PRODAMAS chatbot: A flask and support vector machine based implementation

Riefyal Arshyzha Mustain¹, Muhammad Muhajir²*, Pandri Ferdias³, Nurirwan Saputra⁴

¹,² Department of Statistics, Universitas Islam Indonesia
³ Department of Mathematics, Universitas Lampung
⁴ Department of Informatics, Universitas PGRI Yogyakarta

ABSTRACT

In accelerated and equitable development in the Kediri City, the Kediri City Government launched the Community Empowerment Program called Prodamas. Prodamas aims to develop and encourage community participation in development at the Neighbourhood Level. To increase the dissemination of information about Prodamas, digital technology can be used as an information service provider. One of them is Chatbot. To develop Chatbots, Natural Language Processing, which is a branch of Artificial Intelligence, has become the most frequently used computer program. This Prodamas chatbot development uses the pattern matching method as an answer search algorithm and Support Vector Machine (SVM) classification as a method to see the machine's level of accuracy in answering questions given by users. Furthermore, the chatbot will be connected to WhatsApp so that it is expected to be able to provide and provide information about Prodamas. The results of testing the chatbot response with new questions provide an accuracy of 79%. Then testing the classification of the new question text with SVM. Obtained an accuracy of 88% with a precision value of 91% and a recall of 88%.

INTRODUCTION

Kediri, the third-largest city in East Java, is a modest yet prosperous metropolis with enormous potential across the board. In order to spur economic development, the Kediri City Government has been steadily upgrading the city's infrastructure (Pemerintah Kota Kediri, 2018). According to C.B.S. data, Kediri's city economy grew faster than the country's overall economy between 2016 and 2019 (Badan Pusat Statistika Kota Kediri, 2020).

The Community Empowerment Program is an initiative of the Kediri City Government aimed at promoting rapid and equitable growth in the Kediri City neighborhood (Prodamas). Aiming to foster and stimulate community participation in development, Pro-damas (also known as the 50 million per RT program) was established. Community
feedback was gathered through online forums in each RT, and those ideas were used to shape the final product. After that, the money will be put to good use improving physical and social infrastructure. Kediri’s municipal government has decided to take action in 2019 to maximize the benefits of Prodamas by launching Prodamas Plus, a program with a budget of Rp100 billion ($1 million) that focuses on improving the city’s infrastructure, culture, economy, health care, and youth opportunities (Pemerintah Kota Kediri, 2020).

There has been a recent surge in digital activity, which some have labeled a disruptive era. Being prepared for the technical advances of the future requires, in a roundabout way, that people today be adept at using electronic gadgets. To realize good governance, or the administration of government that is effective, efficient, transparent, and responsible, the City of Kediri is among those prepared to deal with technological advancements by creating a technologically-based integrated service system (Dinas Komunikasi dan Informatika Kota Kediri, 2022).

There are a wide variety of situations that call for someone to be technologically inventive. The government can benefit from digital technology in the realm of public services in areas where engagement with the public is required. PricewaterhouseCoopers (PwC) published a paper titled Future of Government, in which they claimed that the government has improved public services with the use of digital technology and is able to adapt and develop to suit the demands of the people. In addition, digital technology provides new ways for people to collaborate on the development of approaches, designs, and the provision of services that address local needs (Pricewaterhouse Coopers, 2013). The employment of chatbots is an example of how technology can be applied to government work.

A chatbot, sometimes known as a chatterbot, is a type of computer program that mimics human conversation by providing predetermined responses to user-entered keywords. Natural language processing is commonly used in the development of chatbot technology (NLP). Natural language processing (NLP) is a subfield of AI concerned with translating between human and computer speech. This allows websites or chat apps like WhatsApp, Line, Telegram, and others to implement chatbots that may simulate human conversations in a natural way through text, voice, or both (Adamopoulou & Moussiades, 2020; Al Maksur & Muhajir, 2021; Juhn & Liu, 2020).

The Kediri City Prodamas WhatsApp bot is an example of how natural language processing (NLP) technology is being put to use. WhatsApp is the most widely used messaging service in Indonesia right now. Whatsapp’s average monthly usage in Indonesia in 2021 was 31.4 hours, as reported by the Kepios and We Are Social reports in January 2022 (Data Reportal, 2022).

When the WhatsApp bot code is complete, it will be put to the test to see if it can actually do its intended job of responding to users’ inquiries. Categorization is one of the outcomes of natural language processing. Predictions can be made based on the characteristics of a predefined data class using a classification approach. Classification is utilized to evaluate the chatbot’s response quality in this example.

Support Vector Machine is a popular tool for analyzing and categorizing textual information especially on WhatsApp bot implementations because it has several advantages including Good class separation, Good performance with unstructured data, Overfitting tolerance, Support for high-dimensional feature spaces, and Computational efficiency.
The goal of this research was to develop a chatbot system for the Prodamas programming language that could be used with the popular messaging app WhatsApp. You can use the bot to learn more about Kediri and Prodamas. In addition to its primary function, the chatbot serves as a suggestion box for bettering the city of Kediri's Prodamas implementation. The purpose of this chatbot is to help users learn more about Prodamas and the City of Kediri by facilitating conversation and providing answers to their queries; while it cannot fully replace a human, it is hoped that it will be very helpful in these regards.

METHOD

Dataset

This study has two datasets. Secondary data about Kediri and Prodamas is the first dataset. The 2020 Kediri City Government documents and Kediri Mayor Regulation on Prodamas technical implementation provided the data. The City of Kediri and Prodamas Plus websites provide additional information. The City of Kediri's official website, https://www.kedirikota.go.id/, provided further data. The second dataset is primary data: user-submitted queries concerning Kediri and Prodamas obtained by researchers when testing system replies. In this study, the first data is used to develop a chatbot as a reference or database for the chatbot system, and the second data is utilized to test the chatbot’s question-answering abilities.

Chatbot System Description

Natural Language Processing is used in the chatbot design, along with the string matching approach, which will be implemented by the WhatsApp platform. A string matching algorithm compares two strings in search of a match. The matching process is carried out by the string matching algorithm using a string that is similar to the input string or a writing error occurs, namely the string that is matched for the degree of arrangement similarity different characters which the strings share (Ng et al., 2020; Paul et al., 2019; Qaffas, 2019).

This chatbot system was designed to make it simpler for the public to learn about Prodamas and the City of Kediri (Daniel et al., 2020). In the chatbot, there are two primary menus: City of Kediri and Prodamas. There is a sub-menu in this menu that includes the necessary information. In this scenario, chatbots are used to provide information about the meaning, laws, fields, and other aspects of Prodamas.

The acquired data is used to create a chatbot system as well as a knowledge base. Table 1 is a description of the information contained in the knowledge base:

<table>
<thead>
<tr>
<th>No</th>
<th>Input</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>'hello'</td>
<td><img src="image" alt="Chatbot System Description" /></td>
</tr>
<tr>
<td>No</td>
<td>Input</td>
<td>Response</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td>'menu'</td>
<td><img src="image1.png" alt="Image" /> 'menu' button clicked</td>
</tr>
<tr>
<td>3</td>
<td>'END'</td>
<td><img src="image2.png" alt="Image" /> 'END' button clicked</td>
</tr>
<tr>
<td>4</td>
<td>'1'</td>
<td><img src="image3.png" alt="Image" /> '1' button clicked</td>
</tr>
<tr>
<td>5</td>
<td>'2'</td>
<td><img src="image4.png" alt="Image" /> '2' button clicked</td>
</tr>
<tr>
<td>6</td>
<td>'profil kota'</td>
<td><img src="image5.png" alt="Image" /> 'profil kota' button clicked</td>
</tr>
<tr>
<td>7</td>
<td>'wilayah kota'</td>
<td><img src="image6.png" alt="Image" /> 'wilayah kota' button clicked</td>
</tr>
<tr>
<td>8</td>
<td>'brand kota'</td>
<td><img src="image7.png" alt="Image" /> 'brand kota' button clicked</td>
</tr>
<tr>
<td>No</td>
<td>Input</td>
<td>Response</td>
</tr>
<tr>
<td>----</td>
<td>----------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>9</td>
<td>'program unggulan'</td>
<td><img src="image1" alt="Image 1" /></td>
</tr>
<tr>
<td>10</td>
<td>'pengertian'</td>
<td><img src="image2" alt="Image 2" /></td>
</tr>
<tr>
<td>11</td>
<td>‘regulasi’</td>
<td><img src="image3" alt="Image 3" /></td>
</tr>
<tr>
<td>12</td>
<td>‘bidang’</td>
<td><img src="image4" alt="Image 4" /></td>
</tr>
<tr>
<td>13</td>
<td>‘kampung keren’</td>
<td><img src="image5" alt="Image 5" /></td>
</tr>
<tr>
<td>14</td>
<td>‘kube’</td>
<td><img src="image6" alt="Image 6" /></td>
</tr>
<tr>
<td>15</td>
<td>‘bank sampah’</td>
<td><img src="image7" alt="Image 7" /></td>
</tr>
</tbody>
</table>
Chatbot System Diagram
The system is divided into three major components: input, input processing, and output. Users can initiate a discussion with the chatbot by saying ‘hi’ to it. The input will then be evaluated and searched for patterns that fit the chatbot’s knowledge base. If the algorithm has discovered the correct pattern, the bot will seek replies from that input (Oguntosin & Olomo, 2021). The bot will then react with the proper answer after the solution has been located, chatbot system flow can be seen in Figure 1.

Figure 1. Chatbot system flow

Support Vector Machine
Support Vector Machine (SVM) is a classification analysis that classifies using a field separator (hyperplane). Any hyperplane in a space of dimension \( p-1 \) will also have dimension \( p-1 \). A wide range of hyperplanes can be generated via SVM. The most distinguishing field with the greatest margin is the one to use (maximal margin hyperplane). The margin is the average distance detected between the training data and the Hyperplane. The closest observations to the hyperplane are referred to as support vectors (Cervantes et al., 2020; Saraswati & Indradewi, 2022; Styawati et al., 2022).

Using kernel functions, SVM can be adapted to work with non-linear text data. To put it simply, the kernel is a function that transforms data features from their original (lower) dimensions to higher features (even much higher). This method deviates from the standard classification procedure by avoiding the reduction of the starting dimension in favor of a more straightforward computational process and improved prediction accuracy (Gowthul Alam & Baulkani, 2019; Wang et al., 2021). Here’s an illustration for \( n \) data samples: \(((\phi(x_1), y_1); (\phi(x_2), y_2);\ldots; (\phi(x_n), y_n)\) dot product of two vectors \((x_i)\) and \((x_j)\) are denoted as \(\phi(x_i) \cdot \phi(x_i)\). Without knowing the transformation function \(phi\), the dot product value can be computed using the components of the two vectors in the origin dimensional space, as shown in Equation (1) (Dewi & Chen, 2019; Ghosh et al., 2019).

\[
K(x_i , x_t) = \phi(x_i) \cdot \phi(x_t) \quad (1)
\]

\(K(x_i , x_t)\) is a kernel function with a non-linear mapping in feature space. The following are the dataset predictions using the newly developed features, as shown in Equation (2):

\[
f(\phi(x)) = \text{sign}(w \cdot \phi(x_t) + b)
\]
\[
= \text{sign} \left( \sum_{i=1}^{nS} \alpha_i y_i \phi(x_i) \cdot \phi(x_t) + b \right)
\]
\[
= \text{sign} \left( \sum_{i=1}^{nS} \alpha_i y_i K(x_i , x_t) + b \right) \quad (2)
\]

With:
\(nS\) = The amount of data that is a support vector
\(x_i\) = Support Vector
\(x_t\) = Predicted Testing data
The algorithm for PRODAMAS Vector Machine (SVM) can be seen in WhatsAppbot using Flask and Support Figure 2.

RESULTS AND DISCUSSION
Chatbot System Implementation
This chatbot program works as a machine that answers WhatsApp user questions about Prodamas and the City of Kediri. The source code snippet in Figure 3 is used to create a chatbot program in Python that is connected to Flask, Twilio, and Ngrok for receiving, processing, and replying to WhatsApp messages.
WhatsApp bot application features two primary menus, City of Kediri and Prodamas, which will develop into lessons or modules. There is a specific section devoted to Kediri and Prodamas in each menu. Table 2 provides an overview of the available menus and their associated information.

![Source code snippet](image)

**Table 2. List of Chatbot Menus and Sub-menus**

<table>
<thead>
<tr>
<th>Menu</th>
<th>Sub-Menu</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEDIRI</td>
<td>Profil Kota</td>
<td>Provides a quick overview of the City of Kediri</td>
</tr>
<tr>
<td></td>
<td>Wilayah Kota</td>
<td>This page contains information about the administrative area of the City of Kediri</td>
</tr>
<tr>
<td></td>
<td>Brand Kota</td>
<td>Providing Kediri with information about municipal branding</td>
</tr>
<tr>
<td></td>
<td>Program Unggulan</td>
<td>Providing information on the leading programs of the Kediri City Government</td>
</tr>
<tr>
<td>PRODAMAS</td>
<td>Pengertian</td>
<td>Give some background information on Prodamas</td>
</tr>
<tr>
<td></td>
<td>Regulasi</td>
<td>Provide basic legal information about Prodamas implementation</td>
</tr>
<tr>
<td></td>
<td>Bidang</td>
<td>Prodamas Plus’s areas of focus are listed here</td>
</tr>
<tr>
<td></td>
<td>Kampung Keren</td>
<td>Provide information and a list of Kediri’s Cool Villages</td>
</tr>
<tr>
<td></td>
<td>KUBE</td>
<td>Give information on KUBE Kediri</td>
</tr>
<tr>
<td></td>
<td>bank Sampah</td>
<td>Give details about the Kediri Garbage E-Bank</td>
</tr>
</tbody>
</table>

**Observing the Reaction Time of a Chatbot System**

Functional testing is performed to determine whether the process of operating the Prodamas Plus chatbot system is working properly and capable of providing a suitable answer or reply. Chatbot testing makes use of input that is comparable to knowledge base word input. WhatsApp bot users provided 54 queries. The answers to all 54 exam questions are available at [https://bit.ly/userquestion](https://bit.ly/userquestion).

Based on the results of evaluating the response of the present system, it is clear that the bot is functionally sound since it has been able to receive and reply to messages provided to the system. Of the 54 queries asked by the user, 43 were successfully replied with the information required, while 11 were answered with error warnings. As a result, the achieved accuracy value is 79.76%.

The Prodamas category makes up 56% of the 54 questions. The remaining 44% are in the Kediri category. In other words, there are 30 Prodamas questions and 24 City of Kediri questions. The question’s words are visualized using wordcloud in Figure 4. The visualization results show larger words with more words.

**Figure 4. WordCloud visualization**

Figure 4 displays the City of Kediri word cloud in blue and the Prodamas question in red. “Program,” “City,” “Kediri,” “Pemkot,” “Region,” and “Info” occur most commonly in the Kediri category. Prodamas, kediri, kota, kube, village, and cool are common terms in the category.
Classification

Training and testing data are separated before classification. Data testing tests classification models, while training data studies patterns and builds them. The system randomizes data 70:30. 37 data will be utilized for training, while 17 data will be used for testing. Support Vector Machines classified question texts in this study. The system creates an SVM classification model using identified training data. Predicting testing data will test the model. SVM classification results can be seen in Figure 5.

Figure 5. Confusion matrix of svm

The confusion matrix Figure 5. calculations are carried out to determine the accuracy value. An accuracy value of 88.24% was obtained where 7 data predicted as Kediri category questions were classified correctly as Kediri and not Kediri category question data predicted as Prodamas. Then, there are 8 data in the Prodamas category which are correctly predicted as Prodamas and 2 data in the Prodamas category which are predicted to be Kediri.

CONCLUSIONS AND SUGGESTIONS

The WhatsApp chatbot system can receive and answer to questions that are identical to or not identical to those in the knowledge base. The algorithm answered 54 questions with an accuracy of 79.63%. The chatbot program classified user questions with an accuracy of 88.24%, precision of 91%, and recall of 88% based on the Support Vector Machine classification findings.

Researchers can improve the accuracy of SVM in future research by exploring various techniques and approaches. Several studies have proposed methods to improve SVM accuracy, such as feature selection, kernel selection, and parameter optimization. Additionally, combining ensemble techniques such as bagging or boosting with SVM can also result in performance gains.

REFERENCES


https://diskominfo.kedirikota.go.id/information/news/671


https://doi.org/10.30812/matrik.v2i1.1848


https://doi.org/10.1080/24751839.2018.1558378

https://www.kedirikota.go.id/page/kota-kediri


