



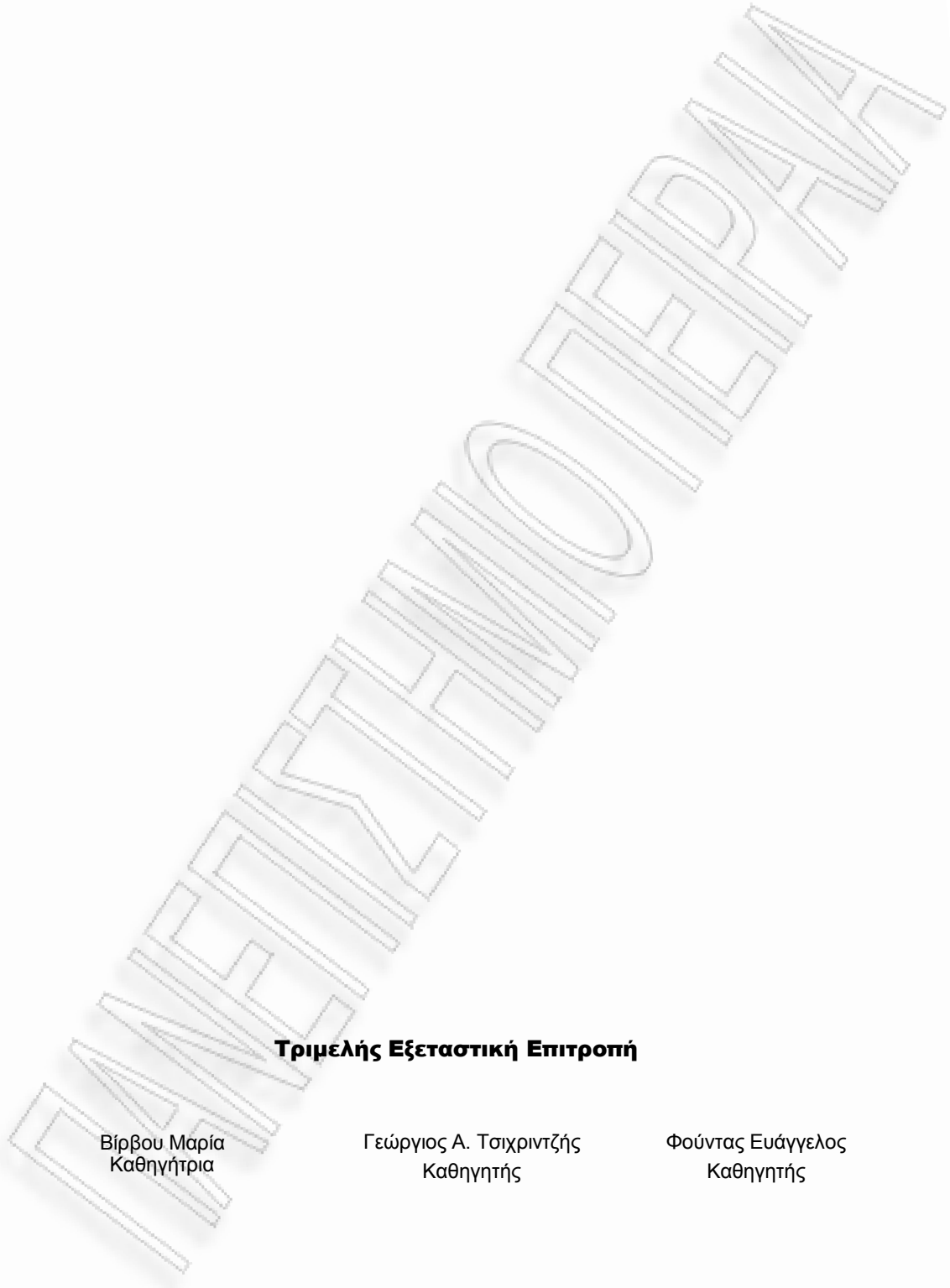
Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής  
Πρόγραμμα Μεταπτυχιακών Σπουδών  
«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	<b>Προσαρμοστικό σύστημα δημιουργίας προτάσεων σε ηλεκτρονικό κατάστημα.</b>
Όνοματεπώνυμο Φοιτητή	<b>Κοντοδιός Ηρακλής του Ελευθερίου</b>
Αριθμός Μητρώου	<b>ΜΠΣΠ/07036</b>
Κατεύθυνση	<b>Τεχνολογία Ενσωματωμένων Υπολογιστικών Συστημάτων</b>
Επιβλέπων	<b>Μαρία Βίβου , Καθηγήτρια</b>

Πανεπιστήμιο Πειραιώς-Τμήμα Πληροφορικής  
Πρόγραμμα Μεταπτυχιακών Σπουδών στα  
Προηγμένα Συστήματα Πληροφορικής

Ημερομηνία Παράδοσης **Απρίλιος 2011**



**Τριμελής Εξεταστική Επιτροπή**

Βίρβου Μαρία  
Καθηγήτρια

Γεώργιος Α. Τσιχριντζής  
Καθηγητής

Φούντας Ευάγγελος  
Καθηγητής

.....  
Ηρακλής Κοντοδιδς  
Πτυχιούχος Μηχανικός Αυτοματισμού

Copyright © Ηρακλής Ε. Κοντοδιδς, 2011  
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πειραιά.

## Περίληψη

Τα τελευταία χρόνια έχει παρατηρηθεί ότι το ηλεκτρονικό εμπόριο κερδίζει συνεχώς έδαφος σε παγκόσμιο επίπεδο. Ολοένα και περισσότερες επιχειρήσεις εκμεταλλεύονται το γεγονός ότι μια μεγάλη μερίδα του πληθυσμού έχει πρόσβαση στο διαδίκτυο και κατασκευάζουν ηλεκτρονικά καταστήματα για την προβολή και πώληση των προϊόντων τους. Ο όγκος πληροφορίας που δέχεται ένας άνθρωπος καθημερινά ολοένα και αυξάνεται, έτσι είναι δύσκολο να σταχυολογήσει την χρήσιμη από την περιττή πληροφορία με αποτέλεσμα να αποθαρρύνεται από το να χρησιμοποιήσει αυτόν τον τρόπο εμπορίου. Για το λόγο αυτό όλο και περισσότεροι ερευνητές τα τελευταία χρόνια έχουν στρέψει την προσοχή τους στο να δημιουργήσουν συστήματα τα οποία βοηθούν τους χρήστες στην τελική τους επιλογή με τα λεγόμενα (recommender systems) Επίσης η ανάγκη για προσαρμοστικό λογισμικό που διαμορφώνεται αναλόγως των απαιτήσεων του χρήστη είναι επιτακτική.

Τα συστήματα δημιουργίας συστάσεων-προτάσεων προϊόντων, χρησιμοποιούνται όλο και περισσότερο από τα ηλεκτρονικά καταστήματα, σκοπός των οποίων είναι η διευκόλυνση των καταναλωτών για εύρεση προϊόντων. Τα συστήματα αυτά μαθαίνουν από τις συνήθειες των καταναλωτών, είτε αυτές είναι αγορές, βαθμολογίες ή επισκέψεις και τους προτείνουν προϊόντα τα οποία θα τους φανούν περισσότερο χρήσιμα.

Σκοπός της μεταπτυχιακής διατριβής είναι η δημιουργία ενός προσαρμοστικού ηλεκτρονικού καταστήματος με χρήση συστημάτων δημιουργίας συστάσεων-προτάσεων(recommender systems), για την αύξηση του ενδιαφέροντος των καταναλωτών.

## Λέξεις Κλειδιά

Ηλεκτρονικό κατάστημα, Μοντελοποίηση χρηστών, Ενοποιημένη γλώσσα μοντελοποίησης, Εξόρυξη δεδομένων, Προσαρμοστικό σύστημα, σύστημα συστάσεων, διαγράμματα uml, αλγόριθμοι ομαδοποίησης, συνδυαζόμενες αγορές, στοχευμένη διαφήμιση, php, mysql, javascript, html.

### **Abstract**

The last years it has been observed that e-commerce is gaining ground worldwide. Much more companies' take advantage the fact that a large amount of population has access to network and build e-shops for projection and sale their products.

Its raising day by day the amount of information people accepts. For this reason it is hard to see the difference between the useful and useless information. So they are discouraged to use that kind of commerce. For this reason the last years much more researchers have turned their attention creating systems that help users to choose recommender systems. The need for adaptive software which can be formed according to user's demands is imperative.

The recommender systems are used increasingly from electronic stores, which their purpose is to help consumers find more valuable and suitable products for them. The systems are learning from the consumers habits, either these are purchases, ratings or visits and they suggest them products which there will be much more useful.

The purpose of this master thesis is the creation of an adaptive e-shop using a recommender system to propose products in order to raise consumers' interest.

### **KeyWords**

E-commerce, e-shop, user modeling, unified modeling language, data mining, adaptive system, recommender system, uml, cross-sales, targeted advertising, clustering, php, sql , JavaScript, Ajax, html.

### **Ευχαριστίες**

Θα ήθελα να ευχαριστήσω όλους όσους με βοήθησαν και με στήριξαν κατά την φοίτησή μου στο Μεταπτυχιακό Πρόγραμμα «Προηγμένα Συστήματα Πληροφορικής» του Πανεπιστημίου Πειραιά, την οικογένεια μου, τους φίλους μου, τους καθηγητές και τους συμφοιτητές μου.

Ιδιαίτερα, θα ήθελα να ευχαριστήσω την καθηγήτρια μου, κ. Βίρβου Μαρία που μου έδωσε την ευκαιρία να αναπτύξω αυτή τη διπλωματική εργασία, μέσα από την οποία απέκτησα ειδικευμένες γνώσεις στο αντικείμενο αυτό.

Επίσης θα ήθελα να ευχαριστήσω ιδιαίτερα τον φίλο και συμφοιτητή μου Ρενιέρη Λουκά για την παραχώρηση γνώσεων και πηγαίου κώδικα για την εκπόνηση της εργασίας μου.

Τέλος, οφείλω κυρίως να ευχαριστήσω τους γονείς μου όπου με στήριξαν και με στηρίζουν από την πρώτη στιγμή που αποφάσισα να κάνω αυτό το μεταπτυχιακό, τόσο οικονομικά, όσο και ψυχολογικά αλλά και για την κατανόηση τους όλο αυτό το διάστημα.

## Περιεχόμενα

1. ΕΙΣΑΓΩΓΗ .....	9
1.1 Αντικείμενο Διπλωματικής.....	9
1.2 Δομή Διπλωματικής.....	9
2. ΑΝΑΣΚΟΠΗΣΗ ΠΕΔΙΟΥ .....	10
2.1 Γενικοί ορισμοί.....	10
2.1.1 Ηλεκτρονικό εμπόριο( E-commerce).....	10
2.1.2 Recommender Systems (Συστήματα δημιουργίας προτάσεων).....	10
2.1.3 Adaptive Systems (Προσαρμοστικά συστήματα).....	10
2.2 Κατηγοριοποίηση Recommender Systems .....	10
2.2.1 Collaborative Filtering.....	10
2.2.2 Model-Based .....	11
2.2.3 Memory-Based .....	11
2.3 Προηγούμενη γνώση .....	12
2.3.1 Amazon.com .....	12
2.3.2 CDNOW.COM.....	13
2.3.3 Drugstore.com.....	16
2.3.4 eBay.....	16
2.3.5 MovieFinder.com.....	18
2.3.6 Recommendation Methods.....	18
3. ΑΝΑΛΥΣΗ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ.....	20
3.1 Εισαγωγή.....	20
3.2 Σχεδιασμός –ανάλυση απαιτήσεων.....	20
3.2.1 Λειτουργικές Απαιτήσεις .....	21
3.2.2 Μη Λειτουργικές Απαιτήσεις .....	23
3.3 Διαγράμματα περιπτώσεων χρήσης (Use case diagrams).....	23
3.3.1 Διάγραμμα χρήσης Σύνδεσης και Εγγραφής χρήστη .....	24
3.3.2 Διάγραμμα χρήσης Περιήγησης και Αναζήτησης .....	25
3.3.3 Διάγραμμα χρήσης Ολοκλήρωσης αγοράς (Check-Out).....	26
3.3.4 Διάγραμμα χρήσης διαχείρισης καλαθιού αγοράς.....	26
3.3.5 Διάγραμμα χρήσης διαχείρισης προτάσεων .....	27
3.3.6 Διάγραμμα χρήσης Δημιουργίας μοντέλου χρηστών .....	28
3.3.7 Διάγραμμα πακέτων υποσυστημάτων.....	30
4. ΠΕΡΙΓΡΑΦΗ ΕΦΑΡΜΟΓΗΣ .....	30
4.1 Πλατφόρμα και Προγραμματιστικά Εργαλεία.....	31
4.1.1 Επιλογή Γλώσσας Προγραμματισμού .....	31
4.1.2 Επιλογή Λειτουργικού Συστήματος .....	33

4.1.3 Επιλογή Εργαλείου Ανάπτυξης Κώδικα και σχεδιασμού βάσης δεδομένων .....	33
4.2 Περιγραφή Βασικών λειτουργιών του e-shop .....	34
4.2.1 Εγγραφή στο ηλεκτρονικό μας κατάστημα.....	34
4.2.2 Προσθήκη προϊόντος στο καλάθι αγορών.....	35
4.2.3 Παρακολούθηση επίσκεψης προϊόντων .....	36
4.2.4 Βαθμολόγηση προϊόντων .....	37
4.2.5 Βαθμολόγηση του ηλεκτρονικού καταστήματος.....	38
4.2.6 Αναζήτηση προϊόντος.....	38
4.3 Περιγραφή εμφάνισης προϊόντων σε δευτερεύουσες σελίδες .....	40
4.3.1 Cross-Sales.....	40
4.3.2 Δημοφιλέστερα προϊόντα ανά υπέρ-κατηγορία.....	41
4.4 Περιγραφή εμφάνισης προϊόντων στην κεντρική σελίδα .....	42
4.4.1 Προϊόντα που δεν βασίζονται στη μοντελοποίηση χρηστών. ....	43
4.4.2 Προϊόντα που βασίζονται στη μοντελοποίηση χρηστών. ....	45
4.5 Εμφάνιση τρόπων προσαρμοστικότητας του ηλεκτρονικού καταστήματος .....	50
4.5.1 Αλλαγή σειράς κατηγοριών στο κεντρικό μενού.....	50
4.5.2 Αυτόματη εμφάνιση ή μη λεπτομερειών ενός προϊόντος.....	51
5. ΧΡΗΣΗ ΑΛΓΟΡΙΘΜΩΝ – ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ .....	52
5.1 Χρήση αλγορίθμων ιεραρχικής ομαδοποίησης.....	52
5.1.1 Η μέθοδος του ελαχίστου (Min ή Single Linkage) .....	52
5.1.2 Η μέθοδος Μέσου όρου ομάδας (Average Linkage) .....	53
5.1.3 Η μέθοδος πλήρους συνδεσιμότητας (Max ή Complete Linkage) .....	54
5.1.4 Η μέθοδος Ward.....	54
5.2 Χρήση αποτελεσμάτων αλγορίθμου.....	55
5.2.1 Δημιουργία βοηθητικού πίνακα αγορών από clustering.....	58
5.2.2 Δημιουργία βοηθητικού πίνακα επισκέψεων από clustering .....	58
5.2.3 Δημιουργία προσωρινού πίνακα βαθμολογίας προϊόντων από clustering.....	59
5.3 Εξόρυξη δεδομένων ανά ομάδα χρηστών .....	60
5.3.1 Εύρεση κατηγορίας με τις περισσότερες αγορές. ....	60
5.3.2 Εύρεση κατηγορίας με τις μεγαλύτερη επισκεψιμότητα.....	61
5.3.3 Εύρεση προϊόντων με τη μεγαλύτερη βαθμολογία.....	61
6. ΕΦΑΡΜΟΓΗ ΠΡΟΤΑΣΕΩΝ .....	61
6.1 Προτάσεις ομαδοποιημένων χρηστών (από Clustering) .....	62
6.1.1 Cross Sales(ομαδοποιημένες αγορές της ομάδας του χρήστη). ....	62
6.1.2 Κατηγορία με τις περισσότερες αγορές της ομάδας του χρήστη.....	63
6.1.3 Κατηγορία με την μεγαλύτερη επισκεψιμότητα της ομάδας του χρήστη.....	63
6.1.4 Προϊόν με τη μεγαλύτερη βαθμολογία της ομάδας του χρήστη. ....	64
6.1.5 Νεότερο προϊόν της κατηγορίας με της περισσότερες αγορές της ομάδας του χρήστη. .....	64
6.2 Προτάσεις από το σύνολο των χρηστών (ανεξαρτήτου clustering).....	64



6.2.1 Cross Sales(ομαδοποιημένες αγορές όλων των χρηστών).....	64
6.2.2 Κατηγορία με τις περισσότερες αγορές όλων των χρηστών.....	65
6.2.3 Κατηγορία με την μεγαλύτερη επισκεψιμότητα όλων των χρηστών.....	66
6.2.4 Προϊόν με τη μεγαλύτερη βαθμολογία με συμμετοχή όλων των χρηστών. ....	66
6.2.5 Νεότερο προϊόν της κατηγορίας με της περισσότερες αγορές όλων των χρηστών. ....	66
7. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ .....	67
7.1 Συνεισφορά.....	67
7.2 Συμπεράσματα.....	67
7.3 Μελλοντικές Επεκτάσεις - Βελτιώσεις.....	68
7.3.1 Μελλοντικές Επεκτάσεις .....	68
7.3.2 Βελτιώσεις.....	68
Βιβλιογραφία.....	69

## 1. ΕΙΣΑΓΩΓΗ

### 1.1 Αντικείμενο Διπλωματικής

Στόχος της παρούσας διατριβής είναι η δημιουργία ενός προσαρμοστικού ηλεκτρονικού καταστήματος με τη χρήση συστημάτων δημιουργίας συστάσεων-προτάσεων (recommender systems), για την αύξηση του ενδιαφέροντος των καταναλωτών. Τα συστήματα αυτά μαθαίνουν από τις συνήθειες των καταναλωτών, είτε αυτές είναι αγορές, βαθμολογίες ή επισκέψεις προϊόντων και τους προτείνουν προϊόντα τα οποία θα τους φανούν περισσότερο χρήσιμα. Επίσης το σύστημα προσαρμόζεται ανάλογα με τις συνήθειες των καταναλωτών για την διευκόλυνση τους στην περιήγηση του ηλεκτρονικού καταστήματος.

Για την υλοποίηση της παραπάνω εφαρμογής χρειάστηκε να σχεδιαστούν και να χρησιμοποιηθούν τα παρακάτω συστήματα:

1. Σύστημα συλλογής δεδομένων και παρακολούθησης των συνηθειών των καταναλωτών. Ποια είδη πληροφοριών του χρήστη είναι χρήσιμα;
2. Εφαρμογή του κατάλληλου αλγόριθμου ομαδοποίησης των χρηστών βάσει συνηθειών.
3. Εκμετάλλευση των αποτελεσμάτων που απορρέουν μετά την εφαρμογή του αλγόριθμου (Πώς μπορούν αυτές οι πληροφορίες να χρησιμοποιηθούν;).
4. Δημιουργία προσαρμοστικού συστήματος με αλλαγές στον τρόπο περιήγησης. (Adaptive system)
5. Δημιουργία συστήματος προτάσεων προϊόντων ξεχωριστά για κάθε ομάδα χρηστών. (Recommender System).

Βάση της βιβλιογραφικής ανασκόπησης που πραγματοποιήθηκε και της αναζήτησης υπαρχόντων ηλεκτρονικών καταστημάτων με ενσωμάτωση recommender system διαπιστώθηκε ότι τα εργαλεία, οι μέθοδοι και τα αποτελέσματα της ερευνητικής εργασίας προσδίδουν μια νέα δυναμική στο τομέα του ηλεκτρονικού εμπορίου. Πιο συγκεκριμένα η χρήση της μοντελοποίησης χρηστών, εξόρυξης δεδομένων και εφαρμογής αλγορίθμων ομαδοποίησης για την υλοποίηση έξυπνων ηλεκτρονικών καταστημάτων με τρόπο αποτελεσματικό και ευέλικτο, δίνει τη δυνατότητα να αναφερόμαστε σε υποδομές που παρέχουν εξελιγμένες υπηρεσίες προς το χρήστη / πελάτη λαμβάνοντας υπόψη το προφίλ του και τα ιδιαίτερα χαρακτηριστικά του που εξάγονται από τη συμμετοχή του χρήστη στη διαδικασία περιήγησης, αγοράς ή ψήφισης προϊόντων.

### 1.2 Δομή Διπλωματικής

Η παρούσα διπλωματική αποτελείται από πέντε κεφάλαια. Στα κεφάλαια αυτά αναλύεται το θεωρητικό πλαίσιο στο οποίο στηρίζεται ο σχεδιασμός και η υλοποίηση του έξυπνου ηλεκτρονικού καταστήματος το οποίο ενσωματώνει ένα σύστημα αυτόματης προσαρμογής.

Το **πρώτο κεφάλαιο** αφορά την εισαγωγή η οποία δίνει μια σύντομη περιγραφή του συστήματος που υλοποιείται, εξηγεί τις ανάγκες που πρέπει να καλυφθούν και παρουσιάζει την δομή της διπλωματικής εργασίας.

Στο **δεύτερο κεφάλαιο** γίνεται η ανασκόπηση πεδίου με αναφορά προηγούμενων γνώσεων και άλλων εφαρμογών που χρησιμοποιούν ίδια συστήματα με αυτά τις διπλωματικής.

Στο **τρίτο κεφάλαιο** γίνεται η ανάλυση και ο σχεδιασμός του ηλεκτρονικού καταστήματος με τη βοήθεια της τεχνολογίας λογισμικού. (Γίνεται χρήση της UML)

Στο **τέταρτο κεφάλαιο** γίνεται η περιγραφή της εφαρμογής, καθώς και των εργαλείων που χρησιμοποιήθηκαν για την υλοποίησή της.

Στο **πέμπτο κεφάλαιο** περιγράφεται η δοκιμή τεσσάρων αλγορίθμων ομαδοποίησης καθώς και η εκμετάλλευση των αποτελεσμάτων μετά την εφαρμογή τους.

Στο **έκτο κεφάλαιο** περιγράφεται ο τρόπος με τον οποίο εφαρμόζονται τα διάφορα είδη προτάσεων προς τους χρήστες του ηλεκτρονικού καταστήματος.

Το **έβδομο κεφάλαιο** που αποτελεί και τον επίλογο της διπλωματικής παρατίθενται τα συνολικά συμπεράσματα που εξήχθησαν και οι μελλοντικές βελτιώσεις της εφαρμογής.

## **2. ΑΝΑΣΚΟΠΗΣΗ ΠΕΔΙΟΥ**

### **2.1 Γενικοί ορισμοί**

#### **2.1.1 Ηλεκτρονικό εμπόριο( E-commerce)**

Ως ηλεκτρονικό εμπόριο ορίζεται το εμπόριο που πραγματοποιείται με ηλεκτρονικά μέσα, αποτελεί δηλαδή μια ολοκληρωμένη συναλλαγή που πραγματοποιείται μέσω διαδικτύου - internet χωρίς να είναι απαραίτητη η φυσική παρουσία των συμβαλλομένων μερών (δηλαδή του πωλητή και του αγοραστή).

Τα σχετικά με το ηλεκτρονικό εμπόριο ζητήματα ρυθμίζονται από το Π.Δ. 131/2003 με το οποίο επήλθε προσαρμογή της ελληνικής νομοθεσίας προς την Οδηγία 31/2000/ΕΚ.[54]

#### **2.1.2 Recommender Systems (Συστήματα δημιουργίας προτάσεων)**

Το σύστημα που παράγει προτάσεις για προϊόντα ή υπηρεσίες προς τους επισκέπτες ή πελάτες μιας επιχείρησης τους οποίους πιθανολογείται ότι τους ενδιαφέρουν, περιγράφεται με τον όρο Recommender System.

Τα recommender systems προϊόντων έχουν ως σκοπό τη διευκόλυνση των καταναλωτών για εύρεση προϊόντων. Τα συστήματα αυτά μαθαίνουν από τις συνήθειες των καταναλωτών, είτε αυτές είναι αγορές, βαθμολογίες ή επισκέψεις και τους προτείνουν προϊόντα τα οποία θα τους φανούν περισσότερο χρήσιμα.

#### **2.1.3 Adaptive Systems (Προσαρμοστικά συστήματα)**

Με την έννοια προσαρμοστικό λογισμικό εννοούμε το λογισμικό που παρέχει αυτόματη προσαρμογή στις ανάγκες ή προτιμήσεις του χρήστη βασιζόμενο σε τεχνικές μοντελοποίησης χρηστών. Επιπλέον μπορεί να παρέχει ευφυή βοήθεια πιο ευέλικτη από την βοήθεια που παρέχουν τα μη προσαρμοστικά προγράμματα. Τα προσαρμοστικά συστήματα λογισμικού δομούν ένα μοντέλο χρήστη με τους στόχους, τις προτιμήσεις και τις γνώσεις του, και χρησιμοποιούν αυτό το μοντέλο με σκοπό να προσαρμοστούν στις ανάγκες αυτού του χρήστη.

## **2.2 Κατηγοριοποίηση Recommender Systems**

Σε αυτό το κεφάλαιο θα εξετάσουμε σε ποιες κατηγορίες διακρίνονται τα Recommender Systems με βάση διάφορες παραμέτρους. Τα Recommender Systems διαχωρίζονται σε τρεις κύριες κατηγορίες, τα Collaborative Filtering Systems, τα Content Based Systems και τα Rule Based Systems.

### **2.2.1 Collaborative Filtering**

Τα συστήματα collaborative filtering (CF) είναι η διαδικασία του φιλτραρίσματος για πληροφορίες ή σχέδια χρησιμοποιώντας τεχνικές που απαιτούν συνεργασία μεταξύ πολλών παραγόντων, απόψεων, πηγές δεδομένων, κλπ. Οι εφαρμογές του collaborative filtering συνήθως περιλαμβάνουν πολύ μεγάλα σύνολα δεδομένων.

Οι μέθοδοι του Collaborative έχουν εφαρμοστεί σε πολλά διαφορετικά είδη δεδομένων, συμπεριλαμβανομένων της ανίχνευσης και παρακολούθησης δεδομένων – όπως είναι η εξερεύνηση μετάλλων, η περιβαλλοντική ανίχνευση πάνω από μεγάλες περιοχές ή η πολλαπλή πληροφόρηση, οικονομικά στοιχεία - όπως είναι τα χρηματοπιστωτικά ιδρύματα παροχής

υπηρεσιών που ενσωματώνουν πολλές πηγές χρηματοδότησης ή στο ηλεκτρονικό εμπόριο και τις εφαρμογές του Web 2.0, όπου η έμφαση δίνεται στα δεδομένα του χρήστη κλπ. Το υπόλοιπο αυτής της συζήτησης εστιάζεται στο collaborative filtering για τα δεδομένα του χρήστη, αν και ορισμένες από τις μεθόδους και τις προσεγγίσεις μπορεί να χρησιμοποιηθούν και για τις άλλες μεγάλες εφαρμογές.

Το Collaborative filtering είναι μια μέθοδος λήψης αυτόματων προβλέψεων (φιλτράρισμα) που αφορά τα ενδιαφέροντα του χρήστη με τη συλλογή πληροφοριών σχετικά με το γούστο από πολλούς χρήστες (collaborating).

Η βασική παραδοχή της προσέγγισης CF είναι ότι αυτοί που συμφώνησαν κατά το παρελθόν τείνουν να συμφωνήσουν και πάλι στο μέλλον.

Για παράδειγμα, ένα collaborative filtering ή ένα recommendation system που αφορά τα γούστα στη τηλεόραση, μπορεί να κάνει προβλέψεις για το ποιό σόου στην τηλεόραση μπορεί να αρέσει σε ένα χρήστη δίνοντας μια μεροληπτική λίστα για τα γούστα του χρήστη (αρέσει ή δεν αρέσει).

Να σημειώσουμε ότι αυτές οι προβλέψεις είναι συγκεκριμένες στο χρήστη, αλλά έχουν χρησιμοποιηθεί πληροφορίες προερχόμενες από πολλούς χρήστες.

Αυτό διαφέρει από την απλή προσέγγιση που δίνει ένα μέσο σκορ (όχι συγκεκριμένο) για κάθε προϊόν ενδιαφέροντος, όπως για παράδειγμα με βάση των αριθμό των ψήφων .

### 2.2.2 Model-Based

Αυτός ο μηχανισμός χρησιμοποιεί την βαθμολογία του χρήστη για να υπολογίσει την ομοιότητα ανάμεσα σε χρήστες και προϊόντα. Αυτός χρησιμοποιείται για να γίνονται οι προτάσεις. Αυτός ήταν ο αρχικός μηχανισμός και χρησιμοποιείται σε πολλά εμπορικά συστήματα. Είναι εύκολο να εφαρμοστεί και είναι αποτελεσματικός. Τυπικά παραδείγματα αυτού του μηχανισμού είναι το neighborhood που βασίζεται στο CF και προϊόν-που βασίζεται/ χρήστης- που βασίζεται στις top-N- προτάσεις.

Ο neighborhood-based αλγόριθμος υπολογίζει την ομοιότητα ανάμεσα σε δύο χρήστες ή προϊόντα, έτσι παράγει μια πρόβλεψη για τον χρήστη παίρνοντας τον σταθμισμένο μέσο όρο από όλες τις βαθμολογίες. Ο όμοιος υπολογισμός ανάμεσα σε προϊόντα ή χρήστες είναι ένα σημαντικό κομμάτι αυτής της προσέγγισης. Πολλαπλοί μηχανισμοί όπως ο Pearson correlation και ο vector cosine με βάση την ομοιότητα χρησιμοποιούνται γι' αυτό.

Ο user based top-N recommendation αλγόριθμος εντοπίζει τα πιο όμοια k στους χρήστες και ενεργούς χρήστες χρησιμοποιώντας την ομοιότητα βασισμένη στο μοντέλο διανύσματος.

Μετά το k βρίσκονται περισσότεροι χρήστες που μοιάζουν, η αντιστοιχία χρήστη-προϊόντος matrices συγκεντρώνονται για να αναγνωρίσουν το σύνολο των προϊόντων που θα προταθεί.

### 2.2.3 Memory-Based

Αυτός ο μηχανισμός χρησιμοποιεί την βαθμολογία του χρήστη για να υπολογίσει την ομοιότητα ανάμεσα σε χρήστες και προϊόντα. Αυτός χρησιμοποιείται για να γίνονται οι προτάσεις. Αυτός ήταν ο αρχικός μηχανισμός και χρησιμοποιείται σε πολλά εμπορικά συστήματα. Είναι εύκολο να εφαρμοστεί και είναι αποτελεσματικός. Τυπικά παραδείγματα αυτού του μηχανισμού είναι το neighborhood που βασίζεται στο CF και προϊόν-που βασίζεται/ χρήστης- που βασίζεται στις top-N- προτάσεις.

ο neighborhood-βασικός αλγόριθμος υπολογίζει την ομοιότητα ανάμεσα σε δύο χρήστες ή προϊόντα, έτσι παράγει μια πρόβλεψη για τον χρήστη παίρνοντας τον σταθμισμένο μέσο όρο από όλες τις βαθμολογίες.

Ο όμοιος υπολογισμός ανάμεσα σε προϊόντα ή χρήστες είναι ένα σημαντικό κομμάτι αυτής της προσέγγισης. Πολλαπλοί μηχανισμοί όπως ο Pearson correlation και ο vector cosine με βάση την ομοιότητα χρησιμοποιούνται γι' αυτό.

Ο αλγόριθμος εντοπίζει τα πιο όμοια k στους χρήστες και ενεργούς χρήστες χρησιμοποιώντας την ομοιότητα βασισμένη στο μοντέλο vector.

Μετά το  $k$  βρίσκονται περισσότεροι χρήστες που μοιάζουν, η αντιστοιχία χρήστη-προϊόντος matrices συγκεντρώνονται για να αναγνωρίσουν το σύνολο των προϊόντων που θα προταθεί.

## 2.3 Προηγούμενη γνώση

Σε αυτή την ενότητα παρουσιάζουμε μερικά από τα καταστήματα ηλεκτρονικού εμπορίου που χρησιμοποιούν μερικές παραλλαγές της τεχνολογίας των recommender systems.

Τα recommender systems χρησιμοποιούνται από τα καταστήματα ηλεκτρονικού εμπορίου για να προτείνουν τα προϊόντα τους στους πελάτες και να παρέχουν στους καταναλωτές πληροφορίες για να τους βοηθήσουν πια προϊόντα να διαλέξουν. Τα προϊόντα που προτείνονται μπορεί να είναι αυτά με τις περισσότερες αγορές, ανάλογα με τα δημογραφικά στοιχεία του κάθε καταναλωτή ή μετά από ανάλυση των παλιών του αγοραστικών συνηθειών, να γίνονται προβλέψεις για τις μελλοντικές του αγορές. Οι μορφές των προτάσεων μπορεί να περιλαμβάνουν εξατομικευμένες πληροφορίες ενός προϊόντος, τη γνώμη της κοινότητας μέσα από την αναφορά κριτικών. Σε γενικές γραμμές, αυτές οι τεχνικές συστάσεων αποτελούν μέρος της εξατομίκευσης ενός ιστότοπου γιατί το βοηθούν να προσαρμοστεί σε κάθε πελάτη.

### 2.3.1 Amazon.com

Θα επικεντρωθούμε εδώ για συστήματα συστάσεων που αφορούν το τμήμα βιβλίου του Amazon.com.

Συνδυαζόμενες αγορές: Όταν βρισκόμαστε στη σελίδα ενός βιβλίου για να δούμε τα χαρακτηριστικά του, αξιολογήσεις, περίληψη κλπ παρατηρούνται δύο τρόποι για την πρόταση προϊόντων. Κατά τον πρώτο τρόπο προτείνονται βιβλία που συνήθως αγοράζονται μαζί με το βιβλίο της τρέχουσας σελίδας. Ενώ ο κατά τον δεύτερο τρόπο προτείνονται συγγραφείς των οποίων τα βιβλία αγοράζονται από πελάτες που έχουν αποκτήσει το συγκεκριμένο βιβλίο.

#### Customers Who Bought This Item Also Bought



Castle Vroman (A Galaxy Unknown, Book 6) by Thomas DePrima

★★★★☆ (48)

\$9.19



Milor! (A Galaxy Unknown, Book 5) by Thomas DePrima

★★★★☆ (40)

\$9.19

Εικόνα 2.1: Συνδυαζόμενες αγορές του amazon.com.

Συστάσεις πελατών: Το amazon ενθαρρύνει την άμεση ανατροφοδότηση από τους πελάτες για τα βιβλία που έχουν ήδη αναγνώσει. Οι πελάτες βαθμολογούν τα βιβλία που έχουν ήδη αναγνώσει με κλίμακα από τον 1 έως το 5 και να αποθηκεύεται αυτή η πληροφορία στο σύστημα. Όταν ο πελάτης θα ζητήσει να του γίνει κάποια πρόταση το σύστημα θα του φέρει σαν αποτέλεσμα βιβλία που ανήκουν στις κατηγορίες που είχε βαθμολογήσει με 5.

**Eyes:** Το χαρακτηριστικό επωνομαζόμενο και ως “Eyes” επιτρέπει στους χρήστες να ενημερώνονται μέσω e-mail για τις καινούριες κυκλοφορίες που προστέθηκαν στον κατάλογο του amazon.com. Εδώ οι πελάτες συμπληρώνουν αιτήσεις με βάση τον συγγραφέα, τίτλο, θέμα ISBN, ή ημερομηνία κυκλοφορίας.

**Amazon.com Delivers:** Είναι μία παραλλαγή του χαρακτηριστικό “eyes”. Οι πελάτες επιλέγουν πλαίσια ελέγχου για να επιλέξουν από μια λίστα των ειδικών κατηγοριών / ειδών (Οprah βιβλία, βιογραφίες, μαγείρεμα). Κατά διαστήματα οι εκδότες του Amazon.com στέλνουν τελευταίες συστάσεις τους μέσω ηλεκτρονικού ταχυδρομείου σε συνδρομητές σε κάθε κατηγορία.

**Ιδέες δώρων του βιβλιοπωλείου:** Το χαρακτηριστικό αυτό επιτρέπει στους πελάτες να λαμβάνουν συστάσεις από τους εκδότες. Οι πελάτες διαλέγουν μια κατηγορία βιβλίων για τα οποία θα ήθελαν κάποιες προτάσεις. Με την πλοήγηση στο εν λόγω τμήμα “Gift Department” μπορούν να δουν έναν γενικό κατάλογο με συστάσεις που δημιουργήθηκε από τους συντάκτες του amazon.com. Επίσης οι πελάτες μπορούν να λαμβάνουν συστάσεις ανώνυμα, αφού δεν χρειάζεται η εγγραφή στην ιστοσελίδα.

**Σχόλια πελατών:** Αυτό το χαρακτηριστικό επιτρέπει στους χρήστες να λαμβάνουν συστάσεις κειμένου βασισμένο στις απόψεις των άλλων πελατών. Οι συστάσεις αυτές βρίσκονται στη σελίδα πληροφοριών ενός βιβλίου, οι οποίες εμφανίζονται με την βαθμολογία και τα γραμμένα σχόλια που παρέχονται από πελάτες που έχουν ήδη αποκτήσει το βιβλίο. Οι πελάτες εάν επιθυμούν λαμβάνουν υπόψη τους σχόλια αυτά για την αγορά του βιβλίου. Επίσης έχουν την δυνατότητα να βαθμολογήσουν ένα σχόλιο επιλέγοντας με ένα ναι ή όχι στο ερώτημα “Σας βοήθησε αυτό το σχόλιο;” Τα αποτελέσματα απεικονίζονται στην εικόνα 2.2 παρακάτω.

### Customer Reviews



### Most Helpful Customer Reviews

11 of 12 people found the following review helpful:

★★★★★ **The cure for boredom**, July 7, 2002

Εικόνα 2.2: Κριτικές πελατών του amazon.com.

**Κύκλος Αγορών:** Επιτρέπει στους πελάτες για να δουν τον κατάλογο με τα δέκα δημοφιλέστερα βιβλία για μια συγκεκριμένη γεωγραφική περιοχή, η εταιρεία, εκπαιδευτικό ίδρυμα, η κυβέρνηση ή άλλη οργάνωση. Για παράδειγμα, ένας πελάτης θα μπορούσε να ζητήσει να δει τι βιβλία είναι τα best sellers για τους πελάτες της Oracle, MIT, ή τους κατοίκους της Νέας Υόρκης. Επίσης οι πελάτες έχουν τη δυνατότητα να ενταθούν σε έναν τομέα αναγνωστών ώστε να λαμβάνουν πιο εξατομικευμένες συστάσεις από το amazon.com

### 2.3.2 CDNOW.COM

**Σύμβουλος αγοράς άλμπουμ:** Ο σύμβουλος αγοράς άλμπουμ του CDNOW λειτουργεί με τρεις διαφορετικούς τρόπους. Οι πρώτοι δυο είναι παρόμοιοι με τις συνδυαζόμενες αγορές του amazon.com. Όταν βρίσκεται ο πελάτης στην σελίδα πληροφοριών ενός άλμπουμ το σύστημα τότε προτείνει δέκα άλμπουμ που συνδέονται με τον καλλιτέχνη του άλμπουμ της σελίδας. Τα αποτελέσματα εμφανίζονται όπως απεικονίζεται στην εικόνα 2.2 και εικόνα 2.3.

## Customers Who Bought This Item Also Bought



Hard Bargain ~ Emmylou Harris  
\$11.99



So Beautiful or So What (CD / DVD Deluxe) ~ Paul Simon  
★★★★★ (10)  
\$12.99

Εικόνα 2.2: Συνδυαζόμενες αγορές στο cdnow.com (πρώτος τρόπος)

## Customers Also Bought Items By



[Emmylou Harris](#)



[Sierra Hull](#)

Εικόνα 2.3: Συνδυαζόμενες αγορές στο cdnow.com (δεύτερος τρόπος)

Ο τρίτος τρόπος δουλεύει σαν σύμβουλος δώρων, όπου ο πελάτης πληκτρολογεί έως τρία ονόματα καλλιτεχνών και το σύστημα του επιστρέφει μια λίστα από δέκα άλμπουμ που θεωρούνται παρόμοια με των καλλιτεχνών που πληκτρολογήθηκαν.

**Σχετικοί καλλιτέχνες :** Αυτό το χαρακτηριστικό του cdnow.com λειτουργεί με βάση την υπόθεση ότι εάν ένας πελάτης ενδιαφέρεται για έναν συγκεκριμένο καλλιτέχνη, τότε επιλέγοντας το σύνδεσμο “σχετικοί καλλιτέχνες”(related artists), λαμβάνει μια λίστα καλλιτεχνών που ανήκουν σε ομάδα με παρόμοιο στυλ μουσικής. Όπως επίσης εμφανίζονται και καλλιτέχνες οι οποίοι αποτέλεσαν έμπνευση για τον επιλεγμένο καλλιτέχνη.

## Related Artists on Tour [\(What's this?\)](#)



See 10 tour dates for Bon Jovi

[Click here for more info](#)



See 3 tour dates for Def Leppard

[Click here for more info](#)



See 33 tour dates for Journey

[Click here for more info](#)

Εικόνα 2.4: Το cdnow προτείνει σχετικούς καλλιτέχνες με τον επιλεγμένο που στο παράδειγμα μας είναι ο Bon Jovi.

**Οδηγός Αγοραστή:** Το χαρακτηριστικό επιτρέπει στους πελάτες να λαμβάνουν συστάσεις οι οποίες βασίζονται σε ένα συγκεκριμένο ύφος μουσικής. Εδώ οι πελάτες περιηγούνται σε μια λίστα με είδη μουσικής που παρέχονται από το cdnow. Επιλέγοντας έναν από τους συνδέσμους από τον κατάλογο αυτό μεταφέρει τους πελάτες σε ένα νέο κατάλογο των άλμπουμ τον οποίο οι συντάκτες του cdnow θεωρούν το ιδανικό για αυτό το είδος μουσικής.

**Artist Picks:** Προηγουμένως στο χαρακτηριστικό σχετικοί καλλιτέχνες, οι λίστες με τις συστάσεις είχαν δημιουργηθεί από τους συντάκτες του cdnow. Στο Artist Picks παρέχονται παρόμοιες συστάσεις, απευθείας από τους καλλιτέχνες. Κάθε εβδομάδα ένας διαφορετικός καλλιτέχνη χαρακτηρίζεται, ο οποίος απεικονίζεται με τα άλμπουμ που διαμόρφωσαν το γούστο του ή, επίσης ακούν σήμερα από το cd player τους.

### This Month's Spotlight Artists



[Sting](#)



[Arcade Fire](#)



[James Taylor](#)



[The Beach Boys](#)



[Ozzy Osbourne](#)



[Wolf Parade](#)

Εικόνα 2.5: Επιλεγμένοι καλλιτέχνες του μήνα που αναφέρουν τη γνώμη τους στο cdnow.com

**Top 100 Best seller:** Είναι η παραδοσιακή μέθοδος που έχει χρησιμοποιηθεί από τα περισσότερα ηλεκτρονικά καταστήματα. Εδώ δίνεται η δυνατότητα στον χρήστη να λάβει μία λίστα εκατό cd η οποία προέρχεται από τα cd με τις μεγαλύτερες πωλήσεις του cdnow. Αυτή η λίστα ανανεώνεται συνεχώς, καθώς οι πωλήσεις αλλάζουν με τον καιρό.

**My CDNOW:** Επιτρέπει στον πελάτη να δημιουργήσει το δικό του μουσικό κατάστημα, βασισμένο στα άλμπουμ και τους καλλιτέχνες που επιθυμεί. Οι πελάτες αναφέρουν ποια άλμπουμ κατέχουν και ποιοι καλλιτέχνες είναι οι αγαπημένοι τους. Οι αγορές από το CDNOW εγγράφονται απευθείας στην λίστα των αποκτηθέντων και όταν οι πελάτες ζητήσουν προτάσεις για το τι να αγοράζουν, τότε το σύστημα προβλέπει έξι άλμπουμ που έχουν την πιθανότητα να ενδιαφέρουν τον πελάτη βασισμένο στη λίστα των αποκτηθέντων.



**Returning customer?**  
[Sign in](#) to get your personalized recommendations.

**New customer?**  
**Make Recommended for You just for you!**  
 It's free and easy and requires no credit card or personal information. [Personalize now](#)

---

**Top Sellers**

- 

**Wasting Light**  
 ~ Foo Fighters (April 12, 2011)  
 Average Customer Review: ★★★★★ (35)  
 In Stock  
[Listen to samples](#)  
**Price: \$9.99**  
 48 used & new from \$9.00 [Add to Cart](#) [Add to Wish List](#)
- 

**Paper Airplane**  
 ~ Alison Krauss and Union Station (April 12, 2011)  
 Average Customer Review: ★★★★★ (9)  
 In Stock  
[Listen to samples](#)  
**Price: \$9.99**  
 37 used & new from \$7.99 [Add to Cart](#) [Add to Wish List](#)

Εικόνα 2.6: Το CDNOW προτείνει την εγγραφή του χρήστη για την προσαρμογή του ηλεκτρονικού καταστήματος και την παροχή προτάσεων για αγορά.

### 2.3.3 Drugstore.com

**Σύμβουλος:** Αυτό το χαρακτηριστικό επιτρέπει στους πελάτες να αναφέρουν τις προτιμήσεις τους όταν αγοράζουν ένα προϊόν από μια κατηγορία όπως “αντηλιακή προστασία” ή “διορθωτικά μέτρα για κρύωμα και γρίπη”. Για παράδειγμα, στην περίπτωση που κάποιος έχει κρύωμα, ο πελάτης αναφέρει τα συμπτώματα (π.χ. καταρροή και το φτέρνισμα) και την ηλικία του ασθενή που πρέπει να χορηγηθεί η το προϊόν. Αφού δοθούν όλα τα στοιχεία στον σύμβουλο τότε επιστρέφεται μια λίστα με προϊόντα που πληρούν τις παραπάνω προϋποθέσεις.

**Test Drives:** Μια ομάδα εθελοντών οι οποία αποτελείται από πελάτες του καταστήματος λαμβάνει τα καινούρια προϊόντα και με τη σειρά τους τα βαθμολογούν και γράφουν κριτικές για αυτά.

### 2.3.4 eBay

**Feedback Profile:** Επιτρέπει στους πωλητές και αγοραστές να συμβάλλουν στην ανατροφοδότηση των προφίλ άλλων πελατών με τους οποίους συναλλάσσονται. Η ανατροφοδότηση γίνεται βαθμολογώντας με κριτήριο τον βαθμό ικανοποίησης της συναλλαγής, παρέχοντας και ένα συγκεκριμένο σχόλιο. Το σύστημα συστάσεων παρέχεται στους αγοραστές οι οποίοι έχουν την δυνατότητα παρακολούθησης του προφίλ κάθε πωλητή. Το προφίλ περιέχει βαθμολογίες τελευταίων ημερών, μηνών και το συνολικό μέσο όρο. (Εικόνα 2.7)

**Watch list:** Επιτρέπει στον πελάτη να αναφέρει το ενδιαφέρον του για κάποιο προϊόν με συγκεκριμένα χαρακτηριστικά, όπως τιμή και για το χρονικό διάστημα που τον ενδιαφέρει να το παρακολουθεί. Το σύστημα κάνει αναζητήσεις ανά τακτά χρονικά διαστήματα και ενημερώνει τον πελάτη στέλνοντας του e-mail με τα αποτελέσματα των αναζητήσεων που εκτελεί κάθε φορά. (Εικόνα 2.8)


**Top-rated seller**

**region\_free\_gamez** ( 2398 ★ )

99.6% Positive feedback

- ✓ Consistently receives highest buyers' ratings
- ✓ Ships items quickly
- ✓ Has earned a track record of excellent service

Save this seller  
See other items

Visit store:  [regionfreegamez](#)

Positive Feedback: 99.6%

Feedback score: 2398

[\[How is Feedback calculated?\]](#)


**Detailed Seller Ratings** (last 12 months) ?

Criteria	Average rating	Number of ratings
Item as described	★★★★★	1784
Communication	★★★★★	1785
Shipping time	★★★★★	1782
Shipping and handling charges	★★★★★	1786

Εικόνα 2.6: Παρατηρούμε το feedback profile που αναφέρθηκε στην παράγραφο αυτή.

**Buying Reminders**


(Last 31 days)

 1 item I'm watching is ending soon.

**Watch List (1)**

◀ **All (1)** | Active (1) | Ended (0)

---

 **ENDING SOON**  
**Sony PlayStation 3 Slim 160 GB Black Console w/ games**  
 salyak1126 ( 52 ★ ) 100.0%

[Add note](#) [Delete](#) [Compare](#) [Add to list](#)

Page 1 of 1

Εικόνα 2.7: Παρατηρούμε τη λίστα παρακολούθησης(watch list) που αναφέρθηκε στην παράγραφο αυτή.

### 2.3.5 MovieFinder.com

Βαθμολογία πελατών/συντακτών: Δίνεται η δυνατότητα στους πελάτες να βαθμολογήσουν τις ταινίες που έχουν δει με βαθμούς (A-F). Επίσης οι ταινίες βαθμολογούνται από τους συντάκτες με τον ίδιο τρόπο και οι πελάτες καθώς επισκέπτονται τη σελίδα μιας ταινίας αντικρίζουν τις δυο αυτές βαθμολογίες, το οποίο τους βοηθάει στην επιλογή της αγοράς ή μη.

Top 10 : το Top 10 χαρακτηριστικό στο E online επιτρέπει στους πελάτες να λαμβάνουν προτάσεις από τους εκδότες στην κατηγορία της επιλογής τους. Οι πελάτες διαλέγουν μια κατηγορία από τις λίστες από προηγούμενες συγκεκριμένες κατηγορίες όπως chick flicks, και ταινίες από βιβλία. Διαλέγοντας μια λίστα ο πελάτης λαμβάνει περιγραφές από τις top 10 ταινίες Στην συγκεκριμένη κατηγορία όπως έχει οριστεί από έναν εκδότη του E.

### 2.3.6 Recommendation Methods

Στην προηγούμενη ενότητα επικεντρωθήκαμε στα στοιχεία που χρησιμοποιούνται και παράγονται κατά τη διαδικασία πρότασης προϊόντων-υπηρεσιών. Στο τμήμα αυτό παρέχουμε μια άλλη πλευρά της συγκεκριμένης διαδικασίας που χρησιμοποιήθηκε σε πραγματικές recommender systems του ηλεκτρονικού εμπορίου. Θα πρέπει να επισημάνουμε ότι μεμονωμένα τα συστήματα μπορούν να χρησιμοποιούν έναν συνδυασμό αυτών των διαδικασιών.

Κάθε κατηγορία που εξετάζεται εδώ αντιπροσωπεύει μια οικογένεια αλγορίθμων και προσεγγίσεων. Είναι πέρα από το πεδίο εφαρμογής του παρόντος έργου να ερευνήσουμε τις διάφορες υλοποιήσεις κάθε διαδικασίας. Ο Breese et al. (1998) συγκρίνουν μια ποικιλία από αλγόριθμους για την παραγωγή σύστασης, και ο Herlocker et al. (1999) παρέχει λεπτομερή σύγκριση αλγορίθμων συσχέτισης από πελάτη σε πελάτη.

Raw retrieval systems "null recommender" παρέχει στους πελάτες μια διεπαφή αναζήτησης μέσω των οποίων μπορούν να κάνουν ερώτηση σε μια βάση δεδομένων από στοιχεία.

Σε αυτή την περίπτωση, η σύσταση είναι "δυσιακή", συντακτική διαδικασία με την οποία το σύστημα "συνιστά" ό, τι ο πελάτης έχει ζητήσει. Αν και δεν είναι τεχνικά μια εφαρμογή συστάσεων, μια τέτοια αίτηση μπορεί να εμφανιστεί στους πελάτες. Για παράδειγμα, όταν ένας πελάτης ζητά από ένα μουσικό site άλμπουμ από τους "The Beatles", το σύστημα δίνει μια λίστα των άλμπουμ των Beatles που μπορεί να είναι χρήσιμη και μπορεί να οδηγήσουν πράγματι τον πελάτη σε ένα άλμπουμ το οποίο πριν δεν γνώριζε. Τα raw retrieval systems είναι πανταχού παρόντα στις εφαρμογές του ηλεκτρονικού εμπορίου.

Οι εφαρμογές που εκτιμούν την προσωπικότητα πέρα από την εξατομίκευση μπορούν να δημιουργήσουν τα σύνολα συστάσεων που έχουν επιλεχτεί χειροκίνητα από τους επεξεργαστές, τους καλλιτέχνες, τους κριτικούς, και άλλους ειδικούς. Αυτοί οι "ανθρώπινοι recommenders" προσδιορίζουν τα αντικείμενα που βασίζονται στις προτιμήσεις, τα ενδιαφέροντα, και τους στόχους τους και δημιουργούν έναν κατάλογο συνιστώμενων αντικειμένων διαθέσιμων στα μέλη της κοινότητας.

Οι συστάσεις τους συνοδεύονται συχνά από σχόλια κειμένων που βοηθούν άλλους πελάτες να αξιολογήσουν και να καταλάβουν τη σύσταση. Παραδείγματος χάριν, οι πελάτες που χρησιμοποιούν τη MovieFinder Top 10 lists επιλέγουν ένα ιδιαίτερο «φύλλο» για το οποίο θα ήθελαν τις συστάσεις, παραδείγματος χάριν «chick flicks». τους παρέχεται λοιπόν ένας κατάλογος που μεταγλωττίζεται χειροκίνητα από έναν εκδότη αναφέροντας τι θεωρεί σαν top 10 chick flicks του συγκεκριμένου χρόνου. Η διαδικασία δεν χρησιμοποιεί καθόλου τον υπολογισμό υπολογιστών, αλλά απλά αναπαράγει τι θα μπορούσε να εμφανιστεί σε έναν κατάλογο σε οποιοδήποτε τον τόχο οποιοδήποτε τηλεοπτικού καταστήματος. Αυτή η διαδικασία ο περισσότερο μιμείται τους παραδοσιακούς κριτικούς και τους εκδότες, και συμπεριλαμβανόμενης της πιθανής διορατικότητας και της πιθανής πόλωσης. Αν και δεν περιλαμβάνεται στα παραδείγματά μας, ένας αυξανόμενος αριθμός sites επιτρέπει σε οποιοδήποτε μέλος μιας κοινότητας να καθιερώσει καταλόγους σύστασης (Art.com).

Σε περιπτώσεις όπου η εξατομίκευση είναι μη πρακτική ή περιττή, εφαρμογές recommender μπορούν πολύ αποτελεσματικά να παρέχουν στατιστικές περιλήψεις της άποψης μιας

κοινότητας. Αυτές οι περιλήψεις περιλαμβάνουν μέσα σε μια κοινότητα τα μέτρα δημοτικότητας (π.χ., το ποσοστό των ανθρώπων που αγοράζει ή του αρέσει ένα αντικείμενο) και τις συνολικές ή συνοπτικές εκτιμήσεις (π.χ., αριθμός ανθρώπων που συστήνουν ένα αντικείμενο, μέσος όρος βαθμολογίας για ένα αντικείμενο). Περιλαμβάνουν συστήματα όπως το feedback πελατών του eBay, η οποία παρέχει τις μέσους όρους βαθμολογιών για τους αγοραστές και για τους πωλητές.

Οι ενδεχόμενοι πωλητές και οι αγοραστές μπορούν να συμβουλευθούν το μέσο όρο και τις μεμονωμένες αξιολογήσεις αλλά δεν μπορούν να δουν την βαθμολογία για "τους πελάτες που έχω συμφωνήσει με." Ενώ αυτές οι περιλήψεις δίνουν μόνο μη προσωποποιημένες συστάσεις, είναι δημοφιλείς επειδή είναι εύκολο να υπολογιστούν, και μπορούν να χρησιμοποιηθούν στα μη-προσαρμοσμένα περιβάλλοντα όπως μπορεί να φανεί σε ένα φυσικό κατάστημα.

Συστάσεις που βασίζονται στις συντακτικές ιδιότητες των αντικειμένων και των ενδιαφερόντων των πελατών χρησιμοποιούν τεχνολογίες που βασίζονται στις ιδιότητες των συστάσεων. Αν και η απλούστερη ιδιότητα-βασισμένη στη σύσταση είναι ακατέργαστη ανάκτηση, οι αληθινί "recommenders" που χρησιμοποιούν τα πρότυπα ιδιότητας πελατών ενδιαφέρονται πέρα από ένα απλό ερωτηματολόγιο. Παραδείγματος χάριν, ένας πελάτης που ψάχνει στο τμήμα "country μουσικής" ενός καταστήματος μουσικής και που έχει πολλά "\$9,99 special" compact disc στο καλάθι αγορών του μπορεί να λάβει συστάσεις για εκπαιωτικά country CDs. Other attribute-based recommenders χρησιμοποιούν προφίλ πελατών που γράφουν αν τους αρέσει κάτι ή όχι έτσι ώστε να κάνουν συστάσεις στον πελάτη. Παραδείγματος χάριν, το ίδιο κατάστημα μουσικής μπορεί να μάθει ότι ένας συγκεκριμένος πελάτης αγοράζει μόνο εκπαιωτικά CDs ή ότι ένας άλλος πελάτης δεν αγοράζει ποτέ τη μουσική από τη δεκαετία του '70.

Άλλες εφαρμογές χρησιμοποιούν το συσχετισμό αντικείμενο-προς-αντικείμενο για να προσδιορίσουν τα αντικείμενα που βρίσκονται συχνά σε συνδυασμό με τα αντικείμενα στα οποία ένας πελάτης έχει εκφράσει ενδιαφέρον. Ο συνδυασμός αυτός μπορεί να βασιστεί στα στοιχεία συνδυασμένων αγορών, προτίμηση από κοινούς πελάτες, ή άλλα μέτρα. Στην απλούστερη εφαρμογή του, ο συσχετισμός αντικείμενο-προς-αντικείμενο μπορεί να χρησιμοποιηθεί για να προσδιορίσει τα "αντικείμενα που ταιριάζουν" για ένα μόνο αντικείμενο, όπως άλλα ρούχα που αγοράζονται συνήθως με ένα ζευγάρι των εσωρούχων. Τα ισχυρότερα συστήματα ταιριάζουν ένα ολόκληρο σύνολο αντικειμένων, όπως αυτά σε ένα καλάθι αγορών ενός πελάτη, για να προσδιορίσουν τα κατάλληλα αντικείμενα για να συστήσουν. Οι recommender εφαρμογές που συσχετίζουν αντικείμενο-προς-αντικείμενο χρησιμοποιούν συνήθως τρέχουσες αγορές ή άλλα τρέχουσα ενδιαφέροντα παρά το μακροπρόθεσμο ιστορικό των πελατών, η οποία τις καθιστά κατάλληλες για να προτείνουν δώρα. Ένας πελάτης πρέπει μόνο να προσδιορίσει μερικά άλλα αντικείμενα που μπορεί να αρέσουν στον παραλήπτη έτσι ώστε να αποσπάσουν συστάσεις δώρων που προσαρμόζονται στον παραλήπτη παρά σε αυτόν που δίνει το δώρο.

Τέλος, τα συστήματα συστάσεων, χρησιμοποιώντας την χρήστη-προς-χρήστη αντιστοιχία συστήσουν προϊόντα σε έναν πελάτη με βάση το συσχετισμό μεταξύ του εν λόγω πελάτη και άλλων πελατών που έχουν αγοράσει προϊόντα από την ιστοσελίδα E-commerce. Η τεχνολογία αυτή καλείται συχνά "collaborative filtering", επειδή δημιουργήθηκε σαν τεχνική για το φιλτράρισμα της πληροφορίας που χρησιμοποιείται για τις απόψεις μιας ομάδας ώστε να προτείνει πληροφορίες για στοιχεία σε μεμονωμένα άτομα (al Resnick et, 1994?. Hill et al, 1995?. Shardanand και Maes, 1995? Konstan et al., 1997). Το CDNow είναι ένα σύστημα που χρησιμοποιεί τους χρήστη-προς-χρήστη συσχετισμούς για να προσδιορίσει μια κοινότητα πελατών που έχουν την τάση να κατέχουν και να τους αρέσει το ίδιο σετ CD. Η αρχή είναι ότι, αν ορισμένα μέλη της κοινότητάς μου έχουν και τους αρέσει το τελευταίο album του Sting, τότε είναι πολύ πιθανό ότι θα αρέσει και σε μένα. Αν και χρησιμοποιούμε την λέξη «συσχετισμός» στο όνομα αυτής της τεχνικής, έτσι παραπέμπουμε σε κοντινότερες-γείτονες-τεχνικές βασισμένες στην γραμμική συσχέτιση, η τεχνική μπορεί να εφαρμοστεί επίσης και με πολλές άλλες τεχνολογίες (Breese et al., 1998).

Ένα σημαντικό θέμα κατά την εξέταση της μεθόδου της σύστασης είναι το αν ο υπολογισμός γίνεται εξ' ολοκλήρου online, ενώ το κατάστημα Web βρίσκεται σε αλληλεπίδραση με τον πελάτη, ή αν κομμάτια του υπολογισμού πρέπει να γίνονται offline για λόγους απόδοσης. Οι online συστάσεις προτιμώνται διότι μπορούν να ανταποκριθούν άμεσα στις προτιμήσεις των

καταναλωτών. Τα περισσότερα από την διαδικασία των συστάσεων, που αναφέρονται παραπάνω μπορούν να γίνουν εξ ολοκλήρου online. Raw retrieval, χειροκίνητη επιλογή, Περιληπτικά στατιστικά, και attribute-based είναι όλα απλοί υπολογισμοί που εκτελούνται συνήθως κατά τη διάρκεια της αλληλεπίδρασης με τους πελάτες. Ο συσχετισμός αντικείμενο-προς-αντικείμενο και ο συσχετισμός χρήστη-προς-χρήστη είναι υπολογιστικά πιο εντατικό και συχνά απαιτεί μια σύνδεση offline για να προετοιμάσει ένα μοντέλο που μπορεί να εκτελεστεί αποτελεσματικά σε online.

Μια πρόκληση στο σχεδιασμό model-building είναι να εξασφαλίσει ότι το προκύπτον online σύστημα ανταποκρίνεται όσο το δυνατόν καλύτερα στη Interactive input από το χρήστη.

### 3. ΑΝΑΛΥΣΗ ΚΑΙ ΣΧΕΔΙΑΣΜΟΣ

#### 3.1 Εισαγωγή

Η πολυπλοκότητα και το μέγεθος, του προσαρμοστικού ηλεκτρονικού καταστήματος το οποίο χρησιμοποιεί ένα recommender system, έκριναν αναγκαία την χρήση της τεχνολογίας λογισμικού και της ενοποιημένης γλώσσας μοντελοποίησης(UML). Πριν το σχεδιασμό και την ανάλυση της διαδικτυακής αυτής εφαρμογής αξίζει να γίνει μια μικρή αναφορά σε αυτή τη μέθοδο σχεδιασμού πολύπλοκων λογισμικών.

Η τεχνολογία λογισμικού ως γνωστό αντικείμενο έχει ως στόχο τη βέλτιστη ανάπτυξη συστημάτων λογισμικού και έχει ως βασικό χαρακτηριστικό το συνδυασμό αρχών και γνώσεις από άλλα γνωστικά αντικείμενα όπως της επιστήμης των υπολογιστών, της διαχείρισης έργων, της ψυχολογίας, της διοίκησης και των μαθηματικών. Το εύρος των απαιτούμενων γνώσεων, σε συνδυασμό με την ποικιλία και την πολυπλοκότητα των σημερινών συστημάτων, κάνουν την ανάπτυξη του λογισμικού μια απαιτητική και σε πολλές περιπτώσεις μη προβλέψιμη διαδικασία.

Τα τελευταία χρόνια η κυρίαρχη προσέγγιση για την ανάπτυξη λογισμικού είναι η αντικειμενοστραφής προσέγγιση, σύμφωνα με την οποία η ανάπτυξη των συστημάτων βασίζεται στα αντικείμενα. Μερικά από τα πλεονεκτήματα της αντικειμενοστραφούς ανάπτυξης συστημάτων λογισμικού είναι η απλότητα στην ανάπτυξη, η καλύτερη ποιότητα συστημάτων, η δυνατότητα επαναχρησιμοποίησης των συστατικών αυτών και αξιοπιστία των παραγόμενων συστημάτων.

Στα πλαίσια αυτά, η ενοποιημένη διεργασία (unified process) έχει αναδυθεί ως μια δημοφιλή μέθοδος ανάπτυξης αντικειμενοστραφών συστημάτων λογισμικού, και η ενοποιημένη γλώσσα μοντελοποίησης (unified modeling language) ως την πλέον δημοφιλή και πρότυπη γλώσσα για την τεκμηρίωση και προσδιορισμό συστημάτων λογισμικού. Η ενοποιημένη διεργασία χρησιμοποιεί την ενοποιημένη γλώσσα μοντελοποίησης (UML) κατά τον σχεδιασμό ενός συστήματος λογισμικού. Συγκεκριμένα η UML αποτελεί αναπόσπαστο κομμάτι της ενοποιημένης διεργασίας αφού αναπτύχθηκαν παράλληλα. Σήμερα χιλιάδες εταιρίες χρησιμοποιούν τη μέθοδο της ενοποιημένης διεργασίας και την UML στην ανάπτυξη των συστημάτων τους. Λόγω των παραπάνω, κρίθηκε απαραίτητη η χρήση της UML για τον σχεδιασμό και την ανάλυση των απαιτήσεων της εφαρμογής που υλοποιήθηκε για την μεταπτυχιακή διατριβή.

#### 3.2 Σχεδιασμός –ανάλυση απαιτήσεων

Η διαδικασία ανάπτυξης λογισμικού χωρίζεται σε τρεις κύριες φάσεις οι οποίες κατά σειρά είναι:

- Συλλογή Απαιτήσεων
- Σχεδιασμός
- Υλοποίηση

Κατά την συλλογή των απαιτήσεων ενός συστήματος ορίζονται οι δραστηριότητες, οι κίνδυνοι και ένα δοκιμαστικό σχέδιο του συστήματος. Σε αυτή τη φάση δηλαδή ορίζεται εν συντομία τι πρέπει να κάνει το σύστημα. Πιο αναλυτικά στο στάδιο αυτό εστιάζουμε στο να

παρέχουμε μία μηχανική περιγραφή των αντικειμένων, των λειτουργιών και των καταστάσεων του συστήματος λογισμικού.

Ο σχεδιασμός του συστήματος καθορίζει το ΠΩΣ το σύστημα θα υλοποιήσει τις λειτουργίες οι οποίες περιγράφονται κατά την συλλογή των απαιτήσεων και γενικά ορίζει την δομή που θα έχει το σύστημα χωρίς ωστόσο να μπαίνει σε λεπτομέρειες υλοποίησης. Δηλαδή η σχεδίαση λογισμικού εστιάζει στον τρόπο με τον οποίο οι απαιτήσεις του λογισμικού, μπορούν να υλοποιηθούν.

Τέλος στη φάση της υλοποίησης παράγεται ο πηγαίος κώδικας και η τεκμηρίωση του συστήματος καθώς επίσης γίνονται και δοκιμαστικοί έλεγχοι του λογισμικού που έχει παραχθεί ώστε αυτό να επικυρωθεί.

### 3.2.1 Λειτουργικές Απαιτήσεις

#### 1. Εγγραφή νέου χρήστη (Sign-In)

**Περιγραφή:** Ο χρήστης ο οποίος θέλει να γίνει μέλος του ηλεκτρονικού καταστήματος, εισάγει τα απαραίτητα στοιχεία που απαιτούνται για να καταχωρηθεί στο σύστημα. Σε περίπτωση που όλα τα στοιχεία που θα εισάγει είναι σωστά, το σύστημα καταχωρεί τον χρήστη σαν νέο πελάτη και δημιουργεί ένα μοναδικό κωδικό για αυτόν. Αν το σύστημα εντοπίσει λάθος στην καταχώρηση των στοιχείων, τότε εμφανίζει το κατάλληλο μήνυμα λάθους.

**Είσοδος:** Στοιχεία χρήστη

**Επεξεργασία:** Το σύστημα ελέγχει εάν υπάρχει χρήστης με το ίδιο user-id, το οποίο επέλεξε ο νέος χρήστης. Εάν δεν υπάρχει και όλα τα υπόλοιπα στοιχεία είναι αποδεκτά, το σύστημα καταχωρεί τον πελάτη στην βάση δεδομένων και ενημερώνει τον χρήστη για την επιτυχή ολοκλήρωση της διαδικασίας. Στην αντίθετη περίπτωση, το σύστημα εμφανίζει κατάλληλο μήνυμα λάθους στον χρήστη.

**Έξοδος:** Νέα εγγραφή στον πίνακα πελατών ή μήνυμα λάθους.

#### 2. Εισαγωγή χρήστη στο κατάστημα (Log-In)

**Περιγραφή:** Ο χρήστης ο οποίος επιθυμεί να κάνει log-in στο ηλεκτρονικό κατάστημα πρέπει να εισάγει το αναγνωριστικό χρήστη (user-id) και το συνθηματικό (password). Το σύστημα προχωράει στην ταυτοποίηση των στοιχείων αυτών. Αν ο χρήστης έδωσε σωστά στοιχεία τότε συνδέεται στο σύστημα, αλλιώς εμφανίζει μήνυμα λάθους.

**Είσοδος:** Αναγνωριστικό χρήστη (user-id), Συνθηματικό χρήστη (password)

**Επεξεργασία:** Το σύστημα αναζητά στην βάση πελατών το αναγνωριστικό χρήστη που έδωσε ο χρήστης και ακολούθως κάνει την ταυτοποίηση του συνθηματικού. Αν η διαδικασία είναι επιτυχημένη τότε το σύστημα καλωσορίζει τον χρήστη και πλέον αυτός έχει πρόσβαση σε όλες τις υπηρεσίες του καταστήματος που είναι διαθέσιμες για τους εγγεγραμμένους χρήστες. Σε διαφορετική περίπτωση το σύστημα εμφανίζει στον χρήστη ένα μήνυμα λάθους και τον οδηγεί στο να επαναεισάγει τα ζητούμενα στοιχεία.

**Έξοδος:** Καταγραφή στο session state της εφαρμογής ότι ο χρήστης είναι πλέον συνδεδεμένος ή μήνυμα λάθους.

#### 3. Επεξεργασία στοιχείων χρήστη

**Περιγραφή:** Ο χρήστης ο οποίος επιθυμεί να αλλάξει τα στοιχεία σύνδεσης (password) ή τα προσωπικά του στοιχεία (όνομα, επώνυμο, διεύθυνση, e-mail κτλ) πρέπει προηγουμένως να έχει κάνει επιτυχή είσοδο στο σύστημα (log-in).

**Είσοδος:** Αναγνωριστικό χρήστη (user-id), Συνθηματικό χρήστη (password), προαιρετικά νέα προσωπικά στοιχεία [όνομα, επώνυμο, διεύθυνση, e-mail].

**Επεξεργασία:** Το σύστημα αναζητά στην βάση πελατών το αναγνωριστικό χρήστη που έδωσε ο χρήστης και ακολούθως κάνει την ταυτοποίηση του συνθηματικού. Αν η διαδικασία είναι επιτυχημένη τότε το σύστημα εμφανίζει στον χρήστη την φόρμα συμπλήρωσης των προσωπικών του στοιχείων. Αφού ο χρήστης συμπληρώσει τα επιθυμητά στοιχεία το σύστημα ελέγχει την ορθότητα τους και αν δεν υπάρχει πρόβλημα εμφανίζει μήνυμα επιτυχούς αλλαγής

των στοιχείων και στέλνει ενημερωτικό e-mail στον χρήστη. Αν τα στοιχεία έχουν κάποιο λάθος τότε εμφανίζει το ανάλογο μήνυμα λάθους και προτρέπει τον χρήστη να το διορθώσει.

Έξοδος: Ανανέωση του πίνακα πελατών στην βάση δεδομένων σε περίπτωση επιτυχημένης ολοκλήρωσης της διαδικασίας και αποστολή ενημερωτικού e-mail ή εμφάνιση μηνύματος λάθους και επανάληψη της διαδικασίας.

#### **4. Προσθήκη προϊόντος στο καλάθι αγοράς**

Περιγραφή: Ο χρήστης επιλέγει να προσθέσει ένα προϊόν στο καλάθι αγοράς.

Είσοδος: Αναγνωριστικό προϊόντος

Επεξεργασία: Το σύστημα ψάχνει αν το id του προϊόντος ώστε να διαπιστωθεί αν το προϊόν υπάρχει ήδη στο καλάθι αγοράς. Αν δεν υπάρχει ήδη, τότε ελέγχει την διαθεσιμότητα του προϊόντος και αν το προϊόν είναι διαθέσιμο ενημερώνει τον πίνακα basket στη βάση, αλλιώς ενημερώνει τον χρήστη. Αν το προϊόν έχει ήδη προστεθεί στο καλάθι, τότε ενημερώνει τους μετρητές για το προϊόν αυξάνοντας τον κατά 1.

Έξοδος: Ενημέρωση της βάσης και ανανέωση στην οθόνη του καλαθιού αγοράς ώστε να απεικονίζει τα σωστά προϊόντα ή ενημέρωση του χρήστη με μήνυμα ότι το προϊόν δεν είναι διαθέσιμο.

#### **5. Αφαίρεση προϊόντος από το καλάθι αγοράς**

Περιγραφή: Ο χρήστης επιλέγει να αφαιρέσει ένα προϊόν από το καλάθι αγοράς.

Είσοδος: Αναγνωριστικό προϊόντος (product\_id), ποσότητα (quantity)

Επεξεργασία: Το σύστημα αναζητά το αναγνωριστικό προϊόντος στον πίνακα basket της βάσης. Αν η ποσότητα του προϊόντος είναι 1 τότε αφαιρεί το προϊόν από το καλάθι αγοράς και απο τον πίνακα basket. Σε περίπτωση που η ποσότητα του προϊόντος είναι μεγαλύτερη του ενός, μειώνεται κατά την ποσότητα που απαίτησε ο χρήστης και ενημερώνεται κατάλληλα ο πίνακας basket.

Έξοδος: Ενημέρωση της βάσης με τις νέες τιμές για το καλάθι αγοράς και ανανέωση της οθόνης ώστε το καλάθι αγοράς να απεικονίζει τα νέα δεδομένα.

#### **6. Ολοκλήρωση Αγοράς προϊόντος (-ων) (Check-out)**

Περιγραφή: Ο χρήστης επιλέγει να ολοκληρώσει τις αγορές του πατώντας τον σύνδεσμο Check-Out.

Είσοδος: Αναγνωριστικό χρήστη (user\_id), συνθηματικό χρήστη (password)

Επεξεργασία: Για να ολοκληρώσει ο χρήστης τις αγορές του απαραίτητη προϋπόθεση είναι να έχει κάνει εισαγωγή στο σύστημα. Το σύστημα ελέγχει αν ο χρήστης έχει συνδεθεί και αν ναι τον ανακατευθύνει στην σελίδα όπου καλείται να επιβεβαιώσει την αγορά, αλλιώς τον παραπέμπει στη σελίδα σύνδεσης για να κάνει log-in.

Έξοδος: Το σύστημα ενημερώνει τους μετρητές των προϊόντων στην βάση δεδομένων και ενημερώνει τον χρήστη ότι η συναλλαγή ολοκληρώθηκε επιτυχώς με μήνυμα στην οθόνη και e-mail επιβεβαίωσης με τα στοιχεία της παραγγελίας. Σε περίπτωση που δεν είναι συνδεδεμένος του εμφανίζει το κατάλληλο μήνυμα που τον προτρέπει να συνδεθεί.

#### **7. Επιλογή προϊόντος**

Περιγραφή: Ο χρήστης κατά την περιήγησή του στο site επιλέγει ένα αντικείμενο κάνοντας κλικ επάνω του.

Είσοδος: Κλικ στον σύνδεσμο ή την εικόνα ενός προϊόντος (product\_id).

Επεξεργασία: Το σύστημα ανακτά το προϊόν που αντιστοιχεί στον σωστό κωδικό και εμφανίζει στον χρήστη λεπτομέρειες για το προϊόν, όπως τα χαρακτηριστικά του, μία σύντομη ή αναλυτική περιγραφή και εικόνες του προϊόντος (με δυνατότητα μεγέθυνσης). Τέλος ενημερώνονται οι μετρητές στην βάση δεδομένων που αντιστοιχούν στο προϊόν για τον συγκεκριμένο χρήστη.

Έξοδος: Ενημέρωση μετρητών επισκεψιμότητας προϊόντος στην βάση δεδομένων και εμφάνιση στην οθόνη των χαρακτηριστικών του προϊόντος.

#### **8. Ψήφιση προϊόντος**

Περιγραφή: Ο χρήστης επιθυμεί να αξιολογήσει ένα προϊόν σε μία κλίμακα 5 σημείων.

Είσοδος: product\_id, user\_id, rating

Επεξεργασία: Το σύστημα ελέγχει αν ο συγκεκριμένος χρήστης έχει ξαναβαθμολογήσει το προϊόν. Αν όχι τότε ενημερώνει τους μετρητές ψήφισης του προϊόντος και εμφανίζει στον χρήστη την μέση βαθμολογία του προϊόντος.

Έξοδος: Εμφάνιση της μέσης βαθμολογίας του προϊόντος και της βαθμολογίας που έδωσε ο χρήστης καθώς και ενημέρωση των μετρητών ψήφισης στην βάση δεδομένων.

### 9. Ψήφιση site

Περιγραφή: Ο χρήστης αξιολογεί το site σε μία κλίμακα 5 σημείων.

Είσοδος: user\_id, rating

Επεξεργασία: Το σύστημα ελέγχει αν ο χρήστης είναι συνδεδεμένος και αν δεν είναι τον παραπέμπει να συνδεθεί. Αφού ψηφίσει ενημερώνει την οθόνη με τον μέσο όρο βαθμολόγησης του site από όλους τους χρήστες και με την δική του βαθμολογία, όπως επίσης και τους μετρητές ψήφισης του site στην βάση δεδομένων.

Έξοδος: Εμφάνιση στην οθόνη της μέσης βαθμολογίας του site και της βαθμολογίας που έδωσε ο χρήστης καθώς και ενημέρωση των μετρητών στην βάση.

### 10. Αναζήτηση προϊόντος

Περιγραφή: Ο χρήστης ψάχνει για κάποιο προϊόν με συγκεκριμένα χαρακτηριστικά.

Είσοδος: περιγραφή, [cluster\_id], [κατηγορία]

Επεξεργασία: Το σύστημα ψάχνει την βάση να δει αν ο χρήστης ανήκει σε κάποιο cluster. Το σύστημα αναζητεί προϊόντα με βάση την περιγραφή και πιθανώς την κατηγορία που δόθηκε από τον χρήστη. Αν ο χρήστης ανήκει σε κάποιο cluster τότε τα προϊόντα που επιστρέφονται στον χρήστη ταξινομούνται με βάση της προτίμησης της ομάδας στην οποία ανήκει, αλλιώς επιστρέφονται ταξινομημένα με αλφαβητική σειρά. Προϊόντα που έχει ήδη αγοράσει ο χρήστης δεν παραλείπονται από τα αποτελέσματα της αναζήτησης.

Έξοδος: Λίστα προϊόντων που δημιουργείται βάσει των κριτηρίων αναζήτησης.

## 3.2.2 Μη Λειτουργικές Απαιτήσεις

### 1. Απαιτήσεις σχεδίασης συστήματος

Η σχεδίαση της εφαρμογής (δημιουργία UML διαγραμμάτων) θα γίνει με χρήση της εφαρμογής Rational Rose.

### 2. Βάση Δεδομένων

Η εφαρμογή να αντλεί και να αποθηκεύει τα δεδομένα που απαιτούνται, από μια βάση δεδομένων MySql η οποία να διαθέτει τους εξής πίνακες:

### 3. Απαιτήσεις Υλοποίησης εφαρμογής

Η κύρια γλώσσα προγραμματισμού που θα χρησιμοποιηθεί για την υλοποίηση της εφαρμογής να είναι η server-side scripting language PHP.

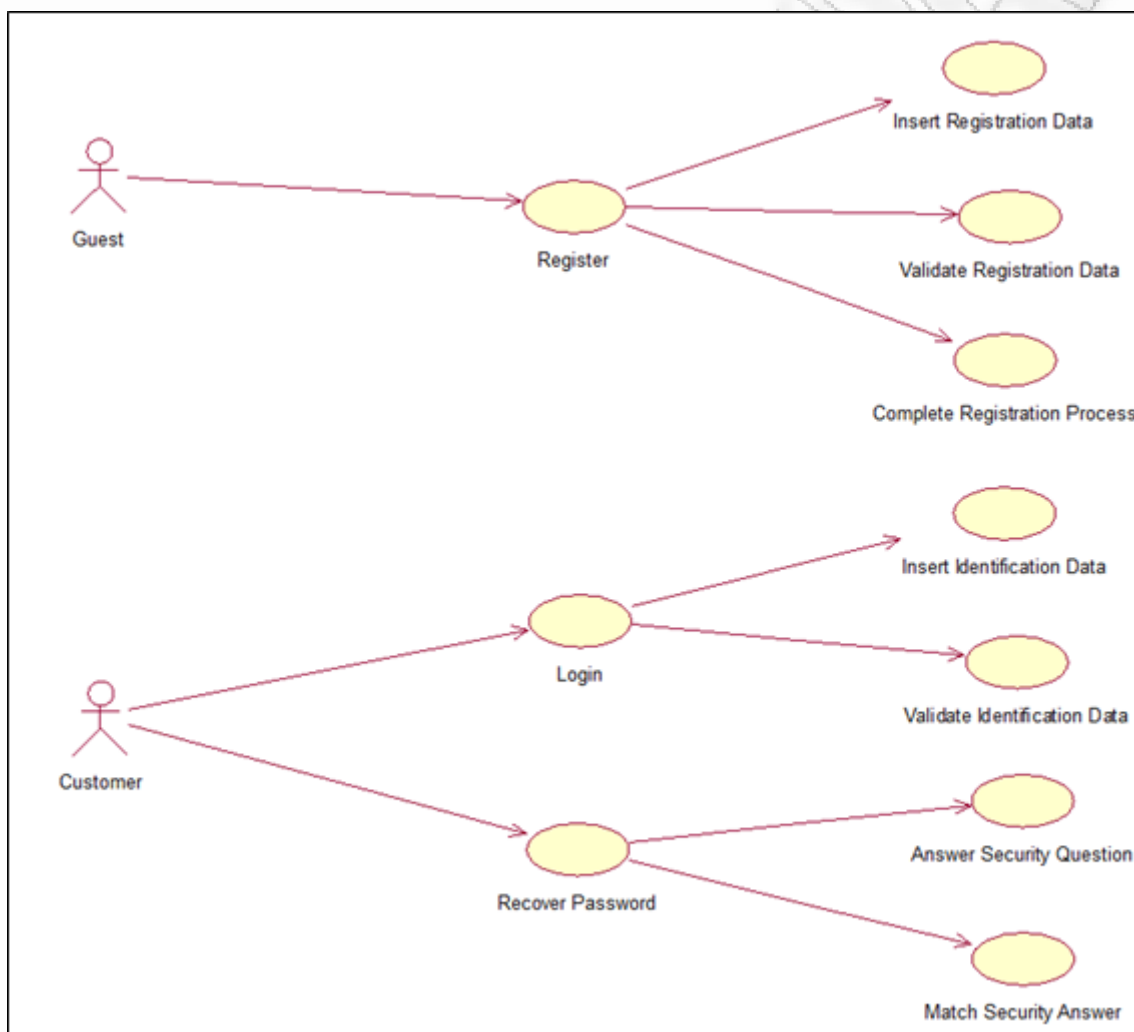
## 3.3 Διαγράμματα περιπτώσεων χρήσης (Use case diagrams)

Το κυριότερο ζήτημα στην επιτυχημένη κατασκευή ενός πληροφοριακού συστήματος είναι η κατασκευή του σωστού συστήματος, δηλαδή ενός συστήματος που να ικανοποιεί τις απαιτήσεις των χρηστών του. Δεν είναι λίγα τα πληροφοριακά συστήματα που αποτυγχάνουν να ικανοποιήσουν τον βασικότατο αυτόν στόχο, και αυτό οφείλεται στην αποτυχία του συστήματος να ικανοποιήσει τις προδιαγραφές του χρήστη. Τα μοντέλα περιπτώσεων χρήσης δίνουν έμφαση στην λειτουργικότητα ενός συστήματος, όπως αυτή μπορεί να γίνει αντιληπτή από τους χρήστες του. Μία περίπτωση χρήσης απεικονίζει στην ουσία μία συναλλαγή (αλληλεπίδραση) ενός χρήστη με το σύστημα. Ένα σενάριο αποτελείται από μια ακολουθία βημάτων τα οποία περιγράφουν την αλληλεπίδραση του χρήστη με το σύστημα. Στην ουσία ένα διάγραμμα



περίπτωσης χρήσης περιγράφει μια σειρά σεναρίων από τα οποία ένα μπορεί να έχει θετική έκβαση για τον χρήστη και τα υπόλοιπα προσπαθούν να περιγράψουν προβληματικές καταστάσεις στις οποίες μπορεί να βρεθεί ο χρήστης είτε λόγω αστοχίας του συστήματος είτε λόγω αβλεψίας του χρήστη. Σε κάθε περίπτωση η λειτουργία ενός συστήματος θα πρέπει να είναι απρόσκοπτη και σε αυτό βοηθάνε τα διαγράμματα περιπτώσεων χρήσης ώστε να προβλέψουμε όλες τις πιθανές κι απίθανες ενέργειες του χρήστη και να μπορούμε να τις διαχειριστούμε.

### 3.3.1 Διάγραμμα χρήσης Σύνδεσης και Εγγραφής χρήστη



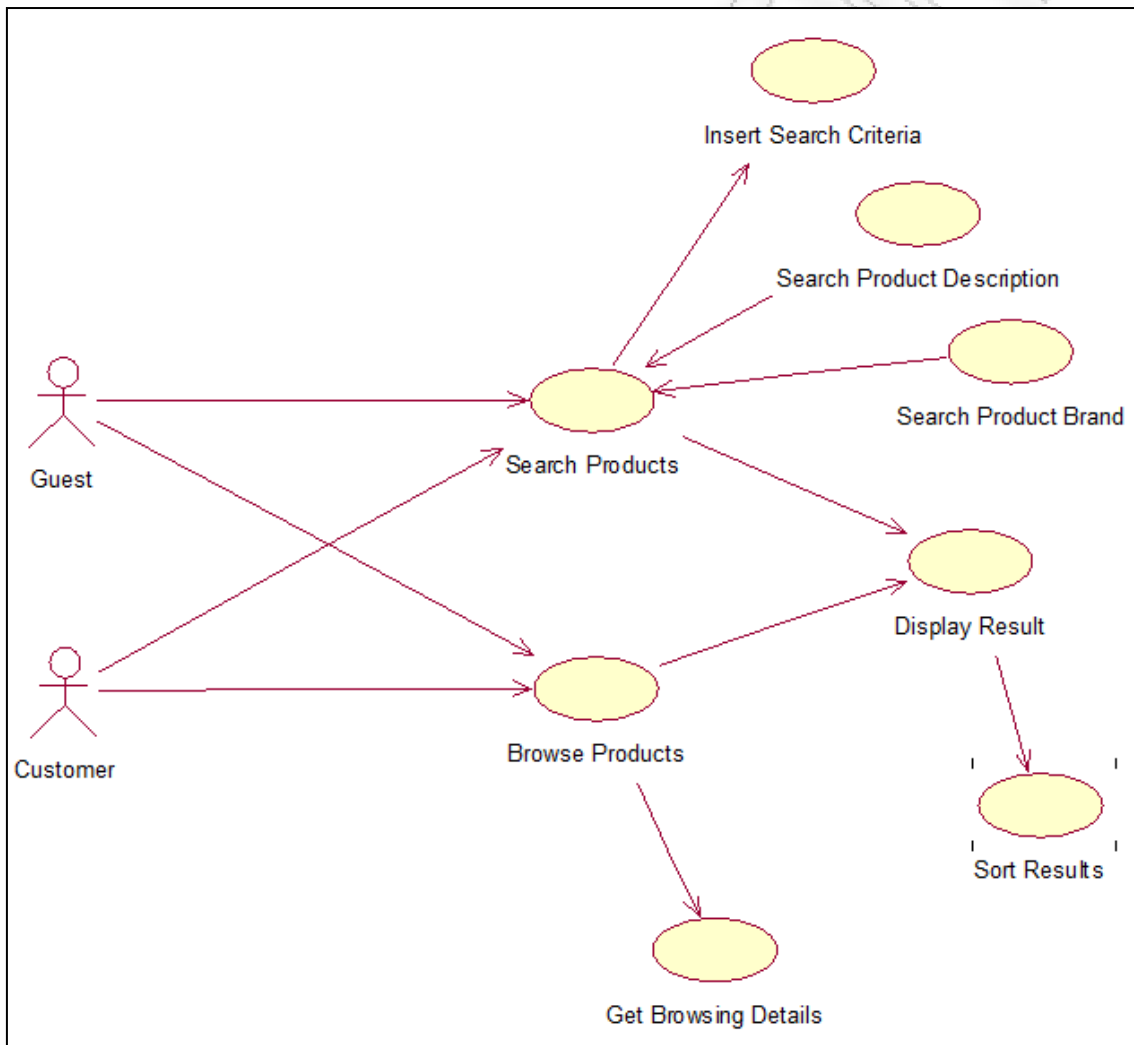
**Εικόνα. 3.1 Use Case 'Customer Register & Log-In'**

Όπως φαίνεται στο παραπάνω διάγραμμα χρήσης (Εικ. 3.1), περιγράφονται τρεις κύριες λειτουργίες του συστήματος. Η μία είναι η διαδικασία Register της οποίας ενεργοποιός είναι ο Guest. Ως Guest ορίζεται ο επισκέπτης του ηλεκτρονικού καταστήματος, ο οποίος όμως δεν έχει γίνει μέλος του. Για να γίνει μέλος πρέπει να εγγραφεί στο σύστημα. Την εγγραφή στο σύστημα απεικονίζει η διαδικασία Register. Η εγγραφή στο ηλεκτρονικό κατάστημα αποτελείται από τρεις διαδικασίες. Πρώτη είναι η διαδικασία εισαγωγής δεδομένων (Insert Registration Data) κατά την οποία ο χρήστης καλείται να συμπληρώσει τα προσωπικά του στοιχεία. Αφού ολοκληρωθεί η συμπλήρωση των στοιχείων ακολουθεί η διαδικασία επαλήθευσης των στοιχείων (Validate Registration Data) ή οποία αν είναι επιτυχημένη ακολουθείται από την διαδικασία Ολοκλήρωσης της εγγραφής (Complete Registration Process), η οποία εισάγει τον χρήστη στο σύστημα ως νέο πιστοποιημένο πελάτη (Customer).

Η διαδικασία Login έχει ως ενεργοποιό τον Customer, ο οποίος συμβολίζει ένα εγγεγραμμένο μέλος του συστήματος. Σύμφωνα με αυτή την διαδικασία ο πελάτης καλείται να εισάγει τα πιστοποιητικά εισόδου στο σύστημα (διαδικασία Insert Identification Data) και αφού το σύστημα τα εγκρίνει (διαδικασία Validate Identification Data), ο χρήστης είναι πλέον αναγνωρισμένος από το σύστημα.

Τέλος περιγράφεται η διαδικασία κατά την οποία ένας χρήστης ζητά την ανάκτηση του συνθηματικού εισόδου και καλείται να εισάγει την απάντηση στην μυστική ερώτηση (διαδικασία Answer Security Question) την οποία το σύστημα εξετάζει αν είναι σωστή ώστε να αποστείλει τα ανακτημένα στοιχεία στον χρήστη (διαδικασία Match Security Answer).

### 3.3.2 Διάγραμμα χρήσης Περιήγησης και Αναζήτησης

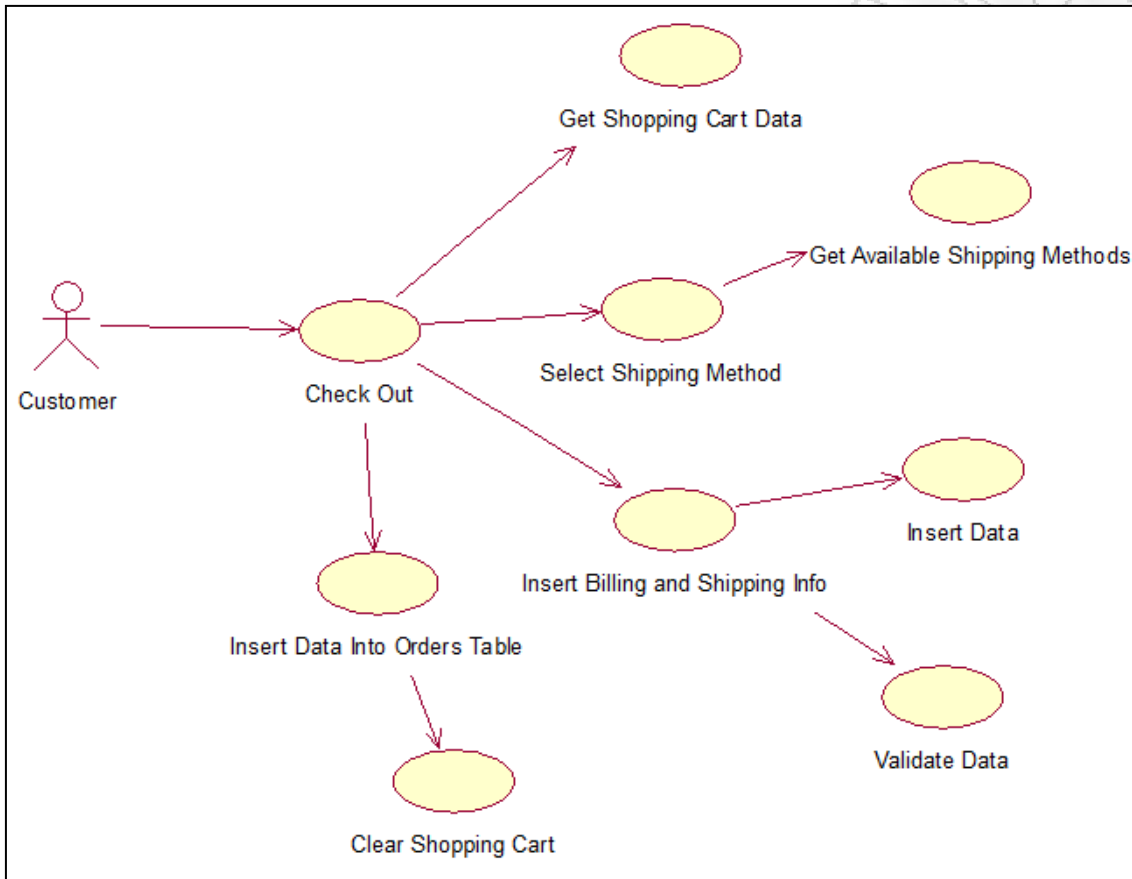


Εικόνα 3.2 Use Case 'Browsing & Search'

Στο παραπάνω διάγραμμα (Εικ. 3.2) περιγράφονται οι λειτουργίες περιήγησης (Browse Products) και αναζήτησης (Search Products). Κύριοι ενεργοποιοί είναι και πάλι ο Guest και ο Customer οι οποίοι συμμετείχαν και στο πρώτο διάγραμμα (Εικ. 3.1). Στην διαδικασία αναζήτησης οι δύο ενεργοποιοί εισάγουν τα κριτήρια αναζήτησης και το σύστημα αναζητά τα κριτήρια αυτά στην περιγραφή ή με βάση την κατηγορία του προϊόντος και επιστρέφει τα αποτελέσματα κατάλληλα ταξινομημένα στην οθόνη.

Στην διαδικασία Browse Products οι δύο ενεργοποιημένοι ζητούν να δουν λεπτομέρειες για ένα προϊόν επιλέγοντάς το και το σύστημα επιστρέφει τις πληροφορίες για το προϊόν στην οθόνη του χρήστη.

### 3.3.3 Διάγραμμα χρήσης Ολοκλήρωσης αγοράς (Check-Out)

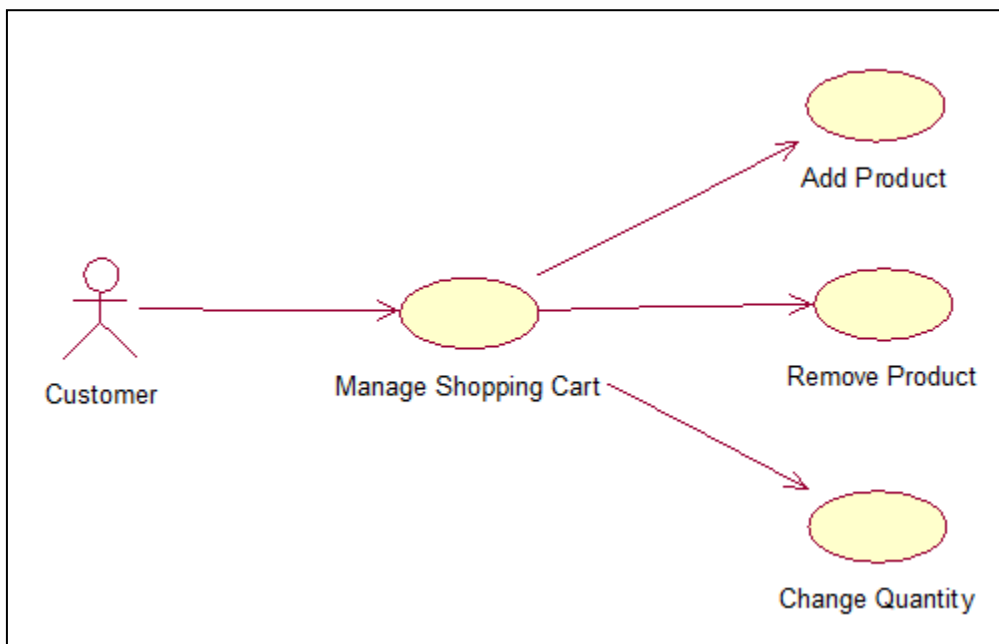


Εικόνα. 3.3 Use Case 'Check-Out'

Το διάγραμμα χρήσης Check-Out (Εικ. 3.3) περιγράφει πώς ένας εγγεγραμμένος χρήστης του ηλεκτρονικού καταστήματος, μπορεί να ολοκληρώσει μία αγορά. Κατά την διαδικασία αυτή, ο χρήστης καλείται να επιλέξει την μέθοδο αποστολής των προϊόντων (Select Shipping Method) και να εισάγει την μέθοδο πληρωμής (Insert Billing and Shipping Info). Το σύστημα αξιολογεί τα στοιχεία αυτά ως προς την ορθότητά τους και εφόσον είναι σωστά ανακτά τα δεδομένα των προϊόντων που υπάρχουν στο καλάθι αγοράς (Get Shopping Cart Data). Αν όλα πάνε καλά όλα τα δεδομένα της παραγγελίας αποθηκεύονται στον πίνακα που διατηρεί το ιστορικό των αγορών για όλους τους πελάτες (orders).

### 3.3.4 Διάγραμμα χρήσης διαχείρισης καλαθιού αγοράς

Το διάγραμμα χρήσης για την διαχείριση του καλαθιού αγοράς (Εικ. 3.4) περιγράφει τις επιλογές που έχει ένας χρήστης σε σχέση με τα περιεχόμενα του καλαθιού. Ο χρήστης έχει την δυνατότητα να προσθέσει ένα ή περισσότερα προϊόντα στο καλάθι αγοράς (Add Product) επιλέγοντας το κουμπί Add to Cart που υπάρχει δίπλα σε κάθε προϊόν. Επίσης ο χρήστης μπορεί να αφαιρέσει εάν επιθυμεί ένα προϊόν από το καλάθι αγοράς (Remove Product). Τέλος μπαίνοντας στη σελίδα διαχείρισης του καλαθιού αγοράς, ο χρήστης μπορεί να αυξομειώσει όσο επιθυμεί τις ποσότητες των προϊόντων που υπάρχουν στο καλάθι.



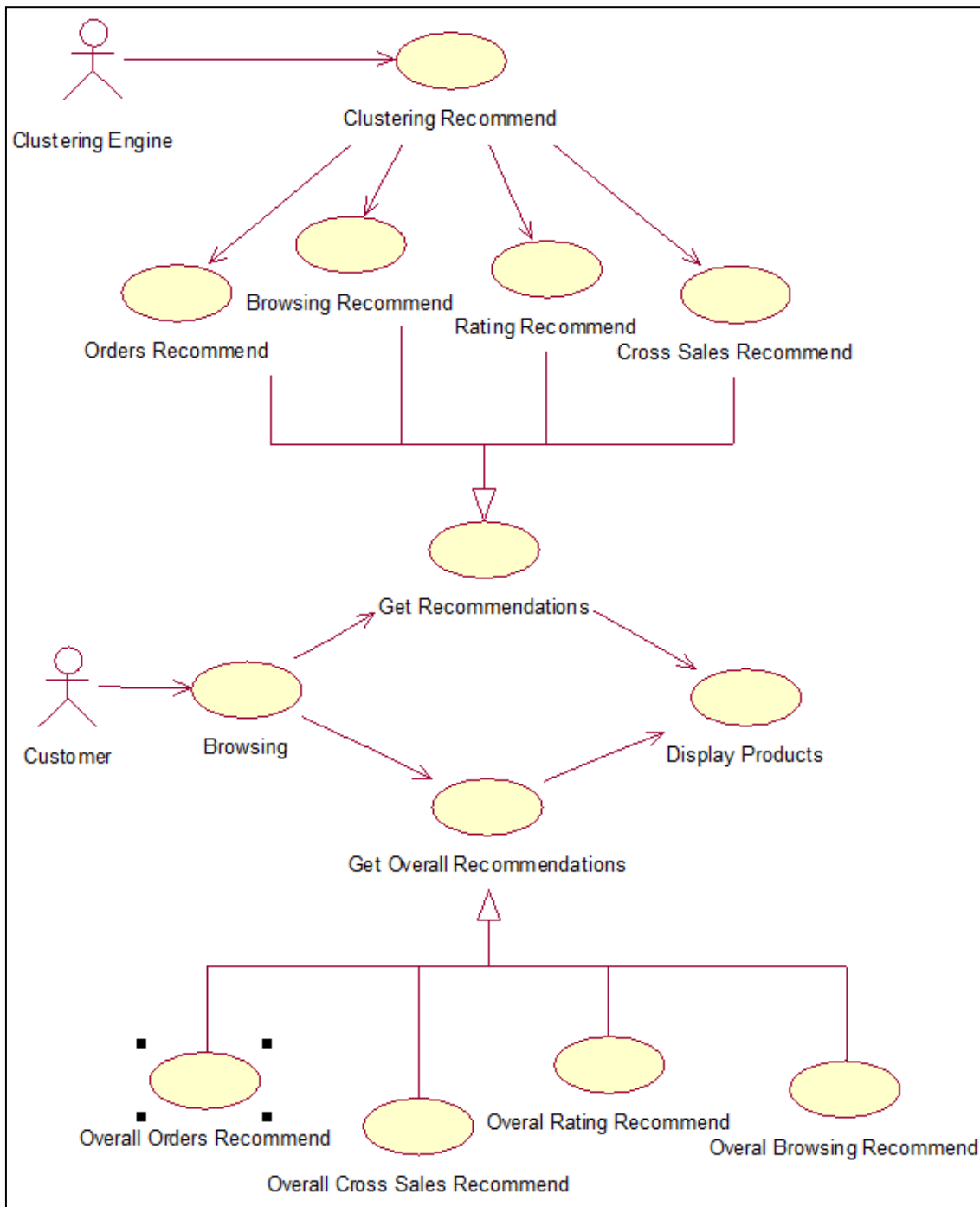
Εικόνα 3.4 Use Case 'Manage Shopping Cart'

### 3.3.5 Διάγραμμα χρήσης διαχείρισης προτάσεων

Στο διάγραμμα χρήσης διαχείρισης προτάσεων (Εικ. 3.5), εξηγείται η διαδικασία με την οποία το σύστημα προτάσεων (Recommender System), προσαρμόζει το περιβάλλον περιήγησης του χρήστη διαμορφώνοντας την διεπαφή του καταστήματος ανάλογα με τις προτάσεις από εξάγονται από τον μηχανισμό ομαδοποίησης.

Όταν ένας χρήστης (ο οποίος έχει κάνει login), περιηγείται στο ηλεκτρονικό κατάστημα και επιλέγει κάποιο προϊόν, τότε το σύστημα συμπληρώνει (Display Products) κάποιες επιλεγμένες θέσεις στην οθόνη (placeholders) με άλλα προϊόντα που πιθανώς να ενδιαφέρουν τον χρήστη στο πλαίσιο της στοχευμένης προώθησης των προϊόντων. Ο ενεργοποιημένος Clustering Engine είναι ο μηχανισμός που διαθέτει το ηλεκτρονικό κατάστημα, ο οποίος είναι υπεύθυνος για την ομαδοποίηση των εγγεγραμμένων χρηστών. Η διαδικασία Clustering Recommends είναι αυτή η οποία εκμεταλλεύεται τα στοιχεία της ομαδοποίησης κάνοντας εξόρυξη δεδομένων (data mining). Τα στοιχεία αυτά χωρίζονται σε τέσσερις κατηγορίες, ανάλογα με την προέλευσή τους. Η πρώτη κατηγορία αφορά δεδομένα που εξάγονται με βάση τις πληροφορίες που αφορούν αγορές τις οποίες έχουν πραγματοποιήσει οι χρήστες της ομάδας. Η δεύτερη κατηγορία προέρχεται από τα δεδομένα της ομάδας που αφορούν στην επισκεψιμότητα των πελατών σε διάφορα προϊόντα. Δεδομένα της ομάδας που αφορούν στην τρίτη κατηγορία προτάσεων. Τέλος η τέταρτη κατηγορία αφορά προϊόντα τα οποία έχουν αγοράσει οι χρήστες της ομάδας, που έχουν αγοράσει το προϊόν το οποίο κοιτάζει ο χρήστης (Cross-Sales).

Τέλος, για κάθε κατηγορία από αυτές που αναλύθηκαν παραπάνω, εξάγονται δεδομένα προτάσεων, συνολικά για όλους τους χρήστες και όχι μόνο για τους χρήστες που ανήκουν στην ομάδα του χρήστη (διαδικασία Get Overall Recommendations).



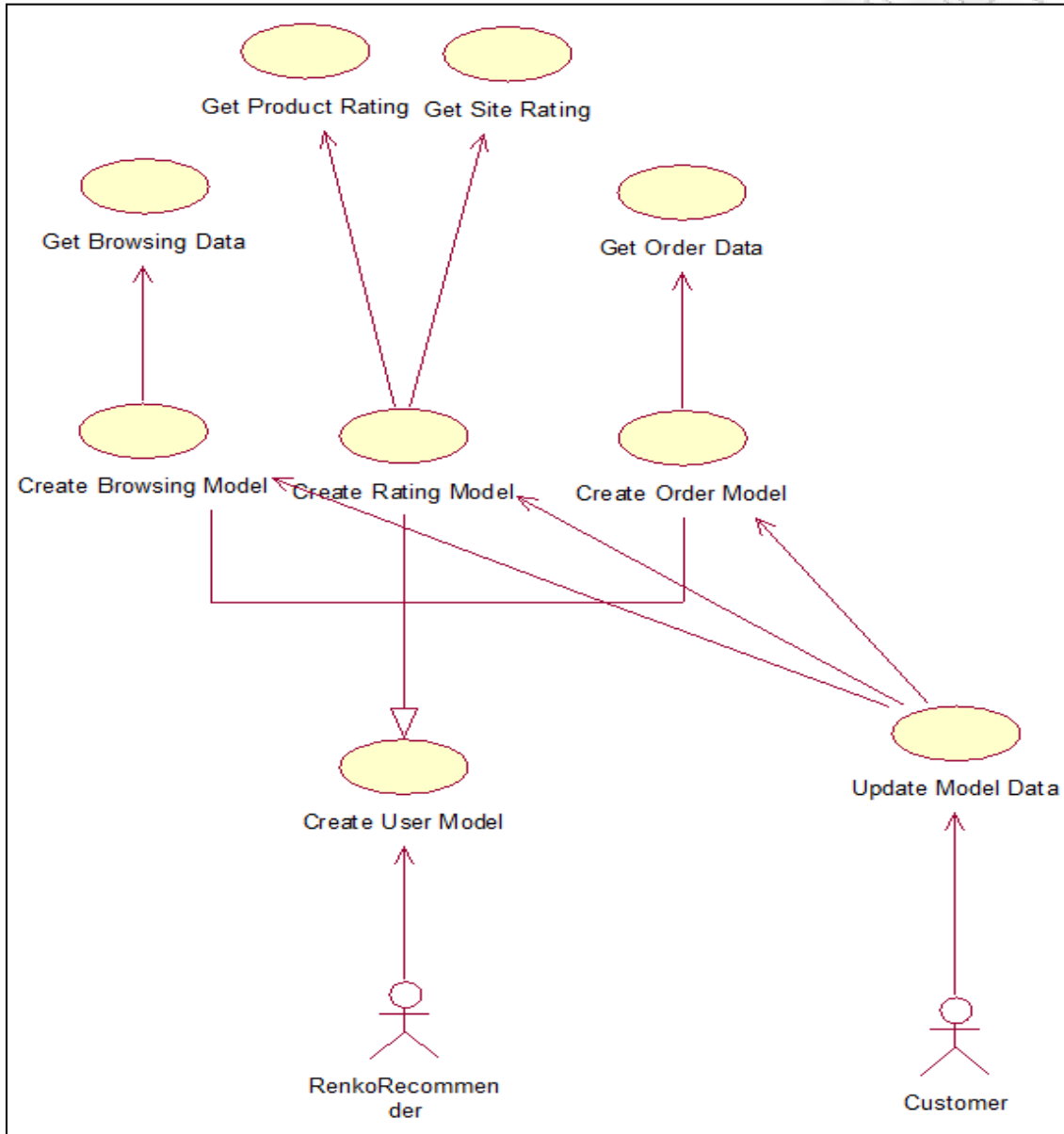
Εικόνα 3.5 Use case ‘Recommend Products’

### 3.3.6 Διάγραμμα χρήσης Δημιουργίας μοντέλου χρηστών

Το διάγραμμα χρήσης δημιουργίας μοντέλου χρηστών (Εικ. 3.6) περιγράφει μία από τις κυριότερες λειτουργίες του συστήματος που περιγράφουμε. Περιγράφει την δημιουργία του μοντέλου χρηστών βάσει των οποίων το σύστημα παράγει προτάσεις για κάθε χρήστη.

Στο σύστημα αυτό παρατηρούμε ότι συμμετέχει ένας νέος ενεργοποιός, ο RenkoRecommender. Έτσι έχουμε ονομάσει το σύστημα το οποίο είναι υπεύθυνο για την παραγωγή προτάσεων. Παρατηρούμε από το διάγραμμα ότι το κύριο μοντέλο χρήστη είναι στην

ουσία σύνθεση τριών μικρότερων μοντέλων (Browsing model, Rating model, Order model) τα οποία παράγονται από την επεξεργασία των δεδομένων αυτών των κατηγοριών που αφορούν σε κάθε ομάδα χρηστών. Σε κάθε ένα από αυτά τα μοντέλα συμμετέχει και ο χρήστης με τις έμμεσες πληροφορίες που παράγει με τις κινήσεις και την εν γένει συμπεριφορά του μέσα στο ηλεκτρονικό κατάστημα.

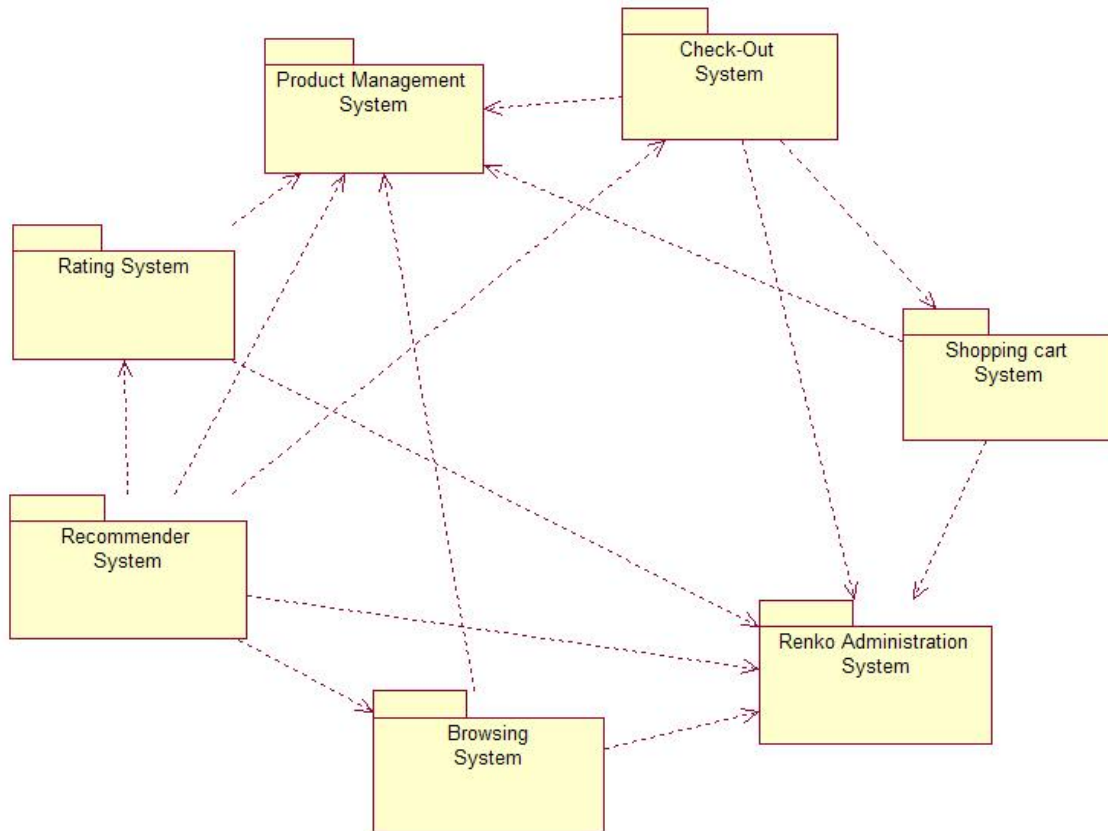


Εικόνα 3.6: Use Case 'Create User Model'

Επίσης είναι εμφανές από το διάγραμμα ότι το μοντέλο που προκύπτει από την βαθμολογία των χρηστών είναι σύνθεση δεδομένων από δεδομένα που αφορούν την ψήφιση προϊόντων και από δεδομένα που αφορούν στην βαθμολόγηση του site. Τέλος όλα τα παραπάνω δεδομένα συνεισφέρουν στην συναρμολόγηση ενός συνολικού μοντέλου χρήστη μέσω του οποίου αποφασίζονται όλες οι προτάσεις που γίνονται στους χρήστες.

### 3.3.7 Διάγραμμα πακέτων υποσυστημάτων.

Στο παρακάτω διάγραμμα παρατηρούνται οι εξαρτήσεις μεταξύ των υποσυστημάτων.



Εικόνα 3.7 Διάγραμμα πακέτων υποσυστημάτων του ηλεκτρονικού καταστήματος.

## 4. ΠΕΡΙΓΡΑΦΗ ΕΦΑΡΜΟΓΗΣ

Στο κομμάτι της υλοποίησης της συγκεκριμένης διατριβής, σαν βασικός στόχος ήταν η δημιουργία ενός ηλεκτρονικού καταστήματος με χρήση ενός αλγορίθμου Clustering, ο οποίος ομαδοποιεί τους χρήστες ανάλογα με τις κινήσεις τους μέσα στον ιστότοπο. Για την υλοποίηση αυτή χρησιμοποιήθηκαν οι γλώσσες προγραμματισμού php, html, xml, javascript οι οποίες είναι από τις πιο διαδεδομένες γλώσσες προγραμματισμού web εφαρμογών με ανοιχτό κώδικα. Επίσης χρησιμοποιήθηκε και η MySql για την δημιουργία μιας βάσης δεδομένων που θα υποστηρίζει το ηλεκτρονικό μας κατάστημα.

Στο συγκεκριμένο κεφάλαιο περιγράφονται τα τεχνικές προγραμματισμού που εφαρμόστηκαν όπως η ajax, προγραμματιστικά εργαλεία που χρησιμοποιήθηκαν για την υλοποίηση της συγκεκριμένης εφαρμογής καθώς και εργαλεία που βοήθησαν στην εξαγωγή στατιστικών και άλλων συμπερασμάτων όπως το matlab.

## 4.1 Πλατφόρμα και Προγραμματιστικά Εργαλεία

### 4.1.1 Επιλογή Γλώσσας Προγραμματισμού

Η γλώσσες προγραμματισμού που επιλέχθηκαν είναι ανοιχτού κώδικα για την διευκόλυνσή μας στην εύρεση εργαλείων και πλατφορμών από το διαδίκτυο. Παρακάτω θα περιγράψουμε μία τις γλώσσες και κάποιες σύγχρονες μεθόδους υλοποίησης που αξίζει να αναφερθούν.

1. Το μεγαλύτερο κομμάτι της εφαρμογής μας υλοποιήθηκε με την PHP 5 η οποία είναι μία γλώσσα προγραμματισμού για τη δημιουργία δυναμικών ιστοσελίδων όπως είναι το ηλεκτρονικό μας κατάστημα. Μια σελίδα PHP για να εμφανιστεί χρειάζεται έναν server τύπου apache (στον οποίο θα αναφερθούμε αργότερα) για την επεξεργασία των εντολών και από έναν φυλλομετρητή ο οποίος απεικονίζει τα αποτελέσματα του apache server. Τα βασικότερα πλεονεκτήματα της συγκεκριμένης γλώσσας είναι ο κύριος λόγος για τον οποίο χρησιμοποιήθηκε:
  - a. Επειδή η PHP είναι μια γλώσσα που εκτελείται από την πλευρά του server, ο χρήστης δεν χρειάζεται να έχει εγκατεστημένη καμία ειδική μηχανή αναζήτησης ή plug-ins για να δει το αποτέλεσμα του κώδικα.
  - b. Τρέχει σε όλα τα λειτουργικά συστήματα Mac, Windows, Linux κ.α., καθώς και όλους τους φυλλομετρητές όπως είναι ο internet explorer, Mozilla firefox, saphari, opera, google chrome κ.α.
  - c. Επειδή είναι script γλώσσα δεν χρειάζεται κάποιον compiler παρά μόνο ένα notepad στο οποίο γράφεις τον κώδικα και στο τέλος το αποθηκεύεις με την κατάληξη .php .
  - d. Είναι ανοιχτού κώδικα με αποτέλεσμα να υπάρχουν πολλά παραδείγματα και manual από την κοινότητα της php, τα οποία διατίθενται δωρεάν.
2. Η HTML σε συνδυασμό με τα CSS(Cascading Style Sheets) είναι η γλώσσες χαρακτηρισμού με τις οποίες δώσαμε σχήμα και χρώμα στο ηλεκτρονικό μας κατάστημα. Χρησιμοποιώντας την html δημιουργήσαμε τη φόρμα εγγραφής στην ιστοσελίδα, τις φόρμες ψηφοφορίας και γενικά της εμφάνισης της ιστοσελίδας.
3. Η javascript είναι άλλη μια γλώσσα προγραμματισμού script η οποία όμως σε αντίθεση με την php τρέχει στον φυλλομετρητή και όχι στον server. Αυτό σημαίνει ότι η επεξεργασία του κώδικα Javascript και η παραγωγή του τελικού περιεχομένου HTML δεν πραγματοποιείται στον server, αλλά στο πρόγραμμα περιήγησης των επισκεπτών. Αυτή η διαφορά έχει και πλεονεκτήματα και μειονεκτήματα για καθεμιά από τις δύο γλώσσες. Συγκεκριμένα, η Javascript δεν έχει καμία απαίτηση από πλευράς δυνατοτήτων του server για να εκτελεστεί (επεξεργαστική ισχύ, συμβατό λογισμικό διακομιστή), αλλά βασίζεται στις δυνατότητες του φυλλομετρητή των επισκεπτών. Επίσης μπορεί να ενσωματωθεί σε στατικές σελίδες HTML. Παρόλα αυτά, οι δυνατότητές της είναι σημαντικά μικρότερες από αυτές της PHP και δεν παρέχει συνδεσιμότητα με βάσεις δεδομένων.
4. Για την διαχείριση των δεδομένων της ιστοσελίδας (πελάτες, προϊόντα, αγορές) και πολλά που θα αναφέρουμε σε άλλη παράγραφο, χρησιμοποιήσαμε άλλη μια γλώσσα ανοιχτού κώδικα, την MySql. Η MySql είναι μία γλώσσα υπολογιστών στις βάσεις δεδομένων, που σχεδιάστηκε για τη διαχείριση δεδομένων, σε ένα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων (Relational Database Management System, RDBMS) και η οποία, αρχικά, βασίστηκε στη σχεσιακή άλγεβρα. Η γλώσσα περιλαμβάνει δυνατότητες ανάκτησης και ενημέρωσης δεδομένων, δημιουργίας και τροποποίησης σχημάτων και σχεσιακών πινάκων, αλλά και ελέγχου πρόσβασης στα δεδομένα.
5. Η XML (αγγλ. αρκτ. από το Extensible Markup Language) είναι μία γλώσσα σήμανσης, που περιέχει ένα σύνολο κανόνων για την ηλεκτρονική κωδικοποίηση κειμένων. Ορίζεται, κυρίως, στην προδιαγραφή XML 1.0 (XML 1.0 Specification), που δημιούργησε ο διεθνής οργανισμός προτύπων W3C (World Wide Web Consortium), αλλά και σε διάφορες άλλες σχετικές προδιαγραφές ανοιχτών προτύπων.



Η XML σχεδιάστηκε δίνοντας έμφαση στην απλότητα, τη γενικότητα και τη χρησιμότητα στο Διαδίκτυο. Είναι μία μορφοποίηση δεδομένων κειμένου, με ισχυρή υποστήριξη Unicode για όλες τις γλώσσες του κόσμου. Αν και η σχεδίαση της XML εστιάζει στα κείμενα, χρησιμοποιείται ευρέως για την αναπαράσταση αυθαίρετων δομών δεδομένων, που προκύπτουν για παράδειγμα στις υπηρεσίες ιστού.

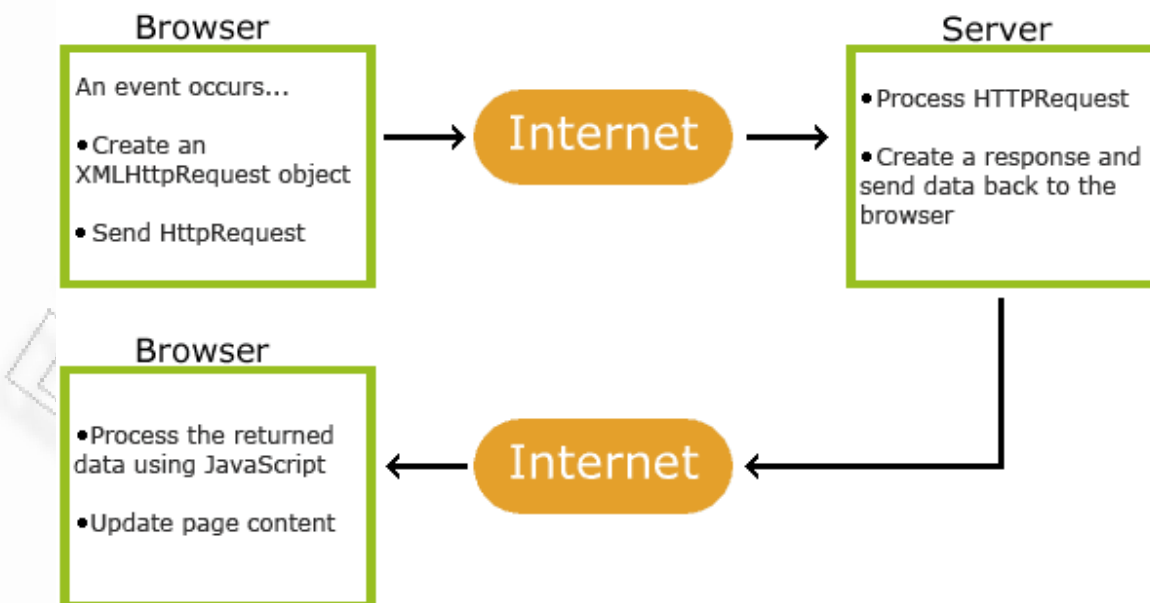
6. Τέλος θα αναφέρουμε και μία μέθοδο προγραμματισμού επονομαζόμενη ως Ajax η οποία είναι πολύ διαδεδομένη στο διαδίκτυο. Η Ajax ( Asynchronous Javascript and XML, είναι ένα σύνολο τεχνικών για την δημιουργία άκρως διαδραστικών ιστοσελίδων και εφαρμογών στο ιντερνέτ. Με την χρήση της Ajax τα δεδομένα μπορούν να μεταφερθούν ασύγχρονα μέσω του αντικειμένου XMLHttpRequest και η σελίδα να συνεχίζει την φόρτωση της χωρίς να περιμένει απάντηση από τον server. Η Ajax είναι ένας από τους πρόσφατους και σημαντικότερους τρόπους βελτίωσης της online εμπειρίας των χρηστών και δημιουργίας νέας και καινοτομικής λειτουργικότητας web. Επιτρέποντας σε συγκεκριμένα μέρη μιας ιστοσελίδας να προβάλλονται χωρίς ανανέωση ολόκληρης της σελίδας, βελτιώνει σημαντικά την εμπειρία των εφαρμογών web. Επιτρέπει επίσης στους προγραμματιστές web να δημιουργούν δαισθητικές και καινοτομικές διαδραστικές διαδικασίες

Η τεχνολογία Ajax, αν και αρκετά νέα, έχει αρχίσει και χρησιμοποιείται από όλο και περισσότερες δικτυακές εφαρμογές κερδίζοντας συνεχώς έδαφος. Πρώτη η Google, η πασίγνωστη μηχανή αναζήτησης, δημιούργησε αρκετή εντύπωση χρησιμοποιώντας την Ajax στην δοκιμαστική υπηρεσία Google Suggest. Έπειτα ακολούθησαν και πάρα πολλές άλλες υπηρεσίες με σημαντικότερες τις Google Maps, Gmail, Youtube, Facebook κ.α. που χρησιμοποιούν σε μεγάλο βαθμό την τεχνολογία Ajax. Επίσης ο νέος ιστότοπος Microsoft Live (<http://www.live.com>) με υπηρεσίες αντίστοιχες της Google, καθώς και η τελευταία έκδοση της σελίδας της Yahoo (<http://www.yahoo.com>), χρησιμοποιούν εκτενώς Ajax.

Το Google Spreadsheets είναι μια νέα εφαρμογή από τα Google Labs και μπορεί να βρεθεί στη διεύθυνση <http://spreadsheets.google.com>. Πρόκειται για μια web εφαρμογή λογιστικών φύλλων με κυριότερο χαρακτηριστικό της ότι είναι φτιαγμένη εξ ολοκλήρου με τη χρησιμοποίηση της Ajax, γεγονός που την καθιστά άκρως ανταγωνιστική.

Η τεχνολογία Ajax περιλαμβάνει τα εξής:

- Παρουσίαση χρησιμοποιώντας XHTML και CSS, με βάση τα standards
- Δυναμική παρουσίαση μέσω του Document Object Model
- Ανταλλαγή και διαχείριση πληροφοριών χρησιμοποιώντας XML
- Ασύγχρονη μεταφορά δεδομένων μέσω του XMLHttpRequest
- Και την Javascript που τα δένει όλα μαζί



Εικόνα 4.1: Πως δουλεύει η AJAX (Πηγή: <http://www.gtsamis.gr/?p=1057>)

#### 4.1.2 Επιλογή Λειτουργικού Συστήματος

Έχοντας να δουλέψουμε με τις παραπάνω γλώσσες είχαμε την επιλογή να αναπτύξουμε τον κώδικά μας σε οποιοδήποτε περιβάλλον και λειτουργικό σύστημα θέλαμε. Επιλέχθηκαν τα Windows XP της εταιρίας Microsoft λόγω της συμβατότητας που προσφέρουν, αφού αποτελούν μακράν την πιο διαδεδομένη πλατφόρμα, αλλά και λόγω της πληθώρας των προγραμμάτων και εργαλείων για ανάπτυξη κώδικα που διατίθενται σε αυτά.

Εδώ να αναφέρουμε ότι ο παραγόμενος κώδικας είναι πολύ εύκολο να χρησιμοποιηθεί σε όλα τα λειτουργικά συστήματα, αρκεί βέβαια να είναι εγκατεστημένος ένας apache server, η php, και η mysql τα οποία εγκαθίστανται πολύ εύκολα και χωρίς ειδικές γνώσεις με την πλατφόρμα wamp.

#### 4.1.3 Επιλογή Εργαλείου Ανάπτυξης Κώδικα και σχεδιασμού βάσης δεδομένων

Για την ανάπτυξη του κώδικα χρησιμοποιήθηκε το εργαλείο notepad++ στο οποίο μπορούν να συνταχθούν όλες οι παραπάνω γλώσσες που αναφέρθηκαν. Για να λειτουργήσει όμως ο κώδικας χρειάζεται εγκατάσταση μίας ομάδας server για την μετάφραση του κώδικα και την δημιουργία της βάσης δεδομένων. Για την εγκατάσταση αυτή χρησιμοποιήσαμε τον wampserver. Ο WampServer είναι ένα web Windows περιβάλλον ανάπτυξης ανοικτού κώδικα, το οποίο μας επιτρέπει να δημιουργήσουμε εφαρμογές web με Apache, PHP και βάση δεδομένων MySQL. Επίσης, για την καλύτερη διαχείριση της βάσης δεδομένων έχει ενσωματωμένο το phpmyadmin .

1. Ο Apache HTTP γνωστός και απλά σαν Apache είναι ένας εξυπηρετητής του παγκόσμιου ιστού (web). Όποτε ένας χρήστης επισκέπτεται ένα ιστότοπο το πρόγραμμα πλοήγησης (browser) επικοινωνεί με έναν διακομιστή (server) μέσω του πρωτοκόλλου HTTP, ο οποίος παράγει τις ιστοσελίδες και τις αποστέλλει στο πρόγραμμα πλοήγησης. Ο Apache είναι ένας από τους δημοφιλέστερους, εν μέρει γιατί λειτουργεί σε διάφορες πλατφόρμες όπως τα Windows, το Linux, το Unix και το Mac OS X. Συντήρεται τώρα από μια κοινότητα ανοικτού κώδικα με επιτήρηση από το Ίδρυμα Λογισμικού Apache (Apache Software Foundation).

Η MySQL είναι ένα σύστημα διαχείρισης σχεσιακών βάσεων δεδομένων (RDBMS) το οποίο μετρά περισσότερες από 11 εκατομμύρια εγκαταστάσεις. Έλαβε το όνομά του από την κόρη του Μόντυ Βιντένιους, την Μάι. Το πρόγραμμα τρέχει έναν εξυπηρετητή (server) παρέχοντας πρόσβαση πολλών χρηστών σε ένα σύνολο βάσεων δεδομένων.

Το phpMyAdmin είναι ένα δωρεάν εργαλείο, λογισμικό γραμμένο σε PHP που προορίζεται για διαχείριση βάσεων δεδομένων της MySQL μέσω του World Wide Web. phpMyAdmin υποστηρίζει ένα ευρύ φάσμα δράσεων με MySQL. Οι πιο συχνά χρησιμοποιούμενες λειτουργίες υποστηρίζονται από το περιβάλλον εργασίας χρήστη (βάσεις δεδομένων διαχείρισης, πίνακες, πεδία, σχέσεις, ευρετήρια, οι χρήστες, δικαιώματα, κ.λπ.), ενώ έχει τη δυνατότητα ο χρήστης να εκτελέσει άμεσα, ερωτήματα SQL.

Ακολουθεί σχηματική αναπαράσταση της βάσης δεδομένων του ηλεκτρονικού καταστήματος που χτίστηκε με τη βοήθεια του phpmyadmin .

	Πίνακας	Ενέργεια	Εγγραφές <sup>1</sup>	Τύπος	Σύνθεση	Μέγεθος	Περίσσεια	
<input type="checkbox"/>	basket		4	InnoDB	latin1_swedish_ci	32,0 KB	-	
<input type="checkbox"/>	categories		18	InnoDB	latin1_swedish_ci	32,0 KB	-	
<input type="checkbox"/>	clicks		305	InnoDB	latin1_swedish_ci	48,0 KB	-	
<input type="checkbox"/>	clustered_clicks		2,175	InnoDB	greek_general_ci	112,0 KB	-	
<input type="checkbox"/>	clustered_orders		2,378	InnoDB	greek_general_ci	128,0 KB	-	
<input type="checkbox"/>	clustered_prod_votes		1,455	InnoDB	greek_general_ci	80,0 KB	-	
<input type="checkbox"/>	clusters		4,970	InnoDB	greek_general_ci	320,0 KB	-	
<input type="checkbox"/>	cross_sales		1	InnoDB	greek_general_ci	16,0 KB	-	
<input type="checkbox"/>	customer		73	InnoDB	greek_general_ci	32,0 KB	-	
<input type="checkbox"/>	customer_type		2	InnoDB	greek_general_ci	16,0 KB	-	
<input type="checkbox"/>	menu		4	InnoDB	latin1_swedish_ci	16,0 KB	-	
<input type="checkbox"/>	orders		198	InnoDB	latin1_swedish_ci	48,0 KB	-	
<input type="checkbox"/>	poll		164	InnoDB	latin1_swedish_ci	16,0 KB	-	
<input type="checkbox"/>	products		216	InnoDB	greek_general_ci	432,0 KB	-	
<input type="checkbox"/>	site_vote		74	InnoDB	greek_general_ci	16,0 KB	-	
15 Πίνακας/Πίνακες			Σύνολο	12,037	MyISAM	greek_general_ci	1,3 MB	0 Bytes

Εικόνα 4.2: Απεικόνιση πινάκων με τη βοήθεια του phpmyadmin.

## 4.2 Περιγραφή Βασικών Λειτουργιών του e-shop

Σε αυτή τη παράγραφο θα περιγράψουμε τις λειτουργίες του ηλεκτρονικού καταστήματος με απεικόνιση τμημάτων κώδικα και εικόνων.

### 4.2.1 Εγγραφή στο ηλεκτρονικό μας κατάστημα.

Κατά την εγγραφή του χρήστη ζητούνται τα παρακάτω στοιχεία: ονοματεπώνυμο, διεύθυνση, φύλλο, user name, e-mail και κωδικός. Δεν έχει προστεθεί κάποιο ερωτηματολόγιο διότι η ομαδοποίηση που εφαρμόζεται είναι βάση των κινήσεων τους οι οποίες είναι οι εξής:

- Αγορά(orders) προϊόντων
- Επίσκεψη προϊόντων(clicks)
- Βαθμολογία προϊόντων(product vote)
- Ενδιαφέρον εμφάνισης περισσότερων χαρακτηριστικών σε κάθε προϊόν.(advanced user)
- Βαθμολογία του ηλεκτρονικού καταστήματος

Ακολουθεί απεικόνιση της φόρμας εγγραφής.

**REGISTER**

First Name:

Last Name:

Street Address:

Street Address Number:

Gender:

User Name:

E-mail:

Password:

Confirm Password:

**Εικόνα 4.3** Φόρμα εγγραφής στο ηλεκτρονικό κατάστημα

Κάνοντας αριστερό κλικ στο κουμπί Register γίνονται έλεγχοι ορθότητας του e-mail,password και συμπλήρωσης όλων των πεδίων. Αν δεν παρουσιαστεί κάποιο σφάλμα τότε εκτελείται ο παρακάτω κώδικας για την εισαγωγή του πελάτη στη βάση δεδομένων.





```
$query="INSERT INTO customer (username, password, email, gender, fname, lname, street_name, street_number)
```

```
VALUES ( '$_POST[username]', '$_POST[password]', '$_POST[email]',  
'$_POST[gender]', '$_POST[fname]', '$_POST[lname]', '$_POST[street_name]',  
'$_POST[street_number]');
```

```
$ret = mysql_query($query);
```

#### **4.2.2 Προσθήκη προϊόντος στο καλάθι αγορών.**

Για την προσθήκη ενός προϊόντος στο καλάθι αγορών χρειάζεται απλά να πατηθεί ο σύνδεσμος Buy ή όπου υπάρχει το εικονίδιο με σχήμα ενός καροτσιού super market.

Product Name	Item Price	Count	Total Price	Del
 Core 2 Quad 9300 Tra	223.00 €	1 ADD	223 €	X
 X1550 256MB TVO-DVI	26.00 €	2 ADD	52 €	X
 DFI Lanparty DK P45	159.00 €	1 ADD	159 €	X
 CoolerMaster Storm Sniper w/Window Case	154.00 €	1 ADD	154 €	X
<a href="#">Return to shopping</a>	<a href="#">Confirm Order</a>			
Total Price:	588 €			

**Εικόνα 4.4** Παράδειγμα καλαθιού αγορών ενός πελάτη.

Στο καλάθι αγορών δίνονται οι εξής δυνατότητες:

1. Επιστροφή στην σελίδα του τελευταίου προϊόντος που επέλεξε για το καλάθι αγορών πατώντας τον σύνδεσμο Return shopping.
2. Επιλογή τεμαχίων ανά προϊόν τοποθετώντας έναν αριθμό στη στήλη Count και πατώντας το κουμπί ADD.
3. Αφαίρεση οποιουδήποτε προϊόντος από το καλάθι αγορών πατώντας το X στη στήλη Del
4. Και τέλος μπορεί να ολοκληρώσει την παραγγελία του πατώντας τον σύνδεσμο Confirm Order. Με την ολοκλήρωση της αγοράς γίνεται προσθήκη των προϊόντων στον πίνακα αγορών (orders) και γίνεται διαγραφή των προϊόντων από το καλάθι αγορών. Παρακάτω παρουσιάζεται ο αντίστοιχος κώδικας που εκτελείται.

```

$query1="SELECT * From basket
        WHERE customer_id = $user_id";
$ret1 = mysql_query($query1);
$num_of_rows = mysql_num_rows($ret1);

while( $row1 = mysql_fetch_array($ret1)){
$product_id = $row1['product_id'];
$count = $row1['count'];

$query = "INSERT INTO orders (cust_id, product_id,count )
        VALUES ('$user_id','$product_id','$count')";
$ret = mysql_query($query);
}

mysql_query("DELETE FROM basket WHERE customer_id= $user_id");

```

#### 4.2.3 Παρακολούθηση επίσκεψης προϊόντων

Με την επίσκεψη της σελίδας ενός προϊόντος, προστίθεται ένα click στο προϊόν για τον συγκεκριμένο πελάτη στη βάση δεδομένων. Με αυτό τον τρόπο μπορούν να ελεγχθούν οι προτιμήσεις ενός πελάτη και να εξαχθούν χρήσιμες πληροφορίες από τη μοντελοποίηση που εφαρμόζεται στο ηλεκτρονικό κατάστημα. Ο κώδικας που εκτελείται είναι ο παρακάτω:

```

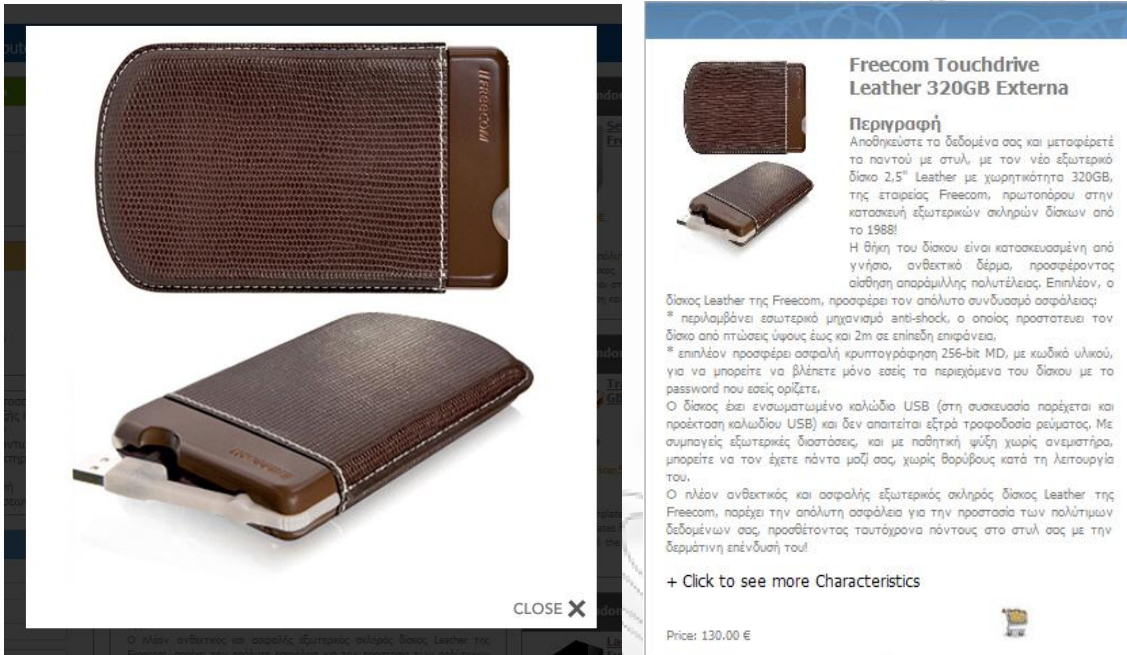
$query = "SELECT * FROM clicks
        WHERE customer_id=$user and product_id=$product_id

$ret = mysql_query($query);
$num = mysql_num_rows($ret);
if($num==0)
{
Εάν δεν έχει ξανά επισκεφτεί το προϊόν γίνεται εισαγωγή καινούριας εγγραφής
$query = "INSERT INTO clicks (customer_id, cat_id, product_id, clicks_number)
        VALUES($user,$cat_id,$product_id, 1)";

$ret = mysql_query($query);
}else{
Εάν έχει ξανά επισκεφτεί το προϊόν γίνεται ενημέρωση της υπάρχουσας εγγραφής προσθέτοντας ένα click στο πεδίο count του πίνακα clicks.
$row = mysql_fetch_array($ret);
$clicks = $row['clicks_number'];
$clicks++;
$query = "UPDATE clicks set clicks_number=$clicks where customer_id=$user and
product_id=$product_id";

```

```
$ret = mysql_query($query);
}
```



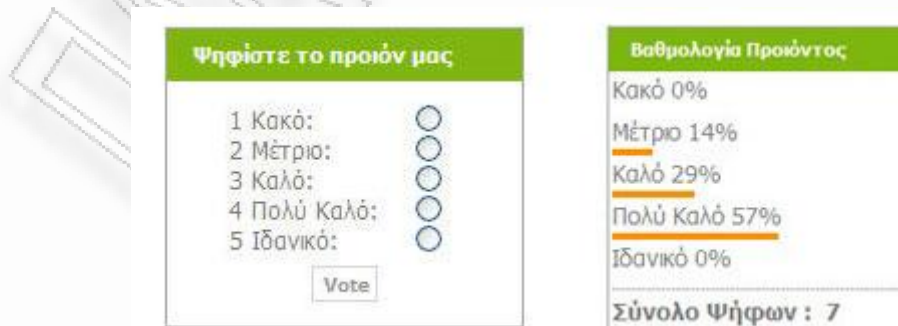
**Εικόνα 4.5: Μεγέθυνση προϊόντος για καλύτερη προεπισκόπηση και πληροφορίες προϊόντος**

Με το που εισέλθει ο χρήστης στην πληροφοριακή σελίδα έχει την δυνατότητα άμεσα να πληροφορηθεί για τα βασικά χαρακτηριστικά του προϊόντος καθώς και την τιμή του. Επίσης αν θέλει να το αγοράσει μπορεί να το προσθέσει στο καλάθι αγοράς πατώντας τον σύνδεσμο Buy now. Επιπρόσθετα μία χρήσιμη λειτουργία είναι η λειτουργία της μεγέθυνσης του προϊόντος (Εικ. 4.5) που προσφέρει μια καλύτερη προεπισκόπηση του στον χρήστη. Τέλος ο χρήστης αν δεν καλύπτεται από την βασική περιγραφή του προϊόντος μπορεί να δει μια αναλυτική περιγραφή του προϊόντος πατώντας το σύνδεσμο Click to see more characteristics (Εικ. 4.5) και αν θέλει να τα αποκρύψει μπορεί να το κάνει πατώντας τον σύνδεσμο Hide Characteristics. Εάν ένας χρήστης έχει πατήσει περισσότερες από 15 φορές τον σύνδεσμο αυτόν τότε η αναλυτική περιγραφή θα είναι ανοιχτή σε κάθε προϊόν από προεπιλογή.

**4.2.4 Βαθμολόγηση προϊόντων**

Στη σελίδα κάθε προϊόντος δίνεται η δυνατότητα στο χρήστη να βαθμολόγηση το προϊόν μόνο μία φορά, και αμέσως μετά εμφανίζεται το αποτέλεσμα της ψηφοφορίας όλων των χρηστών. Η βαθμολόγηση των προϊόντων είναι άλλος ένας τρόπος «παρακολούθησης» των προτιμήσεων του χρήστη, ώστε να εισαχθούν στον αλγόριθμο της μοντελοποίησης χρηστών.

Παρακάτω ακολουθεί απεικόνιση της φόρμας βαθμολόγησης και το αποτέλεσμα αυτής.



**Εικόνα 4.6 Παράδειγμα φόρμας και αποτελέσματος βαθμολογίας προϊόντος.**

Ενδεικτικά παρατίθεται ένα μικρό κομμάτι κώδικα που εκτελείται κατά την βαθμολόγηση του προϊόντος.

```
INSERT INTO poll (product_id, customer_id, vote_rating, ip_address)
VALUES ('$product_id','$user_id','$customers_vote','$customers_ip')
```

#### 4.2.5 Βαθμολόγηση του ηλεκτρονικού καταστήματος.

Έχει δοθεί το δικαίωμα βαθμολόγησης του ηλεκτρονικού καταστήματος σε κάθε εγγεγραμμένο χρήστη μία φορά για κάθε νέα έκδοση του. Έτσι λαμβάνεται η γνώμη των χρηστών και επιτυγχάνεται η συνεχής βελτίωση της ιστοσελίδας. Επίσης η βαθμολογία που δίδεται από κάθε χρήστη, συμπεριλαμβάνεται στη μοντελοποίηση χρηστών για την αναπροσαρμογή της ιστοσελίδας για κάθε ομάδα χρηστών. Παρακάτω παρατίθεται η απεικόνιση της φόρμας βαθμολόγησης και αποτελέσματος της ιστοσελίδας.



Εικόνα 4.7 Παράδειγμα φόρμας και αποτελέσματος βαθμολογίας ιστοσελίδας.

Ενδεικτικά παρατίθεται ένα μικρό κομμάτι κώδικα που εκτελείται κατά την βαθμολόγηση του ηλεκτρονικού καταστήματος.

```
INSERT INTO site_vote (customer_id, site_rate)
VALUES ('$user_id','$customers_vote')
```

#### 4.2.6 Αναζήτηση προϊόντος

Για την πιο γρήγορη εύρεση ενός προϊόντος προστέθηκε ένα παράθυρο αναζήτησης όπως συμβαίνει στις περισσότερες ηλεκτρονικές σελίδες. Ως εκ τούτου δεν θα μπορούσε να λείπει ένα τέτοιο εργαλείο από το ηλεκτρονικό μας κατάστημα. Κατά την αναζήτηση ενός προϊόντος ή ομάδας προϊόντων, εκτελείτε ένα ερώτημα στη βάση δεδομένων και συγκεκριμένα στα πεδία που αφορούν τον πίνακα προϊόντων. Η αναζήτηση ενός προϊόντος μπορεί να γίνει με τρεις τρόπους:

1. Αναζήτηση προϊόντος σε συγκεκριμένη κατηγορία.

Σε αυτήν την περίπτωση ο χρήστης επιλέγει κατηγορία στο παράθυρο της αναζήτησης και μετά προχωράει στην αναζήτηση. Μπορεί επίσης να πληκτρολογήσει μια λέξη προς αναζήτηση και να περιορίσει τα αποτελέσματα επιλέγοντας και κατηγορία.

Παρακάτω ακολουθεί η απεικόνιση του παραθύρου αναζήτησης.

Εικόνα 4.8 Παράδειγμα φόρμας αναζήτησης με επιλογή κατηγορίας.

Παρακάτω ακολουθεί ενδεικτικά ένα μικρό κομμάτι κώδικα για την εκτέλεση της αναζήτησης.

```
SELECT * FROM products
WHERE cat_id = $cat_id
and (text LIKE '%$word%' OR marka LIKE '%$word%')
```

## 2. Αναζήτηση προϊόντος σε όλες τις κατηγορίες

Η αναζήτηση προϊόντος σε όλες τις κατηγορίες είναι η πιο απλή. Το μόνο που έχει να κάνει ο χρήστης είναι να πληκτρολογήσει μια περιγραφή ενός προϊόντος. Παρακάτω ακολουθεί η απεικόνιση του παραθύρου αναζήτησης.

Εικόνα 4.9 Παράδειγμα φόρμας αναζήτησης χωρίς επιλογή κατηγορίας.

Παρακάτω ακολουθεί ενδεικτικά ένα μικρό κομμάτι κώδικα για την εκτέλεση της αναζήτησης.

```
SELECT * FROM products
WHERE product_name
LIKE '%$word%'
OR text LIKE '%$word%'
OR marka LIKE '%$word%'
```

## 3. Αναζήτηση με Autosuggest.

Σε αυτή τη περίπτωση έχει χρησιμοποιηθεί η τεχνική AJAX κατά την οποία, καθώς ο χρήστης πληκτρολογεί ένα γράμμα στέλνεται ασύγχρονα ένα ερώτημα στη βάση δεδομένων και επιστρέφονται άμεσα τα αποτελέσματα του ερωτήματος δίχως να γίνει ανανέωση της ιστοσελίδας. Κάθε φορά που αλλάζει το μήκος της λέξης το περιεχόμενο της λέξης που πληκτρολογεί ο χρήστης επαναλαμβάνεται η διαδικασία. Αυτή η τεχνική χρησιμοποιήθηκε για να υπάρχει ένα πιο σύγχρονο και φιλικό περιβάλλον προς τους χρήστες – πελάτες του ηλεκτρονικού καταστήματος. Παρακάτω ακολουθεί η απεικόνιση του παραθύρου αναζήτησης.





Εικόνα 4.10 Παράδειγμα φόρμας αναζήτησης με AJAX auto suggestion.

Στην παραπάνω εικόνα φαίνεται πως αλλάζει δυναμικά η το παράθυρο με τα αποτελέσματα της αναζήτησης, καθώς αλλάζει το περιεχόμενο της λέξης προς αναζήτηση. Για την υλοποίηση αυτής της τεχνικής έγινε χρήση πολλών γλωσσών προγραμματισμού όπως είναι η xml, JavaScript ,MySQL ,php και html.

### 4.3 Περιγραφή εμφάνισης προϊόντων σε δευτερεύουσες σελίδες

Εκτός από τα προϊόντα που εμφανίζονται ως προτάσεις για αγορά στην αρχική σελίδα, υπάρχουν και άλλου είδους προτάσεις για εμφάνιση προϊόντων τα οποία βρίσκονται σε δευτερεύουσες σελίδες.

#### 4.3.1 Cross-Sales

Στη σελίδα ενός προϊόντος (πχ. `product_page.php?product=65&cat_id=3`) εμφανίζονται στην δεξιά στήλη, όπως απεικονίζεται παρακάτω, δύο προϊόντα τα οποία έχουν τις περισσότερες συνδυαζόμενες αγορές με το συγκεκριμένο προϊόν. Για τον υπολογισμό και εμφάνιση αυτών των δύο προϊόντων αρχικά ελέγχονται οι αγορές των χρηστών που βρίσκονται στην ίδια ομάδα (cluster) με τον χρήστη που έχει κάνει log in. Σε περίπτωση που το προϊόν της σελίδας δεν είναι αρκετά αρεστό από την ομάδα (cluster) του χρήστη, και δεν υπάρχουν συνδυαζόμενες αγορές, τότε ελέγχονται οι αγορές όλων των εγγεγραμμένων χρηστών του ηλεκτρονικού καταστήματος ανεξαρτήτου ομάδας(cluster). Αν πάλι δεν βρεθούν αποτελέσματα συνδυαζόμενων αγορών αυτού του προϊόντος, εμφανίζονται τυχαία προϊόντα της κατηγορίας που ανήκει το προϊόν.

Η μέθοδος αυτή ονομάζεται `cross_sales` και εφαρμόζεται στα περισσότερα ηλεκτρονικά καταστήματα. Στη περίπτωση της διπλωματικής μας χρησιμοποιείται με πιο πρωτοποριακό τρόπο, διότι εφαρμόζεται με μοντελοποίηση χρηστών. Η γενική ιδέα του `cross sales` είναι ότι όποιος έχει αγοράσει αυτό το προϊόν, αγόρασε και ένα άλλο προϊόν. Με αυτόν τον τρόπο παρακινείται ο χρήστης να μεγαλώσει την παραγγελία του βάζοντας στο καλάθι αγορών του περισσότερα προϊόντα κάποια από τα οποία δεν είχε σκεφτεί ή δεν ήταν ακόμα σίγουρος.

Παρακάτω παραθέτουμε ένα μικρό κομμάτι κώδικα με επεξήγηση του `cross-sales`.

1. Μαζεύουμε όλες τις αγορές που εμπεριέχουν αυτό το προϊόν εξαιρώντας τις αγορές που έχει κάνει ο χρήστης που επισκέπτεται τη σελίδα.  
**SELECT \* FROM orders**  
**WHERE product\_id= \$product\_id AND cust\_id != \$user\_id**
2. Συγκεντρώνονται όλες οι συνδυασμένες αγορές των υπόλοιπων χρηστών.  
**SELECT \* FROM orders**  
**WHERE cust\_id = \$cust\_id and product\_id != \$product\_id and Date ='\$date'**
3. Εισάγεται στον πίνακα `cross_sales` ο κωδικός προϊόντος, αριθμός τεμαχίων.  
**INSERT INTO cross\_sales (product\_id, count)**  
**VALUES (\$product\_cross\_id,\$count\_cross)**

4. Υπολογίζεται το προϊόν με τις περισσότερες αγορές και εμφανίζονται με φθίνουσα σειρά ώστε να διαλέξουμε τα δύο πρώτα για την εμφάνιση στην ιστοσελίδα.

**SELECT \*,SUM(count) As A FROM cross\_sales GROUP BY product\_id  
ORDER BY A DESC**

Παρακάτω ακολουθεί απεικόνιση των προϊόντων που προκύπτουν από τη μέθοδο cross-sales(αριστερό μέρος εικόνας 9.χ) και η απεικόνιση των τυχαίων προϊόντων σε περίπτωση μη εύρεσης συνδυασμένης αγοράς (δεξί μέρος εικόνας 9.χ).



Εικόνα 4.11 Παράδειγμα Cross-Sales(αριστερά επιτυχημένο cross-sales,δεξιά τυχαία προϊόντα.)

#### 4.3.2 Δημοφιλέστερα προϊόντα ανά υπέρ-κατηγορία

Κάθε υπέρ-κατηγορία(menu) όπως είναι τα computers, laptop, peripherals, games έχουν τη δική τους σελίδα που εμφανίζονται τα προϊόντα των υποκατηγοριών τους. Σε αυτή τη σελίδα εκτός από τα τυχαία προϊόντα της κάθε υποκατηγορίας εμφανίζονται και άλλα δυο τα όποια είναι τα πιο δημοφιλή αυτής της υπέρ-κατηγορίας.

- **Δημοφιλέστερο βάσει αγορών.**

Στη δεξιά κολώνα των υπέρ-κατηγοριών, το πρώτο προϊόν που εμφανίζεται, είναι το πιο δημοφιλές προϊόν αυτού του μενού βάσει των αγορών. Εμφανίζοντας αυτό το προϊόν στο χρήστη-πελάτη πετυχαίνουμε την προσέλκυση του και την παρότρυνση του, για αγορά ενός από τα πιο δημοφιλή προϊόντα του καταστήματος για τα οποία πιο πριν δεν γνώριζε την απήχηση τους στον υπόλοιπο κόσμο που επισκέπτεται το συγκεκριμένο ηλεκτρονικό κατάστημα. Παρακάτω ακολουθεί ένα μικρό κομμάτι κώδικα που εκτελείται για την εύρεση του προϊόντος με τις περισσότερες αγορές αυτού του menu από την βάση δεδομένων.

**SELECT \*, sum( count )  
FROM orders AS a  
JOIN products AS b ON a.product\_id = b.product\_id  
WHERE menu\_id = \$menu\_id**

**GROUP BY a.product\_id**

**ORDER BY menu\_id, sum( count ) DESC";**

- **Δημοφιλέστερο βάσει επισκεψιμότητας.**

Στη δεξιά κολώνα των υπέρ-κατηγοριών , το δεύτερο προϊόν που εμφανίζεται, είναι το πιο δημοφιλές προϊόν αυτού του μενού βάσει των επισκέψεων που δέχεται από τους εγγεγραμμένους χρήστες. Η επίσκεψη ενός προϊόντος μπορεί να μην έχει την ίδια βαρύτητα με την αγορά του, παρ' όλα αυτά είναι ένας σημαντικός δείκτης για την σημαντικότητα του. Εμφανίζοντας αυτό το προϊόν στο χρήστη-πελάτη πετυχαίνουμε την προσέλκυση και την παρότρυνση του, για αγορά ενός από τα πιο δημοφιλή προϊόντα του καταστήματος για τα οποία πιο πριν δεν γνώριζε την απήχηση τους στον υπόλοιπο κόσμο που επισκέπτεται το συγκεκριμένο ηλεκτρονικό κατάστημα. Παρακάτω ακολουθεί ένα μικρό κομμάτι κώδικα που εκτελείται για την εύρεση του προϊόντος με τις περισσότερες αγορές αυτού του menu από την βάση δεδομένων. Παρακάτω ακολουθεί ένα μικρό κομμάτι κώδικα που εκτελείται για την εύρεση του προϊόντος με τα περισσότερα clicks αυτού του menu από την βάση δεδομένων.

**SELECT \* , sum( clicks\_number )**

**FROM clicks AS a**

**JOIN products AS b ON a.product\_id = b.product\_id**

**WHERE menu\_id = \$menu\_id**

**GROUP BY a.product\_id**

**ORDER BY sum( clicks\_number ) DESC";**



**Εικόνα 4.12 Παράδειγμα προϊόντων most ordered(αριστερά)-most visited (δεξιά) για το menu laptop**

#### **4.4 Περιγραφή εμφάνισης προϊόντων στην κεντρική σελίδα**

Ο καλύτερος τρόπος για την προσέλκυση των πελατών είναι η παρουσίαση των καλύτερων προϊόντων στην αρχική σελίδα του ηλεκτρονικού καταστήματος. Η σελίδα αυτή είναι σαν τα φυλλάδια που μοιράζουν τα super market με τα φθηνότερα προϊόντα «κράχτες», μόνο που στην περίπτωση μας είναι τα πιο δημοφιλή. Επίσης αυτή η σελίδα θα μπορούσε να παρομοιαστεί με τα «stand» των super market που ανάλογα τη θέση τους στο κατάστημα οι εταιρίες κάνουν καλύτερες προσφορές για να το κατοχυρώσουν. Οι λόγοι αυτοί μας οδήγησαν στην προβολή των δημοφιλέστερων ή περισσότερο αγαπητών προϊόντων του καταστήματος στην πρώτη σελίδα. Τα προϊόντα αυτά χωρίζονται σε οχτώ κατηγορίες. Τα τέσσερα κεντρικά βασίζονται στη μοντελοποίηση χρηστών (για τα οποία θα γίνει εκτενέστερη περιγραφή σε άλλη παράγραφο) και είναι :

1. Προϊόν που ανήκει στην κατηγορία με τις περισσότερες αγορές των χρηστών που ανήκουν στο ίδιο "cluster".

2. Προϊόν που ανήκει στην κατηγορία με τη μεγαλύτερη επισκεψιμότητα των χρηστών που ανήκουν στο ίδιο "cluster".
3. Προϊόν με την καλύτερη κατά μέσω όρο βαθμολογία των χρηστών του ίδιου "cluster"
4. Το πιο πρόσφατο προϊόν από την κατηγορία με τις περισσότερες αγορές των χρηστών του ίδιου cluster.

Τα υπόλοιπα τέσσερα προϊόντα δεν βασίζονται στη μοντελοποίηση χρηστών, αλλά είναι τα πιο δημοφιλή προϊόντα όλου του καταστήματος. Σε αυτή τη περίπτωση το ρίσκο αποστροφής του πελάτη είναι μικρό, διότι τα προϊόντα αυτά είναι αρεστά από την πλειοψηφία των πελατών.

Τα προϊόντα αυτά είναι τα εξής:


- Προϊόν με τη μεγαλύτερη βαθμολογία κατά μέσω όρο.
- Προϊόν με τις περισσότερες αγορές όλων των χρηστών( Best Seller).
- Προϊόν με τη μεγαλύτερη επισκεψιμότητα ( Most Visited).
- Η νεότερη άφιξη προϊόντος στο ηλεκτρονικό κατάστημα.

Παρακάτω θα περιγραφούν αναλυτικότερα και τα οχτώ διαφορετικά είδη στοχευμένης διαφήμισης.

#### 4.4.1 Προϊόντα που δεν βασίζονται στη μοντελοποίηση χρηστών.

##### 1. Προϊόν με τη μεγαλύτερη βαθμολογία κατά μέσω όρο.

Στο δεξί πάνω άκρο της πρώτης σελίδας απεικονίζεται το προϊόν με τη μεγαλύτερη βαθμολογία κατά μέσω όρο, που έδωσαν μόνο οι εγγεγραμμένοι χρήστες. Αυτό το προϊόν έχει κρίνεται σημαντικό και άξιο εμφάνισης διότι η βαθμολογία του προέρχεται είτε επειδή ο πελάτης το αγόρασε και εκ του αποτελέσματος το βαθμολόγησε, είτε επειδή προσέλκυσε τον χρήστη να το βαθμολογήσει βάσει χαρακτηριστικών και τιμής. Ο κώδικας για την εμφάνιση του προϊόντος η και πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.

Product With the Most Votes	AVG	Count	Product_id
 790i (s775, DDR3) Price: 325.00 € <a href="#">Buy Now</a> Το Renko πρώτο στην Ελλάδα σας παρουσιάζει την νέα σειρά Motherboards από την φημισμένη eVGA. Η eVGA 790i SLI διακρίνεται...	5.0000	4	11
	5.0000	2	73
	5.0000	2	29
	5.0000	2	40
	5.0000	2	88
	5.0000	1	168
	5.0000	1	93
	5.0000	1	154
	5.0000	1	87
	5.0000	1	156
	5.0000	1	6

Εικόνα 4.13 Παράδειγμα προϊόντος συγκέντρωσης μεγαλύτερης βαθμολογίας (αριστερά η απεικόνιση στην ιστοσελίδα, δεξιά ένα κομμάτι του πίνακα με τη καλύτερη βαθμολογία .

Στον παρακάτω κώδικα sql ζητείται ο καλύτερος μέσος όρος με τους περισσότερους ψήφους.

```
SELECT avg( vote_rating ) , count( vote_rating ) , product_id
FROM poll
GROUP BY product_id
ORDER BY avg( vote_rating ) DESC , count( vote_rating ) DESC
```

Έχοντας το product\_id εκτελούμε την παρακάτω sql για την προσκόμιση όλων των στοιχείων του προϊόντος για την εμφάνιση του.

```
SELECT * FROM products WHERE product_id = 11
```

## 2. Προϊόν με τις περισσότερες αγορές όλων των χρηστών( Best Seller).

Στο κέντρο της δεξιάς στήλης της αρχικής σελίδας απεικονίζεται το προϊόν με τις περισσότερες αγορές χρηστών ανεξαρτήτου ομάδας cluster. Αυτό το προϊόν έχει ιδιαίτερη βαρύτητα διότι αντιπροσωπεύει τις πραγματικές επιθυμίες της πλειοψηφίας του αγοραστικού μας κοινού. Ο κώδικας για την εμφάνιση του προϊόντος και η πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.



Εικόνα 4.14 Παράδειγμα προϊόντος με τις περισσότερες αγορές (αριστερά η απεικόνιση στην ιστοσελίδα, δεξιά ένα κομμάτι του πίνακα με τον αριθμό πωλήσεων και product\_id).

Στον παρακάτω κώδικα sql ζητείται το άθροισμα των αγορών ανά προϊόν σε φθίνουσα σειρά.

```
SELECT product_id Product_id, sum( count ) Count
FROM orders
GROUP BY product_id
ORDER BY sum( count ) DESC
```

Έχοντας τον κωδικό του best seller προϊόντος εκτελούμε την παρακάτω sql για την προσκόμιση όλων των στοιχείων του προϊόντος για την εμφάνιση του.

```
SELECT * FROM products
WHERE product_id = 20
```

## 3. Προϊον με τη μεγαλύτερη επισκεψιμότητα (περισσότερα clicks)

Στο κάτω μέρος της δεξιά στήλης απεικονίζεται το προϊόν με τη μεγαλύτερη επισκεψιμότητα βάσει των click που καταμετρώνται κάθε φορά που ο χρήστης επισκέπτεται ένα προϊόν. Αυτή η κατηγορία διαφήμισης μπορεί να μην έχει την ίδια βαρύτητα με την βαθμολογία και την αγορά ενός προϊόντος, αλλά δείχνει το ενδιαφέρον (έστω και μικρότερο) των πελατών και οφείλουμε να το λάβουμε υπόψιν μας. Ο κώδικας για την εμφάνιση του προϊόντος και η πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.

Στον παρακάτω κώδικα sql ζητείται το άθροισμα των επισκέψεων ανά προϊόν σε φθίνουσα σειρά.

```
SELECT product_id Product, cat_id Category, sum( clicks_number ) Sum_Clicks
FROM clicks
GROUP BY product_id
ORDER BY sum( clicks_number ) DESC
```

Έχοντας τον κωδικό του best visited προϊόντος εκτελούμε την παρακάτω sql για την προσκόμιση όλων των στοιχείων του προϊόντος για την εμφάνιση του.

**SELECT \* FROM products  
WHERE product\_id = 52**

**Product With the Most Clicks**



**AMD Phenom II X4 940**

Price: 309.00 €

[Buy Now](#)

Η AMD παρουσιάζει την εξέλιξη των τετραπύρηνων επεξεργαστών Phenom! Η προηγμένη γενιά Phenom II -με την κωδική ο...

Product	Category	Sum_Clicks
52	1	75
78	11	65
69	10	60
7	2	50
4	1	48
11	2	41
53	1	38
20	10	34
58	2	33
29	11	33

Εικόνα 4.15 Παράδειγμα προϊόντος με τις περισσότερες επισκέψεις (αριστερά η απεικόνιση στην ιστοσελίδα, δεξιά ένα κομμάτι του πίνακα με τον αριθμό επισκέψεων και product\_id,cat\_id) .


#### 4. Η νεότερη άφιξη προϊόντος στο ηλεκτρονικό κατάστημα.

Στο πάνω άκρο της κεντρική στήλης έχει τοποθετηθεί το πιο πρόσφατο προϊόν του καταστήματος. Εμφανίζοντας κάθε μέρα τις καινούριες αφίξεις πετυχαίνουμε την εφαρμογή διαφήμισης όπως συμβαίνει με τα φυλλάδια στα σουπερ μάρκετ ή όπως συμβαίνει στα περισσότερα e-shop. Εμφανίζεται το πιο καινούριο προϊόν για την προσέλκυση με σκοπό την επίσκεψη του και αγορά του.

Στον παρακάτω κώδικα sql ζητείται το προϊόν με την πιο πρόσφατη ημερομηνία καταχώρησης.

**SELECT \* FROM products  
ORDER BY DateAdded DESC**

**Most Recent Product**



**Asus P5Q3 (5775, DDR)**

Το Renko πρώτο στην Ελλάδα σας παρουσιάζει την νέα σειρά Motherboards από την φημισμένη Asus. Το μοντέλο P5Q3 διακρίνεται για το overlocking, την απόδοση, καθώς και το πλούσιο λ...

135.00€

Product	Date
58	2011-04-05 01:15:41
Sort	2011-04-05 00:00:00
122	2010-09-29 16:57:10
121	2010-09-29 16:56:53

Εικόνα 4.16 Παράδειγμα προϊόντος με την πιο πρόσφατη ημερομηνία καταχώρησης (αριστερά όπως απεικονίζεται στο e-shop και δεξιά το αποτέλεσμα αναζήτησης στη βάση δεδομένων).

#### 4.4.2 Προϊόντα που βασίζονται στη μοντελοποίηση χρηστών.

Στο κέντρο της αρχικής σελίδας εμφανίζονται τα προϊόντα που βασίζονται στην μοντελοποίηση χρηστών. Τα προϊόντα αυτά προσαρμόζονται δυναμικά σύμφωνα με τις προτιμήσεις του εκάστοτε χρήστη. Αν με την πάροδο του καιρού ο χρήστης αλλάξει συνήθειες τότε η μοντελοποίηση τον κατατάσσει σε άλλη ομάδα cluster με αποτέλεσμα να αλλάζουν και τα προϊόντα εμφάνισης. Τα προϊόντα αυτά έχουν ιδιαίτερη βαρύτητα διότι κατά μεγάλο ποσοστό ανήκουν στις αγαπημένες κατηγορίες προϊόντων του εγγεγραμμένου χρήστη. Το κεντρικό

σημείο της ιστοσελίδας είναι αυτό που κεντρίζει το ενδιαφέρον του πελάτη και αυτός είναι ο λόγος που έχουν τοποθετηθεί σε αυτό το σημείο. Παρακάτω θα περιγραφεί ο τρόπος με τον οποίο διαλέγονται προς εμφάνιση τα προϊόντα αυτά.

### 1. Προϊόν που ανήκει στην κατηγορία με τις περισσότερες αγορές των χρηστών που ανήκουν στο ίδιο “cluster”.

Το προϊόν αυτό προέρχεται από την κατηγορία με τις περισσότερες αγορές των εγγεγραμμένων χρηστών που ανήκουν στην ίδια ομάδα “cluster”. Κάθε φορά που ανανεώνεται η σελίδα αυτή, εμφανίζεται ένα τυχαίο προϊόν από την κατηγορία αυτή εκτός των προϊόντων που έχουν είδη αποκτηθεί από το συνδεδεμένο χρήστη. Με αυτό τον τρόπο ο χρήστης λαμβάνει διαφημίσεις για προϊόντα που κατά μεγάλο ποσοστό τον ενδιαφέρουν και ενημερώνεται άμεσα για αυτά. Επίσης δεν επαναλαμβάνονται οι διαφημίσεις των είδη αποκτηθέντων προϊόντων με σκοπό την αποφυγή της αποστροφής του χρήστη, αντικρίζοντας συνεχώς προϊόντα που πλέον δεν του προκαλούν κάποιο ενδιαφέρον. Βέβαια σε περίπτωση που επιθυμεί να αποκτήσει ξανά κάποιο προϊόν μπορεί απλά να το επισκεφτεί και να το προσθέσει στο καλάθι αγορών του.

Σε περίπτωση που ο χρήστης δεν έχει καταταχτεί σε κάποια ομάδα “cluster” τότε εμφανίζεται ένα τυχαίο προϊόν ανεξαρτήτου κατηγορίας. Το ίδιο ισχύει και για έναν μη εγγεγραμμένο χρήστη για τον οποίο δεν υπάρχουν στοιχεία στη βάση και δεν μπορεί να πάρει μέρος στη μοντελοποίηση.

Ο κώδικας για την εμφάνιση του προϊόντος η και πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.



Εικόνα 4.17 Παράδειγμα προϊόντος από την best seller κατηγορία των χρηστών του ίδιου cluster.

Στον παρακάτω κώδικα sql ζητείται η κατηγορία με τις περισσότερες αγορές, από τον πίνακα αγορών των user του ίδιου cluster(clustered\_order). Ο πίνακας clustered\_orders δημιουργείται κατά την εκτέλεση του αλγορίθμου clustering, ο οποίος δεν εκτελείται κατά την είσοδο του χρήστη για την αποφυγή καθυστερήσεων φόρτωσης.

```
SELECT cat_id , sum( count ) AS TOTAL
FROM `clustered_orders`
WHERE customer_id = $user
GROUP BY cat_id
ORDER BY `TOTAL` DESC
```

Η κατηγορία αυτή αποθηκεύεται σε μία session μεταβλητή που διαρκεί μέχρι ο χρήστης να αποσυνδεθεί από τη σελίδα.

**\$\_SESSION['best\_cat\_id'];**

Τέλος, γίνεται η αναζήτηση στον πίνακα προϊόντων με παράμετρο τη παραπάνω session μεταβλητή και επιστρέφονται όλα τα στοιχεία του προϊόντος.

**SELECT \* FROM products**

**WHERE cat\_id=\$\_SESSION['best\_cat\_id'];**

## 2. Προϊόν που ανήκει στην κατηγορία με την μεγαλύτερη επισκεψιμότητα των χρηστών που ανήκουν στο ίδιο “cluster”.

Το προϊόν αυτό προέρχεται από την κατηγορία με την μεγαλύτερη επισκεψιμότητα των εγγεγραμμένων χρηστών που ανήκουν στην ίδια ομάδα “cluster”. Κάθε φορά που ανανεώνεται η σελίδα, εμφανίζεται ένα τυχαίο προϊόν από την κατηγορία αυτή εκτός των προϊόντων που έχουν είδη αποκτηθεί από το συνδεδεμένο χρήστη. Με αυτό τον τρόπο ο χρήστης λαμβάνει διαφημίσεις για προϊόντα που κατά μεγάλο ποσοστό τον ενδιαφέρουν και ενημερώνεται άμεσα για αυτά. Επίσης δεν επαναλαμβάνονται οι διαφημίσεις των είδη αποκτηθέντων προϊόντων με σκοπό την αποφυγή της αποστροφής του χρήστη, αντικρίζοντας συνεχώς προϊόντα που πλέον δεν του προκαλούν κάποιο ενδιαφέρον. Βέβαια σε περίπτωση που επιθυμεί να αποκτήσει ξανά κάποιο προϊόν μπορεί απλά να το επισκεφτεί και να το προσθέσει στο καλάθι αγορών του.

Σε περίπτωση που ο χρήστης δεν έχει καταταχτεί σε κάποια ομάδα “cluster” τότε εμφανίζεται ένα τυχαίο προϊόν ανεξαρτήτου κατηγορίας. Το ίδιο ισχύει και για έναν μη εγγεγραμμένο χρήστη για τον οποίο δεν υπάρχουν στοιχεία στη βάση και δεν μπορεί να πάρει μέρος στη μοντελοποίηση.

Ο κώδικας για την εμφάνιση του προϊόντος η και πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.



Εικόνα 4.18 Παράδειγμα προϊόντος από την best visited κατηγορία των χρηστών του ίδιου cluster.

Στον παρακάτω κώδικα sql ζητείται η κατηγορία με τα περισσότερα click, από τον πίνακα click των user του ίδιου cluster(clustered\_clicks). Ο πίνακας clustered\_clicks δημιουργείται κατά την εκτέλεση του αλγορίθμου clustering, ο οποίος δεν εκτελείται κατά την είσοδο του χρήστη για την αποφυγή καθυστερήσεων φόρτωσης.

**SELECT cat\_id, sum( count ) AS TOTAL**

**FROM `clustered\_clicks`**

**WHERE customer\_id = \$user**

**GROUP BY cat\_id**

**ORDER BY `TOTAL` DESC**

Η κατηγορία αυτή αποθηκεύεται σε μία session μεταβλητή που διαρκεί μέχρι ο χρήστης να αποσυνδεθεί από τη σελίδα.



**\$ \_SESSION['best\_click\_cat\_id'];**

Τέλος, γίνεται η αναζήτηση στον πίνακα προϊόντων με παράμετρο τη παραπάνω session μεταβλητή και επιστρέφονται όλα τα στοιχεία του προϊόντος.

**SELECT \* FROM products**

**WHERE cat\_id=\$ \_SESSION['best\_click\_cat\_id'];**

### 3. Προϊόν με την καλύτερη κατά μέσω όρο βαθμολογία των χρηστών του ίδιου “cluster”

Το προϊόν αυτό προέρχεται από την κατηγορία με την μεγαλύτερη κατά μέσω όρο βαθμολογία των εγγεγραμμένων χρηστών που ανήκουν στην ίδια ομάδα “cluster”. Κάθε φορά που ανανεώνεται η σελίδα, εμφανίζεται το ίδιο προϊόν από την κατηγορία αυτή εκτός των προϊόντων που έχουν είδη αποκτηθεί από το συνδεδεμένο χρήστη. Με αυτό τον τρόπο ο χρήστης λαμβάνει διαφημίσεις για προϊόντα που κατά μεγάλο ποσοστό τον ενδιαφέρουν και ενημερώνεται άμεσα για αυτά. Επίσης δεν επαναλαμβάνονται οι διαφημίσεις των είδη αποκτηθέντων προϊόντων με σκοπό την αποφυγή της αποστροφής του χρήστη, αντικρίζοντας συνεχώς προϊόντα που πλέον δεν του προκαλούν κάποιο ενδιαφέρον. Βέβαια σε περίπτωση που επιθυμεί να αποκτήσει ξανά κάποιο προϊόν μπορεί απλά να το επισκεφτεί και να το προσθέσει στο καλάθι αγορών του.

Σε περίπτωση που ο χρήστης δεν έχει καταταχτεί σε κάποια ομάδα “cluster” τότε εμφανίζεται ένα τυχαίο προϊόν ανεξαρτήτου κατηγορίας. Το ίδιο ισχύει και για έναν μη εγγεγραμμένο χρήστη για τον οποίο δεν υπάρχουν στοιχεία στη βάση και δεν μπορεί να πάρει μέρος στη μοντελοποίηση.

Ο κώδικας για την εμφάνιση του προϊόντος η και πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.



Εικόνα 4.19 Παράδειγμα προϊόντος με την μεγαλύτερο μέσω όρο βαθμολογίας που έδωσαν οι χρήστες του ίδιου cluster.

Στον παρακάτω κώδικα sql ζητείται το προϊόν με τη μεγαλύτερη βαθμολογία κατά μέσω όρο, από τον πίνακα roll των user του ίδιου cluster(clustered\_votes). Ο πίνακας clustered\_votes δημιουργείται κατά την εκτέλεση του αλγορίθμου clustering, ο οποίος δεν εκτελείται κατά την είσοδο του χρήστη για την αποφυγή καθυστερήσεων φόρτωσης.

**SELECT \* , avg( vote ) , count( vote )**

**FROM `clustered\_prod\_votes`**

**WHERE customer\_id =\$user**

**GROUP BY product\_id**

**ORDER BY avg( vote ) DESC , count( vote ) DESC**

Το προϊόν αυτό αποθηκεύεται σε μία session μεταβλητή που διαρκεί μέχρι ο χρήστης να αποσυνδεθεί από τη σελίδα.

**\$\_SESSION['clustered\_best\_illected']**

Τέλος, γίνεται η αναζήτηση στον πίνακα προϊόντων με παράμετρο τη παραπάνω session μεταβλητή και επιστρέφονται όλα τα στοιχεία του προϊόντος.

```
SELECT * FROM products
```

```
WHERE product_id = $_SESSION['clustered_best_illected']
```

#### 4. Το πιο πρόσφατο προϊόν από την κατηγορία με τις περισσότερες αγορές των χρηστών του ίδιου cluster.

Το προϊόν αυτό είναι το πιο νεοεισαχθέν της κατηγορίας με τις περισσότερες πωλήσεις των χρηστών που ανήκουν στο cluster του εκάστοτε συνδεδεμένου χρήστη. Αξίζει να σημειωθεί ότι εάν έχει ήδη αποκτηθεί αυτό το προϊόν, τότε εμφανίζεται το αμέσως επόμενο ημερομηνικά από την ίδια κατηγορία. Με αυτό τον τρόπο ο χρήστης λαμβάνει διαφημίσεις για προϊόντα που κατά μεγάλο ποσοστό τον ενδιαφέρουν και ενημερώνεται για τις νέες κυκλοφορίες άμεσα. Επίσης δεν επαναλαμβάνονται οι διαφημίσεις των είδη αποκτηθέντων προϊόντων με σκοπό την αποφυγή της αποστροφής του χρήστη, αντικρίζοντας συνεχώς προϊόντα που πλέον δεν του προκαλούν κάποιο ενδιαφέρον. Βέβαια σε περίπτωση που επιθυμεί να αποκτήσει ξανά κάποιο προϊόν μπορεί απλά να το επισκεφτεί και να το προσθέσει στο καλάθι αγορών του.

Σε περίπτωση που ο χρήστης δεν έχει καταταχτεί σε κάποια ομάδα “cluster” τότε εμφανίζεται ένα τυχαίο προϊόν ανεξαρτήτου κατηγορίας. Το ίδιο ισχύει και για έναν μη εγγεγραμμένο χρήστη για τον οποίο δεν υπάρχουν στοιχεία στη βάση και δεν μπορεί να πάρει μέρος στη μοντελοποίηση.

Ο κώδικας για την εμφάνιση του προϊόντος η και πραγματική του απεικόνιση στην ιστοσελίδα παρατίθενται παρακάτω.



Εικόνα 4.20 Παράδειγμα νεοεισαχθέντος προϊόντος από την best seller κατηγορία των χρηστών του ίδιου cluster..

Στον παρακάτω κώδικα sql ζητείται το νεότερο προϊόν βάσει ημερομηνίας καταχώρισης από την best seller κατηγορία των χρηστών του ίδιου cluster. Ουσιαστικά αναζητά την κατηγορία με τις περισσότερες αγορές στον πίνακα (clustered\_order). Ο πίνακας clustered\_orders δημιουργείται κατά την εκτέλεση του αλγορίθμου clustering, ο οποίος δεν εκτελείται κατά την είσοδο του χρήστη για την αποφυγή καθυστερήσεων φόρτωσης.

```
SELECT cat_id , sum( count ) AS TOTAL
```

```
FROM `clustered_orders`
```

```
WHERE customer_id = $user
```

```
GROUP BY cat_id
```

```
ORDER BY `TOTAL` DESC
```

Η κατηγορία αυτή αποθηκεύεται σε μία session μεταβλητή που διαρκεί μέχρι ο χρήστης να αποσυνδεθεί από τη σελίδα.

```
$_SESSION['best_cat_id'];
```

Τέλος, γίνεται η αναζήτηση στον πίνακα προϊόντων με παράμετρο τη παραπάνω session μεταβλητή και επιστρέφονται όλα τα στοιχεία του προϊόντος.

```
SELECT * FROM products  
WHERE cat_id = $_SESSION['best_cat_id']  
ORDER BY DateAdded DESC
```

## 4.5 Εμφάνιση τρόπων προσαρμοστικότητας του ηλεκτρονικού καταστήματος

Σε αυτή τη παράγραφο περιγράφουμε τους δύο τρόπους με τους οποίους το ηλεκτρονικό κατάστημα αυτό-προσαρμόζεται ανάλογα των συνηθειών τους.

### 4.5.1 Αλλαγή σειράς κατηγοριών στο κεντρικό μενού.

Γνωρίζοντας σε ποια ομάδα ανήκει ο χρήστης που έχει κάνει log in, γίνεται η ταξινόμηση σειράς εμφάνισης των κατηγοριών στο μενού. Με αυτό τον τρόπο διευκολύνεται ο χρήστης ώστε να βρίσκει ευκολότερα τα προϊόντα που τον ενδιαφέρουν περισσότερο. Παρακάτω ακολουθούν εικόνες για την κατανόηση των παραπάνω, καθώς και ένα κομμάτι κώδικα που εκτελείται.

Computers	Laptop	Peripherals	Games	Shopping Cart
-----------	--------	-------------	-------	---------------

Εικόνα 4.21 Αρχικοποιημένη σειρά εμφάνισης του μενού

Games	Laptop	Peripherals	Computers	Shopping Cart
-------	--------	-------------	-----------	---------------

Εικόνα 4.22 Σειρά εμφάνισης μετά από την είσοδο ενός χρήστη X

Laptop	Peripherals	Computers	Games	Shopping Cart
--------	-------------	-----------	-------	---------------

Εικόνα 4.23 Σειρά εμφάνισης μετά από την είσοδο ενός χρήστη Y

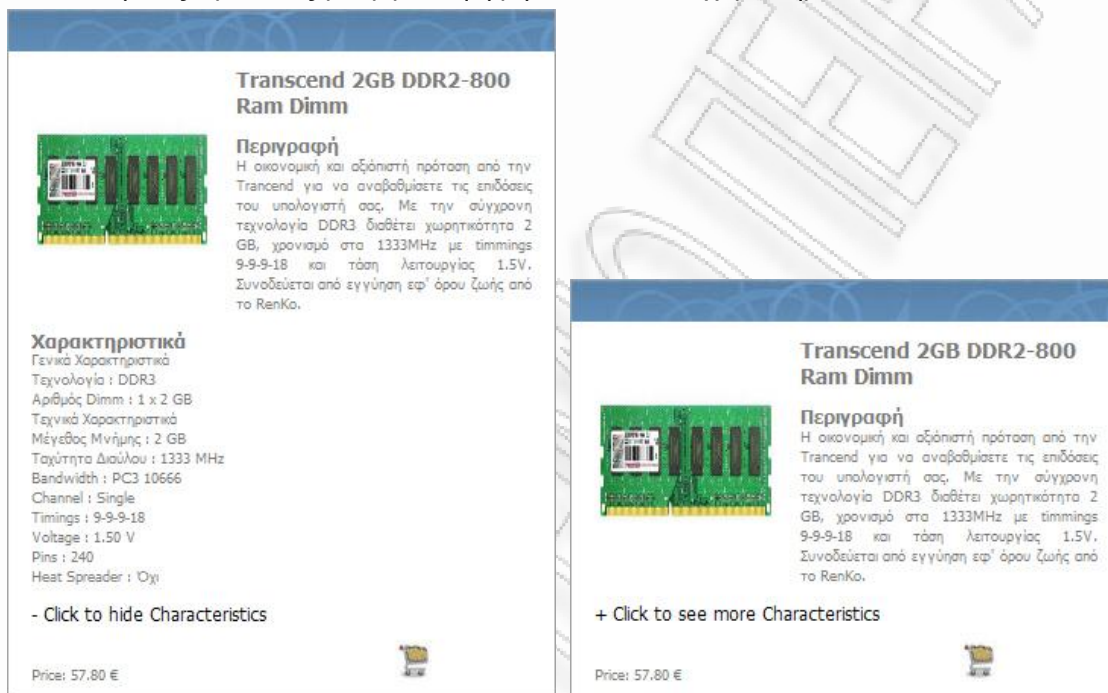
Κατηγορίες εμφανίζονται από αριστερά προς δεξιά με φθίνουσα σημαντικότητα από την οπτική του χρήστη.

```
$query7 = "SELECT menu_id , sum( count ) AS TOTAL  
FROM `clustered_orders`  
WHERE customer_id = $user  
GROUP BY menu_id  
ORDER BY `TOTAL` DESC";
```

```
$ret7 = mysql_query($query7);  
$num_of_rows = mysql_num_rows($ret7);  
$i=0;  
while( $row7 = mysql_fetch_array($ret7))  
{  
    $_SESSION['menu_id'][$i] = $row7['menu_id'];  
// echo "Menu Id : ";    echo $menu_id[$i]; echo "<br />";  
    $i++;  
}
```

#### 4.5.2 Αυτόματη εμφάνιση ή μη λεπτομερειών ενός προϊόντος.

Στην σελίδα με τις πληροφορίες ενός προϊόντος, ο χρήστης έχει την επιλογή εάν επιθυμεί να δει τις πιο λεπτομερείς περιγραφές με αναλυτικότερα στοιχεία. Τα στοιχεία αυτά κρίνονται ως πιο τεχνικά με αποτέλεσμα να μην τραβούν το ενδιαφέρον όλων των πελατών. Για το λόγο αυτό αρχικά οι λεπτομέρειες δεν εμφανίζονται παρά μόνο αν ο χρήστης επιλέξει τον σύνδεσμο “[+ Click to see more Characteristics](#)”. Κάθε φορά που επιλέγεται ο σύνδεσμος αυτός προστίθεται στη βάση +1 στον πεδίο advanced user του πίνακα των customer. Εάν ξεπεράσει έναν εύλογο αριθμό το πεδίο advance\_user ( για παράδειγμα γίνει ίσο με 30) τότε ο χρήστης χρίζεται ως advanced. Από το σημείο αυτό κάθε φορά που επισκέπτεται ένα προϊόν ο χρήστης αυτός, του παρέχονται όλες οι λεπτομέρειες χωρίς να επιλέξει κάποιο σύνδεσμο. Παρακάτω ακολουθεί η απεικόνιση ενός προϊόντος με εμφάνιση ή μη των επιπλέον χαρακτηριστικών.



Εικόνα 4.21 Εμφάνιση ή μη εξειδικευμένων χαρακτηριστικών

Στον κώδικα που ακολουθεί, προστίθεται 1 flag για κάθε επιλογή του συνδέσμου “[+ Click to see more Characteristics](#)”.

```
if ($expand>=1){
    $expand = $expand + 1;
    $query = "UPDATE customer SET expand = $expand
            WHERE id = $user_id";
    $ret = mysql_query($query);
} else {
    $query = "UPDATE customer SET expand = 1
            WHERE id = $user_id";
    $ret = mysql_query($query);
}
```

Στον κώδικα που ακολουθεί ορίζεται ένας χρήστης ως advanced\_user εάν ξεπεράσει τις 30 επιλογές του συνδέσμου “[+ Click to see more Characteristics](#)”.

```
if ($expand_times>=30){
    $query = "UPDATE customer SET advuser = 1 WHERE id = $user_id";
    $ret = mysql_query($query); }
```

## 5. ΧΡΗΣΗ ΑΛΓΟΡΙΘΜΩΝ – ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ

Σε αυτό το κεφάλαιο θα αναλυθούν λεπτομερώς όλα τα βήματα που ακολουθήθηκαν για την εξόρυξη δεδομένων από αλγόριθμους ιεραρχικής ομαδοποίησης, καθώς και η εκμετάλλευση αυτών των δεδομένων για την δημιουργία ενός recommender system.

### 5.1 Χρήση αλγορίθμων ιεραρχικής ομαδοποίησης

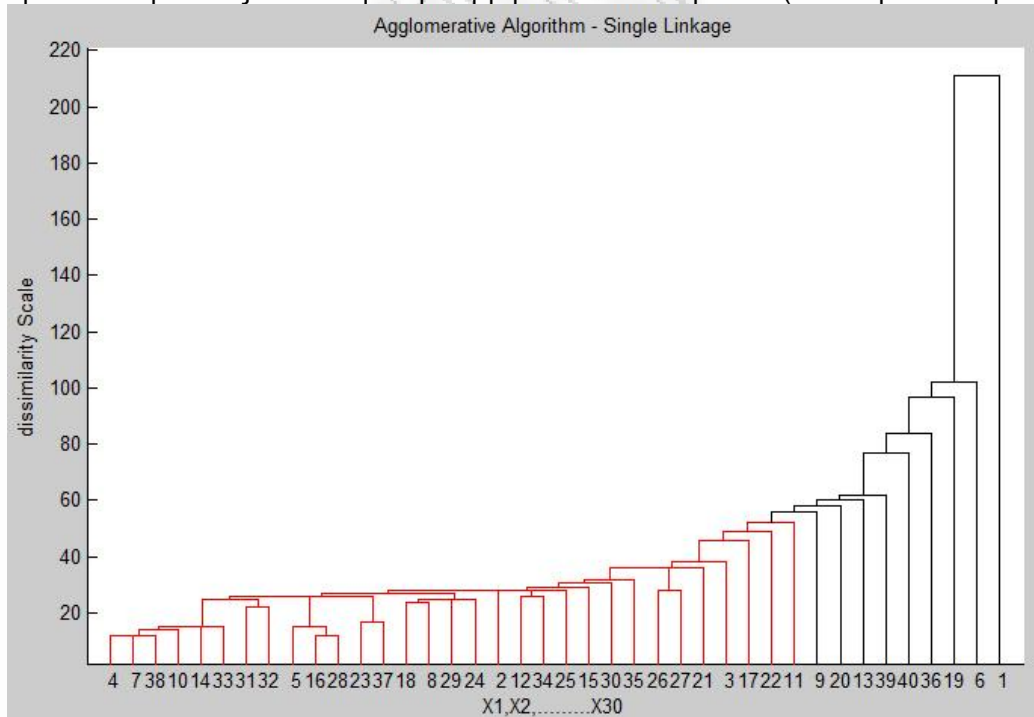
Για την υλοποίηση του παρόντος recommender συστήματος δοκιμάστηκαν τέσσερις αλγόριθμοι ιεραρχικής ομαδοποίησης οι οποίοι οι πιο δημοφιλείς που χρησιμοποιούνται στα περισσότερα στατιστικά πακέτα. Οι αλγόριθμοι αυτοί διαφέρουν μεταξύ τους από τη μέθοδο με την οποία υπολογίζεται η απόσταση μεταξύ των στοιχείων των ομάδων. Οι μέθοδοι αυτοί είναι :

1. Η μέθοδος του ελαχίστου (Min ή Single Linkage),
2. Η μέθοδος πλήρους συνδεσιμότητας (Max ή Complete Linkage),
3. Η μέθοδος Μέσου όρου ομάδας (Average Linkage),
4. Η μέθοδος Ward.

#### 5.1.1 Η μέθοδος του ελαχίστου (Min ή Single Linkage)

Η μέθοδος single linkage, γνωστή και ως ομαδοποίηση κοντινότερου γείτονα (nearest-neighbor clustering), έχει πολλές επιθυμητές θεωρητικές ιδιότητες, ωστόσο στην πράξη φαίνεται να μην έχει το ίδιο επιθυμητά αποτελέσματα με τις άλλες μεθόδους. Στην μέθοδο αυτή, ως απόσταση μεταξύ δύο ομάδων θεωρούμε την απόσταση με την μικρότερη τιμή (από όλους τους πιθανούς συνδυασμούς στοιχείων μεταξύ των δύο ομάδων).

Το βασικό μειονέκτημα της μεθόδου του ελαχίστου είναι η λεγόμενη επίπτωση αλυσίδας (chaining effect), η οποία οφείλεται στον τοπικό χαρακτήρα του κριτηρίου. Έτσι η μέθοδος είναι ευαίσθητη σε δεδομένα θορύβου τα οποία μπορεί να οδηγήσουν στην δημιουργία μεγάλων ομάδων. Η μέθοδος ευνοεί την εύρεση μη ελλειπτικών ομάδων (non-elliptical shape clusters).

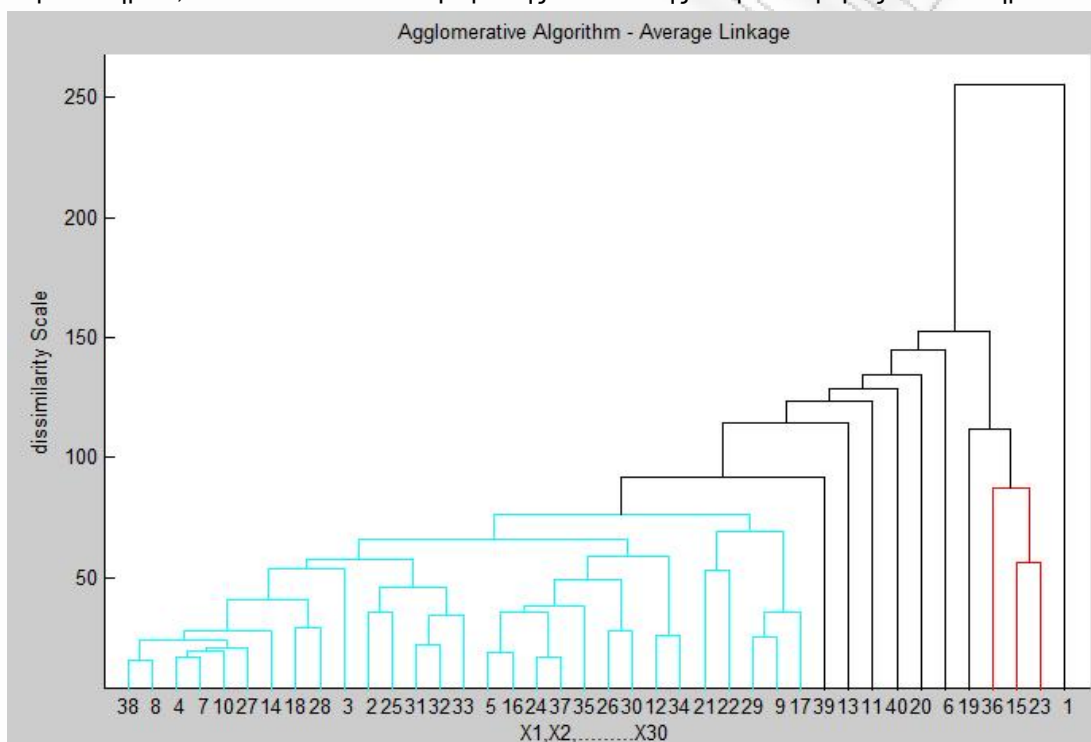


Εικόνα 5.1 Χρήση της μεθόδου ελαχίστου(single-linkage) στο e-shop με δείγμα 72 πελατών

Στην εικόνα 5.1 απεικονίζεται το δένδρογραμμα, της μεθόδου ελαχίστου( ή αλλιώς single linkage) ενός δείγματος 72 πελατών του ηλεκτρονικού μας καταστήματος. Παρατηρείται πολύ εύκολα το chaining effect για το οποίο έγινε αναφορά προηγουμένως, με αποτέλεσμα να μην παρατηρούνται διακριτές ομάδες χρηστών. Για το λόγο αυτό η μέθοδος single linkage απορρίφθηκε.

### 5.1.2 Η μέθοδος Μέσου όρου ομάδας (Average Linkage)

Η μέθοδος average linkage, αποτελεί την μέση οδό μεταξύ των δύο μεθόδων(single και complete linkage) αφού προσπαθεί να συμβιβάσει την ευαισθησία της complete linkage στις ακραίες τιμές και την τάση της single linkage να σχηματίζει μεγάλες μη σφαιρικές ομάδες που δεν συμβαδίζουν ακριβώς με την πραγματικότητα. Ως απόσταση μεταξύ των δύο ομάδων θεωρούμε την μέση απόσταση τους. Η μέθοδος average linkage έχει την τάση να δημιουργεί ομάδες με μικρές αποκλίσεις, παρόλα αυτά μειονεκτεί στο ότι επηρεάζεται λιγότερο από τα ακραία σημεία, διότι κατά τον υπολογισμό της απόστασης παίρνουν μέρος όλα τα σημεία.



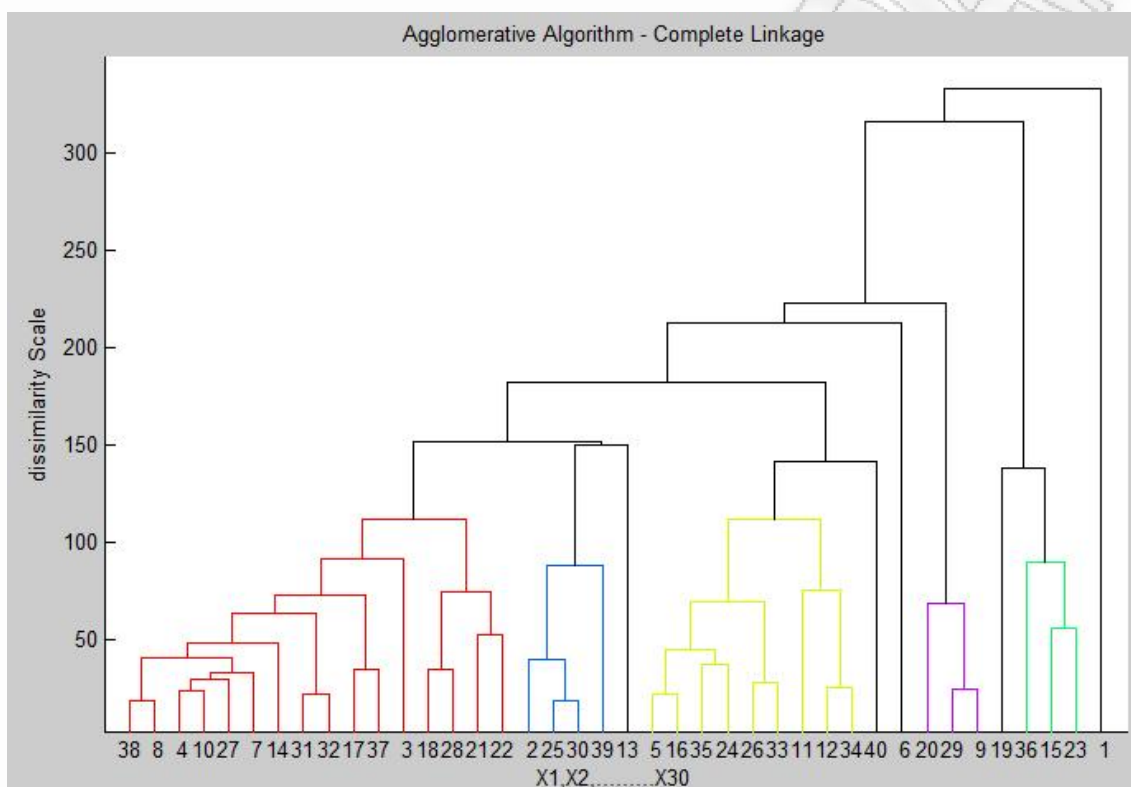
Εικόνα 5.2 Χρήση της μεθόδου μέσου όρου ομάδας (average linkage) στο e-shop με δείγμα 72 πελατών.

Στην εικόνα 5.2 απεικονίζεται το δένδρογραμμα, της μεθόδου μέσου όρου ομάδας( ή αλλιώς average linkage) ενός δείγματος 72 πελατών του ηλεκτρονικού μας καταστήματος. Παρατηρείται ότι υπάρχει μια σαφώς βελτιωμένη εικόνα διακριτών ομάδων χρηστών, ωστόσο μετά από κάποια τιμή ανομοιομορφίας και πάνω δημιουργείται και πάλι το chain effect. Επίσης, λόγω του chain effect αν για παράδειγμα επιλεχθεί ως τιμή ανομοιομορφίας η τιμή 70 τότε πολλοί από τους χρήστες δεν έχουν ενταχθεί σε κάποια ομάδα με κοινά χαρακτηριστικά, με αποτέλεσμα την λανθασμένη πρόταση προϊόντων. Η μέθοδος αυτή δεν παρουσιάζει μεγάλη βελτίωση σε σχέση με την προηγούμενη, για το e-shop της διπλωματικής οπότε απορρίπτεται.

### 5.1.3 Η μέθοδος πλήρους συνδεσιμότητας (Max ή Complete Linkage)

Στη μέθοδο complete linkage η απόσταση μεταξύ δύο ομάδων είναι εκείνη με την μεγαλύτερη τιμή από όλες τις πιθανές αποστάσεις μεταξύ όλων των στοιχείων των ομάδων αυτών.

Ενώ η μέθοδος πλήρους συνδεσιμότητας δεν πάσχει από το chaining effect το οποίο εμφανίζεται στη μέθοδο single linkage, εντούτοις είναι περισσότερο ευαίσθητο σε ακραίες τιμές (outliers). Ανεξάρτητα από αυτό, έχει παρατηρηθεί ότι η συγκεκριμένη μέθοδος παράγει καλύτερες και πιο συμπυκνωμένες ομάδες από τις αντίστοιχες της μεθόδου single linkage.



Εικόνα 5.3 Χρήση της μεθόδου πλήρους συνδεσιμότητας (complete linkage) στο e-shop με δείγμα 72 πελατών.

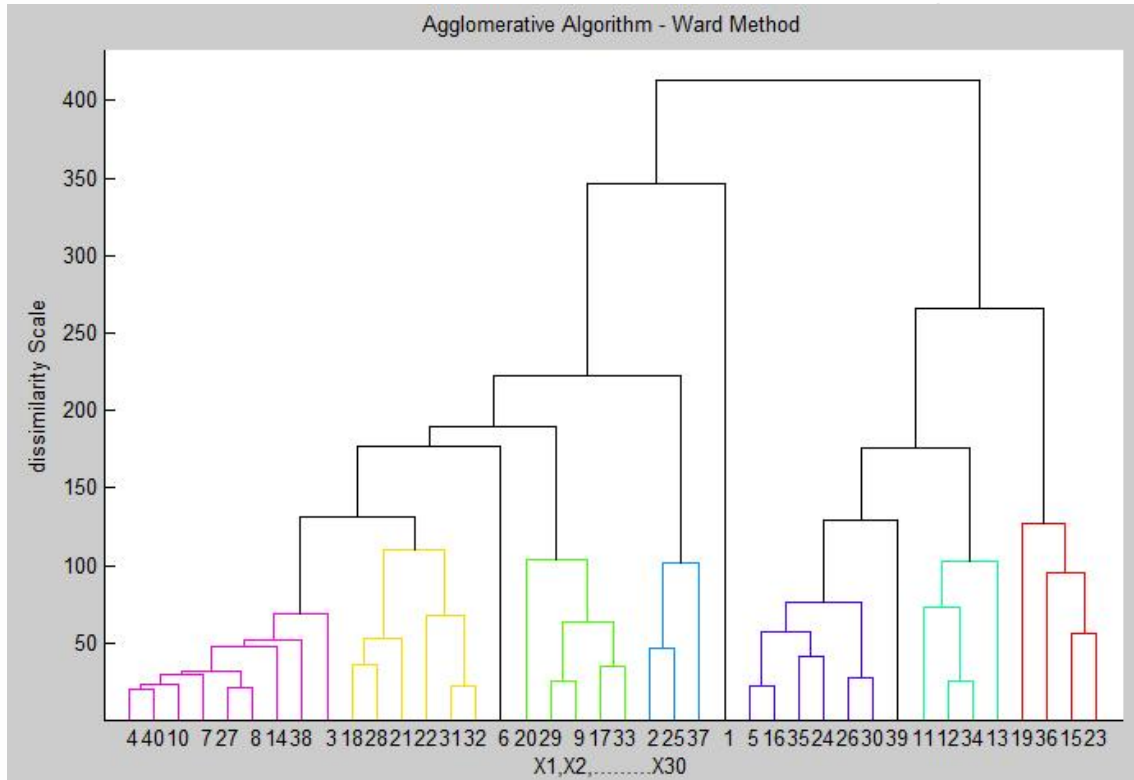
Στην εικόνα 5.3 απεικονίζεται το δένδρογραμμα, της μεθόδου πλήρους συνδεσιμότητας (ή complete linkage) ενός δείγματος 72 πελατών του ηλεκτρονικού μας καταστήματος. Παρατηρείται ότι έχουν δημιουργηθεί συμπαγείς ομάδες χρηστών με κοινά χαρακτηριστικά, το οποίο βοηθά πάρα πολύ στην διαλογή των ομάδων και στην δημιουργία αντιπροσωπευτικότερων προτάσεων-διαφημίσεων σε κάθε ομάδα χρηστών. Τα αποτελέσματα αυτού του αλγόριθμου σαφώς είναι πιο ικανοποιητικά από των άλλων δύο με αποτέλεσμα το recommender system να προτείνει αγαθά ή υπηρεσίες ασφαλέστερα.

### 5.1.4 Η μέθοδος Ward

Στη μέθοδο Ward η απόσταση μεταξύ δύο ομάδων ορίζεται ως η αύξηση στο άθροισμα των τετραγωνικών αποκλίσεων (αποστάσεων) κάθε στοιχείου των ομάδων, που προκύπτει από την ένωση των ομάδων αυτών.

Στην εικόνα 5.4 απεικονίζεται το δένδρογραμμα, της μεθόδου ward ενός δείγματος 72 πελατών του ηλεκτρονικού μας καταστήματος. Η μέθοδος αυτή αποδίδει όσο καμία από τις προαναφερθέντες χωρίς αυτό να σημαίνει ότι απορρίπτεται η complete linkage. Στο παρακάτω δένδρογραμμα παρατηρούνται συμπαγείς και ευκολοδιάκριτες ομάδες χρηστών. Αυτό θα βοηθήσει αργότερα στην εξόρυξη δεδομένων αυτών των ομάδων που ενεργείται για την

τροφοδότηση του recommender system. Αυτή είναι και η μέθοδος που θα χρησιμοποιηθεί για στην παρούσα εργασία.

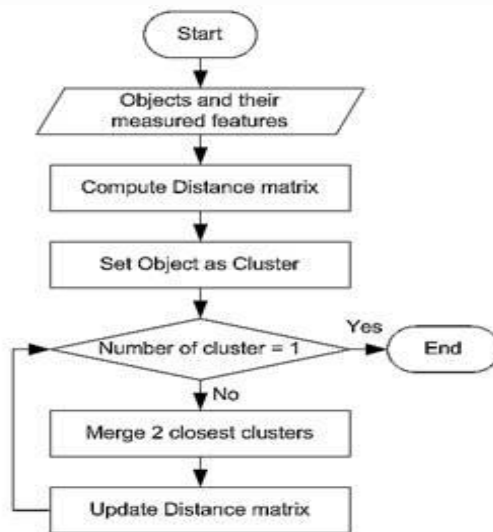


Εικόνα 5.4 Χρήση της μεθόδου ward στο e-shop με δείγμα 72 πελατών.

## 5.2 Χρήση αποτελεσμάτων αλγορίθμου

Ο αλγόριθμος ομαδοποίησης δουλεύει επαναληπτικά τοποθετώντας αρχικά τον κάθε χρήστη σε μία ομάδα cluster με αποτέλεσμα να έχουμε αριθμό ομάδων ίσο με τον αριθμό των χρηστών. Σε κάθε του επανάληψη βρίσκει κοινά σημεία μεταξύ ομάδων και τις συγχωνεύει. Αυτό έχει ως αποτέλεσμα σε κάθε επανάληψη να μειώνεται ο αριθμός των ομάδων έως ότου μείνει μια, από την άλλη να αυξάνεται ο αριθμός των χρηστών που ανήκουν σε κάθε ομάδα έως ότου όλοι ενταχθούν σε μία (όπως φαίνεται στην εικόνα 5.5). Στην παρούσα εργασία θα ορίσουμε κάθε επανάληψη ως το επίπεδο ομαδοποίησης των χρηστών.

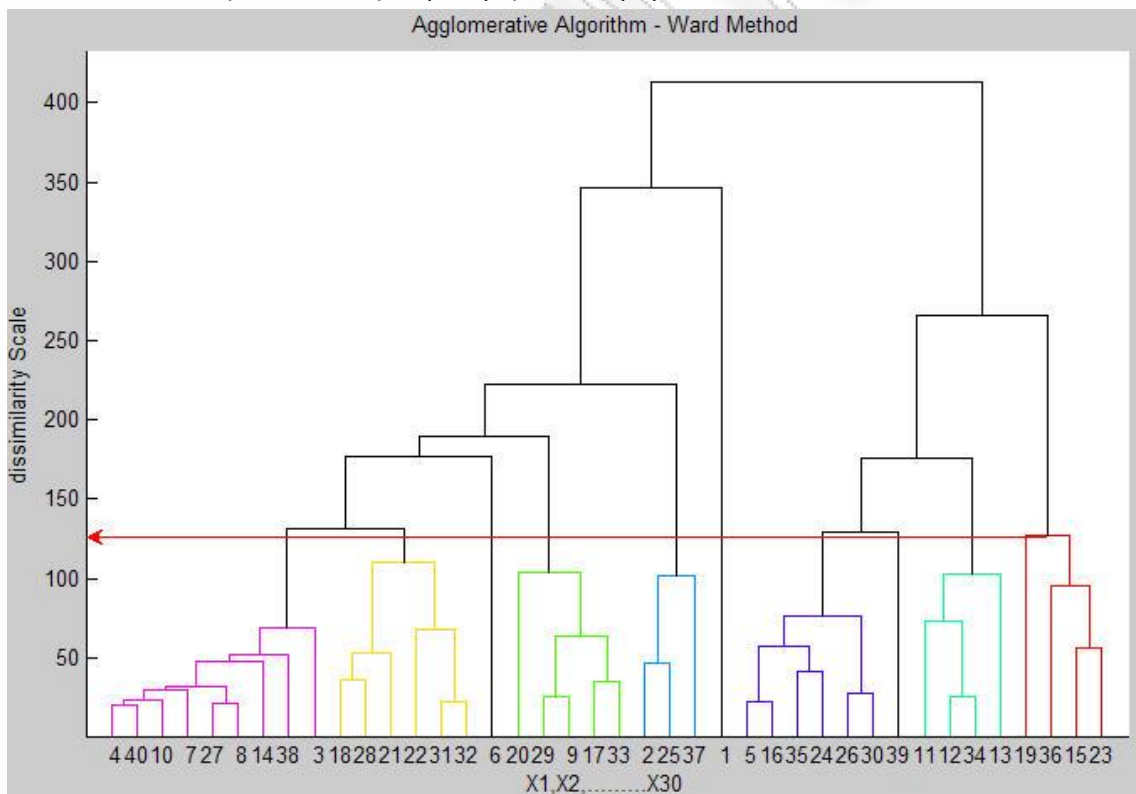




**Εικόνα 5.5** Λογικό διάγραμμα αλγορίθμου ιεραρχικής ομαδοποίησης

Έχοντας έναν αριθμό επιπέδων  $X$  με τον  $Y$  αριθμό ομάδων για το κάθε επίπεδο αντίστοιχα, δημιουργείται το ερώτημα:

Σε ποιο επίπεδο πρέπει να σταματήσει η ομαδοποίηση;



**Εικόνα 5.1** Δενδρόγραμμα με τη χρήση της μεθόδου Ward δείγματος 72 πελατών.

Για την απάντηση της παραπάνω ερώτησης έγινε χρήση του matlab το οποίο έχει συμβουλευτικό χαρακτήρα στην παρούσα εργασία.

Στο δενδρόγραμμα της εικόνας 5.1 παρατηρείται ότι στο επίπεδο με τιμή ανομοιομορφίας περίπου 138 έχουν δημιουργηθεί 10 διακριτές ομάδες χρηστών με κοντινά χαρακτηριστικά. Αν η τιμή της ανομοιομορφίας μικρύνει τότε δημιουργούνται πολλές ομάδες χρηστών και η

εκμετάλλευση του αλγόριθμου ομαδοποίησης λιγοστεύει. Αν πάλι η τιμή της ανομοιομορφίας αυξηθεί τότε οι ομάδες χρηστών λιγοστεύουν με κίνδυνο οι χρήστες του cluster να μην έχουν κοινά χαρακτηριστικά. Κατόπιν συνεχών δοκιμών που έγιναν με διαφορετικό αριθμό εγγεγραμμένων χρηστών στο ηλεκτρονικό κατάστημα κρίθηκε καταλληλότερη τιμή ανομοιομορφίας =138, όπου και σταματάει η ομαδοποίηση.

Ο αλγόριθμος clustering που χρησιμοποιήθηκε είναι ο αλγορίθμων ιεραρχικής ομαδοποίησης μεθόδου ward. Κατά την εκτέλεση του οποίου γεμίζει ένας προσωρινός πίνακας clusters στον οποί καταγράφεται ο αριθμός του cluster, το επίπεδο επανάληψης(cluster\_level), η τιμή ανομοιομορφίας(cluster\_value) και χρήστης (customer\_id). Έχοντας όλα αυτά τα στοιχεία γνωρίζουμε ανά επίπεδο σε ποια ομάδα ανήκει ο κάθε χρήστης. Ακολουθεί η απεικόνιση ενός μικρού μέρους του πίνακα clusters για την κατανόηση των προαναφερθέντων.

cluster_number	customer_id	cluster_level	cluster_value
0	2	62	138
1	3	62	138
1	35	62	138
1	65	62	138
1	71	62	138
1	69	62	138
1	40	62	138
1	43	62	138
1	79	62	138
2	4	62	138
2	5	62	138

**Εικόνα 5.2 Πίνακας clusters έχοντας επιλέξει για cluster\_value την τιμή 138.**

Στην εικόνα 5.2 μπορούμε να παρατηρήσουμε ότι ο customer\_id =2 ανήκει στο πρώτο cluster αριθμό 0 ενώ οι χρήστες με customer\_id =3, 35, 65,71,69, 40,43 ,79 βρίσκονται στο cluster με αριθμό 1. Επίσης οι χρήστες με customer\_id = 4,5 βρίσκονται στο cluster =2, έτσι από αυτό το μικρό παράδειγμα μπορεί να παρατηρηθεί ο τρόπος με τον οποίο ελέγχεται η ομάδα στην οποία ανήκει ο κάθε χρήστες σε κάθε επίπεδο αντίστοιχα.

Ο Πίνακας clusters έχει αρκετά μεγάλο όγκο δεδομένων διότι περιέχει όλα τα επίπεδα του clustering με όλες τις ομάδες και χρήστες ανά επίπεδο. Για το λόγο αυτό ο πίνακας αγορών βάσει clustering(clustered\_orders), ο πίνακας επισκεψιμότητας(clustered\_clicks) και ο πίνακας βαθμολογίας προϊόντων(clustered\_votes) δημιουργούνται κατά την εκτέλεση του clustering η οποία γίνεται ανά τακτά χρονικά διαστήματα στον server και όχι κατά την είσοδο του χρήστη στο ηλεκτρονικό κατάστημα. Με αυτό τον τρόπο δεν ταλαιπωρείται ο χρήστης με καθυστερήσεις φόρτωσης των σελίδων.

Για την εύρεση των ομάδων που ανήκει ο κάθε χρήστης εκτελείται το παρακάτω sql ερώτημα και καταχωρείται σε έναν προσωρινό πίνακα ο οποίος θα χρησιμοποιηθεί στις επόμενες παραγράφους.

```
SELECT cluster_number,customer_id
```

```
FROM clusters
```

```
WHERE cluster_value = 138
```

```
ORDER BY cluster_number,customer_id
```

Θέλοντας να δούμε κάποια από τα χαρακτηριστικά κάθε ομάδας που έχει δημιουργηθεί εκτελούνται μια σειρά ερωτημάτων στη βάση δεδομένων. Αυτά τα χαρακτηριστικά είναι οι αγορές, οι επισκέψεις και οι βαθμολογίες των χρηστών του ηλεκτρονικού καταστήματος.

### 5.2.1 Δημιουργία βοηθητικού πίνακα αγορών από clustering

Σε αυτή την παραγράφω θα αναλύσουμε τον τρόπο με τον οποίο δημιουργούμε έναν βοηθητικό πίνακα αγορών των ομαδοποιημένων χρηστών. Με αυτό τον τρόπο θα διακρίνουμε τις αγοραστικές συνήθειες-προτιμήσεις κάθε ομάδας χρηστών ξεχωριστά ώστε να τροφοδοτήσουμε την κάθε ομάδα με τις ανάλογες προτάσεις-διαφημίσεις.

Με την εκτέλεση του παρακάτω κώδικα γεμίζει ο πίνακας clustered\_orders ο οποίος περιέχει όλες τις αγοραστικές συνήθειες των χρηστών ανά ομάδα και ανά κατηγορία προϊόντων. Όταν εισέρχεται ένας χρήστης στο e-shop τότε γίνεται ανάγνωση αυτού του πίνακα για να δούμε τις αγοραστικές συνήθειες της ομάδας του και να προβούμε στις κατάλληλες προτάσεις

```
$query2 = "SELECT cluster_number,customer_id FROM clusters
          WHERE cluster_level = $cluster_level
          ORDER BY cluster_number,customer_id";

$ret2 = mysql_query($query2);
while ($row2 = mysql_fetch_array($ret2)){
    $cluster_customer = $row2['customer_id'];
    $cluster_number = $row2['cluster_number'];

    $query3 = "SELECT customer_id FROM clusters
              WHERE cluster_number = $cluster_number
              AND cluster_level = $cluster_level
              ORDER BY customer_id";
    $ret3 = mysql_query($query3);
    while($row3 = mysql_fetch_array($ret3)){
        $customer_id = $row3['customer_id'];
        $query4 = "SELECT product_id,count FROM orders
                  WHERE cust_id = $customer_id";
        while($row4 = mysql_fetch_array($ret4)){
            $product_id = $row4['product_id'];
            $count = $row4['count'];

            $query5 = "SELECT menu_id,cat_id FROM products
                      WHERE product_id = $product_id";
            $ret5 = mysql_query($query5);
            $row5 = mysql_fetch_array($ret5);
            $menu_id = $row5['menu_id'];
            $cat_id = $row5['cat_id'];

            $query6="INSERT INTO clustered_orders
                     (menu_id,cat_id,count,customer_id)
                     VALUES ('$menu_id','$cat_id','$count','$cluster_customer');
            }
        }
    }
}
```

### 5.2.2 Δημιουργία βοηθητικού πίνακα επισκέψεων από clustering

Σε αυτή την παραγράφω θα αναλύσουμε τον τρόπο με τον οποίο δημιουργούμε έναν βοηθητικό πίνακα επισκέψεων των ομαδοποιημένων χρηστών. Με αυτό τον τρόπο θα διακρίνουμε τις συνήθειες-προτιμήσεις σε προϊόντα βάσει των επισκέψεων κάθε ομάδας χρηστών ξεχωριστά ώστε να τροφοδοτήσουμε την κάθε ομάδα με τις ανάλογες προτάσεις-διαφημίσεις.

Με την εκτέλεση του παρακάτω κώδικα γεμίζει ο πίνακας clustered\_clicks ο οποίος περιέχει όλες τις επισκεπτικές συνήθειες των χρηστών ανά ομάδα και ανά κατηγορία προϊόντων. Όταν εισέρχεται ένας χρήστης στο e-shop τότε γίνεται ανάγνωση αυτού του πίνακα για να δούμε τις συνήθειες της ομάδας του και να προβούμε στις κατάλληλες προτάσεις.

```
$query2 = "SELECT cluster_number,customer_id FROM clusters
          WHERE cluster_level = $cluster_level
          ORDER BY cluster_number,customer_id";

$ret2 = mysql_query($query2);
while ($row2 = mysql_fetch_array($ret2)){
    $cluster_customer = $row2['customer_id'];
    $cluster_number = $row2['cluster_number'];
    $query3 = "SELECT customer_id FROM clusters
              WHERE cluster_number= $cluster_number AND
              cluster_level=$cluster_level
              ORDER BY customer_id";
    $ret3 = mysql_query($query3);
    while($row3 = mysql_fetch_array($ret3)){
        $customer_id = $row3['customer_id'];
        $query4 = "SELECT * FROM clicks
                  WHERE customer_id = $customer_id";
        $ret4 = mysql_query($query4);
        while($row4 = mysql_fetch_array($ret4)){
            $count = $row4['clicks_number'];
            $cat_id = $row4['cat_id'];
            $query6 = "INSERT INTO clustered_clicks (customer_id,cat_id,count)
                      VALUES ('$cluster_customer','$cat_id','$count)";
        }
    }
}
```

### 5.2.3 Δημιουργία προσωρινού πίνακα βαθμολογίας προϊόντων από clustering

Σε αυτή την παραγράφω θα αναλύσουμε τον τρόπο με τον οποίο δημιουργούμε έναν βοηθητικό πίνακα βαθμολογιών των ομαδοποιημένων χρηστών. Με αυτό τον τρόπο θα διακρίνουμε τις βαθμολογικές συνήθειες-προτιμήσεις κάθε ομάδας χρηστών ξεχωριστά ώστε να τροφοδοτήσουμε την κάθε ομάδα με τις ανάλογες προτάσεις-διαφημίσεις.

Με την εκτέλεση του παρακάτω κώδικα γεμίζει ο πίνακας clustered\_votes ο οποίος περιέχει όλες τις βαθμολογίες των χρηστών ανά ομάδα και ανά προϊόν. Όταν εισέρχεται ένας χρήστης στο e-shop τότε γίνεται ανάγνωση αυτού του πίνακα για να δούμε τις βαθμολογικές συνήθειες της ομάδας του και να προβούμε στις κατάλληλες προτάσεις

```
$query2 = "SELECT cluster_number,customer_id FROM clusters
          WHERE cluster_level = $cluster_level
```

```

ORDER BY cluster_number,customer_id";
$ret2 = mysql_query($query2);
while ($row2 = mysql_fetch_array($ret2)){
    $cluster_customer = $row2['customer_id'];
    $cluster_number = $row2['cluster_number'];
    $query3 = "SELECT customer_id FROM clusters
              WHERE cluster_number = $cluster_number AND
              cluster_level = $cluster_level
              ORDER BY customer_id LIMIT 0,10";
    $ret3 = mysql_query($query3);
    while($row3 = mysql_fetch_array($ret3)) {
        $customer_id = $row3['customer_id'];
        $query4 = "SELECT * FROM poll
                  WHERE customer_id = $customer_id";
        $ret4 = mysql_query($query4);
        while($row4 = mysql_fetch_array($ret4)){
            $vote = $row4['vote_rating'];
            $product_id = $row4['product_id'];
            $query6="INSERT INTO clustered_prod_votes
            (customer_id,product_id,vote)
            VALUES ('$cluster_customer','$product_id','$vote)";
            $ret6 = mysql_query($query6);
        }
    }
}
}
}

```

### 5.3 Εξόρυξη δεδομένων ανά ομάδα χρηστών

Σε αυτή την παράγραφο θα περιγράψουμε τον τρόπο, με τον οποίο βρίσκουμε τις αγοραστικές συνήθειες κάθε ομάδας, κατά την είσοδο κάθε εγγεγραμμένου χρήστη στο ηλεκτρονικό κατάστημα.

#### 5.3.1 Εύρεση κατηγορίας με τις περισσότερες αγορές.

Κατά την είσοδο(log in) ενός χρήστη στην ιστοσελίδα εκτελείται ένα ερωτήμα προς τη βάση και συγκεκριμένα προς τον βοηθητικό πίνακα clustered\_orders που περιγράψαμε στην παράγραφο 5.2. Το ερώτημα που εκτελείται παρατίθεται ακολούθως:

```

SELECT cat_id , sum( count ) AS TOTAL
FROM `clustered_orders`
WHERE customer_id = $user
GROUP BY cat_id
ORDER BY `TOTAL` DESC

```

Στον παραπάνω κώδικα ζητείται από τη βάση η εύρεση της κατηγορία με τις περισσότερες αγορές που έχουν κάνει οι χρήστες της ομάδας του logged in πελάτη. Η κατηγορία αυτή ονομάζεται best\_order\_category. Η κατηγορία αυτή θα παίξει σημαντικό ρόλο στην δημιουργία προτάσεων-διαφημίσεων που θα περιγραφούν στην ενότητα 6.

### 5.3.2 Εύρεση κατηγορίας με τις μεγαλύτερη επισκεψιμότητα

Κατά την είσοδο(log in) ενός χρήστη στην ιστοσελίδα εκτελείται ένα ερώτημα προς τη βάση και συγκεκριμένα προς τον βοηθητικό πίνακα clustered\_clicks που περιγράψαμε στην παράγραφο 5.2. Το ερώτημα που εκτελείται παρατίθεται ακολούθως:

```
SELECT cat_id, sum( count ) AS TOTAL
FROM `clustered_clicks`
WHERE customer_id = $user
GROUP BY cat_id
ORDER BY `TOTAL` DESC"
```

Στον παραπάνω κώδικα ζητείται από τη βάση η εύρεση της κατηγορία με τις περισσότερες επισκέψεις που έχουν κάνει οι χρήστες της ομάδας του logged in πελάτη. Η κατηγορία αυτή θα παίξει σημαντικό ρόλο στην δημιουργία προτάσεων-διαφημίσεων που θα περιγραφούν στην ενότητα 6.

### 5.3.3 Εύρεση προϊόντων με τη μεγαλύτερη βαθμολογία

Κατά την είσοδο(log in) ενός χρήστη στην ιστοσελίδα εκτελείται ένα ερώτημα προς τη βάση και συγκεκριμένα προς τον βοηθητικό πίνακα clustered\_prod\_votes που περιγράψαμε στην παράγραφο 5.2. Το ερώτημα που εκτελείται παρατίθεται ακολούθως:

```
SELECT * , avg( vote ) , count( vote )
FROM `clustered_prod_votes`
WHERE customer_id =$user
GROUP BY product_id
ORDER BY avg( vote ) DESC , count( vote ) DESC
```

Στον παραπάνω κώδικα ζητείται από τη βάση η εύρεση των δημοφιλέστερων προϊόντων βάσει βαθμολογίας που έχουν κάνει οι χρήστες της ομάδας του logged in πελάτη. Στον παραπάνω κώδικα ζητείται από τη βάση η εύρεση της κατηγορία με τις περισσότερες αγορές που έχουν κάνει οι χρήστες της ομάδας του logged in πελάτη. Η κατηγορία αυτή θα παίξει σημαντικό ρόλο στην δημιουργία προτάσεων-διαφημίσεων που θα περιγραφούν στην ενότητα 6.

## 6. ΕΦΑΡΜΟΓΗ ΠΡΟΤΑΣΕΩΝ

Σε αυτή τη παράγραφο θα περιγράψουμε τις προτάσεις που παρέχονται στους εγγεγραμμένους χρήστες του ηλεκτρονικού καταστήματος της παρούσας διπλωματικής. Οι προτάσεις αυτές χωρίζονται σε δύο μεγάλες κατηγορίες για της οποίες θα γίνει ανάλυση στις επόμενες δύο υποπαραγράφους. Οι κατηγορίες αυτές είναι :

1. Προτάσεις που απορρέουν από τις αγοραστικές συνήθειες, επισκεψιμότητα και βαθμολογικές προτιμήσεις ομάδων χρηστών που έχουν προκύψει από την εφαρμογή του του ιεραρχικού αλγόριθμου ομαδοποίησης Ward.
2. Προτάσεις που απορρέουν από τις αγοραστικές συνήθειες, επισκεψιμότητα και βαθμολογικές προτιμήσεις όλων των εγγεγραμμένων χρηστών του ηλεκτρονικού καταστήματος.

Η πρώτη κατηγορία προτάσεων δεν εφαρμόζεται στους απλούς επισκέπτες του e-shop διότι δεν είναι δυνατή η συγκέντρωση ιστορικών στοιχείων για τις προτιμήσεις τους. Για το λόγο αυτό δεν μπορούν να ενταχθούν σε κάποια από τις ομάδες που δημιουργούνται από τον αλγόριθμο clustering. Τουλάχιστον, εμφανίζονται προτάσεις της δεύτερης κατηγορίας για την πρόκληση του ενδιαφέροντος των επισκεπτών. Επίσης γνωστοποιείται στους επισκέπτες ότι με την εγγραφή τους θα απολαμβάνουν τα προνόμια ενός προσαρμοστικού ηλεκτρονικού καταστήματος καθώς θα λαμβάνουν μέρος σε βαθμολογίες προϊόντων.

## 6.1 Προτάσεις ομαδοποιημένων χρηστών (από Clustering)

Σε αυτή την παράγραφο θα αναφερθούμε στις προτάσεις που λαμβάνει ο εγγεγραμμένος χρήστης κατά την διάρκεια της περιήγησης του στο ηλεκτρονικό κατάστημα. Αυτές οι προτάσεις διαφημιστικού χαρακτήρα απορρέουν από το συνδυασμό χρήσης πολύπλοκων τεχνικών όπως είναι η χρήση αλγόριθμου ομαδοποίησης και εξόρυξης δεδομένων. Χρίζουν ιδιαίτερης σημασίας καθώς έχουν μεγάλη βαρύτητα διότι βάσει αυτών των τεχνικών οι προτάσεις που προβάλλονται στον καταναλωτή έχουν μεγαλύτερη πιθανότητα να ταιριάζουν στα γούστα τους. Αυτό οφείλεται στο ότι ανήκουν στην ομάδα με κοινά χαρακτηριστικά βάσει συνηθειών και προτιμήσεων. Οι προτάσεις αυτές παρατίθενται παρακάτω και είναι οι εξής:

1. Cross Sales(ομαδοποιημένες αγορές της ομάδας του χρήστη).
2. Κατηγορία με τις περισσότερες αγορές της ομάδας του χρήστη.
3. Κατηγορία με την μεγαλύτερη επισκεψιμότητα της ομάδας του χρήστη.
4. Προϊόν με τη μεγαλύτερη βαθμολογία της ομάδας του χρήστη.
5. Νεότερο προϊόν της κατηγορίας με τις περισσότερες αγορές της ομάδας του χρήστη.

Ακολουθως θα γίνει εκτενέστερη περιγραφή αυτών των προτάσεων.

### 6.1.1 Cross Sales(ομαδοποιημένες αγορές της ομάδας του χρήστη).

Το δημοφιλές cross\_sales ή αλλιώς “Όσοι αγόρασαν αυτό το προϊόν αγόρασαν και αυτό” δεν θα έπρεπε να λείπει από την εφαρμογή αυτή. Η βασική ιδέα είναι να συγκεντρωθούν οι συνδυαζόμενες αγορές που αφορούν το προϊόν της σελίδας που επισκέπτεται ο καταναλωτής και τα προϊόντα που έχουν λάβει μέρος στις περισσότερες από αυτές να εμφανίζονται. Στη συγκεκριμένη περίπτωση αναζητούνται μόνο οι αγορές που έχουν γίνει από τους χρήστες που ανήκουν στην ομάδα του logged in καταναλωτή (από τον πίνακα clustered\_orders). Πιο συγκεκριμένα θα πρέπει να πούμε ότι στις συνδυαζόμενες αγορές δεν λαμβάνουν χώρα οι αγορές του logged in χρήστη, διότι είναι ανώφελο να του προταθεί ένα προϊόν που έχει ήδη αποκτήσει. Παρακάτω παρατίθεται ο κώδικας με τον οποίο δημιουργείται αυτού του είδους η πρόταση.

Με την εκτέλεση του παρακάτω κομματιού κώδικα γεμίζει ένας βοηθητικός πίνακας cross\_sales με τις συνδυαζόμενες αγορές που προαναφέραμε.

```
$query = "SELECT * FROM clustered_orders
WHERE product_id= $product_id AND cust_id != $user_id";
$ret = mysql_query($query);
while($row = mysql_fetch_array($ret)){
    $cust_id = $row['cust_id'];
    $date = $row['Date'];
    $query2 = "SELECT * FROM orders
WHERE cust_id = $cust_id and product_id != $product_id and
Date = '$date'";
    $ret2 = mysql_query($query2);
    while($row2 = mysql_fetch_array($ret2)){
        $product_cross_id = $row2['product_id'];
        $count_cross = $row2['count'];
        $query_ins = "INSERT INTO cross_sales (product_id, count)
VALUES ($product_cross_id,$count_cross)";
        $ret_ins = mysql_query($query_ins);
    }
}
```

Με το ακόλουθο sql ερώτημα αναζητούνται τα προϊόντα με τις περισσότερες εμφανίσεις στις συνδυαζόμενες αγορές με το προϊόν της σελίδας που επισκέπτεται ο logged in χρήστης.

```
SELECT *,SUM(count) As A
FROM cross_sales
GROUP BY product_id
ORDER BY A DESC";
```

Το id του προϊόντος αυτού αποθηκεύεται στη μεταβλητή \$cross\_product.

Για την εμφάνιση όλων των στοιχείων του προϊόντος αυτού (Όνομα, μάρκα, περιγραφή, τιμή) εκτελείται το παρακάτω sql ερώτημα

```
SELECT * FROM products
WHERE product_id = $cross_product
```

### 6.1.2 Κατηγορία με τις περισσότερες αγορές της ομάδας του χρήστη.

Έχοντας υπολογίσει την κατηγορία με τις περισσότερες αγορές της ομάδας του logged\_in χρήστη από την παράγραφο 5.3.1, δημιουργούμε ένα sql ερώτημα κατά το οποίο λαμβάνει ο χρήστης προϊόντα από την κατηγορία αυτή. Αυτό πιστεύεται ότι είναι από τα προϊόντα με τη μεγαλύτερη βαρύτητα διότι εκφράζει το πραγματικό ενδιαφέρον των καταναλωτών αφού το έχουν ήδη αποκτήσει.

```
$cat_id = $_SESSION['best_cat_id'];
$query="SELECT * FROM products
        WHERE cat_id=$cat_id";
$ret = mysql_query($query);
$num_of_rows = mysql_num_rows($ret);
$rnd = rand(1,$num_of_rows);
for ($i=1; $i<=$num_of_rows;$i++){
    $row = mysql_fetch_array($ret);
    if ($i==$rnd){
        $prod1 = $row['product_id'];
        break;
    }
}
```

### 6.1.3 Κατηγορία με την μεγαλύτερη επισκεψιμότητα της ομάδας του χρήστη.

Έχοντας υπολογίσει την κατηγορία με τις περισσότερες επισκέψεις της ομάδας του logged\_in χρήστη από την παράγραφο 5.3.2, δημιουργούμε ένα sql ερώτημα κατά το οποίο λαμβάνει ο χρήστης προϊόντα από την κατηγορία αυτή. Αυτό το προϊόν είναι αρκετά σημαντικό καθώς δηλώνει την περιέργεια των καταναλωτών .

```
$cat_id = $_SESSION['best_click_cat_id'];
$query="SELECT * FROM products
        WHERE cat_id=$cat_id";
$ret = mysql_query($query);
$num_of_rows = mysql_num_rows($ret);
$rnd = rand(1,$num_of_rows);
for ($i=1; $i<=$num_of_rows;$i++){
    $row = mysql_fetch_array($ret);
    if ($i==$rnd){
```



```

        $prod2 = $row['product_id'];
        break;
    }
}

```

#### 6.1.4 Προϊόν με τη μεγαλύτερη βαθμολογία της ομάδας του χρήστη.

Έχοντας υπολογίσει το προϊόν με την μεγαλύτερη κατά μέσω όρο βαθμολογία της ομάδας του logged\_in χρήστη από την παράγραφο 5.3.3, δημιουργούμε ένα sql ερώτημα κατά το οποίο λαμβάνει ο χρήστης προϊόντα από την κατηγορία αυτή. Αυτό το προϊόν χρίζει ιδιαίτερης σημασίας καθώς προέρχεται από τις βαθμολογικές προτιμήσεις των χρηστών του κάθε cluster.

```

$product = $_SESSION['clustered_best_illected'];
$query = "SELECT * FROM products
        WHERE product_id = $product";
$ret = mysql_query($query);
$row = mysql_fetch_array($ret);
$num = mysql_num_rows($ret);

```

#### 6.1.5 Νεότερο προϊόν της κατηγορίας με της περισσότερες αγορές της ομάδας του χρήστη.

Έχοντας υπολογίσει την κατηγορία με της περισσότερες αγορές της ομάδας του logged\_in χρήστη από την παράγραφο 5.3.1, δημιουργούμε ένα sql ερώτημα κατά το οποίο λαμβάνει ο χρήστης το νεότερο προϊόν από την κατηγορία αυτή. Το προϊόν αυτό πιθανότατα να προκαλεί την μεγαλύτερη περιέργεια των εγγεγραμμένων καταναλωτών, καθώς είναι η νεότερη κυκλοφορία που ανήκει στην αγαπημένη τους κατηγορία.

```

$cat_id = $_SESSION['best_cat_id'];
$query = "SELECT * FROM products
        WHERE cat_id = $cat_id
        ORDER BY DateAdded DESC";
$ret = mysql_query($query);
$row = mysql_fetch_array($ret);
$num = mysql_num_rows($ret);

```

## 6.2 Προτάσεις από το σύνολο των χρηστών (ανεξαρτήτου clustering)

1. Cross Sales (ομαδοποιημένες αγορές όλων των χρηστών).
2. Κατηγορία με τις περισσότερες αγορές όλων των χρηστών.
3. Κατηγορία με την μεγαλύτερη επισκεψιμότητα όλων των χρηστών.
4. Προϊόν με τη μεγαλύτερη βαθμολογία με συμμετοχή όλων των χρηστών.
5. Νεότερο προϊόν της κατηγορίας με της περισσότερες αγορές όλων των χρηστών.

Ακολούθως θα γίνει εκτενέστερη περιγραφή αυτών των προτάσεων.

### 6.2.1 Cross Sales(ομαδοποιημένες αγορές όλων των χρηστών).

Το cross\_sales που περιγράφηκε σε προηγούμενη παράγραφο λειτουργεί με λίγο διαφορετικό τρόπο σε αυτή τη περίπτωση. Εδώ αναζητούνται οι αγορές που έχουν γίνει από όλους χρήστες. Πιο συγκεκριμένα θα πρέπει να πούμε ότι στις συνδυαζόμενες αγορές δεν λαμβάνουν χώρα οι αγορές του logged in χρήστη, διότι είναι ανώφελο να του προταθεί ένα προϊόν που έχει ήδη

αποκτήσει. Παρακάτω παρατίθεται ο κώδικα με τον οποίο δημιουργείται αυτού του είδους η πρόταση.

Με την εκτέλεση του παρακάτω κομματιού κώδικα γεμίζει ένας βοηθητικός πίνακας `cross_sales` με τις συνδυαζόμενες αγορές που προαναφέραμε.

```
$query = "SELECT * FROM orders
WHERE product_id= $product_id AND cust_id != $user_id";
$ret = mysql_query($query);
while($row = mysql_fetch_array($ret)){

    $cust_id = $row['cust_id'];
    $date = $row['Date'];
    $query2 = "SELECT * FROM orders
WHERE cust_id = $cust_id and product_id != $product_id and
Date = '$date'";
    $ret2 = mysql_query($query2);
    while($row2 = mysql_fetch_array($ret2)){
        $product_cross_id = $row2['product_id'];
        $count_cross = $row2['count'];
        $query_ins = "INSERT INTO cross_sales (product_id, count)
VALUES ($product_cross_id,$count_cross)";
        $ret_ins = mysql_query($query_ins);
    }
}
```

Με το ακόλουθο sql ερώτημα αναζητούνται τα προϊόντα με τις περισσότερες εμφανίσεις στις συνδυαζόμενες αγορές με το προϊόν της σελίδας που επισκέπτεται ο logged in χρήστης.

```
SELECT *,SUM(count) As A
FROM cross_sales
GROUP BY product_id
ORDER BY A DESC";
```

Το id του προϊόντος αυτού αποθηκεύεται στη μεταβλητή `$cross_product`.

Για την εμφάνιση όλων των στοιχείων του προϊόντος αυτού( Όνομα, μάρκα, περιγραφή, τιμή) εκτελείται το παρακάτω sql ερώτημα

```
SELECT * FROM products
WHERE product_id = $cross_product
```

### **6.2.2 Κατηγορία με τις περισσότερες αγορές όλων των χρηστών.**

Εξίσου σημαντικά προϊόντα προς πρόταση είναι και τα προϊόντα εκτός ομαδοποίησης διότι εκφράζουν τις προτιμήσεις του ευρύτερου αγοραστικού κοινού του ηλεκτρονικού καταστήματος.

Μπορεί να μην είναι εξατομικευμένη μέθοδος, από την άλλη δεν παύει να είναι ένας έγκυρος τρόπος πρότασης προϊόντων.

Σε αυτή την παράγραφο περιγράφεται ο τρόπος με τον οποίο υπολογίζεται η εμφάνιση της κατηγορίας προϊόντων με τις περισσότερες αγορές όλων των χρηστών του ηλεκτρονικού καταστήματος. Παρακάτω ακολουθούν τα sql ερωτήματα που εκτελούνται για την εμφάνιση του προϊόντος αυτού.

```
$query2="SELECT *,sum(count) FROM orders
GROUP BY product_id
ORDER BY sum(count) DESC ";
```

```

$ret2 = mysql_query($query2);
$row2 = mysql_fetch_array($ret2);
$product_id2 = $row2['product_id'];
$query = "SELECT * FROM products
        WHERE product_id = $product_id2";
$ret = mysql_query($query);
$row = mysql_fetch_array($ret);

```

### 6.2.3 Κατηγορία με την μεγαλύτερη επισκεψιμότητα όλων των χρηστών.

Σε αυτή την παράγραφο περιγράφεται ο τρόπος με τον οποίο υπολογίζεται η εμφάνιση της κατηγορίας προϊόντων με τις περισσότερες επισκέψεις όλων των χρηστών του ηλεκτρονικού καταστήματος.

Παρακάτω ακολουθούν τα sql ερωτήματα που εκτελούνται για την εμφάνιση του προϊόντος αυτού.

```

$query2="SELECT * , sum( clicks_number )
        FROM clicks
        GROUP BY product_id
        ORDER BY sum( clicks_number ) DESC ";
$ret2 = mysql_query($query2);
$row2 = mysql_fetch_array($ret2);
$product_id = $row2['product_id'];
$query = "SELECT * FROM products
        WHERE product_id = $product_id";
$ret = mysql_query($query);
$row = mysql_fetch_array($ret);

```

### 6.2.4 Προϊόν με τη μεγαλύτερη βαθμολογία με συμμετοχή όλων των χρηστών.

Σε αυτή την παράγραφο περιγράφεται ο τρόπος με τον οποίο υπολογίζεται η εμφάνιση του προϊόντος με τη μεγαλύτερη κατά μέσο όρο βαθμολογία όλων των χρηστών του ηλεκτρονικού καταστήματος.

Παρακάτω ακολουθούν τα sql ερωτήματα που εκτελούνται για την εμφάνιση του προϊόντος αυτού.

```

SELECT avg( vote_rating ) , count( vote_rating ) , product_id
FROM poll
GROUP BY product_id
ORDER BY avg( vote_rating ) DESC , count( vote_rating ) DESC
$query_poll = "SELECT * FROM products
        WHERE product_id = $best_elected_product";
$ret_poll=mysql_query($query_poll);
$row = mysql_fetch_array($ret_poll);

```

### 6.2.5 Νεότερο προϊόν της κατηγορίας με τις περισσότερες αγορές όλων των χρηστών.

Οι νέες κυκλοφορίες ενός ηλεκτρονικού καταστήματος δεν θα έπρεπε να λείπουν από τις προτάσεις ενός recommender συστήματος. Σε αυτή την παράγραφο περιγράφεται ο τρόπος με τον οποίο υπολογίζεται η εμφάνιση της κατηγορίας προϊόντων με τις περισσότερες αγορές όλων των χρηστών του ηλεκτρονικού καταστήματος.

Παρακάτω ακολουθεί ένα sql ερώτημα που εκτελείται για την εμφάνιση του προϊόντος αυτού.

```
SELECT * FROM `products`  
ORDER BY `DateAdded` DESC
```

## 7. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Στο τελευταίο κεφάλαιο της παρούσας εργασίας θα επιχειρηθεί μία σύνοψη των κυριότερων σημείων που αναλύθηκαν στα προηγούμενα κεφάλαια καθώς και η συνεισφορά της σε αυτόν τον τομέα έρευνας. Επίσης θα παρατεθούν ορισμένες ιδέες για μελλοντική επέκταση της εργασίας, όπως και θα προταθούν μερικές βελτιώσεις οι οποίες θα συντελέσουν στην αποτελεσματικότερη λειτουργία του συστήματος.

### 7.1 Συνεισφορά

Στην παρούσα εργασία επιχειρήθηκε η δημιουργία ενός προσαρμοστικού ηλεκτρονικού καταστήματος, το οποίο προσαρμόζεται στον εκάστοτε χρήστη, όχι μέσω ενός μοντέλου για κάθε χρήστη αλλά εκμεταλλευόμενοι την ύπαρξη ενός γενικότερου μοντέλου χρηστών. Το μοντέλο αυτό δημιουργήθηκε μέσω τεσσάρων πολύ διαδεδομένων μεθόδων ιεραρχικής ομαδοποίησης και έτσι μπορέσαμε να συγκρίνουμε την αποδοτικότητα των μεθόδων αυτών. Φυσικά το δείγμα πελατών αν και ικανοποιητικό σε αριθμό, επιβάλλεται να γίνει και ποιοτικότερο από άποψη κινήσεων των πελατών. Για το σκοπό αυτό ανεβάσαμε το site σε web-server, ώστε σε σύντομο χρονικό διάστημα να έχουμε ένα μεγαλύτερο dataset από το οποίο να μπορούν να προκύψουν ασφαλέστερα συμπεράσματα. Σε αντίθεση με τα περισσότερα recommender systems δεν ακολουθήσαμε την πεπατημένη, δηλαδή μια item-based προσέγγιση με έτοιμους αλγορίθμους, αλλά προσπαθήσαμε να προσαρμόσουμε τους αλγορίθμους ομαδοποίησης, όπως και την εξόρυξη γνώσης από τα δεδομένα στα μέτρα του ηλεκτρονικού μας καταστήματος, δηλαδή προσανατολισμένα σε προϊόντα πληροφορικής, τα οποία έχουν κάποιες ιδιαιτερότητες σε σχέση με άλλα προϊόντα. Δώσαμε επίσης έμφαση στην διαφάνεια της διαδικασίας της προσαρμοστικότητας του συστήματος. Δεν θέλουμε ένα σύστημα το οποίο να επεμβαίνει συνεχώς κατά τη διάρκεια της αλληλεπίδρασης του χρήστη με αυτό. Για το λόγο αυτό επιχειρήσαμε μια έμμεση προσαρμοστικότητα, που τα αντικείμενα που εμφανίζονται σε συγκεκριμένες θέσεις της σελίδας, μπορεί να προκαλέσουν το ενδιαφέρον του χρήστη. Σίγουρα όμως οι βελτιώσεις που μπορούν να γίνουν στο επίπεδο τόσο της διεπαφής, όσο και της μοντελοποίησης είναι αρκετές, ώστε να προσελκύει όσο το δυνατόν περισσότερους χρήστες. Λόγω της πολυπλοκότητας των αλγορίθμων, το σύστημα αυτό δεν είναι δυνατό να δουλέψει για υπέρτοπο αριθμό χρηστών, ωστόσο μπορεί να βοηθήσει σε καταστήματα που προβλέπεται να έχουν μέχρι 1000 χρήστες και σχετικά περιορισμένη κίνηση. Φυσικά με δεδομένο ότι θα γίνουν κάποιες βελτιστοποιήσεις τις οποίες ήδη επεξεργαζόμαστε, στο σύστημα μπορεί να αυξηθεί η αποδοτικότητα του σε σημαντικό βαθμό. Ο χρόνος και η χρήση του συστήματος θα δείξουν αν η λύση που προτείνουμε εδώ είναι βιώσιμη.

### 7.2 Συμπεράσματα

Το κυριότερο συμπέρασμα που μπορεί να εξαχθεί από την παρούσα εργασία έχει να κάνει με την σύγκριση των τεσσάρων διαφορετικών μεθόδων ιεραρχικής ομαδοποίησης που εφαρμόστηκαν (Κεφ 6). Όπως είδαμε η μέθοδος Ward προκρίνεται ως η καλύτερη λύση καθώς τείνει να οδηγεί στον σχηματισμό ισοσκελισμένων και και ισομερώς καταμερισμένων ομάδων.

Αμέσως μετά έπεται η μέθοδος complete-linkage, η οποία δημιουργεί συμπαγείς ομάδες αλλά όχι τόσο ισοσκελισμένες όσο η Ward, έχοντας ως μειονέκτημα την ευαισθησία σε ακραίες τιμές (outliers).

Τελευταίες στην αξιολόγηση μέθοδοι, είναι κατά σειρά η average-linkage, με χειρότερη όλων την μέθοδο single-linkage. Η πρώτη έχει την τάση να δημιουργεί ομάδες με μικρές αποκλίσεις λόγω του ότι στον υπολογισμό της απόστασης λαμβάνουν μέρος όλα τα σημεία μιας ομάδας και έχει την τάση να επηρεάζεται από τα ακραία σημεία. Αντίθετα η μέθοδος single-linkage πάσχει από το σύνδρομο chain-effect το οποίο ευθύνεται για τη δημιουργία μεγάλων μη σφαιρικών ομάδων.

Ένα άλλο συμπέρασμα το οποίο μπορεί να εξαχθεί είναι ότι οι αλγόριθμοι της ιεραρχικής ομαδοποίησης που εξετάσαμε είναι υπολογιστικά ασύμφοροι για να χρησιμοποιηθούν σε real-time συστήματα με πολύ μεγάλο όγκο χρηστών. Ωστόσο σε μικρότερα συστήματα και πιθανώς με τη χρήση καταμεμημένων συστημάτων, θα μπορούσε να εφαρμοστεί χωρίς ιδιαίτερα προβλήματα.

### 7.3 Μελλοντικές Επεκτάσεις - Βελτιώσεις

Με βάση όλων όσων εξετάστηκαν στην συγκεκριμένη διπλωματική εργασία και λαμβάνοντας υπόψη μας τις ολοένα αυξανόμενες απαιτήσεις ενός προσαρμοστικού ηλεκτρονικού καταστήματος προκύπτουν αρκετά μελλοντικά ζητήματα και επεκτάσεις που αναμένεται να εξεταστούν και να υλοποιηθούν.

#### 7.3.1 Μελλοντικές Επεκτάσεις

- Εισαγωγή δημογραφικών δεδομένων των πελατών στον αλγόριθμο clustering για περαιτέρω ομαδοποίηση των εγγεγραμμένων χρηστών.
- Εύρεση νέων μεθόδων για τη δημιουργία ελκυστικότερων προτάσεων από το σύστημα.
- Υλοποίηση νέων τεχνικών για καλύτερη προσαρμοστικότητα της εφαρμογής στις ανάγκες του καταναλωτή
- Υλοποίηση εργαλείου για εξαγωγή στατιστικών με βάση τα ποσοστά επιτυχίας των προτάσεων που παρέχονται στους καταναλωτές

#### 7.3.2 Βελτιώσεις

- Βελτίωση του εικαστικού περιβάλλοντος της διεπαφής για την προσέλκυση περισσότερων καταναλωτών.
- Βελτιστοποίηση των ερωτημάτων sql για την καλύτερη απόκριση του συστήματος.
- Υιοθέτηση της Ajax σε περισσότερα κομμάτια του ηλεκτρονικού καταστήματος όπως είναι η σελιδοποίηση και το καλάθι αγορών για καλύτερη απόδοση του συστήματος.
- Δημιουργία περισσότερων οντοτήτων κώδικα για την επαναχρησιμοποίηση τους για την μείωση του επαναλαμβανόμενου κώδικα.

**Βιβλιογραφία**

- [1] D.N. Sotiropoulos, G.A. Tsihrintzis, A. Savvopoulos, and M. Virvou, A Comparison of Customer Data Clustering Techniques in an e-Shopping Application , in: Proceedings of 2nd International Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces, Ireland, 2006
- [2] J. Ben Schafer, Joseph Konstan, John Riedl, Recommender Systems in E-Commerce, In EC '99 Proceedings of the 1st ACM conference on Electronic commerce
- [3] M. Virvou, A. Savvopoulos, An intelligent TV-shopping Application that provides Recommendations, 19th IEEE International Conference on Tools with Artificial Intelligence ICTAI 2007, 29-31 October 2007, Patras ,Greece.
- [4] Μ. Βίρβου, 'Σημειώσεις για το μάθημα Μοντελοποίηση Χρηστών'
- [5] Μ. Βίρβου, 'Σημειώσεις για το μάθημα Ειδικά θέματα Τεχνολογίας Λογισμικού'
- [6] Bill Kules, User Modeling for Adaptive and Adaptable Software Systems Department of Computer Science University of Maryland, College Park, MD 20742 USA
- [7] M. Virvou, A. Savvopoulos, D. N. Sotiropoulos, G. A. Tsihrintzis : Constructing Stereotypes for an Adaptive e-Shop using AIN-based Clustering, International Conference on Adaptive and Natural Computing Algorithms ICANNGA 2007, 11-14 April 2007
- [8] [SFW99] Spiliopoulou, M., Faulstich, L. C. and Wilkier, K.: *A data miner analyzing the navigational behavior of Web users*, Proceedings of the Workshop on Machine Learning in User Modeling of the ACAI99, Chania, Greece, pp.54-64, 1999
- [9] [PPK+00] Paliouras, G., Papatheodorou, C., Karkaletsis, V. and Spyropoulos, C. D.:*Clustering the Users of Large Web Sites into Communities*, Proceedings of International Conference on Machine Learning (ICML), Stanford, California, pp. 719-726, 2000
- [10]B. M. Sarwar, G. Karypis, J. A. Konstan, and J.Riedl, Item-based collaborative filtering recommendation algorithms, in Proceedings of the 10th International World Wide Web Conference (WWW10), Hong Kong, May 2001.
- [11]Tom Morrison, Uwe Aickelin, An Artificial Immune System as a Recommender for Web Sites, In Proceedings of the 1st Internal Conference on ARTificial Immune Systems (ICARIS-2002), pp 161-169, Canterbury, UK, 2002.
- [12]Η. Κοντοδός, Προσαρμοστικό σύστημα δημιουργίας προτάσεων σε ηλεκτρονικό κατάστημα.
- [13]Bing Liu, Web data mining: exploring hyperlinks, contents, and usage data
- [14]Marden J. I., 1995, Analyzing and Modeling Rank Data, Chapman & HLL, London
- [15]Everit, B.S., 1978, Graphical Techniques for Multivariate data
- [16]Badrul Saruar, George Karypis, Joseph Konstan and John Riedl, Item-based Collaborative Filtering Recommendation Algorithms
- [17]Hadi Khosravi Farsani, and Mohammadali Nematbakhsh, A Semantic Recommendation Procedure for Electronic Product Catalog
- [18]Al Mamunur Rashid George Karypis John Riedl, Influence in Ratings-Based Recommender Systems: An Algorithm-Independent Approach
- [19]P. Resnick, N. Iacovou, M. Sushak, P. Bergstrom, and J. Riedl, GroupLens: An open architecture for collaborative filtering of netnews, in Proceedings of CSCW 1994, ACM SIG Computer Supported Cooperative Work, 1994.
- [20]Εμμ. Α. Γιακουμάκης, Τεχνολογία λογισμικού, Τόμος Β, Εκδόσεις Α. Σταμούλη 1993.
- [21]J. S. Breese, D. Heckerman and C. Kadie, Empirical analysis of predictive algorithms for collaborative filtering, in Proceedings of the Fourteenth Annual Conference on Uncertainty in AI, July 1998.

- [22] Karamolegkos, P.N., Patrikakis C.Z., Doulamis N.D, Tragos E.Z., User - Profile based Communities Assessment using Clustering Methods in Personal, Indoor and Mobile Radio Communications, 2007, PIMRC 2007. IEEE 18th International Symposium
- [23] J. Ben Schafer, Joseph A. Konstan, John Riedl E-Commerce Recommendation Applications
- [24] P. Domingos and M. Richardson, Mining the Network Value of Customers, Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, 2001. ACM Press
- [25] L. Kaufman and P.J Rousseeuw. Finding Groups in Data, Wiley 1990.
- [26] J. H. Ward. Hierarchical groupings to optimize an objective function. Journal of the American Statistical Association, 58:234 244, 1963
- [27] C. Fraley. Algorithms for Model-Based Gaussian Hierarchical Clustering, 1996
- [28] George Karypis. Evaluation of Item-Based Top-N Recommendation Algorithms, in CIKM '01 Proceedings of the tenth international conference on Information and knowledge management
- [29] R. Burke and J. Sandvig, Model-Based Collaborative Filtering as a Defense Against Profile Injection Attacks, In Proceedings of the 21<sup>st</sup> National Conference on Artificial Intelligence (AAAI '06), Boston, Massachusetts, July 16-20, 2006
- [30] Εμμ. Σκορδαλάκη, Εισαγωγή στην τεχνολογία λογισμικού, 1991
- [31] Y. H Wu and A. L. P. Chen, Index Structures of User Profiles for Efficient Web Page Filtering Services, Proceedings of IEEE Conference on Distributed Computing Systems, pp. 644-651, April 2000.
- [32] T. W. Yan and H. Garcia-Molina, Index Structures for Selective Dissemination of Information under the Boolean Model, ACM Transactions on Database Systems, 19(2): 332-364, 1994.
- [33] M. J. Pazzani, Daniel Billsus, Content-Based Recommendation Systems, Lecture Notes in Computer Science Volume 4321/2007
- [34] Bergstrom, S. Raberg, L. (2003). Adopting The Rational Unified Process: Success With RUP. Addison-Wesley
- [35] Booch, G. Jacobson, I. Rumbaugh, J. (1999). The Unified Software Development Process. Addison-Wesley
- [36] Yanchu Zhang, Guandong Xu, Xiaofang Zhou, A Latent Usage Approach for Clustering Web Transaction and Building User Profile
- [37] Michael W. Trosset, Representing Clusters: K-Means Clustering, Self-Organizing Maps, and Multidimensional Scaling, February 20, 2008
- [38] Daniel Lemire, Anna Maclachlan, Slope One Predictors for Online Rating-Based Collaborative Filtering, In SIAM Data Mining (SDM'05), Newport Beach, California, April 21-23, 2005.
- [39] Umardand Shripad Manikrao, T.V.Prabhakar, Dynamic Selection of Web Services with Recommendation System, International Conference on Next Generation Web Services Practices, August 2005, Seoul, Korea.
- [40] Petros Drineas, Iordanis Kerenidis, and Prabhakar Raghavan, Competitive recommendation systems. In Proc. of the thirty fourth annual ACM symposium on Theory of computing
- [41] Philip Bonhard, Clare Harries, John McCarthy, M. Angela Sasse, Accounting for Taste: Using Profile Similarity to Improve Recommender Systems, In CHI 2006 Proceedings, Social Computing 2
- [42] James Bennett, Stan Lanning, The Netflix Prize
- [43] <http://www.netflixprize.com>

- [44] Herlocker, J, Konstan, J., Terveen, L., and Riedl, J. Evaluating Collaborative Filtering Recommender Systems. *ACM Transactions on Information Systems* 22 (2004), ACM Press
- [45] Distances between Clustering, Hierarchical Clustering, 36-350, *Data Mining*, 14 September 2009
- [46] Μ. Βαρζιγιάννης – Μ. Χαλκίδη, Εξόρυξη γνώσης από βάσεις δεδομένων, 2003
- [47] Ε. Κιουντούζης, Ανάλυση και σχεδιασμός Πληροφοριακών Συστημάτων, 2003
- [48] Ν. Αβούρης, Εισαγωγή στην επικοινωνία ανθρώπου – υπολογιστή, Εκδόσεις διάυλος, 2000
- [49] Ι. Βλαχάβας, Π. Κεφάλας, Ν. Βασιλειάδης, Φ. Κοκκορας, Η. Σακελλαρίου, Τεχνητή Νοημοσύνη Γ' Έκδοση, Εκδόσεις Γκιούρδα, 2006
- [50] Ε. Ι. Γιαννακουδάκης, Σχεδιασμός και διαχείριση Βάσεων Δεδομένων, Εκδόσεις Ευγ. Μπένου, 1999
- [51] <http://people.revoledu.com>
- [52] <http://www.php.net>
- [53] <http://www.mysql.com>
- [54] <http://www.efpolis.gr/el/diasfalisi-oikonomikon-symefronton-katanaloton/ilektroniko-emporio.html>
- [55] [http://en.wikipedia.org/wiki/Collaborative\\_filtering](http://en.wikipedia.org/wiki/Collaborative_filtering)