MACHINE LEARNING BASED REVIEW ANALYSIS OF ELECTRONIC APPLIANCES

Tuan Hoang Vu[*], Minh Tuan Nguyen[+]
[*]ThuyLoi University
[+]Posts and Telecommunications Institute of Technology

Abstract – Sentiment Analysis and Opinion Mining have emerged as highly popular fields for analyzing and extracting valuable information from textual data sourced from diverse platforms like Facebook, Twitter, and Amazon. These techniques hold a crucial role in empowering businesses to actively enhance their strategies by gaining comprehensive insights into customers' feedback regarding their products. The process involves leveraging computational methods to study individuals' buying behavior and subsequently mining their opinions about a company's business entity, which could manifest as an event, individual, blog post, or product experience. This paper focuses on utilizing a dataset obtained from Amazon, comprising reviews spanning various product categories such as laptops, cameras and mobile phones. Following data preprocessing, we employ machine learning algorithms to classify the reviews as either positive or negative sentiment. This classification step enables us to analyze the overall sentiment associated with the products and draw meaningful conclusions.

Keywords—Customer requirement, electronic appliances, machine learning, natural language processing, sentiment analysis.

I. INTRODUCTION

With numerous brands flooding the market, consumers face the challenging task of choosing the right one. The rise of e-commerce has significantly influenced consumer purchasing habits, and they heavily rely on reviews available on e-commerce platforms, including ratings and relevant text summaries, to make informed decisions [1]. In addition to e-commerce platforms, product reviews can also be found on social networking sites [2]. Social networks have experienced immense popularity in recent years, leading to a potential exponential growth in data volume in the future [3, 4]. The continuous influx of user comments has resulted in a vast amount of online data, making it challenging to extract relevant information accurately [5].

Sentiment analysis plays a crucial role in providing valuable insights to both customers and manufacturers by analyzing positive and negative sentiments associated with each product. It is a fundamental task in Natural Language Processing (NLP) [6, 7]. Sentiment or opinion refers to the perspective of customers derived from various sources such as reviews, survey responses, social media, healthcare media, and more [8]. The objective of sentiment analysis is to determine the attitude of a speaker, writer, or subject towards a specific topic or contextual polarity in events, discussions, forums, interactions, or documents. The analysis can be conducted at different levels, including document-level, sentence-level, and aspect-level [9].

At the document-level, sentiment analysis categorizes the entire document as expressing a positive or negative view, making it suitable for analyzing a single product review to determine the opinion about that specific product. However, it may not be applicable when a document contains multiple product reviews as it does not consider different types of reviews. At the sentence-level, individual sentences are analyzed to determine whether they convey a positive, negative, or neutral opinion, like Subjectivity Classification that differentiates between objective and subjective sentences. The aspect-level sentiment analysis, also known as feature-level sentiment analysis, focuses on identifying specific aspects that people liked or disliked, providing a more detailed analysis of sentiment. It directly focuses on the opinions themselves and includes information such as the entity, the specific aspect of that entity, the opinion regarding the aspect, the opinion holder, and the timeframe.

With the widespread use of the internet, sentiment analysis becomes crucial in understanding and extracting insights from the vast amount of opinionated data available online. It is widely applied in analyzing product reviews to understand customer sentiments. By leveraging machine learning (ML) techniques, sentiment analysis helps businesses gather customer insights from various online platforms, including social media, surveys, and e-commerce website reviews. Furthermore, the popularity of smartphones has led to a significant increase in individuals connecting to social networking platforms like Facebook, Twitter, and Instagram. These platforms have become spaces where people freely express their beliefs, opinions,

Contact author: Minh Tuan Nguyen
Email: nmtuan@ptit.edu.vn
Manuscript received: 7/2023, revised: 8/2023, accepted: 9/2023.

emotions, thoughts, experiences, and more, providing additional valuable data for sentiment analysis to understand user sentiments and behaviors.

Most sentiment analysis methods rely on supervised ML. ML approach tends to outperform the computational linguistic approach in terms of performance. Several studies have utilized machine learning and artificial intelligence techniques to conduct sentiment analysis on tweets [10]. In a study [11], various models such as Naive Bayes, support vector machine (SVM), and information entropy-based [12] models were employed to classify product reviews. Another research [13] introduced a hybrid machine learning algorithm based on Twitter opinion mining. Heydari et al. [14] put forth a time series model for analyzing fraudulent sentiment reviewers. Hajek et al. [15] developed a deep feedforward neural network and convolution model to detect fake positive and negative reviews within an Amazon dataset. Long et al. [16] utilized LSTM with a multi-head attention network to predict sentiment-based text using a dataset from Chinese social media. Dong et al. [17] proposed a supervised machine linear regression approach to predict customer sentiment in online shopping data using sentiment analysis learning methods.

Certain conventional approaches, which utilize machine learning techniques, focus on specific aspects of the language used. Pang et al. conducted a study on movie reviews and evaluated the performance of various machine learning algorithms, including Naive Bayes, maximum entropy, and SVM [18]. They achieved an accuracy of 82.9% by employing SVM with unigrams. In the field of NLP, feature extraction for sentiment classification is typically done using NLP techniques. Many NLP strategies primarily rely on N-grams, although the bag-of-words approach is also commonly used [19]. Several studies have shown promising outcomes when employing the bag-of-words technique as a text representation for item categorization [20].

A hybrid approach [21] is employed in this study, which combines both Machine Learning and Lexicon-based methods to enhance the performance and convenience of sentiment classification. The combination of Lexicon-Based and Learning-Based techniques is explored to achieve improved results. Various techniques and tools are discussed in this paper, addressing different aspects of sentiment classification. The purpose of this study is to design an effective and simple algorithm for ML-based sentiment analysis of the electronic products on the E-commerce exchange namely Amazon. The main contributions of our research are as follows:

- The utility of lexicon-based sentiment score, which effectively generate the initial labels for the product reviews of the database.
- Sentiment is improved for the individual words due to combination of the product reviews into a dataframe.
- The use of ML algorithms, which are less complexity but remaining relatively high recognition performance.

The remaining sections of the paper are structured as follows. Section II introduces the data and preprocessing techniques employed. Section III presents the methodology adopted in this study. The simulation and discussion of the method are presented in Section IV. Finally, Section V provides a summary of the research findings.

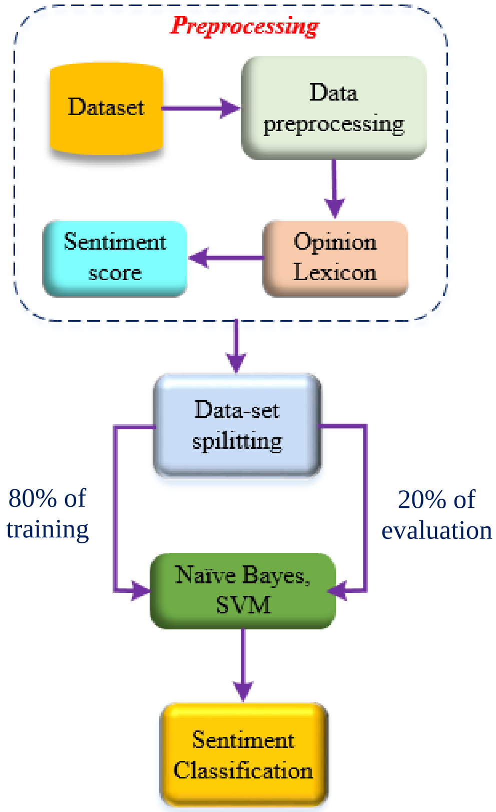

Figure 1. Workflow of the proposed methodology

II. DATA AND PREPROCESSING

A. Dataset

The dataset, collected from Amazon, is in JSON format. Each JSON file comprises a collection of reviews. The dataset includes reviews for various products such as Laptops, Camera and Mobile phones. Amazon is a prominent E-commerce platform with an extensive collection of reviews. In our research, we leveraged the Amazon product data, generously shared in reference [22]. The dataset is structures as follow:

"reviewerID": ID of the reviewer

"asin": ID of the product

"reviewerName": name of the reviewer

"helpful": helpfulness rating of the review

"reviewText": text of the review

"overall": rating of the product

"summary": summary of the review

"unixReviewTime": time of the review (unix time)
"reviewTime": time of the review (raw)

Table 1: The number of reviews for different categories

Categories	Number of Reviews
Laptops	1940
Cameras	3106
Mobile phones	1902

B. Data preprocessing

Preprocessing plays a crucial role in sentiment analysis and opinion mining, involving various steps such as tokenization, stop word removal, stemming, and punctuation mark removal, etc. These steps are performed to transform the text into a bag-of-words representation, which is commonly used in sentiment analysis. Preprocessing ensures that the text data is cleaned and organized in a way that facilitates accurate analysis of sentiment and opinions.

We applied various preprocessing techniques to clean the review texts for ease of processing. As a result, the total of review is 6948 including 1940, 3106, and 1902 review texts of Laptops, Cameras, and Mobile phones, respectively. The following methods is implemented on the entire dataset.

(1) *Lowercasing*: All words in the review text were converted to lowercase.

(2) *Link Removal*: Hyperlinks or URLs are removed.

(3) *Stopword Removal:* Commonly used words in the language, such as "the," "a," "an," "is," and "are," which do not carry significant information for the model, were removed from the review content.

(4) *Punctuation Removal:* All punctuation marks in the review texts were eliminated.

(5) *Elimination of One-Word Reviews:* Reviews containing only one word were discarded.

(6) *Contraction Removal*: Words originally written in a shortened form were replaced with their respective full forms. For example, "I'm" was changed to "I am".

(7) *Tokenization*: Each sentence in the review texts was divided into smaller units or tokens, typically words. Tokenization is the process of breaking down a sequence of strings, which can include words, keywords, phrases, symbols, and other components, into individual units referred to as tokens. These tokens can take the form of single words, short phrases, or even entire sentences. These resulting tokens are then used as input for various processes, including parsing and text mining.

(8) *Part-of-Speech Tagging*: Each word in the sentence was tagged with a part-of-speech (POS) tag, such as "V" for a verb, "ADJ" for an adjective, and "N" for a noun.

(9) *Score Generation:* The sentiment of the review text was evaluated, and a score was generated. This was done by matching the dataset with an opinion lexicon [22], which contains positive and negative words along with their respective scores. The sentiment score for each review text was calculated based on the lexicon scores. If the score was greater than 0, the review text was labeled as positive; otherwise, it was labeled as negative.

(10) *Word Embeddings:* Numerical vectors were computed for each preprocessed sentence in the product review dataset using the "Word embeddings" method. To create word indices, all review text terms were converted into sequences. Subsequently, a unique index was generated for each word in the training and testing sets.

III. METHOD

The proposed methodology for sentiment prediction of reviews relies on the utilization of machine learning algorithms, including dataset collection, data preprocessing, sentiment score generation, polarity calculation, application of the Naïve Bayes and SVM model, evaluation metrics, and result analysis. It is noteworthy that ML methods certainly have advantages in comparison with deep learning algorithms such as less complexity, less time-consuming for training process, simple optimization algorithms for hyper-parameter tuning with respect to the optimal ML structures. The workflow of the proposed methodology used in this research is illustrated in Figure 1.

A. Machine learning model

Naïve Bayes: The Naïve Bayes algorithm is a popular machine learning technique used for classification tasks, including sentiment analysis. It is based on Bayes' theorem and assumes independence among features. The algorithm calculates the probability of a given input belonging to a specific class by multiplying the probabilities of its individual features. Naïve Bayes is known for its simplicity and efficiency, making it well-suited for large-scale text classification tasks. Despite its assumption of feature independence, Naïve Bayes often performs surprisingly well in practice and can handle high-dimensional data efficiently. It is particularly useful in situations where the training data is limited, and it can be trained quickly even with large datasets.

Support vector machine: SVM aims to find an optimal hyperplane that separates data points of different classes with the maximum margin. It works by mapping input data into a high-dimensional feature space and then finding the hyperplane that best separates the classes. SVM is particularly useful for sentiment analysis due to its ability to handle high-dimensional and complex data, as it can capture non-linear relationships through the use of kernel functions. Additionally, SVM is known for its ability to handle small-sized datasets and its robustness against overfitting. It has been successfully applied in sentiment analysis tasks to effectively classify and analyze the sentiment expressed in text data.

B. Evaluating Measures

Evaluation metrics play a significant role in assessing the performance of classification tasks, with accuracy being the most commonly used measure. Accuracy represents the percentage of correctly classified instances in a given test dataset by the classifier. However, in text mining approaches, relying solely on accuracy may not provide a comprehensive understanding for making informed decisions. Therefore, additional metrics such as precision, recall and F1-score are commonly employed to evaluate the performance of classifiers. These measures provide valuable insights into the precision of positive predictions, the recall of actual positive instances, and a combined measure that balances both precision and recall, respectively. The frequency of correct predictions made by a classifier is measured by accuracy (Acc). Precision and Recall parameters show correct document identification and sensitivity of the classifier, respectively. The balance between Precision and Recall is given by F1-score, which is also known as the harmonic mean of those parameters. The following equations are employed for the calculation of above evaluation measures:

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1-score} = \frac{2}{\frac{1}{\text{Precission}} + \frac{1}{\text{Recall}}} \tag{4}$$

Where:

- TP (True Positive) represents the number of positive sentiment data correctly classified.
- FP (False Positive) represents the number of positive sentiment data incorrectly classified as negative sentiments.
- TN (True Negative) represents the number of negative sentiment data correctly classified.
- FN (False Negative) represents the number of negative sentiment data incorrectly classified as positive sentiment data.

IV. SIMULATION RESULTS

In this section, we present the simulation results of the application of the Naïve Bayes and SVM models for the analysis and prediction of sentiment in the E-commerce domain. The evaluation metrics, including accuracy, precision, recall and F1-score were employed to examine the proposed system. The entire dataset is divided into 80% of training data and 20% of evaluation data. Moreover, the grid search method is used to obtain the optimal parameters of the SVM model. As a result, cost (C) of 1.5 and Gaussian kernel (gamma) of 0.5 are selected as the optimal values for the SVM model. It is unnecessary for hyper-parameter tuning of Naïve Bayes model.

Figure 2 illustrates the evaluation parameters for the classifiers applied to the entire dataset. For the SVM classifier, the table shows an accuracy of 90.74%, precision of 90.95%, recall of 99.08%, and F1-score of 94.83%. On the other hand, the Naïve Bayes classifier achieved higher performance with an accuracy of 92.29%, precision of 92.22%, recall of 99.47%, and F1-score of 95.72%. The results indicate that the Naïve Bayes classifier outperforms the SVM classifier in terms of accuracy, precision, recall, and F1-score. It demonstrates the effectiveness of the Naïve Bayes algorithm for sentiment analysis on the entire dataset.

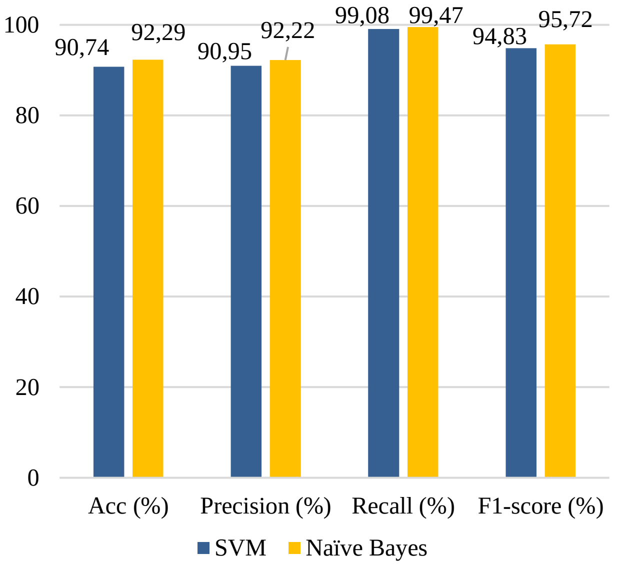

Figure 2: Performance of the ML models on the evaluation data

Table 2: Multiple classification performance of Naïve Bayes model on the evaluation data

Categories	Acc (%)	Precision (%)	Recall (%)	F1-score (%)
Laptops	90.19	90.07	99.88	94.73
Cameras	93.71	94.74	98.59	96.62
Mobile phones	92.98	91.86	99.93	95.72

Table 3: Multiple classification performance of SVM model on the evaluation data

Categories	Acc (%)	Precision (%)	Recall (%)	F1-score (%)
Laptops	88.27	88.49	99.47	93.66
Cameras	91.13	92.70	97.85	95.20
Mobile phones	92.83	91.67	99.93	95.62

Besides, to evaluate the efficiency of the consumer sentiment classification model for each product category, the outcomes are displayed in Tables 2 and 3. Table 2 illustrates the evaluation results of the Naïve Bayes model, while Table 3 displays the evaluation results of the SVM model. Furthermore, Figure 3 provides a visual representation of the results. Based on the comprehensive experimentation, it is clear that the Naïve Bayes algorithm outperformed the SVM model in terms of accuracy across all categories when assessed on the complete dataset.

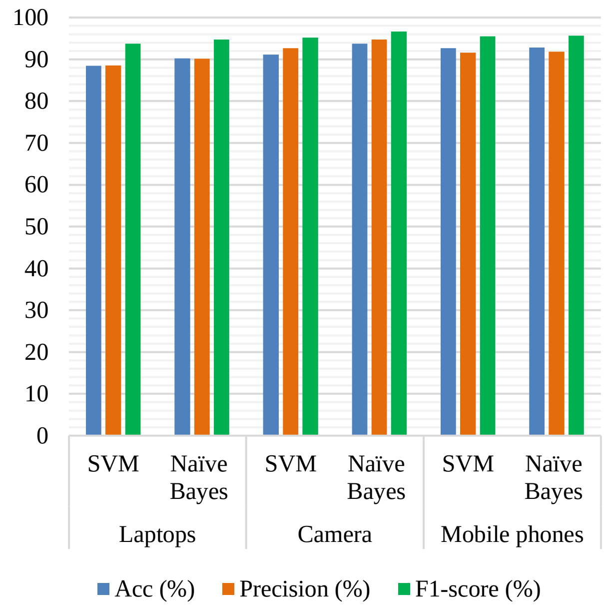

Figure 3: Performance comparisons of ML models in different review categories

V. CONCLUSION

Currently, there is a significant focus on Sentiment Analysis and Opinion Mining research, as it holds great importance for various industries. Industries generate diverse datasets and analyzing this data helps them make informed decisions. The advent of social media has also led to a massive influx of data, which requires analysis to extract meaningful insights.

In this study, a dataset consisting of product reviews from four categories, namely laptops, cameras, and mobile phones, was collected from the Amazon website. The proposed methodology employed a dictionary-based approach within a lexicon-based framework, integrating machine learning techniques. Sentiment analysis was conducted on each product review and subsequently classified using two machine learning algorithms, Naïve Bayes and SVM. The accuracy measurements of these classifiers for the dataset are depicted in Figure 2. Both models achieved an accuracy rate of over 90%, accompanied by precision, recall, and F1-scores also exceeding 90%. Specifically, the Naïve Bayes classifier achieved an accuracy of 92.29%, while the SVM classifier achieved an accuracy of 90.74% for the dataset.

REFERENCES

[1] Verma, J. P., Patel, B., & Patel, A. (2015, February). Big data analysis: recommendation system with Hadoop framework. In *2015 IEEE International Conference on Computational Intelligence & Communication Technology* (pp. 92-97). IEEE.

[2] Choudhary, M., & Choudhary, P. K. (2018, December). Sentiment analysis of text reviewing algorithm using data mining. In *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 532-538). IEEE.

[3] Sasikala, P., & Mary Immaculate Sheela, L. (2020). Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS. *Journal of Big Data*, *7*, 1-20.

[4] Subramaniyaswamy, V., Vijayakumar, V., Logesh, R., & Indragandhi, V. (2015). Unstructured data analysis on big data using map reduce. *Procedia Computer Science*, *50*, 456-465.

[5] Wassan, Sobia, et al. "Amazon product sentiment analysis using machine learning techniques." *Revista Argentina de Clínica Psicológica* 30.1 (2021): 695.

[6] Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." *Journal of Big Data* 2.1 (2015): 1-14.

[7] Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, *10*(2).

[8] Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. *International Journal*, *2*(6), 282-292.

[9] Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177).

[10] Gautam, G., & Yadav, D. (2014, August). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In *2014 Seventh international conference on contemporary computing (IC3)* (pp. 437-442).

[11] Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Berlin, Heidelberg: Springer Berlin Heidelberg.

[12] Khan, F. H., Bashir, S., & Qamar, U. (2014). TOM: Twitter opinion mining framework using hybrid classification scheme. *Decision support systems*, *57*, 245-257.

[13] Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N. (2013). What yelp fake review filter might be doing?. In *Proceedings of the international AAAI conference on web and social media* (Vol. 7, No. 1, pp. 409-418).

[14] Heydari, A., Tavakoli, M., & Salim, N. (2016). Detection of fake opinions using time series. *Expert Systems with Applications*, *58*, 83-92.

[15] Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Computing and Applications*, *32*, 17259-17274.

[16] Long, F., Zhou, K., & Ou, W. (2019). Sentiment analysis of text based on bidirectional LSTM with multi-head attention. *IEEE Access*, *7*, 141960-141969.

[17] Dong, J., Chen, Y., Gu, A., Chen, J., Li, L., Chen, Q., ... & Xun, Q. (2020). Potential Trend for Online Shopping Data Based on the Linear Regression and Sentiment Analysis. *Mathematical Problems in Engineering*, *2020*, 1-11.

[18] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Proceedings of the ACL-02 conference on Empirical methods in natural language processing. *10: 79–86. doi: 10.3115, 1118693*.

[19] Kraus, M., & Feuerriegel, S. (2019). Sentiment analysis based on rhetorical structure theory: Learning deep neural networks from discourse trees. *Expert Systems with Applications*, *118*, 65-79.

[20] Abid, F., Alam, M., Yasir, M., & Li, C. (2019). Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. *Future Generation Computer Systems*, *95*, 292-308.

[21] Ahmad, M., Aftab, S., Ali, I., & Hameed, N. J. I. J. M. S. E. (2017). Hybrid tools and techniques for sentiment analysis: a review. *Int. J. Multidiscip. Sci. Eng*, *8*(3), 29-33.

[22] Bravo-Marquez, F., Frank, E., & Pfahringer, B. (2016). Building a Twitter opinion lexicon from automatically-annotated tweets. *Knowledge-Based Systems*, *108*, 65-78.

PHÂN TÍCH QUAN ĐIỂM KHÁCH HÀNG VỀ SẢN PHẨM ĐIỆN TỬ DỰA TRÊN HỌC MÁY.

Tóm tắt: Phân tích và khai thác quan điểm khách hàng đã được áp dụng phổ biến trong nhiều lĩnh vực nhằm phân tích, trích xuất thông tin hữu ích từ dữ liệu văn bản đến từ các nguồn khác nhau như Facebook, Twitter, Amazon. Kỹ thuật này có vai trò quan trọng trong việc hỗ trợ các doanh nghiệp tích cực nâng cao chiến lược kinh doanh bằng các thu thập thông tin về quan điểm khách hang đối với sản phẩm của họ. Quá trình này bao gồm việc tận dụng các phương pháp tính toán để nghiên cứu hành vi mua hàng của khách hàng và sau đó khai thác ý kiến, phản hồi của họ, được thể hiện dưới dạng sự kiện, cá nhân, bài đăng trên blog hoặc trải nghiệm sản phẩm. Nghiên cứu này tập trung vào việc sử dụng tập dữ liệu thu thập từ Amazon, bao gồm các đánh giá, quan điểm khách hang về các sản phẩm khác nhau như máy tính xách tay, máy ảnh và điện thoại di động. Dữ liệu sau khi được tiền xử lý sẽ được đưa vào các bộ học máy để phân loại các đánh giá theo cảm tính tích cực hoặc tiêu cực. Bước phân loại này cho phép chúng tôi phân tích cảm nhận chung liên quan đến sản phẩm và đưa ra kết luận có ý nghĩa.

Từ khóa: Yêu cầu của khách hàng, thiết bị điện tử, học máy, xử lý ngôn ngữ tự nhiên, phân tích quan điểm khách hàng

Tuan Hoang Vu is B.E student in information technology of ThuyLoi University. His research is gene selection, sentiment analysis, optimization, artificial intelligence.

Minh Tuan Nguyen received the B.S. degree from the Post & Telecommunications Institute of Technology, Hanoi, Vietnam, in 2004, the M.S. degree from Hanoi University of Science and Technology, Hanoi, Vietnam, in 2008, both in electronics and telecommunications engineering, and the Ph.D. degree at the Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2018. He is with Posts and Telecommunications Institute of Technology. His research interests include network security, internet of things, biomedical signal processing, gene analysis, sentiment analysis, brain computer interface, machine learning, deep learning, optimization, and biomedical application design.
Email: nmtuan@ptit.edu.vn