DSBA6188: Homework 3 (8 points)

Scenario:

Imagine you are a historian specializing in early 20th-century cinema, working with a database of films from the 1920s. This collection includes a vast array of silent films, early talkies, and other cinematic works that includes plot descriptions. Your task is to help users, who are scholars, students, and enthusiasts of vintage cinema, navigate this extensive archive to find the specific films they are interested in. Your challenge is to develop a search tool that facilitates an efficient exploration of this rich historical movie database, enhancing users' understanding and appreciation of 1920s cinema.

Your Task:

Your goal is to develop a prototype (POC/MVP) for a search tool that helps users discover relevant movies and shows based on their queries. Your goal isn't to develop a model ready for production, but develop a possible prototype and make a recommendation to your boss for what direction to move towards a production tool.

Dataset:

Use this dataset containing movie descriptions (Plot), genres, actors, and other relevant information. You should focus on querying the text (Plot) and Title. To get the 1920s movies, we'll only look at the frst 1,000 movies. Use this to pull the data:

```python
from datasets import load_dataset


ds = load_dataset("Coder-Dragon/wikipedia-movies", split='train[:1000]')
```

> 	Use only the frst 1,000 examples just like above (the frst 1,000 movies from 1920 or earlier). If you don't follow these instructions, you'll lose 0.5 points.

Part 1: Develop a Baseline Semantic Search Model (2 points total):

- Read this on Semantic Search with sentence transformers.

- Modify this notebook for the data above. Like in the notebook, index by both the Title and Plot (-0.2 points if you don't do this). Use Colab with GPU T4 for runtime (0.5 points).
- Fill in the blue column in the table (0.6 points). You do not need to put this table into the notebook as it will be in the writeup Appendix.

Query	Ground Truth Movie	Your results
"Documentaries	'Nanook of the North'	**1.**0.454 Nanook of the North The documentary follows the lives of an Inuk, Nanook, and his family as they travel, search for food, and trade in the Ungava Peninsula of northern Quebec, Canada. Nanook; his wife, Nyla; and their family are introduced as fearless heroes who endure rigors no other race could survive. The audience sees Nanook, often with his family, hunt a walrus, build an igloo, go about his day, and perform other tasks. **2.** 0.291 The Homesteader The Homesteader involves six principal characters, the leading one being Jean Baptiste (Charles Lucas), a homesteader far off in the Dakotas, living where he alone is black. To this wilderness arrives Jack Stewart, a Scotsman, with his motherless daughter, Agnes (Iris Hall). In Agnes, Baptiste meets the girl of his dreams. Agnes, however, does not know that she is not white. Peculiar fate threw her in the company of the Homesteader, but their

		love is forbidden by the custom of the country. Baptiste eventually sacrifices the love of this girl of his dreams, goes back to his own people and marries the daughter of a preacher. McCarthy, the embodiment of vanity, deceit and hypocrisy, really admires the marriage his daughter has made. He speaks of the "rich" young man she has married, praises him to the highest. Baptiste does not know, however, that McCarthy requires and is in the habit of having people praise him. Baptiste does not do it because he is not of the temperament to do so. Because of this failure grows the tragedy of mismarriage to Orlean (Evelyn Preer), a sweet girl, kind and good, but like her mother, without the strength of her convictions. Baptiste, Orlean having failed him, is persecuted by McCarthy and by Ethel (McCarthy's other daughter), who, like her father, possesses all the evil a woman is capable of; she is married to weak-kneed Glavis. In the end, Orlean, driven insane by the evil she had been the innocent cause of, rights a wrong which causes Baptiste to go back to his land in the Dakotas, where he finds the girl he first discovered. Later, he learns the truth about her race and the story has a beautiful ending.[1]

		3. 0.234 The Goddess of Lost Lake The story is about a young woman who is a quarter Native American Indian, Mary Thorne (Glaum), who returns to the home of her prospector father, Marshall Thorne (Dowling), after completing her education in the East. She has a college degree and an air of refinement. While her father is away hunting for gold at Lost Lake, Mary enjoys the freedom of his mountain cabin. When two hunters on a hunting expedition, Mark Hamilton (Butt) and Chester Martin (Mack), show up and visit the cabin she decides to put on Indian clothing and pretend she is a full-blooded Indian princess for fun. Both men are attracted to the Indian maiden and Hamilton falls deeply in love with her. Martin, however, is contemptuous of her Indian background. When Mary hears him making derisive remarks about the Indian race, she returns to her father's cabin. Martin follows her home, enters her bedroom, and attacks her. Hamilton comes to her rescue and prevents Martin from raping her. He then looks around the room and sees the modern decor. Realizing that Mary is a young woman of culture and education, he becomes angry because she fooled him and leaves. Meanwhile, while Mary's father is searching for gold, which legend has it is at the bottom of Lost Lake, a legend that also says a

white man who once stole some of the gold killed an Indian prince and a white man's blood must fall before anymore gold can be taken, he is killed by an Indian guard at Lost Lake. Mary inherits the gold that her father discovered. Hamilton, who cannot forget her, comes back and they are married.

4.   0.214 Captain January Captain January (Baby Peggy) is a young girl who lives in a lighthouse in Maine with her guardian, Jeremiah "Daddy" Judkins (Hobart Bosworth). Judkins, who is the lighthouse keeper, rescued January from a shipwreck when she was an infant. The only clue to the baby's identity was a locket with a photograph of a woman around her neck, so Judkins adopted her as his own daughter. January helps Judkins with his tasks around the lighthouse. As Judkins' heart begins to fail and his health worsens, these tasks become increasingly more complicated and important. In one instance, January must ascend to the top of the lighthouse by herself to light the lamps. The local townsfolk become skeptical of Judkins' ability to care for the girl, and try to have her taken away. January is saved from the orphanage by a chance meeting with Isabelle Morton (Irene Rich), an affluent young woman who comes to visit the lighthouse. She believes that January looks

		familiar; when she sees the photograph in the locket, she identifies January as her late sister's child. Isabelle wishes to adopt January and reunite her with her blood relatives. Faced with his poor health and the scrutiny of the townspeople, Judkins agrees. However, the girl is miserable in her new surroundings, runs away, and finds her way back to the lighthouse. Judkins and the Morton family finally devise a means to make everyone happy: January returns to the Mortons, and Judkins is employed on the family's yacht, ensuring that he will always be able to visit his former daughter. **5.** 0.205 In the Land of the Head Hunters The following plot synopsis was published in conjunction with a 1915 showing of the film at Carnegie Hall:
showcasing indigenous peoples' survival and daily life in Arctic regions"		
"Western romance"	'The Lucky Horseshoe'	1.0.421 The River Allen John Spender (Charles Farrell) is a virile outdoorsman and Rosalee (Mary Duncan) is his high society sweetheart.[3] **2.** 0.358 Don't Bet on Women On a whim, Herbert Blake proposes a wager with Roger Fallon that he won't be able to get a kiss during the coming 48 hours from the next woman who happens to walk into the room. Fallon takes the bet, whereupon the woman who turns up is Herbert's wife. **3.** 0.343 The Deciding Kiss As described in a film magazine,[3]

Eleanor Hamlin (Roberts), who has been living with an old and impoverished couple, is adopted by two couples, Mr. and Mrs. Sears and Beulah Page (Greenwood) and Peter Bolling (Unterkircher), young people who have read of cooperative parenting and wish to try out the theory. It works very well until Jimmy Sears (Cooley) loses control of himself under the spell of his adopted daughter's kisses. This passes, however, but then Peter falls in love with her. Beulah then tells Eleanor that she is engaged to Peter, and the heart-broken little girl goes back home. After an exhaustive search, Peter fails to find her, and he and Beulah complete their engagement. Eleanor returns, sees the true state of things, and asks God to let her be always their little girl.

4. 0.334 Salomy Jane Rough-and-tumble Gold Rush-era California: a woman (Salomy Jane) is saved from a ruffian (Red Pete) by a heroic stranger (Jack Dart), the latter saved from a lynching when falsely accused of a crime.

5. 0.333 When Knighthood Was in Flower Mary Tudor, Queen of France (Marion Davies), the younger sister of King Henry VIII (Lyn Harding), falls in love with commoner Charles Brandon, 1st Duke of Suffolk (Forrest Stanley). There are other plans for Mary, however; she is supposed to make a politically strategic marriage to the elderly King Louis XII of France (William Norris). Brandon is framed for murder, but Mary, disguised as a boy, helps him to escape. Henry tracks down his sister and her lover at a Bristol Inn, and Mary agrees to wed the French king if Brandon's life is spared. After Brandon is exiled, Mary goes ahead with the wedding, but King Louis, in his attempt to prove he is lively enough for such a pretty young bride, drops dead. His nephew and heir to the throne, Francis (William Powell), wants to wed Mary, but Brandon comes to the rescue. When Henry discovers that his sister and Brandon have married, he remarks, "I should

		have consented in the first place, and saved us all this trouble."
"Silent film about a	'Sahara'	1.0.576 Little Lord Fauntleroy In a shabby New York City side street in the mid-1880s, young Cedric Errol lives with his mother (known only as Mrs. Errol or "Dearest") in genteel poverty after the death of his father, Captain Cedric Errol. One day, they are visited by an English lawyer named Havisham with a message from young Cedric's grandfather, the Earl of Dorincourt, an unruly millionaire who despises the United States and was very disappointed when his youngest son married an American woman. With the deaths of his father's elder brothers, Cedric has now inherited the title Lord Fauntleroy and is the heir to the earldom and a vast estate. Cedric's grandfather wants him to live in England and be educated as an English aristocrat. He offers his son's widow a house and guaranteed income, but he refuses to have anything to do with her, even after she declines his money. However, the Earl is impressed by the appearance and intelligence of his American grandson and is charmed by his innocent nature. Cedric believes his grandfather to be an honorable man and benefactor, and the Earl cannot disappoint him. The

		Earl therefore becomes a benefactor to his tenants, to their delight, though he takes care to let them know that their benefactor is the child, Lord Fauntleroy. Meanwhile, back in New York, a homeless bootblack named Dick Tipton tells Cedric's old friend Mr. Hobbs, a New York City grocer, that a few years prior, after the death of his parents, Dick's older brother Benjamin married an awful woman who got rid of their only child together after he was born and then left. Benjamin moved to California to open a cattle ranch while Dick ended up in the streets. At the same time, a neglected pretender to Cedric's inheritance appears in England, the pretender's mother claiming that he is the offspring of the Earl's eldest son, Cedric, Sr.. The claim is investigated by Dick and Benjamin, who come to England and recognize the woman as Benjamin's former wife. She flees, and the Tipton brothers and the pretender, Benjamin's son, do not see her again. Afterwards, Benjamin goes back to his cattle ranch in California where he happily raises his son by himself. The Earl is reconciled to his American daughter-in-law, realizing that she is far superior to the impostor. The Earl planned to teach his grandson how to be an aristocrat. Instead, Cedric teaches his grandfather that an aristocrat should practice

| | | compassion towards those dependent on him. The Earl becomes the man Cedric always innocently believed him to be. Cedric is happily reunited with his mother, and Mr. Hobbs, who decides to stay to help look after Cedric.
2. 0.566 Sahara Silent film femme fatale, Louise Glaum, portrays the role of Mignon, a Parisian music hall celebrity. Mignon marries a young American civil engineer, John Stanley, portrayed by Matt Moore. Stanley is transferred to Egypt to work on an engineering project in the Sahara. Mignon and her son, portrayed by Pat Moore, join Stanley in the desert.[3][4] Unhappy with life in the desert, Mignon leaves Stanley and her son in the desert and moves to Cairo with the wealthy Baron Alexis, portrayed by Edwin Stevens. Mignon lives in Baron Alexis' palace while Stanley goes blind and becomes addicted to the drug hasheesh. Mignon later encounters Stanley and her son, who have become beggars in the streets of Cairo.[3][4] Mignon returns to the desert to care for her husband, and the two are reconciled.
3. 0.558 A Child for Sale Charles Stoddard (played by Hale) is a poor artist living with his wife and two children in Greenwich Village. Destitute after his wife dies, he is forced to sell one of his children for $1,000 to a childless rich woman. He soon comes to his senses however, and |

		backs out of the deal. From there, the story takes a number of twists and turns involving Ruth Gardner (Leslie) (the wife of Dr. Gardner who treats Stoddard's child for illness) and Ruth's parents -- whose father is also Stoddard's landlord and mother is later revealed to be Stoddard's long-lost mother from a prior marriage.[3]
		4. 0.558 A Woman of Paris Marie St. Clair and her beau, aspiring artist Jean Millet, plan to leave their small French village for Paris, where they will marry. On the night before their scheduled departure, Marie leaves her house for a rendezvous with Jean. Marie's stepfather locks her out of the house, telling her to find shelter elsewhere. Jean invites Marie to his parents' home, but his father also refuses to let her stay. Jean escorts Marie to the train station, and promises to return after going home to pack. When he arrives at home, he discovers his father has died. When Jean telephones Marie at the station to tell her they must postpone their trip, she gets on the train without him. One year later in Paris, Marie enjoys a life of luxury as the mistress of wealthy businessman Pierre Revel. A friend calls and invites Marie to a raucous party in the Latin Quarter. She gives Marie the address but can't remember whether the apartment is in the building on the right or

		the left. Marie enters the wrong building and is surprised to be greeted by Jean Millet, who shares a modest apartment with his mother. Marie tells Jean she would like for him to paint her portrait and gives him a card with her address. Jean calls on Marie at her apartment to begin the painting. Marie notices he is wearing a black armband and asks why he is in mourning. Jean tells Marie his father died the night she left without him. Marie and Jean revive their romance, and Marie distances herself from Pierre Revel. Jean finishes Marie's portrait, but instead of painting her wearing the elegant outfit she chose for the sitting, he paints her in the simple dress she wore on the night she left for Paris. Jean proposes to Marie. Jean's mother fights with him over the proposal. Marie arrives unexpectedly outside Jean's apartment just in time to overhear Jean pacify his mother, telling her that he proposed in a moment of weakness. Jean fails to convince Marie he didn't mean what she overheard, and she returns to Pierre Revel. The following night, Jean slips a gun into his coat pocket and goes to the exclusive restaurant where Marie and Pierre are dining. Jean and Pierre get into a scuffle, and Jean is ejected from the dining room. Jean fatally shoots himself in the foyer of the restaurant.

| | | The police carry Jean's body to his apartment. Jean's mother retrieves the gun and goes to Marie's apartment, but Marie has gone to Jean's studio. Jean's mother returns and finds Marie sobbing by Jean's body. The two women reconcile and return to the French countryside, where they open a home for orphans in a country cottage. One morning, Marie and one of the girls in her care walk down the lane to get a pail of milk. Marie and the girl meet a group of sharecroppers who offer them a ride back in their horse-drawn wagon. At the same time, Pierre Revel and another gentleman are riding through the French countryside in a chauffeur-driven automobile. Pierre's companion asks him, "What ever happened to that Marie St. Clair?" Pierre replies that he doesn't know. The automobile and the horse-drawn wagon pass each other, heading in opposite directions. **5.** 0.552 The Lady A young woman marries the wastrel son of a British aristocrat. Her husband, who has been disinherited by his father, loses what little money he has left gambling in casinos and then dies, leaving her penniless and with an infant son. When her former father-in-law tries to get custody of the child, she leaves him with a couple she trusts, but when she later goes to reclaim her son, she can't find the people she left him with. |

Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo."		
"Comedy film, office	'Ask Father'	**1**.0.632 Mabel's Blunder Mabel's Blunder tells the tale of a young woman who is secretly engaged to the boss's son.[1] The young man's sister comes to visit at their office, and a jealous Mabel, not knowing who the visiting woman is, dresses up as a (male) chauffeur to spy on them. **2.** 0.589 The Saturday Night Kid Set in May 1929, the film focuses on two sisters - Mayme (Clara Bow) and Janie (Jean Arthur) - as they share an apartment in New York City. In daytime, they work as salesgirls at the Ginsberg's department store, and at night they vie for the attention of their colleague Bill (James Hall) and fight over Janie's selfish and reckless behavior, such as stealing Mayme's clothes and hitchhiking to work with strangers. Bill prefers Mayme over Janie and constantly shows his affection for her. This upsets Janie, who schemes to break up the couple. One day at work, Bill is promoted to floorwalker, while Janie is made treasurer of the benefit pageant. Mayme, however, is not granted a promotion, but gets heavily criticized for constantly being late at work by the head of

		personnel, Miss Streeter (Edna May Oliver). **3.** 0.582 Road to Paradise Loretta Young plays the part of an orphan who has been raised by two thieves (Raymond Hatton and George Barraud) and does not know that she has a twin sister who is now a wealthy socialite (Loretta Young as Margaret Waring). One day, while she is dining at a Chinese restaurant with her two guardians, they notice the wealthy socialite and are taken aback at how closely she resembles Young. Hatton and Barraud convince Young that she should impersonate the socialite so that they can enter her house and steal the contents of her safe. Young enters the house and meet Jack Mulhall who senses something different about Waring and immediately falls in love with Young. When night falls, Young lets Hatton and Barraud into the house and they attempt to open the safe. Waring happens to enter the house and is shocked to find a woman that looks like her. She is wounded by Barraud and Young tricks the police into thinking that Waring is an imposter and thief. Even though Mulhall knows the truth, he keeps quiet because he is in love with Young. Eventually Young discovers that Waring is her twin sister when they discover that they have matching lockets. The charges against Waring are dropped and Young accepts Mulhall's proposal of marriage.

		4. 0.577 Which Woman? As described in a film magazine,[2] Doris Standish (Hall), being forced into an unwanted marriage with an aged millionaire, follows the advice of a maid and jumps into a waiting automobile driven by Jimmy Nevin (Sutherland). After an automobile accident that wrecks the car, Doris and Jimmy seek refuge from a storm in a barn. To this same barn come the butler and maid with the stolen wedding presents. Doris transposes bags and goes to a rooming house with Jimmy, but the crooks follow. Doris escapes, but before she can warn her uncle and the millionaire, they are trapped by the crooks. Doris returns to the rooming house and is followed by the police. The crooks are arrested. Jimmy asks the uncle for Doris' hand and the millionaire gives his blessing. **5.** 0.569 On Your Back Putting a son through college, Julianne, owner of a Fifth Avenue dress shop in New York City, is persuaded to supplement her income by providing loans to struggling showgirls. The plan backfires when her son Harvey falls for her business partner's lover Jeanne Burke, who blackmails Julianne.
disguises, boss's daughter, elopement."		
"Lost film, Cleopatra	'Cleopatra'	**1.**0.593 Cleopatra Because the film has been lost, the following

		summary is reconstructed from a description in a contemporary film magazine. Cleopatra (Bara), the Siren of Egypt, by a clever ruse reaches Caesar (Leiber) and he falls victim to her charms. They plan to rule the world together, but then Caesar falls. Cleopatra's life is desired by the church, as the wanton woman's rule has become intolerable. Pharon (Roscoe), a high priest, is given a sacred dagger to take her life. He gives her his love instead and, when she is in need of some money, leads her to the tomb of his ancestors, where she tears the treasure from the breast of the mummy. With this wealth she goes to Rome to meet Antony (Hall). He leaves the affairs of state and travels to Alexandria with her, where they revel. Antony is recalled to Rome and married to Octavia (Blinn), but his soul cries out for Cleopatra. He sends her a message to arm her ships and meet him at Actium, where they battle the opposing forces. They are overpowered, and flee to Alexandria. There they are captured by Octavius (De Vries), and Antony dies in Cleopatra's arms. Before Cleopatra is to be dragged behind the wheels of Octavius' chariot, Pharon the priest, who has never ceased to love her, brings her the serpent that she joyously brings to her breast, dying royally with her crown on her head and

		scepter in her hand as becomes Egypt.
		2. 0.462 Reaching for the Moon As described in a film magazine,[3] Alexis Caesar Napoleon Brown (Fairbanks) learns that his mother was a great princess from the European province of Vulgaria but became an outcast because she did not marry royal blood. Alexis believes that if one concentrates on one thing long enough, it will come true. He is continuously concentrating on the idea some day he will be king of Vulgaria. He tells his ambitions to the girl of his dreams (Percy), who is the "patient listener." After one of his conferences with his patient sweetheart he goes home and dreams he is king of Vulgaria. On all sides his life is threatened by Black Boris (Campeau), who aspires to the throne. Arrangements are made for him to marry the Princess Valentina, but after one glance at her he is ready to run away. However, he is persuaded to remain and it becomes necessary for him to fight a duel with his rival Boris. Alexis, not knowing how to use a sword, puts up a poor fight and after a short struggle is sent flying down a steep precipice. It becomes steeper and steeper until Alexis awakens, having fallen out of bed. He is cured of his desires and is happy in his little home in New Jersey with his "patient listener" as Mrs. Alexis Caesar Napoleon Brown and

		a two-year old to pass his time with.
		3. 0.459 A Splendid Hazard The main charter Karl Breitman played by Henry B. Walthall, thinks he is a descendant of Napoleon and tries bring back to France the French monarchy. As part of his plot he courts Hedda Gobert played by Rosemary Theby as she owns some Napoleon's papers. After winning Hedda haert he takes the documents from she. He travels to America to visit Admiral Killigrew played by Hardee Kirkland. He hopes the stolen papers will lead him to Napoleon wealth. He finds a treasure map in the Admiral's home and then travels to Corsica. Before finding the Napoleon wealth, he comes across someone that mocks him. He challenges them to a duel. In the duel he is mortally wounded. He dies at his love side, Hedda.[3]
		4. 0.449 My Lady's Garter As described in a film magazine,[4] a jeweled garter with an interesting history disappears under mysterious circumstances from the British Museum. The Hawk, a criminal who has never been apprehended even though he obligingly leaves many clues for the police to follow, is suspected. Helen Hamilton (Breamer), daughter of a wealthy American, loses her jewels after throwing them out of a window at Keats Gaunt (Craig), a poet she imagines she is in love with. A tiff with Gaunt follows and she dives into the sea, being

		rescued by a strange gentleman in a yacht who gives his name as Bruce Calhoun (Standing). English detectives suspect him of the robbery and watch him closely. He goes to Helen's home and becomes acquainted with her family, but his mysterious actions raise doubts in the minds of all save Helen, who now loves him. Not even to her, however, will he admit his part in the mysterious proceedings that are occurring continuously until, by a master stroke, he catches the criminal, a rival for Helen's affections, and then reveals that he is an American secret service man and worthy of her love. **5.** 0.442 The Sorrows of Satan Adolphe Menjou stars as Prince Lucio de Rimanez, who is in fact really Satan assuming a human form. When struggling writer Geoffrey Tempest (Ricardo Cortez) is moved to curse God for his misfortunes, Prince Lucio makes a sudden appearance, informing Tempest that he has inherited a fortune. The only proviso is that Tempest must place his fate entirely in the Prince's hands. As he ascends to the uppermost rungs of European society, Tempest is ordered by Lucio to marry Russian Princess Olga (Lya De Putti), even though the writer still loves his sweetheart Mavis Claire (Carol Dempster). Eventually, Prince Lucio reveals his true identity,

		but not before Olga has committed suicide. After rejecting the devil and all his false promises, Tempest lives happily ever after with Mavis.
charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."		

"Denis Gage Deane-Tanner"	'Captain Alvarez'	1.0.320 His Hour Gritzko (John Gilbert) is a Russian nobleman and Tamara (Aileen Pringle) is the object of his desire. 2. 0.272 Captain Alvarez A melodrama about an American who becomes a revolutionary

leader battling evil government spies in Argentina. William Desmond Taylor portrays the title role, and Denis Gage Deane-Tanner, Taylor's younger brother, is thought to have played the small role of a blacksmith.
3. 0.249 Mr. Fix-It As described in a film magazine,[3] because of his ability to fix things Dick Remington (Fairbanks) becomes known as "Mr. Fix-It" and enters the aristocratic home of the Burroughs as their nephew. Before long he has melted the stone hearts of three aunts and one uncle and won the heart of Mary McCullough (Hawley) in addition to setting aright the affairs of pretty Georgiana Burroughs (MacDonald) and Olive Van Tassell (Landis).
4. 0.248 The Black Arrow: A Tale of the Two Roses The novel is set in the reign of "old King Henry VI" (1422–1461, 1470–1471) and during the Wars of the Roses (1455–1487). The story begins with the Tunstall Moat House alarm bell, rung to summon recruits for its absent lord Sir Daniel Brackley, to join the Battle of Risingham; at which the outlaw "fellowship" known as "the Black Arrow" begins to strike with its "four black arrows" for the "four black hearts" of Brackley and three of his retainers: Nicholas Appleyard, Bennet Hatch, and Sir Oliver Oates, the parson. The rhyme posted in explanation of this attack, makes the protagonist Richard ('Dick') Shelton, ward of Sir Daniel, curious about the

death of his father Sir Harry Shelton. Having been dispatched to Kettley, where Sir Daniel was quartered, and sent to Tunstall Moat House by return dispatch, he falls in with a fugitive, Joanna Sedley, disguised as a boy with the alias of John Matcham: an heiress kidnapped by Sir Daniel to obtain guardianship over her and to retain his control over Richard by marrying her to him. As they travel through Tunstall Forest, Joanna tries to persuade Dick to turn against Sir Daniel in sympathy with the Black Arrow outlaws, whose camp they discover near the ruins of Grimstone manor. The next day they are met in the forest by Sir Daniel himself, disguised as a leper and returning to the Moat House after his side was defeated at Risingham. Dick and Joanna then follow Sir Daniel to the Moat House. Here Dick confirms that Sir Daniel is the murderer of his father, and escapes injured from the Moat House. He is rescued by the outlaws of the Black Arrow.
The second half of the novel, Books 3–5, tells how Dick rescues Joanna from Sir Daniel with the help of both the Black Arrow fellowship and the Yorkist army led by Richard Crookback, the future Richard III of England. It centres on Shoreby, where the Lancastrian forces are entrenched. Robert Louis Stevenson inserts seafaring adventure in chapters 4–6 of Book 3, wherein Dick and the outlaws steal a ship and attempt a seaside rescue of Joanna. They are unsuccessful, and after Joanna is moved to Sir Daniel's main quarters in Shoreby, Dick visits her in the guise of a Franciscan friar. Stevenson, the populariser of the tales of the Arabian nights, has Dick tell the

tale of Ali Baba and the Forty Thieves in Book 4, chapter 6 to help him escape from the ruined sea captain Arblaster, whose ship Dick and the outlaws had stolen. While shadowing Sir Daniel, Dick and the outlaws encounter another group of spies interested in Joanna. After a skirmish in which the outlaws prevail, Dick finds that he has conquered Joanna's lawful guardian, Lord Foxham, who promises to give Joanna to Dick in marriage after a contemplated seaside rescue. There is irony in Foxham scolding Dick, who is nobly born, for consorting with outlaws when the outlaws are recruited in Dick and Foxham's plans to rescue Joanna. Wounded in the failed seaside rescue, Foxham writes letters of recommendation for Dick to Richard Crookback, whom Dick must find on the outskirts of Shoreby.
Richard Crookback, Duke of Gloucester, makes his appearance in Book 5. As Dick is leaving Shoreby he sees Crookback holding his own against seven or eight Lancastrian assailants, and assists his victory. Dick's accurate knowledge of the Lancastrian forces in Shoreby aid Crookback in winning the battle that he wages later that day. Dick is also successful as one of Crookback's commanders. Crookback knights Dick on the field of battle and, following their victory, gives him fifty horsemen to pursue Sir Daniel, who has escaped Shoreby with Joanna. Dick succeeds in rescuing Joanna, but loses his men in the process. He, Joanna, and Alicia Risingham travel to Holywood where he and Joanna are married. In this way he keeps his initial pledge to Joanna to convey her safely to Holywood.

In the early morning of his wedding day Dick encounters a fugitive Sir Daniel trying to enter Holywood seaport to escape to France or Burgundy. Because it is his wedding day, Dick does not want to soil his hands with Sir Daniel's blood, so he simply bars his way by challenging him either to hand-to-hand combat or alerting a Yorkist perimeter patrol. Sir Daniel retreats, but is shot by Ellis Duckworth (the outlaws' captain) with the last black arrow. Thereafter Sir Richard and Lady Shelton live in Tunstall Moat House untroubled by the rest of the Wars of the Roses. They provide for both Captain Arblaster and the outlaw Will Lawless by pensioning them and settling them in Tunstall hamlet, where Lawless does a volte face by returning to the Franciscan order, taking the name, Brother Honestus.
5.   0.227   Bulldog Drummond Hugh "Bulldog" Drummond, a demobilised British captain bored with civilian life, places a personal advertisement in The Times offering his services for "any excitement". One of the many replies intrigues him: Phyllis Benton claims she is in great danger. He immediately sets out for the Green Bay Inn, where she has reserved some rooms for him. Unable to persuade him to give up this mad adventure, his friend Algy Longworth follows after, dragging Drummond's valet, Danny, along.
Phyllis turns out to be all Drummond had hoped for: beautiful and desperately in need of help. Her wealthy uncle, John Travers, is being treated in a hospital by a Dr. Lakington for a nervous breakdown, but she is sure there is something wrong about the hospital and Dr.

Lakington, and that she is being watched constantly. She runs away when she spots the outline of two eavesdropping men (Algy and Danny), much to Drummond's annoyance. She is caught and taken to Dr. Lakington's Nursing Home by Carl Peterson, Irma and the doctor.

When Drummond follows, he witnesses Travers' unsuccessful attempt to escape. Drummond drives away, but returns stealthily and rescues Phyllis. Sending her off with Algy and Danny, he sneaks back once more and overhears Irma convince the others to stay and try to get Travers' signature on a document transferring securities and jewels to them. Drummond manages to save Travers.

However, he makes a serious error when he takes Travers back to the inn. The villains soon arrive there. Drummond manages to disguise himself as Travers; the crooks take him back, along with Phyllis. When they realise they have the wrong man they threaten to torture Phyllis.

Drummond tells them Travers is hidden at the inn (whereas he is really being driven to London). While Peterson and Irma go to check, Drummond is freed by Phyllis before Lakington can kill him. He strangles the doctor. Drummond disarms Peterson when he returns, but his gang pose as policemen and take him away. Phyllis persuades Drummond to let them go, telling him she loves him.

In the notebook, calculate Recall@1 and Mean Reciprocal Rank (MRR) for the six examples above (0.4 points). Read [this blog](#) for more details. For MRR, only use your top 5 choices.

Individual Metrics:
Query: Documentaries showcasing indigenous peoples' survival and daily life in Arctic regions
Ground Truth: Nanook of the North
Recall@1: 0
Mean Reciprocal Rank (MRR): 0

Query: Western romance
Ground Truth: The Lucky Horseshoe
Recall@1: 0
Mean Reciprocal Rank (MRR): 0

Query: 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.
Ground Truth: Sahara
Recall@1: 0
Mean Reciprocal Rank (MRR): 0

Query: Comedy film, office disguises, boss's daughter, elopement.
Ground Truth: Ask Father
Recall@1: 0
Mean Reciprocal Rank (MRR): 0

Query: Lost film, Cleopatra charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria.

Ground Truth: Cleopatra
Recall@1: 1
Mean Reciprocal Rank (MRR): 0.008620689655172414

Query: Denis Gage Deane-Tanner
Ground Truth: Captain Alvarez
Recall@1: 1
Mean Reciprocal Rank (MRR): 0.023255813953488372

Answer these questions in your notebook in an "Analysis" Section (0.5 points)

● What type of queries tend to do well? Which not so well?

I personally think that the model tended to perform better with shorter, more concise queries. One could argue that this is indicated by the MRR/Recall measures shown for each of our queries, as the ones that performed best were shorter in length than the ones that generated no scores at all. This would make sense, considering the fact that longer queries might cause the model to get confused by all the different requirements it is being given to satisfy in order to find what it thinks to be a suitable match for the given parameters.

● For the queries that the model didn't perform well, what could be two alternative approaches?

Additional approaches to executing these more complex queries could be performing a very basic version of pre-processing by doing a brief skim through the query to see if there is any extraneous or unnecessary information. By taking away any fluff that is in the query, we can provide the model with a more clear picture of what we are hoping to accomplish and make sure our results are as relevant as possible.

Part 2: Reranker notebook (2 points total)

● Read this blog explaining Reranker.

● Modify this Reranker notebook. You'll calculate BM25, recalculate a similar semantic search, then use a Reranker that combines them. (1 point)

- ○ Make sure to modify so it returns the top 5 results for a given query.

- ○ For the bi-encoder, choose the same (`nq-distilbert-base-v1`) as used in the frst notebook.
- ○ Like the previous, use both Title + Plot as your text. You will need to clean up a few spots. Hint: if passage = ["Title of movie", "Plot of movie"], you can run ": ".join(passage) to get "Title of movie: Plot of movie".

- Recalculate the results table and calculate Recall@1 / MRR for both BM25 and Reranker. Add to the Write up Appendix (1 point)

- Recall@1 for BM25: 0.16666666666666666
- MRR for BM25: 0.0
- Recall@1 for Re-ranker: 0.16666666666666666
- MRR for Re-ranker: 0.0

Query	Ground Truth Movie	Your results

"Documentaries	'Nanook of the North'	Top-3 lexical search (BM25) hits
		    21.136
		Indigenous peoples in Brazil: People have inhabited modern-day Brazil for over 10,000 years. They had developed several distinct cultures before the arrival of Europeans in the 1500's.: Most indigenous people died quickly after European contact due to Old World diseases, like smallpox.: Indigenous peoples may face discrimination and denial of rights because of their ethnicity or status.
		    20.581
		Indigenous peoples: Indigenous people means the first people who lived in any region, and not later immigrants.:

		Indigenous people can also be described as aborigines, native people, first people, first nations and autochthonous.: The United Nations prefers the phrase "Indigenous peoples" to these descriptions. The other descriptions sometimes sound negative (pejorative).: Very often, indigenous people were forced into slavery or badly treated by settlers and immigrants. Indigenous people were oppressed in colonial times. Colonial settlers often robbed the indigenous treasures and lands. Indigenous people were often made slaves, or killed by colonial settlers.: Colonialism often wanted indigenous people to believe that they were lower social class than the new settlers. This is propaganda.: Indigenous people can grow up believing that they are lower social class, or less important. Indigenous people sometimes have to fight for their own land, their own history, human rights and equality. 20.245 Austronesian peoples: Austronesian peoples is a term referring to people that live in Southeast Asia, Oceania and Madagascar, who are speakers of the Austronesian languages. They are thought to have originally come from the indigenous peoples of Taiwan.

		Top-3 Cross-Encoder Re-ranker hits
		    0.339 Nanook of the North: Nanook of the North (also known as Nanook of the North: A Story Of Life and Love In the Actual Arctic) is a 1922 American silent documentary movie directed by Robert J. Flaherty set in the Canadian Arctic. It was one of the first twenty-five films included in the National Film Registry by the Library of Congress in 1989.
		    -3.543 Arctic ecology: Arctic ecology is the science that studies the ecology of the arctic. The arctic is all land and water north of the Arctic Circle (66 33'). This area is very cold and has little rain. In the winter there is almost no sunlight. The growing season, when farmers can grow crops, is very short.: The arctic biomes are taiga, boreal forest, and tundra. Many of the ecosystems in the arctic are fragile, and can be easily destroyed. These ecosystems are being hurt by global warming.: The arctic environment includes both land and water. Two important parts of this environment are sea ice and permafrost.: Sea ice is ice that floats in the ocean. It is made of frozen seawater. It is important because animals live and sleep on the ice, especially in winter. Sea ice lasts all year and never melts completely. There is less ice during

		the summer than during the winter.: Lots of land in the arctic is frozen all year long. Permafrost is a substrate (layer of soil) that has been frozen for two or more years. The dirt on top of the permafrost is called the "active layer". This dirt melts during the summer and is important for plants to grow.: In the arctic region there is less food. The algae is eaten by small insects. These insects are eaten by plankton. These are eaten by small fish these are eaten by big fish. Now, these are eaten by seals. Seals are at last eaten by polar bears.
		-3.559 Norman Hallendy: Norman E. Hallendy is an Arctic researcher, writer and photographer. Hallendy's Arctic journeys began in 1958 at a time when many Inuit, who traditionally lived on the land, were moving to permanent settlements created by the Federal government. Returning north year after year, he was befriended by Inuit elders who shared with him their perception of their material and spiritual world.: His previous books include the best selling Inuksuit:: He has been a lecturer at Oxford, Cambridge, UNESCO (Paris) including numerous universities in Canada and abroad. His most enjoyable presentations were to Inuit elders and their children in several Arctic communities.: Norman Hallendy has been named one of Canada's 100

		greatest modern-day explorers by Canadian Geographic Society.: His contribution to geographical knowledge through documenting Arctic landscapes and ancient sites earned him the coveted Mungo Park Medal of the Royal Scottish Geographical Society in 2007 placing him among such esteemed persons as Professor Jean Malaurie, Count Eigil Knuth and Thor Heyerdahl.: In 2001, the Royal Canadian Geographical Society awarded Hallendy the Gold medal for his internationally acclaimed work on inuksuit and ethnographic work in southwest Baffin Island placing him among such outstanding contributors to geography as Michael Palin, Gilbert Grosvenor and Sir Christopher Ondaatje.: His ethnographic documentation of the Inuit of southwest Baffin Island is regarded as a major contribution to the appreciation of Inuit intellectual culture. Dr. William Fitzhugh Director of the Arctic Studies Centre of the Smithsonian Institution regards Hallendy as a modern day Knud Rasmussen.: The Right Honourable Julie Payette Governor General of Canada has awarded Norman the Polar Medal created by Her Majesty the Queen of England in recognition of his ethnography of the Inuit in the Canadian Arctic.
showcasing indigenous peoples' survival and daily		

life in Arctic regions"		

"Western romance"	'The Lucky Horseshoe'	nput question: Western romance
		Top-3 lexical search (BM25) hits
		    19.727 Western Romance languages: The Western Romance languages are a branch of Romance languages. The main languages in the branch are Spanish, French, and Portuguese. The branch has two parts, Gallo-Romance and Iberian Romance.
		    17.419 Italo-Western languages: The Italo-Western languages are the largest branch of Romance languages. They are made up of two branches, the Italo-Dalmation languages and the Western Romance languages.
		    16.236 Romance languages: The Romance languages (also sometimes called Romanic languages) are a language family in the Indo-European languages. They started from Vulgar Latin (in Latin, "vulgar" is the word for "common" and so "Vulgar Latin" means "Common Latin"). The most spoken Romance languages are Spanish, Portuguese, French, Italian and Romanian.: They are called "Romance languages" because they originate from Latin, the language spoken by the Western Roman Empire. Their grammatical inflection system has been simplified and lost most of the complex case structure of classical Latin.: The area

that the Romance languages are spoken in Europe is mostly extent of the Western Roman Empire. The Greek language superseded Latin in the Eastern Roman Empire. Latin survived in Romania, whose language, Romanian, is a Romance language. In Moldova it is sometimes called Moldovan.: The Romance language family is one of the biggest in the world, and in total, there are almost a billion first and second language speakers.: Spanish is the most widely used Romance language, follow[Portuguese]] and then French.: Spanish is spoken mainly in Spain and Latin America. Portuguese is mostly used/spoken in Brazil, Portugal and parts of Africa. French is spoken in France, Belgium, Switzerland, Canada, and in some areas of Africa. Romanian is spoken in Romania and Moldova.

Top-3 Cross-Encoder Re-ranker hits
    8.889 Western Romance languages: The Western Romance languages are a branch of Romance languages. The main languages in the branch are Spanish, French, and Portuguese. The branch has two parts, Gallo-Romance and Iberian Romance.
    4.533 Italo-Western languages: The Italo-Western languages are the largest branch of

		Romance languages. They are made up of two branches, the Italo-Dalmation languages and the Western Romance languages. 　　1.304 Western (genre): The Western is a genre of fiction. Westerns tell stories that are set mostly in the second half of the 19th century in the American Old West, hence the name. There are also stories about Western-type characters set in other times. These include "Junior Bonner" (set in the 1970s) and "The Three Burials of Melquiades Estrada" (set in the 21st century).: The genre has been popular since the beginning of the movie industry. One of the very first commercial movies was the 1903 silent western, "The Great Train Robbery".: There are Western movies, television programs, books, and other things. Here are some examples:
"Silent film about a Parisian star moving to	'Sahara'	op-3 lexical search (BM25) hits 　　24.433　　Sylvain Sylvain: Sylvain Mizrahi (born February 14, 1951), better known as Sylvain Sylvain, is an American rock musician. He is most notable for being a member of the New York Dolls. He was a guitarist and pianist for the group. He was a member from 1971 to 1976. He has since had a solo career since the

band's first split up.
From 2004 to 2011, the New
York Dolls surviving
members regrouped.:
Sylvain was born in Cairo,
Egypt to a Jewish family.
In the 1950s, his family
fled Egypt to live in
France, before moving to
the U.S. state of New
York. In April 2019,
Sylvain made it known to
the public that he had
cancer.

24.067   Cairo
Governorate: Cairo
Governorate is a
governorate of Egypt. Its
capital, the city of
Cairo, is also the
national capital of Egypt.

23.594   Samir
Farid: Samir Farid ()
(December 1, 1943 – April
4, 2017) was an Egyptian
writer, movie critic,
journalist and movie
historian. He was born in
Cairo, Egypt. He authored
and translated over 60
books since 1966 on
Egyptian, Arab and World
cinema. He was the
consultant for Cinema
affairs in the Bibliotheca
Alexandrina (2001-2016).:
He was a member of jury in
International film
festivals since 1972.
(Among them Oberhausen
Film Festival 1978, DOC
Leipzig Film festival
1984, Annecy International
Animated Film Festival
1998, Torino Film Festival
2001, Controcorrente
(Upstream section) of
Venice Film Festival 2003,
Thessaloniki Film Festival
2003 and Taormina Film
Festival 2010.): He also
was the Head of The Cairo
International Film
Festival 2014.: Farid died

on April 4, 2017 in Cairo from a long-illness at the age of 73.

Top-3 Cross-Encoder Re-ranker hits
        -5.224   Silent movie: A silent movie is a movie made without sound. Silent movies were made from the late 1880s to the early 1930s. They used early cameras that could not record sound. Either the movie was being made before sound cameras were invented, or the movie makers could not afford the expensive new equipment.: Charlie Chaplin was known as one of the best movie stars during "The Silent Era", as it's called today in movie history. Some silent stars also did well in sound movies. Popular silent movies include "", "A Trip to the Moon", "The General", and "Metropolis".
        -5.950   Silent Running: Silent Runnning is a 1972 American science fiction movie directed by Douglas Trumbull and starring Bruce Dern, Cliff Potts, Ron Rifkin, Jesse Vint. It was distributed by Universal Pictures.: The style of the film is "apocalyptic". This is a subgenre of science fiction, science fantasy, dystopian or horror film where the Earth's technological civilization is collapsing or has collapsed.

| | | ```
      -6.069     Les
Enfants du Paradis: Les
Enfants du Paradis
("Children of Paradise")
is a French romantic drama
movie. It was directed by
Marcel Carné and released
in 1945. It stars Arletty
as French courtesan
Garance. She is loved by
four men: a street mime,
an actor, a criminal, and
an aristocrat. The story
is set among the lower
class theatres of Paris
between 1820 and 1840.
Garance is courted by four
men but leaves them all
because they will not
allow her to love in her
own way. The three hour
film has been described as
France's "Gone with the
Wind".
``` |
| Egypt, leaving her husband for a baron, and later reconciling after finding her family in poverty in Cairo." | | |
| "Comedy film, office | 'Ask Father' | ```
Top-3 lexical search
(BM25) hits
     32.074     My Boss's
Daughter: My Boss's
Daughter is a 2003
romantic comedy movie in
which Ashton Kutcher,
Molly Shannon and Tara
Reid star. The subject in
this romantic comedy is
researchers working at a
publishing company. One
researcher, Tom, carries a
crush on the daughter of
the boss. This movie was
released in August 2003.
Its budget was $14
million; it got back over
$18 million. Reviews,
however, were rather
substandard. The original
of this movie was rated
``` |

		PG-13, though an R-rated version of this exists.
		    25.230 Elopement (movie): Elopement is a 1951 American romantic comedy movie directed by Henry Koster and starring Clifton Webb, Anne Francis, Charles Bickford, William Lundigan. It was distributed by 20th Century Fox.
		    22.019 Elopement (marriage): An elopement is a marriage done in secret or private. Elopement may be an alternative to a large expensive wedding. It may also be illegal or against religious customs. In some parts of the world, Bali for example, elopement is the preferred form of marriage. Celebrities often elope to escape the paparazzi.
		
		Top-3 Cross-Encoder Re-ranker hits
		    4.823 My Boss's Daughter: My Boss's Daughter is a 2003 romantic comedy movie in which Ashton Kutcher, Molly Shannon and Tara Reid star. The subject in this romantic comedy is researchers working at a publishing company. One researcher, Tom, carries a crush on the daughter of the boss. This movie was released in August 2003. Its budget was $14 million; it got back over $18 million. Reviews, however, were rather substandard. The original of this movie was rated PG-13, though an R-rated version of this exists.
		    3.033 The Boss (2016 movie): The Boss is a 2016 American comedy movie directed by Ben Falcone and written by Falcone, Melissa McCarthy and Steve Mallory.

		The movie stars McCarthy, Kristen Bell, Ella Anderson, Tyler Labine, Kathy Bates, Annie Mumolo, Timothy Simons, and Peter Dinklage.: The movie is about a wealthy woman who, after being arrested and losing her fortune, uses her assistant's daughter's Dandelion Girls cookie sales to return to the top. The movie was released on April 8, 2016, by Universal Pictures. It made $78 million worldwide.

-1.960   It (1927 movie): It is a 1927 silent movie. It is a romantic comedy movie. The story is about a girl who works in a department store. She sets her sights on her handsome and rich boss. The movie's star Clara Bow became known as the "It girl". Dorothy Parker quipped about Bow's character: "It, hell: She had those." The picture was believed lost, but a copy was found in Prague in the 1960s. In 2001, "It" was selected for preservation in the United States National Film Registry by the Library of Congress as being "culturally, historically, or aesthetically significant".: Betty Lou Spence (Clara Bow) has a crush on her handsome boss, Cyrus Waltham, Jr. (Antonio Moreno). He is the new manager of his father's big department store. They belong to different social classes. He is already romantically involved with socialite Adela Van Norman (Jacqueline Gadsden). Cyrus's silly friend Monty (William Austin) notices Betty. She uses him to get closer to Cyrus.: When she finally gets Cyrus's attention, she convinces him to take her on a date to Coney Island. He has a wonderful time. He tries to kiss her on the way home. Betty Lou slaps his face. She hurries out of his car and into her flat. She then peeks out her window at him as he drives away.: Betty bravely proclaims herself as an unmarried mother to protect her sickly roommate Molly (Priscilla Bonner) from having her baby taken away by a couple of meddling

| | | welfare workers. Monty arrives at just the wrong moment, forcing Betty to continue the charade with him. He tells Cyrus. Although he is in love with her, Cyrus offers her an "arrangement" that includes everything but marriage. Betty Lou, shocked and humiliated, refuses, quits her job, and resolves to forget Cyrus. When she learns from Monty about Cyrus's misunderstanding, she fumes, and vows to teach her former beau a lesson.: When Cyrus hosts a yachting excursion, Betty Lou makes Monty take her along, masquerading as "Miss Van Cortland". Cyrus at first wants to remove her from the ship, but he cannot long resist Betty Lou's "it" factor; he eventually corners her and proposes marriage, but she gets him back, by telling him that she'd "rather marry his office boy", which accomplishes her goal, but breaks her heart.: He then learns the truth about the baby, and leaves Monty at the helm of the yacht to go find her. Monty crashes the yacht into a fishing boat, tossing both Betty Lou and Adela into the water. Betty Lou saves Adela, punching her in the face when she panics and threatens to drown them both. At the end of the movie, she and Cyrus reconcile on the anchor of the yacht, with the first two letters of the ship's name, "Itola", between them. Monty and Adela are upset at losing their friends however, it's implied they pursue a relationship with each other as the movie ends.: Elinor Glynn is thought to have invented the concept of "It". Actually, in 1904, R. Kipling, in the short story "Mrs. Bathurst" introduced "It".: In February 1927 "Cosmopolitan" published a two-part serial story in which Glyn defined "It".: Paramount Pictures paid Glyn $50,000 for the concept. They gave her a small part in the movie as herself. She was given a "story and adaptation" credit.: Hope Loring, Louis D. Lighton and George Marion Jr. (titles) wrote the screenplay. Carl Sandburg noted that |

		Glyn's magazine story was "not at all like the film, not like it in any respect.": This is one of the first examples of a "concept film". It is also one of the earliest examples of product placement. The concept of "It" is referred to throughout the movie. In one scene, Glyn appears as herself to define "It" for Mr. Waltham. "Cosmopolitan Magazine" is featured prominently in another scene. In this scene, the character Monty reads Glyn's story and introduces it to the audience.: Stage actress Dorothy Tree had her first movie role in a small, uncredited part. A young Gary Cooper was cast in a minor role as a newspaper reporter.: "It" was a hit with American audiences. It broke box office records. Critics praised the movie. They said Bow was "a joy to behold". addCode addText
disguises, boss's daughter, elopement."		
"Lost film, Cleopatra	'Cleopatra'	```Top-3 lexical search (BM25) hits 43.256 Cleopatra VII: Cleopatra, Queen of Egypt, was one of the most famous women in history. Her full name was "Cleopatra VII Thea Philopator" (69 BC – 12 August 30 BC). She was the last of the Ptolemaic dynasty of Pharaohs set up in Egypt after the death of Alexander the Great. By descent, she was a Macedonian Princess. After her death, Egypt became the Roman province of "Aegyptus".: The main historical source for her life is Plutarch's "Life```

		of Antony", available in translations.: "Antony and Cleopatra" is the famous tragedy by William Shakespeare, believed to have been written sometime between 1603 and 1607. It was first printed in 1623.: Cleopatra was born in Alexandria, then the capital of Egypt. When she was 18 years old, her father, who was king, died. She and her brother, Ptolemy XIII, became the leaders of Egypt. She was queen and her brother was king. Her brother was only 10 years old, so she was the real leader.: Cleopatra made some enemies amongst the courtiers. The reign of Cleopatra was ended by a cabal of courtiers, led by the eunuch Pothinus. They removed Cleopatra from power, as her being a woman, and made Ptolemy sole ruler, in about 51 to 48 BC. She had to leave the country. Ptolemy was king, but because he was still a boy, Pothinus and his friends were the real leaders of Egypt.: At that time, two Roman generals were fighting to be the leaders of the Roman Republic. They were Pompey and Julius Caesar. Julius Caesar defeated Pompey at the Battle of Pharsalus in Greece, 48 BC. Pompey fled to Egypt from Pharsalus, and was murdered by the orders of Ptolemy.: Ptolemy and his supporters thought Caesar would be pleased, but that was a big mistake. Caesar had pardoned many of the senators who fought against him. His reason

		was to lay the ground for peace in Rome. The murder of Pompey cut right across Caesar's plans. Worse, Pompey was a Roman consul, and the widower of Caesar's only daughter, Julia. Worse still, Pompey was beheaded in front of his fifth wife and children, who were on the ship from which he had just disembarked. This made sure that Pompey's powerful family and supporters would never forget, and would always be enemies of Caesar. She ended up marrying both of their brothers.: Then Caesar came to Alexandria, the capital of Egypt. Economically, Egypt was a bread-basket which could feed the people of Rome. Cleopatra received Caesar in her own unique way. She lay down on a carpet. Then her servants rolled up the carpet with Cleopatra inside. Then the carpet was taken to the palace where Caesar was staying. The guards saw some people carrying a carpet and they did not stop them. They carried the carpet to Caesar. Then Cleopatra came out of the carpet. Caesar fell in love with her and they became lovers. At the time, Cleopatra was 21 years old and Caesar was 52 years old.: Caesar decided to make Cleopatra queen again. Ptolemy did not like that. There was a lot of fighting, but Caesar won. Ptolemy tried to leave, but he fell in the river Nile and drowned. Another younger brother became king, but Cleopatra

		was again the real leader of Egypt. Cleopatra did not marry Julius Caesar, but they did have a son. He was called Caesarion.: Soon after Caesarion was born, Caesar was murdered as he walked to the Senate in Rome. Cleopatra and her entourage were in Rome when Caesar was assassinated on 15 March 44 BC. Cleopatra fled back to Egypt.: Mark Antony, who was Caesar's military commander, and Octavian, Caesar's legal heir, defeated the group who had conspired against Caesar. This group was led by Brutus and Cassius, who took their army to the Eastern Mediterranean, and controlled much of that area. The issue was finally settled by the Battle of Philippi, fought in Macedonia.: Caesar's place was taken by three joint Consuls, the Second Triumvirate. Their names were Octavian, Mark Antony and Lepidus. Lepidus was a minor character, and the long-term fate of the Empire was decided by a conflict between Antony and Octavian. Mark Antony came to see Cleopatra, and they fell in love, though Antony already had a wife. Cleopatra had three more children. Two were twins, and the father of all three was Mark Antony.: The conflict between Octavian and Antony became serious, and soon there was war. Cleopatra was wealthy, and she used her money to help Antony. But Octavian won the war. He became the ruler of the Roman Empire and was known

		as Augustus or Caesar Augustus. He took control of Egypt from Cleopatra. After her kingdom was conquered, she committed suicide by letting an asp (a poisonous snake and small cobra) bite her. Antony also committed suicide.: Ceasarion was murdered by Octavian, and her other children were taken as prisoners to Rome. Two of them died of illness, but her daughter, Cleopatra Selene II, married an African King Juba II of Numidia. The couple were sent to Mauritania to rule by Octavian. They lived a full life there, and had children.: There are many stories told of how rich Cleopatra was, but we cannot be sure if these are true. For example, one story says that she took baths in milk, to make her skin softer. Another story says that she made a bet with Antony that she could make the most expensive meal ever. To win the bet, she took a pearl earring, dissolved it in vinegar, and drank it.
		30.678 Antony and Cleopatra: Antony and Cleopatra is a play by Shakespeare. It is a tragedy. Shakespeare's source was Plutarch's "Lives". The play was first performed between 1606 and 1607. It was probably first printed in the First Folio of 1623. The play describes the romantic love and suicides of Antony and Cleopatra.
		28.349 Mark Antony: Mark Antony (Latin: Marcus Antonius,

		14 January "circa" 82 BC – 1 August 30 BC), was a Roman patrician from an upper-class family. He became a general and politician. He was an important supporter of Julius Caesar as a military commander and administrator.: After Caesar's assassination, Antony, left as sole Consul, surrounded himself with a bodyguard of Caesar's veterans. He forced the senate to transfer to him the province of Cisalpine Gaul, at the time administered by Brutus, one of the conspirators. Brutus refused to surrender the province and Antony set out to attack him in the beginning of 43 BC, besieging him at Mutina.: Encouraged by Cicero, the Senate denounced Antony. In January 43 they granted Octavian "imperium" (commanding power), and sent him to relieve the siege. In April 43, Antony's forces were defeated at the battles of Forum Gallorum and Mutina, forcing Antony to retreat to Transalpine Gaul.: News came that Brutus and Cassius were assembling an army in order to march on Rome. Antony, Octavian and Lepidus joined as allies, in November 43 BC, to stop Caesar's assassins. The trio were the Second Triumvirate. Brutus and Cassius were defeated by Antony and Octavian at the Battle of Philippi in October 42 BC. After the battle, a new arrangement was made: while Octavian

returned to Rome, Antony went on to govern the east of the Republic. Lepidus went to govern Hispania (Spain) and the province of Africa. The triumvirate's enemies were subjected to proscription, including Mark Antony's archenemy Cicero, who was killed on 7 December 43 BC.: Antony followed in Caesar's footsteps by going to Egypt and becoming Cleopatra's lover. They had three children together. His absence from Rome allowed the intelligent Octavian to build up support.: The triumvirate broke up in 33 BC, and disagreement turned to civil war in 31 BC. Antony was defeated by Octavian at the naval Battle of Actium and then at Alexandria. He committed suicide, as did his lover, Cleopatra VII of Egypt, in 30 BC.: William Shakespeare wrote a play ""Antony and Cleopatra"" based on this historical event.

Top-3 Cross-Encoder Re-ranker hits
      -3.612   The Mummy (1932 movie): The Mummy is a 1932 horror movie. It was directed by Karl Freund. It stars Boris Karloff as The Mummy, Zita Johann, David Manners, and Edward Van Sloan. The movie is about an Egyptian mummy that returns to life to stalk a woman who is

		the reincarnation of his lost love.
		-4.504 The Mummy (1999 movie): The Mummy is a 1999 American fantasy adventure horror movie. It is set in Egypt in 1926. Stephen Sommers wrote and directed it. The movie stars Brendan Fraser (acting as Rick O'Connell, who is a person who tries to find treasure so he can sell it), and Rachel Weisz (acting as Evelyn 'Evie' Carnahan, who is a person who learns about the history of Egypt) and Arnold Vosloo, acting as the Mummy. The movie has a lot of talking in the old Egyptian language, which was spoken with the help of a group of people who study it for a job. The movie was first going to be to be part of a few low-budget movies but was later given the money it needed to be made into a big Hollywood movie.: "The Mummy" was seen by people for the first time on May 7, 1999. It earned $43 million from being seen in 3,210 places that weekend. It went on to earn $415 million from being seen in total. Other movies were then made. The first was made in 2001. It is called "The Mummy Returns". A TV programme was made called "". A third movie was then first seen by people on August 1, 2008 and is called "".: The movie won Best Make-Up at Saturn Awards.
		-6.859 Hercules and the Lost Kingdom: Hercules and the Lost Kingdom is a 1994 television movie starring

		Kevin Sorbo and Renee O'Connor. Hercules comes to the aid of a young woman who is seeking the lost city of Troy. Eventually, Hercules leads her to a camp of refugees from the city, which has been taken over by Hera's Blue Monks. Hercules helps the refugees take back the city.: Three men are running along the riverside. One of them stops to take a drink from the river. As he stoops down to quench his thirst a female figure composed of water rises out of the river and pulls him under, drowning him. The other two men run on. While running, one man is hit by a boulder and killed. The remaining man escapes. Elsewhere a young woman is collecting water, when the ground begins to shake. She flees back to the village to warn them that a giant is coming. The giant follows the woman to the village and rips the roof of the tavern. Hercules presents himself to the giant and the giant challenges Hercules to a fight. Hercules goes outside and he beats the giant. As the village celebrates the defeat the man from earlier in the film arrives in the village. He asks for Hercules' help. He explains to Hercules that his people have been driven from their city, the lost city of Troy.: Zeus appears and Hercules asks him if he will help him. He asks his father how to find Troy and he says that Hera made it

		vanish. To find it, Hercules must find the one true compass which will point the way to Troy. Hercules heads off to find the compass. He finds a tribe of men preparing to sacrifice a young virgin to their water god. Hercules saves her from being sacrificed, while unknowingly being watched by a mysterious robed figure. The woman tells Hercules that her name is Deianeira. Later in the evening as they camp for the night, Deianeira tells her a story that she is the daughter of a king. She sees the dark figure and asks who he is and why she is being followed, but the figure does not answer. Later Hercules and Deianeira arrive at the slave market looking for Queen Omphale, the last person to have possessed the compass. In order to get to see the queen, Hercules sells himself as a slave, and the queen buys him. After spending the night with the queen, Hercules gets the compass and he and Deianeira continue searching for Troy. Deianeira gets attacked by some men but the figure in the dark robe saves her and tells her to follow her destiny.: Hercules and Deianeira head on to Troy. They arrive at the ocean and the figure standing on the cliffs summons a sea serpent to do Hera's bidding. It swallows Hercules and Deianeira. Hercules kills the monster from inside and they are washed up on the shore.

		Deianeira sees Troy in the distance and tells Hercules that she now remembers and this is where she is from. Deianeira and Hercules get caught in a trap and taken to the king. The king is ill and he and his daughter are reunited. He tells her that the Cult of the Blue Priests have taken over the city and the people have taken refuge in the woods. He tells Deianeira to rule them well, and then dies.: Hercules tutors the people of Troy and prepares them to battle to take back Troy. Deianeira realises that the people can not beat the Cult of the Blue Priests and goes looking for the Blue Priest. The people notice that Deianeira is missing and Hercules and the people use an underground tunnel to get into the city. As the people fight the cult members Hercules goes looking for Deianeira. He finds her about to be sacrificed to Hera and saves her. The Blue Priest and Hercules fight and Hercules cuts off the priest's head. A huge storm approaches, and Zeus tells Hercules that Hera is coming for Deianeira. Hercules saves her and Hera takes Hercules instead. As Deianeira is crowned Queen of Troy we see Hercules thrown down from out of the sky. A man approaches him asking him for help. Hercules agrees and the two men walk off to the next adventure.

Denis Gage Deane-Tanner

```
Top-3 lexical search
(BM25) hits
     17.872      Gage,
Oklahoma: Gage is a town
of Oklahoma in the United
States.
     16.931      Thomas
Gage: General Thomas Gage
(10 March 1718/19 – 2
April 1787) was a British
Army general officer and
colonial official best
known for his many years
of service in North
America, including his
role as British
commander-in-chief in the
early days of the American
Revolution.: Thomas Gage,
on February 20, 1773,
already communicated to
the governor of Louisiana,
Luis de Unzaga y Amézaga
'le Conciliateur', his
intention to return to the
United Kingdom with his
family, a fact that
occurred 4 months later,
in June Therefore, Gage
was not present when the
Boston Tea Party took
place in December of that
year, a city in which both
Gage and Unzaga left
confidants to be informed
by their respective spy
networks.
     16.606      Gage
County, Nebraska: Gage
County is a county in the
U.S. state of Nebraska. As
of the 2010 census, 22,311
people lived there. The
```

```
county seat is Beatrice.
It was founded in 1855.
```

Top-3 Cross-Encoder Re-ranker hits
    -6.296   Reg Dean: Thomas Reginald "Reg" Dean (November 4, 1902 – January 5, 2013) was a British supercenterian who was the second-oldest person in Europe and was the last living person in Britain who was born in 1902. He was the oldest British-born man in England since Claude Choules.: Dean was born on 4 November 1902 in Tunstall, Staffordshire. He was married three times, one marriage ended in the wife's death and the other two ended in divorce. He had one son, Christopher Dean. Dean died on 5 January 2013 in London, England from natural causes, aged 110.
    -7.337   Candace Cameron Bure: Candace Cameron Bure (born April 6, 1976) is an American actress and author. She is best known for playing DJ Tanner in the sitcom "Full House". She played the part from 1987 and until 1995. She has also been in many television movies over the years. She is the sister of Kirk Cameron. Bure was born near Panorama City, California. She married ice hockey player Valeri Bure in 1996. They have 3 children.
    -9.935   Hugh Laurie: Hugh James Calum Laurie (born 11 June 1959) is an English actor, musician and writer who was born in Oxford. He played Dr. Gregory House in the TV series "House MD" and Frederick Little in "Stuart Little" and "Stuart Little 2". He is also famous in Britain for his job with Stephen Fry and for the BBC series "Blackadder".: Laurie was born in 1959 in Oxford and his family was rich. His mother, Patricia (née Laidlaw), died when he was 29 and his father, W.G.R.M. "Ran" Laurie, was a doctor and an athlete who won

the gold medal in rowing 1948 in London. Hugh has two sisters and a brother and all of them are older than he is.: Laurie studied at Eton and then went to Selwyn College, Cambridge: he took a degree in archeology and social anthropology. He was an athlete like his father, but he became an actor after meeting Emma Thompson, who presented him his future best friend and co-worker Stephen Fry.: In 1989, he married Jo Green and now they live in London with their two sons, Charles and Bill, and daughter Rebecca.: All of the following are nominations or wins for Laurie's role on "House":: People's Choice Awards

addCode
addText

Part 3: RAG notebook (1 point total)

- Create an account for [Fireworks.ai](). It's a platform for running open source LLMs. You'll start with a $1 credit, which isn't much, but it'll allow you 1 million tokens / $0.50 (using Mixtral model).

- After creating a new account, go to API Keys in Colab and save your API key.

- Start with [this notebook](). Open the Notebook in Colab and save your own personal copy.

- Modify the notebook to run on the exact same data (first 1000 movies, along with title+Plot). Fill in the table below (blue column) for the RAG output. Include it in your write up Appendix.

Query	Ground Truth Movie	RAG Output
"Documentaries	'Nanook of the North'	Ground Truth Movie: 'Nanook of the North' The Arctic region is a harsh and unforgiving environment, but it is also home to indigenous peoples who have adapted to survive in this extreme climate. Documentaries have played a crucial role in showcasing the daily lives and survival techniques of these communities. From the Inuit people of Nanook of the North to the Sami people of Lapland, these films have provided a glimpse into the unique cultures and traditions of the Arctic. One such documentary is "The People of the Arctic," which follows the lives of several indigenous communities in the region. The film explores their traditional hunting and fishing practices, as well as their cultural beliefs and customs. Through interviews with community leaders and members, the documentary provides a deep understanding of the challenges and triumphs faced by these people in their daily lives. Another documentary, "The Frozen Frontier," focuses on the survival techniques used by indigenous peoples in the Arctic. The film showcases the incredible resourcefulness and resilience of these communities, as they use their knowledge of the land and sea

		to hunt, fish, and gather food. From igloos to snowmobiles, the film highlights the innovative solutions developed by these people to thrive in the harsh Arctic environment. "Arctic Voices" is a documentary that brings together indigenous leaders from across the Arctic region to discuss the challenges and opportunities facing their communities. The film explores issues such as climate change, resource extraction, and cultural preservation, and provides a powerful voice for these communities to be heard on the global stage. "The Arctic Way of Life" is a documentary that delves into the daily lives of indigenous peoples in the Arctic. The film explores the unique challenges and opportunities faced by these communities, from the harsh climate to the rich natural resources available in the region. Through interviews with community members and leaders, the documentary provides a deep understanding of the cultural and social fabric of these communities. "Arctic Survival" is a documentary that focuses on the survival techniques used by indigenous peoples in the Arctic. The film showcases the incredible resourcefulness and resilience of these communities, as they use their knowledge of the land and sea to hunt, fish, and gather food. From igloos to snowmobiles, the film highlights the innovative solutions

		developed by these people to thrive in the harsh Arctic environment.
showcasing indigenous peoples' survival and daily life in Arctic regions"		
"Western romance"	'The Lucky Horseshoe'	Ground Truth Movie: The Lucky Horseshoe Western romance follows the story of a cowboy named Jack who is on a mission to find his missing brother. Along the way, he meets a beautiful woman named Sarah who is also searching for her missing brother. As they journey together, they discover that they have a lot in common and develop a deep connection. However, their journey is filled with danger and obstacles, including a group of outlaws who are also searching for Jack's brother. Despite the challenges, Jack and Sarah's love for each other and their determination to find their missing siblings keep them going. In the end, they are successful in their mission and find their missing brothers, but they also realize that their love for each other has grown stronger and they decide to stay together.
"Silent film about a	'Sahara'	Sahara The film follows the story of a Parisian star, who, after a series of unfortunate events, finds

		herself moving to Egypt. She leaves her husband for a wealthy baron, but soon realizes that life in Egypt is not all it's cracked up to be. She finds herself struggling to make ends meet and must confront the harsh realities of poverty in the city. However, after a chance encounter with a group of kind-hearted Egyptians, she discovers that her family is still alive and living in poverty in Cairo. With the help of her newfound friends, she sets out to find her family and reconcile with her husband. Along the way, she learns valuable lessons about the importance of family and the resilience of the human spirit.
Parisian star moving to Egypt, leaving her husband for a baron, and later reconciling after fnding her family in poverty in Cairo."		
"Comedy flm, ofce	'Ask Father'	Model Suggestions: The story follows the life of a young woman named Sarah who works in a corporate office. She is the daughter of the boss, and her job is to keep the office running smoothly. However, Sarah is not happy with her life. She feels trapped and unfulfilled, and she longs for something more. One day, Sarah meets a man named Jack who works in the same office. They hit it off immediately, and they start to spend more and more time together. Jack is charming and

		funny, and he makes Sarah laugh. However, Sarah's boss is not happy with their relationship. He sees it as a threat to his authority, and he tries to keep them apart. Despite the obstacles, Sarah and Jack continue to see each other in secret. They plan an elopement, and they run away together. They go to a tropical island and live happily ever after. SUGGESTED_TITLES: 1. "Office Elopement" 2. "Corporate Romance" 3. "Boss's Daughter's Secret Love" 4. "Office Disguises" 5. "The Office Escape"
disguises, boss's daughter, elopement."		
"Lost flm, Cleopatra	'Cleopatra'	The story follows the life of Cleopatra, the last pharaoh of ancient Egypt. Cleopatra is depicted as a cunning and ambitious woman who uses her charm and intelligence to rise to power. She is determined to rule the world and will stop at nothing to achieve her goals. Cleopatra's first major move is to seduce Julius Caesar, the Roman general who has conquered much of the known world. Cleopatra uses her beauty and charisma to win Caesar's heart, and the two become romantically involved. Cleopatra uses this relationship to her advantage, using her influence with Caesar to further her own agenda. As Caesar's power grows, Cleopatra begins to plot her own

		rise to power. She uses her knowledge of Roman politics and her charm to manipulate Caesar's allies and enemies, and eventually becomes the ruler of Egypt.
		Cleopatra's reign is marked by her pursuit of wealth and power. She seeks out treasures from the tombs of her ancestors, including the famous mummy of her father, Ptolemy XIII. Cleopatra uses these treasures to fund her military campaigns and to impress her subjects.
		Cleopatra's reign is also marked by her relationship with Mark Antony, a Roman general who becomes her lover and ally. Together, they revel in their love and their power, but their relationship is ultimately doomed. Cleopatra's ambition and her desire for power lead to her downfall, as she is defeated by Octavian, the future Emperor Augustus, and ultimately meets a tragic end in Alexandria.
		SUGGESTED_TITLES:
		1. Cleopatra: The Last Pharaoh 2. Cleopatra's Rise to Power 3. Cleopatra's Treasures 4. Cleopatra and Caesar: A Love Story 5. Cleopatra's Tragic End
charms Caesar, plots world rule, treasures from mummy, revels with Antony, tragic end with serpent in Alexandria."		

"Denis Gage Deane-Tanner"	'Captain Alvarez'	Captain Alvarez
		`Denis Gage Deane-Tanner is`
		`a former Navy SEAL who has`

		been tasked with protecting a high-ranking government official from a group of terrorists. The official is being held captive in a heavily guarded compound in a remote area of the Middle East.
		As Denis and his team of elite operatives prepare for the mission, they discover that the terrorists have a plan to use a highly potent chemical weapon to assassinate the official. Denis must use all of his skills and experience as a Navy SEAL to infiltrate the compound, neutralize the threat, and rescue the official before it's too late.
		Throughout the mission, Denis faces numerous challenges and obstacles, including treacherous terrain, hostile forces, and unexpected betrayals. But with his unwavering determination and commitment to his mission, he is able to overcome these challenges and ultimately succeed in his mission to protect the government official and prevent a catastrophic terrorist attack.

Submission:

There are two deliverable: a 2 page write up (+ Appendix) and a project folder either as a Zip or GitHub repo (**recommended**). Per the syllabus, late submissions receive a 50% deduction.

Deliverable 1:

Analysis Writeup (2 points):

Write an executive summary (up to 2 pages) + Appendix (whatever length)

Recommendation (1 paragraph): Consider the pros/cons of the three models. Make a recommendation to your boss on which model is the most promising. Justify your decision with your experiments.

- **The first pro/con that comes to mind when thinking about the process of building the first semantic search model is the time it took to actually construct/debug the model to be sure it worked appropriately for our intended dataset and use case. This model required a fair deal of pre-processing and hoops to jump through before it would initialize, including having to authenticate with HuggingFace in order to get the dataset to load properly. Considering that this must be done each time the model is loaded or changed, the other models were much quicker in their setup and initialization process.**
- **This is similar to the pros/cons that we experienced when BM25/Retrieval/Reranker models; although this set of models provided a very comprehensive overview of results for each query, it was extremely time consuming to run. The initialization process was fairly quick and required little preprocessing, but the actual model would take anywhere from 5-10 minutes to run completely. While it was nice to have such a vast selection of results from each model in this deliverable and see the advantages/disadvantages of each, it was extremely time consuming to run and therefore not as efficient compared to our other options.**
- **Lastly, the results of our RAG model seemed to be the most promising; although we spent some time debugging and trying to figure out how to pass along the API key that was necessary for this modeling technique, we spent much less time preprocessing/reworking the data to properly compared to our experience with the other two models.**

Production (1 paragraph): Write up considerations for your boss on deployment (e.g., what would be required to put your model into production?). What challenges are you aware of? Does this alter your recommendation?

- **To put each of these models into production, we would firstly need a machine capable of running whichever we chose at a fast/consistent pace. All of these models required a fair bit of computing power, and also employed the use of a GPU in some cases. Although this is fairly basic hardware, it is still worth noting that your average computer would not be capable enough for these models and we would likely want to procure a dedicated server or virtual environment of some sort. As previously mentioned, there may be some initial challenges with debugging or getting the model optimized for our specific use case, though this is to be expected and can be overcome with trial and error as we begin to figure out how to best employ the model. Additionally, it is worth noting that the model may eventually need access to a wider selection of data in order to improve its capabilities and how it can be applied within our specific environment. Retraining the model on a larger selection of data would likely improve its performance, as it would be able to incorporate a wider variety of trends into its predictive patterns and produce more accurate results. Similarly, it is possible that we could run into a situation down the road where we need to address the scalability of our model; for example, the model is currently formatted to accept/process one query at a time. It may be necessary to expand the utilization of the model in the future, which would require more threads/resources to ensure that many different users are able to utilize the model and query the information concurrently. With adequate computing resources and consideration for the potential of scalability built into our model over time, we can ensure that it remains ready to adapt to the particular volumes of queries we expect at a given time and available for any future use case we see fit.**

RAG/LLM's (1 paragraph): Explain your results using RAG/LLMs. How can you evaluate the model's performance? How does it compare with Semantic Search, BM25 and/or ReRanker? What are major factors that affect the model's performance?

- **Our RAG/LLMs results were pretty accurate with our model suggestions and plot when comparing it to the ground truth. The model successfully outputs relevant suggestions that align with the ground truth and paper title. The plot provides a comprehensive narrative that follows the suggested paper title's description and inherited ideas from the ground truth.**
- **We can evaluate the model's performance by looking at the relevance to how closely the generated text aligns with the prompt given in the query. Diverse responses would be an important evaluation, especially with our prompt for suggesting movie titles. Consistency with responses especially with the flow, grammar and coherence is really important when it comes to evaluating a model's performance.**
- **Semantic search uses methods that rely on understanding the meaning and context but might have a problem with large queries, while BM25 is efficient for keyword-based search but again might have trouble with queries. ReRanker is excellent in generating fine-tuned rankings based on queries, they are tailored to fit what input you give them so it gives the user more control over the outputs. RAG/LLMS have the advantage of generating outputs that are consistent with the query input and can understand natural language queries and reproduce humanistic responses.**
- **Major factors that would affect the models performance would be tasks and context when constructing the query. Then we would have to check the data quantity, quality, size, model structured input, fine-tuning, training data, pre-trained models.**

Fine Tuning (1 paragraph): Let's assume your boss likes the results but has an issue with using pre-trained models like `nq-distilbert-base-v1`. She suggests that that model is not fit for your data and heard about this idea called "fine tuning". Explain (do not need to implement) how you would approach fine tuning. Make a recommendation to your boss whether you think it is worth the effort (Hint: search the sentence-transformers docs).

Fine tuning a model specifically for our dataset would be an especially extensive and time consuming process if we are trying to train the model on a large dataset. This would involve procuring a large amount of text relating to the topic we are trying to model on; while this might sound simple in practice, the most laborious part of the process would come from the extensive data cleaning/pre-processing that would need to be performed to ensure that the dataset was easy for the model to interpret and utilize in its predictive patterns. This initial data cleaning process would likely consist of techniques such as stemming, tokenization and other methodologies to help reduce the amount of clutter within the dataset and ensure that our model is utilizing the most relevant information for its intended purposes. While this may pay off with more accurate results or more efficient modeling techniques, it is worth noting that there are a variety of other ways to improve model performance that could ensure more promising results and make sure we are best optimizing the model for its given use case.

Recommendation

If we have a dataset that isn't fit for our model, maybe because of the quality or quantity, we could adapt pre-trained models that might improve the data for the task and context, we could utilize sentence transformers that offer various pre-trained models designed for our output. We could then narrow down our query parameters like number of batches, learning rate, and batch size. Once we get a model performance evaluated after the first go-around, we can go back and make changes depending on the output performance.

Include in the Appendix the table below on the Experiment Results as well as the tables in Part 2, 3 and 4 (you can combine them to improve the layout).

Experiment results:

	Recall @ 1	Mean Reciprocal Rank
Semantic Search		
BM25		

Reranker

You should also include the tables from Parts 1, 2 and 3.

Deliverable 2: Code/repo/folder

Project (1 point):

Organize your project with this folder structure:

> /notebooks
>
> > /01-semantic-search.ipynb
> >
> > /02-reranker.ipynb
> >
> > /03-rag.ipynb
>
> README.md
>
> .gitignore
>
> requirements.txt

Either use a GitHub repository for your project (recommended), including a README with project overview and setup instructions. Alternatively, you may upload a zip fle with the exact folder. In that case, you don't have to include a .gitignore fle.

- Create a requirements.txt that includes all of the packages used; get their exact version and **pin** them; see this link for more details (-0.25 points if you don't do this).

- Do not include in repo or zip fle the virtual environments (hide venv or .venv) (-0.25 points). This is not needed since it can be reproduced with your requirements.txt.

- Do not include zip fles within your zip fle or GitHub repo. Only provide either a link to your GitHub Repo or a Zip fle with the folder above (-0.25 points if you don't follow these instructions)
- Ensure your code is well-documented and explain your logic and approach. Provide clear instructions for running your code and reproducing results, including dependencies and dataset access (points off for poor documentation and/or not reproducible).