




Article

A Comparative Analysis of Sentence Transformer Models for Automated Journal Recommendation Using PubMed Metadata

Maria Teresa Colangelo ¹, Marco Meleti ², Stefano Guizzardi ¹ , Elena Calciolari ^{2,3}  and Carlo Galli ^{1,*} 

¹ Histology and Embryology Laboratory, Department of Medicine and Surgery, University of Parma, Via Volturno 39, 43126 Parma, Italy; mariateresa.colangelo@unipr.it (M.T.C.); stefano.guizzardi@unipr.it (S.G.)

² Department of Medicine and Surgery, Dental School, University of Parma, 43126 Parma, Italy; marco.meleti@unipr.it (M.M.); elena.calciolari@unipr.it (E.C.)

³ Centre for Oral Clinical Research, Institute of Dentistry, Faculty of Medicine and Dentistry, Queen Mary University of London, London E1 2AD, UK

* Correspondence: carlo.galli@unipr.it

Abstract: We present an automated journal recommendation pipeline designed to evaluate the performance of five Sentence Transformer models—all-mpnet-base-v2 (MpNet), all-MiniLM-L6-v2 (Minilm-l6), all-MiniLM-L12-v2 (Minilm-l12), multi-qa-distilbert-cos-v1 (Multi-qa-distilbert), and all-distilroberta-v1 (roberta)—for recommending journals aligned with a manuscript's thematic scope. The pipeline extracts domain-relevant keywords from a manuscript via KeyBERT, retrieves potentially related articles from PubMed, and encodes both the test manuscript and retrieved articles into high-dimensional embeddings. By computing cosine similarity, it ranks relevant journals based on thematic overlap. Evaluations on 50 test articles highlight mpnet's strong performance (mean similarity score 0.71 ± 0.04), albeit with higher computational demands. Minilm-l12 and minilm-l6 offer comparable precision at lower cost, while multi-qa-distilbert and roberta yield broader recommendations better suited to interdisciplinary research. These findings underscore key trade-offs among embedding models and demonstrate how they can provide interpretable, data-driven insights to guide journal selection across varied research contexts.

Keywords: automated journal recommendation; KeyBERT; PubMed search; Sentence Transformers; semantic similarity



Academic Editor: Yoichi Hayashi

Received: 15 January 2025

Revised: 21 February 2025

Accepted: 11 March 2025

Published: 13 March 2025

Citation: Colangelo, M.T.; Meleti, M.; Guizzardi, S.; Calciolari, E.; Galli, C. A Comparative Analysis of Sentence Transformer Models for Automated Journal Recommendation Using PubMed Metadata. *Big Data Cogn. Comput.* **2025**, *9*, 67. <https://doi.org/10.3390/bdcc9030067>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Selecting the right journal for a scientific manuscript remains a critical step for researchers, regardless of their experience level [1]. A misaligned journal choice can lead to delayed publication, immediate rejection, and reduced visibility within the most relevant research communities [2]. Even seasoned authors sometimes grapple with overlapping scopes and emerging journal niches, while early-career researchers or those working on cross-disciplinary topics often lack clear criteria for identifying the most suitable outlets. The continuous creation of new journals and the evolution of established ones make it even more difficult for authors to keep pace with this shifting array of publication venues [3]. These issues make the case for data-driven systems that can automate and streamline the journal selection process, thereby reducing guesswork, saving time, and potentially improving the likelihood of a favorable peer-review outcome [4].

Historically, automated journal recommendation systems have relied on traditional information retrieval and machine learning techniques [5]. Early approaches, such as

bag-of-words models and term frequency–inverse document frequency (TF–IDF [6]), primarily matched manuscripts to journal scopes based on term co-occurrences and document similarity. For instance, Wang et al. [7] applied softmax regression with TF–IDF-based feature selection to recommend computer science journals, demonstrating modest accuracy. Other studies employed n-gram classification for textual analysis [8] or combined stylistic features with collaborative filtering [9] to identify suitable journals. However, these methods often struggled to capture the semantic richness of scientific texts, especially in fields with rapidly evolving jargon or interdisciplinary overlap.

Various commercial platforms, including those offered by well-known publishing companies such as Springer and Elsevier, aim to assist authors in selecting journals. These tools typically rely on keyword-based matching or limited datasets, which can restrict their applicability across diverse research areas. Beel et al. [10] highlighted key limitations of these platforms, such as their reliance on proprietary data, lack of transparency in recommendation algorithms, and inability to effectively handle nuanced linguistic variations or semantic relationships.

The advent of Transformer-based language models, such as BERT and its variants [11,12], has revolutionized natural language processing by capturing more nuanced, context-sensitive relationships among words. Sentence Transformer models, specifically designed for semantic similarity tasks, have emerged as powerful tools for comparing scientific texts [13]. These models generate dense semantic embeddings that preserve meaningful relationships in a continuous vector space, enabling robust similarity matching even when different terms describe the same concept [14]. Applications of such embeddings range from semantic search and clustering to question answering and recommender systems [13].

In journal recommendation tasks, Transformer embeddings offer significant advantages over traditional methods. Recent studies, including Michail et al. [15], have shown that models such as Sentence Transformers can outperform earlier techniques by encoding article abstracts into high-dimensional vectors and comparing them via cosine similarity. These approaches have proven effective across a variety of domains, from computer science to biomedicine, where capturing the thematic nuances of a manuscript is essential for identifying appropriate journals [14].

Despite this potential and all the pragmatic hurdles described above, a key question remains: Which Sentence Transformer model best captures domain-specific semantic nuances for journal recommendation [16]? Existing models differ in training objectives, corpora, and parameter settings [17]. Domain-general models—such as all-mpnet-base-v2—often excel at broad semantic similarity, although they are not specifically tuned to scientific literature. Alternatively, compact models like all-MiniLM-L6-v2 promise faster runtime but may sacrifice some representational depth. Given these trade-offs, a comparative study of multiple models is needed to determine which embedding approach yields the most accurate, context-sensitive journal recommendations in a biomedical setting.

To address this issue, we propose an automated journal recommendation pipeline for the life sciences and systematically evaluate five Sentence Transformer models: all-mpnet-base-v2 [18], all-MiniLM-L6-v2 [19], all-MiniLM-L12-v2 [20], multi-qa-distilbert-cos-v1 [21], and all-distilroberta-v1 [22]. We started with test articles, which we used to extract keywords through KeyBERT [23]. We then queried PubMed [24] with these KeyBERT-derived keywords to retrieve custom corpora of potentially related articles. We then encoded the titles and abstracts of these articles and the test manuscripts into embeddings, calculated their cosine similarity scores, and generated a data-driven list of recommended journals. In addition to comparing the models' semantic alignment, we assessed their computational efficiency, domain coverage, and alignment with expert-driven expectations.

2. Materials and Methods

2.1. Data Collection

We used NCBI E-utilities to query PubMed for articles potentially relevant to each test manuscript within the 2020–2024 timeframe. Specifically, KeyBERT identified three salient keywords from the manuscript’s title and abstract (see Section 2.2), which we combined with the logical operator AND to perform domain-relevant searches. PubMed often returned up to 20,000 PMIDs for a single query; to maintain computational feasibility, we randomly sampled 5000 articles when more than 5000 PMIDs were retrieved. For each retrieved PMID, we used Biopython’s Entrez.efetch to download metadata (Title, Abstract, Journal) and stored these fields in a pandas DataFrame [25]. The test article’s own PMID was excluded to prevent data leakage.

2.2. Preprocessing and Keyword Extraction

For each test article, we extracted its title and abstract. A keyword extraction step was performed using KeyBERT to identify three salient keywords representing the article’s core content. KeyBERT internally uses pre-trained Transformer embeddings to generate vector representations of both the entire document (title + abstract) and candidate N-grams extracted from the text. It then computes cosine similarity between the document embedding and these candidate N-gram embeddings, selecting the top-scoring ones as the most representative keywords. We chose KeyBERT for its simplicity and transparency, as it readily reveals which terms best capture a manuscript’s thematic essence. These keywords guided PubMed queries to build a dynamically tailored pool corpus for each test article. The number of keywords ($n = 3$) was empirically chosen to capture the core focus of the manuscript while avoiding overly narrow or excessively broad queries.

2.3. Test Article Selection and Evaluation Set

We selected 50 test articles to provide a varied yet tractable dataset for evaluating model performance. These articles span multiple biomedical and life sciences topics, reflecting both clinical and laboratory research published between 2020 and 2024. While this choice focuses on recent scientific discourse, we acknowledge that it may exclude older publications, potentially introducing a bias if journal scopes or key terms have evolved. Additionally, random sampling from PubMed could overlook niche or less frequently indexed research.

2.4. Embedding Models

We employed five Sentence Transformer models to encode textual data into high-dimensional vector embeddings: all-mpnet-base-v2 (mpnet), all-MiniLM-L6-v2 (minilm-l6), all-MiniLM-L12-v2 (minilm-l12), multi-qa-distilbert-cos-v1 (multi-qa-distilbert), and all-distilroberta-v1 (roberta). For each article, we concatenated its title and abstract into a single string and used the model.encode() method to generate embeddings. All embeddings were computed on Google Colab GPUs.

2.5. Similarity Computation

Cosine similarity was calculated between the test article embedding and all embeddings in the pool corpus. Cosine similarity is defined as

$$\text{cosine_similarity}(u, v) = \frac{u \bullet v}{\|u\| \|v\|}$$

where u and v are high-dimensional embedding vectors. Cosine similarity is particularly well-suited for comparing sentence embeddings because it is scale-invariant, focusing on

the angle between vectors rather than their magnitudes. This property makes it robust when embeddings differ in length or magnitude but still capture similar semantic content. Moreover, cosine similarity is widely recognized and interpreted in the NLP community, with values ranging from -1 to $+1$. By contrast, distance metrics such as Euclidean or Manhattan distance can be more sensitive to absolute vector magnitudes and do not necessarily provide an intuitive measure of semantic relatedness in high-dimensional spaces. Using PyTorch's 2.6 `torch.topk` [26], we thus identified the top 10 most similar articles for each model, recording their associated PMIDs and journals.

Alternatively, articles were grouped by journal, and a mean embedding was calculated for each journal by averaging the embeddings of all articles belonging to that journal. The cosine similarity between the test article embedding and each journal's mean embedding was then computed, yielding a ranked list of journals by average thematic alignment.

2.6. Evaluation Metrics

We further quantified the diversity and concentration of recommendation scores using two complementary metrics: Shannon entropy and the Gini coefficient.

2.6.1. Shannon Entropy

Shannon entropy is widely used in information theory to measure the diversity or "spread" of a distribution [27]. We applied Shannon entropy to the cosine similarity scores of the top recommended articles for each test case, aiming to capture how evenly the model distributed its top similarity scores among different articles or journals. Formally, for a given model M and a test article t , let $\{p_i\}$ be the normalized similarity scores (so that $\sum_i p_i = 1$) of the top N recommended items, where $p_i = \frac{s_i}{\sum_j s_j}$ and s_i is the raw similarity for item i . Shannon entropy was calculated using the following formula [28]:

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

A higher entropy indicates a more even distribution of similarity scores—i.e., no single article or journal dominates—while a lower entropy suggests one or a few items have much higher similarity scores than the rest.

2.6.2. Gini Coefficient

The Gini coefficient is commonly used in economics to assess inequality within a distribution. We employed it here to measure how concentrated the similarity scores are among top-ranked articles or journals. Let $\{x_1, x_2, \dots, x_n\}$ represent the raw similarity scores of the top n recommendations, sorted in non-decreasing order. Following the procedure in Dorfman [29], we compute the Gini coefficient G as

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n \sum_{i=1}^n x_i}$$

A G value close to 0 indicates that the model assigns relatively comparable similarity scores across the top recommendations, suggesting a more uniform distribution. By contrast, a G value nearer to 1 implies that a few recommendations dominate, resulting in greater concentration. In other words, a lower Gini coefficient denotes more balanced scoring, while a higher value signals that the model focuses its similarity on a small subset of items. We use the Gini coefficient alongside Shannon entropy to gain deeper insight into how each model distributes its top similarity scores and whether it concentrates recommendations in just a few items or spreads them more evenly.

2.7. Software and Environment

All code was implemented in Python (version 3.10) within a Google Colab environment [30]. The following primary libraries were used:

- Biopython (v1.80) for PubMed queries [31];
- pandas (v1.5+) for DataFrame manipulation [25];
- numpy (v1.23+) and torch (v1.13+) for numerical and tensor operations [32,33];
- sentence-transformers (v2.2+) for encoding text with pre-trained Transformer models;
- seaborn (v0.12+) and matplotlib (v3.6+) for data visualization [34,35].

All code was executed under Google Colab T4 GPU runtime configurations, respecting NCBI E-utilities usage policies (including Entrez.email specification and rate-limiting queries).

2.8. Study Focus and Limitations

Although our pipeline can suggest thematically relevant journals, it does not claim to offer a definitive or exhaustive publication venue recommendation. We acknowledge that real-world journal selection often involves considerations beyond thematic alignment—such as impact factor, review process, and editorial scope. Our results thus solely reflect each model’s ability to identify plausible journal options based on content similarity alone, serving as a starting point for further exploration.

3. Results

3.1. Case Study

In this study, we implemented a pipeline designed to recommend journals for a single test article by dynamically generating a corpus of potentially relevant PubMed articles. Our approach comprised four major steps: (1) extracting keywords from the test article’s title and abstract using KeyBERT, (2) issuing an ad hoc PubMed query to gather a custom pool of papers, (3) encoding the resulting pool with Sentence Transformers to create high-dimensional embeddings, and (4) computing similarity between the test article embedding and the pool embeddings.

To this purpose, we set out to evaluate and compare the recommendations produced by five Sentence Transformer models: *all-mpnet-base-v2*, *all-MiniLM-L6-v2*, *all-MiniLM-L12-v2*, *multi-qa-distilbert-cos-v1*, and *all-distilroberta-v1* (Table 1).

Table 1. This table provides a summary of the main characteristics of the five different Sentence Transformer models used for the present study.

Model	Embedding Dimension	Approx. Parameter Count	Pretraining/Fine-Tuning Corpus	Pooling Strategy
mpnet	68	~110 M	Trained on a >1 B sentence dataset and large-scale general text corpora using the MPNet architecture.	Mean pooling
minilm-l6	384	~22 M	Distilled from a teacher model on a >1 B paraphrase dataset, emphasizing efficiency with fewer layers (six layers).	Mean pooling

Table 1. Cont.

Model	Embedding Dimension	Approx. Parameter Count	Pretraining/Fine-Tuning Corpus	Pooling Strategy
minilm-l12	384	~33 M	Distilled similarly to minilm-l6 but retains more layers (12 layers), trained on the same broad paraphrase/entailment dataset.	Mean pooling
multi-qa-distilbert	768	~66 M	Fine-tuned on 215 M QA datasets to produce universal question-answer embeddings.	Mean pooling
roberta	768	~82 M	Based on DistilRoBERTa-base, further tuned on a large paraphrase corpus (>1 B sentence pairs) for enhanced sentence embedding.	Mean pooling

To get a general overview of the performance of these models, we randomly chose a test article from a corpus of 10,000 scientific articles randomly retrieved from PubMed: “Problems of *Clostridium difficile* infection (CDI) in Polish healthcare units”, published in *Annals of Agricultural and Environmental Medicine (AAEM)* [36].

We began by applying KeyBERT, an unsupervised keyword extraction tool, to the article’s title and abstract. We thus identified the three most representative keywords describing the article’s core content.

Based on these three keywords—*clostridium*, *infection*, and *difficile*—we conducted a PubMed query to locate potentially relevant articles by joining the keywords with the *AND* operator. The number of keywords ($n = 3$) was empirically determined to retrieve a sufficiently large corpus without excessively restricting the query. PubMed returned 15,466 candidate articles matching the query “*clostridium AND infection AND difficile*”. For computational feasibility, we randomly selected 5000 articles from among those 15,466. We then ensured that our test article’s own PMID was not included in this pool, so as not to artificially inflate similarity scores.

Once we generated the 5000-article corpus, we extracted the title, abstract, and journal fields from each article using Biopython’s *Entrez* utilities. The goal was to measure two distinct but complementary aspects of recommendations:

1. Article-Level Similarity: For each model, we sought the top ten articles (out of those 5000) most similar to the test article and identified the journals where they were published.
2. Journal-Level Similarity: We computed the average similarity of each journal’s articles to the test article, and then ranked the journals accordingly.

To ensure consistency, each model was applied systematically. First, the model encoded all 5000 articles, concatenating their titles and abstracts into a single text input per article, and transformed the text into embeddings. Next, the same model encoded the test article’s title and abstract. By comparing each article’s embedding with the test article’s embedding using cosine similarity, we obtained a comprehensive list of similarity scores. The article with the highest similarity score was treated as the *closest article*; its journal can

be found in Table 2 (in the Top Article column). Simultaneously, we grouped articles by their respective journals and computed a mean embedding per journal by averaging the embeddings of all articles belonging to that journal. A second round of similarity checks identified the journal with the highest average similarity to the test article (column Top Avg Journal in Table 2).

Table 2. This table provides a summary of journal recommendations generated by five different Sentence Transformer models when applied to the test article. “False” indicates that the actual journal in which the test article was published is absent from the model’s top 10 recommended journals.

Model	Top Article	Top Avg Journal	Avg Journal Similarity	Journal in Top 10	Encoding Time
minilm-l12	<i>Journal of Preventive Medicine and Hygiene</i>	International journal of environmental research and public health	0.852531	False	9"
minilm-l6	<i>Journal of Hospital Infection</i>	Risk management and healthcare policy	0.848661	False	9"
mpnet	<i>International Journal of Environmental Research and Public Health</i>	International journal of environmental research and public health	0.926116	False	98"
multi-qa-distilbert	<i>European Review for Medical and Pharmacological Sciences</i>	Risk management and healthcare policy	0.858523	False	54"
roberta	<i>Journal of Hospital Infection</i>	International journal of environmental research and public health	0.886216	False	52"

A preliminary comparison of computation time shows the great disparity between mpnet, which required a minute and a half to encode all the 5000 titles and abstracts of the dataset, and the minilm-l6 and minilm-l12 models, which only took 9 s to encode all 5000 titles and abstracts.

3.1.1. Article-Level Findings

Each model produced a list of ten articles deemed most semantically similar to the test article. A list of the journals where these articles appeared can be found in Supplementary Table S1. In general, all models identified articles published in journals that focused on infectious diseases, hospital-acquired infections, or microbial pathogens, confirming that the basic textual content was indeed guiding the retrieval. As shown in Table S1, mpnet consistently produced the highest similarity scores across its top ten recommendations, with articles often exceeding 0.85, and multiple entries from journals such as the *International Journal of Environmental Research and Public Health* and the *Journal of Preventive Medicine and Hygiene*. “Hospital management of *Clostridium difficile* infection: a review of the literature” [37] was mpnet’s suggestion as the closest article, with similarity = 90.62. The overall data suggest mpnet’s robustness in identifying journals strongly aligned with the test article’s thematic content, particularly those related to public health and infection control. Minilm-l6 and minilm-l12 exhibited overlapping trends in their recommenda-

tions, with journals such as the *Journal of Hospital Infection* and the *International Journal of Environmental Research and Public Health* frequently appearing in their top lists. However, minilm-l12 (top choice: “Clostridium difficile infection perceptions and practices: a multicenter qualitative study in South Africa” [38]) often returned higher maximum similarity scores than minilm-l6 (top choice: “The Burden of Clostridium Difficile (CDI) Infection in Hospitals, in Denmark, Finland, Norway And Sweden” [39]) for the same journal, suggesting that the additional model complexity of minilm-l12 might provide a marginally better thematic alignment.

The multi-qa-distilbert model mostly aligned with the results of the other models (top choice: “Ten-year review of *Clostridium difficile* infection in acute care hospitals in the USA, 2005–2014” [40]), although it also diverged from them by identifying broader healthcare-related journals, such as the *European Review for Medical and Pharmacological Sciences* and the *Bulletin of Mathematical Biology*. Despite this, multi-qa-distilbert consistently included journals like the *Journal of Hospital Infection*, indicating its capacity to recognize core infectious disease topics within its broader thematic spectrum. The roberta model demonstrated a strong performance, particularly with journals addressing clinical microbiology and infection control, such as the *Journal of Hospital Infection* (87.07 similarity) and the *International Journal of Environmental Research and Public Health*. Noticeably, roberta identified the same top choice article as mpnet: “Hospital management of *Clostridium difficile* infection: a review of the literature” [37].

3.1.2. Journal-Level Suggestions

To get a broader view of the dataset, we calculated the average similarity of each journal’s articles to the test article. For every journal in the 5000-article pool, we averaged all embeddings belonging to that journal, then measured how close that average embedding was to the test article’s embedding. The results can be found in Supplementary Table S2. Mpnet consistently identified journals strongly aligned with infectious diseases, hospital-acquired infections, and antimicrobial resistance. Its top-ranked journals, such as the *International Journal of Environmental Research and Public Health* and the *Journal of Preventive Medicine and Hygiene*, demonstrated high average similarity scores, with values exceeding 92% for the leading journal, and over 86% similarity for *Infection, Disease & Health*, which ranks 10th and last for this model. This pattern confirms mpnet’s capability to prioritize thematically congruent journals with precise alignment.

Similarly, minilm-l12 showed a strong (but lower than mpnet’s and always lower than 86%) affinity for journals relevant to infection control and epidemiology, with its highest-ranked journal also being the *International Journal of Environmental Research and Public Health*, with 85.2% similarity. Notable overlap was observed between mpnet and minilm-l12, particularly in their top journal selections, which included titles such as the *International Journal of Environmental Research and Public Health*, the *Journal of Preventive Medicine and Hygiene*, the *American Journal of Infection Control* and *Journal of Hospital Infection*. This confirms that these models share a nuanced ability to identify core journals within the test article’s domain.

Minilm-l6, while maintaining a comparable thematic focus, included a broader range of journals, such as *Emerging Infectious Diseases* and *Frontiers in Public Health*. Its slightly lower average similarity scores compared to mpnet and minilm-l12 (always lower than 85%) highlight a broader distribution of relevance, potentially indicating a greater inclusion of interdisciplinary journals. Similarly, multi-qa-distilbert displayed a tendency toward multidisciplinary recommendations. Journals such as *Risk Management and Healthcare Policy* and the *Journal of Global Health* appeared prominently in its rankings. Finally, roberta showcased a balanced profile, with its top-ranked journal being the *International Journal*

of *Environmental Research and Public Health*. While there was some overlap with the other models, such as its inclusion of *Przegląd Epidemiologiczny* and the *Journal of Hospital Infection*, roberta also included journals like the *Journal of Global Health* and *Minerva Medica*.

Figure 1 summarizes these findings, highlighting how the models identify two journals as their most probable candidates, the *International Journal of Environmental Research and Public Health* for mpnet, minilm-l12, and roberta, and *Risk Management and Healthcare Policy* for minilm-l6 and multi-qa-distilbert. Interestingly, none of the models correctly identified the actual journal where the test article was published.

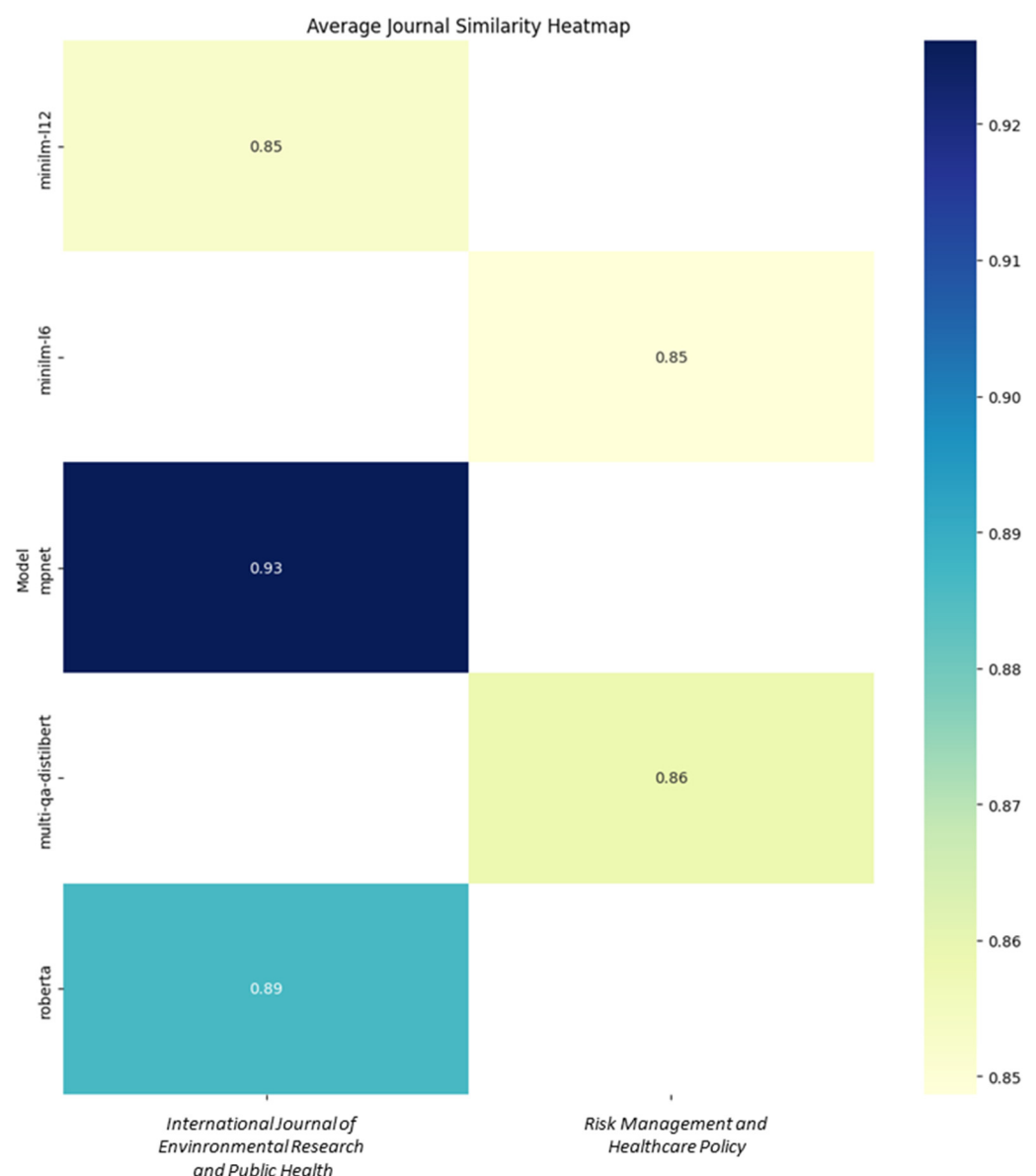


Figure 1. This heatmap visualizes the average journal similarity scores across different Sentence Transformer models. The rows represent the models, while the columns represent the journals that achieved high average similarity scores. The color gradient indicates the similarity score, with darker shades reflecting higher values.

3.1.3. Overlap Among Model Recommendations

To quantify the overlap in recommended journals, we calculated the set intersection and union of the top ten journals from each pair of models, as calculated according to the average similarity, and computed the ratio of the set intersection to the set union.

A heatmap (Figure 2) displayed the percentage overlap, where the color scale ranged from minimal overlap (~30%) to substantial overlap (~60–70%) for certain model pairs; *mpnet* and *minilm-l12* exhibited notably higher overlap, with 66.7% of their recommended journals in common. *Multi-qa-distilbert* frequently diverged, sharing less than 35% with *minilm-12* and *mpnet*, suggesting it captures text semantics or research topics differently. *Minilm-l6* yielded intermediate high overlap percentages with *roberta* (66.7%) and also had a relatively high alignment with *mpnet* or *minilm-l12* (53.8%).

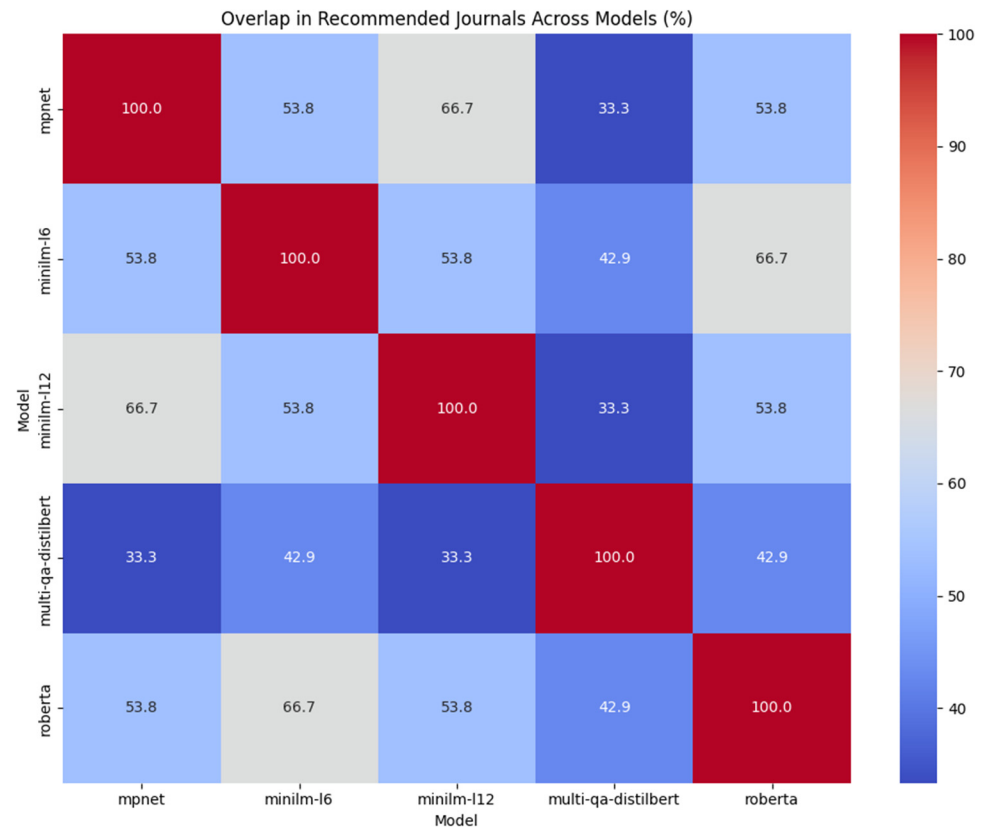


Figure 2. This heatmap illustrates the degree of overlap in journal recommendations among the five models, expressed as percentages.

3.2. Quantitative Comparison

We then extended our initial analysis of a single test article to encompass a broader evaluation involving 50 distinct test articles (Table S3), each randomly selected from the initial test dataset. This expansion aimed to provide a more generalized understanding of how various Sentence Transformer models performed in identifying thematically similar articles and recommending appropriate journals for publication. For each test article, we employed KeyBERT to extract the top three keywords, which were subsequently used to query PubMed, retrieving up to 20,000 PMIDs per query. To maintain computational efficiency and ensure manageable data processing, we randomly sampled 5000 PMIDs (if available) from each pool, carefully excluding the test article's own PMID, if present. The extracted articles were then consolidated into a DataFrame, ensuring the removal of duplicate entries based on PMID, resulting in a final pool size of maximum 5000 unique articles per test case.

Each of the five Sentence Transformer models was systematically applied to encode both the pooled articles and the respective test articles. For every model, we identified the top 10 most similar articles based on these scores. Additionally, we aggregated the similarity scores at the journal level by computing the average similarity of articles within each journal

to the test article, thereby ranking journals according to their thematic alignment with the test case.

The analysis yielded a comprehensive set of metrics for each model across all 50 test articles. Shannon entropy was calculated to assess the diversity of similarity scores among the top-ranked articles, with higher entropy values indicating a more even distribution of similarities and lower values suggesting dominance by one or a few highly similar articles. The maximum similarity score within the top 10 was recorded to measure the peak alignment each model achieved with the most similar article. Furthermore, we evaluated whether the actual journal of the test article was included within the top 10 recommended journals based on average similarity, providing a binary measure of each model's effectiveness in correctly identifying the appropriate publication venue.

Aggregating the results across all test articles (Table 3), we observed that mpnet consistently demonstrated the highest mean similarity scores, averaging 0.71 ± 0.04 , indicating a strong ability to identify journals closely aligned with the test articles' themes. This was followed by roberta, with an average mean similarity of 0.68 ± 0.04 , further showing its proficiency in aligning with relevant journals. The minilm-l12 and minilm-l6 models displayed comparable performance, achieving mean similarity scores of 0.67 ± 0.04 and 0.66 ± 0.04 , respectively, reflecting a solid alignment with test articles but slightly lower than that of mpnet and roberta. Conversely, multi-qa-distilbert exhibited the lowest mean similarity score of 0.63 ± 0.04 , indicating a more moderate alignment with the test articles.

Table 3. This table summarizes the performance of five Sentence Transformer models when applied to 50 test articles. The columns include Top Article, the maximum similarity to an article; Max Similarity, the highest similarity score achieved by the model with the average journal similarity; Mean Sim, the mean similarity identified by the model across the 50 test articles; Shannon Entropy and Gini score, a measure of diversity in similarity scores among the top-ranked articles; Journal in Top 10 (%), i.e., the percentage of test articles for which the actual journal of publication appeared in the top 10 recommended journals.

Model	Top Article	MaxSimilarity	MeanSim	Shannon Entropy	Gini	Journal in Top 10
minilm-l12	0.799	0.765548	0.67 ± 0.04	3.24	0.036077	12.5
minilm-l6	0.775	0.746310	0.66 ± 0.04	3.24	0.034712	14.5
mpnet	0.831	0.794203	0.71 ± 0.04	3.24	0.029997	14.5
multi-qa-distilbert	0.755	0.717609	0.63 ± 0.04	3.24	0.034632	12.5
roberta	0.803	0.773727	0.68 ± 0.04	3.24	0.033623	12.5

Additionally, mpnet achieved the highest Top Article similarity score of 0.83, emphasizing its capacity to match individual articles very closely to the test articles' themes. Roberta followed with a similarity of 0.80, while minilm-l12, minilm-l6, and multi-qa-distilbert achieved Top Article similarity scores of 0.79, 0.77, and 0.76, respectively.

The "JournalinTop10" metric revealed that mpnet and minilm-l6 were the most proficient in correctly identifying the actual journal of the test article within their top 10 recommendations, achieving inclusion rates of 14.5%. This was closely followed by minilm-l12, roberta, and multi-qa-distilbert, all of which had inclusion rates of 12.5%.

All models exhibited identical Shannon entropy values of approximately 3.24, suggesting that the similarity scores were evenly distributed across their recommendations. This low variability reflects a tendency to concentrate recommendations among a few top-ranked journals rather than distributing them broadly. The Gini coefficients were uniformly low across all models, with minilm-l12 showing the highest value of 0.0361, reflecting slightly higher concentration of similarity scores among its top recommendations. Mpnet

and roberta followed closely, with Gini values of 0.03 and 0.0336, respectively, indicating a relatively even distribution of similarity scores across their top journals.

4. Discussion

This study employed a straightforward pipeline for automated journal recommendation, serving primarily as a means to evaluate the performance of five freely available Sentence Transformer models. Specifically, we tested *all-mpnet-base-v2*, *all-MiniLM-L6-v2*, *all-MiniLM-L12-v2*, *multi-qa-distilbert-cos-v1*, and *all-distilroberta-v1*. We examined each model's ability to match manuscripts with thematically appropriate journals, using both a single randomly chosen test article and a larger set of 50 randomly selected articles. Our findings align with the notion that Transformer-based text embeddings are very effective in capturing semantic similarities between texts [13,14].

However, our pipeline operates under the assumption that an author's primary criterion for journal selection is thematic alignment. The final decision on where to submit a manuscript often involves additional considerations. Factors such as journal impact factor, acceptance rate, publication costs, and the expected turnaround time for reviews frequently influence authors' choices [41]. Moreover, authors might initially aim for more prestigious journals and use automated recommendations only after facing rejection from their preferred venues [42]. In our work, we purposefully isolated the scenario in which a researcher needs to identify journals that have recently published similar articles, thus highlighting each model's strengths and weaknesses in semantic matching without conflating these results with more pragmatic concerns.

Our pipeline relied on Transformer-based embeddings, which are numerical representation of sentence semantics. Embeddings allow for easy comparison between the meaning of two or more sentences [43–46]. This pipeline evaluated both article-level and journal-level alignment, illustrating how focusing on individual articles can help authors gauge thematic closeness to previously published work within a given venue, whereas examining journals in aggregate reveals which publication outlets frequently address the subject matter in question.

Consistently with our previous findings [16], mpnet excelled in identifying closely aligned journals, offering the highest average similarity score (MeanSim: 0.71 ± 0.04) when challenged with 50 test articles [36,47–95] and the highest maximum similarity (MaxSimilarity: 0.83) among our 50-test-article dataset. It also occasionally included the actual journal of publication in its top 10 suggestions (14.5% inclusion rate). Yet mpnet's strong performance came with longer encoding times, which could pose challenges in settings requiring swift feedback. Minilm-l12 performed almost as well, reaching a MeanSim of 0.67 ± 0.04 and a MaxSimilarity of 0.79, but required substantially less computational time. Similarly, minilm-l6 struck a balance between speed and accuracy, attaining a MeanSim of 0.66 ± 0.04 and matching mpnet's inclusion rate of 14.5% for the real publication venue. Its compact architecture (and that of minilm-l12) generates embeddings efficiently, suggesting potential utility in large-scale or real-time applications.

While mpnet yielded the highest domain-specific alignment, roberta also scored highly (MeanSim: 0.68), slightly above minilm-l12 (0.67). However, roberta included more interdisciplinary journals in its top recommendations, suggesting it might strike a balance between precision and broader thematic coverage. By contrast, multi-qa-distilbert displayed the lowest average similarity (0.63), reflecting its generalist nature and possibly broader, policy-related scope. This diversity might be advantageous for interdisciplinary manuscripts or exploratory research, where thematic breadth is prioritized over narrow precision; future research should address this issue to optimize the use of this model in the biomedical field.

We used Shannon entropy and the Gini coefficient to examine whether similarity scores were concentrated among a few top suggestions or evenly distributed. Minilm-l12 exhibited slightly higher Gini coefficients (0.036), implying more pronounced concentration at the top, whereas mpnet's lower Gini values (0.029) signaled a relatively equitable spread of scores. These metrics reinforce that each model presents different trade-offs between pinpointing the most thematically aligned options and revealing a broader swath of related venues.

One notable outcome in our single-article case study was that none of the models identified the actual journal where that article was published. This finding highlights a limitation of our evaluation approach, particularly our use of the "JournalinTop10" metric to gauge success. Authors may choose journals for a variety of reasons that are not captured purely by thematic alignment: prestige, speed, editorial focus, and acceptance rates all have significant impacts. Accordingly, we caution that a model's failure to recover the actual publication venue does not necessarily signal poor performance. The real journal choice might have hinged on non-semantic factors or happened to be absent from the sampled dataset. Going forward, these observations point to the need for more comprehensive recommendation strategies that integrate additional context, such as editorial policies, historical acceptance rates, or calls for special issues.

From a methodological standpoint, our use of KeyBERT for keyword extraction and random sampling of up to 5000 articles per test case demonstrates how a scalable NLP pipeline can be easily assembled with open-source tools. However, future iterations might integrate advanced query expansion or domain-specific ontologies like MeSH terms to capture a more exhaustive range of content. Adopting a more targeted or domain-focused search could mitigate any bias introduced by broader sampling, particularly in fields with rapidly changing terminologies.

In sum, this study demonstrates the strengths and limitations of Transformer-based models for automated journal recommendation. Mpnet stands out for high thematic precision but at a higher computational cost; multi-qa-distilbert and roberta offer more diverse, interdisciplinary venues; and the MiniLM variants provide balanced performance with minimal overhead.

5. Conclusions

This study presents an automated pipeline designed principally to assess the performance of Sentence Transformer models in recommending journals aligned with a manuscript's thematic content. By comparing five freely available models—mpnet, minilm-l6, minilm-l12, multi-qa-distilbert, and roberta—we demonstrate how their underlying architectures yield differing trade-offs in precision, computational efficiency, and breadth of recommendations. Mpnet stands out for its strong semantic alignment, whereas minilm-l6 and minilm-l12 offer more rapid processing with marginal compromises in similarity scores. Meanwhile, multi-qa-distilbert and roberta display a more interdisciplinary reach, potentially suiting authors whose research spans multiple domains.

Crucially, our pipeline does not attempt to identify a definitive or best venue for publication. Instead, it supplies a systematic, reproducible method for gauging each model's capabilities under consistent conditions. Real-world journal selection relies on far more than thematic fit, often involving publication fees, prestige, editorial policies, and personal author preferences. Nevertheless, our approach underlines how sentence embeddings can robustly capture scientific content, furnishing a starting point for authors who wish to explore plausible candidate journals grounded in semantic similarity.

Future extensions could fold in practical aspects such as acceptance rates, review timelines, and journal impact metrics. By combining these considerations with refined domain-specific embeddings and advanced query mechanisms, researchers can further

enhance how they evaluate and compare Sentence Transformer models for biomedical or interdisciplinary journal recommendation.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/bdcc9030067/s1>: Table S1: This table provides the list of article-level journal recommendations generated by the five different Sentence Transformer models tested when applied to the test article; Table S2: This table provides the list of journal-level journal recommendations generated by the five different Sentence Transformer models tested when applied to the test article; Table S3: This table contains the list of the 50 test articles used for quantitative appraisal of the model performance.

Author Contributions: Conceptualization, C.G., M.M. and E.C.; methodology, C.G.; software, C.G.; formal analysis, C.G. and M.T.C.; data curation, S.G. and M.T.C.; writing—original draft preparation, C.G. and M.M.; writing—review and editing, S.G. and E.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Welch, S.J. Selecting the Right Journal for Your Submission. *J. Thorac. Dis.* **2012**, *4*, 336–338. [CrossRef]
2. Nicholas, D.; Herman, E.; Clark, D.; Boukacem-Zeghmouri, C.; Rodríguez-Bravo, B.; Abrizah, A.; Watkinson, A.; Xu, J.; Sims, D.; Serbina, G.; et al. Choosing the ‘Right’ Journal for Publication: Perceptions and Practices of Pandemic-era Early Career Researchers. *Learn. Publ.* **2022**, *35*, 605–616. [CrossRef]
3. Larsen, P.O.; von Ins, M. The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index. *Scientometrics* **2010**, *84*, 575–603. [CrossRef]
4. Kreutz, C.K.; Schenkel, R. Scientific Paper Recommendation Systems: A Literature Review of Recent Publications. *Int. J. Digit. Libr.* **2022**, *23*, 335–369. [CrossRef]
5. Park, D.H.; Kim, H.K.; Choi, I.Y.; Kim, J.K. A Literature Review and Classification of Recommender Systems Research. *Expert. Syst. Appl.* **2012**, *39*, 10059–10072. [CrossRef]
6. Qaiser, S.; Ali, R. Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents. *Int. J. Comput. Appl.* **2018**, *181*, 25–29. [CrossRef]
7. Wang, D.; Liang, Y.; Xu, D.; Feng, X.; Guan, R. A Content-Based Recommender System for Computer Science Publications. *Knowl. Based Syst.* **2018**, *157*, 1–9. [CrossRef]
8. Medvet, E.; Bartoli, A.; Piccinin, G. Publication Venue Recommendation Based on Paper Abstract. In Proceedings of the 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, Limassol, Cyprus, 10–12 November 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 1004–1010.
9. Yang, Z.; Davison, B.D. Venue Recommendation: Submitting Your Paper with Style. In Proceedings of the 2012 11th International Conference on Machine Learning and Applications, Boca Raton, FL, USA, 12–15 December 2012; IEEE: Piscataway, NJ, USA; pp. 681–686.
10. Beel, J.; Gipp, B.; Langer, S.; Breiteringer, C. Research-Paper Recommender Systems: A Literature Survey. *Int. J. Digit. Libr.* **2016**, *17*, 305–338. [CrossRef]
11. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.
12. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Volume 30.
13. Liu, Q.; Kusner, M.J.; Blunsom, P. A Survey on Contextual Embeddings. *arXiv* **2020**, arXiv:2003.07278.
14. Noh, J.; Kavuluru, R. Improved Biomedical Word Embeddings in the Transformer Era. *J. Biomed. Inform.* **2021**, *120*, 103867. [CrossRef]

15. Michail, S.; Ledet, J.W.; Alkan, T.Y.; İnce, M.N.; Günay, M. A Journal Recommender for Article Submission Using Transformers. *Scientometrics* **2023**, *128*, 1321–1336. [CrossRef]
16. Galli, C.; Donos, N.; Calciolari, E. Performance of 4 Pre-Trained Sentence Transformer Models in the Semantic Query of a Systematic Review Dataset on Peri-Implantitis. *Information* **2024**, *15*, 68. [CrossRef]
17. Stankevičius, L.; Lukoševičius, M. Extracting Sentence Embeddings from Pretrained Transformer Models. *Appl. Sci.* **2024**, *14*, 8887. [CrossRef]
18. Siino, M. All-Mpnet at Semeval-2024 Task 1: Application of Mpnet for Evaluating Semantic Textual Relatedness. In Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024), Mexico City, Mexico, 20–21 June 2024; pp. 379–384.
19. Yin, C.; Zhang, Z. A Study of Sentence Similarity Based on the All-Minilm-L6-v2 Model With “Same Semantics, Different Structure” After Fine Tuning. In Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024), Singapore, 9–11 August 2024; Atlantis Press: Dordrecht, The Netherlands, 2024; pp. 677–684.
20. He, Y.; Yuan, M.; Chen, J.; Horrocks, I. Language Models as Hierarchy Encoders. *Adv. Neural Inf. Process Syst.* **2025**, *37*, 14690–14711.
21. Bistarelli, S.; Cuccarini, M. BERT-Based Questions Answering on Close Domains: Preliminary Report. In Proceedings of the Proceedings of the 39th Italian Conference on Computational Logic, Rome, Italy, 26–28 June 2024.
22. Wijanto, M.C.; Yong, H.-S. Combining Balancing Dataset and Sentence Transformers to Improve Short Answer Grading Performance. *Appl. Sci.* **2024**, *14*, 4532. [CrossRef]
23. Issa, B.; Jasser, M.B.; Chua, H.N.; Hamzah, M. A Comparative Study on Embedding Models for Keyword Extraction Using KeyBERT Method. In Proceedings of the 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 2 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 40–45.
24. Jin, Q.; Leaman, R.; Lu, Z. PubMed and beyond: Biomedical Literature Search in the Age of Artificial Intelligence. *EBioMedicine* **2024**, *100*, 104988. [CrossRef] [PubMed]
25. McKinney, W. Data Structures for Statistical Computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; pp. 51–56.
26. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; Facebook, Z.D.; Research, A.I.; Lin, Z.; Desmaison, A.; Antiga, L.; et al. Automatic Differentiation in PyTorch. In Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
27. Godden, J.W.; Bajorath, J. Analysis of Chemical Information Content Using Shannon Entropy. *Rev. Comput. Chem.* **2007**, *23*, 263–289.
28. Vajapeyam, S. Understanding Shannon’s Entropy Metric for Information. *arXiv* **2014**, arXiv:1405.2061.
29. Dorfman, R. A Formula for the Gini Coefficient. *Rev. Econ. Stat.* **1979**, *61*, 146. [CrossRef]
30. Bisong, E. Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*; Bisong, E., Ed.; Apress: Berkeley, CA, USA, 2019; pp. 59–64, ISBN 978-1-4842-4470-8.
31. Chapman, B.; Chang, J. Biopython: Python Tools for Computational Biology. *ACM Sigbio Newsl.* **2000**, *20*, 15–19. [CrossRef]
32. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef] [PubMed]
33. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. Available online: <https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html> (accessed on 14 January 2025).
34. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
35. Waskom, M. Seaborn: Statistical Data Visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [CrossRef]
36. Kiersnowska, Z.; Lemiech-Mirowska, E.; Ginter-Kramarczyk, D.; Kruszelnicka, I.; Michałkiewicz, M.; Marczak, M. Problems of Clostridium Difficile Infection (CDI) in Polish Healthcare Units. *Ann. Agric. Environ. Med.* **2021**, *28*, 224–230. [CrossRef]
37. Khanafer, N.; Voirin, N.; Barbut, F.; Kuijper, E.; Vanhems, P. Hospital Management of Clostridium Difficile Infection: A Review of the Literature. *J. Hosp. Infect.* **2015**, *90*, 91–101. [CrossRef]
38. Legenza, L.; Barnett, S.; Rose, W.; Safdar, N.; Emmerling, T.; Peh, K.H.; Coetzee, R. Clostridium Difficile Infection Perceptions and Practices: A Multicenter Qualitative Study in South Africa. *Antimicrob. Resist. Infect. Control* **2018**, *7*, 125. [CrossRef]
39. Nordling, S.; Anttila, V.J.; Norén, T.; Cockburn, E. The Burden of Clostridium Difficile (CDI) Infection in Hospitals, in Denmark, Finland, Norway And Sweden. *Value Health* **2014**, *17*, A670. [CrossRef]
40. Luo, R.; Barlam, T.F. Ten-Year Review of Clostridium Difficile Infection in Acute Care Hospitals in the USA, 2005–2014. *J. Hosp. Infect.* **2018**, *98*, 40–43. [CrossRef]
41. Xu, X.; Xie, J.; Sun, J.; Cheng, Y. Factors Affecting Authors’ Manuscript Submission Behaviour: A Systematic Review. *Learn. Publ.* **2023**, *36*, 285–298. [CrossRef]

42. Gaston, T.E.; Ounsworth, F.; Senders, T.; Ritchie, S.; Jones, E. Factors Affecting Journal Submission Numbers: Impact Factor and Peer Review Reputation. *Learn. Publ.* **2020**, *33*, 154–162. [\[CrossRef\]](#)
43. Worth, P.J. Word Embeddings and Semantic Spaces in Natural Language Processing. *Int. J. Intell. Sci.* **2023**, *13*, 1–21. [\[CrossRef\]](#)
44. Yao, Z.; Sun, Y.; Ding, W.; Rao, N.; Xiong, H. Dynamic Word Embeddings for Evolving Semantic Discovery. In Proceedings of the WSDM 2018—Proceedings of the 11th ACM International Conference on Web Search and Data Mining, Los Angeles, CA, USA, 5–9 February 2018; ACM: New York, NY, USA, 2018; pp. 673–681. [\[CrossRef\]](#)
45. Si, Y.; Wang, J.; Xu, H.; Roberts, K. Enhancing Clinical Concept Extraction with Contextual Embeddings. *J. Am. Med. Inform. Assoc.* **2019**, *26*, 1297–1304. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Gutiérrez, L.; Keith, B. A Systematic Literature Review on Word Embeddings. In *Trends and Applications in Software Engineering, Proceedings of the 7th International Conference on Software Process Improvement (CIMPS 2018)*; Guadalajara, Mexico, 17–19 October 2018; Springer: Berlin/Heidelberg, Germany, 2019; Volume 7, pp. 132–141.
47. Lu, W.-H.; Iuli, R.; Strano-Paul, L.; Chandran, L. Renaissance School of Medicine at Stony Brook University. *Acad. Med.* **2020**, *95*, S362–S366. [\[CrossRef\]](#)
48. Gottlieb, M.; Alerhand, S.; Long, B. Response to: “POCUS to Confirm Intubation in a Trauma Setting”. *West. J. Emerg. Med.* **2020**, *22*, 400. [\[CrossRef\]](#)
49. He, L.; Zhang, Q.; Li, Z.; Shen, L.; Zhang, J.; Wang, P.; Wu, S.; Zhou, T.; Xu, Q.; Chen, X.; et al. Incorporation of Urinary Neutrophil Gelatinase-Associated Lipocalin and Computed Tomography Quantification to Predict Acute Kidney Injury and In-Hospital Death in COVID-19 Patients. *Kidney Dis.* **2021**, *7*, 120–130. [\[CrossRef\]](#)
50. Scheffers, L.E.; Berg, L.E.M.V.; Ismailova, G.; Dulfer, K.; Takkenberg, J.J.M.; Helbing, W.A. Physical Exercise Training in Patients with a Fontan Circulation: A Systematic Review. *Eur. J. Prev. Cardiol.* **2021**, *28*, 1269–1278. [\[CrossRef\]](#)
51. Zhang, J.; Li, Y.; Lai, D.; Lu, D.; Lan, Z.; Kang, J.; Xu, Y.; Cai, S. Vitamin D Status Is Negatively Related to Insulin Resistance and Bone Turnover in Chinese Non-Osteoporosis Patients With Type 2 Diabetes: A Retrospective Cross-Section Research. *Front. Public. Health* **2022**, *9*, 727132. [\[CrossRef\]](#)
52. Gayathri, A.; Ramachandran, R.; Narasimhan, M. Segmental Neurofibromatosis with Lisch Nodules. *Med. J. Armed Forces India* **2023**, *79*, 356–359. [\[CrossRef\]](#)
53. Rice, P.E.; Nimphius, S. When Task Constraints Delimit Movement Strategy: Implications for Isolated Joint Training in Dancers. *Front. Sports Act. Living* **2020**, *2*, 49. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Marques, E.S.; de Oliveira, A.G.e.S.; Faerstein, E. Psychometric Properties of a Modified Version of Brazilian Household Food Insecurity Measurement Scale—Pró-Saúde Study. *Cien Saude Colet.* **2021**, *26*, 3175–3185. [\[CrossRef\]](#)
55. Kumar, P.; Desai, C.; Das, A. Unilateral Linear Capillaritis. *Indian. Dermatol. Online J.* **2021**, *12*, 486–487. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Archer, T.; Aziz, I.; Kurien, M.; Knott, V.; Ball, A. Prioritisation of Lower Gastrointestinal Endoscopy during the COVID-19 Pandemic: Outcomes of a Novel Triage Pathway. *Frontline Gastroenterol.* **2022**, *13*, 225–230. [\[CrossRef\]](#)
57. Ortiz-Martínez, Y.; Fajardo-Rivero, J.E.; Vergara-Retamozo, R.; Vergel-Torrado, J.A.; Esquiaqui-Rangel, V. Chronic Obstructive Pulmonary Disease in Latin America and the Caribbean: Mapping the Research by Bibliometric Analysis. *Indian J. Tuberc.* **2022**, *69*, 262–263. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Taylor, M.A.; Maclean, J.G. Anterior Hip Ultrasound: A Useful Technique in Developmental Dysplasia of the Hips. *Ultrasound* **2021**, *29*, 179–186. [\[CrossRef\]](#)
59. Habigt, M.A.; Gesenhues, J.; Ketelhut, M.; Hein, M.; Duschner, P.; Rossaint, R.; Mechelinck, M. In Vivo Evaluation of Two Adaptive Starling-like Control Algorithms for Left Ventricular Assist Devices. *Biomed. Eng. Biomed. Tech.* **2021**, *66*, 257–266. [\[CrossRef\]](#)
60. Sachser, N.; Zimmermann, T.D.; Hennessy, M.B.; Kaiser, S. Sensitive Phases in the Development of Rodent Social Behavior. *Curr. Opin. Behav. Sci.* **2020**, *36*, 63–70. [\[CrossRef\]](#)
61. Dominguez-Colman, L.; Mehta, S.U.; Mansourkhani, S.; Sehgal, N.; Alvarado, L.A.; Mariscal, J.; Tonarelli, S. Teaching Psychiatric Emergencies Using Simulation: An Experience During the Boot Camp. *Med. Sci. Educ.* **2020**, *30*, 1481–1486. [\[CrossRef\]](#)
62. Roozenbeek, J.; Maertens, R.; McClanahan, W.; van der Linden, S. Disentangling Item and Testing Effects in Inoculation Research on Online Misinformation: Solomon Revisited. *Educ. Psychol. Meas.* **2021**, *81*, 340–362. [\[CrossRef\]](#)
63. Tumyan, G.; Mantha, Y.; Gill, R.; Feldman, M. Acute Sterile Meningitis as a Primary Manifestation of Pituitary Apoplexy. *AACE Clin. Case Rep.* **2021**, *7*, 117–120. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Alade, O.; Ajoloko, E.; Dedeke, A.; Uti, O.; Sofola, O. Self-Reported Halitosis and Oral Health Related Quality of Life in Adolescent Students from a Suburban Community in Nigeria. *Afr. Health Sci.* **2020**, *20*, 2044–2049. [\[CrossRef\]](#)
65. Evans, T.J.; Davidson, H.C.; Low, J.M.; Basarab, M.; Arnold, A. Antibiotic Usage and Stewardship in Patients with COVID-19: Too Much Antibiotic in Uncharted Waters? *J. Infect. Prev.* **2021**, *22*, 119–125. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Bose, S.K.; He, Y.; Howlader, P.; Wang, W.; Yin, H. The N-Glycan Processing Enzymes β -D-N-Acetylhexosaminidase Are Involved in Ripening-Associated Softening in Strawberry Fruit. *J. Food Sci. Technol.* **2021**, *58*, 621–631. [\[CrossRef\]](#) [\[PubMed\]](#)

67. Falsaperla, R.; Saporito, M.A.N.; Pisani, F.; Mailo, J.; Pavone, P.; Ruggieri, M.; Suppiej, A.; Corsello, G. Ocular Motor Paroxysmal Events in Neonates and Infants: A Review of the Literature. *Pediatr. Neurol.* **2021**, *117*, 4–9. [\[CrossRef\]](#)
68. Almutairi, A.F.; Albeshier, N.; Aljohani, M.; Alsinanni, M.; Turkistani, O.; Salam, M. Association of Oral Parafunctional Habits with Anxiety and the Big-Five Personality Traits in the Saudi Adult Population. *Saudi Dent. J.* **2021**, *33*, 90–98. [\[CrossRef\]](#)
69. Lu, Z.; Yang, Y.-J.; Ni, W.-X.; Li, M.; Zhao, Y.; Huang, Y.-L.; Luo, D.; Wang, X.; Omary, M.A.; Li, D. Aggregation-Induced Phosphorescence Sensitization in Two Heptanuclear and Decanuclear Gold–Silver Sandwich Clusters. *Chem. Sci.* **2021**, *12*, 702–708. [\[CrossRef\]](#)
70. Morgun, A.V.; Salmin, V.V.; Boytsova, E.B.; Lopatina, O.L.; Salmina, A.B. Molecular Mechanisms of Proteins—Targets for SARS-CoV-2. *Sovrem Tekhnologii Med.* **2020**, *12*, 98–108. [\[CrossRef\]](#)
71. Do Nascimento, P.A.; Kogawa, A.C.; Salgado, H.R.N. Current Status of Vancomycin Analytical Methods. *J. AOAC Int.* **2020**, *103*, 755–769. [\[CrossRef\]](#)
72. Booth, V. Deuterium Solid State NMR Studies of Intact Bacteria Treated With Antimicrobial Peptides. *Front. Med. Technol.* **2021**, *2*, 621572. [\[CrossRef\]](#)
73. Nguyen, H.; Haeney, O.; Galletly, C. The Characteristics of Older Homicide Offenders: A Systematic Review. *Psychiatry Psychol. Law* **2022**, *29*, 413–430. [\[CrossRef\]](#)
74. Duseja, A.; Chahal, G.; Jain, A.; Mehta, M.; Ranjan, A.; Grover, V. Association between Nonalcoholic Fatty Liver Disease and Inflammatory Periodontal Disease: A Case-Control Study. *J. Indian. Soc. Periodontol.* **2021**, *25*, 47. [\[CrossRef\]](#)
75. Kim, K.Y.; Lee, E.; Cho, J. Factors Affecting Healthcare Utilization among Patients with Single and Multiple Chronic Diseases. *Iran. J. Public. Health* **2020**, *49*, 2367–2375. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Xia, W.; Shangguan, X.; Li, M.; Wang, Y.; Xi, D.; Sun, W.; Fan, J.; Shao, K.; Peng, X. Ex Vivo Identification of Circulating Tumor Cells in Peripheral Blood by Fluorometric “Turn on” Aptamer Nanoparticles. *Chem. Sci.* **2021**, *12*, 3314–3321. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Rathour, M.; Sharma, A.; Kaur, A.; Upadhyay, S.K. Genome-Wide Characterization and Expression and Co-Expression Analysis Suggested Diverse Functions of WOX Genes in Bread Wheat. *Heliyon* **2020**, *6*, e05762. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Hepatobiliary Disease Study Group; Chinese Society of Gastroenterology; Chinese Medical Association. Consensus for Management of Portal Vein Thrombosis in Liver Cirrhosis (2020, Shanghai). *J. Dig. Dis.* **2021**, *22*, 176–186. [\[CrossRef\]](#)
79. Bansal, L.K.; Gupta, S.; Gupta, A.K.; Chaudhary, P. Thyroid Tuberculosis. *Indian J. Tuberc.* **2021**, *68*, 272–278. [\[CrossRef\]](#)
80. Eble, J. Implications of John Kavanaugh’s Philosophy of the Human Person as Embodied Reflexive Consciousness for Conscientious Decision-Making in Brain Death. *Linacre Q.* **2021**, *88*, 71–81. [\[CrossRef\]](#)
81. Rivera, A.; González-Pozega, C.; Ibarra, G.; Fernandez-Ibarburu, B.; García-Ruano, Á.; Vallejo-Valero, A. Punctual Breast Implant Rupture Following Lipofilling: Only a Myth? *Breast Care* **2021**, *16*, 544–547. [\[CrossRef\]](#)
82. Álvarez Hernández, P.; de la Mata Llord, J. Expanded Mesenchymal Stromal Cells in Knee Osteoarthritis: A Systematic Literature Review. *Reumatol. Clínica (Engl. Ed.)* **2022**, *18*, 49–55. [\[CrossRef\]](#)
83. Keche, P.N.; Davange, N.A.; Gawarle, S.H.; Ganguly, S. A Study of Anatomical Variations in the Nasal Cavity in Cases of Chronic Rhinosinusitis. *Indian J. Otolaryngol. Head. Neck Surg.* **2022**, *74*, 943–948. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Igbokwe, M.; Adebayo, O.; Ogunsuji, O.; Popoola, G.; Babalola, R.; Oiwoh, S.O.; Makinde, A.M.; Adeniyi, A.M.; Kanmodi, K.; Umar, W.F.; et al. An Exploration of Profile, Perceptions, Barriers, and Predictors of Research Engagement among Resident Doctors. *Perspect. Clin. Res.* **2022**, *13*, 106–113. [\[CrossRef\]](#)
85. Marraccini, M.E.; Pittleman, C. Returning to School Following Hospitalization for Suicide-Related Behaviors: Recognizing Student Voices for Improving Practice. *School Psych. Rev.* **2022**, *51*, 370–385. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Howe, N. Podcast Extra: Coronavirus—Science in the Pandemic. *Nature* **2020**. *online ahead of print.* [\[CrossRef\]](#)
87. Jain, V.; Sybil, D.; Premchandani, S.; Krishna, M.; Singh, S. Intraoral Bone Regeneration Using Stem Cells-What a Clinician Needs to Know: Based on a 15-Year MEDLINE Search. *Front. Dent.* **2021**, *18*, 40. [\[CrossRef\]](#)
88. Joo, J.Y.; Liu, M.F. Culturally Tailored Interventions for Ethnic Minorities: A Scoping Review. *Nurs. Open* **2021**, *8*, 2078–2090. [\[CrossRef\]](#)
89. Jarmul, S.; Dangour, A.D.; Green, R.; Liew, Z.; Haines, A.; Scheelbeek, P.F. Climate Change Mitigation through Dietary Change: A Systematic Review of Empirical and Modelling Studies on the Environmental Footprints and Health Effects of ‘Sustainable Diets’. *Environ. Res. Lett.* **2020**, *15*, 123014. [\[CrossRef\]](#)
90. Tezuka, M.; Mizusawa, H.; Tsukada, M.; Mimura, Y.; Shimizu, T.; Kobayashi, A.; Takahashi, Y.; Maejima, T. Severe Necrosis of the Glans Penis Associated with Calciphylaxis Treated by Partial Penectomy. *IJU Case Rep.* **2020**, *3*, 133–136. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Reeder, J.A.; Reynolds, T.R.; Gilbert, B.W. Successful Use of Aspirin, Apixaban, and Viscoelastography in a Patient with Severe COVID Disease and Allergy to Porcine Products. *Hosp. Pharm.* **2022**, *57*, 135–137. [\[CrossRef\]](#)
92. Alzer, H.; Kalbounieh, H.; Alsoleihat, F.; Abu Shahin, N.; Ryalat, S.; Alsalem, M.; Alahmad, H.; Tahtamouni, L. Age of the Donor Affects the Nature of in Vitro Cultured Human Dental Pulp Stem Cells. *Saudi Dent. J.* **2021**, *33*, 524–532. [\[CrossRef\]](#)

93. Patel, A.; Muthukrishnan, I.; Kurian, A.; Amalchandra, J.; Sampathirao, N.; Simon, S. Multicentric Primary Diffuse Large B-Cell Lymphoma in Genitourinary Tract Detected on 18F-Fluorodeoxyglucose Positron Emission Tomography with Computed Tomography: An Uncommon Presentation of a Common Malignancy. *World J. Nucl. Med.* **2021**, *20*, 117–120. [[CrossRef](#)]
94. Bittmann, L. Quantum Grothendieck Rings as Quantum Cluster Algebras. *J. Lond. Math. Soc.* **2021**, *103*, 161–197. [[CrossRef](#)] [[PubMed](#)]
95. da Silva Antas, J.; de Holanda, A.K.G.; de Sousa Andrade, A.; de Araujo, A.M.S.; de Carvalho Costa, I.G.; de Lima, S.K.M.; da Fonseca Abrantes Sarmen, P.L. Curva de Aprendizado de Anastomose Arteriovenosa Com Uso de Simulador de Baixo Custo. *J. Vasc. Bras.* **2020**, *19*, e20190144. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.