



## Analisis Metode Representasi Teks Untuk Deteksi Interelasi Kitab Hadis: *Systematic Literature Review*

Amelia Devi Putri A<sup>1</sup>, Chastine Faticah<sup>2</sup>, Agus Zainal Arifin<sup>3</sup>

<sup>1,2,3</sup>Departemen Teknik Informatika, Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember, Surabaya

<sup>1</sup>ameliadev26@gmail.com, <sup>2</sup>chastine@cs.its.ac.id, <sup>3</sup>agusza@cs.its.ac.id

### Abstract

*Hadith is the second source of reference for Islamic law after the Qur'an, which explains the sentences in the Qur'an which are still global by referring to the provisions of the Prophet Muhammad SAW. Classification of text documents can also be used to overcome the problem of interrelation between the Qur'an and hadith. The problem of interrelation between books of hadith needs to be done because some hadiths in certain hadith books have the same meaning as other hadith books. This study aims to analyze the development of text representation and classification methods suitable to overcome similarity meaning problems in detecting interrelationships between hadith books. The research method used is Systematic Literature Review (SLR) sourced from Google Scholar, Science Direct, and IEEE. There are 42 pieces of literature that have been studied successfully. The results showed that contextual embedding as the newest text representation method considered word context and sentence meaning better than static embedding. As a classification method, the ensemble method has better performance in classifying text documents than using only a single classifier model. Thus, future research can consider using a combination of contextual embedding and ensemble methods to detect interrelationships between books of hadith.*

*Keywords: Hadith, Interrelation, Text Document Classification, Text Representation*

### Abstrak

Hadis merupakan sumber rujukan hukum Islam kedua sesudah Al-Qur'an, yang menjelaskan kalimat dalam Al-Qur'an yang masih bersifat global dengan mengacu pada ketetapan Nabi Muhammad SAW. Untuk memudahkan seorang muslim mempelajari Al-Qur'an maupun hadis dapat dilakukan dengan menggunakan klasifikasi dokumen teks. Klasifikasi dokumen teks juga dapat digunakan untuk mengatasi permasalahan interelasi antara Al-Qur'an dan hadis. Namun, sejauh ini belum ada penelitian yang melakukan interelasi antar kitab hadis. Permasalahan interelasi antar kitab hadis perlu dilakukan karena pada beberapa hadis dalam kitab hadis tertentu memiliki kesamaan makna dengan kitab hadis lain sehingga perlu dikelompokkan sesuai dengan topiknya. Tujuan dari penelitian ini adalah menganalisis perkembangan metode representasi teks dan metode klasifikasi supaya dapat ditemukan metode yang cocok untuk mengatasi permasalahan kesamaan makna pada pendeteksian interelasi antar kitab hadis. Metode penelitian yang digunakan adalah *Systematic Literature Review (SLR)* yang bersumber dari *Google Scholar, Science Direct* dan *IEEE*. Setelah melalui proses penilaian kualitas literatur, terdapat 42 literatur yang berhasil dikaji. Hasil penelitian menunjukkan permasalahan kesamaan makna dapat diatasi dengan menggunakan *contextual embedding*. *Contextual embedding* merupakan metode representasi teks terbaru, yang mampu mempertimbangkan konteks kata dan makna kalimat lebih baik daripada *static embedding*. Dalam hal pendeteksian interelasi antar kitab hadis, ditemukan metode yang cocok digunakan sebagai metode klasifikasi adalah metode *ensemble*. Metode *ensemble* memiliki kinerja yang lebih baik dalam mengelompokkan dokumen teks, daripada hanya menggunakan *single classifier* model saja. Dengan demikian, pada penelitian selanjutnya dapat dipertimbangkan penggunaan kombinasi *contextual embedding* dan metode *ensemble* untuk deteksi interelasi antar kitab hadis.

Kata kunci: Hadis, Interelasi, Klasifikasi Dokumen Teks, Representasi Teks

### 1. Pendahuluan

Salah satu dari sumber utama ajaran Islam adalah hadis [1]. Hadis diperlukan agar dalam menjalankan syari'at Islam dapat berjalan dengan baik dan dimengerti secara

keseluruhan [2]. Hal tersebut dikarenakan hadis menjadi sumber rujukan hukum Islam kedua sesudah Al-Qur'an yang menjelaskan kalimat dalam Al-Qur'an yang masih bersifat global dengan merujuk pada perkataan, ketetapan atau perbuatan yang dilakukan oleh Nabi

Muhammad SAW [3][4][5]. Selain itu, semakin berkembangnya teknologi, membuat Al-Qur'an maupun hadis dapat tersedia sepanjang waktu melalui situs website tertentu di internet sehingga dapat memudahkan bagi seorang muslim untuk mempelajari kedua sumber utama ajaran Islam tersebut berdasarkan topik atau kategori yang dapat disesuaikan dengan kebutuhan [1][6].

Dalam hal ini, untuk memudahkan seorang muslim mempelajari Al-Qur'an maupun hadis berdasarkan topik atau kategori yang sesuai dengan kebutuhan, dapat dilakukan dengan menggunakan klasifikasi dokumen teks, yang merupakan proses untuk mengklasifikasikan dokumen teks tidak berlabel ke dalam kategori yang sudah ditentukan sebelumnya [7][8]. Berdasarkan penelitian yang dilakukan oleh Rostam dan Malim [1] untuk mengetahui keterkaitan atau interelasi antara dua sumber utama dalam ajaran Islam yaitu Al-Qur'an dan hadis dapat diatasi dengan menggunakan klasifikasi dokumen teks melalui penggunaan pembobotan kata TF-IDF (*Term Frequency Inverse Document Frequency*) dan algoritma klasifikasi *Support Vector Machine* (SVM), yang memperoleh hasil akurasi lebih baik daripada dengan menggunakan algoritma klasifikasi yang lain seperti *K-Nearest Neighbours* (KNN) maupun *Naïve Bayes*. Penelitian Utomo et al. [9] juga menggunakan pembobotan kata TF-IDF sebagai metode representasi teks dan algoritma SVM digunakan sebagai metode klasifikasi. Kombinasi dari dua metode tersebut bertujuan untuk melakukan interelasi antar ayat-ayat Al-Qur'an. Interelasi yang dimaksud dalam penelitian tersebut adalah memetakan ayat-ayat Al-Qur'an sesuai dengan temanya. Dalam penelitian Mondal et al. [10] juga dilakukan interelasi antara teks review pengguna dengan *rating* bintang yang diberikan oleh pengguna di profil bisnis apapun. Penelitian tersebut menggunakan algoritma klasifikasi KNN untuk menemukan interelasi atau hubungan antara teks review pengguna dengan *rating* bintang tersebut.

Dari beberapa penelitian-penelitian yang sudah ada tersebut, belum ada yang melakukan interelasi antar kitab hadis. Padahal, penting juga untuk mengetahui keterkaitan antar kitab hadis karena pada beberapa hadis dalam kitab hadis tertentu memiliki kesamaan makna dengan kitab hadis lain sehingga perlu dikelompokkan hadis-hadis yang saling berkaitan satu sama lain berdasarkan kesamaan makna ke dalam satu kategori yang sama agar dapat memudahkan seorang muslim dalam mempelajari hadis [3]. Oleh karena itu, tujuan dari penelitian ini adalah menganalisis perkembangan metode representasi teks dan metode klasifikasi supaya dapat ditemukan metode yang cocok untuk mengatasi permasalahan kesamaan makna pada pendeteksian interelasi antar kitab hadis. Sebelum dilakukan analisis perkembangan metode representasi teks dan klasifikasi, terlebih dahulu dilakukan penguatan topik terkait deteksi interelasi antar kitab hadis. Penguatan topik

tersebut dilakukan dengan menganalisis dataset yang banyak digunakan dalam *Arabic natural language processing* sehingga dapat diketahui topik apa saja yang sudah dikembangkan sesuai dengan jenis datasetnya. Kemudian, baru dilakukan analisis perkembangan metode representasi teks dan metode klasifikasi dalam *Arabic natural language processing*, agar dapat diketahui metode representasi teks dan metode klasifikasi yang tepat untuk mendeteksi interelasi antar kitab hadis. Dengan demikian, hasil penelitian ini sangat bermanfaat bagi penelitian selanjutnya dalam menentukan metode yang paling sesuai untuk mendeteksi interelasi antar kitab hadis.

## 2. Metode Penelitian

Pada penelitian ini digunakan metode *systematic literature review* (SLR). Tahapan yang ada pada metode SLR berguna untuk mengidentifikasi, memberikan interpretasi dan evaluasi terhadap beberapa penelitian yang sudah ada dan relevan dengan topik penelitian guna menjawab pertanyaan penelitian, memberi tambahan materi pembelajaran dan mencari celah dari penelitian sebelumnya supaya dapat bermanfaat bagi penelitian selanjutnya [11][12][13]. Penelitian ini mengadopsi dari penelitian SLR yang telah dilakukan oleh Hamid et al. [14]. Tahapan yang dilakukan dalam pembuatan SLR ini ada tiga, yaitu tahap pembuatan rencana awal, tahap pelaksanaan dan tahap pelaporan hasil tinjauan pustaka yang telah didapatkan [15][16].

### 2.1. Tahap Pembuatan Rencana Awal

Dalam tahap ini dilakukan identifikasi tujuan pembuatan penelitian *systematic literature review*, yang sudah dijelaskan pada bagian pendahuluan. Kemudian, dilanjutkan dengan menentukan pertanyaan penelitian berdasarkan tujuan penelitian yang sudah dibuat, supaya dalam proses mereview beberapa penelitian menjadi fokus pada satu tujuan [17][18][19]. Pembuatan pertanyaan penelitian dapat dilakukan dengan menggunakan bantuan pedoman PICO (*Population, Intervention, Comparison dan Outcome*). Pedoman PICO sangat berguna dan memudahkan dalam merumuskan pertanyaan penelitian [20]. Tabel 1 menunjukkan pedoman PICO yang telah dibuat dan pada Tabel 2 menjelaskan rumusan pertanyaan penelitian yang berhasil didapatkan.

### 2.2. Tahap Pelaksanaan

Pada tahap pelaksanaan, terdapat beberapa tahap lagi didalamnya. Beberapa tahap tersebut meliputi tahap pencarian literatur, tahap seleksi dan tahap penilaian kualitas. Tahap pencarian literatur dilakukan pada beberapa database ilmiah digital seperti *Science Direct*, *IEEE* dan *Google Scholar*. Saat melakukan pencarian pada database ilmiah digital digunakan kata kunci "*Arabic text classification*", "*Arabic text categorization*", "*hadith classification*", "*interrelation*

*hadith classification*”, “*word embedding*” dan “*word vector*”. Kata kunci tersebut didapatkan dari proses identifikasi beberapa istilah yang ada di PICO maupun pertanyaan penelitian yang sudah dibuat sebelumnya.

Tabel 1. Pedoman PICO

PICO	Detail
Population	Aplikasi <i>software</i>
Intervention	<i>Arabic text classification</i> , <i>Arabic text categorization</i> dan <i>word embedding</i>
Comparison	Tidak ada
Outcome	Menganalisis dan mengevaluasi metode untuk merepresentasikan teks dan metode klasifikasi dalam <i>Arabic natural language processing</i> yang dapat digunakan untuk mendeteksi interelasi antar kitab hadis

Setelah didapatkan 552 literatur dari proses pencarian, kemudian hasil beberapa literatur tersebut dilanjutkan ke tahap seleksi inklusi dan eksklusi. Kriteria inklusi dan eksklusi yang ditetapkan dapat dilihat pada Tabel 3. Selanjutnya, tahap penilaian kualitas dilakukan setelah tahap seleksi inklusi dan eksklusi. Penilaian kualitas bertujuan untuk menghindari bias dari peneliti serta dapat menjadi acuan dalam menerjemahkan temuan-temuan saat dilakukan sintesis data sehingga dapat memberikan kekuatan pada hasil dan kesimpulan nantinya. Tabel 4 menjelaskan kriteria pertanyaan untuk penilaian kualitas yang ditetapkan pada penelitian ini. Setelah dilakukan proses inklusi/eksklusi dan penilaian kualitas didapatkan sebanyak 42 literatur yang memenuhi penilaian kualitas karena telah menjawab semua poin penting pada kriteria pertanyaan penilaian kualitas literatur. Selain itu, dari 42 literatur tersebut juga relevan dengan tujuan penelitian ini sehingga dapat digunakan untuk tahap pelaporan hasil. Gambar 1 menjelaskan alur metode pengumpulan literatur yang dilakukan dalam penelitian ini supaya lebih mudah dimengerti.

### 2.3. Tahap Pelaporan Hasil

Dalam tahap pelaporan hasil, akan dijelaskan hasil temuan apa saja yang telah didapatkan. Pelaporan hasil dapat berupa rangkuman tiap pertanyaan penelitian dengan disertai tabel-tabel untuk meningkatkan visualisasi penyajian hasil. Tahap pelaporan akan dijelaskan secara detail pada bagian selanjutnya yaitu hasil dan pembahasan.

Tabel 2. Rumusan Pertanyaan Penelitian

ID	Pertanyaan Penelitian	Tujuan
RQ1	Dataset apa saja yang digunakan dalam <i>Arabic natural language processing</i> ?	Menganalisis dataset yang banyak digunakan dalam <i>Arabic natural language processing</i> dan menganalisis topik penelitian apa saja yang sudah dikembangkan sesuai dengan jenis

RQ2	Metode representasi teks apa saja yang digunakan dalam <i>Arabic natural language processing</i> ?	datasetnya guna menguatkan topik deteksi interelasi antar kitab hadis. Mengidentifikasi metode representasi teks yang sudah diusulkan dalam <i>Arabic natural language processing</i> dan menganalisis metode representasi teks yang cocok untuk mendeteksi interelasi antar kitab hadis.
RQ3	Metode klasifikasi apa yang digunakan dalam <i>Arabic natural language processing</i> ?	Mengidentifikasi metode klasifikasi yang sudah digunakan dalam <i>Arabic natural language processing</i> dan menganalisis metode klasifikasi yang cocok untuk mendeteksi interelasi antar kitab hadis.

Tabel 3. Kriteria Inklusi dan Eksklusi

Kriteria Inklusi	Kriteria Eksklusi
Literatur yang diterbitkan antara tahun 2016 hingga 2021	Literatur yang tidak relevan
Literatur yang diterbitkan dalam jurnal nasional maupun internasional dan konferensi internasional	Literatur yang membahas topik diluar batasan dari tujuan penelitian
Literatur yang secara langsung maupun tidak langsung menjawab pertanyaan penelitian	Literatur yang hanya memiliki abstrak saja

Tabel 4. Kriteria Pertanyaan Penilaian Kualitas

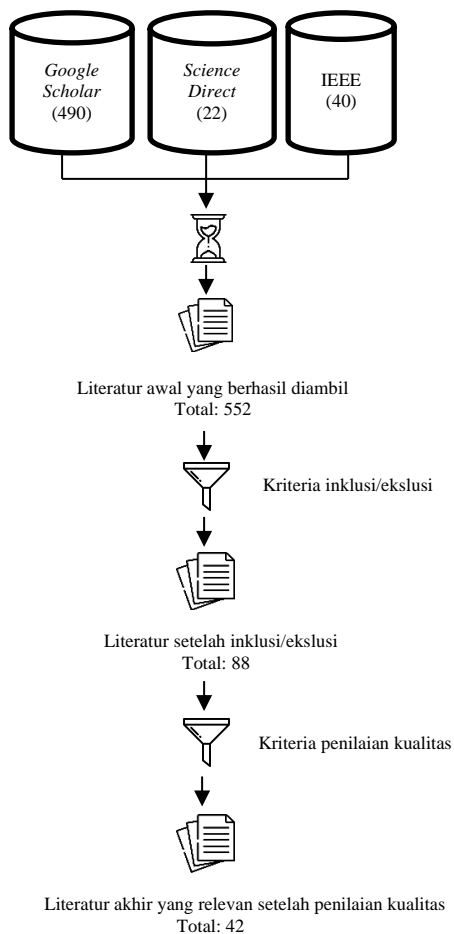
ID	Detail
QA1	Apakah desain penelitian sesuai dengan tujuan penelitian?
QA2	Apakah dataset Arabic dalam penelitian tersebut dilaporkan?
QA3	Apakah metode representasi teks dalam penelitian tersebut dilaporkan?
QA4	Apakah metode klasifikasi dalam penelitian tersebut dilaporkan?
QA5	Apakah ada pernyataan temuan yang jelas?

## 3. Hasil dan Pembahasan

### 3.1. Dataset yang Digunakan Dalam *Arabic Natural Language Processing* (ANLP) Untuk Menguatkan Topik Deteksi Interelasi Antar Kitab Hadis

Berdasarkan 42 literatur yang sudah didapatkan, ada empat jenis dataset yang biasa digunakan dalam *Arabic Natural Language Processing*, yaitu dataset yang diambil dari Twitter, berita berbahasa Arab, Hadis dan Al-Qur'an. Tabel 3 menunjukkan bahwa dari empat jenis dataset tersebut, dataset yang paling banyak digunakan adalah dataset yang diambil dari Twitter. Pengambilan data Twitter dapat dilakukan dengan *crawling* data sendiri menggunakan Twitter API maupun menggunakan dataset publik seperti *Arabic Sentiment Tweets Dataset* (ASTD). Dataset ASTD [21] merupakan kumpulan 10.000 *tweets* yang ditulis dalam

*Modern Standard Arabic* (MSA) dan dialek Mesir. Dataset tersebut dapat diunduh secara bebas di internet, yang terdiri dari 777 *tweets* positif, 1.642 *tweets* negatif, 805 *tweets* netral dan 6.466 *tweets* objektif [22]. Penelitian yang dilakukan oleh Oussous et al. menggunakan dataset ASTD untuk melakukan analisis sentimen pada data twitter berbahasa Arab dan menganalisis penggunaan pra-proses seperti *stemming* yang berbeda-beda [23], sedangkan penelitian Adel dan Wang, menggunakan data Twitter berbahasa Arab yang diambil melalui Twitter API untuk menganalisis komentar masyarakat dalam menanggapi krisis kemanusiaan yang ada seperti krisis kelaparan dan penyakit kolera di Yaman [24]. Kebanyakan penelitian yang menggunakan data Twitter mengambil topik penelitian mengenai analisis komentar masyarakat yang dituangkan dalam *tweets* terkait permasalahan tertentu.



Gambar 1 Alur Metode Pengumpulan Literatur

Selanjutnya, pada penelitian Abuhaiba dan Dawoud [25], dilakukan klasifikasi dokumen teks berbahasa Arab dengan menggunakan dataset publik yang dapat diakses secara bebas yaitu *Open Source Arabic Corpora* (OSAC). Dataset OSAC [26] merupakan kumpulan data standar yang besar untuk klasifikasi dokumen teks berbahasa Arab. Dataset tersebut terdiri dari korpus yang

berbeda, yaitu CNN, BBC dan OSAC. Korpus CNN diambil dari website berita *cnnarabic.com* dengan 5.070 dokumen teks dan enam kategori berita yaitu berita hiburan, bisnis, timur tengah, ilmu pengetahuan dan teknologi, olahraga dan berita dunia. Korpus BBC berasal dari website berita *bbc-arabic.com* yang terdiri dari 4.763 dokumen teks dengan tujuh kategori, dan korpus OSAC berasal dari beberapa website dengan jumlah keseluruhan teks dokumen sebanyak 22.429 yang terdiri dari 10 kategori. Di beberapa penelitian, penggunaan dataset berita berbahasa Arab dipakai untuk menguji performa metode yang diusulkan. Pada penelitian El-Alami et al. digunakan dataset OSAC untuk menganalisis metode yang diusulkan yaitu *pre-trained AraBERT* untuk klasifikasi *multi-class* dokumen teks berbahasa Arab [27]. Selain itu, penelitian Behassine et al. digunakan pula dataset OSAC untuk menguji metode yang diusulkan yaitu perbaikan metode seleksi fitur *Chi-square* guna meminimalkan data dan menghasilkan akurasi tinggi pada klasifikasi dokumen teks berbahasa Arab [28].

Untuk dataset Al-Qur'an, kebanyakan penelitian bertujuan untuk mengklasifikasikan teks Al-Qur'an dengan menggunakan algoritma klasifikasi yang berbeda-beda. Penelitian yang dilakukan Mohamed dan El-Bahaidy menggunakan dataset Al-Quran dan algoritma *ensemble* untuk klasifikasi *multi-label* teks Al-Qur'an menjadi 393 kelas atau topik bahasan [7]. Begitu pula pada penelitian yang dilakukan oleh Arkok dan Zeki, yang melakukan klasifikasi teks Al-Qur'an menjadi delapan kelas atau topik bahasan dengan menggunakan menggunakan algoritma *ensemble* [29]. Arkok dan Zeki juga melakukan penelitian lain dengan menggunakan dataset Al-Qur'an untuk mengatasi permasalahan klasifikasi dengan data tidak seimbang. Dataset Al-Qur'an dianggap sebagai dataset yang tidak seimbang untuk masing-masing jumlahnya. Ketidakseimbangan dataset Al-Qur'an disebabkan ayat-ayat dari topik Al-Qur'an yang jumlahnya tidak sama. Di dalam penelitian tersebut [30], digunakan algoritma *Synthetic Minority Oversampling Technique* (SMOTE), *Random Over Sample* (ROS), dan *Random Under Sample* (RUS) untuk mengatasi permasalahan data tidak seimbang.

Tabel 5. Analisis Penggunaan Dataset dalam *Arabic Natural Language Processing*

Dataset	Peneliti	Jumlah
Twitter	[33], [34], [35], [36], [23], [24], [37], [38], [39], [22], [40], [41], [42], [43], [44], [45], [46]	17
Berita berbahasa Arab	[28], [47], [48], [49], [27], [25], [50], [51]	8
Hadis	[52], [4], [53], [54], [55], [32], [56], [31], [57], [58], [59], [1]	12
Al-Qur'an	[30], [29], [7], [1], [9]	5

Topik penelitian yang biasa dibahas dengan menggunakan dataset hadis adalah mengklasifikasikan dokumen teks hadis berdasarkan keasliannya. Penelitian Abdelaal et al. [31] dan penelitian Ganem et al. [32] dilakukan klasifikasi teks hadis dengan menggunakan metode representasi teks TF-IDF dan mengelompokkannya menjadi empat kategori yang mengindikasikan keaslian suatu hadis, yaitu sah, hasan, da'if dan maudu. Hadis dikatakan sah menandakan bahwa hadis tersebut adalah baik dan sangat sesuai dengan norma-norma yang diucapkan Nabi Muhammad SAW. Hadis hasan berarti bahwa orang yang membawa riwayat hadis tersebut dilaporkan kurang taqwa daripada orang yang melaporkan hadis sah. Hadis da'if adalah hadis yang meleset satu atau dua syarat dari hadis sah maupun hasan, sedangkan hadis maudu menandakan bahwa teks hadis tersebut mengandung kepalsuan [32]. Dari beberapa penelitian dalam *Arabic Natural Language Processing* yang ada, terlebih yang menggunakan dataset hadis, sejauh ini belum ada yang mengambil topik mengenai deteksi interelasi antar kitab hadis, yang mana mengelompokkan hadis-hadis yang saling berkaitan satu sama lain berdasarkan kesamaan makna ke dalam satu kategori yang sama. Hal ini tentunya dapat menjadi penguat ide topik penelitian dan dapat dikembangkan pada penelitian selanjutnya.

### 3.2. Metode Representasi Teks yang Digunakan Dalam ANLP yang Cocok Untuk Deteksi Interelasi Antar Kitab Hadis

Salah satu metode *pre-processing* teks adalah representasi dokumen teks. Metode tersebut merupakan suatu cara untuk merepresentasikan dokumen menjadi vektor atau perubahan dokumen lengkap menjadi vektor seperti *index term*, yang paling banyak digunakan dalam *information retrieval* dan klasifikasi dokumen teks [60]. Berdasarkan Tabel 4 terlihat bahwa metode representasi teks berbahasa Arab yang paling banyak digunakan adalah *Term Frequency Inverse Document Frequency* (TF-IDF). Representasi teks TF-IDF diasumsikan bahwa kata yang jarang muncul pada kumpulan dokumen akan sangat bernilai. Formula untuk menghitung TF-IDF dapat ditulis dalam rumus 1, 2 dan 3. Dari beberapa rumus tersebut menunjukkan N adalah banyaknya dokumen teks, DF<sub>t</sub> atau *document frequency* adalah banyaknya dokumen yang mengandung kata t, W<sub>t,d</sub> adalah pembobotan kata TF-IDF, dan IDF<sub>t</sub> adalah perhitungan dari *inverse document frequency* [61].

$$W_{TF(t,d)} = \begin{cases} 1 + \log_{10} TF_{t,d}, & \text{jika } TF_{t,d} > 0 \\ 0, & \text{sebaliknya} \end{cases} \quad (1)$$

$$IDF_t = \log_{10}(N/DF_t) \quad (2)$$

$$W_{t,d} = W_{TF(t,d)} \times IDF_t \quad (3)$$

Namun, TF-IDF hanya memberikan bobot hubungan kata dari dua sumber tersebut berdasarkan frekuensi kemunculannya dan belum bisa menangkap lokasi dari kata-kata di dalam teks sehingga menyebabkan pembobotan kata TF-IDF menjadi kurang semantik [7][2][62]. Semantik merupakan salah satu cabang ilmu linguistik yang mempelajari representasi makna kata berdasarkan distribusinya dalam teks [63].

Berdasarkan penelitian yang dilakukan oleh Mohamed dan El-Behaidy [7] menyatakan bahwa saat ini teknik *word embedding* banyak digunakan dan dapat meningkatkan kinerja tugas *Natural Language Processing* (NLP) seperti kategorisasi atau klasifikasi teks, sentimen analisis dan *syntactic parsing*, dengan cara mengkonversi kata-kata menjadi vektor semantik sehingga dapat menjaga makna dan hubungan antar kata. Vektor semantik tersebut dapat dihitung dengan berbagai macam pendekatan seperti *word co-occurrence matrix*, representasi konteks berdasarkan kemunculan kata dan *neural networks* [62]. Beberapa vektor semantik atau *embedding* model yang dilakukan berdasarkan *neural networks* yang sangat populer pada tahun 2010-an adalah Word2Vec [64], FastText [65] dan AraVec [66]. *Embedding* model tersebut hanya akan menghasilkan *single static embedding* untuk masing-masing kata dalam kosakata, yang mana *static embedding* model tersebut hanya bisa mempelajari satu *embedding* per bentuk kata, padahal kenyataannya seperti dalam kasus bentuk kata pada bahasa Inggris maupun bahasa Arab secara umum, terdapat banyak *polysemous* yaitu kata yang sama namun memiliki makna yang berbeda-beda tergantung konteks kalimatnya [67].

Tabel 6. Analisis Metode Representasi Teks dalam *Arabic Natural Language Processing*

Metode Untuk Representasi Teks	Peneliti	Jumlah
TF-IDF	[4], [9], [48], [30], [23], [24], [29], [53], [54], [7], [38], [55], [32], [56], [25], [31], [59], [1], [51], [60]	20
Word2Vec	[34], [24], [49], [7], [57], [39], [41], [43], [45], [46]	10
AraVec	[34], [35], [37], [40], [41], [42]	6
FastText	[34], [47], [35], [57], [41], [42], [50], [46]	8
mBERT	[27], [22], [41], [46]	4
AraBERT	[36], [27]	2

Untuk lebih mempertimbangkan konteks kata dan makna kalimat dapat diatasi dengan *dynamic* atau *contextual embedding* yang mempelajari representasi kalimat secara universal [27]. *Contextual embedding* yang pertama kali diusulkan adalah *Embeddings from Language Model* (ELMO) [68] yang mempelajari *contextual embedding* berdasarkan konteks kata meskipun kata tersebut bersifat *polysemous* dengan

membentuk *multilayer Recurrent Neural Network* (RNN). Kemudian, ada penelitian terbaru untuk *contextual embedding* model yaitu *Bidirectional Encoder Representations from Transformers* (BERT) yang menghasilkan *contextual embedding* yang lebih baik dengan menerapkan transformator dua arah guna mendapatkan representasi kata yang lebih akurat [69].

Dalam menangani bahasa Arab terdapat versi BERT yaitu AraBERT, yang dapat mempertimbangkan informasi kontekstual dari teks. Berdasarkan penelitian yang dilakukan oleh El-Alami [27] terbukti bahwa AraBERT mengungguli *static embedding* yaitu Word2Vec. Hal tersebut dikarenakan AraBERT menggunakan representasi *sentence embeddings*, yang dapat menangkap hubungan semantik dalam teks dengan merepresentasikan hubungan antara beberapa kata dan frasa dalam *single dense vector*. AraBERT juga mengungguli *dynamic atau contextual embedding* lain yaitu mBERT dan ELMO. ELMO melakukan *contextual embedding* yang dihasilkan oleh *bidirectional recurrent neural network* secara sederhana dengan menggabungkan konteks kiri dan kanan pada lapisan terakhir saja sedangkan AraBERT menggabungkan konteks kiri dan kanan disemua layer dalam model transformer sehingga mampu mengatasi ambiguitas dan kompleksitas pada bahasa Arab dalam hal klasifikasi dokumen teks [27]. Dengan demikian, guna mendeteksi interelasi antar kitab hadis, *contextual embedding* seperti AraBERT dapat dipertimbangkan untuk digunakan dalam merepresentasikan teks.

### 3.3. Metode Klasifikasi yang Digunakan Dalam ANLP yang Cocok Untuk Deteksi Interelasi Antar Kitab Hadis

Pada Tabel 5 menunjukkan jenis metode klasifikasi yang banyak digunakan dalam *Arabic Natural Language Processing* adalah *machine learning*. Beberapa algoritma pengklasifikasi tunggal (*single classifier*) yang terdapat dalam *machine learning* tersebut adalah *Support Vector Machine* (SVM), *Logistic Regression*, *K-Nearest Neighbor* (KNN), *Naïve Bayes*, *Decision Tree* dan *Random Forest*. Elnagar et al. [70] menyatakan bahwa untuk meningkatkan kinerja algoritma kategorisasi atau klasifikasi dokumen teks dapat dilakukan dengan menggunakan metode *ensemble learning*. Metode *ensemble* dapat meningkatkan kinerja algoritma klasifikasi dokumen teks dengan mengintegrasikan berbagai macam *single classifier* model sehingga dapat membuat prediksi model menjadi lebih baik jika dibandingkan hanya dengan menggunakan *single classifier* model saja [71][72].

Dalam penelitian Oussouss et al. [23] dilakukan analisis pengaruh *stemming* dengan menggunakan metode *ensemble* yang menggabungkan beberapa *single classifier* seperti *Naïve Bayes*, SVM dan *Maximum Entropy*. Dari penelitian tersebut membuktikan bahwa

metode *ensemble* berperforma lebih baik daripada hanya dengan menggunakan *single classifier* saja dengan memperoleh nilai akurasi terbaik sebesar 89.5%. Mengintegrasikan berbagai macam *single classifier* model, mampu memberikan kelebihan bagi metode *ensemble* untuk mengurangi kemungkinan memilih *single classifier* model yang salah atau tidak sesuai. Kesalahan dalam memilih *single classifier* model dapat menyebabkan model menjadi *overfitting* yaitu ketika model tidak bekerja dengan baik pada data yang belum pernah dilihat selama proses pembelajaran atau *testing set*, dan terlalu berperforma baik pada *training set* saja [23][73][74]. Oleh karena itu, pada penelitian selanjutnya, dapat digunakan metode *ensemble* sebagai metode klasifikasi untuk mendeteksi interelasi antar kitab hadis.

Tabel 7. Analisis Metode Klasifikasi dalam *Arabic Natural Language*

Processing			
Jenis Metode Klasifikasi	Metode Klasifikasi Untuk Teks Berbahasa Arab	Peneliti	Jumlah
Machine Learning	SVM	[28], [33], [34], [47], [48], [30], [23], [24], [54], [38], [27], [56], [31], [57], [39], [1], [42], [51], [46], [9]	20
	Logistic Regression	[33], [47], [48], [56], [31], [39], [60], [46]	8
	KNN	[47], [48], [38], [56], [39], [59], [1], [51]	8
	Naïve Bayes	[33], [47], [48], [30], [23], [24], [38], [55], [56], [58], [39], [1], [51], [60], [46]	15
	Decision Tree	[28], [48], [30], [55], [58]	5
Deep Learning	Random Forest	[33], [48], [36], [24], [53], [55], [56]	7
	CNN	[4], [37], [54], [49], [27], [39], [50]	7
	RNN LSTM	[35], [37] [35], [49], [27], [39], [43], [50], [45]	2 7
Ensemble Learning	Gabungan metode <i>machine learning</i> Gabungan metode <i>deep learning</i>	[23], [29], [7], [25], [44] [40], [41]	5 2

Metode *ensemble* sendiri dapat dibagi menjadi dua kelompok berdasarkan struktur yang digunakan untuk membangun model *ensemble*, yaitu metode independen dan dependen. Metode dependen adalah *output* dari suatu *single classifier* digunakan untuk membangun *classifier* selanjutnya, dimana pengetahuan yang didapatkan dari iterasi sebelumnya bisa ditransfer ke

pembelajaran pada iterasi selanjutnya. Metode independen adalah membangun model *ensemble* secara independen dimana *output* dari beberapa *single classifier* digabungkan melalui metode kombinasi seperti *voting modules* atau metode *meta-learning* [75]. *Voting modules* pada metode *ensemble* dibagi menjadi dua teknik, yaitu *soft voting* dan *hard* atau *majority voting*. Pada metode *ensemble soft voting*, probabilitas rata-rata dari setiap *single classifier* akan digunakan sebagai skor untuk *voting* [76]. *Output* dari metode *ensemble* yang menggunakan teknik *soft voting* adalah kategori yang memiliki nilai *voting* tertinggi berdasarkan probabilitas rata-rata dari setiap *single classifier*-nya. Dalam *majority voting*, setiap *single classifier* pada *majority voting* secara independen memiliki satu suara/*vote*, yang mana kategori prediksi akan dipilih berdasarkan perolehan kategori mayoritas yaitu lebih dari separuh suara [73].

Dengan demikian, hasil penelitian ini menunjukkan bahwa berdasarkan jawaban dari pertanyaan penelitian kedua, metode yang cocok untuk mengatasi permasalahan kesamaan makna pada pendeteksian interelasi antar kitab hadis adalah dengan menggunakan metode representasi teks *contextual embedding*. *Contextual embedding* mampu mempertimbangkan makna kalimat secara keseluruhan seperti misalnya pada kalimat “هذا الرجل عينه مصابة” dan “هذا الرجل عينه” akan memberikan representasi teks yang berbeda karena kedua kalimat tersebut memiliki arti yang berbeda. Berbeda halnya dengan *static embedding*, yang akan memberikan representasi teks yang sama pada kedua kalimat tersebut. Kemudian, dikarenakan pendeteksian antar kitab hadis dilakukan dengan mengelompokkan beberapa hadis pada kitab hadis tertentu yang saling berkaitan satu sama lain berdasarkan kesamaan makna menjadi satu kategori yang sama. Maka, perlu dilakukan kombinasi antara metode representasi teks dengan metode klasifikasi. Jadi, setelah dilakukan representasi teks menggunakan *contextual embedding*, langkah selanjutnya adalah mengklasifikasikan dokumen teks hadis untuk dikelompokkan kedalam kategori yang telah ditentukan, agar lebih memudahkan muslim dalam mempelajari hadis sesuai dengan kategori atau topiknya. Berdasarkan jawaban dari pertanyaan penelitian ketiga, ditemukan bahwa metode klasifikasi yang cocok untuk mendeteksi interelasi antar kitab hadis adalah dengan menggunakan metode *ensemble*, yang mampu meningkatkan performa algoritma klasifikasi dokumen teks.

#### 4. Kesimpulan

Hasil penelitian menunjukkan bahwa dalam merepresentasikan teks, metode *contextual embedding* lebih baik daripada *static embedding* karena mampu mempertimbangkan konteks kata dan makna kalimat. Peningkatan kinerja algoritma klasifikasi dokumen teks dapat dilakukan dengan menggunakan metode *ensemble*

yang mampu membuat hasil prediksi model menjadi lebih baik daripada hanya menggunakan *single classifier* saja. Selain itu, sejauh ini penelitian yang menggunakan dataset hadis untuk deteksi interelasi antar kitab hadis masih belum ada. Oleh karena itu, untuk pengembangan penelitian selanjutnya mengenai deteksi interelasi antar kitab hadis, dapat dilakukan kombinasi metode representasi teks dan metode klasifikasi. Penggunaan *contextual embedding* sebagai metode representasi teks dan metode *ensemble* sebagai metode klasifikasi perlu diteliti lebih lanjut.

#### Daftar Rujukan

- [1] N. A. P. Rostam and N. H. A. H. Malim, 2021, “Text categorisation in Quran and Hadith: Overcoming the interrelation challenges using machine learning and term weighting”, *J. King Saud Univ. - Comput. Inf. Sci.*, pp. 658–667, doi: 10.1016/j.jksuci.2019.03.007.
- [2] V. Amrizal, 2018, “Penerapan Metode Term Frequency Inverse Document Frequency (Tf-Idf) Dan Cosine Similarity Pada Sistem Temu Kembali Informasi Untuk Mengetahui Syarah Hadits Berbasis Web (Studi Kasus: Hadits Shahih Bukhari-Muslim)”, *J. Tek. Inform.*, pp. 149–164, doi: 10.15408/jti.v11i2.8623.
- [3] A. T. Ni'mah and A. Z. Arifin, 2020, “Perbandingan Metode Term Weighting terhadap Hasil Klasifikasi Teks pada Dataset Terjemahan Kitab Hadis”, *Rekayasa*, pp. 172–180, 2020, doi: 10.21107/rekayasa.v13i2.6412.
- [4] M. Y. Abu Bakar, Adiwijaya, and S. Al Faraby, 2019, “Multi-Label Topic Classification of Hadith of Bukhari (Indonesian Language Translation) Using Information Gain and Backpropagation Neural Network”, *Proc. 2018 Int. Conf. Asian Lang. Process. IALP 2018*, pp. 344–350, doi: 10.1109/IALP.2018.8629263.
- [5] W. Darmalaksana, C. Slamet, W. B. Zulfikar, I. F. Fadillah, D. S. adillah Maylawati, and H. Ali, 2020, “Latent semantic analysis and cosine similarity for hadith search engine”, *Telkomnika (Telecommunication Comput. Electron. Control.*, pp. 217–227, doi: 10.12928/TELKOMNIKA.V18I1.14874.
- [6] A. Abdi, S. Hasan, M. Arshi, S. M. Shamsuddin, and N. Idris, 2020, “A question answering system in hadith using linguistic knowledge”, *Comput. Speech Lang.*, pp. 1-13, doi: 10.1016/j.csl.2019.101023.
- [7] E. H. Mohamed and W. H. El-Behaidy, 2021, “An Ensemble Multi-label Themes-Based Classification for Holy Qur'an”, *Arab. J. Sci. Eng.*, pp. 3519-3529, doi:10.1007/s13369-020-05184-0.
- [8] A. O. Adeleke, N. A. Samsudin, A. Mustapha, and N. M. Nawi, 2017, “Comparative analysis of text classification algorithms for automated labelling of Quranic verses”, *Int. J. Adv. Sci. Eng. Inf. Technol.*, pp. 1419–1427, doi: 10.18517/ijaseit.7.4.2198.
- [9] F. S. Utomo, N. Suryana, and M. S. Azmi, 2020, “Stemming impact analysis on Indonesian Quran translation and their exegesis classification for ontology instances”, *IJUM Eng. J.*, pp. 33–50, doi: 10.31436/iujmej.v21i1.1170.
- [10] P. Mondal, A. Ghosh, A. Sinha, and S. Goswami, 2020, “A Study of Interrelation Between Ratings and User Reviews in Light of Classification,” *Adv. Intell. Syst. Comput.*, pp. 689–697, doi: 10.1007/978-981-13-7403-6\_60.
- [11] J. C. Carver, E. Hassler, E. Hernandez, and N. A. Kraft, 2013, “Identifying barriers to the systematic literature review process”, *Int. Symp. Empir. Softw. Eng. Meas.*, pp. 203–213, doi: 10.1109/ESEM.2013.28.
- [12] D. C. B. Mariano, C. Leite, L. H. S. Santos, R. E. O. Rocha, and R. C. de Melo-Minardi, 2017, “A guide to performing systematic literature reviews in bioinformatics,” *arXiv*.
- [13] Y. Levy and T. J. Ellis, 2006, “A systems approach to conduct an effective literature review in support of information systems

- research”, *Informing Sci.*, pp. 181–211, doi: 10.28945/479.
- [14] S. Hamid, S. Bukhari, S. D. Ravana, A. A. Norman, and M. T. Ijab, 2016, “Role of social media in information-seeking behaviour of international students: A systematic literature review”, *Aslib J. Inf. Manag.*, pp. 643–666, doi: 10.1108/AJIM-03-2016-0031.
- [15] R. S. Wahono, 2007, “A Systematic Literature Review of Software Defect Prediction: Research Trends, Datasets, Methods and Frameworks”, *J. Softw. Eng.*, pp. 1–16, doi: 10.3923/jse.2007.1.12.
- [16] M. Sobri, M. T. Ijab, and N. M. Nayan, 2018, “Systematic Literature Review untuk Membuat Model Aplikasi Pemantauan Kesehatan Cardiovascular”, *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, pp. 458–464, doi: 10.29207/resti.v2i2.428.
- [17] A. N. Izzati and N. F. Najwa, 2018, “Pengaruh Stakeholder Perspective Dalam Penerapan ERP: A Systematic Literature Review”, *J. Teknol. Inf. dan Ilmu Komput.*, pp. 41, doi: 10.25126/jtiik.201851540.
- [18] Z. Drus and H. Khalid, 2019, “Sentiment analysis in social media and its application: Systematic literature review”, *Procedia Comput. Sci.*, pp. 707–714, doi: 10.1016/j.procs.2019.11.174.
- [19] A. Sari, A. Tosun, and G. I. Alptekin, 2019, “A systematic literature review on crowdsourcing in software engineering”, *J. Syst. Softw.*, pp. 200–219, doi: <https://doi.org/10.1016/j.jss.2019.04.027>.
- [20] B. Kitchenham, E. Mendes, and G. H. Travassos, 2006, “A Systematic Review of Cross- vs. Within-Company Cost Estimation Studies”, pp. 316–329, doi: 10.14236/ewic/ease2006.10.
- [21] M. Nabil, M. Aly, and A. F. Atiya, 2015, “ASTD: Arabic sentiment tweets dataset”, *Conf. Proc. - EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process.*, pp. 2515–2519, doi: 10.18653/v1/d15-1299.
- [22] O. ElJundi, W. Antoun, N. El Droubi, H. Hajj, W. El-Hajj, and K. Shaban, 2019, “hULMonA: The Universal Language Model in Arabic”, pp. 68–77, doi: 10.18653/v1/w19-4608.
- [23] A. Oussous, A. A. Lahcen, and S. Belfkih, 2019, “Impact of text pre-processing and ensemble learning on Arabic sentiment analysis”, *ACM Int. Conf. Proceeding Ser.*, pp. 1–9, doi: 10.1145/3320326.3320399.
- [24] G. Adel and Y. Wang, 2019, “Arabic twitter corpus for crisis response messages classification”, *ACM Int. Conf. Proceeding Ser.*, pp. 498–503, doi: 10.1145/3377713.3377799.
- [25] I. S. I. Abuhaiba and H. M. Dawoud, 2017, “Combining different approaches to improve arabic text documents classification”, *Int. J. Intell. Syst. Appl.*, pp. 39–52, doi: 10.5815/ijisa.2017.04.05.
- [26] M. Saad and W. Ashour, 2010, “OSAC: Open Source Arabic Corpora”, *6th Int. Conf. Electr. Comput. Syst. (EECS’10)*, pp. 118–123, doi: 10.13140/2.1.4664.9288.
- [27] F. zahra El-Alami, S. Ouatik El Alaoui, and N. En Nahnahi, 2021, “Contextual semantic embeddings based on fine-tuned AraBERT model for Arabic text multi-class categorization”, *J. King Saud Univ. - Comput. Inf. Sci.*, pp. 1–7, doi: 10.1016/j.jksuci.2021.02.005.
- [28] S. Bahassine, A. Madani, M. Al-Sarem, and M. Kissi, 2020, “Feature selection using an improved Chi-square for Arabic text classification”, *J. King Saud Univ. - Comput. Inf. Sci.*, pp. 225–231, doi: 10.1016/j.jksuci.2018.05.010.
- [29] B. Arkok and A. M. Zeki, 2021, “Classification of Quranic Topics Using Ensemble Learning”, pp. 244–248, doi: 10.1109/icce50029.2021.9467178.
- [30] B. S. Arkok and A. M. Zeki, 2021, “Classification of Quranic topics based on imbalanced classification”, *Indones. J. Electr. Eng. Comput. Sci.*, pp. 678, doi: 10.11591/ijeecs.v22.i2.pp678-687.
- [31] H. M. Abdelaal, A. M. Ahmed, W. Ghribi, and H. A. Youness Alansary, 2019, “Knowledge Discovery in the Hadith According to the Reliability and Memory of the Reporters Using Machine Learning Techniques”, *IEEE Access*, pp. 157741–157755, doi: 10.1109/ACCESS.2019.2944118.
- [32] M. Ghanem, A. Mouloudi, and M. Mourchid, 2016, “Classification of Hadiths using LVQ based on VSM Considering Words Order”, *Int. J. Comput. Appl.*, pp. 25–28, doi: 10.5120/ijca2016911077.
- [33] R. A. Salama, A. Youssef, and A. Fahmy, 2018, “Morphological Word Embedding for Arabic”, *Procedia Comput. Sci.*, pp. 83–93, doi: 10.1016/j.procs.2018.10.463.
- [34] A. H. Ombabi, W. Ouarda, and A. M. Alimi, 2020, “Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks”, *Soc. Netw. Anal. Min.*, pp. 1–13, doi: 10.1007/s13278-020-00668-1.
- [35] A. Alwehaibi and K. Roy, 2019, “Comparison of Pre-Trained Word Vectors for Arabic Text Classification Using Deep Learning Approach”, *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 1471–1474, doi: 10.1109/ICMLA.2018.00239.
- [36] H. Elgabry, S. Attia, A. Abdel-Rahman, A. Abdel-Ate, and S. Girgis, 2021, “A Contextual Word Embedding for Arabic Sarcasm Detection with Random Forests”, *Proc. Sixth Arab. Nat. Lang. Process. Work.*, pp. 340–344.
- [37] A. M. Bdeir and F. Ibrahim, 2020, “A framework for arabic tweets multi-label classification using word embedding and neural networks algorithms”, *ACM Int. Conf. Proceeding Ser.*, pp. 105–112, doi: 10.1145/3404512.3404526.
- [38] H. Al Saif and T. Alotaibi, 2019, “Arabic text classification using feature-reduction techniques for detecting violence on social media”, *Int. J. Adv. Comput. Sci. Appl.*, pp. 77–87, doi: 10.14569/ijacsa.2019.0100409.
- [39] Y. Albalawi, J. Buckley, and N. S. Nikolov, 2021, “Investigating the impact of pre-processing techniques and pre-trained word embeddings in detecting Arabic health information on social media”, *J. Big Data*, pp. 1–29, doi: 10.1186/s40537-021-00488-w.
- [40] M. Heikal, M. Toriki, and N. El-Makky, 2018, “Sentiment Analysis of Arabic Tweets using Deep Learning”, *Procedia Comput. Sci.*, pp. 114–122, doi: 10.1016/j.procs.2018.10.466.
- [41] H. H. Saeed, T. Calders, and F. Kamiran, 2020, “OSACT4 Shared Tasks: Ensembled Stacked Classification for Offensive and Hate Speech in Arabic Tweets”, *Proc. 4th Work. Open-Source Arab. Corpora Process. Tools, with a Shar. Task Offensive Lang. Detect.*, pp. 71–75.
- [42] M. M. Ashi, M. A. Siddiqui, and F. Nadeem, 2019, “Pre-trained Word Embeddings for Arabic Aspect-Based Sentiment Analysis of Airline Tweets”, *Adv. Intell. Syst. Comput.*, pp. 241–251, doi: 10.1007/978-3-319-99010-1\_22.
- [43] A. Mohammed and R. Kora, 2019, “Deep learning approaches for Arabic sentiment analysis”, *Soc. Netw. Anal. Min.*, pp. 1–12, doi: 10.1007/s13278-019-0596-4.
- [44] B. Haidar, M. Chamoun, and A. Serhrouchni, 2019, “Arabic Cyberbullying Detection: Enhancing Performance by Using Ensemble Machine Learning”, *2019 Int. Conf. Internet Things IEEE Green Comput. Commun. IEEE Cyber. Phys. Soc. Comput. IEEE Smart Data*, pp. 323–327, doi: 10.1109/iThings/GreenCom/CPSCoM/SmartData.2019.00074.
- [45] S. H. Kumhar, M. M. Kirmani, J. Sheetalani, and M. Hassan, 2021, “Word Embedding Generation For Urdu Language Using Word2vec Model”, *Mater. Today Proc.*, doi: 10.1016/j.matpr.2020.11.766.
- [46] S. Alsafari, S. Sadaoui, and M. Mouhoub, 2020, “Hate and offensive speech detection on Arabic social media”, *Online Soc. Networks Media*, pp. 1–15, doi: 10.1016/j.osnem.2020.100096.
- [47] N. Alghamdi and F. Assiri, 2020, “Solving the Cold-Start Problem in Recommender Systems Using Contextual Information in Arabic from Calendars”, *Arab. J. Sci. Eng.*, pp. 10939–10947, doi: 10.1007/s13369-020-04890-z.
- [48] L. Al Qadi, H. El Rifai, S. Obaid, and A. Elnagar, 2019, “Arabic text classification of news articles using contextual supervised classifiers”, *2019 2nd Int. Conf. New Trends Comput. Sci. ICTCS 2019 - Proc.*, pp. 1–6, doi: 10.1109/ICTCS.2019.8923073.
- [49] H. A. Almuzaini and A. M. Azmi, 2020, “Impact of Stemming and Word Embedding on Deep Learning-Based Arabic Text



- Categorization”, *IEEE Access*, pp. 127913–127928, doi: 10.1109/ACCESS.2020.3009217.
- [50] U. Naqvi, A. Majid, and S. Ali Abbas, 2021, “UTSA: Urdu Text Sentiment Analysis Using Deep Learning Methods”, *IEEE Access*, pp. 114085 – 114094, doi: 10.1109/access.2021.3104308.
- [51] R. S. Bennabi and Z. Elberrichi, 2019, “An Empirical Study on the effect of weighting schemes and Machine Learning algorithms on the Arabic text Classification”, pp. 5–8.
- [52] M. Alkhatib, A. A. Monem, and K. Shaalan, 2017, “A Rich Arabic WordNet Resource for Al-Hadith Al-Shareef”, *Procedia Comput. Sci.*, pp. 101–110, doi: 10.1016/j.procs.2017.10.098.
- [53] M. F. Afianto, Adiwijaya, and S. Al-Faraby, 2018, “Text Categorization on Hadith Sahih Al-Bukhari using Random Forest”, *J. Phys. Conf. Ser.*, pp. 1-6, doi: 10.1088/1742-6596/971/1/012037.
- [54] S. Al Faraby, E. R. R. Jasin, A. Kusumaningrum, and Adiwijaya, 2018, “Classification Of Hadith Into Positive Suggestion, Negative Suggestion, And Information”, *J. Phys. Conf. Ser.*, pp. 1-8, doi: 10.1088/1742-6596/971/1/012046.
- [55] H. M. Abdelaal, B. R. Elemary, and H. A. Youness, 2019, “Classification of Hadith According to Its Content Based on Supervised Learning Algorithms”, *IEEE Access*, pp. 152379–152387, doi: 10.1109/ACCESS.2019.2948159.
- [56] F. Haque, A. H. Orthy, and S. Siddique, 2020, “Hadith Authenticity Prediction using Sentiment Analysis and Machine Learning”, *14th IEEE Int. Conf. Appl. Inf. Commun. Technol. AICT 2020 - Proc.*, pp. 1-6, doi: 10.1109/AICT50176.2020.9368569.
- [57] M. Z. Naf’an, Y. Sari, and Y. Suyanto, 2021, “Word Embeddings Evaluation on Indonesian Translation of Al-Quran and Hadiths”, *IOP Conf. Ser. Mater. Sci. Eng.*, pp. 1-10, doi: 10.1088/1757-899x/1077/1/012025.
- [58] H. M. Abdelaal and H. A. Youness, 2019, “Hadith Classification using Machine Learning Techniques According to its Reliability”, *Rom. J. Inf. Sci. Technol.*, pp. 259–271.
- [59] F. A. Setiawati, Q. U. Safitri, A. F. Huda, A. Saepulloh, and W. Darmalaksana, 2019, “Feature Selection using k-Medoid Algorithm for Categorization of Hadith Translation in English”, *Proceeding 2019 5th Int. Conf. Wirel. Telemat. ICWT 2019*, pp. 1-5, doi: 10.1109/ICWT47785.2019.8978221.
- [60] B. R. Elemary, 2021, “The Effect of Clustering Classification and Pre-processing of Text on Improving the Accuracy of Hadith,” *Sci. J. Financ. Commer. Stud. Res.*, pp. 549–575, doi: 10.21608/cfdj.2020.129343.
- [61] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, 2017, “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature”, *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, pp. 1725–1732.
- [62] D. Chandrasekaran and V. Mago, 2020, “Evolution of Semantic Similarity - A Survey”, *Association for Computer Machinery*, pp. 1-37, doi: 10.1145/3440755.
- [63] D. Jurafsky and J. H. Martin, 2019, “Vector Semantics and Embeddings”, *Speech and Language Processing*, pp. 1–31.
- [64] T. Mikolov, 2013, “Learning Representations of Text using Neural Networks (Slides)”, *NIPS Deep Learn. Work.*, pp. 1–31.
- [65] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, 2017, “Enriching Word Vectors with Subword Information”, *Trans. Assoc. Comput. Linguist.*, pp. 135–146, doi: 10.1162/tacl\_a\_00051.
- [66] A. B. Soliman, K. Eissa, and S. R. El-Beltagy, 2017, “AraVec: A set of Arabic Word Embedding Models for use in Arabic NLP”, *Procedia Comput. Sci.*, pp. 256–265, doi: 10.1016/j.procs.2017.10.117.
- [67] J. Turton, D. Vinson, and R. E. Smith, 2020, “Deriving Contextualised Semantic Features from BERT (and Other Transformer Model) Embeddings,” *arXiv*.
- [68] M. Peters *et al.*, 2018, “Deep Contextualized Word Representations”, *Proceedings of the 2018 Conference of the North (A)merican Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 2227–2237, doi: 10.18653/v1/N18-1202.
- [69] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, K. T. Google, and A. I. Language, 2019, “{BERT}: Pre-training of Deep Bidirectional Transformers for Language Understanding”, *Proceedings of the 2019 Conference of the North (A)merican Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, doi: 10.18653/v1/N19-1423.
- [70] A. Elnagar, O. Einea, and R. Al-Debsi, 2019, “Automatic Text Tagging of {A}rabic News Articles Using Ensemble Deep Learning Models”, *Proc. 3rd Int. Conf. Nat. Lang. Speech Process.*, pp. 59–66, doi: 10.17632/57zpx667y9.1.
- [71] M. Fayaz, A. Khan, J. U. Rahman, A. Alharbi, M. I. Uddin, and B. Alouffi, 2020, “Ensemble machine learning model for classification of spam product reviews”, *Complexity*, pp. 1-10, doi: 10.1155/2020/8857570.
- [72] S. Al-Saqqa, N. Obeid, and A. Awajan, 2019, “Sentiment Analysis for Arabic Text using Ensemble Learning”, *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, pp. 1–7, doi: 10.1109/AICCSA.2018.8612804.
- [73] M. A. Fauzi and A. Yuniarti, 2018, “Ensemble method for Indonesian twitter hate speech detection,” *Indones. J. Electr. Eng. Comput. Sci.*, pp. 294–299, doi: 10.11591/ijeecs.v11.i1.pp294-299.
- [74] X. Ying, 2019, “An Overview of Overfitting and its Solutions”, *J. Phys. Conf. Ser.*, pp. 1-7, doi: 10.1088/1742-6596/1168/2/022022.
- [75] A. Onan, S. Korukoğlu, and H. Bulut, 2016, “Ensemble of keyword extraction methods and classifiers in text classification”, *Expert Syst. Appl.*, pp. 232–247, doi: 10.1016/j.eswa.2016.03.045.
- [76] S. Kumari, D. Kumar, and M. Mittal, 2021, “An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier”, *Int. J. Cogn. Comput. Eng.*, pp. 40–46, doi: 10.1016/j.ijcce.2021.01.001.