JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

IMPLEMENTATION OF THE MODIFIED K-NEAREST NEIGHBOR (MKNN) METHOD FOR CLASSIFICATION OF MAJORS CONCENTRATION SELECTION IN LEARNING

Yuda Irawan[1)], Refni Wahyuni[2)], Uci Rahmalisa[3)], Herianto[4)]
[1,3,4]Department of Information System, STMIK Hang Tuah Pekanbaru
[2]Department of Computer Science, STMIK Hang Tuah Pekanbaru
Email: yudairawan89@gmail.com[1]*, refniabid@gmail.com[2], ucirahmalisa89@gmail.com[3], herianto.sy@gmail.com[4]
*Corresponding Author

Abstract

Department of Communication studies UIN SUSKA Riau, Indonesia currently applies the selected of interest (concentration) to students aimed at directed stu-dents to more focus on certain courses based on their interests and academic abili-ties. The concentration that exist in the Department of Communication stuides are public relations, broadcasting and journalistic. Determination of concentration se-lection is carried out by the head of the department. However, to select the right concentration for students, Department have several problems in selection the concentration process. This is because students' data and value of courses must be classified first, then by conducting an interview to determine the level of stu-dent interest. So that with this process requires a long time. For this reason, a system that can help the selection process of concentration is needed. In the datamining, there is a classification to allow users to group data based on their class. In this study classification will used the Modified K-Nearest Neighbor (MKNN) algorithm. In this study, the parameters used are the values of 6 courses namely introduction to public relations, introduction to journalism, the basics of broadcasting, news writing, mass communication and advertising and the interest of the students concerned. The test was carried out on a concentration classifica-tion system that had been built using confusion matrix. Confusion matrix results obtained are the highest accuracy in the 90% data training scenario and 10% data test with k = 2 and k = 3 values that is equal to 78.08%.

Keywords: Communication studies, Concentration, Selecting Department, MKNN, Learning

INTRODUCTION

Sultan Syarif Kasim State Islamic University (UIN) is a State Islamic University in Riau, Indonesia in accordance with the Regulation of the Minister of Religion of the Republic of Indonesia Number 9 of 2013 concerning Organization and Work Procedure of UIN Suska Riau and Regulation of the Minister of Religion of the Republic of Indonesia Number 74 of 2013 concerning Amendments to PMA RI No. 9/2013 on the Organization and Work Procedure of UIN Suska Riau, then UIN Suska Riau has 8 faculties, one of which is the

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

Faculty of Da'wah and Communication Sciences. The Department of

Communication is currently applying the choice of interest (concentration) to students aimed at directing students to focus more on certain subject areas based on their interests and academic abilities. The fields of concentration in the Communication Science department are public relations, broadcasting and journalism.

Based on the objectives of the concentration selection program, it is necessary so that students have high motivation in developing their talents and abilities. According to [1], said that in the learning process students who are interested in learning activities will have tougher efforts than students who are less interested. Based on the results of interviews conducted with the head of the department at the Faculty of Communication Sciences, UIN SUSKA RIAU, he said that the choice of this concentration is very important because it will be a benchmark in the student learning process. By selecting the concentration it is hoped that later students can focus more on developing personal abilities and interests. which are owned. According to him, in the process of choosing the wrong concentration, it also has a big influence or can harm the student[2]. This election is expected to maximize the talents, interests, potentials and existing talents of the individual[3].

The choice of the concentration of the department in Communication Studies is carried out in the odd semester (three) the determination of the concentration selection is made by the head of the department[4]. However, with the implementation of the student concentration selection, the Department of Communication has experienced several obstacles in the process of selecting the concentration[5]. This is because it is still done by classifying all student data and related prerequisite course scores first, then by conducting interviews to see how much interest the student is[6]. So that this process requires a relatively long time[7]. Judging from the problems above, it is necessary to have a solution where later the system to be made is expected to be able to overcome these problems. And can help the department in classifying the choice of concentration of students in the Department of Communication Sciences. Then for this classification the data mining process is very necessary. Data mining is a term used to describe the discovery of knowledge in databases. Data mining is a process that uses statistical techniques, mathematics, artificial intelligence, and machine learning to extract and identify useful information and related knowledge from various large databases[8],[9]. As for the related research that [10] conducted research with the same case using the C4.5 algorithm with 239 data, this data is divided into two, namely 70% is used as training data and 30% as testing data. The results obtained using the C4.5 algorithm tested with the Confusion Matrix obtained an accuracy value of 81.94%. Then in 2015 [11] also conducted research Using the K-Nearest Neighbor Classifier

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

Method. In this study, he used data from majors in 2014 as many as 160 students. Furthermore, the data is divided into two parts, namely training data and test data with a percentage of 60%: 40% and the accuracy results obtained are 79.68%. Whereas in this study the authors used more percentage variations, to see the effect on the level of accuracy, namely 90%: 10%.

In this study, the authors will use data from the selection of student concentrations in the Department of Communication Science UIN SUSKA RIAU batch 2015 and 2016, namely a total of 733 data with details of 181 journalistic concentration selection data, 295 Broadcasting concentration selection data and 257 Public Relations concentration selection data[13]. This data will then be grouped into training data and test data. In the calculation of MKNN all parameters used will be transformed into a number where 3 concentrations are journalism (1), broadcasting (2), public relations (3) then the value of prerequisite courses such as A (4), A- (3,7), B + (3,3), B (3), B- (2,7), C + (2.3), C (2), D (1), E (0)[14]. As well as using the classification method, namely using the Modified K-Nearest Neighbor (MKNN) algorithm to classify the concentration determination because the MKNN algorithm has a very good accuracy value, this can be seen from research conducted[15]. This research uses the K-Nearest Neighbor (KNN) development method, namely the Modified K-Nearest Neighbor (MKNN) where MKNN has the advantage of being able to improve the performance of the KNN method by adding a preprocessing stage to the training data, namely the validity value and processing it with weight voting[16]. Therefore the author raises the title "Application of Modified K-Nearest Neighbor (MKNN) for the Classification of Concentration Selection in the Department of Communication UIN SUSKA RIAU".

METHOD

In this stage the researcher uses the Waterfall method, because this method is a method that is widely used by software developers. According to [17], the waterfall model is a classic model that is systematic, sequential in building software. The name of this model is actually "Linear Sequential Model". This model is often referred to as the "classic life cycle" or the waterfall method. This model is included in the generic model in software engineering and was first introduced by [18] around 1970 so that it is often considered obsolete, but is the most widely used model in Software Engineering (SE). This model takes a systematic and sequential approach. It is called a waterfall because step by step, you have to wait for the completion of the previous stage and run sequentially.

The phases in the Waterfall Model according to the Pressman reference:

Figure 1. Waterfall

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

a. Communication (Project Initiation & Requirements Gathering)
Before starting any technical work, it is necessary to communicate with customers in order to understand and achieve the goals to be achieved. The result of this communication is project initialization, such as analyzing problems encountered and collecting necessary data, as well as helping to define software features and functions. Additional data collection can also be taken from journals, articles, and the internet.

b. Planning (Estimating, Scheduling, Tracking)
The next stage is the planning stage which describes the estimation of technical tasks to be carried out, the risks that may occur, the resources needed to create the system, the work products to be produced, the work scheduling to be carried out, and the tracking of the system work process.

c. Modeling (Analysis & Design)
This stage is the stage of designing and modeling system architecture that focuses on designing data structures, software architecture, interface displays, and program algorithms. The goal is to better understand the big picture of what will be done. System design is carried out by making system designs such as making dataflow diagrams (DFD), databases and interfaces based on problem analysis in classifying the concentration selection of the Communication Science department using the Modified K-Nearest Neighbor method.

d. Construction (Code & Test)
This construction stage is the process of translating the design form into machine-readable code or form / language. After the coding is complete, testing is carried out on the system and also the code that has been created. The goal is to find errors that may occur to be corrected later. In this system the programming language used is the php myadmin language.

e. Deployment (Delivery, Support, Feedback) Deployment stage is the stage of implementing software to customers, periodic software maintenance, software repair, software evaluation, and software development based on the feedback provided so that the system can continue to run and develop according to its function[19]. Testing using the Black Box is testing carried out directly by the User.

Data collection in this study will use observation, interview, and literature study techniques. This method aims to obtain data related to research on the Application of Modified K-Nearest Neighbor (MKNN) for the Classification of Concentration Selection in the Department of Communication UIN SUSKA RIAU.

1. Observation `
It is a method that observes directly in order to obtain the required data such as how to group data manually, how long it takes to group the data manually, and how long the interview process is carried out for concentration selection in the Communication Science department of Uin Suska Riau[20].

2. Interview

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

This method is a method that is carried out by conducting interviews related to the process of choosing the concentration of the Communication Science department of Uin Suska Riau, so that the data obtained is more accurate, in the interview process the party used as a resource is the chairman of the Department of Communication Science, Uin Suska Riau, namely Mr. .Sos.I., MA.

3. Data Analysis

In the analysis of data analysis problems, data collection will be used to analyze the problems that occur and then processed using the Modified K-Nearest Neighbor method. The explanation of the data used is as follows:

a. The data used is data from the Department of Communication, namely data from the selection of concentrations in the UIN SUSKA RIAU Communication Science in 2015 and 2016.
b. The parameters used in this study were 6 (six) grades of related subjects and the interests of these students.

4. Knowledge Discovery in Database (KDD) stages

At this stage, it describes the classification of the results of the student academic ability value data using the Modified K-Nearest Neighbor method. Following are the steps that will be carried out:

a. Data Selection
   This stage is carried out to select the data that will be used in the research. Data selection was carried out so that it was easy to classify the concentration selection
b. Data Cleaning
   At this stage, data cleaning is carried out where unnecessary data will be cleaned. Data cleaning is done by removing unnecessary attributes later in the concentration classification process.
c. Data Transformation
   At this stage the input data used are normalized first so that the data is in the range [0-1] so that the data distribution is not too far away.
d. Classification using Modified K-Nearest Neighbor (MKNN)
   In this study, the method used is the classification method in data mining, namely Modified K-Nearest Neighbor (MKNN). This method classifies the choice of concentration majoring in Communication Studies based on the training data that has the closest distance. To determine the closest distance, it is necessary to calculate using the Euclidean formula, determine the k value, validate all training data and perform weight voting calculations on all test data. After all the calculation process with MKNN is carried out, the results of the concentration choice classification will be obtained, whether based on the value of the course and also the interest of the students themselves obtained from the interview, so that it can be entered into the concentration of Public Relations, Broadcasting or Journalism.

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

The following are the steps in the modified k-nearset neighbor method:

a. Determine the value of k
b. Calculate validity
c. Calculate weight voting
d. Determine the majority class of the k training data with the highest weight voting.
e. Generates a classification model to determine the class of the test data. The output obtained is the classification result of the concentration determination data.

5. Literature study

This method is used to support the observation method that has been carried out in data collection and finding references related to the research carried out such as the theory of data mining, the Modified K-Nearest Neighbor (MKNN) method, the KDD process in data mining, and the programming language used is PHP MyAdmin.

RESULT AND DISCUS

Data Analysis

This research will use data on the results of student concentration selection in the Department of Communication Science UIN SUSKA RIAU batch 2015 and 2016, namely a total of 733 data with details of 181 journalistic concentration selection data, 295 Broadcasting concentration selection data and 257 Public Relations concentration selection data. This data will then be grouped into training data and test data.

In this study, test data and training data are grouped according to the percentage of the amount of training data and test data in a test scenario. In this study, 5 test scenarios were carried out with the following training and test data percentages:

1. Scenario 1 (90% training data: 10% test data), in this scenario, the training data is 660 data while the test data is 73 data. Training data and tested data were taken randomly from 733 available data.
2. Scenario 2 (80% training data: 20% test data), in this scenario, the training data is 586 data while the test data is 147 data. Training data and tested data were taken randomly from 733 available data.
3. Scenario 3 (70% training data: 30% test data), in this scenario, the training data is 513 data while the test data is 220 data. Training data and tested data were taken randomly from 733 available data.
4. Scenario 4 (60% training data: 40% test data), in this scenario, the training data is 440 data while the test data is 1293 data. Training data and tested data were taken randomly from 733 available data.
5. Scenario 5 (50% training data: 50% test data), in this scenario, the training data is 367 data while the test data is 366 data. Training data and tested data were taken randomly from 733 available data.

At the implementation stage, there are several limitations, including the following:

a. The research data used is data on 6 subject values related to the choice of

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

concentration and data on student interest in the desired concentration.

b. The training data and test data used in the classification process were obtained randomly from the transformed concentration data.

The interface implementation will be carried out based on the design that has been done. The interface implementation in this study can be seen as follows:

Implementation of KDD Menu Pages

The interface implementation on the KDD menu consists of 4 sub menus, namely initial data, data selection, data cleaning and data transformation. The following is the implementation of the interface for the initial data sub menu:

Figure 2 Implementation of Initial Data Sub Menu

The following is the interface implementation for the data selection sub menu:

Figure 3 Implementation of the Data Selection sub menu

The following is the interface implementation for the data cleaning sub menu:

Figure 4 Implementation of the Data Cleaning sub menu

Here is the interface implementation for the data transformation sub menu:

Figure 5 Implementation of the data transformation sub menu

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

Implementation of the Classification Menu Interface

The implementation of the interface on the Classification menu consists of 2 sub menus, namely training and testing. The following is the interface implementation for the training sub menu:

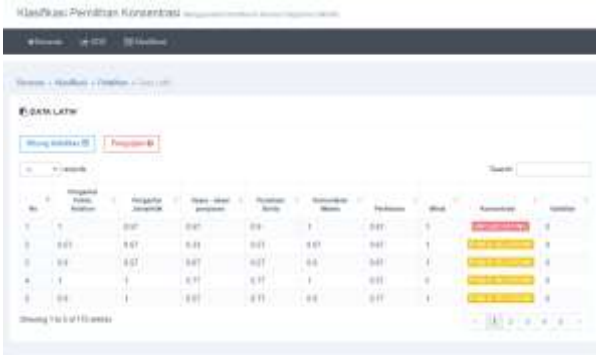

Figure 6 Implementation of the Training sub menu

The following is the interface implementation for the test data sub menu:

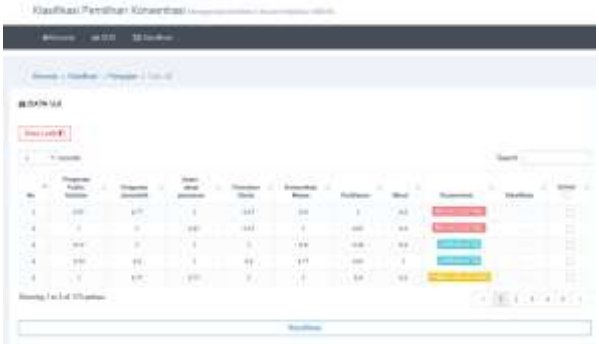

Figure 7 Implementation of the Testing sub menu

Testing the Confusion Matrix

Accuracy testing is done by dividing all existing data into training data and test data. In each test scenario, the k value is tested. Namely k = 1, k = 2, k = 3, k = 4 and k = 5. In this accuracy test, the data are classified into 3 components. Namely the concentration of journalism, broadcasting and public relations.

a. Testing Scenario 1 (90% training data: 10% Test data)

Testing scenario 1, the amount of training data used is 733 data with a specification of 660 training data and 73 test data. The test results in scenario 1 for k = 1 can be seen in table 5.7 below:

Table 1. Testing Scenario 1 (k = 1)

No	Id Data	Concentration	Classification	Result
1	4	*Broadcasting*	*Public Relations*	Wrong
2	12	*Broadcasting*	Journalism	Wrong
3	17	*Broadcasting*	*Broadcasting*	Right
4	29	*Public Relations*	*Public Relations*	Right
5	39	*Journalism*	*Public Relations*	Wrong
6	52	*Journalism*	*Journalism*	Right
7	70	*Journalism*	*Journalism*	Right
8	75	*Public Relations*	*Broadcasting*	Wrong
9	85	*Broadcasting*	*Broadcasting*	Wrong
73	733	*Public Relations*	*Public Relations*	Right

Based on the results of tests carried out on the test data in table 1 above, it can be obtained that the value of the Confusion Matrix for testing scenario 1 against the value of k = 1 is used as follows:

Table 2. Confusion Matrix scenario 1 (K = 1)

	Journa	*Broadca*	*Publi*

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

		lism	*sting*	*c Relati ons*
Classific ation	Journali sm	8	4	0
	Broadca sting	4	24	3
	Public Relation s	4	4	22
Amount		16	32	25

Based on the configuration matrix in table 2 above, the accuracy value for the test results in scenario 1 and with a value of K = 1 can be calculated as follows:

$$Accuracy = \frac{8 + 24 + 22}{8 + 4 + 0 + 4 + 24 + 3 + 4 + 4 + 22} \times 100\%$$

$$Accuracy = 73.97\ \%$$

By using calculations as done above, the classification accuracy can be calculated for the next scenario. So that the overall test results are obtained for all scenarios. The test results are described in more detail in ANNEX B.

b. Test Result Accuracy

In table 3 The following shows the success rate of the 5 test scenarios that have been carried out.

Table 3 The level of accuracy of the results of the 5 test scenarios

No	Value Classification Accuracy	Variation of Testing				
		K=1	K=2	K=3	K=4	K=5
1	Scenario 1 (90% training data : 10% test data)	73.97%	78.08%	78.08%	73.97%	76.71%
2	Scenario 2(80% training data: 20% test data)	68.03%	64.63%	70.07%	68.71%	65.99%
3	Scenario 3(70% training data: 30% test data)	64.09%	63.18%	60.91%	61.82%	60.45%
4	Scenario 4(60% training data: 40% test data)	64.51%	66.89%	66.21%	68.94%	65.19%
5	Scenario 5(50% training data: 50% test data)	60.11%	60.11%	59.56%	61.75%	59.02%

Based on table 3 above, it can be concluded that the average test accuracy for each k value and for each test scenario can be seen in table 4. below:

Table 4. Average Accuracy

Average Accuracy of Value K=1	66.14%
Average Accuracy of Value K=2	66.58%
Average Accuracy of	66.97%

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

JURNAL IPTEKS TERAPAN
Research of Applied Science and Education V15.i2 (127-139)

ISSN : 1979-9292
E-ISSN : 2460-5611

Value K=3	
Average Accuracy of Value K=4	67.04%
Average Accuracy of Value K=5	65.47%
Average Scenario Accuracy 1	76.16%
Average Scenario Accuracy 2	67.49%
Average Scenario Accuracy 3	62.09%
Average Scenario Accuracy 4	66.35%
Average Scenario Accuracy 5	60.11%

Scenario 1 test graph (90% training data: 10% Test data)

The following are the classification results based on the Scenario 1 test chart (90% training data: 10% test data) based on the results of the classification carried out.

Figure 8 Graph of Test Results

CONCLUSION

Based on the results of the tests that have been carried out, the conclusions of testing in this study are obtained, namely: The system built is in accordance with the previous analysis and design. All system components can functionally run properly. And put out the output as expected. The highest testing accuracy is in scenario 1 with the value of k = 2 and k = 3, which is equal to 78.08%. Meanwhile, the highest aver-age accuracy based on the test scenario is in scenario 1 which is 76.16%. The highest average accuracy based on the k value is found at the k = 4 value, which is 67.04%. The test results show that the majority of journalistic concentrations are misclassi-fied. This is because the amount of data on journalistic concentration is less than that of broadcasting and public relations.

AKNOWLEDGEMENT

The author would like to thank the STMIK Hang Tuah Pekanbaru for being willing to provide support in conducting this research.

REFERENCE

[1] LOU, Yi. Storage and Allocation of English Teaching Resources Based on k-Nearest Neighbor Algorithm. *International Journal of Emerging Technologies in Learning (iJET)*, 2019, 14.17: 102-113. https://doi.org/10.3991/ijet.v14i17.11188

[2] AN, Yingbo; XU, Meiling; SHEN, Chen. Classification Method of Teaching Resources Based on Improved KNN Algorithm. *International Journal of Emerging Technologies in Learning (iJET)*, 2019, 14.04: 73-88. https://doi.org/10.3991/ijet.v14.i04.10131

[3] ZAIN, Jasni Mohamad, et al. Data Mining for Education Decision

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

ISSN : 1979-9292
E-ISSN : 2460-5611

Support: A Review. *International Journal of Emerging Technologies in Learning*, 2014, 9.6. http://dx.doi.org/10.3991/ijet.v9i6.3950

[4] A. Kumar, K. Vengatesan, R. Rajesh, M. Parthibhan and A. Singhal, "Review of Gene Subset Selection using Modified K-Nearest Neighbor Clustering Algorithm," 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2018, pp. 570-574. doi: 10.1109/ICSSIT.2018.8748667.

[5] XIANGSHENG, Kong. Big Data X-Learning Resources Integration and Processing in Cloud Environments. *International Journal of Emerging Technologies in Learning*, 2014, 9.5. http://dx.doi.org/10.3991/ijet.v9i5.3783

[6] C. Li, Y. Gu, F. Li and M. Chen, "Moving K-Nearest Neighbor Query over Obstructed Regions," 2010 12th International Asia-Pacific Web Conference, Busan, 2010, pp. 29-35 doi: 10.1109/APWeb.2010.28.

[7] NA, Wei. A Data Mining Method for Students' Behavior Understanding. *International Journal of Emerging Technologies in Learning (iJET)*, 2020, 15.06: 18-32. https://doi.org/10.3991/ijet.v15i06.13175

[8] C. Li, Y. Gu, F. Li and M. Chen, "Moving K-Nearest Neighbor Query over Obstructed Regions," 2010 12th International Asia-Pacific Web Conference, Busan, 2010, pp. 29-35 doi: 10.1109/APWeb.2010.28.

[9] H. Zhang, S. Kiranyaz and M. Gabbouj, "Data Clustering Based on Community Structure in Mutual k-Nearest Neighbor Graph," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, 2018, pp. 1-7 doi: 10.1109/TSP.2018.8441226.

[10] C. Yan and C. Qixiang, "A Novel Parallel Processing for Continuous k-Nearest Neighbor Queries," 2009 International Conference on Environmental Science and Information Application Technology, Wuhan, 2009, pp. 593-596 doi: 10.1109/ESIAT.2009.75.

[11] Okfalisa, I. Gazalba, Mustakim and N. G. I. Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification," 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, 2017, pp. 294-298 doi: 10.1109/ICITISEE.2017.8285514.

[12] Z. Yan, Y. Chen and J. Zhao, "Using improved K-nearest neighbor method to identify anti- and pro-apoptosis proteins," 2015 8th International Conference on Biomedical Engineering and

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021

ISSN : 1979-9292
E-ISSN : 2460-5611

Informatics (BMEI), Shenyang, 2015, pp. 554-559
doi: 10.1109/BMEI.2015.7401566.

[13] Y. T. Guo and B. Luo, "An automatic image annotation method based on the mutual K-nearest neighbor graph," 2010 Sixth International Conference on Natural Computation, Yantai, 2010, pp. 3562-3566
doi: 10.1109/ICNC.2010.5584164.

[14] Z. Yan, Y. Chen and J. Zhao, "Using improved K-nearest neighbor method to identify anti- and pro-apoptosis proteins," 2015 8th International Conference on Biomedical Engineering and Informatics (BMEI), Shenyang, 2015, pp. 554-559
doi: 10.1109/BMEI.2015.7401566.

[15] S. Nutanong, R. Zhang, E. Tanin and L. Kulik, "V*-kNN: An Efficient Algorithm for Moving k Nearest Neighbor Queries," 2009 IEEE 25th International Conference on Data Engineering, Shanghai, 2009, pp. 1519-1522
doi: 10.1109/ICDE.2009.63.

[16] KAUSAR, Samina, et al. Mining Smart Learning Analytics Data Using Ensemble Classifiers. *International Journal of Emerging Technologies in Learning (iJET)*, 2020, 15.12: 81-102.

[17] Y. T. Guo and B. Luo, "An automatic image annotation method based on the mutual K-nearest neighbor graph," 2010 Sixth International Conference on Natural Computation, Yantai, 2010, pp. 3562-3566
doi: 10.1109/ICNC.2010.5584164.

[18] T. Yuwono, A. Franz and I. Muhimmah, "Design of Smart Electrocardiography (ECG) Using Modified K-Nearest Neighbor (MKNN)," 2018 1st International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, 2018, pp. 1-5
doi: 10.1109/CAIS.2018.8441983.

[19] Pressman, R.S.2015. Rekayasa Perangkat Lunak: Pendekatan Praktisi Buku I. Yogyakarta: Andi

[20] Fitriastuti, F., Rahmalisa, U., & Girsang, A. S. (2019, March). Multi-criteria decision making on succesfull of online learning using AHP and regression. In Journal of physics: Conference series (Vol. 1175, No. 1, p. 012071). IOP Publishing.

Submitted : 21 Januari 2021 – Accepted : 20 Mei 2021 – Published : 01 Juni 2021