|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reference | Category | Problem | Method | Key Contribution | Limitations | How it Helps | Evaluation | Implement or Cite |
| Robertson & Zaragoza (2009) – *The Probabilistic Relevance Framework: BM25 and Beyond* | Classical IR | Need a reliable way to rank documents by keyword relevance | Probabilistic ranking function (BM25) | Established BM25 as the standard, interpretable baseline for IR | Doesn’t capture semantics or context, only term frequency & inverse document frequency | Provides a solid baseline for venue recommendation, lets me show improvement with embeddings/LLMs | BM25 is efficient and transparent, making it an ideal benchmark in my experiments. I’ll run it on titles/abstracts to show how classical IR compares against embeddings and hybrid LLM methods. Its simplicity makes evaluation clear and reproducible. | Implement |
| Manning , Raghavan Introduction to information retrieval | Classical Information retrieval | Need a foundation for search and ranking models | Covers vector space model , probabilistic models , evaluation metrics | IR fundamentals | Outdated on neural and embedding models | Useful for background terminology | Good for background on IR and to provide context within my project.  Good to use as an anchor for evaluation metrics and classical IR concepts. | Implement |
| Croft et al | Classical Information retrieval | Need practical methods for building IR systems | Practical textbook on indexing and ranking , evaluation | Implementation of IR | Pre-neural , outdated | Methodological background and support for BM25 baseline | C | Cite |
| Cohan et al. (2020) – *SPECTER: Document-level Representation Learning using Citation-informed Transformers* ACL Anthology Link | Embedding | Need Scientific document embeddings beyond keyword search | Transformer  Trained with citation informed objective | Combines semantics + citation context for stronger paper embeddings | Requires citation networks, weaker with sparse metadata | Provides a strong baseline embedding model for academic venue recommendation |  | Implement |
| MacAvaney et al CEDER, Contextualized embeddings for Document Ranking | Embedding | Need improved ranking that’s better than BM25 | Fine tunes BERT for IR rankings tasks | Shows contextual embeddings outperform static IR methods | High computational cost; limited scalability | Contextual embeddings might be superior to BM25 |  | Cite |