TEKNOSAINS
MEDIA INFORMASI DAN TEKNOLOGI
Journal homepage: https://journal.uin-alauddin.ac.id/index.php/teknosains/

Keterkaitan tematik matan hadis bertema shalat menggunakan algoritma Apriori berbasis *Association Rule Mining*

Aidil Fitra[1], Hastuti[1*], Gafur Rahim[1]
[1]Program Studi Sistem Informasi
Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar
Jl. H. M. Yasin Limpo No. 36, Gowa, Sulawesi Selatan, Indonesia. 92118
*E-mail: hastuti.baharuddin@uin-alauddin.ac.id

Abstrak: Perkembangan digitalisasi koleksi hadis membuka peluang pemanfaatan teknik *text mining* untuk mendukung analisis kandungan hadis secara lebih sistematis. Namun, karakteristik teks hadis berbahasa Arab, seperti keberadaan *sanad*, variasi morfologi, dan perbedaan penulisan, masih menjadi tantangan dalam proses analisis karena dapat menimbulkan *noise* pada data. Penelitian ini bertujuan menganalisis keterkaitan tematik antarmatan hadis bertema shalat menggunakan algoritma Apriori berbasis *Association Rule Minin*g. Penelitian menggunakan pendekatan kuantitatif dengan memanfaatkan dataset hadis Shahih al-Bukhari dan Shahih Muslim yang diperoleh dari Sunnah.com. Tahapan penelitian meliputi ekstraksi matan, normalisasi teks Arab, penghapusan *stopword*, penyaringan tema, transformasi transaksi, serta pembentukan aturan asosiasi berdasarkan nilai *support*, *confidence*, dan *lift*. Hasil penelitian menunjukkan bahwa tahapan *preprocessing* mampu meningkatkan kualitas data sehingga menghasilkan pola asosiasi yang lebih representatif terhadap kandungan matan hadis. Algoritma Apriori berhasil mengidentifikasi keterkaitan tematik antarkonsep dalam matan hadis bertema shalat berdasarkan kemunculan bersama (*co-occurrence*) konsep-konsep utama. Temuan ini menunjukkan bahwa pendekatan *Association Rule Mining* dapat dimanfaatkan sebagai alat bantu dalam kajian ilmu hadis untuk mengeksplorasi hubungan tematik antarmatan hadis secara sistematis serta mendukung analisis terhadap koleksi hadis digital.

Kata Kunci: algoritma Apriori, *Association Rule Mining*, hadis, matan hadis, *text mining*

Abstract: The digitization of hadith collections has created new opportunities for applying text mining techniques to analyze hadith content in a more systematic manner. However, the analysis of Arabic hadith texts remains challenging due to the presence of sanad (chains of transmission), morphological variations, and inconsistencies in writing, which may introduce noise and reduce analytical accuracy. This study aims to analyze the thematic relationships among the matan (text) of prayer-related hadiths using the Apriori algorithm based on Association Rule Mining. A quantitative approach was employed using hadith datasets from Sahih al-Bukhari and Sahih Muslim, obtained from the Sunnah.com repository. The research process consisted of matan extraction, Arabic text normalization, stopword removal, thematic filtering, transaction transformation, and association rule generation based on support, confidence, and lift values. The results indicate that the preprocessing stage effectively reduced noise in the dataset, producing association patterns that more accurately represent the content of the hadith texts. Furthermore, the Apriori algorithm successfully identified thematic relationships among key concepts in the matan of prayer-related hadiths through concept co-occurrence patterns. These findings demonstrate that Association Rule Mining can serve as a computational tool to support hadith studies by systematically exploring thematic relationships among hadith texts while complementing, rather than replacing, established methodologies in hadith scholarship.

Keywords: Apriori algorithm, Association Rule Mining, hadith, matan hadith, text mining

Cara Sitasi:
Fitra, A., Hastuti, H., Rahim, G. (2026). Keterkaitan tematik matan hadis bertema shalat menggunakan algoritma Apriori berbasis *Association Rule Mining*. *Teknosains: Media Informasi Sains dan Teknologi*, 20(2), 285-297. https://doi.org/10.24252/teknosains.v20i2.70426

Diajukan 13 Juli 2026; Ditinjau 1 Agustus 2026; Diterima 30 September 2026; Diterbitkan 30 Oktober 2026
Copyright © 2026. The authors. This is an open access article under the CC BY-SA license

PENDAHULUAN

Hadis merupakan sumber ajaran Islam kedua setelah Al-Qur'an yang memiliki peranan penting dalam menjelaskan berbagai aspek kehidupan umat Islam. Setiap hadis terdiri atas sanad dan matan, di mana matan memuat substansi ajaran Rasulullah SAW sehingga menjadi objek penting dalam memahami tema dan kandungan hadis. Pendekatan tematik memungkinkan hubungan antarmatan hadis dipahami secara lebih komprehensif melalui analisis terhadap hadis-hadis yang memiliki kesamaan tema (Kamala et al., 2025; Mosa, 2023).

Kajian terhadap matan hadis juga terus berkembang seiring munculnya berbagai pendekatan interdisipliner. Penelitian mengenai kritik matan menunjukkan bahwa analisis terhadap matan tidak hanya dilakukan dalam perspektif ilmu hadis, tetapi juga dapat diintegrasikan dengan berbagai disiplin ilmu, termasuk sains dan teknologi, sehingga membuka peluang penerapan pendekatan komputasional dalam memahami kandungan hadis (Firdaus & Suryadilaga, 2020).

Allah Swt. mendorong umat Islam untuk memahami dan mengambil pelajaran dari setiap perkataan yang didengar. Hal tersebut sebagaimana dijelaskan dalam QS. Az-Zumar [39]:18:

ٱلَّذِينَ يَسْتَمِعُونَ ٱلْقَوْلَ فَيَتَّبِعُونَ أَحْسَنَهُۥٓ ۚ أُو۟لَٰٓئِكَ ٱلَّذِينَ هَدَىٰهُمُ ٱللَّهُ ۖ وَأُو۟لَٰٓئِكَ هُمْ أُو۟لُوا۟ ٱلْأَلْبَٰبِ

Artinya: "(Yaitu) orang-orang yang mendengarkan perkataan lalu mengikuti apa yang paling baik di antaranya. Mereka itulah orang-orang yang telah diberi petunjuk oleh Allah dan mereka itulah orang-orang yang mempunyai akal." (QS. Az-Zumar [39]:18) (Lajnah Pentashihan mushaf Al-Qur'an, 2022).

Ayat tersebut menunjukkan pentingnya memahami kandungan suatu perkataan secara cermat untuk memperoleh pengetahuan yang benar. Dalam konteks penelitian ini, prinsip tersebut menjadi landasan dalam mengkaji kandungan matan hadis melalui pendekatan komputasional sehingga hubungan antarkonsep yang terkandung di dalamnya dapat diidentifikasi secara lebih sistematis.

Perkembangan teknologi turut mendorong pemanfaatan ontologi dan berbagai pendekatan komputasional dalam pengelolaan teks hadis. Selain berfungsi sebagai media penyimpanan dan pencarian informasi, representasi pengetahuan berbasis ontologi juga dimanfaatkan untuk mendukung analisis matan hadis, sistem temu kembali informasi, serta pengembangan aplikasi hadis digital(Muhammed et al., 2024). Seiring berkembangnya digitalisasi koleksi hadis, ribuan matan hadis kini tersedia dalam bentuk data digital yang dapat dianalisis menggunakan berbagai pendekatan komputasional. Kondisi tersebut membuka peluang penerapan teknik *text mining* dan *Natural Language Processing* (NLP) untuk mengeksplorasi kandungan matan hadis secara lebih sistematis (Khabir & Ridlo, 2025). Berbagai penelitian telah memanfaatkan *text mining* dan *Natural Language Processing* (NLP) untuk klasifikasi, representasi teks, serta sistem pencarian hadis digital (Grootendorst, 2022; Mosa, 2023; Nadoos & Minaei-Bidgoli, 2025).

Pengembangan metode komputasional tersebut juga diikuti oleh pemanfaatan berbagai teknik *Natural Language Processing* (NLP) dan *machine learning* dalam kajian hadis, baik untuk klasifikasi, representasi teks, analisis sanad, analisis matan, maupun sistem temu kembali informasi. Metode seperti TF-IDF, BERT, *cosine similarity*, *Latent Semantic Analysis*, *Support Vector Machine* (SVM), dan *Convolutional Neural Network* (CNN) terbukti mampu meningkatkan kualitas klasifikasi maupun pencarian hadis digital

(Asy'ari & Hanafi, 2025; Athallah & Lhaksmana, 2025; Darmalaksana et al., 2020; Darmawan et al., 2025; M Yunus, 2020; Surianto et al., 2024). Meskipun demikian, pemanfaatan metode-metode tersebut masih didominasi oleh tugas klasifikasi, representasi teks, dan sistem temu kembali informasi.

Meskipun perkembangan digitalisasi hadis serta penerapan Artificial Intelligence, *machine learning*, dan Natural Language Processing telah meningkatkan kemampuan klasifikasi, representasi teks, serta sistem temu kembali informasi hadis digital, penelitian mengenai keterkaitan tematik antarmatan hadis masih relatif terbatas(Ariyanto et al., 2021; Azwar et al., 2025; Khabir & Ridlo, 2025). Selain itu, algoritma Apriori telah diterapkan untuk menemukan pola hubungan antarrawi pada sanad hadis menggunakan pendekatan *Association Rule Mining*. Pendekatan *Sequential Pattern Mining* menggunakan algoritma SPADE juga telah dimanfaatkan untuk mengidentifikasi pola hubungan berdasarkan urutan sanad hadis. Akan tetapi, kedua penelitian tersebut masih berfokus pada struktur sanad sehingga belum mengeksplorasi keterkaitan tematik antarkonsep dalam matan hadis sebagai objek utama analisis (Septiani et al., 2015; Yotenka, 2017). Kondisi ini menunjukkan bahwa penerapan teknik data mining pada kajian hadis masih didominasi oleh analisis sanad dibandingkan eksplorasi kandungan matan hadis.

Di sisi lain, analisis terhadap teks hadis juga menghadapi tantangan karakteristik bahasa Arab yang memiliki morfologi kompleks, variasi penulisan, dan ambiguitas linguistik sehingga memerlukan tahapan *preprocessing* yang memadai sebelum dilakukan proses penambangan data Oleh karena itu, diperlukan tahapan *preprocessing* yang mampu meminimalkan *noise* pada teks hadis sebelum dilakukan proses penambangan data. Penelitian ini menerapkan ekstraksi matan, normalisasi teks Arab, penghapusan *stopword*, serta penyaringan tema untuk menghasilkan pola asosiasi yang lebih representatif terhadap kandungan matan hadis (Mosa, 2023; Qu et al., 2024; Uriawan et al., 2025).

Penelitian ini menggunakan dataset hadis dari Sahih Bukhari dan Sahih Muslim sebagai objek kajian, mengingat kedua kitab tersebut merupakan sumber hadis yang memiliki tingkat keotentikan tinggi. Melalui tahapan preprocessing teks Arab, ekstraksi kata kunci, dan penerapan algoritma Apriori, penelitian ini bertujuan mengidentifikasi pola keterkaitan tematik antarkonsep dalam matan hadis melalui analisis kemunculan bersama (*co-occurrence*) menggunakan algoritma Apriori.

Berdasarkan kondisi tersebut, penelitian ini bertujuan menganalisis keterkaitan tematik pada matan hadis bertema shalat menggunakan algoritma Apriori berbasis *Association Rule Mining*. Penelitian memanfaatkan tahapan ekstraksi matan, normalisasi teks Arab, *filtering* tema, dan transformasi transaksi untuk mengidentifikasi hubungan antarkonsep dalam matan hadis secara sistematis. Berbeda dengan penelitian sebelumnya yang lebih berfokus pada digitalisasi hadis, klasifikasi, representasi teks dan sistem pencarian (Apriandi, 2022; Ariyanto et al., 2021; Khabir & Ridlo, 2025), penelitian ini menawarkan pendekatan eksploratif untuk mengungkap pola keterkaitan tematik antarmatan hadis melalui analisis kemunculan bersama (*co-occurrence*) konsep-konsep utama menggunakan algoritma Apriori.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksploratif berbasis data. Pendekatan ini bertujuan untuk mengidentifikasi pola keterkaitan antar hadis melalui analisis teks secara komputasional. Selain itu, pendekatan deskriptif

digunakan untuk menjelaskan hasil analisis dalam konteks ilmu hadis, khusus nya dalam menemukan hubungan tematik antarmatan hadis.

1. Sumber data

Sumber data dalam penelitian ini berupa data sekunder yang diperoleh dari koleksi digital hadis pada platform Sunnah.com dalam format *JavaScript Object Notatio*n (JSON). Platform tersebut dipilih karena menyediakan koleksi hadis yang terstruktur, terdokumentasi dengan baik, serta banyak digunakan sebagai sumber data dalam penelitian komputasi hadis. Penelitian ini memanfaatkan dua kitab hadis primer, yaitu Shahih al-Bukhari dan Shahih Muslim, karena keduanya merupakan kitab hadis yang memperoleh pengakuan luas sebagai sumber hadis paling otoritatif setelah Al-Qur'an. Pemilihan kedua kitab ini bertujuan untuk meminimalkan penggunaan hadis yang diperselisihkan kualitasnya sehingga analisis lebih berfokus pada keterkaitan tematik antar matan hadis.

Dataset awal dalam penelitian ini terdiri atas 1.000 hadis yang diperoleh dari Shahih al-Bukhari dan Shahih Muslim. Setelah dilakukan proses *preprocessing* yang meliputi ekstraksi matan, normalisasi teks Arab, penghapusan *stopword*, *filtering* tema shalat, dan pencocokan akar kata (*root matching*) menggunakan daftar kata kunci, diperoleh 366 transaksi yang digunakan sebagai masukan pada algoritma Apriori. Penerapan algoritma dengan nilai minimum *support* sebesar 0,01 dan minimum *confidence* sebesar 0,30 menghasilkan sejumlah aturan asosiasi, kemudian dipilih aturan dengan nilai tertinggi sebagai dasar analisis.

Data yang digunakan berupa teks matan hadis berbahasa Arab, sedangkan informasi sanad, nomor hadis, dan metadata lainnya digunakan sebagai data pendukung pada tahap identifikasi. Penggunaan teks Arab dilakukan untuk mempertahankan keaslian redaksi hadis serta menghindari perubahan makna yang dapat muncul akibat proses penerjemahan. Dengan demikian, pola keterkaitan yang ditemukan berasal langsung dari teks asli hadis.

Tahap awal penelitian diawali dengan mengumpulkan seluruh hadis dari Shahih al-Bukhari dan Shahih Muslim yang tersedia pada dataset. Selanjutnya dilakukan proses ekstraksi matan untuk memisahkan isi hadis dari sanad, kemudian diterapkan *filtering* tematik menggunakan kata kunci yang berkaitan dengan ibadah shalat sehingga hanya hadis yang relevan dengan fokus penelitian yang dipertahankan. Setelah melalui proses seleksi, diperoleh sebanyak 366 hadis yang digunakan sebagai data penelitian. Setiap hadis selanjutnya direpresentasikan sebagai satu transaksi pada proses *Association Rule Mining*.

Pemilihan tema shalat didasarkan pada beberapa pertimbangan. Pertama, shalat merupakan salah satu tema dengan jumlah riwayat yang cukup banyak dalam literatur hadis sehingga menyediakan data yang memadai untuk analisis pola. Kedua, tema ini memiliki keterkaitan dengan berbagai konsep lain seperti imam, jamaah, masjid, kiblat, rakaat, wudu, dan adzan, sehingga memungkinkan terbentuknya aturan asosiasi yang lebih beragam. Ketiga, pembatasan tema dilakukan untuk mengurangi *noise* akibat keberagaman topik hadis sehingga pola keterkaitan yang dihasilkan lebih spesifik dan mudah diinterpretasikan dalam konteks ilmu hadis.

2. Teknik pengumpulan data

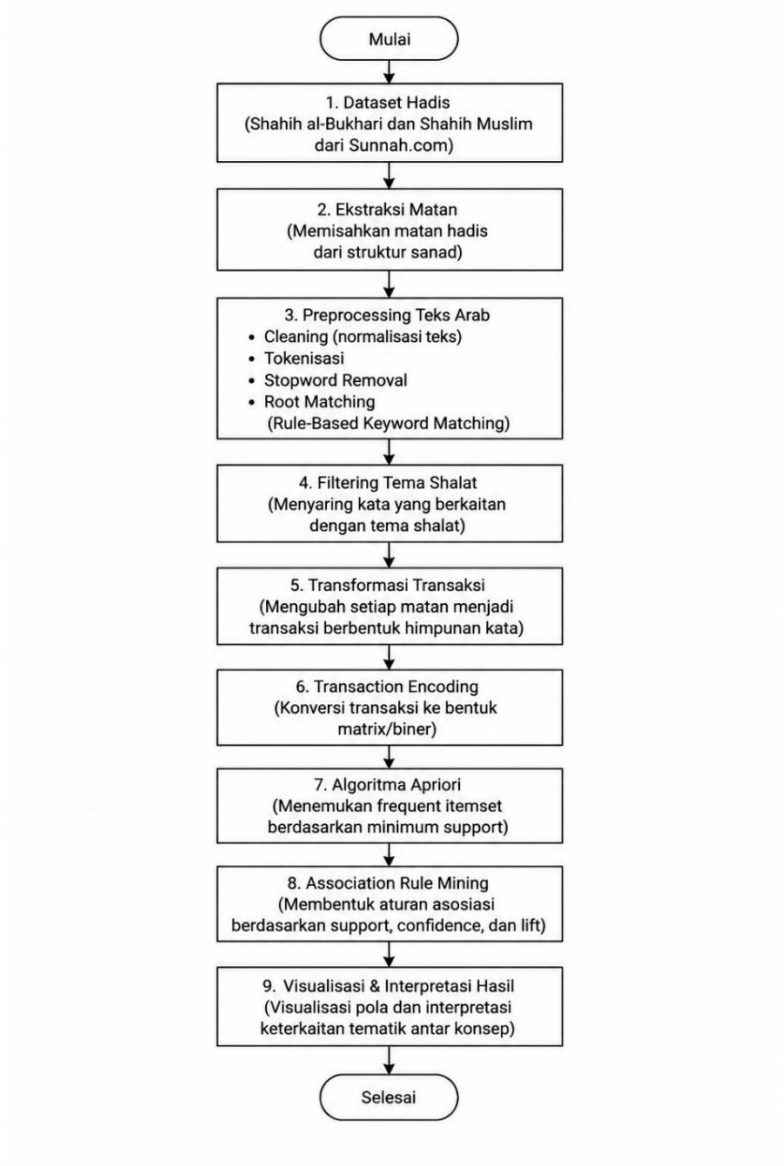
Pengumpulan data dilakukan menggunakan metode dokumentasi dengan memanfaatkan dataset hadis digital yang tersedia pada platform Sunnah.com dalam format JSON. Seluruh data yang diperoleh kemudian dikumpulkan dan diorganisasi sesuai dengan kebutuhan penelitian sebelum memasuki tahap pengolahan. Selanjutnya

dilakukan proses seleksi data dengan memilih hadis yang berasal dari kitab Shahih al-Bukhari dan Shahih Muslim. Data yang telah terseleksi kemudian melalui tahap ekstraksi matan untuk memisahkan isi hadis dari sanad, sehingga analisis difokuskan pada kandungan matan sebagai objek utama penelitian. Setelah itu dilakukan *filtering* berdasarkan tema shalat agar hanya hadis yang relevan dengan ruang lingkup penelitian yang dipertahankan.

Sebelum dianalisis menggunakan algoritma Apriori, seluruh data terlebih dahulu diverifikasi untuk memastikan tidak terdapat data yang rusak (*corrupt*), duplikasi, maupun teks yang tidak lengkap. Dataset yang telah memenuhi kriteria kemudian dikonversi ke dalam format yang dapat diproses menggunakan bahasa pemrograman Python dan digunakan sebagai masukan pada tahap *preprocessing* serta proses *Association Rule Mining*.

3. Teknik analisis data

Analisis data dalam penelitian ini menggunakan metode *Association Rule Mining* dengan algoritma Apriori. Analisis dilakukan melalui beberapa tahapan sebagaimana ditunjukkan pada Gambar 1.



Gambar1. *Flowchart* alur penelitian

Gambar 1 memperlihatkan tahapan penelitian yang terdiri atas sembilan proses utama, yaitu: (1) pengumpulan dataset hadis dari *Shahih al-Bukhari* dan *Shahih Muslim*; (2) ekstraksi matan untuk memisahkan isi hadis dari sanad; (3) *preprocessing* teks Arab yang meliputi normalisasi teks, tokenisasi, penghapusan *stopword*, dan *root matching* (*rule-based keyword matching*) untuk mengidentifikasi kata-kata yang berkaitan dengan tema shalat; (4) penyaringan hadis berdasarkan tema shalat; (5) transformasi matan hadis menjadi transaksi berbentuk himpunan kata; (6) *transaction encoding* untuk mengubah transaksi ke dalam bentuk matriks biner; (7) pembentukan *frequent itemset* menggunakan algoritma Apriori berdasarkan nilai *minimum support*; (8) pembangkitan *association rules* berdasarkan parameter *support*, *confidence*, dan *lift*; serta (9) interpretasi dan visualisasi hasil untuk mengidentifikasi keterkaitan tematik antarkonsep dalam matan hadis.

a. Preprocessing teks

Tahap *preprocessing* dilakukan sebagai langkah awal untuk mempersiapkan data sebelum dianalisis menggunakan algoritma Apriori. Proses ini bertujuan meningkatkan kualitas data dengan mengurangi *noise* serta menyeragamkan representasi teks sehingga pola yang dihasilkan lebih akurat. Tahapan diawali dengan ekstraksi matan hadis, yaitu memisahkan isi hadis dari rangkaian sanad agar analisis difokuskan pada kandungan matan tanpa dipengaruhi oleh nama-nama perawi.

Setelah proses ekstraksi, dilakukan normalisasi teks Arab untuk menyamakan variasi penulisan beberapa karakter sehingga setiap kata memiliki bentuk yang konsisten. Selanjutnya, tanda baca, simbol, angka, dan karakter lain yang tidak berkontribusi terhadap makna dihilangkan agar tidak memengaruhi proses analisis. Teks yang telah dinormalisasi kemudian diproses melalui tokenisasi, yaitu memecah setiap matan hadis menjadi kumpulan kata yang akan digunakan sebagai unit analisis.

Tahap berikutnya adalah penghapusan stopword, yaitu menghilangkan kata-kata umum yang memiliki frekuensi kemunculan tinggi tetapi tidak memberikan informasi penting terhadap pembentukan pola asosiasi. Setelah itu diterapkan filtering tema untuk menyeleksi hadis yang berkaitan dengan tema shalat sehingga analisis difokuskan pada ruang lingkup penelitian dan dapat mengurangi pengaruh topik lain yang tidak relevan.

Sebagai tahap akhir, diterapkan pencocokan akar kata (*root matching*) terhadap token yang berkaitan dengan tema shalat. Proses ini dilakukan dengan mencocokkan setiap token terhadap sejumlah akar kata yang telah ditentukan, seperti سجد, ركع, صل, اذن, وضوء, قبل, مسجد, جمع, امام, dan طهر. Pendekatan tersebut bertujuan memilih token yang mengandung akar kata tertentu. Melalui rangkaian *preprocessing* tersebut, data menjadi lebih terstruktur dan siap dianalisis menggunakan metode *Association Rule Mining*.

b. Ekstraksi kata kunci

Setiap hadis direpresentasikan dalam bentuk kata kunci berdasarkan hasil preprocessing pada matan hadis. Kata-kata yang digunakan difokuskan pada tema shalat, seperti shalat, imam, jamaah, rakaat, kiblat, dan masjid untuk menghasilkan pola keterkaitan yang lebih relevan secara tematik.

c. Transformasi data

Data diubah ke dalam bentuk transaksi, di mana setiap hadis direpresentasikan sebagai kumpulan kata kunci hasil *preprocessing* dan *filtering* tema. Setiap transaksi kemudian digunakan sebagai input dalam proses *Association Rule Mining* menggunakan algoritma Apriori.

d. Penerapan algoritma Apriori

Algoritma Apriori digunakan untuk menemukan *frequent itemset* dan membentuk *association rules* berdasarkan nilai *support*, *confidence*, dan *lift* sebagaimana telah banyak digunakan dalam penelitian data mining (Adila & Fadlisyah, 2025). Penggunaan algoritma Apriori dalam penelitian ini didasarkan pada kemampuannya dalam menemukan pola keterkaitan antar item dalam dataset secara efisien, sebagaimana telah digunakan dalam berbagai penelitian sebelumnya. Dalam penelitian ini, *association rule* digunakan untuk menemukan keterkaitan antar konsep dalam matan hadis bertema shalat berdasarkan nilai *support*, *confidence*, dan *lift*.

Algoritma Apriori digunakan untuk menemukan *frequent itemset* yang menjadi dasar dalam pembentukan aturan asosiasi (*association rules*) berdasarkan pola kemunculan bersama antarkonsep pada matan hadis. Proses ini memanfaatkan tiga parameter utama, yaitu *support*, *confidence*, dan *lift*. Nilai *support* menggambarkan frekuensi kemunculan suatu kombinasi item dalam keseluruhan transaksi, sedangkan confidence menunjukkan tingkat keyakinan bahwa kemunculan suatu konsep diikuti oleh konsep lainnya. Sementara itu, *lift* digunakan untuk mengukur kekuatan hubungan antar konsep, di mana nilai *lift* yang lebih besar dari satu mengindikasikan bahwa hubungan tersebut lebih kuat dibandingkan kemunculan yang terjadi secara acak.

Pada penelitian ini, algoritma Apriori diterapkan menggunakan *minimum support* sebesar 0,01 dan *minimum confidence* (*minimum threshold*) sebesar 0,30. Nilai minimum support sebesar 0,01 dipilih dengan mempertimbangkan karakteristik dataset yang terdiri atas 366 transaksi hasil penyaringan hadis bertema shalat. Nilai tersebut memungkinkan itemset dengan frekuensi kemunculan yang relatif rendah, namun masih memiliki keterkaitan tematik, tetap dapat teridentifikasi. Sebaliknya, penggunaan nilai minimum support yang terlalu tinggi berpotensi menghilangkan pola asosiasi yang relevan karena tidak memenuhi ambang frekuensi kemunculan. Sementara itu, *minimum confidence* sebesar 0,30 digunakan untuk menyaring aturan asosiasi yang memiliki tingkat kepercayaan minimal 30% sehingga aturan yang dihasilkan masih memiliki hubungan yang cukup representatif. Selanjutnya, hanya aturan asosiasi yang memenuhi kedua parameter tersebut serta memiliki nilai *lift* lebih besar dari satu yang digunakan sebagai dasar interpretasi dalam pembahasan hasil penelitian.

4. Rumus analisis *association Rules*

Untuk mengukur kekuatan hubungan antar item, digunakan tiga parameter utama, yaitu *support*, *confidence*, dan *lift*.

a. *Support*

$$Support(A \rightarrow B) = \frac{\text{Jumlah transaksi yang mengandung A}\cup\text{B}}{\text{Total transaksi}}$$

Support menunjukkan seberapa sering kombinasi item muncul dalam dataset.

b. *Confidence*

$$Confidence(A \rightarrow B) = \frac{Support(A \cup B)}{Support(A)}$$

Confidence menunjukkan tingkat kepercayaan bahwa kemunculan item A diikuti oleh item B.

c. *Lift*

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)}$$

Lift digunakan untuk mengukur kekuatan hubungan antar item. Nilai *lift* lebih dari 1 menunjukkan adanya hubungan yang kuat antar item.

5. Interpretasi hasil

Hasil analisis berupa aturan asosiasi kemudian diinterpretasikan untuk mengidentifikasi pola keterkaitan tematik antarmatan hadis. Interpretasi dilakukan dengan menganalisis hubungan antarkonsep berdasarkan kemunculan kata-kata yang muncul secara bersamaan pada matan hadis bertema shalat. Hasil interpretasi tersebut digunakan untuk menjelaskan keterkaitan tema dan konsep yang terkandung dalam matan hadis secara sistematis.

HASIL DAN PEMBAHASAN

1. Hasil Analisis *Association Rule*

Penelitian ini menerapkan algoritma Apriori pada dataset hadis dari Sahih al-Bukhari dan Sahih Muslim untuk mengidentifikasi pola keterkaitan antar kata dalam matan hadis. Sebelum proses *association rule* dilakukan, data terlebih dahulu melalui tahap preprocessing berupa ekstraksi matan, normalisasi teks Arab, penghapusan *stopword*, *filtering* tema, dan pencocokan akar kata (*root matching*) menggunakan daftar kata kunci. Tahapan *preprocessing* dilakukan untuk mengurangi *noise* akibat dominasi sanad dan nama perawi sehingga pola keterkaitan yang dihasilkan lebih merepresentasikan substansi hadis.

Berdasarkan hasil pengolahan data, diperoleh sejumlah aturan asosiasi yang menunjukkan hubungan antar kata dengan nilai support, confidence, dan lift yang bervariasi. Nilai lift digunakan sebagai indikator utama untuk mengukur kekuatan hubungan antar kata, di mana nilai lift lebih dari 1 menunjukkan adanya keterkaitan yang tidak bersifat acak.

Tabel 1. Top *association rules* berdasarkan nilai *lift*

Antecedents	*Consequents*	*Support*	*Confidence*	*Lift*
صلاة ,مسجد	ركعة	0.013	0.333	1.219
امام	صلاة	0.013	1.000	1.020
قبلة	صلاة	0.013	1.000	1.020
جماعة	صلاة	0.026	1.000	1.020
ركعة ,مسجد	صلاة	0.013	1.000	1.020

Berdasarkan Tabel 1, diperoleh beberapa aturan asosiasi yang menunjukkan adanya keterkaitan antarkonsep dalam matan hadis bertema shalat. Hubungan yang terbentuk melibatkan konsep-konsep seperti shalat (صلاة), imam (إمام), jamaah (جماعة), kiblat (قبلة), masjid (مسجد), dan rakaat (ركعة). Nilai *lift* yang lebih besar dari satu pada setiap aturan mengindikasikan bahwa kemunculan konsep-konsep tersebut tidak terjadi secara acak, melainkan memiliki kecenderungan untuk muncul secara bersamaan dalam sejumlah riwayat hadis.

a. Pembahasan hasil *association rule*

Hasil *Association Rule Mining* menggunakan algoritma Apriori menghasilkan lima aturan asosiasi yang menunjukkan keterkaitan antarkonsep pada matan hadis bertema shalat. Seluruh aturan memiliki nilai *lift* lebih besar dari 1, yang menunjukkan bahwa hubungan antaritem tidak terjadi secara acak, melainkan memiliki kecenderungan untuk muncul secara bersamaan dalam transaksi yang dianalisis. Hal ini mengindikasikan bahwa algoritma Apriori mampu mengidentifikasi pola keterkaitan tematik yang terdapat dalam matan hadis.

Aturan pertama menunjukkan bahwa kombinasi konsep صلاة (shalat) dan مسجد (masjid) memiliki hubungan dengan konsep ركعة (rakaat) dengan nilai *support* sebesar

0,013, *confidence* sebesar 0,333, dan *lift* sebesar 1,219. Meskipun nilai *confidence* relatif lebih rendah dibandingkan aturan lainnya, aturan ini memiliki nilai *lift* tertinggi, yang menunjukkan bahwa kemunculan konsep rakaat menjadi lebih mungkin ketika konsep shalat dan masjid muncul secara bersamaan. Secara substantif, hubungan ini dapat dipahami karena pembahasan mengenai pelaksanaan shalat di masjid sering kali disertai dengan penjelasan mengenai tata cara dan jumlah rakaat dalam matan hadis. Temuan ini menunjukkan bahwa algoritma Apriori mampu mengidentifikasi hubungan tematik yang sesuai dengan pembahasan ibadah shalat dalam literatur hadis (Firdaus & Suryadilaga, 2020; Kamala et al., 2025).

Temuan tersebut sejalan dengan hadis Rasulullah SAW yang diriwayatkan dalam Sahih al-Bukhari, Book of Adhan, Hadith No. 465, bahwa:

صَلَاةُ الرَّجُلِ فِي الْمَسْجِدِ تُضَعَّفُ عَلَى صَلَاتِهِ فِي بَيْتِهِ وَفِي سُوقِهِ خَمْسًا وَعِشْرِينَ ضِعْفًا

Artinya:
"Shalat seseorang di masjid dilipatgandakan dibandingkan shalatnya di rumah atau di pasar sebanyak dua puluh lima kali lipat."(Sunnah.com, 2025).

Hadis tersebut menunjukkan bahwa masjid merupakan tempat utama pelaksanaan shalat, sehingga pembahasan mengenai masjid dalam kitab hadis umumnya tidak terlepas dari pembahasan mengenai pelaksanaan shalat beserta tata caranya. Oleh karena itu, keterkaitan antara konsep shalat, masjid, dan rakaat yang ditemukan melalui algoritma Apriori tidak hanya merepresentasikan hubungan statistik, tetapi juga mencerminkan hubungan tematik yang memang terdapat dalam kandungan matan hadis.

Aturan kedua menunjukkan hubungan امام (imam) → صلاة (shalat) dengan nilai *confidence* sebesar 1,000 dan *lift* sebesar 1,020. Nilai tersebut menunjukkan bahwa seluruh transaksi yang memuat konsep imam juga memuat konsep shalat. Temuan ini sejalan dengan karakter penyusunan *Shahih al-Bukhari* dan *Shahih Muslim*, di mana pembahasan mengenai imam hampir selalu ditempatkan dalam konteks pelaksanaan shalat berjamaah. Oleh karena itu, pola asosiasi yang ditemukan tidak hanya menggambarkan hubungan statistik, tetapi juga mencerminkan keterkaitan tematik yang memang terdapat dalam kandungan matan hadis (Kamala et al., 2025).

Temuan tersebut juga didukung oleh hadis Rasulullah SAW:

إِنَّمَا جُعِلَ الإِمَامُ لِيُؤْتَمَّ بِهِ

Artinya:
"Sesungguhnya imam itu dijadikan untuk diikuti." (Sunnah.com, 2025)

Hadis tersebut diriwayatkan dalam Sahih al-Bukhari Book 10 (Book of Adhan), Chapter 51: *The Imam is Appointed to be Followed*, Hadith No. 722, serta Sahih Muslim Book 4 (The Book of Prayer), Chapter: *Following the Imam*, Hadith No. 414. Kandungan hadis tersebut menunjukkan bahwa keberadaan imam selalu berkaitan dengan pelaksanaan shalat berjamaah. Oleh karena itu, hubungan antara konsep imam dan shalat yang diidentifikasi melalui algoritma Apriori tidak hanya memiliki makna statistik, tetapi juga sejalan dengan kandungan matan hadis yang menjadi dasar pelaksanaan shalat berjamaah.

Aturan keempat menunjukkan hubungan جماعة (jamaah) → صلاة (shalat) dengan nilai *confidence* sebesar 1,000 dan *lift* sebesar 1,020. Hubungan ini menunjukkan bahwa konsep jamaah dalam dataset selalu muncul bersamaan dengan pembahasan mengenai shalat. Temuan tersebut konsisten dengan isi hadis dalam *Shahih al-Bukhari* dan *Shahih*

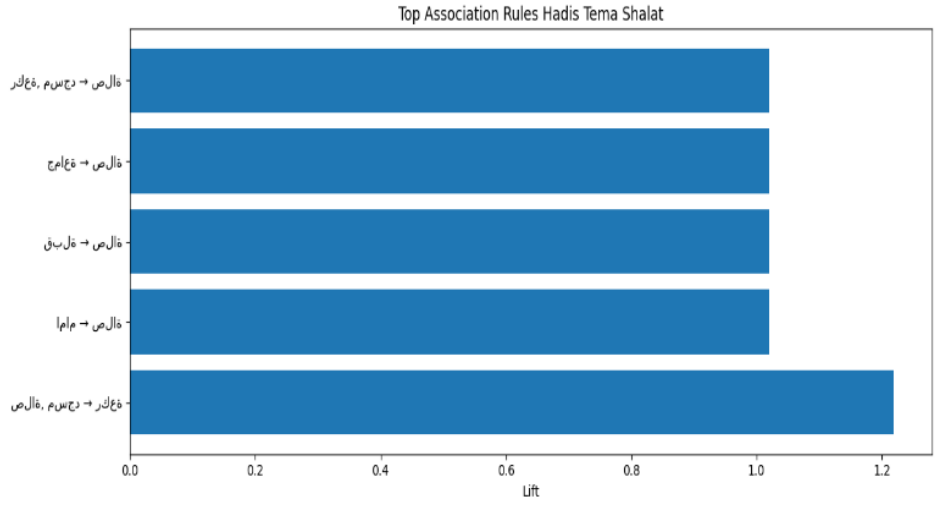
Muslim yang banyak membahas keutamaan shalat berjamaah, tata cara pelaksanaannya, serta hubungan antara imam dan makmum. Dengan demikian, algoritma Apriori berhasil mengidentifikasi pola hubungan yang sesuai dengan tema pembahasan dalam literatur hadis.

Aturan terakhir menunjukkan bahwa kombinasi ركعة (rakaat) dan مسجد (masjid) selalu diikuti oleh konsep صلاة (shalat) dengan nilai *confidence* sebesar 1,000 dan *lift* sebesar 1,020. Hasil tersebut menunjukkan bahwa pembahasan mengenai rakaat di masjid tidak dapat dipisahkan dari konteks pelaksanaan shalat sebagai tema utama. Walaupun nilai *support* relatif kecil, aturan ini tetap memiliki makna tematik karena menggambarkan keterkaitan antara tata cara pelaksanaan shalat dengan tempat pelaksanaannya.

Secara keseluruhan, aturan asosiasi yang dihasilkan menunjukkan bahwa konsep-konsep seperti imam, kiblat, jamaah, masjid, dan rakaat memiliki hubungan yang erat dalam pembahasan hadis mengenai shalat. Pola yang dihasilkan tidak hanya signifikan secara statistik berdasarkan nilai *support*, *confidence*, dan *lift*, tetapi juga konsisten dengan struktur tematik pembahasan dalam *Shahih al-Bukhari* dan *Shahih Muslim*. Temuan ini menunjukkan bahwa *Association Rule Mining* menggunakan algoritma Apriori dapat dimanfaatkan sebagai alat bantu dalam kajian ilmu hadis untuk mengidentifikasi keterkaitan tematik antarmatan hadis secara sistematis. Berbeda dengan penelitian sebelumnya yang menerapkan algoritma Apriori dan *Sequential Pattern Mining* pada analisis sanad atau hubungan antarrawi (Septiani et al., 2015; Yotenka, 2017), penelitian ini menunjukkan bahwa pendekatan yang sama juga dapat diterapkan pada matan hadis untuk mengungkap hubungan antarkonsep yang mendukung eksplorasi tema dalam koleksi hadis digital.

2. Visualisasi Kekuatan Aturan Asosiasi

Untuk memperjelas hasil analisis, dilakukan visualisasi nilai lift dari aturan asosiasi dengan nilai tertinggi.



Gambar 2. Visualisasi nilai *lift* pada top *association rules* hadis tema shalat

.

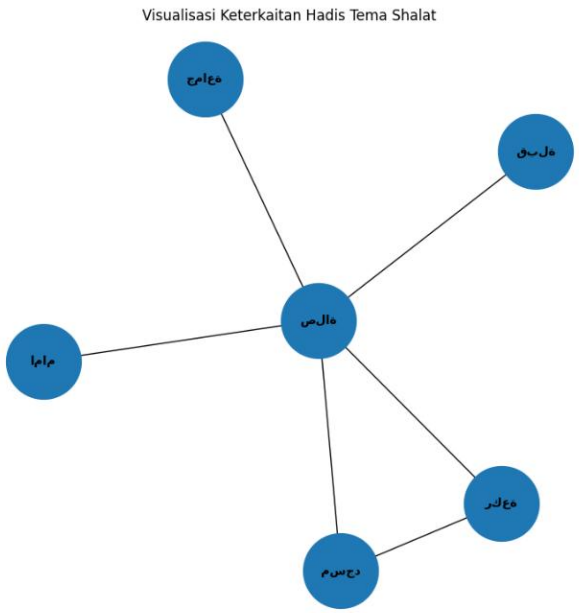
Berdasarkan Gambar 2, terlihat bahwa beberapa aturan asosiasi memiliki nilai *lift* lebih dari 1. Hal tersebut menunjukkan bahwa beberapa konsep dalam hadis bertema shalat memiliki kecenderungan untuk muncul secara bersamaan dalam matan hadis.

Aturan asosiasi dengan nilai lift tertinggi ditemukan pada hubungan antara konsep "صلاة" dan "مسجد" terhadap "ركعة". Selain itu, konsep seperti imam, jamaah, dan kiblat juga menunjukkan keterkaitan dengan konsep shalat.Visualisasi ini membantu memperjelas kekuatan hubungan antar konsep yang ditemukan melalui association rule mining. Hasil tersebut menunjukkan bahwa *preprocessing*, *filtering* tema, dan pencocokan akar kata (*root matching*) menggunakan daftar kata kunci mampu menghasilkan pola keterkaitan yang lebih substantif dibandingkan hasil yang masih didominasi oleh sanad dan nama perawi.

Konsep seperti imam, jamaah, kiblat, dan rakaat terlihat memiliki hubungan dengan konsep shalat dalam berbagai hadis. Hal ini menunjukkan bahwa *association rule* mampu menangkap keterkaitan tematik dalam matan hadis secara sistematis. Visualisasi ini membantu dalam mengidentifikasi aturan yang paling signifikan dan memberikan gambaran yang lebih intuitif dibandingkan representasi tabel.

3. Visualisasi Keterkaitan Kata

Selain visualisasi nilai lift, penelitian ini juga menggunakan *network graph* untuk menggambarkan hubungan antar kata dalam hadis bertema shalat



Gambar 3. *Network graph* keterkaitan konsep pada hadis bertema shalat

Berdasarkan Gambar 3, terlihat bahwa konsep "صلاة" menjadi pusat keterkaitan dengan beberapa konsep lain seperti "مسجد", "قبلة", "جماعة", "امام", dan "ركعة". Hal ini menunjukkan bahwa konsep shalat memiliki hubungan dengan aspek kepemimpinan ibadah, pelaksanaan shalat berjamaah, arah kiblat, dan tata cara pelaksanaan shalat. Visualisasi *network graph* juga menunjukkan bahwa *preprocessing*, *filtering* tema, dan pencocokan akar kata (*root matching*) menggunakan daftar kata kunci berhasil mengurangi dominasi sanad dan nama perawi yang sebelumnya mendominasi hasil association rule. Dengan demikian, pola keterkaitan yang dihasilkan menjadi lebih substantif dan relevan dengan tema hadis yang dianalisis.

KESIMPULAN

Penelitian ini menunjukkan bahwa tahapan *preprocessing* teks Arab, meliputi ekstraksi matan, normalisasi, penghapusan *stopword*, *filtering* tema, dan *root matching*, berhasil meningkatkan kualitas data serta menghasilkan pola asosiasi yang lebih representatif. Algoritma Apriori mampu mengidentifikasi keterkaitan tematik antarkonsep dalam matan hadis bertema shalat, khususnya antara konsep shalat, imam, jamaah, kiblat, masjid, dan rakaat. Nilai *lift* serta visualisasi jaringan menunjukkan bahwa konsep صلاة menjadi pusat hubungan dengan konsep-konsep lainnya. Meskipun terbatas pada hadis Shahih al-Bukhari dan Shahih Muslim bertema shalat serta belum mempertimbangkan aspek linguistik dan konteks hadis secara mendalam, hasil penelitian menunjukkan bahwa *Association Rule Mining* dapat digunakan sebagai pendekatan eksploratif untuk menganalisis keterkaitan tematik pada matan hadis berbasis *text mining* bahasa Arab. Penelitian selanjutnya disarankan menggunakan dataset yang lebih luas dan metode NLP bahasa Arab yang lebih maju untuk meningkatkan kualitas analisis semantik hadis.

UCAPAN TERIMA KASIH

Penulis menyampaikan apresiasi kepada Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar, atas dukungan akademik selama penelitian ini berlangsung. Penulis juga mengucapkan terima kasih kepada seluruh pihak yang telah memberikan masukan dan kontribusi dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

Adila, J., Fadlisyah, & Ansari, S. F. (2025). Klasifikasi tema ayat Al-Qur'an pada juz 30 menggunakan text mining dan algoritma C4.5 berbasis web. *Jurnal Teknologi dan Sistem Informasi Universitas Abdurrab (JTI-UNIVRAB)*, 10(2), 1137–1148. https://doi.org/10.36341/rabit.v10i2.6502.

Apriandi, H. (2025). Integrasi hermeneutik digital dalam studi komparatif hadits: Antara warisan manuskrip dan algoritma modern. *Studi Hadis*, 3(1), 1-10.

Ariyanto, A. D. P., Fatichah, C., & Arifin, A. Z. (2021). Analysis of text representation methods to detect interrelation of hadith books: Systematic literature review. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 5*(5), 992–1000. https://doi.org/10.29207/resti.v5i5.3499.

Asy'ari, A. H., & Hanafi, M. (2025). Pengembangan model klasterisasi topik hadis Bukhari Muslim menggunakan BERT dengan penambahan fitur semantik. *Indonesian Journal Computer Science,* 4(2), 82–90. https://doi.org/10.31294/ijcs.v4i2.8931.

Athallah, M. R., & Lhaksmana, K. M. (2025). Hadith text classification based on topic using convolutional neural network (CNN) and TF-IDF. *Journal of Renewable Energy, Electrical, and Computer Engineering, 5*(1), 30–36. https://doi.org/10.29103/jreece.v5i1.20354.

Azwar, A., Usman, A. H., Abdullah, M. F. R., Khareng, M., & Shamsu, L. S. (2025). The integration of artificial intelligence in hadith scholarship: A bibliometric perspective on emerging research directions. *Nukhbatul 'Ulum: Jurnal Bidang Kajian Islam,* 11(1), 111–139. https://doi.org/10.36701/nukhbah.v11i1.2251.

Darmalaksana, W., Slamet, C., Zulfikar, W. B., Fadillah, I. F., Maylawati, D. S., & Ali, H. (2020). Latent semantic analysis and cosine similarity for hadith search engine. *TELKOMNIKA (Telecommunication Computing Electronics and Control),* 18(1), 217–227. https://doi.org/10.12928/TELKOMNIKA.V18I1.14874.

Darmawan, I., Elmunsyah, H., & Prasetya, D. D. (2025). ALBERTIR: A BERT-based pretraining for Indonesian religious texts using Qur'an and hadith translations. *Engineering, Technology and Applied Science Research,* 15(5), 28307–28312. https://doi.org/10.48084/etasr.12977.

Firdaus, M. T., & Suryadilaga, M. A. (2020). Integrasi keilmuan dalam kritik matan hadis. *TAJDID: Jurnal Ilmu Ushuluddin,* 18(2), 153–176. https://doi.org/10.30631/tjd.v18i2.96

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. arXiv. https://arxiv.org/abs/2203.05794

Kamala, H., Alif, M., & Millah, M. (2025). Konsep dasar dan transformasi historis studi hadis tematik dalam kajian ilmu hadis. *Ikhlas: Jurnal Ilmiah Pendidikan Islam*, 2(3), 246–260. https://doi.org/10.61132/ikhlas.v2i3.1144.

Khabir, N. A. I., & Ridlo, M. A. (2025). Transformasi hadis ke media digital. *Jurnal Syaikh Mudo Madlawan (JSMM): Kajian Ilmu-Ilmu Keislaman, 2*(1), 152–158.

Lajnah Pentashihan Mushaf Al-Qur'an. (2022). *Qur'an Kemenag*. Kementerian Agama Republik Indonesia.

Mosa, M. A. (2023). An exhaustive literature review of hadith text mining. *ACM Transactions on Asian and Low-Resource Language Information Processing,* 22(7), 1-16. https://doi.org/10.1145/3588315.

Muhammed, M., Azab, S., Ali, N., & Gheith, M. (2024). Arabic ontology for hadith texts: A survey. *The Egyptian Journal of Language Engineering,* 11(1), 1–14. https://doi.org/10.21608/ejle.2024.266774.1062

Nadoos, W. L., & Minaei-Bidgoli, B. (2025). CAMELBERT-CNN-BILSTM: Hybrid neural architecture with improved preprocessing to NER of Arabic hadith texts. Lex Localis: Journal of Local Self-Government, 23(S6), 8465-8479. https://doi.org/10.52152/m2wjzh19.

Qu, X., Gu, Y., Xia, Q., Li, Z., Wang, Z., & Huai, B. (2024). A survey on Arabic named entity recognition: Past, recent advances, and future trends. *IEEE Transactions on Knowledge and Data Engineering,* 36(3), 943–959. https://doi.org/10.1109/TKDE.2023.3303136.

Septiani, A., Astuti, B. S. F., & Hakim, R. F. (2015). Sequential pattern mining of hadith narrator with sequential pattern discovery using equivalent classes (SPADE) algorithm. In *International Conference on Waste Management, Ecology and Biological Sciences (WMEBS)* (pp. 147–154).

Sunnah.com. (2025). *Sunnah.com*. https://sunnah.com.

Surianto, D. F., Fajar B, M., Mulia, M. R., & Indanasufya, I. (2024). Comparative analysis of the performance of hadith text classification methods: A case study with ANN and SVM. *Journal of Embedded Systems, Security and Intelligent Systems,* 5(1), 89–98. https://doi.org/10.59562/jessi.v5i1.2942.

Uriawan, W., Priyajie, A. A., Gustian, A., Hidayat, F. N., Rafiudin, S. A., & Zaelani, M. F. (2025). *Unsupervised thematic clustering of hadith texts using the Apriori algorithm*. arXiv. https://arxiv.org/abs/2512.16694.

Yotenka, R. (2017). The implementation of Apriori algorithm and chi-square test in determining pattern of relationship among the rawi hadis. *Jurnal Eksakta UII,* 17(1), 20–32. https://doi.org/10.20885/eksakta.vol17.iss1.art3.

Yunus, B. M. (2020). Similarity detection for hadith of fiqh of women using cosine similarity and Boyer-Moore method. *International Journal of Advanced Trends in Computer Science and Engineering,* 9(1), 65–73. https://doi.org/10.30534/ijatcse/2020/11912020