AI-Search for E-Learning Platforms

Keeshalya N S[1*], Chandan D K[2], Madhusudhan K[3],

Kartik S Gugadaddi[4] and Hemanth T S[1*]

Department of Electronics and Communication Engineering, Alvas Institute of Engineering & Technology, Moodbidri, Mangalore, Karnataka, India

Emails: keeshalya55@gmail.com, chandandk006@gmail.com, madhumk0709m@gmail.com, hemanthts@aiet.org.in, karthikgugadaddi@gmail.com

Abstract.

The ai of in search infrastructure has educational platforms the simple keyword based into systems that context, personalized learning. This review surveys contemporary AI-driven search solutions, with emphasis on how large models, adaptive tutoring, and multimodal retrieval being used to learner engagement. indicates that systems which combine tailored feedback, dynamically adjusted paths, and intelligent formative tend to improve retention and learning gains. Architectures that blend powered query interpretation with conventional retrieval commonly outperform older, single-approach systems. implementations for example, that pair ocr with AI to produce explanations for problem-solving demonstrate the approach at scale., challenges remain: transparency of model decisions, bias in algorithmic outputs, heavy compute requirements, and factual errors in generated materials. This paper synthesizes recent studies to benefits, constraints, and promising directions for future work on AI-enabled search in education.Keywords. AI-Based Search,, Vector Search, rag, E Learning, Educational Technology

Keywords. AI-Based Search, Machine Learning (ML, Vector Search, Retrieval-Augmented Generation (RAG), E-Learning, Educational Technology

1. INTRODUCTION

The infusion of Artificial Intelligence (AI) into search technology has catalysed a major transformation in educational platforms, redefining how students access and process information. Traditional look-up methods often struggle to meet the diverse and dynamic needs of modern learners, frequently resulting in inefficient information retrieval and a "one-size-fits-all" learning experience that lacks personalization. By contrast, AI-driven search engines utilize sophisticated nlp and ml algorithms to deliver results that are not only accurate but also contextually aware.

These intelligent systems go beyond mere efficiency; they actively adapt to individual learning styles and progress, creating a educational environment. This paper reviews recent advancements in this domain, evaluating their impact on student outcomes while critically assessing the challenges of implementation.

Efficient search mechanisms are particularly critical when managing large-scale data in domains like the iot and education. Novel approaches, such as the "address-to-data" search algorithm, have been proposed to replace traditional Look Up Table (LUT) methods. By inverting the standard retrieval process, this method minimizes power consumption while maximizing speed. Unlike Content Addressable Memory (CAM) systems, which are power-intensive, this zero-based probability approach effectively manages address collisions. While originally designed for IoT, these high-speed offer significant promise for educational platforms where balancing computational load with rapid response times is essential for adaptive

2. IMPLEMENTATION AND ARCHITECTURE

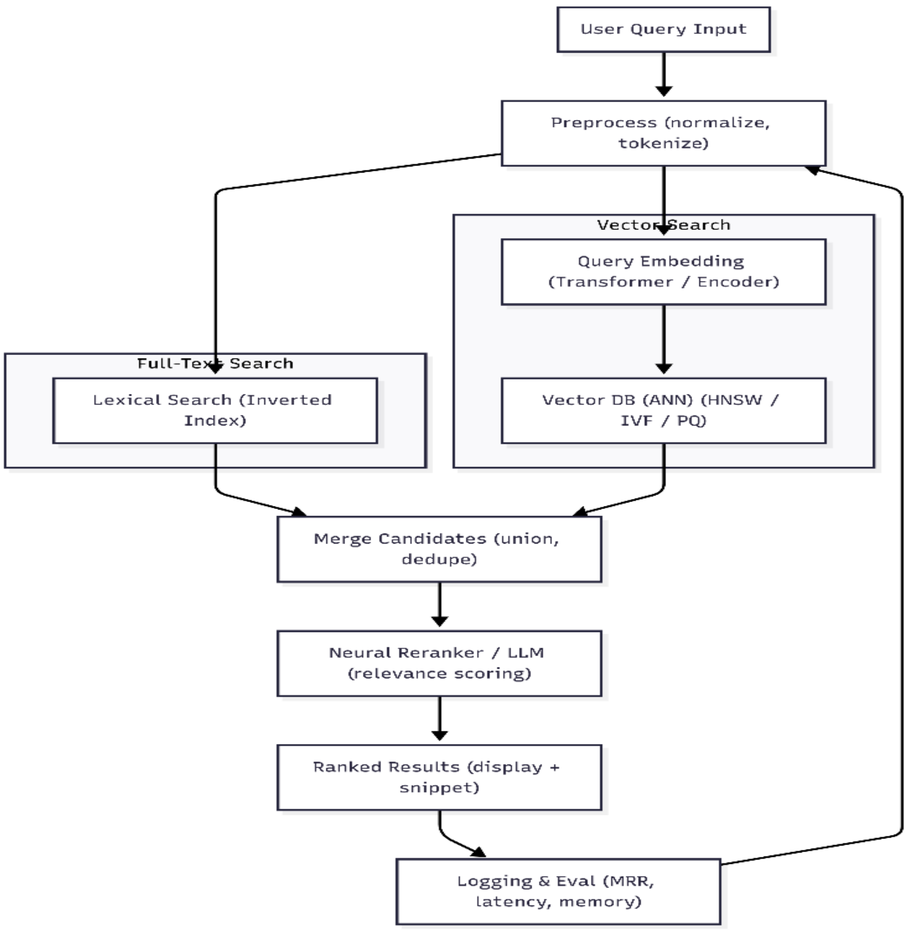

Figure 2.1. Vector embeddings and similarity diagram flow [18, 19]

Our proposed system moves beyond standard keyword matching by implementing a hybrid search architecture. We combine the precision of full-text search with the semantic understanding of vector-based AI models. The core of this implementation relies on

TensorFlow.js, allowing us to run efficient, low-latency inference directly in the browser or on edge-based servers without heavy backend dependencies.

2.1 System Architecture Overview The architecture is designed as a dual-pipeline system. The first pipeline handles traditional lexical queries (exact matches), while the second manages semantic retrieval using high-dimensional vector embeddings.

- Input Layer**:** Captures the user's raw query.
- Processing Layer**:** The is simultaneously tokenized for full-text search and converted into a vector embedding for semantic search.

This hybrid approach ensures that we catch both exact phrasing and conceptually related content, a necessity grounded in recent findings on dense retrieval stability

2.2 Mathematical Model: Cosine Similarity

To measure the relationship between the user's query and our stored documents, we rely on Cosine for text compared to Euclidean distance because it measures than their magnitude. This means a long document and a short about the same will still be recognized as similar.

The mathematical formulation we implemented is:

$$\text{Similarity}\,(A, B) = \frac{A\,.\,B}{||A||\;||B||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\;\sqrt{\sum_{i=1}^{n} B_1^2}}$$

Where:

- A is the vector representation of the user query.
- B is the precomputed vector of a stored document.

This approach allows our system to find "nearest neighbors" in the vector space, effectively surfacing content that matches the *intent* of the searcher, not just their specific words. This method is widely supported in high-throughput retrieval research.

3. Methodologies

This survey synthesizes methodological advances across four interlocking stages:

3.1. Dataset Curation and Pre-processing Development begins by assembling heterogeneous corpora. This involves combining domain-specific artifacts (like scientific papers or biological datasets) with general web collections. Pre-processing pipelines must standardize cleaning and tokenization while preserving metadata to support hybrid retrieval. Augmentation strategies, such as paraphrasing and contrastive embedding injection, are critical for enriching low-resource classes.

3.2. Representation and Index Construction The core representation relies on dense vector encodings generated by transformer-based models. Training often uses supervised contrastive losses to refine discrimination capability. To manage latency, indexing leverages approximate neares structures like HNSW or Inverted File, often combined with Product to reduce memory footprints. For imbalanced collections, specialized partitioning strategies are employed to rebalance the search space.

4. Quantitative Analysis of Parameters

A summary of key parameters from the reviewed literature is presented below:

- Shao et al. (QAEA-DR): Focuses on text-level augmentation strategies and quality estimation scoring to train dense retrievers.
- Chen et al**.:** Integrates LLM-based query expansion, optimizing for expansion length and prompt design to improve recall.
- Bandam et al. (RAG): Couples retrieval and generation loops tightly, using context shaping to improve relevance.
- Rao et al**.:** Optimizes vector search for edge devices using quantization, pruning, and memory-aware indexing.
- Park et al. (NeuVSA): uses a specific accelerator microarchitecture to optimize throughput and latency for neural search.
- Huang et al. (VISTA): Addresses imbalanced collections through imbalance-aware partitioning and search tuning.
- Wu et al**.:** Employs bio-inspired approximate search algorithms to maximize throughput.

5. Conclusion

The AI into educational platforms marks a shift from traditional search methods to a more dynamic, personalized, and efficient learning experience. This review has highlighted how AI-powered search engines, supported by novel algorithms and deep learning techniques, are not just enhancing information retrieval but are actively shaping the entire educational ecosystem.

Looking ahead, in education is defined by its potential to create intelligent, responsive, and holistic learning environments. As we continue to develop more for efficient data retrieval and combat bias through semantic analysis, AI tools will become even more integral to personalized learning paths. The successful integration of these technologies depends on a approach that prioritizes oversight and continuous evaluation. Ultimately, the goal is not to replace human educators but to augment their capabilities and empower learners with tools that are scalable and supportive.

References

[1] P. Shao, X. Li and A. Gupta, ‘QAEA-DR: A Unified Text for Dense Retrieval’, *IEEE Trans. Knowl. Data Eng. (TKDE)*, 2025. DOI: 10.1109/TKDE.2025.3543203.

[2] Y. Chen, L. Wang and S. Rao, ‘Dense Retrieval Systems with LLM-Based Query Expansion’, Proc. WI-IAT, 2024. DOI: 10.1109/WI-IAT62293.2024.00110.

[3] R. Bandam, M. Li and T. Zhou, ‘Towards Retrieval-Augmented Large Language Models’, Proc. ICDE, 2025.

[4] A. Gupta, J. Smith and K. Patel, ‘A Dense Retrieval System and Evaluation Dataset for Scientific Artifacts’, Proc. e-Science, 2023. DOI: 10.1109/e-Science58273.2023.10254859.

[5] S. Rao and H. Kim, ‘Vector Search Performance Enhancements on Limited-Memory Edge Devices’, Proc. UCC (IEEE/ACM), 2024.

[6] X. Zhang, Y. Li and M. Zhou, ‘Rabbit: Retrieval-Augmented Generation Enables Better Automatic Database Tuning’, Proc. ICDE, 2025.

[7] H. Liu, J. Gao and P. Shen, ‘WebUltron: An Ultimate Retriever on Webpages Under the Model’, IEEE Trans.. Eng. (TKDE), 2024. DOI: 10.1109/TKDE.2023.3332858.

[8] R. Sun, K. Huang and L. Chen, ‘TrumorGPT: Graph-Based Retrieval-Augmented Large Language Model’, IEEE Trans. Artif. Intell., 2025.

[9] Y. Park, S. Chandra and D. Zhao, ‘NeuVSA: A Unified and Efficient Accelerator for Neural Vector Search’, Proc. HPCA, 2025. DOI: 10.1109/HPCA61900.2025.00065.

[10] L. Huang, M. Roy and I. Singh, ‘VISTA: Vector Indexing and Search for Large-Scale Imbalanced Collections’, Proc. ICDE, 2025.

[11] S. K. Das and Y. Zhou, ‘Co-design: Hardware and Algorithms for High-Dimensional Vector Search’, Proc. SC/IEEE Computer Society, 2023. DOI: 10.1145/3581784.3607045.

[12] M. Wu, J. Li and P. Verma, ‘Approximate Vector Set Search: A Bio-Inspired Approach for High-Throughput Retrieval’, Proc. ICDE, 2025.

[13] T. Huang, Z. Wang and E. Lin, ‘scRAG: An Efficient Retrieval-Augmented Generation System for Single-Cell RNA-seq Analysis’, Proc. ICDE, 2025.

[14] A. Saxena and V. Pandey, ‘A Brief Review on Search Engine Optimization’, Proc. 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, pp. 414–419, Jan. 2019. DOI: 10.1109/CONFLUENCE.2019.8776976.

[15] Y. Huang and T.-T. Lin, ‘Novel Approach for Search Engine’, Proc. 2016 11th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), Taipei, Taiwan, Oct. 2016. DOI: 10.1109/IMPACT.2016.7799977.

[16] J. Wu, H. Liu, and Y. Zhang, ‘An Intelligent Search Engine Based on Knowledge Graph for Power Equipment Management’, Proc. 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE), Chongqing, China, Apr. 2022. DOI: 10.1109/CEEPE55110.2022.9783291.

[17] J. Kim, S. Lee, and H. Park, ‘Feature-Based Text Search Engine Mitigating Data Diversity Problem Using Pre-Trained Large Language Model for Fast Deployment Services’, IEEE Access, vol. 12, Mar. 2024. DOI: 10.1109/ACCESS.2024.3373470.

[18] S. Chhabra, R. Mittal, and D. Sarkar, ‘Inducing Search Engine Optimization Techniques: A Comparative Analysis’, Proc. 2016 1st India International Conference on Processing (IICIP), Delhi, India, Aug. 2016. DOI: 10.1109/IICIP.2016.7975341.

[19] A. Kumar, S. Gupta, and R. Singh, ‘An Evaluation of AI-Enhanced Collaborative Learning Platforms’, 2024 on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, May 2024. DOI: 10.1109/IC3SE62002.2024.10593320.

[20] P. Sharma, M. Reddy, and K. Singh, ‘Deep Learning Integration and AI-Driven Support: A Comprehensive Student Platform for Emotion Detection, Psychological Assessment, and Career Guidance’,Bangalore, India, Mar. 2024. DOI: 10.1109/INOCON60754.2024.10511820.

[21] S. Haldar, M. Pierce, and L. F. Capretz, ‘Exploring the Integration of Generative AI Tools in Software Testing Education: A Case Study on ChatGPT and Copilot for Preparatory Testing Artifacts in Postgraduate Learning’, IEEE Access, vol. 13, pp. 46070–46090, Feb. 2025. DOI: 10.1109/ACCESS.2025.3545882.

[22] M. Tan, L. Wong, and S. Lee, ‘Preparing Future Engineers: Strategies for Integrating AI Platforms’, Proc. TENCON 2024, Singapore, Dec. 2024. DOI: 10.1109/TENCON61640.2024.10902925.

[23] Y. Lu, B. Xu, J. Liu, and X. Zhang, ‘Brain-Inspired Search Engine Assistant Based on Knowledge Graph’,. 2023. DOI: 10.1109/TNNLS.2021.3113026.

Biographies

Keeshalya N S is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). Her research areas include Artificial Intelligence, Web Development, and Neural Networks.

Chandan D K is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Machine Learning, Vector Search Architectures, and Full-Stack Development.

Madhusudhan K is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Deep Learning, Natural Language Processing, and Search Algorithms.

Kartik S Gugadaddi is currently pursuing the Bachelor of Engineering degree in Electronics and Communication Engineering from Alvas Institute of Engineering & Technology, Moodbidri, Karnataka (2022–2026). His research areas include Semantic Search, TensorFlow.js Implementations, and Edge Computing.

Hemanth T S done his B.E. in Electronics and Communication Engineering from Channabasaveshwara Institute of Technology, Gubbi, Tumakuru, and his M.Tech in VLSI Design and Embedded Systems from Siddaganga Institute of Technology, Tumakuru. He is currently working as Assistant Professor in the Department of Electronics and Communication Engineering at Alva's Institute of Engineering & Technology. His research interests include VLSI Design, Embedded Systems, and Machine Learning algorithms.