



Μάθημα: ΜΕΘΟΔΟΙ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Ακαδημαϊκή περίοδος: Χειμερινό εξάμηνο 2022-2023

2^η Εργασία

Το αρχείο body.csv περιλαμβάνει δεδομένα σχετικά με την ηλικία, το βάρος, το ύψος και διάφορα σωματομετρικά χαρακτηριστικά 252 ανδρών. Συγκεκριμένα, περιλαμβάνονται οι παρακάτω μεταβλητές:

brozek	Ποσοστό σωματικού λίπους βάσει της εξίσωσης Brozek (457/density-414.2)
density	Σωματική πυκνότητα (gm/cm ³)
age	Ηλικία (έτη)
weight	Βάρος (λίβρες)
height	Ύψος (ίντσες)
adipos	Δείκτης παχυσαρκίας (kg/m ²)
neck	Περίμετρος λαιμού (cm)
chest	Περίμετρος στήθους (cm)
abdom	Περίμετρος μέσης (cm)
hip	Περίμετρος ισχίου (cm)
thigh	Περίμετρος μηρού (cm)
knee	Περίμετρος γονάτου (cm)
ankle	Περίμετρος αστραγάλου (cm)
biceps	Περίμετρος δικέφαλου (cm)
forearm	Περίμετρος πήχης (cm)
wrist	Περίμετρος καρπού (cm)

1. Εφαρμόστε κατάλληλες μεθόδους για την επιλογή του βέλτιστου υποσυνόλου επεξηγηματικών μεταβλητών για την πρόβλεψη του σωματικού λίπους.
2. Αξιολογήστε την εφαρμογή του μοντέλου στο οποίο καταλήξατε.
3. Χρησιμοποιώντας τις επεξηγηματικές μεταβλητές στις οποίες καταλήξατε στο ερώτημα (1), εφαρμόστε τεχνικές ταξινόμησης (λογιστική παλινδρόμηση, γραμμική διακριτική ανάλυση, τετραγωνική διακριτική ανάλυση, μέθοδος του κοντινότερου γείτονα) για να προβλέψετε αν ένα άτομο έχει ποσοστό σωματικού λίπους μεγαλύτερο από 24(%). Περιγράψτε αναλυτικά τα αποτελέσματα των διαφορετικών μεθόδων και επιλέξτε τη μέθοδο που κατά την άποψή σας δίνει τα καλύτερα αποτελέσματα.

Τα παραδοτέα της εργασίας περιλαμβάνουν:

- Το script file με τον κώδικά σας,

- Μία γραπτή αναφορά που θα περιλαμβάνει **αιτιολογημένες** απαντήσεις στα παραπάνω ερωτήματα και μία παράγραφο που να συνοψίζει τις τελικές σας σκέψεις και συμπεράσματα. Όπως σε κάθε ανάλυση, θα πρέπει προφανώς να ξεκινήσετε και να συμπεριλάβετε στην παρουσίαση των αποτελεσμάτων σας βασικά περιγραφικά μέτρα και γραφήματα για όλες τις μεταβλητές.

Η εργασία θα βαθμολογηθεί με άριστα το 10 και θα μετρήσει κατά 10% στον τελικό σας βαθμό.

Η εργασία θα πρέπει να αναρτηθεί στο eclass μέχρι τη Δευτέρα 6 Φεβρουαρίου 2023 στις 23:00. Καμία εργασία δε θα γίνει δεκτή μετά από τη συγκεκριμένη ημερομηνία και ώρα.

Καλή επιτυχία!