

Predictii imobiliare

Despre anumite imobile, se cunosc informatii precum: nr bailor, nr dormitoarelor, pretul, suprafata locuibila, suprafata terenului aferent imobilului, nr de etaje, codul postal, conditia, anul construirii, anul ultimei renovari, etc. (a se consulta informatiile din *homeData.csv*)

- a. (300p) Stabiliti suprafata locuibila maxima a imobilelor care au un numar de dormitoare mai mare decat media tuturor imobilelor.
- b. (1000p) Sa se stabileasca eroare medie patratica (mean squared error) a unui model de predictie a pretului imobilelor care foloseste 4 attribute (['bedrooms', 'bathrooms', 'sqft_living', 'sqft_basement']), baza de antrenare fiind formata din primele 1000 de exemple, iar cea de testare din urmatoarele 50 de exemple din fisierul *homeData.csv*.
- c. (1000p) In cazul folosirii unui nou set de date (*homeDataNew.csv*), ce eroare medie patratica ar obtine modelul de predictie precedent (in acelasi settings)? Justificati raspunsul.