

Εφαρμογή Μεθόδων Εξόρυξης Δεδομένων για την Αξιολόγηση της Συναισθηματικής Κατάστασης



Μ. Νικιέλ¹, Σ. Φύτρος¹, Α. Γκέλιου¹, Β. Κεφαλά¹, Π. Μουτοπούλου¹, Ζ. Ξανάκης¹, Ε. Γκιντώνη², Κ. Χαλκιάπουλος^{1,3}, Γ. Αντζουλάτος^{1,3}

¹ Τμήμα Διοίκησης Επιχειρήσεων, ΑΤΕΙ Δυτικής Ελλάδας

² Τμήμα Ψυχολογίας, Πανεπιστήμιο Κρήτης

³ Τμήμα Μαθηματικών, Πανεπιστήμιο Πατρών



ΕΙΣΑΓΩΓΗ

Με την παρούσα ερευνητική εργασία εφαρμόζονται μέθοδοι Μηχανικής Μάθησης και Εξόρυξης Δεδομένων για την αποτύπωση της ψυχολογικής κατάστασης φοιτητών ηλικιακής ομάδας 18 έως 26 χρόνων. Για την καταγραφή, αποτύπωση και αξιολόγηση της ψυχολογικής κατάστασης χρησιμοποιήθηκε η σταθμισμένη κλίμακα ψυχοπαθολογίας Symptom Checklist-90 (SCL-90), η οποία εξετάζει ένα ευρύ φάσμα ψυχολογικών προβλημάτων και συμπτωμάτων ψυχοπαθολογίας. Το SCL-90 αποτελείται από 90 ερωτήσεις και εντοπίζει 9 κλινικά σημεία (υποκλίμακες) και 3 σύνολα απαντήσεων που περιλαμβάνουν τη βαθμολογία των επιμέρους κλιμάκων. Οι υποκλίμακες ορίζονται ως εξής: **σωματοποίηση** (somatization), **ψυχαναγκαστικότητα - καταναγκαστικότητα** (obsessive compulsive), **διαπροσωπική ευαισθησία** (interpersonal sensitivity), **κατάθλιψη** (depression), **άγχος** (anxiety), **θυμός-επιθετικότητα** (anger - hostility), **φοβικό άγχος** (phobic anxiety), **παρανοειδής ιδεασμός** (paranoid ideation), **ψυχωτισμός** (psychotism).

ΑΠΟΤΕΛΕΣΜΑΤΑ

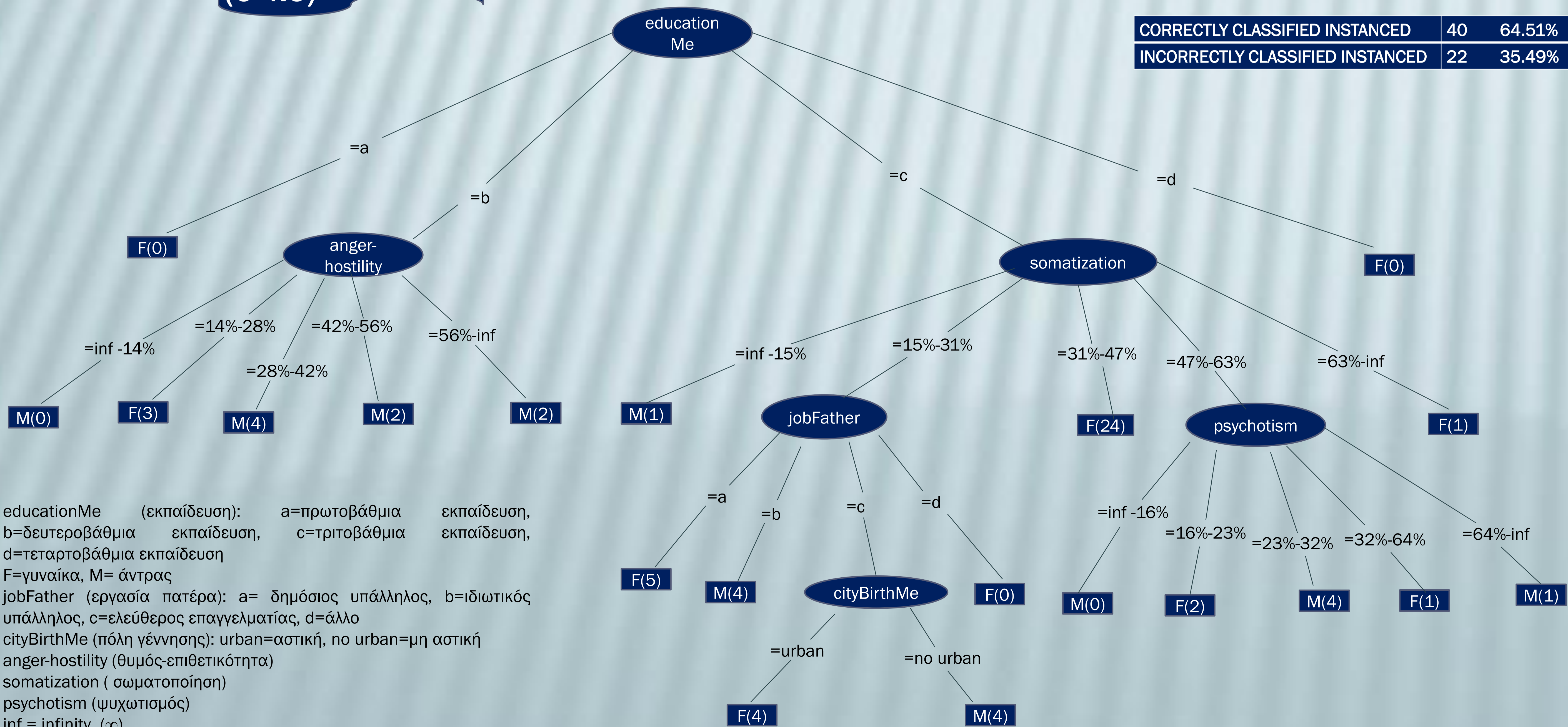
Το ακόλουθο δέντρο απόφασης εμφανίζει ως ρίζα το πως επηρεάζεται η εκπαίδευση των ερωτηθέντων ανάλογα με το φύλο.

- ✓ Στην αριστερή μεριά εμφανίζεται πως καμία γυναίκα με πρωτοβάθμια εκπαίδευση δεν υπάρχει μέσα στον δείγμα.
- ✓ Πιο δεξιά δείχνει το ποσοστό των ερωτηθέντων οι οποίοι έχουν ολοκληρώσει την δευτεροβάθμια εκπαίδευση και εμφανίζουν το σύμπτωμα του θυμού-επιθετικότητας.
- ✓ Δεξιότερα ακολουθεί η τριτοβάθμια εκπαίδευση στην οποία το δείγμα έδειξε ότι οι ερωτηθέντες διακρίνονται από τα εξής συμπτώματα:
 - ✓ **Σωματοποίηση**: Σε αυτό το επίπεδο εμφανίζεται το ποσοστό από 15%-31% το οποίο επηρεάζει 5 γυναίκες του δείγματος οι οποίες ο πατέρας τους εργαζόταν ως δημόσιος υπάλληλος. Επιπρόσθετα επηρεάζονται 4 άντρες του δείγματος όταν το επάγγελμα του πατέρα τους ήταν ιδιωτικός υπάλληλος. Επιπλέον στην περίπτωση που ο πατέρας των ερωτηθέντων εργάζεται ως ελεύθερος επαγγελματίας το δείγμα μας εξαρτάται αν γεννήθηκε σε αστική ή μη αστική πόλη το οποίο επηρεάζει 4 γυναίκες γεννημένες σε αστική πόλη και 4 άντρες γεννημένοι σε μη αστική πόλη. Τέλος καμία γυναίκα δεν βρέθηκε ο πατέρας της να εργάζεται κάτι άλλο. Από 31%-47% βρέθηκαν να επηρεάζονται 24 γυναίκες του δείγματος.
 - ✓ **Ψυχωτισμός**: Αναφέρεται σε ένα ποσοστό της τάξης από 47%-63%. Πιο συγκεκριμένα κανένας άντρας δεν βρέθηκε με βαθμό ψυχωτισμού μικρότερου του 16%, μόνο για 2 γυναίκες ο βαθμός ψυχωτισμού ανέρχεται από 16%-23%, από 23%- 32% επηρεάζονται 4 άντρες, από 32%- 64% ο βαθμός ψυχωτισμού επηρεάζει 1 γυναίκα και για ποσοστό μεγαλύτερο του 64% βρέθηκε να επηρεάζεται 1 μόνο άντρας.
- ✓ Τέλος στον κλάδο της τεταρτοβάθμιας εκπαίδευσης δεν εμφανίζεται καμία γυναίκα στο δείγμα.

TP RATE	FP RATE	PRECISION	RECALL	F-MEASURE	MCC	ROC AREA	RPC AREA	CLASS
0.609	0.225	0.500	0.409	0.450	0.194	0.619	0.459	M
0.775	0.592	0.705	0.775	0.738	0.194	0.619	0.716	F
0.445	0.462	0.632	0.445	0.636	0.194	0.619	0.625	

CORRECTLY CLASSIFIED INSTANCED	40	64.51%
INCORRECTLY CLASSIFIED INSTANCED	22	35.49%

Δέντρο απόφασης (C 4.5)



- educationMe (εκπαίδευση): a=πρωτοβάθμια εκπαίδευση, b=δευτεροβάθμια εκπαίδευση, c=τριτοβάθμια εκπαίδευση, d=τεταρτοβάθμια εκπαίδευση
- F=γυναίκα, M= άντρας
- jobFather (εργασία πατέρα): a= δημόσιος υπάλληλος, b=ιδιωτικός υπάλληλος, c=ελεύθερος επαγγελματίας, d=άλλο
- cityBirthMe (πόλη γέννησης): urban=αστική, no urban=μη αστική
- anger-hostility (θυμός-επιθετικότητα)
- somatization (σωματοποίηση)
- psychotism (ψυχωτισμός)
- inf = infinity (∞)

ΜΕΘΟΔΟΣ

Η συλλογή δεδομένων υλοποιήθηκε σε πρώτη φάση με τη δημιουργία ηλεκτρονικών ερωτηματολογίων μέσω της υπηρεσίας Google Forms, τα οποία αναρτήθηκαν μέσω του διαδικτυακού τόπου <http://www.cicos.gr>. Εν συνεχεία έγινε συλλογή και προεπεξεργασία των δεδομένων από τα ερωτηματολόγια και κατόπιν εισήχθησαν στην πλατφόρμα μηχανικής μάθησης Weka για ανάλυση και εξαγωγή χρήσιμης γνώσης. Πιο συγκεκριμένα, με την χρήση αλγορίθμων κατηγοριοποίησης (ID3, C4.5) παράχθηκαν αντίστοιχα δέντρα απόφασης. Τα δέντρα αποφάσεων είναι ένα ισχυρό μέσο για την αναπαράσταση και τη διευκλίνηση της ανάλυσης καταστάσεων (ψυχολογικών) που περιλαμβάνουν διαδοχικές αποφάσεις και μεταβλητά αποτελέσματα σε οριοθετημένο χρονικό διάστημα. Επιπλέον, με χρήση τεχνικών συσταδοποίησης (πχ k-means) γίνεται προσπάθεια κατάταξης των ερωτώμενων σε συστάδες με βάση τα 9 κλινικά σημεία (υποκλίμακες) της κλίμακας SCL-90. Επίσης, καθορίστηκαν οι παράμετροι των αλγορίθμων αυτών, ανάλογα με τις περιπτώσεις εφαρμογής, ενώ τα αποτελέσματα συσχετίστηκαν με τον τύπο καταγωγής και διαμονής, το εκπαιδευτικό υπόβαθρο τόσο του ερωτηθέντων όσο και των συγγενών α' βαθμού, την επαγγελματική ενασχόληση γονέων κ.α., με σκοπό την αποτίμηση και αξιολόγηση της σημαντικότητας των εξαγομένων κανόνων/συμπερασμάτων.

ΑΛΓΟΡΙΘΜΟΣ APRIORI (ΠΑΡΑΔΕΙΓΜΑ)

somatization_ '(0.63216-inf)'=f 60 ==> interpersonal-sensitivity_ '(0.690720.82086]'=f 59 <conf:(0.98)> lift:(1) lev:(0) [0] conv:(0.48)

ΣΥΜΠΕΡΑΣΜΑΤΑ

- ✓ Το φύλο, η εκπαίδευση, η πόλη διαμονής, η πόλη γέννησης και η εκπαίδευση γονέων δεν επηρεάζει την ψυχολογική κατάσταση των φοιτητών
- ✓ Όταν η σωματοποίηση και ο θυμός-επιθετικότητα είναι μέχρι 15%, τότε η ψυχαναγκαστικότητα - καταναγκαστικότητα είναι μέχρι 18%
- ✓ Όταν η διαπροσωπική ευαισθησία βρίσκεται μεταξύ 69% και 82% και ο θυμός-επιθετικότητα είναι μέχρι 14%, τότε η ψυχαναγκαστικότητα - καταναγκαστικότητα είναι μέχρι 18%
- ✓ Όταν η ηλικία πατέρα του φοιτητή είναι 68 και πάνω ή η ηλικία μητέρας είναι κάτω από τα 40 και η ψυχαναγκαστικότητα - καταναγκαστικότητα είναι μέχρι 18%, τότε η σωματοποίηση είναι μέχρι 15%
- ✓ Όταν η σωματοποίηση είναι πάνω από 63% τότε η διαπροσωπική ευαισθησία βρίσκεται μεταξύ 69% και 82%
- ✓ Όταν η κατάθλιψη είναι μεταξύ 67% με 68%, τότε η σωματοποίηση είναι πάνω από 63%
- ✓ Όταν η ηλικία του πατέρα είναι μεταξύ 61 έως 68 ή όταν η εργασία του πατέρα δεν είναι δημόσιος υπάλληλος, ιδιωτικός υπάλληλος ή ελεύθερος επαγγελματίας, τότε η διαπροσωπική ευαισθησία βρίσκεται μεταξύ 69% και 82%
- ✓ Όταν ο ψυχωτισμός είναι πάνω από 64% τότε το φοβικό άγχος είναι πάνω από 83%
- ✓ Όταν το φοβικό άγχος βρίσκεται πάνω από 83% και ο ψυχωτισμός πάνω από 64% τότε ο παρανοειδής ιδεασμός είναι μέχρι 14%

ΒΙΒΛΙΟΓΡΑΦΙΑ

- ✓ Oded Maimon and Lior Rokach (Editors), "Data Mining and Knowledge Discovery Handbook", 2nd ed., Springer, 2010.
- ✓ Jiawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques", 2nd ed., Morgan Kaufmann Publishers, 2006.
- ✓ Pang-Ning Tan, Michael Steinbach and Vipin Kumar, "Introduction to Data Mining", Addison-Wesley, 2006.
- ✓ Derogatis, L.R., 'The SCL-90 Manual I. Scoring, Administration, and Procedures for the SL-90. Baltimore', MD: John Hopkins University School of Medicine, Clinical Psychometrics Unit., 1977
- ✓ Ντώνιας, Σ., Καραστεργίου, Α., & Μάνος, Ν., 'Στάθμιση της κλίμακας ψυχοπαθολογίας Symptom Checklist-90-R σε ελληνικό πληθυσμό', Ψυχιατρική, 2, 42-48, 1991.