ALGORITHM C4.5 IN CLASSIFYING HEALTH OF CAT

Nurlindasari Tamsir[1], Asrul Syam[2], Nirwana[3],Asmah Akhriana [4], Indo Intan[5], Vivi Rosida[6]

[1,2,3,4,5]Department of Informatics, Dipa Makassar University Makassar, Indonesia
[6]Department of Mathematics, STKIP Andi Matappa Pangkep, Indonesia

[1]nurlindasari.tamsir@dipanegara.ac.id, [2]assyams03@dipanegara.ac.id, [3]nirwana2006@dipanegara.ac.id, [4]asmah.a @dipanegara.ac.id,[5] indo.intan @dipanegara.ac.id, [6]vivirosida@stkip-andi-matappa.ac.id

***Abstract*—One of the activities in the TIU (Technical Implementing Unit) of Office of Animal Husbandry and Animal Health is to carry out the examination of pet health, record and issue an Animal Health Certificate (AHC). The object in this research was cat, where the examination was still performed by laboratory test, using paper as a form of diagnosis carried out by a vet and not using the technology. Therefore, an application to provide decision in determining healthy cat that also implements the algorithm C4.5 based on android system was designed. This system was able to perform diagnosis of cat diseases quickly based on previous medical record, with a training history (training data) as many as 44 cases. Then the classification process would strengthen the results of cat disease diagnosis such as how to deal with disease that had similar symptoms so as to facilitate the Office of Animal Husbandry and Animal Health, as a comparison to issue an Animal Health Certificate. Based on the Black Box testing, the functionality of the module was in accordance with the needs of the system, while the accuracy testing generated the percentage value of 93.18%. This shows that the algorithm C4.5 has a good accuracy to determine healthy cat.**

***Keywords*: Algorithm C4.5; Cat; Healthy.**

I. INTRODUCTION

Technical Implementing Unit (TIU) of the Office of Animal Husbandry and Animal Health, South Sulawesi Province is a government agency whose job is managing the animal husbandry sector concerning all care of animals, treatment of animals, animal health services, prevention of animal diseases, disease rejection, and medic conservation. One of the activities in the TIU of Office of Animal Husbandry and Animal Health is to carry out the examination of pet health, record and issue an Animal Health Certificate (AHC). Animals that have been through health examination are cat, dog, sugar glider, chicken, rabbit, squirrel and animals that are examined directly at the site, namely cow, horse, sheep and goat.

Cat owners/guardians who want to pass the traffic between the Provinces with their pets need AHC. Before issuing a AHC, the animals are vaccinated first by examining the animals' condition. Examination on the TIU of Office of Animal Husbandry and Animal Health has not utilized the information technology, and only use paper media as a material for deciding healthy cat determination. The medical examination is carried out by examining the physical condition and feces through the lab to find out which animal is experiencing infectious symptoms (contagious). If the results of medical examination carried out show that the cat is sick/showing infectious symptoms, TIU of Office of Animal Husbandry and Animal Health gives an option for the cat owners/guardians to do the treatment and care in the Office of Animal Husbandry and Animal Health or in other clinics. If the result of the cat health examination is healthy/does not show any infectious disease symptoms, then the vaccination is given and the Animal Health Certificate (AHC) is issued as a requirement for traffic between Provinces.

Medical record data from March to December 2019 showed that cat was a type of animal that was often examined with the highest frequency of number of cases by 15. Based on the data from the examination, there are 2 (two) types of condition, among others, general condition and physical condition. Common condition includes agile and not agile while the physical condition includes how to walk (disabled, normal, and limping), nostrils (normal wet and dry), eyes (normal, redness and irritation), ears (normal, disabled, dirty, inflammation and infection), fur (normal, flea and excessive hair loss), skin (normal, itching, ringworm, wound and scaly), digestion (not found abnormal (NFA) or normal, vomiting and reduced appetite), respiratory (NFA), heart (NFA), feces (normal, diarrhea, parasite) and body temperature (NFA).

Algorithm C4.5 is an algorithm that performs classification or grouping, predictive so it can be used to reach a decision by assessing several classifications that will determine the outcome of decision [1][2][3] and is applied through Android as an application.

Research on the types of diseases found in cats has been carried out, including the type of Feline Panleukopenia (FP) which is an infectious disease caused by a small virus. This disease causes severe depression, vomiting, diarrhoea, sharp decrease in circulating white blood cells that causes death[4]

In the following year, the type of disease Leishmaniosis was caused by the Leishmania infantum virus that attacks dogs and cats. The virus can be found in wounds on the head and neck and less often on the trunk and legs. Treatment of this disease with antibiotics and corticosteroids[5][6]. Subsequent studies in dogs and cats found pituitary tumors with the type of somatotroph adenoma and corticotrophic adenoma. Symptoms include abdominal enlargement, polyphagia, skin and muscle wasting, and lethargy[7].

. The difference in this study and previous studies is in the method used, which is Algorithm C4.5, based on numerical and categorical attributes consisting of rules to divide a heterogeneous population into smaller with the accuracy-test to measure the accuracy value on the system by comparing the expert diagnosis results with the system diagnosis results so as to produce an accuracy rate of 93.18%.

Algorithm C4.5 or decision tree is a classification algorithm data with numerical and categorical attributes consisting of rules to divide a number of heterogeneous population into smaller, more homogeneous by paying attention to the goal variables[8][9]. The goal variables are grouped and the decision tree model leads to the probability calculation of each record to the categories or classify the record by grouping them in one class. The data in the decision tree are expressed in the form of table with attribute and record. Attribute states a parameter created as a criterion in the tree formation [10][11].

II. METHOD

The research stage was started by collecting data and analyzing the problem which were performed through observation and interview directly, designing using Unified Modelling Language (UML)[12][13], analyzing the data using algorithm C4.5 and testing included Blackbox testing. The research objects were cats, where the medical record data from March to December 2019 with the highest frequency of the number of cases of 15 from 44 data source samples as the training data included the general condition and physical condition outlined in Table 1 below.

TABLE I
Cat Examination Condition Data

Condition		
General	Agile Limp	
Physical	How to walk	Disabled, Normal Limping
	Nostrils	Normal, Wet Dry
	Eyes	Normal, Redness, Irritation
	Ears	Normal, disabled, dirty, inflammation and infection
	Fur	Normal, flea and excessive hair loss
	Skin	Normal, itching, ringworm, wound and scaly
	Digestion	Not find abnormal (TDA) or normal, vomiting and reduced appetite
	Respiratory	TDA
	Heart	TDA
	Body temperature	TDA
	Feces	Normal, diarrhea, parasite

The decision tree was created using algorithm C.45 by changing the form of a table into a tree model based on the Entropy and Gain value[14]

To determine the Entropy value [15] a formula in the following equation was used (1):

$$Entropy(S)=\sum_{i=1}^{n}\left(-p_i*\log_2 p_i\right) \quad (1)$$

Where, S is a set of cases, A is feature, n is the number of partitions, S and p_i is proportion of S_i to S [16]

Next, to determine the Gain value[17][18], it was according to the formula in equation (2)

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

Where S is a set of cases, A as attribute, n as the number of partitions of attribute A, $|S_i|$ is the number of cases on the partition to the-I and |S| is the number of cases in S.

To determine the Gain value, first find the Entropy value (total) then, calculate the Entropy value for each attribute of common conditions which are agile and limp.

a. The total value was calculated by equation (1) as follows:

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

$$\text{Entropy(Total)}=\left(-\frac{27}{44}*\log_2\left(\frac{27}{44}\right)\right)+\left(-\frac{17}{44}*\log_2\left(\frac{17}{44}\right)\right)$$

$$\text{Entropy(Total)}=\left(-\frac{27}{44}*\left(\frac{\log\frac{27}{44}}{\log 2}\right)\right)+\left(-\frac{17}{44}*\left(\frac{\log\frac{17}{44}}{\log 2}\right)\right)$$

$$\text{Entropy(Total)}=\left(-0.61363636364\left(\frac{-2120889123}{0.30102999566}\right)\right)+\left(-0.3863636366\left(\frac{-0.4130037551}{0.30102999566}\right)\right)\text{Entropy(Total)}=\left(-0.61363636364(-0.7045441164)\right)+\left(-0.3863636366(-1.3719687774)\right)$$

$$\text{Entropy(Total)}=0.4323338896+0.5300788414$$

$$\text{Entropy(Total)}=0.962412735$$

The Entropy (total) value obtained was 0.97 based on 44 samples including the unhealthy status by 27 cases and the healthy status by 17 cases.

b. The calculation of general condition (agile) was carried out by equation (1) as follows:

$$\text{Entropy(Agile)}=\left(-\frac{7}{24}*\log_2\left(\frac{7}{24}\right)\right)+\left(-\frac{17}{24}*\log_2\left(\frac{17}{24}\right)\right)\text{Entropy(Agile)}=\left(-\frac{7}{24}*\left(\frac{\log\frac{7}{24}}{\log 2}\right)\right)+\left(-\frac{17}{24}*\left(\frac{\log\frac{17}{24}}{\log 2}\right)\right)$$

$$\text{Entropy(Agile)}=\left(-0.29166666667\left(\frac{-0.5351132017}{0.30102999566}\right)\right)+\left(-0.70833333333\left(\frac{-0.1497623203}{0.30102999566}\right)\right)$$

$$\text{Entropy(Agile)}=\left(-0.29166666667(-1.7776075787)\right)+\left(-0.70833333333(-0.4974996594)\right)$$

$$\text{Entropy(Agile)}=0.5184688771+0.35239559205$$
$$\text{Entropy(Agile)}=0.870864469$$

Data on the number of agile cases of 44 samples were 7 cases with unhealthy status, and 17 cases for healthy status, so the obtained value for agile attribute was 0.87.

c. The calculation of general condition (limp) was carried out based on the equation (1) as follows:

$$Entropy(\text{Limp})=\left(-\frac{20}{20}*\log_2\left(\frac{20}{20}\right)\right)+(-\frac{0}{20}*\log_2\left(\frac{0}{20}\right))$$

$$\text{Entropy(Limp)}=\left(-\frac{20}{20}*\left(\frac{\log\frac{20}{20}}{\log 2}\right)\right)+(-\frac{0}{20}*\left(\frac{\log\frac{0}{20}}{\log 2}\right))$$

$$\text{Entropy(Limp)} = \left(-1\left(\frac{0}{0{,}30102999566}\right)\right) + \left(-0\left(\frac{0}{0{,}30102999566}\right)\right)$$

$$\text{Entropy(Limp)} = \left(-1\left(\frac{0}{0{,}30102999566}\right)\right) + \left(-0\left(\frac{0}{0{,}30102999566}\right)\right)$$

$$\text{Entropy(Limp)} = \left(-1(0)\right) + (0(-0)$$

$$\text{Entropy(Limp)} = 0 + 0$$

$$\text{Entropy(Limp)} = 0$$

The limp entropy value obtained was 0, with a total sample of 44 cases including 20 unhealthy status and 0 healthy status.

The Entropy value above then was substituted into equation (2) to produce the Gain value, so that the obtained value was 0.4, where of the 44 case samples, the agile was 24 cases and the limp was 20 cases.

$$\text{Gain(Total, general condition)} = \text{Entropy(Total)} - \sum_{i=1}^{n} \frac{|\text{general condition}|}{|\text{total}|} * \text{Entropy(general condition)}$$

$$\text{Gain(Total, general condition)} = 0.962412735 - \left(\left(\frac{24}{44} * 0.8708764469\right) + \left(\frac{20}{44} * 0\right)\right)$$

$$\text{Gain(Total, general condition)} = 0.962412735 - (0.5454545454545 \times 0.8708764469) + (0.4545454545 \times 0)$$

$$\text{Gain(Total, kgeneral condition)} = 0.962412735 - (0.47501698309) + (0)$$

$$\text{Gain(Total, general condition)} = 0.487395752$$

Steps in building a decision tree to choose an attribute as the root node is based on the highest Gain, then creating a branch for each value in the attribute and the process will be repeated until all cases in the branches have the same class [19].

After the calculation was carried out using equation (1) and (2), it was showed that the obtained attribute with the highest Gain was a common condition by 0.487395752. So, the general condition became a root node. There were two attribute values of the general condition, namely agile and limp. The value of limp attribute had been classified into case 1, with unhealthy decision, so further calculation was not needed. However, the value of agile attribute was needed to be calculated, as illustrated in the decision tree in Fig 1.

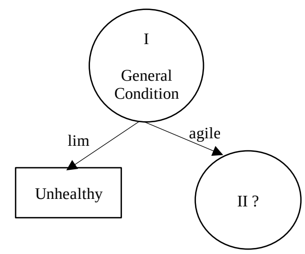

Fig 1. Decision Tree of Node 1 Calculation Result

Calculation of the number of cases for healthy decision, the number of cases for unhealthy decision, and the Entropy of all cases and cases divided by how to walk, nostrils, eyes, ears, fur, skin, digestion, respiratory, heart, feces and body temperature, became the root node of agile attribute value. Then, the Gain calculation was carried out to each attribute, so the highest Gain attribute was ears by 0.372708834. Therefore, ears became the branch node of agile attribute value. There were five ear attribute value including disabled, inflammation, normal, infection, and dirty. From the five attribute values, the disabled attribute value had been classified into case 1, in which the decision was healthy, the inflammation attribute value had been classified into case 1 with unhealthy decision, the infection attribute value had been classified into case 1 with unhealthy decision, the dirty attribute had been classified into case 1 with healthy decision, so there was no need to do further calculation. However, the normal attribute value still needs more calculations, as shown in Fig 2.

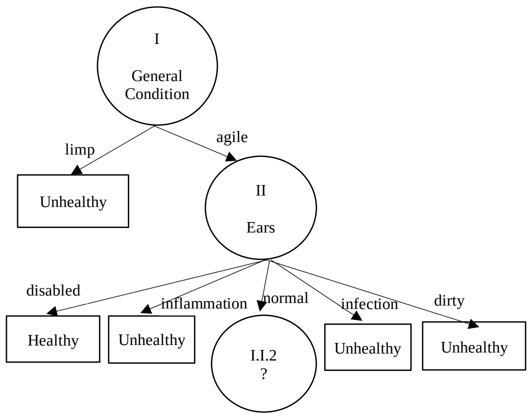

Fig 2. Decision Tree of Node 1.1 Calculation Result

Calculation of the number of cases for normal decision, the number of cases for no decision, and the Entropy of all cases and cases divided based on attributes of way of walking, nostrils, eyes, fur, skin, digestion, respiratory, heart, feces and body temperature became the branch node of normal attribute value. After calculating the Gain for each attribute, it was found that the highest Gain attribute was skin by 0.629249224. Therefore, skin became a branch node of the normal attribute value. There were three skin attribute values including itching, ringworm and normal. From the three attribute values, the itching attribute value had been classified into case 1, with unhealthy decision, the ringworm attribute value had been classified into case 1, with unhealthy decision, the normal attribute value had been classified into case 1, with unhealthy decision, so further calculation was not needed, as shown in Fig 3 below.

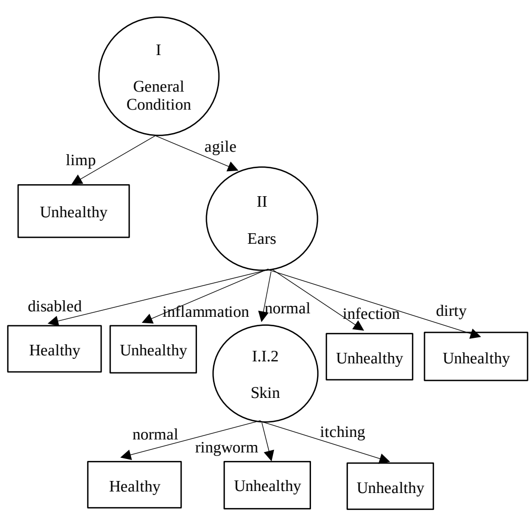

Fig 3. Decision Tree of Node 1.1.2 Calculation Result

Based on the explanation above, it can be concluded that of the rules formed from the calculation result decision tree for the determination of healthy cat are:

a. general condition = limp, then unhealthy,
b. general condition = agile, and ears = disabled, then healthy.
c. If general condition = agile, and ears = inflammation, then unhealthy
d. If general condition = agile, and ears = infection, then unhealthy if general condition = agile, and ears = dirty, then unhealthy
e. If general condition = agile, ears = normal, skin = normal, then healthy
f. If general condition = agile, ears = normal, skin = ringworm, then unhealthy
g. If general condition = agile, ear = normal, skin = itching, then unhealthy

The research design tool using Unified Modelling Language (UML)[20][21], produces (1) Use Case Diagram to describe the interaction of the employee or doctor with the information system to be created, and (2) Class Diagram describes the relationships between classes including: employee login, doctor login, patient/cat data, patient/cat data view, examination input, and examination results.

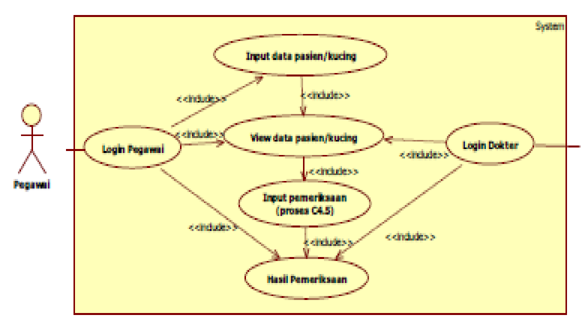

Fig 4. Use Case Diagram

Use Case describes the employee login, performs data entry of patient/cat, sees patient/cat data that has been inputted, inputs the examination of animal health and sees the examination results. While doctor login sees patient /cat, data and examination results as in Fig 4 above.

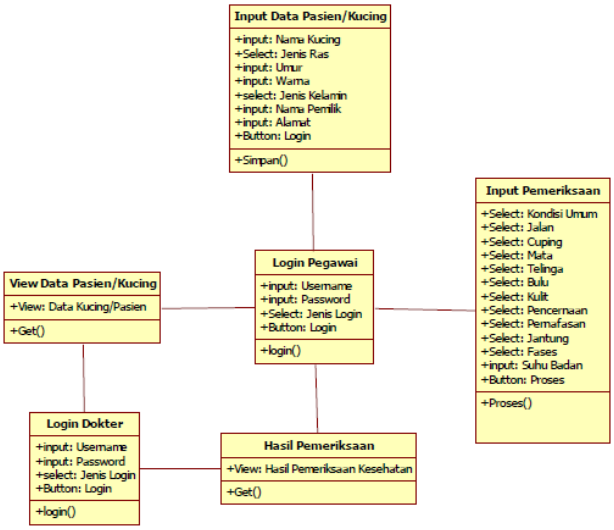

Fig 5. Class Diagram

Class diagram describes classes in a system and the relationship between employee login and doctor login, patient/cat data input, patient data/cat view, examination input, and examination results which are shown in Fig 5[22][23].

III. Result and Discussion

This research results in an application using Java programming language with MySQL database and Black Box and accuracy testing. The output of this application displays the decision in determining the healthy and unhealthy cat.

A. Program Implementation

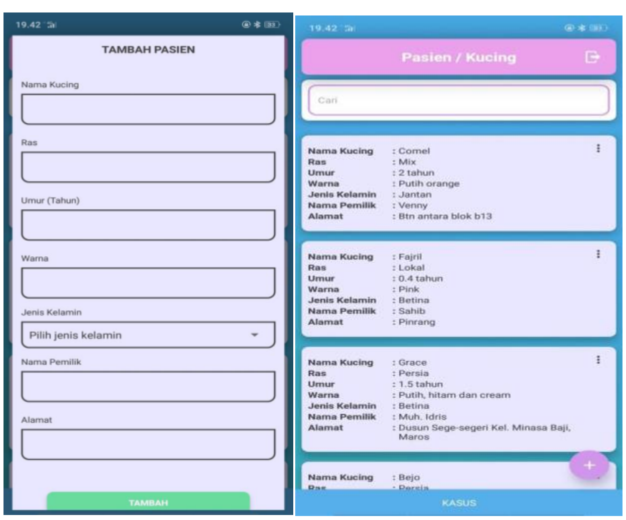

Fig 6. Cat data input form

Fig 6 shows the interface of cat data input (left) and cat data is successfully added (right). To display the examination data of each cat, press one of the case (right) which then displays the form of examination data in Fig 7.

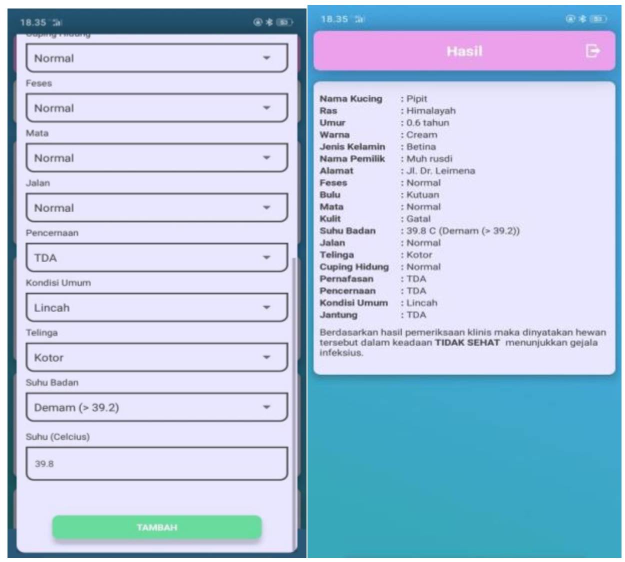

Fig 7. Examination result form

Attribute input data form for each cat is seen in Fig 7 (left) and the determination results of the status is healthy and unhealthy cat (right). In the case example above, the cat with Himalayan race was declared unhealthy based on the attributes of general condition (agile), feces (normal), fur (flea), eyes (normal), skin (itching), body temperature (fever), how to walk (normal), ears (dirty), nostrils (normal), respiratory (NFA), digestion (NFA) and heart (NFA).

B. Black Box Testing

This testing using blackbox testing that generates 5 (five) testing modules, where each module tested the functionality of application, button and suitability of the application results [24][25][26], which are (1) employee login and doctor login to provide access to the system, (2) patient tab serves to add, remove and edit patient/cat data, (3) examination tab displays the form to add the cat case data, (4) result tab has a function to display the healthy or unhealthy result, and (5) log in log out to exit from the system as employee or doctor.

TABLE 2.
Recapitulation of Black Box Testing Results

No.	Module	Description	Successful
1	Testing Employee and Doctor Login	Display the page of patient/cat data	✓
2	Testing Patient Tab	Display the *form* to Add the Patient	✓
3	Testing Examination Tab	Display the *form* of adding case	✓
4	Testing Result Tab	Show healthy or unhealthy results after case input	✓
5	Testing Log Out	Back to *Login* view	✓

Based on black box testing in Table 2 above, it can be concluded that the application has been freed from the functional error, in which the desired form is successfully displayed.

Accuracy testing is used to measure the accuracy value in the system by comparing the results of expert diagnosis and the results of system diagnosis, then the accuracy value will be calculated with the following equation:

$$\text{Accuracy value} = \frac{\text{Number of accuracy data}}{\text{Overall data}} \times 100\%$$

$$\text{Accuracy value} = \frac{41}{44} \times 100\% = 93.18\%$$

Based on the amount of data as many as 44 cases, 41 data of diagnosis results were similar to the expert diagnosis results, so as to produce the percentage value by 93.18%.

IV. Conclusion

The conclusions that can be drawn based on the description of method, design and testing above are:

1. The algorithm C4.5 application on the app showed the output of healthy cat determination that used 44 data samples based on the attribute of general and physical condition.
2. Design tool in this research used UML that generated uses case actor employee and user as well as the class diagram which was the employee login, doctor login, patient/cat data input, patient/cat data view, examination input, and examination results.
3. System testing in this research was Black Box, producing 5 (five) modules which in overall went well and accuracy testing of the expert system for healthy cat determination had a percentage value of 93.18% by comparing the results of expert diagnosis and the results of system diagnosis.

V. Acknowledgment

I would like to thank those who have helped in the completion of this paper.

VI. References

[1] S. J. Lee, Z. Xu, T. Li, and Y. Yang, “A novel bagging C4.5 algorithm based on wrapper feature selection for supporting wise clinical decision making,” *J. Biomed. Inform.*, vol. 78, no. November 2017, pp. 144–155, 2018, doi: 10.1016/j.jbi.2017.11.005.

[2] C. Rajeswari, B. Sathiyabhama, S. Devendiran, and K. Manivannan, “A Gear fault identification using wavelet transform, rough set based GA, ANN and C4.5 algorithm.,” *Procedia Eng.*, vol. 97, pp. 1831–1841, 2014, doi: 10.1016/j.proeng.2014.12.337.

[3] H. Bin Wang and Y. J. Gao, “Research on C4.5 algorithm improvement strategy based on MapReduce,” *Procedia Comput. Sci.*, vol. 183, pp. 160–165, 2021, doi: 10.1016/j.procs.2021.02.045.

[4] R. A. Awad, W. K. B. Khalil, and A. G. Attallah, “Feline panleukopenia viral infection in cats: Application of some molecular methods used for its diagnosis,” *J. Genet. Eng. Biotechnol.*, vol. 16, no. 2, pp. 491–497, 2018, doi: 10.1016/j.jgeb.2018.08.001.

[5] E. Brianti *et al.*, “Treatment and long-term follow-up of a cat with leishmaniosis,” *Parasites and Vectors*, vol. 12, no. 1, pp. 1–7, 2019, doi: 10.1186/s13071-019-3388-9.

[6] M. K. Chatzis *et al.*, “Evaluation of clinicopathological abnormalities in sick cats naturally infected by Leishmania infantum,” *Heliyon*, vol. 6, no. 10, p. e05177, 2020, doi: 10.1016/j.heliyon.2020.e05177.

[7] A. Pereira and C. Maia, “Leishmania infection in cats and feline leishmaniosis: An updated review with a proposal of a diagnosis algorithm and prevention guidelines,” *Curr. Res. Parasitol. Vector-Borne Dis.*, vol. 1, no. April, p. 100035, 2021, doi: 10.1016/j.crpvbd.2021.100035.

[8] X. Wang, C. Zhou, and X. Xu, “Application of C4.5 decision tree for scholarship evaluations,” *Procedia Comput. Sci.*, vol. 151, no. 2018, pp. 179–184, 2019, doi: 10.1016/j.procs.2019.04.027.

[9] A. Joshuva, R. S. Kumar, S. Sivakumar, G. Deenadayalan, and R. Vishnuvardhan, “An insight on VMD for diagnosing wind turbine blade faults using C4.5 as feature selection and discriminating through multilayer perceptron,” *Alexandria Eng. J.*, vol. 59, no. 5, pp. 3863–3879, 2020, doi: 10.1016/j.aej.2020.06.041.

[10] A. P. Muniyandi, R. Rajeswari, and R. Rajaram, "Network anomaly detection by cascading k-Means clustering and C4.5 decision tree algorithm," *Procedia Eng.*, vol. 30, no. 2011, pp. 174–182, 2012, doi: 10.1016/j.proeng.2012.01.849.

[11] K. Bouchard, B. Bouchard, and A. Bouzouane, "A new qualitative spatial recognition model based on Egenhofer topological approach using C4.5 algorithm: Experiment and results," *Procedia Comput. Sci.*, vol. 5, pp. 497–504, 2011, doi: 10.1016/j.procs.2011.07.064.

[12] Y. Xiang, K. Lu, S. L. James, T. B. Borlawsky, K. Huang, and P. R. O. Payne, "K-Neighborhood decentralization: A comprehensive solution to index the UMLS for large scale knowledge discovery," *J. Biomed. Inform.*, vol. 45, no. 2, pp. 323–336, 2012, doi: 10.1016/j.jbi.2011.11.012.

[13] A. H. Khan and I. Porres, "Consistency of UML class, object and statechart diagrams using ontology reasoners," *J. Vis. Lang. Comput.*, vol. 26, pp. 42–65, 2015, doi: 10.1016/j.jvlc.2014.11.006.

[14] I. D. Mienye, Y. Sun, and Z. Wang, "Prediction performance of improved decision tree-based algorithms: A review," *Procedia Manuf.*, vol. 35, pp. 698–703, 2019, doi: 10.1016/j.promfg.2019.06.011.

[15] B. N. Lakshmi, T. S. Indumathi, and N. Ravi, "A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy," *Procedia Technol.*, vol. 24, pp. 1542–1549, 2016, doi: 10.1016/j.protcy.2016.05.128.

[16] K. R. Pradeep and N. C. Naveen, "Lung Cancer Survivability Prediction based on Performance Using Classification Techniques of Support Vector Machines, C4.5 and Naive Bayes Algorithms for Healthcare Analytics," *Procedia Comput. Sci.*, vol. 132, pp. 412–420, 2018, doi: 10.1016/j.procs.2018.05.162.

[17] V. L. Policicchio, A. Pietramala, and P. Rullo, "GAMoN: Discovering M-of-N {¬, ∨} hypotheses for text classification by a lattice-based Genetic Algorithm," *Artif. Intell.*, vol. 191–192, pp. 61–95, 2012, doi: 10.1016/j.artint.2012.07.003.

[18] Y. Han, D. Zhao, and H. Hou, "Oil-immersed Transformer Internal Thermoelectric Potential Fault Diagnosis Based on Decision-tree of KNIME Platform," *Procedia Comput. Sci.*, vol. 83, no. Wtisg, pp. 1321–1326, 2016, doi: 10.1016/j.procs.2016.04.275.

[19] Y. T. Guo *et al.*, "DPI & DFI: A Malicious Behavior Detection Method Combining Deep Packet Inspection and Deep Flow Inspection," *Procedia Eng.*, vol. 174, pp. 1309–1314, 2017, doi: 10.1016/j.proeng.2017.01.276.

[20] A. D. Boyd *et al.*, "Physician nurse care: A new use of UMLS to measure professional contribution: Are we talking about the same patient a new graph matching algorithm?," *Int. J. Med. Inform.*, vol. 113, no. October 2017, pp. 63–71, 2018, doi: 10.1016/j.ijmedinf.2018.02.002.

[21] R. C. Waldemarin and C. R. G. de Farias, "OBO to UML: Support for the development of conceptual models in the biomedical domain," *J. Biomed. Inform.*, vol. 80, no. November 2017, pp. 14–25, 2018, doi: 10.1016/j.jbi.2018.02.015.

[22] J. Geller, Z. He, Y. Perl, C. P. Morrey, and J. Xu, "Rule-based support system for multiple UMLS semantic type assignments," *J. Biomed. Inform.*, vol. 46, no. 1, pp. 97–110, 2013, doi: 10.1016/j.jbi.2012.09.007.

[23] S. Ain El Hayat, F. Toufik, and M. Bahaj, "UML/OCL based design and the transition towards temporal object relational database with bitemporal data," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 32, no. 4, pp. 398–407, 2020, doi: 10.1016/j.jksuci.2019.08.012.

[24] M. N. N. Sitokdana, R. Tanone, and P. F. Tanaem, "Android-based digitalization of number system of traditional, Ngalum, Ketengban, Lepki and Arimtap tribes," *Procedia Comput. Sci.*, vol. 161, pp. 41–48, 2019, doi: 10.1016/j.procs.2019.11.097.

[25] M. N. N. Sitokdana, R. Tanone, and P. F. Tanaem, "Digitalization of the local language dictionary of Pegunungan Bintang," *Procedia Comput. Sci.*, vol. 161, pp. 49–56, 2019, doi: 10.1016/j.procs.2019.11.098.

[26] R. Hermawati and I. S. Sitanggang, "Web-Based Clustering Application Using Shiny Framework and DBSCAN Algorithm for Hotspots Data in Peatland in Sumatra," *Procedia Environ. Sci.*, vol. 33, pp. 317–323, 2016, doi: 10.1016/j.proenv.2016.03.082.