

Εργασία 2022-2023

Στην επιστημονική βιβλιογραφία, υπάρχει αρκετή έρευνα που αφορά στην ταξινόμηση κειμένου σε κατηγορίες, όπως στα εξής θέματα:

- Emotion detection and classification on texts
- **Toxic comment classification**
- Fake news detection and classification

Ανάλογα το θέμα, η κατηγοριοποίηση γίνεται σε δύο ή και περισσότερες κλάσεις. Οι ερευνητικές εργασίες βασίζονται σε δεδομένα που έχουν συλλεχθεί από το διαδίκτυο και συνήθως από το twitter. Αυτές οι συλλογές δεδομένων είναι συνήθως ελεύθερα διαθέσιμες και έτσι μπορούν να επαναχρησιμοποιηθούν για την εφαρμογή και σύγκριση διαφόρων αλγορίθμων.

Γράψτε ένα επιστημονικό ερευνητικό άρθρο για ένα από τα παραπάνω θέματα.

Προδιαγραφές

Το άρθρο σας, μήκους 5000 λέξεων, στην Ελληνική ή την Αγγλική γλώσσα, θα είναι γραμμένο σε επιστημονικό ύφος και θα βασίζεται σε τουλάχιστον 5 επιστημονικά άρθρα (journal articles, conference proceedings) πρέπει συμπεριλαμβάνει εκτός των άλλων τις εξής ενότητες: Περίληψη (**Abstract**), Εισαγωγή (**Introduction**), το κυρίως σώμα (**Literature review**) με ενότητες για τις τρέχουσες τεχνολογίες (state-of-the-art), Μεθοδολογία (**Methodology**), Αποτελέσματα (**Results and Discussion**), Μελλοντικές προκλήσεις (**Future work/suggestions**), Συμπεράσματα (**Conclusions**) και στο τέλος Αναφορές-Βιβλιογραφία (References). Οι ενότητες της εργασίας (εκτός της Περίληψης) θα είναι αριθμημένες (1,2,3, ...) καθώς και οι υπο-ενότητες (1.1, 1.2, 1.2.1., ...) και οι σελίδες αριθμημένες. Προτείνεται να χρησιμοποιήσετε το [πρότυπο](#) αρχείο word που δίνεται από την Elsevier. **Σημείωση:** Στα επιστημονικά άρθρα δεν βάζουμε σελίδα περιεχομένων.

Για το άρθρο σας θα πρέπει να μελετήσετε τις τρέχουσες τεχνολογίες στο θέμα που επιλέξατε και να κάνετε ποιοτική σύγκριση μεταξύ τους, αναφέροντας πλεονεκτήματα και μειονεκτήματα της κάθε μίας. Η δομή του κυρίως σώματος του άρθρου, μπορεί να διαφέρει ανάλογα με τα υπο-θέματα (subtopics) ή τα ερωτήματα που εξετάζονται. Για παράδειγμα, εάν εξετάζετε τρεις διαφορετικές μεθοδολογίες, μπορείτε να διαιρέσετε το κυρίως σώμα του άρθρου σε τρεις ενότητες, καθεμία από τις οποίες συζητά μία από τις μεθόδους. Σε αυτές τις ενότητες, φροντίστε να περιγράψετε τις ερευνητικές μεθόδους και να αξιολογήσετε τον τρόπο διεξαγωγής των μελετών εστιάζοντας στον σχεδιασμό και την ανάλυση της μελέτης, π.χ. την πρόθεση και τον τρόπο αντιμετώπισης αντίθετων προσεγγίσεων/αποτελεσμάτων, τη σύγκριση των μελετών και σχολιασμό για τα αποτελέσματα και τις επιπτώσεις τους.

Παράλληλα, στο πλαίσιο αξιολόγησης του εργαστηριακού μέρους του μαθήματος θα πρέπει να υλοποιηθούν τουλάχιστον τρεις αλγόριθμοι που αφορούν στο θέμα που έχετε επιλέξει εφαρμόζοντάς τους σε δύο τουλάχιστον σύνολο δεδομένων (π.χ. ένα Αγγλικό και ένα Ελληνικό ή δύο Αγγλικά). Τα συγκριτικά αποτελέσματα θα πρέπει να συμπεριληφθούν και να σχολιαστούν στην αντίστοιχη ενότητα.

Στην ενότητα Σύνδεσμοι στο eclass μπορείτε να βρείτε βίντεο με διάλεξη για το πως γίνεται η βιβλιογραφική ανασκόπηση.

Στην παραδοτέα **εργασία-άρθρο** που έχει την μορφή **pdf**, θα συμπεριλαμβάνεται ο τίτλος εργασίας, ΑΜ και ονοματεπώνυμο φοιτήτριας/τή. Επίσης, θα πρέπει να ετοιμάσετε **και να καταθέσετε μαζί με την εργασία-άρθρο**, ένα αρχείο **PowerPoint** για την προφορική παρουσίαση/εξέταση διάρκειας 10'-15' καθώς και τον κώδικα με τα **σύνολα δεδομένων** που χρησιμοποιήσατε για το εργαστηριακό μέρος.

<p>Ημερομηνία κατάθεσης στο eclass: 15/01/2023 Ημερομηνία παρουσιάσεων/εξέτασης: 18/01/2023</p>
--

Κριτήρια αξιολόγησης

Τα κριτήρια αξιολόγησης της εργασίας σας είναι τα ακόλουθα:

- | | |
|--|-----|
| • Ποιότητα και ποσότητα βιβλιογραφίας και σωστή χρήση αναφορών στο κείμενο | 30% |
| • Δομή εργασίας και κριτική ανάλυση | 20% |
| • Επιστημονικό ύφος γραφής | 10% |
| • Πληρότητα ανάπτυξης θέματος | 10% |
| • Ανάλυση αποτελεσμάτων | 10% |
| • Προφορική εξέταση | 20% |