

Task 3_3

hauses.csv ფაილის საფუძველზე რომელშიც მოცემულია ინფორმაცია უძრავ ქონებაზე ამერიკის ერთ რომელიღაც ქალაქში. (ფაილი შეგიძლიათ უკვე ნახოთ ჩვენს დრაივზე მეცხრე ლექციის მასალებში) დაატრენინგეთ რეგრესიის მოდელები LinearRegression, DecisionTreeRegressor, RandomForestRegressor რომლებიც ისწავლიან უძრავი ქონების ფასის პროგნოზირებას შესაბამისი პარამეტრების მიხედვით.

პარამეტრების აღწერა მონაცემთა ბაზაში:

Attribute Information

Input features in order:

- 1) CRIM: per capita crime rate by town
- 2) ZN: proportion of residential land zoned for lots over 25,000 sq.ft.
- 3) INDUS: proportion of non-retail business acres per town
- 4) CHAS: Charles River dummy variable (1 if tract bounds river; 0 otherwise)
- 5) NOX: nitric oxides concentration (parts per 10 million) [parts/10M]
- 6) RM: average number of rooms per dwelling
- 7) AGE: proportion of owner-occupied units built prior to 1940
- 8) DIS: weighted distances to five Boston employment centres
- 9) RAD: index of accessibility to radial highways
- 10) TAX: full-value property-tax rate per \$10,000 [\$ / 10k]
- 11) PTRATIO: pupil-teacher ratio by town
- 12) B: The result of the equation $B = 1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town
- 13) LSTAT: % lower status of the population

Output variable:

- 1) MEDV: Median value of owner-occupied homes in \$1000's [k\$]

Source

StatLib - Carnegie Mellon University

ამ შემთხვევაში მოდელმა უნდა ისწავლოს საშუალო ფასის პროგნოზირება რომელიც არის წარმოდგენილი 'MEDV' სვეტში.

Pandas ბიბლიოთეკის დახმარებით წაიკითხეთ და დაამუშავეთ მონაცემთა ცხრილი. დამუშავების ნაწილში 'MEDV' სვეტს დაარქვით უფრო ლოგიკური სახელი და მისი ყველა ერთეული გაამრავლეთ 1000-ზე

შესაბამისი მოდელები დააიმპორტეთ შემდეგნაირად:

```
from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
```

შეგიძლიათ დაიხმაროთ ის კოდები რომლებიც გუშინ გავარჩეთ როცა რეგრესიის მოდელებს ვატრენინგებდით Netflix კომპანიის აქციების ფასის საპროგნოზირებლად. შესაბამისი .py სკრიპტი და jupyter რვეული უკვე გაქვთ დრაივზე 9 ლექციის მასალებში.

მოდულები ასევე შეაფასეთ საშუალო კვადრატული გადახრის და RR მნიშვნელობის მიხედვით, როგორც ლეციაზე გავაკეთეთ პირდაპირ აიღეთ sklearn ბიბლიოთეკიდან მათი მოდულები

```
from sklearn.metrics import mean_squared_error, r2_score
```

შეაფასეთ ერთმანეთის მიმართ, რამდენად სწორია სამივე მოდელი ამ ორი სტატისტიკის მიხედვით

გაგზავნეთ თქვენი მოდელის კოდი(გინდაც პითონის ფაილი, გინდაც jupyter-notebook) და შედარების მეტრიკები/აღწერილობის ფაილი იმეილზე igugunashvili@cu.edu.ge

subject AI&ML_task3_3 მიუთითეთ

რადგან ჩვენი ლეციები დასასრულს მიუახლოვდა ეცადეთ დავალება გამომიგზავნოთ 5 ნოებრამდე.

თუ რამე კითხვები გექნებათ ნებისმიერ დროს შეგიძლიათ მომწეროთ რა პლატფორმაზეც გირჩევნიათ (fb/skype/mail)