E-ISSN : 2540 - 8984

JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)
Volume 07, Nomor 03, September 2022 : 918–925

IMPLEMENTATION OF QUESTION ENTAILMENT IN QUESTION ANSWERING SYSTEM FOR CHILDREN'S HEALTH TOPIC

Arya Prima Al Aufar[1], Ade Romadhony[], Hasmawati[3]
[1, 2, 3]School of Computing, Informatics, Telkom University
Jl. Telekomunikasi No. 1, Bandung, Indonesia
e-mail: aryaprims@student.telkomuniversity.ac.id[1], aderomadhony@telkomuniversity.ac.id[2], hasmawati@telkomuniversity.ac.id[3]

ABSTRAK

Kesehatan merupakan salah satu hal yang paling penting dalam kehidupan manusia khususnya pada anak-anak, yang memungkinkan anak-anak dapat tumbuh dan berkembang secara baik. Untuk menjaga kesehatan maka orang tua juga harus memiliki informasi yang tepat mengenai cara menjaga kesehatan, pola makan, pola hidup yang baik. Dalam mendapatkan informasi itu dibutuhkan sebuah platform untuk memudahkan pencarian informasi mengenai kesehatan terutama kesehatan anak. Pada penelitian ini diterapkan sebuah sistem tanya jawab atau question answering dengan menerapkan Question Entailment untuk memudahkan mencari pertanyaan mengenai kesehatan anak dan mendapatkan jawaban yang tepat. Dataset yang digunakan dibuat berdasarkan hasil pengumpulan daftar pertanyaan dan jawaban dari buku tanya jawab kesehatan anak dan FAQ di internet. Model Question Entailment dibangun berdasarkan training korpus yang dibuat dan diuji menggunakan algoritma Support Vector Machine(SVM), Logistic Regression, Naïve Bayes, dan J48 Decision Tree. Hasil pengujian menunjukkan bahwa algoritma SVM memberikan performa terbaik dalam mengidentifikasi pertanyaan serupa, dengan metrik precision, recall, dan f1(f-measure).

***Kata Kunci**: Kesehatan anak, question answering, question entailment, training korpus, support vector machine (SVM)*

ABSTRACT

Health is one of the essential things in human life, especially in children, which allows children to grow and develop properly. Parents must also have the correct information about maintaining health, diet, and an excellent lifestyle. In getting that information, a platform is needed to facilitate the search for health information, especially children's health. In this research, a Question Answering System was implemented by applying Question Entailment to make it easier to find questions about children's health and get the correct answers. The dataset is created based on collecting a list of questions and answers from a child health question and answer book and Internet FAQs. The Question Entailment model is built based on a training corpus that has been created and tested using Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, and J48 Decision Tree algorithm. The experimental results show that the SVM method gives the best performance on identifying similar question, with precision, recall, and f1(f-measure) metrics.

***Keywords**: Children's health, question answering, question entailment, training corpus, support vector machine (SVM).*

I. INTRODUCTION

HEALTH is one of the most important things for humans, making it possible for everyone to live socially and economically productive lives. However, since the Coronavirus came to Indonesia in March 2020 [1], the health number of Indonesian citizens has decreased. They are vulnerable to Coronavirus or other diseases, so the Indonesian people must take better care of their health and body hygiene. During the COVID-19 pandemic, the Indonesian people are encouraged to carry out activities only at home. The government is not recommending leaving the house if they do not have essential activities. It is not recommended for children to frequently visit health facilities, especially children at a young age, because they do not have a robust immune system, and the implementation of the Integrated Healthcare Center program may be hampered. So that these limitations indicate the importance of having a Question Answering System that can help parents communicate if they have complaints or questions regarding their children. A Question Answering System should facilitate it.

The Question Answering System is a discipline in Computer Science in information retrieval and natural language processing that builds a system to answer user questions automatically. It is designed to meet human information needs that may arise in situations such as talking to a virtual assistant or interacting with search engines [2]. In this digital era, it is easy to find information related to health, especially children's health, because the data is publicly available on the Internet. However, the information obtained often includes details related to the disease, such as symptoms, actions, prevention, and treatment methods, making it more difficult to understand [3].

Therefore, in such cases, it can be overcome by applying Question Entailment to the Question Answering System, which is a method that seeks answers to a question on existing questions based on similar questions [4]. Moreover, many questions users ask on health-related topics are abstract and open-ended; therefore, traditional search methods are ineffective in such cases [3].

Research on Question Answering has been carried out by various researchers, one of which is Asma Ben et al. [4] with the title Recognizing Question Entailment for Medical Question Answering. This study discusses the Question Entailment method used in the health question and answer system he created to find similar questions based on the Question Entailment method. The data used are obtained from Frequently Asked Questions (FAQ), identical to specific consumer questions, and widely distributed. Many reliable sources, such as National Institutes of Health (NIH) institutes and centers, provide answers to frequently asked questions organized by topic. The NIH offers FAQs for many health problems, such as rare diseases (Rare diseases or Alzheimer's).

This research focuses on implementing the research conducted by Asma Ben et al. [4] by applying the Question Entailment method to the Question Answering System [5], which can return text quotes and even phrases into answers. This method will be assisted by TF-IDF [6] and Bigram [7], which are tested using classification algorithms such as Support Vector Machine (SVM) and the entailment of new questions with dataset questions.

Based on this background, in this research, Question Entailment was applied to the Question Answering System using a dataset containing questions and answers regarding children's health. The dataset collected is obtained from credible experts from health information sources to validate the information obtained. The result will know the entailment between questions in the new question and the existing questions in the dataset.

II. RESEARCH METHOD

A. *System Design*

The stages of system development for this research begin with data collection. Furthermore, there is a division of positive data (combined questions with the correct category) and negative (combined questions with the wrong category). This was followed by preprocessing, looking for cosine similarity values, and selecting features. After that, making five training corpus will be tested using the SVM algorithm to be included in the Question Entailment model. The stages of the system to be built are shown in Fig. 1.

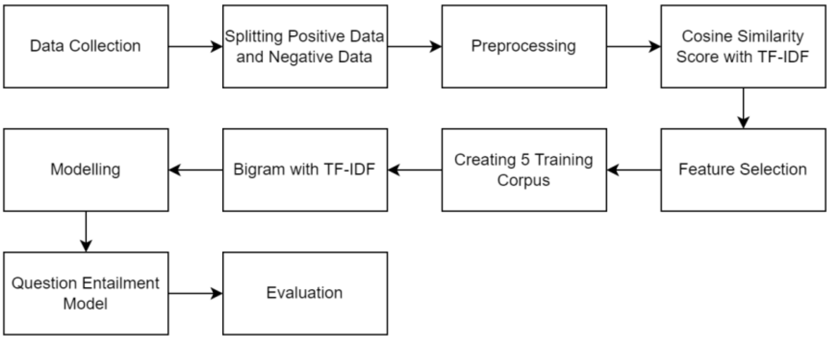
Data Collection
Splitting Positive Data and Negative Data
Preprocessing
Cosine Similarity Score with TF-IDF
Feature Selection
Creating 5 Training Corpus
Bigram with TF-IDF
Modelling
Question Entailment Model
Evaluation

Fig. 1. System development stages

B. *Data Collection*

Data on questions about children's health in Indonesian needed for this research were collected from the Instagram account *@tentanganakofficial*, a book titled "*250 tanya jawab kesehatan anak*" (250 child health Q&A), and several health websites. The amount of data obtained is 300 data. The following list of resources used to build the dataset is shown in Table I. Data from books and other platforms contain information on common questions that pediatricians often ask. It consists of three columns: ' Question,' 'Answer,' and 'Category.' Sample data from the dataset is shown in Table II.

TABLE I
LIST OF RESOURCES FOR QUESTIONS AND ANSWERS ON THE DATASET

Resource	**Platform**
tentanganakofficial	Instagram Account
250 tanya jawab kesehatan anak (250 child health Q&A)	Book
Researchgate.com	Website
Alodokter.com	Website

TABLE II
DATA EXAMPLE FROM DATASET

Attribute Name	**Value**
Question	*Anak saya (10 bulan) terkena campak. Dia belum pernah mendapat imunisasi campak, apakah imunisasi campak masih diperlukan?* (My child (10 months old) has measles. He has never been immunized against measles, is measles immunization still needed?)
Answer	*Kalau sudah dipastikan anak sudah terkena campak, ia tidak perlu imunisasi campak.* (If it's confirmed that the child has already had measles, he/she does not need measles immunization)
Category	*Pengobatan dan Penyakit* (Treatment and Diseases)

C. Data Preprocessing

In this research, we do data preprocessing to transform the data to be more suitable for processing by the Question Entailment algorithm. Because the data used was in the form of text, Text Preprocessing was used. The text Preprocessing [8] includes Case folding, Punctuation Removal, Stopwords Removal, and Stemming.

Case Folding is one of the processes in Text Preprocessing that is carried out to uniform the characters in the data that can change all letters to lowercase. In this research, the case folding process was to make all letters in the 'Question' and 'Category' data small/lowercase. Table III shows the data after going through the Case folding process.

TABLE III
DATA EXAMPLE AFTER CASE FOLDING PROCESS

Attribute Name	**Value**
Question	*anak saya (10 bulan) terkena campak. dia belum pernah mendapat imunisasi campak, apakah imunisasi campak masih diperlukan?* (my child (10 months old) has measles. he has never been immunized against measles, is measles immunization still needed?)
Answer	*Kalau sudah dipastikan anak sudah terkena campak, ia tidak perlu imunisasi campak.* (If it's confirmed that the child has already had measles, he/she does not need measles immunization)
Category	*pengobatan dan penyakit* (treatment and diseases)

Punctuation Removal is one of the processes in Text Preprocessing that is carried out to remove punctuation marks such as commas, periods, and others. In this research, the Punctuation Removal process was carried out to remove punctuation marks and numbers in the 'Question' data. Table IV shows the data after going through the Punctuation Removal process.

TABLE IV
DATA EXAMPLE AFTER PUNCTUATION REMOVAL PROCESS

Attribute Name	**Value**
Question	*anak saya bulan terkena campak dia belum pernah mendapat imunisasi campak apakah imunisasi campak masih diperlukan* (my child month has measles he/she never been immunized against measles is measles immunization still needed)
Answer	*Kalau sudah dipastikan anak sudah terkena campak, ia tidak perlu imunisasi campak.* (If it's confirmed that the child has already had measles, he/she does not need measles immunization)
Category	*pengobatan dan penyakit* (treatment and diseases)

Stopwords Removal [9] is one of the processes in Text Preprocessing, which is carried out for filtering or removing words that have little impact on the text or a question. In this research, the Stopwords Removal process using the library from NLTK is to remove words that have little effect on the text in the 'Question' data. Table V shows the data after going through the Stopwords Removal process.

TABLE V
DATA EXAMPLE AFTER STOPWORDS REMOVAL PROCESS

Attribute Name	**Value**
Question	*anak terkena campak imunisasi campak imunisasi campak* (child has measles measles immunization measles immunization)
Answer	*Kalau sudah dipastikan anak sudah terkena campak, ia tidak perlu imunisasi campak.* (If it's confirmed that the child has already had measles, he/she does not need measles immunization)
Category	*pengobatan dan penyakit* (treatment and diseases)

Stemming is one of the processes in Text Preprocessing that removes words that have affixes and makes them return to basic words. In this research, the Stemming Process is to remove words that have affixes and make them return to the base word in the 'Question' data. Because the data used is Indonesian language data, the library from Sastrawi [10] is used for the Stemming process. Table VI shows the data after going through the Stemming process.

TABLE VI
DATA EXAMPLE AFTER STEMMING PROCESS

Attribute Name	**Value**
Question	*anak kena campak imunisasi campak imunisasi campak* (child has measles measles immunization measles immunization)
Answer	*Kalau sudah dipastikan anak sudah terkena campak, ia tidak perlu imunisasi campak.* (If it's confirmed that the child has already had measles, he/she does not need measles immunization)
Category	*pengobatan dan penyakit* (treatment and diseases)

D. TF-IDF

TF-IDF is used to get the word weight value in 'Question' and 'Category,' and the weight value is used to get the cosine similarity value. TF-IDF is also used to get the word weight values in the train and test data in the Bigram process to fit into the model. The word weight value is obtained using the TF-IDF Vectorizer to form a vector containing the weight values of each word. TF-IDF process to create a term frequency matrix where are distinct terms throughout all documents (1). Compute inverse document frequency (IDF) using formula in (2) equation. Multiply the TF matrix with IDF (3).

$$TF = 1 + log_{10}(f_t, d), f_t > 0 \tag{1}$$

F_t = Term frequency
d = Document / data

$$idf_i = log\left(\frac{n}{df_i}\right) \tag{2}$$

idf_i = IDF score for term i
n = total number of documents
df_i = The number of documents containing term i

$$W_{i,j} = \mathrm{tf}_{i,j} \times idf_i \tag{3}$$

$W_{i,j}$ = TF-IDF score for term i in document j
$tf_{i,j}$ = Term frequency for term i in document j
idf_i = IDF score for i

E. Cosine Similarity

We use Cosine Similarity to measure the similarity between the 'Question' and 'Category' data. Also, when finding out the entailment between questions, the Cosine Similarity method is used between the 'New Question' and

'Dataset Questions' data. This process is done by calculating the cosine [11] value between two texts with the equation (4).

$$similarity = cos(\theta) = \frac{A \cdot B}{||A||||B||} \tag{4}$$

A = Vector A
B = Vector B
||A|| = Vector A length
||B|| = Vector B length

F. Training Corpus

We select the features used in the Question Entailment model stage with a predefined algorithm. The set features are 'Question' data and 'cosine' data (cosine value for each question and category). The 'Question' data is combined with the 'Category' data into 'Combined Question' data (Combination of 'Question' and 'Category' data).

After selecting the features, we create a training corpus that will be used in the model. There is five training corpus [4] that has been developed, which contain a mixture of positive question data (questions with the correct category) and negative questions (questions with the wrong category). Table VII shows the composition of five training corpus.

TABLE VII
COMPOSITION OF FIVE TRAINING CORPUS

Training Corpus	Positive Question	Negative Question
Corpus 1	54.2%	45.8%
Corpus 2	26%	74%
Corpus 3	33.4%	66.6%
Corpus 4	25%	75%
Corpus 5	50%	50%

After divided into five training corpus, the data for each training corpus is split into X and Y. The X data consists of 'question' and 'cosine' data, and the Y data consists of 'label' data which shows a positive or negative label in the form of numbers 0 (negative) and 1 (positive). Later, the library from SKLearn will be used for splitting five train data and five test data based on data from X and Y. The composition of train data is 70%, and test data is 30%.

G. Bigram

We use bigram on five train data and five test data that have been built on the training corpus to process the question sentences on the train data and test data into chunks per 2 words. The bigram used is the TF-IDF Vectorizer by adding the parameter ngram=2. This process produces a vector that will fit into the algorithm and used in the Question Entailment Model.

H. Model Question Entailment

The Support Vector Machine (SVM) [4] algorithm is used to test the five corpus that has been created. In this algorithm, a linear kernel is used to test it, and the highest result is selected based on the precision, recall, and f1 (f-measure) metrics. Because at each test runtime, the results are different, so ten tests were carried out to find the highest precision, recall, and f1 (f-measure) values in the five training corpus.

The training corpus with the highest score is used for further testing using the Logistic Regression, Naïve Bayes, and J48 Decision Tree algorithms [4]. The algorithm that produces the highest precision, recall, and f1(f-measure) metrics will be used as a Question Entailment model [4]. In this Question Entailment model, we examine the entailment of new questions with questions on the dataset.

I. Evaluation

Our evaluation includes three stages. The first stage tests five training corpus with the Support Vector Machine (SVM) algorithm to find the corpus that produces the best value [4]. The second stage tests the best training corpus using the Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, and J48 Decision Tree algorithm to find the algorithm that produces the best value precision, recall, and f1(f-measure) metrics [4]. The third stage tests the best algorithm and training corpus on the Question Entailment model to test Entailment on new questions with dataset questions [4].

III. RESULT AND DISCUSSION

A. Test Result

This test will display the results from Modeling to Testing to get Entailment between questions. In the first stage, testing five training corpus using the Support Vector Machine (SVM) method to get the highest precision, recall, and f1 (f-measure) values were tested ten times with different runtimes. Table VIII shows the results of ten-time test.

TABLE VIII
RESULT OF TEN TESTS IN DIFFERENT RUNTIMES

Runtime	**Training Data**	**Value**		
		Precision	**Recall**	**F1**
1	**1**	**47**	**59.2**	**52.4**
	2	33.6	45.5	38.7
	3	32.6	45.5	37.9
	4	30.5	45.2	40.5
	5	38.5	48.6	43
2	**1**	**48.2**	**51.8**	**50**
	2	40	45	42.3
	3	37.2	42.6	37.1
	4	37.2	47.2	41.6
	5	38.2	59	46.4
3	**1**	**56.2**	**66.6**	**61**
	2	36.3	43.2	39.5
	3	38	41.5	39.3
	4	37.7	44.7	40
	5	35.2	24	28.5
4	**1**	37.8	**63.3**	**47.4**
	2	**41.1**	45.1	43
	3	36.2	38.1	37.1
	4	36.3	43.2	39.5
	5	40.7	44	42.3
5	**1**	**43.5**	**68**	**53.1**
	2	38	46	41.6
	3	29.3	45	35.5
	4	38	41	39.5
	5	36	37.5	36.7
6	**1**	37.5	**50**	**42.8**
	2	**42.8**	44.7	41.4
	3	41.3	47.3	39.3
	4	38.8	44.8	41.6
	5	34.7	32	33.3
7	1	46.1	40	42.8
	2	29.7	36.6	36.3
	3	31.2	36.7	33.7
	4	39.7	43.7	41.6
	5	**48**	**44.4**	**46.1**
8	**1**	**38.7**	42.8	**40.6**
	2	35.5	44.4	39.5
	3	27.6	**46.4**	34.6
	4	32.6	45.4	37.9
	5	36.8	25	29.7
9	**1**	**50**	**49.2**	**49.7**
	2	40.4	41.4	40.9
	3	27.3	38.3	31.9
	4	35.4	48.5	40.9
	5	29	40.9	33.9
10	**1**	**50.1**	**55.5**	**51.7**
	2	38.7	44.7	40.9
	3	49.7	49.8	48.9
	4	37.5	48.6	42.3
	5	33.3	25	28.5

The results of the ten-time runtime test on five training corpus using the Support Vector Machine (SVM) method were able to determine which corpus gave the highest precision, recall, and f1 (f-measure) values. At 9x runtimes, the first corpus produces the highest value, and at 1x other runtimes, the fifth corpus (runtime 7) produces the highest value. So, the first training corpus was chosen for the second test, which was tested with three other algorithms: Logistic Regression, Naïve Bayes, and J48 Decision Tree. The results of the second stage test are shown in Table IX.

TABLE IX
RESULT OF TESTING BEST CORPUS WITH 4 ALGORITHMS

Algorithm	**Precision**	**Recall**	**F1(F-Measure)**
SVM	**65**	**46.5**	**54.1**
Logistic Regression	56.5	46.4	50.9
Naïve Bayes	58.8	35.7	44.4
J48 Decision Tree	61.5	28.5	39

Based on Table IX, it is found that the first training corpus test with these four algorithms shows that the Support Vector Machine algorithm can provide the highest precision, recall, and f1 (f-measure) values than the other three algorithms. By testing the training corpus and the algorithms in Table VIII and Table IX, we get the corpus and the algorithm that produces the highest precision, recall, and f1 (f-measure) values, namely the first training corpus. Moreover, the Support Vector Machine algorithm is used to test the Question Entailment model on the Question Answering System between new questions and existing questions in the dataset. The third-stage test results are shown in Table X.

TABLE X
RESULT OF QUESTION ENTAILMENT OF 2 QUESTIONS

No	**Attribute**	**Value**
1	Q1 (new question)	*Apakah ASI penting bagi bayi yang berumur 1 bulan?* (Is breast milk important for a 1-month-old baby?)
	Q2 (dataset question)	*Mengapa ASI begitu penting bagi bayi?* (Why is breast milk so important for babies?)
	Entailment	True
	Answer	*Tak dapat dipungkiri ASI merupakan makanan ideal untuk bayi, terutama pada bulan-bulan pertama kelahiran. ASI mengandung semua zat gizi yang dibutuhkan bayi, termasuk zat untuk tumbuh-kembang serta zat-zat anti- infeksi. Selain itu, pemberian ASI juga bermanfaat bagi Mama untuk mempercepat pengembalian besarnya rahim ke bentuk dan ukuran semula serta mengurangi insiden kanker payudara di kemudian hari.* (Breast milk is undeniably the ideal food for babies, especially in the first months of birth. Breast milk contains all the nutrients that babies need, including substances for growth and development as well as anti-infective substances. In addition, breastfeeding is also beneficial for Mama to accelerate the return of the size of the uterus to its original shape and size and reduce the incidence of breast cancer in the future.)
2	Q1 (new question)	*Apakah anak yang sedang menderita penyakit yang berinfeksi tidak boleh sekolah?* (Should children with an infectious disease not go to school?)
	Q2 (dataset question)	*Mengapa ASI begitu penting bagi bayi?* (Why is breast milk so important for babies?)
	Entailment	False
	Answer	-
3	Q1 (new question)	*Anak saya berumur 3 bulan sedang mengalami sakit cantengan di jempol kaki nya, kondisinya sudah membengkak, bagaimana cara mengasinya ya dok?* (My 3-month-old child is experiencing a cantengan pain on his big toe, his condition has swollen, how to treat it, doc?)
	Q2 (dataset question)	*Cara mengobati kuku cantengan pada bayi?* (How to treat baby's hangnails?)
	Entailment	True
	Answer	*Kami mengerti kekhawatiran yang Anda rasakan. Penanganan awal di rumah Anda bisa menggunakan air sabun yang hangat, selama 10 menit. Jika sangat nyeri ibu bisa berikan obat pengurang nyeri yang mengandung paracetamol. Jika sudah empat hari tetapi kuku cantengan pada bayi tak kunjung membaik, ada baiknya Anda membawanya ke dokter untuk mendapatkan pengobatan yang tepat. Selain itu, kuku yang memerah, bengkak, dan mengeluarkan nanah harus segera ditangani untuk mencegah infeksi pada kuku bayi.* (We understand your concern. Initial treatment at home you can use warm soapy water, for 10 minutes. If it is very painful, mom can give pain relievers containing para-cetamol. If it has been four days but the baby's nail has not improved, it is better to take him to the doctor to get the right treatment. In addition, nails that are red, swollen, and oozing pus should be treated immediately to prevent infection of the baby's nails.)

B. Test Analysis

The results of the training corpus test with the SVM algorithm showed that the first training corpus with a composition of 54.2% positive and 45.8% negative questions gave the highest value than another training corpus. Most likely, this is due to the many positive questions in the first training corpus. The precision, recall, and f1(f-measure) values given by the first training corpus are higher than other training corpuses but are not high enough for optimal values, probably due to the lack of datasets.

In the testing of four algorithms on the first training corpus, the SVM algorithm produces the highest value compared to other algorithms. However, compared to the Logistic Regression algorithm, the value difference is not too significant. This indicates that the SVM algorithm is slightly superior in processing our data. Testing the Question Entailment model with the SVM algorithm and training corpus 1 gave good results in predicting Entailment between questions, as shown in Table X. This Question Answering System shows that the system can provide results that follow user needs.

IV. CONCLUSION

This research has implemented a Question Entailment in a Question Answering System using a child health dataset. According to the test results, the Question Answering System can show Entailment between questions well and provide answers following user questions. However, the result of precision, recall, and f1(f-measure) values are still not optimal. Precision, recall, and f1 (f-measure) values must be increased. The number of datasets must be increased and equipped with features that can improve performance so this Question Answering System can be used and relied on in real-world scenarios.

REFERENCES

[1] R. F. Nursofwa, M. H. Sukur, B. K. Kurniadi, and . H., “Penanganan Pelayanan Kesehatan Di Masa Pandemi Covid-19 Dalam Perspektif Hukum Kesehatan,” *Inicio Legis*, vol. 1, no. 1, pp. 1–17, 2020, doi: 10.21107/il.v1i1.8822.

[2] M. A. Calijorne Soares and F. S. Parreiras, “A literature review on question answering techniques, paradigms and systems,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 32, no. 6, pp. 635–646, 2020, doi: 10.1016/j.jksuci.2018.08.005.

[3] M. Zhu, W. Wei, A. Ahuja, and C. K. Reddy, “A hierarchical attention retrieval model for healthcare question answering,” *Web Conf. 2019 - Proc. World Wide Web Conf. WWW 2019*, pp. 2472–2482, 2019, doi: 10.1145/3308558.3313699.

[4] A. Ben Abacha and D. F. Dina, “Recognizing Question Entailment for Medical Question Answering,” *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2016, pp. 310–318, 2016.

[5] A. Ben Abacha and D. Demner-Fushman, “A question-entailment approach to question answering,” *BMC Bioinformatics*, vol. 20, no. 1, pp. 1–23, 2019, doi: 10.1186/s12859-019-3119-4.

[6] R. Kustiawan, A. Adiwijaya, and M. D. Purbolaksono, “A Multi-label Classification on Topic of Hadith Verses in Indonesian Translation using CART and Bagging,” *J. Media Inform. Budidarma*, vol. 6, no. 2, pp. 868–875, 2022, doi: 10.30865/mib.v6i2.3787.

[7] K. Bae and Y. Ko, “Improving question retrieval in community question answering service using dependency relations and question classification,” *J. Assoc. Inf. Sci. Technol.*, vol. 70, no. 11, pp. 1194–1209, 2019, doi: 10.1002/asi.24196.

[8] M. R. Choirulfikri, K. M. Lhaksamana, and S. Al Faraby, “A Multi-Label Classification of Al-Quran Verses Using Ensemble Method and Naïve Bayes,” *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 473–479, 2022, doi: 10.47065/bits.v3i4.1287.

[9] H. Veisi and H. F. Shandi, “A Persian Medical Question Answering System,” *Int. J. Artif. Intell. Tools*, vol. 29, no. 6, pp. 1–29, 2020, doi: 10.1142/S0218213020500190.

[10] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.

[11] D. Gunawan, C. A. Sembiring, and M. A. Budiman, “The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents,” *J. Phys. Conf. Ser.*, vol. 978, no. 1, 2018, doi: 10.1088/1742-6596/978/1/012120.