گزارش تمرین (2)

ابتدا برای شروع کار کتابخانه هایی که به آنها نیاز بود از جمله:pandas-numpy-math-sklearn و... را انتخاب و اضافه کردم.

در ادامه با استفاده از لینک درایو،دیتاست مربوطه فراخوانی میشود.

3 فیچر مورد نظرانتخاب میشوند وخروجی بصورت رندوم 5 لاین از سه ستون موردنظر هست.فیچر قیمت بعنوان y تعیین شده است.

در کد5 تعداد اصلی ورودی هارا درمیتوان مشاهده کرد.

کد6:دیتا2 را این بار با چند ویژگی دیگر انتخاب کردم تا در قسمت سوم تمرین استفاده کنم.مربوط یه سوال سه.

دیتا اینفو اطلاعات مربوط به فیچر ها و تعداد و .. را نمایش میدهد.

درادامه با استفاده از mean ,mode مقادیر خالی را سعی میکنیم حذف کنیم.

* Convert string column to numeric

مقادیرstring مثل مدل گرمایش خانه به مقادیر عددی تبدیل میشوند.کد 9 انواع آنها و کد10 خروجی عددی را نمایش میدهد.

* Split Data

دیتاست توسط متد split به دو قسمت شامل train و test تقسیم میشود.

## Prepare train and test for model

## تعیین x\_testو y\_test

## تعیین x\_trainو y\_train

## که x بصورتی قرارداده شده که ستون اول فقط یک درنظر گرفته میشود.

## همین روال برای data2 هم انجام میشود.

## در ادامه یک مدل رگرسیون طراحی شده با استفاده از تابع خطای MSE

## مقدار delta درواقع مقادیر آپدیت شده ضرایب ما هستند.

## Predict with my model

## این قسمت هم با استفاده از فرمول و مقادیر ثابت c[0] c[1] c[2] مقادیرنهایی y را بدست میدهد و در خروجی میتوان آنها رادید.

## Train regression model with sklearn

## با استفاده از پکیج sklearn یک با دیگر مدل را مقایسه میکند.

## ودرنهایت مقادیر y جدید را درخروجی خواهیم داشت و میتوان مقایسه بین دو مدل انجام داد.

## در ادامه با استفاده از یک پلات،با سه رنگ نارنجی،آبی و سبز نقاط بدست آمده را مشخص کردم که نقاط آبی مقادیر واقعی هستند.

## نقاط نارنجی و سبز مقادیر رگرسیون و sklearn هستند و همینطور که درتصویر مشخص است این دو مقدار بسیار نزدیک به هم و برابر هستند.

## و در قسمت بعدی هم رگرسیون هایlasso و ridge پیاده سازی شده اند و با مقایسه نتایج میتوان گفت تفاوت آنها بسیار کم بوده و مقادیر خروجی تقریبا برابر هستند.

## این دو رگرسیون روی دیتا2 هم اجرا شده اند و نتایج آنها هم بصورت پلات نمایش داده شده است.

## در قسمت آخر تسک امتیازی Absolute Error پیاده سازی شده و باز هم بصورت نمودار نتایج نهایی را نمایش داده ام.میتوان مشاهده کرد که مقادیر با مقادیر sklearn تقریبا نزدیک .برابر هستند.

## 