Move Plot Search Analysis

Using semantic search method for the prototype search tool on Wikipedia movie plots provided some noteworthy insight, even for this basic search tool model. To analyze how well a model did with its top five matches, I used both Recall@1 and Mean Reciprocal Rank (MRR) as evaluation methods. Out of the six queries we tested, two of the results were accurate in matching up the top result with the ground truth movie. This resulted in a Recall@1 score of 0.333. Besides the 2 movies that matched as the first result, one of the queries was able to match the ground truth movie with its third result, which means it got an MRR score of 0.389. The queries with no matches at all were the following: "Western Romance", "Comedy film, office disguises, boss's daughter, elopement.", and "Denis Gage Deane-Tanner". It becomes evident that the queries that matched better to the ground truth movies were the ones that were more detailed and were written in a similar sentence structure to that of the plot descriptions. By providing more details, and thus being less vague in the movies they were looking for, the model had a much easier time matching up the correct movie. To improve the queries that did not yield the correct result, the first step is to provide much more detail into the movie they are looking for. The second alternate approach to improve the results would be to write the query in a sentence structure similar to the plot descriptions that the model is using to match; this can fix the issue that the fourth query had were the model had to only work with sentence fragments.

The BM25 model performed better with a Recall@1 of 0.5 and a MRR of 0.556. Furthermore, the reranker model performed much better than the semantic search model. The first performance improvement came from getting 'Saraha' as the correct movie at number one whereas the previous model matched that movie at number three. The second major improvement came from being able to work through the nonstructured query that cause problems in the semantic search model and providing the ground truth movie at number one. This resulted in a Recall@1 and an MRR score of 0.667.

The RAG model uses a large language model from Fireworks.ai to run this search model. It performed generally better than the reranker model; the only downgrade is that the query: "Comedy film, office disguises, boss's daughter, elopement." had the ground truth movie of "Ask Father" at number one for the reranker and number two for RAG. The improvements come from getting the correct ground truth movie for the query: "Denis Gage Deane-Tanner", where the other two models failed to get at all. This resulted in a Recall@1 score of 0.667, the same as the reranker, but a MRR score of 0.75. This makes sense that it is the best out of the three because RAG has been trained on much more complicated and much larger datasets; thus, it is better at handling some ambiguity, like the "Denis Gage Deane-Tanner" query. The RAG method also reduces the possibility of hallucinations.

Out of the three models I would recommend using the RAG/LLM model because of its flexibility and its ability to avoid hallucinations and handle larger sets of more complicated data. The other methods are much simpler to implement but they do not provide more accurate results based on the evaluations of Recall@1 and MRR. Though this method might cost more to implement and it is more complex to run, I believe that the results are worth it while also allowing more flexibility in its parameters and a greater ability to understand context.

Sometimes pre-trained models like 'nq-distilbert-base-v1' are not sufficient enough and are not personalized enough to work with our own data; thus, fine tuning these types of models are required. One aspect of fine tuning we could immediately implement are loss functions like CosineSimilarityLoss; these loss functions aid our embedding model by determining how well it works with out task. Further ways we can fine tune the model are by changing the number of epochs for training, the type of optimizer, and even the learning rate scheduler.

Semantic Search

Query	Ground Truth Movie	Your Results
"Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions"	'Nanook of the North'	1. ==‘Nanook of the North’== 2. ‘David Copperfield’ 3. ‘Straight Shooting’ 4. ‘The Frozen North’ 5. ‘The Boy Friend’
"Western romance"	'The Lucky Horseshoe'	1. ‘The Great Gatsby’ 2. ‘Youth's Endearing Charm’ 3. ‘Frankenstein’ 4. ‘High Society Blues’ 5. ‘Romance’
"Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."	'Sahara'	1. ‘Married in Hollywood' 2. ‘The King on Main Street’ 3. ==‘Sahara’== 4. ‘One Hysterical Night’ 5. ‘Bound in Morocco’
"Comedy film, office disguises, boss's daughter, elopement."	'Ask Father'	1. ‘Youth's Endearing Charm’ 2. ‘Dressed to Kill’ 3. ‘The Boy Friend’ 4. ‘A Little Journey’ 5. ‘This Is Heaven’
"Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."	‘Cleopatra’	1. ==‘Cleopatra’== 2. ‘The Lost World’ 3. ‘The Man Who Lost Himself’ 4. ‘The Golden Louis’ 5. ‘Reaching for the Moon’
"Denis Gage Deane-Tanner"	'Captain Alvarez'	1. ‘Souls for Sale’ 2. ‘The Man in Possession’ 3. ‘The Man from Blankley's’ 4. ‘Blind Youth’ 5. ‘Old Clothes’

BM25

Query	Ground Truth Movie	Your Results
"Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions"	'Nanook of the North'	1. 'I Do' 2. 'Uncharted Seas' 3. 'Power That Pray' 4. 'Rough Waters' 5. 'The King of Kings'
"Western romance"	'The Lucky Horseshoe'	1. 'The Call of the Wild' 2. 'Romance' 3. 'Four Sons' 4. 'A Man's Fight' 5. 'Gun Smoke'
"Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."	'Sahara'	1. 'Sahara' 2. 'He Who Got Slapped' 3. 'A lady from Old Ireland' 4. 'Inspiration' 5. 'The Maltese Falcon'
"Comedy film, office disguises, boss's daughter, elopement."	'Ask Father'	1. 'Mabel's Blunder 2. 'Bucking Broadway' 3. 'Ask Father' 4. 'The Front Page' 5. 'His Wedding Night'
"Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."	'Cleopatra'	1. 'Cleopatra' 2. 'Captain Applejack' 3. 'Fair Lady' 4. 'Reaching for the Moon' 5. 'Tom Sawyer'
"Denis Gage Deane-Tanner"	'Captain Alvarez'	1. 'Captain Alvarez' 2. 'Hangman's House' 3. 'Red Courage' 4. 'The Sea Lion' 5. 'A Sailor-Made Man'

Reranker

Query	Ground Truth Movie	Your Results
"Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions"	'Nanook of the North'	1. ==‘Nanook of the North’== 2. ‘Uncharted Seas’ 3. ‘The Frozen North’ 4. ‘Youth's Endearing Charm’ 5. ‘The Salvation Hunters’
"Western romance"	'The Lucky Horseshoe'	1. ‘Romance’ 2. ‘Married in Hollywood’ 3. ‘Across the Pacific’ 4. ‘A Lady of Chance’ 5. ‘When a Man Loves’
"Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."	'Sahara'	1. ==‘Sahara’== 2. ‘Camille’ 3. ‘Cleopatra’ 4. ‘Peacock Alley’ 5. ‘A woman of Paris’
"Comedy film, office disguises, boss's daughter, elopement."	'Ask Father'	1. ==‘Ask Father’== 2. ‘Mabel's Blunder’ 3. ‘The Garage’ 4. ‘The Venus Model’ 5. ‘Bad Girl’
"Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."	‘Cleopatra’	1. ==‘Cleopatra’== 2. ‘The Magician’ 3. ‘Reaching for the Moon’ 4. ‘The Road to Romance’ 5. ‘The Soul of Broadway’
"Denis Gage Deane-Tanner"	'Captain Alvarez'	1. ‘Caught Plastered’ 2. ‘The Lady Refuses’ 3. ‘The Cheat’ 4. ‘The Man of Possession’ 5. ‘The Pagan’

RAG

Query	Ground Truth Movie	Your Results
"Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions"	'Nanook of the North'	Description: Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions Model Suggestions: Based on the provided movie descriptions and the user query, here are 5 suggested movies: 1. Nanook of the North 2. The Shriek of Araby 3. The Way of All Men 4. The Barker 5. Uncharted Seas
"Western romance"	'The Lucky Horseshoe'	Description: Western romance Model Suggestions: Based on the provided movie descriptions and the user query, here are 5 suggested movies: 1. Bucking Broadway 2. Golden Rule Kate 3. The Rogue Song 4. The Cavalier 5. Love Never Dies
"Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."	'Sahara'	Description: Silent film about a Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo. Model Suggestions: Based on the plot description, here are 5 suggested movies that may be similar to the USER_QUERY: 1. Sahara 2. Foolish Wives 3. Through the Back Door

		4. Long Pants 5. A Child for Sale
"Comedy film, office disguises, boss's daughter, elopement."	'Ask Father'	Description: Comedy film, office disguises, boss's daughter, elopement. Model Suggestions: Based on the provided movie plot and user query, here are 5 suggested movies: 1. Amarilly of Clothes-Line Alley 2. ==Ask Father== 3. The Guardsman 4. The Saturday Night Kid 5. The Front Page
"Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."	'Cleopatra'	Description: Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria. Model Suggestions: Based on the provided movie plot and user query, here are 5 suggested movies: 1. ==Cleopatra== 2. General Crack 3. For Better, For Worse 4. Disraeli 5. Revenge
"Denis Gage Deane-Tanner"	'Captain Alvarez'	Description: Denis Gage Deane-Tanner Model Suggestions: Based on the similarities between the USER_QUERY and the MOVIE_PLOT, here are 5 suggested movies: 1. ==Captain Alvarez== 2. Branded 3. The Wolf Song 4. The Spoilers 5. West Point

Experiment Results:

	Recall@1	Mean Reciprocal Rank
Semantic Search	0.333	0.389
BM25	0.5	0.556
Reranker	0.667	0.667
RAG	0.667	0.75