International Journal of Application on Sciences, Technology and Engineering (IJASTE)
Volume 1, Issue 2, 2023.ISSN:2987-2499

Application for New Student Admission at SMKS Al-Akhyar II Jakarta with *C4.5 Algorithm*

Dwi Eka Meilan[1,a], Sutrisno Tri[1,b], Mulyawan Bagus [1,c].
[1] Informatics Engineering Department, Faculty of Information Technology, Universitas Tarumanagara.

[a] eka.535180139@stu.untar.ac.id, [b] tris@fti.untar.ac.id, [c] bagusm@fti.untar.ac.id

Submitted: January-February 2023, Revised: March 20 2023, Accepted: May 19, 2023

Abstract—

Al-Akhyar II Jakarta Vocational High School (Smk) is a school that has been accredited with A superior. However, the application of new student admissions still has some difficulties in knowing the potential of prospective students who will attend school. This makes schools have to select more students who will be accepted at Smks Al-Akhyar II Jakarta. Therefore, a method is needed to identify prospective new students at school. Data Mining with the C4.5 Algorithm can be used to select and classify prospective new students at school by making a decision tree based on existing data and selecting prospective students. Data on prospective students who register and take tests in the 2019 academic year, 2020 and 2021 will be selected using the C4.5 method which will later see students who are accepted at Smks Al-Akhyar II Jakarta. The results that have been tested using training data and comparing the training data and testing data have an accuracy of 97% using the C4.5 method. Tests were also carried out on 89 respondents who agreed that this design was carried out to overcome problems at Smks Al-Akhyar II Jakarta.

Keywords— Accuracy, C4.5, Test Score, Requirements, Selection of admissions.

I. INTRODUCTION

New Student Admissions at the Al-Akhyar II Jakarta Vocational High School (SMK) level are provided in one route, namely with tests and requirements that the school proposes. In this study, we will discuss the registration path at Smks Al- Akhyar II Jakarta. In the process of accepting new students at Smks Al-Akhyar II Jakarta, it is generally still done manually where prospective students come directly to school and register and take tests at school. So that it allows the results of the selection that is done manually will spend a long time.

The important thing for schools to do when accepting new students is to identify patterns from registration and tests that have been carried out, by classifying the variables that have been determined. One of the classification techniques that are often and easily used by research is the application of the C4.5 algorithm. The C4.5 algorithm used in this study is for the selection of new students at Smks Al-Akhyar II Jakarta.

The C4.5 algorithm used in this study for the selection of new student admissions at Smks Al- Akhyar II Jakarta. The results of this study using the C4.5 method is an application that can classify new students in a tree structure so that it can produce a rule.[1] The aims of this design are:

- This design system will apply the C4.5 method to classify data into the website that will be designed, namely new student admissions.
- The design system will use data on test scores and school conditions that will be entered and processed for the selection of students who will be accepted at Smks Al-Akhyar II Jakarta and to facilitate school staff in the process of accepting new students.

https://doi.org/10.24912/ijaste.v1.i2.379-386

The system designed is a program in the form of a website for new student admissions at Smks Al- Akhyar II Jakarta which will be divided into two categories, namely accepted or not accepted. Prior to the creation of this system, the author had to go directly to Smks Al-Akhyar II Jakarta to request the data needed for testing on a website-based application to be made. The data needed is in the form of student

data who register at Smks Al-Akhyar II Jakarta. Before asking for data, the author asks for a letter of application first to the staff of the faculty of informatics technology to be processed later on to collect the data needed for the design to be made.

II. METHODS

A. *Data Mining*

Data Mining is the science of the filtering data for information and knowledge retrieval, and it has recently developed new album of the applications and engendered an emerging discipline, its called Educational Data Mining.[2]

Data Mining [DM] is a method of using a number of data to obtain information that was previously unknown or found and used for decision making in a problem under study.

Data Mining is a process to explore added value in the form of information that so far it is not known manually from a database by digging with patterns of data whose purpose is to manipulate data into more valuable information obtained by extracting and recognizing important or interesting patterns from

the data contained in the database.[3]

B. *C4.5 Algorithm*

The C4.5 method is a data classification method with a decision tree technique that can process numerical data, C4.5 can also handle the value of an attribute that doesn't exist and produces rules that are easy to interpret and are the fastest algorithms among other algorithms.

The C4.5 algorithm is one of the algorithms that converts large facts into a decision tree. This algorithm is expected to be able to solve the problems that occur at SMKS Al-Akhyar II Jakarta.[4]

Data Mining with C4.5 Algorithm can be used to make predictions and classifications of prospective new students in school by making decision trees based on existing data and predicting new prospective students who want to go to SMKS Al-Akhyar II Jakarta.[5] Using this improved algorithm C4.5, the decision tree model is produced and the classification rules are extracted.[6] In algorithm C4.5 is also compares the accuracy of each algorithm so its have a

good accurancy.[7] There are several stages in making a decision tree in the C4.5 .algorithm:

- Training data is taken from historical data that has happened before or is called past data and has been grouped into certain classes.

- Counting the roots of the tree. The root will be taken from the selected attribute, by calculating the gain value for each attribute, the highest gain value will be the first root. Before calculating the gain value, first calculate the entropy value. To calculate the entropy value can be used the formula:

https://doi.org/10.24912/ijaste.v1.i2.379-386

Information:
S = Set of cases
n= Number of partitions S

ρ_i = Proportion to S

- Calculate the Gain value using equation 2 with the formula:

Information:
S = Set of cases A = Attribute
n = Number of attribute partitions A
$|Si|$= Number of cases on partition i

$|S|$ = Number of cases in S

- Repeat Step 2 and Step 3 until all records are complete.
- The decision tree partitioning process will stop when:
- All records in node N belong to the same class.
- There are no attributes in the partitioned record anymore.
- There are no records in the empty branch.

The C4.5 algorithm is used to determine the graduation of a student who will enroll in Smks Al- Akhyar II Jakarta. Parameters for the selection of graduation are based on categories that are already known to students. Experimental and evaluation results show that the Decision Tree C4.5 Algorithm is accurately applied to determine student graduation with an accuracy rate of 97%.

From the level of accuracy that has been tested, it is known that the design for accepting new students using the c4.5 algorithm has a good level of accuracy and can predict which students are accepted and not accepted at Smks Al-Akhyar II Jakarta. And can help problems that exist at school.

II. RESULT AND DISCUSSION

After all the experiments and testing data, it is known that the testing data has a level of accuracy that is not much different from the training data.

At the initial stage, this student admission system will conduct a trial where the training data used comes from the Al-Akhyar II Jakarta Vocational School. The data will be used for the classification calculation

process. The process of classifying student admissions will be calculated by applying the C4.5 method as a classification method for Data Mining.

After the calculations are done, the results of the student admissions selection will be displayed through the website that will be designed. The results of the selection are in the category of accepted or not

https://doi.org/10.24912/ijaste.v1.i2.379-386

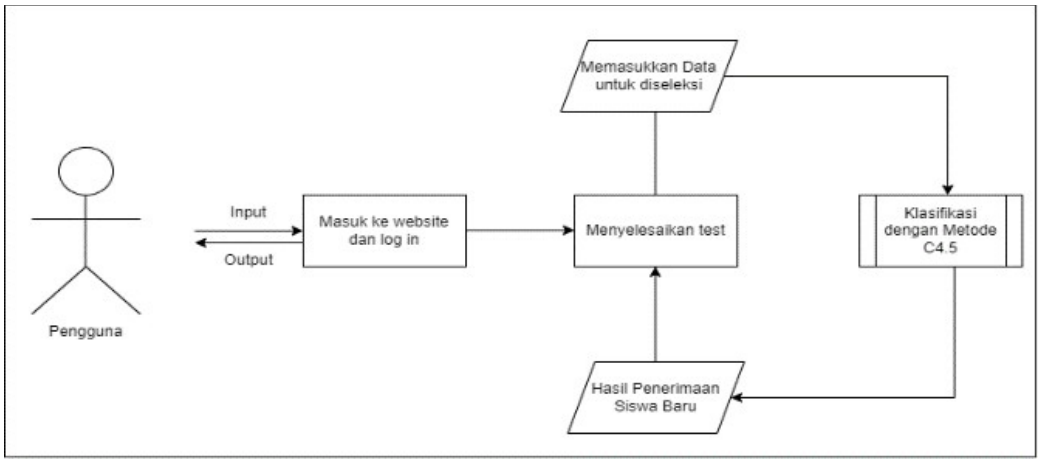

Figure 1. Schematic of the New Student Admission System

At the stage of collecting data samples, it will be done by asking for data from previous batch students who registered and took tests at Smks Al-Akhyar II Jakarta. The following is an example of a sample data used on 10 students.

Tabel 1 Sampel of Dataset New Student Admission

No	Age	Height	Test Result	Phobia Of	Status
1.	13	High enough	Good	No	**Accepted**
2.	15	High enough	Not Good	No	Not Accepted
3.	14	**Not high**	Good	Yes	Not Accepted
4.	13	Very	Good	No	**Accepted**
5.	13	**Not high**	Not Good	No	Not Accepted
6.	15	High enough	Good	No	**Accepted**
7.	14	Very	Good	Yes	**Accepted**
8.	13	**Not high**	Good	Yes	Not Accepted
9.	13	High enough	Not Good	No	Not Accepted
10.	13	**Not high enough**	Good	Yes	Not Accepted

https://doi.org/10.24912/ijaste.v1.i2.379-386

Website Testing

This test on the appearance of the new student admission application website at Smks Al-Akhyar II Jakarta was carried out to see whether the features made were in accordance with the designed system. This test was carried out using the C4.5 method. Testing is carried out on each module by seeing whether each module is appropriate or not.

Testing the appearance based on the new student admissions website at Smks Al-Akhyar II Jakarta was carried out using the C4.5 method. The following are the results of the user interface display on the new student admissions website at Smks Al-Akhyar II Jakarta:

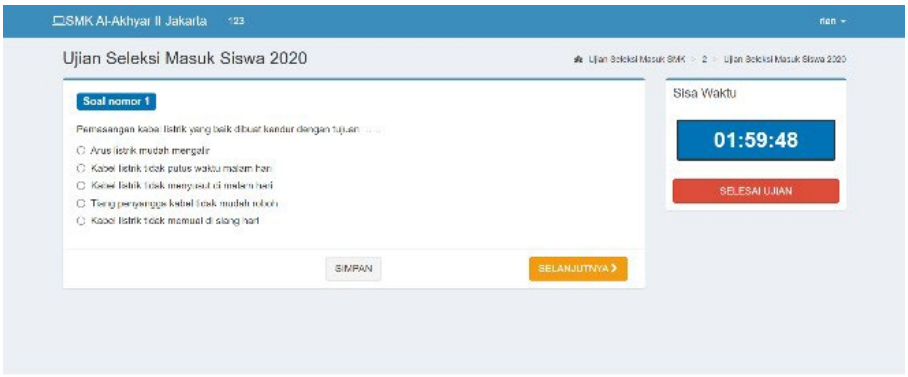

Figure 2. Display Interface Test (User)

Figure 3. Display Interface Form (User)

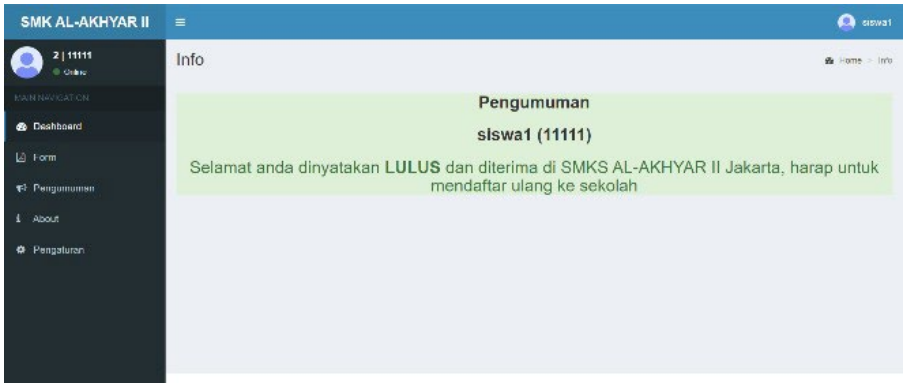

Figure 4. Display Interface Announcement (User)

https://doi.org/10.24912/ijaste.v1.i2.379-386

Apart from the user interface, this design also creates an admin interface and the following are the results of the admin interface on the new student admissions website at Smks Al-Akhyar II Jakarta:

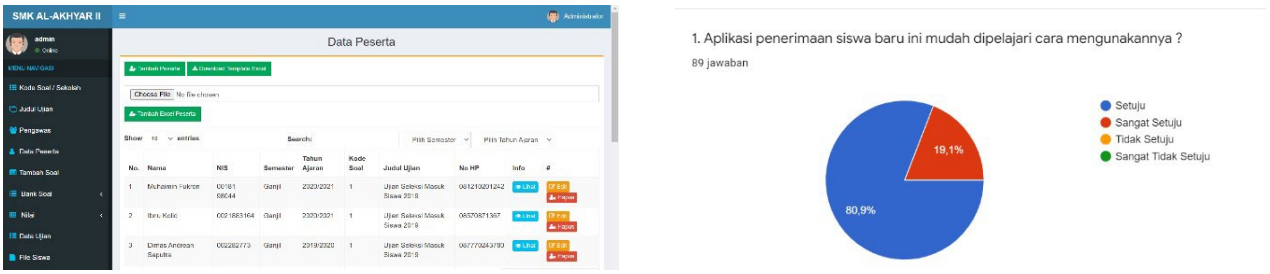

Figure 5. Display Interface Participant Data (Admin)

Figure 6. Display Interface Score (Admin)

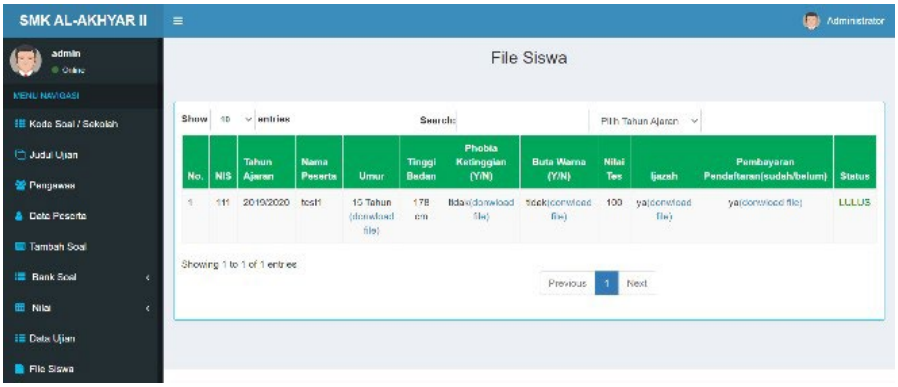

Figure 7. Display Interface Student File (Admin)

https://doi.org/10.24912/ijaste.v1.i2.379-386

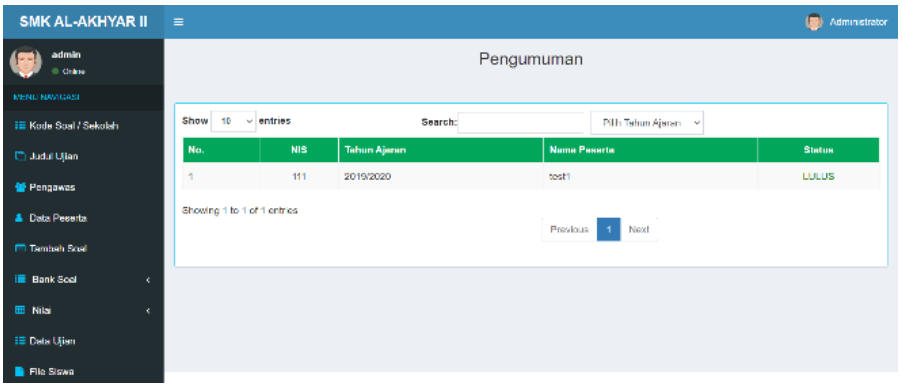

Figure 8. Display Interface Announcemen (Admin)

A. User Testing

To find out whether the design of this website is useful for schools and prospective students, a trial was conducted which received the following responses:

Figure 8. Display Interface Results Responded The display test based on the new student admissions website at Smks Al-Akhyar II Jakarta received 89 respondents, of which 89 people stated that they agreed to the new student admissions plan to be carried out.

IV.CONCLUSION

Based on the results of the calculation data and application testing above, it can be obtained as follows:

- The design of new student admissions that is made can provide answers in the form of results from students who are accepted at Smks Al-Akhyar II Jakarta.
- From the tests conducted by Smks Al-Akhyar II Jakarta, it is possible to find out what categories have a lot of influence on students who are accepted at Smks Al-Akhuat II Jakarta.
- The data used is data from Smks Al-Akhyar II Jakarta in the period 2019, 2020 and 2021. The data used and the results of the website designed have something in common, namely in the form of students who are accepted as students who can meet the requirements category of Smks Al- Akhyar II Jakarta.
- The results of the research above stated that the application designed was very helpful for new student admissions at Smks Al-Akhyar II Jakarta from 89 respondents. The results that have been tested are 97% comparison between fact data and system recommendation data using the C4.5 method.

V. REFERENCES

[1] Darmawan, Erlan. "C4. 5 Algorithm Application for Prediction of Self Candidate New Students in Higher Education." *Jurnal Online Informatika* 3, no. 1 (2018): 22-28.

[2] Yadav, Surjeet Kumar, and Saurabh Pal. "Data mining application in enrollment management: A case study." *International Journal of Computer Applications* 41, no. 5

https://doi.org/10.24912/ijaste.v1.i2.379-386

(2012).

[3] Setyanto, Arief, and Hanif Al Fattah. "Analisis Perbandingan Algoritma Decision Tree (C4. 5) Dan K-Naive Bayes Untuk Mengklasifikasi Penerimaan Mahasiswa Baru Tingkat Universitas."

[4] Hts, Dedek Indra Gunawan. "Penerapan Algoritma C4. 5 Untuk Penerimaan Siswa Baru Pada SMA Al-Azhar Medan." *J-SAKTI (Jurnal Sains Komputer dan Informatika)* 2, no. 2 (2018): 130-139.

[5] Winanjaya, Riki, Faisal Amir, and Rahmad Doni. "Penerapan Data Mining Untuk Memprediksi Penerimaan Peserta Didik Baru Menggunakan Algoritma C4. 5." In *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, pp. 1- 6. 2019.

[6] Gu, Ping, and Qi Zhou. "Student performances prediction based on improved C4. 5 decision tree algorithm." In *Emerging Computation and Information teChnologies for Education*, pp. 1-8. Springer, Berlin, Heidelberg, 2012.

[7] Supianto, Ahmad Afif, Alfi Julisar Dwitama, and Muhammad Hafis. "Decision tree usage for student graduation classification: A comparative case study in faculty of computer science brawijaya university." In *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*, pp. 308-311. IEEE, 2018.

https://doi.org/10.24912/ijaste.v1.i2.379-386