

ΑΣΚΗΣΗ: Ανάλυση και Πρόβλεψη Τιμής Ακινήτων στην Ελλάδα

Δίνεται το αρχείο `greece_housing_data.xlsx`, το οποίο περιέχει στοιχεία αγοραπωλησιών ακινήτων στην Ελλάδα.

Μέρος Α: Κατανόηση & Προεπεξεργασία Δεδομένων

- Φορτώστε το αρχείο σε pandas DataFrame και εμφανίστε:
 - τις πρώτες 5 γραμμές
 - τον αριθμό γραμμών και στηλών
 - τους τύπους δεδομένων (dtypes)
- Θα χρησιμοποιηθούν οι εξής στήλες:

Ανεξάρτητες μεταβλητές (features):

- Τιμή Ζώνης
- Επιφάνεια Κύριων Χώρων (σε τ.μ.)
- Επιφάνεια Βοηθητικών Χώρων (σε τ.μ.)
- Έτος Κατασκευής
- Όροφος
- Κατηγορία Ακινήτου

Εξαρτημένη μεταβλητή (target):

- Τίμημα Δικαιώματος
- Εντοπίστε και διαχειριστείτε:
 - Κενές τιμές (NaN)
 - Ακραίες τιμές (outliers) στο Τίμημα Δικαιώματος
- Μετατρέψτε:
 - Το Έτος Κατασκευής σε **ηλικία ακινήτου**:
Ηλικία=2025-Έτος Κατασκευής
- Μετατρέψτε τις κατηγορικές μεταβλητές σε αριθμητικές:

- Κατηγορία Ακινήτου → One-Hot Encoding
- Εφαρμόστε **StandardScaler** ή **MinMaxScaler** στα αριθμητικά χαρακτηριστικά.

Μέρος Β: Linear Regression

Ορισμός μοντέλου

- Χωρίστε τα δεδομένα σε **train (80%)** και **test (20%)**
 - Εκπαιδεύστε το μοντέλο **Linear Regression**
-

Αξιολόγηση μοντέλου

Υπολογίστε:

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE)
- R² score

Σχολιάστε:

- Πόσο καλά προβλέπει το μοντέλο; Ποιες μεταβλητές (γνωρίσματα) φαίνεται να επηρεάζουν περισσότερο την τιμή;
- Εμφανίστε τους συντελεστές του μοντέλου και εξηγήστε την οικονομική σημασία τους (π.χ. επίδραση επιφάνειας ή τιμής ζώνης)
- Έχει νόημα να κάνουμε ομαλοποίηση Ridge/Lasso ?

Μέρος Γ (Προαιρετικό): Logistic Regression

Γ1. Δημιουργία δυαδικής μεταβλητής

Ορίστε νέα μεταβλητή την Ακριβό Ακίνητο που ορίζεται ως:

Ακριβό Ακίνητο = 1 αν Τίμημα Δικαιώματος > διάμεσος

Ακριβό Ακίνητο = 0 αλλιώς

Γ2. Εκπαίδευση μοντέλου

- Εκπαιδεύστε **Logistic Regression**

- Χρησιμοποιήστε τα ίδια χαρακτηριστικά με πριν
-

Γ3. Αξιολόγηση ταξινόμησης

Υπολογίστε:

- Accuracy
- Precision, Recall, F1-score
- Confusion Matrix
- ROC curve, AUC

Σχολιάστε:

- Πόσο καλά διαχωρίζει τα ακριβά από τα μη ακριβά ακίνητα;
 - Ποια χαρακτηριστικά αυξάνουν την πιθανότητα ένα ακίνητο να είναι «ακριβό»;
-