

Feature	Cosine Similarity (Baseline)	FAISS (Optimized)
Method	Direct pairwise similarity computation using cosine_similarity on the entire matrix.	Builds a vector index (ANN – Approximate Nearest Neighbor) and searches efficiently.
Speed	Very slow on large datasets (because it computes similarity with every vector).	Much faster (sub-linear search) even with millions of vectors.
Memory Usage	Needs to keep the full similarity matrix in memory (N x N).	Stores a compressed index; only searches top-K nearest neighbors.
Accuracy	Exact (100% correct neighbors).	Approximate by default, but very close to exact ; can be tuned for accuracy.
Scalability	Works fine for small CSVs (<10k movies). Becomes unusable for large datasets (>100k).	Designed to handle millions of items efficiently.
Library	scikit-learn	faiss (Facebook AI Similarity Search, optimized in C++/CUDA).