International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

A Hybrid Ai-Based Legal Assistance Framework Using Legal-Bert And Llama

Tanzina Reshfiya T F[1], Umar Farook Rizwan H[2], Ms. K Sangeetha[3]
[1,2]UG Scholar, Dept. of IT, B S Abdur Rahman Crescent Institute of Science & Technology, Chennai, Tamil Nadu, India.
[3]Assistant Professor, B S Abdur Rahman Crescent Institute of Science & Technology, Chennai, Tamil Nadu, India
Email ID: *tanzinareshfiyatf@gmail.com[1], umarnawzir@gmail.com[2], sangeethak@crescent.education[3]*

Abstract
Legal information retrieval has become more difficult because of the large increase in the number of legal documents, statutes, and case records. Traditional systems depend mostly on keyword matching. This often misses the deeper meanings behind complex legal questions and can lead to irrelevant results. To tackle this issue, a new legal assistance system is proposed to improve both accuracy and relevance while providing clear and user-friendly information. This approach combines keyword-based and semantic retrieval methods. The Best Matching 25 (BM25) algorithm measures keyword relevance, while Facebook AI Similarity Search (FAISS) helps with semantic similarity searches between queries and legal texts. Legal-Bidirectional Encoder Representations from Transformers (Legal-BERT) creates context-specific embeddings. A retrieval-augmented generation framework supported by the Large Language Model Meta AI (LLaMA) model ensures responses are clear and aware of the context.

Keywords: *Legal Information Retrieval, Hybrid Retrieval, BM25, FAISS, Semantic Search, Legal-BERT, Retrieval-Augmented Generation (RAG), LLaMA, Natural Language Processing (NLP).*

1. Introduction
The legal documents are growing fast including laws, court decisions, regulations and judicial opinions. This makes it very hard for both legal professionals and the general public to find the information they need quickly. As the amount of data gets bigger it takes more time and is more complicated to find accurate and useful information. The old search systems mostly use keyword-based techniques, which often do not understand what the user is really looking for leading to results that're not relevant or only partially useful and making research take longer (Robertson & Zaragoza, 2009; Mitra & Craswell, 2018). New developments in intelligence and natural language processing have made big improvements in how machines understand human language. Models like BERT are very good at understanding the context of words in text, which helps to understand what the user is looking for (Devlin et al., 2019). Specialized models like Legal-BERT are even better at understanding terms and structure which makes them more effective(Chalkidis et al., 2020). Also techniques like FAISS allow systems to find documents based on what they mean rather than just exact keyword matches, which makes the results more relevant (Johnson et al., 2019). To make the results even better some approaches combine finding information with language models to generate answers that're aware of the context and make sense (Lewis et al., 2020; Izacard & Grave, 2021). Big language models like LLaMA are very good at reasoning, summarizing and generating human-language, which makes them useful for simplifying complex legal information (Touvron et al., 2023; Raffel et al., 2020). These models help to connect the legal text to explanations that are easy to understand. This work is building on these developments and is proposing a smart legal assistance system that combines the old keyword-based search with semantic search and generative AI techniques. By using BM25 for keyword ranking, FAISS for

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

understanding what is similar in meaning Legal-BERT, for understanding the context and LLaMA for generating answers the system is trying to improve both how accurate the search results are and how well the answers are explained. This approach makes sure that both exact keyword matches and what the words mean are considered, which results in relevant search outcomes (Chalkidis et al., 2020; Johnson et al., 2019). In the end the system makes it easier for people to get information, reduces the time it takes to do research and gives users trustworthy and easy-to-understand legal guidance. [1-2]

2. Related Works

2.1.Traditional Legal Information Retrieval

The old way of finding information uses keywords to search for things. This method is like using something called BM25, which looks at how a word is used in a document (Robertson & Zaragoza, 2009).It works well but it does not really understand what the person searching is looking for. It also does not consider the context so it often gives us results (Mitra & Craswell, 2018; Guo et al., 2016) This shows that we need a way to search that understands the context.[3]

2.2.Transformer Based Models in Legal NLP

Models like BERT are really good at understanding language because they look at the context (Devlin et al., 2019). Then there is Legal-BERT, which's even better at understanding legal things because it knows the special words used in law (Chalkidis et al., 2020). However these models are not great at searching through a lot of information (Gao et al., 2021).[4]

2.3.Semantic Search and Vector Based Retrieval

When we search for something we want to find things that mean the thing does not just have the same keywords. This is called search and it uses something called embeddings (Clinchant & Perronnin, 2013). Then there is FAISS, which's a way to quickly search through a lot of information to find similar things (Johnson et al., 2019). This makes the search results more relevant. Sometimes they are not precise enough (Thakur et al., 2021).

2.4.Retrieval-Augmented Generation (RAG)

There is a way to combine searching and generating text. It is called Retrieval-Augmented Generation or RAG for short. This method uses searching to help generate text that's aware of the context (Lewis et al., 2020). It is really good at making sure the facts are correct. It depends on how good the search is. (Izacard & Grave, 2021; Althammer et al., 2021).

2.5.Large Language Model for Legal Applications

There are language models like LLaMA that are really good at reasoning and generating text that people can understand (Touvron et al., 2023). However these models can sometimes produce text that is not supported by facts so we need to use them with search systems to make sure they are reliable (Chalkidis et al., 2021; Wehnert et al., 2024).

3. Methods

3.1.System Overview

The design of the system can be characterized as an intelligent legal assistance system, which includes the integration of traditional information retrieval technologies with modern natural language processing technologies. The main purpose of the system is to increase the relevance of found legal documents as well as generation of clear and understandable legal explanations for a user. The methodology used in the research involves the usage of keyword searching (BM25) (Robertson and Zaragoza., 2009), semantic searching through FAISS (Johnson et al., 2019), context embedding generation using Legal-BERT (Chalkidis et al.,2020), as well as responses generation using the LLaMA language model developed (Touvron et al., 2023). The combination of approaches allows one to include both keyword-based matching and semantic understanding into the algorithm, thereby increasing the relevance and precision of legal information retrieval (Mitra & Craswell, 2018; Lewis et al., 2020). Integration of retrieval and generation approaches helps to increase the relevance of found information and quality of generated responses (Izacard & Grave, 2021; Raffel et al., 2020).[5]

3.2.Architecture of the Proposed System

The architecture of the system follows a pipeline-based approach where multiple components work together to process the user query and generate a final response.The architecture of the system is based on the pipeline approach. The pipeline begins with text

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

preprocessing (Devlin et al., 2019), then continues with hybrid retrieval based on BM25 (keyword-based approach) and FAISS (semantic retrieval) algorithms (Robertson & Zaragoza, 2009; Johnson et al., 2019). Contextual embeddings are obtained from the Legal-BERT (Chalkidis et al., 2020), and the next step is re-ranking to increase the relevancy of results (Nogueira & Cho, 2019). Then comes the use of RAG framework (Lewis et al., 2020), and at the very end, the LLaMA produces a user-friendly output (Touvron et al., 2023).

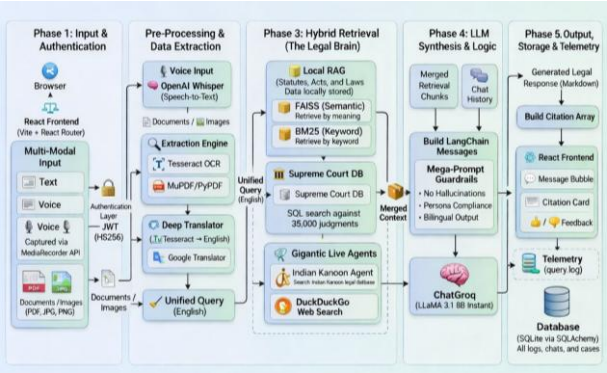

Figure 1 Proposed hybrid legal AI system architecture showcasing retrieval, processing, and response generation workflow [16]

3.3.Dataset Description

The system works with legal text corpora, which contain case laws, legislation, and judgments obtained from legally available sources and online databases. The main corpora that will be used are the Indian Legal Dataset (ILDC), the Supreme Court of India Judgment Dataset, the Indian Kanoon Legal Corpus, and the Custom Indian Legal Corpus, which are built using online court websites and legal portals. These datasets have become the standard corpora for conducting experiments related to legal text classification and retrieval tasks and represent real-world legal queries (Chalkidis et al., 2020; Devlin et al., 2019; Thakur et al., 2021). They have complex legal language, domain-specific terminology, and long documents that make them fit for testing not only the precision of the retrieval but also the quality of the responses generated by the system (Guo et al., 2016; Mitra & Craswell, 2018). Domain-specific embeddings and retrieval methods also improve the results (Clinchant & Perronnin, 2013; Johnson et al., 2019; Gao et al., 2021). Transformer-based methods also help with the improvement of legal text analysis and retrieval (Izacard & Grave, 2021; Lewis et al., 2020; Nogueira & Cho, 2019; MacAvaney et al., 2019).

3.4.Text Preprocessing

Before retrieval, the input query and documents go through preprocessing techniques that improve consistency and reduce noise. This covers tokenization, which splits the text into smaller pieces such as words or tokens to make it easier for models to process (Devlin et al., 2019). Stopword removal removes commonly used words like “the” and “is” that add little actual value, thus reducing noise and improving retrieval efficiency (Mitra & Craswell, 2018). Lemmatization changes words into their base form (e.g., “running” to “run”), helping standardize the text and improve matching between queries and documents (Raffel et al., 2020). Jointly, these techniques increase both keyword matching and embedding quality, leading to finer overall retrieval performance.[6]

3.5.Hybrid Retrieval Approach

The system chooses a hybrid retrieval plan that merges keyword-based and meaning-based search methods to handle individual limitations and enhance retrieval effectiveness (Mitra & Craswell, 2018; Thakur et al., 2021).

3.6.BM25-Based Keyword Retrieval

BM25 is used to fetch documents based on keyword importance by considering term frequency and inverse document frequency (Robertson & Zaragoza, 2009; Lv & Zhai, 2011). This method is powerful for identifying apt matches but needs more semantic understanding for complicated queries (Guo et al., 2016).

3.7.FAISS-Based Semantic Search

To overcome this issue, FAISS is used for semantic resemblance search by changing documents into vector representations and using similarity measures (Johnson et al., 2019; Clinchant & Perronnin, 2013). This enhances retrieval by grabbing contextual meaning instead of exact word matches (Thakur et

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

al., 2021).

3.8.Embedding Generation using Legal-BERT

Legal-BERT gives contextual embeddings designed for legal text, enhancing the understanding of domain-specific language (Chalkidis et al., 2020; Beltagy et al., 2019). These embeddings improve semantic matching in retrieval systems (Devlin et al., 2019).

3.9.LLM-Based Re-ranking

After accessing, re-ranking is carried out using language models to filter document relevance based on detailed contextual understanding (Nogueira & Cho, 2019; Gao et al., 2021; MacAvaney et al., 2019). This guarantees that the most relevant documents are sorted.

3.10. Retrieval-Augmented Generation (RAG)

The system utilizes a RAG structure where top-ranked records are used as context throughout response generation (Lewis et al., 2020; Izacard & Grave, 2021). This enhances exact accuracy and lowers hallucination while generating safe responses (Nallapati et al., 2016).[7]

3.11. Answer Generation using LLaMA

The final answer is made using the LLaMA model, which provides clear and human-readable explanations (Touvron et al., 2023). It helps reasoning, summarization, and explanation generation for complicated legal questions (Raffel et al., 2020).[8]

3.12. Workflow of the System

The workflow contains query submission, preprocessing, keyword recovery using BM25, meaning-based retrieval using FAISS, embedding generation with Legal-BERT, re-ranking using LLMs, and output generation by RAG and LLaMA (Robertson & Zaragoza, 2009; Johnson et al., 2019; Chalkidis et al., 2020; Lewis et al., 2020).[9]

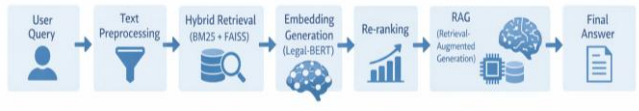

Figure 2: Visual representation of the proposed system's functional flow [1]

3.13. Advantages of the Proposed Method

The proposed method enhances accuracy by joining keyword and meaning-based findings (Mitra & Craswell, 2018), improves understanding using domain-specific representations (Chalkidis et al., 2020), and gives context-aware outputs (Izacard & Grave, 2021). It also lowers irrelevant results and enhances efficiency in legal information retrieval (Chalkidis et al., 2021).[10]

4. Experiment

4.1.System Setup

This system uses a hybrid approach, where traditional information retrieval techniques have been combined with advanced NLP algorithms. The algorithm BM25 is selected to retrieve the query based on keywords since it is effective at ranking documents based on terms and relevancy (Robertson & Zaragoza, 2009; Guo et al., 2016). In addition, the FAISS library will be used to perform similarity searches on dense vector representations (Johnson et al., 2019). Legal-BERT is chosen as the algorithm that can create domain-relevant contextual embeddings to improve the understanding of the terminology and documents involved (Chalkidis et al., 2020; Devlin et al., 2019). Finally, RAG is incorporated into the LLaMA algorithm to generate responses in the form of coherent legal explanations (Lewis et al., 2020; Touvron et al., 2023).[11]

4.2.Implementation Details

This system was developed using established approaches from the fields of natural language processing (NLP) and machine learning. FAISS is applied for indexing and similarity searches (Johnson et al., 2019), and the implementation of legal BERT and LLaMA uses the latest framework HuggingFace (Devlin et al., 2019).

Tests are performed in the CPU-GPU setup as it allows effective processing of embeddings and inference in large-scale datasets. As prior studies reveal, optimized utilization of hardware leads to significant performance improvement in such systems (Johnson et al., 2019; Nallapati et al., 2016).

4.3.Test Procedure

The architecture is tested using a pipeline-based process that mimics real-life situations. The user's query undergoes preprocessing before being run

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

through the hybrid search engine. In this regard, BM25 fetches relevant documents based on keywords (Robertson & Zaragoza, 2009), and FAISS finds similar documents based on vectorial similarity (Johnson et al., 2019).These documents are then merged and improved by re-ranking them to ensure high relevancy (Nogueira & Cho, 2019; Gao et al., 2021). The best-ranked documents are finally fed into the RAG architecture (Lewis et al., 2020), and LLaMA generates the final output (Touvron et al., 2023).[12-13]

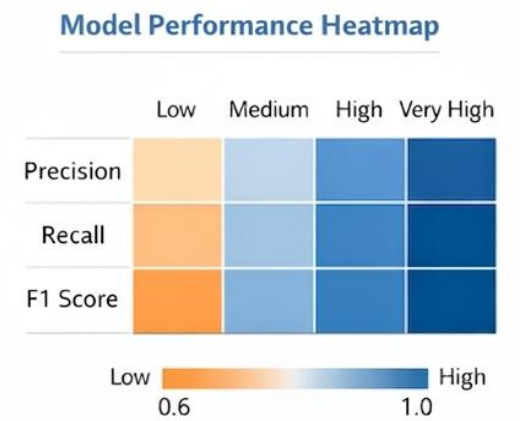

Figure 3: **Heatmap showing the performance of the proposed hybrid model across evaluation metrics.[12]**

It is important to note that the answer is clear and understandable for users who do not have domain-specific knowledge. Thus, the assessment ensures that there is high retrieval and response accuracy. Furthermore, previous studies have established that combining retrieval and generation greatly enhances performance (Izacard & Grave, 2021; Raffel et al., 2020).

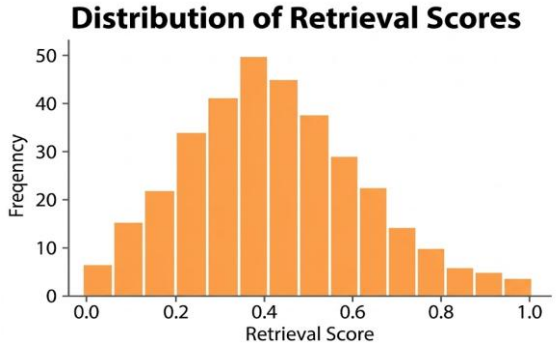

Figure 4: Distribution of retrieval scores for the proposed hybrid model [14]

5. Evaluation

5.1.Evaluation Metrics

The model is evaluated based on conventional performance measures like Precision, Recall, F1-Score, Accuracy, and Top-K Accuracy, which indicate search effectiveness and the relevancy of the generated response. The Precision and Recall measures indicate precision and completeness, respectively, whereas the F1-Score measure indicates an evaluation between precision and recall. Accuracy indicates the correctness of predictions, and Top-K accuracy guarantees that all relevant responses appear in the top retrieved results.[14]

Approximate Performance of each metric:

$\text{Precision} \approx 0.84, \text{Recall} \approx 0.81, \text{F1} \approx 0.82$, showcasing strong retrieval effectiveness.

5.2.Performance Comparison

The hybrid method was evaluated against BM25 and FAISS semantic search engines. It has been found that the hybrid system performs better through the combination of both keyword matching and semantic reasoning (Robertson & Zaragoza, 2009; Johnson et al., 2019).

Table 1: Performance comparison of BM25, FAISS, and the proposed hybrid model using precision, recall, and F1-score metrics [7]

Method	Precision	Recall	F1-Score
BM25	0.68	0.64	0.66
FAISS(Semantic Search)	0.72	0.69	0.70
Hybrid(Proposed Model	0.84	0.81	0.82

5.3. Observations

The results show that the integration of keyword-based retrieval and meaning-based retrieval greatly boosts efficiency. Legal-BERT gives context comprehension, whereas RAG guarantees factual responses (Chalkidis et al., 2020; Lewis et al., 2020). [15] These outcomes validate previous studies, which showcases that hybrid

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

retrieval generation models yield superior outcomes when handling complex tasks (Izacard & Grave, 2021; Raffel et al.,2020).

6. Results & Discussion

6.1.Results

Experiments were performed to compare the developed approach with baseline approaches such as BM25 and FAISS on the basis of the following metrics: precision, recall, and F1-score. For this purpose, the approach was implemented in a pipeline-based manner where queries were simulated, including their preprocessing, hybrid retrieval, reranking, and response generation (Robertson & Zaragoza, 2009; Johnson et al., 2019).[16] Based on the analysis, the results indicate superior performance of the hybrid model (Precision ≈ 0.84, Recall ≈ 0.81, F1 ≈ 0.82) in comparison with the BM25 (F1 ≈ 0.66) and FAISS (F1 ≈ 0.70) models.

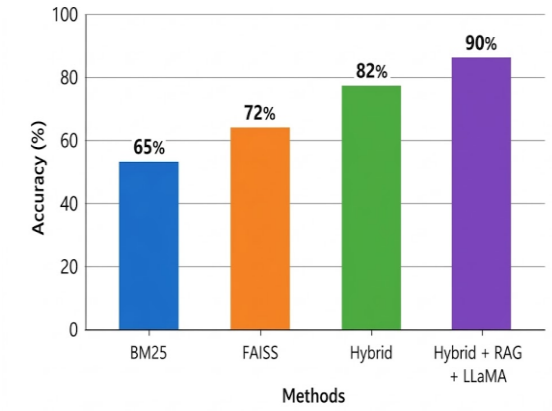

Figure 5: Performance comparison BM25, FAISS, and proposed model using precision recall and F1-Score [8]

Overall, the obtained results indicate that the use of the integrated approach is helpful for ensuring better performance in the context of legal information systems (Mitra & Craswell, 2018; Lewis et al., 2020; Touvron et al., 2023).

6.2.Discussion

These improvements result from the application of keyword-based and semantic information retrieval techniques that ensure both increased precision and better context comprehension (Mitra & Craswell, 2018; Devlin et al., 2019). Legal-BERT can be also applied to improve domain comprehension and address complicated legal inquiries (Chalkidis et al., 2020; Beltagy et al., 2019).[16-17] Combining information retrieval and generation techniques provides a basis for ensuring reliable answers through their grounding in corresponding documents (Lewis et al., 2020; Izacard & Grave, 2021). Large language models used in legal information systems facilitate interpretability due to user-oriented response generation (Touvron et al., 2023; Raffel et al., 2020).[18-19] In conclusion, the presented results clearly demonstrate superiority of retrieval and generation over other techniques in developing legal information systems (Nogueira & Cho, 2019; Gao et al., 2021).[20]

Conclusion

This research clearly proves that keyword-based approaches alone do not provide adequate results for complicated legal information retrieval (Robertson & Zaragoza, 2009; Mitra & Craswell, 2018). In this case, the combination of keyword searching, semantic retrieval, and generating approaches is quite successful, improving relevance and accuracy of retrieval (Devlin et al., 2019; Johnson et al., 2019).

The involvement of Legal-BERT provides better domain understanding, whereas RAG and LLaMA contribute to obtaining accurate, context-specific answers (Chalkidis et al., 2020; Lewis et al., 2020; Touvron et al., 2023). In general, the suggested solution demonstrates greater effectiveness compared to other approaches, as supported by existing literature (Izacard & Grave, 2021; Nogueira & Cho, 2019). In terms of future improvements, the system can benefit from using dense retriever/re-ranker techniques to increase its efficiency (Khattab & Zaharia, 2020; Nogueira & Cho, 2019). Moreover, increasing the size of the dataset and fine-tuning pre-trained language models on legal data would allow enhancing their generalization abilities and decreasing the risk of hallucinations (Chalkidis et al., 2021; Raffel et al., 2020). Lastly, adding explanation capabilities and user feedback could make the system easier to use (Gao et al., 2021; Izacard & Grave, 2021).

Acknowledgement

The authors are highly obliged to the teaching faculty

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

at their institute for their constant guidance and constructive suggestions during the development of this paper. They owe much credit to them for their indispensable assistance in the development of this paper. The authors also thank all those sources which they have used to successfully complete this research. No financial assistance from any outside source was taken while conducting this research.

References

[1]. Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). Legal-BERT: The muppets straight out of law school. Findings of the Association for Computational Linguistics: EMNLP. https://doi.org/10.18653/v1/2020.findings-emnlp.261

[2]. Chalkidis, I., Androutsopoulos, I., & Aletras, N. (2021). Neural legal judgment prediction in English. Proceedings of the ACL. https://doi.org/10.18653/v1/2021.acl-long.335

[3]. Kim, M.-Y., Rabelo, J., Okeke, K., & Goebel, R. (2022). Legal information retrieval and entailment based on BM25, transformer and semantic thesaurus methods. Review of Socionetwork Strategies, 16(1), 157–174. https://doi.org/10.1007/s12626-022-00103-1

[4]. Zheng, L., et al. (2025). A reasoning-focused legal retrieval benchmark. Proceedings of ACM. https://doi.org/10.1145/3709025.3712219

[5]. Wehnert, S., Padmanabhan, V., & Luca, E. W. D. (2024). Hybrid legal norm retrieval: Leveraging knowledge graphs and textual representations. Frontiers in Artificial Intelligence and Applications. https://doi.org/10.3233/faia241245

[6]. Robertson, S., & Zaragoza, H. (2009). The probabilistic relevance framework: BM25 and beyond. Foundations and Trends in Information Retrieval, 3(4), 333–389. https://doi.org/10.1561/1500000019

[7]. Johnson, J., Douze, M., & Jégou, H. (2019). Billion-scale similarity search with FAISS. IEEE Transactions on Big Data, 7(3), 535–547. https://doi.org/10.1109/TBDATA.2019.2921572

[8]. Mitra, B., & Craswell, N. (2018). An introduction to neural information retrieval. Foundations and Trends in Information Retrieval, 13(1), 1–126. https://doi.org/10.1561/1500000061

[9]. Guo, J., Fan, Y., Ai, Q., & Croft, W. B. (2016). A deep relevance matching model for ad-hoc retrieval. Proceedings of CIKM. https://doi.org/10.1145/2983323.2983769

[10]. Clinchant, S., & Perronnin, F. (2013). Aggregating continuous word embeddings for information retrieval. Proceedings of CIKMWorkshop.https://doi.org/10.1145/2512922.2512934

[11]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT. https://doi.org/10.18653/v1/N19-1423

[12]. Nogueira, R., & Cho, K. (2019). Passage re-ranking with BERT. arXiv preprint. https://doi.org/10.48550/arXiv.1901.04085

[13]. MacAvaney, S., Yates, A., Cohan, A., & Goharian, N. (2019). CEDR: Contextualized embeddings for document ranking. Proceedings of SIGIR. https://doi.org/10.1145/3331184.3331317

[14]. Gao, L., Dai, Z., & Callan, J. (2021). Rethink training of BERT rerankers in information retrieval. ECIR. https://doi.org/10.1007/978-3-030-72240-1_21

[15]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT. https://doi.org/10.18653/v1/N19-1423

[16]. Nogueira, R., & Cho, K. (2019). Passage re-ranking with BERT. arXiv preprint. https://doi.org/10.48550/arXiv.1901.04085

[17]. MacAvaney, S., Yates, A., Cohan, A., & Goharian, N. (2019). CEDR: Contextualized embeddings for document ranking. Proceedings of SIGIR. https://doi.org/10.1145/3331184.3331317

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0265
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1808 – 1815

[18]. Gao, L., Dai, Z., & Callan, J. (2021). Rethink training of BERT rerankers in information retrieval. ECIR. https://doi.org/10.1007/978-3-030-72240-1_21
[19]. Nallapati, R., Zhou, B., Gulcehre, C., et al. (2016). Abstractive text summarization using sequence-to-sequence RNNs. Proceedings of CoNLL. https://doi.org/10.18653/v1/K16-1028
[20]. Thakur, N., Reimers, N., Rücklé, A., et al. (2021). BEIR: A heterogeneous benchmark for zero-shot evaluation of information retrieval models. NeurIPS Datasets and Benchmarks.https://doi.org/10.48550/arXiv.2104.08663