Journal of Intelligent Computing and Health Informatics

Check for updates

Development of a website-based system for classification of student characteristics using the prototyping method

Research Article

Lisdi Inu Kencana*, Fauzi Adi Rafrastara (ID), Cinantya Paramita

Lisdi Inu Kencana, Fauzi Adi Rafrastara, and Cinantya Paramita are with the Department of Informatics Engineering, Universitas Dian Nuswantoro Semarang, Jl. Imam Bonjol No. 207, Semarang 50131, Indonesia (emails: ldinuken@gmail.com, {fauziad, cinantya.paramita}@ dsn.dinus.ac.id)

*Corresponding author.

ABSTRACT

Student characteristics are essential attributes to understanding how think and academic abilities the students. In teaching and learning process, appropriate learning strategies must be applied. Hippocrates-Galenus typology categorizes personality types into four different categories, namely sanguine, choleric, melancholic, and apathetic. Classification of characteristics using student approach methods based on experience or intuition can produce inaccurate results and require a lot of time to process. This study proposed a website cognitive system to classify the characteristics of primary school students at SD Wijaya Kusuma 02 Semarang using the prototyping method. The proposed method has several stages: Communication, Quick Plan, Modeling Quick Design, Construction of Prototype, and Deployment Delivery & Feedback. We used the C4.5 algorithm embedded in the system for classifying student characteristics. The results show that tests carried out using C4.5 algorithm obtained an accuracy of 90.08%. It can be concluded the C4.5 algorithm can classify student characteristics well.

Keywords: Student characteristics, classification, prototyping, C4.5 algorithm, data mining.

1. INTRODUCTION

The school becomes a formal educational institution that is important in developing character attitudes in students, especially in the aspects of discipline and responsibility. It cannot be denied that the school is one of the dominating factors in the formation of student characteristics (Suparno, 2018). In learning, it is very important to classify students' characters to identify students in terms of the need to set specifications and qualifications in dealing with problems that occur to students (Sari et al., 2020).

Based on the findings in the field, several students experience problems in learning, where each student requires a different approach. The problems experienced by some students include students who are less able to follow the learning of certain teachers. This is due to the difficulty of teachers in finding learning methods that suit the needs of students. The methods used in schools are generally conventional (Andi Ferawati Jafar, 2021).

The primary school teachers at SD Wijaya Kusuma 02 Semarang have difficulty understand the student's character. The methods used in learning still apply conventional methods, so some students with specific characteristics find it challenging to adapt. Due to the diversity of personalities among these students, it becomes a challenge for teachers to adapt to the prescribed teaching and assessment methods.

SD Wijaya Kusuma 02 Semarang does not yet have a system that can assess the personality of students, so the school has difficulty in giving proper treatment to students in the teaching and learning process in class. In the learning method carried out by the teacher, it is necessary to approach students so that in teaching and learning, all students can understand what is conveyed by the teacher. So far, the teachers of SD Wijaya Kusuma 02 Semarang have tried to understand the characteristics of students only based on their experience or intuition, without a scientific system and approach.

The method that will be used for system development in this research is the prototyping method. In the application of the prototyping development method, the main goal is an iterative process and repeated testing so that the communication and testing process is always carried out by researchers and SD Wijaya Kusuma 02 Semarang so that it can produce the expected system. System development using the prototyping method can save costs and time and have good communication between researchers and SD Wijaya Kusuma 02 Semarang.

To be able to know the characteristics of the students of SD Wijaya Kusuma 02 Semarang, data mining classification techniques are needed. A classification technique is a technique for finding similarities among a collection of underlying data objects and classifying them into various groups according to a certain category model. In this study, the classification algorithm used is the C4.5 algorithm. The C4.5 algorithm is a way to extract data, and the main thing is to find out the input variables or attributes that are related to each other.

From the above problems, the researchers developed a website-based system that can classify the characteristics of the students of SD Wijaya Kusuma 02 Semarang using the prototyping software development method. Later this system can also be applied in other schools, especially to find out the characteristics of students and so that the school can provide the right treatment to students.

2. LITERATURE REVIEW

2.1 Website-based Information System

Web-based information system is an information system that is used as a standard for storing, retrieving, formatting, and displaying data using a client/server architecture. This information system manages various kinds of digital data, including text, hypermedia, and implementation of Graphical User Interface (GUI), which is very simple and easy to use.

A software program can access a web-based information system, which consists of interconnected components that are used to collect, process, store, and transfer information in the form of text, images, sound, and information presented in hypertext form to support creation. HTTP (Hypertext Transfer Protocol) is used by web browsers via web clients to read web pages stored on web servers to translate hypertext documents into human-accessible documents (Prayama et al., 2021)

2.2 Student characteristics

The characteristics of the human personality are the behavior that every human being is born with, and the personality of each human will definitely be different from one another, and no personality is the same as other people. The personality characteristics of a person or human will begin to be seen entering a new environment to find new things that have never been found before (Khalid Aji, 2019).

The teacher's ability to understand the characteristics of students who have differences in physical and non-physical dimensions will certainly affect how well the learning process takes place. Understanding student characteristics is a factor in the success of the learning process. This happens because the teacher is tasked with practicing learning (Janawi, 2019).

Factors Affecting Characteristics, the human personality will develop and grow over time through interaction. There are many factors that influence human character, especially in childhood (Isnainia Solicha & Na'imah, 2020). These factors are:

- Environmental factor
- Genetic and Gender Factors
- Social Factor
- Parental Factor

History of Hippocrates-Galenus Typology, Typology is the science that sorts the diversity of species and forms equations to organize, categorize, and group objects that have the same formal structural qualities and basic properties into certain types. The typological classification aspect focuses on efforts to classify, categorize, and classify objects in the form of typology based on certain aspects (Bali, 2020).

There are several types of typologies, one of which is constitutional typology. This typology was developed by several scholars, including Hippocrates and Galenus. Galenus devised this typology, which was based on the ideas of Hippocrates (460-370 SM), which was influenced by the beliefs of Ependocles (490-435 SM), in which the figure stated that the universe includes four basic elements, namely earth, water, air, and fire (Meilana et al., 2021).

2.3 Prototyping Method
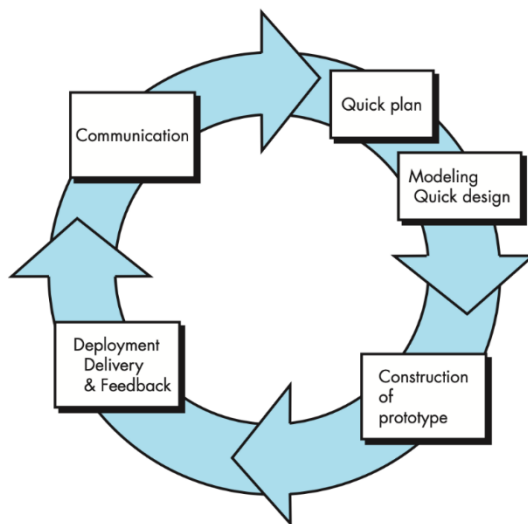
Fig 1. Prototyping model.

The prototyping method is an understanding of the prototype approach (life cycle using prototyping). One approach to the system life cycle based on the concept of a working model is to understand the prototyping method. The prototyping method is used to develop the model into a final system (see Fig. 1). Thus, this system will be developed quickly and at a lower cost (Nurul Renaningtias & Dyah Apriliani, 2021)

2.4 Data Mining Classification

The classification technique is a method to identify common characteristics among a group of database objects and categorize them into different groups according to the recommended categorization model. Decision tree techniques, Bayesian (Naïve Bayesian and Bayesian Belief Networks), Artificial Neural Networks (Backpropagation), concept-based techniques from mining association rules, and other techniques (k-Nearest Neighbor, genetic algorithm, technique with rough set approach and fuzzy) are a classification category in data mining (Devi Udariansyah & Robbi Lesmana, 2022).

The process of classifying data mining by performing new data categories from previously collected data (nominal or ordinal). A confusion matrix occurs if the estimation accuracy of the classification model is accurate. From the resulting matrix, data mining will estimate the accuracy of the process that has been run (Eka Sabna & Yuda Irawan, 2021). The sample data attribute serves as the input for the classification, and the actual class of the sample data serves as the output. Prototyping models with storage can perform categorization procedures. Two stages of classification will take place, namely:

a. Training Process. The training process is a data input process consisting of attributes and labels that will be used as guidelines to determine and find further results.
b. Testing Process. The Testing process is a testing process that uses labels and attributes on the training data to be tested to produce results and determine the accuracy of the results in the training process.

2.5 C4.5 Algorithm

The C4.5 algorithm is a method for extracting data, specifically to determine the relationship between input variables or attributes. The transformed data results in decisions that are easy for users to understand and understand. The following are the components that must be looked for:

a. Entropy (S) is a parameter that measures the variables of each value attribute criterion in a data set with a decision tree. The lower the entropy value, the greater the level of diversity of a data set, and vice versa; the higher the entropy value, the greater the increase in diversity value, follows (1).

$$Entropy(S) = \sum_{i=1}^{n} pi * log_2 pi \qquad (1)$$

where:
S: Set of cases
n: Number of partitions S
pi: proportion of S_i to S

b. *Gain (S, A)* is defined as the difference between the overall entropy value and the entropy value of each value for each criterion multiplied by the proportion of attribute values. The gain value function measures the effectiveness of each characteristic in the categorized data, follows (2). The following is the gain formula:

$$Gain\ (S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_i|}{S} * S_i \qquad (2)$$

where:
S: Set of cases

A: Attributes

n: Number of attribute partitions A

$|S_i|$: Number of cases on the 1st partition $|S|$: Number of cases in S

3. METHODOLOGY

The research method is a scientific way to obtain data with a specific purpose and use (Sugiyono, 2019). Research method refers to the method or approach used in the process of obtaining an accurate response to research. The Fig. 2, show the our research method.
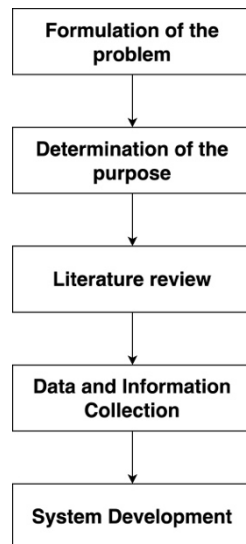

Fig. 2. Research method.

3.1 Formulation of the problem

The system for classifying student characteristics used at SD Wijaya Kusuma 02 Semarang is still using conventional methods. Difficulties experienced by teachers in approaching students with certain characteristics. This can make it difficult for students to adapt to the activities of the learning process.

3.2 Determination of the purpose

Classification of student characteristics of SD Wijaya Kusuma 02 Semarang will use the C4.5 algorithm, which is implemented into a website using the PHP programming language. In classifying the characteristics of students using the website so that teachers can more easily find out the types of characteristics of each student.

The website for classifying the characteristics of students at SD Wijaya Kusuma 02 Semarang has several multi-levels in its use, namely as an admin who can manage user accounts and data on questions to classify students' characteristics, then as students or users can fill in questions to find out the appropriate type of characteristics.

3.3 Literature review

Literature Study is a step in collecting information that is being developed related to research. By using the literature study method to collect data, you can find out about things such as related research that has been done for a long time and research that is currently being used.

Literature studies are used to complement the researcher's data, including the results of studies on data mining that are associated with using the C4.5 algorithm, such as scientific reports, theses, and books that follow the research topic and others.

3.4 Data and information collection

The data collection used in this study was to compile and complete the data in the form of observations, interviews, and questionnaires. Based on the observations, the researchers found a problem for students at SD Wijaya Kusuma 02 Semarang, where each child has a different approach according to the problems experienced, such as lack of literacy and numeracy, reading speed, and other problems, so that each student needs special assistance treatment to find solutions to these problems.

3.5 System development

The development of the system on the website for classifying the characteristics of students at SD Wijaya Kusuma 02 Semarang will use the prototyping method. The following is an explanation of the process phases of the prototyping method:

a. Communication. In the communication process using the prototyping model, users and developers talk about several topics that are significant for the actualization of the software required by the user.

b. Quick Plan. In determining all the specifications that will be used during the prototyping process, the identification of requirements and the tools used is a very important step.

c. Modeling Quick Design. In the quick design modeling stage, the design development for the student characteristics classification system occurs at this stage. Stages of research are carried out by creating a website display, database design, and website system design.

d. Construction of Prototype. During the construction of the Prototype stage, you can observe various software features throughout the prototyping stage, which can later be seen by end users.

e. Deployment Delivery & Feedback. At the Deployment Delivery & Feedback stage, the researcher conducted testing using black box testing. Black box testing is done to see if the application or website is ready for the user or if there are still some problems.

4. RESULTS AND DISCUSSION

Development of A Website-Based System for Classification of Student Characteristics Using The Prototyping Method
p-ISSN 2715-6923, e-ISSN 2721-9186 , Vol. 3, No. 1, March 2022, *pp.9-16*

13

4.1 Use case diagram

The case diagrams explain the description of the system that will be made by the researcher. Fig. 3, shows our use case that explains two actors in development, namely admin and user.
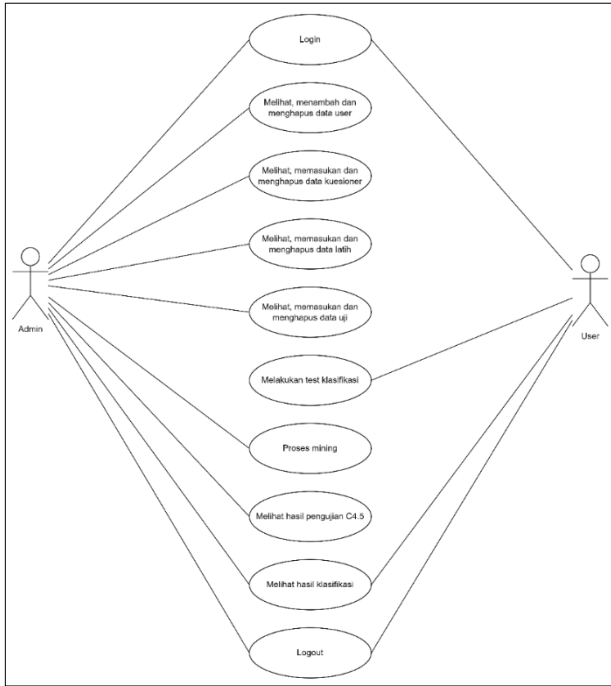

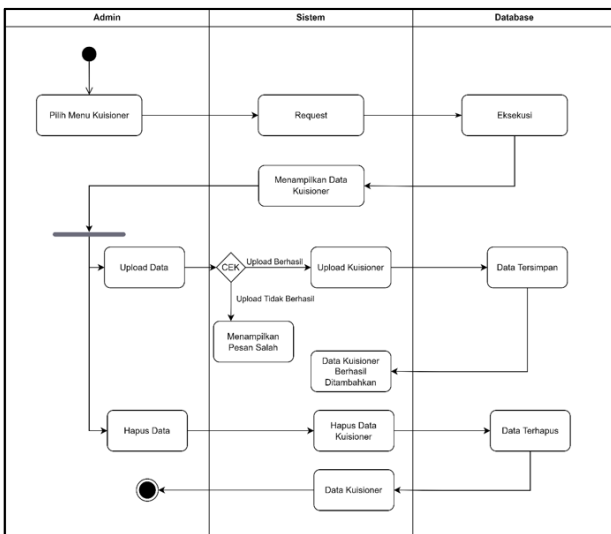Fig. 3. Use case diagram.

4.2 Activity Diagram

a. Questionnaire Activity Diagram

Fig. 4. Questionnaire activity diagram.

The questionnaire activity diagram can see in Fig. 4. As we can see in Fig. 4, the admin can upload data using *.xls file. In addition, also delete data to update the questionnaire that will be used in the process of classifying student characteristics.

b. Data mining activity diagram

In Fig. 5, show the data mining activity diagram. As we can see Inf Fig. 5, admin able first enter the C4.5 menu then the admin will see the training data that will be used for calculating the classification of data mining using the C4.5 algorithm. After the admin adds the training data, it can proceed to the mining process stage to perform calculations and find out the results of the decision tree that will be used to determine the accuracy of the data to be tested. An admin can create, read, delete and upload training data using files by the *.xls extension.
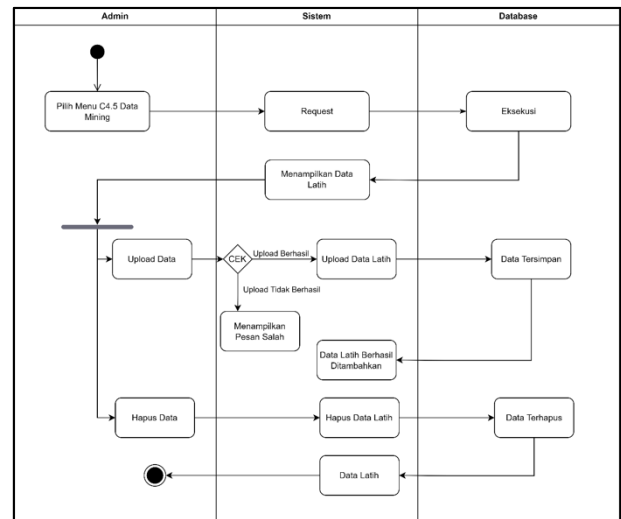
Fig. 5. Data mining activity diagram.

4.3 Class diagram

In the Unified Modeling Language (UML), a class diagram is a special type of structural diagram that describes the structure and description of classes, properties, methods, and relationships between objects.

4.4 Implementation

a. Data mining Page

Fig. 6. Data mining page.

In Fig. 6, shows the data mining page that the admin

will do to upload and delete student data or training data. The admin will upload the data using the .xls file then the data will be displayed on the page.

b. Mining processing page

Fig. 7, shows the mining process page that will be displayed after the data input process is carried out on the data mining page. The mining process will display the amount of data, the number of sanguines, the number of clerics, the number of melancholies, the number of phlegmatics, as well as the entropy and gain of the data that has been inputted.

Fig. 7. Mining processing page.

c. Decision tree page

Fig. 8, shows the decision tree page or what can be referred to as a rule. Rules are formed based on the results of the mining process that has been done. On the decision tree page, the admin can also delete the decision tree that has been formed and then carry out the mining process again. After the admin knows the decision tree that has been formed, the admin can do the rules test.

A decision tree, also known as a decision tree, is a tree that is used as a reasoning technique to arrive at a solution to a problem. If all data features are identified using two types of value categories, a decision tree in the form of a binary tree will be formed. When a feature has more than two value categories, either categorical or numeric, we get a tree shape, which is usually not a binary tree.

Fig. 8. Decision tree page.

d. Decision tree testing page

Fig. 9, shows the decision tree test page that was obtained after the decision tree process was obtained. After the test data is inputted, then the admin can calculate the accuracy to find out the accuracy results of the input test data.

Fig. 9. Decision tree testing page.

5. Classification Result Page

Fig. 10 is a display of the classification results page that is carried out when the user performs a classification test.

Fig. 10. Classification result page.

6. User Classification Page

Fig. 11, is a user classification page display. Users will get characteristic results according to the rules that have been formed.

Fig. 11. User classification page.

Table 1. Black box testing.

No.	Process	Expected Result	Test results
1	Doing admin login	Login to the admin dashboard	In accordance
2	Add user data	User data for students can be saved to the database	In accordance
3	View user data	Displays user data stored in the database	In accordance
4	Delete user data	Delete user data	In accordance
5	Upload questionnaire data	Questionnaire data can be saved to the database	In accordance
6	View questionnaire data	Displaying questionnaire data on the admin or user page	In accordance
7	Delete questionnaire data	Deleting questionnaire data	In accordance
8	Upload training data	Training data can be saved to the database	In accordance
9	View training data	Displaying training data on the data mining page	In accordance
10	Clear training data	Delete training data	In accordance
11	Mining process	See C4.5 calculation from training data	In accordance
12	Decision Tree	Displays the result of the decision tree from the calculation of C4.5 performed	In accordance
13	Delete decision tree	Delete decision tree results to perform the recalculation	In accordance
14	Decision tree test	Perform decision tree test with test data to see accuracy results	In accordance
15	View test data	Displaying test data stored in the database	In accordance
16	Upload test data	Training data can be saved to the database for testing	In accordance
17	Clear test data	Delete test data	In accordance
18	View the classification results	Displays the results of the classification performed by the user	In accordance
19	Admin logout	Log out of system admin	In accordance
20	Doing user login	Log in to the user dashboard	In accordance
21	Classification process	The user answers questionnaire questions	In accordance
22	Classification results	Users see the results of the appropriate characteristic classification	In accordance
23	User logout	Log out of the user system	In accordance

4.5 Black Box Testing

At this stage, after the software prototype is made, the next process is evaluation. The evaluation of the research carried out will use black box testing. These tests can help researchers to determine the specifications of the system being developed. Black box testing is used in the test scenario process. The black box testing is carried out to find out whether the system built is in accordance with the list of system specifications developed previously (Supriyono, 2020).

Table 1 shows the results of black box testing with system requirements that will be used as a guide for the level of testing.

5. CONCLUSION

An information system for classifying student characteristics at SD Wijaya Kusuma 02 Semarang has been completed using the prototyping method. This website-based system has succeeded in classifying the characteristics of students in grades 1-6 at SD Wijaya Kusuma 02 Semarang. The attributes used in the classification process are gender, age, and answers to a questionnaire of 40 questions. In this study, the C4.5 algorithm is used for data classification that has numeric and categorical attributes. From the test results with the data of class 1-6 students amounting to 121 students, get an accuracy of 90.08%. This shows that the C4.5 algorithm can classify student characteristics well.

REFERENCES

Andi Ferawati Jafar. (2021). Penerapan Metode Pembelajaran Konvensional Terhadap Hasil Belajar Fisika Peserta Didik. *Journal of Islamic Education*, *3*(2), 190–199. https://doi.org/10.24252/asma.v3i2.23748

Bali, M. M. E. I. (2020). Tipologi Interaksi Sosial Dalam Meningkatkan Karakter Disiplin Siswa. *Edureligia: Jurnal Pendidikan Agama Islam*, *04*(1), 47–62. https://doi.org/10.33650/edureligia.v4i1.1157

Devi Udariansyah, & Robbi Lesmana. (2022). Implementasi Algoritma C4.5 Pada Jumlah Penduduk Kota Prabumulih menggunakan Metode

Klasifikasi. *Jurnal Pendidikan Dan Konseling (JPDK)*, *4*(4), 2513–2520. https://doi.org/10.31004/jpdk.v4i4.5829

Eka Sabna, & Yuda Irawan. (2021). Data Mining Dengan 2 (Dua) Model Klasifikasi Untuk Prediksi Kinerja Mahasiswa. *Jurnal Ilmu Komputer*, *10*(2), 128–131. https://doi.org/10.33060/JIK/2021/Vol10.Iss2.229

Isnainia Solicha, & Na'imah. (2020). Faktor yang Mempengaruhi Perkembangan Anak Usia Dini. *Jurnal Pelita PAUD*, *4*(2), 197–207. https://doi.org/10.33222/pelitapaud.v4i2.968

Janawi. (2019). Memahami Karakteristik Peserta Didik dalam Proses Pembelajaran . *Tarbawy : Jurnal Pendidikan Islam*, *6*(2), 69–79. https://doi.org/10.32923/tarbawy.v6i2.1236

Khalid Aji. (2019). Sistem Pakar Tes Kepribadian Menggunakan Metode Naive Bayes. *JOINTECS (Journal of Information Technology and Computer Science)*, *4*(2), 75–78. https://doi.org/10.31328/jointecs.v4i2.1010

Meilana, Astuti, Y., Rofni Wulandari, I., Sulistyowati, I., & Arie, B. M. (2021). Algoritma Naive Bayes untuk Mengklasifikasikan Kepribadian Siswa SMP Berdasarkan Tipologi Hippocrates-Galenus. *SISTEMASI: Jurnal Sistem Informasi*, *10*(2), 480–489. https://doi.org/10.32520/stmsi.v10i2.1339

Nurul Renaningtias, & Dyah Apriliani. (2021). Penerapan Metode Prototype Pada Pengembangan Sistem Informasi Tugas Akhir Mahasiswa. *Rekursif: Jurnal Informatika*, *9*(1), 92–98. https://doi.org/10.33369/rekursif.v9i1.15772

Prayama, D., Yuhefizar, & Amelia Yolanda. (2021). Protokol HTTPS, Apakah Benar-benar Aman? *Journal of Applied Computer Science and Technology*, *2*(1), 7–11. https://doi.org/10.52158/jacost.v2i1.118

Sari, N. K., Hutabarat, M. S., Susanti, E., Sopyan, M., & Yudi, A. (2020). Strategi Penanaman Karakter Islami Pada Siswa MTS Swasta Al Manar Medan Johor (Studi Kasus Selama Pembelajaran Daring). *AL-ULUM*, *1*(2), 147–164. https://doi.org/10.56114/al-ulum.v1i2.54

Sugiyono. (2019). *Metode Penelitian Kuantitatif Kualititaf dan R&D*. Alfabeta, CV.

Suparno. (2018). Analisis Faktor-Faktor Pembentuk Karakter SMART Siswa di Sekolah Islam Terpadu. *Jurnal Pendidikan Karakter*, *9*(1), 62–73. https://doi.org/10.21831/jpk.v8i1.21675

Lisdi Inu Kencana received the Bachelor of Computer Science (S.Kom) in Informatics Engineering from Universitas Dian Nuswantoro in 2022. His research interest includes data mining and web development.

Fauzi Adi Rafrastara received a Master degree from the Technical University of Malaysia Malacca (2011). He currently works as a lecturer at Universitas Dian Nuswantoro, Semarang, Indonesia. He published 9 books in Indonesia and Malaysia, and several papers in International Conferences and Journals. His research interest includes data and text processing, high-performance computing, and information security.

Cinantya Paramita received the Master of Electrical and Computer Engineering (M.Eng) degree in Electrical and Computer Engineering from South China Unversity and Technology (2014). Her research interests include the distance measurement method-based binocular vision approach.