ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Εργαστήριο Διοίκησης Πληροφοριακών Συστημάτων

ΔΙΟΙΚΗΣΗ ΨΗΦΙΑΚΗΣ ΕΠΙΧΕΙΡΗΣΗΣ

8ο ΕΞΑΜΗΝΟ

ΔΙΔΑΣΚΟΝΤΕΣ:

ΚΑΘΗΓΗΤΗΣ ΓΡ. ΜΕΝΤΖΑΣ

Υπεύθυνη Άσκησης: Ξένια Παπαδομιχελάκη

email: xpg@central.ntua.gr

Ιούνιος 2021

ΕΡΓΑΣΙΑ 4 ΑΝΑΖΗΤΗΣΗ ΚΑΙ RECOMMENDER SYSTEMS

1. ΑΝΑΖΗΤΗΣΗ – TFIDF

Υποθέστε ότι πραγματοποιείτε μια αναζήτηση σε μια μηχανή αναζήτησης με τους όρους "thriller, mission, planet". Θεωρείστε σαν μοναδικά δεδομένα της μηχανής τις παρακάτω περιγραφές ταινιών.

α. Χρησιμοποιείστε τη μετρική TF-IDF (χωρίς TF normalization) για να βρείτε ποια θα είναι η κατάταξη των αποτελεσμάτων. Να χρησιμοποιηθεί Cosine Similarity και να ληφθούν υπόψη το κείμενο των τίτλων και οι παράγωγες λέξεις (Stemming). Σχολιάστε τα αποτελέσματα.

1.   Sunshine (2007)

R | 1h 47min | Adventure, Sci-Fi, Thriller

As the sun begins to dim along with humankind's hope for the future, it's up to a desperate crew of eight astronauts to reach the dying star and reignite the fire that will bring life back to planet Earth in this tense psychological sci-fi thriller that re-teams 28 Days Later director Danny Boyle with writer Alex Garland and producer Andrew Macdonald. The skies are darkening, and the outlook for planet Earth is grim. A crew of eight men and women has been given a nuclear device designed to literally reignite the sun and sent hurtling through infinity on the most crucial space mission ever attempted. Suddenly, as the crew loses radio contact with mission control, everything begins to fall apart.

2.   Gattaca (1997)

PG-13 | 1h 46min | Drama, Sci-Fi, Thriller

In a futuristic society where commerce has overridden more humanistic concerns, the rich and successful, eager to obtain physical and mental perfection, have taken to genetically engineering their off-spring. Such lab-created babies are known as Valids, while those conceived in the normal, loving fashion are In-Valids and are considered second-class citizens at best. Vincent is an In-Valid while his brother Anton is a Valid. The former brother is short, sickly, and bespectacled, while the latter brother is handsome, healthy and born to succeed. But though Anton seems close to perfection, he lacks the emotional flaws, passion, determination, desire and faith that motivate Vincent, whose strongest desire is to become a space navigator for the Gattaca Aerospace Corporation and travel on an upcoming assignment to the moons of Saturn. Unfortunately, his birth status and a heart defect, relegate him to menial jobs. Unwilling to abandon hope, Vincent determinedly visits DNA broker German (Tony Shalhoub) who is able to create false identities for similar In-Valids. Set in an oppressive, bureaucratic and chillingly plausible early-21st-century world, Andrew Niccol's sci-fi film differs from others in its focus on a morally ambiguous world and on characters rather than gizmos, technobabble and special effects.

3.   Gravity (2013)

PG-13 | 1h 31min | Drama, Sci-Fi, Thriller

Gravity stars Sandra Bullock and George Clooney in a heart-pounding thriller that pulls you into the infinite and unforgiving realm of deep space. Bullock plays Dr. Ryan Stone, a brilliant medical engineer on her first shuttle expedition, with veteran astronaut Matt Kowalsky (Clooney). But on a seemingly routine spacewalk, disaster strikes. The shuttle is destroyed, leaving Stone and Kowalsky completely alone - tethered to nothing but each other and

spiraling out into the blackness. The deafening silence tells them they have lost any link to Earth and any chance for rescue. As fear turns to panic, every gulp of air eats away at what little oxygen is left. But the only way home may be to go further out into the terrifying expanse of space.

4.   Event Horizon (1997)

R | 1h 36min | Horror, Sci-Fi

In this sci-fi/horror scarefest, Dr. William Weir (Sam Neill) is a scientist who has designed a spacecraft called Event Horizon which will explore the outer reaches of space past the planet Neptune; the ship employs a special transport mechanism that, in effect, creates a black hole that the ship can pass through, allowing it to travel tremendous distances in a few seconds. The Event Horizon mysteriously disappears in the midst of a mission with no trace of either the ship or its crew, but it reappears in Neptune's orbit after a seven year absence and it's sending out a distress signal. The spaceship Lewis and Clark, and Dr. Weir, are sent to investigate.
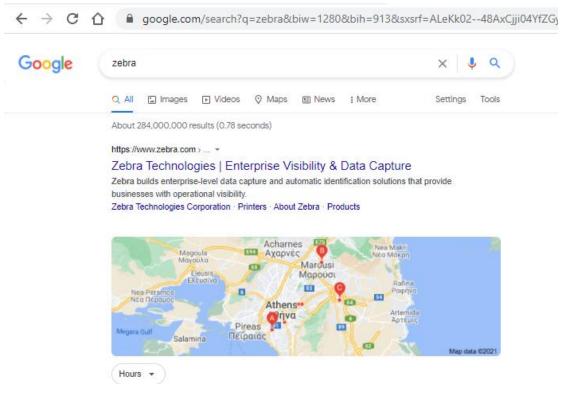
5.   The Arrival (1996)

PG-13 | 1h 55min | Sci-Fi, Thriller

A shy radio telescope operator named Zane Ziminski (Charlie Sheen) picks up a series of regular signals coming from space -- and deliberately pointed toward Earth. Convinced that he has discovered alien transmissions, Ziminski is first chastised and then fired by his boss (Ron Silver). Obsessed, he builds a makeshift radio telescope in his house to find out where the signals were sent. Convinced that they're intended for aliens already hidden on Earth, he tracks them to a bleak, isolated Mexican village, where he joins forces with a female scientist (Lindsay Crouse), who has suspicions of her own after witnessing an acceleration of global warming.

2. ΑΝΑΖΗΤΗΣΗ – PRECISION/RECALL

Ενώ ψάχνετε στοιχεία σχετικά με το ζώο «ζέβρα », η μηχανή αναζήτησης "Google" σας επιστρέφει τα παρακάτω αποτελέσματα.

Zebras - YouTube

One is that they serve as camouflage from the lion, its greatest predator. To a human, the black-and-white ...
1:04 · Jun 12, 2007 · Uploaded by Explore Documentary Films

Zebra | African Wildlife Foundation

Where do plains **zebras** live? They have a wide range in east and southern Africa. They usually live in treeless grasslands and savanna woodlands and are ...

zebra - Wiktionary

NounEdit. **zebra** (plural **zebra** or **zebras**). Any of three species of genus Equus: E. grevyi, E. quagga, or E ...

Zebra Printer | Buy Zebra Label Printers, Mobile Label ...

Buy **Zebra** Label Printers, Mobile Printers, Thermal Transfer Ribbons, Printheads & Labels from **Zebra** Technologies.

Zebras & Libras – ZEBRAS&LIBRAS / SWIMWEAR

Zebras and Libras is a brand inspired by the balance between earth and space. Z&L creates swimwear with high quality textile and new era design cuts.

Zebra εκτυπωτές - Ταμειακές μηχανές

Zebra LP2824 Plus, 8 dots/mm (203 dpi), peeler, RTC, EPL, ZPL, USB, print server (ethernet) (282P-201521-040). Χωρίς ΦΠΑ: 366,89 € Με ΦΠΑ: 454,94 €.

Plains Zebra | National Geographic

Population and Herd Behavior. **Zebras** are social animals that spend time in herds. They graze together, primarily on grass, and even groom one another. Plains ...

Zebra Technologies | LinkedIn

Zebra Technologies | 194845 followers on LinkedIn. **Zebra** empowers those on the front line in retail, health care, transportation and logistics, manufacturing and ...

Zebra Technologies Service & Support | Partners :: Albanis.gr

Zebra is committed to offering our customers the highest quality customer care in the industry, delivering outstanding customer service and providing a global ...

zebra | Size, Diet, & Facts | Britannica

Feb 12, 2021 — **Zebra**, any of three species of strikingly black-and-white striped mammals of the horse family Equidae (genus Equus). All **zebras** are ...

Zebra - Sigma-Hellas.gr

Scanner **Zebra** (Motorola) LS1203 με βάση. Κωδικός: 001034. 74,00 € +φπα. 91,76 € με ΦΠΑ. Αναγνώστης 1D Laser; Ταχύτητα 100 scan/sec; Ενσύρματο (USB) ...

Zebra ZT410 Industrial Barcode/Label Printer | Novatron

Zebra ZT410. Βιομηχανικός θερμικός εκτυπωτής ετικετών-barcode με μέγιστο πλάτος εκτύπωσης 104 mm.

α. Ποια αποτελέσματα αναφέρονται στην αναζήτηση που κάνατε και ποια όχι; Εξηγήστε με συντομία. Ποια από τα αποτελέσματα είναι σωστά(true positive) και ποια λάθος (false positive);

β. Αν γνωρίζουμε ότι υπάρχουν ακόμη 400 αποτελέσματα που σχετίζονται με την αναζήτηση σας και δεν βρέθηκαν – false negative – υπολογίστε τα παρακάτω

i.   Precision

ii.   Recall

iii.   F-Measure

Σχολιάστε τα αποτελέσματα

3. RECOMMENDER SYSTEMS

Έξι άνθρωποι αξιολόγησαν τις ταινίες του ερωτήματος 1 με βαθμολογία από 1 (καθόλου καλή) έως 10 (εξαιρετική) και τα αποτελέσματα φαίνονται στον παρακάτω πίνακα. Αντικαταστήστε όπου X τον τελευταίο ψηφίο του Αριθμού Μητρώου σας, αφού προσθέσετε 1.

	Sunshine	Gattaca	Gravity	Event Horizon	Arrival
Χρήστης 1	X	2		5	8
Χρήστης 2	3	3		5	7
Χρήστης 3	9	2	6	3	X
Χρήστης 4	6	4	4	7	3
Χρήστης 5		4	X	8	3
Χρήστης 6	8	8	7	5	6

a.   Υπολογίστε την ομοιότητα (similarity) μεταξύ των 6 χρηστών χρησιμοποιώντας δυο μεθόδους: Ευκλείδεια απόσταση και Pearson Correlation.

b.   Χρησιμοποιώντας K-Nearest Neighbors με k=2 και weighted average και με τις δύο μετρικές του ερωτήματος a. υπολογίστε πως περιμένουμε να αξιολογήσει την ταινία Gravity ο χρήστης 2. Σχολιάστε τα αποτελέσματα.

c.   Αν υποθέσουμε ότι χρησιμοποιούμε τις προτιμήσεις των χρηστών στις ταινίες για να προτείνουμε φίλους, τότε ποιες σχέσεις σας φαίνονται πιο πιθανές; Εξηγήστε.