International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)
Manuscript Received: 08 January 2022, Received in Revised form: 07 February 2022, Accepted: 12 February 2022 DOI: 10.46338/ijetae0322_15

WhatsApp Chatbot Customer Service Using Natural Language Processing and Support Vector Machine

Malvin[1], Constantine Dylan[2], Abdul Haris Rangkuti[3]
[1,2,3,4]*Informatics Department, School of Computer Science, Bina Nusantara University, Jakarta-11480 Indonesia*

***Abstract*— People have switched professions to become small business entrepreneurs during the COVID-19 pandemic. There are so many UMKMs born in 2020 – 2021 with little experience in how to do business. Moreover, there is a shortage of manpower, where young UMKMs are only run by one person due to the continuity with the pandemic, so owners do not have the time or energy to help answer customer's questions. One way to solve this problem is to use a Chatbot. Usually, a chatbot is used as a tool for humans to communicate by providing digital-based information, such as answering questions about the types of goods sold in a store. Chatbots are used because of their ability to interact with many individuals at once quickly and accurately. So that this research designs and develops a text classification system on chatbots using the Support Vector Machine. The purpose of developing this system is to simplify and speed up the proses of product information exchanges to customers. With an average response of less than 1 second and an average accuracy of 87.75% in classifying text, the chatbot is effective and efficient. The chatbot will classify the text and respond accordingly when given a message. So that the benefits of designing and developing this chatbot are convenience, speed, clarity in the process of customer and seller interaction.**

***Keywords*— Text Classification, Customer Service, Chatbot, Machine Learning, Support Vector Machine**

I. INTRODUCTION

The development of technology is deliberately made to make it easier for humans in all aspects of life. The rapid development of computer technology in the past few years have enabled humans to develop more and more complex systems and tasking jobs for the computers to handle [1] [12]. By using machine learning, humans can be facilitated through the automation process with one of the implementations used in everyday life is in the field of customer service where a bot can answer and provide information about an existing activity or program [2].

In Indonesia there is a lot of Micro, Small, Medium Enterprises (MSME) that have popped up in recent years with the lack funding therefor lack human resources due to the ongoing pandemic in the time of writing this paper resulted in the decreasing of the enterprise's performance and quality [10].

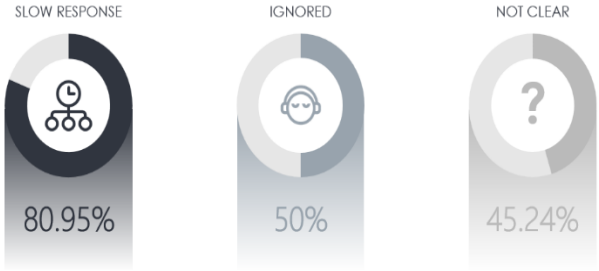

Figure 1.0 Customer difficulties when contacting MSMEs

Figure 1.0 illustrates that over 80% of customers experiences slow responses from MSMEs when contacting and asking about their products and services, 50% is ignored and 45.24% of answers were unclear of irrelevant from the questions. Which is why a chatbot system act as a customer service where it can answer users' question in a fraction of the time.

A well-defined customer service process includes the prompt attention and response to supplier-related questions [7]. With text classification and machine learning algorithm such as natural language processing is used to build models that can understands human language in its natural form. By applying the concept of artificial intelligence which has increased rapidly over time, natural language processing technology allows computers to understand human intention through text. Natural language processing models is used to help people in answering or solving problems regarding their relation to the business.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)

When customer ask for product information from a business, it is important that the trained model recognizes the customer's intention and sends the correct information quickly [7]. These types of customer services or chatbots are rapidly used because of its scaling and concurrent uses [2].

II. Related Study

This system development and research has several references from international journals regarding text classification.

In the first reference Andreas Stöckl builds a machine learning model for classifying inputs that are commonly obtained by chatbots. The sentences that become the training data are 111 sentences with 7 different categories. There are 4 algorithms tested for experiments in this book, namely, K-Nearest Neighbours, Naïve Bayes, Support Vector Machines, and Neural Networks. The research does not give a conclusion what algorithm is the best. [4]

The second reference, Vladimer B. Kobayashi, Stefan T. Mol, Hannah A. Berkers, Gabor Kismihok, and Deanne N. Den Hartog discusses about creating an automatic and accurate document classification system. The author of this journal explains the document classification process by dividing it into several steps, namely data preparation, pre-processing, transformation, application of classification techniques, and validation. To determine the category of a document, this journal uses 3 different models, namely SVM, Random Forest and Naïve Bayes, then selects the majority category from the predictions of these models. The results of the classification will be evaluated by experts in the category of the document.[5]

Based on these references, this research designs and develops a chatbot system based on text classification. This system is made using the Python programming language along with its libraries such as pandas, numpy, sci-kit learn or sklearn, literary, and others. After the categories of questions and the questions asked are known, the system will provide answers that have been prepared by the previous author.

TABLE I
Review and Comparison of Related Study

Reference	Data	Best Algorithm	Accuracy	Precision
Andreas Stöckl (2017)	111	Support Vector Machine	69.70%	-
Vladimer Kobayashi, et al. (2017)	1179	SVM, Random Forest and Naïve Bayes	-	89%
Constantine Dylan, Malvin	**640**	**Support Vector Machine**	**87.81%**	**90.65%**

Table 1 illustrates shows the accuracy and precision of each of our models. Andreas's SVM model that is trained with 111 data points has an accuracy of 69.70% and Vladimer's model using SVN, random forest and naïve bayes that is trained with 1179 data points is 89% precise. Our model using SVM has an accuracy of 87.81% and is 90.65% precise with 640 data points. Our model is better at this type of application which is chatbot system than the other related works.

III. Research Method

The system development method used in this study is the Waterfall Prototyping method which consists of 5 stages, that are communication (project initiation and requirements gathering), planning (estimation, scheduling, and tracking), modelling (analysis and design), construction (code and test), and development (delivery, feedback, and support). The research scheme in developing a customer service chatbot is divided into 5 main stages, starting from:

A. Sentence Classification determination

The development of this system uses 8 different sentence classifications that will be recognized by the system. The corpus that is used in this experiment came from the actual conversation of the owner of MSME and her customers via WhatsApp, which then manually labelled.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)

Each sentence or text that is sent has a different label that corresponds to different responses. The system will use Sci-kit Learn to extract message type from the user. The root of the word known as a lemma will be extracted from the sentences using information extraction process called Stemming. Every sentence will be trained with a Support Vector Machine with Python programming language along with other libraries such as Sci-kit Learn, Pandas, Sastrawi, and Fuzzywuzzy.

TABLE II
SENTENCE CLASSIFICATION THAT THE SYSTEM WILL RECOGNIZE

No	Data	Amount
1	buy	95
2	ask_product	113
3	ask_price	113
4	ask_delivery	68
5	ask_payment	61
6	ask_product_info	57
7	greetings	98
8	thank_you	35
Total		**640**

Table 2 informs that there are 8 different message classifications with the amount of data the classification has for training in the system. The first classification is "buy" where when the user sent a message about how to and buy some fruit with 95 sentences for the system to train. "ask_product", "ask_price", "ask_delivery" and "ask_payment" is when the user sent a message asking about the product, in this case fruits, the fruit price, the delivery and payments correspondingly. The "ask_product_info" classification is when the user sent a message asking about the information for a fruit type, such as "durian monthong" in the durian fruit category. The greetings classification is when the user greets the chatbot system. Last is the "thank_you" classification is when the user says thank you of goodbyes to the system. The total amount of data for the chatbot training is 640 sentences.

B. Data training using Support Vector Machine

Support Vector Machines (SVM) are supervised learning algorithms used for classification, regression, novelty, and anomaly or outlier detection [3]. SVM algorithm can be trained based on this data such that it can assign future data points into one of the two classes.

This algorithm represents the training data samples as points in space such that points belonging to either class can be separated by a wide gap between them, called a hyperplane, and the new data points to be predicted are assigned classes based on which side of this hyperplane they fall into. This process is for a typical linear classification process [8][10].

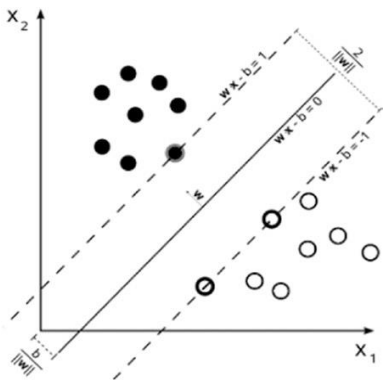

Figure 2.0 Hyperplane and support vectors separating 2 labelled data

Figure 2.0 illustrates that the SVM can separate the 2 labelled data efficiently by maximizing the margin between the data points to the hyperplane and creates the support vectors. SVM can also perform non-linear classification by an interesting approach known as a kernel trick, where kernel functions are used to operate on high-dimensional feature spaces that are non-linear separable. Usually, inner products between data points in the feature space help achieve this [3].

With SVM, the optimum result always looks for a greater margin from hyperplane to the support vector and reducing classifying error. But both of that contradicts with each other which is why there's a hyperparameter that needs to be entered to the system for model training. A hyperparameter will impact the result of the machine learning evaluation. A hyperparameter is a constant that is set to change one aspect of the algorithm. In this case the hyperparameter is C, where C controls the margin between the hyperplane and the support vector. With this hyperparameter in mind, the term SVM is now a C-Support Vector Classifier [8].

For text to work with C-SVC, it needs to be converted into digits that can be calculated. Using Term Frequency – Inverse Document Frequency or TFIDF, words can be converted from string into vector integer that tells how important a word is in a sentence and then stored in binary for it to be used for classification [6].

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)

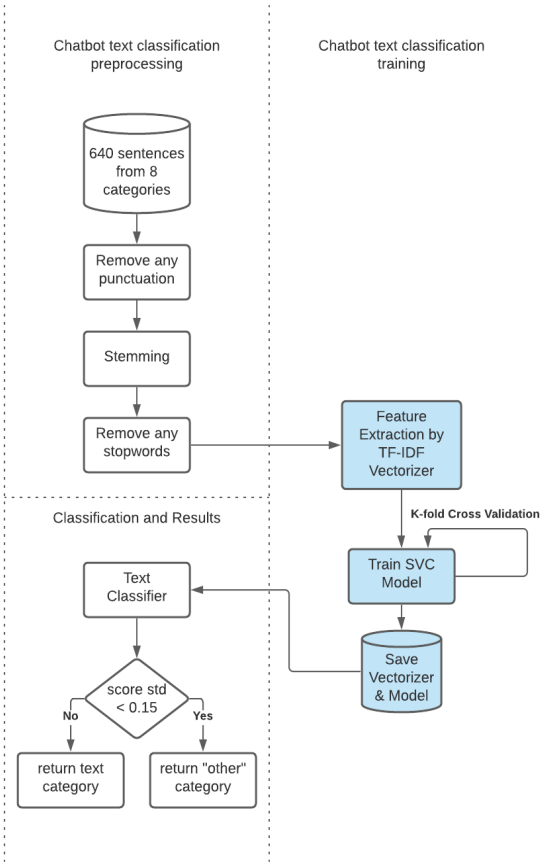

Figure 3.0 Data processing for classification prediction

Figure 3.0 illustrate the data pre-processing, model training and message classification process. Before the model can be trained, it needs clean data to produce a good model. The punctuations from 640 sentences with 8 categories or labelled will be removed and then the sentences will be stemmed resulted in only root stems in the sentences. Stop words are removed afterwards leaving behind the most important words in the sentence.

The remaining words are vectorized into vector integers using the TF-IDF vectorizer. The data will be stored and trained using SVC algorithm and will be validated using K-fold Cross Validation.

K-Fold Cross-Validation is a re-sampling procedure for developing Machine Learning models on limited sample data. This procedure has one parameter (k) which refers to the number of groups to be divided from a data sample. So if the value of k is 5, the number of data samples will be divided into 5 groups. One group will be used as data validation and the rest will be used as training data. After sharing the data, it will be entered into the Machine Learning model and the accuracy value will be calculated. The accuracy value can be calculated from the total correct predictions, divided by the total predictions. This process will be repeated k times with different data sharing. After the procedure is complete, all accuracy values will be averaged to make the final value [13]. The model is then used to classify user sent messages and responds accordingly.

C. Message Classification

The model is created after the training is complete. The system will recognize the sentences that is sent with the following outlines of the working message classification system as a customer service chatbot.

This system uses a C-Support Vector Classifier algorithm for classifying messages. There are 2 types of classifications in this process. The first is the message type classification and the fruit type classification.

Message type classification classifies the type of message the user sent to the system, either greetings, thank you messages or asking for information messages that is in Table 2. The model will receive a message from the user from the WhatsApp app then it will clean the message from punctuations and stop words to its root stem to be classified. The root stems will be put into the model which will classify the corresponding message with one of the sentence classifications in Table 2.

If the first classification outputs a “ask_product_info” then the second process classification begins. The fruit type classification will determine what type of fruit the user is asking about. For example, the user asks “jeruk apa saja yang ada” the first classification will output the “ask_product_info” type and then classifies the fruit that the user is asking which is “jeruk” or an orange. The 2 classifications will determine which response the model outputs to the user from its responses file.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)

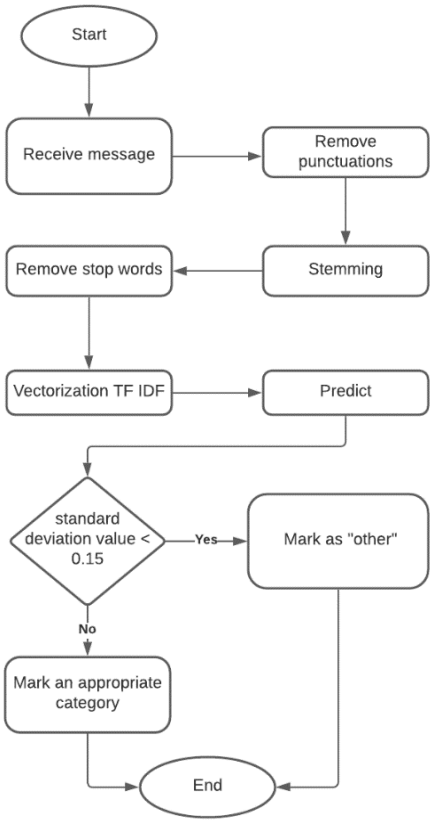

Figure 4.0 Message type classification flowchart

In Figure 4.0, the process starts from receiving the message from the user followed by removing punctuations and stemming to its root stem. After that the system will remove the stop words like "oh", "la", etc. The sentence will then be converted into vector integers by using TF-IDF that will be fed into the model to be predicted. The result is a sentence category that the model predicted and a standard deviation value. To be sure of the prediction's validity, the system ensures it with a standard deviation value below 0.15 for it to predict a particular sentence category. If the system marks a prediction with "other", the system will give out a "try again" message.

D. Data Retrieval using TwiML

TwiML is used to retrieve sent messages from customers. Although TwiML has its own built in machine learning model, for this use case the system only retrieves the text file from the user for the system to process it.

Using Flask micro-framework, we can turn on the local server and open a port to listen for Twilio webhook calls from WhatsApp sent messages.

E. Sending Text Message using TwiML

Using Flask as well from data retrieval, the system also uses it for sending messages using routes that is specified in the code and called automatically by the webhook to process the message that is sent by the user to then be retrieved for it to be sent back as an answer to the user that ask for the information in the first place.

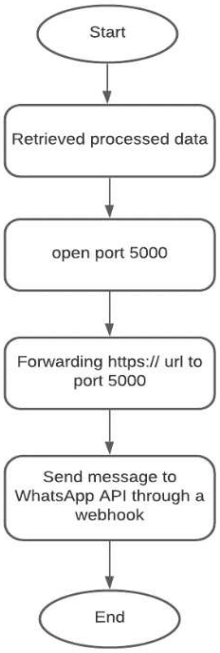

Figure 5.0 Message sending process flowchart

In Figure 5.0 the message sending process starts from the retrieved data from the prediction result of the model. After that the system will open up one of the port on the computer, in this case is port 5000 and an https URL will be forwarded to the port for it to listen to incoming messages through a webhook using WhatsApp API.

IV. Result And Discussion

After the training model is successfully carried out, a model with the .pk file formats will be generated. In this study, after the training is complete, a model accuracy table and best hyperparameter table using 3 kinds of algorithm to get the average accuracy value and standard deviation.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 12, Issue 03, March 2022)

There are Multinomial Naïve Bayes, C-Support Vector Classification and K-Nearest Neighbours. By using them would accept the suitable algorithm because to carry out the highest value. The other study could use the other algorithm to get the optimal of Chabot process.

TABLE III
TABLE EVALUATION OF DECIDING AN ALGORITHM

Model	Algorithm	Average Accuracy	Standard Deviation
Model A	*Multinomial Naïve Bayes*	83.91%	0.04
Model B	*C-Support Vector Classification*	86.25%	0.06
Model C	*K-Nearest Neighbors*	81.87%	0.03

Model A uses the multinomial naïve bayes which produces a model with an accuracy of 83.91%. Model B uses the c-support vector classification which produces a model with an accuracy of 86.25%. Model C uses the k-nearest neighbour algorithm which produces a model with an accuracy of 81.87%. From the result stated in Table III, C-Support Vector Classifier is the best algorithm for predicting a message for a chatbot system with an 86.25% in accuracy.

Hyperparameters needs to be chosen for this system, there are 4 kernel values, 5 C values, and 5 gamma values as it is shown below:

TABLE IV
PARAMETER VALUES

Parameters	**Values**
Kernel	"linear", "poly", "rbf", "sigmoid"
C	0.1, 1, 10, 100, 1000
gamma	1, 0.1, 0.01, 0.001, 0.0001

The first hyperparameter is "kernel", "kernel" is used for choosing the hyperplane type that will separate the data. "kernel" has 2 common types, "linear" and "non-linear" yet all types include "rbf" and "sigmoid". "gamma" is a hyperparameter for non-linear hyperplanes, the greater the value the more overfit the model becomes.

"C" value is used to control the trade-off between smooth decision boundaries and classifies the correct training point [8].

TABLE V
CHOSEN PARAMETER VALUES

Parameters	**Values**
Kernel	"linear"
C	1
gamma	1

Table V illustrates that the optimum hyperparameters for this chatbot system is as follows, "linear" for the "kernel" parameter and the value of "1" for the "C" and "gamma" hyperparameters.

After further evaluation, it concludes that the best values for corresponding hyperparameters are "linear" for the kernel hyperparameter, 1 for C, and 1 for the gamma. Applying the chosen the hyperparameter values to the model training, the model increases its accuracy from 86.25% to 87.81% with a precision of 90.64%

By also using a chatbot system, UMKM can now shorten the time to respond to customers significantly with an average manual response of 18.55 minutes (1113 seconds) to 0.8 seconds automatically using the chatbot system.

TABLE VI
RESPONSE TIME EVALUATION RESULT

Response Type	Average Response Time (seconds)	Standard Deviation	Median (seconds)
Manual	1113	2375.19	240
Chatbot	0.8	0.5231	1

Table VI shows that the system has an average response time of 0.8 seconds which is 18.55 minutes. The system is 1391 times faster than if the UMKM owner responded to a customer.

V. CONCLUSION

1. The model development for this chatbot system for customer service uses the Support Vector Machine and has an accuracy of 87.81% and a precision of 90.65%. This evaluation value can be considered good when compared to other models in the reference

2. The development of this chatbot system for customer service has fulfilled the objectives of this research. The first objective, to simplify the process of interaction between buyers and sellers, has been fulfilled if one looks at the results of user evaluations. 79% of respondents agree that this system facilitates interaction between buyers and sellers. The second objective, accelerating the process of providing information, has also been fulfilled if you look at the results of the system evaluation. The system can cut the response time from 18.55 minutes to 0.8 seconds.
3. The development of this chatbot system for customer service has fulfilled the benefits of this research. These benefits are to simplify the process of obtaining information and speed up the process of providing information.
4. The performance of this chatbot model depends on the amount and variety of data that becomes training data. The only words that can be understood by this system are those in the training data.
5. System errors and failures in answering a question can be traced to several factors. The first factor is the presence of words that are not in the training data. These words can be in the form of abbreviations, foreign languages, or typos. Another factor is the number of words in a question. The ambiguity of questions is also a factor in the errors and failures of this system because the system is made to answer only one category.

REFERENCES

[1] Adamopoulou, E., & Moussiades, L. (2020a). An Overview of Chatbot Technology. IFIP Advances in Information and Communication Technology. https://doi.org/10.1007/978-3-030-49186-4_31

[2] Rahman, A. M., Mamun, A. A., & Islam, A. (2017). Programming challenges of chatbot: Current and future prospective. 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). https://doi.org/10.1109/r10-htc.2017.8288910

[3] Sarkar, D. (2016). Text Analytics with Python: A Practical Real-World Approach to Gaining Actionable Insights from your Data (1st ed.). Apress.

[4] Stöckl, A. (2017). Classification of chatbot inputs. En ligne https://www. researchgate. net/profile/Andreas_Stoeckl2/publication/318661551 _Classification_of_Chatbot_Inputs/links/597642e8a6fdcc8348aa52e4/Classification-of-Chatbot-Inputs. pdf.

[5] Korde, V. (2012). Text Classification and Classifiers:A Survey. International Journal of Artificial Intelligence & Applications, 3(2), 85–99. https://doi.org/10.5121/ijaia.2012.3208

[6] Saadah, M. N., Atmagi, R. W., Rahayu, D. S., & Arifin, A. Z. (2013). SISTEM TEMU KEMBALI DOKUMEN TEKS DENGAN PEMBOBOTAN TF-IDF DAN LCS. JUTI: Jurnal Ilmiah Teknologi Informasi, 11(1), 19. https://doi.org/10.12962/j24068535.v11i1.a16

[7] Doxey, C. H. (2021). The New Accounts Payable Toolkit. The New Accounts Payable Toolkit. https://doi.org/10.1002/9781119700548

[8] Valkov, V. (2020). Hacker's Guide to Machine Learning with Python: Hands-on guide to solving real-world Machine Learning problems with Deep Neural Networks using Scikit-Learn, TensorFlow 2, and Keras [E-book].

[9] Mandot, P. (2018, August 28). What is the Significance of C value in Support Vector Machine? Medium. Retrieved January 30, 2022, from https://medium.com/@pushkarmandot/what-is-the-significance-of-c-value-in-support-vector-machine-28224e852c5a

[10] Annur, C. M. (2021, October 14). Kekurangan Biaya, Penyebab Utama UMKM Tutup Usaha saat Pandemi Covid-19. katadata. Retrieved December 17, 2021, from https://databoks.katadata.co.id/datapublish/2021/10/14/kekurangan-biaya-penyebab-utama-umkm-tutup-usaha-saat-pandemi-covid-19

[11] Fraj, M. B. (2018, August 14). In Depth: Parameter tuning for SVC - All things AI. Medium. Retrieved January 14, 2022, from https://medium.com/all-things-ai/in-depth-parameter-tuning-for-svc-758215394769

[12] Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems (2nd ed.). O'Reilly Media.

[13] Berrar, D. (2019). Cross-Validation. Encyclopedia of Bioinformatics and Computational Biology, 542–545. https://doi.org/10.1016/b978-0-12-809633-8.20349-x