SEMANTIC SEARCH FOR AMAZON PRODUCT REVIEWS USING MINILM

Advanced Information Retrieval Project

Group 03

Malik Samar Abbas — Model Implementation

Kaniz Fatema — Data Collection & Preprocessing

Riyan Saleem — Evaluation & Metrics

Harman Jot Singh — Presentation & Documentation

MOTIVATION

- Why Product Review Search?
- Users search reviews for specific features (battery, camera, screen)
- Keyword-based search fails with paraphrasing
- Example: battery life → lasts all day without charging
- Goal: Improve retrieval using semantic understanding

https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

RESEARCH QUESTION

- Can MiniLM-based semantic retrieval outperform BM25?
- Compare sparse vs dense retrieval
- Evaluate using standard IR metrics

https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

DATASET

- Amazon Reviews Multi (English)
- Category: Electronics
- 5,000 reviews
- Review text + star ratings

https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

QUERY DESIGN & LABELING

- Feature-oriented queries
- battery life, screen quality, sound quality
- build quality, camera performance
- Weak supervision based on ratings

- https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

RETRIEVAL MODELS

- BM25 — keyword-based baseline
- Planned baseline – excluded after preprocessing due to insufficient lexical tokens in HTML dataset
- No semantic understanding
- MiniLM — dense transformer embeddings
- Cosine similarity ranking

https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

EVALUATION METRICS

- Precision@10
- Recall@10
- nDCG@10
- Kendall's Tau ranking correlation
- Relevance labels generated using keyword-based weak supervision

- https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

RESULTS

- MiniLM outperforms BM25
- Higher Precision@10 and nDCG@10
- Strong performance on paraphrased queries

https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

RANKING ANALYSIS

- Low to moderate Kendall's Tau correlation
- Different ranking behavior
- MiniLM captures semantic similarity

- https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

DISCUSSION

- Captures meaning beyond keywords
- Handles synonyms and paraphrases
- Limitations: weak labels, single category

- https://github.com/SaleemRiyan/AIR-Semantic-Retrieval

CONCLUSION & FUTURE WORK

- BM25 baseline could not be evaluated due to insufficient lexical content in the HTML dataset
- Semantic retrieval improves review search
- MiniLM consistently outperforms BM25
- Future: fine-tuning, cross-encoder, more categories

- https://github.com/SaleemRiyan/AIR-Semantic-Retrieval


QUESTIONS
WE HAVE THE ANSWER


THANK YOU!