



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Σύστημα Υποστήριξης Διάγνωσης της Νόσου του Parkinson
με Χρήση Φωνητικών Καταγραφών**

Ευριπίδης Ιωάννης Μήκος

Επιβλέπων

Διονύσιος Κάβουρας, Καθηγητής

ΑΘΗΝΑ

ΙΑΝΟΥΑΡΙΟΣ 2016

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σύστημα Υποστήριξης Διάγνωσης της Νόσου του Parkinson με Χρήση Φωνητικών Καταγραφών

Ευριπίδης Ι. Μήκος
Α.Μ.: ΠΙΒ - 093

ΕΠΙΒΛΕΠΩΝ Διονύσιος Κάβουρας, Καθηγητής

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: Διονύσιος Κάβουρας, Καθηγητής
Εμμανουήλ Σαγκριώτης, Καθηγητής
Παντελεήμων Ασβεστάς, Επίκουρος Καθηγητής

Ιανουάριος 2016

ΠΕΡΙΛΗΨΗ

Σκοπός της εργασίας είναι η ανάλυση δεδομένων προερχόμενα από ασθενείς με νόσο του Parkinson και υγιείς μάρτυρες με χρήση στατιστικών μεθόδων και μεθόδων αναγνώρισης προτύπων, ώστε να δημιουργηθεί ένα σύστημα υποστήριξης της διάγνωσης της συγκεκριμένης νόσου.

Προς αυτή την κατεύθυνση θα χρησιμοποιηθούν δεδομένα, τα οποία έχουν συλλεχθεί από το Πανεπιστήμιο της Οξφόρδης σε συνεργασία με το Εθνικό Κέντρο για την φωνή και την ομιλία του Ντένβερ, Κολοράντο Η.Π.Α.

Το σύνολο δεδομένων αποτελείται από 195 μετρήσεις με 23 χαρακτηριστικά η καθεμία, προερχόμενες από 31 περιστατικά, 23 ασθενείς με νόσο του Parkinson και 8 υγιείς μάρτυρες. Τα χαρακτηριστικά αυτά περιγράφουν ιδιότητες της φωνής που σχετίζονται με μεταβολές στην έντασή της, καθώς και στο φασματικό της περιεχόμενο.

Αρχικά, θα γίνει στατιστική ανάλυση των δεδομένων για να βρεθεί ποια από όλα τα χαρακτηριστικά διαφέρουν στατιστικά σημαντικά ανάμεσα στις δύο ομάδες (υγιείς-πάσχοντες). Στη συνέχεια, για την δημιουργία του συστήματος υποστήριξης της διάγνωσης, θα γίνει επιλογή των χαρακτηριστικών στον ελάχιστο αριθμό αυτών που δίνουν την μέγιστη διάκριση των δύο κατηγοριών. Με αυτά τα χαρακτηριστικά θα γίνει εκπαίδευση του συστήματος με την χρήση διαφόρων ταξινομητών καθώς και συνδυασμούς τους. Στο τέλος με βάση τα υπάρχοντα δεδομένα θα γίνει αξιολόγηση του συστήματος, για την εξαγωγή της ακρίβειάς του και της ευαισθησίας του.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Αναγνώριση Προτύπων

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: νόσος Parkinson, ταξινομητές, πρότυπα, χαρακτηριστικά, κατηγοριοποίηση

ABSTRACT

The aim of this thesis is the analysis of data that are originated from patients with Parkinson's disease and healthy martyrs, with use of statistical and pattern recognition methods, in order an integrated computer aided diagnosis (CAD) system to be fully developed for this particular disease.

Towards this direction, a dataset will be used which is coming from the Oxford University in collaboration with the National Center for Speech and Voice, Colorado USA.

The dataset is consisted of 195 measurements with 23 features each, which are coming from 31 cases, 23 with Parkinson's disease and 8 healthy ones. These features describe speech properties that are linked to intensity variations as well as in spectra content.

Initially, a statistical analysis will take place, in order to find which features are statistical different between two groups (healthy – diseased). After this, for the development of the CAD system, the minimum number of the features consists significant statistical differences will be selected. With the use of these features and combinations of them, the system will be trained. Lastly, these features will participate in system validation, for the extraction of its specificity and sensitivity.

SUBJECT AREA: Pattern Recognition

KEYWORDS: Parkinson's disease, classifiers, patterns, features, classification