Product Review Sentiment Analysis by Artificial Neural Network Algorithm

Tri Astuti [a,1,*], Irnawati Pratika [a,2]

[a] Information System Program, STMIK AMIKOM Purwokerto
[1] tri_astuti@amikompurwokerto.ac.id *; [2] nonnamudanawa@gmail.com
* corresponding author

Abstract

Buying and selling and marketing goods and services are now done online. The online store provides facilities that enable its customers to provide review related products offered. The number of reviews received by the store, online sometimes does not allow the store online to analyze one by one. Thus, it takes the help of machines to assist in the analysis of such sentiments. Analysis of the sentiments of the review the product is done to help the shop get a general overview related to the level of consumer satisfaction. In this study, the ANN algorithm will be used to analyze sentiment for review. A product ANN algorithm used because it can provide high accuracy performance. This research resulted in a reasonably high accuracy performance is 88.2%.

Keywords: Sentiment Analysis; ANN; Product Review.

1. Introduction

The rapid development of technology provides much ease for modern society. One of them is the ease of dealing with buy and sell that can be done over the Internet, or more commonly known by e-commerce. According to research conducted by Ismail in 2017, there are at least 3 online stores that dominate the E-CommerceMarket, namely Alibaba, Amazon, and eBay [1]. Based on a survey conducted by Statista in 2017, the number of active visitors owned by Alibaba reached 488 million visitors [2][3]. While Amazon has 183 million visitors. While eBay only reaches 87 million visitors. The number shows the magnitude of public interest in online stores.

According to Fobes [4], one of the reasons that make the online store is in demand by the community is the ease in obtaining product-related information and the price comparison without having to meet the shop party. Also, the online store provides more convenience than conventional stores, such as free shuttle service. Not infrequently, online stores also provide facilities that allow consumers to provide feedback or review related products offered by the online store. This feedback or reviews can later be used by online store management to analyze consumer feedback on the products offered[5][6].

In general, consumer response can be grouped into two broad groups, namely, satisfying and disappointing. Consumer feedback is said to satisfy when it contains positive sentiments in it. Conversely, the response is said to disappoint when it contains negative sentiments [7]. Nevertheless, the number of responses received by online stores; sometimes, people cannot analyze them one by one. So, it needs the help of the engine to assist with the sentiment analysis process. Thus, a sentiment analysis algorithm is required that can be implemented into the machine so that the machine can recognize the sentiment of the response given by the consumer[8]. The analysis of these sentiments will later help the online store to get an overview of their consumer satisfaction level [9][10].

In this study, the ANN algorithm will be used to perform sentiment analysis on product reviews[11][12]. The implementation of the ANN algorithm on the product review sentiment analysis will result in a more stable classification rule architecture[13]. By setting up a neural network algorithm on the sentiment analysis of product reviews, the machine is expected to have high accuracy in classifying the sentiments of consumer responses related to products offered by online stores[14][15]. As such, management will be helped to get a general idea of their consumer satisfaction level.

2. Research Concept

In this study, a concept was conducted in the same analysis with the following stages:

a. Literary studies

This stage is used to determine an overview of the research related to sentiment analysis. Also, this stage describes the extent to which research is themed sentiment analysis.

b. Problem identification

This stage is used to identify any issues related to the E-Commerce field.

c. Data Collection

This stage is done to collect the necessary data in the research. Data should be collected according to the plans that have been compiled in the previous step.

d. Pre-processing

The pre-processing stage aims to perform pre-processing of data that has been obtained in order to be used in research that will be done.

e. Building an ANN model

This stage aims to design the ANN architecture that is used in research. The architectures used should allow us to have a quite high accuracy but not too much in need of computing resources.

f. Train Model ANN

At this stage, the ANN has been designed to be trained using the training set that has been prepared. ANN is trained in such a way that it has a relatively small error value.

g. Testing ANN Performance

ANN, which has been successfully trained, will be tested its accuracy performance by using a pre-prepared testing set. The test results will be documented and will be the evaluation material of this research.

3. Results and Discussion

3.1. Results

a. Data Collection

The data to be used in this study is product review data from Amazon.com. Datasets have been grouped into two classes, i.e., positive sentiment classes and negative sentiment classes. The Dataset used in this study amounted to 448 reviews consisting of 230 positive sentiment sentences and 218 negative sentiment sentences. Its class-based has labeled each data. The sign (+) is a label of a positive sentiment sentence, while negative sentiment is labeled with a (-) sign. The semicolon (;) (determinant) between the label and the sentence.

b. Pre-processing

1) Punctuation Removal

Punctuation removal aims to remove existing punctuation marks on datasets such as question marks (?), exclamation (!), dots (.), commas (,) and other punctuation marks. Table 1 is an example of data before and after punctuation removal. The punctuation in the review sentence is eliminated so that it only forms a sentence without punctuation.

Table 1 Punctuation Removal

No	Class labels	Before the *punctuation removal process*	After the *punctuation removal process*
1	+ ;	It has good color and price was great.	It has good color and price was great

...	...	....	....
448	- ;	The customer support is very rude.	The customer support is very rude

2) Tokenization

Tokenization is the cutting phase of text documents based on each word they are composing. The piece is called a token or term. At this stage, the dataset will be checked from the first character until the last character shown in table 2.

Table 2 Tokenization

No	Class labels	Instances before the T-tokenization Process	Instances after the tokenization Process
1	+ ;	It has good color and price was great	"It", "has", "good", "color", "and", "price", "was", "great"
...	.....	....	....
448	- ;	The customer support is very rude	"The", "customer", "support", "is", "very", "rude"

3) Single Character Removal

At this stage, the elimination of single characters or letters of that only consists of one character such as a, I, N, and so forth.

4) Stopword Removal

Stopword removal is a process of eliminating words that often appear but has no influence whatsoever in the extraction of a sentiment review. Words include such as a timepiece, a question word, and so on. Examples of and, to and the.

5) Stemming

In the stemming process, it will find the root word and remove the suffix to the word. Stemming aims to reduce variations of words that have the same base word. Examples of "greating" and "greats" have the basic word "great". Table 3 shows changes before and after the stemming process.

Table 3 Stemming

No	Class labels	Instances before the stemming process	Instances after the stemming process
1	+ ;	Good color price Great	Good color price Great
...	.....	....	....
448	- ;	Customer Support Rude	Custom Support Rude

The word for each review is then added to the 2D structure, with each of the columns containing the words of each review. After the words go through the pre-processing process, has syllables are shown in table 4.

Table 4 Review on List of 2d Structures

Review of the								
1	2	3	4	...	...	446	447	448
Good	Monitor	Color	Reali	...	...	Fact	Initi	Custom
Color	Much	Everyth	Enjoy	...	...	Other	Sound	Support
Price	Much	GReat	Product	...	...	Say	Low	Rude
Great	Better			...	...	Bright	Sound	
	Still			...	...	Enough		
	Great			...	...			

6) Vocabulary List

In this process, each word is added to the vocabulary list. The vocabulary is then sorted in alphabetical order. A Review of the 2d list in table 4 will result in a list of 1182 vocabularies shown in table 5.

Table 5 Vocabulary List

No	Vocabulary List
1	Abil
2	Absolute
...	.....
1181	Yet
1182	Youtube

7) Label Vector Training

The training Label for positive or negative is stored in 2 x (Total review) matrix. The number 1 in the first row and the number 0 in the second line is a Negative sentiment review, and the number 1 in the second row with 0 in the first line is a positive review sentiment. Table 6 shows an example of this structure.

Table 6 Label Selection list

Review of the	1	2	3	4	....	.....	....	445	446	447	448
	0	0	0	0	....	....	....	1	1	1	1
	1	1	1	1	....	....	....	0	0	0	0

8) Numerical Training Vector

The size matrix (number of vocab) x (number of reviews) made will be used on ANN for training. The number of rows is the same as the number of words collected during the pre-processing process, so if the word in the review will be labeled 1, if it is not there, it is labeled with a number 0. The variable map is used to determine which line 1 will be placed in. A review of table 4 will produce the input vector shown in table 7.

Table 7 Numerical Vector Equivalent

The	Review of the												
1	1	2	3	4	5	6	7	...	...	445	446	447	448
2	0	0	0	0	0	0	0	...	...	0	0	0	0
3	0	0	0	0	0	0	0	...	...	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1181	0	0	0	0	0	0	0	...	...	0	0	0	0
1182	0	0	0	0	0	0	0	...	...	0	0	0	0

The neural networks are trained using a numerical input vector, which is shown in table 7, and the training label is shown in table 6. The results of the above processes will be saved into 3 different CSV files named Vocabularylist.csv, _vector-training.csv label, and numerical_vector-training.csv. The neural network used is a pattern recognition network provided by Matlab, which is a feedforward network using the conjugate scale gradient method to adjust weights and bias.

9) Results

The ANN development process uses the help of the tool found in the MATLAB, the Nntool (Neural Network tool). DataSet files used as input are the process results of numerical vector training, and the target is numeric results that have been saved as a result of the vector training label. Some inputs to neural networks amounted to 1182 by the amount of vocabulary stored on the Vocabulary list. csv. The output layer consists of two neurons, one neuron represents the probability of being a positive review, and the other is likely to be a negative review.

Preliminary testing is done with 1 hidden layer of 10 neurons used for all experiments. Table 8 shows the accuracy performance generated by neural networks with different data divisions.

Table 8 Results in Accuracy with Different Amount of Data Sharing

Data percentage		Accuracy performance
Data Training	Data Testing	
50%	50%	68.3%
60%	40%	71.9%
70%	30	88.2%
80%	20	85.5%
85%	15	56.9%
90%	10	51.3%

Table 8 shows that the highest accuracy result is 88.2%, with 70% training data sharing and 30% data testing. From the result, the amount of data sharing 70%-30% will be used to perform subsequent tests. Testing is done by entering the number of neurons on different hidden layers ranging from 5 neurons to 20 neurons shown in table 9.

Table 9 Performance Accuracy ANN with Number of Hidden Layers

Number of neurons at Hidden Layer	Accuracy performance
5	69.0%
6	80.4%
7	83.0%
8	78.1%
9	80.1%
10	88.2%
11	60.3%
12	56.3%
13	46.9%
14	68.8%
15	78.1%
16	66.5%
17	75.0%
18	71.5%
19	60.0%
20	63.2%

The difference in the number of neurons in a hidden layer affects the accuracy performance generated by neural networks. The accuracy experienced a drastic decline in testing with the number of neurons 13 in the hidden layer, which is only reached 46.9%.

From table 8 and Table 9, the highest accuracy results gained from testing with 70% of training data distribution and 30% Data testing with the number of neurons on hidden layer 10, which is at 88.2%.

4. Conclusions and Suggestions

Based on the results of the research done, then it can be concluded as follows:

1. It has been applied to neural network algorithms to analyze the sentiment of product reviews on Amazon.com, especially electronic products.

2. By setting up a neural network algorithm to analyze the sentiment of product reviews shows the results of satisfactory accuracy performance of 88.2% with a large number of datasets that are 1182 inputs. This suggests that the artificial neural network algorithm is worth using to analyze sentiment.

References
[1] B. Duncan, and Y. Zhang, "Neural networks for sentiment analysis on twitter". In Cognitive Informatics & Cognitive Computing (ICCI* CC), 2015 IEEE 14th International Conference on (pp. 275-278). IEEE, 2015.

[2] Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages:Feature selection for opinion classification in web forums.ACM Transactions on Information Systems, 26(3), 1–34. 12.

[3] Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of twitter data. InProceeding of ACL HLT conference(pp. 30–38).

[4] Aue, A., & Gamon, M. (2005). Customizing sentiment classifiers to new domains: a case study. InProceeding of the intl. conference on recent advances in natural language processing. Borovets, BG.

[5] Bakshy, E., Hofman, J., Mason, W., & Watts, D. (2011). Everyone's an influencer:Quantifying influence on Twitter. InProceeding of ACM WSDM conference Hong Kong. China.

[6] Barbosa, L., & Feng, J. (2010). Robust sentiment detection on twitter from biased and noisy data. InProceedings of the 23rd international conference on computational linguistics (COLING'10) (pp. 36–44).

[7] Benhardus, J., & Kalita, J. (2013). Streaming trend detection in Twitter.International Journal on Web Based Communities, 9(1), 122–139.

[8] Bermingham, A., & Smeaton, A. (2010). Classifying sentiment in microblogs: Is brevity an advantage? InProceeding of ACM CIKM conference Toronto, Ontario, Canada(pp. 1833–1836).

[9] Bermingham, A., & Smeaton, A. (2011). On using twitter to monitor political sentiment and predict election results. InProceeding of IJCNLP conference, Chiang Mai, Thailand.

[10] Bifet, A., & Frank, E. (2010). Sentiment knowledge discovery in Twitter streaming data. In Proceeding of 13th international conference on Discovery Science Conference(pp. 1–15).

[11] Blitzer, J., Dredze, M., & Pereira, F. (2007). Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. InProceeding of the annual meetings of the association for computational linguistics(pp. 440–447).

[12] Bollen, J., Pepe, A., & Mao, H. (2011b). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In Proceedings of the fifth international aaai conference on weblogs and social media (ICWSM 2011), July, Barcelona, Spain, (pp. 1–10).

[13] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market.Journal of Computational Science, 2(1), 1–8.

[14] Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. (2010). Measuring user influence in Twitter: The million follower fallacy. In Proceeding of 4th AAAI conference on weblogs and social media, Washington DC, (pp. 10–17).

[15] Chung, J., Mustafaraj, E. (2011). Can collective sentiment expressed on twitter predict political elections? InProceedings of the twenty-fifth AAAI conference on artificial intelligence(pp. 1770–1771).