JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

Implementation Of Machine Learning For Web-Based Stroke Probability Prediction

Zuhaira Agustari[1)*], Roberto Kaban[2)], Safarul Ilham[3)]
[1)2)3)]Institut Teknologi dan Bisnis Indonesia, Kab. Deli Serdang, Sumatera Utara, Indonesia
[1)]agustarizuhaira@gmail.com, [2)]roberto.kaban@yahoo.com, [3)]safarulilham@students.usu.ac.id

Submitted:23 Oktober 2025 | **Accepted**: 21 November 2025 | **Psublished**: 30 November 2025

Abstract: In an effort to enhance early detection and prevention of stroke, the implementation of web-based machine learning provides a promising solution. This study focuses on applying machine learning algorithms to predict the likelihood of stroke occurrence based on patient medical data collected online. By using the developed prediction model, the system efficiently analyzes historical data and health risk factors to provide stroke risk estimates. This implementation aims to improve diagnostic accuracy, enable better early detection, and offer appropriate preventive recommendations. The results of this research are expected to assist healthcare professionals and patients in stroke prevention efforts through the utilization of web-based technology.

Keywords: Machine learning; Stroke prediction; Web-based system

INTRODUCTION

Stroke is a leading cause of death and disability worldwide. In Indonesia, the prevalence of stroke continues to increase due to lifestyle changes and an aging population. Stroke occurs when blood flow to part of the brain is disrupted or interrupted, which can be caused by a blockage (ischemic stroke) or a ruptured blood vessel (hemorrhagic stroke). This disease has a significant impact not only on individual health but also on the economic and social burden on society.

Technological advances in the healthcare sector have enabled the collection and analysis of large-scale health data. This opens up opportunities for the development of disease prediction models using machine learning algorithms. By utilizing patient medical history data, including factors such as high blood pressure, diabetes, and cholesterol (40%), family history (10%), age (30%), gender (5%), and lifestyle factors such as smoking, physical activity, and alcohol consumption (15%), a prediction system can help identify individuals at high risk of stroke.

In this paper, the authors developed an innovative system aimed at predicting the likelihood of stroke using machine learning algorithms. This system is designed using data, including the patient's clinical health history and relevant ancestry data, to identify key risk factors contributing to stroke. The clinical health history includes various medical information, such as blood pressure, cholesterol levels, a history of diabetes, and lifestyle factors, while the ancestry data includes genetic information that can influence a person's susceptibility to stroke.

The machine learning algorithm used in this system utilizes data-driven learning techniques to create a predictive model capable of recognizing patterns and relationships between various risk factors and stroke occurrence. The data collected through the patient's medical history and ancestry will be processed to ensure data quality and consistency. The data will be divided into two sets: the dataset and the training data. The training data is used to train the machine learning algorithm, where the algorithm will learn various patterns and relationships in the data to build a predictive model for the system.

This algorithm will identify non-linear relationships and complex interactions between various risk factors, thus providing more accurate predictions. After the model is trained, the training data is used to evaluate its performance by measuring the accuracy, sensitivity, and specificity of the resulting predictions.

Application of the Adboost Method to Optimize Stroke Prediction with the Naive Bayes Algorithm, the researchers designed and built a website to predict stroke using the Naive Bayes algorithm. The stroke prediction test used 11 variables, with 28,500 training data and 572 testing data (results from questionnaires and hospital applications)(Agus Byna,2020).

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

Automatic Prediction Model for Hypertension Types with the Utilization of Artificial Neural Network Machine Learning Algorithms," created a website to automatically predict hypertension types using a machine learning algorithm. The test data consisted of 75 individuals, 70% training data, and 30% testing data, resulting in an accuracy of 85% in hypertension prediction (Purwono, 2022).

In line with research conducted by Sri Hartati (2023), in the Journal of Stroke Detection Prediction Analysis with the EDA Approach and Comparison of Machine Learning Algorithms, researchers created a website to predict stroke detection using machine learning algorithms. This study found that through correlation analysis, age had a strong positive correlation with stroke incidence, indicating that the older a person is, the higher the risk of stroke in this dataset. The SVM model achieved 79.5% accuracy on the test data and 75.6% on the training data. The RFC model achieved 91.0% accuracy on the test data and 100% on the training data.

Referring to previous research, which served as reference material, the author created a stroke prediction system using machine learning to assist medical personnel in processing large amounts of data to determine an individual's stroke risk. The author chose a machine learning algorithm to create a web-based stroke prediction system because of its advantages in ease of implementation and computational efficiency. In healthcare settings that may have limited technical resources and expertise, machine learning can be implemented relatively easily without requiring complex setup and is suited to the needs of the stroke risk prediction system. Therefore, the authors designed and created a system capable of assisting medical personnel and individuals seeking early detection of their stroke risk. This system was built using a machine learning algorithm, which offers advantages in ease of implementation and computational efficiency.

The researchers chose the machine learning algorithm developed by Arthur Samuel because it was considered easier to implement in system development, focusing on each development need at a time and requiring less time. With a stroke prediction system using a machine learning algorithm, it is hoped that those in need can increase efficiency in identifying stroke risk and provide more targeted healthcare services to patients. This system allows the collection of patient medical data through inputs covering various variables such as medical history, age, lifestyle, family history, and gender. By utilizing efficient computer technology, this system can process data quickly and accurately, producing a prediction of a patient's stroke likelihood in a relatively short time.

LITERATURE REVIEW

Machine learning is the application of artificial intelligence that enables a system to automatically learn from a set of data to perform specific tasks without explicit programming (Putra & Santika, 2020). Forward Chaining is a search method that begins the search process with a set of data or facts, and from these facts, a conclusion is sought that provides a solution to the problem at hand. The inference engine searches for rules in the knowledge base whose premises correspond to these facts, then draws a conclusion from these rules. The Forward Chaining method is a decision-making method commonly used in expert systems. The search process with the Forward Chaining method proceeds from left to right, from premises to the final conclusion. This method is often called data-driven, meaning the search is driven by the provided data (Akmaluddin & Dewayanto, 2023).

Stroke is a sudden functional brain disorder with focal or global clinical signs lasting more than 24 hours without signs of non-vascular causes, including signs of subarachnoid hemorrhage, intracerebral hemorrhage, ischemia, or cerebral infarction. Meanwhile, according to (Hariyanti et al., 2020), stroke, often called CVA (Cerebro-Vascular Accident), is a disease/disorder of nerve function that occurs suddenly, caused by disruption of blood flow to the brain.

From the several definitions above, it can be concluded that stroke is a serious medical condition that occurs when blood flow to part of the brain is disrupted, causing damage to brain tissue.

The neurological signs and symptoms that arise in stroke depend on the severity of the blood vessel disorder and its location, including (Ula et al., 2021)1) Sudden facial or limb paralysis (usually hemiparesis), 2). Impaired sensation in one or more limbs (hemisensory impairment), 3). Sudden change in mental status (convulsions, delirium, lethargy, stupor, or coma), 4). Aphysia (slurred speech, lack of speech, or difficulty understanding speech), 5). Dysarthria (slurred or slurred speech), 6). Visual disturbances (hemianopia or monocular) or diplopia, and 7). Ataxia (truncal or limb). h. Vertigo, nausea and vomiting, or headache.

The evolution of website interfaces has undergone significant developments over time, driven by insights gained from server-side logging on most websites. This is because the protocols used to transfer data in web applications do not provide detailed user-interactive information to the server (Stacyana Jesika et al., 2023). The development of JavaScript-based user interfaces and Extensible Markup Language (XML)-based data exchange systems, which already have server-side logging, has been less than satisfactory for users' browsing needs. Today's web page designs have evolved and are capable of transforming user experience, enhancing user experience by displaying

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

web pages with historically distinctive and interactive transitions that involve numerous client-side data transfers within a single website.

The complex functionality of dynamic websites is used for various purposes, such as low-latency input validation, progress reporting, dynamic menus, and providing users with a multimedia experience within a website. The increased interface complexity, compared to the previous concept, allows for more comprehensive logging, allowing for more feedback to end users. Dynamic websites are developed with a focus on comprehensive data-driven information utilization. This data serves to inform page design, marketing focus, and the association's business strategy. Interaction with dynamic interface features utilizes a client-side callback scripting system to handle user mouse and keyboard interactions within dynamic web applications (Jindal & Sardana, 2022).

Prediction is a systematic process of estimating the likelihood of future events based on information from the past and present. The goal of prediction is to reduce errors (the difference between what is predicted and what actually occurs). Predictions do not always provide definitive answers regarding future events, but rather attempt to obtain answers that are as close as possible to what will happen(Hasibuan et al., 2022).

Prediction plays a crucial role in various fields, including economics, health, engineering, and the environment. The application of prediction in certain institutions allows for decision-making or policy-making related to estimates of what might happen in the future based on previous data(Fahrizal et al., 2020).

METHOD

In this thesis, the author analyzes a current system and develops a solution to the problem. In this case, stroke, a frightening threat to some people, especially the elderly, is a solution. The solution to this problem is to develop a system capable of predicting whether someone will suffer a stroke.

Stroke is a leading cause of death and disability worldwide. In Indonesia, the prevalence of stroke continues to increase along with lifestyle changes and an aging population. A stroke occurs when blood flow to part of the brain is disrupted or stopped, which can be caused by a blockage, known as an ischemic stroke, or by a ruptured blood vessel, known as a hemorrhagic stroke. Unhealthy lifestyle changes, such as poor diet and lack of physical activity, as well as an aging population, contribute to the increase in stroke cases in this country. Prompt and appropriate treatment is crucial to reducing the negative impact of this disease.

The Bagus Samudra Pratama Clinic frequently receives patients at risk of stroke. However, due to inadequate resources and equipment, the diagnosis and treatment process based on symptoms is often inaccurate and time-consuming. In the process, medical personnel must gather data from patients, such as their medical history, including high blood pressure, diabetes, and cholesterol (40%), family history (10%), age (30%), gender (5%), and lifestyle factors such as smoking, physical activity, and alcohol consumption (15%). This lack of facilities and resources results in less than optimal treatment and requires greater effort to ensure an accurate diagnosis and provide appropriate care. Several factors contribute to this difficulty, including the following:

1. Limited resources and medical equipment make the process of diagnosing and treating stroke imprecise and slow.
2. Collecting patient medical histories individually is time-consuming and requires significant effort from medical personnel.
3. Lack of facilities leads to inaccuracies in diagnosis and medication administration, resulting in suboptimal patient care.

To address the problems encountered at the Bagus Samudra Pratama Clinic, implementing a stroke prediction system using a machine learning algorithm could be an effective solution. This system can quickly and accurately process and analyze patient data, including medical history, health history such as high blood pressure, diabetes, and cholesterol, as well as other factors such as age, gender, and lifestyle. By using machine learning, medical personnel can easily predict a patient's stroke risk, allowing for more precise and efficient diagnosis and medication administration.

Based on patient data variables, such as medical history, including high blood pressure, diabetes, and cholesterol (40%), family history (10%), age (30%), gender (5%), and lifestyle factors such as smoking, physical activity, and alcohol consumption (15%), along with the aforementioned difficulty factors, the author has developed a stroke diagnosis prediction system using a machine learning algorithm with the forward chaining method. A summary of the manual calculations is shown below.

Based on Table 1 above, a person has factors for each variable, and the calculation weights are shown in the table below:

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

Table 1
Score Table for Each Factor

Number	Factors	Sub-Factors	Score	Symptom Range
1	Medical history	High blood pressure	1	Systolic blood pressure ≥ 140 mmHg
		Diabetes	1	Fasting blood sugar ≥ 126 mg/dL
	Family history	High cholesterol	0.5	Total cholesterol ≥ 240 mg/dL
2	Age	Family history	1	History of stroke in a parent or sibling
3	Gender	Age over 60	1	More susceptible at age ≥ 60 years
4		Male gender	1	Men are at greater risk
	Lifestyle	Female gender	0.5	Women are at lower risk than men
5	Factors	Smoking	1	Current smokers or a history of smoking
		Low physical activity	1	Less than 150 minutes of physical activity per week
		Alcohol consumption	1	Alcohol consumption ≥ 2 drinks per day (for men) or ≥ 1 drink per day (for women)

The table above shows that the scores in this table describe the risk of stroke based on various health and lifestyle factors. A score of 1 is given to factors that significantly influence stroke risk, such as high blood pressure, diabetes, a family history of stroke, age over 60, and unhealthy lifestyles such as smoking, low physical activity, and excessive alcohol consumption. Recent research shows that systolic blood pressure ≥ 140 mmHg and fasting blood sugar levels ≥ 126 mg/dL significantly increase the risk of stroke, so both receive a score of 1. High cholesterol, although also a risk, is considered less influential than other major factors, so it is given a score of 0.5. A family history of stroke significantly increases the risk, as does age over 60. Gender also plays a role, with men being at higher risk than women, who receive a score of 0.5 because the risk of stroke in women is relatively lower. Finally, habits such as smoking, lack of physical activity, and heavy alcohol consumption directly increase the risk of stroke, each receiving a score of 1.

Based on the table above, each score can be calculated using the forward chaining formula that the author has summarized, as shown in the manual calculation below:

Total Score = (Health × 0.4) + (Family × 0.1) + (Age × 0.3) + (Gender × 0.05) + (Lifestyle × 0.15)

With scores for each factor:
Health = High blood pressure + Diabetes + Cholesterol = 1 + 1 + 0.5 = 2.5
Family = 1
Age = 1
Gender = 1
Lifestyle = Smoking + Physical activity + Alcohol consumption = 1 + 1 + 1 = 3

Calculate the total score:
Total Score = (2.5 × 0.4) + (1 (1 × 0.1) + (1 × 0.3) + (1 × 0.05) + (3 × 0.15)
Total Score = 1.0 + 0.1 + 0.3 + 0.05 + 0.45 = 1.95

Based on the manual calculation above, it can be seen that the calculation with a total score of 1.95 as the final ratio yields the following ratio classification:

Table 2
Risk Level Table

Score	Risk Level
Score 0-0.4	Low risk
Score 0.5-0.7	Medium risk
Score 0.8-1	High risk
Score 1.1-1.4	Very high risk
Score 1.5-2	Very high risk
Score 2.1-2.5	Very high risk
Score above 2.5	Very high risk

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

Score	Risk Level
Score 0-0.4	Low risk
Score 0.5-0.7	Medium risk

Based on a total score of 1.95, patients with this score are considered at very high risk for stroke. This stroke diagnosis prediction system is designed to be web-based, utilizing internet technology and can be accessed independently by patients to determine their own predictions or by medical personnel who will diagnose their condition. IT staff will manage the data in the stroke prediction system, ensuring data is always up-to-date and the system runs smoothly.

RESULT

A stroke prediction system uses a machine learning algorithm to analyze and project the risk of stroke in individuals or patients diagnosed with the disease based on various health factors. This system utilizes patient medical history data, including high blood pressure, diabetes, and cholesterol, which contribute 40% to the risk calculation. Furthermore, a family history of previous strokes contributes 10% to the total risk assessment. Age is a key factor, contributing 30%, while gender has a 5% weighting. Lifestyle factors, including smoking habits, physical activity levels, and alcohol consumption, contribute 15% to the risk calculation. By integrating all these variables, this prediction system aims to provide an accurate estimate of a person's likelihood of having a stroke, allowing for more effective preventative measures. This system will enable medical personnel to identify stroke risk more quickly and accurately.

This stroke prediction system, built using the PHP and HTML programming languages and MySQL as a database, is an innovation in the healthcare sector that aims to assist in the early detection of stroke risk in individuals and patients. Using a combination of PHP for programming, HTML for the user interface, and MySQL for the database, this system has been designed to ensure accuracy and reliability. After undergoing a series of tests, the system was successfully implemented without errors and has been piloted on several individuals with satisfactory results. The test results indicate that the system is capable of providing accurate predictions of stroke risk, and it is hoped that it can be used as an effective tool for medical personnel in making decisions regarding stroke prevention and early treatment.

This stroke prediction system makes predictions based on various risk factors, including family history, age, gender, and lifestyle factors such as smoking, physical activity, and alcohol consumption. The system also integrates stroke symptom data obtained from medical personnel at the Bagus Samudra Pratama Clinic. By combining information from various aspects of an individual's health and behavior, the system is able to produce more accurate and relevant predictions. Family history data helps understand a person's genetic predisposition to stroke, while age and gender factors provide additional insights into demographic risk. Lifestyle factors, such as smoking habits, physical activity levels, and alcohol consumption, are also important indicators considered by the system.

In creating a stroke prediction system for patients using the machine learning method that is in accordance with this initial research, it was created using several applications or software were used that were specifically designed for the needs of a stroke prediction system for patients using the machine learning method. The following is the application.

1. XAMPP Software

XAMPP is an open-source PHP-based package developed by an open-source community. XAMPP already provides various requirements, eliminating the need to install other programs. Some of the packages provided include PhpMyAdmin, FileZilla, PHP, MySQL, and Apache.

XAMPP is software that functions as a standalone server (localhost) designed to facilitate the development of web-based applications locally on a computer. This software consists of several main components, such as the Apache HTTP Server as the web server, MySQL as the database management system, and several other supporting programs such as PHP and Perl. Below is an image of the XAMPP application used to build a job candidate prediction system at the Langkat Regency Public Works and Housing Agency, as seen in Figure 1 below.

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

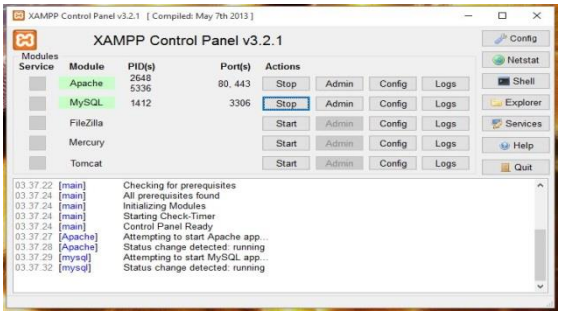

Figure 1. XAMPP Software Image

The image above shows the XAMPP software used to create the MySQL database to be designed. The image below shows the creation of the MySQL database from the initial screen of the application.

2. Sublime Text Software

Sublime Text is a software program that uses HTML and PHP to build user interfaces. Sublime Text is a text editor for the PHP programming language, and it supports other programming languages, with plugins that can extend its functionality.

The following is an image of the Sublime Text application, which the author chose as a tool to design and build a graduate prediction system for prospective workers. This application is used to help determine workforce needs and design the prediction system interface specifically designed for the Langkat Regency Public Works and Housing Agency. This image shows how Sublime Text was utilized in the prediction system development process, from code design to implementation of important features that support the PUPR Agency's needs. For more details, an image of the Sublime Text application can be seen in Figure 2 below, which the author has summarized for your viewing.

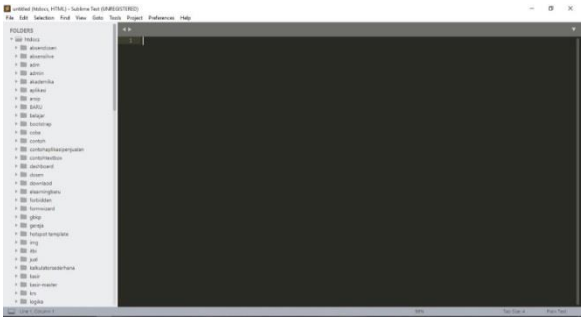
Figure 2. Sublime Text Software

Sublime Text facilitates easy integration between various software components. Developers can quickly access and edit configuration files, ensuring that information can be implemented efficiently and accurately. In addition, the project management features in this application make it easy for developers to organize the project structure well, thus simplifying the maintenance and development of the system in the future. Sublime Text's high reliability and performance also provide additional advantages in information system development. This text editor is capable of handling large projects with good responsiveness, allowing developers to easily explore and edit various parts of the code.

3. Database System Design

Database design is the process of defining the structure, storage, and retrieval of data in a computer system that is efficient and effective in supporting operational and analytical activities. This process involves analyzing data requirements, creating conceptual, logical, and physical data models, and implementing and maintaining the database to ensure data integrity, consistency, and security.

Designing a database for a stroke prediction system is a complex process involving several stages to ensure that the collected data can be analyzed effectively. In this design, data requirements analysis is the primary focus, including identifying the types of data required, such as patient data, medical history, medical test results, and risk factors such as blood pressure and cholesterol levels. Then, a conceptual data model is created to describe the relationships between data entities, followed by a logical data model that defines the structure of tables and relationships within the database. Implementation involves creating a physical database schema in a database

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

management system (DBMS) and populating the database with relevant data. Database maintenance is also crucial to ensure data integrity, consistency, and security. This process allows data to be analyzed using predictive algorithms, such as the Naive Bayes algorithm, to estimate a person's likelihood of having a stroke based on historical data and individual risk factors.

When designing a database for a job candidate prediction system for a Public Works and Housing (PUPR) project, factors such as data security and integration with other systems must be considered. Data security is also a crucial aspect, where security measures such as data encryption and access rights management must be implemented to protect sensitive information. Designing is the process of planning or drawing out a formulation of an activity to be carried out and generating the data that will be needed and used to achieve a desired goal.

According to Gavin Powell, a database is a structured collection of information. A database consists of a collection of data recorded in a computer system. Databases are typically used by organizations to store information related to their business processes, such as employee payroll, customer management, and inventory. [9]. The following is an illustration of the relationships between tables used to build a smart village application, as seen in Figure 3 below.

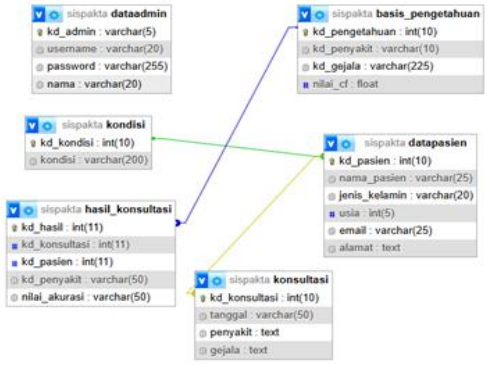

Figure. 3. Relationships Between Tables

This view clearly displays the relationships between tables in the prediction system to provide a deeper understanding of how the data is interconnected. This relationship visualization allows users to easily follow the data flow from query input to prediction results and understand how each element contributes to the system's output. This feature supports a better understanding of the system's structure and dynamics, and simplifies database management and maintenance.

4. System Results

Based on the design results from the system analysis, system design, and testing stages, the stroke probability prediction system using a machine learning algorithm has been tested and performed as intended. The successful test results of the stroke probability prediction system are as follows:

This stroke probability prediction system, built using the PHP and HTML programming languages, and MySQL as a database, is an innovation in the healthcare sector aimed at assisting in early detection of stroke risk in individuals and patients. Using a combination of PHP for programming, HTML for the user interface, and MySQL for the database, this system has been designed to ensure accuracy and reliability. After undergoing a series of tests, the system was successfully implemented without any errors and has been piloted on several individuals with satisfactory results. The test results demonstrate that this system is capable of providing accurate predictions of stroke risk, and it is expected to be used as an effective tool for medical personnel in making decisions regarding stroke prevention and early treatment.

This stroke prediction system makes predictions based on various risk factors, including family history, age, gender, and lifestyle factors such as smoking, physical activity, and alcohol consumption. The system also integrates stroke symptom data obtained from medical personnel at the Bagus Samudra Primary Clinic. By combining information from various aspects of an individual's health and behavior, the system is able to produce more accurate and relevant predictions. Family history data helps understand a person's genetic predisposition to stroke, while age and gender factors provide additional insight into demographic risk. Lifestyle factors, such as smoking habits, physical activity levels, and alcohol consumption, are also important indicators considered by the system. Symptom data obtained directly from the clinic enhances the validity and reliability of the predictions, ensuring that any recommendations are based on accurate and up-to-date medical information. This prediction system not only helps identify stroke risk earlier but also enables more targeted interventions. Based on the design results above, thorough testing and careful monitoring, the author can ensure that this design meets all the expected

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

quality and functionality standards, for more details on the results of testing the entire stroke probability prediction system using a web-based machine learning algorithm, we can see in the image below.

5. Main Page Display

The main page displays health data and prediction results in the form of brief text about the general overview of stroke. This interface is designed for ease of use, with simple navigation and additional guidance for understanding the results. This initial screen also provides users with an overview of how the stroke prediction system works. This includes an explanation of the algorithm used.

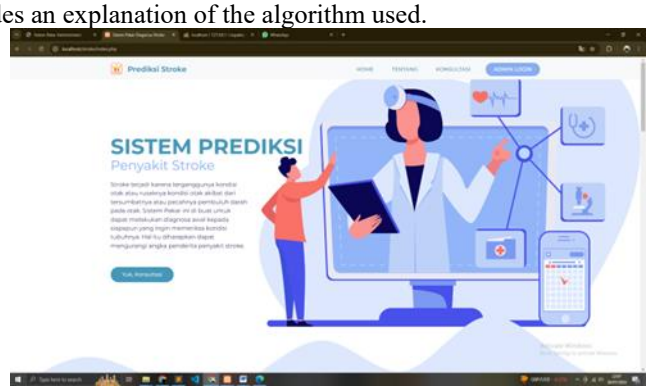

Figure 4. Main Page Display

The following is a screen showing a person's biodata as the initial step in making a prediction. On this screen, users are required to enter various personal data, such as full name, gender, age, email address, and address, which are basic information required to process the prediction accurately. This data not only serves as identification but is also used by the stroke prediction system to generate more targeted, relevant, and personalized results, tailored to each user's individual profile. By providing information such as age and gender, the system can consider certain risk factors that are more likely to occur in certain age groups or genders. Furthermore, the inputted address can also form the basis for additional analysis, for example, to determine environmental conditions that could potentially influence the user's health risks.

The biodata entry interface is specifically designed to be clear, simple, and intuitive, making it easy for users from various backgrounds to understand and fill in information seamlessly. Each input field is designed with informative labels and data validation to ensure only correct and relevant information is accepted by the system. With a clean layout and logical input flow, users can easily complete the process in a short time. This feature aims to ensure a comfortable and efficient user experience, while maintaining the quality of the data received by the system for high-reliability prediction results. The following is a screenshot of the interface, demonstrating how data can be input easily and in an organized manner.

6. Biodata Content Display

The next display displays the stroke symptoms, which the user can fill in after completing the form. This display consists of questionnaire data that the user can select based on their current health status.

Figure 5. Biodata Content Display

7. Stroke Prediction Display

The stroke symptom prediction display provides a questionnaire that users can fill in based on their current health status. Users can select symptoms that best reflect their health status from the list provided, allowing the system to gather relevant information for further analysis.

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

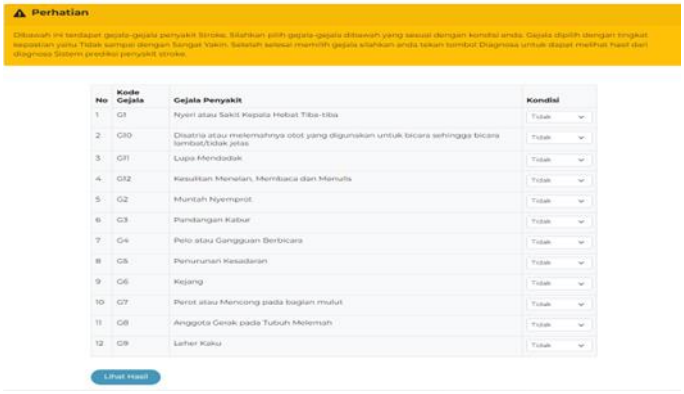

Figure 6. Stroke Prediction Display

8. Stroke Prediction Results Display

In this stroke prediction results display, the system displays analysis results based on the user's answers to the questionnaire. The questionnaire is designed to cover various aspects related to stroke risk factors, such as personal medical history, lifestyle, habits, and relevant genetic or hereditary factors. The system processes the answers to each question using a prediction algorithm, which then produces an estimate of a person's likelihood of having a stroke. This prediction not only provides an overview of the risk level but also includes additional information, such as the main risk factors that influence the prediction results.

The prediction results display is designed to be informative and easy to understand for users. Prediction results are presented in clear formats, for example, as probability numbers or risk categories such as low, medium, or high. Furthermore, users are provided with initial advice or recommendations based on the prediction results to help them take preventive measures, such as seeking medical consultation, making lifestyle changes, or undergoing follow-up health checks. This display also includes details of the user's answers to the questionnaire, so users can understand the relationship between the data they input and the resulting prediction results.

Figure. 7. Stroke prediction results display

9. Main Admin Feature Display

This admin display manages all data and the stroke prediction system. This display can only be accessed by registered administrators. The following is a screenshot:

JCEIT: Journal of Computer Engineering and Information Technology
Volume 2, Number 1, November 2025
Page: 1-11
ISSN 3089-106X
(online)

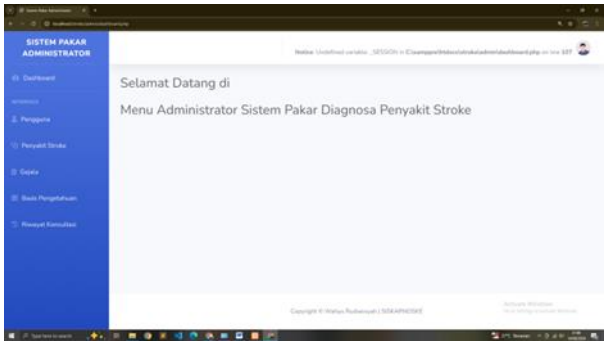

Figure 7. Main Admin Feature Display

This display displays the symptoms data in the stroke prediction system, allowing administrators to manage and analyze the symptom information entered by users. Administrators can monitor selected symptoms, group data by severity, and ensure the accuracy of the information to support more effective predictions.

This stroke prediction result display represents the results of the questionnaire questions completed by the user. These answers will generate a prediction of the likelihood of stroke.

DISCUSSIONS

This individual stroke prediction system was developed using machine learning algorithms to provide an accurate evaluation of stroke risk. The system integrates various types of predictive data, including the patient's family history, age, gender, and lifestyle factors such as smoking habits, physical activity level, and alcohol consumption. The applied machine learning algorithms enable in-depth analysis of this data to predict the likelihood of stroke with a high degree of accuracy.

The prediction system also integrates symptom data identified by medical personnel from the Bagus Samudra Pratama Clinic. Symptoms entered into the system include sudden severe pain or headache, weakness of the muscles used for speech, sudden forgetfulness, difficulty swallowing, reading, and writing, spraying vomit, blurred vision, speech problems, decreased consciousness, seizures, crooked or twisted mouth, weakness of limbs, and neck stiffness. Below are the results of the stroke prediction system trial, summarized by the author.

This prediction system was built using PHP and HTML programming languages, with MySQL as its database. The use of this technology ensures that the system can handle data efficiently and provides a user-friendly interface for both users and administrators. By collecting detailed and relevant symptom data, the system is not only able to provide more accurate predictions regarding the likelihood of stroke, but also enables the creation of appropriate intervention recommendations based on comprehensive data analysis.

From the results and discussion above, there are several advantages and disadvantages of the stroke prediction system. The author has summarized these by considering various aspects and factors that occur in the field. The summary of these advantages and disadvantages is as follows:

1. System Strengths
   a. This prediction system makes it easier for users to continuously monitor their health. Users can see changes in their health data in real time, allowing them to detect problems early and take action before the condition worsens.
   b. This prediction system can be tailored to each person's health condition, providing more precise and appropriate recommendations.
   c. This prediction system uses machine learning algorithms and high-quality health data to accurately predict the likelihood of a stroke.
2. System Weaknesses
   a. Limited resources and technology can limit the system's ability to provide highly accurate predictions.
   b. Incomplete or inaccurate data can result in incorrect predictions, which can lead to results that are inconsistent with actual health conditions.
   c. This prediction system requires regular updates and maintenance to remain effective, which incurs additional costs.

CONCLUSION

The conclusion of the analysis and evaluation of the research entitled "Implementation of Machine Learning for Web-Based Stroke Probability Prediction" is as follows: It can be concluded that: (1) The stroke probability prediction system has been successfully built and designed using the PHP programming language and MySql as

its database. This system will later be able to run smoothly and can help the needs of the Pratama Bagus Samudra clinic. (2) The use of the stroke probability prediction system has been successfully implemented with a machine learning algorithm with patient medical history data, family history, age, gender, and lifestyle such as smoking, physical activity, and alcohol consumption to predict and determine the possibility of patients at the Pratama Bagus Samudra clinic having a stroke.

REFERENCES

Akmaluddin, M., & Dewayanto, T. (2023). Systematic Literature Review: Implementasi Artificial Intelligence Dan Machine Learning Pada Bidang Akuntansi Manajemen. Diponegoro Journal of Accounting, 12(4), 1–11. http://ejournal-s1.undip.ac.id/index.php/accounting.

Byna, A., & Basit, M. (2020). Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naïve Bayes. 09(November), 407–411.

Cahyono, D. S., Nugrahanti, F., & Hendrawan, A. T. (2019). Prosiding Seminar Nasional Teknologi Informasi dan Komunikasi (SENATIK). Aplikasi Pemasaran Berbasis Website Pada Percetakan Morodadi Komputer Magetan, 2(1), 129–134.

Igfirly Mustaib, R., Dwiyansaputra, R., muaidi, M., Desa Sandik Jl Pariwisata, K., & Layar, B. (n.d.). SISTEM INFORMASI COMPANY PROFILE KANTOR DESA SANDIK BERBASIS WEBSITE (Website based Information System of Company Profile for Sandik Village).

Kusuma, A. S., & Nita, S. (2019). Rancang Bangun Media Pembelajaran Pengenalan Tumbuhan Bagi Penyandang Tuna Rungu Pada SDLB Manisrejo Kota Madiun. Seminar Nasional Teknologi Informasi Dan Komunikasi 2019, 281–286.

Metode, M., Di, R. A. D., & Ahmad, S. (2022). No Title. 11(1), 79–85.

Musik dengan Metode Content-Based Filtering. Edumatic : Jurnal Pendidikan Informatika, 4(1), 121–130. https://doi.org/10.29408/edumatic.v4i1.2162

Prediksi, A., Stroke, D., & Pendekatan, D. (2022). Analisis prediksi deteksi stroke dengan pendekatan eda dan perbandingan algoritma machine learning. 02, 355–367.

Purwono, P., Dewi, P., Wibisono, S. K., Dewa, B. P., Informatika, P., Bangsa, U. H., Keperawatan, P., & Bangsa, U. H. (2022). Model Prediksi Otomatis Jenis Penyakit Hipertensi dengan Pemanfaatan Algoritma Machine Learning Artificial Neural Network. 7(2), 82–90.Putra, A. I., & Santika, R. R. (2020). Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. Edumatic : Jurnal Pendidikan Informatika, 4(1), 121–130. https://doi.org/10.29408/edumatic.v4i1.2162

Ula, M., Ulva, A. F., & Mauliza, M. (2021). Implementasi Machine Learning Dengan Model Case Based Reasoning Dalam Mendagnosa Gizi Buruk Pada Anak. Jurnal Informatika Kaputama (JIK), 5(2), 333–339. https://doi.org/10.59697/jik.v5i2.267

Utama, T. P., & Haibuan, M. S. (2023). PENERAPAN ALGORITMA NAÏVE BAYES DAN FORWARD SELECTION UNTUK PREDIKSI PENYAKIT STROKE. 17, 351–357.

Fahrizal, Reynaldi, F. O., & Hikmah, N. (2020). IMPLEMENTASI MACHINE LEARNING PADA SISTEM PETS IDENTIFICATION MENGGUNAKAN PYTHON BERBASIS UBUNTU JISICOM (Journal of Information System , Informatics and Computing) p-ISSN : 2579-5201 (Printed) JISICOM (Journal of Information System , Informatics and. Journal of Information System , Informatic and Computing, 4(1), 86–91.

Hasibuan, E., Informasi, S., Ilmu, F., Informasi, T., Gunadarma, U., Margonda, J., No, R., Cina, P., & Jawa, D. (2022). Implementasi Machine Learning untuk Prediksi Harga Mobil Bekas dengan Algoritma Regresi Linear berbasis Web. Jurnal Ilmiah Komputasi, 21(4), 595–602. https://doi.org/10.32409/jikstik.21.4.3327

Putra, A. I., & Santika, R. R. (2020). Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. Edumatic : Jurnal Pendidikan Informatika, 4(1), 121–130. https://doi.org/10.29408/edumatic.v4i1.2162

Stacyana Jesika, Suci Ramadhani, & Yohanna Permata Putri. (2023). Implementasi Model Machine Learning dalam Mengklasifikasi Kualitas Air. Jurnal Ilmiah Dan Karya Mahasiswa, 1(6), 382–396. https://doi.org/10.54066/jikma.v1i6.1162