JPPIPA 10(Special Issue) (2024)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Natural Language Processing (NLP) Technology for Chatbot Website

Agreianti Puspitasari[1*], Astrid Noviana Paradhita[1], Yohanes Wien Tineka[1], Vivin Sulistyowati[1], Ni Komang Septia Noriska[1], Haryanto[1]

[1] Sebelas Maret University, Indonesia.

Received: February 28, 2024
Revised: June 23, 2024
Accepted: August 25, 2024
Published: August 31, 2024

Corresponding Author:
Agreianti Puspitasari
agreianti@staff.uns.ac.id

DOI: 10.29303/jppipa.v10iSpecialIssue.8241

Abstract: Chatbot technology is one form of application of Natural Language Processing, NLP is one of the branches of science Artificial Intelligence that studying communication between humans and computers through natural language. the purpose of this research will be carried out development of a chatbot application for information about the corporate, making it easier for visitors to find the information needed. This application built with a Natural Language Processing approach. This chatbot application uses medote text mining as a medium of reasoning. System modelling used for this application is using the Unified Model Language. The chatbot that will be developed is a website chatbot. The website chatbot was chosen by the company because one of the platforms that can be used well for export marketing is the website. This chatbot development uses Pyton, JavaScrib, Object notion (Json), Html + CSS. The application used in coding is Virtual Studio Code. Pyton is used to train a chatbot model that uses Natural Language Procecing (NLP). This Chatbot model is trained with the Neural Network Algorithm method, which is an algorithm model that is close to the human brain, which can provide stimulation or stimulation, process, and provide output to find relationships between data sets. JavaScrib is used to create chatbot data sets.

Keywords: Chatbot; Natural language processing; Website

Introduction

Services based on information technology in work will be able to provide convenience in job completion. One form of advances in information technology is the chatbot application (Anagnoste et al., 2021). A chatbot is a computer programme that is designed to simulate a conversation or interactive communication to users (humans) either through text, sound or visuals (Adamopoulou & Moussiades, 2020; Ahmad et al., 2018; Harahap & Fitria, 2020).

A chatbot acts as a conversational agent that can assist or replace the role of a consultant. A chatbot has a knowledge base that can be used to have conversations with customers. Chatbot technology is one form of application of Natural Language Processing, NLP is one of the branches of science Artificial Intelligence that studying communication between humans and computers through natural language (Hormansyah & Utama, 2018).

As a product of technology, NLP and its construction in chatbot platforms are widely used in various human interests, including in this discourse the use for sales. Many businesses currently use chatbot artificial intelligence to support company work, one of which is in the marketing sector. Some research related to chatbot development in business is research on chatbot design that is used to improve business performance by recording orders, order processing, customer records, business locations, and financial transactions (Amalia & Wibowo, 2019).

How to Cite:
Puspitasari, A., Paradhita, A. N., Tineka, Y. W., Sulistyowati, V., Noriska, N. K. S., & Haryanto. (2024). Natural Language Processing (NLP) Technology for Chatbot Website. *Jurnal Penelitian Pendidikan IPA*, *10*(SpecialIssue), 319–324. https://doi.org/10.29303/jppipa.v10iSpecialIssue.8241

Chatbot is also used as a service improvement, one of which is service in hotels. Research results show, from a prototype application developed to improve hotel services, chatbots help personalise services to customers, can accommodate customer needs and as additional facilities from hotels (Putri et al., 2019).

Figure 1. Illustration of NLP (Source: infokomputer.grid.id)

Chatbots are also used to help market agricultural products. The research system developed is a chatbot to connect farmers and consumers, so that farmers will get a good and profitable selling price and consumers can get fresh agricultural products (Kiruthika et al., 2023). In this research, we will also develop a chatbot that supports activities in marketing spice export products at Spicering Rempah Indonesia Ltd. The development of technology in this day and age has has helped humans a lot in carrying out daily activities. This can be seen with the emergence of applications or programmes that really help humans in various aspects of life (Putra, 2020).

Method

The research method using actual and factual situation analysis, as well as considerations based on literature. In this case, it will be possible to find the latest problems on the research topic raised this time, with the literature review method it will get something knowledge that can add insight for researchers. The method used in this research descriptive qualitative research is literature review and actual and factual situation analysis (Sugiono, 2021).

In this second stage by finding a problem, the problem raised in this research is how to implement a smart consumer loyalty system with the chatbot based on web method and internet media. The initial stage in Natural Language Processing (NLP) is to do Text Preprocessing first. Text Preprocessing is the first step in the model building process (Sihombing, 2022).

In this last section, the author conducts research based on data that has been collected through previous research studies, and researchers conduct research so that the reset can produce data and the data can be a proposal for a new system, which will be used in the future by conducting research. Then the problem raised can be solved and a solution can be found. The study used unobtrusive research techniques to analyze objectively the impact of AI (Mhlanga, 2020).

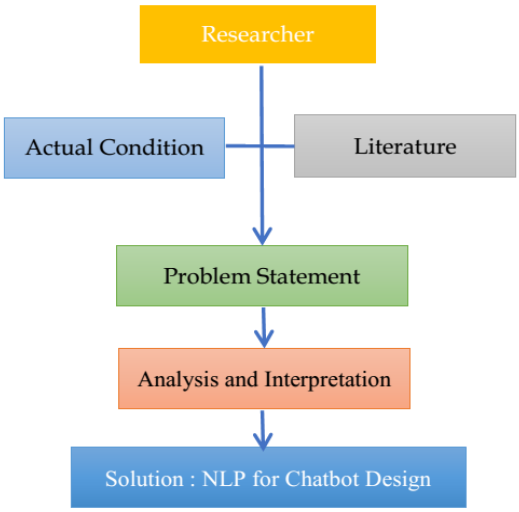

Figure 2. Research logic flow

This research uses qualitative methods by reading previous research, by exploring journals that conduct similar research and reading many books that discuss research like this, qualitative methods by also conducting discussions with various parties who understand this research. And then the author will be design the models as follow Figure 3.

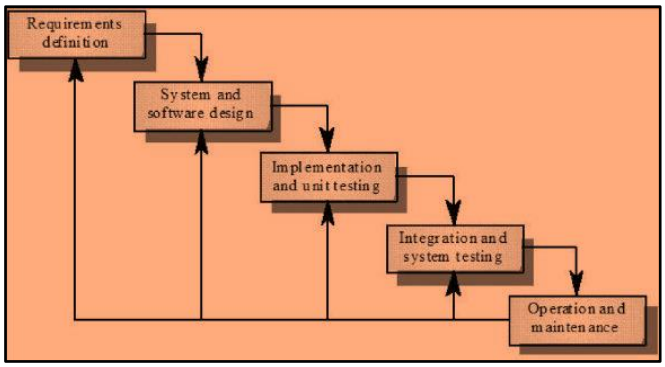

Figure 3. Waterfall model

Result and Discussion

The NLP Term

Natural language processing (NLP) is a machine learning technology that gives computers the ability to interpret, manipulate, and understand human language (Chowdhary, 2020; Fanni et al., 2023). Many organizations today have a vast amount of voice and text data from various communication channels such as

emails, text messages, social media news feeds, video, audio, and more.

Such organizations use NLP software to automatically process this data, analyse the intent or sentiment in messages, and respond to human communication in real time. Natural language processing combines computational linguistics, machine learning and deep learning models to process human language.

Computational Linguistics

Computational linguistics is the science of understanding and building models of human language with computer tools and software. Researchers use computational linguistics methods, such as syntactic and semantic analysis, to create frameworks that help machines understand human language used in conversation. Tools such as language translators, text-to-speech synthesizers, and speech recognition software are based on computational linguistics.

Machine Learning

Machine learning is a technology that trains computers with sample data to improve their efficiency. Human language has a number of features such as sarcasm, metaphors, variations in sentence structure, plus grammar and usage exceptions that take humans years to learn. Programmers use machine learning methods to teach NLP applications to accurately recognise and understand these features right from the start.

Deep Learning

Deep learning is a specific field of machine learning that teaches computers to learn and think like humans. Deep learning involves neural networks consisting of data processing nodes that resemble the operation of the human brain. With deep learning, computers recognise, classify, and correlate complex patterns in input data. When introducing deep learning into the asset management, there are major issues to be aware of (Kato, 2020).

NLP Implementation Steps

Typically, the NLP process starts with collecting and preparing unstructured text or speech data from sources such as cloud data warehouses, surveys, emails, or internal business process applications.

Preprocessing

NLP software uses preprocessing techniques such as tokenisation, stemming, lemmatisation, and stop word removal to prepare the data for various applications. Tokenisation breaks down a sentence into individual word units or phrases. Stemming and lemmatisation simplify words into their root form. For example, this process turns starting into start.

Stop word removal ensures that words that do not add significant meaning to a sentence, such as for and with, are removed. Training, Researchers use pre-processed data to train NLP models with machine learning to perform specific applications based on the textual information provided. Training NLP algorithms requires feeding the software with large data samples to improve its accuracy. Deployment and Inference, Machine learning experts then deploy the model or integrate the model into an existing production environment. The NLP model accepts inputs and predicts outputs for the specific use cases it was designed for. You can run the NLP application on live data and get the required output.

We provide some common approaches to natural language processing below, 1) Supervised NLP. Supervised NLP methods train software with a set of labelled or known inputs and outputs. The programme first processes so much known data and learns how to produce the correct output from unknown inputs. For example, a company trains an NLP tool to categorise documents according to certain labels. 2) Unsupervised NLP. Unsupervised NLP uses statistical language models to predict patterns that occur when fed by unlabelled inputs. For example, auto-complete features in text messages suggest relevant words that make sense for sentences by monitoring user responses. 3). Natural language understanding.

Natural language understanding (NLU) is a subset of NLP that focuses on analysing the meaning behind sentences. NLU allows software to find the same meaning in different sentences or process words that have different meanings. 4) Natural language generation. Natural language generation (NLG) focuses on creating human-like conversational text based on certain keywords or topics. For example, an intelligent chatbot with NLG capabilities can communicate with customers in the same way that customer support staff do. This research aims to build a Chatbot to maximise the automation of FAQs at PT SRI by using the NLP method and applied to chat messenger on the website. So that the chatbot that has been built can contribute to improve services that are able to answer questions automatically (Bock & Garnsey, 2008).

There are strong implications for all businesses, particularly large businesses in competitive industries, where failure to deploy AI in the face of competition from firms who have deployed AI to improve their decision- making could be dangerous (Stone et al., 2020).

NLP and Chatbot Website Platform

Imagine a large company that produces many products every day. In this case, usually not all

customers will be satisfied with the results of the products received. The company must of course provide a means of complaint. The higher the production level, the means of complaints must also increase. If we go back to the old days where the use of the internet and technology was not as massive as it is now, then usually the company will hire a customer service specifically to serve customer complaints.

But along with the times, customer demands are certainly getting higher. Customers want to be served immediately, and of course a customer service has limited energy in working. Nowadays, companies can rely on chatbots, where a chatbot can respond to every customer complaint properly 24 hours a day.

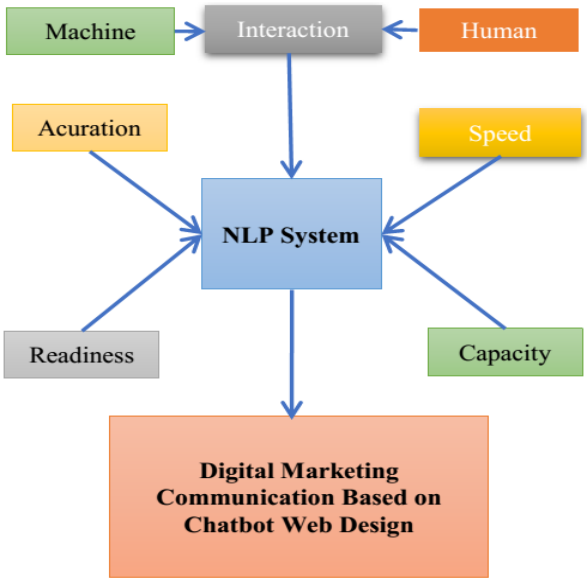

Figure 4. NLP for chatbot web design

Natural Language Processing (NLP) is the processing of language, such as spoken and language processing, such as spoken and written by humans in everyday conversations through computer (Bose, 2004; Parde, 2023). The computational process for processing language processing, must be represented into a series of symbols that fulfil certain rules. In the process, NLP will make computer can understand any commands or standard language that is commonly written or performed by humans. The output of the standard answers entered by the user beforehand already based on the summarized meaning of the input (Lisangan, 2013). The application of chatbot can be applied in the form of NLP (Natural Language Processing) which is one of the field of Artificial Intelligence) to study communication that is by humans with computers through natural language (Aleedy et al., 2019; Prasetyo et al., 2021).

System and software design is the stage of describing and designing the system and interface display, both the letters used and the background as the appearance of the virtual customer services chatbot application (Mulyatun et al., 2021). The basic assumptions of NLP are formulated by Bodenhamer, and other sources that are the foundation of the techniques in NLP, including the following (Wikanengsih, 2012). 1) The map is not the territory. The map is not the territory. What is experienced, seen, heard, felt, is not the real thing, but what the brain interprets it to be. In other words, a territory never changes, it is the meaning of the territory that is constantly changing. This assumption is the basis of several techniques that can be used when dealing with an event. These techniques include swish patterns, mapping accros. People respond according to their internal maps. A person's response is what is in his/her internal map/perception. Meaning is context dependent. A meaning depends on a particular context. We cannot not communicate. Humans are always communicating.

In every situation, humans are basically always communicating. When we are silent, we communicate with ourselves. Although we do not use words verbally, nonverbal language is always used when we do not use verbal words. Therefore, verbal and nonverbal language is a study in NLP. This assumption underlies the use of sensory equity (high sensitivity) or representation system (visual, auditory, kinesthetic) techniques. Wellformed outcome. Express it with positive language. To achieve the goal of an action, language should be used that has a positive meaning. Mind and body are one system and affect each other: the mind and body affect each other. This assumption is the basis of the sensocy acuity technique: sensory sensitivity, matching and mirroring; and state. There are two communication levels: Conscious and Unconscious. There are two levels of communication, namely through the conscious mind and the unconscious mind (Prenga, 2020). Communication done through the unconscious mind is more effective than communication done through the conscious mind. Therefore, techniques that can be used to influence the unconscious mind need to be used. Techniques that can be used include the alpha technique.

The use of artificial intelligence on this chatbot system lies in the pre-processing process, specifically using natural language processing (Pugalenthi et al., 2021; Shah et al., 2017). NLP is a branch of AI concerned with enabling computers to understand words in much the same way as human. The preprocessing process includes tokenization and lemmatization. This method is applied to the process of preparing a bag of words for chatbot training based on the pattern and response file (Christian & Erline, 2022).

This pattern and response file is in json and divided into several tags, such as greeting, thanks, goodbye, no

answer, address, contact, news, event, announcement, scholarship, program, etc. The tag and pattern are obtained from summarizing all information that may be obtained from the website used for web scraping, in this case PT. Spicering Rempah Indonesi website. Meanwhile the response is obtained from the result of web scraping. Thereafter, this method is applied for processing the user's response in advance of doing the pattern matching to find the answer (Christian & Erline, 2022).

Conclusion

There are several things that have been done and can be concluded as follows: In this research has been achieved analysis, design, implementation, testing and publication. Analysis and design in this research using an object-oriented approach. Diagrams for design and modelling using UML. Customer Service can be facilitated in respond to questions from visitors or customers who ask without having to answer them manually. Natural Language Processing approach In the customer services application, conversations and discussion that occurs as if it were done between humans and humans.

Acknowledgments
The author would like to thank the parties who have played a role in this research activity, so that this research can be carried out well.

Author Contributions
This research was supported by equal distribution of roles and contributions of all authors, because each stage was always discussed together.

Funding
This research was funded by the RKAT of PTNBH Universitas Sebelas Maret Class of 2023 through the Research scheme PENELITIAN HIBAH GRUP RISET (HGR-UNS RESEARCH) C with Research Assignment Agreement Number: 228/UN27.22/ PT.01.03/2023".

Conflicts of Interest
In this research, there is no interest and or hidden interests among the researchers.

References

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 373–383. https://doi.org/10.1007/978-3-030-49186-4_31

Ahmad, N. A., Che, M. H., Zainal, A., Abd Rauf, M. F., & Adnan, Z. (2018). Review of chatbots design techniques. *International Journal of Computer Applications*, *181*(8), 7–10. Retrieved from https://shorturl.asia/nvmlq

Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and analyzing chatbot responses using natural language processing. *International Journal of Advanced Computer Science and Applications*, *10*(9). Retrieved from https://doras.dcu.ie/27514/

Amalia, E. L., & Wibowo, D. W. (2019). Rancang Bangun Chatbot Untuk Meningkatkan Performa Bisnis. *Jurnal Ilmiah Teknologi Informasi Asia*, *13*(2), 137–142. https://doi.org/10.32815/jitika.v13i2.410

Anagnoste, S., Biclesanu, I., D'Ascenzo, F., & Savastano, M. (2021). The role of chatbots in end-to-end intelligent automation and future employment dynamics. *Business Revolution in a Digital Era: 14th International Conference on Business Excellence, ICBE 2020, Bucharest, Romania*, 287–302. https://doi.org/10.1007/978-3-030-59972-0_20

Bock, K., & Garnsey, S. M. (2008). Language Processing. *A Companion to Cognitive Science*, *5*(1), 226–234. https://doi.org/10.1002/9781405164535.ch14

Bose, R. (2004). Natural Language Processing: Current state and future directions. *International Journal of the Computer, the Internet and Management*, *12*(1), 1–11. Retrieved from https://shorturl.asia/ICZcS

Chowdhary, K. (2020). Natural language processing. *Fundamentals of Artificial Intelligence*, 603–649. https://doi.org/10.1007/978-81-322-3972-7_19

Christian, Y., & Erline, M. (2022). Web-Based Chatbot With Natural Language Processing and Knuth-Morris-Pratt (Case Study: Universitas Internasional Batam). *JST (Jurnal Sains Dan Teknologi)*, *11*(1), 132–141. https://doi.org/10.23887/jstundiksha.v11i1.43258

Fanni, S. C., Febi, M., Aghakhanyan, G., & Neri, E. (2023). Natural language processing. In *Introduction to Artificial Intelligence* (pp. 87–99). Springer. https://doi.org/10.1007/978-3-031-25928-9_5

Harahap, D. W., & Fitria, L. (2020). Aplikasi Chatbot Berbasis Web Menggunakan Metode Dialogflow. *J-ICOM - Jurnal Informatika Dan Teknologi Komputer*, *1*(1), 6–13. https://doi.org/10.33059/j-icom.v1i1.2796

Hormansyah, D. S., & Utama, Y. P. (2018). Aplikasi chatbot berbasis web pada sistem informasi layanan publik kesehatan di malang dengan menggunakan metode tf-idf. *Jurnal Informatika Polinema*, *4*(3), 224–228. https://doi.org/10.33795/jip.v4i3.211

Kato, Y. (2020). AI/Fintech and Asset Management Businesses. *Public Policy Review*, *16*(4), 1–28. Retrieved from https://ideas.repec.org/a/mof/journl/ppr16_04

_04.html

Kiruthika, V., Pravin, S. C., Rohith, G., Aswin, B., & Ompirakash, S. (2023). A Chatbot-Based Strategy for Regional Language-Based Train Ticket Ordering Using a novel ANN Model. In *Scalable and Distributed Machine Learning and Deep Learning Patterns* (pp. 168–184). IGI Global. https://doi.org/10.4018/978-1-6684-9804-0.ch010

Lisangan, E. A. (2013). Natural Language Processing Dalam Memperoleh Informasi Akademik Mahasiswa Universitas Atma Jaya Makassar. *TEMATIKA: Jurnal Penelitian Teknik Informatika Dan Sistem Informasi*, *1*(1), 39–46. Retrieved from https://ojs.uajm.ac.id/index.php/tematika/article/view/362

Mhlanga, D. (2020). Industry 4.0 in finance: the impact of artificial intelligence (ai) on digital financial inclusion. *International Journal of Financial Studies*, *8*(3), 1–14. https://doi.org/10.3390/ijfs8030045

Mulyatun, S., Utama, H., & Mustopa, A. (2021). Pendekatan Natural Language Processing Pada Aplikasi Chatbot Sebagai Alat Bantu Customer Service. *Journal of Information System Management (JOISM)*, *2*(2), 12–17. https://doi.org/10.24076/joism.2021v3i1.404

Parde, N. (2023). Natural language processing. *The SAGE Handbook of Human--Machine Communication*, 318. Retrieved from https://www.torrossa.com/en/resources/an/5543084#page=365

Prasetyo, V. R., Benarkah, N., & Chrisintha, V. J. (2021). Implementasi natural language processing dalam pembuatan chatbot pada program information technology universitas surabaya. *Jurnal TEKNIKA*, *10*(2), 114–121. Retrieved from https://repository.ubaya.ac.id/39825/

Prenga, K. (2020). *Language learning in the age of Artificial Intelligence* [Thesis). Università Degli Studi Di Milano. Retrieved from https://shorturl.asia/oOyti

Pugalenthi, R., Prabhu Chakkaravarthy, A., Ramya, J., Babu, S., & Rasika Krishnan, R. (2021). Artificial learning companionusing machine learning and natural language processing. *International Journal of Speech Technology*, *24*, 553–560. https://doi.org/10.1007/s10772-020-09773-0

Putra, J. A. (2020). *Penerapan natural language processing dalam aplikasi chatbot sebagai media pencarian informasi dengan menggunakan react: studi kasus Institut Bisnis dan Informatika Kwik Kian Gie/Jasen Aprian Putra/52160253/Pembimbing: Akhmad Budi* [Institut Bisnis dan Informatika Kwik Kian Gie]. Retrieved from http://eprints.kwikkiangie.ac.id/6/

Shah, R., Lahoti, S., & Lavanya, K. (2017). An intelligent chat-bot using natural language processing. *International Journal of Engineering Research*, *6*(5), 281–286. https://doi.org/10.5958/2319-6890.2017.00019.8

Sihombing, D. O. (2022). Implementasi Natural Language Processing (NLP) dan Algoritma Cosine Similarity dalam Penilaian Ujian Esai Otomatis. *Jurnal Sistem Komputer Dan Informatika (JSON)*, *4*(2), 396. https://doi.org/10.30865/json.v4i2.5374

Stone, M., Aravopoulou, E., Ekinci, Y., Evans, G., Hobbs, M., Labib, A., Laughlin, P., Machtynger, J., & Machtynger, L. (2020). Artificial intelligence (AI) in strategic marketing decision-making: a research agenda. *The Bottom Line*, *33*(2), 183–200. https://doi.org/10.1108/BL-03-2020-0022

Sugiono, S. (2021). Pemanfaatan Chatbot Pada Masa Pandemi COVID-19: Kajian Fenomena Society 5.0. *Jurnal Penelitian Komunikasi Dan Pembangunan*, *22*(2), 133–148. Retrieved from https://shorturl.asia/pC0Zc

Wikanengsih, W. (2012). Menerapkan neurolinguistic programming (NLP) dalam pembelajaran. *Semantik*, *1*(1). https://doi.org/10.22460/semantik.v1i1.p%25p