Aspect Based Sentiment Analysis of Product Review Using Memory Network

Hilya Tsaniya Ismet[1]*, Tanzilal Mustaqim[2], Diana Purwitasari[3]

[1,2,3]Department of Informatics Engineering Faculty of Electrical Technology and Intelligent Information Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstract.

Purpose: Consumer opinion is one of the essential keys that affect the success of a product. Sentiment analysis of consumer opinion is needed to find out information about customer satisfaction for companies in the decision-making process. The traditional sentiment analysis process extracts a complete sentiment from a single sentence. However, it does not consist of only one sentiment in one sentence. The total number depends on the number of aspects that make up the sentence. Therefore, a sentiment analysis process is needed to pay attention to aspects.

Methods: This research focuses on product reviews from Indonesian e-commerce on several aspects of sentiment. Uses fastText word embedding to avoid Out of Vocabulary in datasets and Gated Recurrent Units for aspect spread detection. Sentiment classification on aspects using the Memory Network method.

Result: The experiment results showed that aspect-based sentiment classification predictions had an accuracy of 83% compared to 78% overall classification predictions for review texts, indicating that aspect-based sentiment analysis can improve model performance on product review classification predictions.

Novelty: Most product reviews analysis use document-level classification to extract and predict sentiment reviews, aspect-based analysis can be applied to product reviews for better sentiment understanding, using Memory Network to store important information explicitly on aspects and polarity.

Keywords: Aspect Based Sentiment Analysis, Memory Network, Product Review

Received December 2021 / **Revised** January 2022 / **Accepted** April 2022

INTRODUCTION

Technological advances have influenced buying and selling activities in e-commerce to become increasingly popular [1]. E-commerce is becoming popular, supported by easy access via mobile phones. Consumer access becomes easier, more convenient and faster when making buying and selling transactions [2]. Customer satisfaction is an important factor in increasing product attractiveness [3]. E-commerce platforms, in general, have provided a feature of providing feedback or product reviews. This feature is given to consumers who have completed transactions. Product review information can influence potential consumers in determining transaction decisions. The massive number of e-commerce transactions resulted in a massive number of product reviews as well.

Sentiment analysis is one method of product review analysis. Sentiment analysis is a technique or method used to identify how sentiment is expressed using text and how that sentiment can be categorized as polarity [4]. Each part of the document has its own polarity [5]. The application of the sentiment analysis method is very broad in the field of language processing with the aim of analyzing opinions, evaluations, attitudes, judgments and emotions on a topic [6]. There are three types of sentiment analysis: document-level analysis, sentence-level analysis, and aspect level analysis [7]. Sentiment analysis in product reviews generally analyzes polarity at the document level [8]. This approach has limitations because it cannot extract all sentiment polarities contained in the data. To avoid this, sentiment analysis can be narrowed down to an aspect-based level.

Aspect-based sentiment analysis (ABSA) is the process of extracting sentiment from every aspect contained in a sentence. In addition, ABSA also groups aspects based on their semantic suitability [9]. The process of extracting aspect sentiment by paying attention to related words. For example, "This laptop is good, but the

---

price is high", the sentiment aspects of "laptop" and "price" have positive and negative values , respectively. ABSA performs aspect extraction first and then analyzes the polarity of the sentiment. Previous research related to ABSA has been carried out by many researchers with various methods. One of them is the machine learning model method. Naïve Bayes machine learning model with Chi-square feature selection has been used for ABSA in the SemEval-2014 dataset [10]. The research still uses a bag-of-words vectorization process so that it is still vulnerable to out-of-vocabulary (OOV). Research using the Naïve Bayes model, Support Vector Machine (SVM) and Iterative Decision Tree is compared with the vectorization process of Binary Term (BT), Term Frequency (TF) and Term Frequency - Inverse Document Frequency (TF-IDF) [11]. The study is also susceptible to the occurrence of OOV.

Exploration and testing of sentiment analysis in the product review domain continue to grow to the use of deep learning methods. ABSA research with Bidirectional Gated Recurrent Unit (Bi-GRU) and Multi-Level Attention has been carried out on healthcare datasets [12]. The results of the study were able to outperform previous methods such as Gated Convolutional network with Embedding Aspect (GCAE), Sentic Long short-term memory (SLSTM) and Char Sentic LSTM (CSLSTM). GRU has similarities with LSTM in processing text data. The difference is that the GRU does not have a cell state. Bi-Gru also has lower complexity than LSTM while still providing comparable performance. The LSTM model has been used for ABSA research on the evaluation of Amazon e-commerce products [13].

Aspect-based research on Indonesian language data has been carried out on the subject of tourism using the Naïve Bayes machine learning model [14]. Another aspect-based sentiment analysis study was conducted on the subject of restaurants with datasets obtained from zomato and the SVM model used as the basis for learning [15]. Learning the context of information on data and aspects affects the process of extracting aspect and sentence representations separately [16]. The study used interactive attention, which succeeded in increasing the effectiveness of the model with better precision values. The weakness of interactive attention is that the model tends to fail to associate a context with aspects of multi-aspect analysis. Another study related to the Memory Network conducted a trial test on aspect matching and sentiment information[9]. The study succeeded in obtaining better precision on multi-aspect information associations.

Another study related to ABSA uses word embedding as a data representation in vector form. Research on sentiment analysis on aircraft object data sourced from Twitter compares two-word embedding [17]. The study compared fastText and AraVec-Web. fastText, which has been trained from Arabic Wikipedia data, produces better performance than AraVec-Web. Another study uses fastText as word embedding in student feedback analysis [18]. Word embedding used are fastText, GloVe, Word2Vec and MOOC. The results from fastText are able to outperform the other three word embedding from the F1 score measurement. FastText performs the representation process on character-level information.

In this study, aspect-based sentiment analysis will be carried out on product review data from one of the e-commerce sites in Indonesia, with a focus on seeing the multi-aspect effect on the accuracy of sentiment prediction results. To avoid the occurrence of Out of Vocabulary (OOV), the Indonesian language pre-trained fasttext embedding model is used in the Indonesian product review data. The GRU method is used for the extraction and distribution of aspects with the aim of increasing the accuracy of sentiment polarity. As a classification problem, we use the evaluation metrics of Precision, Recall, F-score, and Accuracy to analyze Memory Network's performance for aspect-based sentiment analysis on product review. Contributions to this research are the use of the Memory Network method in aspect-based sentiment analysis to match the aspect value and sentiment to the review sentence and the implementation of sentiment prediction using the Indonesian e-commerce dataset.

Sentiment Aspect
Aspect is the topic of discussion, which is the subject of the sentence. on aspect-based sentiment analysis, the sentiment classification process focuses on each aspect. Before being analyzed sentimentally, aspects are first extracted from the sentence. The aspect extraction process pays attention to the features of the phrases in the related sentences. Several studies related to aspect extraction can use rule-based methods that produce aspects and opinions that are directly related. This method performs aspect extraction based on predetermined rules [19]. The other method performs aspect extraction first and continues with opinion extraction from aspects related to a certain distance [20], [21].

The results of the aspect extraction are then carried out with a sentiment analysis whose analysis process takes into account the supporting words of the aspect. The word supporting the aspect is in the form of an opinion that explains or complements the aspect. The sentiment analysis process from the aspect has been carried out by many methods. Research that combines particle swarm optimization is carried out with feature selection and ensemble learning, the basic algorithms used are Maximum Entropy (ME), Conditional Random Field (CRF) and Support Vector Machine (SVM). The dataset that became the basis of the analysis was SemEval-2014 with the data used, namely restaurants and laptops [22]. The process of analyzing the smart government review data with aspect-based sentiment analysis uses the lexicon-based analysis method with the SVM machine learning algorithm. In this study, pre-processing was carried out to handle negation, intensification, downtoners, repetitive characters, and special cases of opinion-negation rules [23]. The accuracy results obtained outperform the previous analysis by Manek [24] which used a Gini index-based SVM.

Gated Recurrent Units (GRU)
Gated Recurrent Unit (GRU) is a sequential-based deep learning model. This model is the development of the Recurrent Neural Network (RNN). RNN is a form of Artificial Neural Network (ANN) architecture whose analysis process pays attention to sequential data. GRU has the advantage of overcoming the problem of gradient values that disappear in the training process. In the GRU there are two components of information control or can be called a gate, namely the reset gate and the update gate. The reset gate on the GRU will determine the merging of new inputs to existing information, and the update gate will determine a lot of important information to be stored. The GRU was proposed by Cho [25] who proposed an LSTM development method with the aim of obtaining an algorithm with adaptive units to capture dependencies from different timescales. Similar to the units in the LSTM, the GRU has a gating unit or gates that modulate the flow of information within the unit, but without separate memory cells.

The use of GRU as a method of learning aspect context information is carried out by Wang [7] to get a representation of aspects in sentences with better accuracy. Another study used GRU as an aspect detection process in e-commerce datasets by analyzing the results of word embedding [26], and succeeded in outperforming the fully-connected layer which analyzed the vector results from the bag of words in terms of precision measurement, recall and f1- scores. The GRU method in learning context information on aspects has good results with lower complexity in improvising the accuracy of aspect-based sentiment classification predictions [27].

Memory Network
Memory Network is a machine learning method that focuses on long-term memory inference in the process stage for classification prediction. Memory Network has 5 components, namely memory, I as input feature representation, G as memory that will be updated from input feature representation, O as output representation, and R as response to output representation. In the application of text classification, text as input will be received on I which is then stored in the memory of module G, on O the output feature is performed based on the existing memory in module G, on R will be decoded output based on the data stored in memory. The detailed schematic of the Memory Network is shown in Figure 1.
Memory Network proposed by Jason [28], who are members of the Facebook Research AI team with a strategy of using Long-term memory that can be read and written-to processes with the aim of using information in memory to make predictions. Memory Network, a type of neural network that can store context information for words in external memory, Tang [29] first tried to use the Memory Network method
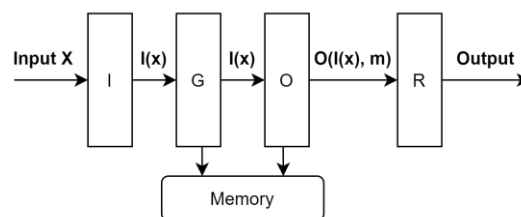

Figure 1. Memory network scheme

in aspect-based sentiment analysis and was able to capture the importance of context in predicting aspect sentiment. The use of the Memory Network with weight learning using the recurrent network model has also been carried out Chen [30] with the ability to capture important information from context to aspects that are better.

METHODS

The research implementation uses a dataset sourced from an Indonesian e-commerce. The e-commerce data used is Lazada. The data can be found publicly on Kaggle under the title "Lazada Indonesian Reviews". The process of retrieving data by scraping using puppeteer with Node.js. The dataset contains product review information with 15 different columns, namely itemId, category, name, rating, originalRating, reviewTitle, reviewContent, likeCount, upVotes, downVotes, helpful, relevanceScore, boughtDate, clientType, and retrievedDate. Of the 15 columns selected 3 columns that show important information related to the review, namely category, Review text, relevanceScore. The data will be word embedding, aspect extraction, and classification, an illustration of the stages of the research method can be seen in Figure 2.

Preprocess Data

Data preprocessing is the process of transforming data from raw datasets into data that can be processed by learning models. The dataset cannot be directly analyzed by the learning model. This is because the format of the dataset is not appropriate. The raw dataset still contains a lot of data that is not needed at the time of analysis. Some unneeded data such as stopwords and non-uniform word character writing. Stopwords are words that appear frequently so that they are considered not to have significant semantic meaning. Writing words on raw data is usually not in accordance with grammatical writing. One example is data that uses capital letters in the middle such as "keRen". The process of data uniformity is changed to the lowercase form.

The dataset used in this study is not yet labeled, so manual labeling is required by the researcher. The amount of data used for testing amounted to 1180 data. The data labeling format process is inspired by SemEval 2015 task 12. The features used are review text, number of aspects, aspect details and aspect sentiment. The results of manual labeling show that each total aspect is between a distance of 1 to 7. The polarity class of aspect sentiment is grouped into neutral, positive and negative. The polarity of the sentiment on the same aspect in different sentences has various values. For example, the "HP battery" aspect has a positive sentiment value in the first sentence and this aspect has a negative sentiment value in the next sentence. The difference in aspect sentiment values can avoid the case of dominant data in one class so as to minimize the occurrence of an unbalanced number of classes. The pre-processing stages of the data in this study were data cleaning, remove stopwords, stemming and labeling. The pre-process stage in this research uses the literary python library. Literature has been equipped with a stopword list in Indonesian and a corpus in Indonesian. The library has been widely used in research in the field of information retrieval systems. Literature has been equipped with a stopword list in Indonesian and a corpus in Indonesian. The library has been widely used in research in the field of information retrieval systems. Literature has been equipped with a stopword list in Indonesian and a corpus in Indonesian. The library has been widely used in research in the field of information retrieval systems.

Data Cleaning

Data cleaning is the process of cleaning the dataset from some incorrect, duplicate, inconsistent and incomplete data. Data cleaning also removes some data that tends to become noise during the analysis process. The purpose of data cleaning is to process raw datasets that are still filled with inappropriate data into clean data. Clean data has a high probability to produce optimal analysis. The data cleaning process in this study includes cleaning and adjusting the characters that are not used in the analysis process.

Case folding also applied to uniform all data into lowercase. Case folding produces cases of all words being the same so that they can optimize the analysis process, especially from the detection of duplicates and data that do not match. The next deleted data are website addresses, phone numbers and random characters. The data deletion process uses the regex feature in the python programming language.

An example of data cleaning by taking one part of the data on the Lazada e-commerce dataset, namely from the data "Barang diterima dengan kondisi yang baik tapi packagingnya kurang" changed to "barang diterima dengan kondisi yang baik tapi packagingnya kurang".

Remove Stopword

Remove stopword is the process of removing words that often appear in the text. These words are articles, prepositions, pronouns, conjunctions and others. Stopword was removed because it has less information than other words. The frequency of occurrence of very high stopwords tends to produce noise during the analysis process. Stopwords are also removed with the aim of saving memory efficiency and optimizing the computational analysis process.

Stopword is deleted based on a pre-compiled word list. If there are words in the dataset that are the same as the list of words in the stopwords, the data will be deleted. The stopword removal process in this study utilizes the word list available in the literary python library.

An example of implementing a remove stopword on a dataset is the data "barang diterima dengan kondisi yang baik tapi packagingnya kurang" changed to "barang diterima kondisi baik packagingnya kurang".

Stemming

Stemming is a process of transformation from the affixed form to its root form. Literally stemming means cutting affixes or changes and basic words. Affixes that are often encountered are prefixes, suffixes, infixes, reduplications, inflections and others. Stemming aims to make data that have the same basic words uniform. This makes the computational analysis process efficient and can optimize the results. The stemming process in this study uses the literary python library. The library uses a dictionary of basic words to match the analysis data.

An example of the implementation of stemming on the dataset is from the data "barang diterima kondisi baik packagingnya kurang" changed to "barang terima kondisi baik packaging kurang".

Labelling

Labeling is the process of identifying raw data by adding information as target data in the analysis process. Raw data labeling analyzes the content and format of the e-commerce product review dataset. The researcher's analysis process is by manually extracting aspects and polarity of sentiment from each product review data. The researcher also transformed the dataset format by removing unused data features. The labeling process was analyzed using an excel spreadsheet as a data processing tool.

Embedding

The dataset used in this study is text data. Text data is very easy to understand by humans. However, the computer cannot understand it directly, computers only understand data in numeric form. Therefore, a conversion process from text data to numeric data is required, which is called embedding. Embedding changes the shape of the text data to a vector form as a representation of the text data. The text data that is changed is the word. The conversion process pays attention to the distance between the vector values and the similarity of the word representation. This optimizes the analysis process. Details of the embedding scheme are shown in Figure 3.
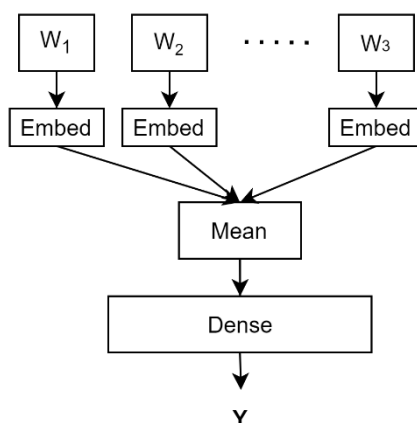

Figure 3. Embedding process scheme

In this study, the embedding model used is FastText. FastText is a library developed by Facebook AI Research (FAIR). FastText can train large amounts of data in a short time. The pre-processed data will be embedding the data using the Indonesian FastText Pre-trained model. The model has a vector representation value that has been trained on the Indonesian wikipedia corpus data. One of the advantages of using FastText is to handle words that have never been encountered before (Out of Vocabulary Word). In the embedding process, words in the data will be given initial weights based on the model used, on phrase weights initialization.

Aspect Extraction

In this study, to see the distribution of aspects of the data, the data representation is processed using the GRU, as the initiation of the input query in the classification model. Because not all words in a sentence have sentiment information, it is necessary to distribute the aspects that represent the sentence. The distribution of aspects of this sentence is carried out using the GRU method by matching the existing aspects with the words in the review sentence. Mathematically it can be written as follows:

$$S = [\, w_1 \oplus a_1,\ w_2 \oplus a_2, \dots, w_n \oplus a_i] \in R^n \tag{1}$$

Next, the sentence representation is used as input to the GRU to study the context of the information from the sentence, the GRU formula notation used can be described as follows:

$$r_t = \sigma\,(W_r s_t + W_r h_{t-1}) \tag{2}$$
$$u_t = \sigma\,(W_u s_t + W_r h_{t-1}) \tag{3}$$
$$h_t = tanh\,(W.[r_t * h_{t-1}] + W s_t) \tag{4}$$
$$a_t = (1 - u_t) * a_{t-1} + u_t * h_t \tag{5}$$

The output of the GRU with the context representation of the aspect information in the sentence will be used as an input query on the Memory Network. The use of GRU can see the spread of aspects to context and other word information in the review sentence. GRU suitable for use for data sequences that are not too large and take less running time with less performance than other learning methods.

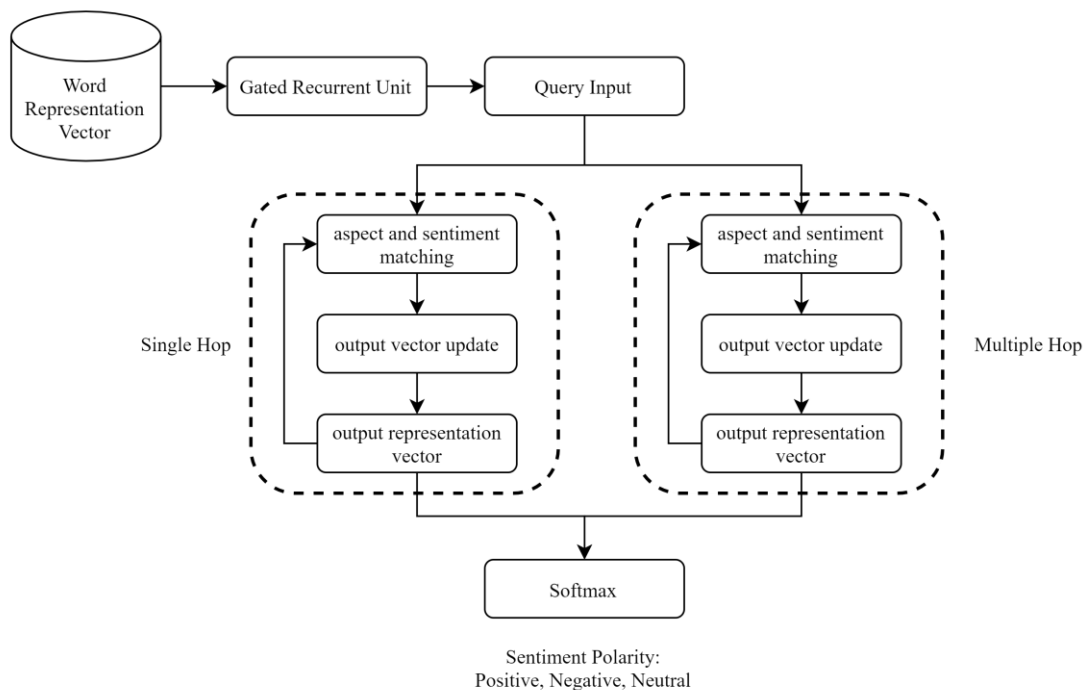**Sentiment Classification**

Figure 4. Memory network model scheme

Memory Network used as a learning model for the classification of sentiment on aspects of seeing the suitability of aspects to the context of the sentence, or the dependence of aspects on other aspects. The schematic model of the Memory Network can be seen in Figure 4. In the Memory Network, the aspect matching stages from the input query to getting the output representation update are called hops. For sentiment with multiple aspects, multi-hop is carried out in the learning process.

From the results of the process of representing aspects of the GRU sentence, it will be used as the initial memory input initiation in the memory module which will be stored and updated as memory output based on the Hop process carried out. Data representation of aspects of the GRU as memory input is formulated as follows:

$$S' = GRU(S) \tag{6}$$

S' is the result of a query from the representation process of learning aspects of the GRU context information. Then the modeling process with the Memory Network can be seen as follows:

$$b = aS \tag{7}$$
$$o = softmax(b) \tag{8}$$
$$r = OS' \tag{9}$$
$$a = a + r \tag{10}$$

The input representation is stored in the output memory as a query S, then aspect matching is performed on the data to get the response vector value, this value will be used as the next hop memory input.

For sentiment classification, for data with a single aspect, the input, matching, and output processes occur in one process or a single hop, while in data with multiple sentiments there is a multiple hop process. Softmax activation function as follows:

$$P = softmax((a + o)Wsmax) \tag{11}$$
$$y = argmax(P[i]) \tag{12}$$

Where W is the weight of the word, and y is the polarity of the sentiment, namely, positive, negative, and neutral. For multiple hops, a similar process is carried out where there is a looping update query in each aspect, and sentiment classification is carried out using the softmax activation function to determine the polarity of the sentiment consisting of, positive, negative, and neutral.

RESULT AND DISCUSSION

The trial uses a product review dataset from an Indonesian e-commerce site obtained using Puppeteer from Node.js with a total of 38071 data lines in the electronic domain. The snippet of the dataset used can be seen in Table 1, where there are 3 columns, namely the category column which is the type of product purchase, the Review text is a direct review of the product buyer, and the relevance score is the relevance of the review by the buyer on the product which means the relationship between the review and the product being purchased, judged by likes and other customer ratings.

Table 1. Dataset of Product Review

Category	Text Review	Relevance Score
Electronics	*Mantap... Sesuai tidak mengecewakan... Actual 127,sekian sekian... Akselerasi copy ok... Suksess selalu ya*	60
Electronics	*mantap tidak ada lecet... semoga awet aja... tp jelek nya gak di paking kayu pengiriman nya... tp TV nya aman tanpa ada rusak*	59.67
Electronics	*baru beli 2 bulan langsung rusak led nya ga nyala, tp suara nya keluar. mau klaim garansi gabisa karena data pelanggan tidak dikirim 14 hr setelah pengiriman dan jauh di jakarta*	62.5
Electronics	*alhamdulillah produk original dan sesuai spesifikasi semua*	64.33
Electronics	*makasih cepet sampenya cuman 2 hari sampe, semoga awet tv nya*	67.32

The relevance value in the review is in the range 0-100, the distribution graph of the data relevance value can be seen in Figure 5, from 38071 data, data lines are filtered with a relevance value above 59 which is the median of the data, the filter results get 1180 data to be used for the modeling and testing process.
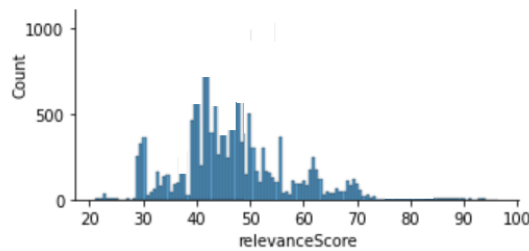

Figure 5. Relevancy score distribution

The dataset is then preprocessed to obtain uniform review text data, clean of characters other than review words. Then, collecting aspects and labeling the polarity values of the sentiment manually, the preprocessing stage of the data before it will be used for modeling can be seen in Table 2. Stemming and stopword processing, using the module from python library from the preprocessed dataset, text will be embedded using fastText in the data review, and aspects, the weight values used are derived from 300-dimensional fastText that has been trained using the Indonesian wikipedia corpus.

Table 2. Preprocessed Data

Stages	Before Preprocess	After Preprocess			
	Review	Review	Number of aspects	aspects	sentiment
Cleaning	Barang sesuai gambar dan pesanan ,barang bagus ,dan berfungsi dengan baik 👍 , cuma packing rada ngeri ngeri CUMA lapis bubble warp.. ☺	Barang sesuai gambar dan pesanan barang bagus dan berfungsi dengan baik cuma packing rada ngeri ngeri CUMA lapis bubble warp			
Casefolding	Barang sesuai gambar dan pesanan barang bagus dan berfungsi dengan baik cuma packing rada ngeri ngeri CUMA lapis bubble warp	barang sesuai gambar dan pesanan barang bagus dan berfungsi dengan baik cuma packing rada ngeri ngeri cuma lapis bubble warp			
Stemming	barang sesuai gambar dan pesanan barang bagus dan berfungsi dengan baik cuma packing rada ngeri ngeri cuma lapis bubble warp	barang sesuai gambar dan pesan barang bagus dan fungsi dengan baik cuma packing rada ngeri ngeri cuma lapis bubble warp			
Stopword	barang sesuai gambar dan pesan barang bagus dan fungsi dengan baik cuma packing rada ngeri ngeri cuma lapis bubble warp	barang sesuai gambar pesan barang bagus fungsi baik cuma packing rada ngeri ngeri cuma lapis bubble warp			
Deteksi aspek	barang sesuai gambar pesan barang bagus fungsi baik cuma packing rada ngeri ngeri cuma lapis bubble warp	barang sesuai gambar pesan barang bagus fungsi baik cuma packing rada ngeri ngeri cuma lapis bubble warp	2	['barang', 'packing']	
Labelisasi sentimen	barang sesuai gambar pesan barang bagus fungsi baik cuma packing rada ngeri ngeri cuma lapis bubble warp	barang sesuai gambar pesan barang bagus fungsi baik cuma packing rada ngeri ngeri cuma lapis bubble warp	2	['barang', 'packing']	['positive', 'negative']

The matrix from embedding process will be used as an initial representation for the distribution of aspects using the GRU. The distribution of aspects on the GRU uses input, output hyperparameters and Hop value limits which can be seen in Table 3.

Table 3. GRU hyperparameters

Hyperparameter	Value
Input Dimension	300
Output Dimension	300
Hop	10

From the value of the embedded results, the aspect vector value will be matched in the review sentence, then the GRU will study the context of the information from the review sentence to get a new matrix representation as input to the Memory Network. From the computational results of the GRU, the distribution of aspects, single aspect and multi aspect distribution of the data can be seen in Table 4.

Table 4. Aspect propagation

Aspect	Train	Test
Single Aspect	813	343
Multi Aspect	1922	312

In the input module, the matrix data representation will be forwarded to the Memory Network model, for each aspect of the review data, a response representation is obtained after updating the memory for one aspect, this process is carried out by a Hop loop until all aspects of the review sentence matched. Ending the hop process, sentiment classification is carried out on the softmax activation function with three polarities, positive, negative, and neutral.

The data is trained with the Memory Network model with a total initial trial iteration of 10 epochs, an early stop iteration is carried out in the 7th epoch because based on the loss and accuracy values of the training data and test data the loss value has increased towards the 8th epoch process, and the value of accuracy decreased in the same position, indicating that the error value increased in the 8th epoch, the graph of the loss and accuracy values can be seen in Figure 6.

Tests were carried out on 4 scenarios, first scenario sentiment classification trials without aspect analysis, then experiment the data with single aspect, the data with multi-aspect, and last by combining single aspects and multi-aspects to see the effect of aspect analysis on sentiment classification. From the test results, an evaluation of the matrix values of precision, recall, f1-score, and accuracy was carried out to see the performance of the Memory Network in sentiment classification in product reviews.
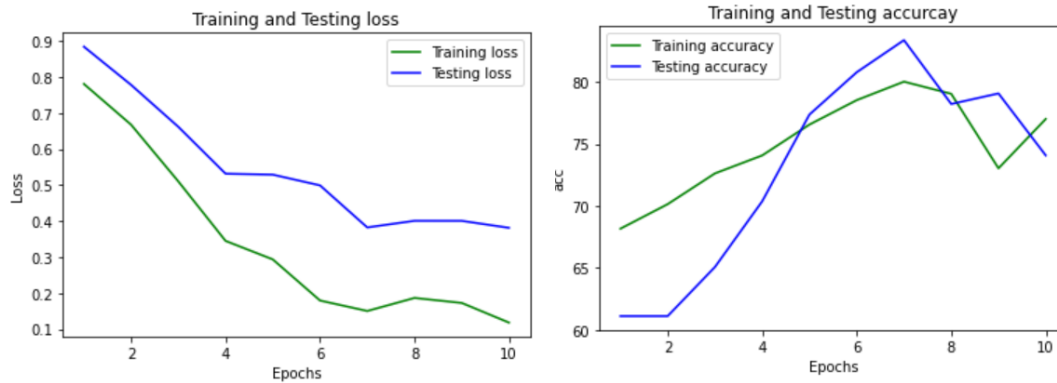

Figure 6. Graph of loss and accuracy

From the evaluation results, sentiment classification without aspect analysis has a matrix evaluation value of 64% precision, 52% recall, 55% f1-score, and 78% accuracy. In the single aspect evaluation, the evaluation value of the precision matrix is 67%, recall is 58%, f1-score is 49%, and accuracy is 78%. In the multi-aspect evaluation results, the evaluation value of the precision matrix is 64%, recall is 58%, f1-score is 49%, and accuracy is 73%. Finally, on the results of sentiment evaluation by combining single and multi-aspect aspects, the evaluation value of the precision matrix is 72%, recall is 64%, f1-score is 67%, and accuracy is 83%. The results of the evaluation matrix can be seen in Table 5.

Table 5. Matrix evaluation of the model

Data	precision	recall	F1-score	Accuracy
Document level	0.64	0.52	0.55	0.78
Single aspect	0.67	0.58	0.49	0.78
Multi aspect	0.64	0.58	0.49	0.73
Aspect based	0.72	0.64	0.67	0.83

From the accuracy results, the highest accuracy was obtained in the product review sentiment classification trial using the Memory Network on data involving aspect analysis, with an accuracy value of 83% and the matrix evaluation value outperforming sentiment classification without aspect, only single aspect, and multi-aspect.

CONCLUSION
Aspect-based sentiment analysis in this study has been carried out using the Gated Recurrent Unit (GRU) method as aspect extraction and the Memory Network on aspect-based product sentiment analysis has better results, compared to ordinary sentiment analysis with 83% and 78% accuracy values. In the comparison of evaluation matrices, aspect-based analysis also has better precision and recall values, both single-aspect and multi-aspect. This shows that the adjustment of weights on aspects by paying attention to the value of information in the context helps improve the performance of the sentiment classification prediction model in product reviews. The results of this study also prove that the assumption of using the Memory Network method in aspect-based sentiment analysis research has a good performance for learning and predicting multi-aspect sentiment classification on the data. There are still many improvements that can be made in the use of the aspect distribution detection method and the use of the Memory Network in aspect-based sentiment analysis, especially with the development of the deep neural network method which is getting better.

REFERENCES
[1] C. Tam, A. Loureiro, and T. Oliveira, "The individual performance outcome behind e-commerce," *Internet Res.*, vol. 30, no. 2, pp. 439–462, 2020.
[2] T. Atalay, "E-commerce ecosystem: Fundamental concepts, business models and cybersecurity issue," *Int. J. Bus. Ecosyst. Strateg.*, vol. 2, no. 2 SE-Articles, pp. 15–24, 2020.
[3] U. Rusmawan, "Analisis Faktor-Faktor Yang Mempengaruhi Minat Beli Produk Digital Secara Online," *J. Agribisnis*, vol. 11, no. 6, pp. 183–201, 2017.
[4] A. Zhao and Y. Yu, "Knowledge-enabled BERT for aspect-based sentiment analysis," *Knowledge-Based Syst.*, vol. 227, 2021.
[5] T. Mustaqim, K. Umam, and M. A. Muslim, "Twitter text mining for sentiment analysis on government's response to forest fires with vader lexicon polarity detection and k-nearest neighbor algorithm," *J. Phys. Conf. Ser.*, vol. 1567, no. 3, p. 32024, 2020.
[6] A. S. M. Alharbi and E. de Doncker, "Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information," *Cogn. Syst. Res.*, vol. 54, pp. 50–61, 2019.
[7] Y. Wang, M. Huang, L. Zhao, and X. Zhu, "Attention-based LSTM for aspect-level sentiment classification," *EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 606–615, 2016.
[8] A. Borg and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert Syst. Appl.*, vol. 162, p. 113746, 2020.
[9] N. Majumder, S. Poria, A. Gelbukh, M. Shad Akhtar, E. Cambria, and A. Ekbal, "IARM: Inter-aspect relation modeling with memory networks in aspect-based sentiment analysis," *Proc. 2018 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2018*, pp. 3402–3411, 2020.
[10] M. Usama, B. Ahmad, E. Song, M. S. Hossain, M. Alrashoud, and G. Muhammad, "Attention-based sentiment analysis using convolutional and recurrent neural network," *Futur. Gener. Comput. Syst.*, vol. 113, pp. 571–578, 2020.
[11] R. Hegde and S. Seema, "Aspect based feature extraction and sentiment classification of review data sets using Incremental machine learning algorithm," *Proc. 3rd IEEE Int. Conf. Adv. Electr. Electron. Information, Commun. Bio-Informatics, AEEICB 2017*, pp. 122–125, 2017.
[12] E. I. Setiawan, F. Ferry, J. Santoso, S. Sumpeno, K. Fujisawa, and M. H. Purnomo, "Bidirectional GRU for targeted aspect-based sentiment analysis based on character-enhanced token-embedding and multi-level attention," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 5, pp. 392–407, 2020.
[13] A. M. Suresha, "Sentiment Analysis on Amazon Product Reviews with Stacked Neural Networks,"

https://www.researchgate.net/publication/344518660_Sentiment_Analysis_on_Amazon_Product_Reviews_with_Stacked_Neural_Networks, 2020.

[14]   S. A. Azzahra and A. Wibowo, "Analisis Sentimen Multi-Aspek Berbasis Konversi Ikon Emosi dengan Algoritme Naïve Bayes untuk Ulasan Wisata Kuliner Pada Web Tripadvisor," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 4, pp. 737, 2020.

[15]   T. J. Pamungkas and A. Romadhony, "Analisis Sentimen Berbasis Aspek Terhadap Ulasan Restoran Berbahasa Indonesia menggunakan Support Vector Machines," *eProceedings Eng.*, vol. 8, no. 4, pp. 4102–4114, 2021.

[16]   D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification," *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 0, pp. 4068–4074, 2017.

[17]   M. M. Ashi, M. A. Siddiqui, and F. Nadeem, "Pre-trained Word Embeddings for Arabic Aspect-Based Sentiment Analysis of Airline Tweets," *Adv. Intell. Syst. Comput.*, vol. 845, pp. 241–251, 2019.

[18]   Z. Kastrati, A. S. Imran, and A. Kurti, "Weakly Supervised Framework for Aspect-Based Sentiment Analysis on Students' Reviews of MOOCs," *IEEE Access*, vol. 8, pp. 106799–106810, 2020.

[19]   Y. Kang and L. Zhou, "RubE: Rule-based methods for extracting product features from online consumer reviews," *Inf. Manag.*, vol. 54, no. 2, pp. 166–176, 2017.

[20]   D. H. Sasmita, A. F. Wicaksono, S. Louvan, and M. Adriani, "Unsupervised aspect-based sentiment analysis on Indonesian restaurant reviews," *Proc. 2017 Int. Conf. Asian Lang. Process. IALP 2017*, vol. 2018-Janua, pp. 383–386, 2018.

[21]   S. Gojali and M. L. Khodra, "Aspect based sentiment analysis for review rating prediction," *4th IGNITE Conf. 2016 Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2016*, 2016.

[22]   M. S. Akhtar, D. Gupta, A. Ekbal, and P. Bhattacharyya, "Feature selection and ensemble construction: A two-step method for aspect based sentiment analysis," *Knowledge-Based Syst.*, vol. 125, pp. 116–135, 2017.

[23]   O. Alqaryouti, N. Siyam, A. A. Monem, and K. Shaalan, "Aspect-based sentiment analysis using smart government review data," *Appl. Comput. Informatics*, 2019.

[24]   A. S. Manek, P. D. Shenoy, M. C. Mohan, and K. R. Venugopal, "Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier," *World Wide Web*, vol. 20, no. 2, pp. 135–154, 2017.

[25]   K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, vol. 281, no. 49, pp. 1724–1734, 2014.

[26]   A. Ilmania, Abdurrahman, S. Cahyawijaya, and A. Purwarianti, "Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis," *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, pp. 62–67, 2019.

[27]   Y. Xing and C. Xiao, "A GRU Model for Aspect Level Sentiment Analysis," *J. Phys. Conf. Ser.*, vol. 1302, no. 3, 2019.

[28]   J. Weston, S. Chopra, and A. Bordes, "Memory networks," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[29]   D. Tang, B. Qin, and T. Liu, "Aspect Level Sentiment Classification with Deep Memory Network(EMNLP2016)," 2014.

[30]   P. Chen, Z. Sun, L. Bing, and W. Yang, "Recurrent attention network on memory for aspect sentiment analysis," *EMNLP 2017 - Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 452–461, 2017.