

Μεταπτυχιακή φοιτήτρια: Τουτουτζή Μαρία

Τίτλος διπλωματικής εργασίας:

Μελέτη μη-ντετερμινιστικής ταξινόμησης και ταξινόμησης μετα-επιπέδου σε ιατρικά δεδομένα

Περίληψη:

Στον τομέα της ιατρικής, το κόστος μιας λάθος διάγνωσης μπορεί να έχει βαρύτατες συνέπειες. Η εφαρμογή αλγορίθμων μηχανικής μάθησης για εύρεση βέλτιστων ταξινομητών μπορεί να διαδραματίσει σημαντικό ρόλο συμβάλλοντας στην ελαχιστοποίηση των λαθών αυτών. Η εργασία αυτή έχει δύο σκέλη. Το πρώτο είναι η διερεύνηση της χρησιμότητας που μπορούν να έχουν σε ιατρικά δεδομένα οι μη-ντετερμινιστικοί ταξινομητές. Οι ταξινομητές αυτοί επιτρέπουν την πρόβλεψη, για κάθε δείγμα, περισσότερων από μια κλάσεις. Δεδομένου ότι η πραγματική κλάση πρέπει να συμπεριλαμβάνεται στις προβλέψεις και ότι ο αριθμός των προβλεφθεισών κλάσεων πρέπει να είναι όσο το δυνατόν μικρότερος, αυτό το είδος ταξινομητών μπορεί να θεωρηθεί ως διαδικασία ανάκτησης πληροφορίας (information-retrieval IR). Ειδικότερα μελετήθηκε ο αλγόριθμος του delCoz για την εύρεση βέλτιστων τέτοιων ταξινομητών, ο οποίος χρησιμοποιεί τις εκ των υστέρων πιθανότητες για να υπολογίσει το υποσύνολο των κλάσεων με τη χαμηλότερη αναμενόμενη απώλεια.

Το δεύτερο σκέλος της εργασίας είναι η διερεύνηση τρόπων συνδυασμένης χρήσης μη-ντετερμινιστικών ταξινομητών με ταξινομητές μετά-επιπέδου, όπως οι ταξινομητές με χρήση ψηφοφορίας και οι ταξινομητές συσσωρευμένης εκμάθησης. Συγκεκριμένα, μελετήθηκαν τρεις διαφορετικοί τρόποι συνδυασμού ντετερμινιστικών και μη ντετερμινιστικών ταξινομητών στο βασικό και στο μετά-επίπεδο αντίστοιχα. Η διερεύνηση αυτή γίνεται για πρώτη φορά στην βιβλιογραφία. Στα πλαίσια της εργασίας, γίνεται εκτενής πειραματική συγκριτική αξιολόγηση των μη-ντετερμινιστικών αλγορίθμων με και χωρίς ταξινομητές μετά-επιπέδου, σε σχέση με τους αντίστοιχους ντετερμινιστικούς, σε ιατρικά δεδομένα.

Τα αποτελέσματα αναδεικνύουν τόσο την χρησιμότητα των μη-ντετερμινιστικών ταξινομητών σε ιατρικά δεδομένα όσο και την εν γένει χρησιμότητα της συνδυασμένης χρήσης μη-ντετερμινιστικών ταξινομητών με ταξινομητές μετά-επιπέδου.

Εξεταστική Επιτροπή

Δρ. Ηλίας Μανωλάκος, Αν. Καθηγητής, Τμήμα Πληροφορικής και Τηλεπικοινωνιών, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Δρ. Σταύρος Περαντώνης, Διευθυντής Ερευνών, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών «Δημόκριτος» (επιβλέπων)

Δρ. Σέργιος Πετρίδης, Ερευνητής, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών «Δημόκριτος»