

# Περιεχόμενα

Πρόλογος. . . . .	xi
<b>Μέρος Α - Εισαγωγή . . . . .</b>	<b>1</b>
<b>Κεφάλαιο 1 - Εισαγωγή . . . . .</b>	<b>6</b>
1.1 Βασικές Εργασίες Εξόρυξης Γνώσης από Δεδομένα . . . . .	6
1.1.1 Κατηγοριοποίηση. . . . .	6
1.1.2 Παλινδρόμηση . . . . .	6
1.1.3 Ανάλυση Χρονοσειρών . . . . .	7
1.1.4 Πρόβλεψη . . . . .	8
1.1.5 Συσταδοποίηση . . . . .	8
1.1.6 Παρουσίαση Συνόψεων . . . . .	9
1.1.7 Κανόνες Συσχέτισης . . . . .	10
1.1.8 Ανακάλυψη Ακολουθιών . . . . .	10
1.2 Η Εξόρυξη ως Στάδιο της Ανακάλυψης Γνώσης σε Βάσεις Δεδομένων . . . . .	11
1.2.1 Η Ανάπτυξη της Εξόρυξης γνώσης από δεδομένα. . . . .	14
1.3 Θέματα Εξόρυξης Γνώσης από Δεδομένα . . . . .	17
1.4 Μέτρα Αξιολόγησης της Εξόρυξης Γνώσης από Δεδομένα . . . . .	19
1.5 Κοινωνικές Επιπτώσεις της Εξόρυξης Γνώσης . . . . .	20
1.6 Η Εξόρυξη Γνώσης από την Σκοπιά των Βάσεων Δεδομένων . . . . .	20
1.7 Το Μέλλον . . . . .	21
1.8 Ασκήσεις . . . . .	23
1.9 Βιβλιογραφικές Σημειώσεις . . . . .	24
<b>Κεφάλαιο 2 - Σχετικές Έννοιες. . . . .</b>	<b>26</b>
2.1 Συστήματα Βάσεων Δεδομένων / Άμεσης Επεξεργασίας Συναλλαγών . . . . .	26
2.2 Ασαφή Σύνολα και Ασαφής Λογική . . . . .	29
2.3 Ανάκτηση Πληροφοριών . . . . .	32
2.4 Συστήματα Υποστήριξης Αποφάσεων . . . . .	35
2.5 Μοντελοποίηση με Διαστάσεις . . . . .	35
2.5.1 Πολυδιάστατα Σχήματα . . . . .	38
2.5.2 Ευρετήρια . . . . .	41
2.6 Αποθήκες Δεδομένων . . . . .	42
2.7 Άμεση Αναλυτική Επεξεργασία - OLAP . . . . .	47
2.8 Μηχανές Αναζήτησης στο Διαδίκτυο. . . . .	49
2.9 Στατιστική. . . . .	50
2.10 Μηχανική Μάθηση. . . . .	51
2.11 Ταίριασμα Προτύπων . . . . .	53
2.12 Σύνοψη . . . . .	53
2.13 Ασκήσεις . . . . .	54
2.14 Βιβλιογραφικές Σημειώσεις. . . . .	54

<b>Κεφάλαιο 3 - Τεχνικές Εξόρυξης Γνώσης . . . . .</b>	<b>55</b>
3.1 Εισαγωγή . . . . .	55
3.2 Η Εξόρυξη Γνώσης από την Σκοπιά της Στατιστικής. . . . .	56
3.2.1 Σημειακή Εκτίμηση . . . . .	56
3.2.2 Μοντέλα Βασιζόμενα στη Σύνοψη . . . . .	61
3.2.3 Θεώρημα του Bayes . . . . .	62
3.2.4 Έλεγχος Υπόθεσης . . . . .	65
3.2.5 Παλινδρόμηση και Συσχέτιση . . . . .	66
3.3 Μέτρα Ομοιότητας. . . . .	68
3.4 Δένδρα Απόφασης. . . . .	70
3.5 Νευρωνικά Δίκτυα. . . . .	73
3.5.1 Συναρτήσεις Ενεργοποίησης . . . . .	77
3.6 Γενετικοί Αλγόριθμοι . . . . .	80
3.7 Ασκήσεις . . . . .	84
3.8 Βιβλιογραφικές Σημειώσεις . . . . .	85
<b>Μέρος Β - Βασικά Θέματα . . . . .</b>	<b>87</b>
<b>Κεφάλαιο 4 - Κατηγοριοποίηση . . . . .</b>	<b>89</b>
4.1 Εισαγωγή . . . . .	89
4.1.1 Ζητήματα Σχετικά με την Κατηγοριοποίηση. . . . .	92
4.2 Αλγόριθμοι Βασισμένοι στη Στατιστική . . . . .	95
4.2.1 Παλινδρόμηση . . . . .	95
4.2.2 Bayesian Κατηγοριοποίηση. . . . .	103
4.3 Αλγόριθμοι Βασισμένοι στην Απόσταση . . . . .	105
4.3.1 Απλή Προσέγγιση . . . . .	106
4.3.2 <i>K</i> Πλησιέστεροι Γείτονες . . . . .	108
4.4 Αλγόριθμοι Βασισμένοι σε Δένδρα Αποφάσεων . . . . .	109
4.4.1 ID3 . . . . .	114
4.4.2 C4.5 και C5.0. . . . .	118
4.4.3 CART . . . . .	120
4.4.4 Κλιμακούμενες Τεχνικές για Δένδρα Απόφασης . . . . .	122
4.5 Αλγόριθμοι Βασισμένοι σε Νευρωνικά Δίκτυα . . . . .	122
4.5.1 Διάδοση . . . . .	125
4.5.2 Εποπτευόμενη Μάθηση του NN . . . . .	126
4.5.3 Δίκτυα Συνάρτησης Ακτινικής Βάσης . . . . .	133
4.5.4 Νευρωνικά Δίκτυα Perceptron. . . . .	134
4.6 Αλγόριθμοι Βασισμένοι σε Κανόνες . . . . .	135
4.6.1 Δημιουργία Κανόνων από ένα Δένδρο Απόφασης. . . . .	136
4.6.2 Δημιουργία Κανόνων από ένα Νευρωνικό Δίκτυο. . . . .	137
4.6.3 Δημιουργία Κανόνων χωρίς DT ή NN . . . . .	137
4.7 Συνδυασμός Τεχνικών . . . . .	141
4.8 Σύνοψη . . . . .	143
4.9 Ασκήσεις . . . . .	144
4.10 Βιβλιογραφικές Σημειώσεις . . . . .	145

<b>Κεφάλαιο 5 - Συσταδοποίηση</b>	<b>148</b>
5.1 Εισαγωγή	148
5.2 Ομοιότητα και Μέτρα Απόστασης	153
5.3 Ακραία Σημεία	155
5.4 Ιεραρχικοί Αλγόριθμοι	156
5.4.1 Συσσωρευτικοί Αλγόριθμοι	158
5.4.2 Διαιρετικοί Αλγόριθμοι Συσταδοποίησης	163
5.5 Διαμεριστικοί Αλγόριθμοι	164
5.5.1 Δένδρο Ελάχιστης Ζευξης	164
5.5.2 Αλγόριθμος Συσταδοποίησης Τετραγωνικού Σφάλλματος	165
5.5.3 Συσταδοποίηση K-Means	166
5.5.4 Αλγόριθμος Πλησιέστερου Γείτονα	168
5.5.5 Αλγόριθμος PAM	169
5.5.6 Αλγόριθμος Ενέργειας Δεσμού	173
5.5.7 Συσταδοποίηση με Γενετικούς Αλγορίθμους	174
5.5.8 Συσταδοποίηση με Νευρωνικά Δίκτυα	175
5.6 Συσταδοποίηση σε μεγάλες Βάσεις Δεδομένων	178
5.6.1 BIRCH	179
5.6.2 Αλγόριθμος DBSCAN	182
5.6.3 Αλγόριθμος CURE	184
5.7 Συσταδοποίηση με Κατηγορικά Γνωρίσματα	187
5.8 Σύγκριση	190
5.9 Ασκήσεις	192
5.10 Βιβλιογραφικές Σημειώσεις	193
<b>Κεφάλαιο 6 - Κανόνες Συσχετίσεων</b>	<b>196</b>
6.1 Εισαγωγή	196
6.2 Συχνά Στοιχειοσύνολα	200
6.3 Βασικοί Αλγόριθμοι	202
6.3.1 Ο Αλγόριθμος Apriori	202
6.3.2 Ο Αλγόριθμος της Δειγματοληψίας	206
6.3.3 Διαμέριση	211
6.4 Παράλληλοι και Κατανεμημένοι Αλγόριθμοι	212
6.4.1 Παραλληλισμός Δεδομένων	213
6.4.2 Παραλληλισμός Εργασιών	215
6.5 Σύγκριση Προσεγγίσεων	216
6.6 Αυξητικοί Κανόνες	219
6.7 Πποηγμένες Τεχνικές Επαγωγής Κανόνων Συσχετίσεων	220
6.7.1 Γενικευμένοι Κανόνες Συσχετίσεων	220
6.7.2 Κανόνες Συσχετίσεων Πολλαπλών Επιπέδων	221
6.7.3 Ποσοτικοί Κανόνες Συσχετίσεων	221
6.7.4 Πολλαπλές Τιμές Ελάχιστης Υποστήριξης	222
6.7.5 Κανόνες (στατιστικής) Συσχέτισης	224
6.8 Μέτρηση της Ποιότητας των Κανόνων	224
6.9 Ασκήσεις	227
6.10 Βιβλιογραφικές Σημειώσεις	228

<b>Μέρος Γ - Προηγμένα Θέματα . . . . .</b>	<b>231</b>
<b>Κεφάλαιο 7 - Εξόρυξη Γνώσης από τον Παγκόσμιο Ιστό . . . . .</b>	<b>233</b>
7.1 Εισαγωγή . . . . .	233
7.2 Εξόρυξη Γνώσης από το Περιεχόμενο του Παγκόσμιου Ίστου . . . . .	235
7.2.1 Web crawlers . . . . .	237
7.2.2 Το Σύστημα Harvest . . . . .	241
7.2.3 Υπερβατική Όψη του Παγκόσμιου Ιστού . . . . .	241
7.2.4 Εξατομίκευση . . . . .	242
7.3 Εξόρυξη Γνώσης από τη Δομή του Παγκόσμιου Ίστου. . . . .	245
7.3.1 PageRank . . . . .	245
7.3.2 Clever . . . . .	246
7.4 Εξόρυξη Γνώσης από τη Χρήση του Παγκόσμιου Ίστου . . . . .	248
7.4.1 Προπαρασκευή . . . . .	250
7.4.2 Δομές Δεδομένων . . . . .	252
7.4.3 Ανακάλυψη Προτύπων . . . . .	254
7.4.4 Ανάλυση Προτύπων . . . . .	261
7.5 Ασκήσεις . . . . .	262
7.6 Βιβλιογραφικές Σημειώσεις . . . . .	263
<b>Κεφάλαιο 8 - Εξόρυξη Χωρικής Γνώσης . . . . .</b>	<b>265</b>
8.1 Εισαγωγή . . . . .	265
8.2 Επισκόπηση των Χωρικών Δεδομένων. . . . .	266
8.2.1 Χωρικές Ερωτήσεις . . . . .	266
8.2.2 Δομές Χωρικών Δεδομένων . . . . .	267
8.2.3 Θεματικοί Χάρτες. . . . .	271
8.2.4 Βάσεις Δεδομένων Εικόνων . . . . .	271
8.3 Βασικές Αρχές Εξόρυξης Γνώσης από Χωρικά Δεδομένα. . . . .	272
8.4 Γενίκευση και Εξειδίκευση . . . . .	274
8.4.1 Προοδευτική Βελτίωση . . . . .	274
8.4.2 Γενίκευση . . . . .	275
8.4.3 Πλησιέστερος Γείτονας . . . . .	277
8.4.4 STING . . . . .	277
8.5 Χωρικοί Κανόνες. . . . .	280
8.5.1 Κανόνες Χωρικών Συσχετίσεων . . . . .	280
8.6 Αλγόριθμοι Χωρικής Κατηγοριοποίησης . . . . .	282
8.6.1 Επέκταση του ID3. . . . .	283
8.6.2 Δένδρο Χωρικής Απόφασης . . . . .	283
8.7 Αλγόριθμοι Χωρικής Συσταδοποίησης. . . . .	284
8.7.1 Επεκτάσεις του CLARANS . . . . .	285
8.7.2 SD(CLARANS). . . . .	286
8.7.3 DBCLASD. . . . .	287
8.7.4 BANG . . . . .	288
8.7.5 WaveCluster . . . . .	289
8.7.6 Προσέγγιση . . . . .	289
8.8 Ασκήσεις . . . . .	291

8.9 Βιβλιογραφικές Σημειώσεις . . . . .	291
---	-----

**Κεφάλαιο 9 - Εξόρυξη Χρονικής Γνώσης . . . . . 293**

9.1 Εισαγωγή . . . . .	293
9.2 Μοντελοποίηση Χρονικών Γεγονότων . . . . .	296
9.3 Χρονολογικές Σειρές. . . . .	301
9.3.1 Ανάλυση Χρονολογικών Σειρών. . . . .	302
9.3.2 Ανάλυση Τάσεων . . . . .	302
9.3.3 Μετασχηματισμός . . . . .	305
9.3.4 Ομοιότητα . . . . .	305
9.3.5 Πρόβλεψη . . . . .	306
9.4 Ανίχνευση Προτύπων . . . . .	308
9.4.1 Ταίριασμα Συμβολοσειρών. . . . .	308
9.5 Ακολουθίες . . . . .	311
9.5.1 AprioriAll . . . . .	313
9.5.2 SPADE . . . . .	314
9.5.3 Γενίκευση . . . . .	316
9.5.4 Εξαγωγή Χαρακτηριστικών . . . . .	318
9.6 Κανόνες Χρονικών Συσχετίσεων . . . . .	318
9.6.1 Κανόνες Συσχετίσεων δια-συναλλαγών. . . . .	319
9.6.2 Κανόνες Επεισοδίων . . . . .	320
9.6.3 Εξαρτήσεις Τάσεων. . . . .	321
9.6.4 Κανόνες Συσχέτισης Ακολουθιών . . . . .	323
9.6.5 Ημερολογιακοί Κανόνες Συσχετίσεων . . . . .	324
9.7 Ασκήσεις . . . . .	325
9.8 Βιβλιογραφικές Σημειώσεις . . . . .	326

**ΠΑΡΑΡΤΗΜΑΤΑ**

<b>A Προϊόντα Εξόρυξης Γνώσης από Δεδομένα . . . . .</b>	<b>327</b>
A.1 Βιβλιογραφικές Σημειώσεις . . . . .	327

<b>B Βιβλιογραφία . . . . .</b>	<b>344</b>
---------------------------------	------------

<b>Γλωσσάρι - Ευρετήριο Όρων. . . . .</b>	<b>358</b>
---	------------

<b>Περί της συγγραφής και των επιμελητών της μετάφρασης . . . . .</b>	<b>365</b>
---	------------

