jurnal5

by dodi mulyadi

General metrics

15,449	2,327	219	9 min 18 sec	17 min 54 sec
characters	words	sentences	reading time	speaking time

Score

This text scores better than 78% of all texts checked by Grammarly

Writing Issues

109	50	59
Issues left	Critical	Advanced

Plagiarism

12 sources

10% of your text matches 12 sources on the web or in archives of academic publications

Writing Issues

73	**Correctness**	
2	Incomplete sentences	
16	Misspelled words	
3	Wrong or missing prepositions	
2	Faulty subject-verb agreement	
1	Unknown words	
1	Pronoun use	
11	Determiner use (a/an/the/this, etc.)	
14	Improper formatting	
1	Incorrect noun number	
12	Punctuation in compound/complex sentences	
7	Comma misuse within clauses	
1	Commonly confused words	
2	Confused words	
15	**Engagement**	
15	Word choice	
21	**Clarity**	
17	Passive voice misuse	
4	Wordy sentences	

Unique Words

Measures vocabulary diversity by calculating the percentage of words used only once in your document

26%

unique words

Rare Words

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

30%

rare words

Word Length

Measures average word length

4.3

characters per word

Sentence Length

Measures average sentence length

10.6

words per sentence

jurnal5

The Analysis of Major Interest at State Senior High School 1 Salem through Logistic Regression and K-Nearst Neighbor

1stTiani Wahyu Utami1, 2ndTestiana Deni Wijayatiningsih2, 3thMartyana Prihaswati3,

4nd Abdul Karim4

{tianiutami@unimus.ac.id1, testiana@unimus.ac.id2 ,

martyana@unimus.ac.id3}

Universitas Muhammaddiyah Semarang1,2,3, Universitas Islam Negeri Wali Songo4, Semarang City Central Java, Indonesia

Abstract. Education is a conscious and planned effort to create a learning atmosphere and learning process so that learners are actively developing their potential. At the high school level, learners will follow a major interest. Majors interest is conducted to provide students with opportunities to choose the subject of interest, deepen the subject matter and develop their potential in a flexible range. The State High School 1 Salem opened 2 majors for the continuity of the students' learning process, namely Science and Social. The classification analysis method used is binary logistic regression and K-NN. Based on the results obtained from logistic regression, the alleged factor to affect the majors' interest in SMA is the score of Science National Examination

and the students' relationship with their friends. The best classification method for majors' interest at State High School 1 Salem is the K-NN method.

Keywords: Senior High School Major, Logistic Regression, K-NN

1 Introduction

Education is a conscious and well-planned endeavor to create a learning atmosphere and learning process so that students actively develop their potential to possess religious spiritual strength, self-control, personality, intelligence, and the skills required by him, society, nation and State (Law of the Republic of Indonesia number 20 year 2003, article 1 paragraph 1). In the process of studying the students can choose the majors in their respective high school. There are high schools that open 3 majors, namely Science, Social, and Language. Besides, there is also high school that only open 2 majors, namely Science and Social. The determination of the opening majors in high school is back in their respective school policies. Schools play an important role in developing their students' potential according to their skills or majors [3]. Further, at the high school level, the learners will follow a major interest. Majors interest is conducted to provide students with opportunities to choose the subject of interest, deepen the subject matter and develop a variety of potential in it flexibly according to common basic skills (intelligence), talents, interests and personality characteristics without being constrained by the partition of the majors that are too rigid [5]. In this research the researchers took the research object at the State High School 1 Salem district of Pekalongan Central Java opened 2 majors for the continuity of the learning process, namely Science and Social. The possibility that will happen if the student has an error in the study is a low student learning achievement or can cause a mismatch with the majors that have been chosen by the student or previous students [1]. According to [6], the factors that influence the learners in

majors interest namely internal factors and external factors. In statistical sciences, many methods can be used to determine the influence of predictor variables on the response variables of a category [16].

Classification is one of the statistical methods to group or classify data that is arranged systematically [14]. Classification problems are often found in everyday life. Whether it's the classification of data in the academic, social, government, and other fields [17]. This classification problem arises when there are a number of measurements consisting of one or several categories that cannot be identified directly but must use a measure [15].

Several classification methods are regression methods of logistic binary and K-Nearest Neighbor. According to Hosmer and Lemeshow [2], the method of logistic regression is a statistical analysis method describing the relationship between a category-scaled response variable that has two or more categories (binaries) with one or more predictor variables. A binary (Variable response) is a response variable that is only 1 for the existence of a characteristic and 0 for the absence of such characteristics. K- Nearest Neighbor (K-NN) is a classification method that specifies categories based on the majority of categories in K-Nearest Neighbor [4]. K-NN is done by looking for group K objects in the training data closest to the object in the data testing [7]. In this research will be conducted analysis of the classification of high school in Salem, with logistic regression and K-NN method.

Since that time, the use of logistic regression has estimation of customer dissatisfaction [8], multivariate logistic regression analysis [9], and quality of logistic regression [10]. Previous research including, the study used K-NN to predict student vgraduation on time [13], recognition number of the vehicle plate using K-NN [12]. A classification system is expected to be able to classify all data sets correctly, but it cannot be denied that the performance of a

system is not 100% correct so a classification system must also measure its performance [18]. Generally, performance measurements are carried out with a cofusion matrix [19]. The comparison of the classifiers and using the most predictive classifier is very important. Each of the classification methods shows different efficacy and accuracy based on the kind of datasets [20]. This study uses the R program, using a logistic regression program package and KNN package [11].

2 Data and Method

2.1 Data and Research Variable

The data which were used in this research is the primary data obtained by distributing the questionnaire to the students of Sate High School 1 Salem Pekalongan Central Java. The total students were 224 students which were consisted of 98 for Social and 126 majoring in Science.

Table 1. Research Variable

Variable	Definition
Variable Response	
Majors interest (Y)	
1= Social	
2=Science	
Variable Predictor	

Math National Examination Score (x1)

Science National Examination Score (x2)

Language National Examination Score (x3)

The relationship of[47] students and their friends (x4)

1= Low

2= Middle

3 = High

The relationship of[48] students and their teachers (x5)

1= Low

2= Middle

3 = High

The relationship of[49] students and their family (x6)

1= Low

2= Middle

3 = High

Self motivation (x7)

1= Low

2= Middle

3 = High

2.2 Research Method

The research analysis steps is[50]:

1. Data retrieval continued with data encoding to become data ready

2. Analyzing data by Binary logistic regression method

3. Analyzing data by K-NN method

4. Comparing the classification results of both methods

3 Result and Discussion

3.1 Logistic Biner Regression

Regression logistic model which had been created:

πx=eg(x)1+eg(x) (1)

with :

gx=-20,4+0.015x1-0.062x2+0.32x3+1,206x4.1+0,987x4.2-0.361x5.1-0.88x5.2+0.271x6.1+0,792+0.008x7.1+22.354x7.2

Furthermore,it continued by testing the significance of the parameters either together or each of the predictor variables.

1. Likelihood Ratio Test

Hipotesis

113 H0 : β1 = β2 = β3 = β4 = β5 = β6 = β7 = β8 = β9 (predictor variable does not affect thr model together)

: minimum there is one βj ≠ 0, with j = 1, 2, ..., 9. (predictor variable affects the model together)

Significance level : α = 5%

Statistical Test : G=-2 Inlikelihood tanpa variabel bebaslikelihood dengan variabel bebas (2)

Test Criteria: H0 is rejected if G > χ2 (0,05;9)

The decision : because G=219,451 > 16,919 χ2 (0,1;7) so H0 is rejected

The conclusion: so, at significance level 5%, it can be concluded that predictor variable influences the model together.

2. Wald Test

Hipotesis

H0 : βj = 0 (variable j does not affect the model)

H1 : βj ≠ 0, untuk j = 1, 2, ..., 9. (variable j affects the model)

Significance level: α = 5%

Statistical Test: W=βjSe (βj) (3)

Test Criteria: H0 is rejected if W > χ2 (0,1;9)=3,481

Wald score for each variable can be seen as follows.

Table 2. Wald Test

Predictor Variable

Wald (W)

Sig

Decision

Math National Exam Score (x1)

0.290

0.590

accepted Ho

Science National Exam Score (x2)

3.688

0.055

rejected Ho

English National Exam Score (x3)

0.575

0.448

accepted Ho

The relationship between students and their friends (x4)

Low (x4.1)

0.897

0.044

rejected Ho

Middle (x4.2)

6.276

0.012

rejected Ho

The relationship between students and their teachers (x5)

Low (x5.1)

0.184

0.668

accepted Ho

Middle (x5.2)

2.009

0.156

accepted Ho

The relationship between students and their family (x6)

Low (x6.1)

0.125

0.723

accepted Ho

Middle (x6.2)

2.581

0.108

accepted Ho

Self Motivation (x7)

Low (x7.1)

0.00

1.00

accepted Ho

Middle (x7.2)

0.00

0.997

accepted Ho

Conclusion: Based on Table 2, the equivalent of 10% significance is concluded that the variables X2 and x4 affect the model[51] whereas the variables x1, x3, X5, X6 and[52] X7 do not affect[53] the model. Factors that influence the major[54] interest of state[55] high school students Salem I is the science national examination score and students' relationship with their friends.[56]

Furthermore, the establishment of the final model uses an influential variable on the model. The final models obtained are:

πx=eg(x)1+eg(x)

with :

gx=1,785-0.027x2+0,72x4.1+0.783x4.2

3.2 Logistic Regression Clasification[57]

By calculating the probability of each observation, then it was obtained the classification result for binary logistic regression methods:

Table 3. Clasification[58] Result

Observed

Predicted

Social Major

Science Major

Social Major

Science Major

48

9

50

117

Overal Percentage

73.7 %

Table 3 exposes that there are 48 students majoring in Social that are predicted to enter the Social Major, 50 students of Social Major who are predicted to enter the Science Major, 9 students majoring in Science, which is predicted to study in Social Science and 117 students of Science majoring in Science. Prediction error of 26.3% and accuracy of prediction of 73.7%. It can be deduced accuracy classification with Logistic Regression method of 73.7%.

3.3 K-NN Clasification

This study used 224 data, 80% data used as training data and 20% used as data testing. To determine whether the student enters where to use 3 nearby data so that the value K = 3. To measure the accuracy of the method used Confussion matrix. Table 4 follows the Matrix confussion table for the K-NN method:

Table 4. Confussion Matrix

Observed

Predicted

Social Major

Science Major

Social Major

Science Major

64

5

34

121

Overal Percentage

82.6 %

Table 4 explicates that there are 64 students majoring in Social that are predicted to enter the Social Major, 34 students of Social who are predicted to enter the Science Major, 5 students majoring in Science and is predicted to study in Social Major and 121 students of Science majoring in Science Majors. The error prediction is 17.4% and the accuracy of prediction is 82.6%. All in all, the accuracy of classification with K-NN method in the School interest data of Salem majors is 82.6%.

4 Conclusion

To sum up the analysis that has been done, there is a conclusion that both using binary logistic regression method, the alleged factor to affect the interest in state high school majors is the Science National Examination score and the relationship of the students with their friends. The interest classification system of State High School 1 Salem which resulted from binary logistic regression method has a precision is 73.7%. The interest classification system of State High School 1 Salem which resulted from the K-NN method has a accuracy of 82.6%. The best classification method for interest in Salem High School is the K-NN method.

References

[1] Departemen Pendidikan Nasional, Panduan Penilaian Penjurusan Kenaikan Kelas dan Pindah Sekolah. Direktorat Pendidikan Menengah Umum, Jakarta (2004).

[2] Hosmer, D. W. and Lemeshow S. Applied Logistic Regression. United States of American: Sons Inc (2000)

[3] Kementrian Pendidikan dan Kebudayaan, Modul Pelatihan Implementasi Kurikulum 2013 untuk Guru BK atau Konselor (2013).

[4] Liu, B. Web Data Mining: Exploring Hyperlinks, Contents and Usage Data. Berlin: Springer (2007)

[5]Sudiarto, A. Pedoman Arah Peminatan Draf 2 (2013) [online]. http://www.akursudianto.com/wp-content/uploads/2019/06/pedoman-arah-peminatan-draf-2.pdf

[6] Syah, M. Psikologi Pendidikan dengan Pendekatan Baru. Bandung: Remaja Rosdakarya (2009)

[7] Wu X, Kumar V. The Top Ten Algorithms in Data Mining. New York:CRC Press. (2009)

[8] Shahin, A. and Janatyan, N. Estimation of Customer Dissatisfaction Based on Service Quality Gaps by Correlation and Regression Analysis in a Travel Agency. International Journal of Business and Management Vol. 6, No. 3. (2011)

[9] Tanboga IH, Kurt M, Isik T, Kaya A, Ekinci M, Aksakal E, et al. Assessment of multivariate logistic regression analysis in articles published in Turkish cardiology journals. Turk Kardiyol Dern Ars 2012;40:129-34. (2012)

[10] Kumar, et al.: Quality of logistic regression in Indian journals Indian Journal of Public Health, Volume 60, Issue 2 (2016)

[11] Trevino,V. dan Falciani,F., GALGO: an R package for multivariate variable selection using genetic algorithms. Bioinformatics, 22, 1154–1156 (2006)

[12] Hidayah, M. R., Aklis, I. & Sugiharti, E. (2017). Recognition Number of The Vehicle Plate Using Otsu Method and K-Nearest Neighbour Classification. Scientific Journal of Informatics, 4(1), 66 – 75 (2017).

[13] Banjarsari, M.A, el all, Penerapan K-Optimal Pada Algoritma Knn untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan IP Sampai Dengan Semester 4 , Jurnal Ilmu Komputer (KLIK) Volume 02, No.02 (2015)

[14] Baby, N & P.L.T. Customer Classification And Prediction Based On Data Mining Technique. International Journal of Emerging Technology and Advanced Engineering, 2(12), 314-18 (2012).

[15] Ferdousy, E.Z., Islam, M.M. & Matin, M.A. Combination of Naive Bayes Classifier and K-Nearest Neighbor (cNK) in the Classification Based Predictive Models. Computer Science and Information Science, 6(3), 48-56 (2013).

[16] Chiang, L. H., dan Pell, R. J. Genetic Algorithms Combined with Discriminant Analysis for Key Variable Identification. Journal of Process Control, 14, 143-155 (2004).

[17] Liao, S. H., Chu, P. H., & Hsiao, P. Y. Data mining techniques and applications - A decade review from 2000 to 2011. Expert Systems with Applications, 39(12), 11303–11311 (2012).

[18] Bujlow, T., Riaz, T., & Pedersen, J. M. A method for classification of network traffic based on C5.0 machine learning algorithm. International Conference on Computing, Networking and Communications, ICNC'12, 237–2419 (2012).

[19] Gorunescu, Florin. (2011). Data Mining: Concepts and Techniques. Verlag berlin Heidelberg: Springer (2011).

[20] Y.S. Kim, Comparision of the decision tree, artificial neural network, and linear regression methods based on the number and types of independent variables and sample size, Journal of Expert Systems with Application, Elsevier, pp. 1227-1234 (2008).

1.	~~Nearst~~ → Nearest	Misspelled Words	Correctness
2.	testiana@unimus.ac.id2-,	Improper Formatting	Correctness
3.	~~major~~ → significant	Word Choice	Engagement
4.	~~Majors~~ → Majors', Major's	Incorrect Noun Number	Correctness
5.	*is conducted*	Passive Voice Misuse	Clarity
6.	, and	Punctuation in Compound/Complex Sentences	Correctness
7.	~~2~~ → two	Improper Formatting	Correctness
8.	religious,	Comma Misuse within Clauses	Correctness
9.	religious-spiritual	Misspelled Words	Correctness
10.	, and	Comma Misuse within Clauses	Correctness
11.	the year	Determiner Use (a/an/the/this, etc.)	Correctness
12.	studying,	Punctuation in Compound/Complex Sentences	Correctness
13.	~~3~~ → three	Improper Formatting	Correctness
14.	a high, or the high	Determiner Use (a/an/the/this, etc.)	Correctness
15.	~~open~~ → opens	Faulty Subject-Verb Agreement	Correctness
16.	~~2~~ → two	Improper Formatting	Correctness
17.	~~important~~ → essential	Word Choice	Engagement
18.	~~major~~ → significant	Word Choice	Engagement
19.	*is conducted*	Passive Voice Misuse	Clarity

20.	~~basic~~ → necessary	Word Choice	Engagement
21.	research,	Comma Misuse within Clauses	Correctness
22.	~~2~~ → two	Improper Formatting	Correctness
23.	, namely	Punctuation in Compound/Complex Sentences	Correctness
24.	*be used*	Passive Voice Misuse	Clarity
25.	*are often found*	Passive Voice Misuse	Clarity
26.	~~. This~~ → , this	Incomplete Sentences	Correctness
27.	~~a number of~~ → several, some, many	Wordy Sentences	Clarity
28.	~~method~~ → process, purpose, practice	Word Choice	Engagement
29.	~~categories~~ → classes, types	Word Choice	Engagement
30.	*is done*	Passive Voice Misuse	Clarity
31.	~~object~~ → purpose	Word Choice	Engagement
32.	an analysis	Determiner Use (a/an/the/this, etc.)	Correctness
33.	~~conducted analysis of~~ → analyzed	Wordy Sentences	Clarity
34.	an estimation	Determiner Use (a/an/the/this, etc.)	Correctness
35.	~~vgraduation~~ → graduation	Misspelled Words	Correctness
36.	*is expected*	Passive Voice Misuse	Clarity
37.	*be denied*	Passive Voice Misuse	Clarity
38.	~~a system~~ → an order	Word Choice	Engagement

39.	, so	Punctuation in Compound/Complex Sentences	Correctness
40.	~~performance~~ → production	Word Choice	Engagement
41.	*are carried*	Passive Voice Misuse	Clarity
42.	~~very important~~ → critical, significant, essential	Word Choice	Engagement
43.	*were used*	Passive Voice Misuse	Clarity
44.	~~Sate~~ → State	Confused Words	Correctness
45.	~~which~~ → who	Pronoun Use	Correctness
46.	*were consisted*	Passive Voice Misuse	Clarity
47.	~~of~~ → between	Wrong or Missing Prepositions	Correctness
48.	~~of~~ → between	Wrong or Missing Prepositions	Correctness
49.	~~of~~ → between	Wrong or Missing Prepositions	Correctness
50.	~~is~~ → are	Faulty Subject-Verb Agreement	Correctness
51.	model,	Punctuation in Compound/Complex Sentences	Correctness
52.	, and	Comma Misuse within Clauses	Correctness
53.	~~affect~~ → modify	Word Choice	Engagement
54.	~~major~~ → primary, significant	Word Choice	Engagement
55.	~~of state~~ → of state	Improper Formatting	Correctness
56.	*Factors that influence the major interest of state high school students Salem I is the science national*	Incomplete Sentences	Correctness

	examination score and students' relationship with their friends		
57.	~~Clasification~~ → Classification	Misspelled Words	Correctness
58.	~~Clasification~~ → Classification	Misspelled Words	Correctness
59.	48 students are majoring	Wordy Sentences	Clarity
60.	*are predicted*	Passive Voice Misuse	Clarity
61.	*are predicted*	Passive Voice Misuse	Clarity
62.	~~9~~ → nine	Improper Formatting	Correctness
63.	*is predicted*	Passive Voice Misuse	Clarity
64.	~~Science and~~ → Science and	Improper Formatting	Correctness
65.	the Logistic	Determiner Use (a/an/the/this, etc.)	Correctness
66.	~~Clasification~~ → Classification	Misspelled Words	Correctness
67.	, and	Punctuation in Compound/Complex Sentences	Correctness
68.	~~where~~ → were	Commonly Confused Words	Correctness
69.	~~use~~ → apply	Word Choice	Engagement
70.	~~3~~ → three	Improper Formatting	Correctness
71.	~~Confussion~~ → Confusion	Misspelled Words	Correctness
72.	~~confussion~~ → confusion, conversion	Misspelled Words	Correctness
73.	~~method~~ → way, process, purpose	Word Choice	Engagement
74.	~~Confussion~~ → Confusion, Confession	Misspelled Words	Correctness
75.	~~Overal~~ → Overall	Confused Words	Correctness

76.	64 students are majoring	Wordy Sentences	Clarity
77.	*are predicted*	Passive Voice Misuse	Clarity
78.	*are predicted*	Passive Voice Misuse	Clarity
79.	~~5~~ → five	Improper Formatting	Correctness
80.	*is predicted*	Passive Voice Misuse	Clarity
81.	, and	Punctuation in Compound/Complex Sentences	Correctness
82.	the K-NN	Determiner Use (a/an/the/this, etc.)	Correctness
83.	up,	Comma Misuse within Clauses	Correctness
84.	*been done*	Passive Voice Misuse	Clarity
85.	the binary	Determiner Use (a/an/the/this, etc.)	Correctness
86.	~~in state~~ → in-state	Misspelled Words	Correctness
87.	, which	Punctuation in Compound/Complex Sentences	Correctness
88.	the binary	Determiner Use (a/an/the/this, etc.)	Correctness
89.	method,	Punctuation in Compound/Complex Sentences	Correctness
90.	, which	Punctuation in Compound/Complex Sentences	Correctness
91.	method,	Punctuation in Compound/Complex Sentences	Correctness
92.	~~method~~ → way, process	Word Choice	Engagement

93.	~~a accuracy~~ → an accuracy	Determiner Use (a/an/the/this, etc.)	Correctness
94.	, and	Comma Misuse within Clauses	Correctness
95.	~~untuk~~ → Untuk	Misspelled Words	Correctness
96.	~~atau~~ → Atau	Misspelled Words	Correctness
97.	, and	Punctuation in Compound/Complex Sentences	Correctness
98.	*dengan*	Unknown Words	Correctness
99.	: CRC	Improper Formatting	Correctness
100.	~~Janatyan~~ → Jonathan	Misspelled Words	Correctness
101.	, V	Improper Formatting	Correctness
102.	, F.	Improper Formatting	Correctness
103.	the Otsu	Determiner Use (a/an/the/this, etc.)	Correctness
104.	~~untuk~~ → Untuk	Misspelled Words	Correctness
105.	~~4.~~,	Improper Formatting	Correctness
106.	~~cNK~~ → NK	Misspelled Words	Correctness
107.	the C5.0	Determiner Use (a/an/the/this, etc.)	Correctness
108.	, and	Comma Misuse within Clauses	Correctness
109.	~~berlin~~ → Berlin	Misspelled Words	Correctness
110.	*Education is a conscious and planned effort to create a learning atmosphere and learning process*	The Professional Competency Teachers Mediate the Influence ... https://files.eric.ed.gov/fulltext/EJ1181987.pdf	Originality

111.	*Since that time, the use of logistic regression has*	Logistic Regression: An Option for a Management Research?	Originality
112.	*The comparison of the classifiers and using the most predictive classifier is very important. Each of the classification methods shows different efficacy and accuracy based on the kind of datasets*	Comparison of Classification Methods Based on the Type of ... https://pdfs.semanticscholar.org/2ce0/664cfcb32461b900dd9e889cbbb2259c503e.pdf	Originality
113.	*H0 : β1 = β2 = β3 = β4 = β5 = β6 = β7 = β8 = β9*	h0 1 2 3 4 5 6 7 8 9 10 11 12 t2882 2817 2752 2687 2622 ... https://www.coursehero.com/file/p5nsf3/h0-1-2-3-4-5-6-7-8-9-10-11-12-t2882-2817-2752-2687-2622-2557-2492-2427-2362/	Originality
114.	*4] Liu, B. Web Data Mining: Exploring Hyperlinks, Contents and Usage Data.*	Opinion Mining on Internet Primary Bank with Online News ... https://link.springer.com/chapter/10.1007/978-3-319-92285-0_41	Originality
115.	*Quality Gaps by Correlation and Regression Analysis in a Travel Agency.*	Quality Gaps by Correlation and Regression Analysis in a ... http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.653.8899	Originality
116.	*Assessment of multivariate logistic regression analysis in articles published in Turkish cardiology journals. Turk Kardiyol Dern Ars*	Archives of the Turkish Society of Cardiology https://www.archivestsc.com/jvi.aspx?un=TKDA-65807	Originality
117.	*Hidayah, M. R., Aklis, I. & Sugiharti, E. (2017). Recognition Number of The Vehicle Plate Using Otsu Method and K-Nearest Neighbour Classification. Scientific Journal of Informatics, 4(1), 66 – 75*	K-Nearest Neighbor and Naive Bayes Classifier Algorithm in ... https://journal.unnes.ac.id/nju/index.php/sji/article/view/9	Originality
118.	*Customer Classification And Prediction Based On Data Mining Technique. International Journal of Emerging Technology and Advanced Engineering, 2(12), 314-18*	K-Nearest Neighbor and Naive Bayes Classifier Algorithm in ... https://journal.unnes.ac.id/nju/index.php/sji/article/view/9	Originality

119.	*Combination of Naive Bayes Classifier and K-Nearest Neighbor (cNK) in the Classification Based Predictive Models. Computer Science and Information Science, 6(3), 48-56*	K-Nearest Neighbor and Naive Bayes Classifier Algorithm in ... https://journal.unnes.ac.id/nju/index.php/sji/article/view/9	Originality
120.	*Data mining techniques and applications - A decade review from 2000 to 2011. Expert Systems with Applications,*	Data mining techniques and applications – A decade review ... https://www.deepdyve.com/lp/elsevier/data-mining-techniques-and-applications-a-decade-review-from-2000-to-D491ES84Z0	Originality
121.	*Y.S. Kim, Comparision of the decision tree, artificial neural network, and linear regression methods based on the number and types of independent variables and sample size, Journal of Expert Systems with Application, Elsevier,*	Comparison of Classification Methods Based on the Type of ... https://pdfs.semanticscholar.org/2ce0/664cfcb32461b900dd9e889cbbb2259c503e.pdf	Originality