



ΔΙΕΘΝΕΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΗΣ ΕΛΛΑΔΟΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Ανάπτυξη εφαρμογής για προβολή και
συστάσεις ταινιών»

MOVIE Theatre

Του φοιτητή
Λαφιώτη Νικόλαου, Α.Μ. 185212

Επιβλέπουσα
Ασδρέ Κατερίνα
Ε.ΔΙ.Π.

31/05/2025

Τίτλος Δ.Ε. Ανάπτυξη εφαρμογής για προβολή και συστάσεις ταινιών

Κωδικός Δ.Ε. 23209

Όνοματεπώνυμο φοιτητή Λαφιώτης Νικόλαος

Όνοματεπώνυμο εισηγητή Ασδρέ Κατερίνα

Ημερομηνία ανάληψης Δ.Ε. 31/03/2023

Ημερομηνία περάτωσης Δ.Ε.

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Λαφιώτη Νικόλαου που την εκπόνησε/αν. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Αφιερωμένη στην οικογένεια μου»

Πρόλογος

Η ενασχόληση με τη δημιουργία εφαρμογών και η κατανόηση του τρόπου με τον οποίο λειτουργούν οι σύγχρονες ψηφιακές πλατφόρμες αποτέλεσαν βασικά ερεθίσματα για την επιλογή του θέματος της παρούσας διπλωματικής εργασίας. Στην εποχή των streaming υπηρεσιών, όπου η εμπειρία του χρήστη καθορίζεται σε μεγάλο βαθμό από τις προτάσεις περιεχομένου και τη σχεδίαση του περιβάλλοντος, το ενδιαφέρον μου επικεντρώθηκε στην κατανόηση του πώς τεχνολογίες όπως τα recommendation systems επηρεάζουν τη σχέση του χρήστη με μια πλατφόρμα.

Η εργασία αυτή αποτέλεσε μια ευκαιρία για έναν συνδυασμό θεωρητικής γνώσης και πρακτικής εφαρμογής, εστιάζοντας όχι μόνο στην τεχνική υλοποίηση αλλά και στην εμπειρία χρήσης. Κατά τη διάρκεια της ανάπτυξης αντιμετωπίστηκαν προκλήσεις, έγιναν πειραματισμοί και δοκιμάστηκαν διαφορετικές προσεγγίσεις, με σκοπό την καλύτερη δυνατή απόδοση μιας λειτουργικής και φιλικής εφαρμογής.

Η διαδικασία αυτή συνέβαλε καθοριστικά στην προσωπική και επαγγελματική εξέλιξη, ενισχύοντας δεξιότητες τόσο στον προγραμματισμό όσο και στη μεθοδολογική προσέγγιση της ανάπτυξης λογισμικού.

Περίληψη

Η εργασία αυτή αφορά την υλοποίηση μιας διαδικτυακής εφαρμογής με το όνομα **MovieTheater** που έχει ως στόχο την προβολή trailers, πληροφοριών και περιεχομένου από ταινίες και σειρές. Ο χρήστης έχει τη δυνατότητα να δημιουργήσει λογαριασμό, να συνδεθεί, να περιηγηθεί σε περιεχόμενο ανάλογα με τις προτιμήσεις του και να οδηγηθεί σε εξωτερικές πλατφόρμες όπως το Netflix για την παρακολούθηση του υλικού.

Στην εφαρμογή έχει ενσωματωθεί ένα βασικό recommendation system, το οποίο προτείνει περιεχόμενο με βάση τη συμπεριφορά και τις προτιμήσεις του χρήστη αλλά και των άλλων χρηστών. Με αυτόν τον τρόπο, η εμπειρία γίνεται πιο προσωποποιημένη και η πλοήγηση πιο εύκολη και λειτουργική.

Για την ανάπτυξη της εφαρμογής χρησιμοποιήθηκαν σύγχρονες τεχνολογίες όπως React και Next.js για το frontend, ενώ η βάση δεδομένων υποστηρίζεται από τη MongoDB σε συνδυασμό με Prisma για τη διαχείριση των δεδομένων. Η υλοποίηση έγινε εξ ολοκλήρου σε TypeScript για μεγαλύτερη αξιοπιστία στον κώδικα.

Μέσα από τη διαδικασία σχεδιασμού και ανάπτυξης της εφαρμογής, δόθηκε έμφαση τόσο στην πρακτική πλευρά των τεχνολογιών όσο και στην κατανόηση του τρόπου λειτουργίας των συστημάτων προτάσεων, όπως αυτά χρησιμοποιούνται σήμερα σε μεγάλες πλατφόρμες ψυχαγωγίας.

«Application development for movie screening and recommendations»

Lafiotis Nikolaos

Abstract

This thesis presents the implementation of a web application designed to display trailers, information, and content related to movies and TV series. Users can create an account, log in, browse content based on their preferences, and be redirected to external platforms such as Netflix to watch the selected content.

A basic recommendation system has been integrated into the application, offering content suggestions based on the user's behavior and preferences, as well as those of other users. This approach helps create a more personalized experience and makes navigation smoother and more effective.

Modern technologies were used for the development of the application, including React and Next.js for the frontend, and MongoDB combined with Prisma for data management. The entire implementation was done using TypeScript to ensure better code reliability.

Throughout the design and development process, emphasis was placed both on the practical application of modern tools and on understanding how recommendation systems work, similar to those used by major entertainment platforms today.

Ευχαριστίες

Πρώτα από όλους θα ήθελα να ευχαριστήσω θερμά την επιβλέπουσα μου κ. Κατερίνα Ασδρέ, για την υπομονή και την εμπιστοσύνη που έδειξε στο πρόσωπό μου, την βοήθεια της και τις πολύτιμες συμβουλές της ώστε να έχω το καλύτερο δυνατό αποτέλεσμα.

Θα ήθελα, επίσης, να ευχαριστήσω θερμά όλους τους διδάσκοντες του Τμήματος. Οι γνώσεις, οι συμβουλές και η καθοδήγηση τους αποτέλεσαν θεμέλιο για την ακαδημαϊκή μου εξέλιξη ενώ παράλληλα και με βοήθησαν να αναπτύξω δεξιότητες και εργαλεία που αποδείχθηκαν πολύτιμα και στην επαγγελματική μου δραστηριότητα.

Πάνω από όλα όμως, θα ήθελα να ευχαριστήσω εγκάρδια την οικογένεια μου για κάθε είδους στήριξης που μου παρείχαν καθ' όλη τη διάρκεια των προπτυχιακών μου σπουδών.

Περιεχόμενα

| | |
|--|------|
| Πρόλογος | i |
| Περίληψη | ii |
| Abstract | iii |
| Ευχαριστίες | iv |
| Περιεχόμενα | v |
| Κατάλογος εικόνων | vii |
| Κατάλογος πινάκων | viii |
| Συντομογραφίες | ix |
| Κεφάλαιο 1 ^ο : Εισαγωγή | 1 |
| Κεφάλαιο 2 ^ο : Εφαρμογές για συστάσεις και προβολή ταινιών | 3 |
| Κεφάλαιο 3 ^ο : Recommendation Systems | 11 |
| 3.1 Εισαγωγή στα Recommendation Systems | 11 |
| 3.2 Evaluation Metrics (Μετρικές αξιολόγησης) | 13 |
| 3.3 Παραδείγματα αποτελεσματικότητας συστημάτων σύστασης | 15 |
| 3.4 Άλλοι τομείς εφαρμογής συστημάτων σύστασης | 16 |
| Κεφάλαιο 4 ^ο : Προβλήματα και προκλήσεις που αντιμετωπίζει κάθε Recommendation System | 18 |
| Κεφάλαιο 5 ^ο : Το Recommendation System του YouTube | 22 |
| 5.1 Το YouTube | 22 |
| 5.2 Η δομή του YouTube Recommendation System | 22 |
| 5.3 Αλγόριθμοι των Recommendation System του YouTube | 23 |
| Κεφάλαιο 6 ^ο : Netflix | 27 |
| 6.1 Το αφεντικό του streaming | 27 |
| 6.2 Ιστορική αναδρομή-μετάβαση από το CD στο streaming | 27 |
| 6.3 Netflix Recommendation System | 29 |
| 6.3.1 Εξέλιξη της αρχιτεκτονικής του Netflix Recommendation System | 29 |
| 6.3.2 Data και tokenizing user interactions | 30 |
| 6.3.3 Στόχος και αρχιτεκτονική του μοντέλου | 31 |
| Κεφάλαιο 7 ^ο : Σύγκριση Netflix-YouTube Recommendation Systems | 33 |
| Κεφάλαιο 8 ^ο : Ηθικά ζητήματα και Ιδιωτικότητα | 34 |
| Κεφάλαιο 9 ^ο : Η εφαρμογή | 36 |
| 9.1 Ανάλυση της εφαρμογής | 36 |
| 9.1.1 Σελίδα σύνδεσης | 36 |
| 9.1.2 Σελίδα επιλογής profile | 37 |
| 9.1.3 Homepage | 38 |
| 9.1.4 Movies | 39 |
| 9.1.5 Series | 40 |
| 9.1.6 My List | 41 |
| 9.1.7 New & Trending | 42 |
| 9.1.8 Others like | 42 |
| 9.1.9 Recommended | 42 |
| 9.1.10 Search page | 43 |
| 9.1.11 Movie details | 45 |

| | |
|---|----|
| Κεφάλαιο 10° : Τεχνολογίες και εργαλεία ανάπτυξης | 46 |
| Κεφάλαιο 11° : Συμπεράσματα και προβληματισμοί | 49 |
| Βιβλιογραφία | 56 |

Κατάλογος Εικόνων

| | |
|---|----|
| Εικόνα 2.1: Αρχική σελίδα IMDB | 4 |
| Εικόνα 2.2: Προβολή trailer μέσω της σελίδας IMDB | 5 |
| Εικόνα 2.3: Αρχική σελίδα JustWatch | 6 |
| Εικόνα 2.3: Αρχική σελίδα Letterbox | 8 |
| Εικόνα 2.4: Αρχική σελίδα Netflix | 9 |
| Εικόνα 2.5: Αρχική σελίδα Track.tv | 11 |
| Εικόνα 5.1: Απεικόνιση ενός: (α) Wide & Deep Learning, (β) Multi-View Deep Neural Network. | 24 |
| Εικόνα 5.2: Παραγοντοποίηση πίνακα αξιολογήσεων σε χρήστες και αντικείμενα | 24 |
| Εικόνα 6.1: Συνδρομητές του Netflix 2011-2016 | 28 |
| Εικόνα 6.2: Η πλατφόρμα Netflix το 2008 | 29 |
| Εικόνα 8.1: Δομή και επιρροές ενός συστήματος συστάσεων | 34 |
| Εικόνα 9.1: Κεντρική σελίδα εισόδου με email, google github | 36 |
| Εικόνα 9.2: Κεντρική σελίδα δημιουργίας λογαριασμού | 37 |
| Εικόνα 9.3: Σελίδα επιλογής profile | 38 |
| Εικόνα 9.4: Κεντρική σελίδα | 39 |
| Εικόνα 9.5: Σελίδα ταινιών | 40 |
| Εικόνα 9.6: Σελίδα σειρών | 41 |
| Εικόνα 9.7: Σελίδα με τις αγαπημένες ταινίες και σειρές του χρήστη | 41 |
| Εικόνα 9.8: Σελίδα με τις δημοφιλείς ταινίες και σειρές | 42 |
| Εικόνα 9.9: Σελίδα με τις προτεινόμενες ταινίες και σειρές για τον χρήστη | 43 |
| Εικόνα 9.10: Σελίδα με Τα αποτελέσματα αναζήτησης του χρήστη | 44 |
| Εικόνα 9.11: Σελίδα ταινίας η σειράς με τις πληροφορίες της και τις δυνατότητες που προσφέρει στον χρήστη | 45 |
| Εικόνα 11.1 Τα 4 στάδια του συστήματος συστάσεων | 50 |
| Εικόνα 11.2: Μελλοντική ροή των API και των υποστηρικτικών πλατφορμών προβολής ταινίας | 50 |
| Εικόνα 11.3: Διάγραμμα συστήματος κοινωνικής σύνδεσης | 51 |
| Εικόνα 11.4: Datasets και συγκριτικές μέθοδοι για συστήματα συστάσεων | 53 |

Κατάλογος Πινάκων

| | |
|---|----|
| Πίνακας 2.1: Σύντομη παρουσίαση παρόμοιων πλατφορμών | 3 |
| Πίνακας 3.1: Συγκριτικός Πίνακας Εφαρμογών Recommender Systems | 17 |
| Πίνακας 11.1 Στατιστικά χαρακτηριστικά των συνόλων δεδομένων (μέγεθος, πυκνότητα, βαθμός δραστηριότητας). | 54 |

Συντομογραφίες

IMDB: Internet Movie Database

API: Application Programming Interface

RNNs: Recurrent Neural Networks

NCF: Neural Collaborative Filtering

MLP: Multilayer Perceptron

FM: Factorization Machine

SVD: Singular value decomposition

PCA: Principal component analysis

Κεφάλαιο 1^ο: Εισαγωγή

Η συνεχής αύξηση του ψηφιακού περιεχομένου, σε συνδυασμό με την ευρεία χρήση διαδικτυακών πλατφορμών ψυχαγωγίας, έχει δημιουργήσει την ανάγκη για εργαλεία που βοηθούν τους χρήστες να εντοπίζουν πιο εύκολα περιεχόμενο που ανταποκρίνεται στα προσωπικά τους ενδιαφέροντα. Στο πεδίο του κινηματογράφου και των διαδικτυακών εφαρμογών που σχετίζονται με αυτόν, παρατηρείται έντονα το φαινόμενο της υπερπροσφοράς επιλογών, γεγονός που δυσκολεύει πολλές φορές την εύρεση κατάλληλων προτάσεων.

Η ανάγκη για φιλτράρισμα περιεχομένου και παροχή εξατομικευμένων επιλογών άρχισε να διαφαίνεται ήδη από τις πρώτες δεκαετίες της ψηφιακής εποχής. Ήδη από τη δεκαετία του 1990, οι ερευνητές ανέπτυσαν πρώιμα μοντέλα collaborative filtering (συνεραργατικό φιλτράρισμα), τα οποία αποτέλεσαν τη βάση για τα σύγχρονα recommendation systems (συστήματα προτάσεων). Στην πορεία, με την εξέλιξη των αλγορίθμων μηχανικής μάθησης και τη ραγδαία αύξηση των δεδομένων χρηστών (big data), τα συστήματα αυτά απέκτησαν νέες δυνατότητες και χρησιμοποιούνται πλέον ευρέως σε πληθώρα εφαρμογών, τόσο στον τομέα της ψυχαγωγίας όσο και σε άλλους τομείς, όπως το ηλεκτρονικό εμπόριο και η εκπαίδευση.

Σε αυτό το πλαίσιο, τα συστήματα προτάσεων (recommendation systems) έχουν αποκτήσει ιδιαίτερη σημασία. Βασική τους λειτουργία είναι η ανάλυση των προτιμήσεων των χρηστών και η παροχή προσαρμοσμένων προτάσεων, προκειμένου να διευκολύνεται η πρόσβαση σε σχετικό και ενδιαφέρον περιεχόμενο. Τέτοια συστήματα χρησιμοποιούνται ευρέως από γνωστές πλατφόρμες όπως το Netflix, το YouTube και το Spotify, συμβάλλοντας σημαντικά στη βελτίωση της εμπειρίας χρήσης.

Τα σύγχρονα συστήματα προτάσεων βασίζονται σε διάφορες τεχνικές, όπως η φιλτραρισμένη συνεργασία (collaborative filtering), η προσέγγιση βάσει περιεχομένου (content-based filtering), αλλά και συνδυαστικά υβριδικά μοντέλα, τα οποία αξιοποιούν δεδομένα συμπεριφοράς, προτιμήσεων και μεταδεδομένων.

Πέραν της χρηστικότητας, τα recommendation systems έχουν βαθύ αντίκτυπο στην ψηφιακή οικονομία, καθώς ενισχύουν τη διατήρηση χρηστών, την αλληλεπίδραση με το περιεχόμενο και τελικά τα έσοδα των πλατφορμών που είναι ο αντικειμενικός σκοπός. Ωστόσο, η μαζική χρήση τέτοιων συστημάτων έχει εγείρει και σημαντικά ζητήματα, όπως η δημιουργία φίλτρων πληροφόρησης (filter bubbles) και η μεροληψία στην προβολή περιεχομένου, που ενδέχεται να περιορίζουν την ποικιλομορφία των πληροφοριών στις οποίες εκτίθεται ένας χρήστης.

Παρά τη σημαντική πρόοδο στον τομέα των recommendation systems, η υλοποίησή τους εξακολουθεί να παρουσιάζει προκλήσεις. Ζητήματα όπως η διαχείριση της ετερογένειας των χρηστών, το λεγόμενο cold start (όταν δεν υπάρχουν αρκετά δεδομένα για νέους χρήστες ή αντικείμενα), καθώς και η ανάγκη για κλιμάκωση σε περιβάλλοντα με μεγάλο όγκο χρηστών, καθιστούν απαραίτητο τον σχεδιασμό ευέλικτων και επεκτάσιμων λύσεων. Η παρούσα εργασία προσπαθεί να προσεγγίσει αυτές τις προκλήσεις μέσα από μια απλή αλλά λειτουργική αρχιτεκτονική.

Στο πλαίσιο της παρούσας διπλωματικής εργασίας υλοποιήθηκε μια διαδικτυακή εφαρμογή με θέμα τις ταινίες, η οποία ενσωματώνει βασικές λειτουργίες αυθεντικοποίησης χρήστη και καταγραφής προτιμήσεων, καθώς και ένα απλό σύστημα προτάσεων. Η εφαρμογή δίνει τη δυνατότητα στους χρήστες να δημιουργούν λογαριασμό, να προσθέτουν ταινίες στις αγαπημένες τους και, βάσει αυτών των δεδομένων, να λαμβάνουν προτάσεις για παρόμοιες ταινίες, τόσο ως προς το είδος όσο και ως προς τους ηθοποιούς που συμμετέχουν. Έχουν επίσης την δυνατότητα να περιηγηθούν μέσα στον μεγάλο όγκο ταινιών που διαθέτει η εφαρμογή, να δουν πληροφορίες, trailer ή ακόμη και σχόλια για τις ταινίες αυτές και φυσικά να τις παρακολουθήσουν.

Η εργασία αυτή εξετάζει αρχικά το θεωρητικό υπόβαθρο των recommendation systems, στις μεθόδους που χρησιμοποιούνται για την δημιουργία αυτών των συστημάτων καθώς και σε πιθανά προβλήματα-προκλήσεις που αντιμετωπίζουν τα recommendation systems καθώς επίσης και τον ρόλο τους στις εφαρμογές που σχετίζονται με τον κινηματογράφο-streaming. Γίνεται αναφορά σε παρόμοιες εφαρμογές καθώς και στα recommendation systems αυτών και αναφέρει πιο λεπτομερώς πληροφορίες για 2 από αυτές τις εφαρμογές

Κεφάλαιο 1

ενώ στο τέλος περιγράφεται η υλοποίηση της εφαρμογής, οι τεχνολογίες που χρησιμοποιήθηκαν, καθώς και τα χαρακτηριστικά της. Ιδιαίτερη έμφαση δίνεται στον τρόπο με τον οποίο τα προσωπικά ενδιαφέροντα του κάθε χρήστη επηρεάζουν το περιεχόμενο που του εμφανίζεται, με στόχο τη δημιουργία μιας πιο προσωποποιημένης και ουσιαστικής εμπειρίας χρήσης.

Κεφάλαιο 2^ο: Εφαρμογές για συστάσεις και προβολή ταινιών

Πριν ξεκινήσει η ανάλυση και η υλοποίηση της εφαρμογής που αναπτύσσεται στη συγκεκριμένη εργασία, είναι σημαντικό να εξεταστούν κάποιες ήδη υπάρχουσες και ευρέως γνωστές εφαρμογές που σχετίζονται με την αναζήτηση, την παρουσίαση και τη σύσταση κινηματογραφικού περιεχομένου. Μέσα από αυτή την έρευνα, εντοπίζονται τα κοινά χαρακτηριστικά, τα δυνατά σημεία αλλά και οι ελλείψεις που υπάρχουν σε αυτές τις εφαρμογές, ώστε να διαμορφωθεί πιο ξεκάθαρα ο στόχος της νέας εφαρμογής και να καλυφθούν συγκεκριμένες ανάγκες των χρηστών. (βλέπε πίνακα 2.1)

Η ανάλυση επικεντρώνεται σε πέντε βασικά σημεία για κάθε εφαρμογή:

- τον τρόπο εγγραφής και αυθεντικοποίησης των χρηστών,
- τον τρόπο με τον οποίο παρουσιάζεται και αναζητείται το περιεχόμενο,
- τη λειτουργία του συστήματος προτάσεων (recommendation system),
- τη διασύνδεση με άλλες πλατφόρμες και υπηρεσίες μέσω APIs,
- και τέλος, τη σύγκριση με τη λειτουργικότητα της εφαρμογής που προτείνεται στην παρούσα εργασία.

Πίνακας 2.1: Σύντομη παρουσίαση παρόμοιων πλατφορμών

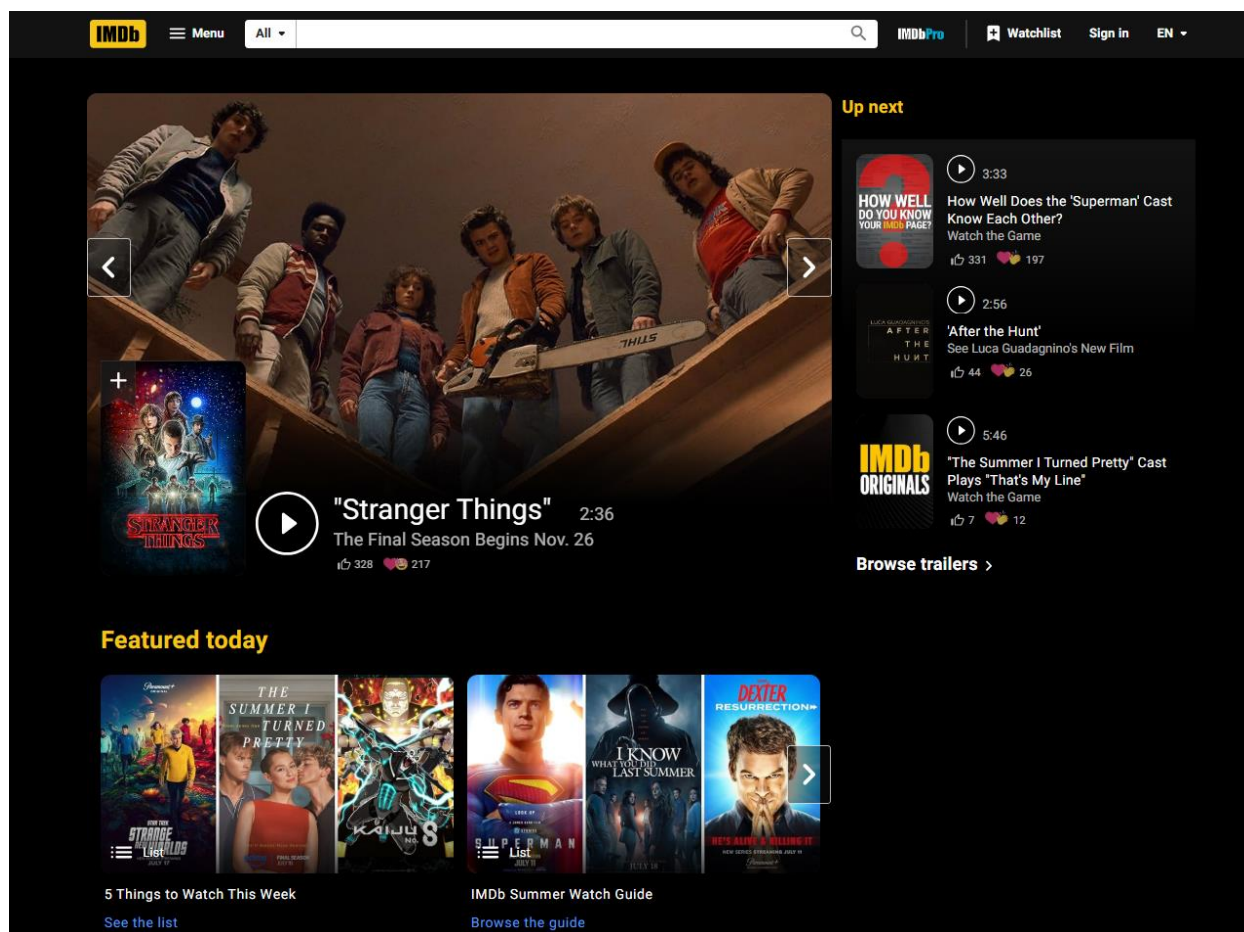
| Πλατφόρμα | Τύπος Υπηρεσίας | Κύρια Χαρακτηριστικά | Δυνατότητες Χρήστη |
|------------|-----------------------------------|---|---|
| IMDb | Βάση δεδομένων για ταινίες/σειρές | Πληροφορίες για cast, ratings, reviews, trivia | Προβολή, βαθμολόγηση, δημιουργία watchlists |
| JustWatch | Μηχανή αναζήτησης Streaming | Αναζήτηση διαθεσιμότητας σε πλατφόρμες streaming | Φιλτράρισμα ανά χώρα, υπηρεσία, ποιότητα |
| Letterboxd | Κοινωνικό δίκτυο σινεφίλ | Ημερολόγιο προβολών σε πλατφόρμες streaming | Κριτικές, αξιολογήσεις κοινωνική διάδραση |
| Netflix | Υπηρεσία streaming περιεχομένου | Παρακολούθηση σειρών/ταινιών, προτάσεις με βάση τις προτιμήσεις | Προσωπικό προφίλ, προτάσεις, λήψεις |
| Track.tv | Σύστημα παρακολούθησης προβολών | Συγχρονισμός με apps/ συσκευές, στατιστικά παρακολούθησης | Scrobbling, ημερολόγιο, λίστες |

IMDb:

Το IMDb (βλέπε εικόνα 2.1) είναι μια πολύ γνωστή ιστοσελίδα και εφαρμογή όπου μπορεί κανείς να βρει πληροφορίες για ταινίες, σειρές, ηθοποιούς, σκηνοθέτες και γενικά οτιδήποτε έχει σχέση με τον κόσμο του σινεμά και της τηλεόρασης. Είναι σαν μια μεγάλη βάση δεδομένων που συγκεντρώνει στοιχεία για εκατομμύρια ταινίες και επεισόδια. Εκεί μπορεί κάποιος χρήστης να δει από περιλήψεις και βαθμολογίες,

Κεφάλαιο 2

μέχρι trailer (σύντομη προβολή σκηνών της ταινίας) (βλέπε εικόνα 2.2) και κριτικές από άλλους χρήστες. Ουσιαστικά, το IMDb βοηθάει τους ανθρώπους να μάθουν περισσότερα για τις ταινίες που τους ενδιαφέρουν και να πάρουν ιδέες για το τι να δουν στη συνέχεια. Είναι μια από τις πιο δημοφιλείς πηγές πληροφοριών για τον κινηματογράφο παγκοσμίως. Πολλές από τις πληροφορίες είναι ανοιχτές για όλους τους επισκέπτες, οπότε και δεν χρειάζεται πάντα να έχει ο χρήστης λογαριασμό.

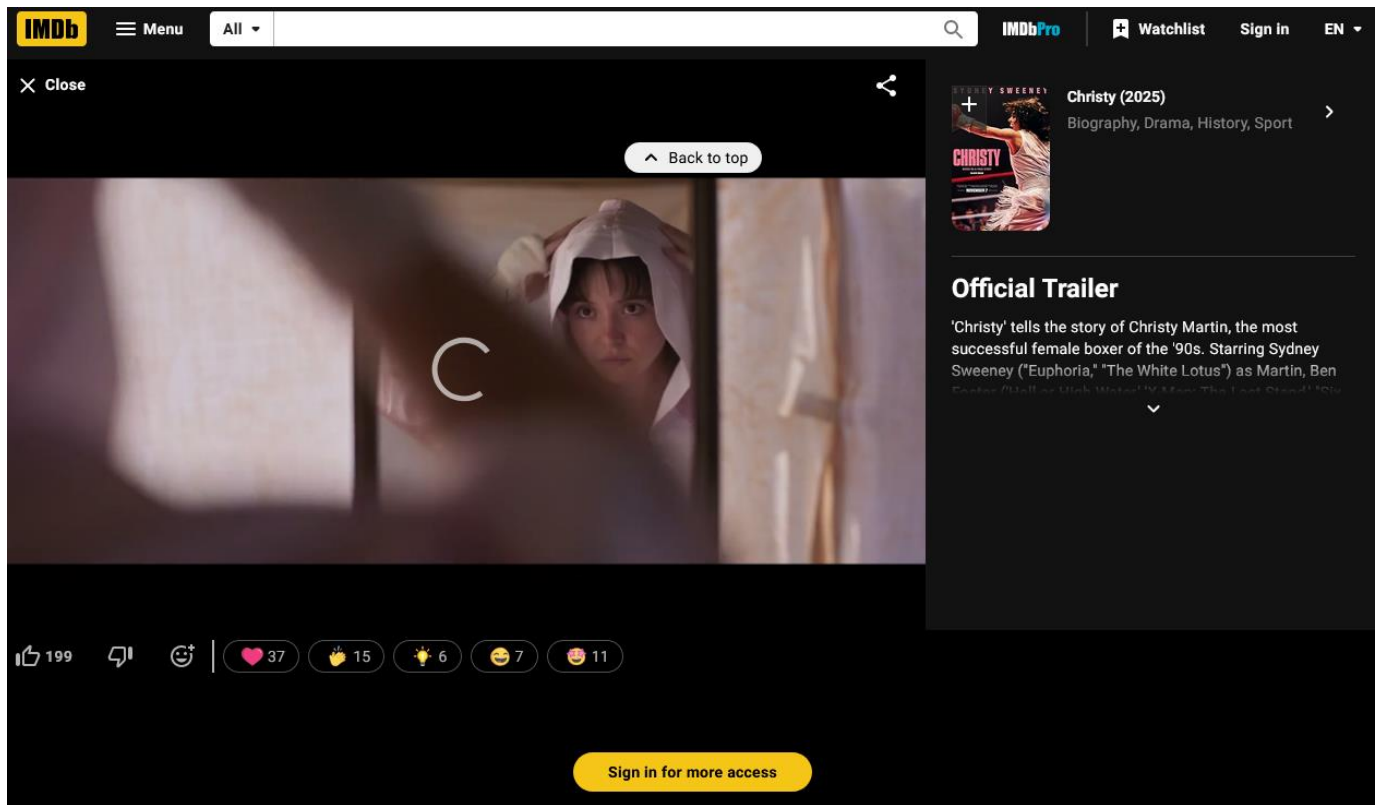


Εικόνα 2.1: Αρχική σελίδα IMDb [20]

Για να χρησιμοποιήσει όμως κάποιος ορισμένες πιο «προχωρημένες» λειτουργίες στο IMDb, όπως να βάλει ταινίες στα αγαπημένα του, να γράψει κριτικές ή να κάνει αξιολογήσεις, πρέπει να έχει λογαριασμό. Η εγγραφή γίνεται με απλό τρόπο, δίνοντας το email του και έναν κωδικό. Επίσης, επειδή το IMDb ανήκει στην Amazon, μπορεί να συνδεθεί και μέσω του λογαριασμού του στην Amazon, αν έχει. Αυτό κάνει πιο εύκολη τη διαδικασία για όσους ήδη χρησιμοποιούν την Amazon. Δεν υποστηρίζει άλλους τύπους σύνδεσης, όπως μέσω Google ή Facebook.

Όσον αφορά την αναζήτηση μέσα στην πλατφόρμα, όταν ο χρήστης θέλει να βρει μια ταινία ή σειρά, απλά πληκτρολογεί το όνομά της στη μπάρα αναζήτησης. Η εφαρμογή του δείχνει αποτελέσματα με εικόνες, τίτλους και σύντομες περιγραφές για να διαλέξει αυτό που ψάχνει. Μόλις επιλέξει κάτι, εμφανίζονται λεπτομέρειες όπως η υπόθεση, οι ηθοποιοί, το είδος της ταινίας, οι βαθμολογίες και συχνά trailer για να πάρει μια γεύση. Επιπλέον, μπορεί να δει κριτικές από άλλους χρήστες και να διαβάσει τι σκέφτηκαν για την ταινία.

Το IMDb έχει ένα σύστημα προτάσεων, δηλαδή προτείνει ταινίες ή σειρές που μπορεί να αρέσουν στον συνδεδεμένο χρήστη. Αυτό γίνεται με βάση τις ταινίες που έχει αξιολογήσει ή έχει προσθέσει σε λίστες. Το σύστημα συγκρίνει τις προτιμήσεις του με άλλων χρηστών που έχουν παρόμοιο γούστο και του δείχνει τι εκείνοι έχουν δει και τους άρεσε. Έτσι, η πρόταση βασίζεται σε συλλογικές προτιμήσεις και όχι μόνο στα χαρακτηριστικά της κάθε ταινίας. Επίσης ένα βασικό στοιχείο είναι πως το IMDb δεν είναι πλατφόρμα streaming, δηλαδή δεν μπορεί να δει ο χρήστης ταινίες μέσα από αυτή.



Εικόνα 2.2: Προβολή trailer μέσω της σελίδας IMDb

Όμως, παρέχει συνδέσμους προς άλλες υπηρεσίες όπου μπορεί να νοικιάσει ή να αγοράσει την ταινία, όπως Amazon Prime Video ή iTunes. Επίσης, δίνει πρόσβαση σε trailer που φιλοξενούνται συχνά στο YouTube. Η επικοινωνία με αυτές τις υπηρεσίες γίνεται μέσω API που επιτρέπουν να αντλούν και να εμφανίζουν τα σχετικά δεδομένα.

Στο επίπεδο σύγκρισης λοιπόν, σε αντίθεση με το IMDb, το MovieTheater έχει ενσωματωμένη τη δυνατότητα να δει κανείς τις ταινίες απευθείας μέσω Netflix, αν υπάρχουν διαθέσιμες μέσω API key. Επιπλέον, το σύστημα προτάσεων είναι πιο δυναμικό και βασισμένο όχι μόνο στις προτιμήσεις του χρήστη, αλλά και στις επιλογές και δραστηριότητα όλης της κοινότητας, μέσα από τα αγαπημένα που βάζουν οι χρήστες. Τέλος, η δυνατότητα σύνδεσης μέσω Google και email θα κάνει την εμπειρία πιο φιλική και γρήγορη για τον χρήστη.

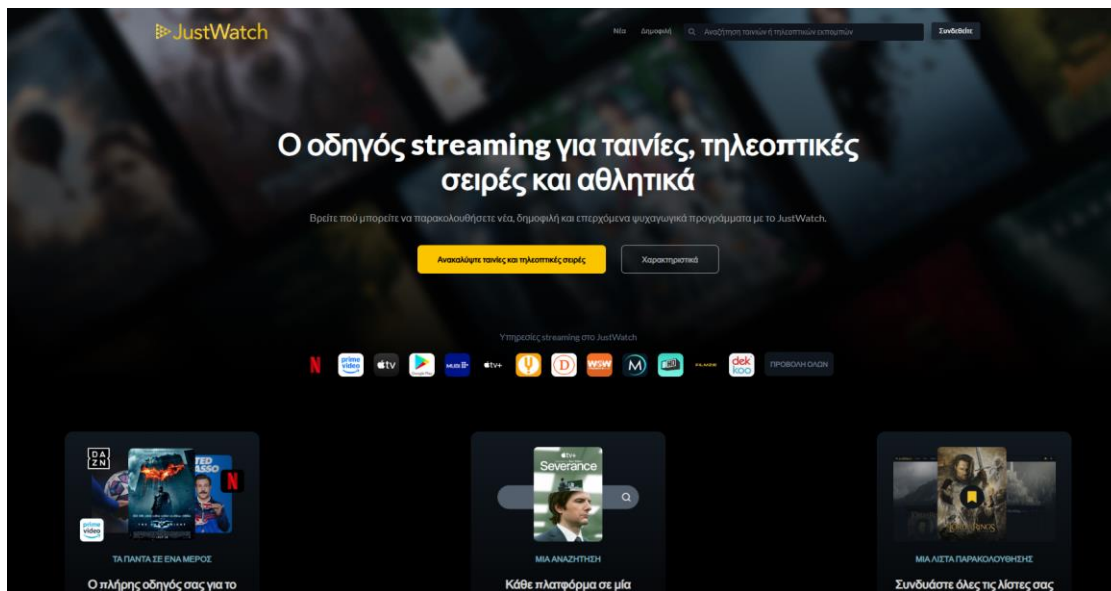
JustWatch:

Το JustWatch (βλέπε εικόνα 2.3) είναι μια εφαρμογή και ιστοσελίδα που βοηθάει τους χρήστες να βρουν σε ποια πλατφόρμα streaming (όπως Netflix, Amazon Prime, Disney+, HBO, κτλ) είναι διαθέσιμη μια συγκεκριμένη ταινία ή σειρά. Στην ουσία, δεν είναι πλατφόρμα για να βλέπει κάποιος ταινίες, αλλά ένας οδηγός που δείχνει από πού μπορεί να τις παρακολουθήσει.

Επίσης, δίνει τη δυνατότητα αναζήτησης, φιλτραρίσματος (για παράδειγμα ανά είδος, χώρα, τιμή) και παρουσιάζει νέες κυκλοφορίες ή δημοφιλή έργα. Το JustWatch επιτρέπει στους χρήστες να δημιουργήσουν λογαριασμό μέσω email ή να συνδεθούν με social accounts όπως Google και Facebook. Η σύνδεση αυτή βοηθάει το σύστημα να αποθηκεύσει τις προτιμήσεις και τις λίστες του χρήστη, όπως τις αγαπημένες ταινίες ή σειρές που θέλει να δει αργότερα.

Επιπλέον, η αυθεντικοποίηση βοηθάει να έχει ο χρήστης μια πιο προσωποποιημένη εμπειρία και να λαμβάνει ειδοποιήσεις για νέες κυκλοφορίες που τον ενδιαφέρουν. Όταν ο χρήστης ψάχνει μια ταινία ή σειρά στο JustWatch, πληκτρολογεί το όνομα στην αναζήτηση και η εφαρμογή εμφανίζει τα διαθέσιμα αποτελέσματα, μαζί με πληροφορίες για το σε ποιες πλατφόρμες streaming μπορεί να τις δεις. Μπορεί να φιλτράρει τα αποτελέσματα ανά είδος, τιμή (αν είναι δωρεάν, ενοικίαση ή αγορά), ποιότητα εικόνας, και άλλα. Για κάθε ταινία, βλέπει περιγραφή, βαθμολογία, trailer και τις πλατφόρμες που την προσφέρουν.

Το JustWatch χρησιμοποιεί δεδομένα από τη συμπεριφορά των χρηστών του, όπως τις αναζητήσεις τους και τις προτιμήσεις που έχουν δηλώσει, για να προτείνει ταινίες και σειρές που πιθανώς να τους αρέσουν. Παρότι δεν είναι τόσο «κοινωνικό» όσο το IMDb ή το Letterboxd, το σύστημά του βασίζεται σε απλές τεχνικές που αξιοποιούν το ιστορικό προβολών και δημοφιλή έργα στη χώρα του χρήστη. Το JustWatch δουλεύει ουσιαστικά σαν «μεσίτης» μεταξύ του χρήστη και πολλών πλατφορμών streaming. Για να ενημερώνεται συνεχώς για τη διαθεσιμότητα και τις τιμές των ταινιών, συνεργάζεται με τα APIs των μεγάλων υπηρεσιών (Netflix, Amazon Prime, Hulu, κτλ) και συγκεντρώνει όλες τις πληροφορίες σε ένα μέρος. Έτσι, ο χρήστης δεν χρειάζεται να ψάχνει ξεχωριστά σε κάθε πλατφόρμα.



Εικόνα 2.3: Αρχική σελίδα JustWatch [21]

Το MovieTheater, πέρα από το να δείχνει αν μια ταινία είναι διαθέσιμη στο Netflix (μέσω API, όπως κάνει το JustWatch), προσφέρει ολοκληρωμένη εμπειρία με προσωπικό σύστημα προτάσεων βασισμένο στα αγαπημένα των χρηστών και τη συμπεριφορά όλης της κοινότητας. Επιπλέον, ο χρήστης μπορεί να δει trailer και πληροφορίες για τις ταινίες μέσα στην ίδια την εφαρμογή, ενώ υπάρχει και η δυνατότητα σύνδεσης μέσω Google ή email για ευκολότερη πρόσβαση. Με άλλα λόγια, συνδυάζει το στοιχείο του οδηγού του JustWatch με το κοινωνικό και εξατομικευμένο σύστημα που υπάρχει σε πιο πολύπλοκες πλατφόρμες.

Letterboxd:

Το Letterboxd (βλέπε εικόνα 2.4) είναι μια εφαρμογή και ιστοσελίδα που λειτουργεί σαν κοινωνικό δίκτυο για τους λάτρεις του κινηματογράφου. Οι χρήστες μπορούν να καταγράφουν τις ταινίες που έχουν δει, να γράφουν κριτικές, να βαθμολογούν, να δημιουργούν λίστες και να παρακολουθούν τη δραστηριότητα άλλων χρηστών. Δεν είναι πλατφόρμα streaming, δεν μπορεί άρα ο χρήστης να δει ταινίες εκεί, αλλά λειτουργεί σαν προσωπικό ημερολόγιο ταινιών, ενώ ταυτόχρονα δίνει και ιδέες για νέες προβολές. Είναι ένα είδος «κινηματογραφικού Instagram», αλλά αντί για φωτογραφίες, το περιεχόμενο είναι ταινίες, βαθμολογίες και

κριτικές.

Στο Letterboxd μπορεί ο χρήστης να φτιάξει λογαριασμό πολύ εύκολα είτε χρησιμοποιώντας email και κωδικό, είτε μέσω σύνδεσης με Google, Apple ID ή Twitter. Αυτή η ευελιξία κάνει πιο εύκολη την εγγραφή για τον χρήστη, χωρίς πολύπλοκα βήματα. Μόλις μπει, μπορεί να φτιάξει το προσωπικό του προφίλ και να ξεκινήσει να προσθέτει ταινίες που έχει δει ή θέλει να δει.

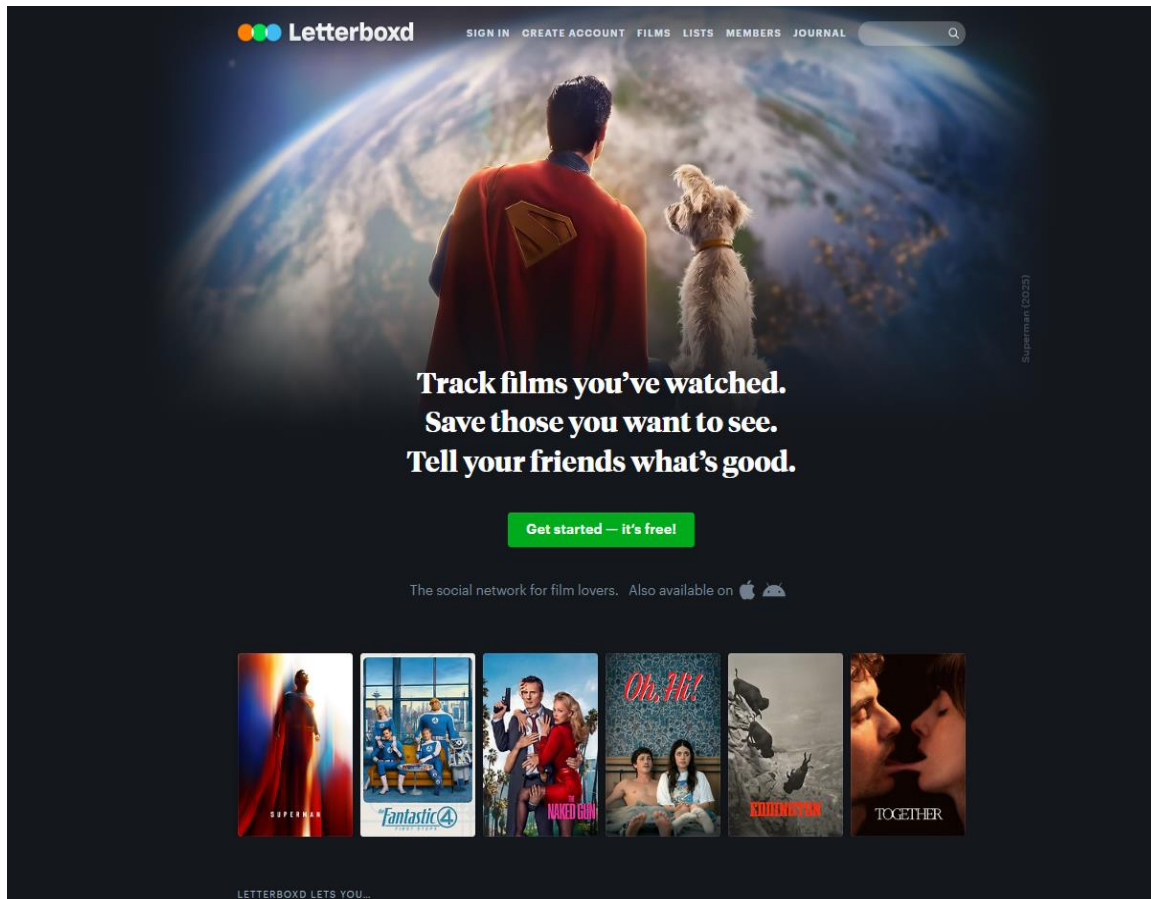
Η αναζήτηση στο Letterboxd είναι πολύ απλή και άμεση. Πληκτρολογεί τον τίτλο της ταινίας και του εμφανίζει το αντίστοιχο αποτέλεσμα με εικόνα, τίτλο και ημερομηνία κυκλοφορίας. Όταν μπαίνει στη σελίδα μιας ταινίας, βλέπει πληροφορίες όπως υπόθεση, σκηνοθέτης, ηθοποιοί, trailer, αφίσες, καθώς και τις κριτικές και τις βαθμολογίες άλλων χρηστών. Είναι σχεδιασμένο έτσι ώστε να προβάλλεται και η “άποψη της κοινότητας”, δηλαδή πώς τη δέχτηκε ο κόσμος και τι συζητιέται γύρω από αυτή.

Το Letterboxd δεν έχει τόσο εξελιγμένο recommendation system όσο άλλες πλατφόρμες, αλλά βασίζεται κυρίως σε κοινωνικές συμπεριφορές. Για παράδειγμα:

- Μπορεί ο χρήστης να δει τι είδαν οι φίλοι του.
- Του προτείνει ταινίες που είναι δημοφιλείς αυτή την περίοδο ή που σχετίζονται με τα γούστα του.
- Οι λίστες που δημιουργούν άλλοι χρήστες (όπως “Top Horror Movies” ή “Underrated Films”) λειτουργούν έμμεσα σαν σύστημα προτάσεων.

Αν και δεν χρησιμοποιεί κάτι πολύπλοκο όπως collaborative filtering, οι προτάσεις που βλέπει κάποιος συχνά επηρεάζονται από το πώς βαθμολογεί, τι προσθέτει στη λίστα του ή τι παρακολουθούν όσοι ακολουθεί. Τέλος το Letterboxd δε δείχνει πού μπορεί να δει μια ταινία, αλλά συνεργάζεται με εξωτερικές πλατφόρμες για να του εμφανίζει trailer από το YouTube και βασικά δεδομένα για τις ταινίες (τα οποία αντλεί από το TMDb – The Movie Database). Δεν είναι σχεδιασμένο για να δείχνει links προς Netflix ή Amazon, άρα δεν παίζει τον ρόλο οδηγού όπως το JustWatch. Παρ' όλα αυτά, με integration APIs μπορεί να δείχνει trailers, posters και άλλα metadata.

Το MovieTheater συνδυάζει πολλά από τα στοιχεία του Letterboxd, αλλά με ένα μεγάλο πλεονέκτημα: δίνει στον χρήστη τη δυνατότητα να δει πού ακριβώς μπορεί να παρακολουθήσει την ταινία (π.χ. μέσω Netflix) και όχι απλά να τη βαθμολογήσει ή να τη συζητήσει. Ακόμα, ο χρήστης μπορεί να δει trailer, πληροφορίες και κατευθείαν να επιλέξει αν θέλει να δει την ταινία στη διαθέσιμη πλατφόρμα. Άρα, το MovieTheater είναι κάτι σαν ένα υβρίδιο Letterboxd + JustWatch + IMDb, με δικό της recommendation engine.



Εικόνα 2.4: Αρχική σελίδα Letterbox [22]

Netflix:

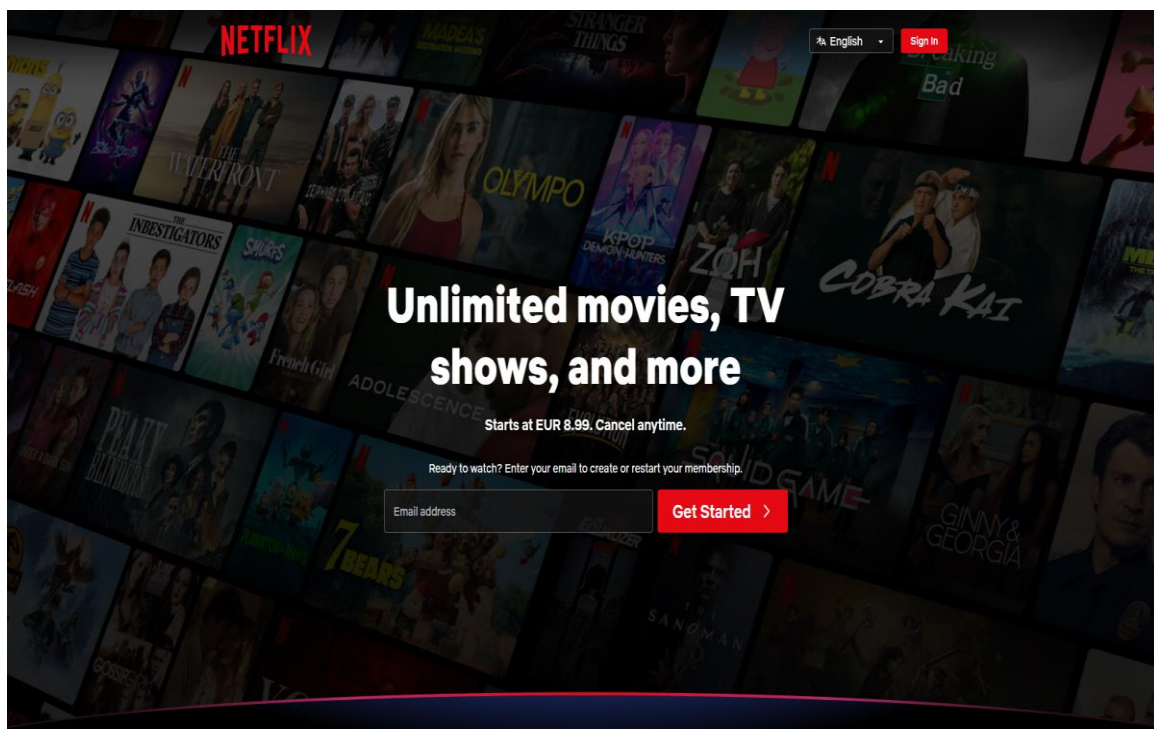
Το **Netflix** (βλέπε εικόνα 2.5) είναι μια από τις πιο γνωστές και δημοφιλείς πλατφόρμες παγκοσμίως για την παρακολούθηση ταινιών, σειρών, ντοκιμαντέρ και άλλου οπτικοακουστικού περιεχομένου μέσω internet. Πρόκειται για μια υπηρεσία streaming, δηλαδή ο χρήστης πληρώνει μηνιαία συνδρομή και μπορεί να παρακολουθεί το περιεχόμενο που προσφέρει ανά πάσα στιγμή, από διάφορες συσκευές (τηλεοράσεις, κινητά, υπολογιστές κ.ά.). Η κύρια δύναμη του Netflix είναι η τεράστια βιβλιοθήκη του, που περιλαμβάνει τόσο ξένες παραγωγές όσο και δικό του αποκλειστικό περιεχόμενο (Netflix Originals), καθώς και το γεγονός ότι προσαρμόζει τις προτάσεις του με βάση τα ενδιαφέροντα του κάθε χρήστη.

Η εγγραφή και η σύνδεση στο Netflix γίνεται με email και κωδικό, ενώ κατά την αρχική εγγραφή απαιτείται και η καταχώρηση στοιχείων πληρωμής, καθώς πρόκειται για συνδρομητική υπηρεσία. Δεν προσφέρει σύνδεση με social media (όπως Google ή Facebook). Ωστόσο, αν χρησιμοποιεί κάποιος Netflix μέσα από εξωτερικές συσκευές ή πλατφόρμες (π.χ. Apple TV), μπορεί να γίνει έμμεση σύνδεση μέσω αυτών. Η αυθεντικοποίηση στοχεύει στο να διασφαλίσει την ταυτότητα του χρήστη και να του προσφέρει ένα προσωπικό προφίλ με βάση τις συνήθειες και τις προτιμήσεις του.

Κεφάλαιο 2

Το Netflix προσφέρει μια δυνατή και εύχρηστη μηχανή αναζήτησης, όπου μπορεί ο χρήστης να πληκτρολογήσει τίτλο, είδος, ηθοποιό, σκηνοθέτη κτλ., και να δει άμεσα αποτελέσματα. Για κάθε ταινία ή σειρά παρουσιάζεται:

- μια σύντομη περιγραφή (υπόθεση),
- το είδος και η ηλικιακή καταλληλότητα,
- το trailer ή teaser,
- η διάρκεια,
- και συστάσεις για παρόμοιο περιεχόμενο.



Εικόνα 2.5: Αρχική σελίδα Netflix [23]

Το recommendation system του Netflix θεωρείται ένα από τα πιο προηγμένα στον κόσμο. Χρησιμοποιεί αλγορίθμους που αναλύουν τη συμπεριφορά του χρήστη: ποιες ταινίες βλέπει, τι είδη προτιμά, πόσο συχνά παρακολουθεί και τι έχει απορρίψει ή σταματήσει να βλέπει. Επιπλέον, λαμβάνει υπόψη τα κοινά χαρακτηριστικά μεταξύ διαφορετικών χρηστών με παρόμοιο γούστο. Ο συνδυασμός αυτών των στοιχείων δημιουργεί ένα εξατομικευμένο περιβάλλον, με προτάσεις που ταιριάζουν όσο το δυνατόν περισσότερο στα ενδιαφέροντα του κάθε χρήστη. Αυτό είναι ένα χαρακτηριστικό που κρατάει τους χρήστες ενεργούς και «δεμένους» με την πλατφόρμα. Το περιεχόμενο είναι οργανωμένο σε κατηγορίες όπως «Δημοφιλή τώρα», «Νέα κυκλοφορία», «Προτάσεις για εσάς», κλπ., με βάση τις προτιμήσεις του χρήστη.

Το Netflix, παρόλο που είναι μια "κλειστή" πλατφόρμα, επιτρέπει ορισμένες εξωτερικές διασυνδέσεις μέσω εσωτερικών APIs ή συνεργασιών. Για παράδειγμα:

- Πλατφόρμες όπως το JustWatch ή το Google TV δείχνουν πότε μια ταινία είναι διαθέσιμη στο Netflix.
- Ορισμένες εφαρμογές ταινιών, όπως η εφαρμογή που περιγράφεται στην εργασία αυτή, μπορούν να αντλούν στοιχεία από APIs που περιλαμβάνουν το Netflix ως πηγή, χωρίς να συνδέονται απευθείας με αυτό.

Επιπλέον, το Netflix συνεργάζεται με συσκευές και λειτουργικά συστήματα (Smart TVs, Android, iOS) ώστε να επιτρέπει streaming από άλλες εφαρμογές ή συστήματα.

Η MovieTheater δεν είναι πλατφόρμα streaming όπως το Netflix, αλλά λειτουργεί συμπληρωματικά: επιτρέπει στον χρήστη να δει αν μια ταινία είναι διαθέσιμη στο Netflix και να μεταφερθεί εκεί για προβολή. Επιπλέον, εστιάζει πολύ περισσότερο στην εξατομίκευση των προτάσεων, βασιζόμενη όχι μόνο στη δική του παρακολούθηση, αλλά και στις προτιμήσεις άλλων χρηστών (collaborative filtering). Τέλος, περιλαμβάνει λειτουργίες όπως αγαπημένα, trailer, προβολή πληροφοριών και διασύνδεση με πολλαπλές υπηρεσίες, κάτι που το Netflix δεν προσφέρει με τον ίδιο τρόπο.

Trakt.tv:

Το Trakt.tv (βλέπε εικόνα 2.6) είναι μια διαδικτυακή υπηρεσία που βοηθά τους χρήστες να παρακολουθούν τι ταινίες και σειρές έχουν δει, να οργανώνουν watchlists και να ανακαλύπτουν νέο περιεχόμενο βασισμένο στα ενδιαφέροντά τους. Αν και δεν προσφέρει προβολή ταινιών, λειτουργεί σαν «ψηφιακό ημερολόγιο» προβολών και ταυτόχρονα σαν ένα κοινωνικό δίκτυο με δυνατότητα σχολιασμού, αξιολόγησης και σύστασης περιεχομένου. Το δυνατό του σημείο είναι η συγχρονισμένη παρακολούθηση με άλλες εφαρμογές (όπως Plex, Kodi, Netflix μέσω τρίτων εργαλείων), επιτρέποντας στον χρήστη να έχει συγκεντρωμένα όλα τα δεδομένα παρακολούθησης σε ένα μέρος.

Η εγγραφή και σύνδεση στο Trakt.tv γίνεται με email και κωδικό, αλλά υποστηρίζει και social login μέσω Google, Twitter, Apple ID και άλλων. Η σύνδεση είναι απαραίτητη για να μπορεί ο χρήστης να κρατά το ιστορικό των προβολών του, να φτιάχνει λίστες, να αξιολογεί ταινίες και να λαμβάνει εξατομικευμένες προτάσεις. Η αυθεντικοποίηση επιτρέπει επίσης συγχρονισμό με εξωτερικές υπηρεσίες: αν για παράδειγμα δει ένα επεισόδιο σε άλλη εφαρμογή, αυτό ενημερώνεται αυτόματα και στον λογαριασμό σου στο Trakt.

Η αναζήτηση στο Trakt.tv είναι αρκετά λειτουργική. Μπορεί να αναζητήσει: ταινίες ή σειρές με βάση τον τίτλο, λίστες που έχουν δημιουργήσει άλλοι χρήστες, και ακόμα και άτομα (ηθοποιούς, σκηνοθέτες). Για κάθε ταινία ή σειρά εμφανίζονται βασικές πληροφορίες όπως η υπόθεση, η διάρκεια, τα επεισόδια (αν είναι σειρά), βαθμολογίες και σύνδεσμοι για παρακολούθηση. Το περιεχόμενο παρουσιάζεται με σύγχρονο και ευανάγνωστο τρόπο και υπάρχουν φίλτρα ανά είδος, έτος κυκλοφορίας, γλώσσα, κτλ.

Το recommendation system του Trakt λειτουργεί κυρίως με βάση τα δεδομένα προβολής του ίδιου του χρήστη, αλλά και την δραστηριότητα άλλων. Υπάρχουν δύο κύριες προσεγγίσεις:

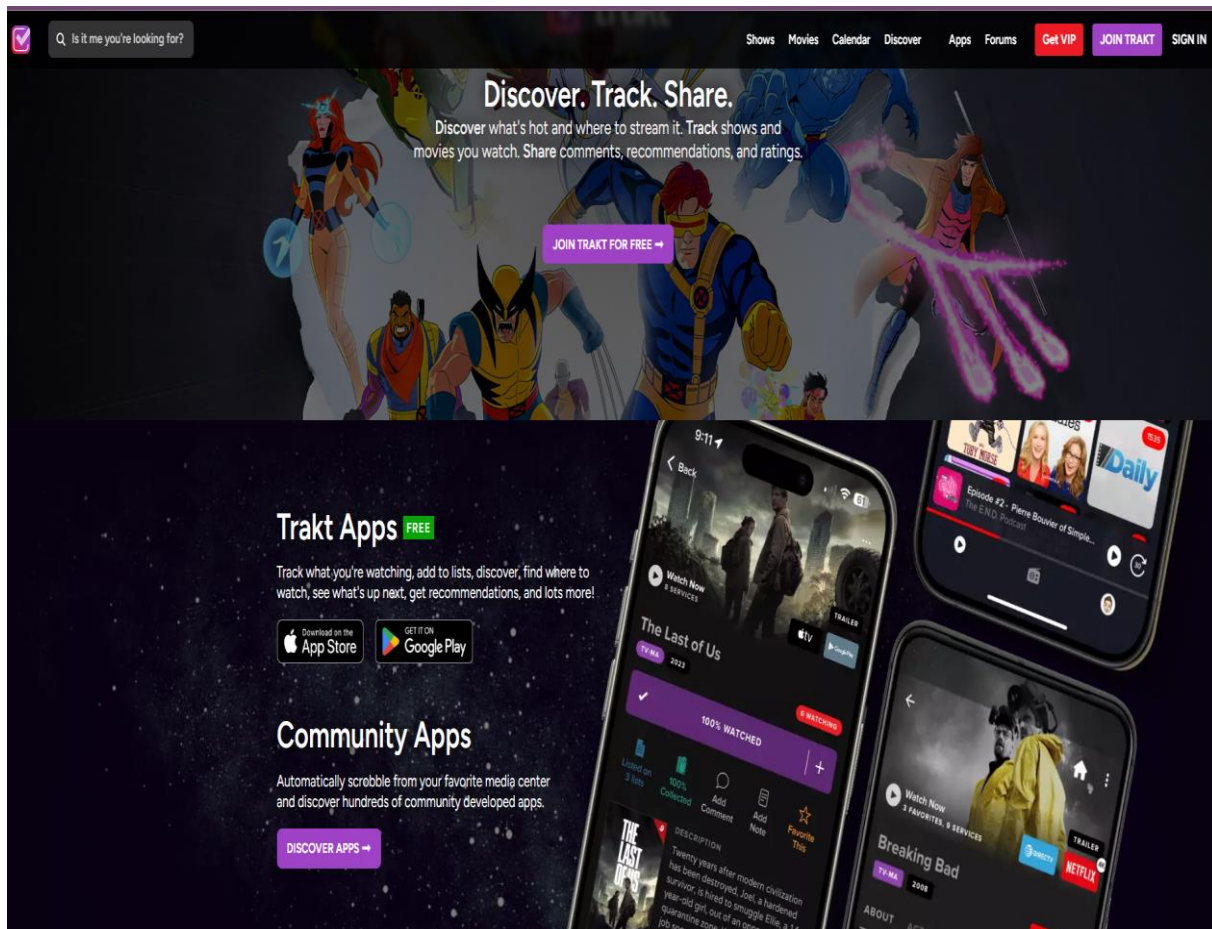
- Content-based filtering: Προτείνει παρόμοιες ταινίες με αυτές που έχει δει ή βαθμολογήσει.
- Collaborative filtering: Σου δείχνει τι έχουν δει ή προτείνει χρήστες με παρόμοιο γούστο.

Επιπλέον, μπορεί ο χρήστης να ανακαλύψει νέες ταινίες και σειρές μέσω λιστών που έχουν δημιουργηθεί από την κοινότητα ή μέσα από τις ημερήσιες/εβδομαδιαίες τάσεις. Το Trakt ξεχωρίζει γιατί είναι πολύ ανοιχτό στη διασύνδεση με άλλες υπηρεσίες. Υποστηρίζει:

- Plex, Kodi, MediaPortal, VLC κ.ά., μέσω ειδικών plugins ή εφαρμογών.
- Netflix, μέσω τρίτων εργαλείων που τραβούν δεδομένα και τα στέλνουν στο Trakt.
- API δικής του ανάπτυξης, που μπορούν να χρησιμοποιήσουν developers για να δημιουργήσουν δικές τους εφαρμογές που αλληλεπιδρούν με το Trakt.

Κεφάλαιο 2

Αυτό επιτρέπει πλήρη αυτοματοποίηση της καταγραφής προβολών, κάτι που πολλές άλλες πλατφόρμες δεν προσφέρουν.



Εικόνα 2.6: Αρχική σελίδα Trakt.tv [24]

Το MovieTheater δεν βασίζεται μόνο σε καταγραφή προβολών, αλλά επικεντρώνεται στην άμεση αναζήτηση και πρόσβαση σε περιεχόμενο μέσω streaming (όπως Netflix), κάτι που το Trakt δεν υποστηρίζει άμεσα. Επιπλέον, το σύστημα συστάσεων είναι σχεδιασμένο να προσφέρει προτάσεις με βάση τα αγαπημένα του χρήστη αλλά και αυτά των άλλων χρηστών, εντός της ίδιας πλατφόρμας, χωρίς να απαιτεί εξωτερικό συγχρονισμό. Με λίγα λόγια, η εφαρμογή απλοποιεί τη διαδικασία για τον τελικό χρήστη: δεν χρειάζεται να συγχρονίσει με άλλες εφαρμογές ή εργαλεία, απλώς φτιάχνει προφίλ, επιλέγει αγαπημένα και λαμβάνει προτάσεις και συνδέσμους για προβολή άμεσα.

Κεφάλαιο 3^ο: Recommendation Systems

3.1 Εισαγωγή στα Recommendation Systems

Τα συστήματα σύστασης (recommendation systems ή recommender systems) είναι τεχνολογίες οι οποίες βοηθούν τους χρήστες της εκάστοτε εφαρμογής να ανακαλύψουν υπηρεσίες, προϊόντα ή περιεχόμενο (π.χ. βιβλία, μουσική, ταινίες, τηλεοπτικά προγράμματα, ειδησεογραφικά άρθρα, ιστοσελίδες, δημοσιεύσεις, ταξιδιωτικές υπηρεσίες, εφαρμογές λογισμικού ή/και άλλα καταναλωτικά αγαθά και υπηρεσίες) που πιθανώς τους ενδιαφέρει. Ακολουθώντας κάποιον αλγόριθμο βασισμένο στην ανάλυση συμπεριφοράς καθώς και των ιδιαίτερων προτιμήσεων που μπορεί να έχει κάθε χρήστης ξεχωριστά, παρέχουν εξατομικευμένες προτάσεις που πιθανότατα να ταιριάζουν στον αντίστοιχο χρήστη και θα του λύσει τα χέρια από το να σπαταλήσει πολύ χρόνο ψάχνοντας μόνος και ίσως τον αποτρέψει και από κάποια εκ των υστέρων 'λανθασμένη' επιλογή.

Τα συστήματα συστάσεων είναι εξαιρετικά σημαντικά στη σύγχρονη ψηφιακή εποχή και παίζουν ρόλο σε πολλούς τομείς αφού ως μια γενική αναφορά βελτιώνουν την εμπειρία του χρήστη, αυξάνουν τα έσοδα, βοηθούν στην καλύτερη διαχείριση τεράστιων ποσοτήτων πληροφοριών, δημιουργούν ανταγωνιστικό πλεονέκτημα και μαθαίνουν από τα δεδομένα (χρήση machine learning).[25]

Δύο είναι οι βασικοί τύποι recommendation system, το content-based filtering και το collaborative filtering. Το content-based filtering ελέγχει τι χαρακτηριστικά έχουν τα ίδια τα αντικείμενα που πρόκειται να προτείνει (τι είδος είναι, τι περιγραφή έχουν, τι χρώμα, τι λέξεις περιέχουν κτλ.) και εμφανίζει αυτά που κάνουν αντιστοίχιση, ταιριάζουν δηλαδή ή μοιάζουν με πράγματα που άρεσαν στον χρήστη προηγουμένως.

Για παράδειγμα, αν ένας χρήστης σε μια εφαρμογή προβολής ταινιών (Netflix) παρακολουθήσει και του αρέσει μια πολεμική ταινία, το σύστημα θα το εμφανίσει παρόμοιες ταινίες (πολεμικές) όχι γιατί τις είδαν και άλλοι χρήστες με τους οποίους μπορεί να συνδέεται ή επειδή έγιναν viral αλλά γιατί αυτές οι ταινίες έχουν παρόμοιο περιεχόμενο.

Τα βήματα λοιπόν που θα ακολουθήσει είναι πρώτα να μάθει τι είναι το κάθε αντικείμενο, στην συνέχεια να μάθει τι αρέσει στον χρήστη και τέλος να ταιριάζει αυτά τα δύο. Ωστόσο ίσως ο όρος content-based να μην είναι απόλυτα σωστός ή τουλάχιστον όχι πάντα καθώς τα συστήματα που χρησιμοποιούν αυτόν τον τύπο δεν κοιτά πάντα αυτό καθ' αυτό το περιεχόμενο (τις περισσότερες φορές βασίζεται στο ίδιο το περιεχόμενο), το ίδιο το βίντεο δηλαδή (ή την μουσική αν πρόκειται για εφαρμογές τύπου Spotify) αλλά κάποια keys που περιγράφουν το περιεχόμενο αυτό (τίτλος, ετικέτες, λέξεις κλειδιά).

Ο τύπος αυτός έχει λοιπόν την δυνατότητα να δείξει πράγματα στον χρήστη που πραγματικά τον ενδιαφέρουν, εξηγεί στον χρήστη τον λόγο που το προτείνει κάτι (π.χ. «επειδή παρακολούθησες το χ κανάλι) ενώ δεν χρειάζεται να γνωρίζει τι κάνουν οι άλλοι χρήστες της εφαρμογής, ενδιαφέρεται ξεχωριστά για τον χρήστη στον οποίο θα προτείνει περιεχόμενο. Υπάρχει όμως και η άλλη όψη του νομίσματος, αυτή με τα μειονεκτήματα που παρουσιάζει ο συγκεκριμένος τύπος ξεκινώντας με το ότι αν σταματήσει σε ένα είδος, υπάρχει το ενδεχόμενο να μην προτείνει τίποτα καινούριο ή διαφορετικό. Επίσης για να λειτουργήσει σωστά, πρέπει να έχει καλές πληροφορίες. Πρώτα πρώτα για το αντικείμενο και δεύτερον για τον χρήστη. Αν ο χρήστης είναι νέος στην εφαρμογή και δεν έχει ιστορικό η το αντικείμενο είναι νέο, δεν μπορεί το σύστημα με αυτόν τον τύπο να τα ταιριάζει.

Από την άλλη το collaborative filtering ελέγχει αν δύο άτομα (ή δύο αντικείμενα) έχουν παρόμοια γούστα ή απόψεις για κάποια πράγματα (π.χ. ενδιαφέρονται για του ίδιου τύπου ταινίες) οπότε και υποθέτει ότι πιθανόν να συμφωνούν και σε άλλα. Έτσι μπορεί και προβλέπει τι θα αρέσει ίσως σε κάποιο χρήστη, με βάση το τι άρεσε σε κάποιον άλλο χρήστη που του "μοιάζει".

Το collaborative filtering χωρίζεται σε κάποιες κατηγορίες οι οποίες με την σειρά τους λειτουργούν με την δική τους λογική.

Στη πρώτη κατηγορία, η οποία ονομάζεται Memory-based collaborative filtering, υπάρχουν δύο μέθοδοι για να προτείνει το σύστημα περιεχόμενο.

- Η πρώτη μέθοδος, user-based, στην οποία η σύγκριση που κάνει το σύστημα γίνεται μεταξύ δύο χρηστών και αν οι δυο τους έχουν κοινά στοιχεία στις αξιολογήσεις τους, ακολουθείται ένα recommendation χιαστί, προτείνει δηλαδή στο έναν χρήστη αυτά που άρεσαν στον άλλο.
- Η δεύτερη μέθοδος, item-based, στην οποία η σύγκριση γίνεται μεταξύ των προϊόντων. Το σύστημα βλέπει ένα προϊόν που άρεσε στον χρήστη και αν βρει κάποιο άλλο που να του μοιάζει, του το προτείνει.

Ο υπολογισμός ομοιότητας που χρησιμοποιεί αυτός ο τύπος γίνεται με μαθηματικές μεθόδους όπως person - cosine similarity. Σε περίπτωση ελλιπών αξιολογήσεων, το σύστημα βάζει μια προκαθορισμένη κριτική (default voting). Με την μέθοδο case amplification (Ενίσχυση τιμών), το σύστημα δίνει περισσότερη βάση σε χρήστες με ισχυρές απόψεις έναντι στους πιο “αδιάφορους”.

Η δεύτερη κατηγορία αυτού του τύπου, η model-based collaborative filtering, δεν συγκρίνει άμεσα τους χρήστες ή τα προϊόντα αλλά δημιουργεί μαθηματικά μοντέλα (παρόμοια με αυτά της στατιστικής ή της τεχνητής νοημοσύνης) όπως clustering, SVD, PCA, bayesian δίκτυα. Αυτά τα μοντέλα μπορούν να λειτουργήσουν καλύτερα σε μεγάλες βάσεις δεδομένων και όταν τα δεδομένα είναι λίγα ή διάσπαρτα, θέλουν όμως αρκετό χρόνο και υπολογιστική ισχύ για να εκπαιδευτούν καθώς επίσης απαιτεί σωστή ισορροπία μεταξύ ακρίβειας και ταχύτητας.

Τέλος υπάρχει μια τρίτη κατηγορία, η hybrid collaborative filtering, στην οποία γίνεται μια μίξη των προηγούμενων δυο. Συνδυάζεται το collaborative filtering με στοιχεία από το προφίλ του χρήστη (προτιμήσεις τους), personality diagnosis, ή με τα χαρακτηριστικά ενός προϊόντος (είδος ταινίας), content-boosted. Είναι πιο ευφύες διότι καλύπτει τα κενά, τις περιπτώσεις που αδυνατούν τα προηγούμενα μοντέλα αλλά είναι πιο δύσκολο στην υλοποίησή τους. [11] [12]

Οι μεγαλύτερες πλατφόρμες δεν επιλέγουν ή τον έναν ή τον άλλο τύπο αλλά έναν συνδυασμό από τους δύο για καλύτερα αποτελέσματα. Το Netflix για παράδειγμα το διάστημα 2006-2009, πραγματοποίησε έναν διαγωνισμό με έπαθλο 1.000.00 ευρώ σε όποια προγραμματιστική ομάδα κατάφερε να αυξήσει την ακρίβεια του recommendation system κατά 10%.

Μια ομάδα πέτυχε το 2009 αλλά το Netflix δεν χρησιμοποίησε ποτέ τον νέο αυτόν αλγόριθμο καθώς ήταν πολύ δαπανηρός (βελτίωση συστήματος < επένδυση) αλλά και απαιτούσε αναθεώρηση του επιχειρηματικού μοντέλου που χρησιμοποιούσε η πλατφόρμα. Εν τέλει αυτό που έκανε ήταν να χρησιμοποιήσει τους δύο τύπους μαζί για να δημιουργήσει ένα ακόμη καλύτερο recommendation system, με τον συνδυασμό να αποδίδει περίπου σε αύξηση της ακρίβειας προτάσεων περίπου ~8.5% [13].

Κύριοι στόχοι των recommendation systems

1. Εξατομίκευση Εμπειρίας: Η βασική ιδέα αυτών των αλγορίθμων και πρώτος στόχος είναι η προσαρμογή ουσιαστικά της εφαρμογής στις ανάγκες της κάθε περίπτωσης ώστε ο χρήστης να βλέπει περιεχόμενο σχετικό με τα προσωπικά του γούστα.
2. Μείωση υπερφόρτωσης πληροφορίας: Τα μεγέθη των πληροφοριών που διαθέτουν οι βάσεις μεγάλων εφαρμογών θα μπορούσαν να χαρακτηριστούν ωκεανοί πληροφοριών. Η χρήση συστημάτων συστάσεων δίνει την δυνατότητα στον χρήστη να βρει αυτό που τον ενδιαφέρει χωρίς να ταλαιπωρηθεί ψάχνοντας μέσα σε αυτόν τον ωκεανό.

3. Αύξηση αλληλεπίδρασης και engagement: Η απλή ανθρώπινη λογική λέει πως όταν ο χρήστης βλέπει περιεχόμενο σχετικό με ό,τι τον ενδιαφέρει, είναι πιθανότερο να παραμείνει περισσότερο στην πλατφόρμα.
4. Νέα περιεχόμενα: Οι αλγόριθμοι δίνοντας φιλτραρισμένο και μικρό όγκο περιεχομένου που θα ταιριάζει στον χρήστη, του δίνει την ευκαιρία να ανακαλύψει-γνωρίσει πράγματα που ίσως να μην τα έψαχνε μόνος του.
5. Έσοδα για τις πλατφόρμες: Αναμφίβολα όλοι οι παραπάνω στόχοι αποσκοπούν μέσω της διευκόλυνσης του χρήστη στην αύξηση των εσόδων της πλατφορμας. Στο e-commerce (Amazon, ebay, Skrutz), στο streaming (Netflix, disney +, spotify) ή στα social media (Youtube, TikTok) όση περισσότερη ώρα περνάει ο χρήστης στην πλατφόρμα, τόσες περισσότερες διαφημίσεις και αγορές άρα και αυξημένα έσοδα.

Που χρησιμοποιούνται τα recommendation systems:

- E-commerce: (Amazon, ebay): Θα προταθούν προϊόντα που ίσως αρέσουν στον χρήστη βάση προηγούμενων αγορών του, αναζητήσεων η και προϊόντων που έχει κρατημένα στην wishlist.
- Social Media (YouTube, Instagram, TikTok): Τα posts (βίντεο η φωτογραφίες) που θα εμφανιστούν στον κάθε χρήστη θα είναι βασισμένα στην δραστηριότητα του κάθε χρήστη, εμφανίζοντας του παρόμοια με αυτά που βλέπει περισσότερο η το τελευταίο διάστημα.
- Online Learning (Memrise, Udacity, Udemy): Οι αλγόριθμοι συστήνουν μαθήματα, γλώσσες, τεχνολογίες που ταιριάζουν στο προφίλ του χρήστη.
- News (Google News): Προτεινόμενα άρθρα σχετικά με τις αναζητήσεις του χρήστη

3.2 Μετρικές Αξιολόγησης (Evaluation Metrics)

Η αξιολόγηση των recommendation systems αποτελεί ένα κρίσιμο στάδιο στη διαδικασία ανάπτυξης και βελτιστοποίησής τους. Καθώς τα μοντέλα προτάσεων επιχειρούν να προβλέψουν τις προτιμήσεις των χρηστών και να προσφέρουν σχετικό και ενδιαφέρον περιεχόμενο, είναι απαραίτητο να υπάρχουν αντικειμενικές μετρικές που να επιτρέπουν τη σύγκριση και την αποτίμηση της απόδοσής τους. Οι μετρικές αυτές εστιάζουν σε διαφορετικές πτυχές της λειτουργίας ενός recommendation system δηλαδή από το πόσο σχετικές είναι οι προτάσεις, μέχρι το πόσο ποικίλες, καινοτόμες ή επιδραστικές αποδεικνύονται για τον τελικό χρήστη. Στην ενότητα που ακολουθεί παρουσιάζονται ορισμένες από τις πιο διαδεδομένες μεθόδους αξιολόγησης, οι οποίες χρησιμοποιούνται ευρέως στη βιβλιογραφία και στη βιομηχανία για τη μέτρηση της ποιότητας και της χρηστικότητας των recommendation systems ανεξάρτητα την εφαρμογή τους.

Precision / Recall (Ακρίβεια / Ανάκληση): Αυτά τα δύο είναι σαν τις δύο όψεις του ίδιου νομίσματος.

- Ακρίβεια (Precision): Το Netflix για παράδειγμα προτείνει στον χρήστη στην πρώτη σειρά 10 ταινίες. Η Ακρίβεια απαντά στην ερώτηση: "Από αυτές τις 10 που μου δείχνει, πόσες όντως φαίνονται καλές και θα τις έβαζε στη λίστα του ο χρήστης;". Αν του αρέσουν οι 3, το recommendation system έχει Ακρίβεια 3/10. Είναι το ποσοστό της ευστοχίας στις προτάσεις που βλέπει μπροστά του. Ένα αξιόπιστο και ποιοτικό σύστημα δεν πρέπει σε καμία περίπτωση να προβάλλει περιεχόμενο χαμηλής ποιότητας

- **Ανάκληση (Recall):** Τώρα ο συνδρομητής (αν πρόκειται για το Netflix ή οποιαδήποτε πλατφόρμα που δουλεύει με συνδρομητές) πρέπει να σκεφτεί κάτι διαφορετικό. Σε ολόκληρη τη βιβλιοθήκη του Netflix, υπάρχουν ίσως 50 ταινίες που θα ταίριαζαν απόλυτα στον χρήστη. Από τις 10 που του πρότεινε, οι 3 ήταν μέσα σε αυτές τις 50. Η Ανάκληση απαντά στην ερώτηση: "Από όλες τις ταινίες που θα του άρεσαν, πόσες κατάφερε το σύστημα να του βρει και να του τις δείξει;". Εδώ είναι 3/50. Μετράει πόσο καλά το σύστημα "ψαρεύει" τις σωστές ταινίες από τον τεράστιο ωκεανό του περιεχομένου.

Το πρόβλημα σε αυτή την μέτρηση είναι πως αν για παράδειγμα το Netflix προτείνει σε έναν συγκεκριμένο χρήστη συνέχεια ταινίες του Adam Sandler επειδή είδε μία και του άρεσε, θα έχει υψηλή Ακρίβεια (πιθανότητα θα τις δει), αλλά χαμηλή Ανάκληση (χάνει όλες τις άλλες κωμωδίες από άλλους ηθοποιούς που θα του άρεσαν).

Hit Rate (Ποσοστό Επιτυχίας): Είναι ένα πολύ βασικό τεστ επιτυχίας.

- **Hit Rate:** Έστω ότι ένας χρήστης άνοιξε το YouTube. Αν στις προτάσεις που του έβγαλε, υπήρχε έστω και ένα βίντεο που του κίνησε το ενδιαφέρον και έκανε κλικ, τότε το YouTube πέτυχε έναν "Hit" μαζί του. Το Hit Rate είναι απλά το ποσοστό των χρηστών που το σύστημα κατάφερε να τους κάνει να πατήσουν "play" έστω σε μία από τις προτάσεις του. Είναι η απόλυτη βάση αν το σύστημα κάνει έστω και το ελάχιστο που πρέπει. Γι' αυτό και θεωρείται βασικό καθώς είναι η αρχή του να θεωρηθεί ένα σύστημα προτάσεων, όπως αυτό που χρησιμοποιεί το YouTube, πετυχημένο.

F1 Score: Αυτός είναι ο μετρητής που προσπαθεί να κρατήσει την ισορροπία.

- **F1 Score:** Επειδή η Ακρίβεια και η Ανάκληση συχνά λειτουργούν αντίθετα σαν 2 μαθηματικές τιμές που είναι αντιστρόφως ανάλογες (αν ανέβει η μία, πέφτει η άλλη), το F1 Score τα συνδυάζει σε ένα νόμμο. Είναι ένας έξυπνος μέσος όρος που δεν αφήνει το σύστημα ούτε να είναι υπερβολικά "ασφαλές" (υψηλή Ακρίβεια, χαμηλή Ανάκληση) ούτε να προτείνει ό,τι να 'ναι με την ελπίδα να πετύχει κάτι (χαμηλή Ακρίβεια, υψηλή Ανάκληση). Στην ουσία, αναγκάζει το σύστημα να είναι ταυτόχρονα εύστοχο και να ανακαλύπτει νέες επιλογές για τον χρήστη.

Coverage (Κάλυψη): Αυτό κοιτάζει ολόκληρη τη "βιβλιοθήκη" της πλατφόρμας.

- **Coverage:** Το Netflix έχει χιλιάδες ταινίες και σειρές. Μέσα σε έναν χρόνο, πόσες από αυτές προτάθηκαν έστω και μία φορά σε κάποιον χρήστη; Αν το σύστημα προτείνει συνέχεια τις ίδιες 500 δημοφιλείς παραγωγές, τότε έχει χαμηλή Κάλυψη. Αυτό σημαίνει ότι χιλιάδες ταινίες (π.χ. παλιές, ξενόγλωσσες, ανεξάρτητες) μένουν "θαμμένες" και κανείς δεν τις ανακαλύπτει. Ένα καλό σύστημα δίνει μια ευκαιρία σε ολόκληρο τον κατάλογο του να φτάσει στους θεατές.

Diversity - Novelty - Serendipity (Ποικιλία - Καινοτομία - Τυχαία Ανακάλυψη): Αυτά τα τρία μετράνε πόσο "ενδιαφέρον" και "έξυπνο" είναι το Netflix (ή όποια άλλη πλατφόρμα για την οποία γίνονται οι μετρήσεις αυτές).

- **Ποικιλία (Diversity):** Έστω ότι ένας συνδρομητής παρακολούθησε το "The Dark Knight". Το σύστημα του μπορεί να προτείνει μόνο άλλες ταινίες με σούπερ-ήρωες. Ή καταλαβαίνει ότι του αρέσουν οι ταινίες με σκοτεινή ατμόσφαιρα και περίπλοκους χαρακτήρες και του προτείνει και το "Se7en" (αστυνομικό θρίλερ) και το "Blade Runner" (ταινίες τρόμου). Η ποικιλία εξασφαλίζει ότι οι προτάσεις του έχουν εύρος και δεν τον κλειδώνουν σε ένα μόνο είδος.

- **Καινοτομία (Novelty):**Μπορεί να του προτείνει το "Stranger Things" που το βλέπουν όλοι ή να ανακαλύψει μια φανταστική, αλλά άγνωστη νοτιοκορεάτικη σειρά τρόμου που δεν ήξερε ότι υπάρχει. Η καινοτομία μετράει την ικανότητα του συστήματος να προτείνει στον πελάτη "κρυμμένα διαμάντια", όχι μόνο τα προφανή hits.
- **Τυχαία Ανακάλυψη (Serendipity):** Αυτή είναι η "μαγεία". Ο ίδιος πάλι χρήστης βλέπει συνέχεια ταινίες δράσης. Ξαφνικά, το σύστημα του προτείνει ένα ντοκιμαντέρ για αγώνες Formula 1, το "Drive to Survive". Φαινομενικά άσχετο, αλλά το βλέπει και εθίζεται, γιατί ανακαλύπτει ότι το στοιχείο της αδρεναλίνη και του ανταγωνισμού, που του αρέσει στη δράση, υπάρχει κι εδώ. Serendipity είναι αυτή η ευχάριστη, απρόσμενη έκπληξη που τον κάνει να πει "Δεν θα το σκεφτόμουν ποτέ, αλλά είναι τέλειο!" [18].

3.3 Παραδείγματα αποτελεσματικότητας συστημάτων σύστασης

1. Το Netflix χρησιμοποιεί εξελιγμένους αλγορίθμους μηχανικής μάθησης για να προτείνει στους χρήστες της πλατφόρμας-εφαρμογής ταινίες και σειρές που πιθανώς θα τους αρέσουν, βασίζόμενο πάντα σε προηγούμενη συμπεριφορά θέασης του κάθε χρήστη, αξιολογείσες αλλά και άλλα δεδομένα [1]. Παράδειγμα του αλγορίθμου, αν ένας χρήστης παρακολουθεί μια σειρά, το σύστημα θα του προτείνει παρόμοιες σειρές ή ταινίες που έχουν υψηλή πιθανότητα να του αρέσουν αυξάνοντας έτσι και την πιθανότητα να παραμείνει στην πλατφόρμα [1]

- **Μετρήσιμα αποτελέσματα:** Το Netflix εκτιμά σύμφωνα με τις μετρήσεις του ότι περίπου κάτι παραπάνω από το 80% του περιεχομένου που βλέπουν οι χρήστες προέρχεται από τις προτάσεις του αλγορίθμου. [2].
- **Επιπτώσεις:** Αυτό βελτιώνει την αφοσίωση του χρήστη (user engagement) προς την εφαρμογή, αυξάνει το χρόνο παρακολούθησης (watch time) και μειώνει την πιθανότητα ακύρωσης της συνδρομής (churn rate) σε περίπτωση δυσαρέσκειας του χρήστη αφού οι αλγόριθμοι συστάσεων πλαγίως βοηθούν να αποφευχθεί κάτι τέτοιο[1].

2. Το Amazon εφαρμόζει επίσης πολύ αποτελεσματικούς αλγορίθμους συστάσεων για τα προϊόντα του που βασίζονται στο ιστορικό αγορών, αναζητήσεων και συμπεριφορών του χρήστη ή και άλλων χρηστών με παρόμοια προφίλ [3]. Παράδειγμα του αλγορίθμου είναι οι προτάσεις του τύπου "Άλλοι χρήστες αγόρασαν επίσης ..." ή "Θα σου ταίριαζε επίσης - Recommended for you". Βοηθούν με αυτόν τον τρόπο στην ανακάλυψη προϊόντων και την τόνωση της ζήτησης[4].

- **Μετρήσιμα αποτελέσματα:** Η εταιρεία σύμφωνα με τις μετρήσεις της, απέδωσε περίπου το 35% των πωλήσεων της στις συστάσεις προϊόντων[3]
- **Επιπτώσεις:** Οι συστάσεις ενισχύουν σε πολύ θετικό για την αγορά τις πωλήσεις μέσω αύξησης του μέσου μεγέθους καλαθιού (average basket size) και της συχνότητας αγορών [3].

3. Το YouTube αξιοποιεί recommendation systems που λαμβάνουν υπόψη το ιστορικό του χρήστη δηλαδή την προηγούμενη θέαση, τις αλληλεπιδράσεις και τις προτιμήσεις του χρήστη για να προτείνουν τελικώς σχετικά βίντεο[5]. Παράδειγμα του αλγορίθμου είναι πως με την ολοκλήρωση της προβολής ενός βίντεο, το σύστημα προτείνει αυτόματα σχετικά βίντεο με σκοπό να κρατήσουν τον τον χρήστη σε επαφή με την πλατφόρμα προσφέροντας του εύκολα αυτό που τον ενδιαφέρει.

- **Μετρήσιμα αποτελέσματα:** Επίσημα το YouTube αναφέρει πως περισσότερο από το 70% των views της εφαρμογής, προέρχονται από προτεινόμενα βίντεο [8].
- **Επιπτώσεις:** Η προβολή περιεχομένου που ενδιαφέρει τον χρήστη αυξάνει τον χρόνο που ο ίδιος θα ξοδέψει στην πλατφόρμα, ενισχύεται η αφοσίωση και βελτιώνεται η διατήρηση των χρηστών[5].

3.4 Άλλοι τομείς εφαρμογής των συστημάτων σύστασης

Παρότι η πιο ορατή και διαδεδομένη χρήση των συστημάτων σύστασης εντοπίζεται στις πλατφόρμες κοινωνικής δικτύωσης και ψυχαγωγίας – όπως το Netflix, το YouTube και το Spotify – η εφαρμογή τους δεν περιορίζεται αποκλειστικά εκεί. Στην πραγματικότητα, τα recommender systems έχουν αρχίσει να διαδραματίζουν κρίσιμο ρόλο και σε άλλους τομείς της καθημερινότητας, όπως το ηλεκτρονικό εμπόριο, η υγειονομική περίθαλψη και οι χρηματοοικονομικές υπηρεσίες.

Κάθε ένας από αυτούς τους τομείς παρουσιάζει ξεχωριστές απαιτήσεις και προκλήσεις: η εξατομίκευση της εμπειρίας του καταναλωτή στο ηλεκτρονικό εμπόριο, η ανάγκη για ακρίβεια και ασφάλεια στην υγεία, ή η διαφάνεια και αξιοπιστία στον χρηματοοικονομικό κλάδο. Τα συστήματα σύστασης, επομένως, οφείλουν να προσαρμόζονται όχι μόνο στις τεχνολογικές απαιτήσεις του κάθε περιβάλλοντος, αλλά και στις ηθικές και νομικές ευαισθησίες που αυτό συνεπάγεται.

Ηλεκτρονικό Εμπόριο (E-commerce):

- Χρησιμοποιούνται για εντοπισμό σχετικών προϊόντων, προσωποποιημένες προτάσεις και upselling/cross-selling.
- Αυξάνεται η ικανοποίηση των πελατών και τα έσοδα· σχετική αύξηση πωλήσεων αποδίδεται στην εφαρμογή RS [32]
- Επιλύονται προβλήματα όπως ο *cold-start* με χρήση υβριδικών και review-based μεθόδων.

Υγεία (Healthcare):

- Εφαρμόζονται στην προσωποποιημένη ιατρική, π.χ. διατροφή, άσκηση, φαρμακευτική αγωγή.
- Ο τομέας απαιτεί ειδική προσοχή (π.χ. content-based, knowledge-based συστήματα) διότι εμπλέκονται θέματα ασφάλειας και επαγγελματικών αποφάσεων.[32]

Χρηματοοικονομία (Finance):

- Χρησιμοποιούνται για προτάσεις επενδύσεων, διαχείριση ρίσκου και εξατομικευμένα οικονομικά προϊόντα.
- Απαιτείται αξιοπιστία, διαφάνεια και έλεγχος συγκρούσεων συμφερόντων γιατί οι αποφάσεις έχουν άμεσες οικονομικές συνέπειες. [32] (βλέπε πίνακα 3.1)

| Τομέας | Inputs | Recommendations | Αντίδραση Χρήστη |
|------------|---|---|--|
| E-commerce | Προηγούμενες αγορές, προβολές προϊόντων, αξιολογήσεις | Προϊόντα, Upsell/Cross-sell, Καλάθια αγορών | Αγορά, Προσθήκη στο καλάθι, χρόνο επίσκεψης |
| Υγεία | Ιατρικό ιστορικό, διατροφικές συνήθειες, data wearables | Πρότυπα άσκησης, διατροφής, θεραπευτικά πλάνα | Δέσμευση, συμμόρφωση, feedback ασθενούς[33] |
| Οικονομικά | Επενδυτικό προφίλ, Συναλλακτική συμπεριφορά, δημοσιονομικά δεδομένα | Χαρτοφυλάκια, επενδυτικά προϊόντα, alerts κινδύνου/ risk- based | Δημιουργία/προσαρμογή λογαριασμών, αλληλεπίδραση με συμβουλές [34] |

Πίνακας 3.1: Συγκριτικός Πίνακας Εφαρμογών Recommender Systems [32]

Κεφάλαιο 4^ο: Προβλήματα και προκλήσεις που αντιμετωπίζει κάθε recommendation system

Παρά τη ραγδαία εξέλιξή τους και τη σημαντική τους συμβολή στη βελτίωση της εμπειρίας χρήστη, τα recommendation systems δεν είναι απαλλαγμένα από αδυναμίες. Στην πράξη, καλούνται να λειτουργήσουν σε περιβάλλοντα με μεγάλα και πολύπλοκα σύνολα δεδομένων, με χρήστες που έχουν ποικίλα και δυναμικά ενδιαφέροντα, αλλά και με συνεχώς μεταβαλλόμενο περιεχόμενο. Η αποτελεσματικότητα αυτών των συστημάτων συχνά περιορίζεται από τεχνικές και θεμελιώδεις προκλήσεις, όπως η έλλειψη δεδομένων για νέους χρήστες ή αντικείμενα (cold start), η αραιότητα των αξιολογήσεων, η υπολογιστική επιβάρυνση (scalability) και ζητήματα μεροληψίας ή κακόβουλης χρήσης. Στην ενότητα που ακολουθεί, παρουσιάζονται τα σημαντικότερα προβλήματα που εντοπίζονται στη λειτουργία των recommendation systems, καθώς και ενδεικτικές προσεγγίσεις για την αντιμετώπισή τους.

Cold Start Problem: Το πρόβλημα εμφανίζεται όταν το σύστημα δεν έχει αρκετά δεδομένα για να βγάλει συμπεράσματα. Συμβαίνει είτε με νέους χρήστες που δεν έχουν βαθμολογήσει τίποτα, είτε με νέες ταινίες που μόλις προστέθηκαν. (επεξηγείται περισσότερο παρακάτω)

- Νέος χρήστης: Μόλις έφτιαξε κάποιος λογαριασμό στο Netflix. Το Netflix τον κοιτάζει σαν έναν άγνωστο και δεν έχει ιδέα τι να του προτείνει. Είναι "κρύο" απέναντί του.
- Νέο αντικείμενο: Μια ολοκαίνουργια ταινία ανεβαίνει στην πλατφόρμα. Κανείς δεν την έχει δει ακόμα. Το Netflix δεν ξέρει αν είναι καλή, κακή, ή σε ποιους θα άρεσε. Η ταινία είναι "κρύα".

Οι λύσεις που προτείνει:

- Να σε ρωτήσει ευθέως τι αρέσει στον νέο χρήστη (π.χ. "Διάλεξε 3 είδη που προτιμάς").
- Να του δείξει μερικές ταινίες και να του ζητήσει να τις βαθμολογήσει.
- Να χρησιμοποιήσει δημογραφικά στοιχεία (ηλικία, χώρα) για μια πρώτη, γενική πρόταση.

Shilling Attack Problem: Κακόβουλοι χρήστες δημιουργούν ψεύτικα προφίλ για να δώσουν ψεύτικες βαθμολογίες. Στόχος τους είναι είτε να ανεβάσουν τεχνητά τη δημοτικότητα μιας ταινίας, είτε να τη "θάψουν". Έστω ένα κινηματογραφικό στούντιο να θέλει να προωθήσει τη νέα του, μέτρια ταινία. Δημιουργεί 10,000 ψεύτικους λογαριασμούς στο Netflix και όλοι τους δίνουν 5 αστέρια στην ταινία. Ξαφνικά, η ταινία φαίνεται σαν τεράστια επιτυχία και το Netflix αρχίζει να την προτείνει σε αληθινούς χρήστες. Είναι απάτη που χαλάει την αξιοπιστία του συστήματος.

Η λύση που προτείνει:

- Το σύστημα πρέπει να λειτουργεί σαν "ντετέκτιβ" για να εντοπίζει αυτά τα ψεύτικα προφίλ και να σβήνει τις ψεύτικες βαθμολογίες τους.

Synonymy Problem (πρόβλημα συνωνυμίας): Το πρόβλημα συμβαίνει όταν το ίδιο πράγμα έχει διαφορετικά ονόματα στο σύστημα (στο Amazon ή στο Skrutz π.χ. "βρεφικά ρούχα" και "παιδικά ενδύματα"). Το σύστημα δεν καταλαβαίνει ότι είναι το ίδιο πράγμα. Έχει δει ο χρήστης το "Star Wars: A New Hope". Το Netflix μπορεί να μην καταλαβαίνει ότι αυτό συνδέεται με την ετικέτα "Sci-Fi" ή "Επιστημονική Φαντασία" ή "ταινίες με τον Luke Skywalker". Αν δεν μπορεί να κάνει αυτές τις συνδέσεις, χάνει την ευκαιρία να του προτείνει σχετικό περιεχόμενο. Είναι σαν να μιλάει μια πολύ αυστηρή και κυριολεκτική γλώσσα, χωρίς να καταλαβαίνει τις αποχρώσεις.

Latency Problem: Το σύστημα αργεί να προτείνει καινούργια αντικείμενα. Αυτό συμβαίνει κυρίως σε συστήματα που περιμένουν πρώτα να βαθμολογηθεί μια ταινία από άλλους, πριν την προτείνουν. Βγαίνει μια τυχαία μέρα το πρώτο επεισόδιο μιας πολυαναμενόμενης σειράς. Το σύστημα όμως περιμένει να δει τι θα πουν οι άλλοι χρήστες. Έτσι, μπορεί να περάσουν ώρες (ή και μέρες) μέχρι να το προτείνει, παρόλο που είναι ακριβώς αυτό που θα ήθελε να δει ο χρήστης. Υπάρχει μια καθυστέρηση (latency) ανάμεσα στο πότε κάτι γίνεται διαθέσιμο και πότε το σύστημα νιώθει αρκετά σίγουρο για να το προτείνει.

Grey Sheep Problem: Το πρόβλημα αφορά χρήστες των οποίων οι προτιμήσεις είναι τόσο μοναδικές που δεν ταιριάζουν με καμία ομάδα χρηστών. Υπάρχει ένας χρήστης, που το γούστο του περιλαμβάνει γαλλικές κωμωδίες του '60, κορεάτικα θρίλερ και ντοκιμαντέρ για τη φύση. Το σύστημα προσπαθεί να τον βάλει σε ένα "κουτάκι" για να του βρει χρήστες με παρόμοιο γούστο, αλλά δεν μπορεί. Δεν ανήκει ούτε στην ομάδα "fans των γαλλικών ταινιών" ούτε στην ομάδα "fans των θρίλερ". Είναι ένα "γκρι πρόβατο" και οι προτιμήσεις του είναι πολύ ιδιαίτερες για να ομαδοποιηθούν. Για αυτόν, οι συστάσεις που βασίζονται σε "άλλους χρήστες" αποτυγχάνουν παταγωδώς.

Scalability Problem: Το πρόβλημα εμφανίζεται όταν η ποσότητα των δεδομένων (χρήστες και ταινίες) αυξάνεται πάρα πολύ γρήγορα. Ας αναλογιστεί κάποιος το μέγεθος του Netflix. Έχει εκατοντάδες εκατομμύρια χρήστες και δεκάδες χιλιάδες τίτλους. Αν ο αλγόριθμος προσπαθούσε να συγκρίνει κάθε χρήστη με κάθε άλλον χρήστη κάθε φορά που ανοίγει κανείς την εφαρμογή, οι σέρβερ θα κατέρρεαν. Το σύστημα απλά δεν θα "κλιμακωνόταν", δεν θα μπορούσε να ανταπεξέλθει σε αυτόν τον τεράστιο όγκο. Πρέπει να βρεθούν πιο έξυπνοι και γρήγοροι τρόποι (όπως το να βάζουν τους χρήστες σε μικρότερες ομάδες-clusters) για να λειτουργεί το σύστημα χωρίς να "κολλάει".

Sparsity Problem: Το πρόβλημα εμφανίζεται όταν οι χρήστες έχουν βαθμολογήσει πολύ λίγα πράγματα. Τα δεδομένα είναι "αραιά", δηλαδή γεμάτα κενά. Κάτι σαν ένα τεράστιο φύλλο Excel στο οποίο κάθε σειρά είναι ένας χρήστης και κάθε στήλη είναι μία ταινία του Netflix. Οι περισσότερες χιλιάδες κελιών θα είναι άδεια, γιατί ο μέσος χρήστης έχει δει ένα ελάχιστο κλάσμα της συνολικής βιβλιοθήκης. Αυτή η "αραιότητα" κάνει πολύ δύσκολο για το σύστημα να βρει χρήστες με παρόμοια γούστα, γιατί απλά δεν υπάρχουν αρκετές κοινές ταινίες που να έχουν βαθμολογηθεί για να γίνει η σύγκριση [19].

Προκλήσεις που αντιμετωπίζει το Recommendation system

Cold start

Όταν θέλει μια πλατφόρμα να χτίσει ένα σύστημα που προτείνει περιεχόμενο, όπως στο Netflix, υπάρχει μια βασική δυσκολία: το τι κάνει με νέες ταινίες ή σειρές που μόλις έχουν προστεθεί στην βάση δεδομένων του συστήματος και δεν έχουν ακόμα καμία προβολή. Δηλαδή, πώς μπορεί να προτείνει κάτι που κανένας δεν έχει δει ακόμη;

Εδώ μπαίνει το λεγόμενο "cold start" πρόβλημα. Για να λυθεί, χρησιμοποιείται ένας συνδυασμός τεχνικών, Incremental training (σταδιακή προπόνηση) και Dealing with unseen entities (αόρατες οντότητες). Αρχικά, δεν χρειάζεται να φτιάχεται το μοντέλο από την αρχή κάθε φορά. Αντίθετα, ανανεώνεται προοδευτικά. Όταν προστίθεται ένας νέος τίτλος, το σύστημα δημιουργεί μια αρχική "εικόνα" του βασισμένη σε παρόμοιους τίτλους - είτε βάζοντας λίγη τυχαιότητα είτε χρησιμοποιώντας πληροφορίες όπως το είδος του ή η γλώσσα.

Στην συνέχεια, για να μπορεί να προτείνει έναν τίτλο που δεν έχει ξαναδεί, το μοντέλο δεν βασίζεται μόνο σε τι έχει δει ο χρήστης, αλλά και σε metadata: δηλαδή τι είδους είναι ο τίτλος, αν είναι κωμωδία ή θρίλερ, ποια γλώσσα έχει, πότε κυκλοφόρησε, κλπ. Ένας έξυπνος τρόπος που χρησιμοποιούν είναι το να δίνουν περισσότερη σημασία στα metadata όταν ο τίτλος είναι καινούργιος, και λιγότερη όσο περνάει ο καιρός και μαζεύονται πραγματικά δεδομένα προβολής.

Κεφάλαιο 4

Έτσι, το μοντέλο μπορεί να προτείνει με νόημα ακόμα και περιεχόμενο που δεν έχει καμία αξιολόγηση ή προβολή - κάτι πολύ χρήσιμο για πλατφόρμες που ανεβάζουν συνέχεια καινούρια πράγματα.

Downstream Applications και προκλήσεις

Το μοντέλο που έχουν φτιάξει στο Netflix έχει ως βασικό στόχο να "καταλαβαίνει" τι προτιμά κάθε χρήστης σε βάθος χρόνου. Το ωραίο είναι ότι αυτό το μοντέλο δεν είναι φτιαγμένο μόνο για ένα συγκεκριμένο πράγμα, αλλά μπορεί να χρησιμοποιηθεί με διάφορους τρόπους με downstream applications, ανάλογα με το τι χρειάζεται μια εφαρμογή.

Ένας τρόπος είναι ο Direct use as a predictive model (Ως προγνωστικό μοντέλο) όπου η βασική του δουλειά είναι να "μαντεύει" τι θα δει μετά ο χρήστης. Μάλιστα, δεν κάνει μόνο μια πρόβλεψη, αλλά έχει και ξεχωριστές "μονάδες" που προβλέπουν πιο συγκεκριμένα πράγματα, όπως τι είδους περιεχόμενο προτιμά κάποιος. Αυτό βοηθάει πολύ για πιο στοχευμένες προτάσεις.

Ο δεύτερος τρόπος ονομάζεται utilizing embeddings όπου το μοντέλο δημιουργεί ποιοτικά embeddings (μαθηματικές αναπαραστάσεις) για χρήστες και οντότητες όπως βίντεο, παιχνίδια και είδη. Αυτά υπολογίζονται σε μαζικές επεξεργασίες (batch jobs) και αποθηκεύονται για χρήση τόσο σε offline όσο και σε online εφαρμογές

Μπορούν να χρησιμοποιηθούν:

- Ως χαρακτηριστικά (features) σε άλλα μοντέλα
- Για εύρεση υποψηφίων τίτλων προς πρόταση (candidate generation)
- Για να γίνουν προτάσεις τύπου "αν σου άρεσε αυτό, τότε ίσως σου αρέσει και αυτό"

Ωστόσο, υπάρχει μια σημαντική δυσκολία: ο "χώρος" των embeddings δεν έχει σταθερή, ερμηνεύσιμη δομή και μπορεί να αλλάζει κάθε φορά που γίνεται επανεκπαίδευση του μοντέλου. Αυτό προκαλεί προβλήματα για άλλες εφαρμογές που χρησιμοποιούν αυτά τα embeddings, γιατί μπορεί να βασίζονται σε παλιές υποθέσεις που δεν ισχύουν πια.

Για να λυθεί αυτό το πρόβλημα, το σύστημα εφαρμόζει έναν ορθογώνιο μετασχηματισμό χαμηλής τάξης, ο οποίος σταθεροποιεί τον χώρο των embeddings. Έτσι, διατηρείται η συνέπεια στη σημασία κάθε διάστασης, ακόμα και όταν το μοντέλο εκπαιδεύεται ξανά ή αλλάζει.

Τρίτος τρόπος είναι η προσαρμογή του μοντέλου με συγκεκριμένα δεδομένα (Fine-tuning) όπου το μοντέλο είναι φτιαγμένο έτσι ώστε να μπορεί να προσαρμοστεί εύκολα. Αν κάποια εφαρμογή θέλει να το φέρει στα μέτρα της, μπορεί να το κάνει χωρίς να χρειάζεται τεράστιους πόρους. Μπορεί να χρησιμοποιήσει ολόκληρο το μοντέλο ή μόνο ένα μέρος του και να το προσαρμόσει στα δικά της δεδομένα. Έτσι πετυχαίνει καλό αποτέλεσμα με λιγότερο κόπο.

Τεχνικές Βελτιστοποίησης/Επεκτάσεις

Πρόκειται για τρόπους με τους οποίους μπορεί να εκπαιδευτεί καλύτερα ένα recommendation system, ώστε να δίνει πιο έξυπνες αξιόπιστες και σχετικές προτάσεις στον χρήστη της εκάστοτε εφαρμογή γρηγορότερα και με μεγαλύτερη ακρίβεια.

Deep Learning Approaches: Με την χρήση νευρωνικών δικτύων (π.χ. autoencoders, RNNs) για πιο πολύπλοκες συσχετίσεις. Το σύστημα κοιτάει πώς αλληλεπιδρά ο χρήστης με προϊόντα ή ταινίες (π.χ. τι βαθμολογεί, τι κάνει click) και μαθαίνει:

- είτε από ρητά δεδομένα (π.χ. βαθμολογία 4/5),

- είτε από σιωπηρά δεδομένα (π.χ. είδες κάτι ή όχι).

Προβλήματα και προκλήσεις που αντιμετωπίζει κάθε recommendation system

Χρησιμοποιεί μαθηματικές συναρτήσεις (λέγονται "συναρτήσεις απώλειας") για να βελτιώνεται. Αν το «μαντέψει» λάθος, προσαρμόζει τα βάρη του ώστε την επόμενη φορά να είναι πιο κοντά στην αλήθεια. Με την μέθοδο NCF χρησιμοποιεί νευρωνικά δίκτυα για να "πιάσει" πιο σύνθετες σχέσεις. Μπορεί να εκπαιδευτεί:

- με διαδοχικά παραδείγματα (π.χ. διάλεξε ανάμεσα σε δύο),
- με επιπλέον δεδομένα (π.χ. από άλλο domain – ταινίες, βιβλία),
- με πιο πλούσιες πληροφορίες εισόδου (π.χ. προφίλ χρηστών αντί για απλό ID).

Ενώ με την μέθοδο DeepFM συνδυάζονται δύο τεχνικές: μία που πιάνει απλές σχέσεις (όπως ποιος είδε τι) και μία που πιάνει πολύπλοκες σχέσεις (με MLP – νευρωνικό δίκτυο). Το αποτέλεσμα είναι ένα σύστημα που αυτοματοποιεί την εύρεση των πιο χρήσιμων χαρακτηριστικών, χωρίς να χρειάζεται να τα σχεδιάσει κάποιος χειροκίνητα.

Τέλος με τα χαρακτηριστικά Negative sampling & ranking loss στις περιπτώσεις που δεν έχει αξιολογήσει κάθε προϊόν ο χρήστης, το σύστημα επιλέγει μερικά «τυχαία» που δεν είδε (negative samples) και τα χρησιμοποιεί για να μάθει να ξεχωρίζει τι προτιμά περισσότερο.

Αν το σύστημα δει ότι ο χρήστης προτιμά το αντικείμενο A από το B, το σύστημα μαθαίνει με βάση την κατάταξη (ranking) και όχι μόνο το αν του άρεσε ή όχι.

Explainability in Recommender Systems: Τα μοντέλα μηχανικής μάθησης, και ιδιαίτερα τα βαθιά νευρωνικά, είναι συχνά «μαύρα κουτιά» αφού παίρνουν δεδομένα και δίνουν αποτέλεσμα, χωρίς να μπορεί να καταλάβει ο χρήστης γιατί. Αυτό δημιουργεί έλλειψη εμπιστοσύνης, ειδικά όταν χρησιμοποιούνται για σημαντικές αποφάσεις (π.χ. ιατρική διάγνωση, εγκληματολογία). Οι χρήστες δεν αρκούνται μόνο στο να βλέπουν προτάσεις αλλά θέλουν και να καταλαβαίνουν από πού προέρχονται. Γι' αυτό τον λόγο, στα συστήματα συστάσεων υπάρχει μια σημαντική ανάγκη: η εξηγησιμότητα.

Όταν το σύστημα δεν δίνει στον χρήστη απλά μια πρόταση, αλλά του δείχνει και τον λόγο πίσω από αυτήν, τότε πρόκειται για το φαινόμενο της εξηγησιμότητας στα recommendation systems. Για παράδειγμα:

- «Σου προτείνουμε αυτό το βιβλίο επειδή σου άρεσαν άλλα του ίδιου συγγραφέα».
- «Αυτό το προϊόν είναι παρόμοιο με όσα έχεις αγοράσει πρόσφατα».

Αυτό βοηθάει πολύ στο να εμπιστευτεί τις προτάσεις του συστήματος και να νιώσει ότι είναι προσωποποιημένες και όχι τυχαίες. Μπορεί επίσης να απορρίψει μια πρόταση, αν καταλάβει ότι βασίστηκε σε λάθος δεδομένα (π.χ. μια λέξη που δεν είχε προσέξει). Τέλος οι δημιουργοί του συστήματος μπορούν να εντοπίσουν προβλήματα και να τα βελτιώσουν.

Πολλές φορές, τα συστήματα σύστασης βασίζονται σε περίπλοκους αλγόριθμους που λειτουργούν σαν «μαύρα κουτιά». Δηλαδή, δίνουν απαντήσεις χωρίς να είναι προφανές πώς τις υπολόγισαν.

Μια έξυπνη ιδέα για να λυθεί αυτό είναι να παρατηρήσει κανείς πώς αλλάζει η πρόταση αν αλλάξουν λίγο τα δεδομένα. Δηλαδή:

- Αν αφαιρεθούν κάποιες πληροφορίες από το προφίλ του χρήστη, αλλάξει το αποτέλεσμα;
- Αν αφαιρεθεί μία λέξη από μια κριτική, αλλάζει η πρόβλεψη;

Με αυτόν τον τρόπο μπορούν να δούνε ποια σημεία επηρεάζουν περισσότερο τη σύσταση και να δημιουργήσουν μια απλή, κατανοητή εξήγηση.[27]

Κεφάλαιο 5 °: Το Recommendation system του YouTube

5.1 Το YouTube

Το YouTube είναι μια από τις πιο δημοφιλείς διαδικτυακή πλατφόρμα κοινής χρήσης και προβολής βίντεο στην οποία, ο χρήστης μπορεί είτε να παρακολουθήσει, είτε να ανεβάσει ο ίδιος δικό του βίντεο. Κυκλοφόρησε στις 14 Φεβρουαρίου 2005 από 3 πρώην υπαλλήλους της PayPal, τους Steve Chen, Chad Hurley και Jawed Karim και έμελλε να γίνει μια από τις πιο κοινές εφαρμογές στην καθημερινότητα του 21ου αιώνα.

Τα διαθέσιμα βίντεο της πλατφόρμας ξεπέρασαν τα 25 εκατομμύρια μόλις έναν χρόνο αργότερα, τον Μάρτιο του 2006 με περίπου κάτι παραπάνω από 20.000 βίντεο να ανεβαίνουν καθημερινά. Η εφαρμογή προσφέρει στον χρήστη την δυνατότητα προβολής βίντεο πάρα πολλών κατηγοριών, μέσα στις οποίες περιέχονται φυσικά trailer ταινιών και μουσική.

Επίσης δεν έχει απλά την δυνατότητα να ανεβάσει ένα βίντεο αλλά να δημιουργήσει κανάλια τα οποία θα περιέχουν συγκεκριμένα βίντεο για καλύτερα οργάνωση της θεματολογίας του. Μέσω αυτών των καναλιών μπορεί κάθε χρήστης να συγκεντρώσει συνδρομητές και να έχει κέρδος με το περιεχόμενο το οποίο προβάλλει. Τα βίντεο που προβάλλονται στην εφαρμογή δεν είναι απλά “Βλέπεις-φεύγεις”, έχει κάθε χρήστης την δυνατότητα να αλληλεπιδρά με άλλους χρήστες μέσω κάποιου βίντεο στο οποίο μπορούν να σχολιάσουν, να βαθμολογήσουν, να κοινοποιήσουν ενώ έχουν και την δυνατότητα να δημιουργήσουν της δικής τους wishlist με βιντεο που θέλουν να παρακολουθήσουν αργότερα.

Τέλος το YouTube έγινε ένα πολύ δυνατό μέσω προώθησης προϊόντων καθώς βασίζονται πάνω στην πλατφόρμα ολόκληρες διαφημιστικές εκστρατείες καθώς η δύναμή της στην προβολή στοχευμένου περιεχομένου την καθιστά ιδανική για marketers και influencers. [10].

Το YouTube, η πλατφόρμα στην οποία εκατομμύρια άνθρωποι καθημερινά ξοδεύουν είτε λίγο είτε πολύ από τον χρόνο τους παρακολουθώντας βίντεο, έχει εξελίξει ένα από τα πιο σύνθετα και δυναμικά recommendation systems στον κόσμο.

Ο αλγόριθμος του βασίζεται σε deep learning και ιδιαίτερα νευρωνικά δίκτυα, για να επιλέξει και να προτείνει τα πιο σχετικά βίντεο για κάθε χρήστη, κάθε στιγμή. Το σύστημα αυτό λαμβάνει υπόψη του εκατοντάδες σήματα, όπως το ιστορικό προβολών, η διάρκεια παρακολούθησης, το engagement (likes, σχόλια, κοινοποιήσεις), καθώς και το είδος περιεχομένου με το οποίο αλληλεπιδρά περισσότερο κάθε χρήστης.

5.2 Δομή του YouTube Recommendation System

Το σύστημα συστάσεων που χρησιμοποιεί το YouTube αποτελείται από δύο βασικά στάδια:

Στάδιο 1: Γεννήτρια Υποψήφιων Βίντεο (Candidate Generation)

Στο πρώτο στάδιο, το σύστημα εντοπίζει έναν περιορισμένο αριθμό βίντεο που ενδέχεται να ενδιαφέρουν τον χρήστη. Αυτό επιτυγχάνεται με την χρήση ενός νευρωνικού δικτύου που εκπαιδεύεται να προβλέπει την πιθανότητα ότι ο χρήστης θα αλληλεπιδράσει με ένα συγκεκριμένο βίντεο.

Το μοντέλο αυτό συνήθως αποτελείται από πλήρως συνδεδεμένα επίπεδα με λειτουργία ενεργοποίησης ReLU (Rectified Linear Unit) μια συνάρτηση ενεργοποίησης που χρησιμοποιείται σε νευρωνικά δίκτυα και βαθιά μάθηση, και η έξοδος του είναι μια διανυσματική αναπαράσταση του χρήστη και του βίντεο.

Η πρόβλεψη γίνεται μέσω της συνάρτησης softmax, η οποία και υπολογίζει την πιθανότητα αλληλεπίδρασης του χρήστη με το βίντεο. Έτσι, το σύστημα ταξινομεί τα υποψήφια βίντεο βάσει της πιθανότητας προβολής ή αλληλεπίδρασης.

Για την εκπαίδευση του συγκεκριμένου μοντέλου, χρησιμοποιούνται δεδομένα όπως το ιστορικό παρακολούθησης του κάθε χρήστη, τις αναζητήσεις του, τις αλληλεπιδράσεις του (likes, σχολία, κοινοποιήσεις, subscribes), δημοφιλή βίντεο που είδαν παρόμοιοι χρήστες καθώς και δημογραφικά χαρακτηριστικά του όπως η ηλικία, το φύλο ή η τοποθεσία του. Η αποτελεσματικότητα του μοντέλου αυτού αξιολογείται μέσω μετρικών όπως η ακρίβεια και 'ανακλητικότητα' (precision and recall) [5].

Στάδιο 2: Κατάταξη (Ranking)

Στο δεύτερο στάδιο, το σύστημα αξιολογεί και κατατάσσει τα υποψήφια βίντεο που προέκυψαν από το πρώτο στάδιο, προκειμένου να επιλέξει τα πιο σχετικά και πιθανώς αρεστά για τον χρήστη. Το στάδιο αυτό είναι κρίσιμο, καθώς καθορίζει ποια βίντεο θα εμφανιστούν τελικά στη σελίδα προτάσεων.

Το μοντέλο κατάταξης συνήθως αποτελείται από ένα βαθύ νευρωνικό δίκτυο που λαμβάνει ως είσοδο ένα συνδυασμό κατηγορηματικών και συνεχών χαρακτηριστικών. Οι κατηγορηματικές μεταβλητές ενσωματώνονται μέσω αναπαραστάσεων, ενώ οι συνεχείς μεταβλητές κανονικοποιούνται πριν την είσοδο στο δίκτυο.

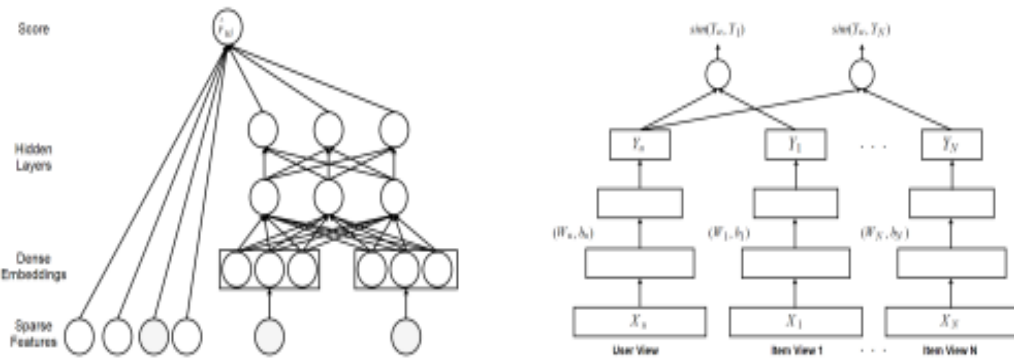
Ο στόχος του μοντέλου είναι να μεγιστοποιήσει τον χρόνο παρακολούθησης του βίντεο από τον χρήστη, χρησιμοποιώντας συνάρτηση απώλειας όπως η λογιστική παλινδρόμηση με σταυροειδή εντροπία (cross-entropy loss), όπου το βάρος των θετικών παραδειγμάτων (βίντεο που παρακολούθηθηκαν) είναι ανάλογο του χρόνου παρακολούθησης.

Η εκπαίδευση του μοντέλου πραγματοποιείται τόσο χωρίς σύνδεση στο διαδίκτυο (βλέπε εικόνα 11.1) με βάση ιστορικά δεδομένα και αξιολογείται σε εσωτερικά σύνολα δοκιμών, όσο και με σύνδεση χρησιμοποιώντας A/B testing, όπου δοκιμάζονται μικρές τροποποιήσεις του μοντέλου σε ζωντανό περιβάλλον και αξιολογείται η απόδοσή τους στην πράξη, για τη συνεχιζόμενη βελτίωση του συστήματος [5].

5.3 Αλγόριθμοι των Recommendation System του YouTube

1. Deep neural networks (DNNs)

Η πλατφόρμα του YouTube χρησιμοποιεί deep neural networks για την επεξεργασία του τεράστιου όγκου δεδομένων χρηστών και περιεχομένων. Τα δίκτυα αυτά εκπαιδεύονται και μαθαίνουν μη γραμμικές σχέσεις μεταξύ της συμπεριφοράς των χρηστών και των χαρακτηριστικών των βίντεο που παρακολουθούν αυτοί, όπως τα μεταδεδομένα, τα ψηφιακά προφίλ χρηστών (user embeddings) και το περιεχόμενο (όπως η ώρα και η συσκευή προβολής) [5]



Εικόνα 5.1: Απεικόνιση ενός: (α) Wide & Deep Learning, (β) Multi-View Deep Neural Network.[26]

2. Ανάλυση Μητρώου (Matrix Factorization)

Η ανάλυση μητρώου είναι μια τεχνική συνεργατικού φιλτραρίσματος που χρησιμοποιείται για την ανακάλυψη “κρυφών” σχέσεων μεταξύ χρηστών και αντικειμένων. Ουσιαστικά το matrix factorization μπορεί να θεωρηθεί ως η εξαγωγή ενσωματωμένων χρηστών και στοιχείων από έναν πίνακα αξιολογήσεων. Το YouTube χρησιμοποιεί αυτήν την τεχνική για την ανάλυση αλληλεπιδράσεων χρηστών με βίντεο, επιτρέποντας την εξατομίκευση των προτάσεων.

$$\begin{bmatrix} 5 & 1 & 4 & 5 & 1 \\ & 5 & 2 & 1 & 4 \\ 1 & 4 & 1 & 1 & 2 \\ 4 & 1 & 5 & 5 & 4 \\ 5 & 3 & 3 & & 4 \\ 1 & 5 & 1 & 1 & 1 \\ 5 & 1 & 5 & 5 & 4 \end{bmatrix} \approx \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1K} \\ u_{21} & u_{22} & \dots & u_{2K} \\ u_{31} & u_{32} & \dots & u_{3K} \\ u_{41} & u_{42} & \dots & u_{4K} \\ u_{51} & u_{52} & \dots & u_{5K} \\ u_{61} & u_{62} & \dots & u_{6K} \\ u_{71} & u_{72} & \dots & u_{7K} \end{bmatrix} \times \begin{bmatrix} v_{11} & v_{21} & v_{31} & v_{41} & v_{61} \\ v_{12} & v_{22} & v_{32} & v_{42} & v_{62} \\ \vdots & & & & \\ v_{1K} & v_{2K} & v_{3K} & v_{4K} & v_{6K} \end{bmatrix}$$

Εικόνα 5.2: Παραγοντοποίηση πίνακα αξιολογήσεων σε χρήστες και αντικείμενα

Ο αλγόριθμος αυτός δουλεύει ως εξής: Έστω λοιπόν ο πρώτος αριστερά πίνακας στον οποίο έχουμε κάποιες αξιολογήσεις. Ο μεσαίος πίνακας περιέχει τους χρήστες ενώ ο δεξιός πίνακας τα προϊόντα. Σκοπός του αλγορίθμου είναι να βρει την καλύτερη αντιστοιχία μεταξύ χρήστη-πελάτη που να ταιριάζει στον αριστερά πίνακα. Στην περίπτωση του YouTube όπου δεν υπάρχουν αξιολογήσεις τύπου ratings (0-10), χρησιμοποιείται η λογική του implicit ratings. Αν ο χρήστης παρακολουθήσει ένα βίντεο, το σύστημα θεωρεί την αξιολογήση θετική και παίρνει την τιμή 1. Στην αντίθετη περίπτωση που δεν παρακολουθήσει ένα βίντεο, η αξιολογήση θεωρείται αρνητική ή ουδέτερη και παίρνει την τιμή 0. Αυτή η μέθοδος βοηθάει στην ομαδοποίηση των δεδομένων και την συσχέτιση έμμεσα των χρηστών μεταξύ τους. Έτσι αν κάτι νέο εμφανιστεί και το δει κάποιος χρήστης, το recommendation system θα το προτείνει και σε άλλους χρήστες χάρη σε αυτή την ομαδοποίηση. Το πρόβλημα με αυτή την μέθοδο είναι ότι δεν μπορεί να αντιμετωπίσει, ή καλύτερα να προτείνει σωστό περιεχόμενο σε έναν νέο χρήστη ο οποίος δεν έχει ιστορικό παρακολούθησης ή αναζήτησης στην πλατφόρμα ούτε έχει συμπεριληφθεί σε κάποια ομαδοποίηση του αλγορίθμου [8].

3. Ενισχυτική Μάθηση (Reinforcement Learning)

Το reinforcement learning είναι ένας αλγόριθμος που εκπαιδεύει το σύστημα να παίρνει αποφάσεις ανάλογα με το system state. Τα recommendation system με reinforcement learning έχουν δύο μορφές. Τα ‘Self-Supervised Reinforcement Learning for Recommender Systems’ και τα ‘Top-K Off-Policy Correction for a REINFORCE Recommender System’.

Στην πρώτη μορφή ο αλγόριθμος έχει δύο επίπεδα.

1. Υπάρχει ένα μοντέλο που εκπαιδεύεται μόνο, χωρίς ρητή επίβλεψη (δεν δέχεται εντολές “αυτό είναι σωστό, αυτό όχι”) και βασίζεται σε δεδομένα που έχει το ίδιο το σύστημα, όπως ποιες ταινίες ή βίντεο πάτησε κάποιος. Το μοντέλο αυτό εκπαιδεύεται με μια τεχνική που λέγεται “απώλεια διασταυρούμενης εντροπίας”, η οποία το βοηθά να μαθαίνει πως να κατατάσσει σωστά τα αντικείμενα π.χ. να προβλέπει ποια βίντεο θα αρέσουν σε κάποιον χρήστη και να του τα εμφανίζει πρώτα. Είναι με απλά λόγια ένα είδος φίλτρου, το οποίο εκπαιδεύεται για να βάζει πρώτα τα βίντεο που πιθανότατα θα πατήσει ο χρήστης.
2. Σε αυτό το επίπεδο μπαίνει ένα δεύτερο “έξυπνο” κομμάτι που βασίζεται στο Reinforcement learning, δηλαδή μάθηση μέσω δοκιμών και λαθών, σαν ένας παίκτης που δοκιμάζει κινήσεις ξανά και ξανά και μαθαίνει ποιες φέρνουν τις νίκες. Αυτό το κομμάτι δεν προσπαθεί απλά να δείξει στον χρήστη οτιδήποτε θα πατήσει εκείνη την στιγμή, αλλά να βελτιώσει και τις μακροπρόθεσμες αντιδράσεις του:
 - Να μείνει περισσότερο χρόνο στην πλατφόρμα (εμπλοκή)
 - Να αγοράσει κάτι
 - Να επιστρέψει ξανά στην εφαρμογή

Δεν επικεντρώνεται δηλαδή στο να κάνεις απλά ένα κλικ αλλά στο να κάνει τον χρήστη να αγαπήσει την πλατφόρμα στο βάθος χρόνου.

Το πρώτο επίπεδο μαθαίνει γρήγορα και προβλέπει τι θα αρέσει στο χρήστη με βάση τα δεδομένα που διαθέτει και το δεύτερο επίπεδο, με ενίσχυση (RL), το βελτιώνει για μακροπρόθεσμα αποτελέσματα, όχι απλώς για το στιγμιαίο ενδιαφέρον.

Για παράδειγμα, ένας χρήστης μπαίνει στην πλατφόρμα του YouTube, στην αρχική σελίδα και του εμφανίζονται 10 βίντεο από τα οποία η εφαρμογή προσπαθεί να διαλέξει ποια να σου δείξει και με ποια σειρά. Αρχικά, στο πρώτο επίπεδο το μοντέλο αυτοεκπαιδεύεται βάση το ιστορικό του χρήστη. Αν π.χ. βλέπει κατα κύριο λόγο βίντεο σχετικά με μαγειρική, θα του προτείνει τέτοιου είδους ενώ αν έχει πατήσει like σε κάποιο τραγούδι, θα του δείξει παρόμοια. Εκπαιδεύεται ναμαντεύει ποιο θα είναι το επόμενο κλικ του χρήστη και χρησιμοποιεί τεχνικές όπως η διασταυρωμένη εντροπία για να μάθει ποια είναι η σωστή σειρά προτίμησης. Αυτή η διαδικασία είναι το γρήγορο φίλτρο όπου το recommendation system του YouTube βρίσκει τα 100 καλύτερα βίντεο για τον χρήστη, βάση του τι κάνει συνήθως.

Εδώ έρχεται το δεύτερο επίπεδο στο οποίο γίνεται χρήση τεχνικών Q-learning και πιο συγκεκριμένα Double Q-learning, μια παραλλαγή που μειώνει το σφάλμα υπερεκτίμησης, όπου το μοντέλο προσπαθεί να καταλάβει ποια από αυτά τα 100 θα κρατήσουν τον χρήστη περισσότερο στην πλατφόρμα. Αν δει ότι ενώ του αρέσουν και οι συνεντεύξεις και η μαγειρική αλλά στην πρώτη περίπτωση μένει στην εφαρμογή περισσότερο χρόνο από όσο στην δεύτερη, θα σου προτείνει περισσότερο βίντεο με συνεντεύξεις.

Επίσης μπορεί να οδηγήσει τον χρήστη σε βίντεο που θα οδηγήσουν σε εγγραφή, σχόλια ή πληρωμένες υπηρεσίες. Το reinforcement learning επίπεδο λειτουργεί και ως κανονικοποιητής, βοηθώντας το σύστημα να αποφεύγει υπερβολικά “εύκολες” προτάσεις και να ενισχύει εκείνες που οδηγούν σε πιο ποιοτική και διαρκή εμπλοκή του χρήστη. Αυτό το δεύτερο πιο έξυπνο φίλτρο προσπαθεί να κάνει τον χρήστη να επιστρέψει ξανά ή να μείνει περισσότερο από την προηγούμενη φορά.

Για αυτή την διαδικασία υπάρχουν δύο προτεινόμενα frameworks, το SQN (Self-Supervised Q Learning) και το SAC (Self-Supervised Actor-critic). Το πρώτο framework συνδυάζει supervised και Reinforcement Learning μέσω τεχνικών Q-Learning. Χρησιμοποιεί replay buffer, ένα είδος ‘μνήμης’ με προηγούμενες εμπειρίες/δράσεις του χρήστη. Ακολουθεί την μέθοδο Co-training, δηλαδή τα δυο layers εκπαιδεύονται ταυτόχρονα και αλληλοενισχύονται και μέσω του Q-learning προβλέπει ποιο βίντεο θα έχει την μεγαλύτερη ανταμοιβή.

Το δεύτερο framework χρησιμοποιεί ένα μοντέλο που ονομάζεται Actor-Critic. Την θέση του Actor παίρνει το self-supervised μοντέλο που προτείνει το τελικό ranking. Την θέση του critic παίρνει το reinforcement learning μοντέλο που αξιολογεί τη χρησιμότητα της πρότασης σε βάθος χρόνου. Αυτό επιτρέπει fine-tuning των προτάσεων με βάση το Long-term engagement.

Στην δεύτερη μορφή το “Top-K Off-Policy Correction for a Reinforcement Recommender System” είναι μια μεθοδολογία, η οποία προτάθηκε για να βελτιώσει τα recommendation systems μεγάλης κλίμακας όπως αυτό του YouTube, χρησιμοποιώντας reinforcement learning. Η κύρια πρόκληση που αντιμετωπίζει είναι η εκπαίδευση ενός μοντέλου που μπορεί να συστήνει ταυτόχρονα πολλαπλά αντικείμενα (βίντεο στην περίπτωση του YouTube) σε χρήστες και να διορθώνει τις προκαταλήψεις που εισάγονται από την εκπαίδευση με δεδομένα που συλλέγονται από προηγούμενες πολιτικές συστήματος.

Η Off-policy correction είναι μια τεχνική η οποία έχει σκοπό να διορθώσει προκαταλήψεις κατά κάποιον τρόπο που εισάγονται από την εκπαίδευση με δεδομένα του χρήστη τα οποία συλλέγονται από προηγούμενες πολιτικές. Αυτό γίνεται υπολογίζοντας τον λόγο των πιθανοτήτων δράσης μεταξύ της τρέχουσας πολιτικής και της πολιτικής που χρησιμοποιήθηκε για την συλλογή των δεδομένων.

Η Top-K Off-policy correction είναι μια εφαρμογή της προηγούμενη off-policy correction σε σύνολα K αντικειμένων (π.χ. τα καλύτερα 10 βίντεο). Έτσι το σύστημα μπορεί να προτείνει πολλαπλά αντικείμενα ταυτόχρονα, λαμβάνοντας υπόψη τις αλληλεπιδράσεις μεταξύ τους και διορθώνοντας τις προκαταλήψεις που εισάγονται από την προηγούμενη πολιτική. [9].

Κεφάλαιο 6 °: Netflix

6.1 Το αφεντικό του streaming

Το Netflix, ένας από τους πρωταγωνιστές μαζί με την Amazon και Disney+ και ίσως ο πιο πετυχημένος στον τομέα της συνδρομητικής διαδικτυακής ψυχαγωγίας, μετράει από το 2023 κάτι παραπάνω από 270 εκατομμύρια συνδρομητές παγκοσμίως και έσοδα αυτές τις συνδρομές που αγγίζουν μηνιαίως το 1 δισεκατομμύριο δολάρια (κάτι παραπάνω από 950 εκατομμύρια).

Μόνο τυχαίο δεν μπορεί να χαρακτηριστεί αυτή η επιτυχία αφού η εταιρία έχει την δική της παραγωγή περιεχομένου και χρησιμοποιεί πολύ αποτελεσματικά recommendation systems που «διαβάζει» τις προτιμήσεις του κάθε χρήστη και του προσφέρει επιλογές κατάλληλες ώστε να συνεχίσει να είναι συνδρομητής και να μην απομακρυνθεί από την πλατφόρμα. Αυτή η επιτυχία, αυτή αύξηση συνδρομητών είχε ως αναμενόμενο αποτέλεσμα την μείωση των συνδρομητών της καλωδιακής τηλεόρασης η οποία έχασε την δυναμική με την άνοδο των streaming υπηρεσιών.

Με απλά λόγια, το Netflix είναι μια πλατφόρμα που λειτουργεί ως μια μεγάλη αποθήκη ή βιβλιοθήκη από ταινίες, σειρές, ντοκιμαντέρ η κινούμενα σχέδια. Υλικό το οποίο είτε κατά κάποιο τρόπο νοικιάζει (πληρώνει άδεια προβολής για κάθε ταινία κλπ.) είτε είναι δική του παραγωγή.

Ο χρήστης με μια μηνιαία συνδρομή μπορεί να έχει πρόσβαση σε αυτό τον ατελείωτο αριθμό περιεχομένου μέσω του smartphone, του tablet, του υπολογιστή του ή ακόμη και μέσω μιας smartTV, ενώ σύμφωνα πάντα και με τις προτιμήσεις του ή τις οικονομικές του δυνατότητες μπορεί να διαλέξει κάποιο από τα τρία πακέτα που προσφέρει η εταιρία. [2].

6.2 Ιστορική αναδρομή – μετάβαση από το CD στο streaming

Ο Marc Randolph, ο ένας από τους δύο ιδρυτές της πλατφόρμας Netflix, υποστηρίζει ότι για να καταλάβει κανείς αν μια μια τρελή ιδέα αξίζει, πρέπει απλά να τολμήσει να την δοκιμάσει. Αυτό ακριβώς λοιπόν έγινε και στην περίπτωση του Netflix, όταν έστειλαν ως δοκιμή ένα CD με το ταχυδρομείο για να δουν αν τελικά φτάσει, κάτι που έγινε την επόμενη μέρα. Με αυτό το επιτυχημένο τεστ, η ιδέα των ιδρυτών για μια συνδρομητική υπηρεσία η οποία θα στέλνει DVD στο σπίτι με το ταχυδρομείο πήρε μορφή.

Από το 1998, αυτή ήταν η βασική ιδέα-υπηρεσία της πασίγνωστης σήμερα πλατφόρμας για 16 χρόνια ώσπου και σταμάτησε, για να φτάσει 27 χρόνια μετά να είναι ένας παγκόσμιος κολοσσός στο streaming.

Την εποχή που το internet είχε γίνει λέξη καθημερινότητας αφού άρχισε σιγά σιγά να κάνει αντιληπτή την εμφάνιση του στην ζωή των ανθρώπων (μέσα προς τέλη 90s), ο Randolph ήθελε να στήσει την δική του διαδικτυακή επιχείρηση χωρίς όμως να γνωρίζει το πως. Με τον συνεργάτη του και συνιδρυτή της υπηρεσίας, Reed Hastings, μοιράζονταν γραφείο και αυτοκίνητο, προσπαθούσαν καθημερινά να βρούνε αυτή την ιδέα, αυτό το μεγάλο που έλειπε από τον κόσμο εκείνη την εποχή, αυτό που θα τους έβαζε στον δρόμο της δόξας, της επιτυχίας και του χρήματος (eshop για σαμπουάν, τροφή για σκύλους, βιταμίνες), ιδέες που τότε φάνταζαν δύσκολες, καινοτόμες.

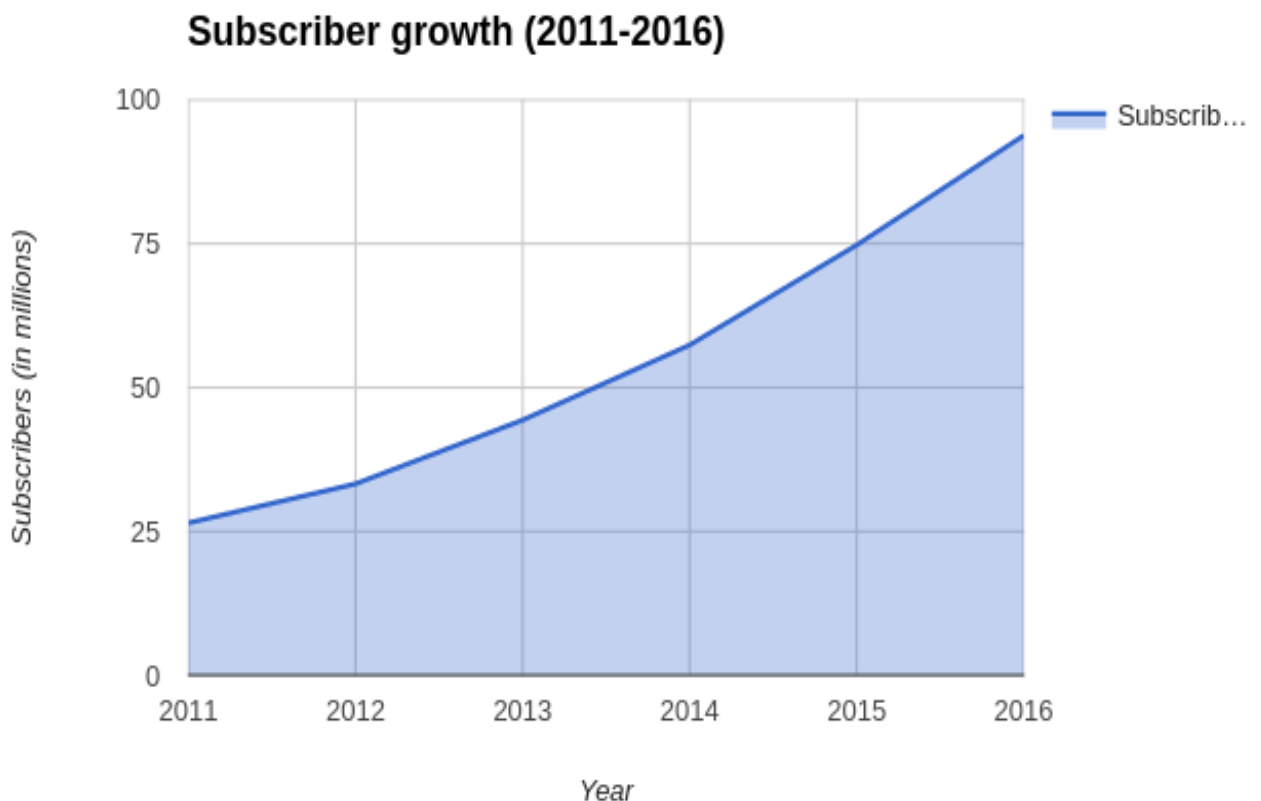
Η ιδέα να στέλνουν βιντεοκασέτες με ταχυδρομείο για να ανταγωνιστούν το Blockbuster δεν έγινε ποτέ πράξη καθώς οι κασέτες ήταν ακριβές στην μεταφορά λόγω μεγέθους. Με την εμφάνιση των DVD, αλλάζουν τα δεδομένα, καθώς σε αντίθεση με την κασέτα, είναι λεπτά, μικρά, ελαφριά και ιδανικά για αποστολή αλλά δυσεύρετα στην αγορά. Randolph και Hastings παίρνουν το επιχειρηματικό αυτό ρίσκο θεωρώντας ότι αυτά τα CD είναι το μέλλον.

“Μπορώ να υποδείξω εκατοντάδες σημεία στον δρόμο που μπορεί να έκαναν τη διαφορά μεταξύ επιτυχίας και αποτυχίας-καθένα από αυτά μου θυμίζει πόσο σημαντική είναι η τύχη σε κάθε ιστορία επιτυχίας” είπε ο Randolph θέλοντας να δείξει πως και η τύχη έπαιξε τον μεγάλο σημαντικό της ρόλο στην επιτυχία του Netflix. Το 1998 το Netflix βγαίνει για πρώτη φορά online, και την πρώτη κιόλας μέρα πέφτουν οι serves λόγω μικρού μεγέθους ετικετών αποστολής, αλλά καταγράφει 200 πελάτες.

Εκείνη την χρονιά, υπάρχει η πρόταση από τον Jeff Bezos (Amazon) να αγοράσει αυτούς και την υπηρεσία τους για 15 εκατομμύρια δολάρια , μια πολύ καλή προσφορά για τα δεδομένα αλλά οι δύο συνεργάτες θα αρνηθούν ελπίζοντας πως έχουν την ευκαιρία για κάτι μεγάλο. Απόφαση σωστή αφού 27 χρόνια αργότερα το Netflix κοστίζει πάνω από 140 εκατομμύρια δολάρια.

Ο Randolph έχοντας δουλέψει πριν σε εταιρείες με ταχυδρομικές πωλήσεις, γνώριζε από πρώτο χέρι πόσο σημαντικό είναι η παράδοση να γίνεται όσο το δυνατόν γρηγορότερα, οπότε και εφάρμοσε αυτή την φιλοσοφία στο Netflix. Η επόμενη ιδέα που μπήκε από νωρίς στην υπηρεσία τους ήταν η χρήση συστάσεων (recommendation system), προτείνοντας στους χρήστες βάση του τι έβλεπαν ή έψαχναν. Αυτό βοήθησε αρκετά στο να κρατηθούν οι πελάτες ευχαριστημένοι και να μην ξεμένουν από τα πιο δημοφιλεί DVD.

Μετά το 2000 όταν δηλαδή και το ίντερνετ έκανε το μεγάλο μπαμ, το Netflix βρέθηκε σε πολύ δύσκολη θέση, σε σημείο να φτάσουν οι δύο ιδρυτές να ζητήσουν 50 εκατομμύρια για την πώληση του 49% της εταιρίας στο Blockbuster και το Netflix να μετατραπεί σε Blockbuster.com. Η απάντηση ήταν αρνητική αλλά και λανθασμένη όπως κρίνει η ιστορία με το Netflix να ξεπερνάει μερικά χρόνια αργότερα (2005) τα 4 εκατομμύρια συνδρομητές ενώ το blockbuster να συρρικνώνεται. [14].



Εικόνα 6.1: Συνδρομητές του Netflix 2011-2016[28]

Το διαδίκτυο εξαπλωνόταν, γινόταν γρηγορότερο και τα DVD φαινόταν πως δεν θα κρατήσουν για πολύ και έτσι το Netflix ξεκίνησε να επενδύει στο streaming. Το 2007 έκανε την εμφάνιση του στον κόσμο του streaming(βλέπε εικόνα 6.2) με μια αρχική βιβλιοθήκη 1000 ταινιών. [29]



Εικόνα 6.2: Η πλατφόρμα Netflix το 2008[29]

Η πραγματική εκτόξευση όμως ήρθε 6 χρόνια μετά. Βλέποντας όμως ότι οι άδειες που υποχρεούνταν για την προβολή σειρών και ταινιών ήταν μια δύσκολη διαδικασία, η εταιρεία κάνει ακόμη ένα τολμηρό βήμα για παραγωγή δικού της περιεχομένου, αρχής γενομένου το 2013 όταν και προβάλλεται η πρώτη δική τους παραγωγή, το House of cards [14].

6.3 Netflix Recommendation System

6.3.1 Εξέλιξη της Αρχιτεκτονικής του Netflix Recommendation System

Η εταιρία χρησιμοποιεί ένα αρκετά πολύπλοκο recommendation system για να προτείνει στους χρήστες της πλατφόρμας τι άλλο θα μπορούσαν να παρακολουθήσουν, το οποίο βασίζεται σε κάποια μοντέλα τεχνητής νοημοσύνης. Κάθε μοντέλο δουλεύει αυτόνομα, έχει δηλαδή τον δικό του ξεχωριστό ρόλο, δηλαδή έστω ότι ένα μοντέλο θα ασχοληθεί με το να υπενθυμίσει στον χρήστη ότι δεν έχει τελειώσει με την ταινία, εμφανίζοντας ειδοποίηση του τύπου “Συνέχισε την παρακολούθηση”. Ένα άλλο μοντέλο θα είναι υπεύθυνο να προτείνει στον χρήστη περιεχόμενο που θα ήταν ενδιαφέρον να το παρακολουθήσει την συγκεκριμένη ημέρα.

Η εταιρία συνέχισε να προσθέτει νέα μοντέλα για να καλύπτει κάθε πιθανή ανάγκη με σκοπό να “δέσει” ακόμη περισσότερο τον χρήστη στην πλατφόρμα με αποτέλεσμα το recommendation system να γίνεται ολοένα και πιο ακριβό ή δύσκολο στην διαχείριση. Το πρόβλημα που αντιμετώπισε το Netflix ήταν πως τα μοντέλα αυτά παρόλο που είχαν να επεξεργαστούν τα ίδια δεδομένα, έπρεπε το καθένα να εκπαιδευτεί ξεχωριστά κάνοντας έτσι την διαδικασία μεταφοράς ιδεών και βελτιώσεων από τον ένα μοντέλο στο άλλο

δυσκολότερη.

Αυτός ήταν και ο λόγος που οδήγησε την εταιρεία στο να ακολουθήσει μια νέα τακτική, με ένα σύστημα το οποίο θα μπορούσε να συγκεντρώσει όλες τις προτιμήσεις των χρηστών σε ένα σημείο κοινό, και τα μοντέλα του recommendation system θα μπορούσαν να συνεργάζονται μεταξύ του καλύτερα προσφέροντας πιο εύστοχες και εύχρηστες προτάσεις. Αυτά τα μοντέλα βασίζονται κυρίως στα πιο πρόσφατα δεδομένα του χρήστη, τις τελευταίες αλληλεπίδρασης του με την πλατφόρμα, ώστε και να προτείνουν κάτι σχετικό σύμφωνα πάντα με τα γούστα του το συγκεκριμένο χρονικό διάστημα.

Βέβαια παρουσιάζουν και κάποιους περιορισμούς εξαιτίας του χρόνου που χρειάζεται για να απαντήσουν-προτείνουν αλλά και εξαιτίας του κόστους που απαιτεί κάθε μοντέλο για την εκπαίδευση του. Έτσι η εταιρεία οδηγήθηκε στην ανάπτυξη ενός νέου recommendation μοντέλου το οποίο θα αναλύει το συνολικό περιεχόμενο της πλατφόρμας σε μεγάλη κλίμακα ταυτόχρονα με την ανάλυση των πρόσφατων αλληλεπιδράσεων του χρήστη με αποτέλεσμα να γίνεται ευκολότερη η μεταφορά δεδομένων από μοντέλο σε μοντέλο είτε απευθείας μέσω της μαθηματικής τεχνικής embedding είτε με επιπλέον εκπαίδευση μέσω κοινών στοιχείων.

Η νέα τάση να χρησιμοποιούνται μεγάλα και γενικά μοντέλα αντί για πολλά μικρά και εξειδικευμένα καθώς και οι αλλαγές στον τομέα της τεχνητής νοημοσύνης και στην επεξεργασία φυσικής γλώσσας σαν αυτή που χρησιμοποιούν τα chatbots όπως αυτά της vodafone, cosmote, οδήγησαν την εταιρία στο να αναπτύξει αυτό το νέο μοντέλο προτάσεων.

Η αλλαγή αυτή είχε ως πρώτο αποτέλεσμα την μεταπήδηση από τα πολλά αυτόνομα μοντέλα με καθένα να έχει τα δικά του ειδικά χαρακτηριστικά σε ένα pool με πολλά ποιοτικά δεδομένα από τα οποία το μοντέλο εκπαιδεύεται μόνο του από την αρχή. Αυτό το σύστημα μπορεί να εξελίσσεται με τον χρόνο να προσαρμόζεται πιο εύκολα στις ανάγκες που αλλάζουν και φυσικά είναι πιο οικονομικό και αποδοτικό μακροπρόθεσμα.

6.3.2 Data και Tokenizing User Interactions

Στο Netflix, οι τρόποι με τους οποίους ένας χρήστης έχει την δυνατότητα αλληλεπίδρασης με την πλατφόρμα ποικίλουν. Κάποιοι χρήστες απλώς περνούν την ώρα τους, βλέποντας τίτλους ταινιών, trailers, περιεχόμενα και περιγραφές, άλλοι βλέπουν ταινίες ενώ υπάρχουν και οι πιο φανατικοί, οι οποία θα παρακολουθήσουν ολόκληρες σειρές.

Με περισσότερους από 300 εκατομμύρια χρήστες μέχρι και το 2024, μπορεί κανείς να φανταστεί για τι αριθμό αλληλεπιδράσεων πρόκειται. Κάτι δισεκατομμύρια ενέργειας καθημερινά γίνονται στην πλατφόρμα. Το γεγονός ότι είναι τόσο μεγάλος ο όγκος των δεδομένων της πλατφόρμας δεν είναι κάτι ανησυχητικό καθώς η ποιότητα υπερσχύει της ποσότητας, ως προς την κατανόηση των προτιμήσεων των χρηστών.

Για να αξιοποιηθεί σωστά αυτός ο τεράστιος όγκος, το Netflix ανέπτυξε μια tokenizing μεθοδο, μια μέθοδο δηλαδή που μετατρέπει σε tokens κάθε ενέργεια των χρηστών. Κάθε κλικ ή κάθε ενέργεια δεν έχει απαραίτητα αξία για την κατανόηση του χρήστη οπότε και γίνεται με την μέθοδο αυτή μια συμπίκνωση των ενεργειών με τελικό σκοπό να κρατηθούν μόνο οι πιο σημαντικές για να μελετηθεί καλύτερα ο χρήστης.

Αν δηλαδή ένας χρήστης x κάνει κλικ να δει μια ταινία. μετά κλικ για να σταματήσει, ξανά κλικ για επανεκκίνηση και μετά ολοκλήρωση της ταινίας όλες αυτές οι ενέργειες δεν έχουν σημασία για την διαμόρφωση του “τύπου” του χρήστη ώστε και να μπορέσει μετά το σύστημα να του προτείνει κάτι κοντά στα θέλω του.

Ωστόσο, εδώ υπάρχει ένα δίλημμα για τον τρόπο που θα το διαχειριστεί το σύστημα: αν κρατήσει πάρα πολλές λεπτομέρειες, η αλληλουχία γίνεται τεράστια και δύσκολη στην επεξεργασία. Αν όμως αφαιρέσει πολλά, μπορεί να χαθούν σημαντικά σημάδια των προτιμήσεων του χρήστη. Άρα πρέπει να βρεθεί μια ισορροπία ανάμεσα στο «πόσα δεδομένα κρατάμε» και στο «πόσο περιεκτικό είναι κάθε token».

Ακόμη και με έξυπνη συμπίκνωση, οι ενεργοί χρήστες μπορεί να έχουν ιστορικά με χιλιάδες ενέργειες. Τα πιο εξελιγμένα μοντέλα τεχνητής νοημοσύνης έχουν περιορισμούς στο πόσα δεδομένα μπορούν να επεξεργαστούν ταυτόχρονα, ειδικά όταν οι υπολογισμοί πρέπει να γίνουν μέσα σε λίγα χιλιοστά του δευτερολέπτου (π.χ. όταν πατάει ο χρήστης “προτείνει μου κάτι”).

Για να αντιμετωπιστεί αυτό, εφαρμόζονται δύο τεχνικές:

1. Αραιή Προσοχή (Sparse Attention): Επιτρέπει στο μοντέλο να διαβάζει μεγαλύτερα ιστορικά δεδομένα χωρίς να "βαραίνει" υπερβολικά, κάνοντας τους υπολογισμούς πιο αποδοτικούς.
2. Δειγματοληψία με Κινούμενο Παράθυρο (Sliding Window Sampling): Κατά την εκπαίδευση του μοντέλου, δεν χρησιμοποιείται όλο το ιστορικό ταυτόχρονα. Αντίθετα, χωρίζεται σε μικρότερα κομμάτια, τα οποία παρουσιάζονται στο μοντέλο σε διαφορετικές στιγμές. Έτσι, το μοντέλο μαθαίνει από όλο το ιστορικό, σταδιακά.

Όταν έρχεται η ώρα να γίνει πρόβλεψη (δηλαδή να προταθεί μια ταινία στον χρήστη), χρησιμοποιούνται έξυπνες τεχνικές για να μην υπολογίζονται ξανά και ξανά τα ίδια πράγματα. Έτσι εξασφαλίζεται ταχύτητα και ακρίβεια.[40]

6.3.3 Στόχος-Αρχιτεκτονική του Μοντέλου

Κανονικά, μοντέλα σαν το GPT εκπαιδεύονται προβλέποντας λέξη-λέξη τι έρχεται μετά, θεωρώντας κάθε λέξη ξεχωριστά το ίδιο σημαντική. Όμως όταν η πλατφόρμα προτείνει περιεχόμενο στους χρήστες της, όσα δουν εν τέλη, δεν είναι όλα το ίδιο σημαντικά. Για παράδειγμα, το ότι κάποιος είδε ένα trailer για 5 λεπτά δεν έχει την ίδια βαρύτητα με το να παρακολούθησε ολόκληρη ταινία.

Αυτό είναι σημαντικό καθώς δείχνει πως παίζει μεγάλη βαρύτητα η ποιότητα των προβολών και όχι μόνο ο αριθμός. Γι' αυτό, αντί να εκπαιδεύει το μοντέλο να προβλέπει μόνο το αμέσως επόμενο περιεχόμενο, του ζητάει να προβλέψει παραπάνω επόμενα βήματα κάθε φορά. Έτσι μαθαίνει καλύτερα ποια πράγματα ακολουθούν μετά από ένα σύνολο επιλογών και δεν μένει σε ασήμαντα πράγματα. Με αυτή την προσέγγιση, το σύστημα κάνει τις προβλέψεις πιο ρεαλιστικές και λιγότερο πρόχειρες.

Επιπλέον, το σύστημα δεν εστιάζει μόνο στο τι είδε ο χρήστης (π.χ. ID της ταινίας), αλλά κοιτάει και άλλα στοιχεία — όπως σε ποιο είδος ανήκει αυτό που είδε ή ποια γλώσσα είχε, ποιοι ηθοποιοί πρωταγωνιστούν ή ποιας εταιρείας παραγωγή είναι. Αυτές οι επιπλέον πληροφορίες βοηθούν το μοντέλο να μάθει καλύτερα τις γενικές προτιμήσεις του κάθε χρήστη, αντί να μαντεύει μόνο από μεμονωμένες επιλογές. Είναι σαν να ρωτάει πρώτα "σου αρέσουν οι κωμωδίες;" και μετά να προτείνει μια συγκεκριμένη κωμωδία που ταιριάζει στα γούστα του.

Όλα αυτά τα γίνονται ώστε το μοντέλο να μάθει να προτείνει πιο έξυπνα, όχι απλώς τι πιθανόν θα τραβήξει την προσοχή του χρήστη μια στιγμή απλώς, αλλά τι θα τον ικανοποιήσει πραγματικά και στο μέλλον.[41]

Κεφάλαιο 7^ο: Σύγκριση Netflix-Youtube Recommendation Systems

Και οι δύο πλατφόρμες έχουν απλό και ευχάριστο περιβάλλον, εύκολο στη χρήση, ακόμα και για άτομα που δεν είναι εξοικειωμένα με την τεχνολογία. Στο YouTube μπορεί κανείς να ψάξει βίντεο ανά κατηγορία, όπως μουσική ή gaming, ενώ στο Netflix η αναζήτηση γίνεται κυρίως βάσει είδους – για παράδειγμα, ταινίες δράσης ή σειρές μυστηρίου. Όσον αφορά το σύστημα προτάσεων και τα δύο συστήματα προσπαθούν να δείξουν στον χρήστη περιεχόμενο που τον ενδιαφέρει. Χρησιμοποιούν δεδομένα από τη συμπεριφορά του, όπως τι βλέπει πιο συχνά, πόσο χρόνο περνάει σε κάθε βίντεο, τι κάνει συνήθως like κ.λπ. Όμως υπάρχει μια βασική διαφορά:

- Το **YouTube** έχει ως βασικό στόχο να κρατήσει τον κάθε χρήστη όσο το δυνατόν περισσότερο στην πλατφόρμα. Θέλει οι χρήστες του να βλέπουν συνεχόμενα βίντεο, γι' αυτό και πολλές φορές τους προτείνει πολύ «ελκυστικά» ή δημοφιλή βίντεο, ακόμα κι αν δεν είναι 100% στις προτιμήσεις τους.
- Το **Netflix**, από την άλλη, προσπαθεί να δείξει ταινίες ή σειρές που ταιριάζουν στις συνήθειές των συνδρομητών του και να τους κρατήσει ικανοποιημένους, ώστε να συνεχίσουν να πληρώνουν τη συνδρομή. Το ενδιαφέρον εδώ είναι ότι η πλατφόρμα δεν επικεντρώνεται τόσο στο τι βλέπουν οι πολλοί, αλλά στο τι μπορεί να στον καθένα ξεχωριστά αρέσει με βάση το ιστορικό του.

Σχετικά με την ποικιλία, και οι δύο έχουν πολύ μεγάλο όγκο περιεχομένου. Όμως το Netflix φαίνεται να υπερτερεί σε αυτόν τον τομέα, γιατί έχει παραγωγές από πολλές χώρες και σε διάφορες γλώσσες. Το YouTube έχει απίστευτο αριθμό καναλιών, αλλά πολλές φορές επικεντρώνεται σε συγκεκριμένα είδη όπως μουσική, vlogs, τεχνολογία ή tutorials.

Και τα δύο χρησιμοποιούν αλγόριθμους που συνεχώς βελτιώνονται. Όμως λειτουργούν με διαφορετικά κριτήρια.

- Το YouTube συχνά δείχνει βίντεο που έχουν πολλές προβολές ή είναι "πιασάρικα", ακόμη κι αν δεν έχουν τόσο ουσιαστικό περιεχόμενο.
- Το Netflix, αντίθετα, λαμβάνει υπόψη του τόσο τις προτιμήσεις των συνδρομητών του, όσο και την ποιότητα των τίτλων, δηλαδή πώς έχουν βαθμολογηθεί από κοινό και ειδικούς.

Συμπερασματικά, και τα δύο συστήματα κάνουν καλά τη δουλειά τους, αλλά με διαφορετική φιλοσοφία. Το YouTube έχει στόχο να "κρατήσει μέσα" στην πλατφόρμα τους χρήστες, να τους κάνει να βλέπουν το ένα βίντεο μετά το άλλο. Το Netflix, από την άλλη, θέλει να τους προτείνει κάτι που θα απολαύσουν, ώστε να παραμείνουν συνδρομητές του. Φυσικά και τα δυο έχουν κοινό και τελικό σκοπό το μεγαλύτερο κέρδος.

Και τα δύο εξελίσσονται συνεχώς και γίνονται πιο έξυπνα, με αποτέλεσμα να καταλαβαίνουν όλο και καλύτερα. Στο μέλλον, οι χρήστες θα βλέπουν προτάσεις ακόμα πιο κοντά στα γούστα τους, σχεδόν σαν να τους ξέρουν προσωπικά [17].

Κεφάλαιο 8ο: Ηθικά Ζητήματα και Ιδιωτικότητα στα Συστήματα Συστάσεων

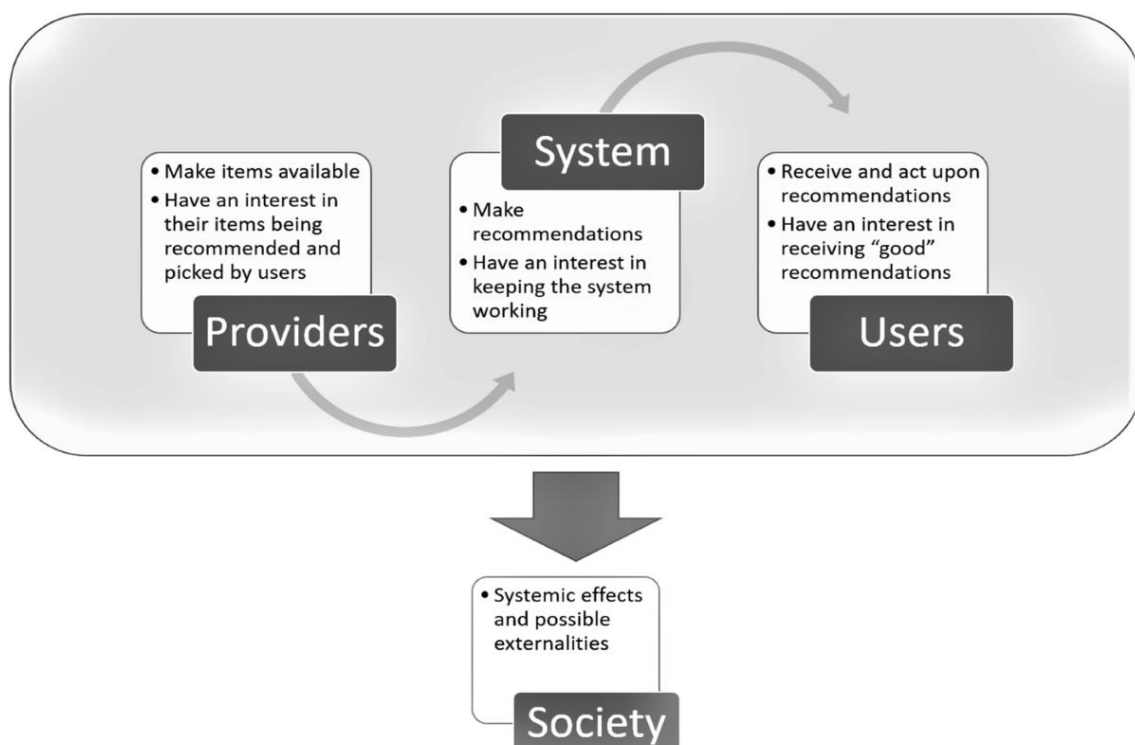
Στην εποχή της ψηφιακής πληροφορίας, τα συστήματα συστάσεων έχουν γίνει βασικό εργαλείο σε πολλές πλατφόρμες που χρησιμοποιεί ο άνθρωπος καθημερινά, από το Netflix και το YouTube, μέχρι τα ηλεκτρονικά καταστήματα και τα μέσα κοινωνικής δικτύωσης. Αυτά τα συστήματα λειτουργούν συλλέγοντας και αναλύοντας δεδομένα από τη συμπεριφορά των χρηστών, ώστε να προβλέψουν τι θα τους ενδιαφέρει περισσότερο. Ωστόσο, όσο πιο «έξυπνα» γίνονται, τόσο μεγαλώνουν και οι ανησυχίες γύρω από την προστασία των προσωπικών δεδομένων, την ιδιωτικότητα και τα γενικότερα ηθικά ζητήματα που εμπλέκονται.

Η ανάγκη για προσωποποιημένες προτάσεις πολλές φορές έρχεται σε σύγκρουση με το δικαίωμα του χρήστη να έχει έλεγχο πάνω στις πληροφορίες που αποκαλύπτει. Επιπλέον, τίθενται ζητήματα όπως η διαφάνεια στους αλγόριθμους, η πιθανή χειραγώγηση των επιλογών του χρήστη, και η δημιουργία «φούσκας φίλτρων» (filter bubbles), που περιορίζει την έκθεση του ατόμου σε διαφορετικές απόψεις και περιεχόμενο.

Οι βασικές ηθικές προκλήσεις που δημιουργούνται από τη χρήση των συστημάτων σύστασης, με έμφαση στην προστασία της ιδιωτικότητας και της ταυτότητας του χρήστη είναι οι εξής:

Προσωπικά δεδομένα και ιδιωτικότητα:

- Οι μηχανές σύστασης συλλέγουν κι αναλύουν τεράστιες ποσότητες δεδομένων (τι βλέπει, τι αγοράζει, πόσο χρόνο αφιερώνει ο χρήστης) για να προβλέψουν τι θα σου αρέσει.
- Αυτό δημιουργεί κίνδυνο παραβίασης της ιδιωτικότητας, καθώς παρεισφρύνουν στη ζωή του χρήστη, πολλές φορές χωρίς να καταλαβαίνει το πώς ή γιατί.[30]



Εικόνα 8.1: Δομή και επιρροές ενός συστήματος συστάσεων [32]

Διαβάθμιση ταυτότητας και αναγνώριση:

- Τα δεδομένα καταγράφουν την «ταυτότητα» του χρήστη – τα ενδιαφέροντα, τις πεποιθήσεις, ακόμα και την πολιτική του τοποθέτηση.
- Αυτό οδηγεί σε μεροληψίες (bias): το σύστημα σε «βρίσκει» απο τη μία, συχνά χρωματίζει τα ενδιαφέροντά του χρήστη και μπορεί να φτάσει σε επίπεδο χειραγώγησης.[30][31]

Αυτονομία και χειραγώγηση:

- Τα συστήματα συστάσεων μπορούν να ωθήσουν τον χρήστη σε συγκεκριμένες επιλογές – π.χ. τι να δει ή να αγοράσει – χωρίς να σκεφτεί ότι «επιηρεάζεται».
- Υπάρχει το φαινόμενο του “nudging” (ψηφιακή ώθηση) όπου δε φαίνεται παράνομο ή επιθετικό, αλλά μπορεί να είναι λεπτό και διασκορπιστικό.[31]

Έλλειψη διαφάνειας (black box):

- Οι αλγόριθμοι είναι συνήθως «μαύρα κουτιά»: δεν ξέρει ο χρήστης τι δεδομένα χρησιμοποιούνται, πώς παίρνονται οι αποφάσεις και πώς μπορείς να αντιδράσει ή να αλλάξει τις ρυθμίσεις.[30][31]

Οι τρόποι που συνήθως προτείνονται για την προστασία των χρηστών, της ιδιωτικότητάς τους και γενικά για την αντιμετώπιση αυτών των φαινομένων όσο πιο αποτελεσματικά γίνεται είναι οι εξής:

- Περισσότερη διαφάνεια: Να είναι ξεκάθαρα εμφανές ποια δεδομένα συλλέγονται και γιατί – όχι καθόλου “κρυφές” πρακτικές ούτε «ψιλά γράμματα».
- Ελέγχει ο χρήστης: Να μπορεί ο χρήστης να επιλέγει τι αποκαλύπτει, να διαγράφει/μπλοκάρει δεδομένα από το σύστημα.
- Συμμετοχικές σχεδιάσεις: Να εμπλέκονται χρήστες, ρυθμιστικές αρχές και η κοινωνία συνολικά στον σχεδιασμό και την κατεύθυνση αυτών των συστημάτων – όχι μόνο οι εταιρείες.[30][31]
- Ανάπτυξη κανόνων και νομοθεσίας: Π.χ. GDPR στην ΕΕ, για να εξασφαλιστεί πως η συλλογή και χρήση δεδομένων γίνεται με σεβασμό και υπευθυνότητα. [30]

Υπάρχει μια ξεκάθαρη ιδέα, με απλό τρόπο, τα συστήματα συνιστούν περιεχόμενο βασισμένα στα προσωπικά σου δεδομένα. Αν δεν υπάρχει διαφάνεια, έλεγχος από τον ίδιο τον χρήστη και συμμετοχή της κοινωνίας, τότε κινδυνεύουν όλοι χρήστες:

- να φιλτράρεται η πληροφόρηση (filter bubbles),
- να χειραγωγείται η συμπεριφορά των χρηστών,
- να υπάρχει διαρροή στην ιδιωτικότητά του χρήστη
- να οδηγηθεί ο χρήστης σε ένα περιβάλλον που δεν περιέχει επιλογή ή ποικιλία.

Με απλά λόγια το συμπέρασμα είναι ότι: Ο χρήστης πρέπει να ξέρει τι γίνεται με τα δεδομένα του, να έχει επιλογή και έλεγχο, και τα συστήματα πρέπει να σχεδιάζονται με τέτοιο τρόπο ώστε να προστατεύουν την ταυτότητα και την ιδιωτικότητά του, χωρίς κρυφές διεργασίες ή χειρισμούς.

Κεφάλαιο 9^ο: Η εφαρμογή

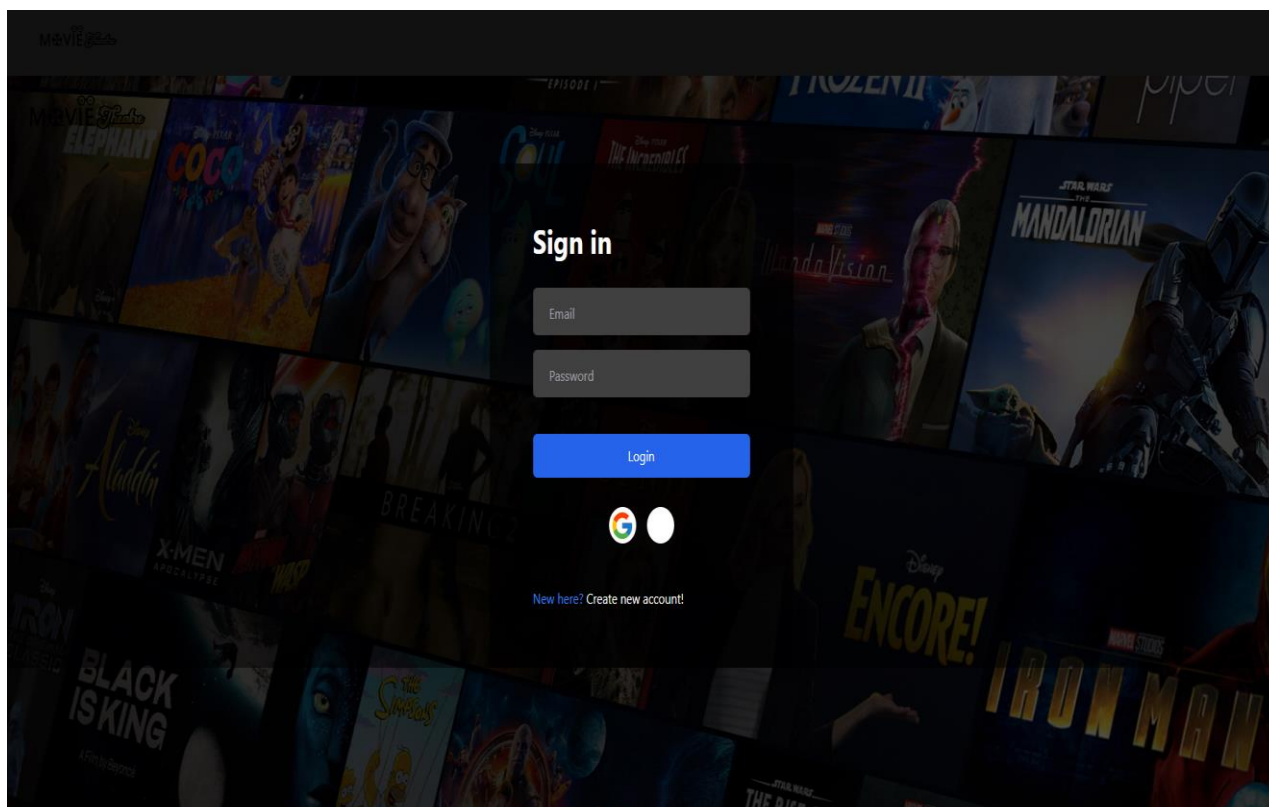
9.1 Ανάλυση της εφαρμογής

Το MovieTheater είναι μια εφαρμογή στην οποία ο χρήστης αφού δημιουργήσει λογαριασμό συνδέεται (ή συνδέεται αμέσως με google-github) και έχει την δυνατότητα να περιηγηθεί και να βρει πληροφορίες για ταινίες και σειρές, να τις παρακολουθήσει μέσω Netflix να βάλει στα αγαπημένα του όποιες θέλει, να τις σχολιάσει και να πάρει προτάσεις από το σύστημα.

9.1.1 Σελίδα σύνδεσης

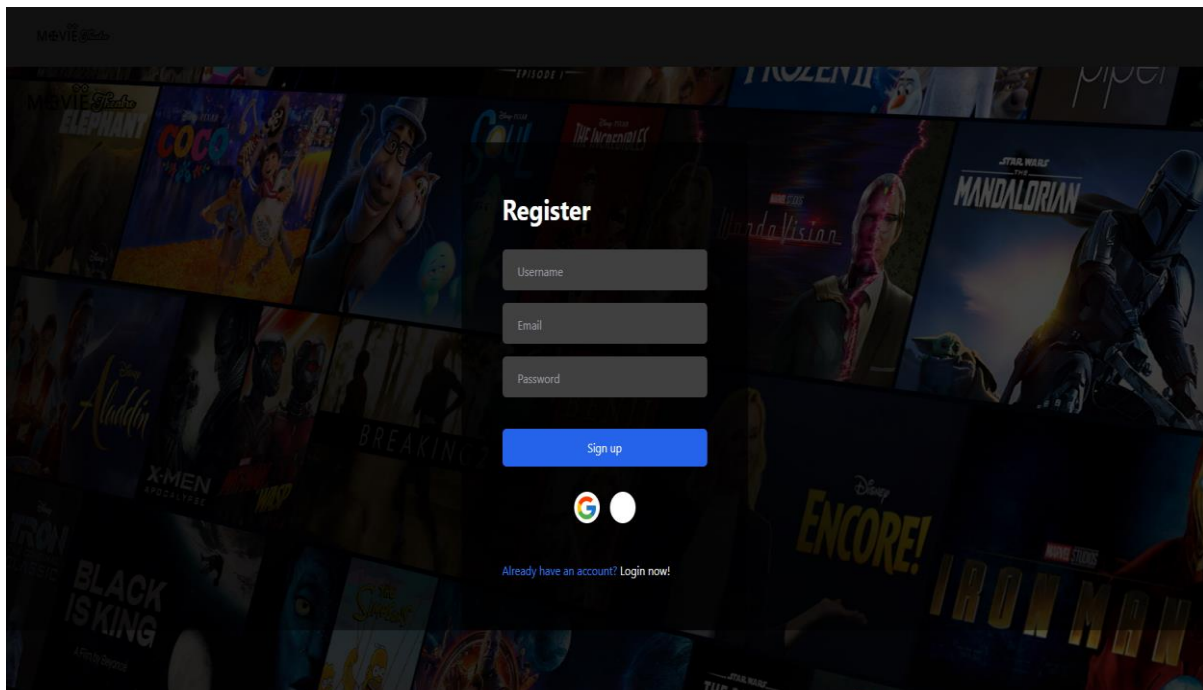
Με το που ο χρήστης μπει στην εφαρμογή, έρχεται σε επαφή απευθείας με την αρχική σελίδα που περιμένει κάθε χρήστη που δεν είναι συνδεδεμένος (βλέπε εικόνα 9.1), την σελίδα δηλαδή σύνδεσης ή δημιουργίας λογαριασμού. Το MovieTheater έχει ένα σύστημα σύνδεσης (authentication system) που δίνει στον χρήστη διάφορους τρόπους για να μπορέσει να συνδεθεί ή να δημιουργήσει λογαριασμό. Ο πιο απλός τρόπος είναι να κάνει είσοδο με το email του και έναν προσωπικό κωδικό που έχει φτιάξει ο ίδιος. Αν έχει ήδη λογαριασμό, απλά βάζει αυτά τα στοιχεία και συνδέεται.

Εκτός από αυτόν τον «παραδοσιακό» τρόπο, υπάρχει και η επιλογή να συνδεθεί μέσω Google, αν έχει Gmail, ή μέσω GitHub χωρίς να περάσει από την διαδικασία δημιουργίας λογαριασμού καθώς γίνεται αυτόματα στην πρώτη σύνδεση με αυτόν τον τρόπο. Αυτοί οι τρόποι κάνουν τη διαδικασία πιο γρήγορη, ειδικά αν κάποιος δεν θέλει να θυμάται ακόμα έναν κωδικό.



Εικόνα 9.1: Κεντρική σελίδα εισόδου με email, google github

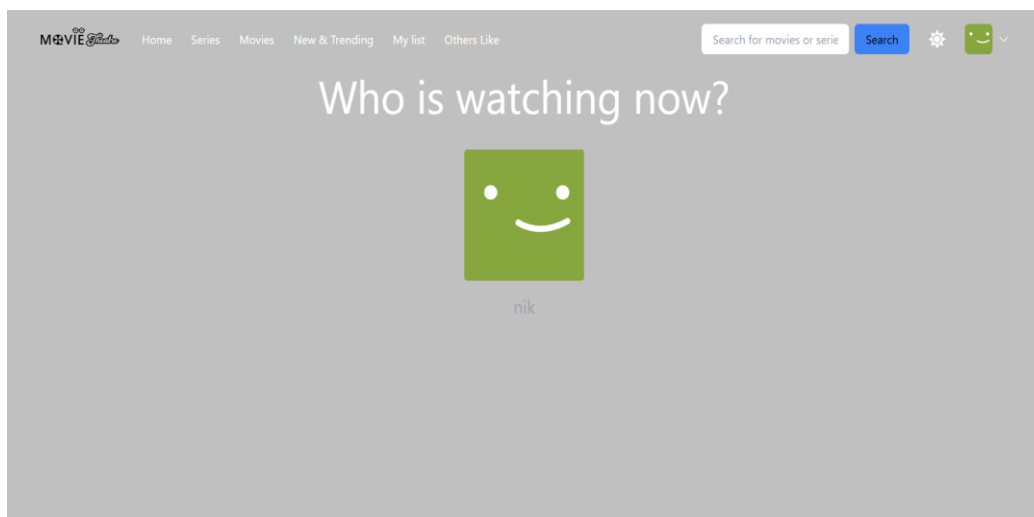
Για τους χρήστες που δεν έχουν λογαριασμό, υπάρχει η δυνατότητα εγγραφής (βλέπε εικόνα 9.2). Εκεί τους ζητάει η εφαρμογή να βάλουν ένα email, ένα username και έναν κωδικό. Το σύστημα πρώτα ελέγχει αν το email που έβαλαν χρησιμοποιείται ήδη από άλλο λογαριασμό. Αν δεν υπάρχει, τότε στέλνεται ένα email επιβεβαίωσης στον χρήστη. Μέσα σε αυτό το email υπάρχει ένας σύνδεσμος τον οποίο πρέπει να πατήσει για να ολοκληρωθεί η εγγραφή. Μόνο όταν γίνει αυτή η επιβεβαίωση, δημιουργείται ο λογαριασμός και μπορεί πλέον ο χρήστης να συνδεθεί κανονικά στην εφαρμογή. Με αυτόν τον τρόπο επιτυγχάνεται να κρατηθεί το σύστημα ασφαλές, χωρίς να δημιουργούνται ψεύτικοι λογαριασμοί από ανύπαρκτα email αλλά και να είναι εύκολο στη χρήση για όλους.



Εικόνα 9.2: Κεντρική σελίδα δημιουργίας λογαριασμού

9.1.2 Σελίδα επιλογής profile

Αφού ο χρήστης κάνει login στην εφαρμογή, το σύστημα τον κατευθύνει αυτόματα στην επόμενη σελίδα, που είναι η επιλογή προφίλ (βλέπε εικόνα 9.3),, εκεί όπου ο χρήστης δεν έχει την δυνατότητα δημιουργίας δεύτερου προφίλ αλλά όλες τα στοιχεία του είναι αποθηκευμένα σε ένα προφίλ, οπότε ουσιαστικά δεν χρειάζεται να επιλέξει κάτι, απλώς βλέπει το όνομά του και πατάει στο avatar για να συνεχίσει. Η ιδέα όμως είναι να μοιάζει με το Netflix, όπου κάθε χρήστης έχει το δικό του "προφίλ", ακόμα κι αν είναι μόνο ένα.



Εικόνα 9.3: Σελίδα επιλογής profile

Στο σημείο αυτό, όταν δηλαδή και γίνει η είσοδος του χρήστη στην εφαρμογή, ο χρήστης μπαίνει στο βασικό περιβάλλον της εφαρμογής, και για πρώτη φορά βλέπει το navigation bar (navbar), το οποίο είναι σταθερό και εμφανίζεται σε όλες τις εσωτερικές σελίδες της εφαρμογής αλλά και άλλες μικρές η πιο σημαντικές λειτουργίες. Από εδώ και πέρα, μπορεί να πλοηγηθεί στις βασικές ενότητες της πλατφόρμας.

Στο navbar υπάρχουν οι εξής επιλογές:

- Logo: Στην αριστερή γωνία του navbar εμφανίζεται το logo της εφαρμογής.
- Home: Η αρχική σελίδα της εφαρμογής
- Movies: Η σελίδα με τις ταινίες
- Series: Η σελίδα με τις σειρές
- MyList: Η σελίδα με τις αγαπημένες ταινίες και σειρές του χρήστη
- New & Trending: Η σελίδα με τις νέες και τις δημοφιλείς ταινίες που κυκλοφορούν
- Others Like: Η σελίδα όπου προτείνονται οι αγαπημένες ταινίες των άλλων
- Recommended: Η σελίδα όπου προτείνει ταινίες παρόμοιες με τις αγαπημένες του χρήστη

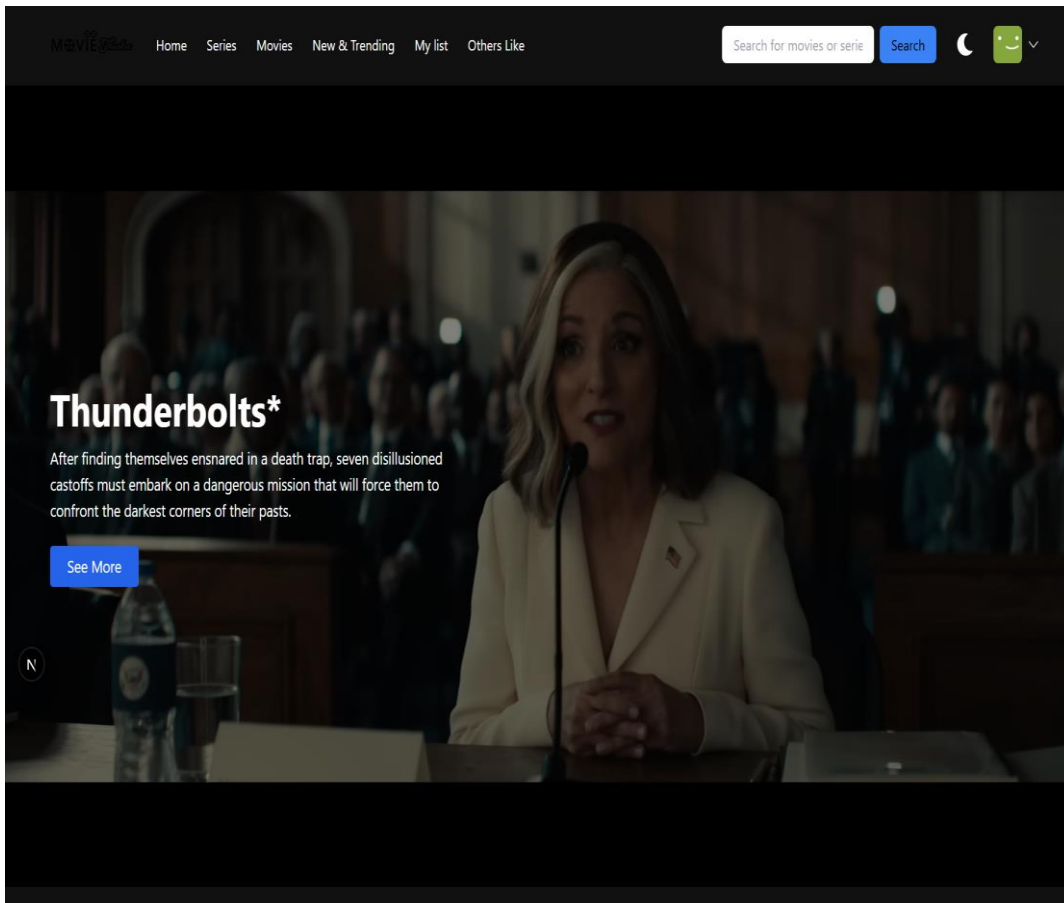
Στην συνέχεια και αφού τελειώσουν οι επιλογές των σελίδων στις οποίες μπορεί να περιηγηθεί ο χρήστης, δίπλα υπάρχει η μπάρα αναζήτησης (search bar) όπου ο χρήστης μπορεί να πληκτρολογήσει ό,τι θέλει και πατώντας enter ή στο εικονίδιο του search, η εφαρμογή τον μεταφέρει σε μια επίσης “κρυφή” σελίδα όπου και εμφανίζει όλα τα αποτελέσματα που θα βγάλει η αναζήτηση. Δεξιά του υπάρχει υλοποιημένο functionality για light-dark mode όπου πατώντας τον ήλιο ή το γεφφάρι (buttons) αλλάζουν τα χρώματα της εφαρμογής σε σκοτεινά ή φωτεινά. Τέλος στην δεξιά γωνία υπάρχει εικονίδιο προφίλ (avatar). Αν ο χρήστης πατήσει εκεί, εμφανίζεται ένα dropdown με επιλογή αποσύνδεσης (logout) ώστε να μπορέσει να βγει από την εφαρμογή. Με την αποσύνδεση (logout) του, ο χρήστης μεταφέρεται αυτόματα στην αρχική σελίδα εισόδου η δημιουργίας λογαριασμού και το navbar κρύβεται πάλι.

9.1.3 Homepage

Η κεντρική σελίδα και η σελίδα που σε μεταφέρει η εφαρμογή αυτόματα μετά την επιλογή προφίλ. Σε αυτή την σελίδα εμφανίζεται κάθε φορά μια καινούρια τυχαία ταινία μέσα από την τεράστιο αριθμό

περιεχομένου που διαθέτει η εφαρμογή, η οποία παίζει αυτόματα με λειτουργία repeat το trailer της μέσω της εφαρμογής YouTube. (βλέπε εικόνα 9.4),

Στο πλαίσιο όπου επαναλαμβάνεται το trailer υπάρχουν μερικές πληροφορίες για την ταινία από την περιγραφή της καθώς και ένα κουμπί το οποίο πατώντας το θα σε μεταφέρει αυτόματα σε άλλη σελίδα για τις λεπτομέρειες της ταινίας (επεξηγείται παρακάτω η σελίδα αυτή).

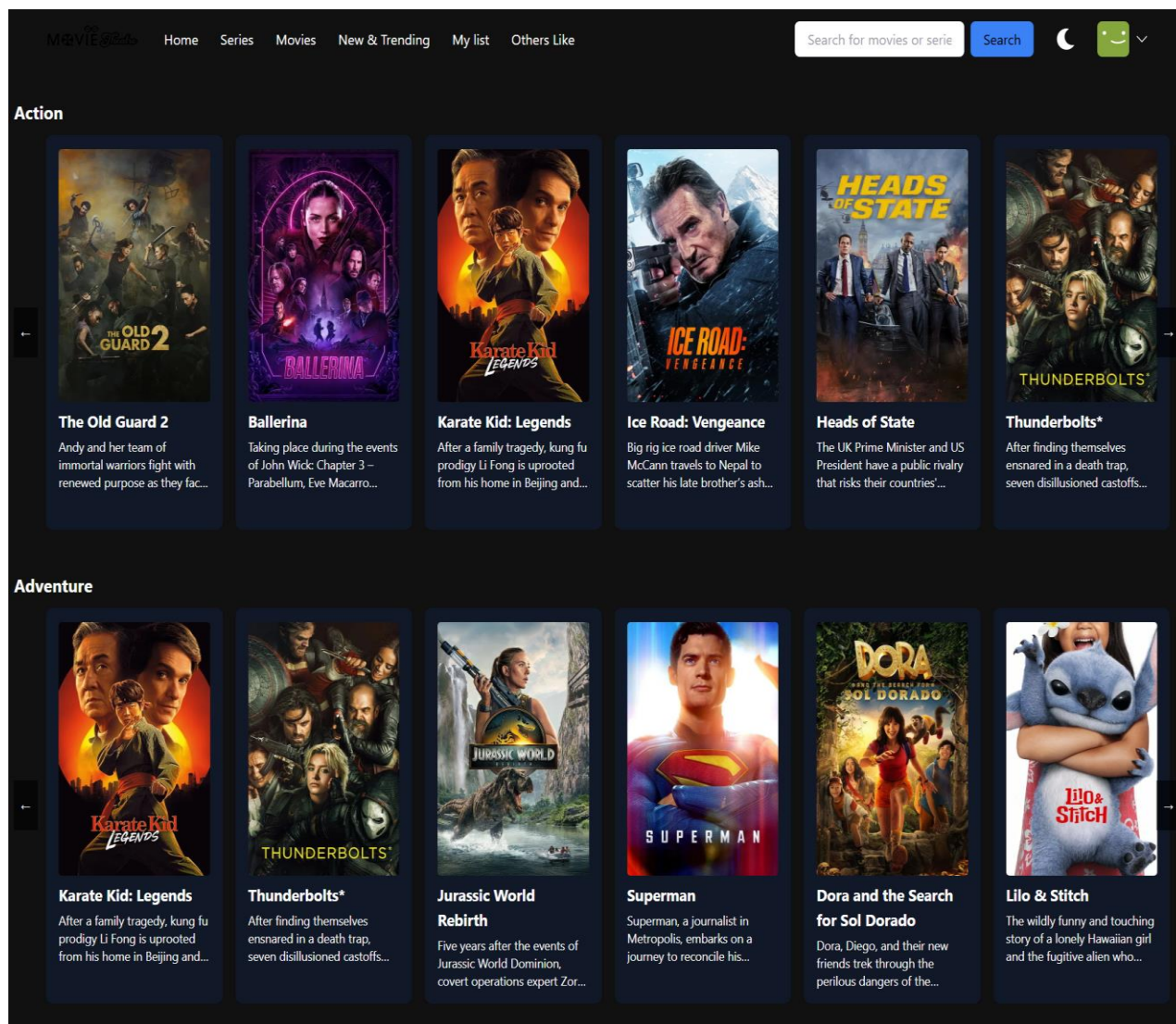


Εικόνα 9.4: Κεντρική σελίδα

9.1.4 Movies

Εδώ ο χρήστης βλέπει μόνο ταινίες. Η σελίδα εμφανίζει τις ταινίες σε μορφή μικρής κάρτας (βλέπε εικόνα 9.5), όπου φαίνεται ο τίτλος της ταινίας και η φωτογραφία (poster) της ταινίας καθώς και ένα κομμάτι από την περιγραφή της ταινίας, ανά κατηγορία-είδος ταινίας (κωμωδία, θρίλερ κλπ.) και κάθε είδος εμφανίζεται σε γραμμή, με τον χρήστη να έχει την δυνατότητα να περιηγηθεί κατά κάποιον τρόπο σε κάθε κατηγορία σκρολάροντας δεξιά-αριστερά για επόμενες ή προηγούμενες ταινίες της συγκεκριμένης κατηγορίας. Πατώντας πάνω σε μια ταινία ο χρήστης μεταφέρεται σε μια σελίδα η οποία δεν υπάρχει στο πανθαρ. Πρόκειται για την movie details σελίδα στην οποία ο χρήστης βλέπει αναλυτικά τις πληροφορίες της ταινίας, όπως το poster, τίτλος, πρωταγωνιστές, είδος, διάρκεια, γλώσσα, περιγραφή της ταινίας κ.α. Επίσης έχει την δυνατότητα να παρακολουθήσει το trailer της ταινίας μέσω YouTube pop-up window. Δίπλα υπάρχει άλλο button με το οποίο ο χρήστης μπορεί να αποθηκεύσει την ταινία στα αγαπημένα (favorites), ώστε να την δει αργότερα, η να χτίσει και το σύστημα μια εικόνα για τον χρήστη για να του προτείνει ταινίες ξεχωριστά βάση του γούστου του. Επιπλέον υπάρχει (όχι πάντα) ένα τρίτο button δίπλα το οποίο μέσω API key δίνει την δυνατότητα αν η ταινία υπάρχει στην βάση του API να παρακολουθήσει την ταινία στο Netflix.

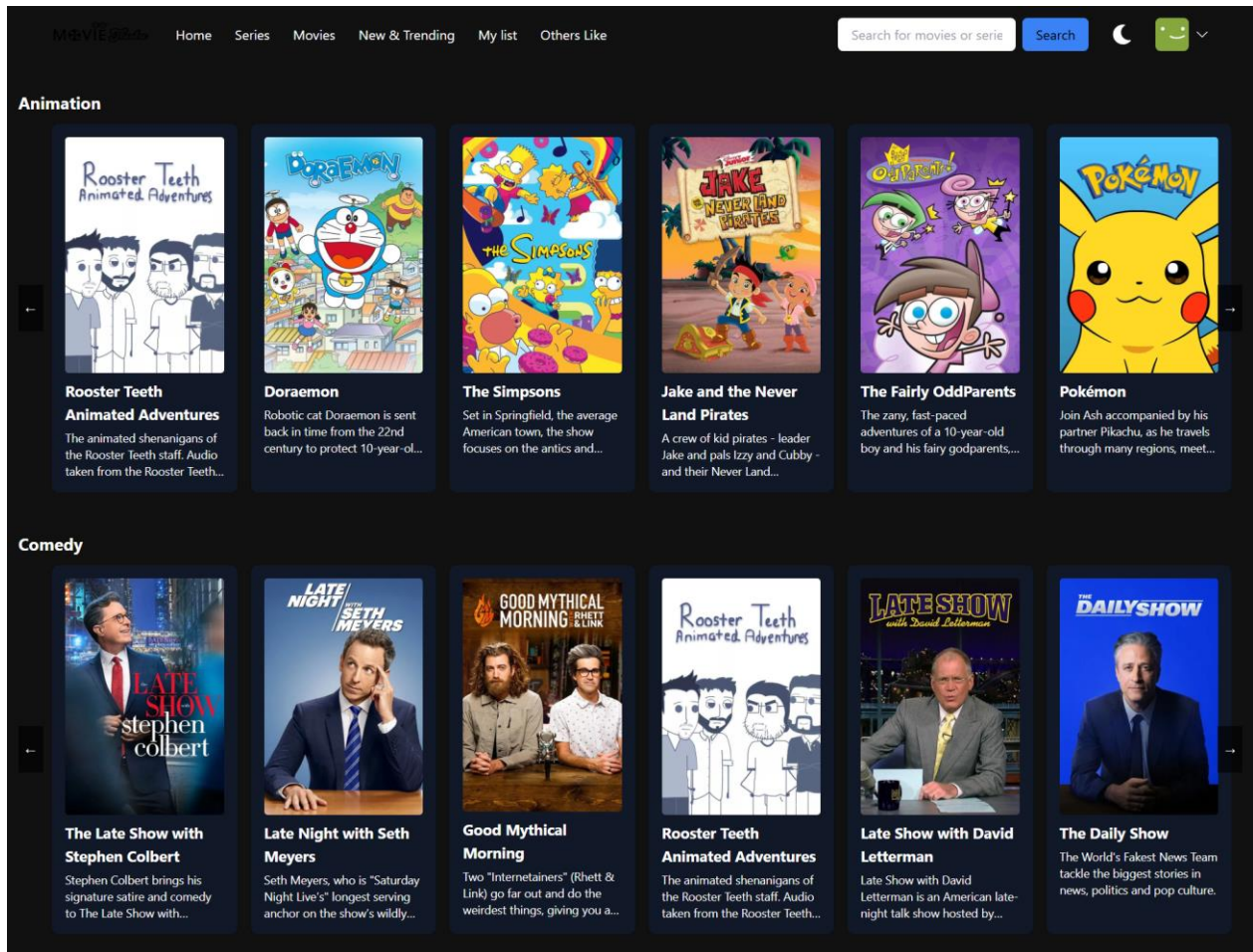
Πατώντας το, ανοίγει μια νέα σελίδα για τον χρήστη στο Netflix, στην συγκεκριμένη ταινία ώστε ο χρήστης να συνδεθεί και να παρακολουθήσει αμέσως την ταινία. Τέλος στο κάτω μέρος της σελίδας, υπάρχει το μέρος όπου οι χρήστες μπορούν να δημοσιεύσουν σχόλια για την ταινία τα οποία και θα εμφανίζονται επώνυμα από κάτω.



Εικόνα 9.5: Σελίδα ταινιών

9.1.5 Series

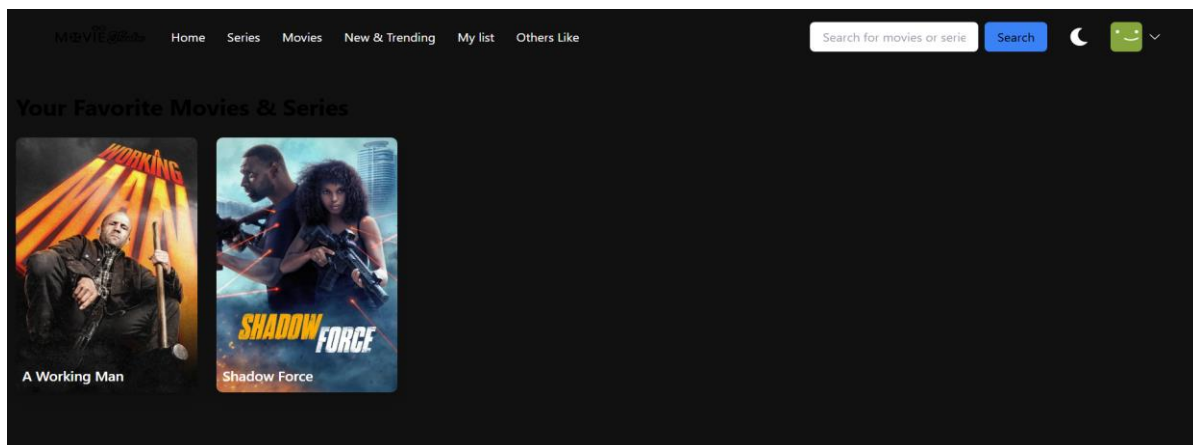
Αντίστοιχα, αυτή η σελίδα προσφέρει τις ίδιες λειτουργίες που θα βρει ο χρήστης στην σελίδα movies αλλά με σειρές, anime κλπ. (βλέπε εικόνα 9.6),



Εικόνα 9.6: Σελίδα σειρών

19.1.6 MyList

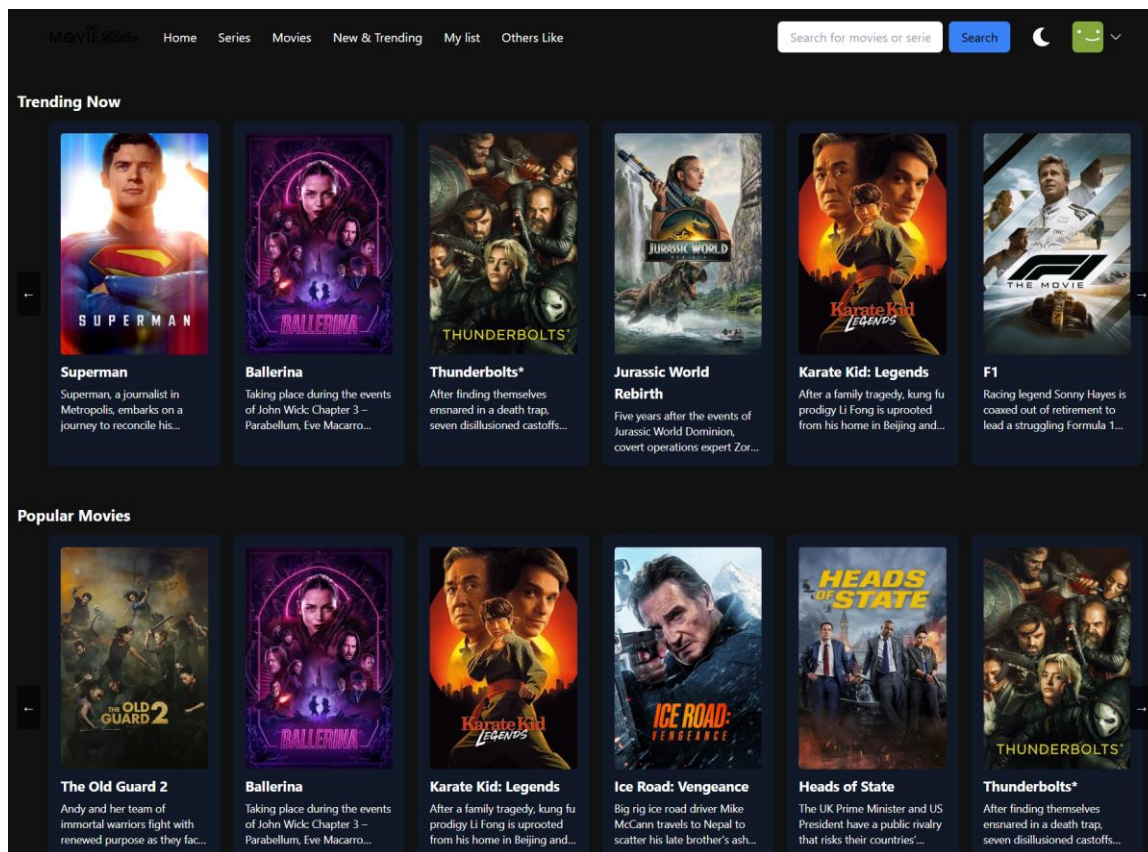
Σε αυτή την ενότητα αποθηκεύονται και εμφανίζονται όλα όσα έχει προσθέσει ο χρήστης στα αγαπημένα του από τις σελίδες movies-series (βλέπε εικόνα 9.7). Είναι μια προσωπική λίστα του χρήστη που χρησιμοποιεί και το σύστημα για να προτείνει στον χρήστη ταινίες παρόμοιες, βασισμένες σε αυτά τα γούστα του χρήστη ώστε να του ταιριάζουν.



Εικόνα 9.7: Σελίδα με τις αγαπημένες ταινίες και σειρές του χρήστη

9.1.7 New & Trending

Η σελίδα αυτή δείχνει τις ταινίες και τις σειρές οι οποίες είτε είναι καινούριες (άρα έχουν ένα λόγο παραπάνω να προταθούν-διαφημιστούν), είτε αποτελούν trending, δηλαδή έχουν πολλές προβολές την συγκεκριμένη χρονική περίοδο. (βλέπε εικόνα 9.8),



Εικόνα 9.8: Σελίδα με τις δημοφιλείς ταινίες και σειρές

9.1.8 Others Like

Αυτή είναι μια πιο "πειραματική" ενότητα όπου προβάλλονται περιεχόμενα που άρεσαν σε άλλους χρήστες. Αν και θεωρείται ένα είδος recommendation, δεν σημαίνει ότι είναι πάντα αποτελεσματικό, καθώς μπορεί ο χρήστης που είναι συνδεδεμένος να μην έχει κανένα απολύτως κοινό ενδιαφέρον με τους υπόλοιπους χρήστες της εφαρμογής, οπότε και να μην βρίσκει τίποτα το ενδιαφέρον στις ταινίες που αρέσουν στους υπόλοιπους χρήστες. Παρόλα αυτά είναι ένας τρόπος να έρχεται εύκολα και χωρίς πολύ αναζήτηση σε επαφή με ταινίες που αρέσουν σε άλλους άρα υπάρχει κάτι ενδιαφέρον σε αυτές και μπορεί κάποια να τον πείσει να την παρακολουθήσει.

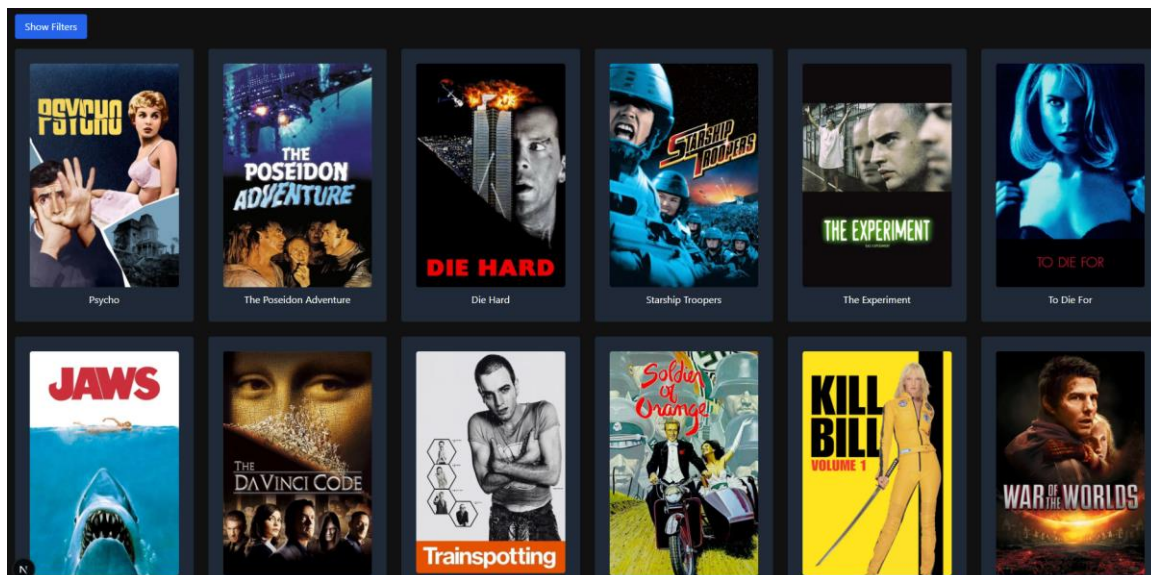
9.1.9 Recommended

Μία από τις βασικές σελίδες της εφαρμογής είναι η σελίδα με προτάσεις ταινιών, η οποία είναι φτιαγμένη έτσι ώστε να δείχνει στο χρήστη ταινίες που ταιριάζουν στα δικά του γούστα. Το σύστημα βασίζεται στις αγαπημένες ταινίες που έχει αποθηκεύσει ο ίδιος ο χρήστης. Για παράδειγμα, αν κάποιος έχει βάλει στις αγαπημένες του αρκετές κωμωδίες ή ταινίες δράσης, τότε σε αυτή τη σελίδα θα του εμφανιστούν προτάσεις από παρόμοια είδη, όπως άλλες κωμωδίες δράσης ή γενικά ταινίες με παρόμοιο ύφος.

Εκτός από το είδος, το σύστημα λαμβάνει υπόψη του και τους ηθοποιούς που εμφανίζονται στις αγαπημένες ταινίες του χρήστη. Αν, για παράδειγμα, κάποιος βλέπει συνέχεια ταινίες με τον ίδιο πρωταγωνιστή ή πρωταγωνίστρια, το σύστημα προτείνει και άλλες ταινίες που παίζει ο ίδιος αυτός ο ηθοποιός. Με αυτόν τον τρόπο, ο χρήστης βλέπει περιεχόμενο που είναι πιο πιθανό να του αρέσει, χωρίς να χρειάζεται να το ψάχνει μόνος του.

Επιπλέον, στη σελίδα αυτή υπάρχει και μια σειρά από φίλτρα, τα οποία βοηθούν τον χρήστη να ψάξει ακόμα πιο συγκεκριμένα αυτό που θέλει. Για παράδειγμα, μπορεί να φιλτράρει τις προτάσεις με βάση το είδος της ταινίας (π.χ. μόνο θρίλερ ή μόνο ρομαντικές). Τα φίλτρα κάνουν την αναζήτηση πολύ πιο εύκολη και πιο γρήγορη, ειδικά όταν ο χρήστης έχει στο μυαλό του κάτι πιο συγκεκριμένο.

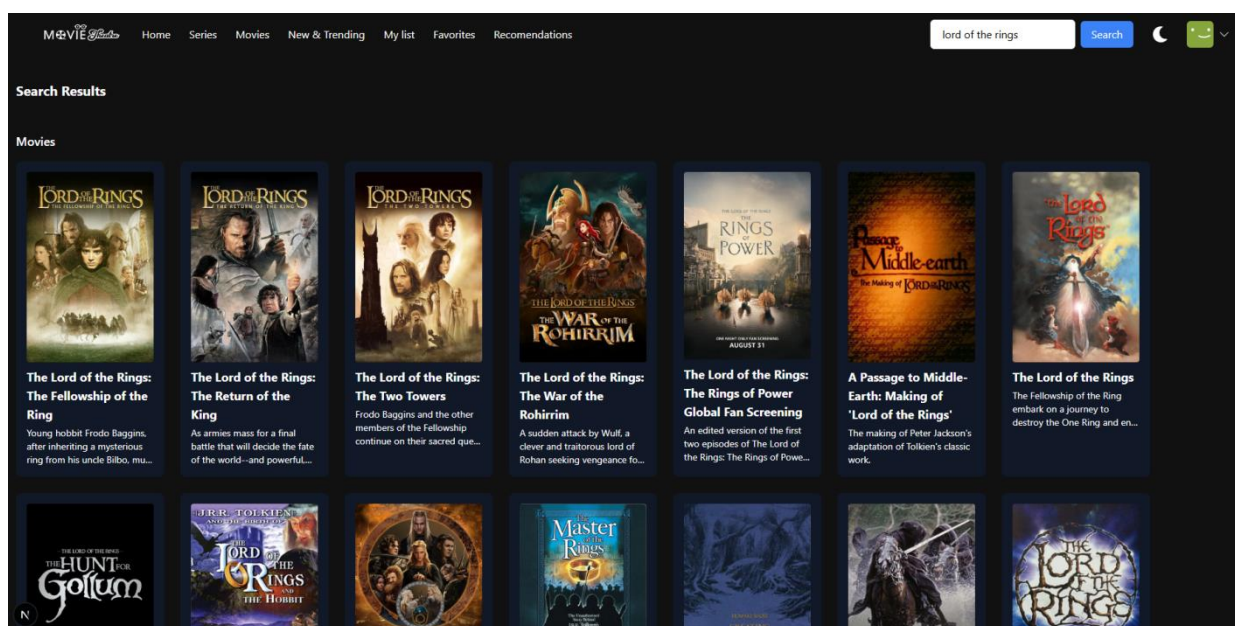
Συνολικά, η σελίδα αυτή βοηθά στο να γίνεται πιο «έξυπνη» η εμπειρία του χρήστη μέσα στην εφαρμογή, αφού του προτείνει περιεχόμενο που πραγματικά τον ενδιαφέρει και του γλιτώνει χρόνο από το να ψάχνει μόνος του. (βλέπε εικόνα 9.9),



Εικόνα 9.9: Σελίδα με τις προτεινόμενες ταινίες και σειρές για τον χρήστη

9.1.10 Search page

Αφού ο χρήστης πληκτρολογήσει στην μπάρα αναζήτησης την ταινία η σειρά που θέλει και πατήσει το κουμπί αναζήτησης, το σύστημα τον μεταφέρει στην σελίδα αναζήτησης. Στην σελίδα αυτή εμφανίζονται τα αποτελέσματα αναζήτησης του χρήστη με ταινίες που αντιστοιχίζουν ακριβώς στον όνομα που έψαξε η έχουν πολλά κοινά με αυτό. (βλέπε εικόνα 9.10),



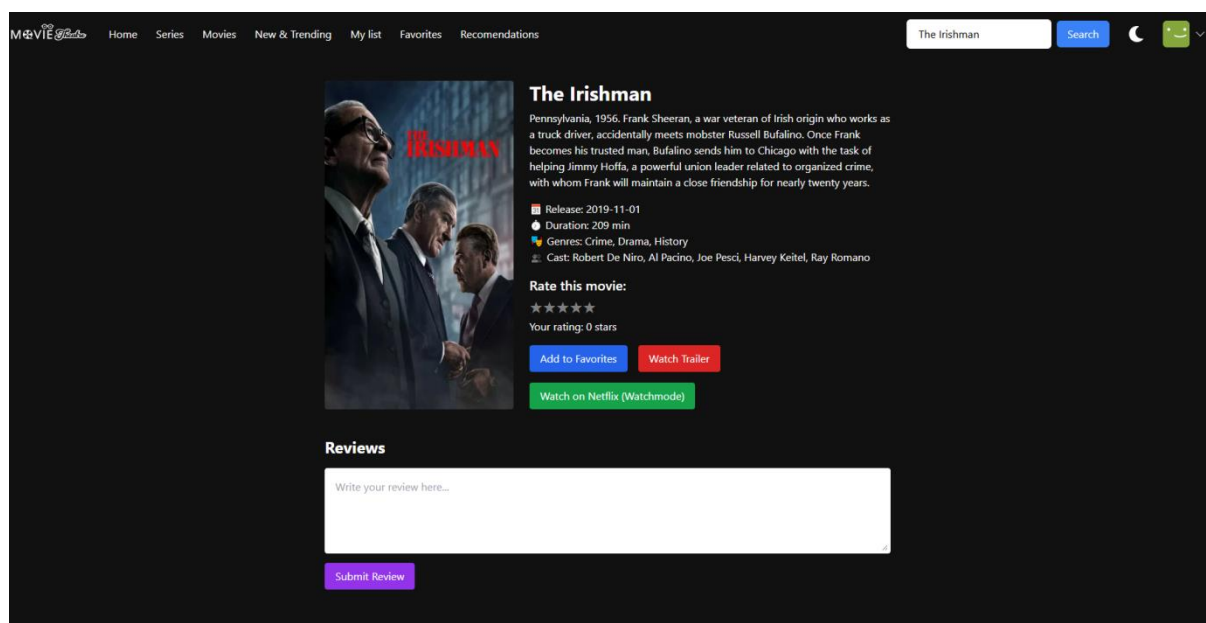
Εικόνα 9.10: Σελίδα με Τα αποτελέσματα αναζήτησης του χρήστη

9.1.11 Movie details

Η σελίδα για την κάθε ταινία η σειρά αποτελεί μια κεντρική σελίδα της εφαρμογής, υπεύθυνη για την παρουσίαση αναλυτικών πληροφοριών σχετικά με μία συγκεκριμένη ταινία-σειρά, καθώς και για την αλληλεπίδραση του χρήστη με αυτήν. Αποτελεί το σημείο όπου ο χρήστης μπορεί όχι μόνο να ενημερωθεί για το περιεχόμενο της ταινίας, αλλά και να αφήσει σχόλια ή να την προσθέσει στα αγαπημένα του. (βλέπε εικόνα 9.11),

Κύρια λειτουργικότητα:

- Εμφάνιση πληροφοριών: Προβάλλονται όλα τα βασικά μεταδεδομένα της ταινίας, όπως:
 - Τίτλος, είδος (genre), έτος κυκλοφορίας, σκηνοθέτης, βασικοί ηθοποιοί
 - Σύνοψη/περιγραφή της πλοκής
 - Εξώφυλλο ή εικόνα αφίσας της ταινίας
- Δυνατότητα σχολιασμού: Οι χρήστες μπορούν να γράψουν ή να διαβάσουν σχόλια για την ταινία, ενισχύοντας την κοινωνική διάσταση της πλατφόρμας.
- Προβολή βίντεο: Στην περίπτωση που είναι διαθέσιμο, παρέχεται ενσωματωμένος player για την αναπαραγωγή του επίσημου trailer. Σε ορισμένες περιπτώσεις, μπορεί να υπάρχει και σύνδεση με πλατφόρμες όπως το Netflix, ώστε ο χρήστης να μπορεί να ξεκινήσει απευθείας την προβολή.
- Προσθήκη στα αγαπημένα: Κουμπί αποθήκευσης της ταινίας στα “Add to Favorites” του χρήστη, ώστε να μπορεί να επιστρέψει εύκολα σε αυτήν σε μελλοντική πλοήγηση.



Εικόνα 9.11: Σελίδα ταινίας η σειράς με τις πληροφορίες της και τις δυνατότητες που προσφέρει στον χρήστη

Κεφάλαιο 10^ο: Τεχνολογίες και Εργαλεία Ανάπτυξης

Κατά την υλοποίηση της εφαρμογής, στόχος ήταν να αξιοποιηθούν μοντέρνα εργαλεία και τεχνολογίες του web, ώστε να δημιουργηθεί μια εμπειρία που να θυμίζει σύγχρονες πλατφόρμες streaming όπως το Netflix, το Amazon Prime ή το YouTube. Το project είναι δομημένο ως full-stack web εφαρμογή, κάτι που σημαίνει ότι περιλαμβάνει τόσο το frontend (δηλαδή το κομμάτι που βλέπει και αλληλεπιδρά ο χρήστης), όσο και το backend (που διαχειρίζεται τη βάση δεδομένων, τα APIs και τη λογική της εφαρμογής).

Frontend:

Η βάση της εφαρμογής είναι η React Js, μια βιβλιοθήκη JavaScript που επιτρέπει την ανάπτυξη διαδραστικών ιστοσελίδων με component-based λογική. Κάθε σελίδα ή στοιχείο της εφαρμογής είναι οργανωμένο σαν ξεχωριστό component, πράγμα που διευκολύνει την οργάνωση του κώδικα και την επαναχρησιμοποίηση λειτουργιών. Επίσης όλη η εφαρμογή είναι δομημένη με React Hooks όπως το useState, useEffect, useContext κ.ά., για να μπορεί να διαχειρίζεται ο developer την κατάσταση του UI και να ανταποκρίνεται άμεσα σε αλλαγές. Παράλληλα, δημιουργήθηκαν ξεχωριστά components για τις βασικές λειτουργίες, ώστε ο κώδικας να είναι οργανωμένος και εύκολα επεκτάσιμος. Επιπλέον, η React συνδυάζεται στην εφαρμογή με TypeScript, που είναι μια παραλλαγή της JavaScript με αυστηρότερο έλεγχο τύπων (types). Ο προγραμματιστής μπορεί να ορίσει τι τύπο δεδομένων περιμένει σε κάθε μεταβλητή, συνάρτηση ή αντικείμενο. Αυτό βοηθάει πιο εύκολα στον εντοπισμό λαθών στον κώδικα κατά τη διάρκεια ανάπτυξης καθώς και πιο ασφαλή και «καθαρό» κώδικα.

Frontend και Backend:

Η εφαρμογή είναι στημένη πάνω στο Next.js, ένα framework που βασίζεται στη React αλλά προσφέρει επιπλέον λειτουργίες, όπως server-side rendering, δυναμική δημιουργία σελίδων και routing. Το σημαντικότερο όμως είναι ότι επιτρέπει στον προγραμματιστή να γράψει και το backend της εφαρμογής μέσα στο ίδιο project, με τη βοήθεια των API routes. Χρησιμοποιώντας τα λεγόμενα API routes του Next.js, μπορώ να δημιουργήσω endpoints που λειτουργούν σαν κανονικό backend — δηλαδή μπορώ να στέλνω αιτήματα (requests), να τραβάω δεδομένα από τη βάση ή να αποθηκεύω νέα στοιχεία. Αυτό σημαίνει ότι δεν χρειάστηκε να φτιάξω ξεχωριστό server, αφού όλα τρέχουν «μέσα» στο Next.js. Με αυτό τον τρόπο, το app είναι full-stack αλλά ταυτόχρονα πιο απλό και ευέλικτο στην υλοποίηση.

Επιπλέον, για το backend μέρος της εφαρμογής, χρησιμοποιείται η TypeScript, η οποία παίζει πολύ σημαντικό ρόλο στην ποιότητα και την ασφάλεια του κώδικα. Η TypeScript είναι υπερσύνολο της JavaScript και προσθέτει στατικό τύπο μεταβλητών (static typing) όπως αναφέρεται και πιο πάνω.

Για παράδειγμα, όταν ο developer φτιάχνει API routes στο Next.js για να επικοινωνήσει με τη βάση δεδομένων του μέσω Prisma, μπορεί να ορίσει με ακρίβεια ποια δεδομένα περιμένει να έρθουν από τον client, ποια θα στείλει πίσω και τι είδους δομή έχουν. Αυτό βοηθάει να αποφεύγονται λάθη όπως το να περάσει ένα λάθος ID ή να λείπει κάποιο απαραίτητο πεδίο.

Επίσης, μέσω TypeScript μπορεί να εκμεταλλευτεί πλήρως το safe-type API του Prisma, που του επιτρέπει να γράφει queries προς τη MongoDB με προβλέψιμο και καθοδηγούμενο τρόπο, με autocompletion και χωρίς να ανησυχεί για syntax errors ή λάθη σε ονόματα πεδίων.

Η χρήση της TypeScript τελικά ενισχύει τη συντηρησιμότητα και την αξιοπιστία του backend, ειδικά σε μεγαλύτερες εφαρμογές, όπου η σαφήνεια και η ασφάλεια στον κώδικα γίνονται όλο και πιο σημαντικές.

Βάση/Διαχείριση Δεδομένων:

Για την αποθήκευση δεδομένων, χρησιμοποίησα τη MongoDB, μια βάση δεδομένων NoSQL που είναι ιδανική για web εφαρμογές. Εκεί αποθηκεύονται πληροφορίες όπως οι χρήστες, τα αγαπημένα τους, ιστορικό, προτάσεις κ.ά. Για να είναι πιο εύκολη η σύνδεση της εφαρμογής με τη βάση δεδομένων, χρησιμοποιήθηκε το Prisma, έναν ORM (Object Relational Mapper). Το Prisma επιτρέπει στον developer να γράφει queries με πιο απλό τρόπο, χωρίς να χρειάζεται να γράφει χειροκίνητα MongoDB queries. Επίσης του προσφέρει αυτόματο έλεγχο των τύπων και δομών δεδομένων, πράγμα που τον βοηθάει να αποφύγει πολλά λάθη.

Styling:

Για το design και το στυλ της εφαρμογής, επιλέχθηκε το Tailwind CSS, ένα σύγχρονο εργαλείο που βασίζεται σε utility classes. Αντί να υπάρχει μεγάλο custom CSS, δίνεται η δυνατότητα για styling στα στοιχεία απευθείας μέσα στον κώδικα HTML/JSX, κάνοντας το σχεδιασμό πολύ πιο γρήγορο και ευέλικτο. Με το Tailwind, ο προγραμματιστής πετυχαίνει ένα απλό αλλά λειτουργικό και καθαρό user interface, παρόμοιο με αυτό γνωστών streaming εφαρμογών. Επίσης το tailwind παρέχει εύκολες responsive design κλάσεις, ενώ είναι και ευκολότερο στην συντήρηση του.

TMDB API και Watchmode API – Εξωτερικά Δεδομένα:

Παρόλο που υπάρχει η βάση δεδομένων, το περιεχόμενο της εφαρμογής (ταινίες, σειρές, trailers, κ.ά.), δεν έρχονται από την βάση δεδομένων παρά μόνο σε συγκεκριμένες σελίδες όπου και τραβάει συγκεκριμένα δεδομένα (favorites). Αντίθετα λοιπόν, χρησιμοποιήθηκαν δύο αξιόπιστες πηγές μέσω εξωτερικών restful APIs:

- Το TMDB API (The Movie Database), από όπου η εφαρμογή τραβάει όλο το περιεχόμενο της, τι ταινίες τις σειρές, τίτλους, περιγραφές, κατηγορίες, εικόνες και trailers αυτών και τις εμφανίζει στις διάφορες σελίδες της.
- Το Watchmode API, το οποίο με την βοήθεια των πληροφοριών που δίνει το TMDB API (όνομα ταινίας), η εφαρμογή κάνει αναζήτηση σε αυτό το API το όνομα της ταινίας την οποία “τσεκάρει” ο χρήστης και αυτό του επιστρέφει μια απάντηση (true-false) αν υπάρχει διαθέσιμο link για να δει ο χρήστης την συγκεκριμένη ταινία-σειρά στο Netflix μέσω ενός netflix key που διαθέτει.

Με αυτά τα APIs, η εφαρμογή φαίνεται «γεμάτη» και δυναμική, χωρίς να χρειαστεί ο developer να κατασκευάσει ή να γεμίσει χειροκίνητα κάποια βάση περιεχομένου.

Ειδικές βιβλιοθήκες

Για την προβολή trailers, χρησιμοποιήθηκαν οι βιβλιοθήκες:

- react-youtube: ενσωματώνει εύκολα βίντεο από το YouTube μέσα στην εφαρμογή.
- watch-trailer: μια βιβλιοθήκη που βοηθάει να εντοπίζεται αυτόματα το σωστό trailer μιας ταινίας ή σειράς ψάχνοντας με το όνομα που επιστρέφει το TMDB API.
- Nodemailer: μια βιβλιοθήκη μέσω της οποίας αποστέλεται αυτόματα email verification στον χρήστη για επιβεβαίωση του λογαριασμού ώστε να μπορεί να συνδεθεί κανονικά μετά την εγγραφή του.

Αυτό προσφέρει μια πολύ πιο ολοκληρωμένη εμπειρία στον χρήστη, καθώς μπορεί να βλέπει αμέσως trailer για κάθε τίτλο χωρίς να φύγει από την εφαρμογή.

Διαχείριση Κώδικα:

Καθ' όλη τη διάρκεια της ανάπτυξης, χρησιμοποιήθηκε το Git για έλεγχο εκδόσεων, και το GitHub για να αποθηκεύεται ο κώδικας online και να υπάρχει backup ή και η δυνατότητα σε παλαιότερες εκδόσεις όταν χρειαζόταν.

Συμπερασματικά με τις παραπάνω τεχνολογίες δημιουργήθηκε μια εφαρμογή που συνδυάζει frontend και backend λογική, υποστηρίζει αποθήκευση χρηστών και περιεχομένου, ενσωματώνει δεδομένα από εξωτερικές πηγές και έχει λειτουργικό user interface. Παρά το γεγονός ότι δεν έχει ξεχωριστό server, η χρήση του Next.js με API routes της επιτρέπει να θεωρείται μια εφαρμογή full-stack, καθώς η backend λειτουργικότητα είναι ενσωματωμένη στο ίδιο σύστημα.

Κάποιες τεχνολογίες από αυτές είναι πολύ δημοφιλείς σε εφαρμογές αυτού του είδους. Αρκεί να δει κανείς τις τεχνολογίες που χρησιμοποιούν οι προαναφερθείσες εφαρμογές IMDb, JustWatch, Letterboxd, Netflix και Trakt.tv.

- Το IMDb χρησιμοποιεί σε ορισμένα σημεία της React.js, ενώ συνδέεται και με RESTful APIs.
- Το JustWatch χρησιμοποιεί React.js, Typescript, RESTful APIs. [46] [47]
- Το Netflix χρησιμοποιεί και αυτό React.js [44] [45]

Κεφάλαιο 11^ο Συμπεράσματα και προβληματισμοί

Δυσκολίες

Κατά την ανάπτυξη της εφαρμογής, όπως λογικό είναι σε κάθε εφαρμογή να συμβαίνει αυτό, έτσι και στο MovieTheater προέκυψαν διάφορες προκλήσεις τόσο σε τεχνικό όσο και σε σχεδιαστικό επίπεδο. Το κυριότερο πρόβλημα λόγω και της μη δωρεάν χρήσης ιστοσελίδων προβολής ταινιών, μονόδρομος αποτέλεσε η επιλογή η προβολή να γίνεται μέσω κάποιας πλατφόρμας streaming η οποία λειτουργεί φυσικά με πληρωμή. Το πρόβλημα δηλαδή ήταν στο Streaming περιεχομένου & API limitations καθώς το Netflix δεν παρέχει πλέον δημόσιο API (όπως κάποια χρόνια πριν) για άμεση προβολή ταινιών ή trailer, γεγονός που δυσχέρανε την υλοποίηση της πλήρους αναπαραγωγής περιεχομένου. Η προσπάθεια που έγινε ήταν η χρήση εναλλακτικών λύσεων δηλαδή η προβολής των πληροφοριών μέσω του TMDb API , η προβολή των trailer μέσω YouTube και η χρήση του watchmode API, το οποίο έδινε την δυνατότητα απλά ο χρήστης σε περίπτωση που ήταν συνδρομητής του Netflix, μέσα από το MovieTheater, αν θέλει να δει κάποια ταινία, να την παρακολουθεί μέσω του λογαριασμού του στο Netflix. Αυτό πρώτον αναγκάζει τον χρήστη ουσιαστικά να είναι συνδρομητής σε μια πλατφόρμα, αλλά επίσης έκανε και την εφαρμογή περίπλοκη στην υλοποίηση της και το έργο του προγραμματιστή δυσκολότερο.

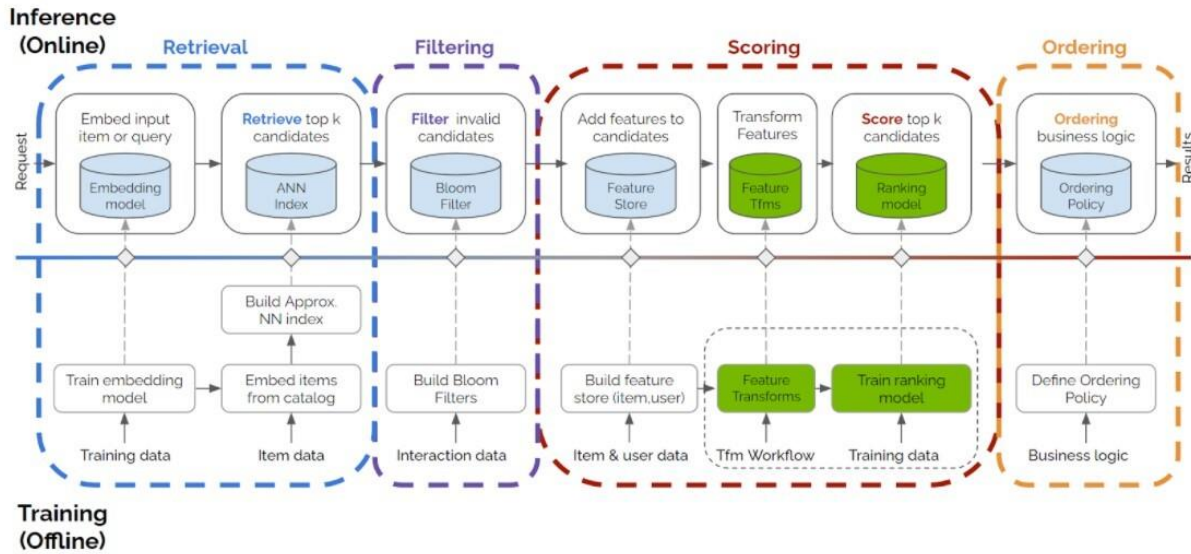
Ένα δεύτερο και σημαντικό πρόβλημα ήταν η δημιουργία συστήματος προτάσεων αποφεύγοντας όμως την δημιουργία κάποιου μοντέλου που θα εκπαιδευτεί μόνο του διότι θα ήθελε μηχανική μάθηση. Επιλέχθηκε λοιπόν μια άλλη λύση, κάπως πιο απλή η οποία βασίστηκε στα θέλω των χρηστών. Οι χρήστες δεν συνδέονται μεταξύ τους οπότε οι προτάσεις γίνονταν καθαρά βάση το τι ταινίες οι χρήστες αποθηκεύουν στις αγαπημένες τους.

Μελλοντικές αλλαγές και προσθήκες

Η εφαρμογή παρουσιάζει πού απλές λύσεις σε πολλά σημεία και γι αυτόν τον λόγο για τη μελλοντική εξέλιξη της εφαρμογής, έχουν προβλεφθεί τα εξής:

1. Προσωποποίηση & Καλύτερη Εξατομίκευση

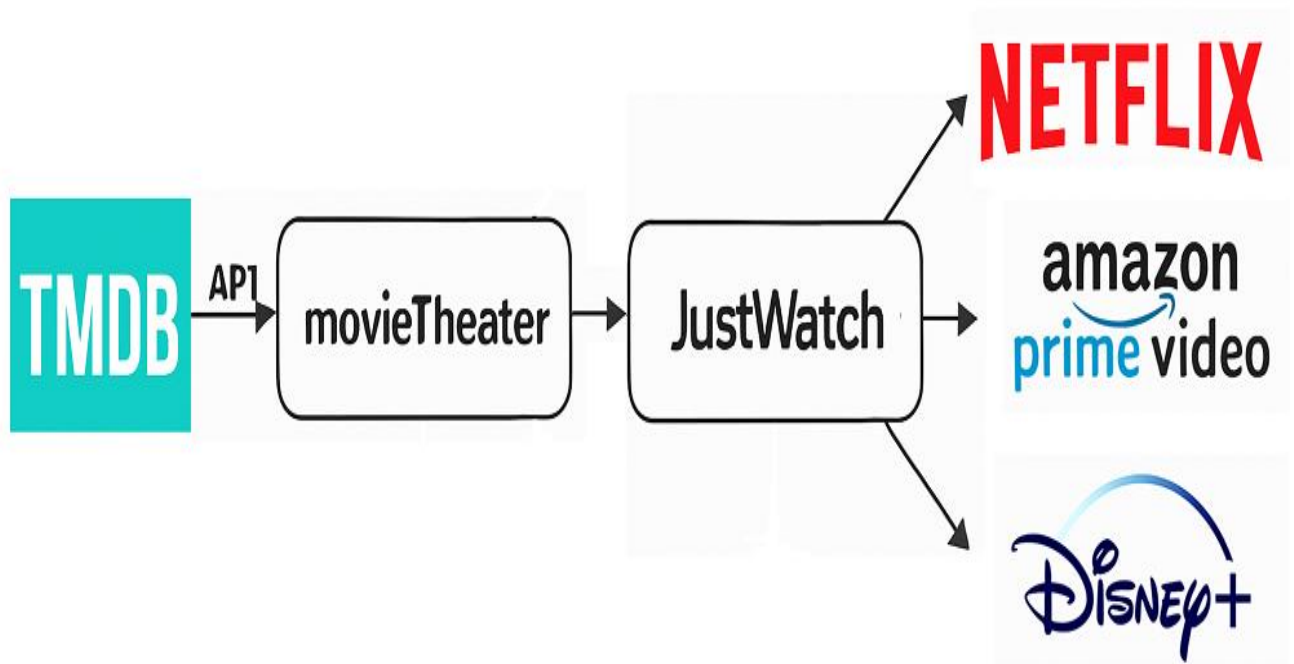
Αρχικά η εφαρμογή χρήζει περαιτέρω βελτίωση των συστημάτων συστάσεων μέσω machine learning (π.χ. deep learning models για πρόβλεψη προτιμήσεων). Στο σημείο αυτό οι προτάσεις του συστήματος βασίζεται σε πολύ απλά δεδομένα όπως τις ταινίες που αρέσουν στους χρήστες και τις αποθηκεύουν στα αγαπημένα τους. Αυτό δίνει την ευκαιρία σε μελλοντική φάση να υλοποιηθούν μοντέλα που να εκπαιδεύονται ανάλογα με τον κάθε χρήστη. Επίσης υπάρχει και η επιλογή για χρήση real-time δεδομένων για πιο άμεσες συστάσεις.



Εικόνα 11.1 Τα 4 στάδια του συστήματος συστάσεων [42]

2. Υποστήριξη Πολλαπλών Πηγών Περιεχομένου

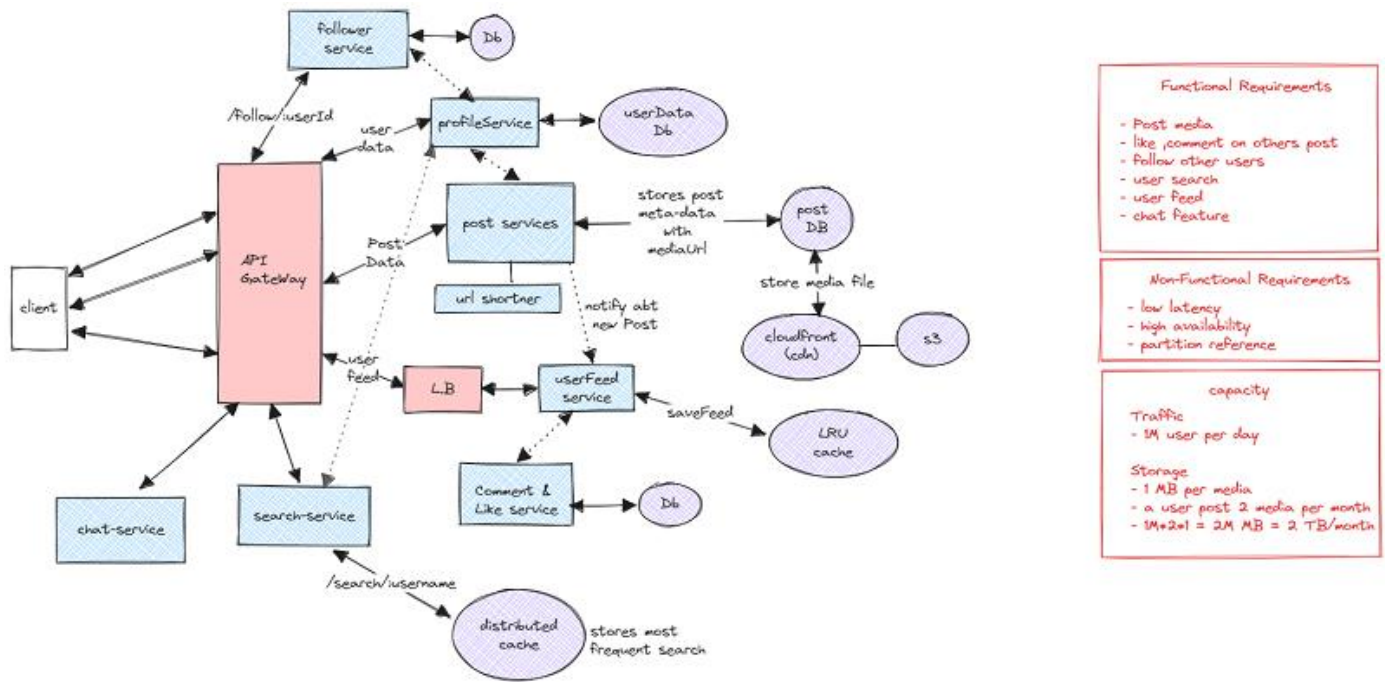
Ενσωμάτωση περιεχομένου και από άλλες streaming υπηρεσίες ή βάσεις δεδομένων, όπως το Amazon prime, Disney+, ώστε η εφαρμογή να γίνει πολυπλατφορμική (βλέπε εικόνα 11.2) και να μπορεί να εξυπηρετηθεί και κάποιος χρήστης που δεν είναι συνδρομητής στο Netflix αλλά σε κάποια άλλη παρόμοια πλατφόρμα.



Εικόνα 11.2 Μελλοντική ροή των API και των υποστηρικτικών πλατφορμών προβολής ταινίας

3. Social Features (κοινωνικά χαρακτηριστικά)

Δυνατότητα δημιουργίας watchlists, κοινής προβολής με φίλους (watch parties), σύνδεση χρηστών μέσω αιτήματος φιλίας άρα και σύνδεση των πληροφοριών του καθενός για υλοποίηση καλύτερου συστήματος προτάσεων βάση και των φίλων, καθώς και δυνατότητα αξιολόγησης των ταινιών.



. Εικόνα 11.3: Διάγραμμα συστήματος κοινωνικής σύνδεσης [48]

4. Βελτίωση User Experience (UX)

Εξέταση για περαιτέρω απλοποίηση του UI, με περισσότερες δυνατότητες σε menu αλλά και πιο φιλικό σε σύγχρονες ανάγκες με την υλοποίηση accessibility το οποίο καλείται να έχει υποχρεωτικά βάση ευρωπαϊκού νόμου κάθε σύγχρονη εφαρμογή.

5. Διαχείριση Περιεχομένου & Νομικά Θέματα

Ενσωμάτωση ροών περιεχομένου που είναι συμβατές με ανοιχτές άδειες ή υπηρεσίες που επιτρέπουν legal streaming. Το YouTube API ή και το Vimeo API θα ήταν τα κατάλληλα ώστε να παρέχω έξτρα υλικό στην εφαρμογή.

6. Σωστότερη διαχείριση κώδικα και folder structure

Αν και δεν είναι κομμάτι που μπορεί κανείς να θεωρήσει ως κάτι νέο, κάτι που θα προστεθεί στην εφαρμογή και θα προσφέρει κάτι καινούριο, η οργάνωση κώδικα και η σωστότερη οργάνωση του project με συγκεκριμένο τρόπο χρήσης φακέλων, με περισσότερα και μικρότερα reusable components, με component composition για να αποφευχθεί το prop drilling και να μην χρειάζεται σε κάθε αρχείο να μοιράζονται συνέχεια δεδομένα είναι σίγουρα μια αναβάθμιση που χρήζει απαραίτητη να γίνει. Αυτό διότι με ένα πιο σωστό folder structure και διανομή κάθε κομματιού κώδικα σε ξεχωριστό αρχείο και φάκελο όπου ανήκει, δηλαδή ξεχωριστά όλα τα types, ξεχωριστά όλα τα reusable components, αλλού τα feature components κλπ, επιτυγχάνεις αρχεία με μικρότερο μέγεθος κώδικα, άρα και πιο εύκολα στην κατανόηση και στην επεξεργασία. Αυτό έχει ως αποτέλεσμα κάθε καινούριο feature που σκοπεύει να βάλει ένας προγραμματιστής, όπως τα παραπάνω τα οποία σκοπεύω να ενσωματώσω στην πλατφόρμα στο μέλλον, να γίνεται ευκολότερα και με λιγότερες πιθανότητες να υπάρξουν ακατανόητα bus-errors τα οποία δεν θα γνωρίζω από που έρχονται.

7. Αυθεντικοποίηση χρήστη

Μία ακόμη μελλοντική αλλαγή που είναι απαραίτητη για το πρόγραμμα και χρήζει άμεση αλλαγή και ισχύ της νέας ιδέας είναι σχετικά με το session. Μέχρι στιγμής το πρόγραμμα δουλεύει με session, δηλαδή ο χρήστης συνδέεται στην πλατφόρμα και αποθηκεύονται εκεί τα στοιχεία σύνδεσης του. Το πρόβλημα σε αυτή την περίπτωση είναι ότι τα στοιχεία σύνδεσης δεν χάνονται, μένουν εκεί. Ακόμη και αν κλείσω την εφαρμογή χωρίς να αποσυνδεθώ, και ανοίξω την πλατφόρμα ξανά μετά από καιρό, θα παραμένω συνδεδεμένος από την τελευταία φορά. Αυτό καταλαβαίνει κανείς είναι ένα μικρό πρόβλημα που χρήζει λύση και αυτή η αλλαγή είναι ένα jwt token το οποίο θα λήγει μετά από κάποιο χρονικό διάστημα έτσι ώστε να ξανά χρειαστεί να συνδεθώ μετά από κάποιο διάστημα αδράνειας.

Συμπεράσματα

Η παρούσα διπλωματική εργασία είχε ως βασικό στόχο τη μελέτη και κατανόηση των συστημάτων συστάσεων, τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο, μέσα από την υλοποίηση της εφαρμογής MovieTheater. Κατά τη διάρκεια της εργασίας έγινε φανερό ότι τα recommendation systems δεν είναι απλώς τεχνικές κατηγοριοποίησης ή πρόβλεψης, αλλά αποτελούν βασικά εργαλεία που καθορίζουν τον τρόπο με τον οποίο λειτουργούν πολλές σύγχρονες ψηφιακές πλατφόρμες.

Σε θεωρητικό επίπεδο, διαπιστώθηκε πως η εξέλιξη των συστημάτων συστάσεων είναι συνεχής και εντυπωσιακή. Από τα απλά μοντέλα συνεργατικού φίλτραρίσματος (collaborative filtering) και τις content-based μεθόδους, περάσαμε σε πιο σύνθετες και δυναμικές λύσεις, που βασίζονται σε συνδυασμούς αλγορίθμων, βαθιά μάθηση, προσωποποίηση σε πραγματικό χρόνο και αξιοποίηση πολύπλοκων δεδομένων χρήστη. Η μελέτη συστημάτων όπως αυτά που εφαρμόζονται στο YouTube ή στο Netflix έδειξε ότι πρόκειται για πολυεπίπεδα οικοσυστήματα, στα οποία κάθε στάδιο (από την επιλογή υποψήφιου περιεχομένου μέχρι την τελική κατάταξη) παίζει σημαντικό ρόλο και προσαρμόζεται δυναμικά στις ανάγκες του κάθε χρήστη.

Από την άλλη πλευρά, η πρακτική υλοποίηση της εφαρμογής MovieTheater αποτέλεσε ένα πολύτιμο πεδίο εφαρμογής των εννοιών που μελετήθηκαν. Μέσα από αυτή τη διαδικασία έγινε πιο κατανοητό πόσο δύσκολη και πολυδιάστατη είναι στην πράξη η ανάπτυξη ενός συστήματος συστάσεων. Παρόλο που χρησιμοποιήθηκαν πιο απλές τεχνικές (όπως το user-based collaborative filtering), τα αποτελέσματα έδειξαν ότι ακόμη και αυτές οι μέθοδοι μπορούν να ενισχύσουν ουσιαστικά την εμπειρία του χρήστη, αρκεί να εφαρμοστούν σωστά και με συνέπεια.

Κατά την υλοποίηση της εφαρμογής, ήρθαν στην επιφάνεια και ορισμένοι σημαντικοί περιορισμοί. Ο πιο βασικός ήταν η έλλειψη πραγματικών δεδομένων χρήστη, κάτι που επηρέασε τόσο την ποιότητα των συστάσεων όσο και την αξιολόγηση της εφαρμογής. Επιπλέον, ο περιορισμένος χρόνος και η υπολογιστική ισχύς δεν επέτρεψαν την ενσωμάτωση πιο σύνθετων τεχνικών ή τη διαχείριση μεγάλου όγκου δεδομένων. Ένα χαρακτηριστικό παράδειγμα αποτελεί η διαχείριση του cold start problem, το οποίο αντιμετωπίστηκε στην εφαρμογή με έναν απλό, αλλά όχι ιδανικό τρόπο: προτείνοντας τυχαίες ταινίες στην αρχική σελίδα. Σε μελλοντικές εκδόσεις, θα μπορούσαν να εξεταστούν πιο στοχευμένες λύσεις, όπως η χρήση δημοφιλών ή πρόσφατων ταινιών ή ακόμα και η συλλογή βασικών προτιμήσεων από τον χρήστη κατά την εγγραφή.

| Dataset | Top Models | Metrics | Extra Training Data | Year |
|----------------|---------------------------------|-------------------|---------------------|------|
| MovieLens 100K | GHRM [90] | 0.887(RSME) | N | 2021 |
| | GLocal-K [91] | 0.889(RSME) | N | 2021 |
| | MG-GAT [92] | 0.890(RSME) | Y | 2020 |
| MovieLens 1M | GLocal-K [91] | 0.823(RSME) | N | 2021 |
| | Sparse FC [93] | 0.824(RSME) | N | 2018 |
| | CF-NADE [94] | 0.829(RSME) | N | 2016 |
| MovieLens 10M | Bayesian timeSVD++ [95] flipped | 0.749(RSME) | N | 2019 |
| | Bayesian timeSVD++ [95] | 0.752(RSME) | N | 2019 |
| | Bayesian SVD++ [95] | 0.756(RSME) | N | 2019 |
| MovieLens 20M | VASP [96] | 0.448(nDCG@100) | Y | 2021 |
| | H+Vamp Gated [97] | 0.445(nDCG@100) | Y | 2019 |
| | RecVAE [98] | 0.442(nDCG@100) | N | 2019 |
| Amazon-Book | SSCF [99] | 0.065(nDCG@20) | N | 2022 |
| | SANSA [100] | 0.064(nDCG@20) | N | 2023 |
| | BSPM-LM [101] | 0.061(nDCG@20) | N | 2022 |
| Netflix | H+Vamp Gated [97] | 0.409(nDCG@100) | N | 2019 |
| | RecVAE [98] | 0.394(nDCG@100) | N | 2019 |
| | EASE [102] | 0.393(nDCG@100) | N | 2019 |
| Last.FM | Ekar [103] | 0.248(HR@10) | N | 2019 |
| | HAKG [104] | 0.093(nDCG@20) | N | 2022 |
| | KGNN-LS [105] | 0.370(Recall@100) | N | 2019 |
| Yelp | DGRec [106] | 0.1427(nDCG) | N | 2019 |
| Gowalla | BSPM-EM [107] | 0.160(nDCG@20) | N | 2022 |
| | MGDCF [108] | 0.159(nDCG@20) | N | 2024 |
| | UltraGCN [109] | 0.158(nDCG@20) | N | 2021 |

Εικόνα 11.4: Datasets και συγκριτικές μέθοδοι για συστήματα συστάσεων

Ένα ακόμη σημαντικό σημείο που αναδείχθηκε μέσα από την εργασία είναι ότι η τεχνική ακρίβεια δεν είναι από μόνη της αρκετή. Η εμπειρία του χρήστη, ο τρόπος με τον οποίο αλληλεπιδρά με την εφαρμογή, η αίσθηση ότι οι προτάσεις "ταιριάζουν" στις προτιμήσεις του, είναι στοιχεία εξίσου καθοριστικά. Στην περίπτωση του MovieTheater, στόχος ήταν να προσφέρεται μια εμπειρία πλοήγησης που, παρόλο τον περιορισμένο όγκο δεδομένων, να φαίνεται πιο "προσωπική", με βάση είτε τη συμπεριφορά άλλων χρηστών είτε τη γενική δημοτικότητα των ταινιών.

| Dataset | U | I | L | density | $E(k_i)$ | $M(k_i)$ | $E(k_\alpha)$ | $M(k_\alpha)$ |
|-----------|--------|--------|---------|---------------------|----------|----------|---------------|---------------|
| MovieLens | 943 | 1,682 | 82,520 | $5.2 \cdot 10^{-2}$ | 88 | 1,018 | 49 | 1,116 |
| Netflix | 10,000 | 6,000 | 824,802 | $1.4 \cdot 10^{-2}$ | 82 | 2,120 | 137 | 9,018 |
| Amazon | 10,000 | 24,403 | 57,037 | $2.3 \cdot 10^{-4}$ | 5.7 | 679 | 2.3 | 1,198 |

Πίνακας 11.1 Στατιστικά χαρακτηριστικά των συνόλων δεδομένων (μέγεθος, πυκνότητα, βαθμός δραστηριότητας). [43]

Η σύγκριση με πλατφόρμες υψηλής κλίμακας, όπως το Netflix και το YouTube, έγινε περισσότερο για να κατανοηθούν οι τεχνικές που χρησιμοποιούν και όχι για να επιχειρηθεί άμεση σύγκριση. Είναι προφανές ότι τέτοιες πλατφόρμες διαθέτουν σχεδόν απεριόριστους πόρους, τεράστιες βάσεις δεδομένων και ολόκληρες ομάδες εξειδικευμένων επιστημόνων που εργάζονται διαρκώς για τη βελτίωση των συστημάτων τους. Ωστόσο, αυτό που προκύπτει ως συμπέρασμα είναι ότι ακόμη και απλούστερες υλοποιήσεις, σε μικρότερη κλίμακα, μπορούν να προσφέρουν αξία και να λειτουργήσουν ως βάση για περαιτέρω εξέλιξη.

Ιδιαίτερα ενδιαφέρον ήταν και το κομμάτι που αφορά τις ηθικές διαστάσεις των recommendation systems. Θέματα όπως η διαφάνεια στις προτάσεις, η πιθανότητα ύπαρξης προκαταλήψεων στους αλγόριθμους, ή ακόμα και η υπερ-εξατομίκευση που μπορεί να οδηγήσει σε φιλτράρισμα διαφορετικών απόψεων (filter bubbles), είναι ζητήματα που αξίζουν ιδιαίτερη προσοχή. Αν και στην παρούσα εφαρμογή αυτά τα θέματα δεν μπορούσαν να εξεταστούν σε βάθος, αποτέλεσαν σημαντικά σημεία προβληματισμού.

Τέλος, από προσωπικής πλευράς, η διαδικασία της υλοποίησης αποτέλεσε μια ευκαιρία ουσιαστικής μάθησης. Στην πορεία χρειάστηκε να αναθεωρηθούν αρχικές προσεγγίσεις και να βρεθούν λύσεις σε πρακτικά προβλήματα, κάτι που με βοήθησε να κατανοήσω καλύτερα τον ρόλο του προγραμματιστή αλλά και του ερευνητή στον χώρο της τεχνολογίας. Η εργασία δεν ήταν απλώς ένα τεχνικό project· ήταν μια εμπειρία που συνδύασε θεωρητική έρευνα, τεχνολογική δημιουργία και πρακτική επίλυση προβλημάτων.

Συνοψίζοντας η διπλωματική εργασία έδειξε ότι τα συστήματα συστάσεων είναι κάτι πολύ περισσότερο από έναν «αλγόριθμο που προτείνει». Είναι ένας πολύπλοκος και ταυτόχρονα καθοριστικός μηχανισμός που επηρεάζει το πώς οι χρήστες ανακαλύπτουν περιεχόμενο, παίρνουν αποφάσεις και αλληλεπιδρούν με τον ψηφιακό κόσμο.

Κεφάλαιο 11

.

Μέσα από τη θεωρητική μελέτη και την πρακτική υλοποίηση του MovieTheater, έγινε κατανοητό τόσο το εύρος των τεχνολογιών που εμπλέκονται όσο και η δυσκολία του να εφαρμοστούν αποτελεσματικά. Παρά τους περιορισμούς, η εργασία κατάφερε να δείξει ότι ακόμα και απλές προσεγγίσεις μπορούν να προσφέρουν σημαντική αξία — και κυρίως, να αποτελέσουν το πρώτο βήμα για κάτι μεγαλύτερο στο μέλλον.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Βιβλία

[30] Sergio Genovesi, Katharina Kaesling, Scott Robbins. *Recommender Systems: Legal and Ethical Issues*

Διαθέσιμο σε: <https://link.springer.com/content/pdf/10.1007/978-3-031-34804-4.pdf>

Internet Sites

[2] Netflix Tech Blog. (2017). How Netflix's Recommendations System Works.

Διαθέσιμο σε: <https://netflixtechblog.com>

[4] Amazon Annual Reports. (2020).

Διαθέσιμο σε: <https://www.amazon.com/ir>

[6] YouTube Official Blog. (2018). How Recommendations Shape What You Watch.

Διαθέσιμο σε: <https://youtube.googleblog.com>

[8] Paul Kim. (2019). *Behind the algorithm: YouTube's Recommendation System*

Διαθέσιμο σε: <https://p-kim.com/2019/02/28/behind-the-algorithm-youtubes-recommendation-system/>

[9] Yogesh Patodia (2022).

Διαθέσιμο σε :

<https://medium.com/@yogesh.patodia/recommender-system-using-reinforcement-learning-849308b615bf>

[10] William L. Hosch (2025) YouTube

Διαθέσιμο σε: <https://www.britannica.com/topic/YouTube>

[12] Jacob Murel Ph.D. , Eda Kavlakoglu (2024) What is content-based filtering?

Διαθέσιμο σε: <https://www.ibm.com/think/topics/content-based-filtering>

[13] Casey Johnston (2012) Netflix Never Used Its \$1 Million Algorithm Due To Engineering Costs

Διαθέσιμο σε: <https://www.wired.com/2012/04/netflix-prize-costs/>

[14] Haje Jan Kamps (2023) How technology unlocks business models: The story of Netflix

Διαθέσιμο σε: <https://techcrunch.com/2023/04/22/netflix-history-25-years/>

[15] Infovestopedia team, Julius Mansa, Vikki Velasquez (2023) How Netflix Is Changing the TV Industry

Διαθέσιμο σε: <https://www.investopedia.com/articles/investing/060815/how-netflix-changing-tv-industry.asp>

[16] Ko-Jen-Hsiao, Yesu Feng, Sudarshan Lamkhede (2025), Foundation Model for Personalized Recommendation

Διαθέσιμο σε : <https://netflixtechblog.com/foundation-model-for-personalized-recommendation-1a0bd8e02d39>

[17] Dr. Subhabaha Pal (2023), Comparison of Recommender Systems of Netflix and YouTube

Διαθέσιμο σε : <https://instadatahelp.com/comparison-of-recommender-systems-of-netflix-and-youtube/>.

[18] Evidently AI Team (2025), 10 metrics to evaluate recommender and ranking systems

Διαθέσιμο σε : <https://www.evidentlyai.com/ranking-metrics/evaluating-recommender-systems>

[20] IMDB (Internet Movie Database)

Διαθέσιμο σε: <https://www.imdb.com>

IMDB API (RapidAPI)

Διαθέσιμο σε: <https://rapidapi.com/blog/imdb-api/>

[21] JustWatch

Διαθέσιμο σε: <https://www.justwatch.com>

Github (JustWatch) API

Διαθέσιμο σε: <https://github.com/dawoudt/JustWatchAPI>

[22] Letterboxd

Διαθέσιμο σε: <https://letterboxd.com>

Letterboxd API

Διαθέσιμο σε: <https://letterboxd.com/api-beta/>

[23] Netflix

Διαθέσιμο σε: <https://www.netflix.com>

[24] Trakt.tv

Διαθέσιμο σε: <https://trakt.tv>

Trakt.tv API

Διαθέσιμο σε: <https://trakt.docs.apiary.io>

[28] Siddhant Jain, History of Netflix- Founding, Model, Timeline, Milestones (2025)

Διαθέσιμο σε: <https://www.vdocipher.com/blog/2017/06/netflix-revolution-part-1-history/>

[48] Social Media Application System Design Explained With System Diagrams & Requirements

Διαθέσιμο σε: <https://blog.calcont.in/2023/06/social-media-application-system-design.html>

[29] Maria Bastero, We delve into the exciting history of Netflix: How a global entertainment leader was forged (2025)

Διαθέσιμο σε: <https://marketing4ecommerce.net/en/history-of-netflix/>

[34] Alon Gubkin, What are Recommender Systems? Use Cases, Types, and Techniques

Διαθέσιμο σε: <https://cologix.com/ai-blog/what-are-recommender-systems-use-cases-types-and-techniques/>

[44] The Tech Behind Netflix's Unstoppable Streaming

Διαθέσιμο σε: <https://iamireneyu.medium.com/the-tech-behind-netflixs-unstoppable-streaming-74b30d630b25>

[45] Why Netflix chose React to build its user interface

Διαθέσιμο σε: <https://rootstack.com/en/blog/why-netflix-chose-react-build-its-user-interface>

[46] Building JustWatch, Part 1: A company culture that really works

Διαθέσιμο σε: <https://www.justwatch.com/blog/post/justwatch-company-culture>

[47] JustWatch docs-devops

Διαθέσιμο σε: https://www.justwatch.com/au/intern_devops

Papers in Conference Proceedings

[5] Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)*, 191–198.

Διαθέσιμο σε : <https://doi.org/10.1145/2959100.2959190>

[26] Shuai Zhang, Deep Learning based Recommender System: A Survey and New Perspectives
Διαθέσιμο σε: <https://arxiv.org/pdf/1707.07435>

[27] Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin, “Why Should I Trust You?” Explaining the Predictions of Any Classifier

Διαθέσιμο σε: <https://arxiv.org/pdf/1602.04938>

[32] Shaina Raza, Mizanur Rahman, Safiullah Kamawal, Armin Toroghi, Ananya Raval, Farshad Navah, Amirmohammad Kazemeini, A Comprehensive Review of Recommender Systems: Transitioning from Theory to Practice (2024)

Διαθέσιμο σε: <https://arxiv.org/pdf/2407.13699>

[33] Jia Sua , Yi Guana , Yuge Lia , Weile Chena , He Lvb , Yageng Yanb, Do recommender systems function in the health domain: a system review

Διαθέσιμο σε: <https://arxiv.org/pdf/2007.13058>

[35] Erion Cano, Maurizio Morisio, Hybrid Recommender Systems: A Systematic Literature Review

Διαθέσιμο σε: <https://arxiv.org/pdf/1901.03888>

[36] Alvisè De Biasio^{a,b}, Nicolò Navarino, Dietmar Jannach, Economic Recommender Systems – A Systematic Review

Διαθέσιμο σε: <https://arxiv.org/pdf/2308.11998>

[37] Franziska Scherpinski, Stefan Lessmann, Personalization in E-Grocery: Top-N versus Top-k Rankings

Διαθέσιμο σε: <https://arxiv.org/pdf/2105.14599>

[38] DIETMAR JANNACH, MICHAEL JUGOVAC, Measuring the Business Value of Recommender Systems

Διαθέσιμο σε: <https://arxiv.org/pdf/1908.08328>

[39] Ziyuan Xia, Anchen Sun, Jingyi Xu, Yuanzhe Peng, Rui Ma, Minghui Cheng, Contemporary Recommendation Systems on Big Data and Their Applications: A Survey

Διαθέσιμο σε: <https://arxiv.org/pdf/2206.02631>

[41] Sergey Fedorov, Chris Pham, Flavio Ribeiro, Chris Newton, and Wei Wei, FM-Intent: Predicting User Session Intent with Hierarchical Multi-Task Learning

Διαθέσιμο σε: <https://noise.getoto.net/author/netflix-technology-blog>

[42] Best Practices for Building and Deploying Recommender Systems

Διαθέσιμο σε: [Best Practices for Building and Deploying Recommender Systems - NVIDIA Docs](#)

[40]Jinseok Seol, Youngrok Ko, Sang-goo Lee , Exploiting Session Information in BERT-based Session-aware Sequential Recommendation

Διαθέσιμο σε: [Exploiting Session Information in BERT-based Session-aware Sequential Recommendation](#)

[43] Fei Yu, An Zeng., Sebastien Gillard, Matus Medo, Network-based recommendation algorithms: A review

Διαθέσιμο σε: <https://arxiv.org/pdf/1511.06252>

Πατέντες

[25] Benoit Lardeux, David Renaudie, Rodrigo Alejandro Acuna Agost, Eoin Thomas, Mourad Boudia, Papa Birame Sane (2022), Recommender systems and methods using cascaded machine learning models

Διαθέσιμο σε: <https://patents.justia.com/patent/11538086>

Journal Articles

[1] Gomez-Uribe, C. A., & Hunt, N. (2015). The Netflix Recommender System: Algorithms, Business Value, and Innovation. ACM Transactions on Management Information Systems, 6(4), 13.

Διαθέσιμο σε :<https://dl.acm.org/doi/pdf/10.1145/2843948>

[3] Smith, B., & Linden, G. (2017). Two Decades of Recommender Systems at Amazon.com. IEEE Internet Computing, 21(3), 12-18.

Διαθέσιμο σε : <https://doi.org/10.1109/MIC.2017.72>

[7] Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. IEEE Computer, 42(8), 30–37.

Διαθέσιμο σε : <https://doi.org/10.1109/MC.2009.263>

[11] Jun Hong (2009) A Survey of Collaborative Filtering Techniques

Διαθέσιμο σε: <https://onlinelibrary.wiley.com/doi/full/10.1155/2009/421425>

[19] Deepjyoti Roy, Mula Dutta (2022) , A systematic review and research perspective on recommender systems

Βιβλιογραφία

Διαθέσιμο σε : <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00592-5>

[31] Silvia Milano, Mariarosaria Taddeo, Luciano Floridi, Recommender systems and their ethical challenges

Διαθέσιμο σε: <https://link.springer.com/article/10.1007/s00146-020-00950-y>