



ΔΙΕΘΝΕΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΗΣ ΕΛΛΑΔΟΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΕΥΦΥΕΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΔΙΚΤΥΟΥ - WEB INTELLIGENCE

**Συστήματα Συστάσεων
για πλατφόρμες ηλεκτρονικής μάθησης**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

Δήμητρας Κυριακού

Επιβλέπων : Κωνσταντίνος Διαμαντάρας
Καθηγητής

Θεσσαλονίκη, Φεβρουάριος 2022



ΔΙΕΘΝΕΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΗΣ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΕΥΦΥΕΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΔΙΚΤΥΟΥ - WEB INTELLIGENCE

**Συστήματα Συστάσεων
για πλατφόρμες ηλεκτρονικής μάθησης**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

Δήμητρας Κυριακού

Επιβλέπων : Κωνσταντίνος Διαμαντάρας
Καθηγητής, ΔΙΠΑΕ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 28^η Φεβρουαρίου 2022.

(Υπογραφή)

.....
Όνομα Επώνυμο
Καθηγητής

(Υπογραφή)

.....
Όνομα Επώνυμο
Καθηγητής

(Υπογραφή)

.....
Όνομα Επώνυμο
Καθηγητής

Θεσσαλονίκη, Φεβρουάριος 2022

(Υπογραφή)

.....

ΔΗΜΗΤΡΑ ΚΥΡΙΑΚΟΥ
Συστήματα Συστάσεων
για πλατφόρμες ηλεκτρονικής μάθησης
© **1111** – All rights reserved

Περίληψη

Στη σύγχρονη εποχή τα συστήματα ηλεκτρονικής μάθησης είναι πλέον ευρέως αποδεκτά και χρησιμοποιούνται όλο και περισσότερο τόσο από Πανεπιστημιακά Ιδρύματα, όσο και από ιδιώτες, αλλά και επιχειρήσεις, που σκοπό έχουν τη διαρκή επιμόρφωση των μελών τους και την ενίσχυση της διά Βίου Μάθησης. Βασικό εργαλείο στην επίτευξη των στόχων αυτών των συστημάτων, αποτελούν τα συστήματα συστάσεων που έρχονται να προσδώσουν προσωπικό χαρακτήρα στη μαθησιακή διαδικασία, καθώς ο κάθε χρήστης έρχεται αντιμέτωπος με έναν κυκεώνα πληροφοριών. Τα Συστήματα Συστάσεων είναι αυτά που προτείνουν εκπαιδευτικό υλικό κατάλληλο και προσαρμοσμένο στις ανάγκες του κάθε χρήστη ξεχωριστά. Είναι ενσωματωμένα σε κάθε πλατφόρμα ηλεκτρονικής μάθησης και αποτελούν αντικείμενο διαρκούς έρευνας και βελτίωσης.

Η παρούσα Μεταπτυχιακή Διπλωματική Εργασία αποτελεί μία συστηματική βιβλιογραφική έρευνα και εστιάζει στην περιοχή των Συστημάτων Συστάσεων στην εκπαίδευση και ειδικότερα στην ηλεκτρονική μάθηση. Μέσα από αυτή την έρευνα γίνεται αναφορά και επεξήγηση για το τι είναι οι πλατφόρμες ηλεκτρονικής μάθησης όπως επίσης και η διερεύνηση σε βάθος τι ορίζεται ως σύστημα συστάσεων, ποιοι αλγόριθμοι εξυπηρετούν τα συστήματα αυτά, ποιες μέθοδοι χρησιμοποιούνται για την υλοποίησή τους και την επιτυχία τους. Στόχος αυτής της εργασίας είναι να παρουσιάσει τα εκπαιδευτικά συστήματα συστάσεων από κάθε πτυχή, πραγματοποιεί μία ολιστική προσέγγιση και τελικά δείχνει που εστιάζουν σήμερα οι έρευνες. Ο ρόλος των συστημάτων αυτών είναι ενισχυμένος. Οι μέθοδοι που χρησιμοποιούνται στην ολοκλήρωση αυτών των συστημάτων αγγίζουν κάθε πτυχή των ευφών τεχνολογιών και προσφέρονται για περαιτέρω μελέτη.

Λέξεις-Κλειδιά: Ηλεκτρονική Μάθηση, εξ αποστάσεως εκπαίδευση, Συστήματα Συστάσεων, εξατομικευμένες συστάσεις, συνεργατικό φιλτράρισμα

Abstract

Nowadays, e-learning systems are now widely accepted and used more and more by both University Institutions and individuals, as well as companies, which aim at the continuous training of their members and the enhancement of Lifelong Learning. A key tool in achieving the goals of these systems, are the recommendation systems that come to give personal character to the learning process, as each user is faced with a barrage of information. Recommendation Systems are those that offer educational material suitable and adapted to the needs of each user individually. They are integrated in every e-learning platform and are constantly researched and improved.

This Master's Thesis is a systematic bibliographic research and focuses on the area of Recommendation Systems in education and in particular in e-learning. Through this research, reference and explanation are made about what e-learning platforms are as well as the in-depth investigation of what is defined as a recommendation system, what algorithms serve these systems, what methods are used for their implementation and their success. The aim of this work is to present the educational systems of recommendations from every aspect, it takes a holistic approach and finally shows where the researches focus today. The role of these systems is enhanced. The methods used to integrate these systems touch on every aspect of intelligent technologies and are offered for further study.

Keywords: E-Learning, Recommendation Systems, personalized recommendations, collaborative filtering

Ευχαριστίες

Θα ήθελα θερμά να ευχαριστήσω τον επιβλέποντα καθηγητή μου Δρ. Κωνσταντίνο Διαμαντάρα για την εργασία που μου ανέθεσε, τη βοήθεια που μου παρείχε και την υποστήριξη που έδειξε στο πρόσωπό μου. Είμαι ευγνώμων.

Θα ήθελα ακόμα να ευχαριστήσω την υποψήφια διδάκτωρ Μαρίνα Δελιανίδη για τις συμβουλές, κατευθύνσεις και διορθώσεις που έκανε στην εργασία μου και υπήρξε αρωγός στην προσπάθειά μου μέχρι τέλους.

Θα ήθελα να ευχαριστήσω όλους τους συναδέλφους που μου παρείχαν ηθική και πρακτική συμπαράσταση σε όλη τη διάρκεια του μεταπτυχιακού προγράμματος σπουδών και ιδιαίτερα την οικογενειακή φίλη και συνάδελφο Δέσποινα Καλαϊτζίδου.

Τέλος, οφείλω ένα τεράστιο ευχαριστώ στην οικογένειά μου που συμμερίστηκε τους κόπους και τις δυσκολίες αυτής της προσπάθειας, στον σύζυγο Τάσο και στα παιδιά μου Μαρία, Ηλία, Ηρώ και Νίκο.

Συντομογραφίες και Ακρωνύμια

AOES	Autonomous Online Education System
BBIRS	Browsing Behavior Personalized Information Recommendation System
BOW	Bag of Words
CLE	Cloud Learning Environment
CF	Collaborative Filtering
DBN	Deep Belief Network
DBNCF	Deep Belief Network Collaborative Filtering
ETL	Extract, Transform, Load (Data)
GIGO	Garbage In Garbage Out
IEEE	Institute of Electrical and Electronics Engineers
IR	Information Retrieval
LOM	Learning Objects Metadata (Standard)
LMS	Learning Management Systems
MAE	Mean Absolute Error
MCU	Multi Control Unit
ML	Machine Learning
MOOC	Massive Open Online Courses
OLS	Ordinary Least Square
PLE	Personal Learning Environment
RBM	Restricted Boltzmann Machine
RNN	Recurrent Neural Networks
RS	Recommendation Systems
TF-IDF	Term Frequency - Inverse Document Frequency
URL	Uniform Resource Locator
VLS	Virtual Management Systems
VSM	Vector Space Model
ΒΔ	Βάση Δεδομένων
ΜΜ	Μηχανική Μάθηση
ΣΣ	Συστήματα Συστάσεων

Πίνακας περιεχομένων

1	Εισαγωγή.....	7
1.1	Πεδίο έρευνας – σκοπός της διπλωματικής εργασίας.....	7
1.2	Αντικείμενο διπλωματικής.....	8
1.2.1	Συνεισφορά.....	9
1.3	Οργάνωση κειμένου.....	9
2	Πλατφόρμες ηλεκτρονικής μάθησης.....	11
2.1	Ηλεκτρονική Μάθηση, E-Learning	13
2.2	Προσαρμοστικά Μαθησιακά Συστήματα (Adaptive Learning Systems)	14
2.3	LMS-Learning Management System.....	16
2.4	MOOC Ecosystems.....	17
3	Συστήματα Συστάσεων(ΣΣ)	19
3.1	Ορισμός ΣΣ.....	19
3.2	Αναγκαιότητα Εκπαιδευτικών Συστημάτων Συστάσεων	20
3.3	Διάγραμμα ροής Προσωποποιημένου Συστήματος Εκπαιδευτικών Συστάσεων	22
3.3.1	Συλλογή πληροφοριών που αφορούν τον χρήστη.....	22
3.3.2	Μοντελοποίηση του χρήστη	23
3.3.3	Φιλτράρισμα ή ταίριασμα χρήστη με αντικείμενο	24
3.3.4	Παραγωγή Συστάσεων.....	24
3.4	Βασικός Διαχωρισμός των ΣΣ	24
3.4.1	Φιλτράρισμα βάσει περιεχομένου	25
3.4.2	Συνεργατικό φιλτράρισμα (Collaborative Filtering-CF).....	26
3.4.3	Υβριδικά Συστήματα.....	29
3.4.4	Δημογραφικές Συστάσεις.....	29
3.4.5	Συστάσεις βάσει Γνώσεων.....	29
3.4.6	Συστήματα συστάσεων με επίγνωση περιβάλλοντος	30
3.4.7	Συγκριτικοί πίνακες των βασικών τεχνικών δημιουργίας συστημάτων συστάσεων	30

3.4.8	Συστήματα συστάσεων με βάση τις οντολογίες	32
3.4.9	Συστήματα Συστάσεων με χρήση τεχνικών εξόρυξης δεδομένων (Data mining).32	
3.5	Προκλήσεις στα ΣΣ.....	34
3.5.1	Πρόβλημα ψυχρής εκκίνησης (Cold-start problem).....	35
3.5.2	Διασπορά δεδομένων (Data Sparsity).....	35
3.5.3	Επεκτασιμότητα (Scalability).....	35
3.5.4	Υπέρ-εξειδίκευση (Over Specialization).....	35
3.5.5	Ιδιωτικότητα (Privacy).....	36
4	Εφαρμογές Εκπαιδευτικών ΣΣ.....	37
4.1	ΣΣ με Vector Space Model και Μέσο Όρο Αξιολόγησης καλών σπουδαστών	37
4.2	ΣΣ με χρήση αλγορίθμων για Συνεργατικό φιλτράρισμα	40
4.2.1	ΣΣ στην ηλεκτρονική μάθηση	40
4.2.2	Συνεργατικό φιλτράρισμα με χρήση k-means αλγορίθμου.....	42
4.2.3	Αυτόνομο Διαδίκτυο Εκπαιδευτικό Σύστημα με ευφρείς συστάσεις	47
4.3	ΣΣ βασισμένο σε Νευρωνικά Δίκτυα Βαθιάς Πίστης για εξατομικευμένα μαθήματα σε περιβάλλον MOOC	53
4.4	ΣΣ με χρήση οντολογιών	57
4.5	ΣΣ με τεχνικές εξόρυξης δεδομένων (Data mining)	59
4.5.1	Σύστημα BBIRS.....	59
4.5.2	ΣΣ με χρήση του αλγορίθμου Apriori.	61
4.6	Συγκεντρωτικός πίνακας ΣΣ στην παρούσα εργασία	65
4.7	Τρόποι αξιολόγησης ΣΣ.....	68
5	Σύγχρονες τάσεις στα ΣΣ	72
5.1	Η έννοια της «ευτυχούς συγκυρίας» -serendipity στα ΣΣ και το πρόβλημα της «φυσαλίδας φίλτρου»- filter bubble.....	72
5.1.1	Μοντέλα και Μεθοδολογία.....	74
5.1.2	Ανάδυση συστάσεων «ευτυχούς συγκυρίας» από αναπαραστάσεις μαθημάτων .	75
5.1.3	Συμπεράσματα της έρευνας	75
5.2	Εξηγήσεις προτάσεων στα ΣΣ	76
5.3	Αξιολόγηση του προτεινόμενου συστήματος.....	77
6	Επίλογος	80

6.1	Σύνοψη και συμπεράσματα.....	80
6.2	Μελλοντικές επεκτάσεις	81
7	Βιβλιογραφία.....	82

Ευρετήριο Εικόνων

Εικόνα 1. Γενική αρχιτεκτονική Συστήματος Συστάσεων [15].....	29
Εικόνα 2. Διάγραμμα ροής προσωποποιημένου ΣΣ [17].....	31
Εικόνα 3. Σχηματικό διάγραμμα συστάσεων βάσει περιεχομένου [17].....	34
Εικόνα 4. Σχηματικό διάγραμμα συστάσεων συνεργατικού φιλτραρίσματος [17].....	36
Εικόνα 5. Σχηματικό διάγραμμα συστάσεων βάσει συσχετιστικών κανόνων [17].....	44
Εικόνα 6. Συνολική αρχιτεκτονική του ΣΣ ηλεκτρονικής μάθησης [16].....	48
Εικόνα 7. Διάγραμμα ροής ΣΣ [16].....	49
Εικόνα 8. Διάγραμμα ροής εκτέλεσης αλγορίθμου k-means [21].....	55
Εικόνα 9. Πλαίσιο λειτουργίας AOES [1].....	60
Εικόνα 10. Αρχιτεκτονική Συστήματος DBNCF [22].....	66
Εικόνα 11. Η πυραμίδα των μαθησιακών οντολογιών [23].....	69
Εικόνα 12. ΣΣ Προσωποποιημένων πληροφοριών συμπεριφοράς περιήγησης (BBIRS) [3]..	71
Εικόνα 13. Βασικό Διάγραμμα Ροής Συστάσεων Βίντεο [20].....	74
Εικόνα 14. Συστάσεις βασισμένες στις θεάσεις βίντεο του χρήστη [20].....	75

Ευρετήριο Πινάκων

Πίνακας 1. Σύγκριση των διάφορων μοντέλων φιλτραρίσματος [4].....	39
Πίνακας 2. Σύγκριση τεχνικών για συστάσεις [12].....	40
Πίνακας 3. Πίνακας αξιολόγησης μαθήματος από σπουδαστή [14].....	51
Πίνακας 4. Μέσες βαθμολογίες μαθητών για προτάσεις μεμονωμένων μαθημάτων από τη μελέτη χρήστη που αναλύονται ανά μοντέλο που χρησιμοποιήθηκε για τη δημιουργία προτάσεων μαθημάτων, μέθοδο που χρησιμοποιήθηκε για τη δημιουργία των επεξηγήσεων και δομή βαθμολογίας. [27].....	88
Πίνακας 5. Αντίκτυπος των ΣΣ και των στρατηγικών εξηγήσεων σύμφωνα με τις OLS παλινδρομήσεις [27].....	89

1

Εισαγωγή

1.1 Πεδίο έρευνας – σκοπός της διπλωματικής εργασίας

Η εξέλιξη του Διαδικτύου και η ανάπτυξη ποικίλων διαδικτυακών υπηρεσιών έφεραν επανάσταση σε πολλούς τομείς όπως η δημόσια διοίκηση, η επικοινωνία, το ηλεκτρονικό εμπόριο, τα τουριστικά επαγγέλματα κα. Σημαντικές αλλαγές προέκυψαν και στο χώρο της εκπαίδευσης. Η χρήση του διαδικτύου και των παρεχόμενων υπηρεσιών του, οδήγησε σε μια κατάσταση μετάβασης από την παραδοσιακή σχολική τάξη στην ηλεκτρονική, εικονική τάξη ή ηλεκτρονικό μαθησιακό περιβάλλον. Μια από τις αλλαγές που συνοδεύουν το ηλεκτρονικό μαθησιακό περιβάλλον είναι η ανάπτυξη συστημάτων ηλεκτρονικής μάθησης. Αυτοί οι διαδικτυακοί χώροι μάθησης παρέχουν πληθώρα πληροφοριών και ποικίλο ψηφιακό υλικό εκπαιδευτικού περιεχομένου όπως βίντεο, διαλέξεις, ασκήσεις, quiz κλπ, πρακτικά για οποιοδήποτε επιστημονικό πεδίο και θέμα που μπορεί να ενδιαφέρει κάποιον σπουδαστή, ερευνητή, εκπαιδευτή ή εκπαιδευόμενο. Το γεγονός αυτό έφερε τεράστιες αλλαγές στον τρόπο μάθησης από την πλευρά των φοιτητών, και στον τρόπο οργάνωσης του εκπαιδευτικού υλικού από την πλευρά των εκπαιδευτών, αλλά και των υπολοίπων εμπλεκόμενων μερών όπως για παράδειγμα των τεχνικών υποστήριξης συστημάτων ηλεκτρονικής μάθησης. Ως εκ τούτου νέοι τρόποι και δικτυακοί χώροι μάθησης εμφανίστηκαν με έναν καταγιστικό ρυθμό παρέχοντας έναν τεράστιο όγκο πληροφορίας, που όμως, είναι δύσκολα διαχειρίσιμος από τον μέσο χρήστη.

Εύλογα δημιουργήθηκε η ανάγκη παροχής συστάσεων και μάλιστα προσωποποιημένων για την εξεύρεση του κατάλληλου εκπαιδευτικού περιεχομένου που αφορά το πεδίο ενδιαφέροντος ή/και έρευνας του κάθε φοιτητή/σπουδαστή. Έτσι, εκτός από τα συστήματα συστάσεων στον εμπορικό ή τον τουριστικό τομέα, αναπτυχθήκαν διάφορων ειδών συστήματα συστάσεων στον τομέα της εκπαίδευσης με σκοπό την στοχευμένη σύσταση

κατάλληλου εκπαιδευτικού περιεχομένου από έγκυρες πηγές και σύμφωνα με τις ανάγκες των εκπαιδευόμενων.

Σκοπός της διπλωματικής εργασίας είναι η βιβλιογραφική έρευνα, η καταγραφή και η μελέτη των συστημάτων συστάσεων που έχουν αναπτυχθεί για κάθε μορφή ηλεκτρονικής μάθησης, όπως επίσης, ο εντοπισμός και η παράθεση καινών σημείων ερευνάς στον συγκεκριμένο τομέα.

1.2 Αντικείμενο διπλωματικής

Το αντικείμενο της διπλωματικής εργασίας επικεντρώνεται στην ανάλυση, περιγραφή, ανάδειξη των συστημάτων συστάσεων στον τομέα της εκπαίδευσης και ειδικότερα της ηλεκτρονικής μάθησης. Τα συστήματα συστάσεων συμβάλλουν στην εξεύρεση πληροφοριών με τα καλύτερα δυνατά αποτελέσματα για τους χρήστες πλατφορμών ηλεκτρονικής μάθησης με την δυνατότητα παροχής κατάλληλων προτάσεων μαθημάτων, εκπαιδευτικού περιεχομένου, ψηφιακού υλικού (βίντεο, διαλέξεις, συγγράμματα από αποθετήρια βιβλιοθηκών Πανεπιστημιακών Ιδρυμάτων) κα. Οι προτάσεις αυτές πολλές φορές είναι προσωποποιημένου χαρακτήρα και παράγονται βάσει κριτηρίων όπως οι προτιμήσεις των χρηστών, του ιστορικού αναζήτησης, η ομοιότητα των χρηστών ή του περιεχομένου κα.

Οι συστάσεις εκπαιδευτικού περιεχομένου μπορούν να χαρακτηριστούν ως σχετικά καινούριο αντικείμενο διότι διαφέρει από άλλες εφαρμογές συστάσεων όπως στον τομέα της ψυχαγωγίας ή ηλεκτρονικού εμπορίου. Η διαπίστωση αυτή πηγάζει από το γεγονός ότι το προτεινόμενο περιεχόμενο θα πρέπει να ενισχύει και να εμπλουτίζει την γνώση του εκπαιδευόμενου η οποία μεταβάλλεται δυναμικά, αυξητικά και με διαφορετικό ρυθμό για τον καθένα. Έτσι, τα εκπαιδευτικά συστήματα συστάσεων διαφοροποιούνται από τα συστήματα συστάσεων άλλων τομέων διότι κομβικό σημείο μιας πετυχημένης σύστασης είναι το τρέχων γνωστικό επίπεδο του εκπαιδευόμενου πάνω σε ένα συγκεκριμένο μαθησιακό αντικείμενο και όχι μόνο το γενικό εκπαιδευτικό πλαίσιο που μελετάει.

Στην παρούσα βιβλιογραφική έρευνα χρησιμοποιήθηκαν ως λέξεις κλειδιά οι όροι «ηλεκτρονική μάθηση», «συστήματα συστάσεων», «εξατομικευμένες συστάσεις», «αλγόριθμοι συστημάτων συστάσεων». Η έρευνα οδήγησε σε διάφορες μεθόδους υλοποίησης των συστημάτων συστάσεων και επιλέχθηκαν τα πιο αντιπροσωπευτικά κάθε μεθόδου. Στόχος της εργασίας είναι να παρουσιαστούν τα συστήματα συστάσεων στην εκπαίδευση και ειδικότερα στην ηλεκτρονική εκπαίδευση, να εξηγηθούν συνοπτικά οι διαφορές ανάμεσα σε διάφορους χώρους ηλεκτρονικής μάθησης, οι διαφορές ανάμεσα στις ανάγκες του κάθε εκπαιδευόμενου και τέλος, γιατί ένα επιτυχημένο εκπαιδευτικό ηλεκτρονικό σύστημα

συνοδεύεται και από ένα σύστημα συστάσεων. Η εργασία λοιπόν αυτή, δίνει τους ορισμούς των συστημάτων συστάσεων και αναδεικνύει τους βασικούς τρόπους υλοποίησης τους. Επίσης, δείχνει τρόπους αξιολόγησης των συστημάτων αυτών ώστε να είναι όσο το δυνατόν πιο επιτυχημένα και εύστοχα στις προτάσεις που δίνουν στους χρήστες. Τέλος, ενημερώνει για τις τρέχουσες εξελίξεις στον τομέα των εξατομικευμένων συστάσεων για εκπαιδευτικά περιβάλλοντα.

Η μελέτη αυτή θα μπορέσει να βοηθήσει κάθε ερευνητή που επιθυμεί να γνωρίσει και να ασχοληθεί με τα συστήματα συστάσεων για πλατφόρμες ηλεκτρονικής μάθησης. Έχει στη διάθεσή του μία εργασία που συγκεντρώνει πληροφορίες από κάθε οπτική γωνία για τα συστήματα αυτά, αλγορίθμους που επιτυγχάνουν βελτιωμένες συστάσεις και πληροφορίες για τις τρέχουσες τάσεις στον χώρο των εκπαιδευτικών συστάσεων.

1.2.1 Συνεισφορά

Η συνεισφορά της διπλωματικής συνοψίζεται ως εξής:

1. Ορισμός συστημάτων συστάσεων και θεωρητική ανάλυση αλγορίθμων.
2. Εφαρμογή διαφόρων αλγορίθμων σε μελέτες περίπτωσης.
3. Μελέτη συστημάτων ηλεκτρονικής μάθησης και παιδαγωγικοί στόχοι.
4. Ορισμός «ευτυχούς συγκυρίας» και επεξηγήσεων συστάσεων.
5. Συμπεράσματα και ανάδειξη της σημαντικότητας των συστημάτων συστάσεων στην εκπαιδευτική διαδικασία.

1.3 Οργάνωση κειμένου

Η εργασία με θέμα: «Συστήματα συστάσεων για πλατφόρμες ηλεκτρονικής μάθησης» προκειμένου να αναλυθεί το θέμα της, κινήθηκε σε δύο άξονες:

Αρχικά, στο κεφάλαιο 2 δόθηκε ορισμός της ηλεκτρονικής μάθησης και αναλύθηκαν θεωρητικά τα διάφορα συστήματα που παρέχουν τέτοιου είδους υπηρεσίες. Επίσης, παρουσιάστηκε γιατί χρειάζεται να παρέχονται εξατομικευμένες συστάσεις στους χρήστες των μέσων αυτών. Στη συνέχεια στο Κεφάλαιο 3 εξηγείται τι ονομάζεται σύστημα συστάσεων, ποιοι είναι οι στόχοι του και ποιοι οι βασικοί τρόποι υλοποίησης ενός τέτοιου συστήματος. Στο Κεφάλαιο 4 δίνονται αντιπροσωπευτικά παραδείγματα αλγορίθμων που εφαρμόστηκαν σε διάφορες πλατφόρμες ηλεκτρονικής μάθησης είτε σε ερευνητικό επίπεδο είτε σε πραγματικές συνθήκες. Στο κεφάλαιο 5 προβάλλονται οι σύγχρονες τάσεις εξέλιξης και βελτίωσης των συστημάτων προσωποποιημένων συστάσεων για εκπαιδευτικά

περιβάλλοντα και δείχνουν που τείνει η έρευνα γύρω από αυτά. Τέλος, στο κεφάλαιο 6 συνοψίζονται τα συμπεράσματα της παρούσας βιβλιογραφικής έρευνας.

2

Πλατφόρμες ηλεκτρονικής μάθησης

Στον 21ο αιώνα, η τεχνολογία της πληροφορίας σαρώνει ολόκληρο τον κόσμο, αλλάζει τον τρόπο ζωής και σκέψης της ανθρώπινης κοινωνίας και παρέχει ισχυρή τεχνολογική βάση για την ριζοσπαστική εξέλιξη της παραδοσιακής εκπαίδευσης. Το εκπαιδευτικό σύστημα του μέλλοντος δημιουργείται από τον συνδυασμό δύο βασικών κινήτρων της οικονομίας, το Διαδίκτυο και την μόρφωση. Η διαδικτυακή εκπαίδευση συνδέει τους δασκάλους και τους μαθητές μεταξύ τους με χιλιάδες ηλεκτρονικά μηνύματα και πραγματώνει με αυτόν τον τρόπο την επικοινωνία πρόσωπο με πρόσωπο. Το γεγονός αυτό φέρνει νέα αντίληψη στο τι είναι διδασκαλία και μάθηση και αλλάζει κατά πολύ την εκπαίδευση. Συνεπώς, η μάθηση μέσω ηλεκτρονικών μέσων αποτελεί τη νέα τάση και απασχολεί ήδη όλο και περισσότερες χώρες [1].

Η ηλεκτρονική μάθηση αποτελεί την σύγχρονη καινοτομία στο χώρο των τεχνολογιών πληροφορίας και διαδραματίζει σημαντικό ρόλο στην εκπαίδευση και την έρευνα [2]. Η ολιστική προσέγγιση διαχείρισης της γνώσης μέσω της ηλεκτρονικής εκπαίδευσης αποτελεί σημαντική πτυχή για τις εκπαιδευτικές κοινότητες ακόμα και για τις επιχειρήσεις. Γίνεται όλο και περισσότερο αποδεκτή από εταιρείες, κοινωνία, ακαδημαϊκά περιβάλλοντα και προσφέρει διάφορα είδη εκπαιδευτικών υπηρεσιών ευέλικτων και έγκαιρων, ανεξάρτητα από γεωγραφικούς, χρονικούς, φυσικούς, κοινωνικούς και οικονομικούς περιορισμούς [3]. Τα τρέχοντα συστήματα διαχείρισης γνώσης περιέχουν μία τεράστια δεξαμενή δεδομένων, τα οποία συλλέγουν από πολλαπλές πηγές [4].

Παρόλο που τα οφέλη και οι δυνατότητες των συστημάτων αυτών είναι προφανή δεν έχει γίνει πλήρης εκμετάλλευση των πλεονεκτημάτων τους. Δεν έχει εκπληρωθεί ακόμα η υπόσχεση να γίνουν τα σημαντικότερα παραδείγματα μάθησης, ιδιαίτερα στο πλαίσιο του αυξανόμενου ρόλου της συνεχούς και δια βίου μάθησης. Παρά την εντυπωσιακή ανάπτυξη που γνωρίζουν, τα συστήματα ηλεκτρονικής μάθησης εξακολουθούν να είναι λιγότερο ελκυστικά τόσο για τους εκπαιδευόμενους όσο και για τους εκπαιδευτές [3].

Στη σύγχρονη βιβλιογραφία ένας μεγάλος αριθμός μελετών αποσκοπεί στην διασαφήνιση αν η εκπαίδευση με τη μορφή της ηλεκτρονικής μάθησης, η μικτή ή υβριδική μάθηση είναι καλύτερη από την παραδοσιακή δασκαλοκεντρική εκπαίδευση. Οι ερευνητές, οι εκπαιδευτές, και οι σχεδιαστές εκπαιδευτικών συστημάτων αναζητούν εναγωνίως ποιες εκπαιδευτικές μορφές και διαδικασίες δίνουν τα καλύτερα αποτελέσματα για τους σπουδαστές και τα εκπαιδευτικά ιδρύματα. Όμως οι συγκριτικές μελέτες δείχνουν διαφορετικά αποτελέσματα γεγονός που υποδεικνύει πώς και άλλες παράμετροι πέρα από τη μορφή του εκπαιδευτικού συστήματος επηρεάζουν το μαθησιακό αποτέλεσμα, την ικανοποίηση του φοιτητή, την παραμονή και αφοσίωση στις σπουδές κλπ [5].

Η βιβλιογραφική έρευνα των [5] προσπαθεί να απαντήσει το βασικό ερώτημα: ποιες είναι εκείνες οι παράμετροι που επηρεάζουν την ηλεκτρονική και την μικτή μάθηση σε σχέση με την ικανοποίηση του φοιτητή και την αφοσίωση του ιδιαίτερα στις ανώτερες εκπαιδευτικές βαθμίδες και την επαγγελματική εκπαίδευση;

Δυστυχώς, το περιεχόμενο του υπάρχοντος διαδικτυακού εκπαιδευτικού συστήματος αποτελεί μάζα συγκεχυμένου εκπαιδευτικού υλικού, χωρίς να λαμβάνονται υπόψη η προσωπικότητα του σπουδαστή και οι συνήθειες που διατηρεί. Απαιτήση είναι η προσαρμογή του φοιτητή στο εκάστοτε σύστημα και όχι το αντίθετο, δηλαδή, να προσαρμόζεται αυτόματα το σύστημα στον μαθητή και να σχεδιάζεται σύμφωνα με το «νόμο» του μαθητή. Επομένως, υπάρχουν πάντα προβλήματα, όπως και φτωχή προσαρμοστικότητα και αναποτελεσματικότητα [1].

Η κύρια πρόκληση που αντιμετωπίζουν τα συστήματα αυτά είναι πως θα καταφέρουν να προτείνουν κατάλληλο υλικό για τον κάθε χρήστη ώστε να μη σπαταλά χρόνο στην αναζήτηση πληροφοριών. Επιπλέον, οι διαφορές στην μαθησιακή ικανότητα του κάθε ατόμου οδηγούν στο συμπέρασμα πως δεν μπορούν όλοι οι άνθρωποι να ακολουθήσουν την ίδια μαθησιακή διαδρομή προκειμένου να κατανοήσουν ένα αντικείμενο [4]. Από την πλευρά των φοιτητών υπάρχουν παράπονα για έλλειψη εργαλείων του συστήματος με ευελιξία, διαμοιρασμό γνώσεων, εξειδίκευση κατά απαίτηση, έγκαιρη βοήθεια και επαρκή καθοδήγηση από τους διδάσκοντες. Από την πλευρά των εκπαιδευτικών το κύριο μειονέκτημα είναι πως η παραγωγή υλικού είναι περισσότερο χρονοβόρα από ότι με τους παραδοσιακούς τρόπους ιδίως αν δεν υπάρχει ικανοποιητική εξοικείωση με τη χρήση τεχνολογιών [4].

Με βάση τους προαναφερθέντες περιορισμούς που συναντώνται στα υπάρχοντα συστήματα ηλεκτρονικής μάθησης χρειάζεται να προσδοθεί ευφυΐα σε αυτά τα μαθησιακά περιβάλλοντα με ενίσχυση εξατομικευμένων μαθησιακών υπηρεσιών και δημιουργία καινοτόμων μηχανισμών ώστε ο φόρτος εργασίας των διδασκόντων στην παραγωγή διδακτικού υλικού να μειωθεί και να διευκολυνθεί με την απόκρυψη τεχνικών λεπτομερειών. Οι παραπάνω παράμετροι είναι το κλειδί για μελλοντική επιτυχία της ηλεκτρονικής εκπαίδευσης [3].

2.1 Ηλεκτρονική Μάθηση, E-Learning

Η ηλεκτρονική μάθηση ορίζεται σε αντίθεση με την κατά πρόσωπον διδασκαλία και με βάση το κύριο χαρακτηριστικό της που είναι η απουσία τάξης, η οποία αντικαθιστάται από τεχνολογίες Διαδικτύου ανεξάρτητες από τον χώρο, το χρόνο και το βηματισμό. Επίσης, ορισμένες φορές ερμηνεύεται ως προσφορά μαθημάτων μόνο μέσω Διαδικτύου. Συνήθως, η ηλεκτρονική μάθηση προσφέρεται μέσα από «Συστήματα Διαχείρισης Μάθησης» (Learning Management Systems–LMS) ή εικονικά εκπαιδευτικά περιβάλλοντα (Virtual Learning Environments-VLE) όπως πχ MOODLE και Blackboard [5].

Τα συστήματα αυτά αποτελούν μία σημαντική εναλλακτική λύση για ανθρώπους που δεν μπορούν να παρακολουθήσουν με φυσική παρουσία μαθήματα είτε λόγω απόστασης, είτε για λόγους υγείας, είτε για οικονομικούς ή κοινωνικούς λόγους. Ο καθηγητής προετοιμάζει και δημοσιοποιεί διδακτικό περιεχόμενο και οι φοιτητές το χρησιμοποιούν όσες φορές επιθυμούν. Ένα μειονέκτημα είναι πως το σύστημα είναι απρόσωπο και αντιμετωπίζει όλους τους σπουδαστές το ίδιο, παρόλο που ο καθένας από αυτούς μπορεί να έχει διαφορετικούς στόχους και ανάγκες. Η έλλειψη ευέλικτων στρατηγικών είναι μία δυσκολία αλλά υπάρχουν και θετικές πλευρές όπως η στατιστική ανάλυση για τις δραστηριότητες κάθε σπουδαστή [6]. Ένα άμεσο πλεονέκτημα για τους φοιτητές είναι το κέρδος του χρόνου (λιγότερο χρονοβόρες διαδικασίες) και ένα έμμεσο πλεονέκτημα είναι ότι περισσότεροι φοιτητές μπορούν να εργάζονται χωρίς πρόσθετα κόστη.

Η χρήση συστημάτων εκπαίδευσης με τεχνολογίες Διαδικτύου εξαπλώνεται γοργά σε όλο τον κόσμο. Το ενδιαφέρον γύρω από αυτές τις τεχνολογίες αυξήθηκε σημαντικά στη σύγχρονη εποχή. Το κυριότερο πλεονέκτημα είναι πως μπορεί κάποιος να αποκτήσει γνώσεις «οποτεδήποτε και οπουδήποτε» δηλαδή σε οποιοδήποτε χρόνο και χώρο [7].

Τα συστήματα ηλεκτρονικής μάθησης παρέχουν υλικό προερχόμενο από διάφορες πηγές και περιεχόμενα όπως βίντεο, ιστολόγια, άρθρα κλπ. Τα συστήματα διαχείρισης μάθησης παρέχουν στους χρήστες ένα περιβάλλον που τους επιτρέπει να διαχειρίζονται και να αναζητούν μικρές ενότητες περιεχομένου για καλύτερη διαδραστική μάθηση [4].

Η παρακολούθηση των ενεργειών ενός χρήστη ενός τέτοιου συστήματος θα βοηθούσε την εκπαιδευτική διαδικασία καθώς θα πρότεινε κατάλληλο περιεχόμενο και θα μπορούσε ακόμα να βοηθήσει τους διδάσκοντες στην παραγωγή κατανοητού περιεχομένου [7]. Ένας δείκτης αξιολόγησης για την κατανόηση της αποτελεσματικότητας ενός συστήματος ηλεκτρονικής μάθησης είναι: «ο λιγότερος χρόνος που χρειάζεται για την εκπαίδευση» [7].

Ως μικτή μάθηση μπορεί να θεωρηθεί ο συνδυασμός των παραδοσιακών μεθόδων με τις σύγχρονες πρακτικές που χρησιμοποιούν τεχνολογίες Διαδικτύου και σε ορισμένες περιπτώσεις αναγνωρίζεται ως ο περισσότερο αποδοτικός τρόπος μάθησης. Κάποιες μελέτες

ενισχύουν αυτή την άποψη, καθώς οι φοιτητές που χρησιμοποιούν προγράμματα μικτής μάθησης φαίνεται να πετυχαίνουν τους στόχους τους λίγο καλύτερα από τους φοιτητές που ακολουθούν μόνο προγράμματα εντός τάξης [5]. Υπάρχουν και μελέτες που οδηγούν στην αντίθετη άποψη καθώς φάνηκε πως ορισμένοι φοιτητές σε συνθήκες «απομόνωσης» καταλήγουν σε αποτυχία. Επίσης, αυτό που ταιριάζει σε έναν σπουδαστή μπορεί να αντενδείκνυται για κάποιον άλλο.

Ο σχεδιασμός των LMS συστημάτων θα πρέπει να αξιοποιεί στο μέγιστο τόσο τις τεχνολογικές εξελίξεις όσο και τη διεπιστημονική γνώση σχετικά με το ακαδημαϊκό περιεχόμενο, τη μάθηση και την καθοδήγηση. Τόσο τα παραδοσιακά συστήματα όσο και τα διαδικτυακά έχουν τα πλεονεκτήματά τους με αντίστοιχες χρήσεις αλλά και τους περιορισμούς τους. Το ένα μοντέλο εκπαίδευσης αλληλοσυμπληρώνει το άλλο και μία ολιστική προσέγγιση είναι αυτή που υιοθετείται ως καλύτερη δυνατή λύση.

Τα μαθησιακά περιβάλλοντα μέσω Διαδικτύου μπορούν να δημιουργήσουν ευκαιρίες για αλληλεπίδραση ανάμεσα σε φοιτητές, φοιτητές και καθηγητές, φοιτητές και περιεχόμενο αν και το αίσθημα του ανήκειν σε μία κοινότητα με κοινά ενδιαφέροντα είναι δύσκολο να εννοηθεί στο διαδικτυακό περιβάλλον [5].

Τρεις βασικές αρχές διέπουν την εκπαίδευση είτε πραγματοποιείται διαδικτυακά, είτε φυσικά:

- Κατάλληλη διδασκαλία και διδακτικά διαστήματα.
- Δημιουργία κοινοτήτων που να υποστηρίζουν τους φοιτητές στις κοινωνικές σχέσεις και τη μαθησιακή εμπειρία.
- Ισχυρή και εξέχουσα αίσθηση ταυτότητας του σπουδαστή [5].

Οι παραπάνω πληροφορίες δίνονται με σκοπό την εμβάθυνση στην κατανόηση του τι είναι μία πλατφόρμα ηλεκτρονικής μάθησης και γιατί δίνεται τόση έμφαση στην υποστήριξη αυτών των συστημάτων με συστήματα συστάσεων.

2.2 Προσαρμοστικά Μαθησιακά Συστήματα (Adaptive Learning Systems)

Σύμφωνα με την βιβλιογραφία [12], υπάρχουν δύο τύποι προσαρμογής για εξατομίκευση ενός μαθησιακού συστήματος:

- **προσωποποίηση συστήματος οδηγούμενη από τον χρήστη** - όταν το σύστημα επιτρέπει στο χρήστη να κάνει αλλαγές σε κάποιες παραμέτρους και να δίνει προσωπικό χαρακτήρα στη συμπεριφορά του και

-
- **προσωποποίηση μέσω συστήματος** - όταν το σύστημα πραγματοποιεί αλλαγές αυτόματα ανάλογα με τις ανάγκες και τις ενέργειες του χρήστη [7].

Σύμφωνα με τους [2] η κύρια διαφορά μεταξύ της προσαρμογής και της εξατομίκευσης είναι ότι η προσαρμογή πραγματοποιείται όταν ο χρήστης μπορεί να διαμορφώσει μια διεπαφή και να δημιουργήσει ένα προφίλ χειροκίνητα, προσθέτοντας και αφαιρώντας στοιχεία στο προφίλ. Στη διαδικασία προσαρμογής, η αντιστοίχιση προφίλ χρήστη είναι σαφής και καθοδηγείται από τον χρήστη. Ενώ στην άλλη περίπτωση της εξατομίκευσης, ο χρήστης θεωρείται ότι είναι παθητικός, ή με λιγότερο έλεγχο. Είναι ευθύνη του συστήματος να παρακολουθεί, να αναλύει και να αντιδρά στη συμπεριφορά. Για παράδειγμα, το περιεχόμενο που προσφέρεται μπορεί να βασίζεται σε αποφάσεις παρακολούθησης και πλοήγησης. Μια εξατομικευμένη υπηρεσία δεν χρειάζεται πάντα να βασίζεται στην ατομική συμπεριφορά του χρήστη ή στη συμβολή του χρήστη.

Τα κύρια καθήκοντα ενός τέτοιου συστήματος είναι τα εξής [7]:

- i. Οργάνωση του περιεχομένου με διάφορους τρόπους προσφέροντας διαφορετικά πλαίσια και προοπτική.
- ii. Αναγνώριση του τρόπου με τον οποίο προτιμά ο χρήστης να μαθαίνει και οργάνωση αποτίμησης των προτιμήσεων.
- iii. Χρήση των αξιολογήσεων για ανατροφοδότηση που ενθαρρύνει και βοηθά τη μεγιστοποίηση της μαθησιακής απόδοσης.

Χαρακτηριστικά και λειτουργίες αυτών των συστημάτων είναι:

- η δοκιμή προηγούμενων γνώσεων με στόχο τη συλλογή χαρακτηριστικών του χρήστη,
- η δυνατότητα να ελέγχει ο χρήστης την ταχύτητα μάθησης και να ακολουθεί τον δικό του ρυθμό,
- η παροχή συνεχούς ανατροφοδότησης σχετικά με την εργασία του χρήστη πχ αν οι απαντήσεις του είναι σωστές ή λανθασμένες και παροχή βοήθειας με τη μορφή προτάσεων επιπρόσθετων ασκήσεων,
- η παρακολούθηση της προόδου του χρήστη και ο τεμαχισμός της εργασίας σε μικρότερα κομμάτια,
- η ενίσχυση του κινήτρου με επιβράβευση μέσω πόντων.

Κύρια στοιχεία ενός συστήματος με δυνατότητα ευελιξίας, σύμφωνα με την ανασκόπηση στο [7], είναι:

1. Ένα μοντέλο περιεχομένου, αναφέρεται στον τρόπο δόμησης των περιεχομένων με λεπτομερή μαθησιακά αποτελέσματα και καθορισμένα καθήκοντα που πρέπει να

-
- διδαχθούν. Η σειρά θα μπορούσε να αλλάξει ανάλογα με την επίδοση του σπουδαστή.
2. Ένα μοντέλο μαθητή, αναφέρεται σε αξιολόγηση του επιπέδου του μαθητή σε διαφορετικές θεματικές, ιχνηλάτηση της επίδοσης και επιβεβαίωση γνώσεων για προηγούμενες γνώσεις.
 3. Ένα καθοδηγητικό μοντέλο, που αποφασίζει με ποιο τρόπο το σύστημα θα προτείνει περιεχόμενο σε ποιον χρήστη ως απόρροια του συνδυασμού των πληροφοριών των δύο προηγούμενων μοντέλων με στόχο την περαιτέρω πρόοδο του μαθητή.

Η προσαρμοστικότητα του συστήματος αντανακλάται στους τρόπους με του οποίους μπορεί ο χρήστης να καταλάβει επαρκώς τα μαθησιακά αντικείμενα. Επίσης, το πόσο καλά οι μέθοδοι επικοινωνίας και αλληλεπίδρασης ταιριάζουν στους χρήστες και τέλος, στο αν προτείνει περιεχόμενο που πραγματικά ικανοποιεί τους στόχους των φοιτητών [4].

Τέλος, οι τεχνολογίες που χρησιμοποιούνται σε προσαρμοστικά μαθησιακά περιβάλλοντα είναι: Σηματολογικός Ιστός, συλλογιστική βάσει κανόνων, δίκτυα Bayes, προσαρμοστικά υπερμέσα, αρχιτεκτονική προσανατολισμένη στις υπηρεσίες, συστήματα βασισμένα σε πράκτορες, μηχανική μάθηση και εξόρυξη δεδομένων, μοντελοποίηση χρηστών και ομάδων, εμπειρική μοντελοποίηση, μοντελοποίηση και πρόβλεψη περιβάλλοντος, διαχείριση ποιότητας υπηρεσιών και υποστήριξη αποφάσεων.

2.3 LMS-Learning Management System

Σύμφωνα με τους συγγραφείς στο [7], «Το LMS είναι ένα λογισμικό διακομιστή που παρέχει ένα σύνολο διεπαφών σε μια βάση δεδομένων που περιλαμβάνει πληροφορίες χρήστη, μαθήματα και περιεχόμενο. Αυτά τα συστήματα επιτρέπουν στα εκπαιδευτικά ιδρύματα να διαχειριστούν έναν τεράστιο αριθμό διαδικτυακών ή μικτών μαθημάτων με διασυνδέσεις και σύνολα πόρων που χρησιμοποιούνται συνήθως. Μαθήματα εκτός σύνδεσης που χρησιμοποιούν το LMS ως υποχρεωτικό ή προηγμένο εργαλείο είναι συχνά διευρυμένα διαδικτυακά μαθήματα.».

Κύριο στόχος ενός LMS είναι να βοηθά τους χρήστες να χρησιμοποιούν κατάλληλες υπηρεσίες και να επιτρέπει ανταλλαγή γνώσεων [4]. Το κύριο πλεονέκτημα της χρήσης LMS στη μαθησιακή διαδικασία είναι η δυνατότητα αποθήκευσης τεράστιου όγκου μέσων που αφορούν τη διάδοση πληροφοριών, την επικοινωνία και την εξέταση γνώσεων, συγκεντρωμένα σε μία πλατφόρμα δημόσια διαθέσιμη μέσω του Ιστού. Επιπλέον, απλοποιείται η πρόσβαση στα μαθήματα, χωρίς να ανησυχεί κανείς για την προστασία του υλικού καθώς η πρόσβαση γίνεται με τα ανάλογα διαπιστευτήρια χρηστών [7].

Κάθε μαθησιακή πηγή έχει δικά της χαρακτηριστικά στον τρόπο παρουσίασης των δεδομένων, στη δομή, στη μορφή του περιεχομένου κα. Η ανάπτυξη μίας πύλης σπουδαστικού περιεχομένου αντιμετωπίζει κυρίως το πρόβλημα στην αρχικοποίηση του προφίλ του χρήστη το οποίο θα πρέπει συνεχώς να ενημερώνεται ώστε να παρακολουθούνται οι αλλαγές στις προτιμήσεις, τα ενδιαφέροντα και τις ανάγκες του χρήστη [4].

Ένα άλλο μειονέκτημα που προκύπτει είναι πως ένας μεγάλος αριθμός εργαλείων του συστήματος αφορούν τη διαχείριση παρά τις ανάγκες των χρηστών. Η υλοποίηση της διαπαφής με το χρήστη δεν είναι πάντα εύκολη και τα μαθήματα αναπαριστούν αποθετήρια ποικίλου περιεχομένου που δεν ικανοποιούν το μαθησιακό στόχο [7].

Ένα θέμα που δεν αναφέρεται συχνά αλλά είναι πολύ σημαντικό είναι η ασφάλεια ενός ηλεκτρονικού περιβάλλοντος από εισβολείς. Οι διαχειριστές καλούνται να διατηρούν τη λειτουργία ενός συστήματος πραγματικού χρόνου αδιάκοπα αλλά δεν μπορούν να παρακολουθούν τη δραστηριότητα του συστήματος συνεχώς. Ένα αυτόνομο σύστημα ανίχνευσης εισβολέων θα μπορούσε να αποτελεί μία δεύτερη γραμμή άμυνας μαζί με άλλα εργαλεία όπως firewall και ανθρώπινη διαχείριση. Σε αυτόν τον τομέα οι τεχνικές εξόρυξης της πληροφορίας δίνουν λύσεις στην ανίχνευση εισβολής. Η πιο συχνά χρησιμοποιούμενη τεχνική είναι η ανίχνευση ανωμαλιών ως αποκλίσεις από τα κανονικά μοτίβα (στατικά ή δυναμικά) της αλληλεπίδρασης του συστήματος με τους χρήστες. Η άλλη προσέγγιση είναι ο εντοπισμός κακής χρήσης με τον εντοπισμό αδύναμων στοιχείων του συστήματος και τον καθορισμό πιθανών μεθόδων επίθεσης [6]. Επίσης, μπορεί να δημιουργηθεί ένα μοντέλο επιθέσεων από ίχνη που έχουν μείνει από επιθέσεις του παρελθόντος στο σύστημα και να αποτελεί ένα μέσο πρόληψης κακόβουλων ενεργειών. Είναι προφανές ότι ο εντοπισμός κακόβουλων ενεργειών και η προστασία από επιθέσεις που αλλοιώνουν τα κανονικά μοτίβα αλληλεπίδρασης των χρηστών, έμμεσα προστατεύουν και το σύστημα συστάσεων από το να προτείνει στους χρήστες περιεχόμενο που υπό την ομαλή λειτουργία του LMS θα ήταν έκτος της λίστας των συστάσεων.

2.4 MOOC Ecosystems

Ένας από τους τομείς που γνώρισε βελτίωση λόγω της εξέλιξης της τεχνολογίας είναι και η εκπαίδευση. Ανάμεσα στις υπάρχουσες επιλογές που αλλάζουν τον τρόπο μάθησης είναι και τα Μαζικά Ανοιχτά Διαδικτυακά Μαθήματα εν συντομία MOOC από το ακρωνύμιο των λέξεων Massive Open Online Courses. Εισάγει μία νέα εκπαιδευτική φιλοσοφία με βασικές αξίες την ευρύτητα, συμμετοχή και τη συνεργασία. Συνιστά κάτι περισσότερο από ανοιχτή εκπαίδευση [8]. Τα MOOC ορίζονται ως «διαδικτυακά μαθήματα προσβάσιμα σε οποιονδήποτε στον Ιστό», όπου «τα ιδρύματα έχουν ενταχθεί σε μια προσπάθεια να κάνουν

την εκπαίδευση πιο προσιτή σε συνεργασία με τους παρόχους MOOC». Η ιδέα του ελεύθερου λογισμικού αποτελεί τη βάση για ένα από τα σημαντικότερα εκπαιδευτικά κινήματα του 21^{ου} αι. την Ανοιχτή Διαδικτυακή Εκπαίδευση [8].

Τα συστήματα αυτά ξεκίνησαν το 2008 και έκτοτε δεν έπαψαν να εξελίσσονται. Θεωρούνται κάποια από τα πλέον επιδραστικά διαδικτυακά συστήματα καθώς επηρεάζουν και τα πιο παραδοσιακά ακαδημαϊκά ιδρύματα. Ένα MOOC μπορεί να είναι επωφελές, να ενισχύει το σχηματισμό συνεργασιών ή και συμμαχιών ανάμεσα σε εταιρείες οικοσυστημάτων. Οι πάροχοι MOOC γνωρίζουν τη δυσκολία που συναντούν οι χρήστες στην εύρεση του κατάλληλου περιεχομένου το οποίο μπορεί να προσφέρεται από πολλούς παρόχους ή ένα μέρος (module) κάποιου προγράμματος να ικανοποιεί τις ανάγκες του χρήστη [8].

Εκτός από την ανάπτυξη ενός βιώσιμου MOOC προκύπτουν και άλλα οφέλη από τα οικοσυστήματα λογισμικού. Τα οικοσυστήματα λογισμικού ορίζονται ως ένα σύνολο επιχειρήσεων που λειτουργούν ως μονάδα και αλληλεπιδρούν σε μία κοινή αγορά λογισμικού και υπηρεσιών. Ενώ οι μεταξύ τους σχέσεις συχνά υποστηρίζονται από μία κοινή τεχνολογική πλατφόρμα ανταλλαγής πληροφοριών, πόρων και έργων. Κατανοεί κανείς πως προωθούν την εξέλιξη του λογισμικού, και την καινοτομία σε οργανισμούς ενισχύοντας την επιτυχία αυτών των πλατφορμών. Βοηθούν την ανάλυση και κατανόηση της αρχιτεκτονικής λογισμικού ώστε να αποφασισθεί ποια πλατφόρμα είναι κατάλληλη να χρησιμοποιηθεί, πραγματοποιείται διαμοιρασμός γνώσης μέσω πολλαπλών και ανεξάρτητων οντοτήτων που εντείνει τη συνεργασία και τέλος υποστηρίζουν την αναγνώριση επιχειρηματικών καθηκόντων, παράγουν αρχιτεκτονικό σχεδιασμό και αναγνωρίζουν ρίσκο [8].

Για να εξελιχθεί μία μαθησιακή κοινότητα MOOC χρειάζεται να γίνουν αντιληπτοί οι ρόλοι και οι σχέσεις μεταξύ τους στο οικοσύστημα λογισμικού. Οι βασικοί ρόλοι είναι φοιτητές, καθηγητές, ανώτερα εκπαιδευτικά ιδρύματα (αν και δεν υπάρχει κοινή ονομασία οπότε σε άλλο περιβάλλον μπορεί να ονομάζονται πχ σπουδαστές, διδάσκοντες και Πανεπιστήμια) ενώ οι σχεδιαστές του συστήματος μπορεί να αναφέρονται ως «αυτοί που φτιάχνουν MOOC». Η συνένωση κάποιων ομάδων είναι δυνατή. Οι ομάδες αλληλεπιδρούν μεταξύ τους διαμέσου των πλατφορμών προσφέροντας κάτι σε αντάλλαγμα [8].

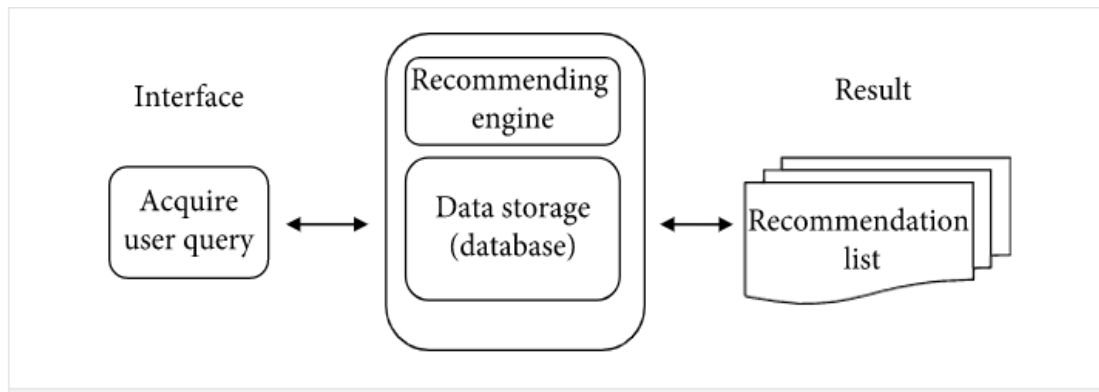
3

Συστήματα Συστάσεων(ΣΣ)

3.1 Ορισμός ΣΣ

«Συστήματα Συστάσεων αποκαλούνται όλα εκείνα τα συστήματα τα οποία έχουν ως βασικό καθήκον την επιλογή αντικειμένων που ικανοποιούν τις απαιτήσεις του χρήστη» πρόκειται για έναν σύντομο ορισμό που αφορά κάθε είδους πληροφορίες [9]. Ως εκ τούτου είναι συστήματα που χρησιμοποιούνται σε διάφορους τομείς όπως ηλεκτρονικό εμπόριο, ψυχαγωγία, ψηφιακές βιβλιοθήκες και έρχονται να επιλύσουν το πρόβλημα της διαχείρισης τεράστιου όγκου πληροφορίας που παράγεται καθημερινά. Παρόμοια προβλήματα αντιμετωπίζει και ο χώρος της εξ αποστάσεως εκπαίδευσης, όπου ο χρήστης καλείται να επιλέξει μαθησιακά αντικείμενα μέσα από ένα εύρος επιλογών [9].

Ένα ΣΣ πρέπει να είναι ικανό να παρουσιάσει προτάσεις για ένα θέμα με κριτήρια όπως οι προηγούμενες προτιμήσεις του ίδιου χρήστη ή άλλων χρηστών με παρόμοιες απαιτήσεις και να βοηθά στη μείωση του όγκου πληροφοριών παρέχοντας παράλληλα προσαρμοσμένη πρόσβαση σε πληροφορίες ενός συγκεκριμένου τομέα [9]. Ο κύριος σκοπός ενός συστήματος συστάσεων είναι να παρέχει προτάσεις, πραγματικά βασισμένες στις προτιμήσεις και τις επιλογές του χρήστη. Έτσι η ποιότητα των συστάσεων, ιδίως όταν ο όγκος των πληροφοριών είναι πολύ μεγάλος, είναι καθοριστικής σημασίας. Για αυτό χτίζονται διαφορετικά ΣΣ με διάφορες μεθόδους και μια πληθώρα από αλγόριθμους ανάλογα τις συστάσεις που χρειάζεται να δημιουργηθούν [15]. Ένα παράδειγμα γενικής αρχιτεκτονικής συστημάτων συστάσεων φαίνεται στην παρακάτω εικόνα:



Εικόνα 1. Γενική αρχιτεκτονική Συστήματος Συστάσεων.

Σύμφωνα με το παραπάνω βασικό σχήμα, σε κάθε ΣΣ στην διεπαφή επικοινωνίας με τον χρήστη, ζητείται το ερώτημα του χρήστη το οποίο πηγαίνει στη μηχανή παραγωγής συστάσεων και σε μία Βάση Δεδομένων και ως αποτέλεσμα παράγεται μία λίστα προτάσεων.

Τα εκπαιδευτικά ΣΣ διαφοροποιούνται σε σχέση με τα ΣΣ άλλων τομέων. Ένας άλλος ορισμός που εστιάζει στο τι είναι τα συστήματα συστάσεων ιδιαίτερα για την ηλεκτρονική μάθηση είναι ο ορισμός που δίνεται στο [16] και αναφέρει πώς ένα σύστημα συστάσεων για την ηλεκτρονική μάθηση είναι διαφορετικό από άλλους τομείς διότι το ΣΣ πρέπει να προτείνει αντικείμενα χωρίς να επηρεάζει την ακολουθία μάθησης και τα αντικείμενα – μαθήματα που προτείνει θα πρέπει να παραμένουν στο τρέχον πλαίσιο μάθησης.

3.2 Αναγκαιότητα Εκπαιδευτικών Συστημάτων Συστάσεων

Στη σύγχρονη εποχή η χρήση συστημάτων διαδικτυακής μάθησης γνωρίζει αξιοσημείωτη ανάπτυξη. Το γεγονός αυτό εγείρει ορισμένες προκλήσεις για τους χρήστες, καθώς η εξαγωγή χρήσιμων και σχετικών πληροφοριών, που να ικανοποιούν τις ανάγκες του χρήστη δεν είναι πάντα εύκολη υπόθεση, λόγω της πληθώρας πληροφοριών που συναντά κανείς στο Διαδίκτυο [9]. Ένα απλό αλλά χαρακτηριστικό παράδειγμα είναι η ανάγκη για καθοδήγηση που έχουν οι φοιτητές είτε από τους καθηγητές τους, είτε από άλλους συμφοιτητές ως προς την επιλογή καταλλήλων για αυτούς μαθημάτων [9]. Τα τρέχοντα συστήματα διαχείρισης ηλεκτρονικής μάθησης περιέχουν μία τεράστια δεξαμενή από δεδομένα που συλλέγονται από διάφορες πηγές όπως βιντεοδιαλέξεις, ιστολόγια, άρθρα κτλ. Ως εκ τούτου, αποτελεί πρόκληση για κάθε τέτοιο σύστημα η παροχή σχετικού ποιοτικού και έγκυρου περιεχόμενου καθώς και η μείωση του χρόνου που απαιτείται από το χρήστη για τη συλλογή του υλικού αυτού [4].

Οι περισσότεροι φοιτητές διαλέγουν μαθήματα «στα τυφλά», με αποτέλεσμα κάποια μαθήματα να μην επιλέγονται λόγω της σειράς μαθημάτων ή λόγω ελλιπούς κατανόησης από τους φοιτητές σχετικά με τη χρησιμότητα τους [10]. Ένας άλλος λόγος που εντείνει την

ανάγκη για παροχή προσωποποιημένων συστάσεων σε ευέλικτα μαθησιακά περιβάλλοντα είναι η μη δυνατότητα των διδασκάλων να δουλέψουν πάνω στις αδυναμίες του κάθε σπουδαστή ξεχωριστά [11].

Για την ικανοποίηση αυτών των αναγκών έχουν αναπτυχθεί διάφορες στρατηγικές συστημάτων συστάσεων που λειτουργούν κατά αναλογία με τις συστάσεις που θα έπαιρνε κάποιος από κάποιο φίλο ή κάποιον με παρόμοια ενδιαφέροντα [9]. Χρειάζεται, λοιπόν, να ενσωματωθούν μαθησιακά χαρακτηριστικά των σπουδαστών στη διαδικασία παραγωγής συστάσεων [12] και να αναλυθούν και να οπτικοποιηθούν οι επιδόσεις των σπουδαστών με σκοπό την στοχευμένη εκπαιδευτική σύσταση [11]. Εξάλλου οι διαφορές στη μαθησιακή ικανότητα του κάθε σπουδαστή, οδηγούν σε διαφορετικά μαθησιακά μονοπάτια κατανόησης ενός συγκεκριμένου περιεχομένου [4].

Κάθε σύστημα ηλεκτρονικής μάθησης έχει ως κύριο σκοπό την παροχή βοήθειας στο σπουδαστή, κάθε φορά που είναι αναγκαίο, ώστε να επιτύχει τους μαθησιακούς του στόχους. Μαθησιακός στόχος μπορεί να είναι η επιθυμία ολοκλήρωσης ενός μαθήματος, η ανάπτυξη κάποιας δεξιότητας, ή ακόμα η συλλογή γνώσεων και δεξιοτήτων με σκοπό την εξειδίκευση πάνω σε ένα αντικείμενο [11].

Βασικό πλεονέκτημα των συστημάτων ηλεκτρονικής μάθησης είναι η δυνατότητα που δίνουν στο χρήστη να μαθαίνει οπουδήποτε, οποτεδήποτε και μάλιστα καλούνται να αντιμετωπίσουν το πρόβλημα της αντιμετώπισης όλων των σπουδαστών με τον ίδιο τρόπο, ενώ οι ανάγκες του κάθε σπουδαστή είναι διαφορετικές [7]. Επιπλέον, ιδανικά, σκοπός ενός τέτοιου συστήματος είναι η βελτιστοποίηση του μαθησιακού αποτελέσματος στον μικρότερο δυνατό χρόνο καθώς παρέχει τις απαραίτητες πληροφορίες στον χρήστη και ανατροφοδότηση ως προς την μαθησιακή διαδικασία [13].

Τα συστήματα συστάσεων για πλατφόρμες ηλεκτρονικής μάθησης προτείνουν μία εργασία στον εκπαιδευόμενο βάσει άλλων προηγούμενων επιτυχώς εκτελεσμένων καθηκόντων του ιδίου [12] ή άλλων σπουδαστών με παρόμοια χαρακτηριστικά. Ένα σύστημα που προσαρμόζεται στις ανάγκες του σπουδαστή αναμένεται να προτείνει εκπαιδευτικές ιδέες με βάση τις επιδόσεις, τις γνώσεις, τις δεξιότητες και τα μαθησιακά μοτίβα και τρόπους του φοιτητή [11]. Με αυτόν τον τρόπο εγκαθίσταται μία σχέση ανάμεσα στο χρήστη και το μαθησιακό περιβάλλον η οποία μπορεί να είναι μακροπρόθεσμη και σταθερή. Γεγονός που σημαίνει για το περιβάλλον μάθησης διατήρηση των χρηστών και μικρότερος ρυθμός διαρροής. Για τον χρήστη σημαίνει μικρότερο χρόνο αναζήτησης, παροχή καλύτερων υπηρεσιών και χρησιμότερων πληροφοριών [14].

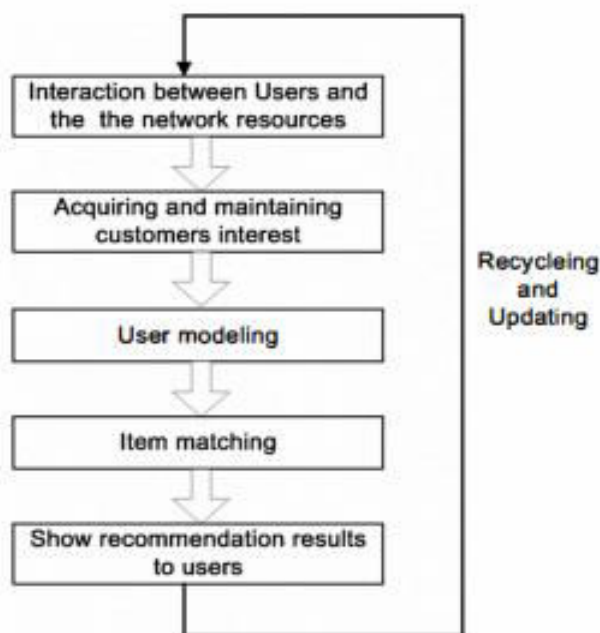
Στον αντίποδα, ένα σύστημα ηλεκτρονικής μάθησης χωρίς προσωποποιημένες συστάσεις μπορεί να εφαρμοστεί, αλλά θα προτείνει μαθήματα σε σπουδαστές με βάση σχόλια και απόψεις των σπουδαστών κατά μέσο όρο για ένα μάθημα. Αυτού του είδους οι συστάσεις

είναι ανεξάρτητες από τον χρήστη με αποτέλεσμα όλοι να παίρνουν τις ίδιες συστάσεις. Είναι αυτόματα συστήματα, είναι εύκολο να παράξουν συστάσεις, αλλά αυτές είναι εφήμερες, καθώς το σύστημα δεν αναγνωρίζει τον σπουδαστή από μία συνεδρία σε επόμενη [14].

3.3 Διάγραμμα ροής Προσωποποιημένου Συστήματος Εκπαιδευτικών Συστάσεων

Τα στάδια τα οποία είναι αναγκαία για την επίτευξη ενός ΣΣ σε ένα εκπαιδευτικό περιβάλλον σύμφωνα με τους [17] είναι:

1. Αλληλεπίδραση του χρήστη με τις πηγές δικτύου.
2. Συλλογή πληροφοριών που αφορούν τα ενδιαφέροντα του χρήστη.
3. Μοντελοποίηση του χρήστη.
4. Ταίριασμα με μαθησιακό αντικείμενο.
5. Προτάσεις στον χρήστη με βάση τα παραπάνω αποτελέσματα.



Εικόνα 2. Διάγραμμα ροής προσωποποιημένου ΣΣ

3.3.1 Συλλογή πληροφοριών που αφορούν τον χρήστη

Η συλλογή των πληροφοριών του χρήστη είναι η πρώτη και πιο βασική εργασία για την επιτυχία ενός προσωποποιημένου ΣΣ αφού ένα τέτοιο σύστημα απευθύνεται προσωπικά στον

κάθε χρήστη για να «εφαρμόζει στα μέτρα του χρήστη» [17]. Δύο είναι οι βασικές μέθοδοι συλλογής: άμεσα και έμμεσα. Κατά την εγγραφή του στο σύστημα ο χρήστης δίνει συγκεκριμένες πληροφορίες και απαντώντας κάποιο ερωτηματολόγιο γίνεται ανατροφοδότηση των πληροφοριών. Ενώ έμμεσα αποκτώνται πληροφορίες κατά την περιήγηση του χρήστη, όπου επιδεικνύει κάποιες συμπεριφορές, όπως μέσα από το ιστορικό, τη συχνότητα επισκέψεων συγκεκριμένης σελίδας, το χρόνο που αφιερώνει σε αυτή κοκ. Επίσης, ο τρόπος με τον οποίο αναζητά κάποια πληροφορία, ποιες λέξεις κλειδιά χρησιμοποιεί, η ανάλυση των αρχείων εγγραφών σε εξυπηρετητές, μπορεί να δώσει χρήσιμες πληροφορίες σε ποιες ιστοσελίδες έχει πρόσβαση ο χρήστης. Τέλος, πληροφορίες αξιολόγησης υποδεικνύουν το βαθμό προσήλωσης, καθώς και πληροφορίες και έγγραφα τα οποία ο χρήστης λαμβάνει και αποθηκεύει [17].

Στον άμεσο τρόπο εξαγωγής πληροφοριών για το χρήστη, γνωστός και ως στατικός, τα στοιχεία που συλλέγονται είναι σαφή και προβλέψιμα αφού ο ίδιος ο χρήστης καλείται σε πολλές περιπτώσεις να συμπληρώσει ερωτηματολόγια ή διαδικτυακές φόρμες και να δώσει απαντήσεις γύρω από τα ενδιαφέροντά του [4].

Ο έμμεσος τρόπος είναι δυναμικός και οι πληροφορίες για τον χρήστη συλλέγονται κατόπιν παρακολούθησης και παρατήρησης των αλληλεπιδράσεων του χρήστη με το σύστημα. Στην προκειμένη περίπτωση πραγματοποιείται ανάλυση συμπεριφοράς και παρακολουθείται η επίσκεψη σε συνδέσμους, ο βαθμός αρεσκείας σε κάποιο περιεχόμενο με «likes» του χρήστη, ακόμα και το υλικό το οποίο κάνει λήψη μπορεί να ανιχνευθεί για να ενισχύσει τις συστάσεις που δίνει το σύστημα [4].

Το υβριδικό μοντέλο δημιουργίας προφίλ συνδυάζει τα πλεονεκτήματα των προηγούμενων τύπων προφίλ, κρατάει αρχείο των στατικών πληροφοριών του χρηστή και των αλληλεπιδράσεών του και ενημερώνει τυχόν αλλαγές στο προφίλ δυναμικά [4].

3.3.2 Μοντελοποίηση του χρήστη

Πρόκειται για τη διαδικασία κατά την οποία τα ενδιαφέροντα του χρήστη εκφράζονται με δομημένα δεδομένα και αποκτούν μία μαθηματική μορφή ικανή για επεξεργασία και εξαγωγή συγκεκριμένων και ποσοτικοποιημένων συμπερασμάτων. Κάποιες από τις μεθόδους είναι: αναπαράσταση θέματος, αναπαράσταση λίστας με λέξεις κλειδιά, αναπαράσταση νευρωνικών δικτύων και αναπαράσταση διανύσματος. Η δημιουργία διανύσματος είναι η πιο συχνά χρησιμοποιούμενη μέθοδος και το μοντέλο χρήστη αναπαριστάται σαν ένα διάνυσμα n διαστάσεων από δυαδικές ομάδες λέξεων κλειδιών και των βαρών τους, όπου το βάρος αντιπροσωπεύει το βαθμό ενδιαφέροντος του χρήστη για τη συγκεκριμένη λέξη [17].

Το διάνυσμα των προτιμήσεων θα μπορούσε να αποδοθεί ως εξής:

$$\text{Preferences} = \{f_1, f_2, \dots, f_n\} = \{(k_1, w_1, t_1), (k_2, w_2, t_2), \dots, (k_n, w_n, t_n)\}$$

Σε αυτή τη συνάρτηση εκτός από τη λέξη κλειδί και το αντίστοιχο βάρος εκφράζεται και ο τελευταίος – πιο πρόσφατος χρόνος ενημέρωσης των βαρών.

3.3.3 Φιλτράρισμα ή ταίριασμα χρήστη με αντικείμενο

Η ανάλυση των δεδομένων με διάφορους αλγορίθμους όπως px συνεργατικό φιλτράρισμα θα οδηγήσει σε συμπεράσματα ως προς την ομοιότητα ανάμεσα σε χρήστες και αντίστοιχα αντικείμενα προτιμητέα έναντι άλλων. Προκύπτει ένας πίνακας $m*n$ αξιολόγησης αντικειμένων όπου m οι χρήστες, n τα αντικείμενα και το στοιχείο r_{ij} είναι κάθε αντικείμενο j για το οποίο ο χρήστης i έχει πραγματοποιήσει αξιολόγηση. Πόσοι χρήστες, δηλαδή, ενδιαφέρονται για συγκεκριμένη πηγή [17].

3.3.4 Παραγωγή Συστάσεων

Απόρροια αυτού του πίνακα θα είναι η παραγωγή συστάσεων κατόπιν εφαρμογής συναρτήσεων πρόβλεψης και ομοιότητας. Η διαδικασία παραγωγής συστάσεων είναι δυναμική, κυκλική και επαναλαμβανόμενη. Οι προτιμήσεις και τα ενδιαφέροντα των χρηστών αλλάζουν στο πέρασμα του χρόνου, ως εκ τούτου η τακτική ενημέρωση του συστήματος με τις πιο πρόσφατες αλλαγές συνιστά το κλειδί της επιτυχίας του ΣΣ [17].

3.4 Βασικός Διαχωρισμός των ΣΣ

Ένα ΣΣ σε ένα μαθησιακό περιβάλλον επιχειρεί να αναθέσει καθήκοντα στον σπουδαστή βάσει προηγούμενων εργασιών του σπουδαστή που ανέλαβε αποτελεσματικά και με επιτυχία. Ένα σύστημα μπορεί να παράγει συστάσεις βάσει αναζητήσεων του χρήστη ή αυτομάτως. Επίσης, οι προτιμήσεις μπορεί να καθορίζονται απευθείας από το χρήστη ή με σύγκριση χρηστών με παραπλήσια χαρακτηριστικά [12]. Τελικά, το σύστημα προτείνει με βάση τις προτιμήσεις των χρηστών, το ιστορικό αναζητήσεων χρηστών με παρόμοια χαρακτηριστικά και ανάλογα ενδιαφέροντα. Η ομοιότητα ανάμεσα σε χρήστες καθορίζεται από κοινά μοτίβα αναζητήσεων ή με χρήση προφίλ χρηστών [12].

Υπάρχουν διάφορες προσεγγίσεις (για την ανάπτυξη των) στην πραγματοποίηση ΣΣ. Οι πιο βασικές είναι οι παρακάτω:

- Συστάσεις βάσει Περιεχομένου (Content Based)
- Συνεργατικό Φιλτράρισμα (Collaborative Filtering)
- Υβριδικά Συστήματα (Hybrid Filtering)

- Δημογραφικές Συστάσεις (Demographic)
- Συστάσεις βάσει Γνώσεων (Knowledge Based)
- Συστήματα Συστάσεων με Επίγνωση του περιβάλλοντος (Context-aware)
- Συστήματα Συστάσεων βάσει Οντολογιών (Ontology)
- Συστήματα Συστάσεων με χρήση τεχνικών εξόρυξης πληροφορίας (Data Mining)

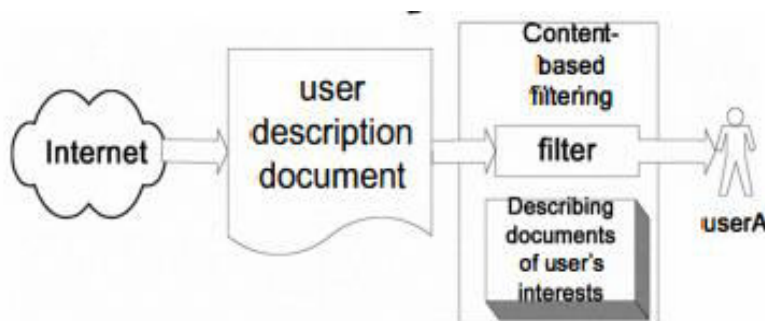
3.4.1 Φιλτράρισμα βάσει περιεχομένου

«Σύστημα στο οποίο οι συστάσεις γίνονται βασισμένες μόνο στο προφίλ του χρήστη, το οποίο δημιουργείται όταν λαμβάνεται υπόψη η ανάλυση περιεχομένου του αντικειμένου το οποίο έχει εκτιμήσει στο παρελθόν ο χρήστης» ορισμός σύμφωνα με τους [9].

Η παραγωγή συστάσεων βασίζεται στο περιεχόμενο των εγγράφων. Η κύρια ιδέα είναι η εξαγωγή του περιεχομένου με ανάκτηση πληροφοριών (Information Retrieval-IR). Σύμφωνα με τα χαρακτηριστικά του περιεχομένου και τις προτιμήσεις του χρήστη, αναγνωρίζεται η απαίτηση του χρήστη και γίνονται προτάσεις με το κατάλληλο περιεχόμενο [18].

Η τεχνική αυτή στηρίζεται στη σύγκριση ανάμεσα στην περιγραφή μαθησιακών αντικειμένων και στα ενδιαφέροντα του χρήστη [17]. Προϋποθέτει ότι ο χρήστης συμπεριφέρεται το ίδιο σε παρόμοιες καταστάσεις [4] και κατά πάσα πιθανότητα αντικείμενα για τα οποία έδειξε ενδιαφέρον στο παρελθόν θα δείξει και στο προσεχές μέλλον [9]. Η διαλογή γίνεται χρησιμοποιώντας προηγούμενες εγγραφές του χρήστη πρόσβασης σε κάποιο περιεχόμενο και συσχέτιση αντικειμένων στη Βάση Δεδομένων [4]. Μόλις το σύστημα δεχθεί αξιολογήσεις από τον χρήστη για περιεχόμενο το οποίο προσπέλασε, το εντάσσει στο νέο περιεχόμενο και ένας δείκτης νέας αξίας για περιεχόμενο που δεν έχει ειπωθεί γεννάται [4]. Οι πληροφορίες περιεχομένου μπορούν να γεφυρώσουν το χάσμα ανάμεσα στον ήδη υπάρχοντα και στον νέο σπουδαστή καθώς και ανάμεσα στα μαθησιακά αντικείμενα [12].

Σχηματική αναπαράσταση των συστάσεων βάσει περιεχομένου φαίνεται στην παρακάτω εικόνα [17]:



Εικόνα 3. Σχηματικό διάγραμμα συστάσεων βάσει περιεχομένου

Δύο είδη συστάσεων βάσει περιεχομένου συναντώνται [12]:

- οι τεχνικές που στηρίζονται σε προηγούμενες περιπτώσεις και
- οι τεχνικές που στηρίζονται σε χαρακτηριστικά

Στην πρώτη περίπτωση συστήνονται τα μαθησιακά αντικείμενα που σχετίζονται με αυτά που προτίμησε ο χρήστης στο παρελθόν. Η τεχνική αυτή δεν επιθυμεί ανάλυση περιεχομένου, αλλά οδηγεί σε υπέρ-εξειδίκευση καθώς προτείνει μόνο ότι έχει τη μεγαλύτερη συσχέτιση [12]. Στο χρήστη προτείνεται πρώτα το περιεχόμενο με τη μεγαλύτερη αξία/εγγύτητα και ακολουθείται φθίνουσα σειρά [4].

Στη δεύτερη περίπτωση γίνεται χαρτογράφηση χαρακτηριστικών με το προφίλ του σπουδαστή. Νέοι χρήστες μπορούν να προστεθούν, καθώς και νέα χαρακτηριστικά, αλλά είναι στατικό μοντέλο που δεν μπορεί να κατανοήσει από τη συμπεριφορά του δικτύου. Απευθείας χαρτογραφεί χαρακτηριστικά σπουδαστών σε μαθησιακά χαρακτηριστικά και έτσι δε χρειάζονται δεδομένα συμπεριφοράς του χρήστη [12].

3.4.2 Συνεργατικό φιλτράρισμα (Collaborative Filtering-CF)

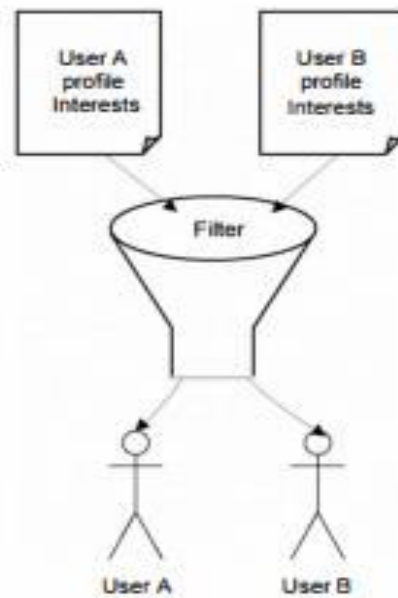
«Σύστημα στο οποίο οι συστάσεις προκύπτουν με βάση το βαθμό ομοιότητας ανάμεσα σε χρήστες» [9]. Πρόκειται για την πιο διαδεδομένη τεχνική παραγωγής συστάσεων, δύο είναι τα βασικά ήδη, συστάσεις με βάση τα αντικείμενα όπου συγκρίνονται οι ομοιότητες των αντικειμένων ή με βάση τους χρήστες, όπου γίνεται σύγκριση ανάμεσα στους χρήστες και παρατηρούνται πιθανές ομοιότητες ανάμεσά τους [17].

Τα συστήματα συσχέτισης μαθητευόμενων προτείνουν μαθήματα σε έναν εκπαιδευόμενο με βάση τη συσχέτιση μεταξύ αυτού του μαθητή και άλλων μαθητών που έχουν παρακολουθήσει μαθήματα από την ίδια σελίδα ηλεκτρονικής μάθησης. Αυτή η τεχνολογία είναι γνωστή ως «συνεργατικό φιλτράρισμα», επειδή ξεκίνησε ως μια τεχνική φιλτραρίσματος πληροφοριών που χρησιμοποιούσε απόψεις ομάδων χρηστών για να προτείνει στοιχεία πληροφοριών σε άτομα [14].

Αυτή η προσέγγιση «χτίζει» το μοντέλο με παραμέτρους όπως: θεάσεις του χρήστη, ιστορικό χρήσης κα. Η τεχνική στηρίζεται στην κατηγοριοποίηση των χρηστών με παρόμοια ενδιαφέροντα [4]. Η βασική αρχή πίσω από αυτή την τεχνική είναι ότι «χρήστες με παρόμοια χαρακτηριστικά πολύ πιθανόν να έχουν και παρόμοιες προτιμήσεις» [12]. Χρήστες που δήλωσαν ότι τους άρεσε κάποιο περιεχόμενο στο παρελθόν, είναι πιο πιθανό να ενδιαφερθούν για παρόμοια αντικείμενα στο μέλλον και χρήστες που ανήκουν στην ίδια ομάδα ενδιαφερόντων είναι πιο πιθανό να συμπεριφερθούν ανάλογα [4].

Αυτή η τεχνική συλλέγει αξιολογήσεις χρηστών για αντικείμενα σε έναν συγκεκριμένο τομέα και αναγνωρίζει κοινά στοιχεία ανάμεσα σε σπουδαστές με σκοπό να παρέχει σχετικές συστάσεις [12][17]. Το προφίλ του σπουδαστή αποτελείται από ένα διάγραμμα μαθησιακών αντικειμένων και των αξιολογήσεών τους. Οι αξιολογήσεις υποδεικνύουν το βαθμό αρεσκείας, ο οποίος μπορεί να είναι δυαδικός ή πραγματικός [12]. Το περιεχόμενο που θα προταθεί στηρίζεται στη σχέση ανάμεσα στους χρήστες. Αν ισχύει η σχέση ανάμεσα σε δυο χρήστες τότε το περιεχόμενο που ενδιαφέρει τον Β χρήστη, προτείνεται και στον Α χρήστη [17].

Ο αλγόριθμος στηρίζεται στον k-πλησιέστερο γείτονα που μπορεί να βρει και να προβλέψει το σημείο ενδιαφέροντος του χρήστη που είναι προς διερεύνηση [17]. Σχηματική αναπαράσταση της μεθόδου αυτής φαίνεται στην παρακάτω εικόνα:



Εικόνα 4. Σχηματικό διάγραμμα συστάσεων συνεργατικού φιλτραρίσματος

Δύο είδη συνεργατικού φιλτραρίσματος συναντώνται:

- βάσει Μνήμης,
- βάσει Μοντέλου.

Βάσει Μνήμης υπάρχει η κατηγοριοποίηση βάσει χρήστη ή βάσει αντικειμένου [12]. Στην πρώτη περίπτωση ο χρήστης ανήκει σε μία ομάδα χρηστών με παρόμοιες συμπεριφορές και βρίσκει ένα σύνολο σπουδαστών με παρόμοιες προτιμήσεις. Παράγει μία λίστα συστάσεων για τον τελικό ενδιαφερόμενο. Στην άλλη περίπτωση, το μοντέλο βάσει αντικειμένου αναγνωρίζει ένα σύνολο μαθησιακών αντικειμένων και βρίσκει τα πιο σχετικά αντικείμενα με τα αντικείμενα που έχει αξιολογήσει ο χρήστης [12][15]. Δύο είναι τα βασικά βήματα αυτού

του αλγορίθμου. Αρχικά, σύμφωνα με των πίνακα αξιολόγησης των μαθημάτων του χρήστη, υπολογίζονται τα σύνολα με παρόμοια μαθήματα. Χρησιμοποιώντας τους πίνακες σκορ όλων των μαθημάτων των χρηστών υπολογίζεται η ομοιότητα μαθημάτων ανάμεσα σε όλα τα μαθήματα στον πίνακα. Στην επόμενη φάση υπολογίζει προβλεπόμενα σκορ χρηστών-στόχων για κάθε μάθημα και όταν το προβλεπόμενο σκορ υπερβαίνει ένα συγκεκριμένο κατώφλι εισάγεται στη λίστα συστάσεων με την προϋπόθεση ότι εξάγει από τη λίστα αυτά με τα οποία έχει έρθει ήδη σε επαφή ο χρήστης [10].

Βελτίωση αυτού του αλγορίθμου αποτελεί η έρευνα των [10] όπου προστίθεται στον αλγόριθμο και μία τεχνική ανάκτησης πληροφοριών TF-IDF (term frequency and inverse document frequency). Τα δεδομένα σε ένα σύστημα συστάσεων συνήθως αναπαριστώνται με κείμενο. Η περιγραφή ενός μαθήματος είναι αποθηκευμένη σε μία εκπαιδευτική βάση δεδομένων Πανεπιστημίου. Ο αλγόριθμος TF-IDF χρησιμοποιείται για να εξάγει από το κείμενο τα χαρακτηριστικά ενός μαθήματος και ένα διανυσματικό μοντέλο χρησιμοποιείται για την αναπαράσταση αυτών των χαρακτηριστικών. Πρόκειται για στατιστική μέθοδο που δείχνει τη βαρύτητα μιας λέξης μέσα σε ένα κείμενο. Προκειμένου να βρεθεί η ομοιότητα ανάμεσα στα μαθήματα χρησιμοποιήθηκε η μέθοδος αυτή όπου κάθε λέξη της περιγραφής ενός μαθήματος αποκτά ένα βάρος και το κείμενο αναπαριστάται ως ένα διάνυσμα με αποτέλεσμα ένα πρόβλημα υπολογισμού ομοιότητας κειμένων μετατρέπεται σε διάνυσμα ομοιότητας [10].

Το μοντέλο [10] στηρίζεται σε 2 παραμέτρους, τη συχνότητα με την οποία εμφανίζεται μία λέξη στο κείμενο TF (Term Frequency) με τον τύπο: $TF_{i,d} = N_{t_i} / \sum_k N_{t_k}$ όπου N είναι ο αριθμός στιγμιτύπων της λέξης t_i στο κείμενο d και ο παρονομαστής είναι ο συνολικός αριθμός των στιγμιτύπων όλων των λέξεων σε αυτό το κείμενο. Ακόμα, στηρίζεται στη συχνότητα που μία λέξη εμφανίζεται σε άλλα κείμενα.

Στην τεχνική βάσει μοντέλου δημιουργούνται συστάσεις υπολογίζοντας στατιστικά μοντέλα για αξιολογήσεις χρήστη. Μία πιθανολογική μέθοδος μπορεί να χρησιμοποιηθεί για τον υπολογισμό της πιθανότητας ότι ο χρήστης θα δώσει συγκεκριμένη αξιολόγηση σε ένα νέο μαθησιακό αντικείμενο βάσει προηγούμενων αξιολογημένων αντικειμένων [12].

«Ο υπολογισμός της ομοιότητας είναι ο πυρήνας του συνεργατικού αλγορίθμου φιλτραρίσματος». Ο συνεργατικός αλγόριθμος συστάσεων φιλτραρίσματος έχει πολλά πλεονεκτήματα, όπως υψηλή ταχύτητα, υψηλή απόδοση, σταθερότητα, και αποτελέσματα συστάσεων πιο ακριβή από άλλους αλγόριθμους [19].

Το μεγαλύτερο πρόβλημα που αντιμετωπίζει αυτή η μέθοδος είναι το πρόβλημα της ψυχρής εκκίνησης. Το ερώτημα που προκύπτει είναι: τι προτείνει στους σπουδαστές που μπαίνουν στο σύστημα για πρώτη φορά και δεν υπάρχουν προηγούμενες αξιολογήσεις ή προτιμήσεις; Επίσης, άλλο πρόβλημα είναι η μεγάλη διασπορά των δεδομένων που αυξάνεται συν τω

χρόνω καθώς αυξάνεται ο αριθμός χρηστών ενός συστήματος οι οποίοι όμως δεν αξιολογούν πάντα το περιβάλλον.

3.4.3 Υβριδικά Συστήματα

Ο συνδυασμός των δυο βασικών μεθόδων δημιουργίας συστάσεων που είναι η τεχνική βάσει περιεχομένου και το συνεργατικό φιλτράρισμα δημιουργεί μία νέα υβριδική τεχνική που προσπαθεί να άρει τα μειονεκτήματα των δυο πρώτων τεχνικών και να οδηγήσει σε ένα βελτιωμένο σύστημα συστάσεων με καλύτερα αποτελέσματα και λιγότερα προβλήματα [12][4]. Η μέθοδος διασφαλίζει τη διαθεσιμότητα προφίλ για κάθε χρήστη καθώς την πρώτη φορά που μπαίνει στο σύστημα δίνει πληροφορίες. Μετά την πρώτη είσοδο το σύστημα ενημερώνεται δυναμικά καθώς παρατηρούνται οι αλληλεπιδράσεις του χρήστη με το σύστημα και αυτός είναι ένας έμμεσος τρόπος εξαγωγής του προφίλ [4].

Μπορούν να χρησιμοποιηθούν και άλλες τεχνικές και να παραχθούν διαφορετικές έξοδοι όπως μεικτές, με βάρη, επαύξηση χαρακτηριστικών ή συνδυασμός αυτών κλπ διευθετώντας κάθε τύπο περιεχομένου και εξερευνώντας νέα δεδομένα με στόχο την εύρεση των ενδιαφερόντων του χρήστη [12].

Το υβριδικό φιλτράρισμα αυξάνει την ορθότητα των συστάσεων και αυξάνει την προσαρμοστικότητα του συστήματος [4].

3.4.4 Δημογραφικές Συστάσεις

Στην τεχνική αυτή η κατηγοριοποίηση γίνεται βάσει δημογραφικών στοιχείων των χρηστών (φύλο, ηλικία κλπ), προϋποθέτει δε, ότι χρήστες που ανήκουν στην ίδια δημογραφική κατηγορία είναι πολύ πιθανό να έχουν ίδια ενδιαφέροντα και προτιμήσεις [12]. Το πλεονέκτημα αυτής της μεθόδου είναι ότι είναι ανεξάρτητη από το ιστορικό αξιολόγησης του χρήστη, λειτουργεί με ανάλογο τρόπο με το συνεργατικό φιλτράρισμα, δημιουργώντας συσχετίσεις μεταξύ χρηστών, αλλά χρησιμοποιεί άλλα δεδομένα [12].

3.4.5 Συστάσεις βάσει Γνώσεων

Η μέθοδος αυτή προτείνει αντικείμενα βάσει των αναγκών του σπουδαστή. «Περιλαμβάνει γνώση του πως ένα συγκεκριμένο μαθησιακό αντικείμενο ικανοποιεί μία συγκεκριμένη μαθησιακή ανάγκη του σπουδαστή» [12]. Δεν σχετίζεται με αξιολογήσεις του χρηστή, ούτε με προσωπικά χαρακτηριστικά κάθε χρήστη αλλά συλλέγει πληροφορίες σχετικά με τους χρήστες γενικά και τα μαθησιακά αντικείμενα και τα εισάγει στο σύστημα συστάσεων [12].

Ένα τέτοιο σύστημα απαιτεί την απόκτηση γνώσεων και μπορεί να συνδυαστεί με άλλες τεχνικές ειδικά για την επίτευξη συστάσεων σε μαθησιακά περιβάλλοντα.

3.4.6 Συστήματα συστάσεων με επίγνωση περιβάλλοντος

Το σύστημα αυτό επικεντρώνεται σε ιδιαίτερα χαρακτηριστικά, πέρα από τις δυο ταυτότητες χρήστης-αντικείμενο, όπως χρόνος, ημερομηνία, τοποθεσία, ψυχολογικοί παράγοντες, συναισθηματική κατάσταση, προσωπικό ιστορικό κα. Δεν αρκείται σε γενικές πληροφορίες αλλά προσπαθεί να εντοπίσει τι θα ενδιέφερε τον χρήστη υπό συγκεκριμένες συνθήκες και να εγείρει την επιθυμία του χρήστη για συστάσεις μέσα από ένα βελτιωμένο μοτίβο ερωταποκρίσεων συστάσεων [12]. Οι πληροφορίες που χρησιμοποιεί ένα τέτοιο σύστημα υποδεικνύουν την κατάσταση μιας οντότητας και αντιμετωπίζει το χρήστη ολιστικά και όχι αποσπασματικά, τον εξετάζει υπό ένα ευρύ φάσμα πληροφοριών που οδηγούν σε κατανόηση της προοπτικής του χρήστη [12].

3.4.7 Συγκριτικοί πίνακες των βασικών τεχνικών δημιουργίας συστημάτων συστάσεων

Ο πίνακας των [4] που φαίνεται παρακάτω δείχνει συγκεντρωτικά τα υπέρ και τα κατά των μεθόδων:

Πίνακας 1. Σύγκριση των διάφορων μοντέλων φιλτραρίσματος

Μέθοδοι Φιλτραρίσματος	Περιγραφή	Πλεονεκτήματα	Μειονεκτήματα
Βάσει Περιεχομένου Content-Based	Το φιλτράρισμα στηρίζεται σε εξαγωγή και ταίριασμα των χαρακτηριστικών περιεχομένου.	Ανάλυση και ομοιότητες ανάμεσα σε ένα μεγάλο σύμπλεγμα πηγών δεδομένων μπορεί να γίνει χωρίς ιδιαίτερη ανάμειξη του χρήστη.	Εξαρτάται σε μεγάλο βαθμό από το περιεχόμενο. Έλλειψη περιεχομένου μπορεί να οδηγήσει σε εσφαλμένες προτάσεις. Η αλληλεπίδραση άλλων χρηστών με το σύστημα δεν λαμβάνεται υπόψη.

Συνεργατικό Φιλτράρισμα Collaborative Filtering	Το φιλτράρισμα στηρίζεται σε ομοιότητες του προφίλ του χρήστη και των προφίλ των άλλων χρηστών της ίδιας ομάδας.	Λειτουργεί ανεξάρτητα από το περιεχόμενο και παρέχει περισσότερο ακριβή αποτελέσματα σε σχέση με φιλτράρισμα βάσει περιεχομένου.	Έχει φτωχές προβλέψεις για νέα αντικείμενα στη βάση δεδομένων. Ο χρήστης θα πρέπει να έχει αρκετές αξιολογήσεις στο προφίλ για να μπορεί να συγκριθεί με προφίλ άλλων χρηστών.
Υβριδικό (Hybrid)	Συνδυάζει τις 2 παραπάνω μεθόδους.	Μειώνει τα σφάλματα σε κάθε μέθοδο και σημειώνει τη δύναμή τους.	Αν δημιουργηθεί επιτόλαιο μπορεί να οδηγήσει σε λανθασμένες συστάσεις.

Και ο πίνακας των [12] δείχνει πλεονεκτήματα και περιορισμούς αλλά για περισσότερες μεθόδους:

Πίνακας 2. Σύγκριση τεχνικών για συστάσεις

Τεχνική	Πλεονέκτημα	Περιορισμός
Collaborative filtering	Οι πληροφορίες τομέα δεν απαιτούνται.	Προκλήσεις όπως ψυχρή εκκίνηση, διασπορά, επεκτασιμότητα
Content-based filtering	Γνώση τομέα δεν είναι αναγκαία.	Υπέρ-εξειδίκευση.
Demographic recommendation	Δεν απαιτεί το ιστορικό αξιολογήσεων του χρήστη.	Ο μαθητευόμενος πρέπει να εισάγει προσωπικά χαρακτηριστικά.
Utility-based recommendation	Παράμετρος μπορεί να είναι χαρακτηριστικά όχι του προϊόντος.	Ο χρήστης πρέπει να βάλει την συνάρτηση χρησιμότητας.
Knowledge-based recommendation	Ανεξάρτητη από αξιολογήσεις χρηστών.	Απόκτηση Γνώσεων.
Hybrid filtering	Δεν παρουσιάζει το πρόβλημα ψυχρής εκκίνησης.	Ζήτημα με την πολυπλοκότητα του χρόνου.
Context aware systems	Με βάση την αλλαγή πλαισίων, οι συστάσεις μπορούν να προσαρμοστούν.	Ενσωμάτωση δεδομένων που υπονοούνται.
Ontology based recommendation	Βελτιώνει την ποιότητα των προσωποποιημένων συστάσεων.	Η Γνώση του τομέα απαιτείται.

3.4.8 Συστήματα συστάσεων με βάση τις οντολογίες

Οντολογία είναι ένας τρόπος έκφρασης και επεξήγησης μιας έννοιας με τρόπο περιγραφικό μέσα από οντότητες, χαρακτηριστικά και σχέσεις μεταξύ τους. Σε ότι αφορά το χρήστη μοντελοποιεί τη γνώση σχετικά με το γνωστικό υπόβαθρο, το αντικείμενο, και τον τομέα. Καθώς και προσωπικές πληροφορίες, τρόπο εκμάθησης και επίπεδο γνώσεων του χρήστη. Ενώ σε ότι αφορά το μαθησιακό αντικείμενο εξετάζει τον τύπο και τη μορφή των πόρων. Η παραγωγή συστάσεων μέσω οντολογιών βελτιώνει την ποιότητα των προσωποποιημένων συστάσεων αλλά συνιστά μία δύσκολη και χρονοβόρα διαδικασία [12].

«Οντολογία είναι μία μέθοδος δόμησης της πληροφορίας» σύμφωνα με τους [4]. Επιτρέπει την ιεραρχική δομή της πληροφορίας η οποία χρησιμοποιείται στη διαλογή των πληροφοριών. Μέσα από αυτή την ιεραρχική δομή προκύπτει το γνωστικό μονοπάτι (Knowledge Path) το οποίο μέσα από τη δομή δέντρου οδηγεί στη ζητούμενη λέξη κλειδί. Τα βήματα που χρειάζονται είναι: η διεύθυνση URL ορίζεται ως είσοδος, η οντολογία με δομή δέντρου δημιουργείται και παράγεται το γνωστικό μονοπάτι, λαμβάνονται υπόψη όλα τα URL που είναι σχετικά με την είσοδο που τέθηκε, σύνδεσμοι, που παρουσιάζονται στην ληφθείσα ιστοσελίδα, εξάγονται και αποτελούν νέο URL. Τέλος, τα βήματα επαναλαμβάνονται μέχρι να αποκτηθεί το απαιτούμενο αποτέλεσμα [4]. Ένα πρόγραμμα ανίχνευσης ιστού με λέξεις κλειδιά μειώνει το χρόνο αναζήτησης και κατεβάζει μόνο τις απαιτούμενες ιστοσελίδες [4].

3.4.9 Συστήματα Συστάσεων με χρήση τεχνικών εξόρυξης δεδομένων (*Data mining*).

Κάθε σπουδαστής αλληλεπιδρά με ένα σύστημα ηλεκτρονικής μάθησης και τα ίχνη των δραστηριοτήτων του αποθηκεύονται σε μία ΒΔ. Στις περισσότερες περιπτώσεις τα δεδομένα αυτά δεν χρησιμοποιούνται παρά μόνο για στατιστικούς λόγους ή συνόψεις των δραστηριοτήτων. Οι συλλογές δεδομένων που προκύπτουν και παραμένουν αναξιοποίητες θα μπορούσαν να αποτελούν πηγές για εξαγωγή πολύ χρήσιμων πληροφοριών [6]. Η ανάλυση αυτών των δεδομένων με τη χρήση τεχνικών εξόρυξης της πληροφορίας μπορεί να δημιουργήσει μοτίβα, ομάδες δεδομένων ή προφίλ που δεν γίνονται φανερά με την πρώτη ματιά και αποτελούν εργαλεία για τη βελτίωση της μαθησιακής διαδικασίας μέσω ενός ηλεκτρονικού περιβάλλοντος. Είναι αυτές οι πληροφορίες που θα οδηγήσουν σε προσωποποιημένες συστάσεις ή θα αποτελέσουν την πηγή ανατροφοδότησης για τον συγγραφέα διδακτικών αντικειμένων ή ακόμα μπορούν να χρησιμοποιηθούν στην ανίχνευση εισβολών στο σύστημα [6]. Τέλος, με αυτές τις τεχνικές μπορούν να εντοπιστούν

περιπτώσεις λογοκλοπής σε εργασίες φοιτητών ή να οικοδομηθούν ομάδες φοιτητών με μεγάλη αλληλεπίδραση.

Αναλυτικότερα, στην εργασία των [6] περιγράφονται διάφοροι τομείς ενός μαθησιακού περιβάλλοντος που μπορούν να βελτιωθούν με χρήση εξόρυξης της πληροφορίας. Κάθε σπουδαστής έχει τους δικούς του στόχους, γνωστικό υπόβαθρο, ικανότητες, προσδοκίες ακόμα και στρατηγικές πρόσληψης της γνώσης. Όλα αυτά οδηγούν στο συμπέρασμα πως ένα μαθησιακό περιβάλλον θα πρέπει να έχει τέτοια χαρακτηριστικά που να ικανοποιεί τις ανάγκες του κάθε χρήστη ξεχωριστά σαν να αφορά προσωπικά εκείνον. Όσο πιο στοχευμένη γνώση παρέχεται τόσο ελαχιστοποιείται ο χρόνος μάθησης και μειώνεται το φορτίο του συστήματος. Το εργαλείο συστάσεων προτείνει το κατάλληλο διδακτικό αντικείμενο και ο σπουδαστής δε σπαταλά χρόνο σε περιττές αναζητήσεις. Γεγονός που συνεπάγεται την αύξηση της αποτελεσματικότητας του διδακτικού περιεχομένου και συνολικά ολόκληρης της μαθησιακής εμπειρίας. Ουσιαστικά προσαρμοζόμενο το σύστημα στις δεξιότητες αλλά και τις δυσκολίες του κάθε σπουδαστή χτίζει ένα προσωπικό μαθησιακό μονοπάτι [6].

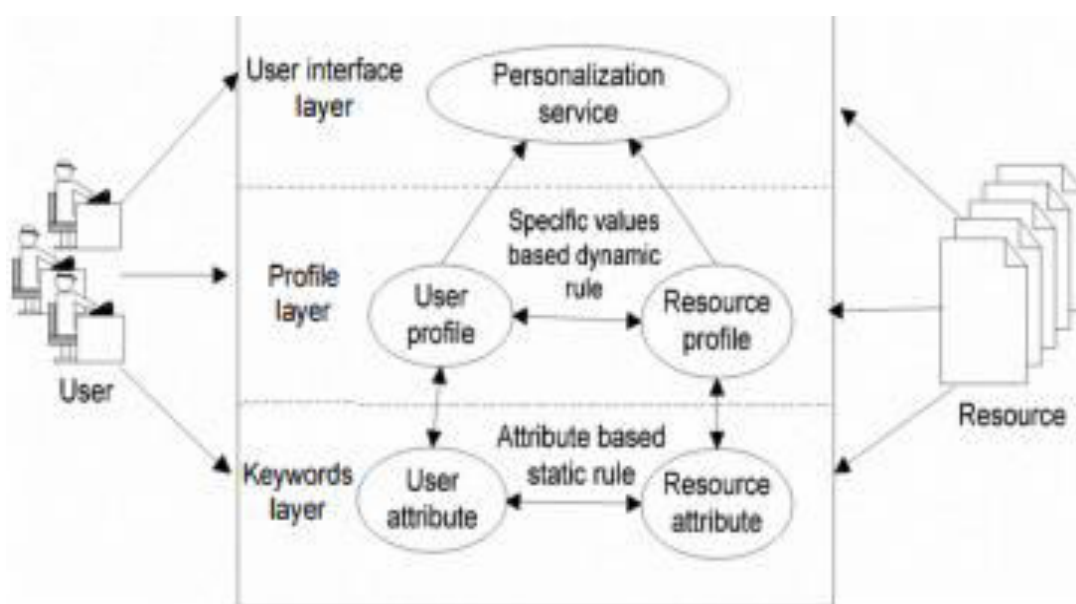
Από την άλλη μεριά ο διδάσκων ενός αντικειμένου αποκτά γνώση του αντίκτυπου που έχει στους φοιτητές του με έμμεσο τρόπο και κατανοεί τις δυσκολίες του μαθήματός του ή το βαθμό κατανόησης από μέρους των φοιτητών. Άλλες πληροφορίες που μπορεί να εξαχθούν είναι η διάρκεια που απαιτείται για κάποια δραστηριότητα του μαθήματος ή η αναζήτηση και ενέργειες όπως αποθήκευση, εκτύπωση κλπ.

Ένα άλλο καθήκον στο οποίο η εξόρυξη πληροφορίας είναι χρήσιμη είναι ο σχεδιασμός ομάδων με κοινά χαρακτηριστικά που θα μπορούσε να αποτελεί μία επιτυχή ομάδα συνεργασίας ή επικοινωνίας. Από την άλλη μεριά θα μπορούσε να αναγνωρίσει στοιχεία πλαγιαρισμού ή αντιγραφής ανάμεσα σε εργασίες φοιτητών, εξετάσεις ή άλλα έργα.

Η τεχνολογία εξόρυξης της πληροφορίας προσφέρει αλγορίθμους για τη δημιουργία συσχετιστικών κανόνων που οδηγούν σε συστάσεις. Η ποιότητα και ποσότητα των κανόνων καθορίζει και την επιτυχία των συστάσεων. Το πιο σημαντικό στοιχείο στην εξόρυξη πληροφοριών είναι η ποιότητα των δεδομένων διότι υπάρχει ο λεγόμενος κανόνας GIGO (Garbage In Garbage Out) [6], που σημαίνει πως ότι ποιότητας δεδομένα εισάγει κανείς ανάλογης ποιότητας συμπεράσματα εξάγει. Σε ότι αφορά λοιπόν τα δεδομένα υπάρχουν 5 στάδια: η προ-επεξεργασία, ο μετασχηματισμός, η εξόρυξη, η οπτικοποίηση και η ερμηνεία. Στις 2 πρώτες φάσεις επιλέγεται και απαλλάσσεται από «σκουπίδια» η συλλογή δεδομένων. Στην επόμενη φάση εξορύσσεται η πληροφορία χωρίς περιττά μοτίβα δεδομένων, με ταξινόμηση, συσταδοποίηση, ανακάλυψη ή πρόβλεψη χρονικής σειράς. Τέλος, το μοντέλο που παράγεται από την εξόρυξη αναπαριστάται γραφικά και ερμηνεύεται ως προς τη σημασία του [6]. Κύριος στόχος της όλης προσπάθειας είναι η εξαγωγή χρήσιμων και ασφαλών

συμπερασμάτων από την ανάλυση δεδομένων, που με τη σειρά τους θα επιφέρουν βελτιώσεις στο μαθησιακό περιβάλλον.

Όσο περισσότεροι οι κανόνες, τόσο δυσκολότερη η διαχείριση ενός συστήματος συστάσεων [17]. Κανόνες μπορούν να δημιουργηθούν είτε από στατικές πληροφορίες του χρήστη, είτε από δυναμικές. Ανάλογα με το τι περιεχόμενο έχει διαβάσει ο χρήστης και έχει ενδιαφερθεί μπορεί να γίνει πρόβλεψη βάσει των κανόνων που παράγονται [17]. Σχηματικό διάγραμμα αυτών των συστημάτων εικονίζεται παρακάτω:



Εικόνα 5. Σχηματικό διάγραμμα συστάσεων βάσει συσχετιστικών κανόνων.

Η ανάλυση χαρακτηριστικών του σπουδαστή είναι πρωταρχικής σημασίας για την επίτευξη ενός προσωποποιημένου συστήματος συστάσεων. Παράγονται συστάσεις για διαφορετικό περιεχόμενο όπως κείμενα, βιβλία, άρθρα, εικόνες, ιστοσελίδες, βίντεο κλπ. Οι προτιμήσεις του χρήστη συγκρίνονται με κάποια ιδιαίτερα χαρακτηριστικά και γίνεται πρόβλεψη της πρόθεσης του χρήστη για περιεχόμενο που δεν έχει αξιολογηθεί από τον ίδιο [20].

Οι τεχνικές εξόρυξης δεδομένων προσφέρουν όλους εκείνους τους αλγορίθμους που μπορούν να ανακαλύψουν σχέσεις με σημασία και νόημα μέσα από μεγάλο μέγεθος δεδομένων [20].

3.5 Προκλήσεις στα ΣΣ

Τα συστήματα συστάσεων χρησιμοποιούνται όλο και περισσότερο στη σύγχρονη εποχή γεγονός που εγείρει προκλήσεις ως προς την βελτιστοποίηση των διαδικασιών και την εφαρμογή περισσότερων και ακριβέστερων αλγορίθμων [8]. Καμία τεχνική δεν είναι άμοιρη

μειονεκτημάτων και κάθε φορά χρειάζεται αντιμετώπιση ή εξεύρεση λύσεων για την επίτευξη καλύτερων συστάσεων.

Τα σημαντικότερα προβλήματα που αντιμετωπίζουν τα ΣΣ είναι τα εξής:

3.5.1 Πρόβλημα ψυχρής εκκίνησης (Cold-start problem)

Για έναν νέο χρήστη που χρησιμοποιεί πρώτη φορά ένα σύστημα και δεν έχει προηγούμενο αξιολόγησης δεν μπορεί να γίνει πρόβλεψη καθώς δεν υπάρχουν πλησιέστεροι γείτονες βάσει των οποίων θα μπορούσαν να γίνουν κάποιες συστάσεις. Επίσης, ένα καινούριο μαθησιακό αντικείμενο για το οποίο δεν υπάρχουν αξιολογήσεις δεν μπορεί να κατηγοριοποιηθεί και συνιστά και αυτό πρόβλημα ψυχρής εκκίνησης [12][15].

3.5.2 Διασπορά δεδομένων (Data Sparsity)

Το πρόβλημα αυτό εντοπίζεται όταν έχουμε πολύ μικρό αριθμό σπουδαστών που έχουν αξιολογήσει κάποιο αντικείμενο, εν προκειμένω μαθησιακό και έτσι δεν υπάρχει μέτρο σύγκρισης ανάμεσα στα μαθήματα. Πολλά μαθήματα, λίγες αξιολογήσεις, συνεπάγεται ακατάλληλο δείγμα για τη δημιουργία συστάσεων ιδιαίτερα σε ότι αφορά το συνεργατικό φιλτράρισμα αφού, όταν δεν υπάρχει ικανός αριθμός αξιολογήσεων, μπορεί να οδηγήσει σε μηδενικές ομοιότητες ανάμεσα σε σπουδαστές [12][15].

3.5.3 Επεκτασιμότητα (Scalability)

Μπορεί το σύστημα συστάσεων να οδηγηθεί σε ανακριβή αποτελέσματα, όταν ο αριθμός σπουδαστών και μαθησιακών αντικειμένων αυξάνεται γραμμικά ή είναι πολύ μεγάλος ο αριθμός των δεδομένων για να γίνει επεξεργασία [12][15]. Ένας τρόπος να βελτιωθεί το πρόβλημα της επεκτασιμότητας είναι να συνδυαστεί ο αλγόριθμος συστάσεων με ανάλυση συναισθήματος η οποία επιτρέπει την επεξεργασία δεδομένων με έναν επαρκή και οικονομικό τρόπο [15].

3.5.4 Υπέρ-εξειδίκευση (Over Specialization)

Προτείνοντας συνέχεια αντικείμενα βάσει προτιμήσεων του χρήστη το σύστημα αποτρέπει εναλλακτικές προτάσεις και νέα μαθήματα με αποτέλεσμα τη μονομέρεια και την έλλειψη ποικιλίας στις προτάσεις. Το πρόβλημα αυτό συναντάται συνήθως στις συστάσεις βάσει περιεχομένου και λύση αποτελεί ο συνδυασμός και άλλων μεθόδων προκειμένου να αποφευχθεί η υπέρ-εξειδίκευση και έλλειψη ευελιξίας από πλευράς του συστήματος [12].

3.5.5 *Ιδιωτικότητα (Privacy)*

Τέλος, σε όποιο σύστημα απαιτούνται προσωπικά στοιχεία του χρήστη, όπως δημογραφικά και τοποθεσία, προκειμένου να επιτευχθούν περισσότερο ακριβείς συστάσεις, η ιδιωτικότητα κινδυνεύει να διαρραγεί καθώς ευαίσθητα προσωπικά δεδομένα κοινοποιούνται [12].

Τα προβλήματα αντιμετωπίζονται είτε με τη χρήση συνδυασμού τεχνικών, είτε με μεθόδους Μηχανικής Μάθησης. Η ανάλυση συναισθήματος με MM χρησιμοποιείται για να ταξινομεί και να παρέχει συστάσεις στους χρήστες [15]. Το σύστημα μαθαίνει από τα δεδομένα εισόδου και τα χρησιμοποιεί ώστε να κατηγοριοποιήσει νέες συστάσεις [15]. Το σημαντικό σε αυτή τη διαδικασία είναι η συλλογή των χαρακτηριστικών (προς κατηγοριοποίηση). Όσο πιο κατάλληλα τα χαρακτηριστικά, τόσο πιο καλή η εκπαίδευση του συστήματος, τόσο πιο επιτυχής η απόδοση της ταξινόμησης [15]. Το συνεργατικό φιλτράρισμα με βάση τη δημιουργία κλάσεων μειώνει τον υπολογιστικό χρόνο και επικεντρώνεται στη βελτίωση του χρόνου καθώς η κατηγοριοποίηση γίνεται εκτός του Διαδικτύου [15].

4

Εφαρμογές Εκπαιδευτικών ΣΣ

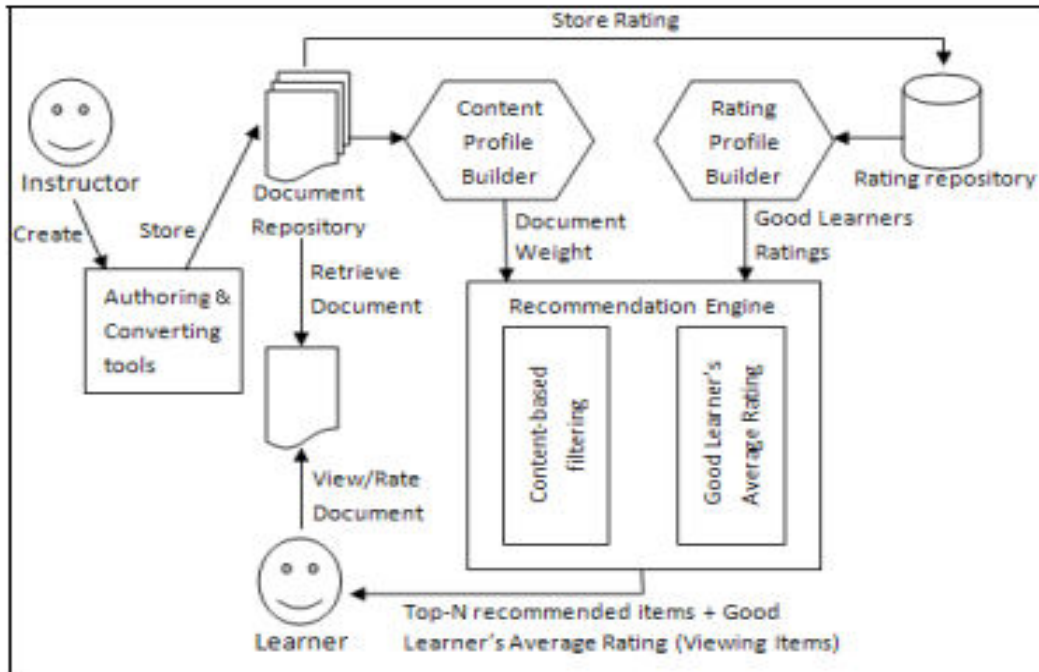
4.1 ΣΣ με Vector Space Model και Μέσο Όρο Αξιολόγησης

καλών σπουδαστών

«Σύμφωνα με την κοινωνική θεωρία μάθησης οι άνθρωποι μαθαίνουν παρατηρώντας τη συμπεριφορά άλλων καθώς και το αποτέλεσμα αυτής της συμπεριφοράς.» Η ιδέα αυτή υποστηρίζει σθεναρά την πρόταση των [16] για ΣΣ. Σε αυτή την πρόταση ο μέσος όρος αξιολογήσεων σε συγκεκριμένα αντικείμενα από καλούς φοιτητές, όπου όλοι γνωρίζουν το αποτέλεσμα της συμπεριφοράς τους, και εστιάζοντας σε αντικείμενα με υψηλή βαθμολογία από τους καλούς φοιτητές, μπορεί να οδηγήσει σε βελτίωση και της δικής τους επίδοσης. Είναι δύσκολο να γνωρίζει κανείς στον πραγματικό κόσμο ποια μαθησιακά υλικά έχουν αξιολογηθεί με υψηλή βαθμολογία.

Δύο τύπους συστάσεων προτείνει η έρευνα των [16] : α) βάσει περιεχομένου, β) με μέσο όρο αξιολογήσεων των καλών φοιτητών. Στόχος του πρώτου είναι να προσθέσει στο τρέχων μαθησιακό πλαίσιο, μαθησιακούς πόρους παραπλήσιους με αυτούς που βλέπει ήδη ο χρήστης. Ο δεύτερος έρχεται να βοηθήσει και να καθοδηγήσει τους φοιτητές στην επιλογή μαθησιακών πηγών με στόχο τη βελτίωση της μαθησιακής διαδικασίας. Χωρίς να παραβιάζεται η σειρά μαθημάτων και πάντα εντός πλαισίου.

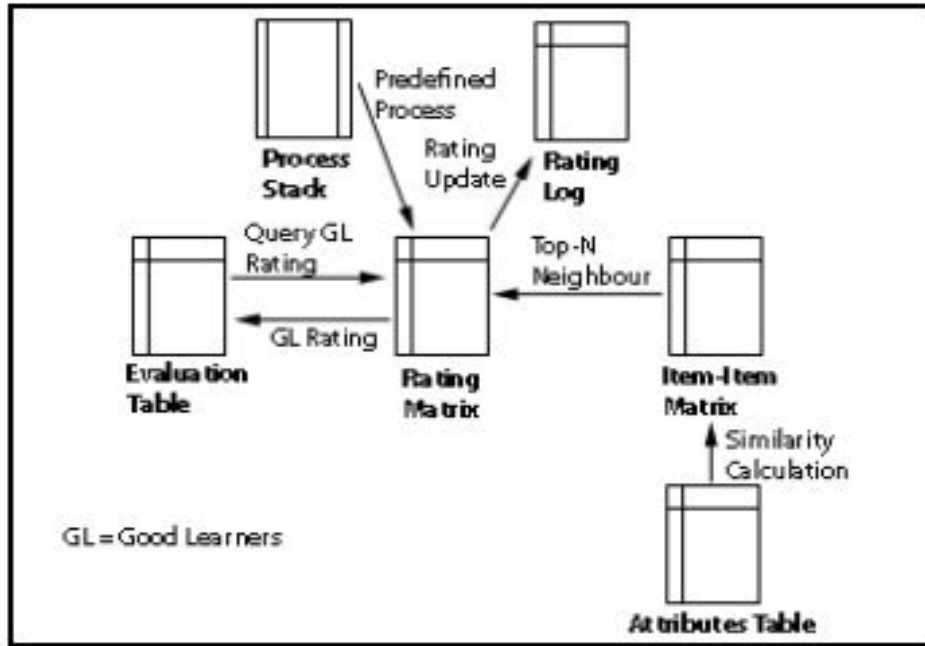
Η προτεινόμενη μεθοδολογία [16] σχηματικά απεικονίζεται ως εξής:



Εικόνα 6. Συνολική αρχιτεκτονική του ΣΣ ηλεκτρονικής μάθησης

Ο καθηγητής είναι υπεύθυνος για την δημιουργία εκπαιδευτικού υλικού και χρησιμοποιεί το εργαλείο μετατροπής και συγγραφής. Το εργαλείο αυτό μετατρέπει κείμενα και παρουσιάσεις σε εικόνες. Η μετατροπή σε εικόνες διασφαλίζει την ακεραιότητα των αρχικών κειμένων και εξοικονομεί χρόνο κατά την ανατύπωση σε άλλα συγγραφικά εργαλεία. Το συγγραφικό εργαλείο ενσωματώνει την εικόνα σε προκαθορισμένο πρότυπο html σελίδας. Όλα τα έγγραφα αποθηκεύονται σε αποθετήριο εγγράφων. Ο σπουδαστής μπορεί να δει και να αξιολογήσει το περιεχόμενο το οποίο ανακαλείται από το αποθετήριο. Ο χώρος όπου οικοδομείται το προφίλ περιεχομένου (Content Profile Builder) υπολογίζει τα βάρη των εγγράφων μέσω των οποίων θα υπολογιστεί ο βαθμός ομοιότητας αντικειμένου. Το αποθετήριο αξιολογήσεων διατηρεί τις αξιολογήσεις των καλών φοιτητών. Τελικά, εκεί όπου οικοδομείται το προφίλ αξιολογήσεων θα τεθούν ερωτήματα προς το αποθετήριο αξιολογήσεων. Τα βάρη εγγράφων μαζί με τις αξιολογήσεις γίνονται είσοδοι στη μηχανή παραγωγής συστάσεων όπου υπολογίζεται η ομοιότητα των αντικειμένων και ο μέσος όρος αξιολόγησης των καλών φοιτητών.

Το διάγραμμα ροής της συγκεκριμένης αρχιτεκτονικής φαίνεται στην εικόνα:



Εικόνα 7. Διάγραμμα ροής ΣΣ [16]

Ο πίνακας χαρακτηριστικών χρησιμοποιείται στον υπολογισμό ομοιότητας ανάμεσα στα αντικείμενα και τα αποτελέσματα αποθηκεύονται σε έναν πίνακα item-item matrix. Η ομοιότητα υπολογίζεται με χρήση του μοντέλου διανυσματικού χώρου. Τα N πρώτα προτεινόμενα αντικείμενα, που ξεπερνά η τιμή ομοιότητας το κατώφλι, θα χρησιμοποιηθούν για τον υπολογισμό των προβλέψεων αξιολογήσεων. Ο πίνακας αξιολογήσεων αρχικά αποθηκεύει τις προβλεφθείσες αξιολογήσεις και θα αντικατασταθούν μόλις ο χρήστης αξιολογήσει αντικείμενα. Η προβλεφθείσα αξιολόγηση τότε αποθηκεύεται στο αρχείο αξιολογήσεων. Η διαδικασία στοίβα περιλαμβάνει ένα σύνολο προκαθορισμένων εντολών-οδηγιών που θα εκτελεστεί κατά την διάρκεια ενός από τα ακόλουθα συμβάντα: εισαγωγή νέου αντικειμένου, διαγραφή υπάρχοντος, ενημέρωση χαρακτηριστικού αντικειμένου, αξιολόγηση και επαναξιολόγηση του αντικειμένου. Ο πίνακας επιδόσεων που αποθηκεύει τους βαθμούς των καλών φοιτητών θα ρωτήσει τον πίνακα αξιολογήσεων για τυχόν αξιολογήσεις από τους καλούς φοιτητές. Οι μέσοι όροι αξιολογήσεων χρησιμοποιούνται για τις συστάσεις.

Το Διανυσματικό μοντέλο (Vector Space Model-VSM) αφορά δύο κύριες φάσεις: τον υπολογισμό των βαρών, τον υπολογισμό της ομοιότητας συνημιτόνου. Τα βάρη υπολογίζονται με τον TF-IDF με κανονικοποιημένη συχνότητα όπως φαίνεται από τον παρακάτω τύπο:

$$w_{i,j} = \frac{f_{i,j}}{\max_z f_{z,j}} * \log\left(\frac{D}{d_i}\right)$$

Όπου $f_{i,j}$ επισημαίνει τη συχνότητα που ένας όρος i συναντάται σε ένα έγγραφο j . Το μέγιστο δείχνει τη μέγιστη συχνότητα ανάμεσα σε όλες τις λέξεις κλειδιά z που εμφανίζονται σε ένα έγγραφο j . D είναι ο συνολικός αριθμός εγγράφων που μπορούν να προταθούν στους σπουδαστές. d_i είναι ο αριθμός των εγγράφων που περιλαμβάνουν τον όρο i .

Το αποτέλεσμα της παραπάνω εξίσωσης χρησιμοποιείται στον υπολογισμό του βαθμού ομοιότητας ανάμεσα σε δύο αντικείμενα σύμφωνα με τον τύπο:

$$\cos(\vec{w}_c, \vec{w}_s) = \frac{\vec{w}_c \cdot \vec{w}_s}{\|\vec{w}_c\| \|\vec{w}_s\|}$$

Όπου τα βάρη αντιμετωπίζονται ως διάνυσμα με βάση το περιεχόμενο του προφίλ του χρήστη c και το περιεχόμενο του εγγράφου s . Οι απόλυτες τιμές των βαρών μαζί είναι το

μέγεθος των διανυσμάτων \vec{w}_c και \vec{w}_s .

Ο μέσος όρος αξιολογήσεων ορίζεται ως εξής:

$$R_{i,j} = \frac{\sum_{i=1}^N r_{i,j}}{N_j}$$

Όπου $r_{i,j}$ είναι η αξιολόγηση του καλού σπουδαστή i στο αντικείμενο j . Ο αριθμός N_j είναι ο συνολικός αριθμός καλών σπουδαστών που αξιολόγησαν το αντικείμενο j .

4.2 ΣΣ με χρήση αλγορίθμων για Συνεργατικό φιλτράρισμα

4.2.1 ΣΣ στην ηλεκτρονική μάθηση

Η μέθοδος συνεργατικού φιλτραρίσματος βάσει χρηστών επιλέγεται ως η καλύτερη μέθοδος συστάσεων για την ηλεκτρονική εκπαίδευση από τους [14] και χωρίζεται σε 5 ενότητες:

- A) Συλλογή Δεδομένων
- B) Εξαγωγή, Μετασχηματισμός, Μεταφορά (Extract, Transform, Load – ETL) Δεδομένων
- Γ) Παραγωγή μοντέλου
- Δ) Διαμόρφωση
- E) Παροχή υπηρεσιών

Ο συσχετισμός ομάδων χρηστών μπορεί να γίνει με δύο τρόπους, ο πρώτος απευθείας με αξιολόγηση των μαθημάτων από τον χρήστη, οπότε πρόκειται για φυσική συμμετοχή στο σύστημα. Ο άλλος τρόπος είναι να συλλεχθούν πληροφορίες σχετικές με τη μελέτη, μαθησιακό μοτίβο ή ροή ενεργειών κλικ, οπότε το σύστημα λειτουργεί αυτόματα [14]. Επίσης, η αναλογία χρόνου διαβάσματος με τις συνολικές ώρες μαθήματος καταγράφεται σαν σκορ και μετατρέπεται σε αξιολογική κλίμακα από το 1-5. Οι βαθμολογίες καταχωρούνται σε έναν πίνακα $m \times n$ όπου m είναι ο αριθμός χρηστών και n ο αριθμός μαθημάτων και οι τιμές $R_{j,k}$ είναι βαθμοί που δόθηκαν σε μάθημα k από σπουδαστή j . Σχηματικά ο πίνακας απεικονίζεται παρακάτω:

	Course ₁	...	Course _k	...	Course _n
Learner ₁	R _{1,1}	...	R _{1,k}	...	R _{1,n}
...
Learner _j	R _{j,1}	...	R _{j,k}	...	R _{j,n}
...
Learner _m	R _{m,1}	...	R _{m,k}	...	R _{m,n}

Πίνακας 3. Πίνακας αξιολόγησης μαθήματος από σπουδαστή

Το σημαντικότερο βήμα στα ΣΣ με συνεργατικό φιλτράρισμα είναι ο υπολογισμός ομοιότητας ανάμεσα σε σπουδαστές και χρησιμοποιείται στη διαμόρφωση εγγύτητας βάσει γειτόνων ανάμεσα σε έναν χρήστη-στόχο και έναν αριθμό χρηστών με παρόμοιες απόψεις. Ο κύριος σκοπός σχηματισμού γειτονιάς είναι να βρεθεί για κάθε χρήστη u μία λίστα από l χρήστες $N = \{N_1, N_2, \dots, N_l\}$ όπου ο u είναι διαφορετικός από τον N και η ομοιότητα $\text{sim}(u, N_1)$ είναι μέγιστη, η ομοιότητα $\text{sim}(u, N_2)$ είναι η αμέσως επόμενη μεγαλύτερη κοκ. Ο σχηματισμός γειτονιάς χρησιμοποιεί δύο μεγέθη το μέτρο εγγύτητας και τον αλγόριθμο σχηματισμού γειτονιάς.

Στη μελέτη των [14] χρησιμοποιείται η εγγύτητα ανάμεσα σε δύο σπουδαστές a και b , ενώ για τον υπολογισμό χρησιμοποιείται η συσχέτιση Pearson coor_{ab} που δίνεται από τον τύπο:

$$\text{coor}_{ab} = \frac{\sum_i (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\sum_i (r_{ai} - \bar{r}_a)^2 \sum_i (r_{bi} - \bar{r}_b)^2}}$$

Το τελικό βήμα του CF είναι η ανάδυση των κορυφαίων top- N συστάσεων από τη «γειτονιά» των σπουδαστών. Για το σκοπό αυτό χρησιμοποιούνται οι συσχετιστικοί κανόνες. Για να παραχθούν οι συσχετιστικοί κανόνες δεν χρησιμοποιείται όλος ο πληθυσμός σπουδαστών αλλά ένα μέρος l γειτόνων. Βέβαια, εδώ προκύπτει το ζήτημα πως αν το δείγμα l είναι μικρό οι κανόνες που θα παραχθούν θα είναι αναποτελεσματικοί και δε θα μπορούν να δώσουν

ικανή λίστα προτάσεων μαθημάτων. Δυσκολία που αίρεται αν χρησιμοποιηθεί ένα υποσύνολο συχνά χρησιμοποιούμενων μαθημάτων.

Οι [14] σχεδιάζουν το ΣΣ ανεξάρτητα από σύστημα ηλεκτρονικής μάθησης ώστε να μειωθεί περαιτέρω βάρος στην ιστοσελίδα κάτι το οποίο, φυσικά, δεν γίνεται αντιληπτό από τους χρήστες. Ο χρήστης στέλνει τα αιτήματά του και λαμβάνει απαντήσεις, εν προκειμένω, σχετικές σελίδες, χωρίς να γνωρίζει τις τεχνολογίες που χρησιμοποιούνται στην υλοποίηση του περιβάλλοντος.

Το διάγραμμα ροής του ΣΣ των [14] αποτελείται από 5 στάδια:

- 1) Συλλογή Δεδομένων: Συλλέγονται οι πληροφορίες μέσω του ιστορικού, συχνά επισκεπτόμενες σελίδες, αρχεία καταγραφής και αξιολογήσεις μαθημάτων.
- 2) ETL: Οι χρήσιμες πληροφορίες από τις πηγές δεδομένων εξάγονται, τροποποιούνται και αποθηκεύονται στην αποθήκη δεδομένων.
- 3) Παραγωγή Μοντέλου: ανάλογα με τις απαιτήσεις χρησιμοποιούνται και οι κατάλληλοι αλγόριθμοι. Τα μοντέλα που παράγονται αποθηκεύονται στη ΒΔ μοντέλων.
- 4) Διαμόρφωση; Διαφορετικά είδη υπηρεσιών εξυπηρετούνται από διαφορετικές στρατηγικές συστάσεων.
- 5) Παροχή υπηρεσιών: Το ΣΣ αναλύει τις απαιτήσεις του χρήστη, τρέχει τους κατάλληλους αλγόριθμους, γεννά αποτελέσματα και τα επιστρέφει στο περιβάλλον ηλεκτρονικής μάθησης.

Η διαδικασία αυτή των 5 σταδίων επαναλαμβάνεται συνεχώς. Όταν τα δεδομένα στη ΒΔ συναλλαγών ενημερώνονται τότε και οι συλλογές δεδομένων και τα μοντέλα στην αποθήκη δεδομένων πρέπει να ενημερωθούν ώστε να ανιχνεύονται οι αλλαγές στη συμπεριφορά του χρήστη. Η συχνότητα ενημέρωσης διαφέρει από σύστημα σε σύστημα ανάλογα με τις απαιτήσεις, αλλά συνήθως, αφορά σταθερό χρονικό διάστημα.

4.2.2 Συνεργατικό φιλτράρισμα με χρήση *k-means* αλγορίθμου

Η συμπεριφορά των εκπαιδευομένων μέσα σε ένα κοινωνικό δίκτυο με στόχο την ηλεκτρονική μάθηση μπορεί να διαφέρει σημαντικά όποτε εγείρεται η ανάγκη κατηγοριοποίησης των χρηστών σε ομάδες με διακριτά χαρακτηριστικά. Η ταξινόμηση των χρηστών με ορθά ορισμένα κριτήρια πριν την παραγωγή συστάσεων μπορεί να επιτευχθεί με εργαλεία που παρέχει η Μηχανική Μάθηση. Μία προσέγγιση για επιτυχή κατηγοριοποίηση των χρηστών είναι η έρευνα των [21] που χρησιμοποιεί τον αλγόριθμο *k-means*. Στόχος του αλγορίθμου στη συγκεκριμένη έρευνα είναι η ανάλυση της συμπεριφοράς των χρηστών με πολλαπλά κριτήρια με απώτερο σκοπό την παραγωγή επιτυχών συστάσεων. Άλλοι λόγοι για

αυτού του είδους την κατηγοριοποίηση είναι η αναγνώριση των περισσότερο ενεργών χρηστών από αυτούς που χρησιμοποιούν το ηλεκτρονικό μαθησιακό περιβάλλον λιγότερο ή να ταξινομηθούν οι δραστηριότητες των χρηστών ανάλογα με κάποιες ενέργειές τους. Στην έρευνα των [21] συνδυάζεται ο k-means με συσχετίσεις και συνυπάρξεις ως προκαταρκτικό βήμα. Δύο είναι τα βασικά κριτήρια: ο βαθμός δραστηριοποίησης στο σύστημα και ο βαθμός υστέρησης ανάμεσα σε δύο συμβάντα. Κάθε κλάση αντιμετωπίζεται ξεχωριστά αλλά με βάση την συσχέτιση και την συνύπαρξη. Ο συνδυασμός αυτός οδηγεί σε βελτιωμένες συστάσεις.

Ο k-means εντάσσεται στους αλγόριθμους μηχανικής μάθησης χωρίς επίβλεψη και είναι ένας από τους πιο δημοφιλής αλγόριθμους. Το θεωρητικό πλαίσιο παρουσιάζεται συνοπτικά ως εξής: «Είναι ένας τύπος μη ιεραρχικής ομαδοποίησης. Παρόμοια δεδομένα ομαδοποιούνται στην ίδια κλάση. Ο αλγόριθμος k-means λειτουργεί με τον υπολογισμό της υπάρχουσας απόστασης μεταξύ των δεδομένων με καθορισμένες παραμέτρους. Όσο πιο κοντά είναι η απόσταση, τόσο μεγαλύτερη πιθανότητα δύο σημεία να βρίσκονται στην ίδια ομάδα.

Από τις γνωστές αποστάσεις:

- Ευκλείδεια απόσταση
- Απόσταση Μανχάταν

Είσοδοι στον αλγόριθμο αποτελούν οι k ομάδες και ο πίνακας δεδομένων. Οι συστάδες που προκύπτουν στο τέλος αντιπροσωπεύουν τις εξόδους του αλγόριθμου. Ο αλγόριθμος περιστρέφεται μέχρι να συγκλίνει ή να σταθεροποιηθεί.»

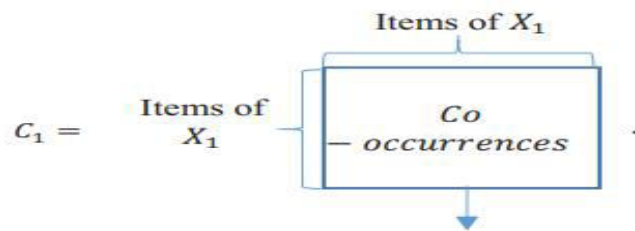
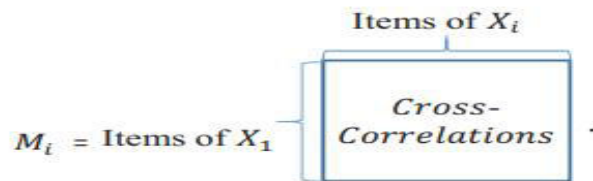
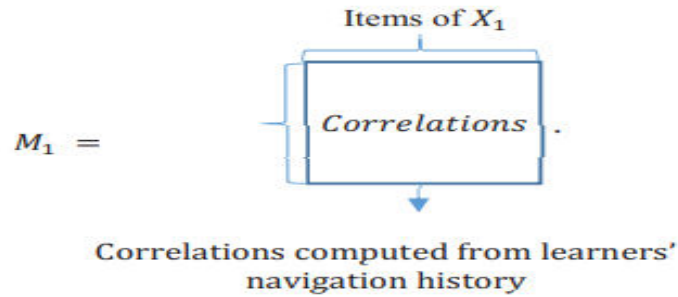
Σε πολλές έρευνες που αφορούν ηλεκτρονική μάθηση γίνεται χρήση αυτού του αλγόριθμου. Στα μαθησιακά περιβάλλοντα ως κριτήρια ταξινόμησης μπορούν να χρησιμοποιηθούν οι επιδόσεις των φοιτητών, οι δραστηριότητες εντός του συστήματος, οι αλληλεπιδράσεις και άλλα κριτήρια και να παραχθούν συστάσεις σχετικές και κατάλληλες για τους εκπαιδευόμενους. Τελικός στόχος είναι να γίνεται η κατηγοριοποίηση με εκείνα τα κριτήρια που θα οδηγήσουν σε συστάσεις κατάλληλες για την κάθε ομάδα ξεχωριστά. Ο κάθε εκπαιδευόμενος ανάλογα με το που ανήκει, σε ποια κλάση θα δέχεται και διαφορετικές προτάσεις.

Ο υπολογισμός της συσχέτισης ανάμεσα σε γεγονότα που εκτελέστηκαν από τους χρήστες και των συνυπάρξεων τους ταυτόχρονα δίνουν μία βαθμολογία συστάσεων. Υπολογίζεται το κάθε μέγεθος ξεχωριστά και το άθροισμα των δύο τιμών δίνει την τελική βαθμολογία.

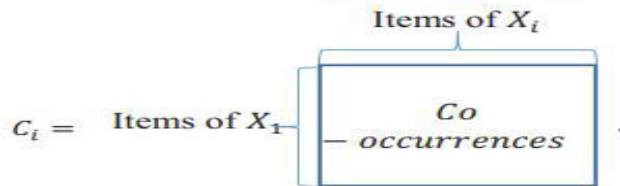
Σχηματικά:

$$r_1(X_1) = M_1 \times h(X_1) + \sum_{i=2}^n M_i \times h(X_i). \quad (1)$$

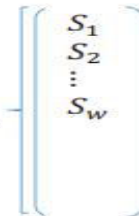
$$r_2(X_1) = C_1 \times H(X_1) + \sum_{i=2}^n C_i \times H(X_i). \quad (2)$$



Co-occurrences computed from learners' navigation history



$$\Rightarrow r_1(u, X_1) + r_2(u, X_1) = \text{recommendations scores for } X_1 \text{ items}$$



Such as :

$w = \text{Number of } X_1 \text{ items}$

Όπου:

$w =$ αριθμός από X_1 αντικείμενα

$h(X_i) =$ διάνυσμα ιστορικού εκπαιδευομένου σχετικού με το γεγονός X_i

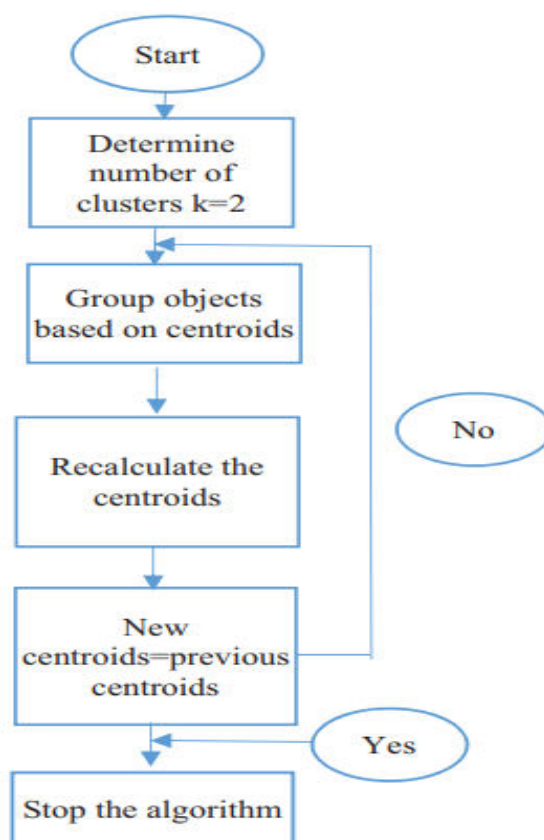
$h(X_j) =$ διάνυσμα ιστορικού εκπαιδευομένου σχετικού με το γεγονός X_j

$H(X_1)$ = δυαδικό διάνυσμα ιστορικού εκπαιδευομένου σχετικού με το γεγονός X_1

$H(X_j)$ = δυαδικό διάνυσμα ιστορικού εκπαιδευομένου σχετικού με το γεγονός X_j

Στη συλλογή δεδομένων από την οποία θα γίνει η άντληση των δεδομένων για συσταδοποίηση συμπεριλαμβάνεται η συχνότητα των γεγονότων. Ο k-means με δυο κριτήρια που χρησιμοποιούν τη συχνότητα θα προχωρήσει σε διαχωρισμό σε συστάδες. Το ένα κριτήριο είναι ο βαθμός ενασχόλησης του χρήστη στο διαδικτυακό περιβάλλον και υπολογίζεται από το άθροισμα των στιγμιότυπων που συνυπάρχουν και το άλλο είναι ο ρυθμός των κενών ανάμεσα σε κάποια γεγονότα του κάθε χρήστη, ο μέσος όρος χρονικής απόστασης ανάμεσα σε διαφορετικά γεγονότα. Δηλαδή, στην προκειμένη περίπτωση ο βέλτιστος αριθμός $k = 2$ με 15 επαναλήψεις για να υπάρξει σύγκλιση.

Το διάγραμμα ροής της εκτέλεσης του αλγορίθμου [21] φαίνεται στην εικόνα 8.



Εικόνα 8. Διάγραμμα ροής εκτέλεσης αλγορίθμου k-means.

Η έρευνα των [21] οδηγεί στο συμπέρασμα ότι η χρήση του k-means αλγορίθμου οδηγεί σε βελτιωμένα και περισσότερο ισχυρά αποτελέσματα στην παραγωγή συστάσεων. Η σύγκριση

έγινε σε 2 ΒΔ, που περιελάμβαναν περισσότερο και λιγότερο ενεργούς χρήστες, και υπολογίστηκαν ως μεγέθη η ορθότητα (accuracy) και η ακρίβεια (precision). Σαφώς ο αλγόριθμος λειτουργεί όταν υπάρχει μεγάλη διαφοροποίηση ανάμεσα στους χρήστες και όχι όταν έχουν παρόμοια συμπεριφορά. Σε μία τέτοια περίπτωση δεν συνεισφέρει στο αποτέλεσμα.

Η εργασία των [19] παρουσιάζει τη χρήση του αλγορίθμου συνεργατικού φιλτραρίσματος δίνοντας έμφαση στα χαρακτηριστικά των χρηστών. Στοχεύει στην εξάλειψη προβλημάτων που παρουσιάζει ο αλγόριθμος του συνεργατικού φιλτραρίσματος όπως η ψυχρή εκκίνηση και η αραιότητα δεδομένων. Σε αυτή την προσπάθεια δεν χρησιμοποιείται μόνο η ανάλυση συμπεριφοράς του χρήστη και η συνάρτηση του χρόνου που δείχνει ενδιαφέρον ο χρήστης αλλά αφού υπολογιστεί ο βαθμός ομοιότητας προστίθεται ο βαθμός προτίμησης και ο βαθμός εμπιστοσύνης για τη βελτίωση της ακρίβειας των αποτελεσμάτων.

Οι παραδοσιακοί αλγόριθμοι συστάσεων, σε γενικές γραμμές, αντλούν τη βαθμολογία χρήσης του χρήστη ή το σχόλιο για το έργο ως την κύρια πηγή για τη δημιουργία του συνόλου των «εγγύτερων γειτόνων». Αυτές οι συμπεριφορές που είναι εύκολο να συλλεχθούν και εκφράζουν απευθείας και με σαφήνεια τις προτιμήσεις των χρηστών. Ωστόσο, εξακολουθούν να υπάρχουν ορισμένες συμπεριφορές χρηστών, όπως η αυθαίρετη βαθμολογία ή η ελλιπής αξιολόγηση, που εκφράζουν έμμεση αξιολόγηση. Τα σχόλια των χρηστών δεν αντιπροσωπεύουν πλήρως ή με μοναδικό τρόπο την προτίμηση του χρήστη. Αποτέλεσμα αυτού είναι η κατασκευή ενός μοντέλου, πλησιεστέρων γειτόνων, ανακριβούς που μπορεί να οδηγήσει σε αστοχία συστάσεων.

Η τροποποιημένη ομοιότητα συνημιτόνου που συμπεριλαμβάνει και τις έμμεσες συμπεριφορές των χρηστών ενισχύει την ακρίβεια και παράγει καλύτερες συστάσεις. Οι έμμεσες συμπεριφορές φαίνονται από τις καταγραφές αναζητήσεων. Δημιουργείται ένας πίνακας καταγραφών και σε κάθε αντικείμενο του πίνακα δίνεται μία βαρύτητα. Κατά αυτόν τον τρόπο η τελική βαθμολογία κάθε χρήστη παράγεται από το άθροισμα των βαθμολογιών και των άμεσων και των έμμεσων συμπεριφορών. Ο βελτιστοποιημένος τύπος συνημιτόνου ομοιότητας σύμφωνα με τους [19] είναι:

$$sim'(u,v) = \frac{\sum_{i \in I_{uv}} (R'_{ui} - \bar{R}'_u) \times (R'_{vi} - \bar{R}'_v)}{\sqrt{\sum_{i \in I_u} (R'_{ui} - \bar{R}'_u)^2} \sqrt{\sum_{i \in I_v} (R'_{vi} - \bar{R}'_v)^2}}$$

Όπου:

R_{ui} είναι η βαθμολογία του χρήστη "u" στο στοιχείο "i", R_{vi} είναι η βαθμολογία του χρήστη «v» στο στοιχείο «i».

'R u και 'R v είναι οι μέσοι όροι βαθμολογίας χρήστη «u» και χρήστη «v», αντίστοιχα.

I_u και I_v αντιπροσωπεύουν τα σύνολα βαθμολογίας του χρήστη "u" και του χρήστη "v" αντίστοιχα. I_{uv} αντιπροσωπεύει το κοινό σύνολο βαθμολογίας του χρήστη "u" και χρήστη "v".

Μία άλλη παράμετρος που χρειάζεται να λαμβάνεται υπόψη είναι η μεταβολή των διαθέσεων των χρηστών στην πάροδο του χρόνου. «Στην πραγματική ζωή, τα περισσότερα από τα ενδιαφέροντα των χρηστών θα αλλάξουν με την αλλαγή του περιβάλλοντος, της ηλικίας και του χρόνου, γεγονός που οδηγεί στη διαφορά των βαθμολογιών του έργου σε διαφορετικές περιόδους» [19]. Μία χρονογραμμή αποτυπώνει τις διαφορετικές αξιολογήσεις ενός χρήστη στο πέρασμα του χρόνου και εκφράζεται μέσω μιας συνάρτησης που παράγει μία τιμή T(u,i) η οποία συμμετέχει στον τύπο του συνημιτόνου ομοιότητας πολλαπλασιαζόμενη στον αριθμητή.

Τα σχόλια σε μορφή κειμένου είναι αυτά που εκφράζουν τις προτιμήσεις ενός χρήστη. Το περιεχόμενο των σχολίων γύρω από ένα θέμα σχετικό με τα ενδιαφέροντα του χρήστη αναλύεται στη μορφή διανύσματος πιθανοτήτων.

Όλες οι πληροφορίες που εκμαιεύονται από τους χρήστες άμεσα ή έμμεσα και συμπεριλαμβάνονται στους υπολογισμούς καθώς και η παράμετρος του χρόνου οδηγούν σε μεγαλύτερη ακρίβεια του αλγορίθμου σχηματισμού συστάσεων. Η αξιολόγηση της απόδοσης του συστήματος συστάσεων γίνεται με υπολογισμό του μέσου απόλυτου σφάλματος (MAE), που βρίσκει τη μέση απόκλιση μεταξύ της προβλεπόμενης βαθμολογίας και της πραγματικής βαθμολογίας. Όσο μικρότερη είναι η τιμή του MAE, τόσο μικρότερη είναι η απόκλιση, τόσο μεγαλύτερη είναι η ακρίβεια της πρόβλεψης και τόσο καλύτερη είναι η ακρίβεια της σύστασης [19].

4.2.3 Αυτόνομο Διαδικτυακό Εκπαιδευτικό Σύστημα με ευφείς συστάσεις

Το προτεινόμενο σύστημα από την εργασία των [1] με την ονομασία Autonomous Online Education System (AOES) μπορεί να ανιχνεύσει τα βήματα μελέτης κάθε εκπαιδευόμενου, να καταγράφει σχετικές πληροφορίες και να φτιάχνει το αρχείο μελέτης του χρήστη. Παρέχει προτάσεις μελέτης δυναμικά ανάλογα με αυτό το αρχείο. Γεγονός που βελτιώνει το μαθησιακό αποτέλεσμα και την πρωτοβουλία του μαθητή/σπουδαστή.

Το AOES χωρίζεται σε τρία κομμάτια, το στάδιο της ευφούς μελέτης, ευφυή ανάλυση και ευφείς συστάσεις. Προκειμένου να αντληθούν τα ενδιαφέροντα και οι ικανότητες του χρήστη καταγράφονται πληροφορίες μελέτης, αποτελέσματα εξετάσεων κα. Συνεχίζει ανάλυση αυτών των στοιχείων και εναρμόνισή τους με τις κατάλληλες μαθησιακές τακτικές. Τελικά, παράγονται συστάσεις σύμφωνες με τα γνωστικά χαρακτηριστικά και το επίπεδο του φοιτητή [1].

Η είσοδος του χρήστη για πρώτη φορά σημαίνει εγγραφή όπου ζητούνται οι βασικές πληροφορίες του χρήστη όπως όνομα, φύλο, όνομα χρήστη και κωδικός εισόδου. Επιπλέον τομέας, επίπεδο σπουδών, μαθησιακοί στόχοι και κάποιες λεπτομέρειες για τον χρήστη. Κατόπιν, δίνονται διαγνωστικά τεστ στον χρήστη όπου διαπιστώνονται οι γνώσεις, ο βαθμός, το στυλ και οι συνήθειες μάθησης. Το αποτέλεσμα αυτής της δοκιμής θα δώσει την αρχική γνωστική ικανότητα του χρήστη η οποία θα αποτελέσει το εφευρητικό βάσει του οποίου το σύστημα θα δώσει την προκαταρκτική προσαρμοσμένη μελέτη. Στη συνέχεια, το σύστημα παρουσιάζει υλικό για μελέτη, λαμβάνοντας υπόψη τη γνωστική ικανότητα, το ιστορικό μελέτης, την αμοιβαία κατάσταση με το σύστημα κατά τη διάρκεια της μελέτης και το αποτέλεσμα του τεστ προσαρμογής για τον χρήστη. Ο χρήστης μπορεί να επιλέγει όποια τακτική μάθησης επιθυμεί. Ένας προχωρημένος χρήστης θα επιλέξει την διερεύνηση ή μεθόδους συνεργασίας, ενώ ένας περισσότερο αδύναμος σπουδαστής θα επιλέξει περισσότερο παραδοσιακές μεθόδους. Το σύστημα μετά από κατάλληλους συνδυασμούς προτείνει ενδιαφέρον περιεχόμενο. Για την απόκτηση μιας γνώσης ο στόχος είναι κοινός για όλους αλλά οι τρόποι έκφρασης αλλάζουν ανάλογα με το γνωστικό επίπεδο. Είτε ο φοιτητής πετυχαίνει το στόχο είτε όχι, το σύστημα αλλάζει δυναμικά τη γνωστική ικανότητα του χρήστη η οποία κρίνεται με δοκιμές. Τα τεστ αυτά ενσωματώνονται στη μαθησιακή διαδικασία, δεν είναι ανεξάρτητα αυτής. Δηλαδή, μελέτη και δοκιμές γίνονται μία αδιαίρετη μονάδα [1].

Το πλαίσιο σχεδιασμού του συστήματος αποτελείται από τη βάση δεδομένων πόρων, την επεξεργασία των δεδομένων, την ανάλυση δεδομένων και την εφαρμογή. Ο εκπαιδευόμενος καλείται να υλοποιήσει όλες τις δραστηριότητες μελέτης στην εφαρμογή. Οι πληροφορίες περιήγησης, λήψη αρχείων. Ερωτήματα και δοκιμές κατά τη διάρκεια της μελέτης συλλέγονται και στο επίπεδο επεξεργασίας δεδομένων. Στο στάδιο αυτό η εξαγωγή χαρακτηριστικών του χρήστη θεμελιώνεται με εξόρυξη της πληροφορίας και αποθηκεύεται στη βάση δεδομένων πόρων. Παράλληλα πραγματοποιείται ανάλυση δεδομένων με βάση την χαρακτηριστική τιμή του χρήστη στη βάση δεδομένων και το κέντρο οργάνωσης πηγών αναλαμβάνει το ταίριασμα και τον προγραμματισμό των μαθησιακών πόρων στη βάση δεδομένων πόρων-πηγών. Μετά από όλα αυτά τα στάδια προκύπτουν ευφυείς προτάσεις μαθησιακών αντικειμένων ανάλογα με το χρήστη [1].

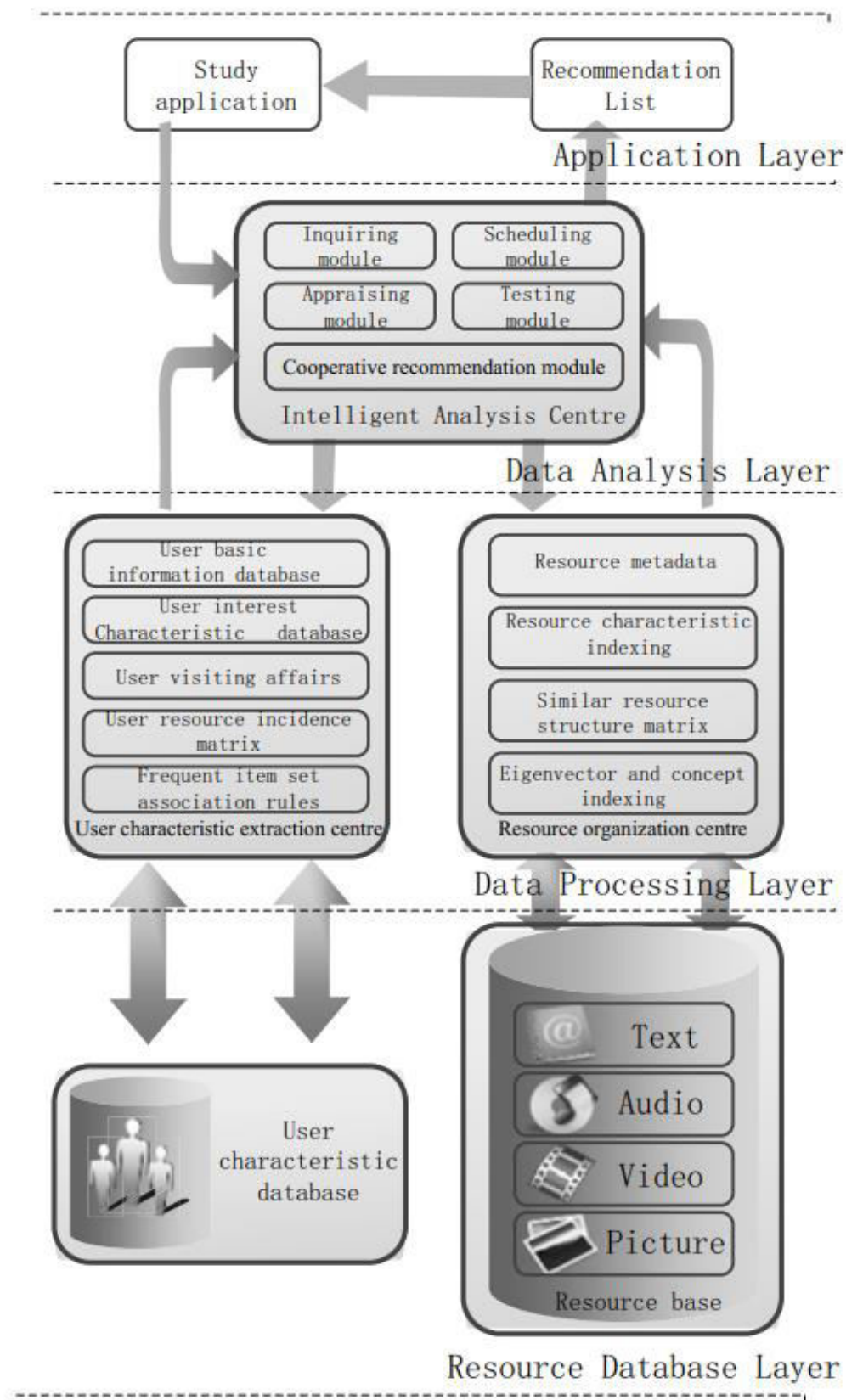
Η ΒΔ πόρων περιλαμβάνει τη ΒΔ χρήστη και τη ΒΔ πηγών. Ο χρήστης αποθηκεύει τα χαρακτηριστικά του και κάθε είδους προσωπικές πληροφορίες: γνωστική ικανότητα, εξειδίκευση κλπ. Η ΒΔ πηγών περιλαμβάνει διάφορα είδη σχετικά με τις πηγές όπως οντολογία γνώσης, λεξικό, εργαλεία λίστας λέξεων. Οι πηγές μπορεί να είναι εικόνες, ήχος, βίντεο και ηλεκτρονικά έγγραφα. Η οντολογία γνώσης, το λεξικό και τα εργαλεία λίστας λέξεων παρέχουν τη σημασιολογική έκφραση του γνωστικού πεδίου.

Το επίπεδο επεξεργασίας δεδομένων διαχειρίζεται την προ-επεξεργασία πόρων, βαθμολογεί, περιγράφει αντικείμενα, παράγει σχετικά ευρετήρια χρησιμοποιώντας ανάλογη τεχνολογία και εξάγει χαρακτηριστικά. Επιπλέον, αποθηκεύει το αποτέλεσμα επεξεργασίας σε βάση δεδομένων που περιλαμβάνει και πόρους μετά-δεδομένων, χαρακτηριστικά των ευρετηρίων αντικειμένων, παρόμοιας δομής πόρους, ευρετήριο εννοιών σε διανυσματική μορφή κλπ.

Το κέντρο ευφυούς ανάλυσης οργανώνει πληροφορίες, τις αξιολογεί, τις δοκιμάζει, ερευνά και συνεργάζεται με τον τομέα παραγωγής συστάσεων. Ο τομέας οργάνωσης πληροφοριών συλλέγει τις προσωπικές πληροφορίες του χρήστη και στέλνει δεδομένα και πόρους. Το τμήμα αξιολόγησης δομεί ένα σύστημα αξιολογήσεων και πραγματοποιεί στατιστική ανάλυση της μαθησιακής διαδικασίας. Την ίδια ώρα ο τομέας δοκιμών πραγματοποιεί διορθώσεις στην αξία των χαρακτηριστικών. Η μηχανή αναζήτησης του συστήματος στηρίζεται σε υπολογισμούς ομοιότητας και σχετικούς κανόνες. Δεν στηρίζεται μόνο σε ταίριασμα λέξεων- κλειδιών αλλά και σε σημασιολογία εννοιών. Με τον όρο αυτό εννοούμε εύρεση επεκτάσεων, επιπτώσεων και συσχετίσεων με τα ερωτήματα του χρήστη με γνώμονα τον οντολογικό συλλογισμό. Το σύστημα συνεργατικών συστάσεων προτείνει διάφορες πηγές καθ' όλη τη διάρκεια διαδικασίας μελέτης ανάλογα με το ιστορικό επισκέψεων και το βαθμό συνάφειας. Ο χρήστης περιηγείται στη λίστα συστάσεων και επιλέγει πηγές. Επιπροσθέτως, η κατάσταση οπτικής περιήγησης παρουσιάζει τις οντολογίες σε μορφή δικτυωτής δομής όπου ο χρήστης μπορεί να επιλέγει κατάλληλες λέξεις κλειδιά [1].

Η εφαρμογή δεν είναι άλλο από ένα περιβάλλον αλληλεπίδρασης ανθρώπου-μηχανής που δίνει πληροφορίες και πηγές στον χρήστη. Επιπλέον περιλαμβάνει μία σειρά από λειτουργίες όπως είσοδος, εγγραφή, μελέτη, προσαρμογή, περιήγηση, λήψη αρχείων, αξιολόγηση, θέση ερωτημάτων κα. Η λίστα συστάσεων και τα αποτελέσματα της ανάλυσης δεδομένων παρουσιάζονται σε αυτό το κομμάτι διεπαφής [1].

Το πλαίσιο λειτουργίας του AOES φαίνεται στην παρακάτω εικόνα:



Εικόνα 9. Πλαίσιο λειτουργίας AOES

Κατά τη διάρκεια της μελέτης πολλές ενέργειες του χρήστη μπορεί να δώσουν χρήσιμες πληροφορίες: η αναζήτηση με λέξεις κλειδιά, ο σελιδοδείκτης σε αγαπημένη σελίδα, η περιήγηση και τα κλικ την ώρα μελέτης, αρχεία συστήματος και αυτόματες καταγραφές στον κεντρικό εξυπηρετητή, σελίδες που ο χρήστης κάνει λήψη ή αποθηκεύει, κανάλια ή προγράμματα που παρακολουθεί ο χρήστης. Εκτός από την άμεση συλλογή στοιχείων υπάρχει και η έμμεση που εξάγεται με εξόρυξη πληροφορίας και αφορά ερωτήματα που θέτει ο σπουδαστής και περιεχόμενο μελέτης. Ένα σημαντικό στοιχείο είναι η συχνότητα με την οποία ο χρήστης συναντά συγκεκριμένο περιεχόμενο. Σύνδεσμοι, θέματα, λέξεις κλειδιά, περιλήψεις επιστημονικής και τεχνολογικής βιβλιογραφίας πιστοποιούν τη βαρύτητα και σημασία εισόδου σε κάποιο περιεχόμενο. Καταγραφή φορών σε ένα σύνδεσμο δίνει το βάρος της συχνότητας και τα βαθμό σημασίας κάθε συνδέσμου. Υπολογισμός του χρόνου περιήγησης σε σελίδες ή έγγραφα δίνει το βαθμό σπουδαιότητας ανάλογα με τη σχετικότητα του κειμένου. Δράσεις του χρήστη επάνω στο κείμενο όπως αντιγραφή παραγράφου χρησιμοποιούνται στην ανάλυση δεδομένων. Τέλος, ένα άλλο στοιχείο που υποδεικνύει το βαθμό σπουδαιότητας είναι η αποθήκευση και λήψη ενός αρχείου [1].

Ο κλασικός αλγόριθμος συνεργατικού φιλτραρίσματος μπορεί να βελτιωθεί με διάφορους τρόπους. Η ακρίβεια του αλγορίθμου βελτιώνεται αν η σχέση των αντικειμένων μεταξύ τους ληφθεί υπόψη κατά τον υπολογισμό ομοιότητας χρηστών έτσι ώστε να βρεθεί με μεγαλύτερη ακρίβεια ο πλησιέστερος γείτονας. Άλλο ένα στοιχείο που χρησιμοποιείται για να αντικατοπτριστούν οι αλλαγές στη διάθεση του χρήστη στη διάρκεια του χρόνου, εισάγεται μία συνάρτηση χρόνου (δίνει βαρύτητα στην παράμετρο χρόνος) και βοηθά τη διαδικασία προβλέψεων. Το ενδιαφέρον ενός χρήστη για ένα αντικείμενο τείνει να σβήνει σε βάθος χρόνου. Η διαπίστωση αυτή βοηθά την ευαισθησία στα νέα ενδιαφέροντα του χρήστη και δίνει ορθότερες συστάσεις σε πραγματικό χρόνο.

Ο βελτιωμένος αλγόριθμος συνεργατικού φιλτραρίσματος περιγράφεται ακολούθως:

Είσοδος: χρήστης-στόχος U^T , αντικείμενα αριθμού k , σύνολα χρηστών αριθμού m , γείτονες αριθμού n , σύνολα αντικειμένων για πρόβλεψη I_{ir} .

Έξοδος: k προτεινόμενα αντικείμενα για τον χρήστη-στόχο U^T

- 1) Υπολογίζεται η απόσταση ανάμεσα στον χρήστη-στόχο και το κέντρο της συστάδας, ταξινομείται με βάση το μέγεθος και επιλέγονται τα πρώτα m σύνολα χρηστών ως ο χώρος αναζήτησης του γείτονα.
- 2) Βρίσκεται το αντικείμενο για το οποίο ο χρήστης U^T έχει προβλεφθεί μαζί με κάθε χρήστη U_i και φτιάχνεται ένας πίνακας $U_i[]$.
- 3) Επιλέγεται ένα αντικείμενο να προβλεφθεί I^T κατά βούληση από το σύνολο αντικειμένων I_{iT} που είναι για πρόβλεψη. Ανάλογα με τον πίνακα χαρακτηριστικών

των αντικειμένων και το αντικείμενο κλάσης υπολογίζεται η σχετικότητα ανάμεσα σε όλα τα αντικείμενα με τον τύπο:

$$rel(i_1, i_2) = Sim(i_1, i_2) = \begin{cases} \frac{\sum_{j=1}^k i_{1j} \wedge i_{2j}}{|attr(i_1) \cup attr(i_2)|} & i_{1j} \& i_{2j} \text{ belong to different cluster} \\ & i_{1j} \& i_{2j} \text{ belong to same cluster} \end{cases}$$

- 4) Υπολογίζεται ο βαθμός ομοιότητας ανάμεσα στο χρήστη-στόχο και τους υπόλοιπους χρήστες με τον παρακάτω τύπο:

$$Sim^{i_T}(u_a, u_b) = \frac{\sum_{i \in I_{ab}} (r_{a,i} - \bar{r}_a)(r_{b,i} - \bar{r}_b) \bullet rel(i, i_T)}{\sqrt{\sum_{i \in I_{ab}} (r_{a,i} - \bar{r}_a)^2 \bullet rel(i, i_T)} \bullet \sqrt{\sum_{i \in I_{ab}} (r_{b,i} - \bar{r}_b)^2 \bullet rel(i, i_T)}}$$

Η ταξινόμηση γίνεται με το μέγεθος της συστάδας και επιλέγονται οι n πρώτοι πλησιέστεροι γείτονες.

- 5) Με βάση τον βαθμό ομοιότητας γίνεται πρόβλεψη ενδιαφέροντος P του χρήστη-στόχου για αντικείμενο με τον τύπο:

$$P'_{u_T, i_T} = \bar{r}_{u_T} + \frac{\sum_{u \in Neighbor_{u_T}} Sim^{i_T}(u_T, u) \times (r_{a, i_T} - \bar{r}_u) \times f(t_{ui_T})}{\sum_{u \in Neighbor_{u_T}} Sim^{i_T}(u_T, u) \times f(t_{ui_T})}$$

Στον τύπο αυτό t_{ui_T} είναι οι φορές που έδειξε ενδιαφέρον ο χρήστης u για το αντικείμενο i_T .

- 6) Τακτοποιούνται τα αντικείμενα σύμφωνα με την πρόβλεψη ενδιαφέροντος και προτείνονται τα k πρώτα στον χρήστη ως τα αντικείμενα για τα οποία θα δείξει το μεγαλύτερο ενδιαφέρον.

Προκειμένου να επιτευχθεί γρήγορη, άνετη, σταθερή, ασφαλής διαδικτυακή εκπαίδευση χρειάζεται μία μονάδα πολλαπλών ελέγχων (Multi Control Unit-MCU). Πρόκειται για έναν εξυπηρετητή που συνδέει μέσα ανάμεσα σε πολλές διευθύνσεις. Μπορεί να διεκπεραιώνει επικοινωνία ανάμεσα σε μερικές θέσεις διάσκεψης ταυτόχρονα. Έχει τη δυνατότητα να ανταλλάσσει διαφορετικού τύπου σήματα όπως βίντεο, ήχος και δεδομένα. Διαχωρίζει τα σήματα, και εξάγει πληροφορίες συγχρόνως από κάθε διάσκεψη και τα στέλνει στη μονάδα επεξεργασίας ενώ συνεχίζει η μίξη και μετατροπή βίντεο ή ήχου, η μετάδοση δεδομένων, η επιλογή διαδρομής, ο έλεγχος της διάσκεψης κα. Εν κατακλείδι, η μονάδα ελέγχου συνδυάζει όλα τα είδη πληροφοριών που είναι απαραίτητα

σε κάθε συνεδρία και τα στέλνει στον εξοπλισμό τερματικού συστήματος. Επιπλέον, το MCU έχει τη λειτουργία της αυτόματης ενοποίησης του ρυθμού μεταφοράς, έτσι ώστε κάθε σύστημα τερματικού στην ίδια διάσκεψη να λειτουργεί με την ίδια ταχύτητα.

Η εφαρμογή του συστήματος AOES οδηγεί στο συμπέρασμα πως είναι ένας νέος τρόπος βελτίωσης του διαδικτυακού εκπαιδευτικού χώρου. Χρησιμοποιεί ευφυείς τεχνολογίες για την παραγωγή συστάσεων με απώτερο σκοπό την προσαρμοστικότητα του συστήματος και την τακτική ενημέρωσή του για αλλαγές στις προθέσεις των χρηστών.

4.3 ΣΣ βασισμένο σε Νευρωνικά Δίκτυα Βαθιάς Πίστης για εξατομικευμένα μαθήματα σε περιβάλλον MOOC

Ζούμε στην εποχή των Big Data, δηλαδή σε μία εποχή όπου οι υπολογιστές καλούνται να αναλύσουν εξαιρετικά μεγάλα σύνολα δεδομένων και μέσα από αυτά να αποκαλύψουν σχέσεις, τάσεις, μοτίβα και αλληλεπιδράσεις. Τα συστήματα MOOC τα οποία εξηγήσαμε σε προηγούμενο κεφάλαιο γίνονται όλο και περισσότερο δημοφιλή. Ο αριθμός διαδικτυακών πόρων είναι τεράστιος, ενώ η γνωστική δομή του κάθε χρήστη διαφέρει. Αυτό έχει σαν αποτέλεσμα να υπάρχει δυσκολία στην αναγνώριση του κατάλληλου περιεχομένου μαθημάτων. Οι εξατομικευμένες συστάσεις βελτιώνουν τη μαθησιακή αποτελεσματικότητα [22].

Πολλά συστήματα MOOC χρησιμοποιούν το συνεργατικό φιλτράρισμα για να επιτύχουν συστάσεις μαθημάτων αλλά η λογική του είναι απλή και δε βοηθά στην εμπάθυνση του μοντέλου για περισσότερες συστάσεις και εύρεση περαιτέρω σχέσεων μεταξύ χρηστών και μαθημάτων [22].

«Η βαθιά εκμάθηση μπορεί να αντιμετωπίσει αραιά δεδομένα υψηλών διαστάσεων και έχει καλή απόδοση για εξαγωγή χαρακτηριστικών και μπορεί να χρησιμοποιηθεί για την αναπαράσταση των εσωτερικών χαρακτηριστικών της δομής δεδομένων σε μια σχετικά σύνθετη μορφή». Η έρευνα λοιπόν των [22] έρχεται να προσθέσει τα χαρακτηριστικά ενός μοντέλου βαθιάς μάθησης (Deep Belief Network-DBN) στα συστήματα συστάσεων που αφορούν πλατφόρμες MOOC με σκοπό την επίλυση των προβλημάτων που προκαλεί η μαζικότητα στη χρήση των MOOC και δεν είναι άλλα από τον μεγάλο αριθμό χαρακτηριστικών και τη διασπορά των δεδομένων.

Τρεις τομείς περιλαμβάνει η εργασία των [22]:

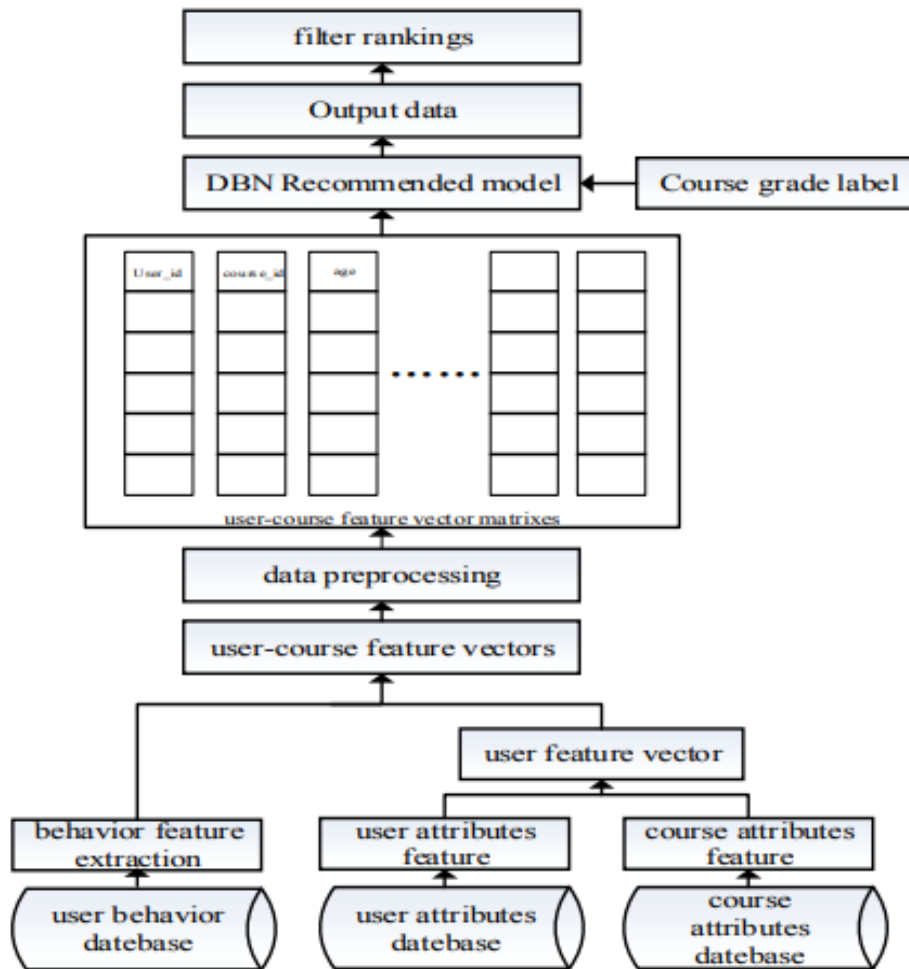
A) Συνδυάζοντας τα χαρακτηριστικά του χρήστη, με τη συμπεριφορά του και τα χαρακτηριστικά των μαθημάτων, επηρεάζεται η ενημέρωση των παραμέτρων του μοντέλου. Γεγονός που επιλύει την ανακριβή εκδήλωση ενδιαφέροντος για κάποιο μάθημα και τα αραιά δεδομένα σε συνεργατικές συστάσεις.

B) Η βαθιά μάθηση στόχο έχει την ανακάλυψη του πραγματικού ενδιαφέροντος του χρήστη και αυτό το πετυχαίνει με τη μείωση των διαστάσεων (του διανύσματος των χαρακτηριστικών) και τη δυνατότητα που έχει να εξάγει χαρακτηριστικά από αραιά δεδομένα. Θέτει ως ετικέτα κλάσης του DBN μοντέλου μάθησης με επίβλεψη, το βαθμό στο μάθημα. Ολόκληρο το δίκτυο ρυθμίζεται με προπόνηση χωρίς επίβλεψη και ανατροφοδοτείται με επίβλεψη.

Γ) Το μοντέλο DBNCF που συνδυάζει βαθιά μάθηση με συνεργατικό φιλτράρισμα έχει υψηλή αποτελεσματικότητα και καλό ποσοστό σύγκλισης. Γεγονός που πιστοποιείται και μετά από σύγκριση και ανάλυση με περισσότερα παραδοσιακές μεθόδους.

Ως συλλογή δεδομένων χρησιμοποιήθηκε το ιστορικό της συμπεριφοράς των χρηστών σε συγκεκριμένη πλατφόρμα MOOC, συνδυασμένη με χαρακτηριστικά των χρηστών και πληροφορίες αληθινών επιλογών του χρήστη. Με αυτό τον τρόπο αποκτάται ένας διανυσματικός πίνακας χαρακτηριστικών χρήστη-μαθημάτων όπου κάθε διάνυσμα αναπαριστά το βαθμό προτίμησης του χρήστη ως προς ένα μάθημα. Η βαθμολογία μαθήματος αντιπροσωπεύει το βαθμό ενδιαφέροντος από μέρος του χρήστη. Η διαδικασία παραγωγής συστάσεων αποτελεί ένα πρόβλημα ταξινόμησης πρόβλεψης το οποίο χρησιμοποιεί ως ετικέτα κλάσης τη βαθμολογία μαθήματος από τους χρήστες στην υπό επίβλεψη εκπαίδευση του συστήματος. Ο αλγόριθμος που χρησιμοποιείται είναι ο back propagation (μετάδοση προς τα πίσω ή ανάστροφη μετάδοση) ο οποίος διαδίδει μηνύματα σφάλματος από πάνω προς τα κάτω σε κάθε στρώμα του δικτύου και ρυθμίζει ολόκληρο το δίκτυο βαθιάς μάθησης εκπαιδύοντας ένα εξαιρετικό μοντέλο συστάσεων.

Η αρχιτεκτονική του συστήματος φαίνεται στην παρακάτω εικόνα:



Εικόνα 10. Αρχιτεκτονική Συστήματος DBNCF

Στη διάρκεια της προ-επεξεργασίας των δεδομένων παράγονται τρία μέρη πληροφοριών.

- Το πρώτο μέρος αφορά τα χαρακτηριστικά του χρήστη (ταυτότητα, ηλικία, φύλο, επίπεδο εκπαίδευσης, περιοχή) $U_a = \{ user_id, age, gender, loe, area \}$.
- Το δεύτερο τα δεδομένα συμπεριφοράς του χρήστη (ιστορικό περιήγησης, σχόλια, συλλογές, διαμοιρασμό, προεπισκόπηση, θεάσεις τουλάχιστον μέχρι τα μισά του μαθήματος, αριθμός αλληλεπιδράσεων, αριθμός συμμετοχών στο μάθημα ανά ημέρα, αριθμός αναρτήσεων στο φόρουμ συζητήσεων του μαθήματος) $U_b = \{ browse, comment, collection, share, viewed, explored, nevents, ndays_{act}, nplayvideo, nforum_posts \}$.
- Το τρίτο μέρος πληροφορίες σχετικά με το μάθημα (ταυτότητα μαθήματος, όνομα μαθήματος, τίτλο μαθήματος, κατηγορία μαθήματος) $C_a = \{ course_id, course_name, course_title, classification \}$. Πραγματοποιούνται διαδικασίες αφαίρεσης του θορύβου δηλαδή άχρηστων πληροφοριών, διαγραφή εγγραφών με σημαντική απόκλιση από τις κανονικές τιμές. Ακολουθεί ποσοτικοποίηση των διακριτών τιμών και

κανονικοποίηση σύμφωνα με τον τύπο $x^* = (x - \min) / (\max - \min)$. Τα δεδομένα διαχωρίζονται ανάλογα με το αν υπάρχει βαθμολογία και τα δεδομένα με βαθμό χρησιμοποιούνται ως το σύνολο εκπαίδευσης του DBNCF μοντέλου με σύνολο δοκιμής επιλεγμένα δεδομένα με βαθμολογίες και δεδομένα χωρίς βαθμολογίες αναμειγμένα.

Σε αυτή τη μελέτη των [22] υπάρχουν αυστηρές απαιτήσεις στον αριθμό επαναλήψεων, στο ρυθμό μάθησης, στα στρώματα RBM-Restricted Boltzmann Machine (ένα παραγωγικό στοχαστικό τεχνητό νευρωνικό δίκτυο που μπορεί να μάθει μια κατανομή πιθανοτήτων στο σύνολο των εισόδων του) και στον αριθμό νευρώνων σε κάθε στρώμα. Ο αριθμός χαρακτηριστικών στο διάνυσμα χρήστη-μάθημα θεωρείται ως ο αριθμός ορατών στρωμάτων νευρώνων. Ανιχνεύοντας τους αμοιβαίους περιορισμούς ανάμεσα στις παραμέτρους και μετά από επαναλαμβανόμενα πειράματα καταλήγει κανείς στο συμπέρασμα πως ο αριθμός επαναλήψεων είναι 200, ο ρυθμός μάθησης 0,05, τα ορατά στρώματα νευρώνων 19, ο αριθμός RBM στρωμάτων είναι 3 και ο αριθμός νευρώνων τύπου DBN για κάθε στρώμα είναι 19, 39, 79, 159, 5.

Τα αποτελέσματα του μοντέλου στηρίζονται στη βαθμολογία των μαθημάτων από τους χρήστες και κάθε βαθμός ανήκει σε μία κατηγορία (0 αστέρια, 1 αστέρι, 2 αστέρια, μέχρι το άριστα που είναι ο βαθμός 5). Για να γίνουν πλουσιότερα τα δεδομένα και να μειωθούν τα σφάλματα, οι διαφορές στα στοιχεία αξιολόγησης θα πρέπει να αντικατοπτρίζονται πλήρως.

Δύο είναι τα στάδια εκπαίδευσης του μοντέλου: η προπόνηση και η ρύθμιση των παραμέτρων. Με το πρώτο στάδιο επιδιώκεται βελτιστοποίηση των παραμέτρων ώστε να χρησιμοποιηθεί ο αλγόριθμος back propagation για τη ρύθμιση των παραμέτρων. Αυτό που χρειάζεται πραγματικά είναι να ρυθμιστεί κατά την εκπαίδευση η παράμετρος θ ώστε να ταιριάζει με ένα δοσμένο δείγμα και να αποκτηθεί η βέλτιστη τιμή της θ . Η διαφορά ανάμεσα στις δύο πιθανές κατανομές υπολογίζεται από την απόσταση Kullback-Leibler. Εκφράζεται ως $KL(P||P')$ και ο τύπος είναι:

$$CD_n = KL(P^0 || P_\theta^\infty) - KL(P_\theta^n || P_\theta^\infty)$$

Όπου η απόσταση θα τείνει στο μηδέν μετά από k φορές υπολογισμών της κλίσης και διόρθωσης της παραμέτρου θ . Ο τύπος ενημέρωσης του διανύσματος βαρών weight vector $\theta = (W, b, c)$ είναι:

$$\theta_{i+1} = \theta_i + u \frac{\partial \log_p(v, h)}{\partial \theta} \theta_i$$

Το διάνυσμα χαρακτηριστικών αφαιρούμενο από τον RBM γίνεται είσοδος στον αλγόριθμο

back propagation με τη φόρμουλα $X_j^k = \left| \sum_i W_{ij}^{k-1} X_i^{k-1} \right|$ και παράγει τη

$$X_j^{k'} = \begin{cases} X_j^k & X_j^k \geq \varphi \\ 0 & X_j^k < \varphi \end{cases}$$

φόρμουλα ρύθμισης των παραμέτρων

Η σύγκριση της αρχικής ετικέτας κλάσης με το υπολογισμένο σφάλμα ανά στρώμα νευρώνων

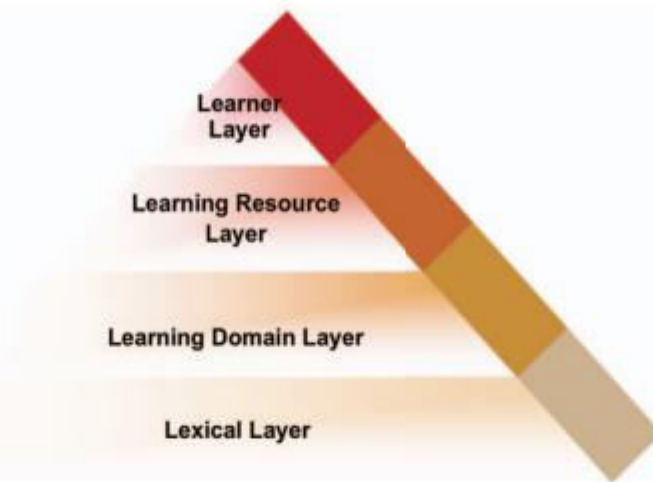
διαμέσου της φόρμουλας βαρών $W'_{ij} = W_{ij} + e_i$ ενημερώνει αντίστροφα τα βάρη για κάθε στρώμα νευρώνων.

Ένα καλά εκπαιδευμένο μοντέλο DBNCF δίνει συστάσεις για μαθήματα στον χρήστη, χρησιμοποιώντας τυχαία συνάρτηση για την εξαγωγή μερικών συνόλων δεδομένων δοκιμής. Κάθε διάνυσμα χαρακτηριστικών μαθημάτων-χρήστη παράγει ένα επίπεδο ταξινόμησης (ένα σκορ) μέσω δοκιμών. Στη συνέχεια, υπάρχει η υψηλή προς χαμηλή ταξινόμηση σύμφωνα με τις προβλεπόμενες βαθμολογίες, αφαιρώντας τα επιλεγμένα από τον χρήστη μαθήματα και ορισμένα μαθήματα κακής ποιότητας, τα οποία προτείνονται στον χρήστη μετά την ταξινόμηση [22].

Η έρευνα των [22] θέτει τα βαθιά νευρωνικά δίκτυα για πρώτη φορά στα εξατομικευμένα ΣΣ για μαθήματα και καταλήγει στο συμπέρασμα πως ο αλγόριθμος έχει καλή απόδοση στην κατηγοριοποίηση προβλέψεων. Σε συνδυασμό δε με παραδοσιακές μεθόδους συστάσεων η αποτελεσματικότητα του αλγορίθμου είναι αξιοσημείωτη [22].

4.4 ΣΣ με χρήση οντολογιών

Η έρευνα των [23] προτείνει μία οργάνωση, των μετά-δεδομένων που σχετίζονται με διάφορες πλευρές της εκπαιδευτικής διαδικασίας, σε έναν αριθμό μαθησιακών οντολογιών προκειμένου να επιτευχθεί καλύτερη διαχείριση των διαδικασιών αυτών. Στην πυραμίδα αυτή, που σχηματικά φαίνεται στην παρακάτω εικόνα:



Εικόνα 11. Η πυραμίδα των μαθησιακών οντολογιών.

τα κατώτερα επίπεδα αναπαριστούν περισσότερο γενικού σκοπού οντολογίες, ενώ οι οντολογίες στα ανώτερα επίπεδα είναι προσαρμοσμένες σε συγκεκριμένες χρήσεις σε ένα περιβάλλον μάθησης όπως είναι τα προσωπικά μαθησιακά περιβάλλοντα (Personal Learning Environments-PLEs) ή τα περιβάλλοντα μάθησης στο σύννεφο (Cloud Learning Environments-CLEs). Καθώς προχωρά κανείς από τα κατώτερα στα ανώτερα στρώματα οντολογιών, φαίνεται πως πραγματοποιείται επαναχρησιμοποίηση και επέκταση των εννοιών από κάθε κομμάτι της πυραμίδας. Οι οντολογίες στην κορυφή σχηματοποιούν το προφίλ του χρήστη που εμπλέκεται στην εκπαιδευτική διαδικασία σύμφωνα με τα ενδιαφέροντα, τις επιδιώξεις, τις προτιμήσεις και τις δεξιότητες. Το επόμενο επίπεδο μοντελοποιεί τους πόρους που χρησιμοποιεί ένα PLE ή CLE. Στο σχηματισμό των οντολογιών χρησιμοποιούνται ετικέτες (tags) των χρηστών ή σχόλια που παράγονται αυτόματα κατά τη διαδικασία εξαγωγής πληροφοριών ή με τεχνικές επεξεργασίας φυσικής γλώσσας. Ένα πρότυπο που χρησιμοποιείται τόσο στο προφίλ χρήστη, όσο και στους εκπαιδευτικούς πόρους είναι το IEEE Learning Objects Metadata Standard (LOM)¹ το οποίο ορίζει πρότυπα για μαθησιακά αντικείμενα, πολυμεσικό περιεχόμενο, εκπαιδευτικό λογισμικό κλπ.

Στο επόμενο επίπεδο περιγράφεται γενικά ένας τομέας επιστημονικού ενδιαφέροντος και τέλος η γενικότερη μορφή οντολογιών είναι η οντολογία λεξικό όπου πρακτικά περιγράφεται οποιαδήποτε έννοια μπορεί να έχει ερευνητικό ενδιαφέρον και γλωσσολογικοί και άλλοι κανόνες αφορούν αυτού του είδους τις οντολογίες, κάτι το οποίο δεν συμβαίνει στα άλλα επίπεδα.

¹ http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

Οι οντολογίες αποτελούνται από έννοιες που συνδέονται μεταξύ τους με σχέσεις είτε εσωτερικές του ίδιου επιπέδου, είτε σχέσεις ανάμεσα στα επίπεδα. Η χαρτογράφηση των οντολογιών δείχνει ποιες σχέσεις είναι σημαντικές και μπορούν να οδηγήσουν συστάσεις προς τον χρήστη για κάποιο μαθησιακό αντικείμενο είτε προτάσεις ποιιοι χρήστες θα μπορούσαν να συνδεθούν μαθησιακά μέσω των προφίλ τους με σκοπό την συνεργασία ή αλληλοβοήθεια. Τέλος, εξίσου σημαντική είναι η ενοποίηση πληροφοριών στις οντολογίες τομέα καθώς μπορεί να οδηγήσει σε έναν διευρυμένο σημασιολογικό ιστό.

Οι οντολογίες αποτελούν εργαλείο βελτίωσης της διαχείρισης των πληροφοριών μέσα σε ένα διαδικτυακό μαθησιακό περιβάλλον και να δώσουν συστάσεις με νόημα στους εκπαιδευόμενους σε έναν κατά τα άλλα αχανή χώρο.

4.5 ΣΣ με τεχνικές εξόρυξης δεδομένων (Data mining)

4.5.1 Σύστημα BBIRS

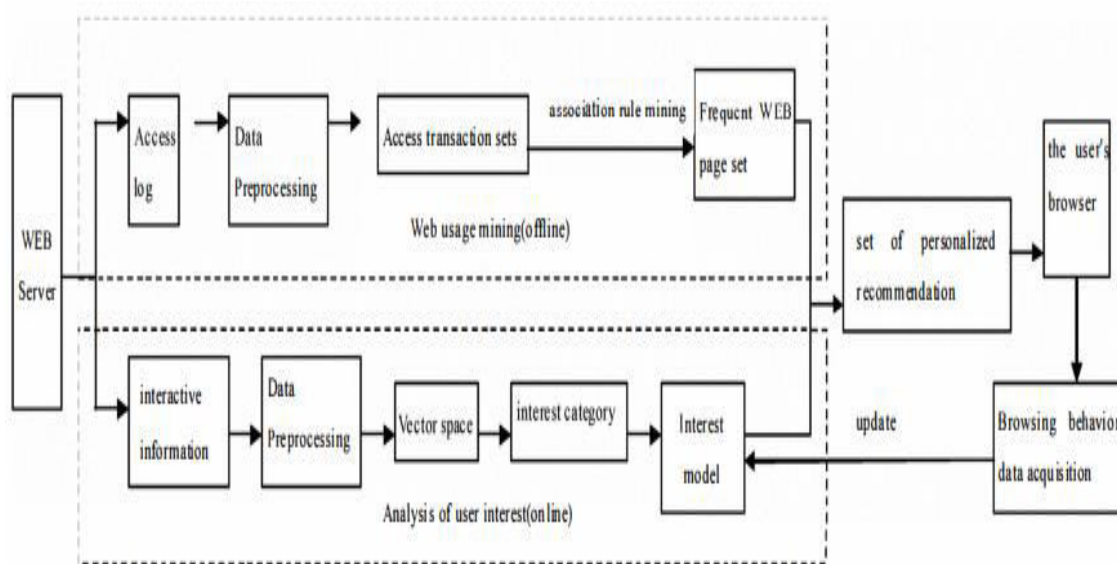
Στο σύστημα BBIRS (Browsing Behavior Personalized Information Recommendation System) που αποτελεί έρευνα των [3] παρουσιάζεται μία στρατηγική προσωποποιημένων συστάσεων με ανάλυση της συμπεριφοράς του χρήστη κατά την περιήγησή του στο Διαδίκτυο. Εξορύσσονται τα αρχεία καταγραφής ιστού και άλλα διαδραστικά δεδομένα με ή χωρίς σύνδεση στο Διαδίκτυο και υπολογίζεται ο βαθμός ενδιαφέροντος σε μία σελίδα από την ταχύτητα περιήγησης. Χρησιμοποιούνται 2 αλγόριθμοι παραγωγής συστάσεων ανάλογα με το βαθμό ενδιαφέροντος του χρήστη.

Το σύστημα χωρίζεται σε δύο κομμάτια: με σύνδεση και χωρίς σύνδεση στο Διαδίκτυο. Χρησιμοποιεί εξόρυξη πληροφοριών από τη χρήση του Ιστού, αναλύει τη συμπεριφορά του χρήστη και παράγει συστάσεις. Κύριο καθήκον στη φάση της εξόρυξης πληροφορίας είναι η προ-επεξεργασία των δεδομένων ώστε να παραχθούν σύνολα συναλλαγών από αναγνωρισμένους χρήστες μέσα από ένα τεράστιο αριθμό διαδικτυακών καταγραφών. Τα σύνολα συναλλαγών γίνονται το αντικείμενο επεξεργασίας διαφόρων μεθόδων εξόρυξης πληροφορίας. Για παράδειγμα, ο αλγόριθμος ομαδοποίησης μπορεί να χρησιμοποιηθεί για πρόσβαση σε υπηρεσίες στο σύμπλεγμα και ο αλγόριθμος παραγωγής συσχετιστικών κανόνων που βασίζεται στο δέντρο αποφάσεων μπορεί να χρησιμοποιηθεί για την επεξεργασία των συνόλων συναλλαγών.

Η συμπεριφορά που επιδεικνύει ο χρήστης κατά την περιήγηση καθώς και το περιεχόμενο των σελίδων που επισκέπτεται και η συχνότητα επισκέψεων παράγουν ένα σύνολο προτεινόμενων σελίδων που μπορεί να προσαρμόζεται στις αλλαγές ενδιαφέροντος του

χρήστη. Κύριο χαρακτηριστικό του συστήματος όπως φαίνεται στην εικόνα 12 είναι οι παράλληλοι δρόμοι επεξεργασίας των δεδομένων με αλγορίθμους εξόρυξης της πληροφορίας. Διαφορετική επεξεργασία για τα αρχεία καταγραφής και άλλη για τα δεδομένα συναλλαγών. Αυτό αποσκοπεί στην απλότητα του συστήματος και ταυτόχρονα αντιμετωπίζει ανεξάρτητα τις διαφορετικές πηγές δεδομένων.

Επιπλέον, το σύστημα υιοθετεί την ανατροφοδότηση του χρήστη αφού ξεκινά από τον χρήστη και επιστρέφει σε αυτόν με στόχο την προσαρμογή του συστήματος στις όποιες αλλαγές ενδιαφέροντος, είναι δηλαδή επικεντρωμένο στο χρήστη και όχι σε ένα σύνολο χρηστών. Από παθητικό το σύστημα γίνεται ενεργητικό ώστε να επιτυγχάνεται όσο το δυνατόν περισσότερο προσωποποιημένο αποτέλεσμα συστάσεων και ο χρήστης να έχει την αίσθηση πως η εκπαιδευτική διαδικασία αφορά αποκλειστικά τον ίδιο.



Εικόνα 12. ΣΣ Προσωποποιημένων πληροφοριών συμπεριφοράς περιήγησης (BBIRS).

Οι πληροφορίες που σχετίζονται με τη συμπεριφορά του χρήστη είναι ποικίλες και δίνονται από διαφορετικές πηγές όπως ο χρόνος διατήρησης μιας σελίδας, ο αριθμός επισκέψεων που αφορούν την συμπεριφορά περιήγησης αλλά και πληροφορίες όπως αναθέσεις εργασιών, εξετάσεις, επικοινωνία με τον καθηγητή ή κάποιο συμφοιτητή που αφορούν την αλληλεπίδραση με το εξ αποστάσεως εκπαιδευτικό σύστημα και άρα δεν μπορεί οι ίδιες μέθοδοι να είναι κατάλληλες για τη διαχείριση αυτών των δεδομένων. Οι αλλαγές στο ενδιαφέρον του χρήστη αντικατοπτρίζονται καλύτερα από την ταχύτητα κατά την περιήγηση.

Έστω ότι ο χρήστης περιηγείται N αριθμό σελίδων σε χρονικό διάστημα b_1, b_2, \dots, b_N . Καθώς ο χρήστης μπορεί να επισκέπτεται την ίδια σελίδα περισσότερες από μία φορές $n \leq N$. Οι επισκέψεις στις σελίδες στο ίδιο χρονικό διάστημα εκφράζονται ως p_1, p_2, \dots, p_n , η συχνότητα επισκέψεων για μία σελίδα p_i ($i=1, 2, \dots, n$) ως $Fre(p_i)$, ενώ η διάρκεια επίσκεψης

για αυτή τη σελίδα $Duration(p_i)$. Ο υπολογισμός του ενδιαφέροντος με βάση την ταχύτητα περιήγησης δίνεται από τον τύπο:

$$Interest(p) = \frac{1 / Speed(p)}{\max_{v \in P} (1 / Speed(v))}$$

Όπου η ταχύτητα ($Speed$) υπολογίζεται από τον τύπο

$$Speed(p) = \frac{Size(p)}{Duration(p)}$$

ορίζεται ως ο αριθμός των χαρακτήρων που θεάται ο χρήστης στη μονάδα του χρόνου.

Ο αλγόριθμος του συνοψίζεται στο παρακάτω σχήμα:

Set a threshold as α

If $(Interest(p) > \alpha)$ then

 Update the users interest database;

 Recommended by way of analyzing the user interest;

Else

 Recommended by WEB usage mining methods

End

Αξίζει να σημειωθεί ότι ο χρόνος που εκφράζει πραγματικό ενδιαφέρον είναι μεγαλύτερης διάρκειας από το κατώφλι α ($Interest(p) > \alpha$). Αλλά αυτό συμβαίνει σε δύο περιπτώσεις, είτε όταν απασχολεί πραγματικά τον χρήστη η σελίδα, είτε όταν συναντά το περιεχόμενο για πρώτη φορά οπότε και το διερευνά, όμως τότε οι φορές πρόσβασης είναι λιγότερες οπότε πρέπει να εξεταστεί και το αρχείο καταγραφής. Το τελικό συμπέρασμα είναι πως θα πρέπει να ενημερώνονται ταυτόχρονα η ανάλυση ενδιαφέροντος του χρήστη και η ΒΔ όπου αποθηκεύεται το ενδιαφέρον που έχει επιδειξει ο χρήστης. Με αυτόν τον τρόπο διασφαλίζεται η προσαρμοστικότητα του συστήματος σε τυχόν αλλαγές στην διάθεση και συμπεριφορά του χρήστη.

4.5.2 ΣΣ με χρήση του αλγορίθμου *Apriori*.

Μία άλλη έρευνα που παρουσιάζει ενδιαφέρον είναι η εργασία των [20] που προτείνει ένα σύστημα συστάσεων για τη δημιουργία λίστας με εκπαιδευτικά βίντεο και χρησιμοποιεί τον *Apriori* αλγόριθμο δημιουργίας συσχετιστικών κανόνων. Ο *Apriori* αλγόριθμος είναι τυπικός του είδους του και οδηγεί τα υποψήφια σύνολα αντικειμένων που απαντούν την ελάχιστη

υποστήριξη να βρουν τα στοιχειοσύνολα με υψηλή συχνότητα και με τη χρήση των συχνών συνόλων θέτει αξιόπιστους κανόνες [20]. Ο πυρήνας του αλγορίθμου στηρίζεται στην ιδέα των συχνών επαναλήψεων.

Μόλις ο αλγόριθμος συναντήσει δύο ανθρώπους με έναν συγκεκριμένο βαθμό συνοχής αμέσως μπορεί να παρουσιάζει στους χρήστες μία λίστα με βίντεο σχετικά με τα ενδιαφέροντά τους. Δύο είναι οι βασικοί στόχοι του προτεινόμενου συστήματος η παροχή της λίστας συστάσεων με βίντεο ανάλογα με τις προτιμήσεις του χρήστη και η γρήγορη ενημέρωση των προτιμήσεων του χρήστη.

Προκειμένου να επιτευχθούν αυτοί οι στόχοι το σύστημα βασίζεται σε τρία δομικά στοιχεία. Ένα περιβάλλον όπου ο χρήστης θέτει το αίτημά του και αυτό μεταφέρεται σε έναν ενδιάμεσο χώρο όπου πραγματοποιείται αναζήτηση τόσο των προτιμήσεων του χρήστη όσο και πολυμεσικού περιεχομένου με κάποια χαρακτηριστικά και τέλος παράγεται η λίστα συστάσεων ως απάντηση στο αίτημα του χρήστη. Ο χρήστης αποφασίζει αν θα δει ή όχι το πολυμεσικό περιεχόμενο αφού διαβάσει μία περιγραφή.

Συλλέγονται προσωπικά στοιχεία του χρήστη όπως φύλο, γνωστικός τομέας, βαθμός, καθώς και προσωπικές εγγραφές του χρήστη στην εκπαιδευτική πλατφόρμα όπως εγγραφή, είσοδος, παρακολούθηση βίντεο με ποιο θέμα, περιγραφή, τύπο και άλλες λειτουργίες. Ο διαχειριστής της ΒΔ με βίντεο είναι υπεύθυνος για την προσθήκη, αλλαγή, διαγραφή χρηστών ή πληροφοριών που αφορούν τα μέσα. Η μηχανή αναζήτησης γρήγορα προσαρμόζει τις παραμέτρους και γεννά συσχετιστικούς κανόνες. Το σύστημα παραγωγής συστάσεων χρησιμοποιεί κυρίως ταξινόμηση μέσω συσχετιστικών κανόνων και συνδυάζει δύο τύπους πληροφοριών το προφίλ του χρήστη και το αρχείο εγγραφών του χρήστη. Όταν το σύστημα έχει πληροφορίες για νέα βίντεο ειδοποιεί τους χρήστες μέσω ηλεκτρονικής αλληλογραφίας με απόσταση ενός μήνα από την τελευταία ενημέρωση [20].

Οι συσχετιστικοί κανόνες ορίζονται ως εξής:

{(user information) – (video information), (click video record)}

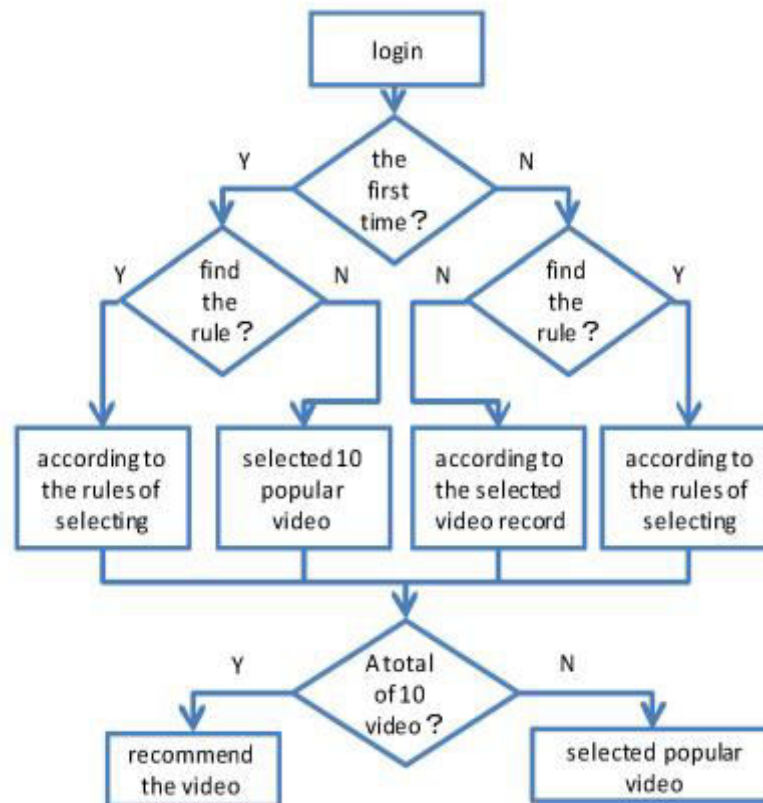
Το αρχείο καταγραφής «χτυπημάτων» στο βίντεο μπορεί να κατηγοριοποιηθεί ανάλογα με το βαθμό των «κλικ». Σε πρώτη φάση ένα απλό κλικ δείχνει μικρό ενδιαφέρον, σε ένα δεύτερο επίπεδο που χαρακτηρίζεται ως προεπισκόπηση ο χρήστης δείχνει ενδιαφέρον για το περιεχόμενο και τέλος με το χαρακτηρισμό «θέαση» ο χρήστης παρακολουθεί όλο το περιεχόμενο και αποτελεί αυτού του είδους το βίντεο στόχο για συστάσεις.

Παράδειγμα πολυδιάστατου συσχετιστικού κανόνα φαίνεται στο παρακάτω σχήμα:

{ (gender, male), (professional, software engineering), (grade one), (degree, undergraduate course)} -- {(categories, application cases), (course, web page design), (quality, HD)}

και ερμηνεύεται ως εξής άνδρες, προπτυχιακοί φοιτητές με εξειδίκευση στη μηχανική λογισμικού παρακολουθούν βίντεο που αφορά τη σχεδίαση ιστοσελίδων.

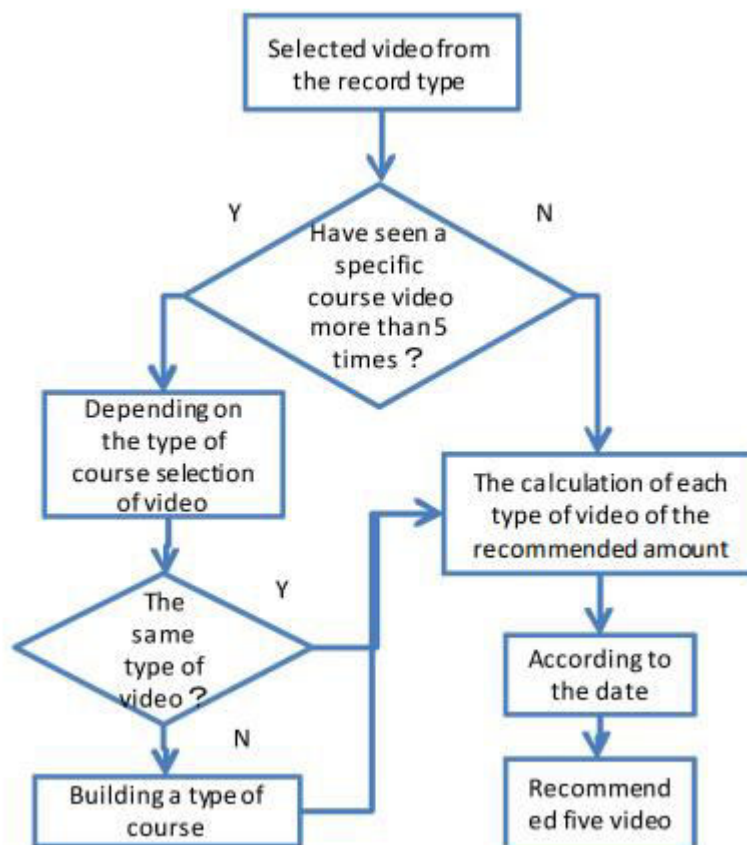
Η προγραμματιστική λογική φαίνεται με ένα διάγραμμα ροής που δείχνει τι συμβαίνει κατά την πρώτη είσοδο του χρήστη και τι τις επόμενες.



Εικόνα 13. Βασικό Διάγραμμα Ροής Συστάσεων Βίντεο

Κατά την πρώτη είσοδο αναζητείται συσχετιστικός κανόνας για τον χρήστη και αν βρεθεί προτείνονται 10 βίντεο βάσει του κανόνα. Αν δεν βρεθεί κανόνας προτείνονται τα 10 πιο δημοφιλή βίντεο. Αν ο χρήστης είναι ήδη εγγεγραμμένος αναζητείται ο κανόνας και αν βρεθεί ο κανόνας αναζητώνται 10 βίντεο για προτάσεις. Αν δεν βρεθούν 10 βίντεο προτείνει τα επιλεγμένα δημοφιλή βίντεο. Αν δεν βρεθεί ο κανόνας αναζητά το αρχείο καταγραφών επιλογών του χρήστη και προτείνει 10 βίντεο.

Η παραγωγή συστάσεων σχηματικά φαίνεται με την παρακάτω εικόνα:



Εικόνα 14. Συστάσεις βασισμένες στις θεάσεις βίντεο του χρήστη

Ελέγχονται τα συμπεράσματα των συσχετιστικών κανόνων στη δεξιά πλευρά. Όταν δίνουν μία απλή πληροφορία το σύστημα προχωρά σε συστάσεις και επιλέγει τα 5 πρώτα βίντεο. Αν οι κανόνες δίνουν πολλαπλές πληροφορίες τότε συνδυάζονται οι πληροφορίες και δίνουν μία λίστα με 10 βίντεο από την οποία τελικά θα επιλεγθούν τα 5 πρώτα.

Οι πειραματικές δοκιμές έδειξαν πως για τα ίδια δεδομένα ο χρόνος παραμένει μάλλον σταθερός. Τα σύνολα συχνών αντικειμένων όπου εμφανίζεται η προσωπικότητα του χρήστη συχνά σε μεγαλύτερα σύνολα, έχει υψηλότερη ακρίβεια στην παροχή συστάσεων βίντεο. Όσο το πείραμα με μεγαλύτερο μήκος χρησιμοποιεί το ελάχιστο όριο για να παράγει σύνολα συχνών αντικειμένων, τόσο μεγαλύτερη είναι και η ακρίβεια στην πρόβλεψη των προτεινόμενων βίντεο.

Η εργασία των [20] παρουσιάζει ένα ευφύες σύστημα συστάσεων βίντεο εκπαιδευτικού περιεχομένου με εξατομικευμένα χαρακτηριστικά ανάλογα με το βαθμό αρεσκείας του χρήστη απέναντι σε μία κατηγορία βίντεο. Παρέχει τις καλύτερες δυνατές μαθησιακές πηγές στους χρήστες, τους δίνει την πρωτοβουλία να χτίσουν το δικό τους μαθησιακό περιεχόμενο και να βελτιώσει την απόδοση συνολικά του εκπαιδευτικού περιβάλλοντος.

4.6 Συγκεντρωτικός πίνακας ΣΣ στην παρούσα εργασία

Έρευνα των	Είδος ΣΣ	Μέθοδος	Κριτήρια (ομοιότητας)	Στόχος	Αποτέλεσμα
[16] Khairil Imran Bin Ghauth et al.	Content based recommendations	Vector Space Model + Average ratings	cosine similarity	Προτείνει παρόμοια μαθήματα με αυτά που βλέπει ο χρήστης αλλά επιπλέον έχουν αξιολογηθεί με υψηλή βαθμολογία από τους καλούς φοιτητές.	Βελτίωση της μαθησιακής διαδικασίας για τον φοιτητή.
[14] Huiyi Tan et al.	Collaborative Filtering	User-based CF	Pearson correlation	Προτείνει μαθήματα με εξατομικευμένο τρόπο υπολογίζοντας την ομοιότητα ανάμεσα σε χρήστες με συσχέτιση προσώπων.	Βελτίωση της μαθησιακής διαδικασίας μέσω εξατομικευμένων προτάσεων.
[21] Sonia Souabi et al.		CF με k-means	elbow method για k=2	Συστάσεις ανά ομάδα χρηστών με κοινή διαδικτυακή συμπεριφορά, που έχουν ταξινομηθεί με τον k-means.	Δίνει σχετικότερες συστάσεις στους φοιτητές.
[1] Qingsong Tu et al.		AOES	$Sim^{iT}(u_a, u_b)$	Προσαρμόζει τις τακτικές μελέτης ανάλογα με την κατάσταση του μαθητευόμενου. Αναλύει πληροφορίες του χρήστη μετά από καταγραφή την ώρα της μελέτης.	Βελτιώνει την αποτελεσματικότητα της Διαδικτυακής σπουδής
[22] Hao Zhang et al.		DBNCF	Kullback-Leibner Distance	Συνδυάζει το CF με τα συστήματα βαθιάς Πίστης	Επιτυγχάνει καλές επιδόσεις στην πρόβλεψη κατανομής και

				των Νευρωνικών Δικτύων.	έχει αξιοσημείωτη αποτελεσματικό- τητα στις συστάσεις.
[23] Alexander Mikroyan nidis et al.	Onto- logy	Οντο- λογίες	IEEE LOM Standard	Χρησιμοποιεί την πυραμίδα μαθησιακών οντολογιών.	Βελτιώνει τον τρόπο με τον οποίο γίνεται μοντελοποίηση και διαχείριση των πληροφοριών σε ένα σύστημα ηλεκτρονικής μάθησης.
[3] Rong Shan et al.	Data mining	BBIRS	συσχετιστικοί κανόνες βάσει δέντρων αποφάσεων	Δημιουργεί ένα σύστημα που ανιχνεύει το ενδιαφέρον του χρήστη με ολιστικό τρόπο.	Παρέχει εξατομικευμένες προτάσεις και προσαρμόζεται στις αλλαγές του ενδιαφέροντος.
[20] Ying Liang et al.		Apriori	συσχετιστικοί κανόνες	Προτείνει εξατομικευμέ- νες πηγές βίντεο. Αναλύει το προφίλ του χρήστη με τον αλγόριθμο εξόρυξης πληροφορίας Apriori.	Ανεβάζει τη διδασκτική αξία της πλατφόρμας καθώς προτείνει αξιόλογες πηγές.

Στην εργασία των [16] προτείνεται ένα σύστημα συστάσεων για ηλεκτρονική μάθηση στο οποίο λαμβάνεται η ομοιότητα των προτάσεων με τα αντικείμενα που έχει ήδη δει ο χρήστης και επίσης, στο σχηματισμό των συστάσεων λαμβάνονται υπόψη οι αξιολογήσεις «καλών» φοιτητών. Οι συστάσεις αυτές χρειάζονται ώστε να επιτευχθεί βελτίωση της μαθησιακής διαδικασίας.

Οι [14] επέλεξαν τον αλγόριθμο συνεργατικού φιλτραρίσματος βάσει χρήστη για να παράξουν συστάσεις καθώς δίνουν θετικά αποτελέσματα και σε άλλα συστήματα συστάσεων όπως πχ. Συστάσεις ταινιών, μουσικής κλπ. Η εγγύτητα ανάμεσα στους χρήστες υπολογίζεται με τον τύπο συσχέτισης του Pearson και από τη γειτονιά που προκύπτει εξάγονται οι 10 πρώτες προτάσεις. Το διαδικτυακό εκπαιδευτικό σύστημα που χρησιμοποιεί αυτού του είδους τις συστάσεις, είναι περισσότερο αλληλεπιδραστικό και προτείνει διάφορα είδη αντικειμένων.

Το άρθρο των [21] χρησιμοποιεί τον αλγόριθμο k-means ως ένα προκαταρκτικό βήμα και αναφέρεται στην συνοχή και συσχέτιση ανάμεσα στις ενέργειες των χρηστών. Η υλοποίηση

με τη χρήση του αλγορίθμου δίνει περισσότερο επαρκή αποτελέσματα από ότι χωρίς. Η αξιολόγηση του συστήματος χρησιμοποίησε τα μεγέθη της ακρίβειας και της ορθότητας και πιστοποίησε πως παράγονται σχετικότερες συστάσεις.

Το σύστημα AOES των [1] βελτιώνει τα υπάρχοντα εκπαιδευτικά διαδικτυακά συστήματα. Χρησιμοποιεί για το σκοπό αυτό ευφυείς τεχνολογίες συστάσεων με στόχο την προσαρμογή του συστήματος στην μαθησιακή κατάσταση του μαθητευόμενου. Αυτό το πετυχαίνει με καταγραφή και ανάλυση των δεδομένων που συλλέγονται κατά τη διάρκεια της μελέτης, αποκτά τη γνωστική ικανότητα του χρήστη και το ενδιαφέρον γύρω από ένα θέμα. Με αυτή τη γνώση δημιουργεί συστήματα συστάσεων για τακτικές μελέτης και εκπαιδευτικές πηγές. Επιπλέον, προσαρμόζει τις προτάσεις στο επίπεδο γνώσεων του χρήστη και ενημερώνεται ταυτόχρονα.

Η εργασία των [22] φέρνει τα Νευρωνικά Δίκτυα Βαθιάς Πίστης στα εξατομικευμένα συστήματα συστάσεων για πρώτη φορά. Χρησιμοποιεί δεδομένα από πραγματική, ανοιχτή πλατφόρμα MOOC που προσφέρει υπηρεσίες μάθησης και επικοινωνίας. Η σύγκριση με άλλους αλγορίθμους δείχνει ότι η ποιότητα συστάσεων με RBM και DBNCF είναι ανώτερη, ενώ η μέθοδος DBNCF είναι ανώτερη όλων σε ταχύτητα σύγκλισης. Το πείραμα δείχνει πως ο προτεινόμενος αλγόριθμος από τους [22] έχει καλύτερο ποσοστό επιτυχίας.

Οι οντολογίες όπως περιγράφονται στους [23] δίνουν σημαντικές ευκαιρίες βελτίωσης του τρόπου που οι πληροφορίες μοντελοποιούνται και διευθετούνται μέσα σε ένα προσωποποιημένο εκπαιδευτικό περιβάλλον. Η διαστρωμάτωση οντολογιών και η ολοκλήρωσή τους προς αυτή την κατεύθυνση, αποσκοπούν στην παροχή συστάσεων στους σπουδαστές με νόημα και στην παροχή βοήθειας στην ανακάλυψη νέων πηγών και ανθρώπων.

Οι παραδοσιακές μέθοδοι προσωποποιημένων τεχνολογιών Διαδικτύου βασίζονται σε στατικές μεθόδους και δεν συμβάλλουν στην αντανάκλαση των αλλαγών στα ενδιαφέροντα περιήγησης του χρήστη. Το σύστημα BBIRS που προτείνουν οι [3] υπολογίζει το βαθμό ενδιαφέροντος του χρήστη για κάποια σελίδα με βάση την ταχύτητα περιήγησης, κατόπιν συγκρίνει το ενδιαφέρον με ένα κατώφλι και με κατάλληλες τεχνικές εξόρυξης πληροφορίας δίνει συστάσεις προσωποποιημένες και με προσαρμογή στις αλλαγές ενδιαφέροντος.

Το σύστημα συστάσεων προσωποποιημένων βίντεο των [20] χρησιμοποιεί την εξόρυξη δεδομένων σε συνδυασμό με το συνεργατικό φιλτράρισμα και προτείνει βίντεο. Θέτει αυτόματα μία κατάσταση αυτό-εκπαίδευσης μέσω διαρκούς μάθησης και ανάλυσης των χαρακτηριστικών θέασης του χρήστη. Προσφέρει τη λειτουργία ενεργών συστάσεων αφού επιτρέπει στο χρήστη να διερευνά τις ενημερωμένες πληροφορίες για τα πολυμέσα ελέγχοντας τις νεότερες πληροφορίες. Ο βαθμός αρεσκείας του χρήστη σε συνδυασμό με τις καλύτερες μαθησιακές πηγές που προσφέρει η πλατφόρμα, βελτιώνουν το ενδιαφέρον του

χρήστη για μάθηση και του δίνουν την πρωτοβουλία να δημιουργήσει το δικό του μαθησιακό περιεχόμενο. Με αυτόν τον τρόπο ανυψώνεται η αποτελεσματικότητα της εκπαιδευτικής πλατφόρμας συνολικά.

4.7 Τρόποι αξιολόγησης ΣΣ

Η αξιολόγηση ΣΣ καλύπτει τρεις τύπους πειραμάτων κατά τους [24]:

- Πειράματα εκτός σύνδεσης που χρησιμοποιούν προεπιλεγμένα δεδομένα ή σύνολα δεδομένων ειδικά φτιαγμένα για δοκιμή της απόδοσης των υποψηφίων αλγορίθμων,
- Μελέτες περίπτωσης, όπου μία μικρή ομάδα χρηστών χρησιμοποιούν το σύστημα σε ένα ελεγχόμενο περιβάλλον και καταθέτουν τις εντυπώσεις τους.
- Δοκιμή στην πραγματική ζωή, όπου το σύστημα εξετάζεται υπό ρεαλιστικές συνθήκες κατά τη διάρκεια της πραγματικής του λειτουργίας από πραγματικούς χρήστες.

Υπάρχουν ορισμένες μετρικές που αποκαλύπτουν τα αποτελέσματα χρήσης ενός συστήματος και κατά πόσον αυτά είναι αποδεκτά από τους χρήστες στους οποίους απευθύνεται το εκάστοτε σύστημα. Η αξιολόγηση χρησιμεύει στο να φανερωθεί τι χρειάζεται να βελτιωθεί πριν το σύστημα αρχίσει να χρησιμοποιείται ευρέως [15]. Σε διαφορετική περίπτωση είναι πολύ πιθανόν το σύστημα να παρουσιάσει διάφορα προβλήματα όπως πχ αργή απόκριση στις απαιτήσεις των χρηστών [15].

Τρεις μετρικές χρησιμοποιούν οι [15] είναι:

- Ακρίβεια (Precision) που αποδίδεται στον τύπο:

$$\text{precision} = \frac{A}{C + A} \times 100\%.$$

Όπου A είναι ο συνολικός αριθμός προτεινόμενων αντικειμένων που άρεσαν στον χρήστη, C ο αριθμός αντικειμένων που δεν άρεσαν και άρα το άθροισμα των A + C αναπαριστά το σύνολο των προτάσεων.

- Ανάκληση (Recall) που ορίζεται ως εξής:

$$\text{Recall} = \frac{A}{B + A} \times 100\%.$$

Όπου το B αναπαριστά τα αντικείμενα που επιλέχθηκαν ως πιθανά για σύσταση αλλά τελικά δεν προτάθηκαν στον χρήστη, ενώ το άθροισμα B+A περιλαμβάνει όλα όσα είναι πιθανόν να αρέσουν στον χρήστη.

Η Ακρίβεια και η ανάκληση είναι μεγέθη αντικρουόμενα. Αν το άριστα στην ανάκληση είναι «1» τότε σημαίνει πως η ακρίβεια θα είναι χαμηλά. Υψηλή ακρίβεια με χαμηλή ανάκληση δίνει εξαιρετικά ακριβείς συστάσεις [15].

- Τρίτη μετρική είναι η F-measure όπως φαίνεται στον τύπο:

$$F - \text{measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Αποτελεί τον αρμονικό μέσο της ακρίβειας και της ανάκλησης και αξιολογεί την ορθότητα και επάρκεια του προτεινόμενου συστήματος συστάσεων.

Στην εργασία των [10] η ακρίβεια για έναν χρήστη u που ανήκει σε ένα σύνολο U αναπαριστάται από τον τύπο:

$$Precision_u = \frac{|L_u \cap B_u|}{|L_u|}$$

Όπου L_u είναι το αποτέλεσμα του αλγορίθμου για τον χρήστη u στο σύνολο εκπαίδευσης και B_u το ρεαλιστικό ενδιαφέρον του χρήστη στο σύνολο δοκιμής. Η αξία του μεγέθους κυμαίνεται μεταξύ 0 και 1. Η ακρίβεια συνολικά του συστήματος φαίνεται με τον τύπο:

$$Precision_{all} = \frac{\sum_{u \in U} |L_u \cap B_u|}{\sum_{u \in U} |L_u|}$$

Ακολουθεί η ανάκληση που μετρά το βαθμό πραγματικού ενδιαφέροντος του χρήστη στην λίστα με προτάσεις και ορίζει την αναλογία ανάμεσα στα ορθώς προτεινόμενα αντικείμενα και στα αντικείμενα για τα οποία υπήρξε αληθινό ενδιαφέρον:

$$Recall_u = \frac{|L_u \cap B_u|}{|B_u|}$$

Αντίστοιχα, η συνολική ανάκληση αποτυπώνεται παρακάτω:

$$Recall_{all} = \frac{\sum_{u \in U} |L_u \cap B_u|}{\sum_{u \in U} |B_u|} \quad \text{N}$$

Τέλος, η μετρική F1-score συνδυάζει την ακρίβεια με την ανάκληση προκειμένου να αξιολογήσει τον αλγόριθμο και υπολογίζεται από τον τύπο:

$$F_1 = \frac{2 \times Precision_{all} \times Recall_{all}}{Precision_{all} + Recall_{all}}$$

Η εργασία των [10] υπολογίζει την ομοιότητα κειμένου με χρήση tf-idf και με αλγόριθμο συνεργατικού φίλτραρίσματος φαίνεται στα πειράματα να πετυχαίνει καλύτερη επίδοση στις τρεις μετρικές που χρησιμοποιεί.

Επίσης, στην εργασία των [2] η ανάκτηση/σύσταση αξιολογείται με 86 ερωτήματα σε διαφορετικούς χρήστες. Και σε αυτή την περίπτωση χρησιμοποιείται η ακρίβεια και η ανάκληση και αποτυπώνονται παρακάτω:

$$Precision = \frac{\text{Number of retrieved documents that are relevant}}{\text{Total number of document retrieved}}$$

$$Recall = \frac{\text{Number of retrieved documents that are relevant}}{\text{Total number of relevant documents}}$$

Η ακρίβεια ορίζεται ως το πηλίκο του αριθμού των σχετικών ανακτημένων κειμένων προς το συνολικό αριθμό ανακτημένων κειμένων ενώ η ανάκληση ως το πηλίκο του αριθμού των σχετικών ανακτημένων κειμένων προς το συνολικό αριθμό σχετικών κειμένων.

Τα πειραματικά δεδομένα που χρησιμοποίησαν οι [19] από μία συλλογή δεδομένων αποτελείται από τέσσερα μέρη: χρήστες, αντικείμενα, βαθμοί-σκορ και χρόνοι. Η κλίμακα βαθμολογίας είναι ένας ακέραιος από 1-5. Ο χρόνος έχει αντίκτυπο στη συγκεκριμένη έρευνα των [19] και επισημαίνεται ταξινομώντας τις βαθμολογίες με αύξουσα χρονική σειρά. Η επίδοση του συστήματος αξιολογείται με το μέσο απόλυτο σφάλμα (Mean Absolute Error – MAE). Το μέγεθος αυτό χρησιμοποιείται στον υπολογισμό του μέσου όρου απόκλισης ανάμεσα στα προβλεφθέντα σκορ και τα πραγματικά σκορ έτσι ώστε να υπολογιστεί η ορθότητα του προβλεφθέντος σκορ. Όσο μικρότερη η αξία του MAE, τόσο μικρότερη η απόκλιση, υψηλότερη η ακρίβεια του συστήματος και άρα καλύτερες προτάσεις για τον χρήστη. Στον παρακάτω τύπο φαίνεται ο συλλογισμός των [19]:

$$MAE = \frac{\sum_{u,i \in T} |R_{ui} - \hat{R}_{ui}|}{|T|}$$

Όπου T ο συνολικός αριθμός αντικειμένων στο σύνολο δοκιμής και \hat{R}_{ui} είναι το πραγματικό σκορ του χρήστη u στο αντικείμενο i .

Στη διαδικασία υπολογισμού ομοιότητας η παράμετρος χρόνος-ενδιαφέρον εισάγεται ως βάρος και η αξία του MAE αυξάνεται καθώς αυξάνεται αυτό το βάρος.

Το MAE είναι μια τεχνική για τη σύγκριση των αποκλίσεων μεταξύ των αποτελεσμάτων μιας πρόβλεψης με πραγματικές αξιολογήσεις από μαθητές ενός μαθησιακού περιεχομένου. Όσο μικρότερο είναι το MAE που παράγεται τόσο χειρότερη θα είναι η επίδοση του συστήματος [25].

5

Σύγχρονες τάσεις στα ΣΣ

5.1 Η έννοια της «ευτυχούς συγκυρίας» -*serendipity* στα ΣΣ

και το πρόβλημα της «φουσαλίδας φίλτρου»- filter bubble.

Οι αλγόριθμοι συνεργατικού φιλτραρίσματος ακόμα και αυτοί που περιέχουν αναδρομικά νευρωνικά δίκτυα (RNN-Recurrent Neural Networks) τείνουν να προβλέπουν συνεχώς σύμφωνα με προηγούμενες συμπεριφορές του χρήστη. Το γεγονός αυτό οδηγεί σε περιορισμό των συστάσεων, σε ένα μικρό σύνολο προτάσεων στο οποίο απουσιάζει η έννοια της «ευτυχούς συγκυρίας» και τοποθετείται ο χρήστης χωρίς να το επιθυμεί σε ένα πλαίσιο απομόνωσης από άλλες προτάσεις γνωστό ως «φουσαλίδα φίλτρου». Τα ζητήματα αυτά πραγματεύεται ο σχεδιασμός των [26] για ένα ΣΣ Δημοσίου Πανεπιστημίου.

Όταν ένας φοιτητής καλείται να επιλέξει μαθήματα από έναν εξαντλητικά μεγάλο κατάλογο, στόχος των ΣΣ είναι να φέρει στην επιφάνεια μαθήματα χρήσιμα για τον σπουδαστή που διαφορετικά δεν θα τα έβλεπε [27]. Σκοπός είναι να προταθούν μαθήματα νέα, αναπάντεχα που όμως είναι σχετικά με τις ανάγκες του φοιτητή. Η φύση αυτού του τύπου συστάσεων έχει περιγραφεί με τον όρο στα αγγλικά *serendipity* ο οποίος στα ελληνικά δεν έχει ακριβή απόδοση και αποδίδεται με τη φράση «ευτυχής συγκυρία» καθώς θέλει να εκφράσει το στοιχείο της έκπληξης, του μη αναμενόμενου αλλά επιτυχούς και επιθυμητού από τον χρήστη [27]. Η μεγιστοποίηση και των δυο στοιχείων που αποτελούν την “*serendipity*” είναι δύσκολο να επιτευχθεί αν αναλογιστεί κανείς πως μαθήματα σχετικά μπορεί να προσφέρονται και από άλλα τμήματα του ίδιου πανεπιστημίου αλλά οι περιγραφές στους καταλόγους να διαφέρουν σημαντικά και έτσι να μη λαμβάνονται υπόψη και να μην υιοθετούνται από τον φοιτητή-χρήστη του ΣΣ [27].

Στόχος των [26] είναι να προτείνουν ένα ΣΣ που θα δίνει αποτελέσματα απροσδόκητα αλλά εξίσου σχετικά με τα ενδιαφέροντα του χρήστη. Για το σκοπό αυτό χρησιμοποιούν δύο μοντέλα: ένα με περιγραφές καταλόγου μαθημάτων (BOW-Bag of Words) και ένα μοντέλο που χρησιμοποιεί το ιστορικό εγγραφών (course2vec) και τα συγκρίνουν. Τα αποτελέσματα της εργασίας έδειξαν έλλειψη πρωτοτυπίας στα RNN συστήματα και με δυσκολία κατορθώνουν να δώσουν την αίσθηση της «ευτυχούς συγκυρίας». Ενώ, τα μοντέλα μηχανικής μάθησης course2vec απέδωσαν καλύτερα στα καθήκοντα αξιολόγησης εκτός σύνδεσης Διαδικτύου. Οι φοιτητές αξιολόγησαν συστάσεις με απλές δέσμες λέξεων ως περισσότερο συμπτωματικές.

Ανάμεσα στις αξίες ενός ακαδημαϊκού ιδρύματος είναι και η έκθεση των φοιτητών σε έναν πλουραλισμό ιδεών και προοπτικών μέσω των μαθημάτων. Τα υπάρχοντα μοντέλα δίνουν εξατομικευμένες συστάσεις με βάση συμπεριφορές του χρήστη σε ένα χρονικό πλαίσιο. Δεν ανεβάζουν στην επιφάνεια μαθήματα που μπορεί να είναι ενδιαφέροντα για τον χρήστη, διότι δεν έχουν εξερευνηθεί από πολλούς χρήστες με επιλογές παρόμοιων μαθημάτων στο παρελθόν. Χρειάζεται λοιπόν, ένα σύστημα που να προάγει την ποικιλομορφία και τη διαφοροποίηση σε ότι αφορά τις συστάσεις. Στην αντίθετη περίπτωση το σύστημα, ιδιαίτερα στα κοινωνικά δίκτυα, οδηγείται στο πρόβλημα που είναι γνωστό ως «φυσαλίδα φίλτρου», μία κατάσταση δηλαδή περιορισμού των προτάσεων σε στερεοτυπικές επαναλήψεις. Ο φοιτητής οδηγείται σε ένα μονοπάτι γνώσης από το οποίο δεν μπορεί να ξεφύγει. Ο σχεδιαστής ενός ΣΣ θα πρέπει να έχει επίγνωση αυτών των προβλημάτων, να προνοεί και να σχεδιάζει ανάλογα [26].

Οι [26] ως αντίδραση στη «φυσαλίδα φίλτρου» εισάγουν μία παραλλαγή του course2vec στην παραγωγή ΣΣ σχεδιάζοντας ένα σύστημα που να ανακαλύπτει απροσδόκητες συστάσεις. Ο course2vec εφαρμόζει μία τεχνική μάθησης χωρίς επίβλεψη που βρίσκει τις πιο κοντινές λέξεις μιας δοσμένης λέξης και ονομάζεται skip-gram. Η τεχνική αυτή προβλέπει τις λέξεις που πλαισιώνουν την λέξη-στόχο και χρησιμοποιήθηκε στα ιστορικά εγγραφών αντί για φυσική γλώσσα στην αναπαράσταση μαθημάτων. Ο φοιτητής εκφράζει ενδιαφέρον σε μία πρόταση που ξεφεύγει από τα τετριμμένα. Συμβαίνει συχνά πολλά πανεπιστημιακά τμήματα να περιλαμβάνουν μαθήματα με παρόμοια εννοιολογικά θέματα αλλά επειδή χρησιμοποιούν διαφορετικές περιγραφές να μην κατανοούν οι φοιτητές πόσο κοινά είναι αυτά μεταξύ τους. Ένα καλά ρυθμισμένο διάνυμα αναπαράστασης των μαθημάτων είναι ικανό να συλλάβει περισσότερη έμμεση σημασιολογία των μαθημάτων και να δομήσει ομοιότητες μαθημάτων με μεγαλύτερη ακρίβεια. Η ενσωμάτωση γενικότερων χαρακτηριστικών των μαθημάτων από διάφορα τμήματα επιτυγχάνεται με την τεχνική skip-gram καθώς μαθαίνει ταυτόχρονα και από τα τμήματα και από τους εκπαιδευτές. Η προσέγγιση των [26] προσθέτει κύρος

ρυθμίζοντας το μοντέλο έναντι συνόλων επικύρωσης δημιουργημένα με επιστημονική γνώση και επιμέλεια του πανεπιστημίου.

Η σύγκριση των μοντέλων οδήγησε στο συμπέρασμα πως το σύστημα RNN είναι ανεπαρκές σε τύχη μαζί με επιτυχία, δηλαδή, «ευτυχή συγκυρία». Το μοντέλο `course2vec` επιτυγχάνει ουσιαστικά παραπάνω σε ακρίβεια αλλά το μοντέλο BOW (δέσμες λέξεων) επιτυγχάνει καλύτερα αποτελέσματα σε συμπτώσεις («ευτυχείς συγκυρίες»). Ένας πιθανός λόγος για αυτή την ασυμφωνία είναι η φύση των πληροφοριών που παρουσιάζονται στους φοιτητές. Στους φοιτητές παρουσιάζονται περιγραφές καταλόγου των μαθημάτων και αποτελούσαν την μόνη πηγή πληροφόρησης ως προς την επιλογή μαθημάτων. Η μεταφορά λανθανόντων χαρακτηριστικών από τη μηχανή μάθησης μπορεί να είναι αναγκαία στην πραγμάτωση οφέλους. Σε σχετικές εργασίες αναφέρεται πως μία μικρή μείωση στην ακρίβεια μπορεί να δώσει καλά αποτελέσματα σε «ευτυχείς συγκυρίες» [26].

5.1.1 Μοντέλα και Μεθοδολογία

Για να παραχθούν οι αναπαραστάσεις στην έρευνα των [26] χρησιμοποιήθηκαν τα παρακάτω ανταγωνιστικά μοντέλα:

- `Course2vec` για μάθηση αναπαραστάσεων μαθημάτων από προτάσεις εγγραφών. Στα γλωσσικά μοντέλα τα διανύσματα δύο λέξεων θα είναι όμοια αν μοιράζονται παρόμοιες προτάσεις νοηματικά. Στον τομέα ενός Πανεπιστημίου, μαθήματα με κοινές εγγραφές είναι πιθανότερο να είναι κοντά στον διανυσματικό χώρο.
- Παραλλαγή `course2vec` που μαθαίνει από σαφώς καθορισμένα χαρακτηριστικά (καθηγητής και τμήμα) εκτός από τις αναπαραστάσεις μαθημάτων. Η επιρροή των καθηγητών που έχουν διδάξει ένα μάθημα κατά καιρούς μπορεί να ενυπάρχει αναμειγμένη στις αναπαραστάσεις των μαθημάτων. Πρόκειται για ένα ενισχυμένο μοντέλο αναπαραστάσεων που περιλαμβάνει πολλές παραμέτρους. Ο αντικειμενικός σκοπός ενός `skip-gram` μοντέλου είναι να βρίσκει λέξεις που είναι χρήσιμες στην πρόβλεψη γειτονικών λέξεων σε μία πρόταση ή έγγραφο. Λέξεις μέσα σε ένα συγκεκριμένο εύρος, πριν και μετά την τρέχουσα λέξη, διαμορφώνουν ένα γλωσσικό, νοηματικό πλαίσιο.
- Ένα διάνυσμα BOW αυστηρά δομημένο από τις περιγραφές καταλόγου. Κάθε κείμενο αναπτύσσεται σε ένα διάνυσμα πραγματικών αριθμών όπου κάθε αριθμός αναπαριστά το βάρος κάθε όρου, πχ βάρος θα μπορούσε να λογίζεται η συχνότητα εμφάνισης μίας λέξης σε ένα έγγραφο ή η δυαδική τιμή 1 για ύπαρξη τη λέξης και 0 για απουσία της ή ένα σχήμα `tf-idf` που διαχωρίζει τις συχνά εμφανιζόμενες λέξεις σε ένα κείμενο από τις συχνά εμφανιζόμενες λέξεις σε οποιοδήποτε κείμενο.

-
- Συνδυασμός αναπαράστασης `course2vec` και διανύσματος BOW για καλύτερα αποτελέσματα.

5.1.2 Ανάδυση συστάσεων «ευτυχούς συγκυρίας» από αναπαραστάσεις μαθημάτων

Οι συστάσεις χρειάζεται να είναι ενδιαφέρουσες αλλά απροσδόκητες. Βρίσκεται ένα μάθημα c_j που παρουσιάζει τις περισσότερες ομοιότητες με το αγαπημένο μάθημα c_i ενός φοιτητή. Για κάθε τμήμα d_j επιτρέπεται ένα μόνο αποτέλεσμα, για διαφοροποίηση των αποτελεσμάτων και αποδίδεται από τον τύπο:

$$c_j^* = \arg \max_{c, d(c)=d_j} \cos(c, c_i)$$

Όπου $d(c)$ σημαίνει ότι το τμήμα d με το μάθημα c . Τότε όλα τα αντίστοιχα μαθήματα c_j^* σε όλα τα τμήματα ταξινομούνται σύμφωνα με το συνημίτονο $\cos(c_j^*, c_i)$ και $j=1,2,\dots, k$. Μπορούν να χρησιμοποιηθούν και νευρωνικές αναπαραστάσεις και δέσμες λέξεων σε αυτή τη μέθοδο για την αναπαραγωγή παρόμοιων μαθημάτων σε κάθε τμήμα.

5.1.3 Συμπεράσματα της έρευνας

Η ενδεδειγμένη έρευνα των [26] σε ένα μεγάλο πανεπιστημιακό ίδρυμα οδήγησε σε κάποια συμπεράσματα. Η ανάδειξη μαθημάτων που ενδιαφέρουν τους φοιτητές αλλά δεν το γνωρίζουν, επεκτείνει τις γνώσεις των φοιτητών και την κατανόηση όσων προσφέρει το Πανεπιστήμιο. Όσο περισσότερο εκτίθενται σε μαθήματα που απομακρύνονται από το τμήμα τους και το πλέγμα κατανόησής τους, τόσο λιγότερο οικείες με περιγραφές γίνονται οι συστάσεις και γίνεται δύσκολο να καταφέρει κανείς να συνδεθεί με τέτοιου είδους συστάσεις. Γεγονός που υπογραμμίζει τη δυσκολία παραγωγής συστάσεων με χαρακτήρα έκπληξης και συνάφειας ταυτόχρονα. Στο ερώτημα που τέθηκε στους φοιτητές αν έχει νόημα να τους προτείνονται μαθήματα για τα οποία είναι ήδη ενημερωμένοι, δύο είδη απαντήσεων δόθηκαν. Μία απάντηση που προκρίνει την ποικιλομορφία και τη διαφοροποίηση και μία άλλη που προτιμά την περπατημένη οδό. Το συμπέρασμα που εξάγεται είναι πως, ανάλογα την περίπτωση του φοιτητή ή του αντικειμένου σπουδών, και οι δύο περιπτώσεις είναι σημαντικές. Όταν ο πλουραλισμός και η διερεύνηση είναι προτεραιότητες τότε χρειάζεται βελτιστοποίηση της «ευτυχούς συγκυρίας» και ανακάλυψη νέων μαθημάτων χωρίς στερεοτυπικές προκαταλήψεις που πηγάζουν από το ιστορικό και την προηγούμενη εμπειρία. Αυτή η προσέγγιση αγγίζει τους φοιτητές που είναι έτοιμοι να φτιάξουν το δικό τους

γνωστικό μονοπάτι. Όταν όμως ο στόχος είναι η προσήλωση στην απόκτηση του βασικού πτυχίου και η ολοκλήρωση των σπουδών με επιτυχία, τότε για τους φοιτητές που χρειάζονται μεγαλύτερη καθοδήγηση, οι συστάσεις με τα υπάρχοντα συστήματα είναι ικανοποιητικές. Ένα μείγμα των δύο οπτικών, δηλαδή, ενθάρρυνση για εξερεύνηση αλλά εντός του πεδίου μαθημάτων, θα ήταν ίσως ιδανική.

5.2 Επεξηγήσεις προτάσεων στα ΣΣ

Μία πρόκληση που φαίνεται να αντιμετωπίζουν τα ΣΣ είναι πώς να εξηγήσει κανείς σε φοιτητές γιατί τους προτάθηκαν συγκεκριμένα μαθήματα και όχι κάποια άλλα [27]. Στο πλαίσιο ενός πανεπιστημίου, είναι πολύτιμο για τα ΣΣ να εισάγουν μαθήματα σε φοιτητές που είναι σχετικά με τα ενδιαφέροντά τους αλλά οι ίδιοι δεν το γνωρίζουν και ως εκ τούτου δεν θα εγγράφονταν στα μαθήματα αυτά. Όσο λιγότερο εξοικειωμένοι είναι με τα μαθήματα, τόσο μικρότερη ικανότητα έχουν να κρίνουν τη σχετικότητα ενός μαθήματος. Η εργασία των [27] πραγματεύεται τρόπους εξοικείωσης των φοιτητών με συστάσεις με τρεις τύπους επεξηγήσεων σχεδιασμένους με διάφορα επίπεδα εξατομίκευσης. Οι φοιτητές αξιολογούν κάθε προτεινόμενο μάθημα σε επίπεδο νεωτερικότητας, έκπληξης και συνάφειας με πρόθεση ο φοιτητής να εγγραφεί, γεγονός που συνιστά επιτυχία της σύστασης. Όταν εξηγούνται οι συστάσεις ενός νέου μαθήματος χρησιμοποιώντας λέξεις κλειδιά από άλλα μαθήματα στα οποία ο φοιτητής συμμετέχει, τότε αυξάνεται η «ευτυχής συγκυρία»(serendipity) του μαθήματος.

Στην έρευνα των [27] γίνεται αποδεκτή η υπόθεση πως η ενσωμάτωση επεξηγήσεων στις συστάσεις μαθημάτων μπορεί να βελτιώσει την «ευτυχή συγκυρία» ή «εύνοια» του συστήματος. Με περισσότερες εξηγήσεις οι φοιτητές εξοπλίζονται με πληροφορίες που τους βοηθούν να κρίνουν τη χρησιμότητα ενός μαθήματος και είναι πιθανότερο να μην το απορρίψουν λόγω άγνοιας του αντικειμένου. Προκειμένου να επιβεβαιωθεί η αρχική υπόθεση των ερευνητών [27] χρησιμοποιήθηκαν τρεις διαφορετικές στρατηγικές. Ένας φοιτητής καλείται να επιλέξει ένα αγαπημένο μάθημα. Μία μηχανή αναζήτησης προτείνει 5 παραπλήσια σύμφωνα με αλγορίθμους αναπαραστάσεων μαθημάτων βάσει περιεχομένου ανάμεσα σε διαφορετικά τμήματα. Οι εξηγήσεις διαφέρουν σε επίπεδο εξατομίκευσης. Οι τρεις στρατηγικές περιγράφονται ακολούθως:

1. Πραγματοποιείται συνένωση προτεινόμενων μαθημάτων με «συναγόμενες λανθάνουσες λέξεις» και λειτουργεί ανεξάρτητα από τον χρήστη.
2. Παράγονται εξατομικευμένες επεξηγήσεις με βάση το επιλεγμένο αγαπημένο μάθημα, επισημαίνοντας τις κοινές λέξεις κλειδιά ανάμεσα στο προτεινόμενο μάθημα και το επιλεγμένο.

-
3. Η περισσότερο προσωποποιημένη στρατηγική καθώς υπερτονίζει τις λέξεις κλειδιά στα προτεινόμενα μαθήματα και σε όσα μαθήματα έχει εγγραφεί ήδη ο φοιτητής. Αυτή η στρατηγική φάνηκε να είναι η περισσότερο ελκυστική για τους φοιτητές καθώς ήταν ήδη εξοικειωμένοι με προηγούμενα μαθήματα και γνώριζαν τις λέξεις κλειδιά, ώστε να προσανατολιστούν σε ένα νέο μάθημα. Στη διάρκεια της μελέτης των [27] αποδείχθηκε η περισσότερο αποτελεσματική, αυξάνοντας στατιστικά σημαντικά, την έκκληξη, τη νεωτερικότητα και άρα την «ευτυχή συγκυρία».

Προηγούμενες προσπάθειες ενσωμάτωσης εξηγήσεων στα ΣΣ έδειξαν πως οι επεξηγήσεις αυξάνουν την αντίληψη της διαφάνειας του συστήματος και την εμπιστοσύνη σε αυτό, που οδηγεί σε αυξημένη κατανόηση της σχετικότητας και αποδοχή των συστάσεων. Η παρούσα εργασία των [27] επεκτείνει προηγούμενες και αξιολογεί πως διαφορετικοί τύποι επεξηγήσεων επηρεάζουν την κρίση των φοιτητών στα προτεινόμενα αντικείμενα.

5.3 Αξιολόγηση του προτεινόμενου συστήματος

Στο τέλος της έρευνας των [27] ζητήθηκε από τους φοιτητές να αξιολογήσουν τις συστάσεις με τρεις προτάσεις:

- Το μάθημα ήταν μη αναμενόμενο.
- Με ενδιαφέρει να γραφτώ στο μάθημα.
- Δεν γνώριζα νωρίτερα σχετικά με αυτό το μάθημα.

Με βάση τις απαντήσεις μετρώνται διαφορετικές διαστάσεις των συστάσεων. Η πρώτη μετρά την έκκληξη, η δεύτερη την επιτυχία, η τρίτη την καινοτομία. Ο συνδυασμός των στοιχείων αυτών κυρίως της μη προσδοκίας μαζί με την επιτυχία αποτελούν την «ευτυχή συγκυρία» του συστήματος.

Οι συμμετέχοντες στην έρευνα των [27] αφιέρωσαν περίπου 10 λεπτά. Ελήφθησαν 329 αξιολογήσεις από 67 φοιτητές όλων των ετών φοίτησης από πρωτοετείς μέχρι τελειόφοιτους. Οι συμμετέχοντες εξέτασαν και τις συστάσεις και τις επεξηγήσεις των συστάσεων. Τα αποτελέσματα των αξιολογήσεων των 10 συνθηκών φαίνονται στον πίνακα 4:

Πίνακας 4. Μέσες βαθμολογίες μαθητών για προτάσεις μεμονωμένων μαθημάτων από τη μελέτη χρήστη που αναλύονται ανά μοντέλο που χρησιμοποιήθηκε για τη δημιουργία προτάσεων μαθημάτων, μέθοδο που χρησιμοποιήθηκε για τη δημιουργία των επεξηγήσεων και δομή βαθμολογίας.

Condition	Model		Explaining Keywords			Average Student Ratings				
	BOW	Analogy	Inferred	Anchored	Taken	Unexpectedness	Successfulness	Serendipity	Novelty	Ratings
1A	●	○	○	○	○	3.571	2.771	3.171	3.714	35
2A	●	○	●	○	○	3.096	2.741	2.919	3.354	31
3A	●	○	○	●	○	3.114	2.771	2.943	3.800	35
4A	●	○	○	○	●	3.700	2.666	3.183	3.900	30
5A	●	○	●	●	●	3.025	2.975	3.000	3.575	40
1B	○	●	○	○	○	3.200	2.320	2.760	3.080	25
2B	○	●	●	○	○	3.244	2.311	2.688	2.644	40
3B	○	●	○	●	○	2.767	3.046	2.907	3.279	43
4B	○	●	○	○	●	3.650	2.950	3.300	3.450	20
5B	○	●	●	●	●	3.500	2.766	3.133	3.533	30

Κάθε γραμμή αναπαριστά τις μέσες αξιολογήσεις αποτελεσμάτων όλων των φοιτητών στους οποίους ανατέθηκε μία συγκεκριμένη συνθήκη. Ο σχεδιασμός του συστήματος έχει επίδραση στις συστάσεις και τις επεξηγήσεις. Το μοντέλο BOW είχε υψηλό αντίκτυπο στην νεωτερικότητα σε αυτή τη μελέτη με μεγάλη διαφορά σε μέσες και διάμεσες αξιολογήσεις από το Αναλογικό μοντέλο. Επιπλέον, πραγματοποιείται ανάλυση για να δοκιμαστούν διαφορές ανάμεσα σε χωρίς-εξηγήσεις και με-εξηγήσεις όπου λαμβάνονται υπόψη και οι τρεις τύποι επεξηγήσεων. Δεν βρέθηκαν ιδιαίτερες διαφορές στα αποτελέσματα, γεγονός που υποδεικνύει ότι διαφορετικοί τύποι επεξηγήσεων επηρεάζουν με διαφορετικούς τρόπους τους χρήστες [27]

Για την καλύτερη κατανόηση της επίδρασης των μοντέλων συστάσεων και των διαφόρων επεξηγήσεων στις μετρήσεις χρησιμοποιήθηκε η μέθοδος ελαχίστων τετραγώνων (Ordinary Least Square-OLS) για κάθε μία από τις μετρήσεις. Τα αποτελέσματα φαίνονται στον πίνακα 5.

Πίνακας 5. Αντίκτυπος των ΣΣ και των στρατηγικών εξηγήσεων σύμφωνα με τις OLS παλινδρομήσεις

Variable	Unexpectedness		Successfulness	
	Coefficient	P-value	Coefficient	P-value
Analogy Model	-0.0759	0.595	-0.1039	0.431
Inferred Keyword	-0.1982	0.180	-0.1293	0.327
Anchored Keyword	-0.3501	0.013 (*)	0.2873	0.037 (*)
Taken Keyword	0.4161	0.006 (**)	0.1182	0.442
Variable	Serendipity		Novelty	
	Coefficient	P-value	Coefficient	P-value
Analogy Model	-0.0899	0.292	-0.5240	0.001 (**)
Inferred Keyword	-0.1637	0.063	-0.4454	0.011 (*)
Anchored Keyword	-0.0314	0.722	0.2275	0.161
Taken Keyword	0.2671	0.005 (**)	0.3701	0.039 (*)

* p-value<0.05, ** p-value<0.01

- Το αναλογικό μοντέλο είχε αρνητική επίδραση στις μετρήσεις. Αυτό σημαίνει πως το μοντέλο BOW δίνει καλύτερα αποτελέσματα ακόμα και αν δεν είναι στατιστικά σημαντικά.
- Η μέθοδος με συναγόμενες λέξεις κλειδιά είχε επίσης, αρνητική επίδραση.
- Η μέθοδος με συνδεδεμένες λέξεις κλειδιά είχε θετικό αντίκτυπο στην επιτυχία αλλά αρνητικό στην έκπληξη.
- Η μέθοδος με ληφθείσες λέξεις κλειδιά είχε θετική επίδραση σε τρία από τα τέσσερα μεγέθη με κυριότερο την serendipity.

Οι μετρήσεις έδειξαν πως ο συνδυασμός των δυο στοιχείων που αποτελούν την serendipity είναι δύσκολο να αυξηθούν ταυτόχρονα. Η μέθοδος που έδωσε τα καλύτερα αποτελέσματα είναι αυτή στην οποία χρησιμοποιήθηκαν λέξεις κλειδιά από όλα τα μαθήματα των προηγούμενων ετών του φοιτητή, με μεγαλύτερο αντίκτυπο στην «ευτυχή συγκυρία».

Συμπερασματικά, μπορεί να ειπωθεί, πως υπάρχει πλέον μία μέθοδος παραγωγής επιτυχών επεξηγήσεων που βελτιώνουν σημαντικά όλες τις επιθυμητές μετρικές του συστήματος και χρειάζεται να αναζητηθούν περαιτέρω οφέλη σε περισσότερους τομείς συστάσεων [27].

6

Επίλογος

Η παρούσα διπλωματική εργασία επιχειρεί να ερευνήσει τη συνεισφορά των Συστημάτων Συστάσεων στην εκπαιδευτική διαδικασία και συγκεκριμένα μέσω της ηλεκτρονικής μάθησης. Εξηγεί συνοπτικά την έννοια και τα χαρακτηριστικά της ηλεκτρονικής μάθησης, κάνει ανασκόπηση των τρόπων αξιοποίησης εκπαιδευτικών συστάσεων μέσω αντιπροσωπευτικών συστημάτων συστάσεων. Τα ΣΣ αποτελούν εργαλεία βελτίωσης της μαθησιακής εμπειρίας και αναπόσπαστο πλέον κομμάτι των μηχανών αναζήτησης κάθε συστήματος παροχής υπηρεσιών και πληροφοριών.

6.1 Σύνοψη και συμπεράσματα

Ο όρος ηλεκτρονική μάθηση είναι ένας γενικός όρος –ομπρέλα, που περιγράφει τη μάθηση που αποκτάται σε έναν υπολογιστή συνδεδεμένο στο Διαδίκτυο και ο οποίος αίρει πολλούς περιορισμούς πρόσβασης στη γνώση. Η ηλεκτρονική μάθηση είναι πλέον ευρέως αποδεκτή και μπορεί να αποτελεί μία πλούσια και πολύτιμη μαθησιακή εμπειρία ενίοτε περισσότερο και από την φυσική παρουσία σε τάξη. Με τα μοναδικά της χαρακτηριστικά οδηγεί σε κατανόηση και επάρκεια γνώσεων ακριβώς όπως και ο παραδοσιακός εταίρος [14].

Η διδασκαλία μέσω των νέων τεχνολογιών έχει τελειοποιηθεί ενσωματώνοντας τις βασικές αρχές διδασκαλίας με πολλά οφέλη για τους σπουδαστές. Αυτό έχει σαν αποτέλεσμα κολλέγια, πανεπιστήμια, επιχειρήσεις και οργανισμοί σε όλο τον κόσμο να προσφέρουν προγράμματα σπουδών επαγγελματικής κατάρτισης και δια βίου μάθησης με πλήρη αναγνώριση και σε αφθονία μέσω Διαδικτύου. Σε πολλές ιστοσελίδες είναι διαθέσιμα αμέτρητα ανεπίσημα μαθήματα που καλύπτουν ένα εύρος θεμάτων συμπεριλαμβανομένης της διοίκησης επιχειρήσεων, πληροφορικής, δεξιοτήτων για τη ζωή, γλώσσες κα σε όλα τα επίπεδα [14].

Τα συστήματα αυτά παρέχουν πλήθος πληροφοριών που δεν είναι εύκολο να προσπελαστούν χωρίς βοήθεια από ένα σύστημα συστάσεων. Τα ΣΣ προσφέρουν εξατομικευμένες πληροφορίες, διευκολύνουν την πρόσβαση του χρήστη σε χρήσιμα μαθήματα, τον απαλλάσσουν από χρονοβόρες και μάταιες αναζητήσεις και βελτιώνουν συνολικά το μαθησιακό αποτέλεσμα για τον σπουδαστή που συμβουλευέται τέτοια συστήματα. Υλοποιούνται με χρήση ευφών τεχνολογιών, με τεχνικές μηχανικής μάθησης και εξόρυξης δεδομένων αναλύονται οι συμπεριφορές των χρηστών και βάσει αυτών και άλλων κριτηρίων που συζητήθηκαν στην εργασία αυτή δίνουν προσωποποιημένες συστάσεις.

Τα Συστήματα Συστάσεων σε ένα εκπαιδευτικό περιβάλλον διαφέρουν από τα συστήματα συστάσεων σε άλλους τομείς (τουρισμού, ηλεκτρονικού εμπορίου κ.α.) διότι παράγονται βάσει των μαθησιακών αναγκών των εκπαιδευομένων. Πρόκειται για πολύ σημαντικά συστήματα που αγκαλιάζουν κάθε πλευρά ενός εκπαιδευτικού ηλεκτρονικού συστήματος και επιπλέον χρησιμοποιούν κάθε σύγχρονη πλευρά των ευφών τεχνολογιών Διαδικτύου. Η ιδιαιτερότητα των εκπαιδευτικών ΣΣ συνοψίζεται στο γεγονός ότι η γνώση είναι δυναμική και μεταβάλλεται με διαφορετικό ρυθμό για τον κάθε εκπαιδευόμενο. Η μοντελοποίηση της γνώσης με διάφορες μεθόδους και η σύσταση κατάλληλου προσωποποιημένου εκπαιδευτικού περιεχομένου μπορεί να βελτιώσει την μαθησιακή επίδοση των εκπαιδευομένων.

6.2 Μελλοντικές επεκτάσεις

Οι σύγχρονες τάσεις στα ΣΣ επιβάλλουν συνεχή διερεύνηση των συστημάτων αυτών και υλοποίηση τεχνικών με διάφορους συνδυασμούς, ώστε να δοθούν όσο το δυνατόν επιτυχέστερες συστάσεις. Επιτυχείς είναι οι συστάσεις που αφορούν πραγματικά τον χρήστη, του δίνουν κατάλληλες συμβουλές, τον συνοδεύουν στο μονοπάτι γνώσης που επιθυμεί να αποκτήσει, χωρίς να τον εγκλωβίζουν σε στερεοτυπικές, επαναλαμβανόμενες προτάσεις δίχως πρωτοτυπία και ώθηση στην καινοτομία. Ο όρος serendipity θα απασχολεί όλο και περισσότερο τα συστήματα αυτά και χρειάζεται περισσότερο να διερευνηθεί καθώς και η παραγωγή επεξηγήσεων ακόμα και με χρήση τεχνητής νοημοσύνης.

Εν κατακλείδι συστήματα όπως τα ΣΣ που χρησιμοποιούν τεχνικές ανάκτησης πληροφοριών, εξόρυξης δεδομένων, οντολογίες, αλγορίθμους μηχανικής μάθησης μέχρι τεχνητή νοημοσύνη, δε μπορεί παρά να απασχολήσουν τους ερευνητές για πολλά χρόνια ακόμα, ιδίως αν αναλογιστεί κανείς πως η εκπαίδευση και το ταξίδι στη γνώση δεν τελειώνουν ποτέ.

7

Βιβλιογραφία

- [1]. Qingsong Tu and Jian Liu, “Research on autonomous online education system based on intelligent recommendation,” in *2011 IEEE International Symposium on IT in Medicine and Education*, Guangzhou, China, 2011, pp. 410–413. doi: 10.1109/ITiME.2011.6132136
- [2]. R. Ponnusamy and T. V. Gopal, “A User-Adaptive Self-Proclamative Multi-Agent Based Recommendation System Design for E-Learning Digital Libraries,” in *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Jun. 2006, pp. 1–7. doi: 10.1109/ICCIS.2006.252279
- [3]. R. Shan and Z. Ren, “Research on personalized recommendation system in E-learning,” in *2010 2nd International Conference on Education Technology and Computer*, Jun. 2010, vol. 4, pp. V4-182-V4-184. doi: 10.1109/ICETC.2010.5529705.
- [4]. T. Kulkarni, M. Kabra, and R. Shankarmani, “User Profiling Based Recommendation System for E-Learning,” in *2019 IEEE 16th India Council International Conference (INDICON)*, Dec. 2019, pp. 1–4. doi: 10.1109/INDICON47234.2019.9028982.
- [5]. A.-M. Nortvig, A. K. Petersen, and S. H. Balle, “A Literature Review of the Factors Influencing E-Learning and Blended Learning in Relation to Learning Outcome, Student Satisfaction and Engagement,” *Electronic Journal of e-Learning*, vol. 16, no. 1, p. pp46-55-pp46-55, Feb. 2018, Accessed: Jul. 24, 2021. [Online]. Available: <https://academic-publishing.org/index.php/ejel/article/view/1855>
- [6]. Myszkowski, Paweł & Kwasnicka, Halina & Markowska-Kaczmar, Urszula. (2008). Data Mining Techniques in e-Learning CelGrid System. Proceedings - 7th Computer Information Systems and Industrial Management Applications, CISIM 2008. 315-319. 10.1109/CISIM.2008.35.
- [7]. V. Filatov, O. Zolotukhin, A. Yerokhin, and M. Kudryavtseva, “Personalized Adaptation of Learning Environments,” in *2019 IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL)*, Sep. 2019, pp. 584–587. doi: 10.1109/CAOL46282.2019.9019525.
- [8]. R. Campos, R. Pereira dos Santos, and J. Oliveira, “Web-Based Recommendation System Architecture for Knowledge Reuse in MOOCs Ecosystems,” in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, Salt Lake City, UT, 2018, pp. 193–200. doi: 10.1109/IRI.2018.00036

-
- [9]. Sanjuán, Oscar & Pelayo García-Bustelo, B. & Gonzalez Crespo, Ruben & Franco, Enrique. (2009). Using Recommendation System for E-learning Environments at degree level. *International Journal of Interactive Multimedia and Artificial Intelligence*. 1. 67-70.
- [10]. Z. Chen, X. Liu, and L. Shang, "Improved course recommendation algorithm based on collaborative filtering," in *2020 International Conference on Big Data and Informatization Education (ICBDIE)*, Apr. 2020, pp. 466–469. doi: 10.1109/ICBDIE50010.2020.00115.
- [11]. A. Sharma, "A proposed e-learning system facilitating recommendation using content tagging and student learning styles," in *2017 5th National Conference on E-Learning E-Learning Technologies (ELELTECH)*, Aug. 2017, pp. 1–6. doi: 10.1109/ELELTECH.2017.8074989
- [12]. A. S. Nath and S. E., "A Pragmatic Review on Different Approaches Used in E-Learning Recommender Systems," in *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*, Dec. 2018, pp. 1–4. doi: 10.1109/ICCSDET.2018.8821088.
- [13]. A. Das and M. A. Akour, "Intelligent Recommendation System for E-Learning using Membership Optimized Fuzzy Logic Classifier," in *2020 IEEE Pune Section International Conference (PuneCon)*, Dec. 2020, pp. 1–10. doi: 10.1109/PuneCon50868.2020.9362416.
- [14]. H. Tan, J. Guo, and Y. Li, "E-learning Recommendation System," in *2008 International Conference on Computer Science and Software Engineering*, Dec. 2008, vol. 5, pp. 430–433. doi: 10.1109/CSSE.2008.305.
- [15]. B. Ramzan *et al.*, "An Intelligent Data Analysis for Recommendation Systems Using Machine Learning," *Scientific Programming*, Oct. 31, 2019. <https://www.hindawi.com/journals/sp/2019/5941096/> (accessed May 16, 2021).
- [16]. K. I. B. Ghauth and N. A. Abdullah, "Building an E-learning Recommender System Using Vector Space Model and Good Learners Average Rating," in *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, Jul. 2009, pp. 194–196. doi: 10.1109/ICALT.2009.161.
- [17]. Y. Jiang, H. Zhan, and Q. Zhuang, "Application research on personalized recommendation in distance education," in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, Taiyuan, China, 2010, pp. V13-357-V13-360. doi: 10.1109/ICCASM.2010.5622798.
- [18]. R.-M. Chao, J.-T. Huang, and C.-W. Yang, "The study of knowledge service-oriented recommendation mechanism - a case of e-learning platform," in *2005 International Conference on Machine Learning and Cybernetics*, Aug. 2005, vol. 4, pp. 2228-2233 Vol. 4. doi: 10.1109/ICMLC.2005.1527315.
- [19]. R. Ji, Y. Tian, and M. Ma, "Collaborative Filtering Recommendation Algorithm Based on User Characteristics," in *2020 5th International Conference on Control, Robotics and Cybernetics (CRC)*, Oct. 2020, pp. 56–60. doi: 10.1109/CRC51253.2020.9253466.
- [20]. Ying Liang and Hanrong Chen, "The research of video resource personalized recommendation system based on education website," in *2014 9th International Conference on Computer Science & Education*, Vancouver, BC, Canada, 2014, pp. 898–902. doi: 10.1109/ICCSE.2014.6926592.
- [21]. S. Souabi, A. Retbi, M. K. Idrissi, and S. Bennani, "A Recommendation Approach in Social Learning Based on K-Means Clustering," in *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Jun. 2020, pp. 1–5. doi: 10.1109/ISCV49265.2020.9204203.

-
- [22]. H. Zhang, H. Yang, T. Huang, and G. Zhan, "DBNCF: Personalized Courses Recommendation System Based on DBN in MOOC Environment," in *2017 International Symposium on Educational Technology (ISET)*, Jun. 2017, pp. 106–108. doi: 10.1109/ISET.2017.33.
- [23]. A. Mikroyannidis, P. Lefrere, and P. Scott, "An Architecture for Layering and Integration of Learning Ontologies, Applied to Personal Learning Environments and Cloud Learning Environments," in *2010 10th IEEE International Conference on Advanced Learning Technologies*, Jul. 2010, pp. 92–93. doi: 10.1109/ICALT.2010.33.
- [24]. Chan, Selena. "Digitally Enabling Learning by Doing in Vocational Education: Enhancing Learning as Becoming Processes." *SpringerBriefs in Education* (2021).
- [25]. A. F. Hidayat, D. D. J. Suwawi, and K. A. Laksitowening, "Learning Content Recommendations on Personalized Learning Environment Using Collaborative Filtering Method," in *2020 8th International Conference on Information and Communication Technology (ICoICT)*, Jun. 2020, pp. 1–6. doi: 10.1109/ICoICT49345.2020.9166371.
- [26]. Pardos, Z. A., & Jiang, W. (2020, March). Designing for serendipity in a university course recommendation system. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 350-359).
- [27]. Yu, R., Pardos, Z., Chau, H., & Brusilovsky, P. (2021, June). Orienting Students to Course Recommendations Using Three Types of Explanation. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 238-245).