecting Consumer’s Online Shopping Buying Behavior. In Proceedings of the 10th International Conference on Digital Strategies for Organizational Success, Gwalior, India, 5–7 January 2019.
12. Koch, J.; Frommeyer, B.; Schewe, G. Online Shopping Motives during the COVID-19 Pandemic—Lessons from the Crisis. Sustainability 2020, 12, 247. [CrossRef]
13. Neger, M.; Uddin, B. Factors affecting consumers’ internet shopping behavior during the COVID-19 pandemic: Evidence from Bangladesh. Chin. Bus. Rev. 2020, 19, 91–104.
14. Lien, N.T.B.; Trang, N.T.X. Factors affecting the online shopping intention of customers in Ho Chi Minh City during the Covid-19 pandemic. Ind. Trade Mag. 2021, 19, 1–9.
15. Litsa, T. How a Customer Reviews Strategy Can Impact SEO. Search Engine Watch, 6 September 2016. Available online: https://www.searchenginewatch.com/2016/09/06/how-a-customer-reviews-strategy-can-impact-seo/ (accessed on 14 June 2020).
16. Goldsmith, R.E.; Horowitz, D. Measuring Motivations for Online Opinion Seeking. J. Interact. Advert. 2006, 6, 2–14. [CrossRef]
17. Ngai, E.; Xiu, L.; Chau, D. Application of data mining techniques in customer relationship management: A literature review and classification. Expert Syst. Appl. 2009, 36, 2592–2602. [CrossRef]
18. Zeyer, A.; Doetsch, P.; Voigtlaender, P.; Schluter, R.; Ney, H. A Comprehensive Study of Deep Bidirectional LSTM RNNS for Acoustic Modeling in Speech Recognition. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, New Orleans, LA, USA, 5–9 March 2017.
19. Le, H.; Lee, J.; Lee, H.-K. Purchase Process Aspect-based Opinion Mining—An Application for Online Shopping Mall. J. Internet Electron. Commer. Res. 2015, 15, 15–28.
20. Vo, T.H.; Nguyen, T.T.; Pham, H.A.; Le, T.V. An Efficient Hybrid Model for Vietnamese Sentiment Analysis; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 227–237.
21. Le, H.S.; Le, T.V.; Pham, T.V. Aspect Analysis for Opinion Mining of Vietnamese Text. In Proceedings of the 2015 International Conference on Advanced Computing and Applications (ACOMP), Ho Chi Minh City, Vietnam, 23–25 November 2015.
22. Nguyen-Thanh, T.; Tran, G.T. Vietnamese Sentiment Analysis for Hotel Reviews based on Overfitting Training and Ensemble Learning. In Proceedings of the SoICT 2019: The 10th International Symposium on Information and Communication Technology, Hanoi-Halong Bay, Vietnam, 4–6 December 2019.

23. EU-Vietnam Business Network. *E-Commerce Industry in Vietnam—Edition 2018*; EU-Vietnam Business Network: Ho Chi Minh, Vietnam, 2018.

24. Le, H.S.; Lee, J.H.; Lee, H.K. *Hotel Services Preferences Across Cultures: A Case Study of Applying Opinion Mining on Vietnamese and American Online Reviews*; Korean Society for Management Information System: Seoul, Korea, 2016; pp. 137–149.

25. Nguyen, C.V.; Le, K.H.; Nguyen, B.T. A Novel Approach for Enhancing Vietnamese Sentiment Classification. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Kuala Lumpur, Malaysia, 26–29 July 2021; Springer: Cham, Switzerland, 2021; pp. 99–111.

26. Van Thin, D.; Nguyen, N.L.T.; Truong, T.M.; Le, L.S.; Vo, D.T. Two New Large Corpora for Vietnamese Aspect-based Sentiment Analysis at Sentence Level. *Trans. Asian Low-Resour. Lang. Inf. Process.* 2021, 20, 1–22. [CrossRef]

27. Robinson, R.; Goh, T.T.; Zhang, R. Textual factors in online product reviews: A foundation for a more influential approach to opinion mining. *Electron. Commer. Res.* 2012, 12, 301–330. [CrossRef]

28. Liu, B. *Sentiment Analysis and Opinion Mining.* *Synth. Lect. Hum. Lang. Technol.* 2012, 5, 1–167. [CrossRef]

29. Abulaish, M.; Jahiruddin; Doja, M.N.; Ahmad, T. Feature and Opinion Mining for Customer Review Summarization. In *Pattern Recognition and Machine Intelligence*; Chaudhury, S., Mitra, S., Murthy, C.A., Sastry, P.S., Pal, S.K., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5909.

30. Oggunmolu, O.; Gu, X.; Jiang, S.; Gans, N. Nonlinear Systems Identification Using Deep Dynamic Neural Networks. *arXiv* 2016, arXiv:1610.01439.

31. Chowdhary, K.R. Natural Language Processing. In *Fundamentals of Artificial Intelligence*; Springer: New Delhi, India, 2020; p. 645.

32. Sheikh, T.H. Text Mining and its Applications. *IAPRR Int. Peer Rev. Refereed J.* 2017, 4, 1–8.

33. Blokdyk, G. Tokenization Strategy and Managing Projects, Criteria for Project Managers. In *Tokenization Strategy, A Complete Guide—2020 Edition*; The Art of Service: Brendale, Australia, 2020; p. 127.

34. Jerome, T.; Connor, R.; Martin, D. Recurrent Neural Networks. *IEEE Trans. Neural Netw.* 1995, 5, 240–254.

35. Dissen, Y.; Goldberger, J.; Keshet, J. Formant estimation and tracking: A deep learning approach. *J. Acoust. Soc. Am.* 2019, 145, 642. [CrossRef]

36. Yin, C.; Zhu, Y.; Fei, J.; He, X. A Deep Learning Approach for Intrusion Detection using Recurrent Neural Networks. *IEEE Access* 2017, 5, 2069–3356. [CrossRef]

37. Tsoi, A.C. Recurrent Neural Network Architectures: An Overview. In *Adaptive Processing of Sequences and Data Structures*; Lecture Notes in Computer Science; Giles, C.L., Gori, M., Eds.; Springer: Berlin/Heidelberg, Germany, 1998; Volume 1387.

38. Ramos, J. Using TF-IDF to Determine Word Relevance in Document Queries. *arXiv* 2013, arXiv:2001.09896.

39. Wu, H.C.; Luk, R.W.P.; Wong, K.F.; Kwok, K.L. Interpreting TF-IDF term weights as making relevance decisions. *ACM Trans. Inf. Syst.* 2008, 26, 13. [CrossRef]

40. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* 2013, arXiv:1301.3781.

41. Thu, H.N.T.; Ngoc, V.H.; Nhat, A.N.; Huynh, H.X. Developing Vietnamese Sentiment Lexicon from Social Reviews Corpus Based on Support Measurement. In Proceedings of the International Conference on Intelligent Systems & Networks, Hanoi, Vietnam, 19 March 2021; pp. 52–58.