



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Σύστημα συστάσεων τουριστικών προορισμών με χρήση
γνωσιακών γράφων**

**Χούντα Ιωάννα
Α.Μ. 18390133**

Εισηγητής: Χρήστος Τρούσσας, Επίκουρος Καθηγητής

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Σύστημα συστάσεων τουριστικών προορισμών με χρήση
γνωσιακών γράφων**

**Χούντα Ιωάννα
Α.Μ. 18390133**

Εισηγητής:

Χρήστος Τρούσσας, Επ. Καθηγητής

Εξεταστική Επιτροπή:

- 1. Χρήστος Τρούσσας, Επ. Καθηγητής**
- 2. Ακριβή Κρούσκα, Μέλος ΕΔΙΠ**
- 3. Παναγιώτα Τσελέντη, Μέλος ΕΔΙΠ**

Ημερομηνία εξέτασης: Μάρτιος 2024

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ/ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η κάτωθι υπογεγραμμένη Χούντα Ιωάννα του Παναγιώτη, με αριθμό μητρώου 18390133, φοιτήτρια του Πανεπιστημίου Δυτικής Αττικής της Σχολής Μηχανικών του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της πτυχιακής/διπλωματικής εργασίας και ότι 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».

Ο/Η Δηλών/ούσα



ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία αντιπροσωπεύει το αποτέλεσμα εκτενών προσπαθειών και αφοσίωσης σε έναν τομέα που με απασχόλησε βαθιά, τον χώρο ανάπτυξης συστημάτων προτάσεων. Την προσπάθειά μου αυτή υποστήριξε ο επιβλέπων καθηγητής μου, Χρήστος Τρούσσας, τον οποίο θα ήθελα να ευχαριστήσω.

ΠΕΡΙΛΗΨΗ

Σε έναν διαδικτυακό κόσμο που οι πληροφορίες ολοένα και αυξάνονται είναι σημαντική η εύρεση ενός μοντέλου προτεινόμενων που θα παρέχει στον χρήστη τις κατάλληλες συστάσεις. Το πεδίο του τουρισμού αναδύεται ως ένα από τα πιο υποσχόμενα πεδία που μπορεί να προσαρμοστεί ένα σύστημα προτάσεων. Έχοντας ως βάση το μοντέλο WSM (Weighted Sum Model) και εκμεταλλευόμενοι τις μεθοδολογίες ανάλυσης αποφάσεων, το σύστημα προτάσεων για τον τουρισμό προτείνει τις κατάλληλες ταξιδιωτικές επιλογές στους χρήστες ανάλογα με τις προτιμήσεις και τα ενδιαφέροντά τους. Το σύστημα στηρίζεται σε έναν γνωσιακό γράφο που περιλαμβάνει πληροφορίες σχετικές με τον χρήστη και τους ταξιδιωτικούς προορισμούς.

ABSTRACT

In an online world where information is constantly expanding, finding a recommendation model to provide users with appropriate suggestions is crucial. The field of tourism emerges as one of the most promising domains where a recommendation system can be adapted. Based on the Weighted Sum Model (WSM) and using decision analysis methods, the tourism recommendation system suggests appropriate travel options to users according to their preferences and interests. The system is based on a knowledge graph containing information about the user and the destinations.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ.....	10
Σημασία της εξατομίκευσης	10
Σκοπός της διπλωματικής εργασίας.....	10
Δομή της διπλωματικής εργασίας.....	12
ΚΕΦΑΛΑΙΟ 2 - ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	13
Γράφημα	13
Μοντέλα γραφημάτων	13
Γνωσιακός γράφος.....	14
Λήψη αποφάσεων	15
Συστήματα Προτάσεων	17
Τύποι συστημάτων συστάσεων.....	18
ΚΕΦΑΛΑΙΟ 3 - ΜΕΘΟΔΟΛΟΓΙΑ	20
Περιγραφή της μεθοδολογίας.....	20
Επιλογή θέματος.....	20
Καθορισμός βασικών οντοτήτων.....	21
Αναζήτηση Δεδομένων	21
Δημιουργία προφίλ χρήστη.....	22
Συλλογή πιθανών ενδιαφερόντων χρήστη.....	23
Κατασκευή Γνωσιακού Γράφου.....	24
Ενσωμάτωση Budget	24
Ενσωμάτωση Μοντέλου Weighted Sum	25
ΚΕΦΑΛΑΙΟ 4 - Αρχιτεκτονική Συστήματος.....	27
Γνωσιακός γράφος με το Neo4j.....	27
Μοντελοποίηση	29
Σύστημα προτάσεων με χρήση του WSM	31
Το ενσωματωμένο Μοντέλο WSM	32
Σύστημα προτεινόμενων	36
Παρουσίαση διεπαφής.....	39
Προβολή προτεινόμενων.....	40
Αναζήτηση προορισμών	41
ΚΕΦΑΛΑΙΟ 5 - ΑΞΙΟΛΟΓΗΣΗ	43
Αξιολόγηση Δεδομένων.....	43
ΚΕΦΑΛΑΙΟ 6 - ΣΥΜΠΕΡΑΣΜΑΤΑ.....	48

ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ

Σημασία της εξατομίκευσης

Ο Παγκόσμιος Ιστός, ευρέως γνωστός ως "World Wide Web" ή απλά "Internet," αποτελεί ένα εκτεταμένο δίκτυο από έγγραφα και υπηρεσίες που δίνει τη δυνατότητα στους χρήστες να αναζητούν, να μοιράζονται και να αλληλεπιδρούν με πληροφορίες μέσω του Διαδικτύου. Στην παρούσα δεκαετία, ο αριθμός των χρηστών και ο όγκος των πληροφοριών έχουν αυξηθεί σημαντικά, επιφέροντας νέες προκλήσεις και ευκαιρίες.

Κάθε χρήστης μπορεί να συνδεθεί στο διαδίκτυο με ένα κλικ, και στη στιγμή εκείνη, να έχει πρόσβαση σε εκατοντάδες αποτελέσματα σχετικά με την αναζήτησή του. Οι μηχανές αναζήτησης συνεχώς βελτιστοποιούνται με στόχο την αποτελεσματικότερη παρουσίαση των αποτελεσμάτων και την εξαγωγή πληροφοριών που ικανοποιούν τις ανάγκες του χρήστη.

Καθώς ο όγκος των πληροφοριών συνεχώς αυξάνεται, η ανάγκη για αποδοτική αναζήτηση και παροχή σχετικών πληροφοριών γίνεται ολοένα και πιο σημαντική. Συνεπώς, παρατηρούνται σημαντικές προσπάθειες στην ανάπτυξη υπηρεσιών που επιδιώκουν να παρέχουν στους χρήστες του διαδικτύου όσο το δυνατόν πιο ακριβείς και εξατομικευμένες πληροφορίες. Ο αναπτυσσόμενος τομέας αυτός έχει πρωταρχικό στόχο την προσαρμογή του περιεχομένου στις ανάγκες, τα ενδιαφέροντα και τις προτιμήσεις κάθε χρήστη.

Σήμερα, οι χρήστες έχουν υψηλές προσδοκίες όσον αφορά την εξατομίκευση των πληροφοριών που λαμβάνουν. Σε αυτό το πλαίσιο, όλες οι υπηρεσίες που έχουν αναπτυχθεί εξελίσσονται προκειμένου να ανταποκριθούν στις απαιτητικές ανάγκες των χρηστών και να ξεχωρίσουν από άλλες υπηρεσίες με παρόμοιο στόχο.

Ένα χαρακτηριστικό παράδειγμα αυτής της τάσης είναι η εξέλιξη των πλατφορμών όπως το Netflix και το YouTube. Αυτές οι πλατφόρμες χρησιμοποιούνται για την αναζήτηση ταινιών και μουσικής αντίστοιχα. Ειδικά στον τομέα της ψυχαγωγίας, οι εφαρμογές που έχουν αναπτυχθεί είναι αμέτρητες, με ορισμένες από αυτές να ανταγωνίζονται σκληρά μεταξύ τους. Παράλληλα, διαρκώς εμφανίζονται νέες εφαρμογές που απειλούν την κυριαρχία των προηγούμενων, όπως το Disney+ και το Spotify.

Σε αυτό τον ανταγωνιστικό τομέα, η συνεχής καινοτομία και η ικανότητα προσαρμογής στις εξελίξεις της αγοράς είναι ζωτικής σημασίας. Οι πλατφόρμες πρέπει όχι μόνο να προσφέρουν εξαιρετική εξατομίκευση και ποιότητα περιεχομένου, αλλά και να αντιλαμβάνονται τις μεταβαλλόμενες προτιμήσεις των χρηστών. Η συνεχής προσπάθεια επίτευξης της ακρίβειας και της προσαρμοστικότητας στις ανάγκες του κάθε χρήστη αποτελεί τον πυρήνα της εξέλιξης σε αυτόν τον τόσο απαιτητικό και ανταγωνιστικό τομέα χωρίς όμως να παραλείπεται η αξία της εμπειρίας του χρήστη.

Σκοπός της διπλωματικής εργασίας

Τα συστήματα προτεινόμενων (Recommendation Systems), γνωστά και ως συστήματα συστάσεων, είναι λογισμικό σχεδιασμένο να παρουσιάζει περιεχόμενο που μπορεί να ενδιαφέρει τους χρήστες. Δεδομένης της σημασίας της εξατομίκευσης στον κόσμο της πληροφορίας, τα συστήματα προτάσεων προσπαθούν να προσαρμόσουν το περιεχόμενο που παρέχουν στους χρήστες, ανάλογα με τις συγκεκριμένες τους ανάγκες.

Το πεδίο του τουρισμού είναι ένα από τα πιο υποσχόμενα πεδία που μπορούν να εφαρμοστούν τα συστήματα προτεινόμενων. Το TripAdvisor, ένα από τα πιο διαδεδομένα site που υπάρχουν στο διαδίκτυο που σχετίζονται με το πεδίο του τουρισμού, αποτελεί έναν έμπιστο οδηγό που μπορεί να χρησιμοποιήσει ένας χρήστης προκειμένου να πληροφορηθεί για προορισμούς τους οποίους πρόκειται να επισκεφθεί. Οι διαθέσιμες πληροφορίες περιλαμβάνουν πρόσβαση, διαμονή, εστίαση, σημεία ενδιαφέροντος όπως μνημεία, αξιοθέατα αλλά και εκδρομές για να γνωρίσει τον πολιτισμό της πόλης που επισκέπτεται ή για να ανακαλύψει την ομορφιά της φύσης σε τοποθεσίες κοντινές του προορισμού του.

Αντικείμενο αυτής της εργασίας είναι λοιπόν η ανάπτυξη ενός συστήματος το οποίο λαμβάνοντας υπόψη τις προτιμήσεις και τα ενδιαφέροντα ενός χρήστη θα παρέχει προσωποποιημένες συστάσεις για ταξιδιωτικούς προορισμούς. Το σύστημα δεν επεκτείνεται σε όλο το εύρος των παροχών που συναντάμε στο Trip Advisor αποτελεί όμως μια προσπάθεια συγκέντρωσης προορισμών από όλο τον πλανήτη σε έναν εννιαίο γνωσιακό γράφο (Knowledge Graph), από τον οποίο προκύπτουν τα προτεινόμενα στοιχεία. Επικεντρώνεται στην δυνατότητα του χρήστη να αλληλεπιδρά με τους προορισμούς, εμβαθύνοντας στα ενδιαφέροντα του και παρέχοντας πληροφορίες για προορισμούς που ταιριάζουν με το προσωπικό του προφίλ.

Κύριος στόχος της διπλωματικής είναι να συνδυάσει τεχνικές από τον τομέα Θεωρίας Αποφάσεων (Decision Theory - DT) όπως η Πολυκριτηριακή Ανάλυση Αποφάσεων (Multiple Criteria Analysis – MCDA) και συγκεκριμένα το μοντέλο βαρυτικού αθροίσματος (Weighted Sum Model - WSM) και την πιο συνηθισμένη προσέγγιση στην ανάπτυξη των συστημάτων προτάσεων, το συνεργατικό φιλτράρισμα (Collaborative Filtering). Το μοντέλο του βαρυτικού αθροίσματος που είναι ενσωματωμένο στο σύστημα προτάσεων δίνει την δυνατότητα στο σύστημα να παρέχει προτάσεις για προορισμούς που συνδυάζουν τα ενδιαφέροντα του χρήστη και άλλες παραμέτρους όπως για παράδειγμα τον προϋπολογισμό του (budget).

Τα δεδομένα του συστήματος προτάσεων έχουν δομηθεί ως ένας γνωσιακός γράφος (Knowledge Graph), στα πλαίσια διερεύνησης της μοντελοποίησης ενός προβλήματος με χρήση του γνωσιακού γράφου αντί μίας σχεσιακής βάσης δεδομένων, τύπου SQL ή MariaDB. Για τις ανάγκες αποθήκευσης των δεδομένων, έχει επιλεγεί η βάση δεδομένων Neo4j, η οποία βασίζεται αποκλειστικά σε γράφους. Η Neo4j είναι μια προηγμένη γραφική βάση δεδομένων που επιτρέπει την αποθήκευση και ανάκτηση δεδομένων με τρόπο που αντανακλά τις σχέσεις μεταξύ τους.

Τέλος για την παρουσίαση των προτεινόμενων ταξιδιωτικών προορισμών στον χρήστη έχει αναπτυχθεί μια εφαρμογή κινητής συσκευής Android. Η εφαρμογή αυτή αποσκοπεί στην εγκαθίδρυση ενός τρόπου επικοινωνίας μεταξύ του χρήστη και του συστήματος προτάσεων, ο οποίος να είναι εύκολος, φιλικός και να ενισχύει την εμπειρία του χρήστη.

Δομή της διπλωματικής εργασίας

Το κεφάλαιο 2, με τίτλο "Θεωρητικό Υπόβαθρο," παρέχει εμβάθυνση στις έννοιες που σχετίζονται με τον γνωσιακό γράφο. Έννοιες όπως γραφήματα και γνώση συνδυάζονται γεννώντας τους γνωσιακούς γράφους οι οποίοι αποτελούν ανερχόμενο πεδίο έρευνας. Εξετάζονται επίσης οι μεθοδολογίες ανάλυσης αποφάσεων και οι εφαρμογές τους, με ειδική αναφορά στο μοντέλο βαρυτικού αθροίσματος (Weighted sum model - WSM). Το κεφάλαιο 2 κλείνει με τον ορισμό των συστημάτων προτάσεων και τους τύπους συστημάτων που υφίστανται σήμερα.

Στο κεφάλαιο 3, με τίτλο "Μεθοδολογία," παρουσιάζονται λεπτομερώς οι βασικές μέθοδοι που επιλέχθηκαν για τον σχεδιασμό και την ανάπτυξη του συστήματος προτάσεων. Αναφέρονται επίσης οι προκλήσεις και οι δυσκολίες που αντιμετωπίστηκαν κατά την διαδικασία της ανάπτυξης του λογισμικού.

Στο κεφάλαιο 4, παρουσιάζεται αναλυτικά η αρχιτεκτονική του συστήματος που αναπτύχθηκε. Στο τεχνικό μέρος, τρεις οντότητες συνεργάζονται προκειμένου να προβληθούν τα προτεινόμενα στον χρήστη μέσω γραφικής διεπαφής κινητής συσκευής. Αυτές οι οντότητες είναι ο γνωσιακός γράφος, το μοντέλο προτεινόμενων και το Flask API που αναλαμβάνει την επικοινωνία μεταξύ γραφικής διεπαφής και του λογισμικού που έχει δημιουργηθεί στην πλευρά του server.

Στο κεφάλαιο 5 παρέχεται λεπτομερής ανάλυση της απόδοσης του συστήματος μέσω εκτέλεσης διαφόρων σεναρίων δοκιμών. Αξιολογείται η αποτελεσματικότητα και η απόδοση του συστήματος προτάσεων με ενσωματωμένο το μοντέλο WSM, χρησιμοποιώντας δεδομένα που αναπαρίστανται με την μορφή γράφου. Η απόδοση ενός συστήματος εξαρτάται τόσο από την λογική του μοντέλου με το οποίο λειτουργεί αλλά σημαντικό επίσης παράγοντα αποτελεί η ποιότητα των δεδομένων.

Το κεφάλαιο 6, "Συμπεράσματα" συνοψίζει τα βασικά ευρήματα της εργασίας. Εκφράζονται προτάσεις για μελλοντικές επεκτάσεις του έργου, προτείνοντας βελτιώσεις στο σύστημα ή πεδία έρευνας που μπορούν να εξερευνηθούν στο μέλλον.

ΚΕΦΑΛΑΙΟ 2 - ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Γράφημα

Ένα γράφημα είναι μια μαθηματική δομή που αποτελείται από ένα σύνολο κορυφών και ένα πλήθος ακμών που συνδέουν αυτές τις κορυφές [7]. Συνήθως συμβολίζεται ως G , και ορίζεται από το ζεύγος των συνόλων V και E , όπου:

V είναι το σύνολο των κορυφών και περιλαμβάνει μη κενές στοιχεία που αντιπροσωπεύουν τα σημεία του γράφου.

E είναι το σύνολο των ακμών και περιλαμβάνει τα στοιχεία που αντιπροσωπεύουν τις συνδέσεις μεταξύ των κορυφών.

Στο πεδίο των γραφημάτων, υπάρχει μια διάκριση μεταξύ ομοιογενών και ετερογενών γραφημάτων. Ένα ομοιογενές γράφημα χαρακτηρίζεται από κόμβους που ανήκουν σε έναν μοναδικό τύπο οντότητας, ενώ οι ακμές συμβολίζουν μια ομοιόμορφη σχέση. Αντίθετα, ένα ετερογενές γράφημα περιλαμβάνει διαφορετικούς κόμβους και ακμές, υποδηλώνοντας διάφορους τύπους οντοτήτων και σχέσεων [4].

Μια ακμή που συνδέει ένα διατεταγμένο ζεύγος κορυφών, που ονομάζεται αρχική (αρχική ή πηγή) κορυφή και μια τερματική (τελική ή στόχος) κορυφή, ονομάζεται κατευθυνόμενη ακμή. Τα γραφήματα είναι ευέλικτα όσον αφορά τον τύπο των ακμών και τους τρόπους σύνδεσης των κορυφών. Στην βιβλιογραφία, συναντάμε πολλούς ορισμούς για τον όρο γράφημα. Όταν μιλάμε για γραφήματα, μπορεί να αναφερόμαστε σε γραφήματα με κατευθυνόμενες ή μη κατευθυνόμενες ακμές, με ή χωρίς βρόχους και με ή χωρίς πολλαπλές ακμές [7].

Το γράφημα παρουσιάζει τη δυνατότητα είτε να είναι σταθμισμένο είτε μη σταθμισμένο. Στην περίπτωση του σταθμισμένου γραφήματος, χρησιμοποιείται μια συνάρτηση βάρους, συμβολιζόμενη ως w , η οποία αναθέτει συγκεκριμένο βάρος, w_{ij} , σε κάθε ακμή που συνδέει τις κορυφές v_i και v_j . Αυτό σημαίνει ότι δίνεται έμφαση σε συγκεκριμένες συνδέσεις με τη χρήση της συνάρτησης βάρους, ενισχύοντας ή αποδυναμώνοντας την επίδρασή τους στη δομή του γραφήματος. Αντίθετα, σε μη σταθμισμένο γράφημα, όλες οι ακμές θεωρούνται ισοδύναμες χωρίς διάκριση βάρους, προσφέροντας έναν πιο ομοιόμορφο χαρακτήρα στη σύνδεση μεταξύ των κορυφών [4].

Μοντέλα γραφημάτων

Κατά καιρούς, διαθέτουμε πληροφορίες που περιλαμβάνουν πολλά πράγματα συνδεδεμένα μεταξύ τους με περίπλοκους τρόπους. Είναι ουσιώδες να κατανοήσουμε πώς αυτά τα πράγματα αλληλεπιδρούν μεταξύ τους.

Μία από τις λύσεις που μπορούν να δοθούν στην απεικόνιση πολύπλοκων σχέσεων είναι η αναπαράσταση των πληροφοριών μέσω ενός μοντέλου που έχει την δομή ενός γράφου. Οι γράφοι αντιπροσωπεύουν έναν τρόπο να εμφανίσουμε πώς διάφορα πράγματα συνδέονται χρησιμοποιώντας σημεία (κόμβους) και συνδέσεις (ακμές).

Γενικά, τα γραφήματα χρησιμοποιούνται για να αναπαραστήσουν διάφορα συστήματα και δομές, όπως δίκτυα, κοινωνικές σχέσεις, δρόμους και πολλά άλλα. Είναι βασικό εργαλείο στη θεωρία γράφων και χρησιμοποιούνται ευρέως σε πολλούς τομείς της επιστήμης των υπολογιστών και της πληροφορικής. [7]

Γνωσιακός γράφος

Ορισμός γνωσιακού γράφου

Οι γνωσιακοί γράφοι είναι σημασιολογικά μοντέλα που αναπαριστούν την γνώση μέσω μιας ιεραρχικής δομής που βασίζεται σε κλάσεις οντοτήτων, στις μεταξύ τους σχέσεις και στα χαρακτηριστικά στοιχεία αυτών [9].

Τυπικά αποτελούν μοντέλα γραφημάτων στα οποία έχει προστεθεί γνώση. Οι γνωσιακοί γράφοι δίνουν τη δυνατότητα αναπαράστασης γνώσης μέσω των κόμβων και των στοιχείων τους αλλά και την δυνατότητα εξαγωγής αυτής της γνώσης μέσω των σχέσεων που συνδέουν τους κόμβους [6]. Με την χρήση των γνωσιακών γράφων ενισχύεται η κατανόηση των πολύπλοκων σχέσεων και συνδέσεων στα δεδομένα.

Εφαρμογές γνωσιακών γράφων

Οι γνωσιακοί γράφοι, γνωστοί και ως Knowledge Graphs (KGs), αποτελούν ένα πολύ σημαντικό εργαλείο στο πλαίσιο της έρευνας που συναντάται τόσο στον κόσμο της βιομηχανίας όσο και στον ακαδημαϊκό ορίζοντα. Κατα κύριο λόγο αποτελούν θεμέλιο πολλών πληροφοριακών συστημάτων που για την λειτουργία τους χρειάζεται να έχουν πρόσβαση σε γνώση [12].

Οι γνωσιακοί γράφοι έχουν συνδεθεί στενά με έννοιες όπως μηχανική μάθηση και η τεχνητή νοημοσύνη. Συγκεκριμένα, η δυνατότητα των γνωσιακών γράφων να συνδυάζουν ετερογενείς πληροφορίες από πολλές και διαφορετικές πηγές είναι ένας από τους λόγους που έχουν χρησιμοποιηθεί στις έρευνες σε εφαρμογές της τεχνητής νοημοσύνης [6]. Μαζί με το RDF και το OWL παρέχουν την βάση, πάνω στην οποία στηρίζονται συστήματα που διαχειρίζονται γνώση [9].

Οι γνωσιακοί γράφοι έχουν μια ευρεία γκάμα εφαρμογών και χρησιμοποιούνται για διάφορους σκοπούς σε πολλά πεδία. Παράλληλα χάρη στην ικανότητα τους να παρέχουν σημασιολογικά δομημένες πληροφορίες, προσφέρουν σημαντικές λύσεις σε εφαρμογές όπως οι εικονικοί βοηθοί (virtual assistants) και τα συστήματα

συστάσεων (recommendation systems) [12]. Ακόμα, είναι ευρέως διαδεδομένοι σε πεδία όπως η ανάκτηση πληροφορίας (Information Retrieval).

Στη βιομηχανία, τα γραφήματα γνώσης έχουν εφαρμοστεί για να εξηγήσουν χρηματοοικονομικά αποτελέσματα, να βελτιώσουν συστήματα συστάσεων, να ενισχύσουν την αναζήτηση προϊόντων, να μοντελοποιήσουν δίκτυα μεταφορών, να αναλύσουν συστήματα ενέργειας, να ανιχνεύουν κυβερνοαπειλές, να προσαρμόζουν τη μάθηση, να διευκολύνουν την ολοκλήρωση δεδομένων και τη διαχείριση της γνώσης, καθώς επίσης και να ενισχύσουν τις μηχανές αναζήτησης. Τα γραφήματα γνώσης έχουν χρησιμοποιηθεί επίσης στην εξερεύνηση, οπτικοποίηση και ανάλυση δεδομένων, καθώς και σε διαδικασίες κοινοποίησης δεδομένων, συλλογισμού και λήψης αποφάσεων [6,9]

Λήψη αποφάσεων

Η καθημερινή ζωή συχνά συναντά πολλά πρακτικά προβλήματα και προκλήσεις που απαιτούν αποφάσεις για την αντιμετώπισή τους. Η διαδικασία λήψης αποφάσεων είναι μία πολύ κρίσιμη διαδικασία που απαιτεί πρώτα από όλα σαφή καθορισμό του προβλήματος, αξιολόγηση των δεδομένων και των εναλλακτικών και στην συνέχεια επιλογή της πιο υποσχόμενης εναλλακτικής.

Στην βιβλιογραφία, μπορούν να βρεθούν οι έννοιες "Multiple Criteria Decision Making" (MCDM) και "Multiple Criteria Decision Analysis" (MCDA). Η MCDM αναφέρεται στην ευρύτερη έννοια της διαδικασίας λήψης αποφάσεων όπου πολλαπλά κριτήρια χρησιμοποιούνται για την επιλογή μιας εναλλακτικής [4]. Η MCDA επικεντρώνεται περισσότερο στις αναλυτικές μεθόδους και τεχνικές που χρησιμοποιούνται για να υποστηρίξουν τη διαδικασία λήψης αποφάσεων με πολλαπλά κριτήρια. Επικεντρώνεται κυρίως στην ανάλυση και στην αξιολόγηση των εναλλακτικών παρά στην επιλογή της καλύτερης [2].

Πολυκριτηριακή ανάλυση αποφάσεων (MCDA)

Συνήθως, η MCDA περιλαμβάνει μαθηματικές μοντελοποιήσεις, τεχνικές βαρύτητας, αξιολογήσεις επιδόσεων, και άλλα εργαλεία για να βοηθήσει στη λήψη αποφάσεων.

WSM

Η μέθοδος του συνολικού βαρυτικού άθροισματος (Weighted Sum Method - WSM) είναι μια πολυκριτηριακή μέθοδος ανάλυσης αποφάσεων που χρησιμοποιείται για να αξιολογήσει πολλαπλές εναλλακτικές ως προς πολλά κριτήρια. Συγκεκριμένα, η WSM είναι η πιο γνωστή και απλούστερη μέθοδος πολυκριτηριακής ανάλυσης αποφάσεων [10].

Στην WSM, κάθε εναλλακτική αξιολογείται βάσει πολλών κριτηρίων. Κάθε κριτήριο έχει ένα σχετικό βάρος που υποδηλώνει τη σημασία του. Στη συνέχεια, υπολογίζεται το συνολικό σκορ της εναλλακτικής με βάση τον συνδυασμό των αποδόσεων της σε κάθε κριτήριο, λαμβάνοντας υπόψη τα βάρη των κριτηρίων [11].

Η μέθοδος WSM είναι ιδιαίτερα χρήσιμη όταν όλα τα δεδομένα εκφράζονται στην ίδια μονάδα μέτρησης [8] και εμφανίζει προβλήματα όταν χρησιμοποιείται σε πολυδιάστατα μοντέλα.

Η μαθηματική φόρμουλα που ακολουθεί το μοντέλο WSM είναι η εξής:

$$WSM = \sum_{i=1}^n w_i \cdot x_i$$

Συστήματα Προτάσεων

Ορισμός

Η αναζήτηση στο διαδίκτυο, σήμερα είναι αναπόσπαστο κομμάτι της καθημερινότητας όλων των ατόμων που θέλουν να αναζητήσουν πληροφορίες και να ενημερωθούν άμεσα για θέματα που τους αφορούν. Βρισκόμαστε στην εποχή των Big Data όπου τα δεδομένα ρέουν από παντού σε τέτοιο βαθμό που πλέον μιλάμε για ύπαρξη και συνεχή παραγωγή δεδομένων που περιγράφουν τα ήδη υπάρχοντα δεδομένα, τα λεγόμενα μετα-δεδομένα ή αλλιώς metadata.

Τα πλεονεκτήματα της ύπαρξης και της διαθεσιμότητας δεδομένων με τόσο εύκολο και άμεσο τρόπο είναι πολλά, όμως σε όλη αυτή την διαδικασία υπάρχουν συνέπειες και οι συνέπειες αυτές αφορούν τον τρόπο διαχείρισης αυτών των πληροφοριών από τους χρήστες. Οι χρήστες του διαδικτύου είναι αυτοί για τους οποίους προορίζονται όλες οι πληροφορίες όμως τόσο ο όγκος των δεδομένων όσο και η παρουσίαση τους σε πολλές και διάφορες μορφές προκαλεί σύγχυση. Η υπερφόρτωση των χρηστών με πληροφορίες είναι σήμερα αδιαμφισβήτητο γεγονός και έρευνες έχουν στραφεί προς την αντιμετώπιση αυτού του προβλήματος, υιοθετώντας διάφορες τεχνικές όπως η βελτιστοποίηση υπάρχοντων συστημάτων ανάκτησης πληροφορίας όπως είναι για παράδειγμα οι περιηγητές (browsers). Οι προσπάθειες όμως δεν σταματούν εκεί καθώς είναι επιτακτική η ανάγκη της δημιουργίας ενός συστήματος το οποίο θα δρα ως φίλτρο στις ήδη υπάρχουσες πληροφορίες και θα παρέχει προτάσεις στους χρήστες λαμβάνοντας υπόψη τις προτιμήσεις και τα ενδιαφέροντα τους [1]. Έτσι, δημιουργήθηκαν τα συστήματα προτάσεων ή αλλιώς Recommendation Systems (RS).

Μιλώντας για την διαφορά ανάμεσα στα συστήματα συστάσεων και των συστημάτων ανάκτησης πληροφορίας ο Burke [3] ανέφερε πως η ειδοποιός διαφορά μεταξύ των δύο συστημάτων είναι ο τύπος της αλληλεπίδρασης του εκάστοτε συστήματος με τον χρήστη. Συγκεκριμένα, στα συστήματα προτάσεων η αλληλεπίδραση με τον χρήστη είναι πιο δυναμική. Το σύστημα προτείνει αυτόματα αντικείμενα ή πληροφορίες, λαμβάνοντας υπόψη τις προηγούμενες προτιμήσεις και συμπεριφορές του χρήστη. Αυτό σημαίνει ότι οι προτάσεις προσαρμόζονται σταδιακά στις ανάγκες και τις προτιμήσεις του χρήστη, προσφέροντας ένα εξατομικευμένο επίπεδο εξυπηρέτησης, δρώντας περισσότερο ως πράκτορας (information agent) παρά ως ως ένα εργαλείο που παρέχει βασικές απαντήσεις σε συγκεκριμένα ερωτήματα.

Σύστημα Ανάκτησης Πληροφοριών (IR):

Τα συστήματα ανάκτησης πληροφοριών είναι σχεδιασμένα για την αναζήτηση και την εξαγωγή πληροφοριών από μεγάλα σύνολα δεδομένων. Συνήθως χρησιμοποιούνται όταν ο χρήστης έχει ένα συγκεκριμένο ερώτημα ή ανάγκη για

πληροφορίες. Τα συστήματα ανάκτησης πληροφοριών επιστρέφουν ένα σύνολο εγγραφών που πιθανώς να ανταποκρίνονται στο ερώτημα του χρήστη.

Σύστημα Προτάσεων (RS):

Τα συστήματα προτάσεων επικεντρώνονται στην παροχή συστάσεων στον χρήστη, συχνά με σκοπό τη βελτίωση της εμπειρίας του ή την αύξηση της απόδοσης του σε κάποιον τομέα. Αυτά τα συστήματα αξιοποιούν συνήθως προσωπικά δεδομένα του χρήστη, όπως ιστορικό αγορών, αξιολογήσεις ή προτιμήσεις, προκειμένου να προτείνουν περιεχόμενο ή προϊόντα που πιθανώς να τον ενδιαφέρουν.

Τύποι συστημάτων συστάσεων

Υπάρχουν διάφορες μέθοδοι για την υλοποίηση των συστημάτων προτάσεων. Οι δύο πιο βασικές προσεγγίσεις που συναντούνται πιο συχνά στην βιβλιογραφία είναι το συνεργατικό φιλτράρισμα (collaborative filtering) και το φιλτράρισμα βάση περιεχομένου (content-based filtering). Οι υπόλοιπες προσεγγίσεις συχνά προσπαθούν να συνδυάσουν τα πλεονεκτήματα αυτών των δύο προσεγγίσεων, συνδυάζοντας αυτές [3] ή ενσωματώνοντας και άλλες τεχνικές.

Συνεργατικό Φιλτράρισμα

Το συνεργατικό φιλτράρισμα αναφέρεται στην προσέγγιση όπου τα προτεινόμενα προκύπτουν με βάση την γνώμη άλλων χρηστών που έχουν παρόμοιες προτιμήσεις [3.5].

Στον τομέα του συνεργατικού φιλτραρίσματος, έχουν αναπτυχθεί συστήματα που βοηθούν στη σύσταση αντικειμένων με βάση τη συμπεριφορά του χρήστη.

Τα συστήματα βασισμένα στη μνήμη χρησιμοποιούν τις αξιολογήσεις των χρηστών για να εντοπίσουν ομοιότητες και να προτείνουν αντικείμενα. Ωστόσο, η επιλογή του αριθμού των παρόμοιων χρηστών που θα ληφθούν υπόψη μπορεί να αποτελέσει μια πρόκληση [5].

Άλλα συστήματα ομαδοποιούν χρήστες με παρόμοια μοτίβα αξιολόγησης και χρησιμοποιούν πολύπλοκες τεχνικές για να μάθουν από αυτά τα δεδομένα [5]. Ωστόσο, η προσαρμογή του συστήματος και η λειτουργία του σε διάφορες καταστάσεις μπορεί να είναι δύσκολη.

Φιλτράρισμα Περιεχομένου

Το φιλτράρισμα περιεχομένου χρησιμοποιεί το ιστορικό του χρήστη προκειμένου να παρέχει προτάσεις με παρόμοια χαρακτηριστικά των προτιμήσεων του χρήστη. [3].

Προφίλ στα Συστήματα Προτάσεων

Η βάση για τα προσωποποιημένα συστήματα προτάσεων αποτελεί η δημιουργία προφίλ. Ο στόχος της δημιουργίας ενός προφίλ είναι η συλλογή περιγραφικών κριτηρίων σχετικά με τις πληροφορίες που χειρίζεται το σύστημα, προκειμένου να παρέχει εξατομικευμένες συστάσεις που προσαρμόζονται στα ειδικά χαρακτηριστικά κάθε χρήστη.

Το προφίλ στα συστήματα προτάσεων αναφέρεται στη διαδικασία συλλογής και ανάλυσης πληροφοριών χρήστη για τη δημιουργία ενός προφίλ που αντικατοπτρίζει τις προτιμήσεις, τα ενδιαφέροντα και τα χαρακτηριστικά του.

Υπάρχουν δύο μορφές συλλογής πληροφοριών χρήστη για το προφίλ: η άμεση και η έμμεση συλλογή δεδομένων.

Η άμεση συλλογή δεδομένων περιλαμβάνει την έκφραση των ενδιαφερόντων και των προτιμήσεων του χρήστη, όπως η αξιολόγηση αντικειμένων ή η παροχή ανατροφοδότησης για την εκτίμηση ενός αντικειμένου.

Η έμμεση συλλογή δεδομένων, από την άλλη, περιλαμβάνει την παρατήρηση και ανάλυση της συμπεριφοράς του χρήστη στην εφαρμογή χωρίς να ζητείται ρητά η άποψη του χρήστη.

Το προφίλ βοηθάει στη βελτίωση της ακρίβειας και της συνάφειας των προτάσεων κατανοώντας τις προτιμήσεις του χρήστη και προσαρμόζοντας ανάλογα τις συστάσεις.

ΚΕΦΑΛΑΙΟ 3 - ΜΕΘΟΔΟΛΟΓΙΑ

Το παρόν κεφάλαιο επικεντρώνεται στη μεθοδολογία που ακολουθήθηκε για την ανάπτυξη του συστήματος προτάσεων με τη χρήση γνωσιακού γράφου και του μοντέλου Weighted Sum.

Περιγραφή της μεθοδολογίας

Επιλογή θέματος

Η επιλογή του τουρισμού ως κύριου θέματος πραγματοποιήθηκε λόγω του εύρους των πληροφοριών που κυκλοφορούν στο διαδίκτυο και της ανάγκης αξιοποίησης τους προς όφελος του χρήστη.

Ο τουρισμός, ήταν ανέκαθεν ένας από τους βασικότερους πυλώνες της οικονομίας και πολλές αναζητήσεις από αυτές που πραγματοποιούνται καθημερινά στο διαδίκτυο από τους χρήστες σχετίζονται με την εύρεση του καταλληλότερου τουριστικού προορισμού, ενός προορισμού που θα καλύπτει τα περισσότερα, αν όχι όλα, από τα ενδιαφέροντα και τις προτιμήσεις τους.

Οι χρήστες του διαδικτύου βρίσκονται καθημερινά αντιμέτωποι με πληθώρα πληροφοριών και πολλούς και διαφορετικούς τύπους δεδομένων. Τα δεδομένα είναι κατά βάση αδόμητα καθώς οι περισσότερες πληροφορίες παρέχονται σε μορφή κειμένου. Τον τελευταίο καιρό, η κατάσταση έχει βελτιωθεί και οι χρήστες βρίσκουν την απάντησή τους εύκολα λόγω της εξέλιξης των μηχανών αναζήτησης και της πρόσφατης ανάπτυξης του τομέα της τεχνητής νοημοσύνης, που ενσωματώνει στα αποτελέσματα αναζήτησης την παραγωγή αυτοματοποιημένου κειμένου με σύντομη και εύστοχη διατύπωση στην ερώτησή τους (chatbots).

Η αναζήτηση τουριστικών πληροφοριών είναι μία από τις πιο δύσκολες αναζητήσεις στο διαδίκτυο λόγω της πληθώρας πληροφοριών και του τρόπου αναπαραγωγής τους. Οι χρήστες επιθυμούν να πληροφορηθούν για ζητήματα όπως είναι το ποια είναι η καλύτερη περίοδος να επισκεφθεί κάποιος έναν συγκεκριμένο προορισμό, τι αξίζει να δει κάποιος επισκεπτόμενος για πρώτη φορά ένα μέρος καθώς και ποιες δραστηριότητες είναι αυτές που θα του δώσουν το 100% της ταξιδιωτικής εμπειρίας που αναζητά. Όλες οι απαντήσεις που μπορούν να δωθούν σε αυτά τα ερωτήματα στηρίζονται σε θέματα εμπειριών οι οποίες πέραν του ότι είναι υποκειμενικές, τις περισσότερες φορές δεν παρουσιάζονται με πλήρη τρόπο.

Αυτό που λείπει από την διαδικασία αναζήτησης είναι η προσωποποίηση των αποτελεσμάτων. Επομένως, η επιλογή του συγκεκριμένου θέματος πραγματοποιήθηκε έπειτα από πρόκληση να δημιουργηθεί μία εφαρμογή που μέσα από ένα μοντέλο που λαμβάνει υπόψη όσο το δυνατόν περισσότερες παραμέτρους του χρήστη, θα προσφέρει μία ξεχωριστή για τον καθένα απάντηση, προσαρμοσμένη στις προτιμήσεις και τις απαιτήσεις του.

Ακόμη, είναι ένας τομέας στον οποίο υπάρχει ζήτηση, επιθυμία για έμπιστη πληροφόρηση ενώ η ανάγκη για παροχή εξατομικευμένου περιεχομένου συνεχώς αυξάνεται, καθώς αυξάνονται οι προσδοκίες του χρήστη μέσα και από την εμπειρία του από άλλες ήδη υπάρχουσες εφαρμογές που παρέχουν εξατομικευμένο περιεχόμενο και έχουν ενσωματωθεί στην καθημερινότητά του.

Καθορισμός βασικών οντοτήτων

Κάθε σύστημα λογισμικού δημιουργείται προκειμένου να εξυπηρετήσει σε ένα σκοπό. Τα συστήματα συστάσεων έχουν δημιουργηθεί προκειμένου να εξυπηρετήσουν τους χρήστες. Λόγω του ανθρωποκεντρικού χαρακτήρα των συστημάτων προτάσεων μία από τις βασικές οντότητες του συστήματος που δημιουργήθηκε είναι ο άνθρωπος-χρήστης για τον οποίο προορίζεται.

Με τον καθορισμό του χρήστη σαν κεντρική οντότητα του συστήματος, επόμενη οντότητα στην ιεραρχία αποτελεί η οντότητα που θα προτείνεται στους χρήστες. Η οντότητα αυτή είναι οι ταξιδιωτικοί προορισμοί οι οποίοι αφορούν τοποθεσίες σε διάφορες χώρες, οι οποίες ξεχωρίζουν για τις εμπειρίες που προσφέρουν στους επισκέπτες τους.

Σε λογική συνέχεια της ύπαρξης της οντότητας 'Ταξιδιωτικός Προορισμός' προστέθηκαν στις οντότητες η οντότητα Χώρα στην οποία μπορεί να ανήκει κάθε προορισμός και η οντότητα Ήπειρος η οποία μπορεί να περιλαμβάνει πολλές χώρες.

Αναζήτηση Δεδομένων

Στο διαδίκτυο υπάρχουν άφθονες πληροφορίες σχετικές με ταξιδιωτικούς προορισμούς που αξίζουν την προσοχή όσων είναι λάτρεις των ταξιδιών. Υπάρχουν τόσο ενημερωτικά site ξεχωριστά για κάθε προορισμό, όπως είναι για παράδειγμα το site <https://www.thisisathens.org/> όπου προσφέρει μία σφαιρική εικόνα όλων όσων προσφέρει η Αθήνα στους επισκέπτες της, αλλά και site που αναφέρονται αποκλειστικά σε χώρες όπως είναι το <https://www.discovergreece.com/>, που αποτελεί κτήμα του φορέα που ασχολείται με τον τουρισμό εντός των συνόρων της Ελλάδας. Οι επιλογές όσον αφορά την αναζήτηση τουριστικών πληροφοριών είναι αμέτρητες και πολλές από τις πηγές που παρουσιάζονται στο διαδίκτυο είναι προϊόν διαφήμισης των τουριστικών προορισμών τόσο για δημόσιο όσο και για ιδιωτικό συμφέρον.

Αναφέρονται ενδεικτικά διαδικτυακές πλατφόρμες, γνωστές στον χώρο του τουρισμού, οι οποίες χρησιμοποιήθηκαν στα πλαίσια διερεύνησης των παρεχόμενων τουριστικών πληροφοριών στους χρήστες του διαδικτύου:

Trip Advisor

Το TripAdvisor είναι μια διαδικτυακή πλατφόρμα που προσφέρει πληροφορίες σχετικά με ταξίδια και αξιολογήσεις για ξενοδοχεία, εστιατόρια, αξιοθέατα, και δραστηριότητες σε διάφορες πόλεις και προορισμούς.

Συνολικά, το TripAdvisor παρέχει μια ολοκληρωμένη πλατφόρμα για τους ταξιδιώτες που τους βοηθά να προγραμματίσουν, να αξιολογήσουν και να μοιραστούν τις ταξιδιωτικές τους εμπειρίες. Τα TripAdvisor Travelers' Choice Awards είναι ετήσια βραβεία που απονέμονται σε δημοφιλείς προορισμούς, ξενοδοχεία, εστιατόρια και αξιοθέατα, βασισμένα στις αξιολογήσεις και τις κριτικές των χρηστών του TripAdvisor. Τα βραβεία διακρίνουν τους κορυφαίους σε διάφορες κατηγορίες και παρέχουν έναν πολύτιμο οδηγό για τους ταξιδιώτες.

Booking

Το Booking.com είναι μια διαδικτυακή πλατφόρμα που παρέχει εκτενείς πληροφορίες και κριτικές σχετικά με καταλύματα, όπως ξενοδοχεία, διαμερίσματα, βίλες και άλλα τύπους καταλυμάτων, σε διάφορα μέρη του κόσμου. Οι χρήστες μπορούν να βρουν πληροφορίες σχετικά με τιμές, διαθεσιμότητα, ανέσεις και φωτογραφίες, παρέχοντας μια ολοκληρωμένη εικόνα για κάθε κατάλυμα. Ακόμα, το Booking.com επίσης προσφέρει ευέλικτες επιλογές κρατήσεων για συμμετοχή σε δραστηριότητες που μπορεί να ταιριάζουν με τα ενδιαφέροντα που έχει ο κάθε χρήστης. Μερικές από αυτές περιλαμβάνουν, περιηγήσεις με θέμα την γαστρονομία, πεζοπορίες στην ύπαιθρο, εισιτήρια για συναυλίες ή άλλα ψυχαγωγικά θεάματα.

Δημιουργία προφίλ χρήστη

Τα ενδιαφέροντα του κάθε χρήστη ποικίλουν από άτομο σε άτομο ενώ οι προορισμοί οι οποίοι μπορούν να υποστηρίξουν δραστηριότητες που καλύπτουν κάθε πτυχή των ενδιαφερόντων ενός ταξιδιώτη είναι δυσεύρετοι.

Κατανοώντας ότι οι ταξιδιώτες έχουν διάφορες προτιμήσεις και ανάγκες, είναι σημαντικό να ληφθούν υπόψη διάφορα προφίλ χρηστών κατά την ανάπτυξη ενός συστήματος πληροφοριών για ταξίδια. Ενδεικτικά, αναφέρονται κάποια από τα βασικότερα προφίλ που μπορούν να ταυτιστούν με λάτρεις των ταξιδιών.

Περιπετειώδεις Ταξιδιώτες:

Οι περιπετειώδεις ταξιδιώτες αναζητούν αδρεναλίνη και δραστηριότητες όπως ορειβασία και άλλα extreme sport.

Πολιτισμικοί Εξερευνητές:

Οι πολιτισμικοί εξερευνητές ενδιαφέρονται για μνημεία, μουσεία και τοπικές παραδόσεις. Ψάχνουν αυθεντικές πολιτιστικές εμπειρίες που τους επιτρέπουν να εξερευνηθούν τον τοπικό τρόπο ζωής.

Λάτρεις της Φύσης:

Οι λάτρεις της φύσης αναζητούν εκπληκτικά φυσικά τοπία, εθνικά πάρκα και δραστηριότητες στη φύση.

Γαστρονόμοι:

Οι γαστρονόμοι αναζητούν τοπικές γεύσεις και γαστρονομικές εμπειρίες. Ενδιαφέρονται για εστιατόρια υψηλής κουζίνας, τοπικά προϊόντα και γαστρονομικές εκδηλώσεις.

Ταξιδιώτες Χαλάρωσης:

Οι ταξιδιώτες χαλάρωσης αναζητούν παραλίες, θερινά θέρετρα και χαλαρωτικές δραστηριότητες.

Αυτά είναι μόνο μερικά από τα προφίλ ταξιδιωτών που αναγνωρίστηκαν. Οι ταξιδιώτες μπορεί να έχουν πολλαπλές πτυχές και ενδιαφέροντα, και τα προφίλ αυτά δεν είναι απολύτως διακριτά. Οι συνδυαστικοί ταξιδιώτες είναι ένα παράδειγμα του πώς κάποιοι επιθυμούν να συνδυάσουν πολλές εμπειρίες κατά τη διάρκεια του ταξιδιού τους.

Συλλογή πιθανών ενδιαφερόντων χρήστη

Η κατάταξη των προορισμών ανά κατηγορία ενδιαφερόντων αποτελεί ένα κρίσιμο βήμα για την ανάπτυξη ενός συστήματος πληροφοριών για ταξίδια. Η διαίρεση των προορισμών σε κατηγορίες σύμφωνα με τα ενδιαφέροντα των χρηστών επιτρέπει την παροχή εξατομικευμένων και στοχευμένων πληροφοριών.

Από τα προφίλ που αναλύθηκαν παραπάνω προκύπτουν τα εξής ενδιαφέροντα:

Φαγητό: Οι χρήστες που ενδιαφέρονται για το φαγητό αναζητούν νέες γευστικές εμπειρίες. Μπορεί να αγαπούν τη γαστρονομία διάφορων κουζινών, ψάχνοντας για συνταγές, εστιατόρια, και νέα φαγητά που προκαλούν τις αισθήσεις τους.

Κουλτούρα: Οι λάτρεις της κουλτούρας εκτιμούν την τέχνη, τη λογοτεχνία, το θέατρο, τη μουσική και τις τέχνες γενικά. Είναι πιθανό να επισκέπτονται μουσεία, θεατρικές παραστάσεις, συναυλίες και να αναζητούν νέες δημιουργικές εκδηλώσεις.

Φύση και περιπέτεια: Όσοι αγαπούν τη φύση μπορεί να είναι φανατικοί των υπαίθρων. Απολαμβάνουν πεζοπορία, κάμπινγκ, ποδηλασία και γενικά τις δραστηριότητες που τους φέρνουν κοντά στο φυσικό περιβάλλον.

Κατασκευή Γνωσιακού Γράφου

Σε πρώτο στάδιο, ο καθορισμός των βασικών οντοτήτων ήταν απαραίτητος προκειμένου να καθοριστούν οι βασικοί στόχοι της διπλωματικής εργασίας. Με την αύξηση του αριθμού των οντοτήτων αυξάνεται και ο αριθμός των πληροφοριών που πρέπει να είναι διαθέσιμες προκειμένου να ξεκινήσει η υλοποίηση του συστήματος.

Η αναγνώριση των ενδιαφερόντων και η κατηγοριοποίηση τους οδήγησε στην δημιουργία μιας νέας οντότητας που συνδυάζοντας τα ενδιαφέροντα του χρήστη και τους ταξιδιωτικούς προορισμούς μπορεί να εξασφαλίσει περισσότερο εξατομικευμένες προτάσεις και συστάσεις.

Επομένως, με την ενσωμάτωση των πληροφοριών που παρείχαν οι πλατφόρμες TripAdvisor, Booking και σε συνδυασμό με τις πληροφορίες που παρείχαν άλλες αντίστοιχες ταξιδιωτικές σελίδες σχηματίστηκε ο γνωσιακός γράφος με βασικές οντότητες να αποτελούν ο χρήστης, οι ταξιδιωτικοί προορισμοί και τα ενδιαφέροντα.

Ακολουθεί η λίστα με τα ενδιαφέροντα τα οποία ενσωματώθηκαν στον γνωσιακό γράφο και συμπλήρωσαν τους τύπους οντοτήτων που συναντιούνται σε αυτόν:

Γαστρονομία	Ανακαλύπτοντας την τοπική κουζίνα και τα εστιατόρια.
Πολιτισμός	Ανακαλύπτοντας την τοπική τέχνη, παραδόσεις και εκδηλώσεις.
Νυχτερινή Ζωή	Απολαύστε τη νυχτερινή ζωή, μπαρ και κλαμπ.
Ιστορία	Ανακαλύπτοντας την ιστορία και την πολιτιστική κληρονομιά της πόλης.
Φύση	Ανακαλύπτοντας τη φυσική ομορφιά της πόλης.
Αγορές	Ανακαλύπτοντας τα τοπικά καταστήματα και τις αγορές.
Περιπέτεια	Ανακαλύπτοντας τις υπαίθριες δραστηριότητες και τα αθλήματα.
Οικογένεια	Ανακαλύπτοντας τις δραστηριότητες και τα αξιοθέατα φιλικά προς την οικογένεια.
Ρομαντικό	Ανακαλύπτοντας τα ρομαντικά μέρη και δραστηριότητες.
Χριστούγεννα	Ανακαλύπτοντας τις αγορές και τις εκδηλώσεις των Χριστουγέννων.

Ενσωμάτωση Budget

Όταν ένας χρήστης κάνει την επιλογή του επόμενου ταξιδιωτικού προορισμού του, υπάρχουν διάφορα πράγματα που θα μπορούσαν να τον προτρέψουν (ή και να τον αποτρέψουν) από το να επισκεφθεί τελικά αυτόν τον προορισμό. Τέτοια είναι το κόστος του ταξιδιού το οποίο ανάλογα και με το μεταφορικό μέσο που πρόκειται να ταξιδέψει ο χρήστης μπορεί να είναι ανάλογο της χιλιομετρικής απόστασης. Ακόμα, πολύ σημαντικό είναι το επίπεδο της ασφάλειας της χώρας της οποίας επισκέπτεται.

Σε αυτό το πλαίσιο, για κάθε χώρα πραγματοποιείται κατηγοριοποίηση ανάλογα με το κόστος. Κάθε χώρα μπορεί να κατηγοριοποιείται στις 3 παρακάτω βασικές κατηγορίες:

Middle Range (Μεσαίο Εύρος):

Προορισμοί σε αυτήν την κατηγορία παρέχουν ικανοποιητικό επίπεδο ποιότητας υπηρεσιών και εγκαταστάσεων με μέτριες τιμές. Κατάλληλοι για ταξιδιώτες που επιθυμούν ισορροπία μεταξύ ποιότητας και οικονομίας.

Low Budget (Χαμηλό Προϋπολογισμό):

Προορισμοί με χαμηλό κόστος που προσφέρουν βασικές υπηρεσίες και εγκαταστάσεις με χαμηλές τιμές. Ιδανικοί για ταξιδιώτες που θέλουν να εξοικονομήσουν χρήματα και να επικεντρωθούν σε πιο οικονομικές δραστηριότητες.

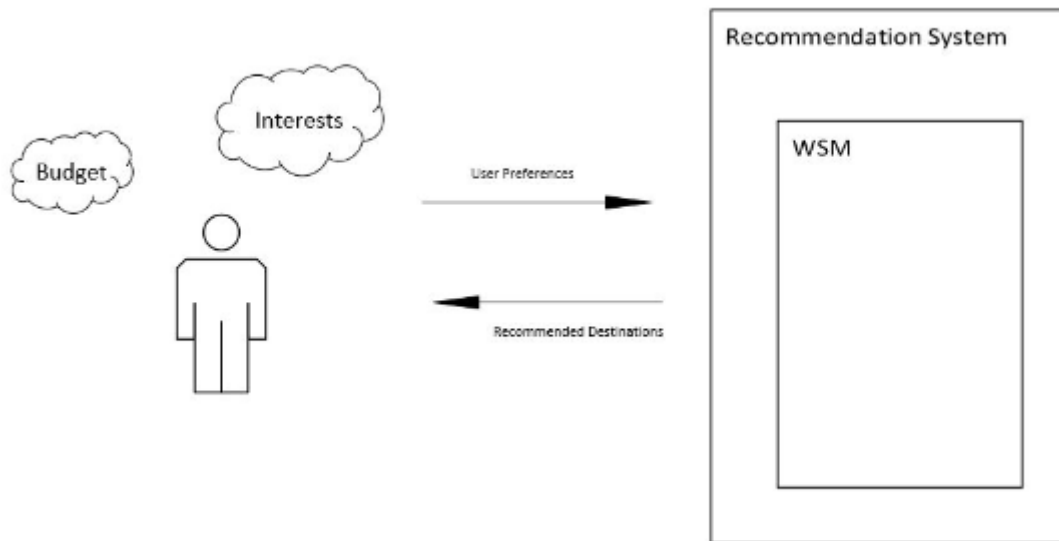
High Range (Υψηλό Εύρος):

Προορισμοί που προσφέρουν υψηλής ποιότητας υπηρεσίες, πολυτέλεια και εγκαταστάσεις με αντίστοιχα υψηλό κόστος. Κατάλληλοι για ταξιδιώτες που αναζητούν πολυτέλεια, άνεση και προνομιακές εμπειρίες.

Στόχος παντα είναι να παρέχεται στον χρήστη ένα ευρύ φάσμα προορισμών που ανταποκρίνονται στις προσωπικές του προτιμήσεις και διαθέσιμες επιλογές, δημιουργώντας μια εξατομικευμένη εμπειρία.

Ενσωμάτωση Μοντέλου Weighted Sum

Το μοντέλο Weighted Sum θα ενσωματωθεί στο σύστημα προτάσεων για τον υπολογισμό των συνολικών βαρών για κάθε πρόταση. Αυτό θα επιτρέψει τη βελτίωση της ακρίβειας και της εξατομικεύσης των προτάσεων.



Αποφασίστηκε ότι το σύστημα θα πρέπει με κάποιο τρόπο να αλληλεπιδρά με τον χρήστη. Ο χρήστης αποφασίζει ο ίδιος ποια πολιτική απόφασης θα ακολουθήσει ορίζοντας τα βάρη προτίμησης του προς τα ενδιαφέροντα και καθορίζοντας ποιος τύπος budget θέλει να ληφθεί υπόψη από το σύστημα στην προβολή των προτεινόμενων.

ΚΕΦΑΛΑΙΟ 4 - Αρχιτεκτονική Συστήματος

Στο παρόν κεφάλαιο παρουσιάζεται η αρχιτεκτονική του αναπτυγμένου συστήματος με επίκεντρο τη λεπτομερή ανάλυση των βασικών δομικών του στοιχείων. Ανάμεσα σε αυτά συγκαταλέγονται το μοντέλο, πάνω στο οποίο στηρίζεται το σύστημα για την παροχή συστάσεων με ενσωματωμένο το μοντέλο WSM, και η διεπαφή του χρήστη, η οποία παρέχει τη δυνατότητα αλληλεπίδρασης με το σύστημα προτάσεων και τον εντοπισμό νέου περιεχομένου. Ακόμα βασικό στοιχείο αποτελεί ο γνωσιακός γράφος, ο οποίος έχει δημιουργηθεί από την αρχή, χρησιμοποιώντας δεδομένα από διάφορες πηγές στο διαδίκτυο και έχει την δομή που περιγράφηκε στο κεφάλαιο 3, με τίτλο Μεθοδολογία.

Η διεπαφή χρήστη υλοποιήθηκε ως μια εφαρμογή κινητής συσκευής Android, χρησιμοποιώντας τη γλώσσα προγραμματισμού Kotlin, η οποία αποτελεί την επίσημη γλώσσα προγραμματισμού της Google για την ανάπτυξη κινητών συσκευών Android.

Ο γνωσιακός γράφος, αναπαριστάται με την βοήθεια του εργαλείου Neo4j. Το Neo4j είναι ένα σύστημα διαχείρισης γραφημάτων που αναπτύχθηκε από την Neo4j, Inc.

Τα στοιχεία που αναφέρθηκαν αποτελούν ανεξάρτητες μονάδες, που όμως για την επιτυχή λειτουργία του συστήματος χρειάζεται να επικοινωνούν μεταξύ τους. Η επικοινωνία, σε ότι αφορά την πλευρά συστήματος συστάσεων – γνωσιακού γράφου πραγματοποιείται χρησιμοποιώντας τους drivers της βάσης Neo4j. Στην πλευρά διεπαφής χρήστη – συστήματος συστάσεων τα δεδομένα μεταφέρονται με χρήση ενός REST API. Το σύστημα συστάσεων είναι γραμμένο στην γλώσσα προγραμματισμού Python, που παρέχει βιβλιοθήκες όπως το Flask, για την φιλοξενία του συστήματος συστάσεων σε έναν server αλλά και άλλες βιβλιοθήκες όπως είναι η NumPy που βοηθά στην ενσωμάτωση του μοντέλου WSM.

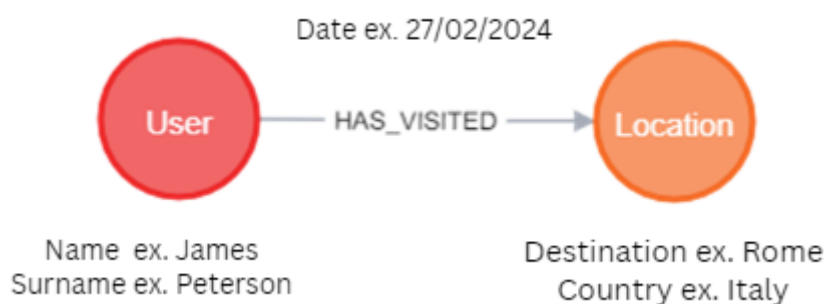
Γνωσιακός γράφος με το Neo4j

Η Neo4j είναι μια βάση δεδομένων γράφου, η οποία είναι ιδανική για αναπαραστάσεις δεδομένων με σχέσεις που μπορούν να παρουσιαστούν καλύτερα ως γράφοι. Η επιλογή της βάσης Neo4j έναντι μιας σχεσιακής βάσης δεδομένων, όπως η SQL όπου τα δεδομένα αποθηκεύονται με την μορφή πινάκων [βιβλιοβασεων], οφείλεται στην διευκόλυνση που προσφέρει η βάση Neo4j αφού εκεί τα δεδομένα αποθηκεύονται σε μορφή γράφου [neo4j].

Η αναπαράσταση ενός μοντέλου σε γράφο αποτελεί μία φιλική προς αναπαράσταση δομή. Οι κόμβοι αντιπροσωπεύουν τις οντότητες ενός συστήματος ενώ οι σχέσεις μεταξύ των οντοτήτων αναπαρίστανται χρησιμοποιώντας ακμές. Στους κόμβους δίνεται η δυνατότητα να αποθηκευτούν δεδομένα που είναι χαρακτηριστικά της κάθε οντότητας που αναπαρίστανται [neo4j]. Στις ακμές μπορούν να χρησιμοποιηθούν ετικέτες προκειμένου να μπορούν να διακριθούν εύκολα από άλλες ακμές. Φυσικά, μπορούν να χρησιμοποιηθούν ακμές

διαφορετικού τύπου. Ο τύπος της κάθε ακμής καθορίζεται τόσο από το λεκτικό της ετικέτας δηλαδή τον τύπο της, όσο και από τα επιπρόσθετα χαρακτηριστικά που μπορούν να αποδωθούν σε αυτήν.

Για παράδειγμα, μιλώντας για έναν γράφο όπου αναπαρίσταται το ιστορικό ενός χρήστη σε ταξιδιωτικούς προορισμούς, έχουμε ως βασικούς κόμβους τις κύριες οντότητες του μοντέλου που είναι ο χρήστης ή οι χρήστες και ο προορισμός ή οι προορισμοί στους οποίους αναφερόμαστε. Ως σύνδεσμο έχουμε την ακμή τύπου 'HAS VISITED'.



Όπως βλέπουμε, για τον κόμβο του χρήστη (User) και κάθε κόμβο τύπου χρήστη, έχουν αποθηκευτεί τα βασικά στοιχεία του όπως είναι το όνομα και το επίθετο του. Αντίστοιχα, ο προορισμός μοντελοποιείται ως κόμβος με κύρια στοιχεία να είναι το όνομα του προορισμού καθώς και η χώρα στην οποία βρίσκεται. Όσον αφορά την ακμή που συνδέει αυτές τις δύο κορυφές, αντιπροσωπεύεται από την ετικέτα 'HAS VISITED' και άμεσα συσχετίζει το στιγμιότυπο της κλάσης-κόμβου Χρήστη με το αντίστοιχο στιγμιότυπο του κόμβου Προορισμός. Είναι λογικό στην συγκεκριμένη σχέση να θέλουμε να έχουμε ως πληροφορία την ημερομηνία επίσκεψης του χρήστη στον συγκεκριμένο προορισμό. Αυτό μπορεί εύκολα να πραγματοποιηθεί ενσωματώνοντας το πεδίο 'date' στην ακμή.

Η βάση Neo4j για την ανάκτηση των δεδομένων χρησιμοποιεί ερωτήματα (queries) όπως και η βάση SQL. Όμως, πρόκειται για διαφορετικού είδους δομές ερωτημάτων λόγω της φύσης των δεδομένων. Συγκεκριμένα στην βάση Neo4j αποθηκεύονται κυρίως ημιδομημένα δεδομένα. Για παράδειγμα, για την ανάκτηση των δεδομένων της εικόνας 4.1 θα χρησιμοποιούσαν το query:

```
MATCH (user:User)-[visited:HAS_VISITED]->(destination:Destination)
```

```
RETURN user, visited.date, destination;
```

Η δομή του παραπάνω query κάνει την βάση Neo4j ιδανική για ανάκτηση δεδομένων με σύνθετες και πολλαπλές ιδιότητες.

Μοντελοποίηση

Όπως ακριβώς ορίζεται στο προηγούμενο παράδειγμα (εικόνα 4.1) η διμελής σχέση μεταξύ Χρήστη-Προορισμού για την μοντελοποίηση του ιστορικού επίσκεψης ενός χρήστη σε προορισμούς, έτσι μπορεί να οριστεί οποιαδήποτε άλλη διμελής σχέση μεταξύ στιγμιοτύπων των κόμβων, λαμβάνοντας υπόψη ότι ο γράφος μπορεί να επεκταθεί προσθέτοντας και άλλους τύπους κόμβους και συνδέσεων μεταξύ αυτών.

Η μοντελοποίηση είναι ένα από τα βασικότερα στάδια στην διαδικασία επίλυσης ενός προβλήματος και κατά συνέπεια στην διαδικασία ανάπτυξης ενός συστήματος συστάσεων.

Ακολουθεί μια συνοπτική περιγραφή των κόμβων οντοτήτων που έχουν χρησιμοποιηθεί:

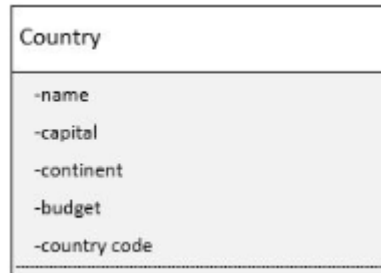
Η κλάση Προορισμών - Destination

Ξεκινώντας με την κλάση προορισμών, η οντότητα Προορισμός πρεσβεύει αυτό που τελικά προτείνει το σύστημα προτεινόμενων στον χρήστη. Πρόκειται για προορισμό κατάλληλο για τον χρήστη, που πληρεί τα κριτήρια του και που υπάρχουν δραστηριότητες που καλύπτουν τα ενδιαφέροντα του.



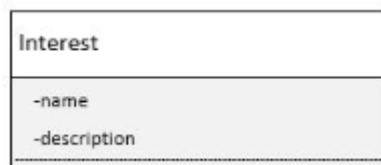
Η κλάση Χωρών - Country

Η κλάση χώρα απεικονίζει την γεωγραφική τοποθεσία του κάθε προτεινόμενου προορισμού. Κάθε χώρα σχετίζεται με διαφορετικούς προορισμούς. Τα χαρακτηριστικά της χώρας είναι:



Η κλάση Ενδιαφερόντων - Interest

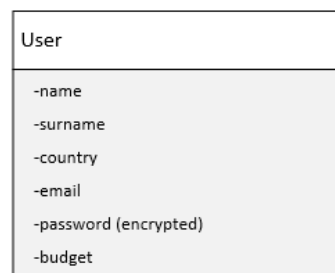
Η κλάση Interest απεικονίζει τα ενδιαφέροντα του χρήστη. Έχουν υλοποιηθεί σε ξεχωριστή οντότητα προκειμένου να διευκολυνθεί η σύνδεση τους με το βάρος που τους αποδίδει ο κάθε χρήστης. Ακόμα σημαντική είναι η σύνδεση του ενδιαφέροντος με τον κάθε προορισμό καθώς αποτελεί έναν από τους παράγοντες που συντελούν στον υπολογισμό των μετρικών με σκοπό την βελτίωση των αποτελεσμάτων στην προβολή των προτεινόμενων.



Η κλάση Χρηστών - User

Η κλάση Χρήστης αποτελεί την πιο σημαντική οντότητα του συστήματος. Η οντότητα του χρήστη είναι σχεδιασμένη έτσι ώστε να μπορεί να υλοποιηθεί η αυθεντικοποίηση του χρήστη και η εφαρμογή να είναι στα πρότυπα μιας σύγχρονης εφαρμογής με μοντέρνα διεπαφή που παράλληλα υποστηρίζει την δημιουργία ενός προσωπικού λογαριασμού για τον χρήστη και την δυνατότητα σύνδεσης σε αυτόν.

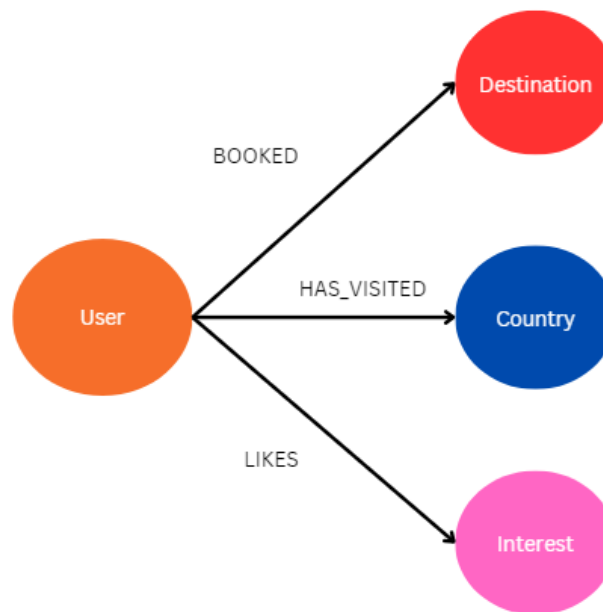
Περιλαμβάνει τα εξής χαρακτηριστικά:



Μέσα από την δημιουργία ενός λογαριασμού ο χρήστης έχει την δυνατότητα να ρυθμίσει παραμέτρους στα ενδιαφέροντα του, να μπορέσει να φιλτράρει τα

αποτελέσματα που του εμφανίζονται στα προτεινόμενα με κριτήρια όπως τον προτιμώμενο προϋπολογισμό.

Η οντότητα του χρήστη θεωρείται η οντότητα με τον μεγαλύτερο αριθμό εξερχόμενων ακμών. Συγκεκριμένα, λόγω του ότι το σύστημα έχει στηθεί για χρήση από τους χρήστες, οι ακμές που ξεκινούν από τον κόμβο του χρήστη και καταλήγουν σε γειτονικούς κόμβους χρησιμοποιούνται για την καταγραφή πληροφοριών που αφορούν κάθε χρήστη.



Όπως φαίνεται στην εικόνα, ο χρήστης μπορεί να συνδεθεί με πολλούς κόμβους.

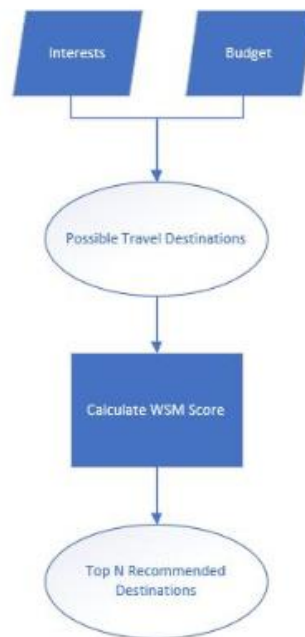
- Η βασικότερη ακμή είναι η ακμή 'LIKES' που συσχετίζει τον κάθε χρήστη με το ενδιαφέρον το οποίο φέρεται να έχει. Έχει μια επιπρόσθετη ιδιότητα που εκφράζει το ποσοστό προτίμησης του συγκεκριμένου ενδιαφέροντος από τον κάθε χρήστη. Η ακμή LIKES αποτελεί μέρος του μοντέλου WSM που αναλύεται στην συνέχεια.
- Η ακμή 'VISITED' καταγράφει το ιστορικό επισκέψεων του χρήστη σε κάθε χώρα.
- Η ακμή 'BOOKED' αναφέρεται στην πρόθεση του χρήστη να ταξιδέψει σε συγκεκριμένο προορισμό στο άμεσο μέλλον. Η ημερομηνία επίσκεψης του χρήστη στον συγκεκριμένο προορισμό καθώς και η διάρκεια της παραμονής του σε αυτόν ενσωματώνεται σαν πληροφορία στα στοιχεία-πεδία που συνοδεύουν την κάθε ακμή.

Σύστημα προτάσεων με χρήση του WSM

Ο σκοπός ενός συστήματος προτεινόμενων είναι να παράγει προτάσεις για τους χρήστες. Εάν εξετάσουμε το πρόβλημα της παραγωγής προτάσεων ως την

αναζήτηση των ταξιδιωτικών προορισμών που ταιριάζουν καλύτερα με τα ενδιαφέροντα του χρήστη, τότε απαιτείται σε πρώτο στάδιο η εύρεση ενός τρόπου μέτρησης του κατά πόσο ένας προορισμός είναι σχετικός με τα ενδιαφέροντα του χρήστη και ανταποκρίνεται σε αυτά. Άρα το πρόβλημα αυτό μπορεί να αναχθεί σε πρόβλημα ταξινόμησης του κάθε προορισμού με βάση αυτή την μετρική.

Στην εικόνα φαίνεται καθαρά η δομή του συστήματος προτάσεων:



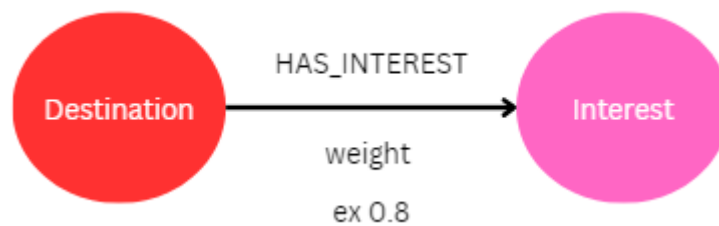
Το ενσωματωμένο Μοντέλο WSM

Η εφαρμογή του Multiple Criteria Decision Analysis (MCDA) στα Recommender Systems είναι μια ενδιαφέρουσα προσέγγιση που μπορεί να βοηθήσει στην προσαρμογή των συστημάτων συστάσεων στις προτιμήσεις των χρηστών με βάση πολλαπλά κριτήρια.

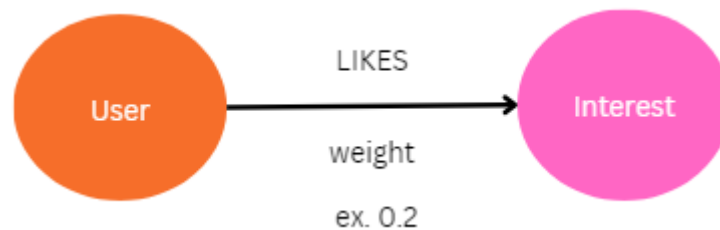
Το μοντέλο WSM, όπως περιγράφεται στο κεφάλαιο 2, αποτελεί μία μέθοδο ανάλυσης αποφάσεων. Στο πλαίσιο του WSM, υπολογίζεται το σταθμισμένο άθροισμα. Το σταθμισμένο άθροισμα αποτελεί έναν τρόπο να λαμβάνεται υπόψη η σημασία των διάφορων παραγόντων κατά την υπολογιστική διαδικασία, καθιστώντας το μοντέλο πιο εξατομικευμένο και προσαρμοσμένο στις προτιμήσεις του χρήστη. Αυτό σημαίνει ότι το μοντέλο λαμβάνει ως είσοδο τις προτιμήσεις των χρηστών και χρησιμοποιεί αυτές τις προτιμήσεις για να υπολογίσει τη σχέση μεταξύ των προτεινόμενων επιλογών και των ενδιαφερόντων του χρήστη.

Στον γνωσιακό γράφο που έχει δημιουργηθεί υπάρχουν πληροφορίες που μπορούν να χρησιμοποιηθούν για να υπολογιστεί η συσχέτιση χρήστη-προορισμού. Οι ακμές με ετικέτες 'LIKES' και 'HAS INTEREST', περιέχουν ως στοιχεία τιμές που είναι πολύ χρήσιμες για το WSM και μπορούν να χρησιμοποιηθούν ως παράγοντες στον υπολογισμό του σταθμισμένου αθροίσματος. Αυτές οι τιμές είναι το ποσοστό προτίμησης του χρήστη προς κάθε ενδιαφέρον με το οποίο συσχετίζεται ο χρήστης αλλά και το ποσοστό συνάφειας του κάθε προορισμού με το αντίστοιχο ενδιαφέρον.

Συγκεκριμένα χρησιμοποιούνται οι σχέσεις:



Σε αυτή την σχέση, το βάρος έχει αποδοθεί με στατικό τρόπο, ακολουθώντας την διαδικασία της κανονικοποίησης που αναλύθηκε στο κεφάλαιο της Μεθοδολογίας.



Η ιδιότητα της συγκεκριμένης ακμής αποδίδεται δυναμικά μέσω της διεπαφής. Εξαρτάται από τον χρήστη και δηλώνει την επιθυμία του χρήστη να του προταθούν προορισμοί που σχετίζονται με το συγκεκριμένο ενδιαφέρον.

Ο κώδικας που μοντελοποιεί το WSM, είναι ο εξής:

```

class RecommendationSystem:
    def __init__(self, user_preferences, weights, destination_labels):
        self.user_preferences = user_preferences
        self.weights = weights
        self.destination_labels = destination_labels

    def calculate_scores(self):
        scores = np.dot(self.weights, self.user_preferences)
        return scores

    def recommend_items(self, num_recommendations):
        scores = self.calculate_scores()
        # Rank destinations based on scores
        destination_scores = {label: score for label, score in
                               zip(self.destination_labels, scores)}
        sorted_destinations = sorted(destination_scores.items(),
                                     key=lambda x: x[1], reverse=True)
        # Display top recommendations with labels and scores
        top_recommendations_with_scores = [(destination, score) for
                                           destination, score in sorted_destinations]
        return top_recommendations_with_scores[:num_recommendations]

```

Τα κύρια δομικά στοιχεία της κλάσης WSM είναι οι προτιμήσεις του χρήστη (user preferences) και η αντιστοιχία βαρών-προορισμών (weights). Το τρίτο στοιχείο δηλαδή οι ετικέτες των προορισμών (destination labels) έχουν συμπληρωματικό ρόλο.

Η μέθοδος calculate_scores υπολογίζει τα σκορ με βάση τα βάρη και τις προτιμήσεις του χρήστη. Αυτό επιτυγχάνεται με τον πολλαπλασιασμό του διανύσματος των βαρών (self.weights) με το διάνυσμα των προτιμήσεων του χρήστη (self.user_preferences) χρησιμοποιώντας τον πίνακα γινομένου (dot product) της βιβλιοθήκης NumPy. Τα αποτελέσματα αυτού του πολλαπλασιασμού επιστρέφονται ως το σύνολο των σκορ. Συνοπτικά, η μέθοδος υλοποιεί έναν αλγόριθμο για τον υπολογισμό σκορ βάσει των βαρών και των προτιμήσεων του χρήστη.

Η μέθοδος recommend_items υλοποιεί έναν αλγόριθμο σύστασης για την επιλογή κορυφαίων προτάσεων προορισμών βάσει των υπολογισμένων σκορ από την προηγούμενη μέθοδο calculate_scores.

Συγκεκριμένα, η μέθοδος καλεί την calculate_scores για να πάρει τα σκορ για κάθε προορισμό. Στη συνέχεια, δημιουργεί ένα λεξικό (destination_scores) που αντιστοιχεί τα ονόματα των προορισμών με τα αντίστοιχα σκορ. Τα σκορ ταξινομούνται σε φθίνουσα σειρά, και οι κορυφαίες προτάσεις αποθηκεύονται σε μια λίστα (top_recommendations_with_scores), η οποία περιέχει ζεύγη προορισμού και σκορ.

Τέλος, η μέθοδος επιστρέφει τις κορυφαίες προτάσεις με βάση τον αριθμό που περνιέται ως παράμετρος (num_recommendations).

Τυπικά, το WSM αντιπροσωπεύεται από την μέθοδο calculate_scores. Για να λειτουργήσει το WSM χρειάζεται οι πίνακες που παίρνουν σαν παράμετροι να έχουν

συγκεκριμένη δομή. Ο λόγος για τους πίνακες με ονόματα `user_preferences`, `weights` και `destination labels` αντίστοιχα. Η δομή αυτή εξηγείται μέσω του αμέσως επόμενου παραδείγματος.

Έστω ότι υπάρχει ένας χρήστης με τις εξής προτιμήσεις:

Budget -> Flexible

Romance -> 0.6

History -> 0.7

Gastronomy -> 0.8

Ένα πιθανό σύνολο προορισμών που θα μπορούσαν να καλύψουν τα ενδιαφέροντα του χρήστη είναι:

Destination Labels

['Budapest', 'Vienna', 'Naples', 'Florence', 'Paris', 'Rome']

Σημειώνεται ότι ο πίνακας `destination labels` δεν περιέχει όλους τους δυνατούς προορισμούς αλλά ένα μικρό σύνολο εξ αυτών.

Προκειμένου να λειτουργήσει το σύστημα, τα βάρη προτίμησης του χρήστη προς το κάθε ενδιαφέρον του πρέπει να είναι μία λίστα η οποία στην πρώτη θέση θα πρέπει να έχει το βάρος που αναφέρεται στο πρώτο ενδιαφέρον, στην δεύτερη το δεύτερο κλπ.

Έτσι ο πίνακας `User Preferences` δομείται ως εξής:

User Preferences

[0.6, 0.7, 0.8]

Σχετικά με τον πίνακα `weights` έχει ακολουθηθεί η εξής λογική. Κάθε γραμμή αντιπροσωπεύει τα βάρη για έναν προορισμό. Ανα στήλη έχουμε συγκεκριμένα το βάρος της σύνδεσης του κάθε προορισμού προς το αντίστοιχο ενδιαφέρον. Αν δεν υπάρχει ακμή που να συνδέει προορισμό με ενδιαφέρον τότε η τιμή που συμπληρώνεται στον πίνακα στην αντίστοιχη θέση είναι 0.

Weights

	Gastronomy	History	Romance
Rome	1.0	0.0	0.0
Florence	0.8	0.8	0.4
Paris	0.7	0.0	0.6
Naples	0.4	0.0	0.0
Vienna	0.0	0.3	0.0
Budapest	0.0	0.0	0.1

Οι υπολογισμοί που λαμβάνουν χώρα είναι οι εξής:

1. Προορισμός 1 (Ρώμη)
 $Y_1 = 0.1 * 0.6 + 0.0 * 0.7 + 0.0 * 0.8 = 0.06$
2. Προορισμός 2 (Βιέννη)
 $Y_2 = 0.0 * 0.6 + 0.3 * 0.7 + 0.0 * 0.8 = 0.21$
3. Προορισμός 3 (Νάπολι)
 $Y_3 = 0.0 * 0.6 + 0.0 * 0.7 + 0.4 * 0.8 = 0.32$
4. Προορισμός 4 (Φλωρεντία)
 $Y_4 = 0.4 * 0.6 + 0.8 * 0.7 + 0.8 * 0.8 = 1.44$
5. Προορισμός 5 (Παρίσι)
 $Y_5 = 0.6 * 0.6 + 0.0 * 0.7 + 0.7 * 0.8 = 0.92$
6. Προορισμός 6 (Βουδαπέστη)
 $Y_6 = 0.0 * 0.6 + 0.0 * 0.7 + 1.0 * 0.8 = 0.8$

Χρησιμοποιώντας την φόρμουλα του μοντέλου WSM για κάθε προορισμό ξεχωριστά προκύπτει το σκορ για κάθε προορισμό. Αν ταξινομηθούν τα αποτελέσματα σε φθίνουσα σειρά προκύπτει ότι οι προτεινόμενοι προορισμοί είναι κατά σειρά:

- Φλωρεντία
- Παρίσι
- Ρώμη
- Νάπολι
- Βιέννη
- και,
- Βουδαπέστη

Σύστημα προτεινόμενων

Ένα υπογράφημα του γνωσιακού γράφου είναι συσχετισμένο με τον χρήστη. Σε αυτό το πλαίσιο όταν πρόκειται να παραχθούν προτάσεις για έναν συγκεκριμένο

χρήστη, τότε συγκεντρώνονται όλα τα στοιχεία που αφορούν τον χρήστη και που συμμετέχουν στην διαδικασία παραγωγής των προτάσεων. Αυτά τα στοιχεία είναι η προτίμηση του χρήστη στο budget και τα ποσοστά προτίμησης του χρήστη σε αντιστοιχία με τα ενδιαφέροντα του.

Το σύστημα αποκτά πρόσβαση σε αυτά τα δεδομένα μέσα από τους drivers του Neo4j, εκτελώντας queries που αναζητούν όλες τις σχέσεις HAS_INTEREST μεταξύ του χρήστη στόχου και των πιθανών προορισμών.

Το query που χρησιμοποιείται είναι το εξής:

```
MATCH (i:Interest {name: $interest})<-[r:HAS_INTEREST]-(d:Destination)
RETURN i.name AS interest, d.name AS destination, r.weight AS weight
```

Αν δούμε λίγο αναλυτικότερα την δομή του παραπάνω query ξεχωρίζουμε δύο βασικές εντολές που χρησιμοποιούνται κατά την διαδικασία σχηματισμού ενός query που απευθύνεται στην βάση neo4j. Αυτές είναι:

```
MATCH (u:User {email: $user_email})-[l:LIKES]->(i:Interest):
```

Η εντολή αυτή εντοπίζει έναν κόμβο χρήστη (`u`) με βάση το email που περνάμε ως παράμετρο (`\$user_email`). Το email αποτελεί το στοιχείο της βάσης που ξεχωρίζει έναν χρήστη από έναν άλλον. Στη συνέχεια, εντοπίζει τις σχέσεις `LIKES` μεταξύ αυτού του χρήστη και των κόμβων ενδιαφέροντος (`i`).

```
RETURN u.email as user_email, i.name as interest_name, l.weight as weight:
```

Αυτή η εντολή καθορίζει ποιες πληροφορίες πρέπει να επιστραφούν ως αποτέλεσμα. Επιστρέφει το email του χρήστη ως `user_email`, το όνομα του ενδιαφέροντος ως `interest_name` και το βάρος της σχέσης `LIKES` ως `weight`.

Συνολικά, το ερώτημα αυτό χρησιμοποιείται για να ανακτήσει τις προτιμήσεις ενός χρήστη σχετικά με τα ενδιαφέροντά του, περιλαμβάνοντας το email του χρήστη, το όνομα του ενδιαφέροντος και το βάρος της σχέσης `LIKES`.

Ένας ακόμα παράγοντας που λαμβάνει υπόψη του το σύστημα προτάσεων με ενσωμάτωση του μοντέλου WSM είναι το budget.

Ανακτάται με την συγκεκριμένη εντολή:

```
MATCH (u:User {email: $email})
```

```
RETURN u.budget
```

Αφού ολοκληρωθεί η συλλογή των δεδομένων από την πλευρά του χρήστη, το σύστημα χρειάζεται τα δεδομένα που αφορούν τους κομβους-προορισμούς. Αυτό επιτυγχάνεται μέσω του παρακάτω query:

```
MATCH (i:Interest {name: $interest})<-[r:HAS_INTEREST]-(d:Destination)
```

```
RETURN i.name AS interest, d.name AS destination, r.weight AS weight
```

Συνοπτικά, το ερώτημα αυτό αναζητά προορισμούς που σχετίζονται με ένα συγκεκριμένο ενδιαφέρον, επιστρέφοντας το όνομα του ενδιαφέροντος, το όνομα του προορισμού και το βάρος της σχέσης μεταξύ τους. Με βάση αυτές τις σχέσεις προκύπτει η λίστα των πιθανών προορισμών προς πρόταση.

Αντίστοιχα με την δομή του προηγούμενου ερωτήματος, αυτό το ερώτημα εκτελεί τις παρακάτω ενέργειες:

```
MATCH
```

Αναζητεί έναν κόμβο ενδιαφέροντος (`Interest`) με το όνομα που ορίζεται από τη μεταβλητή `\$interest`.

```
<-[:HAS_INTEREST]-
```

Βρίσκει σχέσεις (`HAS_INTEREST`) που πηγαίνουν από τον κόμβο ενδιαφέροντος προς έναν κόμβο προορισμού (`Destination`).

```
RETURN
```

Επιστρέφει τα δεδομένα του ενδιαφέροντος, του προορισμού και του βάρους της σχέσης. Ορίζονται συνώνυμα (aliases) για τα ονόματα των οντοτήτων για να είναι πιο ευανάγνωστο το αποτέλεσμα.

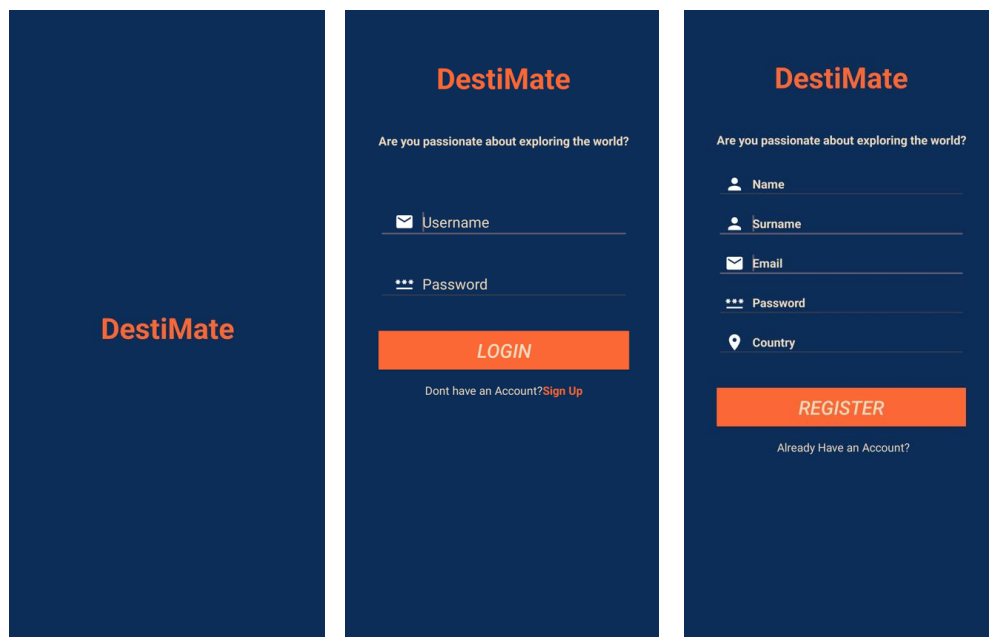
Αν ο χρήστης επιλέξει να μην υπάρχει φίλτρο στα οικονομικά τότε το σύστημα δεν λαμβάνει υπόψη την παράμετρο budget και επικεντρώνεται στα ενδιαφέροντα.

Παρουσίαση διεπαφής

Στα πλαίσια της δυνατότητας που δίνεται στον χρήστη να αλληλεπιδράσει με το σύστημα συστάσεων υλοποιήθηκε μια διεπαφή κινητής συσκευής Android, η οποία υποστηρίζει διάφορες λειτουργίες.

Αυθεντικοποίηση Χρήστη

Ξεκινώντας, μία από τις βασικότερες λειτουργίες είναι αυτή της αυθεντικοποίησης. Ο χρήστης έχει την δυνατότητα να εισέρχεται στην εφαρμογή με εισαγωγή των στοιχείων του λογαριασμού στον οποίο έχει εγγραφεί. Σε περίπτωση που ο χρήστης δεν έχει ήδη δημιουργήσει έναν λογαριασμό, μπορεί να το κάνει μεταφερόμενος στην οθόνη Sign Up, καταχωρώντας τα στοιχεία του.

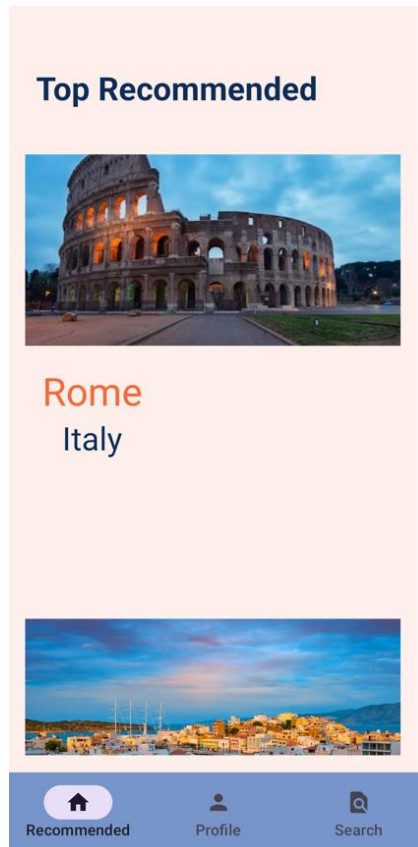


Τα στοιχεία που πληκτρολόγησε ο χρήστης στέλνονται μέσω API στον server που έχει δημιουργηθεί το σύστημα συστάσεων και όπου πραγματοποιούνται όλες οι επικοινωνίες τόσο μεταξύ android UI και συστήματος συστάσεων όσο και συστήματος συστάσεων με την βάση Neo4j όπου βρίσκεται ο γνωσιακός γράφος που περιέχει όλα τα δεδομένα στα οποία στηρίζεται το όλο σύστημα.

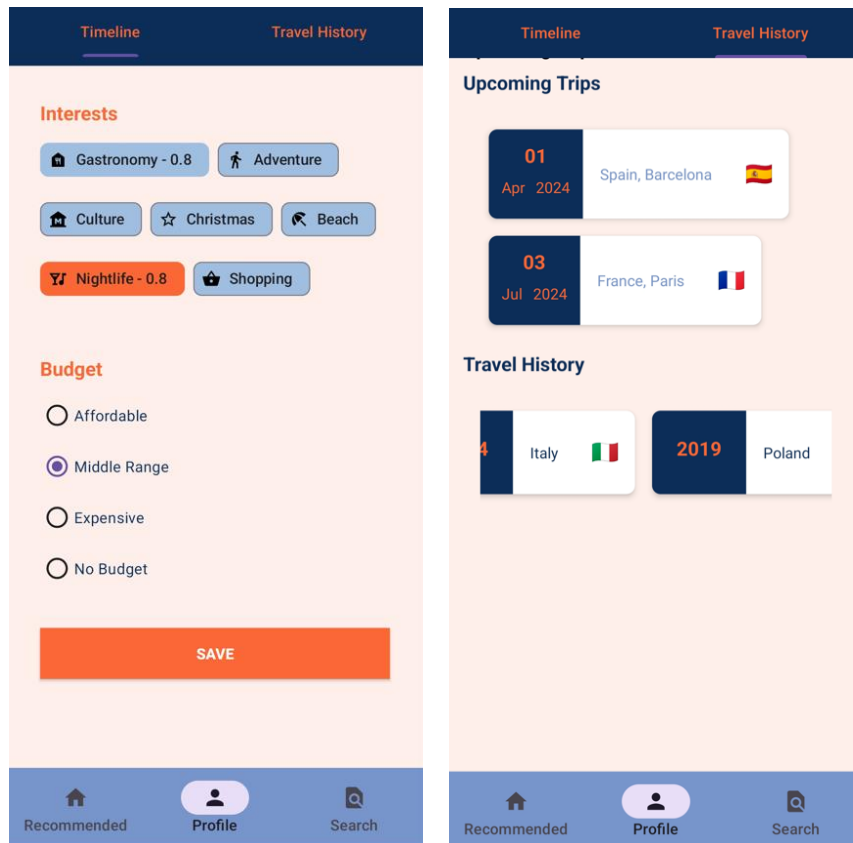
Το κομμάτι της αυθεντικοποίησης χειρίζεται η οντότητα User.

Προβολή προτεινόμενων

Με την είσοδο του χρήστη στην εφαρμογή, η πρώτη οθόνη που του εμφανίζεται αφορά τους προτεινόμενους προορισμούς. Ο χρήστης βλέπει προτεινόμενους προορισμούς με βάση τις προτιμήσεις του και τα ενδιαφέροντά του.



Η οθόνη αυτή στο κάτω μέρος της έχει 3 views. Πατώντας το view που αφορά το προφίλ του χρήστη, η εφαρμογή προβάλλει τα ενδιαφέροντα του χρήστη και την προτίμηση του στο budget όπως αυτά αποτυπώνονται στον γνωσιακό γράφο.

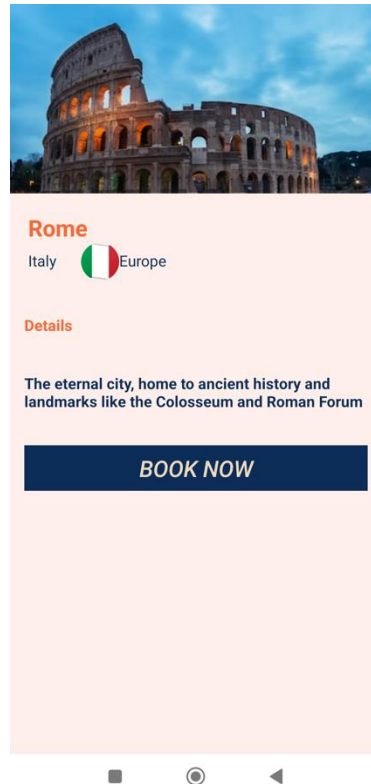


Στο δεύτερο tab στο view του προφίλ του χρήστη, περιγράφονται κάποιοι από τους επερχόμενους προορισμούς ενώ επίσης κρατείται ιστορικό επισκέψεων ανά χώρα.

Αναζήτηση προορισμών

Στο τρίτο View ο χρήστης μπορεί να αναζητήσει ονόματα προορισμών. Επιλέγοντας κάποιο από τα ονόματα των προορισμών για τους οποίους ενδιαφέρεται μεταφέρεται στην οθόνη που προβάλλονται οι πληροφορίες για τον προορισμό που ζήτησε μεμονωμένα.

Η Οθόνη που εμφανίζεται στον χρήστη, όταν αναζητά έναν προορισμό είναι η ακόλουθη:



Προγραμματισμός επόμενου ταξιδιού

Ο χρήστης μπορεί να προγραμματίσει το επόμενο ταξίδι του, επιλέγοντας προορισμό και ημερομηνία αναχώρησης. Ο προορισμός αυτός εισάγεται στην λίστα των upcoming trips ενώ ταυτόχρονα αποθηκεύεται στον γνωσιακό γράφο με την σχέση User -BOOKED-> Rome.

Αυτές οι λειτουργίες παρέχουν στους χρήστες μια εξατομικευμένη εμπειρία ταξιδιού μέσω της εφαρμογής Android.

ΚΕΦΑΛΑΙΟ 5 - ΑΞΙΟΛΟΓΗΣΗ

Ο γνωσιακός γράφος όπως αναλύθηκε στα κεφάλαια 2 και 3, έχει κατασκευαστεί με δεδομένα που κυκλοφορούν στο διαδίκτυο αναφορικά με τους καλύτερους προορισμούς ανά ενδιαφέρον χρήστη. Στον χρήστη δίνεται η δυνατότητα να λαμβάνει εξατομικευμένες προτάσεις που ταιριάζουν με τα προσωπικά του ενδιαφέροντα, επιλέγοντας από μία λίστα από διαφορετικά χαρακτηριστικά του προορισμού που τον ενδιαφέρουν όπως είναι η ιστορία, η γαστρονομία, η φυσική ομορφιά και άλλα. Ακόμα, δίνεται η δυνατότητα προσαρμογής του budget, μέσω διαβάθμισης των προορισμών σε high range, middle range και low range.

Το σύστημα από την στιγμή δημιουργίας ενός λογαριασμού χρήστη δεν παρέχει προσωποποιημένες συστάσεις καθώς απαιτείται το ελάχιστο η επιλογή ενός ενδιαφέροντος χρήστη από την διαθέσιμη λίστα ενδιαφερόντων ή ο ορισμός του budget. Αυτές οι δύο δράσεις του χρήστη επιτρέπουν στο σύστημα να γίνει πιο προσωποποιημένο και άρα και πιο αποδοτικό. Η απόδοση του συστήματος λοιπόν εξαρτάται από την μία πλευρά από το πόσο διατεθειμένος είναι ο χρήστης να ορίσει τα ενδιαφέροντα του και να είναι εύστοχος στον καθορισμό των βαρών που πρόκειται να λάβει υπόψη το σύστημα. Από την άλλη πλευρά, είναι σημαντικό να εξεταστούν ζητήματα όπως το πώς διανέμεται ο κάθε προορισμός ανά ενδιαφέρον και το πόσο τελικά επιδραστική είναι η προσθήκη του μοντέλου WSM.

Αξιολόγηση Δεδομένων

Στον γνωσιακό γράφο συναντάμε πληροφορίες από 65 προορισμούς που είναι διαθέσιμοι προς πρόταση. Όλοι οι προορισμοί τοποθετούνται γεωγραφικά στην Ευρώπη.

Από τα διαθέσιμα δεδομένα προκύπτει η ομαδοποίηση των προορισμών με βάση κοινά τους χαρακτηριστικά. Αρχικά οι προορισμοί εισήχθησαν στον γνωσιακό γράφο καταταγμένοι ανάλογα με το πόσο σχετικοί είναι με το συγκεκριμένο ενδιαφέρον. Η ύπαρξη ακμής ανάμεσα σε έναν κόμβο ενδιαφέροντος και σε έναν κόμβο προορισμού στον γνωσιακό γράφο υποδεικνύει ότι υπάρχει σύνδεση μεταξύ αυτών των δύο στιγμιοτύπων ενώ η ιδιότητα που έχει αποδωθεί στην ακμή τους, φέρει το βάρος σχετικότητας. Αν δεν υπάρχει αυτή η ακμή σημαίνει ότι η σχετικότητα είναι μηδενική και ο προορισμός δεν συμπεριλαμβάνεται στις προτάσεις. Αυτό είναι σημαντικό για τη διασφάλιση του ότι οι προτάσεις είναι στοχευμένες και σχετικές με τα ενδιαφέροντα του χρήστη. Σε περιπτώσεις όπου όλοι οι προορισμοί έχουν κοινά χαρακτηριστικά όπως ιστορία ή φυσική ομορφιά, η αποσαφήνιση της σχετικότητας μπορεί να βοηθήσει στην παροχή πιο εξειδικευμένων προτάσεων.

Υπάρχουν 10 διαθέσιμες κατηγορίες ενδιαφερόντων. Για κάθε κατηγορία υπάρχουν 10 προορισμοί οι οποίοι συνδέονται με τον αντίστοιχο κόμβο. Για παράδειγμα, 10 πιθανοί προορισμοί για έναν χρήστη που τον ενδιαφέρει μόνο η γαστρονομία είναι:

Κατάταξη	Χώρα	Προορισμός	Σκορ
01	Italy	Rome	0.80
02	Greece	Crete	0.72
03	Spain	Barcelona	0.68
04	Italy	Florence	0.64
05	France	Paris	0.64
06	Portugal	Lisbon	0.40
07	Italy	Naples	0.32
08	Spain	San Sebastian	0.24
09	Spain	Madrid	0.18
10	Spain	Ibiza	0.16

Το σύστημα συστάσεων, έχει στην διάθεση του συνολικά 25 χώρες από τις οποίες προκύπτουν οι προτεινόμενοι προορισμοί.

Σενάριο 1

Ένας από τους χρήστες του συστήματος, ο Νίκος, έχει εκφράσει το ενδιαφέρον του μέσω εφαρμογής προς την γαστρονομία και την νυχτερινή ζωή ενός προορισμού. Τα βάρη προτίμησης που έχει θέσει ο ίδιος είναι 0.2 και 0.8 για την παρουσία νυχτερινής ζωής και το φαγητό αντίστοιχα.

Nightlife	0.2
Gastronomy	0.8

Ο χρήστης Νίκος δεν έχει εκφράσει κάποια προτίμηση στον προϋπολογισμό του. Έτσι το σύστημα λαμβάνει υπόψη όλους τους διαθέσιμους προορισμούς. Έτσι με την τιμή στο πεδίο Budget να είναι 'Flexible' τα προτεινόμενα που εμφανίζονται σε αυτή την περίπτωση είναι τα εξής:

Κατάταξη	Χώρα	Προορισμός	Σκορ
01	Italy	Rome	0.80
02	Greece	Crete	0.72
03	Spain	Barcelona	0.68
04	Italy	Florence	0.64
05	France	Paris	0.64
06	Portugal	Lisbon	0.40
07	Italy	Naples	0.32
08	Spain	San Sebastian	0.24
09	Spain	Madrid	0.18
10	Spain	Ibiza	0.16

Η Ρώμη καταλαμβάνει την πρώτη θέση, ακολουθούμενη από την Κρήτη και τη Βαρκελώνη. Η Ιταλία παρουσιάζει τρεις εκπροσώπους στη λίστα, ενώ η Ελλάδα έχει μόνο την Κρήτη. Η Ισπανία έχει τη Βαρκελώνη ως εκπρόσωπο στη δεύτερη θέση, ενώ έχει και άλλους τρεις προορισμούς με χαμηλότερα σκορ στο τέλος της λίστας. Η Γαλλία και η Πορτογαλία συμπληρώνουν τις επιλογές στην κατάταξη.

Ο χρήστης λαμβάνει προτάσεις από 5 διαφορετικές χώρες. Έτσι μπορούμε να καταλήξουμε στο ότι το σύστημα συστάσεων παρέχει ποικιλία επιλογών σε επίπεδο χωρών με τις πρώτες επιλογές να αποτελούνται από εκπροσώπους διαφορετικών χωρών ενώ οι πρώτες 5 επιλογές να καλύπτουν τις 4 από τις 5 χώρες που προτείνονται συνολικά.

Ακολουθούν αναλυτικά οι τιμές της παρουσίας κάθε ενδιαφέροντος σε κάθε προορισμό.

Προορισμός	Gastronomy	Nightlife
Rome	1.0	
Crete	0.9	
Barcelona	0.6	1.0
Florence	0.8	
Paris	0.7	0.4
Lisbon	0.5	
Naples	0.4	
San Sebastian	0.3	
Madrid		0.9
Ibiza		0.8

Παρατηρώντας τα βάρη της γαστρονομίας προς τους αντίστοιχους προορισμούς βλέπουμε ότι ακολουθούν φθίνουσα πορεία με εξαίρεση τις θέσεις 3 έως 5 όπου υπάρχει μία ανακατάταξη. Η ανακατάταξη οφείλεται στην παρουσία βάρους έστω και

αν αυτό έχει ελάχιστη τιμή (0.2) στην προτίμηση του χρήστη προς την νυχτερινή ζωή.

Η αναλογία που έχει δώσει ο χρήστης ανάμεσα στα δύο ενδιαφέροντα είναι 1 προς 4, όπου 0.8 το βάρος της γαστρονομίας και 0.2 το βάρος της νυχτερινής ζωής. Παρότι, ο χρήστης έχει δώσει πολύ μεγαλύτερη τιμή βάρους στην γαστρονομία, βλέπουμε ότι το μοντέλο WSM δίνει βάση και στις 2 προτιμήσεις του χρήστη, καταφέροντας να παρουσιάσει στις πρώτες 5 θέσεις της λίστας προορισμούς οι οποίοι να μην καλύπτουν το πρώτο ενδιαφέρον του αλλά δεν αγνοεί την ύπαρξη του δεύτερου.

Το σύστημα, παρείχε εξατομικευμένες συστάσεις για τον χρήστη Νίκο, λαμβάνοντας υπόψη τα ενδιαφέροντά του για το φαγητό ενός προορισμού (γαστρονομία) αλλά και για την νυχτερινή ζωή.

Σενάριο 2

Το σύστημα πραγματοποίησε συστάσεις για τον χρήστη με το όνομα Maggie Smith, λαμβάνοντας υπόψη τα ενδιαφέροντά του σε γαστρονομία, κουλτούρα και ιστορία. Ο ίδιος ο χρήστης έχει ορίσει τα ενδιαφέροντα του αναθέτοντας βάρη που υποδεικνύουν μεγαλύτερη προτίμηση προς τη γαστρονομία και την κουλτούρα, με μικρότερο ενδιαφέρον προς την ιστορία των ταξιδιωτικών προορισμών. Συγκεκριμένα, για την παροχή εξατομικευμένων συστάσεων τα δεδομένα που έχει στην διάθεση του το σύστημα είναι:

Budget	High Range
Ενδιαφέρον	Βάρος
History	0.6
Culture	0.8
Gastronomy	0.8

Σύμφωνα με τα προφίλ χρηστών, η Maggie Smith κατατάσσεται στην ευρύτερη κατηγορία χρήστη που του αρέσει η πολιτισμική εξερεύνηση, που επιδιώκει να κατανοήσει βαθύτερα τον πολιτισμό και την ιστορία των τόπων που επισκέπτεται.

Οι καταλληλότεροι προορισμοί σύμφωνα με το μοντέλο WSM είναι:

Κατάταξη	Χώρα	Προορισμός	Score
01	France	Paris	1.04
02	England	London	0.56
03	Spain	Barcelona	0.48
04	Turkey	Istanbul	0.42
05	Spain	Bilbao	0.32
06	Spain	San Sebastian	0.24
07	France	Lyon	0.16
08	Netherlands	The Hague	0.16
09	Switzerland	Basel	0.08
10	England	York	0.06

Φαινομενικά οι προορισμοί με υψηλότερη βαθμολογία στα χαρακτηριστικά που τον ενδιαφέρουν περιλαμβάνουν το Παρίσι, το Λονδίνο και τη Βαρκελώνη. Όμως παρατηρείται μεγάλη διαφορά στο σκορ μεταξύ του πρώτου και του δεύτερου προτεινόμενου προορισμού. Αυτό οφείλεται στα δεδομένα που χρησιμοποιούνται από το μοντέλο WSM αφού η επιλογή του budget μειώνει τον αριθμό των διαθέσιμων προορισμών.

ΚΕΦΑΛΑΙΟ 6 - ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα συστήματα προτάσεων αποτελούν ένα πολύ χρήσιμο εργαλείο στα χέρια του χρήστη. Η ανάγκη για εξατομίκευση αυξάνεται όσο αυξάνονται και τα δεδομένα τα οποία υπάρχουν διαθέσιμα προς όλους στο διαδίκτυο. Τόσο ο όγκος των δεδομένων όσο και η παρουσία τους σε πολλές και διαφορετικές μορφές, δημιουργούν στον σημερινό χρήστη σύγχυση. Ο χρήστης σήμερα, δεν δυσκολεύεται τόσο στο να βρει μια πληροφορία καθώς υπάρχει διαθεσιμότητα πληροφορίας, δυσκολεύεται όμως, να βρει την πληροφορία η οποία να ανταποκρίνεται στα ενδιαφέροντα του, εύκολα και γρήγορα. Τον σκοπό αυτό επιτελούν τα συστήματα προτάσεων και συγκεκριμένα τα συστήματα που βασίζονται σε δεδομένα που αφορούν τον ίδιο τον χρήστη.

Η δημιουργία ενός προφίλ για τον χρήστη αποτελεί ένα βήμα προς την υλοποίηση ενός προσωποποιημένου συστήματος προτάσεων. Η κεντρική ιδέα αυτής της διπλωματικής εργασίας ήταν η ανάπτυξη ενός τέτοιου συστήματος, το οποίο με βάση τα ενδιαφέροντα και τις προτιμήσεις του χρήστη θα παρείχε προτάσεις για τον επόμενο πιθανό προορισμό τον οποίο θα μπορούσε να επισκεφθεί ο χρήστης. Η επιλογή του τουρισμού ως θέματος αυτής της εργασίας παρείχε την δυνατότητα ενσωμάτωσης πολλών και διαφορετικών προφίλ χρηστών πάνω στα οποία δομήθηκαν τα δεδομένα τα οποία χρησιμοποιήθηκαν στον γνωσιακό γράφο. Η εφαρμογή του τουρισμού, λοιπόν, ως κεντρικού θέματος έκανε δυνατή την μοντελοποίηση των ενδιαφερόντων που είναι πιθανό να έχει στην φαρέτρα του ένας ταξιδιώτης. Ξεκινώντας από την μοντελοποίηση των ενδιαφερόντων και καταλήγωντας στην αποτίμηση του τι μπορεί να αρέσει πραγματικά σε έναν χρήστη και στην παρουσίαση αυτής της πρότασης, το σύστημα προσέγγισε την διαδικασία παροχής συστάσεων μπαίνοντας στην θέση του χρήστη, υιοθετώντας τα ενδιαφέροντα του, τις προτιμήσεις του και σε ένα γενικότερο πλαίσιο την οπτική του.

Αυτή η προσέγγιση οδήγησε στην παροχή εξατομικευμένων συστάσεων, αφού το σύστημα προτάσεων ήταν πλέον σε θέση να συνυπολογίσει προεκτάσεις του χρήστη όπως είναι τα ενδιαφέροντα του τα οποία είναι πολύ δύσκολο να οριστούν χωρίς την ενεργή συμμετοχή του χρήστη. Το ποσοστό συμμετοχής του χρήστη στην διαδικασία ορισμού των ενδιαφερόντων του, έμμεσα ή άμεσα με ενεργή ή παθητική δηλαδή συμμετοχή του χρήστη, μπορεί να αποτελέσει μια πρόκληση για το μέλλον, οδηγώντας πιθανόν σε βελτίωση της απόδοσης του συστήματος.

Ο γνωσιακός γράφος αποτελεί μία δομή που είναι εύκολα επεκτάσιμη. Κατά την διάρκεια ανάπτυξης του συστήματος ήταν πολύ εύκολο να προστεθούν ή και να αφαιρεθούν δεδομένα χωρίς να αλλοιωθεί η συνολική εικόνα των δεδομένων που αντιπροσώπευε ο γνωσιακός γράφος. Η χρήση ετικετών στις ακμές μεταξύ δύο οντοτήτων παρείχε εξαιρετική δύναμη στο να ενσωματωθούν νέες πληροφορίες, ενώ παράλληλα άνοιγε την πόρτα στην ενσωμάτωση ακόμα περισσότερων πληροφοριών. Αυτή η ευελιξία του γνωσιακού γράφου θα μπορούσε να φανεί χρήσιμη και στην επεξηγησιμότητα των ταξιδιωτικών προτάσεων. Το κλειδί άλλωστε ενός συστήματος συστάσεων, προκειμένου οι προτάσεις που παρέχει να

είναι επιτυχείς, είναι να καταλάβει όχι τόσο το τι αρέσει στον χρήστη αλλά γιατί του αρέσει.

Η χρήση του Neo4j, ως εργαλείου διαχείρισης του γνωσιακού γράφου παρείχε πολύτιμη βοήθεια. Το neo4j έχει το πλεονέκτημα ότι διευκολύνει την αναζήτηση και την ανάκτηση πληροφοριών που βασίζονται σε σχέσεις, καθιστώντας το ιδανικό για εφαρμογές που απαιτούν πολυπλοκότερες σχέσεις μεταξύ των δεδομένων. Οι ετικέτες των ακμών χρησιμοποιούνται ως λέξεις κλειδιά για την αναζήτηση των δεδομένων, μειώνοντας τον χρόνο αναζήτησης.

Η εφαρμογή του Multiple Criteria Decision Analysis (MCDA) στα Recommender Systems είναι μια ενδιαφέρουσα προσέγγιση που μπορεί να βοηθήσει στην προσαρμογή των συστημάτων συστάσεων στις προτιμήσεις των χρηστών με βάση πολλαπλά κριτήρια. Το μοντέλο WSM, το οποίο χρησιμοποιήθηκε σε συνδυασμό με τον γνωσιακό γράφο πάνω στον οποίο στηρίχθηκε το σύστημα συστάσεων, λειτούργησε ομαλά και ήταν εύκολο να ενσωματωθεί αφού ήταν επίσης εύκολη η σύνδεση των οντοτήτων και η ενσωμάτωση ιδιοτήτων στις ακμές με τέτοιο τρόπο ώστε να αποτυπωθούν στον γνωσιακό γράφο τόσο οι προτιμήσεις του χρήστη όσο και τα κριτήρια επιλογής του κάθε ταξιδιωτικού προορισμού.

Παρόλα αυτά, μια δυσκολία η οποία αντιμετωπίστηκε ήταν ο σαφής καθορισμός του βαθμού συσχητικότητας μεταξύ ενός προορισμού και ενός ενδιαφέροντος. Είναι γεγονός ότι η επιτυχής λειτουργία του συστήματος εξαρτάται από την ποιότητα των δεδομένων. Το μοντέλο WSM λειτουργεί όπως αναλύθηκε, λαμβάνοντας υπόψη κριτήρια, τα οποία έχουν αποδωθεί χρησιμοποιώντας μόνο αριθμητικές τιμές. Το να ληφθεί υπόψη μόνο μία διάσταση ήταν περιοριστικό καθώς η ενσωμάτωση πολλών προορισμών κάνει δύσκολη την απόδοση βαρών-κριτηρίων τα οποία είναι αντιπροσωπευτικά για κάθε ταξιδιωτικό προορισμό. Για αυτό τον λόγο τουλάχιστον σε αυτή την διπλωματική εργασία επιλέχθηκε η Ευρώπη ως η ήπειρος στην οποία θα ανήκουν οι προορισμοί που προτείνονται στους χρήστες της εφαρμογής. Σαν πιθανή μελλοντική επέκταση της παρούσας διπλωματικής αφήνεται το να βρεθεί μία προσέγγιση η οποία θα αποδίδει πιο εύστοχα βάρη στην σύνδεση μεταξύ ενός ενδιαφέροντος και ενός προορισμού, παρέχοντας στον χρήστη πιο ακριβείς προτάσεις σε συνδυασμό με την δυνατότητα εύρεσης ταξιδιωτικών προορισμών από όλο τον πλανήτη.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1]

Mashaël Aldayel et al. «Collaborative Filtering-Based Recommendation Systems for Touristic Businesses, Attractions, and Destinations». In: Electronics 12 (Sept. 2023), p. 4047. doi: 10.3390/electronics12194047.

[2]

Martin Aruldoss, T. Lakshmi, and V. Venkatesan. «A survey on multi-criteria decision-making methods and its applications». In: American Journal of Information Systems 1 (Jan. 2013), pp. 31–43. doi: 10.12691/ajis-1-1-5.

[3]

Robin Burke and Robin. «Hybrid Web Recommender Systems». In: LNCS (2007). doi: 10.1007/978-3-540-72079-9_12.

[4]

Ines Chami et al. «Machine Learning on Graphs: A Model and Comprehensive Taxonomy». In: arXiv (Cornell University) (May 2020). url: <http://export.arxiv.org/pdf/2005.03675>.

[5]

Ioannis Konstas, Vassilios Stathopoulos, and Joemon M. Jose. «On social networks and collaborative recommendation». In: SIGIR '09. Boston, MA, USA: Association for Computing Machinery, 2009, pp. 195–202. isbn: 9781605584836. doi: 10.1145/1571941.1571977. url: <https://doi.org/10.1145/1571941.1571977>.

[6]

Enayat Rajabi and Kobra Etminani. «Knowledge-graph-based explainable AI: A systematic review». In: Journal of Information Science (Sept. 2022), p. 016555152211128. doi: 10.1177/01655515221112844.

[7]

Kenneth H. Rosen. ΔΙΑΚΡΙΤΑ ΜΑΘΗΜΑΤΙΚΑ και Εφαρμογές τους (8η έκδοση). Εκδόσεις Τζιόλα, 2018.

[8]

Paulus Santosa, Ari Cahyono, and Karina Auliasari. «A multi-criteria recommender system for tourism destination». In: Proceedings - Pacific Asia Conference on Information Systems, PACIS 2012 (Jan. 2012).

[9]

Victor Telnov and Yu Korovin. «Programming of Knowledge Graphs, Reasoning on Graphs». In: PROGRAMMNAYA INGENERIA 10 (Feb. 2019), pp. 59–68. doi: 10.17587/prin.10.59-68.

[10]

Evangelos Triantaphyllou. «Fuzzy Multi-Criteria Decision Making». In: Jan. 2000, pp. 241–262. isbn: 978-1-4419-4838-0. doi: 10.1007/978-1-4757-3157-6_13.

[11]

Evangelos Triantaphyllou and Alfonso Vargas Sánchez. «A sensitivity analysis approach for some deterministic Multi-Criteria Decision-Making methods*». In: Decision Sciences 28.1 (Jan. 1997), pp. 151–194. doi: 10.1111/j.1540-5915.1997.tb01306.x. url: <https://doi.org/10.1111/j.1540-5915.1997.tb01306.x>.

[12]

Xiaohan Zou. «A Survey on Application of Knowledge Graph». In: (2020). doi: 10.1088/1742-6596/1487/1/012016.