International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse262942020.pdf
https://doi.org/10.30534/ijatcse/2020/262942020

Thematic Textual Hadith Classification: An Experiment in Rapidminer using Support Vector Machine (SVM) and Naïve Bayes Algorithm

Norzihani Yusof [1], Siti Aishah Rosidi [2], Nuzulha Khilwani Ibrahim [3], Ahmed ElMogtaba Banga Ali [4]
[1,2,3]BIOCORE Research Group, Centre for Advanced Computing Technologies (C-ACT), Fakulti Teknologi Maklumat dan Komunikasi, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.
[4] Department of Quran Sunnah (QS), Kuliyyah of Islamic Revelation Knowledge and Human Science (KIRKHS), International Islamic University Malaysia, PO Box 10, 50728 Kuala Lumpur, Malaysia.
[1]norzihani@utem.edu.my
[2]aishahrosidi96@gmail.com
[3]nuzulha@utem.edu.my
[4]elmogtaba@iium.edu.my

ABSTRACT

There are many existing problems in *Hadith* studies trending in the study field. The issues are changeable from the digitalization of the *Hadith* data to an exact case study of estimation of narrators' chain for a particular *Hadith*. However, in this paper, we are not concentrating on the such learning of estimating, confirming or authenticating a *Hadith*. It focuses more on the data mining use to the *Hadith* dataset. We put on the *Hadith* dataset onto one of machine learning tools which is text classification. The *Hadith* dataset is put into experiment for *Hadith* textual classification. It concentrates on the thematic classification based on the themes and words occurrences from the *Hadith* text (*matn*). The *Hadith* textual classification does not trace on the *hukm* and position or class of *Hadith*. This research does not categorize the *Hadith* into *hukm* Sahih, Hasan, Dhaif, or Mawdhoo'. However, the Hadith thematic dataset of this study use only *Hadith* from Sahih Bukhari, where all *Hadith* in the Book is categorized as sahih by Imam Al-Bukhari. The classification for this thematic Hadith dataset is implemented using Rapidminer, a machine learning tool using Naïve Bayes and Support Vector Machine (SVM) methods. From the results, the different value of accuracy for both SVM and Naïve Bayes Algorithm was 2.4%. The Naïve Bayes Algorithm displayed better result comparing to SVM. We believe that the result could be better by improving the data, algorithms, algorithm tuning or ensemble methods for the future experiments.

Key words : Machine learning, Naïve Bayes, Rapidminer Support Vector Machine.

1. INTRODUCTION

In Arabic, the noun of *Hadith* (حـديث) means report, account or narrative. *Hadith* in Arabic plural is *Ahadith* (أحاديـث). Speech of a person also refers to *Hadith*. In Islamic terminology, according to Encyclopedia of Islam by Juan Campo et al, *Hadith* refers to prophet Muhamad report of statement or actions, or his tacit approval or criticism of something said or done in his presence. Ibn Hajar Al-Asqalani, a classical *Hadith* specialist says in [1] that *Hadith* in religious tradition is something attributed to prophet Muhammad that is not found in the Holy Quran.

After the Holy Quran, *Hadith* is the second source that become guidance for Muslims. Hadith are important textual textual source of law, tradition and teaching in Islamic world. *Hadith* is derived from the Arabic word "Hadatha" meaning news or story. According to Sunni, a *Hadith* is any discussion, action, approval and physical or moral description to the prophet Muhammad, whether supposedly or truly [2].

To take a close look on *Hadith*, components of *Hadith* divided by two parts which are *Isnad* and *Matn*. *Isnad* is the chain or sequence of narrators who narrate the *matn*. *Matn* is the narration or the words of the prophet. The authenticity of the Hadith is depending on the reliability of the components and the linkage among them. The components of *Hadith* are presented in Figure 1.

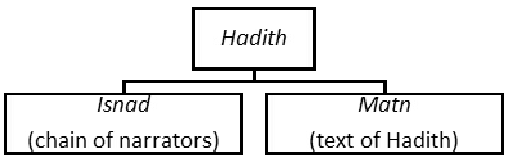


Figure 1: Structure of Hadith

There are many issues in *Hadith* studies as it has been summarized into 4 levels of *Hadith* studies in Ibrahim et al. [3, 4] as depicted in Figure 2.

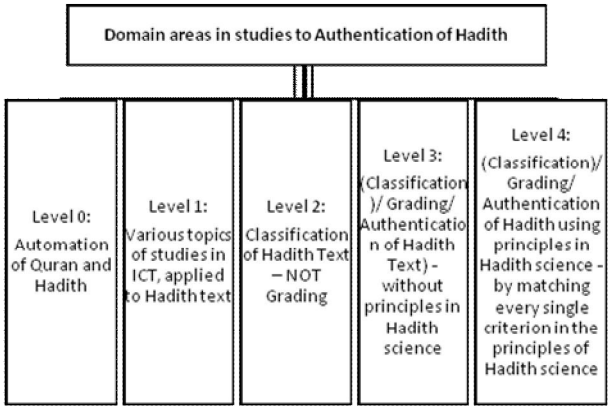


Figure 2: Classification of *Hadith* studies from Ibrahim et al. [3, 4]

The subjects are changeable from the digitalization of the Hadith data [5-15] to an exact case study of calculation of narrators' chain for an exact Hadith, as it is today trending in present study educations [16-21]. In details, for example, the Prophet Muhammad passed away 14 centuries ago, how can we make sure that the *Hadith* is not interrupted to someone who unfamiliar to the science of *Hadith*. The compilation of *Hadith* maybe inaccurate or misleading. So that there is systematic approach in Islamic science to identify valid *Hadith* from an invalid one, included in '*Hadith* science' [22]. This is one of the issues catered in *Hadith* studies at a larger scope.

However, in this paper, we are not concentrating on the such learning of valuing, endorsing or authenticating a *Hadith*, or precisely estimating the *hukm*, rating, position or dependability point of a *Hadith*. This paper concentrates further on the data mining use to the *Hadith* thematic dataset which we apply the *Hadith* thematic dataset on one of machine learning implements that is text classification.

The *Hadith* thematic dataset is put into test for *Hadith* textual sorting, but it concentrates on the thematic classification based on the themes and words occurrences from the *Hadith* text. The *Hadith* thematic textual classification does not trace on the *hukm* and position or group of *Hadith*. This research does not group the *Hadith* into hukm Sahih, Hasan, Dhaif, or Mawdhoo'. However, the *Hadith* thematic dataset from this study use only Hadith from Sahih Bukhari, where all *Hadith* in the Book is categorized as sahih by Imam Al-Bukhari.

Hadith classification is an innovative research studies in computing fields that use different Data mining methods with a list of various options for the approach and algorithms such as decision tree, support vector machine (SVM), K-nearest neighbor (KNN), and Naive Bayes probabilistic classifier [23-30].

For this limited scope of study, our input dataset will be a list of *Hadith* text (*matn*) to be applied for thematic *Hadith* classification. Some of the existing studies in the *Hadith* thematic studies can be found in the literatures [31-34]. The dataset that we use is in Bahasa Indonesia. Regarding the problem in our project, we are classifying the *matn* into 5 themes, which is kitab 'ilmu', 'jual beli', 'makanan', 'minuman', 'sakit'.

2. MATERIALS AND METHODS

In preprocessing phase, the initial process was started by retrieving the dataset using retrieve operator. While selecting attribute, two important attributes were chosen for classification. The two important attributes were Kitab (Theme) and Matan (Text of Hadith). After that, nominal to text operator was selected. This operator would change all nominal dataset into string or text. There were three subprocesses inside the process documents operator. The first subprocess was tokenize. It changed the dataset into token. If there was no letter inside the dataset it considered as one token. The second sub process was the transform case. This operator changed all uppercases into lowercases. The third sub process was called filter stop words (Dictionary). It would remove all the stop words inside dataset that could affect the result of classification.

In modelling process there was cross validation operator. Cross validation was a nested operator. It was divided into two sub processes. They were training and testing parts. In the training sub process Support Vector Machine (SVM) and Naïve Bayes were chosen. In the testing phase, the trained model was applied, and the accuracy of the performance would be obtained. The design was divided into two phases, preprocessing and modeling as in Figure 3. Preprocessing is an important phase in text classification. The model could predict the text of Hadith into their theme and show the dataset prediction based on value of accuracy.

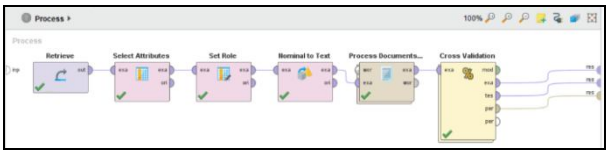


Figure 3: The full design for text of Hadith classification

Preprocessing is the first phase in the text of *Hadith* classification. First, is to retrieve the dataset. The dataset consists of two attributes. They are *kitab* and *matan.* Next step, attributes are selected. Figure 4 shows the data that have been retrieved into Rapidminer tools.

Figure 4: Retrieved Data into Rapidminer

Then, Nominal to Text was chosen. It converted all nominal attributes to string attributes. After that the Process documents from data was picked. This was a nested operator. There were three subprocesses operator inside the process documents from data as in Figure 5.

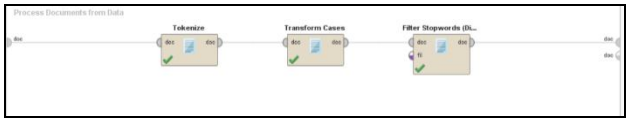


Figure 5: The subprocess operator inside the process documents from data

In tokenization step, the document text then has been split into a sequence of tokens. The mode of this operator was non-letter as in Figure 6. Whenever the operator encountered non-letter in the data it would be considered as token.

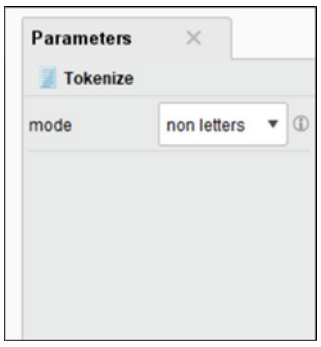


Figure 6: The parameters for 'Tokenize' operator

In Figure 7 shows the dataset after tokenization phase.

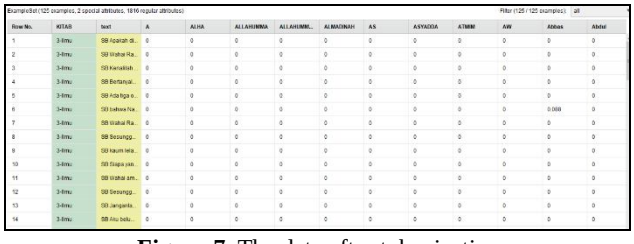


Figure 7: The data after tokenization

Transform cases are used to transform all letters to lowercases. In Figure 8 showed the dataset after being processed by transform cases operator.

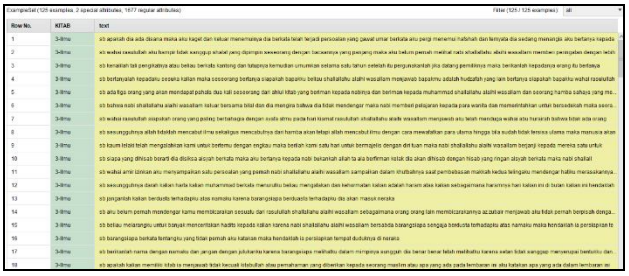


Figure 8: Dataset after being processed by 'transform cases' operator

Filter Stopwords (Dictionary) operator was used to remove all the stopwords. The input file for operator was obtained from GitHub.com. The parameter of filter stopwords (Dictionary) was shown in Figure 9.

Figure 9: The parameters for 'Filter stopwords (Dictionary)'

In figure 10 showed the dataset after filtering stopwords process was applied.

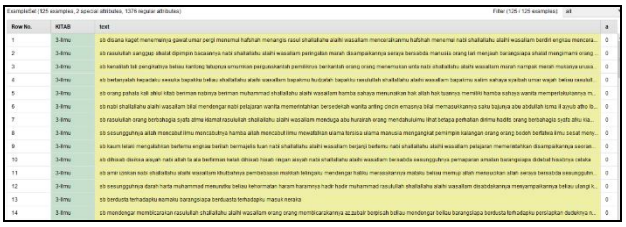


Figure 10: The dataset after 'filtering stopwords' process been applied

All the operators in preprocessing phase are demonstrated in Figure 11.

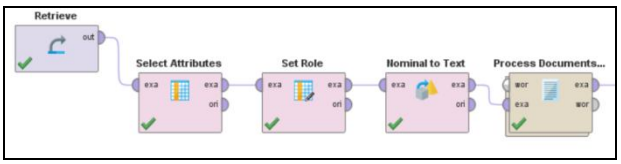


Figure 11: All the preprocessing phase operators

In modelling phase, cross validation operator is used to estimate the accuracy of the dataset. Cross validation is a nested operator. It is divided by two subprocesses, training and testing. In training, SVM and Naïve Bayes are used as the model. The trained model is then applied to the testing

subprocess. The performance of the model is measured in the testing phase.

In this study, multiclass classification is used. Because of SVM model is a binominal classification, therefore classification by regression operator is added as Figure 12.

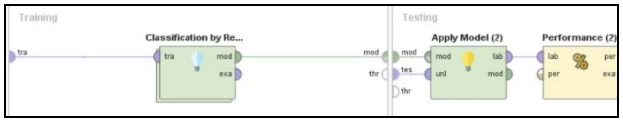


Figure 12: Subprocess in cross validation operator for SVM Classifier

The SVM operator was used as subprocess inside Classification by regression. So that SVM operator could do multiclass classification. The operator was shown in figure 13.

Figure 13: Subprocess in classification by regression

In testing subprocess, there were two operators been applied onto the data text. They are 'apply model' operator and 'performance (classification)' operator. In this scope of study, the criteria value of classification task is accuracy where the performance is measured by the precision and recall value of the classification. For Naïve Bayes sequence operator could be shown as in Figure 14.

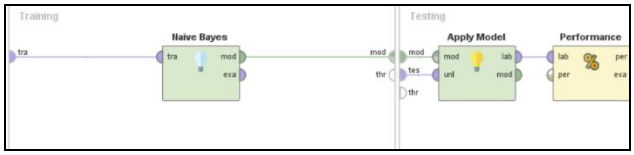
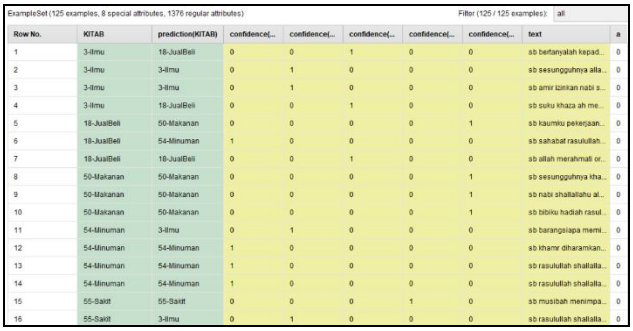


Figure 14: Subprocess in cross validation operator for Naïve Bayes Classifier

3. RESULTS AND DISCUSSION

The result of the classification text of Hadith based on accuracy value could be analyzed. There are two output of the classification. They are table of prediction and table of accuracy. Figure 15 shows the table of prediction example.



ExampleSet (125 examples, 8 special attributes, 1376 regular attributes) Filter (125 / 125 examples): all

Row No.	KITAB	prediction(KITAB)	confidence(...	confidence(...	confidence(...	confidence(...	confidence(...	text	a
1	3-Ilmu	18-JualBeli	0	0	1	0	0	sb bertanyalah kepad...	0
2	3-Ilmu	3-Ilmu	0	1	0	0	0	sb sesungguhnya alla...	0
3	3-Ilmu	3-Ilmu	0	1	0	0	0	sb amir izinkan nabi s...	0
4	3-Ilmu	18-JualBeli	0	0	1	0	0	sb suku khaza ah me...	0
5	18-JualBeli	50-Makanan	0	0	0	0	1	sb kaumku pekerjaan...	0
6	18-JualBeli	54-Minuman	1	0	0	0	0	sb sahabat rasulullah...	0
7	18-JualBeli	18-JualBeli	0	0	1	0	0	sb allah merahmati or...	0
8	50-Makanan	50-Makanan	0	0	0	0	1	sb sesungguhnya kha...	0
9	50-Makanan	50-Makanan	0	0	0	0	1	sb nabi shallallahu al...	0
10	50-Makanan	50-Makanan	0	0	0	0	1	sb bibiku hadiah rasul...	0
11	54-Minuman	3-Ilmu	0	1	0	0	0	sb barangsiapa memi...	0
12	54-Minuman	54-Minuman	1	0	0	0	0	sb khamr diharamkan...	0
13	54-Minuman	54-Minuman	1	0	0	0	0	sb rasulullah shallalla...	0
14	54-Minuman	54-Minuman	1	0	0	0	0	sb rasulullah shallalla...	0
15	55-Sakit	55-Sakit	0	0	0	1	0	sb musibah menimpa...	0
16	55-Sakit	3-Ilmu	0	1	0	0	0	sb rasulullah shallalla...	0

Figure 15: The table of prediction example

Percentage value of accuracy is the number of correctly predicted class divided by total testing class multiplied by one hundred as given in Eq. (1) as it has been discussed in Juan [37].

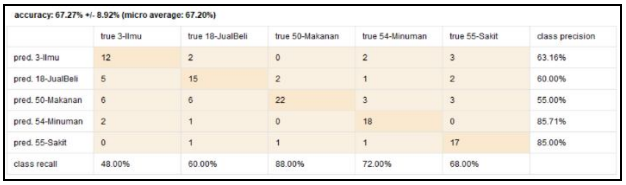
$$\%\ of Accuracy = \left(\frac{correctly\ predicted\ class}{total\ testing\ class}\right) \times 100 \quad (1)$$

The value of accuracy in SVM algorithm is 67.20%. This algorithm predicted 84 from 125 text of Hadith correctly. There were 12 text of Hadith out of 25 from *kitab Ilmu*, 15 text of Hadith out of 25 from *kitab JualBeli*, 22 text of Hadith out of 25 from *kitab Makanan*, 18 text of Hadith out of 25 from *kitab Minuman and* 17 text of Hadith out of 25 from *kitab Sakit*. The calculation of value of accuracy in this model were stated as follows.

$$\%\ of Accuracy\ of\ SVM = \left(\frac{84}{125}\right) \times 100$$

$$= \pm 67.20\ \%$$

Figure 16 shows the table of accuracy from SVM Algorithm.



accuracy: 67.27% +/- 8.92% (micro average: 67.20%)

	true 3-Ilmu	true 18-JualBeli	true 50-Makanan	true 54-Minuman	true 55-Sakit	class precision
pred. 3-Ilmu	12	2	0	2	3	63.16%
pred. 18-JualBeli	5	15	2	1	2	60.00%
pred. 50-Makanan	6	6	22	3	3	55.00%
pred. 54-Minuman	2	1	0	18	0	85.71%
pred. 55-Sakit	0	1	1	1	17	85.00%
class recall	48.00%	60.00%	88.00%	72.00%	68.00%	

Figure 16: The table of accuracy from SVM Algorithm

The value of accuracy in Naïve Bayes algorithm was 69.60% as stated in Figure 17. This algorithm predicted 87 from 125 texts of Hadith correctly. There were 11 texts of Hadith out of 25 from *kitab Ilmu*, 16 texts of Hadith out of 25 from *kitab JualBeli*, 21 texts of Hadith out of 25 from *kitab Makanan*, 21 texts of Hadith out of 25 from *kitab Minuman and* 18 texts of Hadith out of 25 from *kitab Sakit*. The calculation of value of accuracy in this model were stated as follows.

$$\%\ of Accuracy\ of\ Naïve\ Bayes = \left(\frac{87}{125}\right) \times 100$$

$$= \pm 69.60\ \%$$

accuracy: 69.68% +/- 15.65% (micro average: 69.60%)

	true 3-Ilmu	true 18-JualBeli	true 50-Makanan	true 54-Minuman	true 55-Sakit	class precision
pred. 3-Ilmu	11	2	0	2	2	64.71%
pred. 18-JualBeli	4	16	1	1	1	69.57%
pred. 50-Makanan	1	4	21	0	2	75.00%
pred. 54-Minuman	4	3	1	21	2	67.74%
pred. 55-Sakit	5	0	2	1	18	69.23%
class recall	44.00%	64.00%	84.00%	84.00%	72.00%	

Figure 17: The table of accuracy from Naïve Bayes Algorithm

Or, we can summarize the result for both SVM and Naïve Bayes algorithms as in Table 1.

Table 1: Recall and precision values from SVM and Naïve Bayes algorithms

Theme/ Kitab	SVM		Naïve Bayes	
	Recall	Precision	Recall	Precision
Ilmu	48.0 %	63.16 %	44.0 %	64.71 %
Jualbeli	60.0 %	60.0 %	64.0 %	69.57 %
Makanan	88.0 %	55.0 %	84.0 %	75.0 %
Minuman	72.0 %	85.71 %	84.0 %	67.74 %
Sakit	68.0 %	85.0 %	72.0 %	69.23 %

Next section will conclude the result from this scope of study.

4. CONCLUSION

From the results, the different value of accuracy for both SVM and Naïve Bayes Algorithm was 2.4%. The Naïve Bayes Algorithm gives better result compared to SVM. We believe that the result could be better by improving the data, algorithms, algorithm tuning or ensemble methods for the future experiments.

ACKNOWLEDGEMENT

The authors would like to gratefully acknowledge the assistance, support and funding made available by the Universiti Teknikal Malaysia Melaka (UTeM), Faculty of Information and Communication Technology (FTMK), Centre for Research and Innovation Management (CRIM), and the Centre for Advanced Computing Technologies (C-ACT).

REFERENCES

1. Ibn Ḥajar al-ʿAsqalānī and Murad, A. **Selections from the Fatḥ al-Bārī.** (commentary on Ṣaḥ īḥ al-Bukhārī) followed by twenty fatwas on life after death. Cambridge: Muslim Academic Trust, 2000.
2. Batyrzhan M, Kulzhanova BR, Abzhalov SU, Mukhitdinov RS. **Significance of the hadith of the Prophet Muhammad in Kazakh proverbs and sayings.** Proced Social Behav Sci 116:4899–4904. doi:10.1016/j. sbspro.2014.01.1046, 2014.
3. Ibrahim, N. K., Noordin, M. F., Samsuri, S., Seman, M. S. A., & Ali, A. E. B. **Isnad Al-Hadith Computational Authentication: An Analysis Hierarchically.** In Information and Communication Technology for The Muslim World (ICT4M), 2016 6th International Conference on (pp. 344-348). IEEE. (2016a, November). https://doi.org/10.1109/ICT4M.2016.075
4. Ibrahim, N., Noordin, M., Samsuri, S., Abu Seman, M., Ali, A., & Hasan Basari, A. **A Review and Analysis for A Hierarchy from Computational Hadith to Isnad Authenticity Examination.** International Journal On Islamic Applications In Computer Science And Technology, 5(3). Retrieved from http://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/1686, 2017a.
5. Eljazzar, M. M., Abdulhamid, M., Mouneer, M., & Salama, A. **Hadith Web Browser Verification Extension.** arXiv preprint arXiv:1701.07382, 2017.
6. Ibrahim, N. K., Ali, A. E. B., Samsuri, S., Seman, A., Sadry, M., Kartiwi, M., ... & Samad, A. **Takhreej e-Guide: an innovation of electronic guide for Takhreej Al-Isnad Al-Hadith-based on physical structure of Isnad and principles of Hadith science**. Advanced Materials Research, (na), 129-131, 2017b.
7. Aldhlan, K. A., Zeki, A. M., & Zeki, A. M. **Datamining and Islamic knowledge extraction: alhadith as a knowledge resource.** In Information and Communication Technology for the Muslim World (ICT4M), 2010 International Conference on (pp. H-21). IEEE, 2010, December.
8. Rahman, N. A., Alias, N., Ismail, N. K., bin Mohamed Nor, Z., & b Alias, M. N. **An identification of authentic narrator's name features in Malay hadith texts.** In Open Systems (ICOS), 2015 IEEE Conference on (pp. 79-84). IEEE, August 2015.
9. Alias, N., Rahman, N. A., Ismail, N. K., Nor, Z. M., & Alias, M. N. **Searching Algorithm of Authentic Chain of Narrators' in Shahih Bukhari Book**, 2016.
10. Alias, N., Rahman, N. A., Ismail, N. K., Nor, Z. M., & Alias, M. N. **Graph-based text representation for Malay translated hadith text. In Information Retrieval and Knowledge Management (CAMP)**, 2016 Third International Conference on IEEE, August, 2016, pp. 60-66.
11. Alias, N., Rahman, N. A., Ismail, N. K., Nor, Z. M., & Alias, M. N. **Pengkelasan Teks Hadis dalam Kitab Shahih Bukhari berdasarkan kepada Perawi dengan menggunakan Teori Graf.** Persidangan Autentikasi Al-Quran dan Al-Hadith (SAHIH), 2016, pp. 61-65.
12. Rahman, N. A., Ismail, N. K., Nor, Z. M., Alias, M. N., Kamis, M. S., & Alias, N. **Tagging narrator's names in Hadith text.** Journal of Fundamental and Applied Sciences, 9(5S), 2017, pp. 295-309.
13. Najib, S. R. M., Rahman, N. A., Ismail, N. K., Alias, N., Nor, Z. M., & Alias, M. N. **Comparative Study of Machine Learning Approach on Malay Translated Hadith Text Classification based on Sanad.** In MATEC Web of Conferences (Vol. 135, p. 00066). EDP Sciences, 2017.
14. Luthfi, E. T., Suryana, N., & Basari, A. H. **Digital Hadith Authentication: A Literature Review and**

Analysis. Journal of Theoretical & Applied Information Technology, 2018, 96 (15).

15. Mahmood, A., Khan, H. U., Alarfaj, F. K., Ramzan, M., & Ilyas, M. **A Multilingual Datasets Repository of the Hadith Content.** International Journal Of Advanced Computer Science And Applications, 9(2), 2018, pp. 165-172.
16. Aldhlan, K. A., Zeki, A. M., & Zeki, A. M. **Enhanced mechanism to handle missing data of Hadith classifier,** 2011.
17. Aldhaln, K., Zeki, A., Zeki, A., & Alreshidi, H. **Improving knowledge extraction of Hadith classifier using decision tree algorithm.** In Information Retrieval & Knowledge Management (CAMP), 2012 International Conference on (pp. 148-152). IEEE, March, 2012a.
18. Aldhlan, K. A., Zeki, A. M., Zeki, A. M., & Alreshidi, H. A. (2012b, November). **Novel mechanism to improve hadith classifier performance.** In Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on (pp. 512-517). IEEE.
19. Aldhaln, K. A., Zeki, A. M., & Zeki, A. M. (2012c). **Knowledge extraction in hadith using data mining technique.** Int J Inf Technol Comput Sci, 2, 13-21.
20. Ibrahim, N. K., Samsuri, S., Seman, M. S. A., Ali, A. E. B., & Kartiwi, M. (2016b, November). **Frameworks for a Computational Isnad Authentication and Mechanism Development.** In Information and Communication Technology for The Muslim World (ICT4M), 2016 6th International Conference on (pp. 154-159). IEEE.
21. Ibrahim, N., Samsuri, S., Abu Seman, M., Ali, A., Kartiwi, M., & Basari, A. **Design and Frameworks with Experiment for A Basic Guide of Theoretical Isnad Al-hadith Authenticity Examination.** International Journal On Islamic Applications In Computer Science And Technology, 5(3). Retrieved from http://www.sign-ific-ance.co.uk/index.php/IJASAT/article/view/1685, 2017c. https://doi.org/10.4314/jfas.v9i5s.21
22. Melchert, C. **Early renunciants as Hadīth transmitters. The Muslim World**, 92(3-4), 2002, pp. 407-418.
23. Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In European conference on machine learning (pp. 137-142). Springer, Berlin, Heidelberg. https://doi.org/10.1007/BFb0026683
24. Yang Y, Liu X (1999) **A re-examination of text categorization methods.** In: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval, ACM, New York, NY, USA, pp 42–49. doi:10.1145/312624.312647
25. Mining, W. I. D. **Data Mining: Concepts and Techniques.** Morgan Kaufinann, 2006.
26. Hand, D. J. **Principles of data mining. Drug safety**, 30(7), 2007, pp. 621-622.
27. Kotu, V. and Deshpande, B. **Predictive analytics and data mining.** Waltham, MA: Morgan Kaufmann, 2015.
28. Search Enterprise AI. **What is machine learning (ML)?** - Definition from WhatIs.com. [online] Available at: https://searchenterpriseai.techtarget.com/definition/machine-learning-ML [Accessed 1 Aug. 2018].
29. Brownlee, J. **Supervised and Unsupervised Machine Learning Algorithms.** [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/supervised-and-unsupervised-machine-learningalgorithms/ [Accessed 8 Aug. 2018].
30. Statsoft.com. **Naive Bayes Classifier**. [online] Available at: http://www.statsoft.com/textbook/naive-bayes-classifier [Accessed 8 Aug. 2018].
31. Al-Kabi, M. N., and AL-Shalabi "**ALHadith Text Classifier**", Journal of applied sciences, 5(3), 2005, pp. 548-587. https://doi.org/10.3923/jas.2005.584.587
32. Alkhatib, M. (2010, April). **Classification of Al-Hadith Al-Shareef using data mining algorithm.** In European, mediterranean and middle eastern conference on information systems, EMCIS2010, Abu Dhabi, UAE (pp. 1-23).
33. Jbara, K. **Knowledge discovery in Al-Hadith using text classification algorithm.** Journal of American Science, 6(11), 2010, pp. 409-419.
34. Al-Kabi, M. N., Wahsheh, H. A., Alsmadi, I. M., & Moh'd Ali Al-Akhras, A. **Extended Topical Classification of Hadith Arabic Text.** Int. J. on Islamic Applications in Computer Science and Technology, 3(3), 2015, pp. 13-23.