A Multi-label Classification on Topic of Hadith Verses in Indonesian Translation using CART and Bagging

Rendi Kustiawan*, Adiwijaya, Mahendra Dwifebri Purbolaksono

Fakultas Informatika, Program Studi Informatika, Telkom University, Bandung, Indonesia
Email: 1rendik@student.telkomuniversity.ac.id, 2adiwijaya@telkomuniversity.ac.id, 3mahendradp@student.telkomuniversity.ac.id
Email Penulis Korespondensi: rendikstwn@gmail.com

Abstract—Hadis merupakan sumber hukum bagi umat muslim setelah al-qur’an yang di dalamnya terdapat petunjuk baik benar mereka, perbuatan, sikap, dan lain-lain. Hadis wajib dipelajari dan diamalkan oleh umat muslim, lalu dijadikan sebagai pedoman hidup setelah al-qur’an. Melakukan klasifikasi pada hadis merupakan cara untuk memudahkan umat muslim dalam mempelajari hadis dengan cara melihat pola teks pada terjemahan hadis bukhari berdasarkan tiga kelas atau kategori berdasarkan anjuran, larangan, dan informasi. Klasifikasi yang dilakukan adalah klasifikasi multi-label. Pada proses klasifikasi menggunakan N-gram dan TF-IDF sebagai ekstraksi fitur, lalu menggunakan CART dan bagging sebagai metode klasifikasi dan hamming loss sebagai metode evaluasi. Bagging digunakan untuk menutupi kekurangan dari CART yaitu model CART kurang stabil yang mana jika terdapat sedikit perubahan pada data latih akan berpengaruh besar pada model pembelajaran yang dihasilkan. Pada penelitian ini dilakukan beberapa metode pengujian untuk mendapatkan nilai hamming loss terbaik. Berdasarkan beberapa pengujian yang telah dilakukan didapatkan nilai hamming loss terbaik sebesar 0.1914 atau sebesar 80.86%. Hasil tersebut menunjukkan bahwa penggunaan bagging dapat membantu meningkatkan akurasi sebesar 5%.

Kata Kunci: Klasifikasi; Hadis Bukhari; Preprocessing; Ekstraksi Fitur; CART; Bagging

Abstract—Hadith is a source of law for Muslims after the al-qur’an, in which there are instructions in the form of words, actions, attitudes, and others. Hadith must be studied and practiced by Muslims, then used as a way of life after the al-qur’an. Classifying hadith is a way to make it easier for Muslims to learn hadith by looking at the text pattern in the translation of Bukhari hadith based on three classes or categories based on suggestions, prohibitions, and information. The classification carried out is a multi-label classification. The classification process uses N-gram and TF-IDF as feature extraction, CART and bagging as classification methods, and hamming loss as evaluation methods. Bagging is used to cover the shortcomings of CART, namely, the CART model is less stable, which, if there is a slight change in the training data, will have a significant effect on the resulting learning model. Several testing methods were carried out to obtain the best hammer loss value in this study. Based on several tests that have been carried out, the best hammering loss value is 0.1914 or 80.86%. These results indicate that the use of bagging can help increase accuracy by 5%.

Keywords: Classification, Hadith Bukhari, Preprocessing, Feature Extraction, CART, Bagging

1. INTRODUCTION

Hadith is a source of law for Muslims after the al-qur’an, in which there are instructions in the form of words, actions, attitudes, and others. As a Muslim, hadith must be studied and practiced by Muslims, then used as a way of life after the al-qur’an [1][2][3]. There are several hadith experts who have narrated many hadiths, one of which is Imam Bukhari, who has the highest level of validity [3]. Imam Bukhari’s full name is Abu Abdullah Muhammad bin Ismail bin Ibrahim bin Al-Mugrihah Al-Ju’fi who lived between 194 to 256 Hijri. In the authentic hadith of Imam Bukhari, there are recommendations that we need to carry out, there are prohibitions that we need to avoid or leave behind, and there is also information that we need to practice and implement in our daily lives. To make it easier for Muslims to study hadith, grouping or classification is needed by knowing the text pattern from the translation of the Bukhari hadith.

Text classification itself is a process to obtain data or information that has good quality by separating less important data and entering important data into a certain category or class. Meanwhile, multi-label text classification is a classification that has more than one class or category, where one fragment of hadith in the form of words or sentences can be entered into several classes or categories based on the characteristics of the word or sentence.

There are several relevant studies on the classification of multi-label text that has been previously carried out. One of them is a study conducted by Prasetya et al [1] with the title Multi-label Classification of Bukhari Hadith in Indonesian Translation using Mutual Information and Backpropagation Neural Network. This study uses several methods, including mutual information as a feature selection method, TF-IDF as a feature extraction method. K-fold validation to share test data and training data, backpropagation neural network as a classification method, and hamming loss as an evaluation method. The neural network is a classification method because it can classify data with a large and varied number of features. At the experimental stage, the test results concluded that the use of mutual information as a feature selection in terms of hamming loss performance is not good, but in terms of the computational time, it is better with a time difference of 5284.8 s compared to without mutual information. And the best hamming loss is 0.0892 by involving the stemming process, mutual information, and the best learning rate value.

A study conducted by Wiraguna et al [2] with the title Multi-label Classification of Bukhari Hadith in Indonesian Translation using Random Forest. This study uses several methods, including TF-IDF as a feature
selection method, random forest as a classification method, and hamming loss as an evaluation method. The researcher uses the Problem Transformation approach, namely binary relevance and labels powerset, to adapt random forest in building a multi-label text classification system. From the experimental results, the best hamming loss value is 0.0663 using binary relevance and without using stemming.

In a study conducted by Ilham Kurnia et al [3] with the title Multi-label Text Classification in Indonesian Translated Hadith based on Recommendations, Prohibitions and Information using TF-IDF and KNN. This study uses several methods, including N-gram as feature extraction, TF-IDF as feature selection, KNN as a classification method, and hamming loss as an evaluation method. There are several conclusions to get the best k value in this text classification by conducting experiments on the use of feature extraction, adding a threshold value to feature selection, making changes to the preprocessing stage, and using odd k values. From the experimental results, the best hamming loss value is 0.1461.

In a study conducted by Saba Bashir et al [8] with the title An Efficient Rule-based Classification of Diabetes Using ID3, C4.5 & CART Ensembles. In this study, several methods were used to obtain efficient results. The method used in this process consists of basic classification and ensemble classification. The basic classification method consists of ID3, C4.5, and CART, while the ensemble classification consists of majority voting, AdaBoost, Stacking, Bayesian Boosting, and Bagging. From several tests on ensemble bagging classification, it shows good performance compared to other classification techniques.

In a study conducted by Ahmad Rusandi et al [6] with the title Bagging and Boosting Techniques in the CART Algorithm for Classification of Student Study Periods. In this study, several methods were used to obtain efficient results. The method used in this process consists of basic classification and ensemble classification. The basic classification method consists of ID3, C4.5, and CART, while the ensemble classification consists of majority voting, AdaBoost, Stacking, Bayesian Boosting, and Bagging. From several tests on ensemble bagging classification, it shows good performance compared to other classification techniques. By referring to the problems above and several related studies, it is necessary to build a text classification that consists of several stages, namely preprocessing, feature extraction, classification methods, and evaluation methods. The preprocessing used includes punctual removal, case folding, stopword removal, and stemming. Then the feature extraction used is TF-IDF and N-gram. The classification method used is CART and Bagging. CART has several advantages compared to other classification algorithms, including having a decision tree that is easy to interpret, has fairly good accuracy and has a faster calculation. It was proven in research [6] using CART to produce a fairly good accuracy with an average accuracy of 81%. But behind the advantages possessed by the CART algorithm, there are several drawbacks, including that it is very dependent on the number of samples and the model is less stable, which if there is a slight change in the training data, will have a major effect on the resulting learning model. To cover this shortcoming, bagging can be a solution because it can make the classification model more stable and improve classification accuracy. In research [7] the bagging algorithm was able to increase accuracy by 9%.

2. RESEARCH METHODOLOGY

2.1 Research Flow

This study consists of several stages to build a multi-label classification system for the Indonesian translation of the Bukhari hadith. Build this classification system consists of several stages, including preparing the dataset, preprocessing, feature extraction, data classification, and evaluation of the classification model. The description of the system to be built can be seen in Figure 1.

Figure 1 Stages of the system to be built
2.1.1 Dataset

The data used is Bukhari hadith, which amounts to 7000 hadith data, where each hadith has a label. Label one indicates that the text belongs to that category or class, and label zero indicates that the text does not fall into that category or class. The category or class label used consists of recommendations, prohibitions, and information. In the recommended category or class, there are 5665 hadith labeled 0 and 1335 hadith labeled 1, then in the prohibited category or class, there are 6156 labeled 0 and 844 labeled 1, and in the information category or class, there are 366 labeled 0 and 6634 labeled 1.

2.1.2 Preprocessing

Preprocessing is the earliest stage carried out in the classification of text data before being used in the next process. The preprocessing stage itself aims to eliminate words that are not needed in building text classification [2]. At this stage, it will change the data into better data so that it can produce information in the text with good quality and ready to be used in the next process. In this study, the preprocessing techniques used include punctual removal, case folding, stopword removal, stemming, and tokenization [9]. Punctual removal is the process of removing punctuation marks such as commas, periods, and others. Case folding is the process of changing uppercase letters into lowercase letters in text or sentences. Stopword removal is the process of removing text that has no effect or has only a small effect on a sentence. Stemming is the process of changing words with affixes into basic words.

The example of the preprocessing process starting from the inputted data to the data that has produced good quality can be seen in Table 1.

| Process Name         | Sentence Input                  | Sentence Output                  |
|----------------------|---------------------------------|----------------------------------|
| Punctuation Removal  | Ya Allah, ajarkanlah dia Kitab. | Ya Allah ajarkanlah dia Kitab   |
| Case Folding         | Ya Allah ajarkanlah dia Kitab   | ya allah ajarkanlah dia kitab   |
| Stopword Removal     | ya allah ajarkanlah dia kitab   | ya allah ajarkanlah kitab       |
| Stemming             | ya allah ajarkanlah kitab       | ya allah ajar kitab             |

2.1.3 N-gram Feature Extraction

At the feature extraction stage using the N-gram method. Feature extraction is a process to produce features that can describe information that will later be used in the classification process [3]. While the N-gram is a chunk of N-characters taken from the string [10] and the n-gram will separate a word based on the order of words that are in a sentence. In this study, the n-gram used is n which is worth 1 or commonly referred to as a unigram and n is worth two or commonly called a bigram.

| N-gram      | Application Results |
|-------------|---------------------|
| Unigram     | ‘ya’, ‘allah’, ‘ajarkanlah’, ‘dia’, ‘kitab’ |
| Bigram      | ‘ya allah’, ‘allah ajarkanlah’, ‘ajarkanlah dia’, ‘dia kitab’ |

2.1.4 TF-IDF Word Weighting

At the feature extraction stage or giving weights using the TF-IDF method. TF-IDF is a feature extraction method that is quite popular and is often used in text classification research as in research [1][2][3]. The TF-IDF feature extraction method will form a matrix where the rows of this matrix are data while each column of the matrix is a feature.

TF-IDF is a word weighting method that combines two concepts at once, namely TF and IDF. TF is the frequency of occurrence of a term(words)/term(words) the most in a sentence, and IDF is the number of sentences in which a certain term(word)(t) [3]. The following is the TF-IDF equation:

\[ W_{t,d} = tf \times idf = tf_{t,d} \times \log \frac{N}{df} \]  

Description:
- \( W_{t,d} \) = the weight of the word tj against the document N
- \( tf_{t,d} \) = The number of occurrences of the word term in a document
- \( N \) = Number of documents
- \( df \) = Number of occurrences of words in a document N

2.1.5 Classification and Regression Tree (CART)

Classification and Regression Tree (CART) is a classification method for building a decision tree. The CART method was developed by Breiman, Friedman, Olshen, and Stone in their paper entitled “Classification and Regression Tree” in 1984. The CART algorithm is a classification technique using a binary recursive partitioning
algorithm [11][12]. Binary means the selection of a group of data that has been collected, which is called a node, and if two groups of data are collected, it is called a child node. Recursive means a binary restriction step that is done repeatedly. Partitioning means the classification process is carried out by selecting a data set into several parts.

There are three stages of the CART algorithm, including the formation of a classification tree, pruning the classification tree, and the formation of an optimal classification tree [11]. In the first step, namely the formation of a classification tree, consists of three stages that require a learning sample L. The first step determines the selection of sorts, at this stage the data used is data from the learning sample. The resulting subset must be more homogeneous than the previous sorting. The homogeneous function that is commonly used is the Gini function because it will always separate the class with the most important member to be used as a node first. The following is the equation of the Gini function [11]:

$$\Phi (s)|t) = \Delta i (s, t) = i(t) - P_L i(t_L) - P_R i(t_R)$$  \hspace{1cm} (3)

Description:
- $\Phi (|s)t)$ = criteria goodness of split (decreased heterogeneity)
- $P_L i(t_L)$ = the proportion of observations from node $t$ to the left node
- $P_R i(t_R)$ = the proportion of observations from node $t$ to the right node

The next step is to determine the terminal node. Node $t$ can be used as a terminal node if it fulfills the conditions, including there is no significant decrease in heterogeneity in the sorting and there is only one observation (n=1) in each child node. Then the next step is labeling each terminal node according to the rules for the largest number of class members. The following is the equation [11]:

$$P(j |t) = \max_j p(j |t) = \frac{\sum_{i=1}^{N(t)} N_j(t)}{N(t)}$$  \hspace{1cm} (4)

Description:
- $P(j |t)$ = class j proportion at vertex $t$
- $N_j(t)$ = number of class observations at $j$ at node $t$
- $N(t)$ = number of observations at node $t$
- $j_a$ = gives the estimated value of the largest $t$ node classification

In the second step, namely the pruning of the classification tree that was formed in the first stage with several rules including sorting rules and criteria for goodness of split. This classification tree pruning is done to prevent the formation of a very large and complex classification tree that will cause overfitting and also to prevent the formation of a very small classification tree that will cause overfitting. To prevent this, a calculation is carried out based on cost complexity pruning in order to get a proper classification tree. Then the magnitude of the replacement estimate for the tree $T$ on the complexity parameter can be seen in equation [11]:

$$R_a(T) = R(T) + a|T|$$  \hspace{1cm} (5)

Description:
- $R_a(T)$ = resubstitution a tree $T$ on complexity $a$
- $R(T)$ = cost -complexity parameter for adding one node to the end of the tree $T$
- $|T| = \text{the number of terminal nodes of the tree } T$

Then the next step is to perform cost complexity pruning by determining the subtree $T(a)$ which minimizes $R_a(T)$ in the entire subtree for each value of $a$. Later, the complexity value of $a$ will increase during the pruning process. Then it will be minimized to $R_a(T)$ by searching the tree section $T(a) < T_{\max}$. The following is the equation [11]:

$$R_a((T(a))) = \min_{T<T_{\max}} R_a(T)$$  \hspace{1cm} (6)
After pruning, the most optimum classification tree is obtained with a very feasible or simple size but provides a fairly small replacement. Then according to research [11] a substitute estimator that is often used if the observations are not too large is the Cross-Validation Estimate. The estimator is divided into two, namely the test sample estimate and the V-fold cross-validation (Cross Validation V-fold Estimate).

2.1.6 Bagging (Bootstrap Aggregating)

Bagging is known as one of the easiest methods of adaptive reweighting and combining, a general term that refers to the reuse or selection of data to improve accuracy. The bagging method consists of two stages, namely bootstrap and aggregating. At the bootstrap stage, it is done by taking samples from the training data they have, and this is called resampling [13]. At the aggregating stage, it will combine many predictive values into one predictive value [13]. One way to get the predicted value in the classification is by means of a majority vote (majority vote).

This bagging is used to improve classification accuracy and avoid overfitting, which later results from individual classifications being combined to produce better classes. The meaning of the individual here is the CART algorithm. All individual classifiers were trained from the training set to sampling. In the study [8] [14], bagging was able to improve the classification performance better because bagging was able to eliminate model instability. The classification of the bagging model can be seen in Figure 2.

![Bagging model classification framework](image)

**Figure 2** Bagging model classification framework [6]

2.1.7 Evaluation

Hamming loss will calculate the number of prediction errors in the text classification. The smaller the value of the hammering loss, the better it means that the error is quite small, and the accuracy is quite good. The following is the equation of the hammering loss [2]:

$$hloss(h) = \frac{1}{P} \sum_{i=1}^{P} \frac{1}{Q} |h(X_i)\Delta Y_i|$$

Keterangan:

- \(P\) = Amount of data
- \(Q\) = Number of classes
- \(|h(X_i)\Delta Y_i|\) = The number of errors or errors in the classification that occurred

3. RESULTS AND DISCUSSION

Testing the classification of the text of the Bukhari Hadith translated into Indonesian, there are 7000 datasets which are divided into three categories or classes, including recommendations, prohibitions, and information. The test
scenario of this final project focuses on the preprocessing stage and the classification stage. The first test scenario is testing preprocessing techniques, namely punctual removal, case folding, stopword removal, and stemming. The purpose of this test is to find out the best combination of preprocessing techniques and to find out from preprocessing to classification performance. The second test scenario is a feature extraction test with the addition of n-grams. The aim is to determine the effect of the application of n-gram on classification performance. The third scenario is testing the classification method. The purpose of this test is to determine the effect of bagging on the classification of hadith texts using the CART classification method.

### 3.1 Results and discussion of the effect of preprocessing

In the first scenario, testing is conducted on the preprocessing process, which focuses on the effect of using stopword removal and stemming on accuracy results. Stopword removal is the process of removing text that has no effect or only has a small effect on a sentence while stemming is the process of changing words with affixes into basic words.

To get the results of the stopword removal and stemming process on the hadith data used, the author uses a literary library. In this test, the system uses different training data splits and different test data in order to determine the effect of split data. This test also uses TF-IDF and Unigram as feature extraction and uses CART and bagging as classification methods. This test was conducted to determine the effect of preprocessing on the value of hammering loss. The test results in the first scenario can be seen in Table 3.

**Table 3 Effect of preprocessing**

| Preprocessing                        | Preprocessing Accuracy against Split Data (Data train : Data test) |
|--------------------------------------|---------------------------------------------------------------|
|                                      | 60:40 | 70:30 | 80:20 | 90:10 |
| No Stemming                          | 0.2465 | 0.2215 | 0.2194 | 0.2346 |
| No Stopword Removal                  | 0.2367 | 0.2042 | 0.2042 | 0.1971 |
| No Stemming dan Stopword Removal     | 0.2275 | 0.2028 | 0.2 | 0.1914 |
| Full Preprocessing                   | 0.2490 | 0.2310 | 0.2380 | 0.2374 |

Based on the results of the tests that have been carried out, it can be seen in table 3, and it can be concluded that the effect of data splits and different preprocessing techniques can produce different hammering loss values. In the tests that have been carried out, the best hammering loss value is 0.1914, or about 80.86% of the multi-label data, which is classified correctly. The best accuracy is obtained from the preprocessing process without stemming and stopword removal by dividing the training data by 90% and the test data by 10% from the dataset.

In the split test data with 60% train data and 40% test data with preprocessing punctual removal, case folding, stopword removal, and case folding (Full preprocessing), the lowest hammering loss value is 0.2490 or about 75.1. And the 90: 10 data split test resulted in the best hammering loss value compared to other data splits.

Judging from the best hammering loss value, stemming and stopword removal processes cannot be carried out on every word in the hadith. This is because the stemming process can change the basic word form in the hadith. For example, the sentence "pray standing up" will change to "pray standing up" if it is done in the stemming process. Eliminating the word -lah in the prayer sentence fragment can then change the meaning of the sentence fragment, which originally meant a suggestion would turn into information. Meanwhile, stopword removal is the process of removing text that has no effect or only has a small effect on a sentence. The stopword removal used comes from the literary library, with a total of 758 words removed. For example, the sentence "don't be angry" will turn into "angry" if the stopword removal process is carried out using a literary library. Eliminate the word don't, and you can reduce information from a sentence so that it can reduce the accuracy of hammering loss. However, the use of stopword removal can speed up computing time because the fewer words that are processed in a sentence, the faster the system will process the words in a sentence.

### 3.2 Results and Discussion of the influence of the Unigram and Bigram

In the second scenario, testing is carried out by applying additional feature extraction with unigram and bigram using preprocessing techniques without stemming and stopword removal. TF-IDF and N-gram as feature extraction and C4.5 Decision Tree and bagging as classification methods. This test was conducted to determine the effect of unigram and bigram on the value of hammering loss. The test results in the first scenario can be seen in Table 4.

**Table 4 Effect of N-gram**

| Split Data (Data train : Data test) | N-gram accuracy |
|------------------------------------|-----------------|
|                                    | Unigram | Bigram  |
| 60:40                              | 0.2275   | 0.2303  |
Based on the results of the tests that have been carried out, it can be seen in Table 4, and it can be concluded that feature extraction using unigram has a better level of accuracy than bigram. This is because the use of features on unigram only has a single feature so that it is easier to find and is more widely used in sentences in the data train. While the bigram feature is made up of two words, this causes the bigram to be more difficult to find in a sentence that is in the data train, and the bigram also produces more features than the unigram so that it requires a longer computational process compared to the unigram.

Based on the results of preprocessing testing using punctual removal and case folding with feature extraction using TF-IDF and n-gram, the best hamming loss value is 0.1914 or equivalent to 80.86 using unigram and split data train as much as 90% and test data is 10%.

3.3 Results and Discussion Effect of bagging

In the third scenario, testing is carried out by applying different classification techniques. At this stage, the classification uses only CART and CART+bagging. The aim is to find out how influential the bagging algorithm is on the accuracy of text classification and to find out the comparison using CART. In this test, the system uses different training data splits and different test data in order to determine the effect of split data. This test also uses preprocessing punctual removal and case folding, then uses TF-IDF and unigram as feature extraction. This test was conducted to determine the effect of bagging on the classification results. The results of the second scenario test can be seen in Table 5.

| Split Data (Data train: Data test) | N-gram accuracy |
|-----------------------------------|-----------------|
|                                  | Unigram | Bigram |
| 70:30                            | 0.2028   | 0.2209 |
| 80:20                            | 0.2000   | 0.2178 |
| 90:10                            | **0.1914** | 0.2028 |

Based on the results of the tests that have been carried out, it can be seen in Table 7, and it can be concluded that the use of bagging can improve accuracy. This is because the use of bagging in this study uses 300 model samples from the CART algorithm and trains each model with different samples from the same dataset. Then after training all the models, the predicted value from the 300 samples will be obtained, and the predicted value is voted on by majority voting to get the final predictive value.

Based on the results of the tests that have been carried out with 90% train data split and 10% test data that the CART algorithm produces the best hamming loss of 0.2342 or about 76.26, while the CART + bagging algorithm produces the best hamming loss of 0.1914 or 80.86. Thus, it can be seen that the use of bagging can increase the classification accuracy on CART by about 5%.

4. CONCLUSION

Based on the tests that have been carried out using several scenarios for the classification of Hadith al-Bukhari translated into Indonesian using CART and Bagging, it can be concluded that the best system performance is produced by a combination of preprocessing punctual removal and case folding (without stemming and stopword removal) using feature extraction TF-IDF and unigram, then with 90% training data split and 10% test data resulted in a hamming loss accuracy of 0.1914 or 80.86%. The use of stemming in preprocessing should be avoided because it can change the meaning of a sentence that is in the hadith dataset used, and the use of stopword removal in the literary library should also be avoided because it can reduce information from a sentence of the hadith dataset used so that it can reduce the accuracy of hamming loss. Then the best use of feature extraction is obtained from the results of the analysis using TF-IDF and unigram. This can be proven by comparing the Hamming Loss value in the classification with the Bigram feature selection. Split data can increase the accuracy slightly because the higher the test data, the lower the classification results, and the lower the test data, the higher the classification results. This is due to the spread of data on the test data so as to produce different classification values. For further research are the need to review the categories or classes in the dataset used so that there is no imbalance data and besides that, try to use a more diverse ensemble method to find out the extent of the ensemble's influence on the classification results.
REFERENCES

[1] H. Prasetyo, A. Adiwijaya, and W. Astuti, “Klasifikasi Multi-label Pada Hadis Bukhari Dalam Terjemahan Bahasa Indonesia Menggunakan Mutual Information Dan Backpropagation Neural Network,” e-Proceedings Eng., vol. 6, no. 2, 2019.

[2] A. Wiraguna, S. Al Faraby, and Adiwijaya, “Klasifikasi Topik Multi Label pada Hadis Bukhari dalam Terjemahan Bahasa Indonesia Menggunakan Random Forest,” e-Proceedings Eng., vol. 6, no. 1, pp. 2144–2153, 2019.

[3] I. K. Syuriadi, Adiwijaya, and W. Astuti, “Klasifikasi Teks Multi Label pada Hadis dalam Terjemahan Bahasa Indonesia Berdasarkan Anjuran, Larangan dan Informasi menggunakan TF-IDF dan KNN,” vol. 6, no. 2, pp. 9121–9132, 2019.

[4] S. Al Faraby, E. R. R. Jasim, A. Kusumaningrum, and Adiwijaya, “Classification of hadith into positive suggestion, negative suggestion, and information,” J. Phys. Conf. Ser., vol. 971, no. 1, 2018, doi: 10.1088/1742-6596/971/1/012046.

[5] H. Fauzan, A. Adiwijaya, and S. Al-Faraby, “Pengklasifikasian Topik Hadis Terjemahan Bahasa Indonesia Menggunakan Latent Semantic Indexing dan Support Vector Machine,” J. Media Inform. Budidarma, vol. 2, no. 4, p. 131, 2018, doi: 10.30865/mib.v2i4.948.

[6] A. R. Arrahimi, M. K. Ihsan, D. Kartini, M. R. Faisal, and F. Indriani, “Teknik Bagging Dan Boosting Pada Algoritma CART Untuk Klasifikasi Masa Studi Mahasiswa,” J. Sains dan Inforn., vol. 5, no. 1, pp. 21–30, 2019, doi: 10.34128/jsi.v5i1.171.

[7] D. A. N. A. Lgoritma, C. U. Ntuk, and M. E. K. Etidakseimbangan, “PENERAPAN T EKNIK B AGGING P ADA A LGORITMA N AIVE B AYES,” vol. 1, pp. 41–48, 2020.

[8] S. Bashir, U. Qamar, F. H. Khan, and M. Y. Javed, “An Efficient Rule-Based Classification of Diabetes Using ID3, C4.5, & CART Ensembles,” Proc. - 12th Int. Conf. Front. Inf. Technol. FIT 2014, pp. 226–231, 2015, doi: 10.1109/FIT.2014.50.

[9] L. G. Irlham, A. Adiwijaya, and U. N. Wisesty, “Klasifikasi Berita Bahasa Indonesia Menggunakan Mutual Information dan Support Vector Machine,” J. Media Inform. Budidarma, vol. 3, no. 4, p. 284, 2019, doi: 10.30865/mib.v3i4.1410.

[10] X. Zheng, P. Li, Z. Chu, and X. Hu, “A Survey on Multi-Label Data Stream Classification,” IEEE Access, vol. 8, pp. 1249–1275, 2020, doi: 10.1109/ACCESS.2019.2962059.

[11] N. Indah Prabawati, Widodo, and H. Ajie, “Kinerja Algoritma Classification And Regression Tree (Cart) dalam Mengklasifikasikan Lama Masa Studi Mahasiswa yang Mengikuti Organisasi di Universitas Negeri Jakarta,” PINTER J. Pendidik. Tek. Inform. dan Komput., vol. 3, no. 2, pp. 139–145, 2019, doi: 10.21009/pinter.3.2.9.

[12] R. Irmanita, Sri Suryani Prasetyowati, and Yuliant Sibaroni, “Classification of Malaria Complication Using CART (Classification and Regression Tree) and Naïve Bayes,” J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 1, pp. 10–16, 2021, doi: 10.29207/resti.v5i1.2770.

[13] E. Prasetyo, B. Prasetyo, U. N. Semarang, and P. Korespondensi, “INCREASED CLASSIFICATION ACCURACY C4.5 ALGORITHM USING BAGGING TECHNIQUES IN DIAGNOSING HEART DISEASE,” vol. 7, no. 5, pp. 1035–1040, 2020, doi: 10.25126/jtiik.202072379.

[14] R. Richman and M. V. Wüthrich, “Nagging predictors,” Risks, vol. 8, no. 3, pp. 1–26, 2020, doi: 10.3390/risks8030083.

[15] R. Latifah, E. S. Wulandari, and P. E. Kreshna, “Model Decision Tree Untuk Prediksi Jadwal Kerja Menggunakan Scikit-Learn,” J. Univ. Muhammadiyah Jakarta, pp. 1–6, 2019. [Online]. Available: https://journal.umj.ac.id/index.php/semntast/article/download/5239/3517.

[16] A. Rahayu, L. T. Sianturi, and T. Zebua, “Implementasi Algoritma Cart untuk Mengklasifikasikan Buku Yang Paling Sering Dibaca (Studi Kasus: Yayasan Cinta Baca),” Inf. dan Teknol. Ilm., vol. 13, pp. 223–227, 2018.

[17] R. Pujianto, Adiwijaya, and A. A. Rahmawati, “Analysis Ekstraksi Fitur Principle Component Analysis pada Klasifikasi Microarray Data Menggunakan Classification And Regression Trees,” Proceedings …, vol. 6, no. 1, pp. 2368–2379, 2019.

[18] Adiwijaya, Said Al Faraby, Mohamad Syahrul Mubarok, and Mahendra Dwilebri Purbolaksono. 2021. Indonesian Translation of the Hadith of Bukhari (Single-label), Dataverse Telkom University, DOI: https://doi.org/10.34820/FK2/GWSEWB.