



Οδηγίες για την υποβολή

- Το cover sheet θα πρέπει να είναι το πρώτο φύλλο στην εργασία σας όταν την παραδώσετε
- Όλες οι σελίδες θα πρέπει να είναι αριθμημένες με συνεχόμενη αύξουσα σειρά
- Το σύνολο της εργασίας θα πρέπει να είναι έτοιμο για υποβολή κατά την άφιξη στο Registry. Καμία φόρμα δεν θα είναι διαθέσιμη για τους φοιτητές σε έντυπη μορφή από το Registry

| | |
|--|---|
| Module code | DS7003 |
| Module title | Advanced Decision Making: Predictive Analytics & Machine Learning |
| Module leader | Stylianos Kavalaris |
| Assignment tutor | Stylianos Kavalaris |
| Τίτλος εργασίας | Σύγκριση κατηγοριών αλγορίθμων ML – Άσκηση λήψης αποφάσεων σε επιλογή από ορισμένα dataset. |
| Αριθμός Εργασίας | 1 (6000 λέξεις) |
| Βαρύτητα | 100% |
| Ημερομηνία Ανάθεσης | 5 Μαΐου 2020 |
| Ημερομηνία Υποβολής | Έως και 14 Ιουνίου 2020 23:59 |
| Μαθησιακά Αποτελέσματα που αξιολογούνται από την εργασία (δείτε το εγχειρίδιο του μαθήματος) | 1,2,3,4,5,6,7 και 8 |
| Απαιτήση για την υποβολή μέσω Turnitin | OXI |
| Additional information | Ανατροφοδότηση Εργασίας – Η ανατροφοδότηση για την εργασία σας θα είναι διαθέσιμη σε τέσσερις εργάσιμες εβδομάδες από την ημέρα υποβολής. Παρακαλούμε ανατρέξτε στις σελίδες του μαθήματος στο UEL+ για περαιτέρω λεπτομέρειες. |

Είδος εργασίας:☒ Ατομική Εργασία☐ Ομαδική Εργασία

Για τις **ομαδικές εργασίες** που απαιτούν την υποβολή και ατομικών και ομαδικών τμημάτων της εργασίας, η εργασία θα πρέπει να υποβληθεί ως:

☐ Ένα ενοποιημένο έγγραφο☐ Διαχωρισμένο για κάθε μέλος**Αριθμός απαιτούμενων αντιγράφων της εργασίας:**

1



2



Άλλο

Διευκρινίστε

Η εργασία θα παρουσιαστεί με την ακόλουθη μορφή:

Υποβολή στην ιστοσελίδα



Συραμμένο απλά στην επάνω αριστερή γωνία



Βιβλιοδεσία με κόλλα



Βιβλιοδεσία με σπιράλ



Σε ντοσιέ A4 με κρίκου (χωρίς μηχανισμό)

Σημείωση: Για τους φοιτητές που υποβάλλουν εργασίες σε χαρτόνια A3/A2, η εργασία θα πρέπει να καλύπτεται με κατάλληλη προστατευτική θήκη για την αποφυγή φθορών στην εργασία.

Ηλεκτρονικό αντίγραφο:

CD (επισυνάπτεται σε φάκελο ή ειδική θήκη στο τελευταίο φύλλο)



USB (επισυνάπτεται σε φάκελο ή ειδική θήκη στο τελευταίο φύλλο) Δεν



απαιτείται ηλεκτρονικό αντίγραφο

Σημείωση για όλους τους φοιτητές

Κάθε εργασία (ηλεκτρονική ή έντυπη κατάθεση) θα πρέπει απαραίτητως να συνοδεύεται από τις παρακάτω 2 φόρμες (με τη σειρά που εμφανίζονται):

1. Φόρμα UEL (Είναι προσυμπληρωμένη και διαθέσιμη στη ιστοσελίδα του μαθήματος στο Moodle)

2. Έντυπο Γνησιότητας (Λειτουργεί ως αποδεικτικό ότι ο εκάστοτε φοιτητής παρέδωσε την εργασία)

Το σύνολο της εργασίας θα πρέπει να είναι έτοιμο για υποβολή κατά την άφιξη στο Registry. Καμία φόρμα ή αναλώσιμα (συμπεριλαμβανομένων και των συρραπτικών) δεν θα είναι διαθέσιμα από το Registry. Αυτό σημαίνει ότι οι φοιτητές δεν θα μπορούν να συρράψουν καμία φόρμα στο Registry.

1. Essay :

Συγκρίνετε τους αλγόριθμους που βασίζονται σε δέντρα αποφάσεων (tree-based algorithms) με αυτούς των πιθανοτήτων υπό συνθήκη (conditional probability algorithms). Αναπτύξτε τον τρόπο λειτουργίας κάθε κατηγορίας, περιλαμβάνοντας ομοιότητες και διαφορές, πλεονεκτήματα και περιορισμούς αλλά και την χρήση τους. Οι απαντήσεις σας πρέπει να είναι τεκμηριωμένες, να στηρίζονται σε βιβλιογραφικές πηγές και να είναι εμφανές που γίνεται αναφορά σε αυτές και που εκφράζετε δικές σας θέσεις και συμπεράσματα.

[30 βαθμοί]

2. Decision Making Project

Μέσα από ένα project σας θα πρέπει να επιδείξετε ικανότητα στη χρήση Machine Learning, χρησιμοποιώντας γλώσσα Python (ή R).

Επιλέξτε ένα dataset (σύνολο δεδομένων)* από το UCI Machine Learning Repository στην ηλεκτρονική διεύθυνση <https://archive.ics.uci.edu/ml/datasets.php>

Θα πρέπει να πραγματοποιήσετε και να συγκρίνετε δύο μεθόδους εποπτευόμενης (supervised learning) ή / και μη εποπτευόμενης μάθησης (unsupervised learning) ή χρονικής πρόβλεψης (temporal forecasting) για να επιτύχετε το καλύτερο δυνατό αποτέλεσμα από την ανάλυση.

Η εργασία θα πρέπει να παρέχει σαφή κατανόηση του ιστορικού και των θεωρητικών θέσεων που υπογραμμίζουν την ανάλυση, αιτιολόγηση για τη μορφή της ανάλυσης που πραγματοποιήθηκε και τις τεχνικές που χρησιμοποιήθηκαν, καθώς και αξιολόγηση και παρουσίαση των αποτελεσμάτων. Θα πρέπει να υπάρχει βιβλιογραφία σύμφωνα με αποδεκτές ακαδημαϊκές συμβάσεις και πρότυπα. Η αναφορά σας πρέπει να περιλαμβάνει τις ακόλουθες ενότητες:

- Τίτλο
- Περίληψη (Abstract)
- Εισαγωγή που περιλαμβάνει εξήγηση του ιστορικού του θέματος και ανασκόπηση της σχετικής βιβλιογραφίας
- Κριτική έκθεση της μεθοδολογικής προσέγγισής σας.
- Περιγραφή των πηγών δεδομένων σας και από τι συνίστανται τα δεδομένα (διερευνητική ανάλυση)
- Παρουσίαση της ανάλυσης ML, των αποτελεσμάτων και της αξιολόγησής τους ή της ανάλυσης ευαισθησίας (sensitivity analysis)
- Συμπερασματική αξιολόγηση των ευρημάτων, και κριτική ανασκόπηση / σύγκριση των τεχνικών που χρησιμοποιήθηκαν
- Λίστα (References) με αναφορές σε στυλ Χάρβαρντ
- Προσάρτημα (Appendix) του κώδικα σε Python (ή R) που χρησιμοποιήσατε.

*Εξαιρούνται τα ακόλουθα dataset, τα οποία δεν μπορείτε να χρησιμοποιήσετε:

Breast Cancer Wisconsin (Diagnostic), Breast Cancer Wisconsin (Original), Breast Cancer Wisconsin (Prognostic), Challenger USA Space Shuttle O-Ring , Concrete Compressive Strength, Concrete Slump Test, Credit Approval, Letter Recognition, Mushroom, Semeion Handwritten Digit, SMS Spam Collection, Spambase, Wine Quality.

[70 βαθμοί]