

# Prediction analysis of student interest in design learning using Naïve Bayes method

Jefri Junifer Pangaribuan<sup>1)</sup>, Ferawaty<sup>2)</sup>

<sup>1,2)</sup> Universitas Pelita Harapan, Medan, Indonesia

<sup>1)</sup>[Jefri.pangaribuan@uph.edu](mailto:Jefri.pangaribuan@uph.edu), <sup>2)</sup>[gold3n27@gmail.com](mailto:gold3n27@gmail.com)

**Submitted** : Oct 12, 2020 | **Accepted** : Mar 12, 2021 | **Published** : Apr 1, 2021

**Abstract**— Students' interest in a field is usually seen from the grades resulting from students' learning in the classroom. This is a common thing for parents and teachers to do. In the research conducted by researchers this time is in the field of design studied by students in the classroom in one semester, namely the subjects Photoshop and CorelDraw. The grades taken are grades that include the value of theory, practical grades, assignments, and quiz assignments obtained by students in the classroom. The four grades will be calculated until they get a provision on whether or not the student graduates in the subject. These values will be studied by researchers using the Naïve Bayes method, so that it can be known how much can be said to pass in this design lesson. Researchers conducted research using Rapid Miner program, where the data will be divided into 2 parts, namely by using some existing data as learning samples and the rest as test data. The results obtained from the experiment were 146 students graduated in the field of design, and 119 students failed. This suggests the experiment using the Naïve Bayes method was successful if the experiment data was entered a lot.

**Keywords**— CorelDraw; Design; Naïve Bayes; Rapid Miner; Photoshop

## BACKGROUND

To Solve The Problems That Exist In This World, So Many Researchers Use Artificial Intelligence To Assist Humans In Solving These Problems. From Easy Problems To Difficult Problems. Therefore, The Computer Is Needed To Help Solve The Problems. In Order For Computers To Assist Humans In Solving Problems, Many Researchers Use Artificial Intelligence (Ai) Assisted By Existing Methods To Help Solve Them. Researchers Conduct Studies, Research Related To Human Nature, Human Way Of Thinking, Etc. So That Computers Can Think And Move And Make Decisions To Humans.

In The Research Conducted By Researchers, Researchers Will Make Decisions About Determining Pupils' Interest In Design Which Is One Of The Types Of Subjects Studied In 2 Semesters In School. The Interests Of Students Are Commonly Determined By Parents Or Teachers, Sometimes Appropriately, Sometimes Not. Because The Interests That Teachers And Parents Decide Are Always Based On The Values And Daily Attitudes Shown In Students In Their Lives. Mistakes In The Determination Of Students' Interests, Will Sometimes Make Students Burdened In Their Lives. That's Because Students Will Try To Pursue What Is Expected Of Them To Achieve Success.

Researchers Tried To Use The Naïve Bayes Method To Help Determine Students' Interest In Design. The Naïve Bayes Method Is One Of The Methods That Researchers Use In Artificial Intelligence. The Naïve Bayes Method Is Widely Used By Researchers To Make Decisions Such As One Of The Journals Used By Ms. Ferawaty Titled "Analysis Of Accuracy Comparisons In Autism Identification With Svm And Naïve Bayes".

Due To Problems In The Determination Of Students' Interests In Design, The Researchers Created A Study With The Title "Analysis Of Junior High School Students In Design Lessons Using Naïve Bayes"

## LITERATURE REVIEW

Below can be seen the research that has been done by other researchers using the Naïve Bayes method:

1. Research conducted by Mr. Husni Naparin with the title "Classification of Interest of High School Students Using Naïve Bayes Method. The data used by the researchers is the average grades of

\*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Mathematics, Science, and IPS from students of SMAN 2 Banjarmasin in the 5th semester, the school curriculum in 2013. Mr. Husni's research using the Naïve Bayes method achieved a level of accuracy of 99.47% and an AUC value of 1,000.

2. Research conducted by Mr. Riszki Wijayatun Pratiwi and Mr. Yusuf Sulisto Nugroho with the title "Film Rating Prediction using Naïve Bayes Method". The study conducted assessments for use in film ratings based on film genre, film actor, language, color, film duration, country and others, which the results of the scattered research can be used by the director as a benchmark value for making films.
3. Research conducted by Mrs Ferawaty with the title "Accuracy Comparison Analysis In Autism Identification with SVM and Naïve Bayes". This study conducted a comparison to determine the accuracy of autism identification that occurs in children aged 3 years and under using the SVM and Naïve Bayes methods. The results of the study were obtained on average using SVM is 93.12% and the average result using the Naïve Bayes method is 73.34%.
4. Research conducted by Ms Dewi Simanjuntak with the title "Expert System for Detection of Malnutrition of Toddlers With Naïve Bayes Method". This study aims to detect malnutrition in toddlers aged 1-3 years using the Method Naïve Bayes Clasifier. In this study, there were 3 types of disease based on symptoms, namely Kwarshiorkor (P1), Marasmik-Kwarshiorkor (P2), Marasmus (P3) with 24 symptoms of malnutrition. The diagnosis is influenced by the symptoms of malnutrition that appear. From the cases discussed, the calculation of Naive Bayes Classifier on the diagnosis of malnutrition patients with 3 types of diseases, has symptoms number G3, G6, G12, G13 obtained the highest multiplication classification of value 0.0013168617, type kwarshiorkor disease.
5. Research conducted by Mrs. Naisha Rahma Indraswari and Mr. Yogiek Indra Kurniawan with the title "Birth Age Prediction Application With Naïve Bayes Method". This research was conducted to make a prediction application that will be able to help patients in knowing the age of birth and anticipate unwanted things in the future. The method used is Naïve Bayes method with variables included are factors experienced by pregnant women, including: maternal age, blood pressure, number of babies, delivery history, history of abortus / cureve, malnutrition, congenital diseases before pregnancy and problems during pregnancy. The result of this study is an application that can predict the age of birth with the highest application accuracy value at 78.69%, the highest precision value is at 70.14% and the highest recall value is at 63.64%.

### METHOD

The data obtained is data of students in grades 9, semester 1 and 2. The data obtained are theoretical data, and practice data. The total data obtained is 882 data where the data consists of 2 semesters, for photoshop and CorelDraw subjects.

The data processed at the beginning will be done data cleaning, where researchers do data cleaning by dumping empty or null data in the data obtained. The data that has been cleaned, will be determined whether the student graduated according to manual calculations, where the minimum value used is a value of 85. Determination of graduation for manual calculation can be seen in table 1 below.

Table 1. Determination of Information According to Original Data

29680	ADRIEL JONATHAN DAELI	90,76	Lulus
29681	ANDRE FERUZZI MARCO	91,00	Lulus
29682	ANNABELLE JURIYAN SOETIJOSO	84,68	Gagal
29683	AUDREY ELVARETTA CONG	87,40	Lulus
29684	AXEL WENDY	83,56	Gagal
29685	BERWYN DELROY	85,96	Lulus
29686	CALISTA	87,72	Lulus
29688	EDBERT VINCENT SALIM	89,44	Lulus
29689	ELBERT CHAILES	91,88	Lulus
29690	ERICA TANDY	92,20	Lulus

After determining the manual graduation of the initial data received by the researchers, it was carried out several stages conducted in this study using the formula naïve bayes method as below.

$$P(y | x) = \frac{P(X|y) \cdot P(y)}{P(X)}$$

Description:

y : The number of students who pass or failed

X : Subjects tested

P(X|y) : Probability hypothesis y based on condition X

The author performs a calculation of P(y) in which to determine the student who graduated or failed in the manual calculation of the entire data tested. The results obtained from Photoshop subjects were as many as 77 students who pass and 70 students who failed. As for Coreldraw subjects, the students who pass were as many as 75 students who graduated and 72 students who failed.

From the results of previous calculations, the subjects tested in this study, obtained probability with the results of:

$$P(X|y) = P(\text{MP} = \text{Photoshop} | y = \text{Pass}) = 77 / 441$$

$$P(X|y) = P(\text{MP} = \text{Photoshop} | y = \text{Failed}) = 70 / 441$$

$$P(X|y) = P(\text{MP} = \text{CorelDraw} | y = \text{Pass}) = 75 / 441$$

$$P(X|y) = P(\text{MP} = \text{CorelDraw} | y = \text{Failed}) = 72 / 441$$

From the calculation results obtained, the author multiplies the subjects with the provisions y = Pass, as well as subjects whose provisions y = failed, will be multiplied for both subjects tested. The calculations carried out are as follows

$$P(\text{MP} = \text{Photoshop} | y = \text{Pass}) \times P(\text{MP} = \text{CorelDraw} | y = \text{Pass}) = (77 / 441) \times (75 / 441) = 0.0297$$

$$P(\text{MP} = \text{Photoshop} | y = \text{Failed}) \times P(\text{MP} = \text{CorelDraw} | y = \text{Failed}) = (70 / 441) \times (72 / 441) = 0.0259$$

From the results tried, the researchers conducted an experiment on a student with a Photoshop subject grade = 85 who was declared pass, and a CorelDraw = 67.2 subject score that was considered a failure. From the results of the probability calculation obtained from the student's grades, obtained for Photoshop subjects with the category y = Pass = 1/441 and y = Failed = 0 / 441, while in the subjects CorelDraw, obtained the opposite value.

Then from the results of the previous probability, multiplication of subjects tested with the provision y = Pass, obtained a value = 0, as well as subjects with the provisions y = Failed, obtained also the value = 0. Because the result is 0, then compared with the multiplication result in the previous learning data, and the value is smaller, then Jane is considered "Failed" in design lessons.

## RESULT

In each study, the research data will be divided into 2, namely learning data and test data. The results of the study obtained from five tests with different data resulted in a table such as table 1 below. The table is a comparison of the original result with the prediction result using the Naïve Bayes method.

Table 1. Manual Results and Prediction Results Using Naïve Bayes Method

Trial	Manual Results		Prediction Results Using Naïve Bayes Method	
	Pass	Failed	Pass	Failed
1	7	2	9	0
2	22	5	0	27
3	45	45	28	62
4	101	87	70	118
5	139	126	146	119

The result of table 1 above can be viewed using the chart as shown figure 7 below.

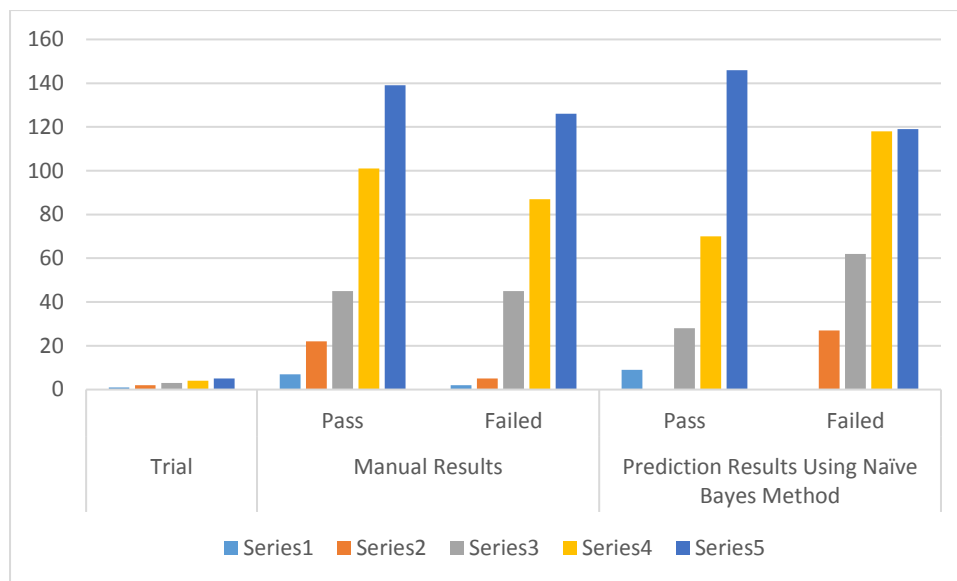


Figure 7. Comparison Graph

### DISCUSSION

Researchers conducted test 1. From the data entered, researchers divided the data into 90% for learning and 10% for testing. From the data shared, it was tested using the Naïve Bayes method. The results of manual calculations obtained as many as 7 people who graduated and 2 people who failed in the field of design. while from the results using the Method Naïve Bayes, can be seen the results obtained is all passed to enter into the design.

The researchers then conducted the 2nd test. Just like the first test, the data is also divided into 2 parts. 90% for learning, and 10% for testing. The results of manual testing obtained by students who graduated or interested in the field of design is 22 people, while students who failed or are not interested in the field of design is as many as 5 people. The results of the test using the Naïve Bayes method stated that all students graduated or were interested in the field of design.

Researchers conducted the 3rd test with the addition of more data from the 2nd experiment. The results of manual testing obtained that passed or interested in the field of design is 45 people, but who failed also as many as the successful, namely as many as 45 people. While the prediction results using the Method Naïve Bayes obtained the results of the pass is as many as 28 people, and who failed or not interested in the field of design is as many as 62 people. In the 3rd test, from the score data obtained, Naïve Bayes method stated that students who do not fit in the field of design was increased drastically, more than experiment 1 and experiment 2.

Researchers conducted the 4th test, with the addition of data more than tests 1, 2, and 3. The results of the manual calculation for test 4 were those that passed as many as 101 people, while the failed was as many as 87 people. But in testing using the Naïve Bayes method, the pass was as many as 70 people, while the failed was as many as 118 people. In the 4th test using the Naïve Bayes method also obtained more failed results than passed.

In the 5th test, the tester adds data to the test data. data obtained from the results of manual calculations, which passed as many as 139 people, failed as many as 126 people. While the predicted results are using the Method Naïve Bayes is 146 people who graduated and 119 people who failed.

### CONCLUSION

From the results of experiments conducted using the Naïve Bayes method, it can be seen, the more experimental data, the more successful the experiment is done. This can be seen in the 5th test, namely for students who graduated is as many as 146 people, and for students who have no interest in design is as many as 119 people. So experiments using the Naïve Bayes Method will be successful if more data is entered.

### REFERENCES

Annur, Haditsyah. 2018. KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES. ILKOM Jurnal Ilmiah. Vol 10 No 2

- Astuti, Yani Parti, Usman Sudibyo, Achmad Wahid Kurniawan, Yuniarsi Rahayu. 2018. ALGORITMA NAÏVE BAYES DENGAN FITUR SELEKSI UNTUK MENGETAHUI HUBUNGAN VARIABEL NILAI DAN LATAR BELAKANG PENDIDIKAN. SIMETRIS. Vol 9 No 1
- Basri, Hasan. 2019. IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MENGETAHUI POTENSI SISWA DALAM PEMAHAMAN MS. EXCEL. Jurnal SISTEMASI. Vol 8 No 3
- Chen, Shenglei, Geoffrey I. Webb, Linyuan Liu, Xin Ma. 2020. A NOVEL SELECTIVE NAÏVE BAYES ALGORITHM. Knowledge-Based Systems. Vol 192
- Ferawaty. 2016. ANALISIS PERBANDINGAN AKURASI DALAM IDENTIFIKASI AUTISM DENGAN SVM DAN NAÏVE BAYES. JSM. Vol 17 No 2
- Hadi, Abdul, Karim Karim, Agni Danaryanti. 2020. ANALISIS SELFT CONCEPT, KEAKTIFAN, DAN HASIL BELAJAR MATEMATIKA SISWA SMP. UrbanGreen Proceeding :Konferensi Nasional Pendidikan I. Vol 1
- Hartshorn, Scott. 2016. BAYES THEOREM EXAMPLES : A VISUAL GUIDE FOR BEGINNERS.
- Indraswari, Naisha Rahma, Yogie Indra Kurniawan. 2018. APLIKASI PREDIKSI USIA KELAHIRAN DENGAN METODE NAÏVE BAYES. Jurnal SIMETRIS. Vol 9 No 1
- Lasarudin, Alter, Wahyudin Hasim, Febriyanto Sako. 2019. SISTEM INFORMASI JURNAL UNIVERSITAS MUHAMMADIYAH GORONTALO DENGAN MENERAPKAN METODE NAÏVE BAYES. SemanTECH. Vol 1 No 1
- M. Kaviza. 2020. PERSEPSI KENDIRI MURID TERHADAP TUGASAN ANALISIS TEKS DALAM TEKS DALAM MATA PELAJARAN SEJARAH. e-BANGI : Jurnal Sains Sosial dan Kemanusiaan. Vol 17. No 5
- Manalu, Effrida, Fricles A. Sianturi, Mamed R. Manalu. 2017. PENERAPAN ALGORITMA NAÏVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTERIES. Jurnal Mantik Penusa. Vol 1 No2
- Manik, Fuzy Yustika, Kana Saputra Saragih. 2017. KLASIFIKASI BELIMBING MENGGUNAKAN NAÏVE BAYES BERDASARKAN FITUR WARNA RGB. IJCCS. Vol 11 No 1
- Mueller, John Paul, Luca Massaron. 2016. MACHINE LEARNING FOR DUMMIES. For Dummies.
- Mujilawati, Siti, Febrianti Widyahastuti, Nur Nafi'iyah. 2018. PERFORMANSI KLASIFIKASI KELAS SISWA MENGGUNAKAN NAÏVE BAYES DAN DECISION. Seminar Nasional Unisla
- Naparin, Husni. 2016. KLASIFIKASI PEMINATAN SISWA SMA MENGGUNAKAN METODE NAÏVE BAYES. SYSTEMIC. Vol 2 No 1
- Pratiwi, Rizki Wijayatun, Yusuf Sulianto Nugroho. 2016. PREDIKSI RATING FILM MENGGUNAKAN METODE NAÏVE BAYES. Jurnal Teknik Elektro. Vol 8 No 2
- Purnamasari, Indah, Karnita Afnisari. 2018. PERFORMANSI KLASIFIKASI DOSEN BERPRESTASI MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER. PARADIGMA Jurnal Komputer dan Informatika. Vol 20 No 2
- Purnamasari, Indah. 2020. ANALISIS KLASIFIKASI LOYALITAS SISWA LEMBAGA PENDIDIKAN TARI DENGAN METODE NAÏVE BAYES. Jurnal Ilmiah KOMPUTASI. Vol 19 No 1
- Richards, Ken. 2018. MACHINE LEARNING : FOR BEGINNERS – YOUR COMPREHENSIVE GUIDE FOR MARKOV MODELS, REINFORCES LEARNING, MODEL EVALUATION, SVM, NAÏVE BAYES, CLASSIFIER (VOLUME 3). CreateSpace Independent Publishing Platform
- Theobald, Oliver. 2017. MACHINE LEARNING FOR ABSOLUTE BEGINNERS. Scatterplot press
- Simanjuntak, Dewi, Anita Sindar. 2019. SISTEM PAKAR DETEKSI GIJI BURUK BALITA DENGAN METODE NAÏVE BAYES CLASSIFIER. Jurnal Infokar. Vol 1 No 2
- Syarli, Asrul Ashari Muin. 2016. METODE NAÏVE BAYES UNTUK PREDIKSI KELULUSAN (STUDI KASUS: DATA MAHASISWA BARU PERGURUAN TINGGI). Jurnal Ilmiah Ilmu Komputer. Vol 2 No 1
- Valdiviezo-Diaz, Priscila, Fernando Ortega, Eduardo Cobos, Raúl Lara-Cabrera. 2019. A COLLABORATIVE FILTERING APPROACH BASED ON NAÏVE BAYES CLASSIFIER. IEEE. Vol 7
- Wibowo, Indera Cahyo, Abd Charis Fauzan, Marshella Dwi Putri Yustiana, Fiqih Ainul Qhabib. 2019. KOMPARASI ALGORITMA NAÏVE BAYES DAN DECISION TREE UNTUK MEMPREDIKSI LAMA STUDI MAHASISWA. ILKOMNIKA: Jurnal of Computer Science and Applied Informatics. Vol 1 No 2