



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

**Ανάπτυξη συστήματος συστάσεων τουριστικών
προορισμών**

ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ: 2019 – 2020

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ ΦΟΙΤΗΤΗ

ΑΒΕΛΛΑ ΑΣΤΕΡΙΟΥ

ΑΜ:2012010023

ΕΠΙΒΛΕΠΩΝ :

ΜΑΤΣΑΤΣΙΝΗΣ ΝΙΚΟΛΑΟΣ

Καθηγητής Πολυτεχνείου Κρήτης

Τριμελής εξεταστική επιτροπή:

Καθηγητής ΝΙΚΟΛΑΟΣ ΜΑΤΣΑΤΣΙΝΗΣ (Επιβλέπων)

Επίκουρος Καθηγητής ΣΤΕΛΙΟΣ ΤΣΑΦΑΡΑΚΗΣ

Δρ. ΕΥΑΓΓΕΛΙΑ ΚΡΑΣΑΔΑΚΗ

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία αναπτύσσεται η μεθοδολογία ενός πολυκριτήριου συστήματος σύστασης για ξενοδοχειακές μονάδες. Το σύστημα λαμβάνει υπόψη τις προτιμήσεις των χρηστών (πελατών) καθώς και τα χαρακτηριστικά των ξενοδοχείων και στη συνέχεια αξιολογεί και προτείνει στο χρήστη τα κατάλληλα ξενοδοχεία που του ταιριάζουν.

Για την άντληση πληροφοριών σχετικά με τις προτιμήσεις των χρηστών πραγματοποιήθηκε συλλογή ερωτηματολογίων. Παράλληλα συλλέχτηκαν δεδομένα για τα χαρακτηριστικά και τις υπηρεσίες των ξενοδοχειακών μονάδων από την ιστοσελίδα Booking.com. Στη συνέχεια έγινε καθαρισμός των δεδομένων για την αποφυγή ελλειπών τιμών.

Αξιοποιήσαμε τα δεδομένα αυτά χρησιμοποιώντας τη μεθοδολογία του agent allocator. Με τη βοήθεια ενός κώδικα, ο οποίος αναπτύχθηκε σε γλώσσα προγραμματισμού Python, δημιουργήσαμε προφίλ χρηστών και υλοποιήσαμε τον πολυκριτήριο πίνακα.

Στη συνέχεια έγινε χρήση του αλγορίθμου της UTASTAR σε ένα αντιπροσωπευτικό σύνολο των εναλλακτικών, για τον υπολογισμό των βαρών. Τέλος χρησιμοποιώντας τα βάρη των κριτηρίων και τις μερικές χρησιμότητες, υπολογίζουμε τις ολικές χρησιμότητες για όλους τους δυνατούς συνδυασμούς Ξενοδοχείων - Πελατών και οδηγούμαστε στα κατάλληλα ξενοδοχεία για κάθε χρήστη.

ABSTRACT

In this thesis we have the development of a multi-criteria recommendation system on hotels. The particular system takes into consideration the users' (clients) preferences as well as the various characteristics of the hotels and, in addition evaluates and recommends it's users', the appropriate hotels.

There has been a collection of questionnaires, in order to gather information regarding the users' preferences. At the same time, there was a acquisition of data that had to do with the characteristics and the services of the hotels, taken directly from [Booking.com](https://www.booking.com). Afterwards, data cleaning has been done in order to avoid incomplete or false values.

The specific data were utilized using the Agent Allocator method. We created profiles for each user and generated a multi-criteria data matrix with the assistance of a python developed a code.

Furthermore, we used UTASTAR algorithm on a representative aggregate of hotels, in order to calculate the weight of each criteria. In the end, we achieved the best possible recommendation for each user, using the weight of criteria and the marginal utilities, to calculate the global utilities for every possible Hotel-User combination.

Ευχαριστίες

Για την εκπόνηση αυτής της εργασίας θα ήθελα να ευχαριστήσω θερμά τον Καθηγητή μου Νικόλαο Ματσατσίνη για την συνεχή του καθοδήγηση και το ενδιαφέρον του, όχι μόνο για την εκπόνηση αυτής της εργασίας, αλλά και για την εις βάθος κατανόηση της ουσίας του προβλήματος.

Ευχαριστίες επίσης, θα ήθελα να εκφράσω, στο κύριο Σακελλάρη Αλκαίο καθώς και στην κυρία Κρασαδάκη Ευαγγελία που με ενδιαφέρον απάντησαν σε όλες τις ερωτήσεις μου, και με βοήθησαν στις δυσκολίες της διπλωματικής μου εργασίας.

Τέλος, επειδή με την εργασία αυτή, ολοκληρώνονται και οι σπουδές μου ως προπτυχιακού φοιτητή θα ήθελα να ευχαριστήσω την οικογένειά μου, που με υποστήριξε σε όλες μου τις αποφάσεις με κάθε τρόπο.

Χανιά , Ιούλιος 2020

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ	4
ABSTRACT	5
ΕΥΧΑΡΙΣΤΙΕΣ	6
ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ.....	8
ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ	8
ΠΙΝΑΚΑΣ ΔΙΑΓΡΑΜΜΑΤΩΝ.....	9
ΚΕΦΑΛΑΙΟ 1	10
1.1 ΣΚΟΠΟΣ ΕΡΓΑΣΙΑΣ.....	10
1.2 ΣΥΣΤΗΜΑΤΑ ΣΥΣΤΑΣΕΩΝ	10
1.3 ΥΦΙΣΤΑΜΕΝΗ ΚΑΤΑΣΤΑΣΗ	12
ΚΕΦΑΛΑΙΟ 2 : ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	15
2.1 ΒΑΣΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΣΥΣΤΗΜΑΤΩΝ ΣΥΣΤΑΣΗΣ	15
2.1.1 Φιλτράρισμα με βάση το περιεχόμενο (Content-based Filtering).....	15
2.1.2 Συνεργατικό φιλτράρισμα (Collaborative Filtering)	16
2.1.3 Φιλτράρισμα με βάση τη γνώση (Knowledge-based Filtering).....	19
2.1.4 Δημογραφικό φιλτράρισμα (Demographic Filtering).....	20
2.1.5 Υβριδικά συστήματα συστάσεων (Hybrid Recommendation Systems)	21
2.1.6 Πολυκριτήρια συστήματα συστάσεων (Multi-criteria Recommendation Systems) ..	22
2.2 ΔΕΙΓΜΑΤΟΛΗΨΙΑ	23
2.3 ΣΤΑΤΙΣΤΙΚΑ ΜΕΤΡΑ	24
2.4 ΑΝΑΛΥΤΙΚΑ ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ. ΟΙ ΜΕΘΟΔΟΙ UTA ΚΑΙ UTASTAR	25
2.4.1 Η μέθοδος UTA.....	25
2.4.2 Η μέθοδος UTASTAR.....	27
2.5 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	30
2.6 ΕΞΟΥΡΥΞΗ ΔΕΔΟΜΕΝΩΝ	32
ΚΕΦΑΛΑΙΟ 3: ΠΡΟΤΕΙΝΟΜΕΝΗ ΜΕΘΟΔΟΛΟΓΙΑ	34
3.1 ΕΙΣΑΓΩΓΗ	34
3.2 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ	35
3.2.1 Επεξεργασία δεδομένων.....	36
3.3 ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΣΥΣΤΑΣΗΣ	40
3.4 ΑΠΟΔΟΣΗ ΤΙΜΩΝ	43
3.5 ΠΟΛΥΚΡΙΤΗΡΙΟΣ ΠΙΝΑΚΑΣ	52
3.6 UTASTAR	53
ΚΕΦΑΛΑΙΟ 4 : ΑΠΟΤΕΛΕΣΜΑΤΑ UTASTAR.....	54
ΚΕΦΑΛΑΙΟ 5 : ΣΥΜΠΕΡΑΣΜΑΤΑ.....	62
ΒΙΒΛΙΟΓΡΑΦΙΑ	63
ΠΑΡΑΡΤΗΜΑ	65
1. ΚΩΔΙΚΑΣ ΡΥΘΜΩΝ ΓΙΑ ΤΗ ΔΗΜΙΟΥΡΓΙΑ ΠΟΛΥΚΡΙΤΗΡΙΟΥ ΠΙΝΑΚΑ.....	65
2. ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ	85

ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1 : Αναπαράσταση ενός συστήματος βασιζόμενο στο φιλτράρισμα με βάση το περιεχόμενο.....	15
Σχήμα 2 : Αναπαράσταση ενός συστήματος συνεργατικού φιλτραρίσματος βασιζόμενο στο χρήστη.....	17
Σχήμα 3 : Αναπαράσταση ενός συστήματος φιλτραρίσματος βασιζόμενο στη γνώση.....	20
Σχήμα 4: Αναπαράσταση ενός υβριδικού συστήματος σύστασης.....	22
Σχήμα 5: Αναπαράσταση ενός πολυκριτήριου συστήματος σύστασης	22
Σχήμα 6 : Σχέση μεγέθους δείγματος και σφάλματος εκτίμησης.....	24
Σχήμα 7: Η κανονικοποιημένη μερική συνάρτηση αξίας	26
Σχήμα 8: Καμπύλη μονότονης παλινδρόμησης.....	28
Σχήμα 9 : Μέθοδος UTASTAR	29
Σχήμα 10 : Βήματα εξόρυξης γνώσης.....	32
Σχήμα 11 : Αρχιτεκτονική Μεθοδολογικού Πλαισίου	35

ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ

Πίνακας 1 : Δεδομένα Χρηστών από Ερωτηματολόγιο.....	35
Πίνακας 2 : Δεδομένα Χρηστών από Ερωτηματολόγιο.....	35
Πίνακας 3 : Δεδομένα Ξενοδοχείων από Booking.....	36
Πίνακας 4 : Δεδομένα Ξενοδοχείων από Booking.....	37
Πίνακας 5 : Πίνακας Δειγματοληψίας	37
Πίνακας 6 : Πίνακας Δειγματοληψίας	38
Πίνακας 7 : Κωδικοποίηση Ποιοτικών Δεδομένων για τον Υπολογισμό Στατιστικών Μέτρων	38
Πίνακας 8: Στατιστικά Μέτρα για δεδομένα χρηστών.....	39
Πίνακας 9: Στατιστικά Μέτρα για δεδομένα χρηστών.....	39
Πίνακας 10 : Χαρακτηριστικά Ξενοδοχείων – Απαιτήσεις Χρηστών.....	41
Πίνακας 11 : Μοντελοποίηση κριτηρίων	43
Πίνακας 12 : Απόδοση τιμών στο κριτήριο Τοποθεσία(1).....	44
Πίνακας 13 : Απόδοση τιμών στο κριτήριο Τοποθεσία(2).....	44
Πίνακας 14 : Κλίμακα και Κωδικοποίηση κριτηρίου Τοποθεσία	45
Πίνακας 15 : Απόδοση τιμών στο κριτήριο Καθαριότητα (1).....	45
Πίνακας 16 : Απόδοση τιμών στο κριτήριο Καθαριότητα(2).....	45
Πίνακας 17 : Κλίμακα και Κωδικοποίηση κριτηρίου Καθαριότητα	46
Πίνακας 18 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(1).....	46
Πίνακας 19 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(2).....	46
Πίνακας 20 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(3).....	46
Πίνακας 21 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(4).....	47
Πίνακας 22 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(5).....	47
Πίνακας 23 : Κλίμακα και Κωδικοποίηση κριτηρίου Παροχές Δωματίου.....	47

Πίνακας 24 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(1)	48
Πίνακας 25 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(2)	48
Πίνακας 26 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(3)	48
Πίνακας 27 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(4)	48
Πίνακας 28 : Κλίμακα και Κωδικοποίηση κριτηρίου Παροχές Ξενοδοχείου.....	49
Πίνακας 29 : Απόδοση τιμών στο κριτήριο Σχέση Ποιότητας Τιμής (1).....	49
Πίνακας 30 : Κλίμακα και Κωδικοποίηση κριτηρίου Σχέσης Ποιότητας Τιμής.....	49
Πίνακας 31 : Απόδοση τιμών στο κριτήριο Άνεση (1)	50
Πίνακας 32 : Απόδοση τιμών στο κριτήριο Άνεση (2)	50
Πίνακας 33 : Απόδοση τιμών στο κριτήριο Άνεση (3)	50
Πίνακας 34 : Απόδοση τιμών στο κριτήριο Άνεση (4)	51
Πίνακας 35 : Κλίμακα και Κωδικοποίηση κριτηρίου Άνεση.....	51
Πίνακας 36 : Κλίμακα και Κωδικοποίηση κριτηρίου Προσωπικό	51
Πίνακας 37 : Κλίμακα και Κωδικοποίηση κριτηρίου Εικόνα Μονάδας.	52
Πίνακας 38 : Πολυκριτήριο Πίνακας για User1	53
Πίνακας 39 : Πολυκριτήριο Πίνακας για User2	53
Πίνακας 40 : Πίνακας μεταδεδομένων κριτηρίων	54
Πίνακας 41 : Συντελεστής συσχέτισης.....	54
Πίνακας 42 : Ολικές Χρησιμότητες για User1.....	57
Πίνακας 43 : Ολικές Χρησιμότητες για User2	60
Πίνακας 44 : Τελικές Συστάσεις Ξενοδοχείων για User1 και User2.....	62

ΠΙΝΑΚΑΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 1 : Μερικές Χρησιμότητες των κριτηρίων Τοποθεσία και Καθαριότητα	55
Διάγραμμα 2 : Μερικές Χρησιμότητες των κριτηρίων Παροχές Δωματίου και Παροχές Ξενοδοχείου	55
Διάγραμμα 3 : Μερικές Χρησιμότητες των κριτηρίων Σχέση Ποιότητας Τιμής και Άνεση	55
Διάγραμμα 4 : Μερικές Χρησιμότητες των κριτηρίων Προσωπικό και Εικόνα Μονάδας	56
Διάγραμμα 5 : Βάρη κριτηρίων.....	56
Διάγραμμα 6 : Ολικές Χρησιμότητες εναλλακτικών.....	57

ΚΕΦΑΛΑΙΟ 1

1.1 Σκοπός Εργασίας

Ο τομέας του τουρισμού αποτελεί μια από τις σημαντικότερες βιομηχανίες τόσο στην Ελλάδα όσο και σε παγκόσμιο επίπεδο. Με την άνθηση του διαδικτύου, ο όγκος των πληροφοριών έγινε τεράστιος και κατά συνέπεια η διαδικασία επιλογής του κατάλληλου ξενοδοχείου(προορισμού) πολύ δύσκολη. Αυτό κατέστησε αναγκαία την ανάπτυξη συστημάτων συστάσεων που θα προτείνουν στους πελάτες τις κατάλληλες προς αυτούς εναλλακτικές ανάλογα με τις προτιμήσεις τους.

Σκοπός της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη μεθοδολογίας ενός πολυκριτήριου συστήματος συστάσεων για ξενοδοχειακά καταλύματα. Αρχικά γίνεται άντληση δεδομένων από την ιστοσελίδα booking.com σχετικά με τις αξιολογήσεις, τις υπηρεσίες και τα χαρακτηριστικά των ξενοδοχείων. Επίσης συλλέγουμε πληροφορίες για τις προτιμήσεις των χρηστών με την υλοποίηση ερωτηματολογίου. Έγινε προεπεξεργασία των δεδομένων και στη συνέχεια προχωρήσαμε στην καταγραφή απαιτήσεων των χρηστών και των χαρακτηριστικών των ξενοδοχείων. Έπειτα έγινε ο καθορισμός των κριτηρίων τα οποία αποτελούνται από συνδυασμούς των απαιτήσεων χρήστη – χαρακτηριστικών ξενοδοχείων. Κατόπιν, σύμφωνα με την τεχνική ανάθεσης εργασιών σε πράκτορα²⁸ γίνεται αντιστοιχία των βαθμολογιών των χαρακτηριστικών(ξενοδοχεία) και των απαιτήσεων(χρήστες). Τέλος, γίνεται ο υπολογισμός του πολυκριτήριου πίνακα και με τη χρήση του αλγορίθμου της UTASTAR καταλήγουμε στα κατάλληλα ξενοδοχεία για τον χρήστη.

1.2 Συστήματα συστάσεων

Η εκρηκτική αύξηση του αριθμού των διαθέσιμων ψηφιακών πληροφοριών και ο αριθμός των επισκεπτών στο Διαδίκτυο έχουν δημιουργήσει μια πιθανή

πρόκληση υπερφόρτωσης των πληροφοριών που παρεμποδίζει την έγκαιρη πρόσβαση σε αντικείμενα ενδιαφέροντος στο διαδίκτυο. Τα συστήματα ανάκτησης πληροφοριών, όπως το Google, το DevilFinder και το Altavista, έχουν επιλύσει εν μέρει αυτό το πρόβλημα, αλλά η απουσία των στοιχείων ήταν η ιεράρχηση και η εξατομίκευση (όπου ένα σύστημα καταγράφει το διαθέσιμο περιεχόμενο στα συμφέροντα και τις προτιμήσεις του χρήστη). Αυτό έχει αυξήσει τη ζήτηση για συστήματα συστάσεων περισσότερο από ποτέ.

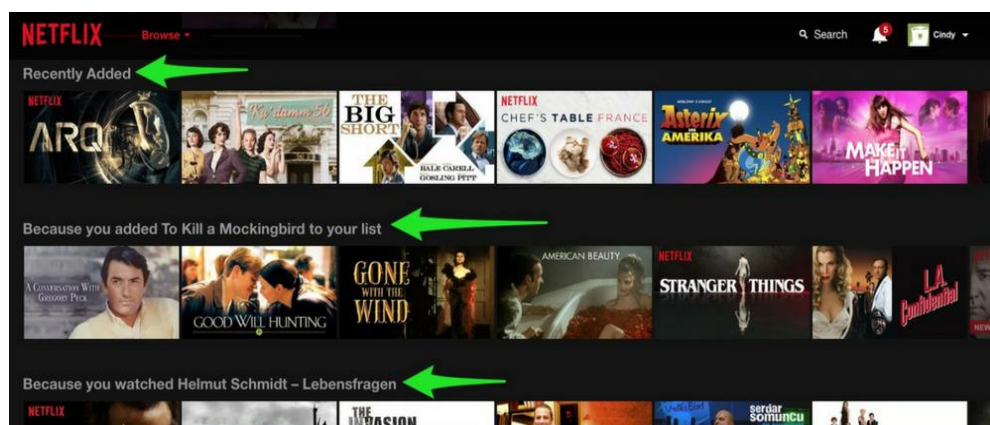
Τα συστήματα συστάσεων ορίστηκαν αρχικά ως συστήματα στα οποία οι άνθρωποι παρέχουν τις συστάσεις ως εισροές τις οποίες το σύστημα συγκεντρώνει και κατευθύνει στους κατάλληλους αποδέκτες (Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56-58.

Ένας ευρύτερος και περισσότερο γενικός ορισμός αναφέρει ότι τα συστήματα συστάσεων είναι συστήματα τα οποία έχουν την επίδραση να καθοδηγήσουν το χρήστη με έναν εξατομικευμένο τρόπο σε ενδιαφέροντα ή χρήσιμα αντικείμενα μέσα από ένα μεγάλο όγκο πιθανών επιλογών (Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.

Επίσης σύμφωνα με τους (Resnick, P., & Varian, H. R. (1997). Recommender systems. Communications of the ACM, 40(3), 56-58.

1. Burke, R. (2002). Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 12(4), 331-370.

Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems—A survey. Knowledge-Based Systems, 123, 154-162. τα συστήματα συστάσεων ορίζονται ως εργαλεία που δημιουργούν μια λίστα από αντικείμενα, μέσα από μια μεγάλη γκάμα επιλογών, που ενδιαφέρουν τον χρήστη χωρίς να απαιτείται μεγάλη αλληλεπίδραση με αυτόν.



Εικόνα 1: Παράδειγμα σύστασης για ένα συγκεκριμένο χρήστη στο Netflix

Πηγή: <https://uxdesign.cc/netflix-system-design-ef5802426ad4>

1.3 Υφιστάμενη κατάσταση

Για τη δημιουργία καλύτερων συστάσεων, πολλές μελέτες επικεντρώνονται στο να αναπτύξουν μεθοδολογίες οι οποίες εκμεταλλεύονται με τον πιο αποδοτικό τρόπο την πληροφορία που λαμβάνεται μέσω των πολυκριτήριων βαθμολογιών. Οι (Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56-58.

2. Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4), 331-370.

3. Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems—A survey. *Knowledge-Based Systems*, 123, 154-162.

Adomavicius, G., & Kwon, Y. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems*, 22(3), 48-55. προτείνουν δύο μεθοδολογίες. Στη πρώτη μεθοδολογία υπολογίζεται ένα μέτρο ομοιότητας μεταξύ των χρηστών, μέσω της απόστασης (π.χ. απόλυτης διαφοράς) που έχουν οι βαθμολογήσεις των χρηστών στα k κριτήρια. Μέσω της μεθοδολογίας αυτής προσεγγίζεται με μεγαλύτερη ακρίβεια η έννοια της ομοιότητας μεταξύ χρηστών. Σύμφωνα με τη δεύτερη μεθοδολογία, το πρόβλημα των k κριτηρίων αντιμετωπίζεται ως k διαφορετικά προβλήματα. Αρχικά υπολογίζουν την ομοιότητα μεταξύ των χρηστών λαμβάνοντας υπ' όψη μόνο το συγκεκριμένο κριτήριο. Έπειτα, έχοντας ως δεδομένο τις ομοιότητες μεταξύ χρηστών με βάση τα k κριτήρια, υπολογίζουν με τη χρήση τυπικών αλγορίθμων των συστημάτων συστάσεων, πιθανές βαθμολογίες των χρηστών σε κάθε κριτήριο. Το τελικό αποτέλεσμα στηρίζεται στη λογική ότι η συνολική προτίμηση του κάθε χρήστη και οι k πολυκριτήριες βαθμολογίες που υπολογίσθηκαν συνδέονται μέσω μιας συνθετικής συνάρτησης. Οι δύο μεθοδολογίες εξετάστηκαν με τη χρήση ενός συστήματος σύστασης ταινιών, με δεδομένα από την Yahoo! Movies τα οποία περιέχουν αξιολογήσεις ταινιών με μία συνολική βαθμολογία αλλά και βαθμολογία σε 5 επιπλέον κριτήρια: σενάριο, ηθοποιοί, σκηνοθεσία και οπτικά εφέ. Το αποτέλεσμα έδειξε ότι και οι δύο μεθοδολογίες προσφέρουν μεγαλύτερη ακρίβεια στις συστάσεις τους, σε σχέση με τις συστάσεις που θα έκανε ένα σύστημα το οποίο θα χρησιμοποιούσε μόνο τη συνολική βαθμολογία των χρηστών.

Στις περισσότερες υφιστάμενες μεθοδολογίες, οι αναλυτές στοχεύουν στη δημιουργία διαφορετικών ομάδων χρηστών με παρόμοιες προτιμήσεις και ανάγκες ή ομάδων παρόμοιων αντικειμένων. Η λογική πίσω από αυτό

βασίζεται στην έννοια της τμηματοποίησης της αγοράς. Η υποδιαίρεση της συνολικής αγοράς σε ομοιογενή τμήματα πελατών, που το καθένα μπορεί εύκολα να επιλεγεί και να αντιμετωπιστεί ως μια μικρότερη εξειδικευμένη αγορά, μπορεί να οδηγήσει στη καλύτερη κατανόηση των αναγκών και προτιμήσεων των πελατών. Ξέροντας ότι οι χρήστες δίνουν διαφορετική βαρύτητα στα διάφορα κριτήρια όταν τους ζητείται να αξιολογήσουν ένα αντικείμενο, η δημιουργία ομάδων χρηστών με παρόμοιες προτιμήσεις μπορεί να οδηγήσει στη δημιουργία περισσότερο εξατομικευμένων συστημάτων. Ένα σύστημα το οποίο επιλύει το πρόβλημα συστάσεων με αυτή τη λογική είναι αυτό των (Lakiotaki, K., Matsatsinis, N. F., & Tsoukias, A. (2011). Multicriteria user modeling in recommender systems. *IEEE Intelligent Systems*, 26(2), 64-76. Σύμφωνα με αυτό το σύστημα, δημιουργούνται διαφορετικές ομάδες από προφίλ χρηστών και στη συνέχεια εφαρμόζοντας αλγορίθμους συνεργατικού φιλτραρίσματος εξάγονται συστάσεις. Τα προφίλ που δημιουργούνται είναι αποτέλεσμα μιας διαδικασίας μοντελοποίησης των χρηστών η οποία βασίζεται στη πολυκριτήρια ανάλυση αποφάσεων. Το σύστημα εξετάσθηκε επίσης με τα δεδομένα από την Yahoo! Movies, ενώ επέφερε καλύτερα αποτελέσματα σε σχέση με τα συστήματα αξιολόγησης που χρησιμοποιούν αξιολογήσεις σε ένα χαρακτηριστικό, αλλά και σε σχέση με άλλες μεθοδολογίες που εκμεταλλεύονται τις πολυκριτήριες αξιολογήσεις.

Παρόμοιο σύστημα το οποίο αναπτύχθηκε και εφαρμόζει τεχνικές ομαδοποίησης των χρηστών με στόχο την εξατομίκευση, είναι αυτό των (Liu, L., Mehandjiev, N., & Xu, D. L. (2011, October). Multi-criteria service recommendation based on user criteria preferences. In *Proceedings of the fifth ACM conference on Recommender systems* (pp. 77-84). Σε αυτό το σύστημα εφαρμόζεται συσταδοποίηση (clustering) των χρηστών, με βάση τα κριτήρια που θεωρούν πιο σημαντικά (significant criteria) για να καταλήξουν στη τελική αξιολόγηση. Για την εύρεση των σημαντικών κριτηρίων εφαρμόζεται ένα μοντέλο γραμμικής παλινδρόμησης (linear regression model), που στόχο έχει το προσδιορισμό μιας συνθετικής συνάρτησης. Μέσω αυτής της συνάρτησης προκύπτουν συντελεστές βαρών οι οποίοι προσδιορίζουν τη σημασία που δίνει ο κάθε χρήστης σε κάθε κριτήριο.

Οι (Jannach, D., Karakaya, Z., & Gedikli, F. (2012, June). Accuracy improvements for multi-criteria recommender systems. In *Proceedings of the 13th ACM conference on electronic commerce* (pp. 674-689). χρησιμοποίησαν support vector regression (SVR) και feature selection για την περαιτέρω βελτίωση της ακρίβειας των πολυκριτήριων συστημάτων σύστασης. Η δουλειά τους βασίστηκε στα user-based και item-based μοντέλα προσέγγισης των CF συστημάτων σύστασης. Αξιολόγησαν τη μέθοδο χρησιμοποιώντας πενταπλάσια επικύρωση στα δεδομένα από την πλατφόρμα Yahoo! Movies

και τον τομέα του τουρισμού. Οι (Premchaiswadi, W., & Poompuang, P. (2013). Hybrid profiling for hybrid multicriteria recommendation based on implicit multicriteria information. *Applied Artificial Intelligence*, 27(3), 213-234. ανέπτυξαν δύο υβριδικές μεθόδους σύστασης πολλαπλών κριτηρίων, User-Item Matching (UIM) και User-Attribute-Based (UAB), για να επωφεληθούν από τα σιωπηρά δεδομένα για τη βελτίωση της ακρίβειας των συστημάτων συστάσεων πολλαπλών κριτηρίων στην περίπτωση του μέγεθος ενός γείτονα. Πραγματοποίησαν πειράματα στο σύνολο δεδομένων του MovieLens και αξιολόγησαν τις μεθόδους χρησιμοποιώντας Precision και MAE μετρήσεις. Τα αποτελέσματα των εργασιών τους έδειξαν ότι η ακρίβεια των προτεινόμενων μεθόδων ήταν σημαντικά υψηλότερη από ότι η CF προσεγγίζει τη βιβλιογραφία.

Οι (Xu, Y., Li, K., Hu, J., & Li, K. (2014). A genetic algorithm for task scheduling on heterogeneous computing systems using multiple priority queues. *Information Sciences*, 270, 255-287. πρότειναν μια προσέγγιση πολλαπλών κριτηρίων CF με τη χρήση της σχέσης αδιαφορίας ως ένα μέτρο ομοιότητας και εφάρμοσαν το Genetic Algorithm (GA) για τη μεγιστοποίηση της ακρίβειας στην κορυφή-N. Για την αξιολόγηση της μεθόδου, τα δεδομένα συγκεντρώθηκαν σε ιστότοπο αγοράς στην Ταϊβάν από δείγμα 211 προπτυχιακών φοιτητών που ήταν εξοικειωμένοι με το group-buying. Ο συγγραφέας χρησιμοποίησε μετρήσεις Precision (ακρίβεια σε κορυφή-1, ακρίβεια σε κορυφή-3 και ακρίβεια στην κορυφή-5) για να δείξει την απόδοση του συστήματος σύστασης. Οι (Bilge, A., & Kaleli, C. (2014, May). A multi-criteria item-based collaborative filtering framework. In 2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 18-22). IEEE. ανέπτυξαν ένα πολυκριτηριακό CF item-based πλαίσιο χρησιμοποιώντας regression-based προσέγγιση. Χρησιμοποίησαν τη MAE με πενταπλάσια εγκυρότητα επικύρωσης και αξιολόγησαν τη μέθοδο στο σύνολο δεδομένων της πλατφόρμας Yahoo! Movies.

Η χρησιμότητα των συστημάτων συστάσεων έχει αποδειχθεί έμπρακτα στον τομέα των επιχειρήσεων (e-commerce sites) με αποτέλεσμα να δίδεται μεγάλη έμφαση στην ανάπτυξη αλγορίθμων και μεθοδολογιών που είναι σε θέση όχι μόνο να προσφέρουν αξιόπιστες συστάσεις αλλά και να ελαχιστοποιούν το χρόνο και το υπολογιστικό κόστος. Πολλοί αλγόριθμοι αποδεδειγμένα προσφέρουν αξιόπιστες προβλέψεις, αλλά στη πράξη, αποδίδουν πολύ αργά για να εφαρμοστούν σε ιστοσελίδες ηλεκτρονικών καταστημάτων που λειτουργούν σε πραγματικό χρόνο. Ένας από τους αλγόριθμους που έχει αναπτυχθεί λαμβάνοντας υπόψη την ελαχιστοποίηση του υπολογιστικού κόστους και χρόνου, είναι αυτός των (Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering

recommendation algorithms. In Proceedings of the 10th international conference on World Wide Web (pp. 285-295). Στον αλγόριθμο αυτό υπολογίζονται ομοιότητες και εξετάζονται συσχετίσεις μεταξύ αντικειμένων. Οι συστάσεις που εξάγονται είναι σετ αντικειμένων τα οποία έχουν τη μεγαλύτερη ομοιότητα με τα αντικείμενα τα οποία έχει αξιολογήσει θετικά ο χρήστης στο παρελθόν. Τα μεγάλα ηλεκτρονικά καταστήματα αλλά και οι ιστοσελίδες που παρέχουν υπηρεσίες (Amazon, E-bay, Booking, Trip - Advisor), έχουν πολλούς περισσότερους χρήστες απ' ότι προϊόντα. Σε αυτές τις περιπτώσεις ένας item-based αλγόριθμος έχει πιο γρήγορη ανταπόκριση στην εξαγωγή συστάσεων, ειδικά όταν οι σχέσεις μεταξύ των αντικειμένων έχουν υπολογισθεί εκ των προτέρων. Άλλοι αλγόριθμοι με πρακτικό ενδιαφέρον για online εφαρμογή σε ηλεκτρονικά καταστήματα χρησιμοποιούν τεχνικές μείωσης των διαστάσεων (π.χ. principal component analysis, singular value decomposition) για να εξάγουν ακριβείς συστάσεις (Billsus, D., & Pazzani, M. J. (1998, July). Learning Collaborative Information Filters. In Icml (Vol. 98, pp. 46-54). Αυτές οι τεχνικές βασίζονται στη λογική ότι ο πίνακας χρηστών \times αντικειμένων, μπορεί να έχει επιπλέον, ανεξάρτητες από τα υπόλοιπα δεδομένα διαστάσεις οι οποίες "κρύβουν" πληροφορίες για τις προτιμήσεις των χρηστών. Με τη χρήση αυτών των τεχνικών, μειώνονται οι διαστάσεις των δεδομένων, ώστε να περιέχουν μόνο την απαραίτητη πληροφορία σχετικά με τις σχέσεις μεταξύ χρηστών ή αντικειμένων. Η διαδικασία μείωσης των διαστάσεων θεωρείται σχετικά χρονοβόρα, εφόσον όμως πραγματοποιηθεί όλες οι παραδοσιακές τεχνικές και αλγόριθμοι συστάσεων εξάγουν αποτελέσματα γρηγορότερα.

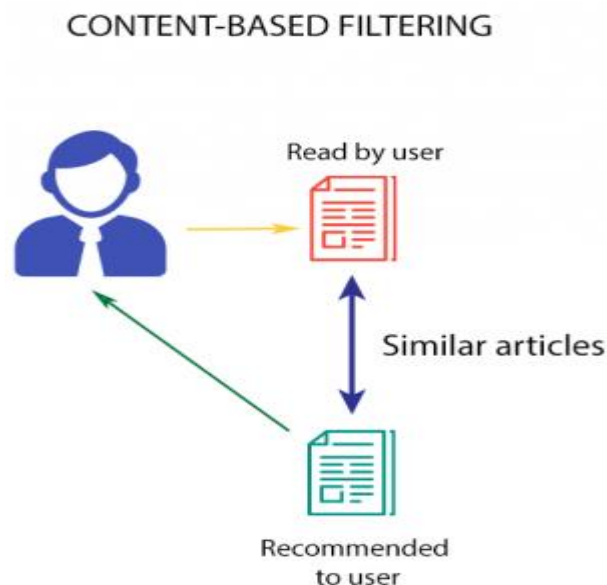
Επίσης, μεγάλο ενδιαφέρον παρουσιάζουν τα συστήματα σύστασης για m-commerce platforms. Οι (Li, Q., Wang, C., & Geng, G. (2008, April). Improving personalized services in mobile commerce by a novel multicriteria rating approach. In Proceedings of the 17th international conference on World Wide Web (pp. 1235-1236). χρησιμοποίησαν την τεχνική της Multilinear Singular Value Decomposition (MSVD) για την ανάπτυξη ενός πολυκριτήριου συστήματος σύστασης για μια mobile-commerce platform. Για να αξιολογήσουν τη μέθοδο, πραγματοποίησαν αρκετά πειράματα χρησιμοποιώντας δεδομένα ενός εστιατορίου. Τα αποτελέσματα της μελέτης τους έδειξαν ότι οι υψηλότερες τιμές ακρίβειας και ανάκλησης μπορεί να αποκτηθούν όταν συγκρίνονται με τις μεθόδους που βασίζονται αποκλειστικά σε ανατροφοδότηση με ενιαία αξιολόγηση.

ΚΕΦΑΛΑΙΟ 2 : ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 Βασικοί αλγόριθμοι συστημάτων σύστασης

2.1.1 Φιλτράρισμα με βάση το περιεχόμενο (Content-based Filtering)

Οι μέθοδοι φιλτραρίσματος βάσει περιεχομένου βασίζονται σε μια περιγραφή του στοιχείου και ενός προφίλ των προτιμήσεων του χρήστη. Σε ένα σύστημα σύστασης με βάση το περιεχόμενο, χρησιμοποιούνται λέξεις-κλειδιά για να περιγράψουν τα στοιχεία και δημιουργείται ένα προφίλ χρήστη για να υποδείξει τον τύπο αντικειμένου που θέλει ο χρήστης. Με άλλα λόγια, αυτοί οι αλγόριθμοι προσπαθούν να προτείνουν αντικείμενα που είναι παρόμοια με αυτά που άρεσαν στο χρήστη στο παρελθόν. Συγκεκριμένα, τα διάφορα υποψήφια στοιχεία συγκρίνονται με αντικείμενα που έχουν προηγουμένως βαθμολογηθεί από το χρήστη και συνιστώνται αυτά που ταιριάζουν καλύτερα. Αυτή η προσέγγιση έχει τις ρίζες της στην ανάκτηση πληροφοριών και την έρευνα φιλτραρίσματος πληροφοριών.



Σχήμα 1 : Αναπαράσταση ενός συστήματος βασισμένο στο φιλτράρισμα με βάση το περιεχόμενο

Πηγή: www.marutitech.com/recommendation-engine-benefits

Προβλήματα Των Συστημάτων Βασισμένα Στο Περιεχόμενο

- **Περιορισμένη ανάλυση περιεχομένου:**

Τα συστήματα τα οποία λειτουργούν με βάση το περιεχόμενό τους δυσκολεύονται να αξιολογήσουν το περιεχόμενο ενός βίντεο ή μιας φωτογραφίας. Η αξιολόγηση τους γίνεται με την χρήση αυτοματοποιημένων αλγόριθμων και το αποτέλεσμα τους δεν είναι σίγουρο πως μπορεί να είναι πάντα σωστό. Επιπλέον σε κάποια κείμενα τα οποία χαρακτηρίζονται με το ίδιο προφίλ, τα συστήματα συστάσεων με βάση το περιεχόμενο δεν μπορούν να ξεχωρίσουν πιο κείμενο είναι αυτό που χρειάζεται ο χρήστης και πιο όχι.

- **Πρόβλημα νέου χρήστη:**

Τα συστήματα με βάση το περιεχόμενο, όπως και στο συνεργατικό φιλτράρισμα προτείνουν αντικείμενα τα οποία έχουν παρόμοιο περιεχόμενο με τα αντικείμενα που ο χρήστης επέλεξε στο παρελθόν. Όταν ο χρήστης δεν έχει δηλώσει την προτίμησή του σε κάποια αντικείμενα, το σύστημα δεν μπορεί να εξαγάγει το περιεχόμενο τους και να προτείνει παρόμοια.

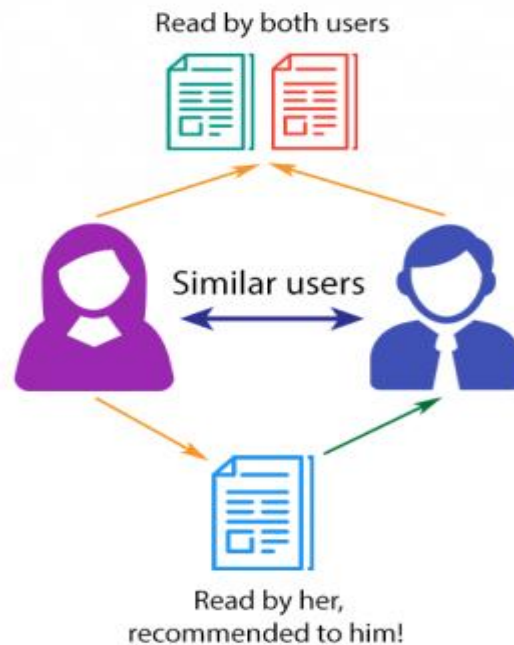
- **Το πρόβλημα της υπεριδίκευσης:**

Επιπλέον τα συστήματα αυτά επιλέγουν πληροφορίες οι οποίες έχουν υψηλό βαθμό συσχέτισης με το προφίλ του χρήστη. Έτσι αν ένας χρήστης αναζητήσει κάτι το οποίο δεν ταιριάζει με το προφίλ που είχε δημιουργήσει τότε τα αποτελέσματα θα ήταν πολύ περιορισμένα σε σύγκριση με άλλους χρήστες. Για την αντιμετώπιση αυτού του προβλήματος έχουν δημιουργηθεί αλγόριθμοι οι οποίοι προσθέτουν τυχαιότητα στις προτάσεις του συστήματος.

2.1.2 Συνεργατικό φιλτράρισμα (Collaborative Filtering)

Το συνεργατικό φιλτράρισμα αποτελεί έναν πολύ διαδεδομένο αλγόριθμο σύστασης σύμφωνα με τον οποίο οι συστάσεις για κάθε χρήστη προέρχονται με βάση τις προτιμήσεις που έχουν άλλοι χρήστες με παρόμοια συμπεριφορά. Στα σύστημα συστάσεων βασισμένα στο συνεργατικό φιλτράρισμα τα δεδομένα σχετικά με τις προτιμήσεις των χρηστών καταγράφονται σε ένα πίνακα χρηστών \times αντικειμένων του οποίου οι γραμμές αναφέρονται σε διαφορετικούς χρήστες και κάθε εγγραφή υποδεικνύει τη προτίμηση του χρήστη u για το αντικείμενο i . Οι τεχνικές λοιπόν που ανήκουν σε αυτή τη κατηγορία φιλτραρίσματος χρησιμοποιούν τις βαθμολογήσεις των χρηστών πάνω σε ορισμένα αντικείμενα, για να καθορίσουν τη σχέση μεταξύ τους και να εξαγάγουν προβλέψεις σε βαθμολογήσεις που θα έδινε κάθε χρήστης σε κάποιο νέο αντικείμενο.

COLLABORATIVE FILTERING



Σχήμα 2 : Αναπαράσταση ενός συστήματος συνεργατικού φιλτραρίσματος βασισμένο στο χρήστη

Πηγή: www.marutitech.com/recommendation-engine-benefits

- **Προβλήματα των Συστημάτων Βασισμένα στην Συνεργασία**

Τα συστήματα που βασίζονται στην συνεργασία έχουν ορισμένους περιορισμούς στην λειτουργία τους οι οποίοι είναι παρόμοιοι με τους περιορισμούς που έχουν και τα συστήματα βασισμένα στο περιεχόμενο αλλά έχουν και κάποιους ακόμα. Οι πιο σημαντικοί περιορισμοί είναι:

- **Το πρόβλημα των αραιών δεδομένων (Sparsity Problem):**

Ένα πρόβλημα το οποίο κάνει συχνά την εμφάνισή του είναι το γεγονός πως εμφανίζεται ένας μεγάλος αριθμός εναλλακτικών αποτελεσμάτων, το πρόβλημα αυτό χαρακτηρίζεται ως πρόβλημα χαμηλής κάλυψης. Το ίδιο πρόβλημα έχουν ακόμα και οι χρήστες που έχουν αρκετές αξιολογήσεις. Για παράδειγμα αν ένας χρήστης με πολλές αξιολογήσεις κάνει μια αναζήτηση σε μια βιβλιοθήκη με μεγάλο αριθμό βιβλίων τα αποτελέσματα που θα πάρει θα είναι πάρα πολλά για να μπορέσει να τα ψάξει όλα. Αυτό συμβαίνει γιατί τα συστήματα βασισμένα στην συνεργασία δεν μπορούν να εντοπίσουν γειτονικούς χρήστες και έτσι δεν μπορούν να εξαγάγουν ορθά αποτελέσματα.

Για να λυθεί αυτό το πρόβλημα προτάθηκε μια λύση η οποία αναφέρει πως η ύπαρξη περισσότερων στοιχείων για τους χρήστες όπως το φύλο, η ηλικία, τα ενδιαφέροντα και άλλα μπορούν να δώσουν την δυνατότητα στα συστήματα να κάνουν καλύτερους συσχετισμούς στα δεδομένα τους και να εμφανίζουν πιο αξιόπιστα δεδομένα (Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, 13(5-6), 393-408.

Υπάρχουν και άλλες προτεινόμενες λύσεις οι οποίες προϋποθέτουν την ύπαρξη ενός αλγορίθμου. Μια τέτοια λύση είναι και η χρήση του αλγορίθμου SVD (Singular Value Decomposition) ο οποίος χρησιμοποιείται για να εντοπίσει σχέσεις μεταξύ των χρηστών και των εναλλακτικών με σκοπό να χρησιμοποιηθούν για τον υπολογισμό των εναλλακτικών για τους χρήστες. Αφού γίνει αυτό τότε ο αλγόριθμος περιορίζει τα αποτελέσματα σε έναν αριθμό όπου να μπορούν να γίνουν οι κατάλληλες συστάσεις στους χρήστες **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**

- **Το πρόβλημα της Ψυχρής Εκκίνησης (Cold Start Problem):**

Το πρόβλημα της ψυχρής εκκίνησης είναι όμοιο με το πρόβλημα των αραιών δεδομένων. Το πρόβλημα αυτό εμφανίζεται όταν εμφανίζονται νέοι χρήστες ή νέες εναλλακτικές (Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002, August). Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 253-260).

- **Το πρόβλημα Του Νέου Χρήστη (New User Problem):**

Κατά την εγγραφή και την είσοδο ενός νέου χρήστη στο σύστημα πρέπει αυτός να δημιουργήσει ένα προφίλ το οποίο θα περιέχει στοιχεία γι' αυτόν που θα βοηθούν το σύστημα να του κάνει αξιόπιστες προτάσεις. Όμως ακόμα και αν ο χρήστης δώσει στο σύστημα ότι στοιχεία χρειάζεται τα αποτελέσματα και πάλι δεν θα είναι αξιόπιστα μέχρι ο χρήστης να κάνει έναν αριθμό αξιολογήσεων ώστε να μπορούν να γίνουν οι κατάλληλες συγκρίσεις. Το πρόβλημα αυτό ονομάζεται το πρόβλημα του νέου χρήστη.

Για την αντιμετώπιση αυτού του προβλήματος έχουν αναφερθεί πολλές διαφορετικές λύσεις όπως η χρήση των δημογραφικών προβλέψεων προτιμήσεων και άλλες οι οποίες χρησιμοποιούν έναν συνδυασμό από

τεχνικές συστημάτων περιεχομένου και συστημάτων βασισμένα στην συνεργασία.

Ακόμα υπάρχουν και άλλες τεχνικές οι οποίες δεν βοηθούν στο να γίνουν καλύτεροι οι υπολογισμοί των εναλλακτικών αλλά βοηθούν στην συμπλήρωση του προφίλ του χρήστη με πιο γρήγορο τρόπο. Οι τεχνικές αυτές χρησιμοποιούν διάφορες μεθόδους για να πετύχουν τον σκοπό τους όπως την χρήση των εναλλακτικών που εμφανίζουν μεγάλη δημοτικότητα και άλλες (Rashid, A. M., Albert, I., Cosley, D., Lam, S. K., McNee, S. M., Konstan, J. A., & Riedl, J. (2002, January). Getting to know you: learning new user preferences in recommender systems. In Proceedings of the 7th international conference on Intelligent user interfaces (pp. 127-134).).

- **Το Πρόβλημα του Νέου Αντικειμένου (New Item Problem):**

Στα συστήματα όπου υπάρχουν προτάσεις, το φαινόμενο να υπάρχουν νέα αντικείμενα όπως νέα βιβλία, νέα τραγούδια και άλλα είναι πολύ συχνό. Όμως βάσει των συστημάτων τα οποία βασίζονται στην συνεργασία για να εμφανιστεί μια εναλλακτική ως αποτέλεσμα της αναζήτησης ενός χρήστη θα πρέπει να υπάρχει τουλάχιστον μια αξιολόγηση γι' αυτήν την εναλλακτική. Αυτό έχει σαν αποτέλεσμα οι νέες εναλλακτικές να μην εμφανίζονται ως αποτελέσματα στο σύστημα μέχρι να αξιολογηθούν (Breese, J. S., Heckerman, D., & Kadie, C. (2013). Empirical analysis of predictive algorithms for collaborative filtering. arXiv preprint arXiv:1301.7363

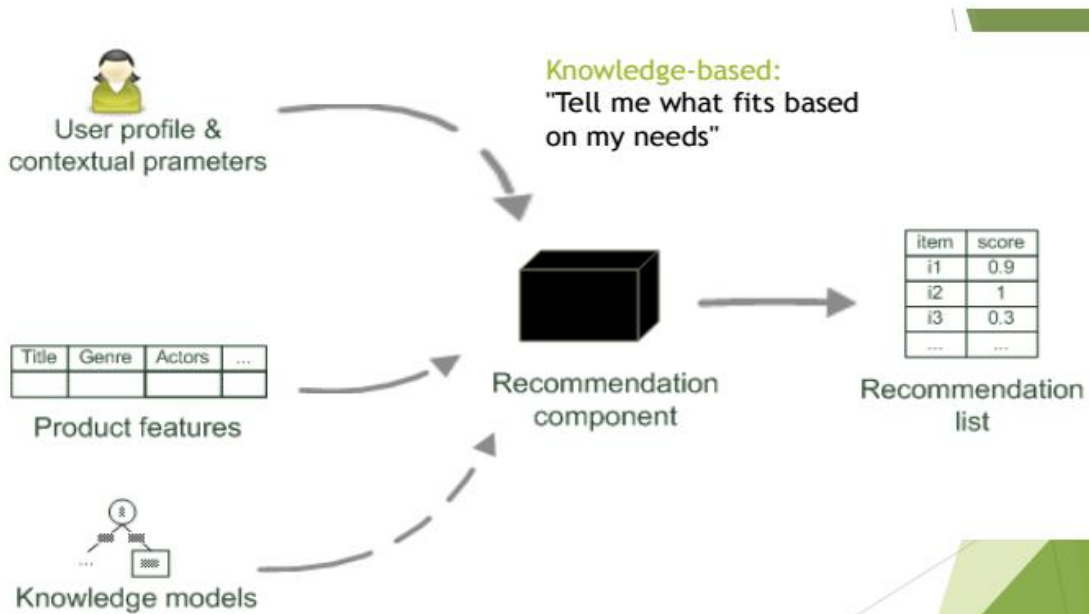
Burke, R. (2003, August). Hybrid systems for personalized recommendations. In IJCAI Workshop on Intelligent Techniques for Web Personalization (pp. 133-152). Springer, Berlin, Heidelberg..

Για να αντιμετωπιστεί αυτό το πρόβλημα θα πρέπει να χρησιμοποιηθεί ένας συνδυασμός από τα συστήματα με βάση το περιεχόμενο και τα συστήματα με βάση την συνεργασία.

2.1.3 Φιλτράρισμα με βάση τη γνώση (Knowledge-based Filtering)

Τα συστήματα που χρησιμοποιούν φιλτράρισμα με βάση τη γνώση, απαιτούν γνώση για τις ανάγκες του κάθε χρήστη καθώς και γνώση για το πως κάθε αντικείμενο μπορεί να καλύψει αυτές τις ανάγκες (π.χ. όταν

αγοράσουμε ένα κινητό τηλέφωνο, αυτά τα συστήματα θα κάνουν την σύσταση να αγοράσουμε μια θήκη ή ακουστικά). Έπειτα καθορίζεται η σχέση μεταξύ των αναγκών που μπορεί να καλύψει το αντικείμενο και των αναγκών του χρήστη. Η χρήση συστημάτων με βάση τη γνώση επιλέγεται μόνο όταν οι παραδοσιακές προσεγγίσεις (με βάση το περιεχόμενο και συνεργατικό φιλτράρισμα) δεν μπορούν να εφαρμοστούν αποτελεσματικά και να μας δώσουν ακριβείς συστάσεις.



Σχήμα 3 : Αναπαράσταση ενός συστήματος φιλτραρίσματος βασισμένο στη γνώση

Πηγή: <https://petrlorenc.github.io/Recommendation-systems/>

2.1.4 Δημογραφικό φιλτράρισμα (Demographic Filtering)

Τα συστήματα συστάσεων με δημογραφική προσέγγιση χρησιμοποιούν τα δημογραφικά χαρακτηριστικά των χρηστών (την ηλικία, το φύλο, το επάγγελμα, κ.α.) ως βασικούς παράγοντες πάνω στους οποίους βασίζονται οι συστάσεις που εξάγουν. Τα συστήματα αυτά ακολουθούν τη λογική του συνεργατικού φιλτραρίσματος, ότι δηλαδή παρόμοιοι χρήστες είναι πολύ πιθανόν να προτιμούν τα ίδια αντικείμενα .

Το βασικό τους πλεονέκτημα που αποτελεί και την κύρια διαφορά τους με τα συνεργατικά, είναι ότι η δημογραφική προσέγγιση δεν χρησιμοποιεί αξιολογήσεις αντικειμένων από χρήστες για τη δημιουργία προφίλ, αλλά δημογραφικά στοιχεία. Τα κύρια μειονεκτήματα των συστημάτων που χρησιμοποιούν δημογραφική προσέγγιση, απορρέουν από το ότι

χρησιμοποιούν μια πολύ γενικευμένη λογική για τη συλλογή των ενδιαφερόντων των χρηστών και στο ότι δεν είναι όλοι οι χρήστες διατεθειμένοι να μοιραστούν τα στοιχεία τους έτσι εύκολα. Επιπλέον σε τέτοιου είδους συστήματα, υπάρχει αδυναμία προσαρμογής στις αλλαγές των προτιμήσεων των χρηστών στο πέρασμα του χρόνου. Παρόλα αυτά, τα δημογραφικά χαρακτηριστικά μπορούν να αποτελέσουν χρήσιμη πληροφορία εάν συμπεριληφθούν σε άλλες προσεγγίσεις.

2.1.5 Υβριδικά συστήματα συστάσεων (Hybrid Recommendation Systems)

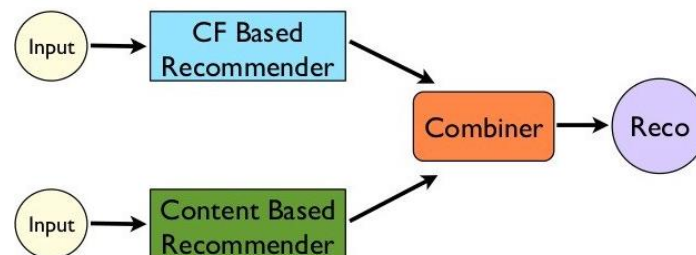
Τα υβριδικά συστήματα αποτελούν ένα συνδυασμό δύο ή περισσότερων τεχνικών σύστασης που έχουν σκοπό την επίτευξη καλύτερων επιδόσεων, ελαχιστοποιώντας τα εκάστοτε μειονεκτήματα τους. Ένα καλό παράδειγμα χρήσης υβριδικών συστημάτων σύστασης είναι το Netflix. Ο ιστότοπος κάνει συστάσεις συγκρίνοντας τις συνήθειες παρακολούθησης και αναζήτησης παρόμοιων χρηστών (Συνεργατικό φιλτράρισμα) καθώς και προσφέροντας ταινίες που μοιράζονται χαρακτηριστικά με ταινίες που ο χρήστης έχει αξιολογήσει υψηλά (φιλτράρισμα βάσει περιεχομένου).

Σύμφωνα με τον (Burke, R. (2003, August). Hybrid systems for personalized recommendations. In IJCAI Workshop on Intelligent Techniques for Web Personalization (pp. 133-152). Springer, Berlin, Heidelberg.. υπάρχουν επτά τρόποι σύμφωνα με τους οποίους μπορούν να συνδυαστούν συστήματα συστάσεων σε ένα υβριδικό πλαίσιο:

- Εφαρμόζοντας ξεχωριστά τα διαφορετικά συστήματα και παρουσιάζοντας τα αποτελέσματά τους είτε μαζί, είτε σε ξεχωριστές λίστες.
- Σταθμίζοντας τα αποτελέσματα των επιμέρους συστημάτων για να εξάγουμε ένα ενιαίο αποτέλεσμα.
- Επιλέγοντας με βάση κάποιο κριτήριο να χρησιμοποιήσουμε τα αποτελέσματα από ένα μόνο σύστημα.
- Βελτιστοποιώντας τις εξαγόμενες συστάσεις του ενός συστήματος με χρήση των άλλων.
- Συνδυάζοντας δεδομένα από διαφορετικές πηγές και αναλύοντάς τα από ένα σύστημα συστάσεων.

- Χρησιμοποιώντας σαν είσοδο του ενός συστήματος τις εξαγόμενες συστάσεις του άλλου.
- Δημιουργώντας ένα μοντέλο βασισμένο σε ένα σύστημα, το οποίο μετά χρησιμοποιείται σαν είσοδο στο δεύτερο.

Hybrid Recommendations

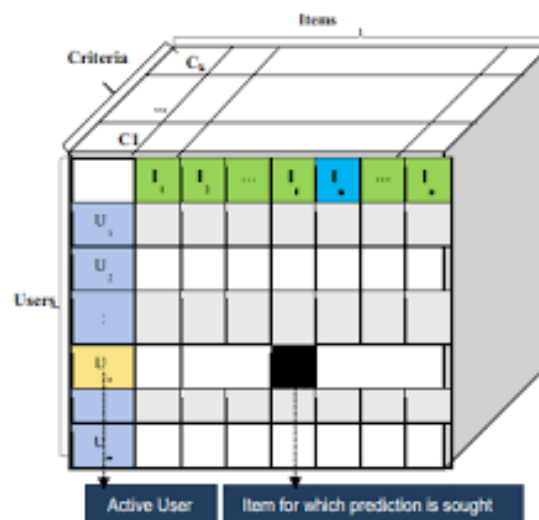


Σχήμα 4: Αναπαράσταση ενός υβριδικού συστήματος σύστασης

Πηγή: <https://dataconomy.com/2015/03/an-introduction-to-recommendation-engines/>

2.1.6 Πολυκριτήρια συστήματα συστάσεων (Multi-criteria Recommendation Systems)

Τα πολυκριτήρια συστήματα συστάσεων ορίζονται ως συστήματα συστάσεων στα οποία οι προτιμήσεις των χρηστών καθορίζονται πάνω σε πολλαπλά κριτήρια (Σχήμα 5: Αναπαράσταση ενός πολυκριτήριου συστήματος σύστασης).



Σχήμα 5: Αναπαράσταση ενός πολυκριτήριου συστήματος σύστασης

Πηγή: Mehrbakhsh N., Othman I. (2015).

Οι συνηθισμένες τεχνικές φιλτραρίσματος που αναφέραμε παραπάνω χρησιμοποιούν βαθμολογίες πάνω σε ένα χαρακτηριστικό για να προσδιορίσουν τη συνολική ικανοποίηση του χρήστη από ένα αντικείμενο και κατ' επέκταση να εξάγουν συστάσεις. Εντούτοις, σε πολλές περιπτώσεις, η αξιολόγηση ενός μόνο χαρακτηριστικού δεν αντικατοπτρίζει πλήρως τη συνολική προτίμηση των χρηστών στα αντικείμενα.

Η πολυκριτήρια προσέγγιση, εν αντιθέσει με τις άλλες τεχνικές, στοχεύει στον προσδιορισμό μιας βαθμολογίας για κάθε νέο, ως προς το χρήστη αντικείμενο, χρησιμοποιώντας τις προτιμήσεις του πάνω σε πολλαπλά κριτήρια.

Τα πολυκριτήρια συστήματα συστάσεων αν και στις περισσότερες περιπτώσεις συνδυάζονται με τις παραδοσιακές τεχνικές αποτελούν πλέον ξεχωριστή περίπτωση. Αυτό διότι έχει αναπτυχθεί πληθώρα αλγορίθμων και τεχνικών (προερχόμενων κυρίως από τον τομέα της Επιστήμης των Αποφάσεων) για να εκμεταλλευτούν πλήρως τη πληροφορία που λαμβάνεται από τις βαθμολογήσεις σε περισσότερα από ένα χαρακτηριστικά.

2.2 Δειγματοληψία

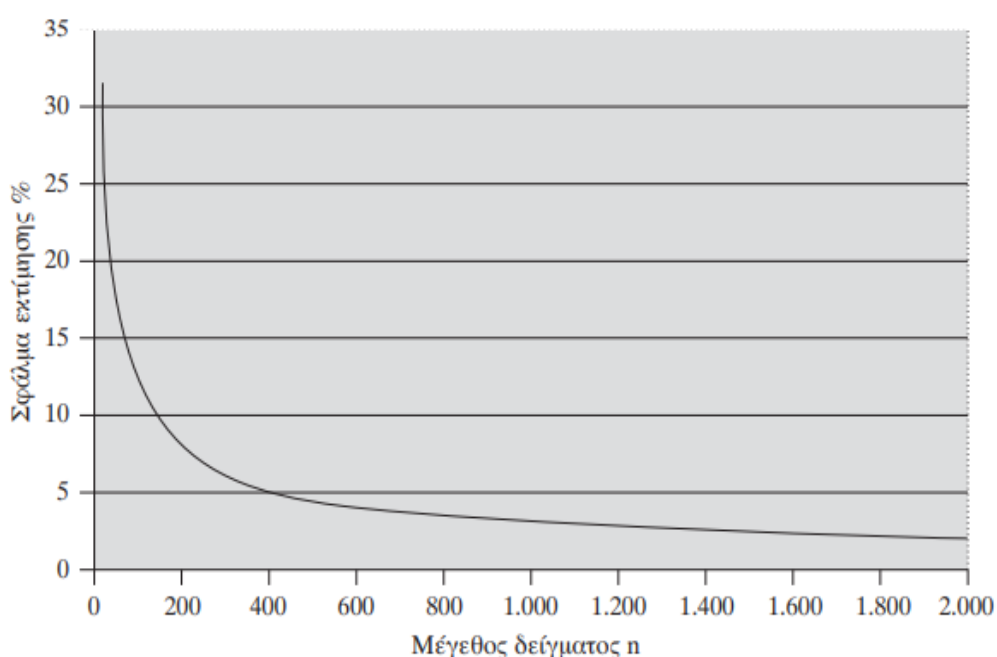
Η διαδικασία σχεδιασμού του δείγματος μιας έρευνας αποτελεί τον ακρογωνιαίό λίθο της. Το μέγεθος και οι ιδιότητες του συνολικού πληθυσμού, η μέθοδος δειγματοληψίας καθώς και το βέλτιστο μέγεθος του δείγματος, είναι οι κυριότερες παράμετροι της διαδικασίας.

Η δειγματοληψία αφορά τη λήψη ενός τμήματος από κάποιο ευρύτερο σύνολο . Γενικά , η δειγματοληψία θεωρείται επιτυχής όταν η επιλογή του δείγματος παράγει αποτελέσματα, δείκτες και μετρήσεις που είναι γενικεύσιμα και όσο το δυνατόν ακριβέστερα, δηλαδή βρίσκονται πιο κοντά στις αντίστοιχες παραμέτρους του ευρύτερου συνόλου, δηλαδή του πληθυσμού . Ο πληθυσμός είναι το ευρύ σύνολο των υποκειμένων για τα οποία εξάγουμε συμπεράσματα.

Το πλήθος των μονάδων του δείγματος λέγεται μέγεθος δείγματος (sample size). Η διεξαγωγή αποτελεσμάτων σε μια έρευνα επηρεάζεται σημαντικά από το μέγεθος του δείγματος. Είναι αυτονόητο, ότι όσο πιο μεγάλο είναι το μέγεθος του δείγματος, τόσο περισσότερο αξιόπιστα θα είναι τα συμπεράσματα.

Πόσο μεγάλο πρέπει να είναι το δείγμα ουσιαστικά καθορίζεται από την ακρίβεια εκτίμησης που θέλουμε να επιτύχουμε. Δηλαδή, καθορίζεται από τον βαθμό σφάλματος που ο ερευνητής μπορεί να ανεχτεί στην έρευνά του.

Το μέγιστο σφάλμα εκτίμησης ενός ποσοστού σε μια έρευνα είναι ίσο με $e = \frac{1}{\sqrt{n}}$ όπου n είναι το μέγεθος του δείγματος που χρησιμοποιείται. Καλείται σφάλμα εκτίμησης σε επίπεδο σημαντικότητας 95% ή σε στάθμη σημαντικότητας 95% και εκφράζει το μέγιστο σφάλμα που μπορεί να έχουμε εξαιτίας της δειγματοληψίας από τις πραγματικές παραμέτρους, και αυτό εμφανίζεται όταν το ποσοστό είναι 50% στο δείγμα. Όταν το ποσοστό παίρνει πιο ακραίες τιμές, όπως 20% ή 80%, το σφάλμα εκτίμησης είναι μικρότερο. Γενικά, σε έρευνες με ερωτηματολόγιο συνηθίζεται να αναφέρουμε τη μέγιστη τιμή σφάλματος που αφορά όλες τις ερωτήσεις. Για το δείγμα μεγέθους $n = 500$ αυτή η τιμή είναι $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{500}} 0,0447 = 4,47\%$ σε επίπεδο σημαντικότητας 95%.



Πηγή: Ψαρρού και Ζαφειρόπουλος, 2001.

Σχήμα 6 : Σχέση μεγέθους δείγματος και σφάλματος εκτίμησης

2.3 Στατιστικά Μέτρα

Μετά τη συλλογή των δεδομένων κρίνεται αναγκαία η εύρεση διαδικασιών με τις οποίες θα καταφέρουμε να οργανώσουμε και να παρουσιάσουμε απλούστερα τα δεδομένα. Αυτό το επιτυγχάνουμε με την περιγραφική στατιστική (descriptive statistics), μια διαδικασία που περιέχει τεχνικές για την συνοπτική και αποτελεσματική παρουσίαση των δεδομένων. Η πιο διαδεδομένη τεχνική είναι ο υπολογισμός της μέσης τιμής (mean) και της

τυπικής απόκλισης (standard deviation). Επίσης παρουσιάζονται οι μέγιστη και η ελάχιστη τιμή καθώς και ο διαχωρισμός του δείγματος σε 4 ίσα ποσοστιαία σημεία (τεταρτημόρια), τα P_{25} , P_{50} , P_{75} .

Τα ποσοστιαία σημεία (ποσοστημόρια) του δείγματος αποτελούν μια γενίκευση της διαμέσου και δίνουν αναλυτικότερη περιγραφή της θέσης της κατανομής των παρατηρήσεων. Το ποσοστημόριο P_a είναι το σημείο της κατανομής για το οποίο το $a\%$ των παρατηρήσεων είναι μικρότερες ή ίσες από αυτό και το υπόλοιπο $(1-a)\%$ των παρατηρήσεων είναι μεγαλύτερες ή ίσες από αυτό. Τα ποσοστημόρια υποδιαιρούν την κατανομή των παρατηρήσεων σε «ίσα» τμήματα, όχι με όρους μονάδων μέτρησης των παρατηρήσεων (δηλαδή απόστασης), αλλά με όρους ποσοστών. Δηλαδή, τα τμήματα αυτά είναι «ίσα» με την έννοια ότι περιέχουν ίσα ποσοστά παρατηρήσεων.

2.4 Αναλυτικά μοντέλα αποφάσεων. Οι μέθοδοι UTA και UTASTAR

2.4.1 Η μέθοδος UTA

Σύμφωνα με τους (Jacquet-Lagrange and Siskos, 1982; Siskos, Y., Grigoroudis, E., & Matsatsinis, N. F. (2016). UTA methods. In Multiple criteria decision analysis (pp. 315-362). Springer, New York, NY. η μέθοδος UTA (Utility Additives), είναι από τις πιο γνωστές πολυκριτήριες μεθοδολογίες λήψης αποφάσεων και θέτει τις βάσεις της σύγχρονης Αναλυτικής-Συνθετικής προσέγγισης. Έχει ως στόχο την εκτίμηση μιας ή περισσότερων προσθετικών συναρτήσεων αξίας από μία προδιάταξη ενός συνόλου αναφοράς A_R . Η μέθοδος χρησιμοποιεί ειδικές τεχνικές γραμμικού προγραμματισμού για να καθορίσει τις συγκεκριμένες συναρτήσεις, έτσι ώστε η κατάταξη που αποκτάται μέσω αυτών των συναρτήσεων στο A_R να είναι όσο το δυνατόν πιο συμβατή με την αρχική προ διάταξη.

Το μοντέλο σύνθεσης των κριτηρίων (μοντέλο απόφασης) στη μέθοδο UTA είναι μία προσθετική συνάρτηση αξίας της ακόλουθης μορφής:

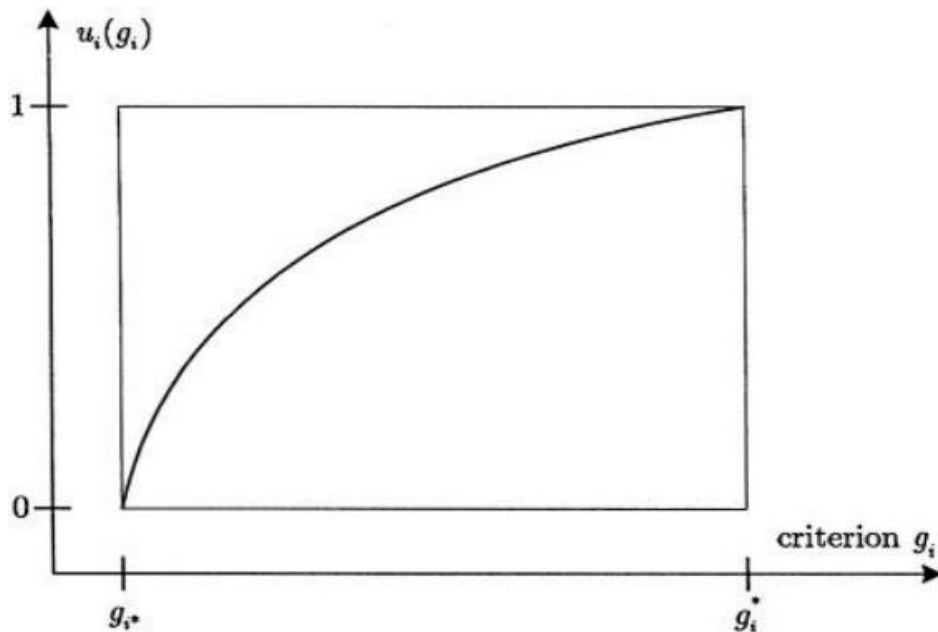
$$u(g) = \sum_{i=1}^n p_i u_i(g_i) \quad (1)$$

υπό τους περιορισμούς κανονικοποίησης:

$$\begin{cases} \sum_{i=1}^n p_i = 1 \\ u_i(g_i^*) = 0, u_i(g_i^*) = 1, \forall i = 1, 2, \dots, n \end{cases} \quad (2)$$

όπου $u_i = 1, 2, \dots, n$ είναι αύξουσες συναρτήσεις των g_i που καλούνται περιθώριες ή μερικές συναρτήσεις αξίας (marginal value functions), κανονικοποιημένες στο διάστημα $[0, 1]$ και p_i τα βάρη των u_i . Και οι μερικές και οι ολικές συναρτήσεις αξίας έχουν την ιδιότητα της μονοτονίας του πραγματικού κριτηρίου. Για παράδειγμα, στην περίπτωση της συνάρτησης ολικής αξίας ισχύουν οι παρακάτω ιδιότητες :

$$\begin{cases} u[g(a)] > u[g(b)] \Leftrightarrow a > b (\text{προτίμηση}) \\ u[g(a)] = u[g(b)] \Leftrightarrow a \sim b (\text{αδιαφορία}) \end{cases} \quad (3)$$



Σχήμα 7: Η κανονικοποιημένη μερική συνάρτηση αξίας

Πηγή: Σίσκος Ι. (2008). ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ, Μεθοδολογία Επιχειρησιακής Έρευνας, Θεωρία Πολυκριτήριας Ανάλυσης, Εφαρμογές σε Επιχειρήσεις & Οργανισμούς, Εκδόσεις Νέων Τεχνολογιών, Αθήνα, σελ. 309.

Η μέθοδος UTA θεωρεί μια μορφή της προσθετικής συνάρτησης αξίας χωρίς βάρη, ισοδύναμη της μορφής (1) , (2) :

$$u(g) = \sum_{i=1}^n u_i(g_i) \quad (4)$$

υπό τους περιορισμούς κανονικοποίησης:

$$\begin{cases} \sum_{i=1}^n u_i(g_i) = 1 \\ u_i(g_i) = 0, \forall i = 1, 2, \dots, n \end{cases} \quad (5)$$

Η ύπαρξη ενός τέτοιου μοντέλου προϋποθέτει φυσικά τη προτιμησιακή ανεξαρτησία των κριτηρίων (preferential independence) για τον αποφασίζοντα. Χρησιμοποιώντας το προσθετικό μοντέλο (4) και (5) και λαμβάνοντας υπ' όψη τις σχέσεις προτίμησης (3), η αξία κάθε εναλλακτικής $a \in A_R$ μπορεί να γραφεί ως εξής:

$$u'[g(a)] = \sum_{i=1}^n u_i[g_i(a)] + \sigma(a), \forall a \in A_R \quad (6)$$

όπου $\sigma(a)$ είναι το ενδεχόμενο σφάλμα σε σχέση με το $u'[g(a)]$. Για την εκτίμηση των αντίστοιχων μερικών συναρτήσεων αξίας σε μια γραμμική κατά τμήματα μορφή, οι (Jacquet-Lagrange, E., & Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. European journal of operational research, 10(2), 151-164. προτείνουν τη χρήση της γραμμικής παρεμβολής. Έτσι για κάθε κριτήριο, το διάστημα $[g_{i*}, g_i^*]$ χωρίζεται σε $a_i - 1$ ίσα διαστήματα και τα τελικά σημεία g^j δίνονται από τη σχέση:

$$g_i^j = g_{i*} + \frac{j-1}{a_i-1} (g_i^* - g_{i*}), \forall j = 1, 2, \dots, a_i \quad (7)$$

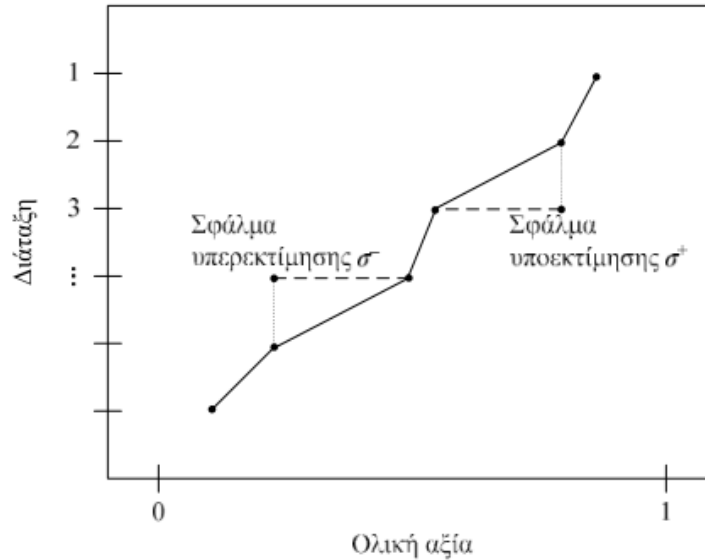
Η μερική αξία μιας εναλλακτικής a υπολογίζεται με χρήση γραμμικής παρεμβολής :

$$u_i[g_i(a)] = u_i(g_i^j) + \frac{g_i(a) - g_i^j}{g_i^{j+1} - g_i^j} [u_i(g_i^{j+1}) - u_i(g_i^j)], \text{ για } g_i(a) \in [g_i^j, g_i^{j+1}] \quad (8)$$

2.4.2 Η μέθοδος UTASTAR

Η μέθοδος UTASTAR προτάθηκε από τους (Siskos, Y., & Yannacopoulos, D. (1985). UTASTAR: An ordinal regression method for building additive value functions. Investigacao Operacional, 5(1), 39-53. και αποτελεί μια βελτιωμένη έκδοση της πρωτότυπης μεθόδου UTA. Στην αρχική έκδοση της μεθόδου UTA, για κάθε μία δράση $a \in A_R$ ορίζεται ένα μοναδικό σφάλμα $\sigma(a)$. Αυτή η συνάρτηση σφάλματος δεν είναι επαρκής για την ελαχιστοποίηση της ολικής διασποράς των σημείων στη μονότονη καμπύλη του Σχήμα 8. Το πρόβλημα αφορά τα σημεία που βρίσκονται δεξιά της καμπύλης, από τα οποία θα ήταν

προτιμότερο να αφαιρεθεί μια ποσότητα αξίας χωρίς να αυξηθούν οι αξίες των άλλων (παράδειγμα της ποιοτικής ή μονότονης παλινδρόμησης, ordinal regression paradigm, Σχήμα 8).



Σχήμα 8: Καμπύλη μονότονης παλινδρόμησης

Πηγή: Σίσκος Ι. (2008). ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ, Μεθοδολογία Επιχειρησιακής Έρευνας, Θεωρία Πολυκριτήριας Ανάλυσης, Εφαρμογές σε Επιχειρήσεις & Οργανισμούς, Εκδόσεις Νέων Τεχνολογιών, Αθήνα, σελ. 313.

Στη μέθοδο UTASTAR, οι Siskos & Yannacopoulos εισάγουν μια διπλή θετική συνάρτηση σφάλματος και έτσι ο τύπος (6) γίνεται:

$$u'[g(a)] = \sum_{i=1}^n u_i[g_i(a)] - \sigma^+(a) + \sigma^-(a), \forall a \in A_R \quad (9)$$

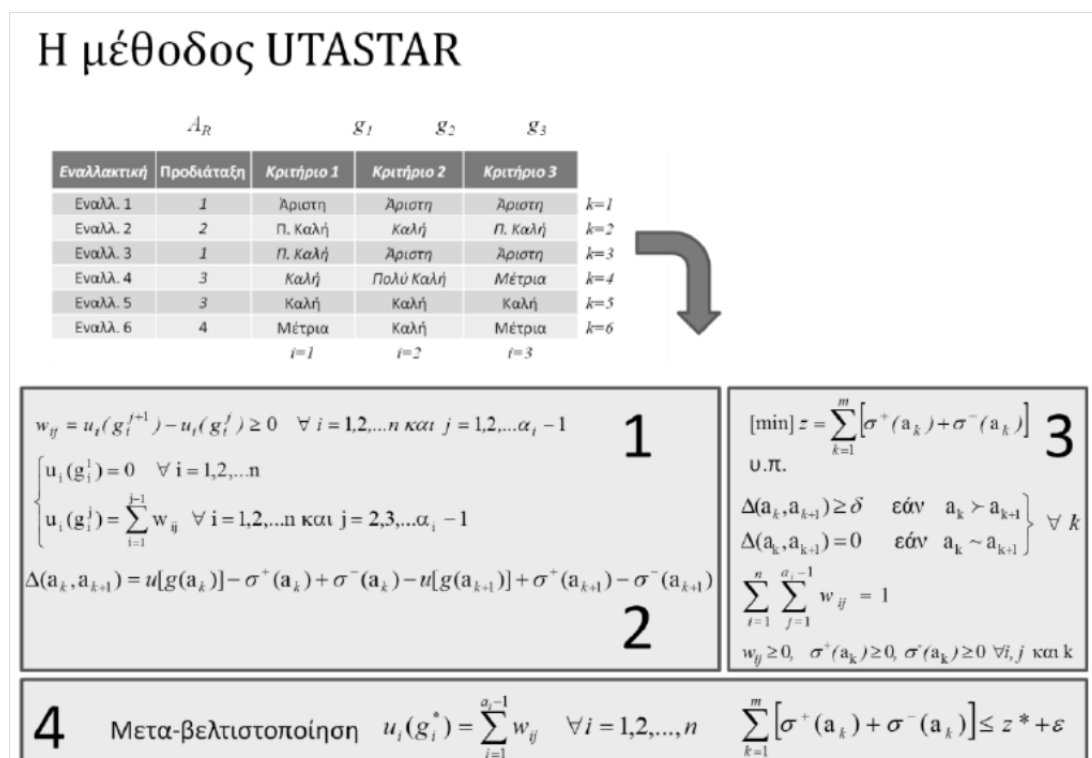
όπου όπου σ^+ και σ^- είναι τα σφάλματα υποεκτίμησης και υπερεκτίμησης, αντίστοιχα. Επιπρόσθετα, μια άλλη τροποποίηση αφορά τους περιορισμούς μονοτονίας των κριτηρίων, οι οποίοι μοντελοποιούνται με τη βοήθεια των ακόλουθων μετασχηματισμών των μεταβλητών:

$$w_{ij} = u_i(g_i^{j+1}) - u_i(g_i^j), \forall i = 1, 2, \dots, n \text{ και } j = 1, 2, \dots, a_i - 1 \quad (10)$$

Με αυτό τον τρόπο, οι συνθήκες μονοτονίας:

$$u_i(g_i^{j+1}) - u_i(g_i^j) \geq s_i \quad \forall i = 1, 2, \dots, n \text{ και } j = 1, 2, \dots, a_i - 1$$

μπορούν να αντικατασταθούν από περιορισμούς μη αρνητικότητας των μεταβλητών w_{ij} . Συνεπώς ο αλγόριθμος UTASTAR συνοψίζεται στα ακόλουθα βήματα:



Σχήμα 9 : Μέθοδος UTASTAR

Πηγή: Lakiotaki et al., (2011).

Βήμα 1: Η ολική αξία των δράσεων $u[g(a_k)]$, $k=1,2,\dots,m$, εκφράζεται αρχικά ως συνάρτηση των μερικών αξιών $u_i(g_i)$ και στη συνέχεια των μεταβλητών w_{ij} , σύμφωνα με την εξίσωση (10), μέσω των ακόλουθων σχέσεων:

$$\begin{cases} u_i(g_i^1) = 0, \forall i = 1, 2, \dots, n \\ u_i(g_i^j) = \sum_{l=1}^{j-1} w_{il}, \forall i = 1, 2, \dots, n \text{ και } j = 2, 3, \dots, \alpha_i - 1 \end{cases} \quad (12)$$

Βήμα 2: Εισάγονται δύο συναρτήσεις σφάλματος σ^+ και σ^- στο A_R , γράφοντας για κάθε ζεύγος διαδοχικών δράσεων στη προδιάταξη τις αναλυτικές εκφράσεις:

$$(\alpha_k, \alpha_{k+1}) = u[g(\alpha_k)] - \sigma^+(\alpha_k) + \sigma^-(\alpha_k) - u[g(\alpha_{k+1})] + \sigma^+(\alpha_{k+1}) - \sigma^-(\alpha_{k+1}) \quad (13)$$

Βήμα 3: Επιλύεται το ακόλουθο γραμμικό πρόγραμμα:

$$\left\{ \begin{array}{l} [min]z = \sum_{k=1}^m [\sigma^+(\alpha_k) + \sigma^-(\alpha_k)] \\ \text{υπό τους περιορισμούς :} \\ \Delta(\alpha_k, \alpha_{k+1}) \geq \delta, \alpha_k > \alpha_{k+1} \\ \Delta(\alpha_k, \alpha_{k+1}) = 0, \alpha_k \sim \alpha_{k+1} \end{array} \right\} \forall k \quad (14)$$

$$\sum_{i=1}^n \sum_{j=1}^{\alpha_i-1} w_{ij} = 1$$

$$w_{ij} \geq 0, \sigma^+(\alpha_k) \geq 0, \sigma^-(\alpha_k) \geq 0 \quad \forall i, j \text{ και } k$$

Βήμα 4: Ελέγχεται η ύπαρξη πολλαπλών βέλτιστων ή ημιβέλτιστων λύσεων στο γ.π. (11), υπολογίζοντας το βαρύκεντρο των προσθετικών συναρτήσεων αξίας που μεγιστοποιούν τις ακόλουθες αντικειμενικές συναρτήσεις :

$$u_i(g_i^*) = \sum_{j=1}^{\alpha_i-1} w_{ij} \quad \forall i = 1, 2, \dots, n \quad (15)$$

στο υπερπολύεδρο των περιορισμών του γ.π. (11) που περιορίζεται από τον επόμενο νέο περιορισμό:

$$\sum_{k=1}^m [\sigma^+(\alpha_k) + \sigma^-(\alpha_k)] \leq z^* + \varepsilon \quad (16)$$

όπου z^* είναι η βέλτιστη τιμή (σφάλμα) του γ.π. και ε είναι ένας πολύ μικρός θετικός αριθμός ή μηδέν.

Οι Siskos & Yannacopoulos απέδειξαν, σε ένα σύνολο πειραματικών δεδομένων, ότι η UTASTAR δίνει καλύτερα αποτελέσματα από το πρωτότυπο αλγόριθμο UTA.

Οι πολυκριτήριες μέθοδοι που αναλύθηκαν παραπάνω μπορούν να χρησιμοποιηθούν για τη δημιουργία προφίλ χρηστών μέσω της μοντελοποίησης των προτιμήσεών τους. Για την αναπαράσταση των προτιμήσεων των χρηστών μπορούν να χρησιμοποιηθούν π.χ. οι συναρτήσεις αξίας που προκύπτουν απ' την εφαρμογή των αλγορίθμων για κάθε χρήστη ή τα βάρη των κριτηρίων.

2.5 Μηχανική μάθηση

Η μηχανική μάθηση είναι μια τεχνική ανάλυσης δεδομένων που διδάσκει τους υπολογιστές να κάνουν αυτό που έρχεται φυσικά στους ανθρώπους: να μαθαίνουν από την εμπειρία χωρίς να προγραμματίζονται ρητά. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούν υπολογιστικές μεθόδους για να

"μαθαίνουν" τις πληροφορίες απευθείας από τα δεδομένα χωρίς να στηρίζονται σε μια προκαθορισμένη εξίσωση ως μοντέλο. Οι αλγόριθμοι βελτιώνουν προσαρμοστικά την απόδοσή τους καθώς αυξάνεται ο αριθμός των διαθέσιμων δειγμάτων για μάθηση [29].

Εποπτευόμενη Μηχανική Μάθηση

Η εποπτευόμενη μηχανική μάθηση κατασκευάζει ένα μοντέλο που κάνει προβλέψεις βασισμένες σε αποδεικτικά στοιχεία παρουσία αβεβαιότητας. Ένας αλγόριθμος εποπτευόμενος εκμάθησης παίρνει ένα γνωστό σύνολο δεδομένων εισόδου και γνωστών απαντήσεων στα δεδομένα (έξοδος) και εκπαιδεύει ένα μοντέλο για να παράγει λογικές προβλέψεις για την απάντηση σε νέα δεδομένα. Η εποπτευόμενη μάθηση χρησιμοποιεί τεχνικές ταξινόμησης και παλινδρόμησης για την ανάπτυξη προγνωστικών μοντέλων.

Οι τεχνικές ταξινόμησης προβλέπουν διακριτές απαντήσεις - για παράδειγμα, εάν ένα μήνυμα ηλεκτρονικού ταχυδρομείου είναι γνήσιο ή ανεπιθύμητο, ή αν ένας όγκος είναι καρκινικός ή καλοήθεις. Τα μοντέλα ταξινόμησης ταξινομούν τα δεδομένα εισόδου σε κατηγορίες. Τυπικές εφαρμογές είναι ιατρική απεικόνιση και η αναγνώριση ομιλίας. Κοινή αλγόριθμοι για την εκτέλεση ταξινόμησης είναι τα ενισχυμένα δέντρα απόφασης, *k*-nearest neighbor, Naïve Bayes, διακριτική ανάλυση, λογιστική παλινδρόμηση, και νευρωνικά δίκτυα.

Μη εποπτευόμενη Μηχανική μάθηση

Η μη εποπτευόμενη μάθηση βρίσκει κρυμμένα μοτίβα ή εγγενείς δομές στα δεδομένα. Χρησιμοποιείται για την εξαγωγή συμπερασμάτων από σύνολα δεδομένων που αποτελούνται από δεδομένα εισόδου χωρίς επισημασμένες απαντήσεις.

Η συσταδοποίηση είναι η πιο κοινή τεχνική μάθησης χωρίς επίβλεψη. Χρησιμοποιείται για διερευνητική ανάλυση δεδομένων για την εύρεση κρυφών μοτίβων ή ομαδοποιήσεων σε δεδομένα. Οι εφαρμογές για την ανάλυση συμπλέγματος περιλαμβάνουν ανάλυση αλληλουχίας γονιδίων, έρευνα αγοράς και αναγνώριση αντικειμένων.

Για παράδειγμα, εάν μια εταιρεία κινητής τηλεφωνίας θέλει να βελτιστοποιήσει τις τοποθεσίες όπου τοποθετεί κεραίες κινητών τηλεφώνων, μπορεί να χρησιμοποιήσει τη μηχανική μάθηση για να εκτιμήσει τον αριθμό των συστάδων των ατόμων που βασίζονται στις κεραίες τους. Ένα τηλέφωνο

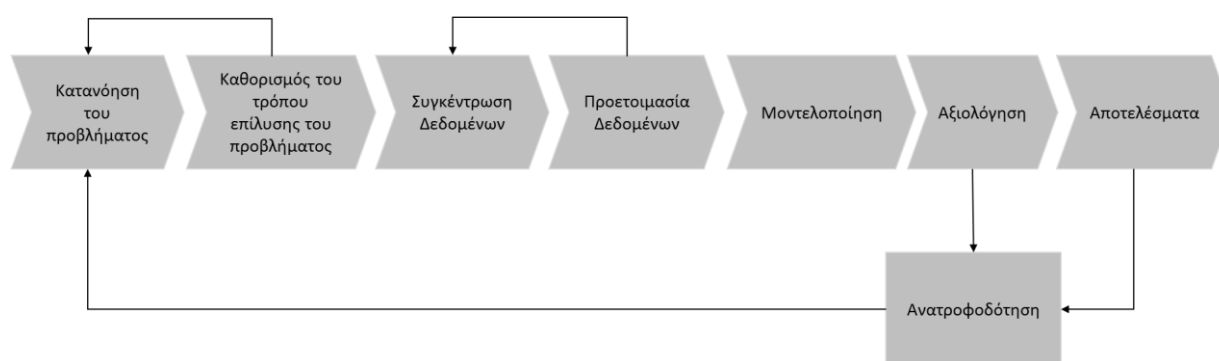
μπορεί να μιλήσει μόνο σε μια κεραία κάθε φορά, οπότε η εταιρία χρησιμοποιεί αλγόριθμους συσταδοποίησης για να σχεδιάσει την καλύτερη τοποθέτηση κεραιών για να βελτιστοποιήσει τη λήψη σήματος για τους πελάτες της.

Οι πιο συνηθισμένοι αλγόριθμοι συσταδοποίησης είναι οι: k-means and k-medoids, hierarchical clustering, Gaussian mixture models, hidden Markov models, self-organizing maps, fuzzy c-means clustering and subtractive clustering.

2.6 Εξόρυξη Δεδομένων

Η εξόρυξη δεδομένων είναι η διαδικασία ανάλυσης κρυφών προτύπων δεδομένων σύμφωνα με διαφορετικές προοπτικές κατηγοριοποίησης σε χρήσιμες πληροφορίες, οι οποίες συλλέγονται και συναρμολογούνται σε κοινόχρηστους χώρους, όπως αποθήκες δεδομένων, για αποτελεσματική ανάλυση, αλγόριθμους εξόρυξης δεδομένων, διευκόλυνση της λήψης επιχειρηματικών αποφάσεων και άλλες πληροφορίες απαιτήσεις για την τελική μείωση των δαπανών και την αύξηση των εσόδων.

Η ανεύρεση γνώσης είναι μια επαναληπτική διαδικασία που αποτελείται από μια σειρά βημάτων (Σχήμα 10), τα οποία οδηγούν από τη συλλογή των δεδομένων στην ανακάλυψη και εξαγωγή χρήσιμης πληροφορίας από αυτά [30]. Τα βήματα από τα οποία αποτελείται η διαδικασία ανεύρεσης γνώσης είναι τα ακόλουθα:



Σχήμα 10 : Βήματα εξόρυξης γνώσης

1) Κατανόηση προβλήματος

Ίσως το σημαντικότερο μέρος της διαδικασίας αφού πρέπει να κατανοηθεί το πρόβλημα. Στο στάδιο αυτό συμμετέχουν όσα μέρη εμπλέκονται στο μέρος του προβλήματος

2) Καθορισμός επίλυσης προβλήματος

Στο βήμα αυτό επιλέγεται η στρατηγική που θα ακολουθηθεί για την επίλυση του προβλήματος είναι άμεσα συνδεδεμένο με το πρώτο βήμα.

3) Συγκέντρωση δεδομένων

Εδώ τα δεδομένα συγκεντρώνονται από διάφορες πηγές, πολλές φορές ανομοιογενή μεταξύ τους και ενσωματώνονται σε μια κοινή βάση δεδομένων.

4) Προετοιμασία δεδομένων

Στο βήμα αυτό, αφαιρούνται από τη βάση δεδομένων αυτά τα δεδομένα που παράγουν θόρυβο, δηλαδή όλα εκείνα τα στοιχεία που μπορούν να επηρεάσουν ή και να διαστρεβλώσουν το αποτέλεσμα. Επίσης από όλα εκείνα τα δεδομένα που έχουμε στη διάθεση μας, επιλέγονται προσεκτικά εκείνα που είναι σχετικά και χρήσιμα για την ανάλυση που θα ακολουθήσει. Το στάδιο αυτό περιλαμβάνει και την τροποποίηση δεδομένων έτσι ώστε η μορφή τους να είναι κατάλληλη για την διαδικασία της εξόρυξης.

5) Μοντελοποίηση

Είναι το σημαντικότερο από τα βήματα της διαδικασίας και αυτό γιατί στο συγκεκριμένο στάδιο, ποικίλες εξελιγμένες τεχνικές χρησιμοποιούνται για την εξαγωγή δυνητικά χρήσιμων προτύπων.

6) Αξιολόγηση

Στο προτελευταίο βήμα αξιολογούνται τα αποτελέσματα με βάσει κάποιους δείκτες αξιολόγησης που έχουν αποφασιστεί στο πρώτο βήμα της διαδικασίας.

7) Αποτελέσματα

Στο τελικό αυτό στάδιο, η γνώση που έχει ανακαλυφθεί παρουσιάζεται στον χρήστη, βοηθώντας τον έτσι να κατανοήσει και να ερμηνεύσει τα αποτελέσματα της εξόρυξης δεδομένων.

8) Ανατροφοδότηση

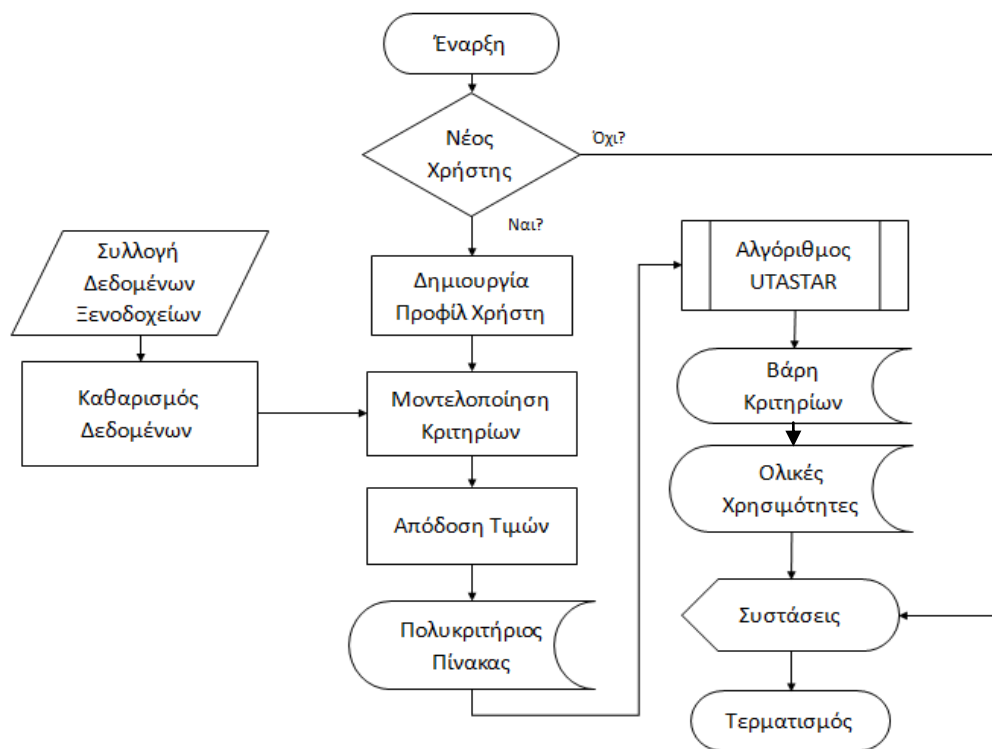
Ο χρήστης, εκμεταλλευόμενος την επαναληπτική μορφή της διαδικασίας ανεύρεσης γνώσης, έχει την δυνατότητα να τροποποιήσει τα μέτρα αξιολόγησης, να τελειοποιήσει την διαδικασία της εξόρυξης, να επιλέξει νέα δεδομένα, να τροποποιήσει περαιτέρω τα ήδη υπάρχοντα ή να ενσωματώσει στη βάση νέα από καινούργιες πηγές, με τελικό στόχο την εξαγωγή διαφορετικών και ακόμη πιο κατάλληλων αποτελεσμάτων.

ΚΕΦΑΛΑΙΟ 3: ΠΡΟΤΕΙΝΟΜΕΝΗ ΜΕΘΟΔΟΛΟΓΙΑ

3.1 Εισαγωγή

Θα προσεγγίσουμε το πρόβλημα εφαρμόζοντας τη μεθοδολογική προσέγγιση του Agent Allocator³¹ (Matsatsinis, N. F., & Delias, P. (2003, September). AgentAllocator: An agent-based multi-criteria decision support system for task allocation. In International Conference on Industrial Applications of Holonic and Multi-Agent Systems (pp. 225-235). Springer, Berlin, Heidelberg. μια πολυκριτήρια μέθοδο η οποία υλοποιείται για την υποστήριξη της απόφασης στην κατανομή εργασιών. Αντιμετωπίζουμε το πρόβλημα με αντίστοιχο τρόπο, όπου οι χρήστες θα είναι οι πράκτορες και τα ξενοδοχεία θα είναι οι εργασίες. Ως κριτήρια χρησιμοποιήσαμε ορισμένα από τα κριτήρια που χρησιμοποιεί η ιστοσελίδα booking.com καθώς και κάποια που δημιουργήσαμε εμείς.

Στο Σχήμα 11 παρουσιάζεται η αρχιτεκτονική του μεθοδολογικού πλαισίου του συστήματός μας, αναλυτικά. Στην πρώτη φάση έγινε η συλλογή και ο καθαρισμός των δεδομένων μας. Ξεκινώντας η διαδικασία ελέγχεται αν ο χρήστης είναι καινούργιος ή όχι. Εφόσον είναι νέος χρήστης, δημιουργείται το προφίλ του, και στη συνέχεια προχωράμε στη μοντελοποίηση των κριτηρίων. Έπειτα αποδίδονται τιμές στα κριτήρια σύμφωνα με την τη μεθοδολογική προσέγγιση που αναφέραμε στην αρχή της προηγούμενης παραγράφου και καταλήγουμε στον πολυκριτήριο πίνακα. Στη συνέχεια τρέχουμε τον αλγόριθμο της Utastar ο οποίος μας δίνει ως αποτέλεσμα τα βάρη των κριτηρίων και προχωράμε στον υπολογισμό των ολικών χρησιμότητων. Τέλος μετά τον υπολογισμό των ολικών χρησιμότητων καταλήγουμε στις τελικές συστάσεις για τον χρήστη. Στη συνέχεια του κεφαλαίου αναπτύσσονται όλα τα βήματα αναλυτικά.



Σχήμα 11 : Αρχιτεκτονική Μεθοδολογικού Πλαισίου

3.2 Συλλογή Δεδομένων

Προκειμένου να δημιουργήσουμε συστάσεις για κάθε χρήστη, είναι απαραίτητο να έχουμε αρκετή πληροφόρηση τόσο για τους χρήστες όσο και για τις εναλλακτικές (ξενοδοχεία). Χρησιμοποιήσαμε δεδομένα από την ιστοσελίδα booking.com για την συλλογή πληροφοριών σχετικά με τα χαρακτηριστικά και τις παροχές των ξενοδοχείων (Πίνακας 3 : Δεδομένα Ξενοδοχεί). Με τη βοήθεια κώδικα συλλέξαμε δεδομένα σε μορφή xlsx για 319 ξενοδοχειακές μονάδες (γραμμές) στην πόλη των Χανίων. Για κάθε ξενοδοχείο έχουμε 48 στήλες με πληροφορίες σχετικά με τα χαρακτηριστικά τους.

Στη συνέχεια, με την δημιουργία ερωτηματολογίου (βλέπε Παράρτημα), συλλέξαμε πληροφορίες από ένα δείγμα 130 χρηστών σχετικά με τις προτιμήσεις τους πάνω σε χαρακτηριστικά των ξενοδοχειακών μονάδων που τους επηρεάζουν για την επιλογή ξενοδοχείου, προκειμένου να δημιουργήσουμε προφίλ χρηστών. Παρακάτω στους Πίνακας 1 και Πίνακας 2 παρουσιάζονται τα δεδομένα που αναφέραμε παραπάνω για τους χρήστες.

Πίνακας 1 : Δεδομένα Χρηστών από Ερωτηματολόγιο

Φύλλο	Ηλικία	Σκοπός Ταξιδιού	Προτίμηση απόστασης ξενοδοχείου από κέντρο	Προτίμηση απόστασης ξενοδοχείου από Μέσα	Επιθυμητό Κόστος	WiFi
Γυναίκα	20	Διακοπές	400-900	0-150	Από 76 έως 100	Δωρεάν
Αντρας	19	Διακοπές	0-400	0-150	Από 101 έως 130	Δωρεάν
Γυναίκα	36	Διακοπές	400-900	150-350	Από 101 έως 130	Δωρεάν
Αντρας	36	Επαγγελματικό	0-400	0-150	έως 50	Δωρεάν
Αντρας	25	Διακοπές	0-400	0-150	έως 50	Δωρεάν

Πίνακας 2 : Δεδομένα Χρηστών από Ερωτηματολόγιο

Air Condition (Βαθμός Προτίμησης)	Τηλέοραση (Βαθμός Προτίμησης)	Πάρκινγκ	Πρωινό Γεύμα	Ψυγείο/Καφετιέρα (Βαθμός Προτίμησης)	Ποιότητα Καθαρισμού (Βαθμός Προτίμησης)	Ποιότητα Ύπνου (Βαθμός Προτίμησης)
9	6	Όχι	Στην Τιμή	9	10	10
8	5	Δωρεάν	Όχι	7	10	9
10	5	Δωρεάν	Στην Τιμή	9	10	10
10	10	Όχι	Στην Τιμή	3	9	10
9	7	Δωρεάν	Με Πληρωμή	9	10	10

Όπως μπορούμε να δούμε παραπάνω οι πληροφορίες που συλλέξαμε για τους χρήστες έχουν να κάνουν με τις προτιμήσεις τους σε διάφορες παροχές που προσφέρουν τα ξενοδοχεία, αν για παράδειγμα προσφέρονται δωρεάν ή όχι, καθώς και σε τι βαθμό επηρεάζονται από την ύπαρξη συγκεκριμένων παροχών και υπηρεσιών. Επίσης, τα χρήματα που είναι διατεθειμένοι να ξοδέψουν και την επιθυμητή απόσταση του ξενοδοχείου σε σχέση με κάποιες συγκεκριμένες τοποθεσίες. Στη συνέχεια της μεθοδολογίας αναλύεται πως αξιοποιούνται τα δεδομένα αυτά.

3.2.1 Επεξεργασία δεδομένων

Στα δεδομένα που συλλέχθηκαν παρατηρήσαμε ότι μέρος απαντήσεων εμπεριείχαν κενές τιμές (No values) οι οποίες θα επηρέαζαν τον αλγόριθμό UTAStar και τα τελικά αποτελέσματα. Μετά τον καθαρισμό συγκεντρώσαμε 104 περιπτώσεις ξενοδοχείων στην πόλη των Χανίων.

Επίσης από τις αξιολογήσεις των ξενοδοχείων που είχαμε στη διάθεση μας, επιλέξαμε 19 στήλες – κατηγορίες δεδομένων από τις 48, καθώς ήταν τα μόνα χρήσιμα για την ανάλυση που θα ακολουθήσει στη μεθοδολογία μας. Όπως βλέπουμε και στους Πίνακες 3 και Πίνακας 4, τα δεδομένα μας

περιέχουν πληροφορίες για τη θέση και την τοποθεσία των ξενοδοχείων, διάφορες παροχές και ανέσεις που προσφέρει κάθε μονάδα, ορισμένες αξιολογήσεις κριτηρίων καθώς και την τιμή/κόστος κάθε ξενοδοχείου. Ως αναφορά χρησιμοποιήσαμε τα δίκλινα δωμάτια, όπου αυτό ήταν εφικτό, για να κινούμαστε στο ίδιο μήκος κύματος για όλες τις εναλλακτικές. Επίσης, ως απόσταση από Μέσα Μαζικής Μεταφοράς πήραμε την κοντινότερη στάση στο εκάστοτε ξενοδοχείο.

Πίνακας 3 : Δεδομένα Ξενοδοχείων από Booking

Name	Απόσταση από Μέσα Μαζικής Μεταφοράς(km)	Απόσταση από Κέντρο(km)	Κόστος(€)	Διαστάσεις δωματίου(m ²)	Staff	Total
Guest house Casa Kasteli	0,395	0,3292	70	16	9,3	8,8
Hotel Samaria	0,050	0,3635	139	20	9,6	9,4
Hotel Kriti	0,293	0,4444	84	15	9,4	9
Hotel Porto Veneziano	0,578	0,5474	108	15	9,6	9,5
Hotel Kydon	0,050	0,3765	113	18	9,7	9,4

Πίνακας 4 : Δεδομένα Ξενοδοχείων από Booking

Name	Wi-Fi	Χώρος Στάθμευσης	Θέρμανση/Κλιματισμός	Τηλεόραση	Εστιατόριο/Σνακ Μπαρ	Πρωινό
Guest house Casa Kasteli	Δωρεάν	Πληρωμή	Διαθέτει	Διαθέτει	Δεν Διαθέτει	Including Breakfast
Hotel Samaria	Δωρεάν	Διαθέτει	Διαθέτει	Διαθέτει	Διαθέτει	Including Breakfast
Hotel Kriti	Δωρεάν	Διαθέτει	Διαθέτει	Διαθέτει	Διαθέτει	Including Breakfast
Hotel Porto Veneziano	Δωρεάν	Πληρωμή	Διαθέτει	Διαθέτει	Διαθέτει	Including Breakfast
Hotel Kydon	Δωρεάν	Πληρωμή	Διαθέτει	Διαθέτει	Διαθέτει	Optional Breakfast

Στο υποκεφάλαιο 2.2 Δειγματοληψία στη διαδικασία της δειγματοληψίας. Τον πληθυσμό της έρευνας αποτέλεσε το σύνολο των ξενοδοχείων της πόλης των Χανίων. Για την παρούσα έρευνα επιλέχθηκε ως δείγμα (ξενοδοχείων) από το σύνολο του εξεταζόμενου πληθυσμού, ένα δείγμα ξενοδοχείων από την πλατφόρμα τουριστικών καταλυμάτων booking.com. Παρακάτω παρουσιάζονται τα στοιχεία που αφορούν τα δεδομένα στα οποία εφαρμόζουμε τη μεθοδολογία μας.

Στην δική μας περίπτωση ο πληθυσμός των ξενοδοχείων στην πόλη των Χανίων είναι 802 ξενοδοχεία, σύμφωνα με την ιστοσελίδα booking.com. Το μέγεθος του δείγματός μας ύστερα και από τον καθαρισμό των δεδομένων είναι 104 ξενοδοχεία. Με επίπεδο σημαντικότητας 95% έχουμε ως αποτέλεσμα το σφάλμα της εκτίμησης να είναι $\pm 8,98\%$.

Πίνακας 5 : Πίνακας Δειγματοληψίας

Confidence Level:	95%
Sample Size:	104
Population Proportion:	50%
Population Size	802
Margin of Error:	$\pm 8,98\%$.

Σχετικά με το δείγμα των χρηστών που συλλέξαμε από τη δημιουργία ερωτηματολογίου, ως πληθυσμό της έρευνας θεωρήσαμε τον αριθμό των επισκεπτών στο αεροδρόμιο των Χανίων για το χρονικό διάστημα από τον Ιανουάριο του 2019 έως τον Οκτώβριο του 2019, σύμφωνα με την Fraport Greece. Ο πληθυσμός ήταν 2.865.765 επισκέπτες και το δείγμα μας μετά των καθαρισμό των δεδομένων ήταν 130 χρήστες. Συνεπώς με επίπεδο σημαντικότητας 95% το σφάλμα εκτίμησης είναι $\pm 8,59\%$.

Πίνακας 6 : Πίνακας Δειγματοληψίας

Confidence Level:	95%
Sample Size:	130
Population Proportion:	50%
Population Size	2.865.765
Margin of Error:	$\pm 8,59\%$.

Σε αυτό το σημείο παρουσιάζονται τα στατιστικά μέτρα των δεδομένων μας για τα οποία αναφερθήκαμε στο υποκεφάλαιο 2.3 Στατιστικά Μέτρα με σκοπό την εξαγωγή συμπερασμάτων.

Στους παρακάτω πίνακες παρουσιάζονται τα στατιστικά μέτρα για το σύνολο του δείγματος των χρηστών όπως αυτά προέκυψαν χρησιμοποιώντας τη συνάρτηση `describe()` της προγραμματιστικής γλώσσας Python. Παρατηρώντας τους πίνακες χρηστών (Πίνακας 1 : Δεδομένα Χρηστών από Ερωτηματολόγιο Πίνακας 1, Πίνακας 2) βλέπουμε ότι ορισμένα δεδομένα είναι ποιοτικά και ορισμένα ποσοτικά. Η παραπάνω συνάρτηση μας δίνει τα στατιστικά μέτρα μόνο για ποσοτικά δεδομένα, όποτε αναγκαστήκαμε να μετατρέψουμε τα ποιοτικά δεδομένα σε ποσοτικά. Η κωδικοποίηση έγινε με τη χρήση της συνάρτησης `LabelEncoder()`.

Πίνακας 7 : Κωδικοποίηση Ποιοτικών Δεδομένων για τον Υπολογισμό Στατιστικών Μέτρων

Απόσταση από Σημεία Ενδιαφέροντος(m)	Κωδικοποίηση	Απόσταση από Μέσα Μεταφοράς(m)	Κωδικοποίηση	Πρόσβαση Τντερνετ (WiFi)	Κωδικοποίηση
0-400	0	0-150	0	Όχι	0
400-900	1	150-350	1	Δωρεάν	1
900-πάνω	2	350-πάνω	2		
Τιμή Δωματίου(€)	Κωδικοποίηση	Πρωινό Γεύμα	Κωδικοποίηση	Πάρκινγκ	Κωδικοποίηση
έως 50	0	Όχι	0	Όχι	0
Από 51 έως 75	1	Με πληρωμή	1	Δωρεάν	1
Από 76 έως 100	2	Στην Τιμή	2		
Από 101 έως 130	3				

Σύμφωνα με την κωδικοποίηση οι κλίμακες κάθε περίπτωσης παίρνουν τις τιμές 0,1,2,3 αντίστοιχα. Για παράδειγμα αν ο μέσος όρος της απόστασης από τα σημεία ενδιαφέροντος είναι ίσος με ένα αυτό σημαίνει ότι οι περισσότεροι χρήστες προτιμούν να απέχουν από 400 έως 900 μέτρα.

Πίνακας 8: Στατιστικά Μέτρα για δεδομένα χρηστών

	Απόσταση από Σημεία Ενδιαφέροντος	Απόσταση από Μέσα Μεταφοράς	Τιμή Δωματίου	Πρόσβαση Τντερνετ (WiFi)	Πρωινό Γεύμα	Πάρκινγκ
count	130	130	130	130	130	130
mean	0,676923	0,438462	1,130769	0,969231	1,807692	0,615385
std	0,625249	0,584076	1,014534	0,17336	0,514837	0,504009
min	0	0	0	0	0	0
25%	0	0	0	1	2	0
50%	1	0	1	1	2	1
75%	1	1	2	1	2	1
max	2	2	4	1	2	2

Πίνακας 9: Στατιστικά Μέτρα για δεδομένα χρηστών

	Air Condition	Τηλεόραση	Ψυγείο/ Καφετιέρα	Ποιότητα Καθαρισμού	Ποιότητα Υπνου	Χρηματοκιβώτιο
count	130	130	130	130	130	130
mean	7,90000	5,41539	6,39231	9,50769	9,39231	3,66154
std	2,04162	3,14207	2,72939	1,09436	1,11714	2,77462
min	1	1	1	3	4	1
25%	7	2,25	5	9,25	9	1
50%	8	5	7	10	10	3
75%	10	8	8	10	10	6
max	10	10	10	10	10	10

Στους πίνακες 8 και 9 παρουσιάζονται τα στατιστικά μέτρα των δεδομένων των χρηστών για το σύνολο του δείγματος. Συγκριμένα στον Πίνακα 8 παρατηρούμε ότι οι χρήστες προτιμούν να διαμένουν όσο πιο κοντά γίνεται στα μέσα μεταφοράς καθώς επίσης η συντριπτική πλειοψηφία προτιμά να παρέχεται δωρεάν Wifi και το πρωινό να περιλαμβάνεται στην τιμή. Αξιοσημείωτο είναι ότι οι περισσότεροι χρήστες είναι διατεθειμένοι να ξοδέψουν από 51 έως 75 ευρώ για μια διανυκτέρευση.

3.3 Μοντελοποίηση Σύστασης

Ύστερα από ενδελεχή έρευνα στις δημοφιλέστερες πλατφόρμες τουριστικών καταλυμάτων (booking.com, airbnb, Expedia.com, Tripadvisor), σχετικά με τις σημαντικότερες προτιμήσεις των πελατών και τα χαρακτηριστικά των καταλυμάτων, καταλήξαμε στις κυριότερες απαιτήσεις των πελατών καθώς και στα σημαντικότερα χαρακτηριστικά των ξενοδοχειακών μονάδων.

Οι παρακάτω απαιτήσεις των χρηστών καθώς και τα χαρακτηριστικά των μονάδων έχουν χωριστεί σε διαστήματα/περιπτώσεις με σκοπό να κατηγοριοποιηθούν.

Από τα πιο σημαντικά χαρακτηριστικά των ξενοδοχείων είναι η τοποθεσία και κατ' επέκταση η απόστασή τους από διάφορα σημεία ενδιαφέροντος. Επίσης σημαντικό ρόλο στην επιλογή ξενοδοχείου παίζουν οι παροχές και οι ανέσεις που διαθέτει κάθε μονάδα. Τέλος, το κόστος της διαμονής επηρεάζει σε μεγάλο βαθμό τους χρήστες στην επιλογή ξενοδοχείου.

Πίνακας 10 : Χαρακτηριστικά Ξενοδοχείων – Απαιτήσεις Χρηστών

Απαιτήσεις Χρηστών	Χαρακτηριστικά Μονάδας
<ul style="list-style-type: none"> • Απόσταση (m) από σημεία ενδιαφέροντος [0-400, 400-900, 900-πάνω] • Απόσταση (m) από Μέσα Μαζικής Μεταφοράς [0-150, 150-350, 350-πάνω] • Πρόσβαση Ιντερνέτ(WiFi) [Δωρεάν, Πληρωμή, ΟΧΙ] • Air Condition [ΝΑΙ, ΟΧΙ] • Τηλεόραση [ΝΑΙ, ΟΧΙ] • Παρκινγκ [Δωρεάν, Πληρωμή, ΟΧΙ] • Ποιότητα Ύπνου [Πολύ καλή, Καλή, Μέτρια, Κακή, Πολύ Κακή] • Ψυγείο/Καφετιέρα [ΝΑΙ, ΟΧΙ] • Πρωινό [Περιλαμβάνεται στη Τιμή, Με πληρωμή] • Τιμή/Κόστος [Πολύ Ακριβό, Ακριβό, Όχι πολύ Ακριβό, Όχι πολύ Φθηνό, Φθηνό] • Ποιότητα Καθαρισμού [Υψηλή, Μέτρια, Κακή] 	<ul style="list-style-type: none"> • Θέση ξενοδοχείου [Καλή, Μέτρια, Κακή] • Wi-Fi [Δωρεάν, Πληρωμή, Δεν Διαθέτει] • Χώρος Στάθμευσης [Διαθέτει, Δεν Διαθέτει] • Θέρμανση /Κλιματισμός [Διαθέτει, Δεν Διαθέτει] • Τηλεόραση [Διαθέτει, Δεν Διαθέτει] • Σνακ Μπαρ /Εστιατόριο [Διαθέτει, Δεν Διαθέτει] • Υπηρεσία καθαριότητας [Καθημερινή, Όχι Καθημερινή] • Μεγάλο κρεβάτι [Διαθέτει, Δεν Διαθέτει] • Κόστος [Ακριβό, Όχι πολύ ακριβό, Φθηνό] • Πρωινό [Δωρεάν, Πληρωμή, Δεν Διαθέτει] • Καφετιέρα/Καφές [Διαθέτει, Δεν Διαθέτει] • Safety Deposit Box [Διαθέτει, Δεν Διαθέτει] • Βεράντα/Θέα [Διαθέτει, Δεν Διαθέτει] • Ψυγείο [Διαθέτει, Δεν Διαθέτει] • Διαστάσεις δωματίου (m²) [Πολύ Μεγάλο, Μεγάλο, Μικρό]

Μοντελοποίηση Κριτηρίων

Στόχος μας είναι να δημιουργήσουμε την καλύτερη σύσταση για κάθε χρήστη. Για αυτό επιλέξαμε οχτώ κριτήρια, τα οποία χρησιμοποιούνται για την εκτίμηση του ζεύγους Ξενοδοχείο - Χρήστης. Τα κριτήρια αυτά είναι η *τοποθεσία* του ξενοδοχείου, η *καθαριότητα* του, οι *παροχές* του ξενοδοχείου και του *δωματίου*, η *σχέση της ποιότητας* του ξενοδοχείου με την *τιμή* του, η *άνεση* του, το *προσωπικό* (εργαζόμενοι) της μονάδας και η γενική *Εικόνα* του ξενοδοχείου από τους πελάτες. Κάθε χρήστης συνδυάζεται με το σύνολο των υπό μελέτη ξενοδοχείων με στόχο την επιλογή του βέλτιστου ζεύγους που θα αποτελέσει την τελική σύσταση του συστήματος.

Η μοντελοποίηση αυτών των κριτηρίων υλοποιείται παρακάτω στον

Τοποθεσία	Καθαριότητα	Παροχές δωματίου	Παροχές ξενοδοχείου	Σχέση ποιότητας/ Τιμής	Ανεση	Προσωπικό	Εικόνα Μονάδας
Απόσταση από σημεία ενδιαφέροντος - Θέση ξενοδοχείου	Ποιότητα Καθαρισμού - Υπηρεσία καθαριότητας	Πρόσβαση Ιντερνέτ(WiFi) -WiFi(Δωρεάν/ Πληρωμή)	Παρκινγκ- Χώρος Στάθμευσης Σνακ Μπαρ /Εστιατόριο	Τιμή/Κόστος- Κόστος	Ποιότητα ύπνου- Μεγάλο κρεβάτι	(Staff Rating από Αξιολογήσεις ξενοδοχείου)	(Total Rating από Αξιολογήσεις ξενοδοχείου)
Απόσταση από Μέσα Μαζικής Μεταφοράς - Θέση ξενοδοχείου	Υπηρεσία καθαριότητας	Τηλεόραση- Τηλεόραση Air Condition- Θέρμανση/ Air Condition Ψυγείο/Καφετι	Πρωινό - Πρωινό(Δωρεάν/ Πληρωμή) Ψυγείο/Καφετιέρα - Καφετιέρα/Καφές		Διαστάσεις δωματίου Θέρμανση /Κλιματισμός Βεράντα Θέα		

Όπως φαίνεται στον πίνακα τα στοιχεία των υποκριτηρίων μπορεί να είναι είτε απαιτήσεις χρηστών είτε χαρακτηριστικά των ξενοδοχείων ή και τα δύο. Για τα κριτήρια **Προσωπικό** και **Εικόνα Μονάδας** θα χρησιμοποιήσουμε τις τιμές που συλλέξαμε από την ιστοσελίδα booking.com. Τα κριτήρια αυτά θα είναι οι στήλες στον πολυκριτήριο πίνακα.

Πίνακας 11 : Μοντελοποίηση κριτηρίων

Τοποθεσία	Καθαριότητα	Παροχές δωματίου	Παροχές ξενοδοχείου	Σχέση ποιότητας/ Τιμής	Ανεση	Προσωπικό	Εικόνα Μονάδας
Απόσταση από σημεία ενδιαφέροντος - Θέση ξενοδοχείου	Ποιότητα Καθαρισμού - Υπηρεσία καθαριότητας	Πρόσβαση Ιντερνέτ(WiFi) -WiFi(Δωρεάν/ Πληρωμή)	Παρκινγκ- Χώρος Στάθμευσης Σνακ Μπαρ /Εστιατόριο	Τιμή/Κόστος- Κόστος	Ποιότητα ύπνου- Μεγάλο κρεβάτι	(Staff Rating από Αξιολογήσεις ξενοδοχείου)	(Total Rating από Αξιολογήσεις ξενοδοχείου)
Απόσταση από Μέσα Μαζικής Μεταφοράς - Θέση ξενοδοχείου	Υπηρεσία καθαριότητας	Τηλεόραση- Τηλεόραση Air Condition- Θέρμανση/ Air Condition Ψυγείο/Καφετι	Πρωινό - Πρωινό(Δωρεάν/ Πληρωμή) Ψυγείο/Καφετιέρα - Καφετιέρα/Καφές		Διαστάσεις δωματίου Θέρμανση /Κλιματισμός Βεράντα Θέα		

3.4 Απόδοση τιμών

Ο αποφασίζων καλείται να αποδώσει τιμές στα κριτήρια ανάλογα με τις προσωπικές του προτιμήσεις. Για κάθε επιμέρους κριτήριο δίνεται μια τιμή από 0 έως 1 σε κάθε πιθανό συνδυασμό που υπάρχει.

Προκειμένου να υπολογίσουμε τον πολυκριτήριο πίνακα πρέπει πρώτα να υπολογίσουμε κάθε κριτήριο ξεχωριστά συνεπώς και κάθε υποκριτήριο από τα οποία αποτελείται το καθένα. Για να τα υπολογίσουμε χρησιμοποιούμε την παρακάτω διαδικασία :

- Σύμφωνα με τα συστατικά κάθε υποκριτηρίου βρίσκουμε την σωστή αντιστοιχία μεταξύ των βαθμολογιών των χαρακτηριστικών των ξενοδοχείων και των απαιτήσεων των χρηστών.
- Για κάθε κριτήριο αθροίζουμε τις τιμές που συγκέντρωσε ο συνδυασμός σε κάθε υποκριτήριο και υπολογίζουμε το μέσο όρο.
- Επίσης στα δύο τελευταία κριτήρια Προσωπικό και Εικόνα Μονάδας, τοποθετούμε τις τιμές από τα ratings των ξενοδοχείων που αντλήσαμε από την ιστοσελίδα booking.com όπως έχουμε προαναφέρει.
- Τέλος, γίνεται αντιστοίχιση τις τιμές του κάθε κριτηρίου, όπως υπολογίστηκαν παραπάνω από τους πίνακες, με τον χαρακτηρισμό του κριτηρίου και κατ' επέκταση την αντίστοιχη κωδικοποίηση.

Παρακάτω παρουσιάζονται αναλυτικά οι πίνακες συμφώνα με τους οποίους δίνονται οι τιμές στα υποκριτήρια και κατ' επέκταση στα κριτήρια, καθώς και οι κλίμακες μέτρησης του κάθε κριτηρίου.

Τοποθεσία

Υπολογισμός τιμής κριτηρίου: Τοποθεσία

Η τοποθεσία αποτελείται από :

Πίνακας 12 : Απόδοση τιμών στο κριτήριο Τοποθεσία(1)

Απόσταση από Σημεία Ενδιαφέροντος (m) – Θέση ξενοδοχείου	Καλή 0-400	Μέτρια 400-900	Κακή 900-πάνω
0-400	1	0,8	0,5
400-900	0,9	1	0,7
900-πάνω	0,5	0,8	1

Για παράδειγμα αν ο χρήστης προτιμάει να απέχει το ξενοδοχείο από το σημείο που τον ενδιαφέρει 300 μέτρα, σημαίνει ότι βρισκόμαστε στη πρώτη γραμμή του πίνακα **(0-400)**. Επίσης αν το ξενοδοχείο απέχει από το σημείο ενδιαφέροντος 500 μέτρα σημαίνει ότι βρισκόμαστε στη δεύτερη στήλη του πίνακα **Μέτρια (400-900)**. Άρα η τιμή του πρώτου υποκριτηρίου είναι $A_1 = 0,8$. Παρόμοια υπολογίζονται οι τιμές όλων των υποκριτηρίων.

Τιμή πρώτου υποκριτηρίου : A_1

Πίνακας 13 : Απόδοση τιμών στο κριτήριο Τοποθεσία(2)

Απόσταση από Μέσα Μαζικής Μεταφοράς (m) – Θέση ξενοδοχείου	Καλή 0-150	Μέτρια 150-350	Κακή 350-πάνω
0-150	1	0,8	0,5
150-350	0,9	1	0,7
350-πάνω	0,5	0,8	1

Τιμή δεύτερου υποκριτηρίου : A_2

Τιμή Κριτηρίου : $K_1 = (A_1 + A_2)/2$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του. Δηλαδή ανάλογα με την τιμή που παίρνει παραπάνω το κριτήριο ανήκει σε ένα συγκεκριμένο διάστημα της κλίμακας. Έτσι σύμφωνα με την κλίμακα μέτρησης, δίνεται στο κριτήριο ο αντίστοιχος χαρακτηρισμός και σύμφωνα με την κωδικοποίηση δίνεται η τελική τιμή του κριτηρίου.

Για παράδειγμα αν παραπάνω είχαμε υπολογίσει την **Τιμή Κριτηρίου** ίση με **0,9** θα αντιστοιχούσε στο τρίτο διάστημα. Άρα η **Τοποθεσία** του συγκεκριμένου ξενοδοχείου θα ήταν **Πολύ Καλή** και το κριτήριο **Τοποθεσία** θα έπαιρνα τελικά την τιμή **3**.

Πίνακας 14 : Κλίμακα και Κωδικοποίηση κριτηρίου Τοποθεσία

Τοποθεσία	Κλίμακα Μέτρησης	Κωδικοποίηση
Κακή	0-0,7	1
Μέτρια	0,71-0,9	2
Πολύ Καλή	0,91-1	3

Καθαριότητα

Υπολογισμός τιμής κριτηρίου : Καθαριότητα

Η καθαριότητα αποτελείται από :

Πίνακας 15 : Απόδοση τιμών στο κριτήριο Καθαριότητα (1)

Ποιότητα Καθαρισμού - Υπηρεσία καθαριότητας	Καθημερινή	Όχι Καθημερινή
Υψηλή (8-10)	1	0
Μέτρια (5-7)	0,8	0,7
Κακή (1-4)	0,6	1

Τιμή πρώτου υποκριτηρίου : A1

Πίνακας 16 : Απόδοση τιμών στο κριτήριο Καθαριότητα(2)

Υπηρεσία καθαριότητας	Υπηρεσία καθαριότητας
Καθημερινή	1
Όχι Καθημερινή	0

Τιμή δεύτερου υποκριτηρίου : A2

Τιμή Κριτηρίου : $K_2 = (A_1 + A_2)/2$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του.

Πίνακας 17 : Κλίμακα και Κωδικοποίηση κριτηρίου Καθαριότητα

Καθαριότητα	Κλίμακα Μέτρησης	Κωδικοποίηση
Κακή	0-0,39	1
Μέτρια	0,4-0,65	2
Ικανοποιητική	0,66-0,94	3
Πολύ Καλή	0,95-1	4

Παροχές Δωματίου

Υπολογισμός τιμής κριτηρίου : Παροχές Δωματίου

Η παροχές δωματίου αποτελούνται από :

Πίνακας 18 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(1)

Πρόσβαση Ιντερνέτ(WiFi) -WiFi	Δωρεάν	Πληρωμή	Δεν διαθέτει
Δωρεάν	1	0,8	0
Πληρωμή	0,8	1	0
ΟΧΙ	0,6	0,5	1

Τιμή πρώτου υποκριτηρίου : A1

Πίνακας 19 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(2)

Τηλεόραση- Τηλεόραση	Διαθέτει	Δεν διαθέτει
ΝΑΙ (6-10)	1	0
ΟΧΙ (1-5)	0,6	1

Τιμή δεύτερου υποκριτηρίου : A2

Πίνακας 20 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(3)

Air Condition- Θέρμανση/ Κλιματισμός	Διαθέτει	Δεν διαθέτει
ΝΑΙ (7-10)	1	0
ΟΧΙ (1-6)	0,6	1

Τιμή τρίτου υποκριτηρίου : A3

Πίνακας 21 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(4)

Ψυγείο/Καφετιέρα- Ψυγείο	Διαθέτει	Δεν διαθέτει
ΝΑΙ (5-10)	1	0
ΟΧΙ (1-4)	0,6	1

Τιμή τέταρτου υποκριτηρίου : A4

Πίνακας 22 : Απόδοση τιμών στο κριτήριο Παροχές Δωματίου(5)

Safety Deposit Box	Safety Deposit Box
Διαθέτει	1
Δεν διαθέτει	0

Τιμή πέμπτου υποκριτηρίου : A5

Τιμή Κριτηρίου : $K_3 = (A_1 + A_2 + A_3 + A_4 + A_5)/5$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του.

Πίνακας 23 : Κλίμακα και Κωδικοποίηση κριτηρίου Παροχές Δωματίου

Παροχές Δωματίου	Κλίμακα Μέτρησης	Κωδικοποίηση
Ανεπαρκείς	0-0,59	1
Μέτριες	0,6-0,79	2
Πολύ Καλές	0,8-0,92	3
Εξαιρετικές	0,93-1	4

Παροχές Ξενοδοχείου

Υπολογισμός τιμής κριτηρίου : Παροχές Ξενοδοχείου

Η παροχές ξενοδοχείου αποτελούνται από :

Πίνακας 24 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(1)

Παρκινγκ - Χώρος Στάθμευσης	Διαθέτει	Δεν διαθέτει
Δωρεάν	1	0
Πληρωμή	1	0
ΟΧΙ	0,6	1

Τιμή πρώτου υποκριτηρίου : A1

Πίνακας 25 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(2)

Σνακ Μπαρ /Εστιατόριο	Σνακ Μπαρ /Εστιατόριο
Διαθέτει	1
Δεν διαθέτει	0

Τιμή δεύτερου υποκριτηρίου : A2

Πίνακας 26 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(3)

Πρωινό – Πρωινό	Περιλαμβάνεται στη Τιμή	Με Πληρωμή	Δεν διαθέτει
Περιλαμβάνεται στη Τιμή	1	0,7	0
Με πληρωμή	0,7	1	0
ΟΧΙ	0	0,8	1

Τιμή τρίτου υποκριτηρίου : A3

Πίνακας 27 : Απόδοση τιμών στο κριτήριο Παροχές Ξενοδοχείου(4)

Ψυγείο/Καφετιέρα-Καφετιέρα/Καφές	Διαθέτει	Δεν διαθέτει
ΝΑΙ (5-10)	1	0
ΟΧΙ (1-4)	0,6	1

Τιμή τέταρτου υποκριτηρίου : A4

Τιμή Κριτηρίου : $K_4 = (A_1 + A_2 + A_3 + A_4)/4$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του.

Πίνακας 28 : Κλίμακα και Κωδικοποίηση κριτηρίου Παροχές Ξενοδοχείου

Παροχές Ξενοδοχείου	Κλίμακα Μέτρησης	Κωδικοποίηση
Ανεπαρκείς	0-0,5	1
Μέτριες	0,51-0,7	2
Πολύ Καλές	0,71-0,89	3
Εξαιρετικές	0,9-1	4

Σχέση Ποιότητας Τιμής

Υπολογισμός τιμής κριτηρίου : Σχέση Ποιότητας Τιμής

Η σχέση ποιότητας τιμής αποτελείται από :

Πίνακας 29 : Απόδοση τιμών στο κριτήριο Σχέση Ποιότητας Τιμής (1)

Κόστος(€)- Τιμή/Κόστος	Πολύ Μεγάλο >131	Μεγάλο >100 και <=131	Όχι πολύ Μεγάλο >75 και <=100	Όχι πολύ Μικρό >50 και <=75	Μικρό <=50
Πολύ Ακριβό >131	1	0,8	0,3	0,1	0
Ακριβό >100 και <=131	0,9	1	0,8	0,3	0
Όχι πολύ ακριβό >75 και <=100	0,3	0,8	1	0,8	0,3
Όχι πολύ Φθηνό >50 και <=75	0,1	0,3	0,6	1	0,6
Φθηνό <=50	0	0	0,3	0,8	1

Τιμή Κριτηρίου : $K_5 = A_1$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του.

Πίνακας 30 : Κλίμακα και Κωδικοποίηση κριτηρίου Σχέσης Ποιότητας Τιμής

Σχέσης Ποιότητας Τιμής	Κλίμακα Μέτρησης	Κωδικοποίηση
Πολύ Κακή	0-0,5	1
Κακή	0,51-0,7	2
Μέτρια	0,71-0,8	3
Καλή	0,81-0,9	4
Πολύ καλή	0,9-1	5

Άνεση

Υπολογισμός τιμής κριτηρίου : Άνεση

Η άνεση αποτελείται από :

Πίνακας 31 : Απόδοση τιμών στο κριτήριο Άνεση (1)

Ποιότητα ύπνου - Μεγάλο κρεβάτι	Διαθέτει	Δεν διαθέτει
Πολύ καλή (9-10)	1	0
Καλή (6-8)	0,9	0,3
Μέτρια (5-6)	0,8	0,5
Κακή (2-4)	0,7	0,8
Πολύ Κακή (0-2)	0,6	1

Τιμή πρώτου υποκριτηρίου : A1

Πίνακας 32 : Απόδοση τιμών στο κριτήριο Άνεση (2)

Βεράντα/Θέα	Βεράντα/Θέα
Διαθέτει	1
Δεν διαθέτει	0

Τιμή δεύτερου υποκριτηρίου : A2

Πίνακας 33 : Απόδοση τιμών στο κριτήριο Άνεση (3)

Θέρμανση/ Κλιματισμός	Θέρμανση/ Κλιματισμός
Διαθέτει	1
Δεν διαθέτει	0

Τιμή τρίτου υποκριτηρίου : A3

Πίνακας 34 : Απόδοση τιμών στο κριτήριο Άνεση (4)

Διαστάσεις δωματίου (m ²)	Διαστάσεις δωματίου (m ²)
Πολύ Μεγάλο >30	1
Μεγάλο >20 και <=30	0,8
Μικρό <=20	0,5

Τιμή τέταρτου υποκριτηρίου : A4

Τιμή Κριτηρίου : $K_6 = (A_1 + A_2 + A_3 + A_4)/4$

Σε αυτό το σημείο γίνεται η κωδικοποίηση του κριτηρίου σύμφωνα με την κλίμακα μέτρησης του.

Πίνακας 35 : Κλίμακα και Κωδικοποίηση κριτηρίου Άνεση

Άνεση	Κλίμακα Μέτρησης	Κωδικοποίηση
Κακή	0-0,6	1
Ικανοποιητική	0,61-0,85	2
Πολύ Καλή	0,86-1	3

Τέλος και η κωδικοποίηση για τα κριτήρια Προσωπικό και Εικόνα Μονάδας.

Πίνακας 36 : Κλίμακα και Κωδικοποίηση κριτηρίου Προσωπικό

Προσωπικό	Κλίμακα Μέτρησης	Κωδικοποίηση
Κακό	0-0,6	1
Μέτριο	0,61-0,86	2
Καλό	0,87-0,95	3
Πολύ καλό	0,96-1	4

Πίνακας 37 : Κλίμακα και Κωδικοποίηση κριτηρίου Εικόνα Μονάδας.

Εικόνα Μονάδας	Κλίμακα Μέτρησης	Κωδικοποίηση
Κακή	0-0,65	1
Μέτρια	0,66-0,87	2
Καλή	0,88-0,95	3
Εξαιρετική	0,96-1	4

3.5 Πολυκριτήριος Πίνακας

Ο πολυκριτήριος πίνακας για κάθε χρήστη αποτελείται από τις εναλλακτικές και από τα κριτήρια που χρησιμοποιούμε. Κάθε γραμμή αντιστοιχεί σε ένα ζεύγος πελάτη με ένα ξενοδοχείο και σε κάθε στήλη υπάρχει η συγκεκριμένη βαθμολογία που παίρνει το κάθε κριτήριο για το συγκεκριμένο συνδυασμό πελάτη - ξενοδοχείου. Στην τελευταία στήλη του πολυκριτήριου πίνακα βρίσκεται το Ranking. Για την δημιουργία του πίνακα αναπτύξαμε

προχωρήσαμε στην ανάπτυξη κώδικα (βλέπε 1. Κώδικας Python για τη δημιουργία πολυκριτήριου πίνακα. Ο κώδικας υλοποιήθηκε σε γλώσσα Python , στο περιβάλλον PyCharm και εφαρμόζει την μεθοδολογία απαιτήσεις-χαρακτηριστικά που αναλύσαμε στο 3.4 Απόδοση . Στην ουσία ο κώδικας αξιοποιεί τα δεδομένα που συλλέξαμε στο 3.2 Συλλογή Δεδομένων και εφαρμόζοντας την παραπάνω μεθοδολογία αποδίδονται οι τιμές στα κριτήρια και δημιουργείται ο πολυκριτήριος πίνακας.

Στους παρακάτω πίνακες (Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.) παρουσιάζονται οι πολυκριτήριοι πίνακες του παραδείγματος μας. Περιέχουν βαθμολογίες των κριτηρίων για ένα αντιπροσωπευτικό σύνολο αναφοράς ξενοδοχείων που επιλέξαμε. Από κάθε αποφασίζοντα ζητήθηκε να κατατάξει τα ξενοδοχεία ανάλογα με τις προτιμήσεις του, γνωρίζοντας τις καλύτερες και τις χειρότερες τιμές που μπορεί να πάρει κάθε κριτήριο.

Πίνακας 38 : Πολυκριτήριος Πίνακας για User1

Users	Alt/cr	Topothesia	Kathariotita	ParoxesDwmatiou	ParoxesKsenodoxeiu	SxesiPoiotitas/Timis	Anesi	Proswpiko	EikonaMonadas	Ranking
User1	Hotel2	3	4	4	4	1	3	3	2	1
User1	Hotel1	3	4	4	1	4	3	3	3	2
User1	Hotel7	3	4	4	2	2	1	3	2	3
User1	Hotel9	3	3	4	1	1	3	1	1	4
User1	Hotel13	3	1	1	1	1	1	2	2	5
User1	Hotel4	2	4	4	2	5	3	4	4	6
User1	Hotel3	2	4	4	4	4	3	4	3	7
User1	Hotel8	2	2	4	2	2	2	4	3	8
User1	Hotel12	1	2	2	1	1	2	4	3	9
User1	Hotel5	2	4	3	3	5	3	3	4	10
User1	Hotel11	2	2	3	1	3	2	2	2	10
User1	Hotel10	3	3	2	1	1	2	2	2	10
User1	Hotel6	2	4	4	3	4	3	3	4	11
User1	Hotel14	2	2	3	1	1	2	3	2	12

Πίνακας 39 : Πολυκριτήριος Πίνακας για User2

Users	Alt/crit	Topothesia	Kathariotita	ParoxesDwmatiou	ParoxesKsenodoxeiu	SxesiPoiotitas/Timis	Anesi	Proswpiko	EikonaMonadas	Ranking
User2	Hotel2	2	3	4	4	4	3	3	2	1
User2	Hotel7	2	3	4	2	5	1	3	2	2
User2	Hotel1	2	3	4	1	4	3	3	3	3
User2	Hotel9	2	3	4	1	3	3	1	1	3
User2	Hotel10	2	3	2	1	3	2	2	2	4
User2	Hotel14	2	2	3	1	4	2	3	2	5
User2	Hotel13	2	2	1	1	3	1	2	2	6
User2	Hotel3	3	3	4	4	1	3	4	3	7
User2	Hotel5	2	3	3	3	2	3	4	3	8
User2	Hotel11	2	2	3	1	1	2	2	2	9
User2	Hotel8	1	4	4	2	5	2	4	3	10
User2	Hotel12	1	4	2	1	4	2	4	3	11
User2	Hotel4	1	4	4	2	2	3	4	4	12
User2	Hotel6	1	4	4	3	2	3	4	4	12

3.6 UTASTAR

Η Utastar δέχεται ως είσοδο το πολυκριτήριο πίνακα καθώς και τον πίνακα με τα χαρακτηριστικά των κριτηρίων. Στο προηγούμενο βήμα δημιουργήσαμε ένα αντιπροσωπευτικό σύνολο αναφοράς ξενοδοχείων, με σκοπό να υπολογίσουμε τα βάρη των κριτηρίων και τις μερικές χρησιμότητες και στη συνέχεια να επεκτείνουμε τον υπολογισμό των ολικών χρησιμότητων για το σύνολο των συνδυασμών Ξενοδοχείο – Χρήστης.

Στην πρώτη στήλη του πίνακα μεταδεδομένων των κριτηρίων βρίσκονται τα κριτήρια που έχουμε επιλέξει να χρησιμοποιήσουμε. Στη δεύτερη στήλη καθορίζεται η μονοτονία κάθε κριτηρίου, 1 για φθίνουσα ή 0 για αύξουσα. Η τρίτη στήλη περιγράφει τον είδος των κριτηρίων, 1 για ποιοτικό και 0 για ποσοτικό. Στην επόμενη δύο στήλες υπάρχουν οι χειρότερες και οι καλύτερες τιμές που μπορεί να πάρει κάθε κριτήριο αντίστοιχα. Τέλος, υπάρχει το α που μας δίνει τα διαστήματα τιμών κάθε κριτηρίου ($\alpha-1$).

Πίνακας 40 : Πίνακας μεταδεδομένων κριτηρίων

Cri/attributes	Monocity	Type	Worst	Best	α
location	0	1	1	3	3
cleanness	0	1	1	4	4
room services	0	1	1	4	4
hotel services	0	1	1	4	4
value	0	1	1	5	5
comfort	0	1	1	3	3
staff	0	1	1	4	4
reputation	0	1	1	4	4

Τρέξαμε τον αλγόριθμο της Utastar δύο φορές, για τους χρήστες User1 και User2 αντίστοιχα. Κάθε φορά η Utastar δέχεται ως είσοδο τον πολυκριτήριο πίνακα του εκάστοτε χρήστη (Πίνακας 38, Πίνακας 39) και τον πίνακα των μεταδεδομένων (Πίνακας 40).

ΚΕΦΑΛΑΙΟ 4 : ΑΠΟΤΕΛΕΣΜΑΤΑ UTASTAR

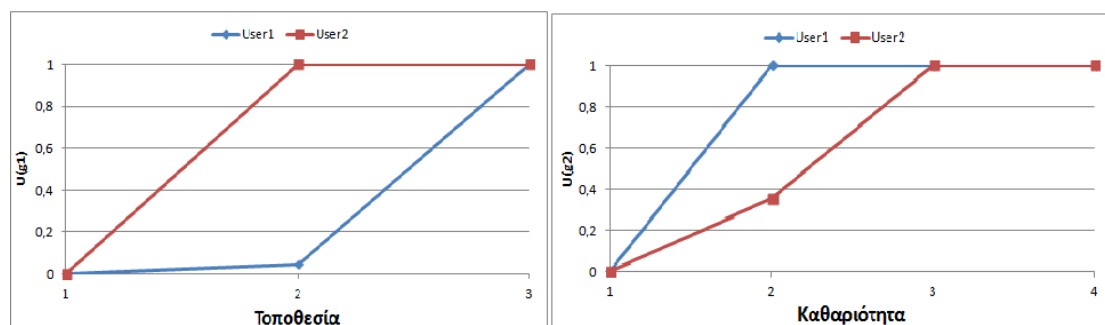
Τρέχοντας την Utastar για τους χρήστες User1 και User2 παίρνουμε ως αποτελέσματα τις μερικές χρησιμότητες των κριτηρίων, τα βάρη τους, τις ολικές χρησιμότητες των εναλλακτικών του καθώς και το συντελεστή συσχέτισης t (Kendall's tau) αντιπροσωπευτικού συνόλου.

Πίνακας 41 : Συντελεστής συσχέτισης

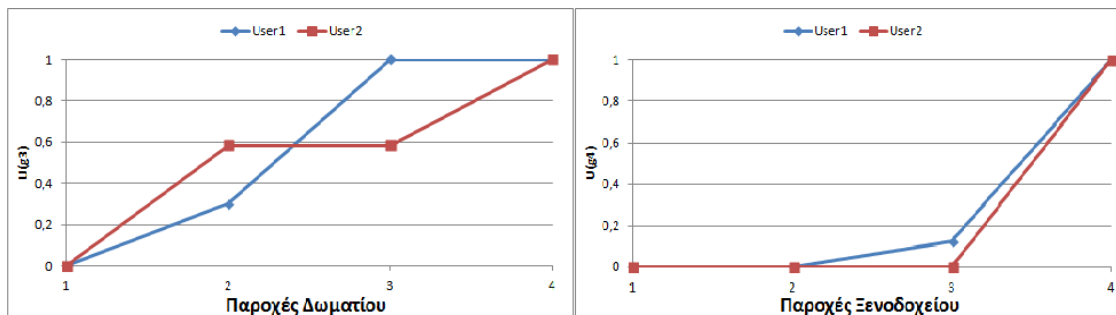
User	Kendall's tau
User1	0.84091
User2	0.99443

Ο συντελεστής συσχέτισης t μας δείχνει κατά πόσο ο χρήστης (αποφασίζων) και το μοντέλο μας έχουν ταξινομήσει τις εναλλακτικές πανομοιότυπα. Όπως μπορούμε να διακρίνουμε και στον Πίνακα 41, ο δεύτερος χρήστης έχει πολύ υψηλό δείκτη t του Kendall, το οποίο δείχνει ότι οι συστάσεις θα είναι αξιόπιστες.

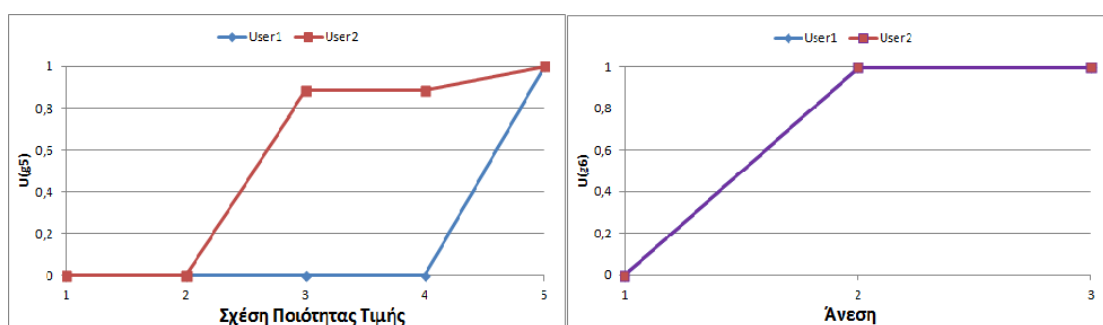
Παρακάτω απεικονίζονται τα γραφήματα με τις μερικές χρησιμότητες αξίας (marginal utilities) των κριτηρίων. Οι μερικές χρησιμότητες αντιπροσωπεύουν την προτίμηση του χρήστη από την λιγότερο επιθυμητή τιμή του κριτηρίου όπου η μερική χρησιμότητα είναι 0, έως την περισσότερη επιθυμητή τιμή του κριτηρίου όπου η τιμή της μερικής χρησιμότητας είναι $u_i(g_i^*)$.



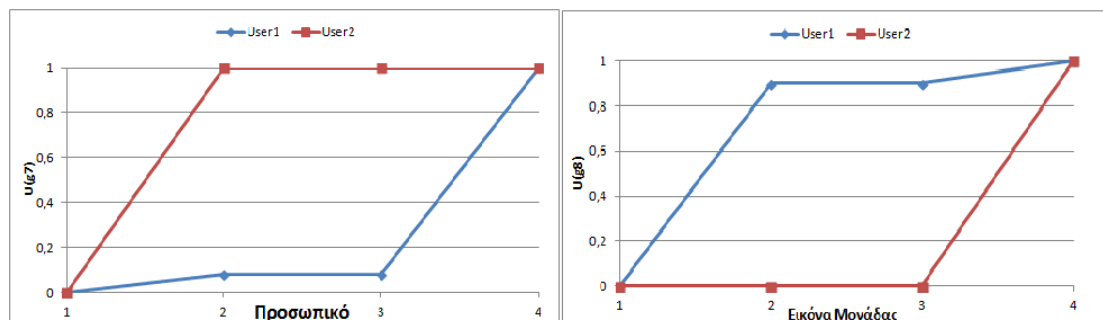
Διάγραμμα 1 : Μερικές Χρησιμότητες των κριτηρίων Τοποθεσία και Καθαριότητα



Διάγραμμα 2 : Μερικές Χρησιμότητες των κριτηρίων Παροχές Δωματίου και Παροχές Ξενοδοχείου



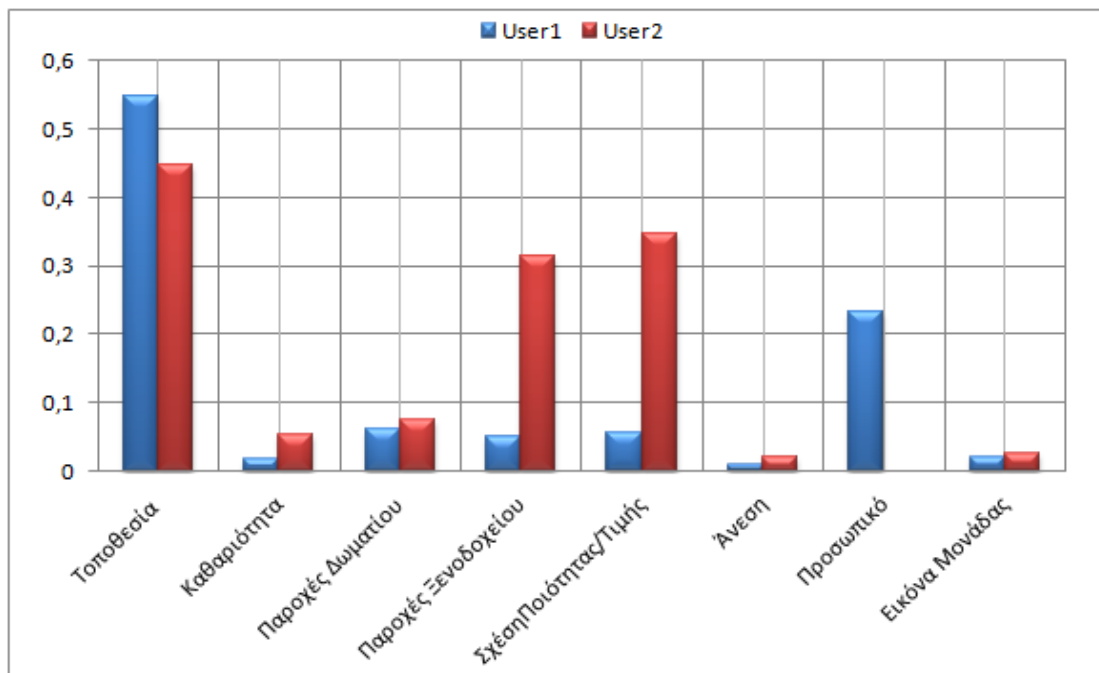
Διάγραμμα 3 : Μερικές Χρησιμότητες των κριτηρίων Σχέση Ποιότητας Τιμής και Άνεση



Διάγραμμα 4 : Μερικές Χρησιμότητες των κριτηρίων Προσωπικό και Εικόνα Μονάδας

Από τα παραπάνω διαγράμματα των μερικών χρησιμοτήτων παρατηρούμε μεγάλες αποκλίσεις για τους δύο χρήστες User1 και User2 στα περισσότερα. Τα κριτήρια Καθαριότητα, Παροχές Δωματίου και Παροχές Ξενοδοχείου παρουσιάζουν μικρότερες αποκλίσεις και τέλος στο κριτήριο Άνεση οι μερικές χρησιμότητες συμπίπτουν και για τους δύο χρήστες.

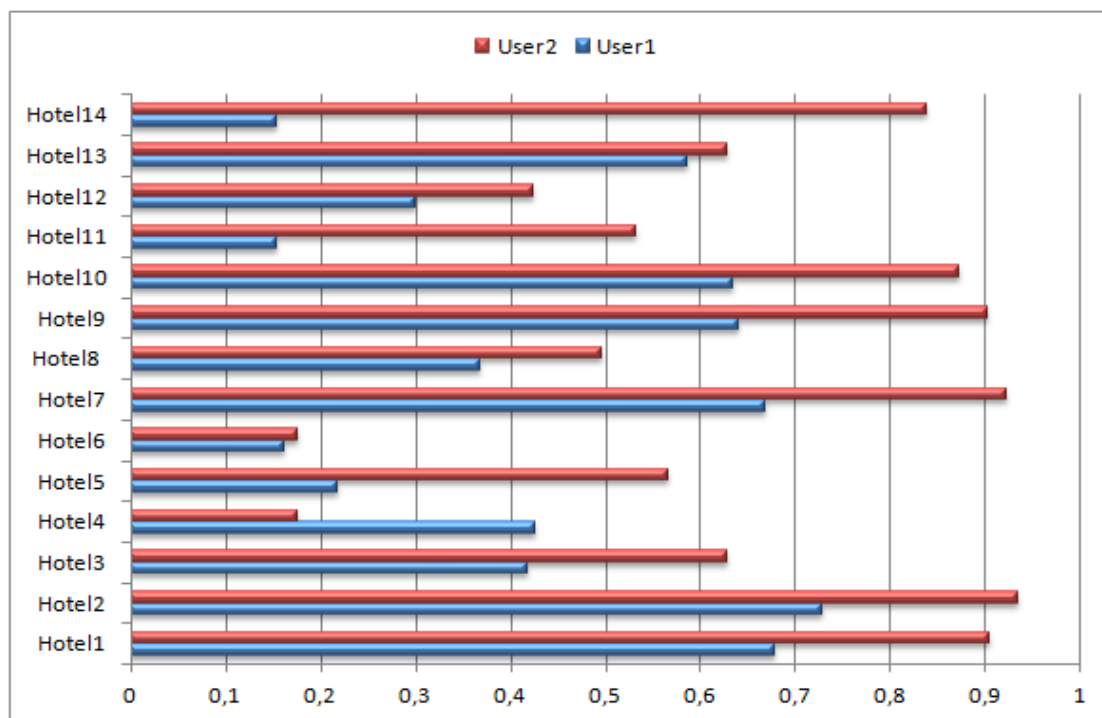
Στη συνέχεια παρουσιάζεται το διάγραμμα με τα βάρη των κριτηρίων για τους δύο χρήστες.



Διάγραμμα 5 : Βάρη κριτηρίων

Παρατηρώντας το διάγραμμα βαρών βλέπουμε ότι και οι δυο χρήστες δίνουν τη μεγαλύτερη έμφαση στο κριτήριο της Τοποθεσίας. Ο πρώτος χρήστης φαίνεται να δίνει αρκετό βάρος για την επιλογή ξενοδοχείου τόσο στη Σχέση Ποιότητας Τιμής όσο και στις Παροχές Ξενοδοχείου σε αντίθεση με τον δεύτερο που μετά την Τοποθεσία δίνει έμφαση στο κριτήριο Προσωπικό.

Τέλος απεικονίζεται το διάγραμμα των ολικών χρησιμοτήτων (Global Utilities) των εναλλακτικών και για τους δύο χρήστες.



Διάγραμμα 6 : Ολικές Χρησιμότητες εναλλακτικών

Στο παραπάνω διάγραμμα βλέπουμε ότι οι τιμές των ολικών χρησιμότητων των δύο χρηστών διαφέρουν μεταξύ τους και αυτό είναι αποτέλεσμα των διαφορετικών προτιμήσεων τους καθώς και των διαφορετικών βαρών που δίδει ο καθένας τους σε κάθε κριτήριο.

Τέλος, για τον υπολογισμό των ολικών χρησιμότητων στο σύνολο των ξενοδοχείων, δημιουργώντας το σύνολο των δυνατών συνδυασμών Ξενοδοχείου - Χρήστη, χρησιμοποιούμε τα βάρη και τις μερικές χρησιμότητες που πήραμε ως αποτελέσματα της UTASTAR, από το αντιπροσωπευτικό σύνολο αναφοράς που χρησιμοποιήσαμε. Με τη βοήθεια του κώδικα (βλέπε Παράρτημα) υπολογίσαμε τις ολικές χρησιμότητες των δύο χρηστών και τις παρουσιάζουμε παρακάτω ταξινομημένες.

Πίνακας 42 : Ολικές Χρησιμότητες για User1

Hotel	Global Utilities	Hotel	Global Utilities	Hotel	Global Utilities
Hostel Cocoon City	0,89799963	Apartment Vranas Studios	0,67720528	Hotel Christina	0,20333512
Guest house Viaggio Elegant Rooms	0,8938	Holiday home Aphrodite House	0,66780528	Hotel Ellinis	0,20333512
Aparthotel Palazzo di Irene	0,8938	Hotel Elia Dorotheou	0,66780528	Hotel El Greco	0,20333512
Apartment Odyssey Studios	0,8938	Hotel Apollon	0,66780528	Apartment Caza Latina	0,18453512
Apartment Irene Maisonettes	0,8938	Guest house Chania Rooms	0,65840528	Apartment Loucerna Art	0,15959512
Hotel Ionas Boutique	0,89172463	Guest house Kydonia Rooms	0,6396	Aparthotel Palazzo Duca	0,15959512
Hotel Minoos Hotel	0,89172463	Apartment Angelika Studios	0,63348028	Hotel Irida Hotel	0,15959512
Guest house Katerina Traditional Rooms	0,89172463	Guest house Narkissos	0,63348028	Guest house Casa Kasteli	0,15333512
Holiday home Zervos Chania	0,73340528	Guest house Shalom Luxury Rooms	0,62408028	Apartment Honolulu Beach Apartments	0,15333512
Guest house Elia Palatino	0,73340528	Guest house Lito Rooms	0,58650528	Hotel Nefeli	0,15333512
Apartment Bluebell Luxury Apartments	0,73340528	Hotel Casa Delfino Hotel & Spa	0,43031447	Apartment Elia Estia	0,15333512
Hotel Kriti	0,72720528	Hotel Porto Veneziano Hotel	0,43031447	Hotel Elia Zampeliou Boutique Hotel	0,15333512
Guest house Castle Inn	0,72720528	Hotel Kydon The Heart City Hotel	0,43031447	Hotel Arkadi	0,15333512
Apartment Bluebell Luxury Suites	0,68346528	Hotel Madonna del Mare	0,42612984	Guest house Elia Kentron	0,15333512

Hotel Elia Bettolo	0,67720528	Aparthotel Trianon Luxury Apartments & Suites	0,42405447	Aparthotel Pantheon Apartments	0,15333512
Apartment C&M Residence Chania Old Town	0,67720528	Apartment Sea & City Apartments	0,42405447	Apartment Stone House NF1	0,15333512
Hotel Elia Portou Due	0,67720528	Hotel Samaria	0,41785447	Hotel Tellus City	0,15333512
Aparthotel 11City Rooms	0,67720528	Hotel Domus Renier Boutique	0,37618984	Hotel TheJoy City Suites	0,15333512
Apartment Bluebell Minaret Suites	0,67720528	Guest house La Maison Ottomane	0,37618984	Hotel Irene Hotel	0,15333512
Guest house Dream Art Studios	0,67720528	Hotel Avra City Hotel	0,37411447	Hotel Filoxenia	0,15333512
Apartment Horizon Deluxe House	0,67720528	Hotel Casa Leone Boutique Hotel	0,37411447	Guest house Pension Eva	0,15333512
Guest house Elia Daliani Suites	0,67720528	Aparthotel Residenza Vranas Boutique Hotel	0,36992984	Apartment Constantino's Studios	0,15333512
Hotel Porto Antico	0,67720528	Apartment Nival Luxury Suites	0,36992984	Hotel TheJoy Hotel	0,15333512
Hotel Vilelmine	0,67720528	Hotel Diamond Luxury Suites	0,36992984	Apartment Eleonora	0,15333512
Hotel Doge Traditional	0,67720528	Apartment Nonnalena Boutique Hotel	0,36992984	Aparthotel Alena Apartments	0,15333512
Guest house Sofita Suites	0,67720528	Guest house Shalom Luxury Rooms Daliani	0,36992984	Hotel Tina Hotel	0,15333512
Apartment Artemide city places 2	0,67720528	Apartment Stoa Suites Chania	0,36992984	Apartment Smaragdi	0,15333512
Guest house Mosaic	0,67720528	Apartment Studio Miltos Aspasia 1	0,36992984	Aparthotel Katerina	0,14393512
Apartment Elia Potie	0,67720528	Hotel Belmondo	0,36785447	Apartment Ifigenia	0,13453512

Πίνακας 43 : Ολικές Χρησιμότητες για User2

Hotel	Global Utilities	Hotel	Global Utilities	Hotel	Global Utilities
Apartment Bluebell Luxury Suites	0,9438	Aparthotel Alena Apartments	0,903143	Guest house Lito Rooms	0,771904
Hotel Elia Portou Due	0,9438	Hotel Ellinis	0,900104	Aparthotel Kleri Beloni Rooms	0,771904
Apartment Bluebell Minaret Suites	0,9438	Apartment Caza Latina	0,900104	Hotel Samaria	0,6282
Guest house Dream Art Studios	0,9438	Hotel Nefeli	0,878215	Hotel Civitel Akali	0,6282
Hotel Vilelmine	0,9438	Hotel Irida	0,871898	Hotel Christina	0,6282
Apartment Elia Potie	0,9438	Guest house Elia Daliani Suites	0,871898	Aparthotel Residenza Vranas Boutique Hotel	0,6219
Apartment Artemide city places	0,9438	Guest house Sofita Suites	0,871898	Hotel Domus Renier Boutique	0,6219
Aparthotel Deluxe Hotel	0,9438	Aparthotel Pantheon Apartments	0,871898	Hotel Porto Veneziano	0,5969
Hotel Kriti	0,934443	Apartment Angelika Studios	0,871898	Hotel Kydon The Heart CityHotel	0,5969
Guest house Castle Inn	0,934443	Hotel Tellus CityHotel	0,871898	Hotel Avra City	0,5969
Guest house Viaggio Elegant Rooms	0,928143	Hotel TheJoy City Suites	0,871898	Hotel Casa Leone Boutique	0,5969
Aparthotel Palazzo di Irene	0,928143	Hotel Irene	0,871898	Hotel Ionas Boutique	0,5969
Guest house Shalom Luxury Rooms Daliani	0,928143	Hotel Filoxenia	0,871898	Aparthotel Trianon Luxury Apartments & Suites	0,5969
Apartment Stoa Suites Chania	0,928143	Hotel TheJoy	0,871898	Guest house Elia Palatino	0,5969

Apartment Odyssey Studios	0,928143	Apartment Eleonora	0,871898	Apartment Horizon Deluxe House	0,5969
Apartment Studio Miltos Aspasia 1	0,928143	Hotel Porto Antico	0,868804	Apartment Bluebell Luxury Apartments	0,5969
Apartment Irene Maisonettes	0,928143	Hotel Arkadi	0,868804	Apartment Sea & City Apartments	0,5969
Hotel Elia Dorotheou	0,9219	Guest house Chania Rooms	0,868804	Hotel Doge Traditional Hotel	0,5969
Aparthotel 11City Rooms	0,912555	Apartment Ifigenia Rooms	0,868804	Hotel Casa Delfino Hotel & Spa	0,565655
Hotel Minoos	0,912555	Guest house Pension Eva	0,868804	Apartment C&M Residence Chania Old Town	0,56256
Guest house Mosaic	0,912555	Guest house Shalom Luxury Rooms	0,849998	Apartment Honolulu Beach Apartments	0,531315
Apartment Vranas Studios	0,912555	Aparthotel Katerina Rooms for Rent	0,849998	Holiday home Zervos Chania	0,531315
Apartment Artemide city places 2	0,90946	Holiday home Aphrodite House	0,846904	Apartment Nonnalena Boutique Hotel	0,5219
Hotel El Greco	0,903198	Guest house Ifigenia by Captain Michalis	0,837559	Guest house Casa Kasteli	0,4969
Cocoon City Hostel	0,903198	Hotel Lefka Ori	0,837559	Hotel Belmondo	0,4969
Apartment Loucerna Art	0,903143	Apartment Constantino's Studios	0,837559	Apartment Nival Luxury Suites	0,481243
Hotel Elia Bettolo	0,903143	Guest house Katerina Traditional	0,837559	Apartment Elia Estia	0,456243

Από τους πίνακες των ολικών χρησιμοτήτων μπορούμε να διακρίνουμε ποια ξενοδοχεία είναι κατάλληλα για τον κάθε χρήστη. Οι εναλλακτικές με χρησιμότητα κοντά στη μονάδα είναι οι καταλληλότερες άρα αυτές θα είναι και οι συστάσεις για εκείνον. Συνεπώς οι συστάσεις για τους χρήστες User1 και User2 παρουσιάζονται στους παρακάτω πίνακες.

Πίνακας 44 : Τελικές Συστάσεις Ξενοδοχείων για User1 και User2

Συστάσεις Ξενοδοχείων για User1	Συστάσεις Ξενοδοχείων για User2
Cocoon City	Apartment Bluebell Luxury Suites
Guest house Viagio Elegant Rooms	Hotel Elia Portou Due
Aparthotel Palazzo di Irene	Apartment Bluebell Minaret Suites
Apartment Odyssey Studios	Guest house Dream Art Studios
Apartment Irene Maissonetes	Hotel Vilelmine
Hotel Ionas Boutique	Apartment Elia Potie
Hotel Minoos	Apartment Artemide city places
Guest house Katerina Traditional Rooms	Aparthotel Deluxe Hotel
Holiday home Zervos Chania	Hotel Kriti
Guest house Elia Palatino	Guest house Castle Inn

ΚΕΦΑΛΑΙΟ 5 : ΣΥΜΠΕΡΑΣΜΑΤΑ

Σκοπός της εργασίας ήταν η ανάπτυξη ενός αξιόπιστου πολυκριτήριου συστήματος συστάσεων που θα προτείνει στους χρήστες τα κατάλληλα ξενοδοχειακά καταλύματα για αυτούς. Για να υλοποιηθεί το συγκεκριμένο σύστημα χρησιμοποιήθηκαν διάφορες τεχνικές. Έγινε εξόρυξη δεδομένων για την εξαγωγή χρησίμων πληροφοριών. Επίσης υλοποιήθηκε κώδικας Python για την απόδοση τιμών στα κριτήρια για όλους τους συνδυασμούς Ξενοδοχείων – Χρηστών καθώς και χρήση της πολυκριτήριας μεθοδολογίας για λήψη αποφάσεων UTASTAR.

Τα αποτελέσματα που πήραμε και για τους δύο χρήστες ήταν αρκετά αξιόπιστα καθώς επιτύχαμε υψηλές τιμές του δείκτη συσχέτισης t του Kendall. Συνεπώς καταλήγουμε στο συμπέρασμα ότι οι συστάσεις μας προς τους χρήστες ήταν αξιόπιστες.

Για την βελτίωση του συστήματος, θα ήταν χρήσιμη η βελτιστοποίηση του κώδικα συλλογής των δεδομένων, ώστε να συλλέγονται μόνο τα απαραίτητα δεδομένα, στη σωστή μορφή και να αξιοποιούνται κατάλληλα.

Σχετικά με τις μελλοντικές προεκτάσεις που μπορούν να γίνουν στην εργασία, είναι η εξέλιξη του υπάρχοντος συστήματος σε ένα πολυκριτήριο σύστημα συστάσεων συνεχούς μάθησης. Κατά συνέπεια, η ανάπτυξη ενός δυναμικού συστήματος υποστηρίζοντας την προσθήκη, επεξεργασία και διαγραφή ξενοδοχειακών μονάδων, των υπηρεσιών και παροχών τους καθώς και τη συνεχή εκπαίδευση του μοντέλου σύμφωνα με τα νέα δεδομένα χρηστών.

Βιβλιογραφία

4. Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3), 56-58.
5. Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4), 331-370.
6. Kunaver, M., & Požrl, T. (2017). Diversity in recommender systems—A survey. *Knowledge-Based Systems*, 123, 154-162.
7. Adomavicius, G., & Kwon, Y. (2007). New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems*, 22(3), 48-55.
8. Lakiotaki, K., Matsatsinis, N. F., & Tsoukias, A. (2011). Multicriteria user modeling in recommender systems. *IEEE Intelligent Systems*, 26(2), 64-76.
9. Liu, L., Mehandjiev, N., & Xu, D. L. (2011, October). Multi-criteria service recommendation based on user criteria preferences. In *Proceedings of the fifth ACM conference on Recommender systems* (pp. 77-84).
10. Jannach, D., Karakaya, Z., & Gedikli, F. (2012, June). Accuracy improvements for multi-criteria recommender systems. In *Proceedings of the 13th ACM conference on electronic commerce* (pp. 674-689).
11. Premchaiswadi, W., & Poompuang, P. (2013). Hybrid profiling for multicriteria recommendation based on implicit multicriteria information. *Applied Artificial Intelligence*, 27(3), 213-234.
12. Xu, Y., Li, K., Hu, J., & Li, K. (2014). A genetic algorithm for task scheduling on heterogeneous computing systems using multiple priority queues. *Information Sciences*, 270, 255-287.
13. Bilge, A., & Kaleli, C. (2014, May). A multi-criteria item-based collaborative filtering framework. In *2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 18-22). IEEE.
14. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001, April). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (pp. 285-295).
15. Billsus, D., & Pazzani, M. J. (1998, July). Learning Collaborative Information Filters. In *Icml* (Vol. 98, pp. 46-54).
16. Li, Q., Wang, C., & Geng, G. (2008, April). Improving personalized services in mobile commerce by a novel multicriteria rating approach. In *Proceedings of the 17th international conference on World Wide Web* (pp. 1235-1236).
17. Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, 13(5-6), 393-408.
18. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Application of dimensionality reduction in recommender system—a case study (No. TR-00-043). Minnesota Univ Minneapolis Dept of Computer Science.
19. Schein, A. I., Popescul, A., Ungar, L. H., & Pennock, D. M. (2002, August). Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 253-260).

20. Rashid, A. M., Albert, I., Cosley, D., Lam, S. K., McNee, S. M., Konstan, J. A., & Riedl, J. (2002, January). Getting to know you: learning new user preferences in recommender systems. In Proceedings of the 7th international conference on Intelligent user interfaces (pp. 127-134).
21. Breese, J. S., Heckerman, D., & Kadie, C. (2013). Empirical analysis of predictive algorithms for collaborative filtering. arXiv preprint arXiv:1301.7363
22. Burke, R. (2003, August). Hybrid systems for personalized recommendations. In IJCAI Workshop on Intelligent Techniques for Web Personalization (pp. 133-152). Springer, Berlin, Heidelberg..
23. Mehrbakhsh N., Othman I. (2015). A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS, Electronic Commerce Research and Applications.
24. Σίσκος Ι. (2008). ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ, Μεθοδολογία Επιχειρησιακής Έρευνας, Θεωρία Πολυκριτήριας Ανάλυσης, Εφαρμογές σε Επιχειρήσεις & Οργανισμούς, Εκδόσεις Νέων Τεχνολογιών, Αθήνα, σελ. 309.
25. Σίσκος Ι. (2008). ΜΟΝΤΕΛΑ ΑΠΟΦΑΣΕΩΝ, Μεθοδολογία Επιχειρησιακής Έρευνας, Θεωρία Πολυκριτήριας Ανάλυσης, Εφαρμογές σε Επιχειρήσεις & Οργανισμούς, Εκδόσεις Νέων Τεχνολογιών, Αθήνα, σελ. 313.
26. Siskos, Y., Grigoroudis, E., & Matsatsinis, N. F. (2016). UTA methods. In Multiple criteria decision analysis (pp. 315-362). Springer, New York, NY.
27. Jacquet-Lagrange, E., & Siskos, J. (1982). Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. European journal of operational research, 10(2), 151-164.
28. Siskos, Y., & Yannacopoulos, D. (1985). UTASTAR: An ordinal regression method for building additive value functions. Investigacao Operacional, 5(1), 39-53.
29. “Artificial Intelligence and Machine Learning: Policy Paper.” Internet Society, April (2017). <https://www.internetsociety.org/wp-content/uploads/2017/08/ISOC-AI-Policy-Paper-2017-04-27-0.pdf>
30. Dunham, M. H. (2006). *Data mining: Introductory and advanced topics*. Pearson Education India.
31. Matsatsinis, N. F., & Delias, P. (2003, September). AgentAllocator: An agent-based multi-criteria decision support system for task allocation. In International Conference on Industrial Applications of Holonic and Multi-Agent Systems (pp. 225-235). Springer, Berlin, Heidelberg.

Παράρτημα

1. Κώδικας Python για τη δημιουργία πολυκριτήριου πινάκα

app

//////////

```
from rw_utils import read_excel
```

```
from rw_utils import write_results
```

```
from ranking import rank
```

```
from conversion import conversion
```

```
from xrisimotites import xrisimotites
```

```
import pandas
```

```
# Initialize all necessary variables
```

```
# The file name of the Booking dataset
```

```
source_file_name = "assets/chania_town_18_2_2019_NEW.xlsx"
```

```
# Define the columns you want to extract
```

```
# from this source excel file
```

```
columns = ["Name", "Distance Center", "Απόσταση από Μ.Μ.Μ", "Clean", "Value",  
"Wi-Fi", "Χώρος στάθμευσης", "Θέρμανση /Κλιματισμός", "Τηλεόραση", "Σνακ Μπαρ  
/Εστιατόριο", "Υπηρεσία καθαριότητας", "Μεγάλο κρεβάτι", "Κόστος", "Διαστάσεις  
δωματίου(m2)", "Πρωινό", "Καφετιέρα/Καφές", "Safety Deposit Box", "Staff",  
"Βεράντα Θέα", "Ψυγείο", "Total"]
```

```
# The file name of user preferences dataset
```

```
user_pref_file_name = "assets/Έρευνα απαιτήσεων πελατών για την επιλογή  
ξενοδοχείου (Απαντήσεις).xlsx"
```

```
# Define the columns you want to extract
```

```
# from this user preferences excel
```

```
pre_columns = ["Dinstance1;", "Dinstance2;", "Price", "Πρόσβαση Ίντερνετ (WiFi)", "Air Condition", "Τηλεόραση", "Πάρκινγκ", "Πρωινό Γεύμα", "Ψυγείο/Καφετιέρα", "Ποιότητα Καθαρισμού", "Ποιότητα Ύπνου", "Safety Deposit Box (Χρηματοκιβώτιο)"]
```

```
# The file name for the quantitative results
```

```
res_file_name = "assets/results.xlsx"
```

```
# The file name for the qualitative results
```

```
res_1_file_name = "assets/results_1.xlsx"
```

```
# The file name for the marginal utilities results
```

```
res_2_file_name = "assets/xrisimotites.xlsx"
```

```
res = read_excel(source_file_name, columns)
```

```
res1 = read_excel(user_pref_file_name, pre_columns)
```

```
#print(res1.describe())
```

```
#res2 = read_excel(user_pref_file_name, pre_columns)
```

```
writer = pandas.ExcelWriter(res_file_name, engine='xlsxwriter')
```

```
writer_1 = pandas.ExcelWriter(res_1_file_name, engine='xlsxwriter')
```

```
writer_2 = pandas.ExcelWriter(res_2_file_name, engine='xlsxwriter')
```

```
for index, row in res1.iterrows():
```

```
    df = rank(row, res)
```

```
    df_1 = conversion(df)
```

```
    df_2 = xrisimotites(df_1)
```

```
    df.to_excel(writer, sheet_name="user id" + str(index))
```

```
    df_1.to_excel(writer_1, sheet_name="user id" + str(index))
```

```
    df_2.to_excel(writer_2, sheet_name="user id" + str(index))
```

```
writer.save()
writer_1.save()
writer_2.save()

////////////////////////////////////////////////////////////////
```

ranking

```
////////
```

```
import pandas
```

```
def rank(user, hotels):

    df = pandas.DataFrame(columns=['hotel', 'location', 'cleanness', 'roomservices',
    'hotelservices', 'value', 'comfort', 'staff', 'reputation']) #, 'overall'])

    for index, row in hotels.iterrows():

        global l_1, l_2, w, c, tv, clima, f_c, s_b, p, breakfast, fridge, value, sp, sp_1,
        sp_2, sp_3

        # ranking according to the distance from the center

        if row["Distance Center"] < 0.4:

            if user["Dinstance1;"] == "0-400":

                l_1 = 1

            elif user["Dinstance1;"] == "400-900":

                l_1 = 0.9

            else:

                l_1 = 0.5

        elif row["Distance Center"] >= 0.4 and row["Distance Center"] < 0.9:

            if user["Dinstance1;"] == "0-400":

                l_1 = 0.8

            elif user["Dinstance1;"] == "400-900":
```

```

        l_1 = 1
    else:
        l_1 = 0.8
elif row["Distance Center"] >= 0.9:
    if user["Dinstance1;"] == "0-400":
        l_1 = 0.5
    elif user["Dinstance1;"] == "400-900":
        l_1 = 0.7
    else:
        l_1 = 1

# ranking according to the distance from the airport
if row["Απόσταση από Μ.Μ.Μ"] < 150:
    if user["Τι απόσταση (m) προτιμάτε να έχει το ξενοδοχείο από τα Μέσα
    Μαζικής Μεταφοράς;"] == "0-150":
        l_2 = 1
    elif user["Dinstance2;"] == "150-350":
        l_2 = 0.9
    else:
        l_2 = 0.5
elif row["Απόσταση από Μ.Μ.Μ"] >= 150 and row["Απόσταση από Μ.Μ.Μ"]
< 350:
    if user["Dinstance2;"] == "0-150":
        l_2 = 0.8
    elif user["Dinstance2;"] == "150-350":
        l_2 = 1
    else:
        l_2 = 0.8

```

```

elif row["Απόσταση από Μ.Μ.Μ"] >= 350:

    if user["Dinstance2;"] == "0-150":

        l_2 = 0.5

    elif user["Dinstance2;"] == "150-350":

        l_2 = 0.7

    else:

        l_2 = 1

# ranking according to the cleanness

if row["Υπηρεσία καθαριότητας"] == "Καθημερινή":

    if user["Ποιότητα Καθαρισμού"] >= 8:

        c = 1

    elif user["Ποιότητα Καθαρισμού"] >= 5 and user["Ποιότητα Καθαρισμού"]
< 8:

        c = (0.8 + 1) / 2

    else:

        c = (0.6 + 1) / 2

elif row["Υπηρεσία καθαριότητας"] == "Όχι Καθημερινή":

    if user["Ποιότητα Καθαρισμού"] >= 8:

        c = (0 + 0) / 2

    elif user["Ποιότητα Καθαρισμού"] >= 5 and user["Ποιότητα Καθαρισμού"]
< 8:

        c = (0.7 + 0) / 2

    else:

        c = (1 + 0) / 2

### ranking according to the room amenities ###

# Wi-fi

```

```

if row["Wi-Fi"] == user["Πρόσβαση Ίντερνετ (WiFi)"]:
    w = 1
elif row["Wi-Fi"] == "Δωρεάν":
    if user["Πρόσβαση Ίντερνετ (WiFi)"] == "Πληρωμή":
        w = 0.8
    elif user["Πρόσβαση Ίντερνετ (WiFi)"] == "ΟΧΙ":
        w = 0.6
elif row["Wi-Fi"] == "Πληρωμή":
    if user["Πρόσβαση Ίντερνετ (WiFi)"] == "Δωρεάν":
        w = 0.8
    elif user["Πρόσβαση Ίντερνετ (WiFi)"] == "ΟΧΙ":
        w = 0.5
elif row["Wi-Fi"] == "Δεν διαθέτει":
    if user["Πρόσβαση Ίντερνετ (WiFi)"] == "Δωρεάν":
        w = 0
    elif user["Πρόσβαση Ίντερνετ (WiFi)"] == "Πληρωμή":
        w = 0
    elif user["Πρόσβαση Ίντερνετ (WiFi)"] == "ΟΧΙ":
        w = 1

# TV
if row["Τηλεόραση"] == "Διαθέτει":
    if user["Τηλεόραση"] < 5:
        tv = 0.6
    else:
        tv = 1
else:
    if user["Τηλεόραση"] < 5:

```



```

        tv = 1
    else:
        tv = 0

# Air condition
if row["Θέρμανση /Κλιματισμός"] == "Διαθέτει":
    if user["Air Condition"] < 6:
        clima = 0.6
    else:
        clima = 1
else:
    if user["Air Condition"] < 6:
        clima = 1
    else:
        clima = 0

# Fridge / Coffee machine
if row["Καφετιέρα/Καφές"] == "Διαθέτει":
    if user["Ψυγείο/Καφετιέρα"] < 4:
        f_c = 0.6
    else:
        f_c = 1
else:
    if user["Ψυγείο/Καφετιέρα"] < 4:
        f_c = 1
    else:
        f_c = 0

```

```

# Safety Box

if row["Safety Deposit Box"] == "Διαθέτει":
    s_b = 1
else:
    s_b = 0

### ranking according to the Hotel Services ###

# Parking

if row["Χώρος στάθμευσης"] == "Διαθέτει":
    if user["Πάρκινγκ"] == "Δωρεάν" or user["Πάρκινγκ"] == "Πληρωμή":
        p = 1
    else:
        p = 0.6
else:
    if user["Πάρκινγκ"] == "Δωρεάν" or user["Πάρκινγκ"] == "Πληρωμή":
        p = 0
    else:
        p = 1

# Bar - Restaurant

if row["Σνακ Μπαρ /Εστιατόριο"] == "Διαθέτει":
    bar = 1
else:
    bar = 0

# Breakfast

if row["Πρωινό"] == "Including Breakfast":
    if user["Πρωινό Γεύμα"] == "Στην Τιμή":

```

```

        breakfast = 1
    elif user["Πρωινό Γεύμα"] == "Με Πληρωμή":
        breakfast = 0.7
    else:
        breakfast = 0
elif row["Πρωινό"] == "Optional Breakfast":
    if user["Πρωινό Γεύμα"] == "Στην Τιμή":
        breakfast = 0.7
    elif user["Πρωινό Γεύμα"] == "Με Πληρωμή":
        breakfast = 1
    else:
        breakfast = 0.8
else:
    if user["Πρωινό Γεύμα"] == "Στην Τιμή":
        breakfast = 0
    elif user["Πρωινό Γεύμα"] == "Με Πληρωμή":
        breakfast = 0
    else:
        breakfast = 1

# Fridge / Coffee machine
if row["Ψυγείο"] == "Διαθέτει":
    if user["Ψυγείο/Καφετιέρα"] < 4:
        fridge = 0.6
    else:
        fridge = 1
else:
    if user["Ψυγείο/Καφετιέρα"] < 4:

```

```

fridge = 1

else:

    fridge = 0

### Value / money ###

if row["Κόστος"] <= 50:

    if user["Price"] == "έως 50":

        value = 1

    elif user["Price"] == "Από 51 έως 75":

        value = 0.8

    elif user["Price"] == "Από 76 έως 100":

        value = 0.3

    elif user["Price"] == "Από 101 έως 130":

        value = 0

    else:

        value = 0

elif row["Κόστος"] > 50 and row["Κόστος"] <= 75:

    if user["Price"] == "έως 50":

        value = 0.6

        elif user["Πόσα χρήματα (ευρώ) είστε διατεθειμένοι να ξοδέψετε για κάθε
διανυκτέρευση (τιμή δωματίου);"] == "Από 51 έως 75":

            value = 1

        elif user["Price"] == "Από 76 έως 100":

            value = 0.6

        elif user["Price"] == "Από 101 έως 130":

            value = 0.3

    else:

        value = 0.1

```

```

elif row["Κόστος"] > 75 and row["Κόστος"] <= 100:
    if user["Price"] == "έως 50":
        value = 0.3
    elif user["Price"] == "Από 51 έως 75":
        value = 0.8
    elif user["Price"] == "Από 76 έως 100":
        value = 1
    elif user["Price"] == "Από 101 έως 130":
        value = 0.8
    else:
        value = 0.5
elif row["Κόστος"] > 100 and row["Κόστος"] <= 131:
    if user["Price"] == "έως 50":
        value = 0
    elif user["Price"] == "Από 51 έως 75":
        value = 0.3
    elif user["Price"] == "Από 76 έως 100":
        value = 0.8
    elif user["Price"] == "Από 101 έως 130":
        value = 1
    else:
        value = 0.9
else:
    if user["Price"] == "έως 50":
        value = 0
    elif user["Price"] == "Από 51 έως 75":
        value = 0.1
    elif user["Price"] == "Από 76 έως 100":

```

```

        value = 0.3
    elif user["Price"] == "Από 101 έως 130":
        value = 0.8
    else:
        value = 1

### Spaciousness ###
if row["Μεγάλο κρεβάτι"] == "Διαθέτει":
    if user["Ποιότητα Ύπνου"] >= 9:
        sp = 1
    elif user["Ποιότητα Ύπνου"] >= 6 and user["Ποιότητα Ύπνου"] < 9:
        sp = 0.9
    elif user["Ποιότητα Ύπνου"] >= 5 and user["Ποιότητα Ύπνου"] < 6:
        sp = 0.8
    elif user["Ποιότητα Ύπνου"] >= 2 and user["Ποιότητα Ύπνου"] < 5:
        sp = 0.7
    elif user["Ποιότητα Ύπνου"] < 4:
        sp = 0.6
else:
    if user["Ποιότητα Ύπνου"] >= 9:
        sp = 0
    elif user["Ποιότητα Ύπνου"] >= 6 and user["Ποιότητα Ύπνου"] < 9:
        sp = 0.3
    elif user["Ποιότητα Ύπνου"] >= 5 and user["Ποιότητα Ύπνου"] < 6:
        sp = 0.5
    elif user["Ποιότητα Ύπνου"] >= 2 and user["Ποιότητα Ύπνου"] < 5:
        sp = 0.8
    elif user["Ποιότητα Ύπνου"] < 4:

```

```

    sp = 1

    if row["Βεράντα Θέα"] == "Διαθέτει":
        sp_1 = 1
    else:
        sp_1 = 0

    if row["Θέρμανση /Κλιματισμός"] == "Διαθέτει":
        sp_2 = 1
    else:
        sp_2 = 0

    if row["Διαστάσεις δωματίου(m2)"] > 30:
        sp_3 = 1
    elif row["Διαστάσεις δωματίου(m2)"] > 20 and row["Διαστάσεις
    δωματίου(m2)"] <= 30:
        sp_3 = 0.8
    else:
        sp_3 = 0.5

    ranking = [row["Name"], (l_1 + l_2) / 2, (c + row["Clean"] / 10) / 2, (w + tv
    + clima + fridge + s_b) / 5, (
        p + bar + breakfast + f_c) / 4, (value + row["Value"] / 10) / 2, (sp
    + sp_1 + sp_2 + sp_3) / 4, row["Staff"] / 10,
        row["Total"] / 10]

    #((l_1 + l_2) / 2 + c + (w + tv + clima + fridge + s_b) / 5 + (
        #p + bar + breakfast + f_c) / 4 + value + (sp + sp_1 + sp_2 +
    sp_3) / 4 + row["Staff"] +
        #
        row["Total"] / 8]

```

```

df.loc[index] = ranking

return df

////////////////////////////////////

conversion

//////////

import pandas

def conversion(df):

    df_1 = pandas.DataFrame(columns=['hotel', 'location', 'cleanness',
    'roomservices', 'hotelservices', 'value', 'comfort', 'staff', 'reputation'])

    for index, row in df.iterrows():

        global location, cleanness, roomservices, hotelservices, value, comfort, staff,
        reputation, overall

        if row['location'] <= 0.7:

            location = 1

        elif row['location'] > 0.7 and row['location'] <= 0.9:

            location = 2

        else:

            location = 3

        if row['cleanness'] <= 0.39:

            cleanness = 1

        elif row['cleanness'] > 0.39 and row['cleanness'] < 0.65:

            cleanness = 2

        elif row['cleanness'] >= 0.66 and row['cleanness'] <= 0.94:

```



```

        cleanness = 3
else:
    cleanness = 4

if row['roomservices'] < 0.59:
    roomservices = 1
elif row['roomservices'] >= 0.6 and row['roomservices'] < 0.8:
    roomservices = 2
elif row['roomservices'] >= 0.8 and row['roomservices'] <= 0.92:
    roomservices = 3
else:
    roomservices = 4

if row['hotelservices'] <= 0.5:
    hotelservices = 1
elif row['hotelservices'] > 0.5 and row['hotelservices'] <= 0.7:
    hotelservices = 2
elif row['hotelservices'] > 0.7 and row['hotelservices'] < 0.9:
    hotelservices = 3
else:
    hotelservices = 4

if row['value'] <= 0.5:
    value = 1
elif row['value'] > 0.5 and row['value'] <= 0.7:
    value = 2
elif row['value'] > 0.7 and row['value'] <= 0.8:
    value = 3

```

```
elif row['value'] > 0.8 and row['value'] <= 0.9:
    value = 4
else:
    value = 5

if row['comfort'] <= 0.6:
    comfort = 1
elif row['comfort'] > 0.6 and row['comfort'] <= 0.85:
    comfort = 2
else:
    comfort = 3

if row['staff'] <= 0.6:
    staff = 1
elif row['staff'] > 0.6 and row['staff'] <= 0.86:
    staff = 2
elif row['staff'] > 0.86 and row['staff'] <= 0.95:
    staff = 3
else:
    staff = 4

if row['reputation'] <= 0.65:
    reputation = 1
elif row['reputation'] > 0.65 and row['reputation'] <= 0.87:
    reputation = 2
elif row['reputation'] > 0.87 and row['reputation'] <= 0.95:
    reputation = 3
else:
```

```
reputation = 4
```

```
ranking = [row["hotel"], location, cleanness, roomservices, hotelservices,  
value, comfort, staff, reputation]
```

```
df_1.loc[index] = ranking
```

```
return df_1
```

Global Utilities

```
//////////
```

```
import pandas
```

```
def xrisimotites(hotels):
```

```
    df_2 = pandas.DataFrame(columns=['hotel', 'Ug']) #'location', 'cleanness',  
'roomservices', 'hotelservices', 'value', 'comfort', 'staff', 'reputation'])
```

```
for index, row in hotels.iterrows():
```

```
    global u1g1, u2g2, u3g3, u4g4, u5g5, u6g6, u7g7, u8g8
```

```
    if row['location'] == 1:
```

```
        u1g1 = 0
```

```
    elif row['location'] == 2:
```

```
        u1g1 = 1
```

```
    else:
```

```
        u1g1 = 1
```

```
    if row['cleanness'] == 1:
```

```
        u2g2 = 0
```

```
    elif row['cleanness'] == 2:
```

```

    u2g2 = 0.3533
elif row['cleanness'] == 3:
    u2g2 = 1
else:
    u2g2 = 1

if row['roomservices'] == 1:
    u3g3 = 0
elif row['roomservices'] == 2:
    u3g3 = 0.5834
elif row['roomservices'] == 3:
    u3g3 = 0.5834
else:
    u3g3 = 1

if row['hotelservices'] == 1:
    u4g4 = 0
elif row['hotelservices'] == 2:
    u4g4 = 0
elif row['hotelservices'] == 3:
    u4g4 = 0
else:
    u4g4 = 1

if row['value'] == 1:
    u5g5 = 0
elif row['value'] == 2:
    u5g5 = 0

```

```
elif row['value'] == 3:
```

```
    u5g5 = 0.8828
```

```
elif row['value'] == 4:
```

```
    u5g5 = 0.8828
```

```
else:
```

```
    u5g5 = 1
```

```
if row['comfort'] == 1:
```

```
    u6g6 = 0
```

```
elif row['comfort'] == 2:
```

```
    u6g6 = 1
```

```
else:
```

```
    u6g6 = 1
```

```
if row['staff'] == 1:
```

```
    u7g7 = 0
```

```
elif row['staff'] == 2:
```

```
    u7g7 = 1
```

```
elif row['staff'] == 3:
```

```
    u7g7 = 1
```

```
else:
```

```
    u7g7 = 1
```

```
if row['reputation'] == 1:
```

```
    u8g8 = 0
```

```
elif row['reputation'] == 2:
```

```
    u8g8 = 0
```

```
elif row['reputation'] == 3:
```

```

        u8g8 = 0

    else:

        u8g8 = 1

    ug = row["hotel"], 0.4469 * u1g1 + 0.0531 * u2g2 + 0.075 * u3g3 + 0.0313
    * u4g4 + 0.3469 * u5g5 + 0.0219 * u6g6 + 0 * u7g7 + 0.025 * u8g8

    df_2.loc[index] = ug

    print(ug)

    return df_2

```

rw_utils

```

//////////

```

```

import pandas

```

```

import xlswriter

```

```

import openpyxl

```

```

import xlrd

```

```

def read_excel(source_file_name, columns):

    df = pandas.read_excel(source_file_name)

    # Parse just the desired columns

    df_selected = df[columns]

    return df_selected

```

```

def write_results(df, user):

    # Create an pandas excel writer based on xlswriter.

    writer = pandas.ExcelWriter('assets/results.xlsx', engine='xlswriter')

```

```
# Write the dataframe to another worksheet.
```

```
df.to_excel(writer, sheet_name=user)
```

```
writer.save()
```

2. Ερωτηματολόγιο

Έρευνα απαιτήσεων πελατών για την επιλογή ξενοδοχείου

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

Το παρών ερωτηματολόγιο εκπονείται στα πλαίσια διπλωματικής εργασίας και ερευνά τις απαιτήσεις και τις προτιμήσεις των πελατών για την επιλογή Ξενοδοχειακού Καταλύματος. Σκοπός είναι η συλλογή δεδομένων για την ανάπτυξη ενός συστήματος σύστασης για τουριστικούς προορισμούς. Ο χρόνος συμπλήρωσης δεν ξεπερνάει τα 2 λεπτά και η συμβολή σας θα είναι σημαντική. Επισημαίνεται ότι σε καμία περίπτωση δεν θα αναγράφονται τα προσωπικά δεδομένα των συμμετεχόντων. Σας ευχαριστώ πολύ εκ των προτέρων για τον χρόνο σας.

* Απαιτείται

1. Φύλλο

☐ Άντρας ☐ Γυναίκα

Ηλικία

Η απάντησή σας

2. Σκοπός Ταξιδιού

☐ Επαγγελματικό ☐ Διακοπές

3. Με ποιόν ταξιδεύετε;

☐ Μόνος ☐ Οικογένεια

4. Χώρα προέλευσης

Η απάντησή σας

5. Τόπο που επιθυμείτε να επισκεφτείτε

Η απάντησή σας

6. Χρονική περίοδος ταξιδιού (Από:)

Ημερομηνία

Χρονική περίοδος ταξιδιού (Έως:)

Ημερομηνία

7. Τι απόσταση (m) προτιμάτε να έχει το ξενοδοχείο από τα σημεία ενδιαφέροντος σας; *

☐ 0-400 ☐ 400-900 ☐ 900-πάνω

8. Τι απόσταση (m) προτιμάτε να έχει το ξενοδοχείο από τα Μέσα Μαζικής Μεταφοράς; *

☐ 0-150 ☐ 150-350 ☐ 350-πάνω

9. Πόσα χρήματα (ευρώ) είστε διατεθειμένοι να ξοδέψετε για κάθε διανυκτέρευση (τιμή δωματίου); *

☐ έως 50 ☐ Από 51 έως 75 ☐ Από 76 έως 100

☐ Από 101 έως 130 ☐ Πάνω από 131

10. Πρόσβαση Ίντερνετ (WiFi) *

☐ Δωρεάν ☐ Με Πληρωμή ☐ Όχι

11. Πάρκινγκ *

☐ Δωρεάν ☐ Με Πληρωμή ☐ Όχι

12. Πρωινό Γεύμα *

☐ Στην Τιμή ☐ Με Πληρωμή ☐ Όχι

**Πόσο σας επηρεάζει το κάθε ένα από τα παρακάτω
χαρακτηριστικά/απαιτήσεις σας στην επιλογή ξενοδοχείου;**

1. Air Condition *

Βαθμολογείτε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Τηλεόραση *

Βαθμολογείτε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Πρωινό Γεύμα *

Βαθμολογείτε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Πάρκινγκ *

Βαθμολογείτε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Πρόσβαση Ίντερνετ (WiFi) *

Βαθμολογείτε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Ψυγείο/Καφετιέρα *

Βαθμολογείστε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Safety Deposit Box (Χρηματοκιβώτιο) *

Βαθμολογείστε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Ποιότητα Καθαρισμού *

Βαθμολογείστε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Ποιότητα Ύπνου *

Βαθμολογείστε τα παρακάτω προϊόντα σε δεκαβάθμια κλίμακα (όπου το 10 δηλώνει την μεγαλύτερη προτίμηση)

1	2	3	4	5	6	7	8	9	10
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>