



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Εφαρμογή πρότασης πλάνου δραστηριοτήτων με βάση το
προφίλ του εκάστοτε χρήστη για τουρίστες στην Ελλάδα**

Αλέξανδρος Αργυρός
A.M. 18390047

Εισηγητής: Χρήστος Τρούσσας, Επίκουρος Καθηγητής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Εφαρμογή πρότασης πλάνου δραστηριοτήτων με βάση το προφίλ του
χρήστη για τουρίστες στην Ελλάδα**

**Αλέξανδρος Αργυρός
Α.Μ. 18390047**

Εισηγητής:

Χρήστος Τρούσσας, Επίκουρος Καθηγητής

Εξεταστική Επιτροπή:

Παναγιώτα Τσελέντη, μέλος Ε.ΔΙ.Π.

Ακριβή Κρούσκα, μέλος Ε.ΔΙ.Π.

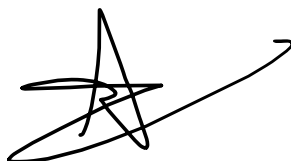
Ημερομηνία εξέτασης : Μάρτιος 2024

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Βεβαιώνω ότι είμαι συγγραφέας αυτής της Διπλωματικής εργασίας και κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου. Επιπλέον βεβαιώνω ότι έχω αναφέρει ή παραπέμψει σε αυτή, ρητά και συγκεκριμένα, όλες τις πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, προτάσεων ή λέξεων, είτε αυτές μεταφέρονται επακριβώς (στο πρωτότυπο ή μεταφρασμένες) είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για την συγκεκριμένη διπλωματική εργασία.

Ο Δηλών

Αλέξανδρος Αργυρός



ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία ολοκληρώθηκε μετά από επίμονες προσπάθειες, σε ένα ενδιαφέρον γνωστικό αντικείμενο, όπως αυτό των συστημάτων προτάσεων. Την προσπάθειά μου αυτή υποστήριξε ο επιβλέπων καθηγητής μου Χρήστος Τρούσσας, τον οποίο θα ήθελα να ευχαριστήσω.

Ακόμα θα ήθελα να ευχαριστήσω και την οικογένειά μου για την συμπαράσταση της κατά την διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία αναλύει και υλοποιεί ένα σύστημα προτάσεων με στόχο τη βελτίωση της εμπειρίας των τουριστών στην Ελλάδα. Εστιάζοντας στη δημιουργία εξατομικευμένων προφίλ για κάθε χρήστη, το σύστημα προτάσεων αναλύει τις προτιμήσεις και τα ενδιαφέροντα του κάθε τουρίστα, παρέχοντας σχεδιασμένα πλάνα δραστηριοτήτων.

Με τη χρήση γενετικού αλγορίθμου, η εφαρμογή παρέχει προτάσεις που προσαρμόζονται δυναμικά στις ατομικές προτιμήσεις του κάθε χρήστη. Αναλύοντας τις χώρες που έχει επισκεφθεί στο παρελθόν αλλά και τις προτιμήσεις του για το επόμενο του ταξίδι στην Αθήνα, το σύστημα προτάσεων δημιουργεί εξατομικευμένα πλάνα που προάγουν την ανακάλυψη νέων και ενδιαφέροντων δραστηριοτήτων.

Η εφαρμογή αναδεικνύει τη σημασία της τεχνολογίας στη βελτίωση της τουριστικής εμπειρίας, παρέχοντας ένα προηγμένο εργαλείο που ενισχύει τη σχεδίαση προσαρμοσμένων προγραμμάτων για τους επισκέπτες στην Αθήνα.

Λέξεις Κλειδιά: Συστήματα προτάσεων, Γενετικός Αλγόριθμος

ABSTRACT

The present thesis analyzes and implements a recommendation system aimed at enhancing the experience of tourists in Greece. Focusing on creating personalized profiles for each user, the recommendation system analyzes the preferences and interests of each tourist, providing curated activity plans.

Using a genetic algorithm, the application offers dynamically tailored suggestions based on the individual preferences of each tourist. By examining the countries visited in the past and preferences for the upcoming trip to Athens, the recommendation system generates personalized plans that promote the discovery of new and interesting activities.

The application highlights the significance of technology in improving the tourist experience, offering an advanced tool that enhances the design of customized programs for visitors in Athens.

Keywords: Recommendation systems, Genetic Algorithm

Πίνακας Περιεχομένων

ΚΕΦΑΛΑΙΟ 1 - ΕΙΣΑΓΩΓΗ	14
1.1 Εισαγωγή.....	14
1.2 Σκοπός	14
1.3 Δομή.....	14
1.4 Συντομογραφίες.....	15
ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΚΑΙ ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	16
2.1 Θεωρητικό Υπόβαθρο	16
2.1.1 Εισαγωγή στα Συστήματα προτάσεων	16
2.1.2 Τρόποι Συλλογής δεδομένων από τους χρήστες.....	16
2.1.3 Φιλτράρισμα Βάση Περιεχομένου(Content-Based Filtering)	17
2.1.4 Συνεργατικό Φιλτράρισμα(Collaborative Filtering)	19
2.1.5 Υβριδικά Συστήματα(Hybrid Systems)	20
2.1.6 Συστήματα Γνώσης(Knowledge-Based RS)	21
2.1.7 Δημογραφικά Συστήματα(Demographic RS)	21
2.1.8 Τεχνητή Νοημοσύνη και Συστήματα προτάσεων	22
2.1.9 Ομαδοποίηση(ML Classification).....	22
2.1.10 Βαθιά Μάθηση(Deep Learning).....	22
2.1.11 Οι δυσκολίες των Συστημάτων προτάσεων	23
2.1.12 Αναπτυσσόμενες Τεχνικές στα Συστήματα προτάσεων	23
2.2 Βιβλιογραφική Ανασκόπηση.....	26
ΚΕΦΑΛΑΙΟ 3 – ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΕΡΓΑΛΕΙΑ	30
3.1 Εισαγωγή.....	30
3.2 Βήματα σχεδίασης του αλγορίθμου.....	30
3.3 Γενετικοί Αλγόριθμοι.....	31
ΚΕΦΑΛΑΙΟ 4 – ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΛΟΓΙΚΗ ΣΥΣΤΗΜΑΤΟΣ	34
4.1 Εισαγωγή.....	34
4.2 Αναπαράσταση των χρηστών στη Βάση Δεδομένων	34
4.3 Αναπαράσταση των δραστηριοτήτων στη Βάση Δεδομένων	34
4.4 Ο αλγόριθμος του συστήματος προτάσεων.....	35
ΚΕΦΑΛΑΙΟ 5 – ΠΑΡΑΔΕΙΓΜΑΤΑ ΧΡΗΣΗΣ	43
5.1 Εισαγωγή.....	43
5.2 Παραδείγματα χρήσης.....	43
ΚΕΦΑΛΑΙΟ 6 – ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΕΠΕΚΤΑΣΕΙΣ	49

6.1 Εισαγωγή.....	49
6.2 Σχολιασμός αποτελεσμάτων αλγορίθμου.....	49
6.3 Μελλοντικές επεκτάσεις της εφαρμογής.....	49
6.4 Συμπεράσματα	50
ΚΕΦΑΛΑΙΟ 7 - ΒΙΒΛΙΟΓΡΑΦΙΑ.....	51
7.1 Βιβλιογραφία.....	51

Κατάλογος Εικόνων

Εικόνα 1: Μοντέλο Πρόβλεψης [11]

Εικόνα 2: Μοντέλο Κατάταξης [11]

Εικόνα 3: Γενική Εικόνα ΣΠ [10]

Εικόνα 4: Content-Based Filtering [14]

Εικόνα 5: Collaborative Filtering [12]

Εικόνα 6: Hybrid Systems [12]

Εικόνα 7: Είδη ΣΠ [13]

Εικόνα 8: Συστάσεις Συνομιλίας [14]

Εικόνα 9: Συστήματα Συστάσεων με Γνώση του Περιβάλλοντος [14]

Εικόνα 10: Συστάσεις Πολλαπλών Εργασιών [14]

Εικόνα 11: Γονίδιο – Χρωμόσωμα – Πληθυσμός σε έναν Γενετικό Αλγόριθμο [18]

Εικόνα 12: Crossover [18]

Εικόνα 13: Mutation [18]

Εικόνα 14: Αναπαράσταση των Χρηστών στη Βάση Δεδομένων

Εικόνα 15: Αναπαράσταση των Δραστηριοτήτων στη Βάση Δεδομένων

ΚΕΦΑΛΑΙΟ 1 - ΕΙΣΑΓΩΓΗ

1.1 Εισαγωγή

Καθώς η επιστήμη της πληροφορικής και το διαδίκτυο εξελίσσονται, μεγαλώνει ο όγκος της πληροφορίας και των δεδομένων σε διάφορους τομείς όπως η υγεία(e-health), η εκπαίδευση(e-learning), η ψυχαγωγία(entertainment), ο τουρισμός κ.α. Τα δεδομένα μπορούν να κατηγοριοποιηθούν σε τρεις βασικές κατηγορίες : δομημένα, μερικώς δομημένα και μη δομημένα, βάζοντας έτσι σε πρόκληση τους προγραμματιστές να τα διαχειριστούν στα συστήματα που αναπτύσσουν. Τα συστήματα προτάσεων αντιμετωπίζουν την πρόκληση αυτή και παρέχουν εξατομικευμένες συστάσεις που ανταποκρίνονται στις ανάγκες και τις προτιμήσεις των χρηστών μέσα σε αυτόν των τεράστιο όγκο πληροφοριών.

Τα συστήματα προτάσεων έχουν παίξει καθοριστικό ρόλο σε τομείς όπως το ηλεκτρονικό και το λιανικό εμπόριο, Τέτοια λαμπρά παραδείγματα αποτελούν τα συστήματα που έχουν ενσωματώσει εταιρείες όπως η Amazon, η eBay, και η Flipkart, όπου οι υπηρεσίες προσαρμόζονται με τη χρήση συστημάτων προτάσεων για μια εξατομικευμένη εμπειρία αγορών, αναδεικνύοντας έτσι τη σημασία δημιουργίας ευφών συστημάτων.

Βασική αρχή των συστημάτων προτάσεων είναι να αναγνωρίζουν την εξάρτηση που υπάρχει μεταξύ χρήστη και προϊόντων. Για παράδειγμα ας θεωρήσουμε το Netflix, στο οποίο διαφορετικοί χρήστες παρακολουθούν διαφορετικά είδη ταινιών. Στη περίπτωση αυτή, ένας χρήστης που παρακολουθεί μια ταινία δράσης είναι πιο πιθανόν να παρακολουθήσει μια ακόμα ταινία δράσης και όχι μια κωμωδία. Σαν αυτό το παράδειγμα μπορούν να υπάρξουν διάφορες συσχετίσεις μεταξύ ενός χρήστη και ενός αντικειμένου που μπορούν να χρησιμοποιηθούν από ένα σύστημα προτάσεων.

1.2 Σκοπός

Σκοπός της παρούσας διπλωματικής είναι η ανάλυση των συστημάτων προτάσεων που υπάρχουν σήμερα και σε δεύτερη φάση η ανάλυση και κατανόηση τους συστήματος πρότασης δραστηριοτήτων που ενσωματώνεται στην εφαρμογή που έχει ως σκοπό να βελτιώσει την εμπειρία των τουριστών στην Ελλάδα.

1.3 Δομή

Η διπλωματική εργασία χωρίζεται σε 6 διακριτά κεφάλαια.

Το πρώτο κεφάλαιο αποτελεί την εισαγωγή όπου περιγράφεται η σημασία ύπαρξης συστημάτων προτάσεων στις μέρες μας, η δομή και ο σκοπός της εργασίας καθώς και οι συντομογραφίες που χρησιμοποιούνται.

Στο δεύτερο κεφάλαιο εξετάζεται η έρευνα που έχει γίνει πάνω στα συστήματα προτάσεων καθώς και το θεωρητικό υπόβαθρο αυτών.

Στο τρίτο κεφάλαιο περιγράφεται η μεθοδολογία που χρησιμοποιείται για την εύρεση του καταλληλότερου πλάνου δραστηριοτήτων για τον εκάστοτε χρήστη στην εφαρμογή «πρότασης πλάνου δραστηριοτήτων για τουρίστες στην Ελλάδα».

Στο τέταρτο κεφάλαιο αναλύεται η αρχιτεκτονική της παραπάνω εφαρμογής

Στο πέμπτο κεφάλαιο παρουσιάζονται και συγκρίνονται κάποια παραδείγματα χρήσης του συστήματος.

Στο έκτο κεφάλαιο αναφέρονται τα συμπεράσματα για το σύστημα αλλά και πιθανές μελλοντικές επεκτάσεις αυτού.

1.4 Συντομογραφίες

ΣΠ – Σύστημα Προτάσεων

RS – Recommendation System

CF – Collaborative Filtering

ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΚΑΙ ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

2.1 Θεωρητικό Υπόβαθρο

2.1.1 Εισαγωγή στα Συστήματα προτάσεων

Υπάρχουν δύο βασικά μοντέλα που μπορεί να ακολουθήσει κανείς για να κατασκευάσει ένα ΣΠ, ονομαστικά αυτά είναι το μοντέλο πρόβλεψης (prediction model) και το μοντέλο κατάταξης (ranking model).

Σε ένα τυπικό μοντέλο πρόβλεψης ενός ΣΠ, ένας πίνακας από χρήστες και αντικείμενα χρησιμοποιείται όπως φαίνεται στην εικόνα 1. Στην εικόνα αυτή, παρουσιάζεται ένα σύνολο από πέντε χρήστες και πέντε αντικείμενα, τα οποία βαθμολογούνται σε κλίμακα από 1 έως 5 (το 1 αντιστοιχεί στη χαμηλότερη και το 5 στην υψηλότερη βαθμολογία). Ο βασικός στόχος του ΣΠ είναι να προβλέψει όσο το δυνατό καλύτερα τις τιμές που είναι κενές (?).

	Π1	Π2	Π3	Π4	Π5
ΧΡΗΣΤΗΣ 1	?	3	2	4	5
ΧΡΗΣΤΗΣ 2	2	5	?	?	5
ΧΡΗΣΤΗΣ 3	3	?	1	3	2
ΧΡΗΣΤΗΣ 4	?	3	?	2	3
ΧΡΗΣΤΗΣ 5	5	?	?	1	2

Εικόνα 1

Σε ένα μοντέλο κατάταξης ενός ΣΠ, ο κύριος στόχος είναι να προσδιοριστούν τα πέντε καλύτερα αντικείμενα για έναν ή ακόμα και για όλους τους χρήστες. Ένα τέτοιο παράδειγμα παρουσιάζεται στην εικόνα 2. Εδώ αντί να λαμβάνεται υπόψη η βαθμολογία του χρήστη για ένα αντικείμενο, γίνεται καταγραφή των αντικειμένων που αγοράστηκαν από αυτόν.

ΧΡΗΣΤΕΣ	ΑΝΤΙΚΕΙΜΕΝΑ ΠΟΥ ΑΓΟΡΑΣΑΝ
ΧΡΗΣΤΗΣ 1	Π1,Π2
ΧΡΗΣΤΗΣ 2	Π3,Π4,Π5
ΧΡΗΣΤΗΣ 3	Π1,Π2,Π2
ΧΡΗΣΤΗΣ 4	Π3
ΧΡΗΣΤΗΣ 5	Π4,Π5

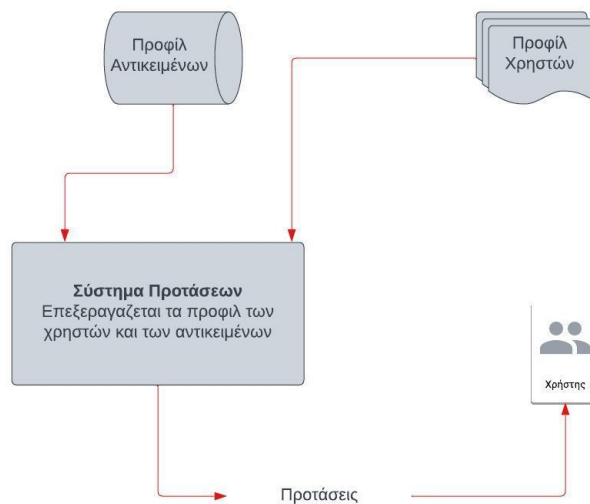
Εικόνα 2

2.1.2 Τρόποι Συλλογής δεδομένων από τους χρήστες

Για τη κατασκευή ενός ΣΠ μια πληροφορία όπως το αν ένα προϊόν αρέσει ή δεν αρέσει σε κάποιο χρήστη είναι πολύ βασική για τη κατασκευή του, έτσι ώστε αυτό με τη σειρά του να μπορεί να του προτείνει άλλα προϊόντα που μπορεί να τον ενδιαφέρουν. Πληροφορίες σαν τη παραπάνω μπορούν να συλλεχθούν από τους χρήστες με δύο τρόπους.

Ο πιο άμεσος τρόπος είναι μια πληροφορία που μπορεί να δώσει απευθείας ο ίδιος ο χρήστης για ένα προϊόν. Αποτελεί και το πιο δύσκολο τρόπο, για αυτό και οι διάφορες εταιρείες που ενσωματώνουν συστήματα προτάσεων στις υπηρεσίες τους δοκιμάζουν με διάφορους τρόπους να αποκτήσουν τέτοιου είδους πληροφορίες για αυτές. Μερικοί από αυτούς είναι τα μου αρέσει(like) ή δεν μου αρέσει(dislike) κουμπιά, οι αξιολογήσεις με αστεράκια ακόμα και οι αξιολογήσεις με απλά πλαίσια κειμένου(text box).

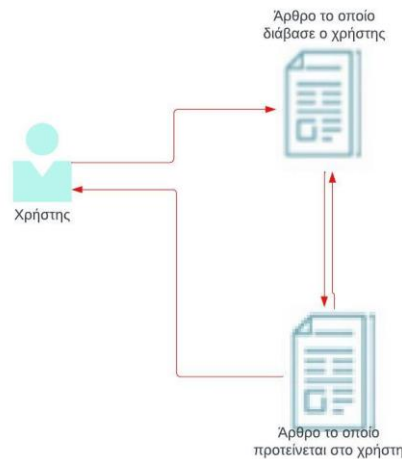
Ένας πιο έμμεσος τρόπος είναι οι πληροφορίες που μπορεί ένας χρήστης να αφήσει εν αγνοία του με τη δραστηριότητα του στο διαδίκτυο. Συγκεκριμένα οι ιστοσελίδες που έχει επισκεφθεί, τα προϊόντα που έχει δει, ο αριθμός επισκέψεων μιας ιστοσελίδας και γενικότερα ένα σύνολο από τις δραστηριότητες του κάθε χρήστη στο διαδίκτυο μπορεί να αναλυθεί στο ενδιαφέρον που έχει για κάποια προϊόντα. Τέτοιου είδους πληροφορίες γενικότερα καταγράφονται αυτόματα μέσω των μπισκότων(cookies) και του ιστορικού αναζήτησης χωρίς να χρειαστεί ο χρήστης να κάνει κάποια ενέργεια.



Εικόνα 3

2.1.3 Φιλτράρισμα Βάση Περιχομένου(Content-Based Filtering)

Η μέθοδος αυτή χρησιμοποιείται σε συστήματα τα οποία προτείνουν αντικείμενα παρόμοια με αυτά που ο χρήστης είτε είχε ήδη επιλέξει είτε είχε δείξει κάποιο ενδιαφέρον στο παρελθόν. Όπως φαίνεται και στην εικόνα 4 ένα νέο άρθρο προτείνεται στο χρήστη βασισμένο στο περιεχόμενο του άρθρου που είχε διαβάσει.



Εικόνα 4

Ένα ακόμα ενδιαφέρον παράδειγμα τέτοιου συστήματος που αξίζει να εξετασθεί σε βάθος είναι η πλατφόρμα του Netflix. Η πλατφόρμα αυτή, αποθηκεύει τη δραστηριότητα του κάθε χρήστη, σε ένα διάνυσμα. Ο πίνακας αυτός περιέχει πληροφορίες όπως τις ταινίες που έχει παρακολουθήσει ο χρήστης στο παρελθόν, ποιες από αυτές του άρεσαν, ποιες από αυτές δε του άρεσαν, τα είδη των ταινιών αυτών και άλλες τέτοιου είδους πληροφορίες. Ταυτόχρονα υπάρχει ένα ακόμα διάνυσμα στο οποίο βρίσκονται αποθηκευμένες πληροφορίες για όλες τις ταινίες που είναι διαθέσιμες στη πλατφόρμα. Οι πληροφορίες αυτές μπορεί να είναι ο τίτλος της ταινίας, οι ηθοποιοί της, η γλώσσα της, η διάρκειά της, η περίληψή της κ.α.

Το παραπάνω σύστημα για να βρει την ομοιότητα αυτών των δύο διανυσμάτων, υπολογίζει το συνημίτονο της μεταξύ τους γωνίας. Υποθέτοντας ότι A είναι το διάνυσμα του χρήστη και B το διάνυσμα των ταινιών, η ομοιότητα τους ισούται με το παρακάτω συνημίτονο.

$$\text{sim}(A, B) = \cos(\theta) = \frac{A \cdot B}{|A| |B|}$$

Το αποτέλεσμα κυμαίνεται πάντα από -1 έως 1, και αυτή η τιμή υπολογίζεται πολλές φορές κρατώντας σταθερό το διάνυσμα του χρήστη και παραμετροποιώντας το διάνυσμα των ταινιών. Στη συνέχεια οι ταινίες κατατάσσονται με βάση την ομοιότητα τους και έπειτα εμφανίζονται κατάλληλα στο χρήστη χρησιμοποιώντας έναν από τους δύο παρακάτω τρόπους.

- **N Καλύτερες (Top N)** – Σε αυτή τη προσέγγιση οι καλύτερες N ως προς τη βαθμολογία ταινίες, είναι αυτές που θα προταθούν στο χρήστη, όπου N μια ακεραία τιμή.
- **Κλίμακα Αξιολόγησης (Rating scale)** – Σε αυτή τη μέθοδο ορίζεται ένα κατώτατο όριο τιμής ομοιότητας. Όλες οι ταινίες που έχουν τιμή ομοιότητας μεγαλύτερη από αυτό το όριο θα προταθούν στο χρήστη.

Υπάρχουν και άλλοι μέθοδοι που χρησιμοποιούνται για τον υπολογισμό ομοιότητας οι οποίοι παρουσιάζονται παρακάτω.

Ευκλείδεια απόσταση είναι η απόσταση μεταξύ δύο σημείων η οποία μετρείται από την ευθεία που συνδέει τα δύο αυτά σημεία. Έτσι αν αναπαρασταθούν με κατάλληλο τρόπο ο χρήστης και τα αντικείμενα στο n -διάστατο χώρο, η ομοιότητα τους ισούται με τη μεταξύ τους απόσταση. Όσο πιο

κοντά βρίσκονται τα αντικείμενα αυτά στο χώρο, τόσο πιο όμοια είναι. Οπότε τα αντικείμενα που βρίσκονται πιο κοντά στο χρήστη, είναι και αυτά που θα του προταθούν. Ο παρακάτω τύπος υπολογίζει την ευκλείδεια απόσταση.

$$\text{Ευκλείδεια Απόσταση} = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

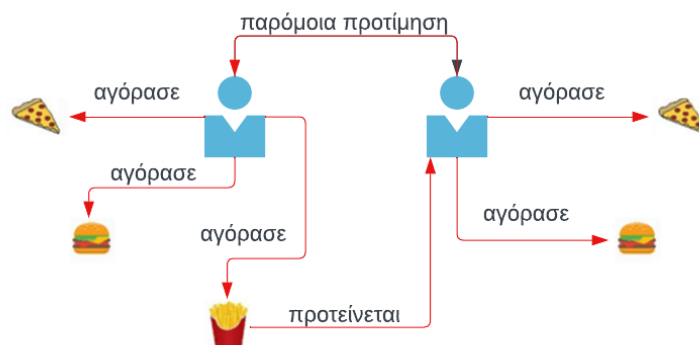
Η συσχέτιση Pearson αναφέρεται στο πόσο όμοια είναι δύο αντικείμενα. Όσο μεγαλύτερη είναι η συσχέτιση μεταξύ τους, τόσο μεγαλύτερη είναι και η ομοιότητα τους. Ο τύπος της συσχέτισης του Pearson παρουσιάζεται παρακάτω.

$$\text{sim}(u, v) = \frac{\sum(r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum(r_{ui} - \bar{r}_u)^2} \sqrt{\sum(r_{vi} - \bar{r}_v)^2}}$$

Το κύριο πρόβλημα ενός μηχανισμού φιλτραρίσματος που βασίζεται στο περιεχόμενο είναι πως όλες οι προτάσεις που δημιουργεί προς το χρήστη ανήκουν στην ίδια κατηγορία, κάνοντας έτσι την εμπειρία του χρήστη κάπως μονότονη. Εφόσον οι προτάσεις αυτές είναι βασισμένες στο τι άρεσε στο χρήστη ή τι έχει δει ο χρήστης στο παρελθόν, δεν πρόκειται ο χρήστης να έχει κάποια πρόταση για κάτι που δεν έχει εξερευνήσει στο παρελθόν. Σε ένα σύστημα ταινιών για παράδειγμα αν ο χρήστης βλέπει αποκλειστικά ταινίες μυστηρίου το ΣΠ θα του προτείνει μόνο ταινίες μυστηρίου. Για να αντιμετωπιστεί αυτό το ζήτημα, το ΣΠ δεν θα πρέπει απλώς να βασίζεται στο περιεχόμενο που αρέσει στον χρήστη, αλλά θα πρέπει επίσης να λαμβάνει υπόψη τη συμπεριφορά των χρηστών και τον τρόπο με τον οποίο άλλοι χρήστες με παρόμοια συμπεριφορά αλληλοεπιδρούν.

2.1.4 Συνεργατικό Φιλτράρισμα(Collaborative Filtering)

Ένα σύστημα που ενσωματώνει CF ψάχνει να βρει την όμοιους χρήστες, και όχι όμοια προϊόντα όπως γινόταν με τα συστήματα της προηγούμενης παραγράφου. Με απλά λόγια ένα CF προτείνει ένα προϊόν σε ένα χρήστη A με βάση τα ενδιαφέροντα ενός άλλου όμοιου χρήστη B. Στη παρακάτω εικόνα φαίνεται, πως λειτουργεί ένας τέτοιος μηχανισμός.



Εικόνα 5

Η ομοιότητα των χρηστών υπολογίζεται πάλι με τους ίδιους τρόπους που αναφέρθηκαν προηγουμένως. Κοιτάζοντας πάλι το παράδειγμα του Netflix υπάρχει πάλι ένα διάνυσμα για κάθε χρήστη ξεχωριστά, στο οποίο αποθηκεύονται πληροφορίες όπως ταινίες που έχει παρακολουθήσει στο

παρελθόν και του άρεσαν, οι κριτικές που έχει κάνει σε ταινίες, τα είδη ταινιών που παρακολουθεί πιο συχνά και άλλες τέτοιες πληροφορίες οι οποίες χρησιμοποιούνται για να συγκριθεί η ομοιότητα μεταξύ των χρηστών. Όταν το σύστημα βρει χρήστες με παρόμοια ενδιαφέροντα για έναν χρήστη, προτείνει σε αυτόν ταινίες που δεν έχει παρακολουθήσει ακόμη, αλλά άλλοι χρήστες με παρόμοια ενδιαφέροντα έχουν ήδη δει.

Αυτός ο τύπος φιλτραρίσματος είναι αρκετά δημοφιλής καθώς βασίζεται μόνο στη συμπεριφορά του χρήστη και δεν χρειάζεται κάποια επιπλέον πληροφορία. Χρησιμοποιείται ευρέως από μεγάλες εταιρείες συμπεριλαμβανομένου της Amazon, του Netflix και ενός American Express.

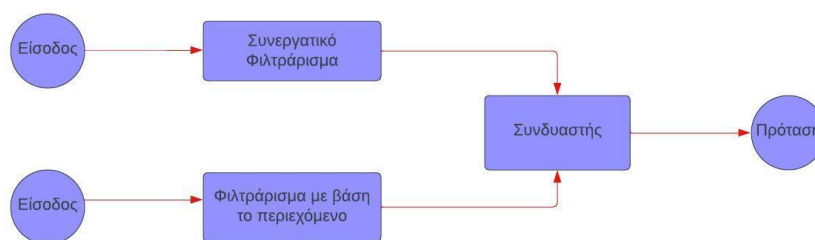
Τα συστήματα συνεργατικού φιλτραρίσματος χωρίζονται σε δύο κύριες κατηγορίες :

Στα χρήστη-χρήστη(user-user) CF, όπου οι προτάσεις για ένα χρήστη γίνονται με βάση τις προτιμήσεις που έχουν κάνει όμοιοι χρήστες στο παρελθόν. Παρόλο που αυτός ο τρόπος είναι αρκετά αποτελεσματικός, απαιτεί πολλούς πόρους από το σύστημα για να συγκρίνει τα δεδομένα μεταξύ των χρηστών. Ειδικά για συστήματα με μεγάλες βάσεις πελατών αυτός ο τρόπος είναι πολύ ακριβός.

Στα αντικείμενο-αντικείμενο(item-item) CF, οι προτάσεις γίνονται βάσει της ομοιότητας των αντικειμένων(προϊόντων) και όχι της ομοιότητας των χρηστών. Δημιουργείται ένα διάνυσμα για όλα τα αντικείμενα που ο χρήστης έχει επιλέξει στο παρελθόν και βάσει αυτού του πίνακα παρόμοια αντικείμενα θα του προταθούν. Αυτή η τεχνική είναι με διαφορά υπολογιστικά φθηνότερη της προηγούμενης καθώς το διάνυσμα των αντικειμένων παραμένει σχετικά αμετάβλητο σε μέγεθος. Το αρνητικό της μεθόδου αυτής είναι ότι δεν μπορεί να προτείνει ένα αντικείμενο το οποίο δεν έχει αξιολογήσει.

2.1.5 Υβριδικά Συστήματα(Hybrid Systems)

Μέχρι στιγμής αναλύθηκαν τα συστήματα που βασίζονται στο περιεχόμενο και τα συνεργατικά συστήματα, καθώς επίσης τα πλεονεκτήματα και μειονεκτήματα τους. Υπάρχουν όμως και συστήματα τα οποία συνδυάζουν και τους παραπάνω μηχανισμούς. Τα συστήματα αυτά ονομάζονται υβριδικά (Hybrid Systems) και μπορούν να εξαλείψουν το κύριο μειονέκτημα που έχουν και τα δύο προαναφερθέντα συστήματα σχετικά με τη δυσκολία τους να λειτουργήσουν στη περίπτωση όπου υπάρχει έλλειψη δεδομένων στη σχέση μεταξύ ενός χρήστη και ενός προϊόντος. Στη παρακάτω εικόνα παρουσιάζεται ο μηχανισμός ενός τέτοιου υβριδικού συστήματος.



Εικόνα 6

Τα υβριδικά συστήματα προτάσεων μπορούν να υλοποιηθούν με διάφορους τρόπους.

Ένας τρόπος είναι ο συνδυασμός των προτάσεων στο τέλος που θα δημιουργηθούν ξεχωριστά από ένα σύστημα βασισμένο στο περιεχόμενο και από ένα συνεργατικό σύστημα. Ένας άλλος τρόπος είναι η προσθήκη των δυνατοτήτων ενός συνεργατικού συστήματος σε ένα σύστημα με βάση το περιεχόμενο.

Σαφώς μπορεί να γίνει και το αντίστροφο, δηλαδή η προσθήκη ενός συστήματος βασισμένο στο περιεχόμενο σε ένα συνεργατικό σύστημα.

Πολλές μελέτες έχουν δείξει ότι τα υβριδικά συστήματα σε γενικές γραμμές έχουν καλύτερες επιδόσεις συγκριτικά με το αν τα συστήματα λειτουργούσαν το καθένα ξεχωριστά, καθώς επίσης οι προτάσεις που κάνουν είναι περισσότερο ακριβείς.

2.1.6 Συστήματα Γνώσης(Knowledge-Based RS)

Σε κάποιες περιπτώσεις, οι προτάσεις στο χρήστη δεν χρειάζεται να είναι βασισμένες στις αξιολογήσεις που έχει κάνει αυτός στο παρελθόν για ένα προϊόν ή ακόμα και στις πληροφορίες ενός προϊόντος. Τέτοια παραδείγματα αποτελούν η αγοραπωλησία ακινήτων, ο τουρισμός και η χρηματοοικονομικές υπηρεσίες. Σε αυτούς τους τομείς, οι μηχανισμοί δημιουργίας προτάσεων αντί να βασίζονται σε αξιολογήσεις και κριτικές, βασίζονται στη γνώση που έχουν για το χρήστη και έτσι οι προτάσεις που κάνουν εξελίσσονται συνεχώς στο πέρασμα του χρόνου. Προφανώς για έναν νέο χρήστη είναι δύσκολο ως και αδύνατο να χρησιμοποιηθεί ένα τέτοιο σύστημα, καθώς δεν υπάρχει καθόλου γνώση για αυτόν. Υπάρχουν δύο τεχνικές που ακολουθούνται για τη δημιουργία ενός τέτοιου συστήματος.

Η μια τεχνική είναι η δημιουργία ενός ΣΠ που ακολουθεί κανόνες χρησιμοποιώντας ένα καλά καθορισμένο σύνολο κανόνων ή φίλτρων, οι οποίοι αντιπροσωπεύουν τις ανάγκες του χρήστη. Αυτοί οι κανόνες αποθηκεύονται στη βάση γνώσης του συστήματος προτάσεων και με τη χρήση αυτών καθορίζεται ένα σύνολο αντικειμένων που ικανοποιούν αυτούς τους κανόνες.

Η δεύτερη τεχνική είναι τα συστήματα που χρησιμοποιούν αρχές της πολυπλοκότητας προσομοίωσης(Complex Adaptive Systems - CAS). Με αυτή τη τεχνική η λύση σε ένα πρόβλημα δίνεται από την αναζήτηση των λύσεων παρόμοιων προβλημάτων που έχουν επιλυθεί στο παρελθόν.

2.1.7 Δημογραφικά Συστήματα(Demographic RS)

Τέτοιου είδους συστήματα χρησιμοποιούν δημογραφικές πληροφορίες, δηλαδή προσωπικές πληροφορίες ενός χρήστη. Οι προτάσεις γίνονται με βάση της δημογραφικές ομοιότητες των χρηστών χωρίς να χρειάζεται κάποιο ιστορικό ενεργειών από το χρήστη.



Εικόνα 7

2.1.8 Τεχνητή Νοημοσύνη και Συστήματα προτάσεων

Στις μέρες, η τεχνητή νοημοσύνη έχει γίνει αναπόσπαστο κομμάτι στο τομέα της αυτοματοποίησης, ωστόσο και ο τομέας των συστημάτων προτάσεων δεν αποτελεί εξαίρεση. Οι μέθοδοι που βασίζονται στη μηχανική μάθηση είναι το μέλλον στα συστήματα προτάσεων καθώς όλο και περισσότερες εταιρείες αρχίζουν να προσαρμόζονται στον ερχομό της τεχνητής νοημοσύνης.

Οι μέθοδοι μηχανικής μάθησης είναι δύο: χωρίς επίβλεψη (unsupervised) και με επίβλεψη (supervised). Παρακάτω θα εξεταστεί μέθοδος μάθησης χωρίς επίβλεψη, που είναι η μέθοδος βασισμένη σε συστάδες (ML-clustering). Η τεχνική μάθησης χωρίς επίβλεψη χρησιμοποιεί αλγόριθμους ML για να εντοπίσει κρυμμένα μοτίβα στα δεδομένα προκειμένου να τα ομαδοποιήσει (unlabeled data). Η ομαδοποίηση σε συστάδες είναι η ομαδοποίηση παρόμοιων αντικειμένων σε συστάδες. Συνήθως, ένα αντικείμενο που ανήκει σε μια συστάδα είναι περισσότερο παρόμοιο με ένα αντικείμενο της ίδιας συστάδας παρά με ένα αντικείμενο που ανήκει σε μια άλλη συστάδα.

Στα συστήματα προτάσεων, η ομαδοποίηση χρησιμοποιείται για τον σχηματισμό ομάδων χρηστών που είναι παρόμοιοι μεταξύ τους. Μπορεί επίσης να ομαδοποιήσει παρόμοια αντικείμενα ή προϊόντα επίσης. Όπως αναφέρθηκε παραπάνω η ομαδοποίηση όμοιων χρηστών και αντικειμένων μέχρι στιγμής γινόταν χρησιμοποιώντας μαθηματικούς τύπους (ομοιότητα συνιμητόνου κ.α.), οι οποίοι είχαν διάφορα μειονεκτήματα όπως το υπολογιστικό τους κόστος και ότι χρειαζόταν ένας χρήστης να κάνει αρκετές αξιολογήσεις σε προϊόντα. Για αυτό μια προσέγγιση βασισμένη σε συστάδες συνήθως χρησιμοποιείται για να ομαδοποιηθούν παρόμοιοι χρήστες εξαλείφοντας αυτά τα ζητήματα. Εάν ένας χρήστης θεωρείται παρόμοιος με μια συστάδα χρηστών, αυτός ο χρήστης προστίθεται σε εκείνη τη συστάδα. Εντός της συστάδας, όλοι οι χρήστες μοιράζονται κοινά ενδιαφέροντα, και τους παρέχονται προτάσεις που είναι βασισμένες σε αυτούς.

2.1.9 Ομαδοποίηση (ML Classification)

Η συσταδοποίηση παρόλο αυτά έχει και τα μειονεκτήματα της. Για αυτό λύση αποτελεί η ένα ΣΠ βασισμένο στην ομαδοποίηση (classification-based recommendation system). Σε ένα τέτοιο σύστημα λαμβάνονται υπόψη χαρακτηριστικά και από τα αντικείμενα-προϊόντα και από τους χρήστες για να προβλεφθεί η συμπεριφορά του χρήστη. Μια εφαρμογή του εν λόγω συστήματος είναι το μοντέλο προδιάθεσης αγοραστή (buyer propensity model). Το μοντέλο αυτό προβλέπει τη πιθανότητα που έχει ένας χρήστης να αγοράσει ένα προϊόν ή ακόμα και τη πιθανότητα που έχει ένας χρήστης για τον αν θα γίνει πελάτης. Μερικοί περιορισμοί του συστήματος είναι ότι η συλλογή δεδομένων από διάφορους χρήστες και προϊόντα και ταυτόχρονα ο συνδυασμός αυτών, μερικές φορές είναι δύσκολο. Επίσης το να γίνει ομαδοποίηση όλων αυτών από μόνο του είναι κάτι απαιτητικό για το σχεδιαστή του συστήματος.

2.1.10 Βαθιά Μάθηση (Deep Learning)

Η βαθιά μάθηση (deep learning) είναι ένα κομμάτι της μηχανικής μάθησης, το οποίο είναι ισχυρότερο των ML αλγορίθμων και καταλήγει σε καλύτερα αποτελέσματα. Προφανώς, έχει και αυτό τους περιορισμούς του καθώς χρειάζεται μεγάλο όγκο δεδομένων και σωστή επεξήγηση για να λειτουργήσει σωστά, πράγμα το οποίο καλείται ο διαχειριστής του συστήματος να ξεπεράσει. Πολλές εταιρείες χρησιμοποιούν νευρωτικά δίκτυα (deep neural networks) για να βελτιώσουν την εμπειρία χρήσης των χρηστών-πελατών τους, ειδικά στις περιπτώσεις που υπάρχουν μη δομημένα δεδομένα

όπως εικόνες και κείμενα. Οι τρεις τύποι βαθιάς μάθησης (deep learning) συστημάτων προτάσεων είναι τα Restricted Boltzmann, Autoencoder based, Neural attention-based.

2.1.11 Οι δυσκολίες των Συστημάτων προτάσεων

Παρακάτω παρουσιάζονται οι δυσκολίες που υπάρχουν κατά την ανάπτυξη ενός οποιουδήποτε ΣΠ σχετικά με την απόδοση και την ακρίβειά τους.

Έλλειψη Δεδομένων

Τα ΣΠ χρησιμοποιούν παλαιότερα δεδομένα αξιολόγησης των χρηστών για τη δημιουργία νέων προτάσεων. Ωστόσο, οι χρήστες στη πλειοψηφία τους αποφεύγουν να παρέχουν αξιολογήσεις, πράγμα το οποίο δυσκολεύει τα συστήματα προτάσεων να παρέχουν προτάσεις.

Πρόβλημα Εκκίνησης (Cold Start)

Το πρόβλημα εκκίνησης σχετίζεται με την απουσία δεδομένων αξιολόγησης στους νέους χρήστες, κάνοντας τη δημιουργία προτάσεων σε αυτούς πολύ προβληματική.

Ανάλυση Συμπεριφοράς Χρήστη

Η ανάλυση των μοτίβων συμπεριφοράς των χρηστών χρησιμοποιώντας δεδομένα αξιολόγησης σε χρονική σειρά είναι κάτι πολύ σημαντικό για τη δημιουργία προτάσεων με ακρίβεια. Όπως είναι προφανές η αναγνώριση τέτοιων μοτίβων από αξιολογήσεις μεμονωμένων προϊόντων αποτελεί σίγουρα μια πρόκληση.

Ευαισθησία στο Χρόνο

Οι προτιμήσεις των χρηστών αλλάζουν με το χρόνο. Για παράδειγμα, κάθε εποχή του χρόνου υπάρχει ανάγκη για διαφορετικά προϊόντα από ένα διαδικτυακό κατάστημα. Είναι σημαντικό ένα ΣΠ να αναγνωρίζει τέτοιες αλλαγές μέσα στο χρόνο. Τα τρέχοντα ΣΠ αντιμετωπίζουν αυτή τη δυσκολία.

2.1.12 Αναπτυσσόμενες Τεχνικές στα Συστήματα προτάσεων

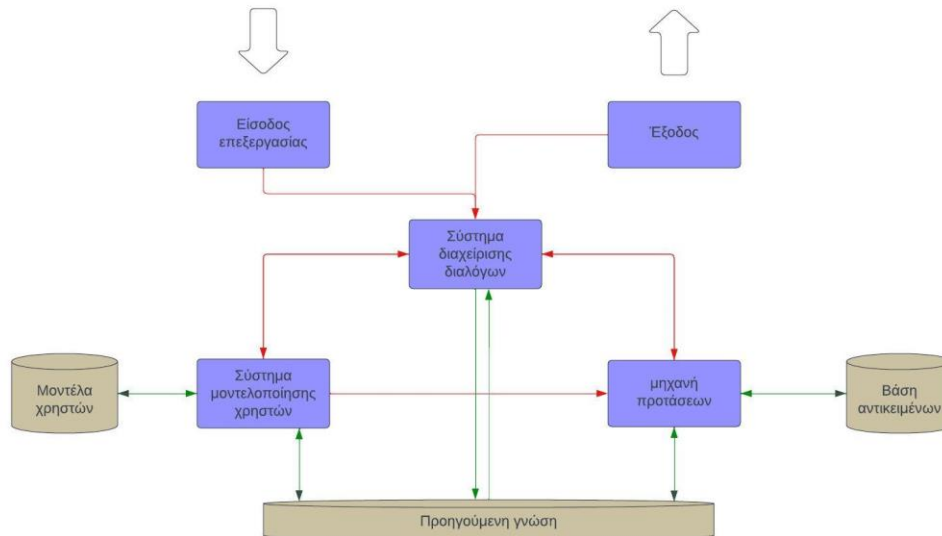
Συστάσεις Πραγματικού Χρόνου

Σε γενικές γραμμές, οι μαζικές συστάσεις (συστάσεις που γίνονται σε παρόμοιους χρήστες) είναι υπολογιστικά οικονομικές και προτιμώνται επειδή μπορούν να δημιουργούνται καθημερινά. Επίσης, είναι πολύ πιο απλές στη λειτουργία τους. Ωστόσο, πρόσφατα έχει τεθεί μεγαλύτερη έμφαση στην ανάπτυξη συστάσεων πραγματικού χρόνου. Οι συστάσεις πραγματικού χρόνου είναι γενικά ακριβότερες υπολογιστικά, διότι πρέπει να δημιουργούνται ξεχωριστά για κάθε χρήστη και βασίζονται στις αλληλεπιδράσεις των χρηστών σε πραγματικό χρόνο, πράγμα που τις καθιστά επίσης πολύπλοκες. Σημαντικό τους πλεονέκτημα είναι ότι βοηθούν στο να μη χάσει ο χρήστης το ενδιαφέρον του, καθώς οι προτάσεις του συστήματος θα αλλάζει σε πραγματικό χρόνο και επιπλέον ότι οι προτάσεις είναι περισσότερο ακριβής καθώς είναι ακόμα πιο εξατομικευμένες για το κάθε χρήστη.

Συστάσεις συνομιλίας

Τα τελευταία χρόνια, γίνονται προσπάθειες για τη δημιουργία πιο συνομιλητικών συστημάτων. Υπάρχει η πεποίθηση ότι αυτό θα ανατρέψει τον τρόπο που λαμβάνουν χώρα οι αλληλεπιδράσεις ανάμεσα σε ανθρώπους και υπολογιστές στο μέλλον. Τα συστήματα συνομιλητικών συστάσεων στοχεύουν στο να παράγουν προτάσεις από κείμενα ή προφορικές συνομιλίες, έτσι ώστε οι χρήστες να

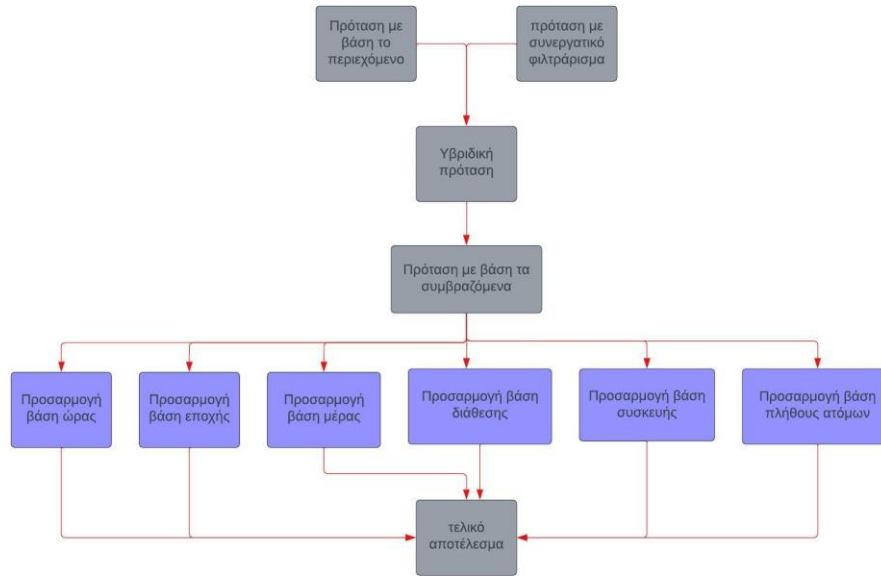
μπορούν να αλληλεπιδρούν με έναν υπολογιστή με όσο το δυνατό πιο φυσικό τρόπο. Πρόσφατα έχει γίνει εξαιρετικά δημοφιλές και χρησιμοποιείται εκτενώς σε φωνητικούς βοηθούς και σε μποτς συνομιλίας(chatbots). Χρησιμοποιεί τη φυσική κατανόηση γλώσσας (input) και τη γεννητική(output). Οι διάφορες ενέργειες που πρέπει να γίνουν αναλόγως τη συνομιλία πραγματοποιούνται με τη βοήθεια ενός συστήματος διαχείρισης συνομιλίας.



Εικόνα 8

Συστήματα συστάσεων με γνώση του περιβάλλοντος

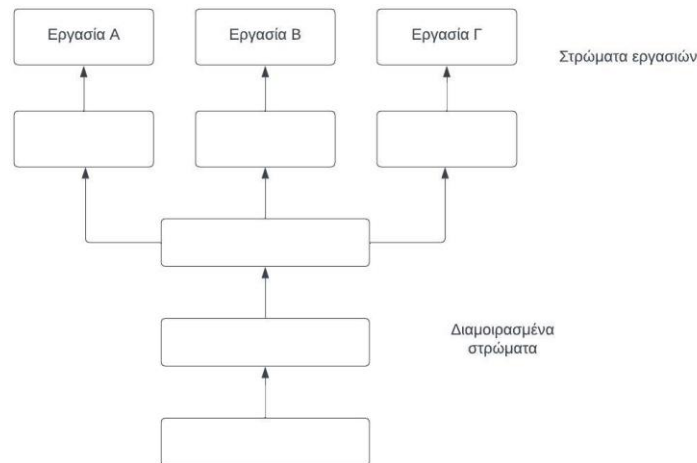
Οι ερευνητές και οι επαγγελματίες πληροφορικής αναγνωρίζουν τη σημασία των πληροφοριών του περιβάλλοντος σε πολλούς τομείς, όπως το ηλεκτρονικό εμπόριο, η ανάκτηση πληροφοριών, η εξόρυξη δεδομένων και το μάρκετινγκ. Παρόλο που έχει γίνει πολλή έρευνα στα συστήματα προτάσεων, πολλές υπάρχουσες προσεγγίσεις δεν λαμβάνουν υπόψη τις πληροφορίες του περιβάλλοντος, όπως είναι ο χρόνος, ο τόπος ή η παρουσία άλλων ατόμων στο χώρο για να προτείνουν πιο σχετικό περιεχόμενο. Υπάρχει μια αυξανόμενη τάση ότι οι συστάσεις πρέπει να λαμβάνουν υπόψη κάποιες πληροφορίες από το περιβάλλον και ότι είναι σημαντικό να τις λαμβάνουμε υπόψη κατά την παροχή συστάσεων. Τα συστήματα προτάσεων με γνώση του περιβάλλοντος αντιπροσωπεύουν ένα νέο τομέα πειραματισμού και έρευνας που στοχεύει στην παροχή πιο ακριβούς περιεχομένου βασισμένο στο περιβάλλον που βρίσκεται ο χρήστης οποιαδήποτε στιγμή. Για παράδειγμα, βρίσκεται ο χρήστης στο σπίτι ή κινείται; Χρησιμοποιεί μεγάλη ή μικρή οθόνη; Είναι πρωί ή βράδυ; Χρησιμοποιώντας κάθε στιγμή τα διαθέσιμα δεδομένα ενός συγκεκριμένου χρήστη, το σύστημα πρέπει να μπορεί να παρέχει συστάσεις που ο χρήστης είναι πιθανόν να χρειάζεται σε οποιοδήποτε από αυτά τα σενάρια.



Εικόνα 9

Συστάσεις πολλαπλών εργασιών

Σε πολλούς τομείς υπάρχουν πολλές σημαντικές πηγές για ανατροφοδότηση(feedback) που μπορούν να χρησιμοποιηθούν κατά την κατασκευή ενός ΣΠ. Για παράδειγμα, οι ιστοσελίδες ηλεκτρονικού εμπορίου καταγράφουν συνήθως τις επισκέψεις των χρηστών που έχει ένα συγκεκριμένο προϊόν, τα κλικ των χρηστών, τα καλάθια των χρηστών, και όλες τις αλληλεπιδράσεις που πραγματοποιούνται σε επίπεδο χρήστη και προϊόντος. Επίσης, καταγράφονται δεδομένα και μετά την αγορά όπως αξιολογήσεις και επιστροφές. Η ενοποίηση αυτών των διαφορετικών μορφών feedback είναι κρίσιμη για την κατασκευή συστημάτων που παρέχουν καλύτερα αποτελέσματα έναντι των μοντέλων που εξειδικεύονται σε μια μόνο κατηγορία ανατροφοδότησης(feedback). Ιδιαίτερα στις περιπτώσεις όπου κάποια δεδομένα δεν είναι επαρκής(αξιολογήσεις) ενώ κάποια άλλα είναι επαρκής με το παραπάνω(όπως τα κλικ). Σε αυτά τα σενάρια, ένα κοινό μοντέλο μπορεί να χρησιμοποιηθεί για να καλύψει το κενά στις ελλιπείς κατηγορίες χρησιμοποιώντας τη γνώση από τις άλλες κατηγορίες που έχουν παραπάνω δεδομένα, βελτιώνοντας έτσι τις προβλέψεις του συστήματος. Η μάθηση πολλαπλών εργασιών είναι μια προσέγγιση μηχανικής μάθησης, η οποία έχει χρησιμοποιηθεί με επιτυχία κυρίως στην επεξεργασία φυσικής γλώσσας και την όραση υπολογιστών. Τελευταία, η χρήση αυτής της μεθόδου για την κατασκευή αξιόπιστων συστημάτων προτάσεων έχει τραβήξει το ενδιαφέρον. Η κατασκευή νευρωτικών δικτύων βασισμένων στη μάθηση πολλαπλών εργασιών χρησιμοποιείται ευρέως λόγω των πολλαπλών της πλεονεκτημάτων όπως είναι η παροχή ερμηνεύσιμων αποτελεσμάτων κατανοώντας έτσι τις συστάσεις του συστήματος, καθώς επίσης και την επέκταση των δεδομένων αντιμετωπίζοντας έτσι το πρόβλημα της έλλειψης δεδομένων.



Εικόνα 10

2.2 Βιβλιογραφική Ανασκόπηση

Υπάρχουν διάφοροι αλγόριθμοι και προσεγγίσεις που παρέχουν προτάσεις στους χρήστες. Όλες αυτές οι τεχνικές επικεντρώνονται είτε στα ατομικά χαρακτηριστικά ενός ατόμου, είτε στα χαρακτηριστικά του προϊόντος ή της υπηρεσίας που αγοράζουν. Εκτός από αυτά, υπάρχει επίσης η δυνατότητα συνδυασμού αυτών των τεχνικών που εκμεταλλεύεται τόσο τα στοιχεία του χρήστη όσο και του προϊόντος, γνωστή και ως υβριδική τεχνική.

Πολλοί ερευνητές βασισμένοι σε διάφορες τεχνικές εργάστηκαν πάνω στην κατασκευή ενός συστήματος προτάσεων υψηλής ποιότητας. Ο Vanetti κ.ά. [1] διατύπωσαν ένα σύστημα προτάσεων βασισμένο στο Content Filtering, που προτείνει ένα στοιχείο σε έναν χρήστη με βάση την περιγραφή του στοιχείου και των ενδιαφερόντων του χρήστη που υπάρχουν στο προφίλ του. Σελίδες ιστού, τηλεοπτικά προγράμματα και σελίδες ειδήσεων χρησιμοποιούν αυτά τα συστήματα για προτάσεις. Ο Yang κ.ά. [2] σχεδίασαν ένα σύστημα προτάσεων βασισμένο στο CF, το οποίο λειτουργεί συλλέγοντας το feedback του χρήστη σε μορφή αξιολογήσεων για στοιχεία ενός συγκεκριμένου τομέα και στη συνέχεια ταιριάζει τους χρήστες που έχουν ομοιότητες στο τρόπο με τον οποίο αξιολογούν αυτά τα στοιχεία έτσι ώστε να τους κάνει τη κατάλληλη πρόταση. Η Τσικρίκα κ.ά. [3] εργάστηκαν σε μια υβριδική προσέγγιση συστήματος προτάσεων που λειτουργεί συνδυάζοντας πολλές τεχνικές προτάσεων μαζί. Πολλοί ερευνητές έχουν προσπαθήσει να συνδυάσουν CF συστήματα και συστήματα βασισμένα στο περιεχόμενο προκειμένου να εξομαλύνουν τα μειονεκτήματά του καθενός και να επιτύχουν καλύτερη απόδοση στις τελικές τους προτάσεις.

Ο Anil κ.ά. [4] στο άρθρο τους ανέφεραν διάφορες τεχνικές που χρησιμοποιούνται στα συστήματα προτάσεων, όπως το CF, φιλτράρισμα με βάση το περιεχόμενο (Content Filtering), κοινωνικό φιλτράρισμα (Social Filtering), δημογραφικό φιλτράρισμα (Demographic Filtering), φιλτράρισμα γνώσης (Knowledge Filtering). Εξέτασαν επίσης τα πλεονεκτήματα και τα μειονεκτήματά τους. Επίσης ασχολήθηκαν και με άλλους τομείς που δεν είχαν την απαραίτητη προσοχή από την επιστημονική κοινότητα όπως η τηλεόραση, τα εστιατόρια και την εύρεση εργασίας. Τέλος, κατέληξαν στο άρθρο τους ότι το φιλτράρισμα με βάση το περιεχόμενο και το CF έχουν καταφέρει να έχουν ευρεία αποδοχή και εκτεταμένη χρήση από την ερευνητική κοινότητα σε σχέση με άλλες προσεγγίσεις.

Οι Paul και Kundu [5] έκαναν μια έρευνα για το πως τα συστήματα προτάσεων που χρησιμοποιούν διάφορες γνωστές πλατφόρμες μουσικής όπως το Spotify, Apple Music, Pandora προτείνουν περιεχόμενο στους χρήστες τους. Σε γενικές γραμμές όλες οι πλατφόρμες χρησιμοποιούν κάποιο υβριδικό μοντέλο ή κάποιο μοντέλο που συνδυάζει μία ή περισσότερες τεχνικές. Αυτό που έχει αξία να σημειωθεί είναι ότι αυτές οι πλατφόρμες πέρα από τις κλασσικές τεχνικές φιλτραρίσματος που αναφέρονται παραπάνω, χρησιμοποιούν και κάποιες άλλες πρωτόγνωρες τεχνικές. Μια από αυτές είναι το φιλτράρισμα με βάση το συναίσθημα (Emotion-Based Filtering), η οποία όπως αναφέρει και το όνομα της αναλύει διάφορα χαρακτηριστικά ενός τραγουδιού, για να προσδιορίσει τα συναισθήματα που προκαλεί το τραγούδι αυτό στον χρήστη. Το μειονέκτημα αυτής της τεχνικής είναι ότι απαιτεί μεγάλα datasets και επίσης ότι δεν είναι τελείως ακριβής καθώς το ίδιο τραγούδι μπορεί να προκαλεί διαφορετικά συναισθήματα στο κάθε χρήστη. Μια ακόμα μέθοδος είναι αυτή που βασίζεται στο ιστορικό ακρόασης του εκάστοτε χρήστη, καθώς ένας χρήστης είναι πιο πιθανό να ακούει παρόμοια ή ακόμα και ίδια τραγούδια με αυτά που έχει ακούσει στο παρελθόν.

Κάποιοι άλλοι συγγραφείς [6],[7] εξετάζοντας τα συστήματα προτάσεων από μια πιο εμπορική προοπτική καταλήγουν στα παρακάτω συμπεράσματα. Ένα ΣΠ με προοπτική ευκολίας (Convenience Perspective RS) είναι μια προσέγγιση από τη πλευρά των πωλητών για τη σύσταση συχνά αγοραζόμενων προϊόντων λαμβάνοντας υπόψη μόνο την πιθανότητα αγοράς. Ένα ΣΠ με σκοπό τη συνεργασία (Collaborative Filtering RS) προτείνει προϊόντα μόνο βάσει της πιθανότητας αγοράς από παρόμοιους πελάτες. Ένα ΣΠ με σκοπό την ευκολία και την κερδοφορία (Convenience Plus Profitability RS) συνιστά προϊόντα με βάση την πιθανότητα αγοράς και την κερδοφορία του προϊόντος. Ένα υβριδικό ΣΠ κάνει προτάσεις συνδυάζοντας και τις δύο προοπτικές, τη συνεργασία και την κερδοφορία.

Το TechLens+[8] είναι ένα υβριδικό ΣΠ που συνδυάζει CF και φιλτράρισμα με βάση το περιεχόμενο (Content Filtering) για να προτείνει ερευνητικά άρθρα στους χρήστες. Μέσα από κάποια πειράματα, δείχνεται ότι το σύστημα παρέχει καλές προτάσεις. Ωστόσο, απαιτεί από τους χρήστες να δίνουν βαθμούς στα άρθρα.

Η Amazon[9] χρησιμοποιεί αλγόριθμους προτάσεων για να εξατομικεύσει το ηλεκτρονικό κατάστημα της για κάθε πελάτη ξεχωριστά. Η διαθέσιμες επιλογές αλλάζουν δραστικά με βάση τα ενδιαφέροντα του κάθε πελάτη. Η Amazon χρησιμοποιεί έναν CF σύστημα προτάσεων αντικειμένου προς αντικείμενο, αλλά απαιτεί επίσης από τους χρήστες να βαθμολογούν τα προϊόντα της.

Ο El Mabrouk κ.α.[10] πρότειναν ένα υβριδικό ΣΠ για πλατφόρμες διαδικτυακής μάθησης που χρησιμοποιεί εξόρυξη δεδομένων. Αυτό το σύστημα έχει στόχο να προτείνει το πιο κατάλληλο περιεχόμενο σε έναν χρήστη διαδικτυακής πλατφόρμας μάθησης και δίνει στους χρήστες τη δυνατότητα να επικεντρώνονται, κάνοντας το περιεχόμενο πιο προσβάσιμο. Επιπλέον, επειδή τα περισσότερα ΣΠ βασίζονται στην ανάλυση των προφίλ των μαθητών, το παραπάνω σύστημα έχει δημιουργηθεί με τέτοιο τρόπο έτσι ώστε να λαμβάνει υπόψη πολλούς παράγοντες κατά την παροχή προτάσεων .

Ο Herath κ.α. πρότειναν μια αρχιτεκτονική για ένα εξατομικευμένο ΣΠ μάθησης. Σε αυτή την υλοποίηση η ηλεκτρονική αξιολόγηση χρησιμοποιείται για να προσδιορίσει το επίπεδο κατανόησης του μαθητή. Το μοντέλο μάθησης, το μοντέλο τομέα (domain), το μοντέλο ηλεκτρονικής αξιολόγησης και το μοντέλο προτάσεων αποτελούν τα τέσσερα συστατικά του ΣΠ .

Ο Yanes κ.α. πρότειναν ένα ΣΠ βασισμένο στη μηχανική μάθηση για τη βελτίωση της μάθησης των φοιτητών. Δημιούργησαν ένα σύστημα που χρησιμοποιεί διάφορους αλγόριθμους για να προβλέπει το πιο κατάλληλο μαθησιακό περιεχόμενο για έναν συγκεκριμένο φοιτητή. Οι συγγραφείς κατέγραψαν δεδομένα από φοιτητές διάφορων τμημάτων.

Ο Fernandez κ.α. επικεντρώθηκαν στην ανάπτυξη ενός ΣΠ για να βοηθήσει τους φοιτητές να επιλέξουν μαθήματα. Οι συγγραφείς χρησιμοποίησαν μια υβριδική προσέγγιση προτάσεων που συνδυάζει τεχνικές βασισμένες στο περιεχόμενο και στο CF. Το σύστημα σχεδιάστηκε για να λαμβάνει υπόψη αρκετούς παράγοντες, συμπεριλαμβανομένου του ακαδημαϊκού υποβάθρου του φοιτητή, των ενδιαφερόντων του και της σχετικότητας του μαθήματος με το πτυχίο του.

Από τη βιβλιογραφική ανασκόπηση βγαίνει το συμπέρασμα ότι παρόλο που έχουν αναπτυχθεί πολλές τεχνικές δημιουργίας προτάσεων στο παρελθόν, καμία δεν είναι τέλεια καθώς όλες αυτές έχουν ένα ή περισσότερα προβλήματα. Δεν υπάρχει κάποια τεχνική που να ταιριάζει σε όλες τις περιπτώσεις, άρα ο καθένας πρέπει να προσαρμόσει το ΣΠ του στις δικές του ανάγκες.

ΚΕΦΑΛΑΙΟ 3 – ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΕΡΓΑΛΕΙΑ

3.1 Εισαγωγή

Στο κεφάλαιο αυτό, θα αναλυθεί η μεθοδολογία και τα εργαλεία που χρησιμοποιήθηκαν για το σχεδιασμό και την υλοποίηση της εφαρμογής της παρούσας εργασίας. Παρέχεται μια ολοκληρωμένη εικόνα για τα βήματα και τις διαδικασίες που ακολουθήθηκαν κατά την ανάπτυξή της, καθώς επίσης και τα εργαλεία που χρησιμοποιήθηκαν για αυτή.

3.2 Βήματα σχεδίασης του αλγορίθμου

Κάτι που μπορεί να αντιμετωπίσει οποιοσδήποτε, επισκεπτόμενος μια ξένη χώρα, είναι η άγνοια που θα νιώθει σχετικά με τις δραστηριότητες που θα πρέπει να κάνει σε αυτή τη χώρα. Από το ποια μέρη πρέπει να δει (αρχαιολογικά και μη), μέχρι το που θα πρέπει να φάει και που μπορεί να πάει να διασκεδάσει. Αυτή η απλή ανάγκη που μπορεί να έχει ο καθένας αποτέλεσε έμπνευση για τη δημιουργία μιας τέτοιας εφαρμογής που θα προτείνει σε τουρίστες ένα ολοκληρωμένο πλάνο δραστηριοτήτων (διαφόρων ειδών δραστηριότητες) εξατομικευμένο στο κάθε χρήστη. Η εφαρμογή βέβαια περιορίζεται μόνο στη πόλη της Αθήνας αλλά θα μπορούσε να επεκταθεί και για άλλες χώρες ή πόλεις.

Το κομμάτι της σχεδίασης του αλγορίθμου ήταν το πιο δύσκολο κομμάτι για την ανάπτυξη της εφαρμογής. Αρχικά το σύστημα είχε σχεδιαστεί έτσι ώστε να κάνει προτάσεις στο χρήστη μόνο με βάση κάποια φίλτρα αναζήτησης του χρήστη, τα οποία αντιπροσώπευαν τις επιθυμίες που είχε αυτός κάθε φορά που έκανε μια νέα αναζήτηση. Στη πορεία όμως, καθώς ο όγκος των εγγραφών της βάσης μεγάλωνε, το σύστημα προτάσεων έχανε τον ανθρωποκεντρικό του χαρακτήρα. Αυτό είχε ως αποτέλεσμα να παρουσιάζεται κάθε φορά διαφορετική λίστα προτάσεων, ακόμα και αν χρήστης ήταν ο ίδιος και είχε βάλει τα ίδια φίλτρα αναζήτησης. Για να αντιμετωπισθεί αυτή η αδυναμία ενσωματώθηκε στον αλγόριθμο η δυνατότητα να διαφοροποιεί τους χρήστες με βάση κάποια χαρακτηριστικά τους(θα αναλυθεί σε επόμενο κεφάλαιο),λαμβάνοντας υπόψη όμως σε μικρό βαθμό και τα φίλτρα αναζήτησης, τα οποία αντιπροσωπεύουν τις τωρινές επιθυμίες που έχει ο χρήστης για το επόμενο του ταξίδι στην Αθήνα.

Η δημιουργία ενός συστήματος προτάσεων από την αρχή αποτελεί μια πολύπλοκη διαδικασία, καθώς οι δυσκολίες που πρέπει να αντιμετωπιστούν είναι πολλές (βλ. Κεφάλαιο 2). Παρόλο που έχει γίνει αρκετή έρευνα σε αυτό το επιστημονικό πεδίο και έχουν δημιουργηθεί αμέτρητα συστήματα προτάσεων, μερικά εκ των οποίων αναλύθηκαν στη βιβλιογραφική ανασκόπηση στο κεφάλαιο 2, σε αυτή την εργασία προτιμήθηκε η δημιουργία ενός τέτοιου συστήματος από την αρχή. Το σύστημα αυτό βασίζεται στη λειτουργία των γενετικών αλγορίθμων, καθώς χρησιμοποιεί γενετικό αλγόριθμο με σωστά προσαρμοσμένη συνάρτηση καταλληλότητας έτσι ώστε να εξυπηρετεί τις ανάγκες ενός τέτοιου συστήματος.

Κρίσιμο σημείο για την δημιουργία του συστήματος αποτέλεσε η επιλογή του περιβάλλοντος και της γλώσσας ανάπτυξης της εφαρμογής. Επιλέχθηκε το Visual Studio για περιβάλλον ανάπτυξης καθώς επιτρέπει τη χρήση βιβλιοθηκών με πολύ εύκολο τρόπο καθώς επίσης προσφέρει ευκολία στην ανάπτυξη γραφικού περιβάλλοντος. Το γραφικό περιβάλλον έγινε με τη χρήση WPF(Windows Presentation Foundation) σε συνδυασμό με τη βιβλιοθήκη MahApp. Η επιλογή της γλώσσας C# έγινε καθαρά για το λόγο της προσωπικής μου εξοικείωσης με αυτή καθώς και για την άποψη συνεργασία

της με το περιβάλλον Visual Studio. Επίσης αποτελεί μια ισχυρή γλώσσα προγραμματισμού με πολλές δυνατότητες.

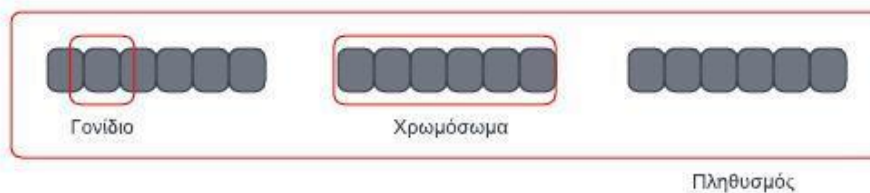
3.3 Γενετικοί Αλγόριθμοι

Οι γενετικοί αλγόριθμοι αποτέλεσαν ένα σημαντικό εργαλείο για τη δημιουργία του συστήματος της εργασίας, για αυτό το λόγο καλό θα ήταν να γίνει μια μικρή αναφορά για αυτούς.

Αρχικά, οι γενετικοί αλγόριθμοι είναι προσαρμοστικοί ευρετικοί αλγόριθμοι αναζήτησης που ανήκουν στο μεγαλύτερο μέρος των εξελικτικών αλγορίθμων. Οι γενετικοί αλγόριθμοι βασίζονται στις ιδέες της φυσικής επιλογής και της γενετικής. Πρόκειται για έξυπνη εκμετάλλευση της τυχαίας αναζήτησης που παρέχεται με ιστορικά δεδομένα για να κατευθύνει την αναζήτηση στην περιοχή καλύτερης απόδοσης στο χώρο λύσεων. Χρησιμοποιούνται συνήθως για τη δημιουργία λύσεων υψηλής ποιότητας για προβλήματα βελτιστοποίησης και προβλήματα αναζήτησης. Οι γενετικοί αλγόριθμοι προσομοιώνουν τη διαδικασία της φυσικής επιλογής, που σημαίνει ότι εκείνα τα είδη που μπορούν να προσαρμοστούν στις αλλαγές στο περιβάλλον τους, μπορούν να επιβιώσουν και να αναπαραχθούν και να πάνε στην επόμενη γενιά. Με απλά λόγια, προσομοιώνουν την «επιβίωση του ισχυρότερου» μεταξύ ατόμων διαδοχικής γενιάς για την επίλυση ενός προβλήματος. Κάθε γενιά αποτελείται από έναν πληθυσμό ατόμων και κάθε άτομο αντιπροσωπεύει ένα σημείο στον χώρο αναζήτησης και μια πιθανή λύση. Κάθε άτομο αναπαρίσταται ως μια συμβολοσειρά char/int/float/bit. Αυτή η συμβολοσειρά είναι ανάλογη με το χρωμόσωμα.

Χώρος Αναζήτησης

Ο πληθυσμός των ατόμων διατηρείται εντός του χώρου αναζήτησης. Κάθε άτομο αντιπροσωπεύει μια λύση στο χώρο αναζήτησης για δεδομένο πρόβλημα. Κάθε άτομο κωδικοποιείται ως πεπερασμένο διάνυσμα μήκους συστατικών. Αυτά τα μεταβλητά συστατικά είναι ανάλογα με τα γονίδια. Έτσι ένα χρωμόσωμα (άτομο) αποτελείται από πολλά γονίδια (μεταβλητά συστατικά).



Εικόνα 11

Βαθμολογία φυσικής κατάστασης(Fitness Score)

Μια βαθμολογία δίνεται σε κάθε άτομο που δείχνει την ικανότητα ενός ατόμου να «ανταγωνίζεται». Αναζητείται το άτομο που έχει τη βέλτιστη βαθμολογία φυσικής κατάστασης (ή σχεδόν βέλτιστη). Ο γενετικός αλγόριθμος διατηρεί ένα πληθυσμό n ατόμων (χρωμόσωμα/λύματα) μαζί με τη βαθμολογία φυσικής κατάστασής(fitness score) τους. Στα άτομα που έχουν καλύτερη βαθμολογία

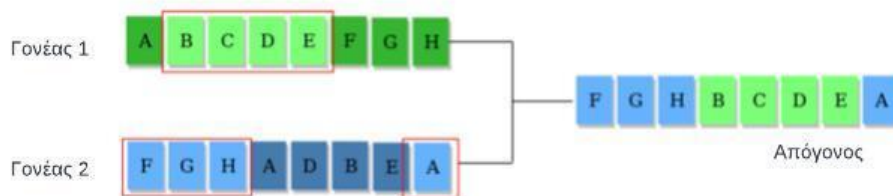
φυσικής κατάστασης δίνονται περισσότερες πιθανότητες να αναπαραχθούν από άλλα. Επιλέγονται τα άτομα με καλύτερη βαθμολογία, τα οποία ζευγαρώνουν και παράγουν καλύτερους απογόνους συνδυάζοντας τα χρωμοσώματα των γονέων. Το μέγεθος του πληθυσμού είναι στατικό, έτσι μερικά άτομα πεθαίνουν και αντικαθίστανται από νέες άτομα, δημιουργώντας τελικά μια νέα γενιά. Στόχος είναι να έρθουν καλύτερες λύσεις με τις διαδοχικές γενιές. Κάθε νέα γενιά έχει κατά μέσο όρο περισσότερα «καλύτερα γονίδια» από το άτομο (λύση) προηγούμενων γενεών. Έτσι, κάθε νέα γενιά έχει καλύτερες «μερικές λύσεις» από τις προηγούμενες γενιές. Μόλις οι απόγονοι που παρήχθησαν δεν έχουν σημαντική διαφορά από τους απογόνους που παρήχθησαν από προηγούμενους πληθυσμούς, ο πληθυσμός συγκλίνει. Ο αλγόριθμος λέγεται ότι συγκλίνει σε ένα σύνολο λύσεων για το πρόβλημα.

Τελεστής Επιλογής

Η ιδέα εδώ είναι να δοθεί προτίμηση στα άτομα με καλή βαθμολογία φυσικής κατάστασης και αυτά με τη σειρά τους να περάσουν τα γονίδιά τους σε διαδοχικές γενιές.

Τελεστής Crossover

Αντιπροσωπεύει το ζευγάρωμα μεταξύ ατόμων. Δύο άτομα επιλέγονται χρησιμοποιώντας τον τελεστή επιλογής και τα σημεία διασταύρωσης επιλέγονται τυχαία. Στη συνέχεια, τα γονίδια σε αυτές τις θέσεις διασταύρωσης ανταλλάσσονται δημιουργώντας έτσι ένα εντελώς νέο άτομο (απόγονο).



Εικόνα 12

Τελεστής Μετάλλαξης (Mutation)

Η βασική ιδέα είναι να εισαχθούν τυχαία γονίδια στους απογόνους για να διατηρηθεί η ποικιλομορφία στον πληθυσμό για να αποφευχθεί η πρόωρη σύγκλιση.



Εικόνα 13

Συνοψίζοντας ο αλγόριθμος λειτουργεί ως εξής :

- 1) Τυχαία αρχικοποίηση πληθυσμού p
- 2) Προσδιορισμός της καταλληλότητας του πληθυσμού
- 3) Έως ότου επαναληφθεί η σύγκλιση:
 - α) Επιλογή γονέων από τον πληθυσμό
 - β) Διασταύρωση και δημιουργία νέου πληθυσμού
 - γ) Εκτέλεση μετάλλαξης σε νέο πληθυσμό
 - δ) Υπολογισμός καταλληλότητας για νέο πληθυσμό

ΚΕΦΑΛΑΙΟ 4 – ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΛΟΓΙΚΗ ΣΥΣΤΗΜΑΤΟΣ

4.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα συζητηθούν περισσότερο τεχνικές λεπτομέρειες και χαρακτηριστικά της εφαρμογής καθώς επίσης θα παρουσιαστούν κομμάτια κώδικα του αλγορίθμου. Επίσης θα παρουσιαστεί και θα εξηγηθεί η δομή του συστήματος.

4.2 Αναπαράσταση των χρηστών στη Βάση Δεδομένων

Όταν δημιουργείται ένας νέος χρήστης ζητούνται από αυτόν να συμπληρώσει μια φόρμα με κάποια βασικά στοιχεία όπως username, password, email, gender. Αυτά τα στοιχεία δεν χρησιμοποιούνται από τον αλγόριθμο για τη δημιουργία πρότασης πλάνου δραστηριοτήτων. Υπάρχει ένα πεδίο όμως που ονομάζεται visited countries, το οποίο αντιστοιχεί στις χώρες που έχει ήδη ταξιδέψει ο χρήστης (καθώς υπάρχει και η επιλογή none). Το πεδίο αυτό είναι αρκετά σημαντικό καθώς ο αλγόριθμος το λαμβάνει υπόψη κατά τη διαδικασία δημιουργίας προτάσεων. Στην εικόνα 4.1 παρουσιάζεται η αναπαράσταση των χρηστών στη βάση δεδομένων.

ID	Username	Password	Email	Gender	Region	VisitedCountries	
14	bill	bill	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	bill@gmail.com	Male	Africa	France, Italy,
15	alex	alex	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	alex@gmail.com	Male	America	France, Italy,
NULL	NULL	NULL	NULL	NULL	NULL	NULL	

Εικόνα 14

Η στήλη ID είναι ένας auto increment πεδίο και είναι primary key του πίνακα αυτού, κάνοντας έτσι το κάθε ID μοναδικό για κάθε χρήση.

4.3 Αναπαράσταση των δραστηριοτήτων στη Βάση Δεδομένων

Κάτι σημαντικό για τον αλγόριθμο είναι το πλήθος και ο τρόπος αναπαράστασης των δραστηριοτήτων στη βάση δεδομένων. Στην εικόνα 4.2 παρουσιάζεται ο πίνακας των δραστηριοτήτων.

ID	Category	ActivityName	Image	ActivityDescription	ActivityType
1	Museum	Acropolis Museum	C:\ProjectApp\DBimages\acropolis_museum.jpg	Modern, up-to-date building on the foot of the Ac...	History
2	Museum	Benaki Museum	C:\ProjectApp\DBimages\benaki_museum.jpg	The Benaki Museum is among the most extensiv...	Art
3	Museum	Museum of Cycladic Art	C:\ProjectApp\DBimages\museum_of_cycladic...	The Museum of Cycladic Art focuses on promoti...	Art
4	Museum	National Archaeological Museum	C:\ProjectApp\DBimages\national_archaeologic...	The National Archaeological Museum of Athens i...	History
5	Sightseeing	Acropolis	C:\ProjectApp\DBimages\acropolis.jpg	This ancient citadel, built in the fifth century B...	
6	Sightseeing	Mount Lycabettus	C:\ProjectApp\DBimages\lycabettus.jpg	Walk or take the funicular tramway to the top o...	
7	Sightseeing	Panathenaic Stadium	C:\ProjectApp\DBimages\panathenaic.jpg	The Panathenaic Stadium dates back to the 4th ...	
8	Sightseeing	Plaka	C:\ProjectApp\DBimages\plaka.jpg	Wander through the narrow cobblestone alleys ...	
9	Sightseeing	Monastiraki	C:\ProjectApp\DBimages\monastiraki.jpg	Lively flea market neighborhood with lots of sm...	
10	Restaurant	Dinner in the Sky	C:\ProjectApp\DBimages\dinner_in_the_sky.jpg	Mediterranean, European, Greek cuisine €€€€	Greek
11	Restaurant	A Little Taste Of Home	C:\ProjectApp\DBimages\a_little_taste_of_hom...	Mediterranean, European, Turkish cuisine €€€	Turkish
12	Restaurant	Μπαϊρακτάρης	C:\ProjectApp\DBimages\bairaktaris.jpg	Mediterranean, BBQ cuisine €	Greek
13	Entertainm...	The Clu...	C:\ProjectApp\DBimages\the_clu...	Cocktail Bar	€

Εικόνα 15

Στο πίνακα αυτό υπάρχει η στήλη ID, η οποία είναι auto increment και primary key, κάνοντας έτσι τη κάθε δραστηριότητα μοναδική. Έπειτα έχει γίνει κατηγοριοποίηση δύο επιπέδων για κάθε δραστηριότητα. Στο πρώτο επίπεδο εξετάζεται το είδος της δραστηριότητας και αναφέρεται στη στήλη Category. Τα είδη είναι τέσσερα και ονομαστικά είναι museum, sightseeing, restaurant, entertainment. Στο δεύτερο επίπεδο κατηγοριοποίησης εξετάζεται το είδος της δραστηριότητας με έναν πιο ειδικό όρο και αναπαρίσταται στη στήλη ActivityType. Για παράδειγμα όπως φαίνεται στην εικόνα παραπάνω ένα μουσείο μπορεί να διαφέρει με ένα άλλο μουσείο βάση του περιεχομένου των εκθεμάτων του. Έτσι τα μουσεία μπορεί να κατηγοριοποιηθούν σε μουσεία ιστορίας, μουσεία τέχνης, μουσεία τεχνολογίας κ.α. Αντίστοιχα στα εστιατόρια έχει γίνει κατηγοριοποίηση με βάση τη κουζίνα του καθενός εστιατορίου, έτσι έχουμε ελληνικά εστιατόρια, γαλλικά εστιατόρια, ιταλικά εστιατόρια κ.α. Ωστόσο, όπως φαίνεται και από τη εικόνα δεν υπάρχει δεύτερου επιπέδου κατηγοριοποίηση στα αξιοθέατα καθώς θεωρήθηκε περιττή.

Τέλος, υπάρχουν άλλες δύο στήλες στο πίνακα των δραστηριοτήτων οι οποίες δεν έχουν κάποια σημασία για το σύστημα πρότασης της εφαρμογής καθώς αναπαριστούν τη περιγραφή της κάθε δραστηριότητας και το path της εικόνας κάθε δραστηριότητας (ActivityDescription, Image).

4.4 Ο αλγόριθμος του συστήματος προτάσεων

Παρακάτω παρουσιάζεται το αρχείο Test1.cs της εφαρμογής, το οποίο περιέχει τον αλγόριθμο του συστήματος προτάσεων.

Κατασκευαστής της κλάσης

Αρχικά παρουσιάζεται ο ψευδοκώδικας του κατασκευαστή της κλάσης Test1 και έπειτα η υλοποίηση της σε C#

ΚΛΑΣΗ Test1

Μέθοδος Αρχικοποίησης(ΟλικόΜέγεθοςΓονιδίων, πιθανότητα_A, πιθανότητα_B, πιθανότητα_C, πιθανότητα_D, ταυτότητα_χρήστη)

1. Αρχικοποίηση της ταυτότητας του χρήστη (ID) με το ταυτότητα_χρήστη
2. Δημιουργία νέου αντικειμένου τύπου Random
3. Κλήση της μεθόδου ConnectToDB για σύνδεση με τη βάση δεδομένων
4. Αρχικοποίηση των μεταβλητών ΟλικόΜέγεθοςΓονιδίων, πιθανότητα_A, πιθανότητα_B, πιθανότητα_C, πιθανότητα_D
5. Δημιουργία νέου αντικειμένου GeneticAlgorithm με παραμέτρους populationSize, ΟλικόΜέγεθοςΓονιδίων, random, GetRandomElement, FitnessFunction, elitism, mutationRate
6. Δημιουργία και αρχικοποίηση του χρονοδιακόπτη (DispatcherTimer) με ένα διάστημα 100ms
7. Σύνδεση του χρονοδιακόπτη με τη συνάρτηση UpdateTimer_Tick και εκκίνηση του χρονοδιακόπτη

ΤΕΛΟΣ ΚΛΑΣΗΣ

```

1. public Test1(int TotalGenesSize, double propability_A, double propability_B, double
propability_C, double propability_D,int usr_id)
2. {
3.     ID = usr_id;
4.     random = new System.Random();
5.     ConnectToDB();
6.     this.TotalGenesSize = TotalGenesSize;
7.     this.propability_A = propability_A;
8.     this.propability_B = propability_B;
9.     this.propability_C = propability_C;
10.    this.propability_D = propability_D;
11.    ga = new GeneticAlgorithm<Dictionary<string,string>>(populationSize, TotalGenesSize,
random, GetRandomElement, FitnessFunction, elitism, mutationRate);
12.    updateTimer = new DispatcherTimer();
13.    updateTimer.Interval = TimeSpan.FromMilliseconds(100);
14.    updateTimer.Tick += UpdateTimer_Tick;
15.    updateTimer.Start();
16.    enabled = true;
17. }

```

Στο κατασκευαστή της κλάσης γίνεται αρχικοποίηση κάποιων βασικών μεταβλητών του αλγορίθμου, οι τιμές των οποίων δίνονται κατά τη δημιουργία του αντικειμένου. Αυτές οι μεταβλητές είναι το id του χρήστη, οι τιμές από τα φίλτρα αναζήτησης του χρήστη, καθώς και το πλήθος των προτάσεων που θα παρουσιαστούν στο χρήστη, το πλήθος των οποίων προκύπτει από τις μέρες επίσκεψης του χρήστη στην Αθήνα. Επίσης γίνεται αρχικοποίηση κάποιων βασικών μεταβλητών του γενετικού αλγορίθμου.

Σύνδεση με τη Βάση Δεδομένων

Αρχικά παρέχεται ο ψευδοκώδικας για τη συνάρτηση που υλοποιεί τη σύνδεση με τη βάση δεδομένων και έπειτα η υλοποίηση του σε C#

ΣΥΝΑΡΤΗΣΗ ConnectToDB()

1. Δημιουργία σύνδεσης με τη βάση δεδομένων με χρήση των παραμέτρων σύνδεσης
2. Άνοιγμα της σύνδεσης
3. Ορισμός ερωτήματος SQL για την εξαγωγή των δεδομένων των δραστηριοτήτων (ActivityName, Category, ActivityType)
4. Εκτέλεση του ερωτήματος SQL και χρήση του αποτελέσματος για την αρχικοποίηση των λεξικών dataDictionary και activityTypeD
5. Ορισμός δεύτερου ερωτήματος SQL για την εξαγωγή των επισκεπτόμενων χωρών από την πίνακα χρηστών βάσει του δοθέντος ID
6. Εκτέλεση του δεύτερου ερωτήματος SQL και αποθήκευση του αποτελέσματος στη μεταβλητή visitedC
7. Κλείσιμο της σύνδεσης

ΤΕΛΟΣ ΣΥΝΑΡΤΗΣΗΣ

```

1. public void ConnectToDB()
2. {
3.     string connectionString =
"Server=localhost;Port=3306;Database=sys;Uid=root;Pwd=1234567890;";

4.     using (MySQLConnection connection = new MySQLConnection(connectionString))
5.     {
6.         connection.Open();
7.         string query = "SELECT ActivityName,Category, ActivityType FROM activities";
8.         using (MySQLCommand command = new MySQLCommand(query, connection))
9.         {
10.            using (MySQLDataReader reader = command.ExecuteReader())
11.            {
12.                while (reader.Read())
13.                {
14.                    dataDictionary[reader.GetString(0)] =
reader.GetString(1);
15.                    activityTypeD[reader.GetString(0)] =
reader.GetString(2);
16.                }
17.            }
18.        }
19.        string query2 = "SELECT VisitedCountries FROM users WHERE ID = @id";
20.        using (MySQLCommand command = new MySQLCommand(query2, connection))
21.        {
22.            command.Parameters.AddWithValue("@id", ID);
23.            using (MySQLDataReader reader = command.ExecuteReader())
24.            {
25.                while (reader.Read())
26.                {
27.                    visitedC = reader.GetString(0);
28.                }
29.            }
30.            connection.Close();
31.        }
32.    }
33. }
34. }
35.

```

Σε αυτό το σημείο γίνεται σύνδεση με τη βάση δεδομένων έτσι ώστε να αποθηκευτούν σε τοπικό επίπεδο οι στήλες ActivityName, Category, ActivityType του πίνακα activities. Η αποθήκευση των στηλών αυτών γίνεται σε δομή dictionary για λόγους ευκολίας και απόδοσης. Επίσης μια δεύτερη κλήση στη βάση δεδομένων λαμβάνει χώρα για την αποθήκευση της στήλης visitedC του πίνακα users. Η στήλη αυτή περιέχει τις χώρες που έχει επισκεφθεί ο χρήστης στο παρελθόν.

Συνάρτηση GetRandomElement()

Αρχικά παρέχεται ο ψευδοκώδικας για τη συνάρτηση GetRandomElement() και στη συνέχεια παρέχεται η υλοποίησή της σε C#

ΣΥΝΑΡΤΗΣΗ GetRandomElement()

1. Δημιουργία ενός ακέραιου i , ο οποίος είναι ένας τυχαίος αριθμός μεταξύ 0 και του αριθμού των κλειδιών στο dataDictionary

2. Δημιουργία μιας λίστας `keysList` που περιέχει όλα τα κλειδιά από το `dataDictionary`
3. Επιλογή του κλειδιού `key` από τη λίστα `keysList` βάσει του δείκτη `i`
4. Δημιουργία ενός νέου λεξικού `result`
5. Προσθήκη του ζεύγους κλειδιού-τιμής στο λεξικό `result`, όπου το κλειδί είναι το επιλεγμένο `key` και η τιμή είναι η τιμή από το `dataDictionary` για αυτό το κλειδί
6. Επιστροφή του λεξικού `result`

ΤΕΛΟΣ ΣΥΝΑΡΤΗΣΗΣ

```

1. private Dictionary<string,string> GetRandomElement()
2. {
3.     int i = random.Next(dataDictionary.Keys.Count);
4.     List<string> keysList = dataDictionary.Keys.ToList();
5.     string key = keysList[i];
6.     Dictionary<string,string> result = new Dictionary<string,string>();
7.     result.Add(key, dataDictionary[key]);
8.     return result;
9. }
10.

```

Η συνάρτηση αυτή χρησιμοποιείται από το γενετικό αλγόριθμο για την αρχικοποίηση του αρχικού πληθυσμού. Συγκεκριμένα επιστρέφει τυχαία στοιχεία από το dictionary των δραστηριοτήτων.

Συνάρτηση `UpdaterTime_Tick()`

Αρχικά παρέχεται ο ψευδοκώδικας για τη συνάρτηση `UpdateTimer_Tick()` και στη συνέχεια παρέχεται η υλοποίηση της σε C#

ΣΥΝΑΡΤΗΣΗ `UpdateTimer_Tick(sender, e)`

1. Δημιουργία νέας γενιάς στο αντικείμενο `ga`
2. Έλεγχος αν ο αριθμός των γενεών στο αντικείμενο `ga` έχει ξεπεράσει το 200
 - α. Αν ναι, τότε σταμάτησε τον χρονοδιακόπτη (`updateTimer`)
 - β. Δημιουργία μιας νέας λίστας `tmp` για προσωρινή αποθήκευση των κλειδιών
 - γ. Για κάθε στοιχείο μέχρι το συνολικό μέγεθος των γονιδίων (`TotalGenesSize`):
 - i. Εξαγωγή της λίστας των κλειδιών από τα καλύτερα γονίδια της *i*-οστής θέσης από το `ga` και αποθήκευση της στην λίστα `tmp`
 - ii. Προσθήκη της λίστας `tmp` στη λίστα αποτελεσμάτων (`result`)

ΤΕΛΟΣ ΣΥΝΑΡΤΗΣΗΣ

```

1. private void UpdateTimer_Tick(object sender, EventArgs e)
2. {
3.
4.     ga.NewGeneration();
5.     if (ga.Generation > 200)
6.     {
7.         updateTimer.Stop();
8.         List<string> tmp = new List<string>();
9.         for (int i = 0; i < TotalGenesSize; i++)

```

```

10.         {
11.             tmp = ga.BestGenes[i].Keys.ToList();
12.             result.Add(tmp);
13.         }
14.     }
15.
16.
17. }
18.

```

Η παραπάνω συνάρτηση περιέχει δύο σημαντικές λειτουργίες για τον γενετικό αλγόριθμο. Αρχικά καλεί τη συνάρτηση `NewGeneration()` του αλγόριθμου, η οποία δημιουργεί νέα γενιά πληθυσμού για το γενετικό αλγόριθμο. Επίσης η συνάρτηση αυτή, ελέγχει αν το σύστημα πληροί τη τερματική συνθήκη(αριθμός γενεών μεγαλύτερος του 200), έτσι ώστε να σταματήσει να τρέχει ο γενετικός αλγόριθμος. Για να επιτευχθούν τα παραπάνω, η συνάρτηση αυτή καλείται ανά 100ms. Τέλος, αξίζει να αναφερθεί ότι ο αριθμός 200 στη τερματική συνθήκη έχει επιλεγεί, διότι μετά από διάφορες δοκιμές παρατηρήθηκε ότι περίπου στις 200 γενιές ο αλγόριθμος βρίσκει τη βέλτιστη λύση.

Συνάρτηση `FitnessFunction()`

Λόγω του μεγέθους της συνάρτησης αυτής, η ανάλυση της γίνεται τμηματικά. Πρώτα παρουσιάζεται ο ψευδοκώδικας του κάθε τμήματος που αναλύεται και στη συνέχεια δίνεται η υλοποίηση του κάθε τμήματος σε C#

ΣΥΝΑΡΤΗΣΗ `FitnessFunction()`(δείκτης)

1. Αρχικοποίηση του συνολικού σκορ και των επιμέρους σκορ (`scoreA`, `scoreB`, `scoreC`, `scoreD`) σε 0
2. Ανάκτηση του DNA από τον πληθυσμό του γενετικού αλγορίθμου με βάση τον δείκτη
3. Για κάθε γονίδιο στο DNA:
 - α. Έλεγχος για διπλότυπα:
 - i. Αν βρεθεί διπλότυπο, επιστροφή σκορ 0.01
 - β. Υπολογισμός σκορ ανάλογα με την κατηγορία του γονιδίου (`Museum`, `Sightseeing`, `Restaurant`, `Entertainment`)
 - i. Αν το γονίδιο αντιστοιχεί σε "Museum", αύξηση του `scoreA` κατά 1
 - ii. Αν το γονίδιο αντιστοιχεί σε "Sightseeing", αύξηση του `scoreB` κατά 1
 - iii. Αν το γονίδιο αντιστοιχεί σε "Restaurant", αύξηση του `scoreC` κατά 1
 - iv. Αν το γονίδιο αντιστοιχεί σε "Entertainment", αύξηση του `scoreD` κατά 1
- 4.Επιστροφή του συνολικού σκορ ως το άθροισμα των επιμέρους σκορ `scoreA`, `scoreB`, `scoreC`, `scoreD`

```

1. private double FitnessFunction(int index)
2. {
3.     double score = 0;
4.     double scoreA = 0;
5.     double scoreB = 0;
6.     double scoreC = 0;
7.     double scoreD = 0;
8.     DNA<Dictionary<string, string>> dna = ga.Population[index];
9.     List<string> tmp = new List<string>();
10.    List<string> tmp2 = new List<string>();
11.    double tmpScore=0;

```

```

12.
13.
14.     for(int j = 0; j < dna.Genes.Length; j++)
15.     {
16.         //check for duplicate
17.         tmp = dna.Genes[j].Keys.ToList();
18.         for(int i=0; i<j; i++)
19.         {
20.             if(dna.Genes[i].ContainsKey(tmp[0]))
21.             {
22.                 score = 0.01;
23.                 return score;
24.             }
25.         }
26.
27.         tmp2 = dna.Genes[j].Values.ToList();
28.         if (tmp2[0] == "Museum")
29.         {
30.             scoreA += 1;
31.         }
32.         else if (tmp2[0] == "Sightseeing")
33.         {
34.             scoreB += 1;
35.         }
36.         else if (tmp2[0] == "Restaurant")
37.         {
38.             scoreC += 1;
39.         }
40.         else if (tmp2[0] == "Entertainment")
41.         {
42.             scoreD += 1;
43.         }
44.

```

Αρχικά, όπως φαίνεται γίνεται η δήλωση και η αρχικοποίηση κάποιων βασικών μεταβλητών της συνάρτησης(3-11). Στη συνέχεια γίνεται έλεγχος για διπλοεγγραφές(18-25) έτσι ώστε να αποφευχθούν στις επόμενες γενιές λύσεων. Έπειτα, γίνεται η πρώτη φάση αξιολόγησης της υποψήφιας λύσης, η οποία γίνεται βάση των γενικών κατηγοριών δραστηριοτήτων (Museum,Sightseeing,Restaurant,Entertainment). Τα score αυτά (scoreA,scoreB,scoreC,scoreD) θα χρησιμοποιηθούν στον υπολογισμό του τελικού score και θα αντιπροσωπεύουν το πόσο κοντά ή πόσο μακριά βρίσκεται η υποψήφια λύση στις προτιμήσεις που έχει δηλώσει ο χρήστης στη φόρμα αναζήτησης.

5. Ελέγξτε αν το ενεργό γονίδιο αντιστοιχεί σε κάποια δραστηριότητα στο λεξικό activityTypeD
 - α. Αν ναι, αποθηκεύστε τον τύπο της δραστηριότητας στη μεταβλητή activity
6. Υπολογισμός του tmpScore με βάση το αν το activity ταιριάζει με τα επισκεπτόμενα κράτη:
 - α. Για κάθε κράτος (π.χ. France, Italy, United States, κτλ.):
 - i. Ελέγξτε αν το activity ταιριάζει με τον τύπο της δραστηριότητας που σχετίζεται με το κράτος
 - ii. Αν ταιριάζει, αυξήστε το tmpScore κατά 1
7. Υπολογισμός του τελικού σκορ με βάση το συνδυασμό των επιμέρους σκορ scoreA, scoreB, scoreC, scoreD και του tmpScore
8. Επιστροφή του τελικού σκορ

```

1. if(activityTypeD.ContainsKey(tmp[0]))
2. {

```

```

3. //apothikese to activity type
4. string activity = activityTypeD[tmp[0]];
5.
6. if(visitedC.Contains("France"))
7. {
8.     if(activity.Equals("History") || activity.Equals("Art") || activity.Equals("French"))
9.     {
10.         tmpScore += 1;
11.     }
12. }
13. if(visitedC.Contains("Italy"))
14. {
15.     if(activity.Equals("History") || activity.Equals("Italian"))
16.     {
17.         tmpScore += 1;
18.     }
19. }
20. if(visitedC.Contains("United States"))
21. {
22.     if(activity.Equals("Technology") || activity.Equals("American"))
23.     {
24.         tmpScore += 1;
25.     }
26. }
27. if(visitedC.Contains("Spain"))
28. {
29.     if (activity.Equals("Art") || activity.Equals("Spanish"))
30.     {
31.         tmpScore += 1;
32.     }
33. }
34. if(visitedC.Contains("England"))
35. {
36.     if(activity.Equals("England"))
37.     {
38.         tmpScore += 1;
39.     }
40. }
41. if(visitedC.Contains("Germany"))
42. {
43.     if(activity.Equals("German") || activity.Equals("History"))
44.     {
45.         tmpScore += 1;
46.     }
47. }
48. if(visitedC.Contains("Japan"))
49. {
50.     if (activity.Equals("Japanese") || activity.Equals("History") ||
51.         activity.Equals("Art") || activity.Equals("Technology"))
52.     {
53.         tmpScore += 1;
54.     }
55. if(visitedC.Contains("Turkey"))
56. {
57.     if(activity.Equals("Turkish") || activity.Equals("History"))
58.     {
59.         tmpScore += 1;
60.     }
61. }
62. }
63.

```

Σε αυτό το σημείο εξετάζεται ποιες δραστηριότητες από την υποψήφια λύση περιέχουν ActivityType που να ταυτίζεται με τα ενδιαφέροντα του χρήστη. Τα ενδιαφέροντα του χρήστη σε αυτό το σημείο προκύπτουν από τη μεταβλητή VisitedC που συμβολίζει τις χώρες που έχει επισκεφθεί ο εκάστοτε χρήστης στο παρελθόν. Για παράδειγμα (6-11), αν ο χρήστης έχει επισκεφθεί τη Γαλλία και

κάποια από τις δραστηριότητες της υποψήφιας λύσης έχει ActivityType History ή Art ή French τότε η μεταβλητή tmpScore θα αυξηθεί κατά μία μονάδα. Αυτή η μεταβλητή θα χρησιμοποιηθεί για τον υπολογισμό του τελικού score.

9. Αν ο χρήστης δεν έχει επισκεφθεί καμία χώρα (δηλαδή, αν η λίστα visitedC περιέχει "None"):

α. Υπολογισμός της απόκλισης (deviation) μεταξύ των πραγματικών σκορ και των αναμενόμενων σκορ (πολλαπλασιασμένα τα ποσοστά πιθανότητας με το TotalGenesSize) για κάθε κατηγορία (A, B, C, D)

β. Υπολογισμός του τελικού σκορ ως την αντίστροφη σχέση της απόκλισης ($1.0 / (1.0 + deviation)$)

γ. Επιστροφή του τελικού σκορ

```

1. if (visitedC.Contains("None"))
2. {
3.     deviation = Math.Abs(scoreA - propability_A * TotalGenesSize) +
4.     Math.Abs(scoreB - propability_B * TotalGenesSize) +
5.     Math.Abs(scoreC - propability_C * TotalGenesSize) +
6.     Math.Abs(scoreD - propability_D * TotalGenesSize);
7.
8.     score = 1.0 / (1.0 + deviation);
9.     return score;
10. }
11.

```

Στη περίπτωση που ο χρήστης δεν έχει επισκεφθεί κάποια χώρα (η μεταβλητή VisistedC περιέχει τη λέξη None) το τελικό score υπολογίζεται αποκλειστικά από την απόσταση της εκάστοτε υποψήφιας λύσης από τα φίλτρα αναζήτησης που έχει επιλέξει ο χρήστης στη φόρμα αναζήτησης.

10. Διαφορετικά (αν δεν εφαρμόζεται η περίπτωση "None"):

α. Υπολογισμός του βάρους sliderWeight βάσει των σκορ κάθε κατηγορίας (A, B, C, D) και των αντίστοιχων πιθανοτήτων (propability_A, propability_B, propability_C, propability_D)

β. Υπολογισμός του τελικού σκορ ως συνδυασμός του sliderWeight και του tmpScore με βάρη 0.1 και 0.9 αντίστοιχα

γ. Επιστροφή του τελικού σκορ

ΤΕΛΟΣ ΣΥΝΑΡΤΗΣΗΣ

```

1. else
2. {
3.
4.     double sliderWeight = (scoreA * propability_A) +
5.     (scoreB * propability_B) +
6.     (scoreC * propability_C) +
7.     (scoreD * propability_D);
8.
9.     score = 0.1 * sliderWeight + 0.9 * tmpScore;
10.    return score;
11.
12.
13. }
14.

```

Σε αντίθετη περίπτωση το score υπολογίζεται κατά 10% από τα φίλτρα αναζήτησης που έχει δηλώσει ο χρήστης στη φόρμα αναζήτησης και κατά 90% από το score που έχει προκύψει στη

μεταβλητή `tmpScore`, η οποία βασίζεται στο ενδιαφέρον του χρήστη βάση των χωρών που έχει επισκεφθεί στο παρελθόν.

ΚΕΦΑΛΑΙΟ 5 – ΠΑΡΑΔΕΙΓΜΑΤΑ ΧΡΗΣΗΣ

5.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα παρουσιαστούν και θα εξηγηθούν τα αποτελέσματα της εφαρμογής. Συγκεκριμένα θα δημιουργηθούν διάφορα προφίλ χρηστών με διαφορετικές και ίδιες προτιμήσεις, έτσι ώστε να συγκριθούν τα αποτελέσματα αυτών.

5.2 Παραδείγματα χρήσης

Παράδειγμα 1

Σε αυτό το παράδειγμα θα συγκριθούν δύο όμοιοι χρήστες, δηλαδή δύο χρήστες που έχουν επισκεφθεί ακριβώς τις ίδιες χώρες.

ID	Username	Password	Email	Gender	Region	VisitedCountries
14	bill	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	bill@gmail.com	Male	Africa	France, Italy,
15	alex	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	alex@gmail.com	Male	America	France, Italy,
18	manolis	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	manolis@gmail.com	Male	Asia	Spain,England,Germany,
19	george	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	george@gmail.com	Male	Africa	None

Παρουσιάζονται τα αποτελέσματα του χρήστη bill με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Mount Lycabettus, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Jasmine Skylounge, Spondi, Tartate, SPIROS&VASILIS, Ovio, Oroscopo, Belpaese
--	--

Παρουσιάζονται τα αποτελέσματα του χρήστη alex με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Panathenaic Stadium, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Spondi, Tartate, SPIROS&VASILIS, Ovio, Oroscopo, Belpaese, THE JAMES JOYCE
--	---

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που έχουν επισκεφθεί ακριβώς τις ίδιες χώρες και έχουν επιλέξει ακριβώς τα ίδια preferences, θα έχουν κατά 87% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

Παράδειγμα 2

Σε αυτό το παράδειγμα θα συγκριθούν πάλι οι χρήστες του προηγούμενου παραδείγματος, με τη διαφορά ότι θα έχουν διαφορετικά preferences.

Παρουσιάζονται τα αποτελέσματα του χρήστη bill με τιμές στα preferences museum=1, sightseeing=3, restaurant=2, entertainment=4

Date: 29/12/2023-31/12/2023	
Preferences: Museums: 1 Sightseeing: 3 Restaurants: 2 Entertainment: 4	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, The Clumsies, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Iguazu Athens, Spondi, Tartate, SPIROS&VASILIS, Ovio, Oroscopto, Belpaese

Παρουσιάζονται τα αποτελέσματα του χρήστη alex με τιμές στα preferences museum=1, sightseeing=5, restaurant=5, entertainment=1

Date: 29/12/2023-31/12/2023	
Preferences: Museums: 1 Sightseeing: 5 Restaurants: 5 Entertainment: 1	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Acropolis, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Spondi, Tartate, SPIROS&VASILIS, Juicy Grill, Ovio, Oroscopto, Belpaese

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που έχουν επισκεφθεί ακριβώς τις ίδιες χώρες και έχουν επιλέξει διαφορετικά preferences, θα έχουν κατά 80% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

Παράδειγμα 3

Σε αυτό το παράδειγμα θα συγκριθούν δύο ανόμοιοι χρήστες, δηλαδή δύο χρήστες που έχουν επισκεφθεί διαφορετικές χώρες.

ID	Username	Password	Email	Gender	Region	VisitedCountries
14	bill	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	bill@gmail.com	Male	Africa	France, Italy,
15	alex	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	alex@gmail.com	Male	America	France, Italy,
18	manolis	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	manolis@gmail.com	Male	Asia	Spain,England,Germany,
19	george	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	george@gmail.com	Male	Africa	None

Παρουσιάζονται τα αποτελέσματα του χρήστη alex με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023	
Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Panathenaic Stadium, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Spondi, Tartate, SPIROS&VASILIS, Ovio, Oroscopto, Belpaese, THE JAMES JOYCE

Παρουσιάζονται τα αποτελέσματα του χρήστη manolis με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, National Observatory of Athens, Theater of Dionysus, Arch Beer House, Barley Cargo, Cerdo Negro, Granada, THE JAMES JOYCE, Molly Malone's
--	---

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που έχουν επισκεφθεί διαφορετικές χώρες και έχουν επιλέξει τα ίδια preferences, θα έχουν κατά 60% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

Παράδειγμα 4

Σε αυτό το παράδειγμα θα συγκριθούν πάλι οι χρήστες του προηγούμενου παραδείγματος, με τη διαφορά ότι θα έχουν διαφορετικά preferences.

Παρουσιάζονται τα αποτελέσματα του χρήστη alex με τιμές στα preferences museum=1, sightseeing=1, restaurant=1, entertainment=1

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 1 Restaurants: 1 Entertainment: 1	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Jasmine Skylounge, Spondi, Tartate, SPIROS&VASILIS, Ovio, Oroscoipo, Belpaese, Cerdo Negro
--	--

Παρουσιάζονται τα αποτελέσματα του χρήστη manolis με τιμές στα preferences museum=1, sightseeing=3, restaurant=5, entertainment=5

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 3 Restaurants: 5 Entertainment: 5	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Dinner in the Sky, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Arch Beer House, Barley Cargo, Cerdo Negro, Granada, Selim Bey, THE JAMES JOYCE, Molly Malone's
--	--

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που έχουν επισκεφθεί διαφορετικές χώρες και έχουν επιλέξει διαφορετικά preferences, θα έχουν κατά 53% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

Παράδειγμα 5

Σε αυτό το παράδειγμα θα συγκριθούν δύο ανόμοιοι χρήστες και συγκεκριμένα ένας χρήστης που έχει επισκεφθεί χώρες του εξωτερικού και ένας χρήστης που δεν έχει επισκεφθεί καμία χώρα. Για αυτό το παράδειγμα όμως θα έχουν τα ίδια preferences.

ID	Username	Password	Email	Gender	Region	VisitedCountries
14	bill	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	bill@gmail.com	Male	Africa	France, Italy,
15	alex	8bb0cf6eb9b17d0f7d22b456f121257dc1254e1f...	alex@gmail.com	Male	America	France, Italy,
18	manolis	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	manolis@gmail.com	Male	Asia	Spain,England,Germany,
19	george	ef797c8118f02dfb649607dd5d3f8c7623048c9c...	george@gmail.com	Male	Africa	None

Παρουσιάζονται τα αποτελέσματα του χρήστη manolis με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, National Observatory of Athens, Theater of Dionysus, Arch Beer House, Barley Cargo, Cerdo Negro, Granada, THE JAMES JOYCE, Molly Malone's
--	---

Παρουσιάζονται τα αποτελέσματα του χρήστη george με τιμές στα preferences museum=1, sightseeing=2, restaurant=1, entertainment=2

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 2 Restaurants: 1 Entertainment: 2	Acropolis Museum, Acropolis, Mount Lycabettus, Plaka, Monastiraki, Μπαϊρακτάρης, Mo-Mix Bar, War Museum, Museum of Illusions Athens, Ancient Agora of Athens, Couleur Locale Athens, THE JAMES JOYCE, Molly Malone's, Amigos, Dos Hermanos
--	--

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που ο ένας εκ το οποίων δεν έχει ταξιδέψει πουθενά αλλά έχουν ίδια preferences, θα έχουν κατά 27% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

Παράδειγμα 6

Σε αυτό το παράδειγμα θα συγκριθούν πάλι οι χρήστες του προηγούμενου παραδείγματος, με τη διαφορά ότι θα έχουν διαφορετικά preferences.

Παρουσιάζονται τα αποτελέσματα του χρήστη manolis με τιμές στα preferences museum=1, sightseeing=1, restaurant=1, entertainment=1

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 1 Restaurants: 1 Entertainment: 1	Acropolis Museum, Benaki Museum, Museum of Cycladic Art, National Archaeological Museum, Byzantine and Christian Museum, Hellenic Motor Museum, War Museum, Arch Beer House, Barley Cargo, Wynwood, Juicy Grill, Cerdo Negro, Granada, THE JAMES JOYCE, Molly Malone's
--	--

Παρουσιάζονται τα αποτελέσματα του χρήστη george με τιμές στα preferences museum=1, sightseeing=5, restaurant=3, entertainment=5

Date: 29/12/2023-31/12/2023 Preferences: Museums: 1 Sightseeing: 5 Restaurants: 3 Entertainment: 5	Mount Lycabettus, Monastiraki, Dinner in the Sky, Hellenic Motor Museum, National Observatory of Athens, Ancient Agora of Athens, National Garden, Plateia Syntagmatos, Arch Beer House, Barley Cargo, Iguazu Athens, Jasmine Skylounge, Tartate, Molly Malone's, Dos Hermanos
--	--

Όπως φαίνεται από το παραπάνω αποτέλεσμα δύο χρήστες που ο ένας εκ το οποίων δεν έχει ταξιδέψει πουθενά και ταυτόχρονα έχουν διαφορετικά preferences, θα έχουν κατά 27% ίδιο προτεινόμενο πλάνο δραστηριοτήτων.

ΚΕΦΑΛΑΙΟ 6 – ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΕΠΕΚΤΑΣΕΙΣ

6.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα συζητηθούν τα αποτελέσματα της εφαρμογής του προηγούμενου κεφαλαίου, καθώς επίσης και το πως θα μπορούσε η συγκεκριμένη εφαρμογή να επεκταθεί και να βελτιωθεί σε μελλοντικό χρόνο. Τέλος, συνοψίζοντας θα αναφερθούν κάποια συμπεράσματα που προκύπτουν από την παρούσα διπλωματική εργασία για τα συστήματα προτάσεων.

6.2 Σχολιασμός αποτελεσμάτων αλγορίθμου

Στο προηγούμενο κεφάλαιο παρουσιάστηκαν κάποια παραδείγματα χρήσης του συστήματος προτάσεων της εφαρμογής, από αυτά θα μπορούσε κανείς να διεξάγει κάποια συμπεράσματα.

Ένα συμπέρασμα είναι ότι αλλάζοντας τα preferences των χρηστών δεν επηρεάζεται σημαντικά το τελικό πλάνο δραστηριοτήτων που θα προκύψει. Αυτό συμβαίνει καθώς τα preferences επηρεάζουν το τελικό αποτέλεσμα σε βαθμό 10% έναντι του 90% που επηρεάζει το ιστορικό των χωρών που έχει επισκεφθεί ο χρήστης.

Ακόμα μία παρατήρηση προκύπτει από την ομοιότητα που έχουν τα πλάνα δραστηριοτήτων δύο χρηστών που έχουν επισκεφθεί διαφορετικές χώρες μεταξύ τους. Στη συγκεκριμένη περίπτωση η ομοιότητα των πλάνων κυμαίνεται από 53% έως 60%. Αυτό συμβαίνει, διότι κάποιες χώρες μπορεί από τη μία να είναι διαφορετικές, από την άλλη όμως μπορεί να έχουν κοινά. Για παράδειγμα η Γαλλία και η Ιταλία φημίζονται και οι δύο για την ιστορία τους και τη τέχνη τους, οπότε είναι λογικό να παρουσιάζονται μουσεία ιστορίας και τέχνης στους χρήστες που έχουν επισκεφθεί κάποια από αυτές τις χώρες.

6.3 Μελλοντικές επεκτάσεις της εφαρμογής

Η εφαρμογή θα μπορούσε να επεκταθεί σε πολλά επίπεδα. Αρχικά θα μπορούσε να επεκταθεί ο πληθυσμός των διαθέσιμων δραστηριοτήτων της βάσης δεδομένων (έναντι των 50 που υπάρχουν τώρα). Αυτό θα μείωνε ακόμα περισσότερο την ομοιότητα των πλάνων στις περιπτώσεις όπου δύο χρήστες δεν έχουν κάποια ομοιότητα μεταξύ τους. Για παράδειγμα ένας χρήστης που δεν έχει ταξιδέψει πουθενά με έναν χρήστη που έχει ταξιδέψει σε κάποιες χώρες και τα preferences τους είναι διαφορετικά, είναι δύο χρήστες που δεν έχουν κάποια ομοιότητα μεταξύ τους, όμως το σύστημα προτάσεων της εφαρμογής τους προτείνει πλάνα με ομοιότητα 27%.

Ένας ακόμα τομέας που θα μπορούσε να επεκταθεί η εφαρμογή θα ήταν να μην περιορίζεται μόνο στην Αθήνα, αλλά να επεκταθεί για όλα τα μέρη της Ελλάδας ή ακόμα καλύτερα επεκταθεί για όλες τις χώρες στο κόσμο.

6.4 Συμπεράσματα

Μετά από τη παρούσα διπλωματική και έπειτα από το σχεδιασμό μιας εφαρμογής που ενσωματώνει ένα σύστημα προτάσεων διεξάγεται το ίδιο συμπέρασμα με αυτό του δεύτερου κεφαλαίου. Συγκεκριμένα, όσο αναφορά τα συστήματα προτάσεων δεν υπάρχει το τέλειο σύστημα πρότασης, δηλαδή ένα σύστημα που να μην έχει μειονεκτήματα, ούτε υπάρχει κάποιο σύστημα που ταιριάζει για όλες τις περιπτώσεις, καθώς κάθε πρόβλημα και κάθε σύστημα είναι μοναδικό και απαιτεί διαφορετική αντιμετώπιση.

ΚΕΦΑΛΑΙΟ 7 - ΒΙΒΛΙΟΓΡΑΦΙΑ

7.1 Βιβλιογραφία

- [1] Vanetti, M., Binaghi, E., Carminati, B., Carullo, M., & Ferrari, E. (2011). Content-based filtering in on-line social networks. In C. Dimitrakakis, A. Gkoulalas-Divanis, A. Mitrokotsa, V. S. Verykios & Y. Saygin (Eds.), *Privacy and Security Issues in Data Mining and Machine Learning. PSDML 2010. Lecture Notes in Computer Science (Vol. 6549)*. Berlin: Springer.
- [2] Yang, B., Lei, Y., Liu, J., & Li, W. (2017). Social collaborative filtering by trust. *Transactions on Pattern Analysis and Machine Intelligence*, 39(8), 1633–1647. IEEE.
- [3] Tsikrika, T., Symeonidis, S., Gialampoukidis, I., Satsiou, A., Vrochidis, S., & Kompatsiaris, I. (2018). A hybrid recommendation system based on density-based clustering. In S. Diplaris, A. Satsiou, A. Følstad, M. Vafopoulos & T. Vilarinho (Eds.), *Internet Science. INSCI 2017. Lecture Notes in Computer Science (Vol. 10750)*. Cham: Springer.
- [4] Anil, P., Neev, P., Tanvi, B., & Rekha, S. (2014). Non-personalized recommender systems and user-based collaborative recommender systems. *International Journal of Applied Information Systems*, 6(9), 22–27.
- [5] Dip Paul and Subhradeep Kundu (2020). A Survey of Music Recommendation Systems with a Proposed Music Recommendation System. *Emerging Technology in Modelling and Graphics, Advances in Intelligent Systems and Computing* 937.
- [6] Chen, L., Hsu, F., Chen, M., Hsu, Y., “Developing recommender systems with the consideration of product profitability for sellers,” *Information Sciences* 178, 4, pp. 1032-1048, 2008.
- [7] Kobsa, A., Koenemann, J., Pohl, W., “Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships,” *The Knowledge Eng. Review* 16, 2, pp. 111-155, 2001.
- [8] Torres, R., et al.: Enhancing Digital Libraries with TechLens+. In: Proc.2004 Joint ACM/IEEE Conference on Digital Libraries(JCDL'04), pp. 228–236 (2004)
- [9] Greg, L., Brent, S., Jeremy, Y.: Amazon.com Recommendations. Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, pp. 76–80 (January/February 2003)

- [10] Aicha Er-Rafyq & Abdellah Idrissi (2023). Review of Recommendation Systems and their Applications in E-learning Part of the Studies in Computational Intelligence book series (SCI,volume 1102)
- [11] K. G. Srinivasa • Siddesh G. M. Srinidhi H. Network Data Analytics A Hands-On Approach for Application Development
- [12] Harish Sharma • Kannan Govindan • Ramesh C. Poonia • Sandeep Kumar • Wael M. El-Medany Editors Advances in Computing and Intelligent Systems Proceedings of ICACM 2019
- [13] Brijesh Iyer • Tom Crick • Sheng-Lung Peng Editors Applied Computational Technologies Proceedings of ICCET 2022
- [14] Akshay Kulkarni Adarsha Shivananda Anoosh Kulkarni V Adithya Krishna Applied Recommender Systems with Python Build Recommender Systems with Deep Learning, NLP and Graph-Based Techniques
- [15] Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland Advances in Intelligent Systems and Computing
- [16] Junyi Gu, Zhiping Wu & Xin Wang Home Frontier Computing Conference paper Research and Practice of Genetic Algorithm Theory
- [17] Seyedali Mirjalili, Jin Song Dong, Ali Safa Sadiq & Hossam Faris Genetic Algorithm: Theory, Literature Review, and Application in Image Reconstruction
- [18] <https://www.geeksforgeeks.org/genetic-algorithms/>