

Question Answering Berbasis Transformer pada Koleksi Hadis Kutubut Tis'ah untuk Pembelajaran Pendidikan Agama Islam

Ana Tsalitsatun Ni'mah^{1*}, Riva Ananda Yuanova Kamajaya², Ariesta Kartika Sari³

^{1,2,3} Pendidikan Informatika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Trunojoyo Madura

¹ana.tsalits@trunojoyo.ac.id, ²raykamajayaa@gmail.com, ³ariestakartika@trunojoyo.ac.id

ABSTRAK

Perkembangan teknologi Natural Language Processing (NLP) berbasis Transformer telah mendorong inovasi dalam sistem Question Answering (QA). Dalam konteks pembelajaran hadis di Pendidikan Agama Islam, akses terhadap koleksi Kutubut Tis'ah masih menghadapi kendala dalam pencarian informasi yang cepat dan kontekstual. Oleh karena itu, diperlukan sistem cerdas yang mampu memahami pertanyaan pengguna dan memberikan jawaban yang relevan secara otomatis. Penelitian ini mengembangkan sistem Question Answering berbasis Transformer dengan memanfaatkan model pra-latih BERT (Bidirectional Encoder Representations from Transformers). Tahapan penelitian meliputi pengumpulan dan pra-proses data hadis Kutubut Tis'ah, anotasi dataset, perancangan arsitektur model, pelatihan, serta pengujian sistem. Dataset dibagi menjadi data latih dan data uji dengan proporsi 80:20. Evaluasi dilakukan menggunakan metrik Exact Match (EM) dan F1-score untuk mengukur performa model. Hasil pengujian menunjukkan bahwa model yang dikembangkan memperoleh nilai Exact Match (EM) sebesar 78,6% dan F1-score sebesar 86,3%. Sistem mampu memberikan jawaban yang relevan dan kontekstual terhadap berbagai variasi pertanyaan pengguna, serta berhasil mengidentifikasi bagian teks hadis yang sesuai sebagai jawaban dengan tingkat akurasi yang baik. Penerapan model Transformer terbukti efektif dalam meningkatkan kualitas pencarian informasi hadis dibandingkan metode konvensional berbasis pencocokan kata kunci. Sistem ini berpotensi menjadi media pembelajaran hadis yang interaktif, efisien, dan mudah diakses.

Kata kunci: *Question Answering, Transformer, Kutubut Tis'ah, Pembelajaran Hadis, Pendidikan Agama Islam*

ABSTRACT

The development of Transformer-based Natural Language Processing (NLP) technology has driven innovation in Question Answering (QA) systems. In the context of hadith learning in Islamic Religious Education, access to the Kutubut Tis'ah collection still faces obstacles in quickly and contextually searching for information. Therefore, an intelligent system is needed that can understand user questions and provide relevant answers automatically. This study developed a Transformer-based Question Answering system by utilizing the BERT (Bidirectional Encoder Representations from Transformers) pre-trained model. The research stages include collecting and pre-processing Kutubut Tis'ah hadith data, dataset annotation, model architecture design, training, and system testing. The dataset was divided into training and test data with a proportion of 80:20. Evaluation was carried out using the Exact Match (EM) and F1-score metrics to measure model performance. The test results showed that the developed model obtained an Exact Match (EM) value of 78.6% and an F1-score of 86.3%. The system is capable of providing relevant and contextual answers to a variety of user questions and successfully identifying appropriate sections of hadith text as answers with a good degree of accuracy. The application of the Transformer model has proven effective in improving the quality of hadith information retrieval compared to conventional keyword-matching methods. This system has the potential to become an interactive, efficient, and easily accessible hadith learning medium.

Keywords: *Question Answering, Transformer Models, Kutubut Tis'ah, Hadis Studies, Islamic Religious Education*



PENDAHULUAN

Perkembangan teknologi informasi telah mendorong transformasi besar dalam cara manusia memanfaatkan pengetahuan keagamaan (Ananda et al., 2025). Kebutuhan sistem pencarian informasi yang cepat, akurat, dan kontekstual saat ini semakin meningkat begitu pula dalam bidang studi Islam. Salah satu sumber utama ajaran Islam setelah Al-Qur'an adalah hadis yang berisi perkataan, perbuatan, dan ketetapan Nabi Muhammad saw. Keberadaan hadis sebagai pedoman hidup menjadikannya objek kajian penting dalam pendidikan, penelitian, maupun kehidupan sehari-hari umat Islam. Oleh karena itu, diperlukan inovasi teknologi yang mampu mempermudah masyarakat dalam memahami kandungan hadis secara efektif (Sartika et al., 2025).

Koleksi Kutubut Tis'ah merupakan himpunan sembilan kitab hadis utama yang memiliki kedudukan penting dalam tradisi keilmuan Islam. Kitab-kitab tersebut mencakup karya para imam besar seperti Sahih Bukhari, Sahih Muslim, Sunan Abu Dawud, Sunan Tirmidzi, Sunan Nasa'i, Sunan Ibnu Majah, Musnad Ahmad, Muwatta Malik, dan Sunan Darimi (Ni'mah & Yunitarini, 2024). Koleksi ini menjadi rujukan luas dalam pembelajaran hadis karena memuat ribuan riwayat dengan tema yang beragam, mulai dari akidah, ibadah, muamalah, hingga akhlak. Namun, besarnya jumlah data dalam Kutubut Tis'ah juga menimbulkan tantangan tersendiri dalam proses pencarian dan penemuan informasi yang relevan (Afabih et al., 2025).

Secara konvensional, pencarian hadis dilakukan melalui indeks kitab, kitab syarah, atau bantuan guru dan ulama. Metode tersebut memiliki nilai akademik yang tinggi, tetapi sering kali membutuhkan waktu lama serta kemampuan literasi teks Arab yang memadai. Bagi pelajar, mahasiswa, maupun masyarakat umum, proses menemukan hadis yang sesuai dengan pertanyaan tertentu dapat menjadi hambatan dalam pembelajaran. Terlebih lagi, pertanyaan yang diajukan sering bersifat natural language atau bahasa sehari-hari, bukan berupa kata kunci formal sebagaimana tercantum dalam kitab hadis (Tsalitsatun et al., 2024).

Dalam konteks tersebut, sistem Question Answering (QA) menjadi solusi yang relevan. QA merupakan cabang kecerdasan buatan yang dirancang untuk memberikan jawaban langsung terhadap pertanyaan pengguna berdasarkan sumber data tertentu (Maxutova, 2025)(Shao, 2019). Berbeda dengan mesin pencari tradisional yang hanya menampilkan daftar dokumen, sistem QA berusaha memahami maksud pertanyaan dan menyajikan jawaban paling sesuai. Pada domain pendidikan Islam, sistem QA dapat dimanfaatkan untuk membantu pengguna menemukan hadis secara cepat berdasarkan pertanyaan tematik maupun factual (Kim et al., 2022)(Alsubhi et al., 2022).

Perkembangan terbaru dalam Natural Language Processing (NLP) menunjukkan bahwa model berbasis Transformer memiliki performa unggul dalam berbagai tugas pemrosesan bahasa, termasuk question answering (Wang et al., 2023)(Ahmed et al., 2023). Arsitektur Transformer memungkinkan model memahami hubungan antar kata dalam kalimat melalui mekanisme attention, sehingga mampu menangkap konteks secara lebih baik dibanding pendekatan sebelumnya (Zhu et al., 2022)(Yang et al., 2023). Model-model seperti BERT, RoBERTa, dan variannya telah terbukti efektif dalam tugas ekstraksi jawaban, klasifikasi teks, dan pencarian semantic (Mitchell et al., n.d.)(Souza et al., 2023).

Penerapan model Transformer pada teks hadis menjadi bidang penelitian yang menarik sekaligus menantang. Bahasa hadis memiliki karakteristik khas, seperti struktur bahasa Arab klasik, variasi periwayatan, istilah keagamaan, serta konteks historis tertentu (Abdallah et al.,

2023)(Shen & Id, 2023). Selain itu, teks hadis sering kali terdiri atas sanad dan matan yang perlu dipahami secara berbeda. Oleh sebab itu, pengembangan sistem QA berbasis Transformer untuk koleksi Kutubut Tis'ah memerlukan strategi prapemrosesan data, representasi teks, dan penyesuaian model yang tepat.

Dalam pembelajaran hadis di Pendidikan Agama Islam, keberadaan sistem QA berbasis Transformer berpotensi meningkatkan efektivitas proses belajar. Peserta didik dapat mengajukan pertanyaan seperti "hadis tentang kejujuran", "larangan riba", atau "keutamaan menuntut ilmu", lalu memperoleh jawaban dari koleksi hadis yang relevan (Tsalitsatun et al., 2025). Hal ini dapat mempercepat eksplorasi materi, meningkatkan minat belajar, dan mendukung pembelajaran mandiri. Guru dan dosen juga dapat memanfaatkan sistem tersebut sebagai media pendukung dalam menyusun materi ajar maupun diskusi kelas.

Meskipun demikian, implementasi sistem QA pada sumber keagamaan memerlukan perhatian terhadap aspek akurasi, validitas, dan etika. Jawaban yang dihasilkan harus tetap merujuk pada sumber hadis yang sahih dan dapat ditelusuri kembali. Sistem tidak cukup hanya memberikan teks jawaban, tetapi juga perlu menyertakan informasi kitab, nomor hadis, serta konteks riwayat. Dengan demikian, teknologi berfungsi sebagai alat bantu ilmiah, bukan pengganti otoritas keilmuan para ahli hadis.

Penelitian mengenai Question Answering berbasis Transformer pada koleksi Kutubut Tis'ah menjadi penting karena menggabungkan kemajuan teknologi modern dengan kebutuhan pembelajaran Islam kontemporer. Kajian ini dapat memberikan kontribusi pada pengembangan digital humanities Islam, khususnya dalam pemanfaatan kecerdasan buatan untuk studi hadis. Selain itu, penelitian ini juga membuka peluang pengembangan dataset bahasa Arab dan Indonesia yang berkaitan dengan teks keagamaan.

Berdasarkan uraian tersebut, penelitian ini bertujuan merancang dan mengevaluasi sistem Question Answering berbasis Transformer pada koleksi hadis Kutubut Tis'ah untuk pembelajaran Pendidikan Agama Islam. Fokus utama penelitian meliputi proses pengolahan data hadis, pemilihan model Transformer yang sesuai, pengujian performa sistem, serta analisis manfaatnya dalam konteks pendidikan. Hasil penelitian diharapkan mampu menjadi solusi inovatif dalam meningkatkan aksesibilitas, efisiensi, dan kualitas pembelajaran hadis di era digital.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan Research and Development (R&D) dengan metode eksperimen komputasional untuk mengembangkan dan mengevaluasi sistem Question Answering (QA) berbasis Transformer pada koleksi Kutubut Tis'ah. Pendekatan ini dipilih karena penelitian tidak hanya berfokus pada analisis, tetapi juga menghasilkan produk berupa sistem cerdas yang dapat dimanfaatkan dalam pembelajaran hadis. Proses penelitian dilakukan secara bertahap mulai dari pengumpulan data hingga evaluasi performa model.

Objek penelitian berupa teks hadis yang bersumber dari koleksi Kutubut Tis'ah, yaitu sembilan kitab hadis utama dalam Islam, yang mencakup Shahih Bukhari, Shahih Muslim, Sunan Abu Dawud, Sunan Tirmidzi, Sunan Nasa'i, Sunan Ibnu Majah, Musnad Ahmad, Muwatta Malik, dan Sunan Darimi. Data diperoleh dari sumber digital terpercaya dalam bentuk teks Arab dan terjemahan bahasa Indonesia. Setiap data hadis terdiri atas informasi nomor hadis,

nama kitab, bab, sanad, dan matan, yang kemudian digunakan sebagai basis dalam penyusunan dataset Question Answering.

Tahap awal penelitian adalah pengumpulan dan praproses data. Pengumpulan data dilakukan melalui teknik dokumentasi terhadap sumber digital hadis. Selanjutnya dilakukan praproses yang meliputi pembersihan teks (cleaning), normalisasi karakter, tokenisasi, penghapusan data duplikat, serta pemisahan antara sanad dan matan. Proses ini bertujuan untuk meningkatkan kualitas data sehingga sesuai dengan kebutuhan model berbasis Transformer.

Setelah praproses, dilakukan penyusunan dataset dalam bentuk pasangan pertanyaan dan jawaban (question-answer pairs). Pertanyaan disusun berdasarkan kebutuhan pengguna dengan berbagai variasi, seperti pertanyaan faktual, tematik, dan natural language. Jawaban diambil dari bagian teks hadis yang relevan. Dataset kemudian dianotasi untuk menentukan posisi jawaban dalam teks konteks, yang diperlukan dalam proses pelatihan model QA berbasis ekstraksi.

Model yang digunakan dalam penelitian ini adalah BERT (Bidirectional Encoder Representations from Transformers) sebagai representasi arsitektur Transformer. Model BERT yang telah dilatih sebelumnya (pre-trained) kemudian disesuaikan (fine-tuning) untuk tugas Question Answering menggunakan dataset yang telah disiapkan. Pemilihan BERT didasarkan pada kemampuannya dalam memahami konteks dua arah (bidirectional) sehingga lebih efektif dalam menangkap makna teks hadis.

Dataset dibagi menjadi data latih dan data uji dengan proporsi 80:20. Proses pelatihan dilakukan menggunakan parameter seperti learning rate, batch size, dan jumlah epoch yang disesuaikan untuk memperoleh performa optimal. Selama pelatihan, model belajar untuk memprediksi posisi awal (start index) dan posisi akhir (end index) dari jawaban dalam teks konteks.

Rumus Precision dan Recall di atas digunakan untuk mengukur performa model klasifikasi atau sistem Question Answering dalam menilai ketepatan dan kelengkapan hasil prediksi pada setiap kelas ke- i . Kedua metrik ini sangat penting dalam mengevaluasi kemampuan sistem seperti KutubBot dalam memberikan jawaban yang relevan berdasarkan pertanyaan pengguna.

Nilai Precision pada kelas ke- i dihitung dengan membandingkan jumlah data yang diprediksi benar pada kelas tersebut terhadap seluruh data yang diprediksi sebagai kelas ke- i . Secara matematis dinyatakan sebagai:

$$Precision_i = \frac{M_{ii}}{M_{ii} + \sum_{j=1 \text{ to } n, i \neq j} M_{ij}} \quad (1)$$

Pada rumus tersebut, M_{ii} menyatakan jumlah prediksi yang benar pada kelas ke- i (true positive), sedangkan M_{ij} menunjukkan jumlah data dari kelas ke- i yang salah diprediksi menjadi kelas lain. Dengan demikian, semakin tinggi nilai Precision, semakin sedikit kesalahan sistem dalam memberikan prediksi pada kelas tersebut. Dalam konteks sistem QA hadis,

Precision yang tinggi menunjukkan bahwa jawaban yang diberikan sistem cenderung tepat dan relevan terhadap pertanyaan pengguna.

Sementara itu, Recall pada kelas ke-*i* digunakan untuk mengukur kemampuan sistem dalam menemukan seluruh data yang seharusnya termasuk ke dalam kelas tersebut. Rumus Recall dinyatakan sebagai:

$$Recall_i = \frac{M_{ii}}{M_{ii} + \sum_{j=1 \text{ to } n, i \neq j} M_{ji}} \tag{2}$$

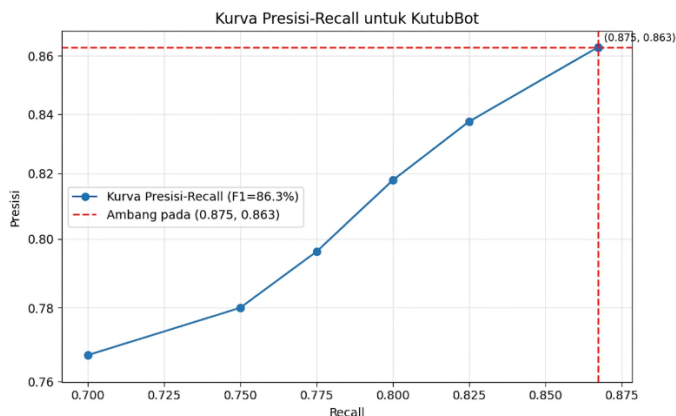
Pada persamaan ini, *M_{ii}* tetap menyatakan jumlah prediksi benar pada kelas ke-*i*, sedangkan *M_{ji}* menunjukkan jumlah data dari kelas lain yang sebenarnya termasuk kelas ke-*i* tetapi gagal dikenali sistem. Nilai Recall yang tinggi menunjukkan bahwa sistem mampu menemukan sebagian besar jawaban yang benar dan tidak banyak melewatkan informasi penting.

Secara umum, Precision berfokus pada ketepatan hasil prediksi, sedangkan Recall berfokus pada kelengkapan hasil yang berhasil ditemukan. Dalam pengembangan sistem KutubBot, kedua metrik ini digunakan secara bersamaan untuk memastikan bahwa sistem tidak hanya memberikan jawaban yang benar, tetapi juga mampu menemukan jawaban yang relevan secara menyeluruh dari teks hadis Kutubut Tis'ah.

Evaluasi sistem dilakukan menggunakan metrik Exact Match (EM) dan F1-score, yang umum digunakan dalam sistem Question Answering. Exact Match mengukur persentase jawaban yang sama persis dengan jawaban referensi, sedangkan F1-score mengukur tingkat kesamaan berdasarkan overlap kata antara jawaban sistem dan jawaban referensi.

HASIL DAN PEMBAHASAN

Pada tahap pengujian, analisis kuantitatif dilakukan untuk mengevaluasi performa algoritma BERT dalam sistem pencarian KutubBot. Proses analisis kinerja model difokuskan pada metrik *Precision*, *Recall*, dan *F1-Score*. Hasil evaluasi ditampilkan dalam bentuk grafik *precision-recall*, yang menunjukkan performa sistem dalam memberikan hasil pencarian yang relevan.



Gambar 1. Kurva *Precision-Recall* KutubBot

Gambar 1 menyajikan Kurva Presisi-Recall (PR Curve) yang menggambarkan efektivitas sistem Question Answering KutubBot dalam mengidentifikasi jawaban yang relevan dari koleksi hadis Kutubut Tis'ah. Kurva menunjukkan kecenderungan positif, yaitu nilai Presisi meningkat seiring dengan meningkatnya nilai Recall. Hal ini menunjukkan bahwa model BERT yang digunakan mampu menjaga kualitas jawaban yang dihasilkan tanpa mengurangi kemampuan sistem dalam menemukan informasi yang sesuai. Dengan kata lain, sistem dapat memberikan jawaban yang tepat sekaligus mencakup informasi yang dibutuhkan pengguna secara lebih luas.

Titik performa optimal pada grafik terlihat pada koordinat sekitar (0,87 ; 0,86) yang merepresentasikan pencapaian terbaik model selama proses evaluasi. Berdasarkan hasil penelitian, sistem memperoleh nilai F1-score sebesar 86,3%, yang divisualisasikan melalui garis ambang merah putus-putus pada grafik. Nilai F1-score merupakan rata-rata harmonik antara Presisi dan Recall, sehingga menunjukkan bahwa KutubBot memiliki keseimbangan yang sangat baik antara ketepatan jawaban dan kelengkapan informasi. Selain itu, sistem juga memperoleh nilai Exact Match (EM) sebesar 78,6%, yang menunjukkan bahwa sebagian besar jawaban yang diberikan model identik secara tepat dengan jawaban referensi pada dataset.

Keberadaan garis merah yang berpotongan di bagian akhir kurva menandakan ambang batas performa terbaik sistem. Dengan capaian F1-score 86,3%, dapat disimpulkan bahwa KutubBot mampu memanfaatkan teknologi Natural Language Processing (NLP) berbasis Transformer secara efektif dalam konteks pembelajaran hadis. Sistem tidak hanya bekerja berdasarkan pencocokan kata kunci seperti metode konvensional, tetapi juga mampu memahami konteks pertanyaan pengguna dan mengekstraksi jawaban yang relevan dari teks hadis dengan tingkat akurasi yang tinggi. Oleh karena itu, KutubBot berpotensi menjadi media pembelajaran hadis yang interaktif, efisien, dan mudah diakses oleh pengguna.

KESIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan teknologi Natural Language Processing (NLP) berbasis Transformer melalui model BERT berhasil diimplementasikan pada sistem Question Answering KutubBot untuk koleksi hadis Kutubut Tis'ah. Sistem yang dikembangkan mampu memahami pertanyaan pengguna dalam bahasa alami dan mengekstraksi jawaban yang relevan secara otomatis dari teks hadis. Hal ini menunjukkan bahwa pendekatan Transformer efektif digunakan dalam pengolahan teks keagamaan, khususnya untuk mendukung pembelajaran hadis secara digital. Hasil evaluasi menunjukkan performa yang baik dengan nilai Exact Match (EM) sebesar 78,6% dan F1-score sebesar 86,3%. Nilai tersebut menandakan bahwa sistem memiliki kemampuan tinggi dalam memberikan jawaban yang akurat, relevan, dan sesuai konteks pertanyaan. Selain itu, kurva Precision-Recall memperlihatkan keseimbangan yang baik antara ketepatan jawaban dan kemampuan sistem dalam menemukan informasi yang dibutuhkan pengguna. Dengan demikian, KutubBot terbukti lebih unggul dibandingkan metode pencarian konvensional berbasis pencocokan kata kunci. Secara praktis, sistem KutubBot berpotensi menjadi media pembelajaran hadis yang interaktif, efisien, dan mudah diakses oleh pelajar, mahasiswa, pendidik, maupun masyarakat umum. Pengguna dapat memperoleh informasi hadis secara cepat tanpa harus melakukan pencarian manual pada banyak kitab. Meskipun demikian, kinerja sistem masih dipengaruhi oleh kualitas dataset, proses anotasi data, dan keterbatasan pemahaman konteks bahasa Arab maupun terjemahan. Oleh karena itu, penelitian selanjutnya

disarankan melakukan penambahan dataset, optimasi model, serta integrasi fitur penjelasan makna hadis agar sistem semakin akurat dan bermanfaat.

DAFTAR PUSTAKA

- Abdallah, A., Piryani, B., & Jatowt, A. (2023). Exploring the state of the art in legal QA systems. *Journal of Big Data*. <https://doi.org/10.1186/s40537-023-00802-8>
- Afabih, A., Roziqi, A., & Maulana, A. (2025). *The Impact Of Scholar ' S Differences In Determining The Status Of Hadith On The Law*. 6(1), 125–152. <https://doi.org/10.55987/njhs.v6i1.219>
- Ahmed, M., Khan, H., Iqbal, T., & Alarfaj, F. K. (2023). *On solving textual ambiguities and semantic vagueness in MRC based question answering using generative pre-trained transformers*. 1–31. <https://doi.org/10.7717/peerj-cs.1422>
- Alsubhi, K., Jamal, A., & Alhothali, A. (2022). *Deep learning-based approach for Arabic open domain question answering*. <https://doi.org/10.7717/peerj-cs.952>
- Ananda, F., Rakha, G., Ardiansyah, R., & Syirajuddin, M. (2025). *Design and Development of a Soccer Shoes Recommendation Application Using NLP Model Implementation and Content-Based Filtering*. 15(3), 300–307. <https://jurnal.polinema.ac.id/index.php/jartel/article/view/6965>
- Kim, Y., Bang, S., Sohn, J., & Kim, H. (2022). Question answering method for infrastructure damage information retrieval from textual data using bidirectional encoder representations from transformers. *Automation in Construction*, 134(October 2021), 104061. <https://doi.org/10.1016/j.autcon.2021.104061>
- Maxutova, K. (2025). *Development of a Hybrid Span-QA Model With Ontology Integration for Semantic Enrichment of Answers*. August. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=11159216>
- Mitchell, J. R., Szepietowski, P., Howard, R., Reisman, P., Jones, D., Lewis, P., Fridley, B. L., & Rollison, D. E. (n.d.). *A Question-and-Answer System to Extract Data From Free-Text Oncological Pathology Reports (CancerBERT Network): Development Study Corresponding Author: 24*. <https://doi.org/10.2196/27210>
- Ni'mah, A. T., & Yunitarini, R. (2024). Relevance of the Retrieval of Hadith Information (RoHI) using Bidirectional Encoder Representations from Transformers (BERT) in religious education media. *BIO Web of Conferences*, 146. <https://doi.org/10.1051/bioconf/202414601041>
- Sartika, D., S, M. N., Putri, S., & Mulyani, R. (2025). *Prevention Strategies for Cyberbullying Based on Hadiths: A Thematic Analysis of Hadiths Pertaining to Ethical Communication*. 49(2), 165–182. <https://doi.org/10.24014/an-nida.v49i2.38192>
- Shao, T. (2019). *Collaborative Learning for Answer Selection in Question Answering*. 7, 7337–7347. <https://ieeexplore.ieee.org/abstract/document/8648373>
- Shen, K., & Id, M. K. (2023). *Quantifying confidence shifts in a BERT-based question answering system evaluated on perturbed instances*. 1–21. <https://doi.org/10.1371/journal.pone.0295925>

- Souza, F. C., Nogueira, R. F., & Lotufo, R. A. (2023). BERT models for Brazilian Portuguese: Pretraining, evaluation and tokenization analysis. *Applied Soft Computing*, 149(PA), 110901. <https://doi.org/10.1016/j.asoc.2023.110901>
- Tsalitsatun, A., Yahya, A., Winata, S., Ramansyah, W., Nawawi, S., Yuanova, R. A., Mansur, A. M., & Muhammed, A. (2025). *Implementation of a transformer-based question answering model in KutubBot for the Kutubut Tis ' ah Hadith Corpus*. 01036. https://www.epj-conferences.org/articles/epjconf/abs/2025/29/epjconf_aiptec2025_01036/epjconf_aiptec2025_01036.html
- Tsalitsatun, A., Yunitarini, R., & Nawawi, S. (2024). *Pembelajaran Pendidikan Agama Islam melalui Sistem Pencarian Artikel Berbasis Bidirectional Encoder Representations from Transformers (BERT)*. 4(2), 187–195. <https://nuris.ac.id/journal/jeis/article/view/181>
- Wang, C., Zhang, L., & Yan, W. (2023). Relation Extraction Based on BERT and BGRU in the Chinese Music Scene. *Procedia Computer Science*, 225, 2429–2438. <https://doi.org/10.1016/j.procs.2023.10.234>
- Yang, J., Yang, X., Li, R., Luo, M., Jiang, S., Zhang, Y., & Wang, D. (2023). BERT and hierarchical cross attention-based question answering over bridge inspection knowledge graph. *Expert Systems with Applications*, 233(June), 120896. <https://doi.org/10.1016/j.eswa.2023.120896>
- Zhu, X., Chen, Y., Gu, Y., & Xiao, Z. (2022). *SentiMedQAer: A Transfer Model for Biomedical Question Answering*. 16(March), 1–11. <https://doi.org/10.3389/fnbot.2022.773329>