

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Ζερβού Μιχαέλα - Αρετή
Μεταπτυχιακή Φοιτήτρια**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης
Επόπτης Μεταπτ. Εργασίας: Καθηγητής, Π. Τσακαλίδης**

Δευτέρα, 01/07/2019, 16:00

Αίθουσα K206, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

“ Αυτοματοποιημένη δυναμική ανάλυση πολυπαραγοντικών φυσιολογικών δεδομένων για εξατομικευμένη υγειονομική περίθαλψη”

ΠΕΡΙΛΗΨΗ

Οι φορητοί αισθητήρες σώματος σε συνεργασία με τις τεχνικές ανάλυσης δεδομένων έχουν αρχίσει να χρησιμοποιούνται πρόσφατα στον τομέα της εξατομικευμένης υγειονομικής περίθαλψης, δεδομένου ότι η φορητή τεχνολογία μπορεί να παρακολουθεί τη φυσική δραστηριότητα, να συλλέγει δεδομένα και να παρέχει πληροφορίες σε πραγματικό χρόνο σε άτομα ή εμπειρογνώμονες. Ωστόσο, υπάρχουν αρκετές προκλήσεις που σχετίζονται με την παρακολούθηση της απόδοσης του ασθενούς με τη χρήση δεδομένων από αισθητήρες σώματος. Συγκεκριμένα, τα ληφθέντα δεδομένα μπορεί να είναι πολύ εξατομικευμένα με βάση τον ασθενή και ανεπαρκή για τη διαφοροποίηση της φυσιολογικής και της ανώμαλης συμπεριφοράς για την μοντελοποίηση ενός προβλήματος ταξινόμησης. Έτσι, απαιτείται μια εξατομικευμένη προσέγγιση για την ανάλυση της δραστηριότητας του ασθενούς με ένα μη εποπτευόμενο ή ημι-εποπτευόμενο τρόπο για την αναγνώριση των προτύπων, η οποία θα είναι επίσης εφαρμόσιμη σε πολλαπλούς ετερογενείς αισθητήρες που καταγράφουν διαφορετικά δεδομένα.

Η πλειοψηφία των υφιστάμενων τεχνικών διάγνωσης χρησιμοποιεί χαρακτηριστικά που εξάγονται από συμβολικές αναπαραστάσεις ή μελετώντας την συχνότητα των σχετικών

δεδομένων, ενώ είτε αγνοούν πλήρως τη συμπεριφορά του υποκείμενου δυναμικού συστήματος που παράγει τα δεδομένα, είτε δεν έχουν την ικανότητα ταυτόχρονης επεξεργασίας πολλαπλών διαστάσεων. Για να επισημάνει αυτό το πρόβλημα, η εργασία αυτή προτείνει μια νέα αυτόματη αρχιτεκτονική για την εξαγωγή χαρακτηριστικών, μοντελοποιώντας άμεσα την εγγενή δυναμική των δεδομένων των πολυδιάστατων χρονοσειρών που παράγονται από φορητούς αισθητήρες, σε υψηλότερης διάστασης χώρους φάσεων όπου στη συνέχεια κωδικοποιούνται οι επανεμφανίσεις των καταστάσεων που αφορούν μια συγκεκριμένη διαδικασία. Στην πραγματικότητα, η εργασία αυτή επεκτείνει την γνωστή από τη βιβλιογραφία μέθοδο Ανάλυσης Ποσοτικής Επαναληψιμότητας (RQA) σε πολυδιάστατα δεδομένα που βασίζεται σε διανύσματα καταστάσεων, σε μια αρχιτεκτονική που βασίζεται σε πίνακες καταστάσεων. Οι πίνακες καταστάσεων θεωρούνται περισσότερο κατάλληλες για την περιγραφή πολυδιάστατων σημάτων από τη μαθηματική σκοπιά, επιτρέποντάς τους να μοντελοποιήσουν τους συσχετισμούς όχι μόνο εντός ενός σήματος, αλλά και μεταξύ διαφορετικών σημάτων. Αξίζει να σημειωθεί ότι σε αντίθεση με την ήδη υπάρχουσα προσέγγιση της βιβλιογραφίας, το παρόν προτεινόμενο πλαίσιο εφαρμόζεται άμεσα σε διάφορες δομές δεδομένων, όπως εικόνες, υπερφασματικά δεδομένα ή ροές βίντεο.

Η μελέτη αυτή χρησιμοποιεί επιπλέον τεχνικές μηχανικής μάθησης για να σχεδιάσει ένα αποδοτικό σύστημα εξαγωγής χαρακτηριστικών για την ανακάλυψη πλούσιων σε πληροφορία προτύπων που καταγράφουν την υποκείμενη δυναμική των δεδομένων πολυδιάστατων χρονοσειρών για την ανίχνευση της δυσλεξίας. Η δυσλεξία είναι μια αναπτυξιακή διαταραχή εκμάθησης της ακρίβειας ή / και της ευχέρειας ανάγνωσης λέξεων, με την έρευνα να κατευθύνεται προς την κατανόηση της συμβολής του οπτικού συστήματος. Παρόλο που η δυσλεξία δεν είναι μια οφθαλμοκινητική ασθένεια, οι αναγνώστες με δυσλεξία έχουν δείξει διαφορετικές κινήσεις των ματιών από ότι οι τυπικά ανεπτυγμένοι μαθητές κατά την ανάγνωση κειμένου. Η πειραματική αξιολόγηση σε πραγματικά δεδομένα παρακολούθησης των ματιών για τη διάγνωση της δυσλεξίας καταδεικνύει μια βελτιωμένη απόδοση της μεθόδου μας όσον αφορά την ακρίβεια ταξινόμησης, το F-score και την ευρωστία παρουσία θορύβου σε σύγκριση με το μονοδιάστατο και το πολυδιάστατο μοντέλο RQA της βιβλιογραφίας.

Zervou Mixaela- Areti

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Professor, P. Tsakalides

Monday, 01/07/2019, 16:00

Room K206, Computer Science Dept., University of Crete

“Automated dynamical analysis of multivariate physiological data for personalized healthcare”

ABSTRACT

Wearable body sensors along with data analytics have been recently employed in the domain of personalized healthcare, since wearable technology can monitor physical activity, collect data and deliver real-time feedback to individuals or experts. However, there are several challenges associated with the monitoring of patient performance using body sensor data. Specifically, the captured sensor data might be too patient-specific and inadequate for the differentiation of normal and abnormal behavior, which is formulated as a classification problem. Thus, a generic approach is required to analyze patient activity in an unsupervised or semi-supervised manner for identifying patterns and also being scalable across multiple heterogeneous sensors capturing different data.

Multivariate time series analysis is considered as a difficult problem due to the complexity of the data types. The main challenges of processing time series data involve the high dimensionality, the presence of noise and redundancy in the data. The majority of existing diagnosis techniques employ features extracted from symbolic or frequency-domain representations of the associated data, whilst ignoring completely the behavior of the underlying data generating the dynamical system. Moreover, those techniques lack the capability of concurrently processing multiple dimensions. To address this problem, this thesis proposes a novel self-tuned architecture for feature extraction, by modeling directly the inherent dynamics of multidimensional wearable sensor time series data in higher-dimensional phasespaces, which encode state recurrences considering a specific task. In essence, our work extends the state-of-the-art multidimensional recurrence quantification analysis (RQA), that is based on state vectors, to a more generic state matrix-based architecture. State matrices are considered more appropriate for describing multi-dimensional signals from a mathematical perspective, enabling them to model the correlations not only within a signal but also between different signals. It is worth mentioning that in contrast to the state-of-the-art, the herein proposed framework is directly applicable to several data structures such as images, hyper-spectral data or video streams.

Furthermore, this study employs machine learning techniques to design an efficient feature extraction scheme for the discovery of information-rich patterns that best capture the underlying data dynamics of multidimensional time series data for dyslexia detection. Dyslexia is a developmental learning disorder of single word reading accuracy and/or fluency, with compelling research directed towards understanding the contributions of the visual system. While dyslexia is not an oculomotor disease, readers with dyslexia have shown different eye movements than typically developing students during text reading. Experimental evaluation on real data of eye-tracking for diagnosing dyslexia demonstrates an improved performance of our method in terms of classification

accuracy, F-score and robustness in the presence of noise when compared against a state-of-the-art unidimensional and vector-based multidimensional RQA.