

Optimasi Pertanyaan Menggunakan Refined Query Dalam Sistem Tanya Jawab Kitab Hadis

Andy Huang Wijaya¹, Nazruddin Safaat Harahap², Muhammad Irsyad³, Febi Yanto⁴

¹Universitas Islam Negeri Sultan Syarif Kasim, 12050113595@students.uin-suska.ac.id, Panam, Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

²Universitas Islam Negeri Sultan Syarif Kasim, nazruddin.safaat@uin-suska.ac.id, Panam, Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

³Universitas Islam Negeri Sultan Syarif Kasim, irsyadtech@uin-suska.ac.id, Panam, Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

⁴Universitas Islam Negeri Sultan Syarif Kasim, febiyanto@uin-suska.ac.id, Panam, Jl. HR. Soebrantas No.Km. 15, RW.15, Simpang Baru, Kota Pekanbaru, Riau, Indonesia

Informasi Makalah

Submit : May 3, 2024
Revisi : May 24, 2024
Diterima : Juni 1, 2024

Kata Kunci :

Refined Query;
Hadis;
Large Language Models;
Sistem Tanya Jawab;
Generative Pre-training Transformer.

Abstrak

Penelitian ini bertujuan untuk meningkatkan Sistem Tanya Jawab untuk teks Hadis dengan menggabungkan teknik Refined Query dan Large Language Models (LLMs), khususnya GPT-4 dari OpenAI. Memanfaatkan dataset sebanyak 62.169 Hadis dari sembilan buku penting. Penelitian ini mengikuti metodologi komprehensif yang mencakup pengumpulan, analisis, dan pra-pemrosesan data, serta integrasi LangChain dan Large Language Models (LLMs) OpenAI untuk kueri yang dioptimalkan. Evaluasi kinerja sistem dilakukan melalui analisis komparatif sebelum dan sesudah penerapan Refined Query, BERTScore untuk kualitas teks, dan penilaian kualitas berbasis pengguna. Hasilnya menunjukkan bahwa Refined Query secara signifikan meningkatkan kapasitas sistem untuk menghasilkan respons yang akurat dan relevan secara kontekstual. Menerapkan Refined Query tidak hanya meningkatkan ketepatan jawaban, tetapi juga memfasilitasi pembuatan tanggapan yang sebelumnya tidak tersedia. BERTScore rata-rata sebesar 0,80351 dan tingkat kualitas jawaban dari pengguna dengan skor rata-rata 87,3% untuk pengujian terhadap mahasiswa dan 90,3% pengujian kepada ahli hadis semakin memvalidasi kemampuan sistem.

Abstract

This research aims to enhance a Question-Answering System for Hadith texts by incorporating Refined Query techniques and Large Language Models (LLMs), specifically OpenAI's GPT-4. Utilizing a dataset of 62,169 Hadith from nine significant books, the study follows a comprehensive methodology that covers data collection, analysis and preprocessing, and the integration of LangChain and OpenAI's Chat Model for optimized querying. The evaluation of the system's performance was conducted through comparative analysis before and after the application of Refined Query, BERTScore for text quality, and user-based quality assessments. Results demonstrate that Refined Query significantly improves the system's capacity to produce

accurate and contextually relevant responses. Implementing Refined Query not only enhanced answer precision but also facilitated the generation of responses where none were previously available. The average BERTScore of 0.80351 and the quality of user responses with an average score of 87.3% for the student test and 90.3% for the hadith expert test further validate the efficacy of the system.

1. Pendahuluan

Akses terhadap informasi merupakan hal yang sangat penting dan bisa dikatakan sebagai kebutuhan pokok, terutama pada kebutuhan pendidikan, penelitian, dan pengambilan keputusan. Hadis adalah salah satu sumber hukum Islam selain Al-Qur'an, yang menjelaskan kalimat dalam Al-Qur'an yang mengacu pada ucapan, perbuatan, atau penetapan dalam kehidupan sehari-hari yang diambil dari ajaran Nabi Muhammad SAW (Abdul Aziz et al., 2021; Ariyanto et al., 2021; Ni'mah & Arifin, 2020). Kitab Hadis yang paling populer adalah Kitab Hadis Sunni yang terdiri dari 9 Kitab oleh 9 Imam, yaitu Shahih Muslim, Shahih Bukhari, Sunan at-Turmudzi, Sunan Abu Dawud, Sunan Ibnu Majah, Muwatta Malik, Sunan an-Nasa'i, Sunan Darimi, dan Musnad Ahmad (Ni'mah & Arifin, 2020). Dalam aspek literasi data hadis, hanya satu indikator dari beberapa indikator yang mampu direpresentasikan oleh lebih dari 80% peserta, menunjukkan bahwa analisis data dalam hadis memerlukan keterampilan yang cermat dan kritis (Supriyadi et al., 2020). Oleh karena itu, penting untuk memahami dan menganalisis data dalam hadis secara mendalam untuk memastikan informasinya.

Penggunaan sistem tanya jawab ini bertujuan untuk memenuhi kebutuhan informasi pengembang secara *on-demand* dengan memberikan respons mesin yang disesuaikan dengan konteks perangkat lunak dan pertanyaan individu (Bansal et al., 2021). Terdapat 3 tahapan utama dalam menghasilkan jawaban atas pertanyaan pengguna. Tahap pertama adalah formulasi *query*, yang mencakup dua sub-tahap yaitu pra-pemrosesan pertanyaan yang masuk dan tahap formulasi untuk mendapatkan kueri

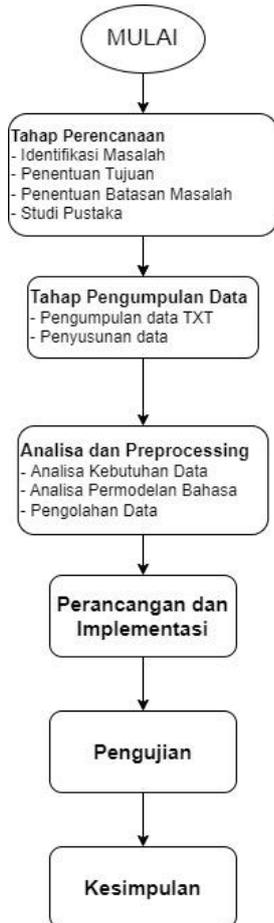
yang lebih baik. Tahap kedua adalah tahap pencarian informasi, saat di mana kueri yang telah dibuat dicocokkan dengan data vektor yang relevan dalam *database*. Tahap ketiga adalah penggalian dan pengolahan jawaban yang ditemukan (Maraoui et al., 2021; J. Wang et al., 2021).

Dalam perancangan sistem tanya jawab dibutuhkan pengolahan data hadis yang sudah ada agar representasi vektor padat dalam ruang berdimensi tinggi yang menangkap makna semantik dari kalimat atau teks, proses tersebut dinamakan dengan *embedding* (Saxena et al., 2020; Yu et al., 2023). Data vektor yang diperoleh kemudian disimpan pada *database* vektor, dalam penelitian ini akan menggunakan *ChromaDB* (González-Santamarta et al., 2023; Jeong, 2023). Kemudian dalam penelitian ini juga akan memanfaatkan *Large Language Models (LLMs)* seperti *GPT* dalam *Question Answering System* agar fokus penelitian menjadi lebih signifikan (Dao, 2023). *LLMs* memainkan peran penting dalam meningkatkan kualitas jawaban yang dihasilkan oleh sistem tersebut (Huo et al., 2023). Bantuan *Langchain* sebagai sebuah *framework* orkestrasi untuk mengelola dan menyediakan akses ke berbagai layanan bahasa alami, seperti penerjemahan, ringkasan teks, jawaban pertanyaan, pembangkitan teks, dan inferensi bahasa alami (Jeong, 2023; Topsakal & Akinci, 2023). Penggunaan *LLMs* seperti model chat *GPT* berfungsi untuk memahami pertanyaan dengan konteks yang lebih luas dan menyediakan jawaban yang lebih kontekstual (Y. Wang et al., 2023). Dengan memanfaatkan teknologi *Refined Query* ini diharapkan dapat membantu pengembangan sistem tanya jawab dapat mencapai tingkat kinerja yang lebih tinggi dalam menyediakan

jawaban yang sesuai dengan pertanyaan pengguna, terutama dalam konteks data hadis yang memiliki kompleksitas linguistik dan kontekstualitas yang tinggi.

2. Metode Penelitian

Alur yang digunakan pada penelitian ini terdiri dari beberapa proses, berikut proses dari penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

2.1. Perencanaan

Pada tahap awal ini terdapat 4 proses yang akan dilakukan untuk menunjukkan tujuan dari penelitian. Proses pertama adalah identifikasi masalah yang dirancang untuk mengenali pencarian dari isu penelitian. Proses kedua adalah penentuan tujuan, bertujuan agar dapat menentukan target yang ingin dicapai dalam penelitian. Proses selanjutnya adalah menentukan batasan masalah, guna agar ruang lingkup dalam pembahasan tidak keluar dari tujuan

penelitian. Proses terakhir adalah studi pustaka, dimana proses ini bertujuan untuk mencari referensi untuk membantu dalam penelitian dan juga untuk mendukung penelitian.

2.2. Pengumpulan Data

Tahapan ini melibatkan pengumpulan data hadis yang akan dijadikan sumber utama. Jumlah data yang dikumpulkan untuk penelitian ini sebanyak 5.362 hadis pada Kitab Shahih Muslim, 7.008 hadis pada Kitab Shahih Bukhari, 3.891 hadis pada Kitab Sunan at-Turmudzi, 4.590 hadis pada Kitab Sunan Abu Dawud, 4.332 hadis pada Kitab Sunan Ibnu Majah, 1.594 hadis pada Kitab Muwatta Malik, 5.662 hadis pada Kitab Sunan an-Nasa'I, 3.367 hadis pada Kitab Sunan Darimi, 26.363 hadis pada Kitab Musnad Ahmad, dengan total data 62.169 hadis yang akan digunakan pada penelitian ini (Rosdi et al., 2023). Pengumpulan data hadis adalah kunci untuk menyediakan konten yang akan diproses dan dijawab oleh sistem dalam bentuk *TXT*.

2.3. Analisa dan *Preprocessing*

Tahapan selanjutnya adalah menganalisa bagaimana format data yang dibutuhkan oleh sistem agar dapat diproses oleh model yang akan digunakan. Pada penelitian ini akan membutuhkan data dalam bentuk vektor agar dapat diproses. Kemudian untuk model bahasa yang akan digunakan pada penelitian adalah model *GPT* dari *OpenAI*.

Selanjutnya proses *preprocessing* yang akan dimulai dengan pengolahan data. Pada penelitian ini pengolahan data akan dilakukan proses *embedding* pada data hadis agar mesin dapat memahami makna dari kata-kata dalam pertanyaan dan dokumen yang dijadikan acuan (Huang et al., 2019).

2.4. Perancangan dan Implementasi

Pada tahap perancangan berfokus pada bagaimana *LangChain* dan *OpenAI Chat*

Model akan diintegrasikan untuk meningkatkan kemampuan sistem dalam menjawab pertanyaan dengan baik. Proses selanjutnya adalah implementasi yang melibatkan *coding*, integrasi sistem, dan pengaturan *database*. Keseluruhan proses ini bertujuan untuk menciptakan sistem yang mampu menjawab pertanyaan tentang hadis secara interaktif

2.5. Pengujian

Setelah sistem sudah diimplementasikan tahapan selanjutnya adalah melakukan pengujian yang bertujuan untuk menemukan dan memperbaiki *bug* serta memastikan kualitas jawaban yang diberikan sistem. Pengujian akan dilakukan dengan 3 tahapan pengujian. Tahapan pertama adalah penilaian dengan membandingkan jawaban saat menggunakan *Refined Query* dan tanpa *Refined Query*. Tahapan kedua adalah *BERTScore* yang bertujuan untuk mengevaluasi metrik secara otomatis untuk *text generation* (Zhang et al., 2019). Tahapan terakhir adalah Pengujian kualitas jawaban dilakukan kepada orang yang sering berinteraksi dengan Hadis seperti mahasiswa Fakultas Ushuluddin Universitas Islam Negeri Sultan Syarif Kasim dan beberapa ahli hadis.

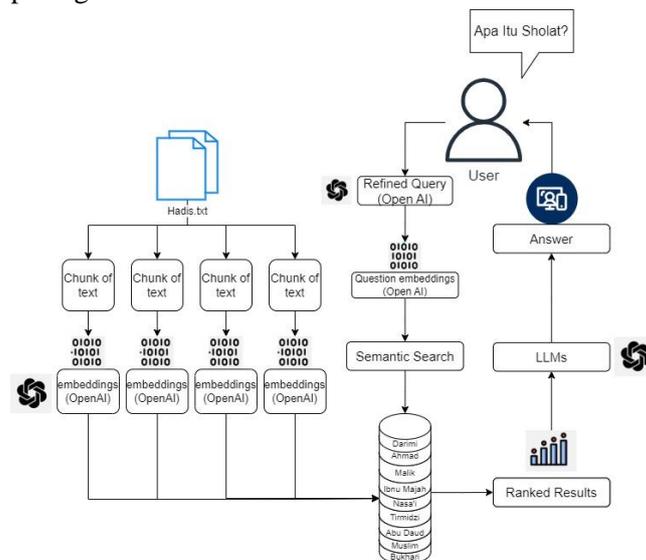
3. Hasil dan Pembahasan

Pada bab ini akan dilakukan 2 tahapan yaitu pengimplementasian hasil sistem yang sudah dirancang dan membahas tentang pengujian yang akan dilakukan.

3.1. Implementasi Sistem

Penelitian ini menghasilkan sebuah sistem tanya jawab untuk membantu pengguna untuk mencari informasi dalam hadis yang diharapkan dapat memberikan kemudahan bagi pengguna. Dimana sistem yang dibangun ini menggunakan program *Python*, *library Langchain*, *framework Streamlit*, *OpenAI chat model*, *database ChromaDB*,

dan *Tiktoken*. Alur untuk sistem dapat dilihat pada gambar 2 di bawah.



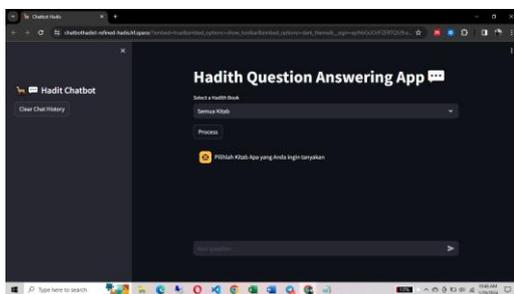
Gambar 2. Alur Sistem

Gambar 2 menjelaskan gambaran alur sistem yang terdiri dari 2 tahapan utama. Tahap pertama adalah penyimpanan data hadis ke dalam database. Pada tahapan ini peneliti membaginya menjadi 3 proses dimana proses awal adalah mengelompokkan data ke dalam bentuk *chunk*. Setelah itu dilanjutkan ke proses selanjutnya yaitu proses *embeddings* dimana proses ini merepresentasi vektor dari kata, frasa, atau kalimat dalam bentuk angka (Feng et al., 2022). Data vektor yang sudah terkumpul kemudian dimasukan ke dalam *database ChromaDB*.

Tahap kedua adalah tahapan sistem tanya jawab agar dapat menjawab pertanyaan dimana pada tahapan ini dibagi kedalam beberapa proses. Proses pertama ialah pengajuan kueri atau pertanyaan yang ingin disampaikan ke sistem. Selanjutnya kueri atau pertanyaan tersebut dilakukan proses *Refined Query* dimana proses ini adalah mekanisme penyempurnaan pertanyaan yang diajukan oleh pengguna diubah atau diperbaiki untuk meningkatkan kemungkinan mendapatkan jawaban yang lebih akurat dari sistem (Kim et al., 2022). Proses ketiga adalah *embedding* agar kueri atau pertanyaan yang sudah disempurnakan agar dapat

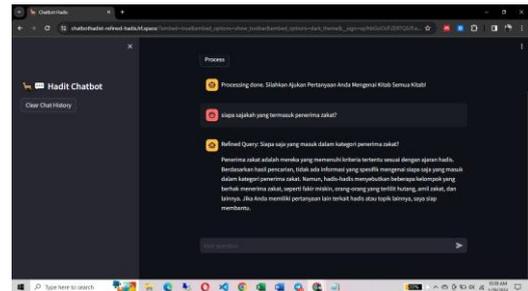
dicocokkan menggunakan *semantic search* ke dalam *database*. *Semantic search* merupakan proses pencarian informasi yang memahami makna dokumen dan pertanyaan di luar kata kunci (Muennighoff, 2022).

Setelah ditemukan kecocokan didalam *database* proses selanjutnya adalah *rank result* dimana data yang cocok diberikan peringkat berdasarkan tingkat kecocokannya. Setelah rank sudah ditentukan maka data yang cocok tersebut diolah lagi oleh *LLMs* dari *OpenAI* agar bisa mendapatkan jawaban yang akan ditampilkan kepada pengguna.



Gambar 3. Halaman Awal Sistem

Gambar 3 menampilkan halaman awal pada sistem tanya jawab, pengguna dapat memilih kitab hadis yang diinginkan pada *dropdown*.



Gambar 4. Tampilan setelah memberikan pertanyaan

Gambar 4 menampilkan halaman awal pada sistem tanya jawab dan sudah berisi percakapan tanya jawab antara pengguna dan *chatbot* yang sudah ditampilkan *Refined Query*.

3.2. Pengujian Sistem

Pengujian sistem dilakukan 3 tahap pengujian yang dimulai dengan pengujian terhadap *Refined Query* dan di lanjutkan dengan pengujian *BERTScore* dan ditutup dengan pengujian kualitas jawaban

3.2.1. Pengujian Terhadap *Refined Query*

Tabel 1 dibawah merupakan tabel yang menunjukkan hasil dari *Refined Query* yang sudah diuji mengubah *query* awal menjadi *query* yang lebih baik lagi format pertanyaannya.

Tabel 1. Perbandingan *Query* Awal dan *Refined Query*

Query Awal	Refined Query
Jika saya menemukan barang di jalan, apakah saya boleh mengambilnya?	Apakah diizinkan mengambil barang yang ditemukan di jalan?
Ketika saya sedang berpuasa, saya tidak sengaja meminum air karena saya lupa, bagaimana dengan puasa saya? apakah puasa saya akan diterima?	Apakah puasa seseorang masih sah jika ia secara tidak sengaja minum air karena lupa saat berpuasa?
Saya akan melakukan perjalanan jauh keluar kota dari kota pekanbaru ke jakarta, apakah saya boleh tidak melaksanakan puasa?	Apakah diperbolehkan tidak berpuasa ketika melakukan perjalanan jauh, seperti dari Pekanbaru ke Jakarta?
Saat sedang berpuasa saya muntah, bagaimana dengan puasa saya?	Apakah puasa saya batal jika saya muntah saat sedang berpuasa?
Pada saat shalat saya ingin membuang angin, namun karena sedang shalat saya menahannya, bagaimana dengan shalat saya?	Apakah shalat saya masih sah jika saya menahan buang angin saat sedang melaksanakan shalat?
Saya sedang melaksanakan sholat dan kemudian ada orang yang lewat didepan saya apa yang perlu saya lakukan ?	Apa yang harus saya lakukan jika ada orang yang lewat di depan saya saat saya sedang melaksanakan sholat?

Tabel Perbandingan <i>Query</i> Awal dan <i>Refined Query</i>	
<i>Query</i> Awal	<i>Refined Query</i>
Saya memberi makan orang yang berpuasa, seperti apa pahala yang saya dapatkan?	Apa pahala yang diperoleh ketika memberi makan kepada orang yang berpuasa?
Ketika saya sedang shalat, kemudian ada orang yang mengucapkan salam kepada saya, apa yang harus saya lakukan?	Apa yang harus dilakukan jika seseorang mengucapkan salam kepada kita saat kita sedang shalat?
Bagaimana cara membersihkan kaki yang telah terkena air liur anjing?	Bagaimana prosedur yang benar untuk membersihkan kaki yang telah tercemar oleh air liur anjing?
Keutamaan Bulan Ramadhan?	Apa saja keutamaan Bulan Ramadhan?

Dapat dilihat dari beberapa hasil dari *Refined Query* terdapat beberapa kata yang terdapat kesalahan tulisan namun masih dapat dimengerti dan diperbaiki dengan baik oleh sistem.

Dapat dilihat juga pada tabel diatas ada beberapa pertanyaan yang sifatnya cerita dapat diubah dengan baik menjadi pertanyaan yang lebih baku tanpa mengubah maksud dari pertanyaan yang diberikan pengguna

Tabel 2. Perbandingan Hasil Jawaban Tanpa *Refined Query* dan Menggunakan *Refined Query*

Tabel Hasil Jawaban Tanpa <i>Refined Query</i> dan Menggunakan <i>Refined Query</i>	
Jawaban tanpa <i>Refined Query</i>	Jawaban Sesudah <i>Refined Query</i>
Jika Anda menemukan barang di jalan, Anda harus mempersaksikannya kepada orang yang adil dan tidak boleh mengubah atau menyembunyikannya. Kemudian, umumkanlah barang temuan tersebut sampai setahun lamanya untuk mencari pemiliknya. Jika pemilik barang tersebut tidak ditemukan, maka barang tersebut adalah harta Allah yang Dia berikan kepada siapa saja yang dikehendaki-Nya.	Menurut Hadith yang disampaikan, barang temuan boleh diambil dengan syarat tertentu. Pertama, orang yang menemukan harus mengumumkannya sampai setahun, jika tidak ada yang mengenali maka kenalilah tali dan wadahnya dan boleh dimakan atau digunakan. Namun, jika pemiliknya datang maka barang tersebut harus dikembalikan kepada pemiliknya.
Ya, puasa Anda masih sah. Jika Anda lupa dan tidak sengaja makan atau minum saat berpuasa, Anda harus meneruskan puasa Anda karena Allah telah memberi Anda makan dan minum.	Ya, puasa masih dianggap sah jika seseorang tidak sengaja makan atau minum karena lupa saat sedang berpuasa. Rasulullah shallallahu 'alaihi wasallam bersabda bahwa jika seseorang lupa lalu makan atau minum ketika sedang berpuasa maka hendaklah dia meneruskan puasanya karena Allah telah memberinya makan dan minum. Hal ini menunjukkan bahwa puasa tidak batal karena lupa makan atau minum.
Berdasarkan hadis yang disampaikan, Rasulullah shallallahu 'alaihi wasallam bersabda bahwa berpuasa saat dalam perjalanan bukan termasuk kebaikan dan hendaknya umat Islam mengambil keringanan yang telah Allah berikan. Anda diizinkan untuk tidak berpuasa ketika melakukan perjalanan jauh dari Pekanbaru ke Jakarta.	Berdasarkan hadis yang disampaikan, Rasulullah SAW menegaskan bahwa bukanlah termasuk kebajikan berpuasa dalam perjalanan dan mengajarkan kita untuk menerima rukhsah (keringanan) yang diberikan Allah SWT. Oleh karena itu, seseorang yang melakukan perjalanan jauh, seperti dari Pekanbaru ke Jakarta, diperbolehkan untuk tidak menjalankan ibadah puasa. Rasulullah SAW mengajarkan untuk mengambil keringanan yang Allah SWT berikan dalam situasi perjalanan.

<p>Puasa Anda masih sah jika Anda muntah tanpa disengaja saat sedang berpuasa. Namun, jika Anda sengaja muntah, maka Anda wajib mengqadha puasa tersebut.</p>	<p>Berdasarkan hadis yang disampaikan, puasa Anda tidak batal jika Anda muntah tanpa sengaja saat sedang berpuasa. Namun, jika Anda sengaja muntah, maka Anda wajib mengganti (mengqadha') puasa tersebut. Hal ini ditegaskan dalam beberapa hadis yang disampaikan oleh Abu Hurairah dari Nabi shallallahu 'alaihi wasallam.</p>
<p>Shalat Anda sah selama tidak ada suara buang angin atau Anda tidak mencium bau yang keluar dari tubuh Anda. Jika Anda ragu dan tidak yakin apakah telah buang angin atau tidak tanpa adanya suara atau bau, maka shalat Anda tetap sah dan tidak perlu keluar dari masjid atau mengulangi shalat. Namun, jika terbukti buang angin dengan mendengar suara atau mencium bau, maka Anda harus berwudhu dan mengulangi shalat tersebut.</p>	<p>Berdasarkan hadis yang disampaikan, shalat tetap sah selama tidak ada keluarnya angin (kentut) yang dibuktikan dengan mendengar suara atau mencium bau. Jika seseorang buang angin ketika sedang shalat, maka ia harus berwudhu dan mengulangi shalatnya. Namun, jika hanya merasa ingin buang angin tetapi tidak ada bukti suara atau bau, maka tidak perlu membatalkan shalatnya.</p>
<p>Jika seseorang mencoba lewat di depan Anda saat Anda sedang shalat, Anda harus menghalanginya semampunya. Jika orang tersebut menolak untuk berhenti, maka hendaklah Anda memeranginya karena dia dianggap sebagai setan. Anda tidak boleh membiarkan seseorang lewat di hadapan Anda saat Anda shalat.</p>	<p>Berdasarkan hadis yang disampaikan, apabila ada orang yang mencoba lewat di depan seseorang yang sedang sholat, maka orang yang sholat tersebut hendaknya menghalanginya semampunya. Jika orang tersebut menolak untuk tidak lewat, maka orang yang sholat diperbolehkan untuk memeranginya karena dianggap orang tersebut membawa qarin atau setan. Rasulullah SAW juga menyampaikan bahwa sebaiknya orang yang akan lewat di depan orang yang sholat mengetahui bahwa lebih baik ia menunggu atau berdiri selama empat puluh (tidak diketahui apakah hari, bulan, atau tahun) daripada lewat di depannya.</p>
<p>Jika Anda memberi makan orang yang berpuasa untuk berbuka, maka Anda akan mendapatkan pahala seperti pahala orang yang berpuasa tersebut tanpa mengurangi pahala orang yang berpuasa sedikitpun.</p>	<p>Orang yang memberi makan kepada orang yang berpuasa akan mendapatkan pahala yang sama seperti pahala orang yang berpuasa tersebut, tanpa mengurangi pahala dari orang yang berpuasa sedikitpun. Ini berlaku untuk setiap kali ia memberi makan orang yang berpuasa untuk berbuka. Pahala yang diberikan ini mencerminkan nilai dan keutamaan dari membantu dan mendukung orang lain dalam ibadah mereka.</p>
<p>Maaf, saya tidak menemukan informasi tentang ini dalam konteks Hadis yang saya miliki. Apakah ada pertanyaan lain yang bisa saya bantu?</p>	<p>Saat kita sedang melaksanakan shalat dan ada seseorang yang mengucapkan salam, kita tidak perlu menjawab salam tersebut secara lisan. Hal ini didasarkan pada praktek Nabi Muhammad SAW yang tidak menjawab salam saat beliau sedang shalat setelah kembali dari negeri raja Najasyi, karena dalam shalat terdapat suatu kesibukan yang membutuhkan konsentrasi penuh. Seorang sahabat, Ibrahim, menyebutkan bahwa ia menjawabnya dalam hati.</p>

Tabel Hasil Jawaban Tanpa Refined Query dan Menggunakan Refined Query	
Jawaban tanpa <i>Refined Query</i>	Jawaban Sesudah <i>Refined Query</i>
<p>Berdasarkan konteks Hadis yang diberikan, tidak ada informasi spesifik tentang proses pembersihan kaki jika terkena air liur anjing. Maaf, saya tidak menemukan informasi tentang ini dalam konteks Hadis yang saya miliki. Apakah ada pertanyaan lain yang bisa saya bantu?</p>	<p>Prosedur yang benar untuk membersihkan kaki yang tercemar oleh air liur anjing adalah dengan mencucinya sebanyak tujuh kali, salah satunya menggunakan tanah. Ini berdasarkan hadis yang menyatakan jika bejana dijilat anjing, harus dicuci tujuh kali dan salah satunya dengan tanah. Meskipun konteks hadis mengacu pada bejana, prinsip yang sama dapat diterapkan pada kaki yang tercemar.</p>
<p>Bulan Ramadhan memiliki keutamaan yang besar dalam Islam, di mana Rasulullah SAW menyebutkan bahwa bulan ini lebih baik dari bulan-bulan lainnya, dan di dalamnya terdapat malam yang lebih baik dari seribu bulan. Orang yang berpuasa dan shalat malam di bulan Ramadhan dengan keimanan dan mengharap pahala akan diampuni dosa-dosanya seperti hari ia dilahirkan. Selain itu, di bulan Ramadhan pintu-pintu surga dibuka, pintu-pintu neraka ditutup, dan setan-setan dibelenggu.</p>	<p>Bulan Ramadhan memiliki keistimewaan dimana pintu-pintu surga dibuka dan pintu-pintu neraka ditutup, serta setan-setan dibelenggu. Ada malam di dalamnya yang lebih baik daripada seribu bulan, dikenal sebagai Lailatul Qadar. Barangsiapa yang berpuasa dan shalat malam dengan keimanan dan mengharap pahala, dosa-dosanya akan diampuni seperti hari ia dilahirkan.</p>

Dari hasil pengujian dapat disimpulkan bahwa setelah menggunakan *Refined Query* hasil jawaban menjadi lebih baik dan beberapa pertanyaan yang sebelumnya tidak ada jawaban menjadi ada jawabannya.

3.2.2. Pengujian *BERTScore*

Tabel 2 di bawah merupakan tabel yang menunjukkan hasil dari perhitungan menggunakan *BERTScore* yaitu *BERTScore* menghitung nilai dari *Precision*, *Recall*, dan *F1*.

$$R_{BERT} = \frac{1}{|x|} \sum_{x_i \in x} \max_{\hat{x}_j \in \hat{x}} x_i^T \hat{x}_j \quad (1)$$

Persamaan (1) di atas menjelaskan bahwa R_{BERT} merupakan rumus yang menghitung

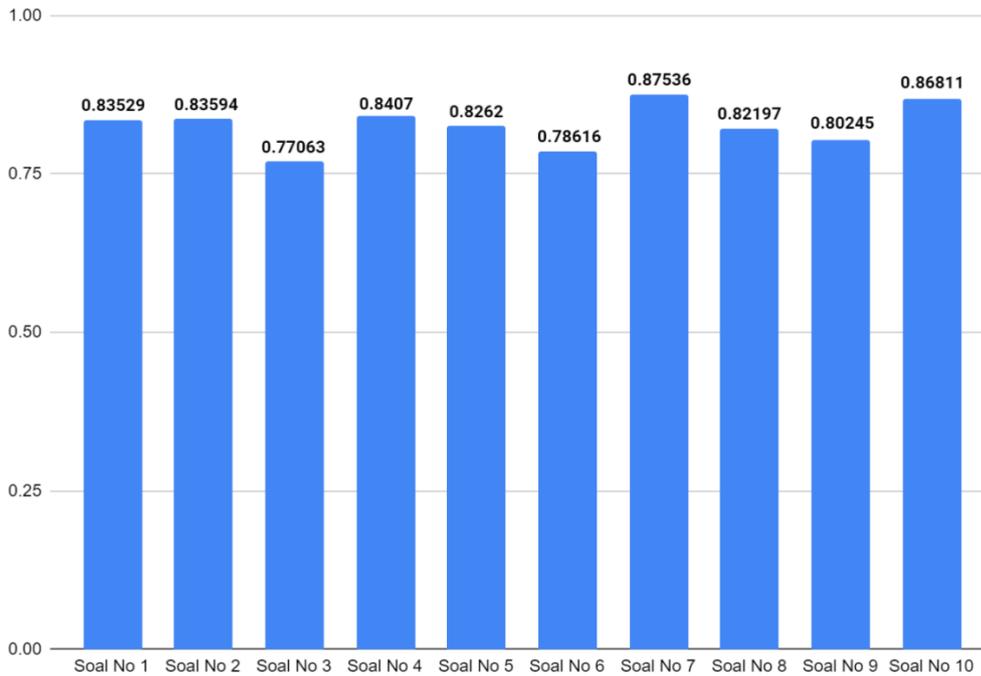
nilai korelasi *Pearson (Recall)* antara skor *BERT* dan kalimat referensi.

$$P_{BERT} = \frac{1}{|x|} \sum_{\hat{x}_j \in \hat{x}} \max_{x_i \in x} x_i^T \hat{x}_j \quad (2)$$

Persamaan (2) di atas menjelaskan bahwa P_{BERT} merupakan rumus yang mengukur nilai presisi sejauh mana kesamaan antara kalimat referensi dan kalimat kandidat.

$$F_{BERT} = 2 \frac{P_{BERT} \cdot R_{BERT}}{P_{BERT} + R_{BERT}} \quad (3)$$

Persamaan (3) di atas menjelaskan bahwa F_{BERT} merupakan rumus yang menghitung nilai *F1 score* yang merupakan perbandingan harmonis antara presisi (P_{BERT}) dan recall (R_{BERT})

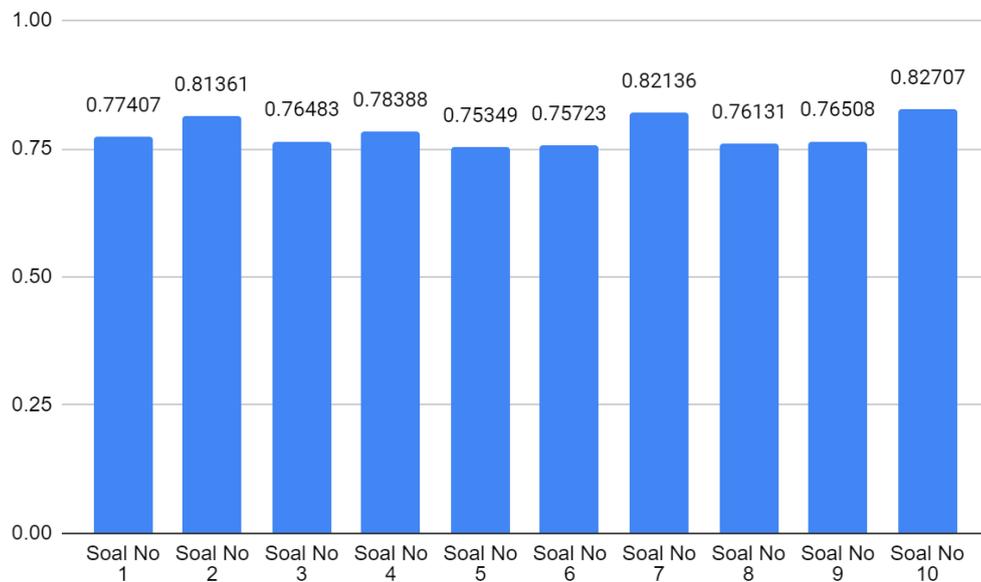


Gambar 5 Grafik Nilai *Precision*

Dapat dilihat pada gambar grafik diatas soal nomor 7 merupakan soal dengan nilai *precision* sedangkan soal dengan nilai *precision* terendah adalah soal nomor 3

Dapat disimpulkan bahwa soal nomor 3 mendapatkan nilai paling rendah karena

terdapat nama kota yang tidak ada didalam database hadis dan nomor 7 mendapatkan nilai tertinggi karena jawaban yang diberikan banyak teks dari sumber hadis yang diambil.



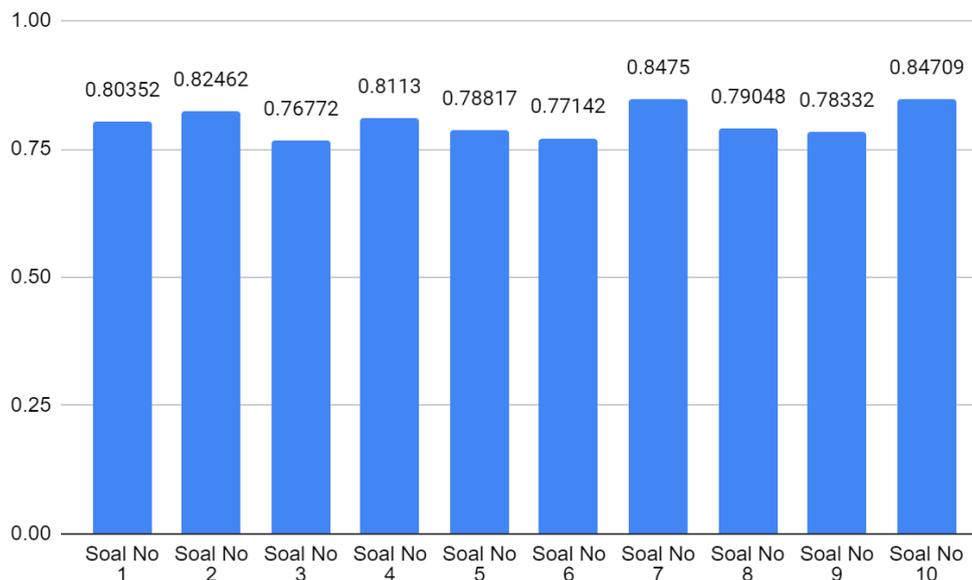
Gambar 6 Grafik Nilai *Recal*

Dapat dilihat pada gambar grafik diatas soal nomor 10 merupakan soal dengan nilai *recall*

sedangkan soal dengan nilai *recall* terendah adalah soal nomor 5

Dapat disimpulkan bahwa soal nomor 10 mendapatkan nilai tertinggi karena banyak hanya membuat kesimpulan dari sedikit data hadis dari database dan nomor 5

mendapatkan nilai paling rendah karena jawaban yang diberikan menyimpulkan dari banyak sumber hadis dalam database



Gambar 7 Grafik Nilai FI

Dapat dilihat pada gambar grafik diatas soal nomor 7 merupakan soal dengan nilai FI sedangkan soal dengan nilai FI terendah adalah soal nomor 3.

3.2.3. Pengujian Kualitas Jawaban

Pengujian ini dilakukan dengan bantuan 20 responden mahasiswa dan 6 ahli hadis yang akan menilai kualitas jawaban dari 10 soal dan jawaban yang sudah dipertanyakan sebelumnya menggunakan sistem yang telah dibangun.

Tabel 3. Pilihan Jawaban dan Bobot Penilaian

Tabel Pilihan Jawaban dan Bobot Penilaian		
Pilihan Jawaban	Keterangan	Nilai Bobot
STS	Sangat Tidak Sesuai	1
TS	Tidak Sesuai	2
N	Netral	3
S	Sesuai	4
SS	Sangat Sesuai	5

Tabel 4. Jumlah Jawaban Kuesioner Mahasiswa

Tabel Jumlah Hasil Kuesioner Mahasiswa							
No Soal	Jumlah Jawaban					Total Nilai	%
	SS	S	N	TS	STS		
1	6	11	2	1	0	82	82%
2	15	5	0	0	0	95	95%
3	8	6	6	0	0	82	82%
4	13	4	0	3	0	87	87%
5	10	7	2	1	0	86	86%
6	6	7	4	3	0	76	76%
7	13	7	0	0	0	93	93%
8	8	11	1	0	0	87	87%
9	12	8	0	0	0	92	92%
10	10	15	3	2	0	93	93%

Hasil pengolahan data dari jawaban kuesioner bagi mahasiswa Fakultas Ushuluddin Universitas Islam Negeri Sultan Syarif Kasim sebagai responden menunjukkan nilai tertinggi ada pada soal nomor 2 dengan nilai 95% dan soal dengan nilai terendah ada pada soal nomor 6 dengan nilai 76% dengan rata-rata nilai 87,3%.

Tabel 5. Jumlah Hasil Kuesioner Ahli Hadis

Tabel Jumlah Hasil Kuesioner Ahli Hadis							
No Soal	Jumlah Jawaban					Total Nilai	%
	SS	S	N	TS	STS		
1	4	2	0	0	0	28	93,3%
2	4	2	0	0	0	28	93,3%
3	4	2	0	0	0	28	93,3%
4	4	2	0	0	0	28	93,3%
5	2	4	0	0	0	26	86,7%
6	4	2	0	0	0	28	93,3%
7	2	3	1	0	0	25	83,3%
8	4	1	0	1	0	26	86,7%
9	3	2	1	0	0	26	86,7%
10	4	2	0	0	0	28	93,3%

Hasil pengolahan data dari jawaban kuesioner kepada ahli hadis sebagai responden menunjukkan nilai tertinggi ada pada soal nomor 1, 2, 3, 4, 6, dan 10 dengan nilai 93,3% dan soal dengan nilai terendah ada pada soal nomor 7 dengan nilai 83,3% dengan rata-rata nilai 90,3%.

4. Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan *Refined Query* dengan pengembangan sistem tanya jawab hadis telah menunjukkan hasil yang baik. Sistem ini mampu membantu memperbaiki kueri sehingga sistem dapat memberikan jawaban yang akurat dan relevan terhadap pertanyaan-pertanyaan seputar hadis, dengan memanfaatkan model *GPT-4-1106-preview* dari *OpenAI*.

Pengujian awal dilakukan dengan membandingkan jawaban dari sistem sebelum dan sesudah menggunakan *Refined Query*, pengujian ini menghasilkan kesimpulan bahwa setelah menggunakan *Refined Query* hasil jawaban menjadi lebih baik dan beberapa pertanyaan yang sebelumnya tidak ada jawaban menjadi ada jawabannya

Pengujian selanjutnya dilakukan dengan menggunakan matrik *BERTScore* yang merupakan salah satu metode evaluasi yang digunakan untuk mengukur kualitas teks yang dihasilkan oleh model bahasa. Hasil pengujian menunjukkan bahwa nilai rata-rata *BERTScore* adalah 0.80351.

Selain itu, pengujian lebih lanjut dilakukan dengan mengevaluasi terhadap kualitas jawaban yang diberikan oleh sistem menunjukkan hasil yang memuaskan, dengan perolehan nilai rata-rata sebesar 87.3% untuk pengujian kepada mahasiswa dan nilai rata-rata 90,3% untuk pengujian kepada ahli hadis. Angka ini mengindikasikan bahwa jawaban yang dihasilkan oleh sistem memiliki kualitas yang tinggi terhadap pertanyaan yang diajukan oleh pengguna.

Pencapaian ini semakin memperkuat bukti bahwa sistem tanya jawab hadis yang dikembangkan bantuan *Refined Query* menghasilkan performa yang baik dalam menghasilkan jawaban yang sesuai dan berkualitas.

5. Referensi

- Abdul Aziz, A., Salleh, D. M., Fadylawaty, S., Abdullah, S., & Norazmi Bin Nordin, M. (2021). Analysis Of Literature Review On Spiritual Concepts According To The Perspectives Of The Al-Quran, Hadith And Islamic Scholars. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(9), 3152–3159. <https://turcomat.org/index.php/turkbilmat/article/view/4790>
- Ariyanto, A. D. P., fatichah, C., & Arifin, A. Z. (2021). Analisis Metode Representasi Teks Untuk Deteksi Interelasi Kitab Hadis: Systematic Literature Review. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 992–1000. <https://doi.org/10.29207/RESTI.V5I5.3499>
- Bansal, A., Eberhart, Z., Wu, L., & Mcmillan, C. (2021). A Neural Question Answering System for Basic Questions about Subroutines. *ArXiv Preprint*. arXiv:2101.03999v1
- Dao, X.-Q. (2023). Performance Comparison of Large Language Models on VNHSGE English Dataset: OpenAI ChatGPT,

- Microsoft Bing Chat, and Google Bard. *ArXiv Preprint*. arXiv:2307.02288
- Feng, F., Yang, Y., Cer, D., Arivazhagan, N., & Wang, W. (2022). Language-agnostic BERT Sentence Embedding. *Proceedings of the Annual Meeting of the Association for Computational Linguistics, 1*, 878–891. <https://doi.org/10.18653/V1/2022.ACL-LONG.62>
- González-Santamarta, M. A., Rodríguez-Lera, F. J., Angel, ´, Guerrero-Higueras, M., & Matellán-Olivera, V. (2023). *Integration of Large Language Models within Cognitive Architectures for Autonomous Robots*. <https://arxiv.org/abs/2309.14945v1>
- Huang, X., Zhang, J., Li, D., & Li, P. (2019). Knowledge graph embedding based question answering. *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining*, 105–113. <https://doi.org/10.1145/3289600.3290956>
- Huo, S., Arabzadeh, N., & Clarke, C. L. A. (2023). Retrieving Supporting Evidence for LLMs Generated Answers. *ArXiv Preprint*. <https://arxiv.org/abs/2306.13781v1>
- Jeong, C. (2023). A Study on the Implementation of Generative AI Services Using an Enterprise Data-Based LLM Application Architecture. *Advances in Artificial Intelligence and Machine Learning, 03(04)*, 1588–1618. <https://doi.org/10.54364/AAIML.2023.1191>
- Kim, R., Research, A. G., Webster, K., Research, G., Collins, M., & Narayan, S. (2022). Query Refinement Prompts for Closed-Book Long-Form Question Answering. *ArXiv Preprint*. arXiv:2210.17525v1
- Maraoui, H., Haddar, K., & Romary, L. (2021). Arabic Factoid Question-Answering System for Islamic Sciences Using Normalized Corpora. *Procedia Computer Science, 192*, 69–79. DOI : 10.1016/J.PROCS.2021.08.008
- Muennighoff, N. (2022). SGPT: GPT Sentence Embeddings for Semantic Search. *ArXiv Preprint*. <https://arxiv.org/abs/2202.08904v5>
- Ni'mah, A. T., & Arifin, A. Z. (2020). Perbandingan Metode Term Weighting terhadap Hasil Klasifikasi Teks pada Dataset Terjemahan Kitab Hadis. *Rekayasa, 13(2)*, 172–180. <https://doi.org/10.21107/REKAYASA.V13I2.6412>
- Rosdi, A. Z., Hassan, S. N. S., Muhamad, N. A. F., Zainuzi, N. I. H. M., & Mahfuz, M. S. (2023). Panduan Asas Kaedah Kenal Pasti Status Hadis: Kajian Diskriptif Penggunaan Ensiklopedia Hadis 9 Imam. *Journal Of Hadith Studies, 46–54*. <https://doi.org/10.33102/JOHS.V8I1.225>
- Saxena, A., Tripathi, A., & Talukdar, P. (2020). Improving Multi-hop Question Answering over Knowledge Graphs Using Knowledge Base Embeddings. *Proceedings of the Annual Meeting of the Association for Computational Linguistics, 4498–4507*. <https://doi.org/10.18653/V1/2020.ACL-MAIN.412>
- Supriyadi, T., Julia, J., Aeni, A., Learning, E. S.-I. J. of, & 2020, undefined. (2020). Action Research in Hadith Literacy: A Reflection of Hadith Learning in The Digital Age. *Ijltter.Net, 19(5)*, 99–124. <https://doi.org/10.26803/ijltter.19.5.6>
- Topsakal, O., & Akinci, T. C. (2023). Creating Large Language Model Applications Utilizing LangChain: A Primer on Developing LLM Apps Fast. *All Sciences Proceedings*. <https://doi.org/10.59287/icaens.1127>
- Wang, J., Yi, X., Guo, R., Jin, H., Xu, P., Li, S., Wang, X., Guo, X., Li, C., Xu, X., Yu, K., Yuan, Y., Zou, Y., Long, J., Cai, Y., Li, Z., Zhang, Z., Mo, Y., Gu, J., ... Xie, C. (2021). Milvus: A Purpose-Built Vector Data Management System. *Proceedings of the ACM SIGMOD International Conference on Management of Data, 2614–2627*. <https://doi.org/10.1145/3448016.3457550>
- Wang, Y., Ma, X., & Chen, W. (2023). Augmenting Black-box LLMs with Medical Textbooks for Clinical Question Answering. *ArXiv Preprint*. <https://arxiv.org/abs/2309.02233v1>
- Yu, C.-W., Chuang, Y.-S., Lotsos, A. N., & Haase, C. M. (2023). Decoding Affect in Dyadic Conversations: Leveraging Semantic Similarity through Sentence Embedding. *ArXiv Preprint*. arXiv:2309.12646v1
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). BERTScore: Evaluating Text Generation with BERT. *8th International Conference on Learning Representations, ICLR 2020*. <https://arxiv.org/abs/1904.09675v3>