تمرین_ تعداد درخواست جهت اجاره دوچرخه

به عنوان تمرین دوم، از شما می خواهیم با استفاده از مجموعه داده مربوط به تعداد درخواست جهت اجاره دوچرخه ($^{\circ}$ عدد فایل ضمیمه شده) با استفاده از آموختههایتان تا این مرحله از دوره، $^{\circ}$ و بصری سازی را اجرا نمایید. سوال اصلی این است که مدلی جهت پیش بینی تعداد درخواستها با استفاده از این دیتاست تهیه نمایید.









توضیحات فیچرها به صورت زیر میباشد:

datetime - hourly date + timestamp

season - 1 = spring, 2 = summer, 3 = fall, 4 = winter

holiday - whether the day is considered a holiday

workingday - whether the day is neither a weekend nor holiday

weather - 1: Clear, Few clouds, Partly cloudy, Partly cloudy

2: Mist + Cloudy, Mist + Broken clouds, Mist + Few clouds, Mist

3: Light Snow, Light Rain + Thunderstorm + Scattered clouds, Light Rain + Scattered clouds

4: Heavy Rain + Ice Pallets + Thunderstorm + Mist, Snow + Fog

temp - temperature in Celsius

atemp - "feels like" temperature in Celsius

humidity - relative humidity

windspeed - wind speed

casual - number of non-registered user rentals initiated

registered - number of registered user rentals initiated

count - number of total rentals

قدم اول:

• Import نمودن كتابخانه هاى مورد نياز (كتابخانه هايى نظير Numpy و Pandas و ...)

قدم دوم:

- بارگذاری و خواندن دیتاست با استفاده از کتابخانه Pandas
 - اکتشاف دیتاست و بررسی داده به صورت کامل (EDA)

قدم سوم:

• شناسایی دادههای گم شده و Missing Value ها

قدم چهارم:

• ترسیم نمودارهای مختلف جهت درک بهتر از دادهها با استفاده از Seaborn و یا کتابخانههای دیگر

قدم پنجم:

- آموزش مدل با استفاده از روش های رگرسیون آموخته شده :(با استفاده از فایل train.csv)
 - Linear Regression o
 - Multiple Regression o
 - KNN Regressor o
 - Decision Tree Regressor o
 - Random Forest Regressor o
 - ... 0

قدم ششم:

• ذخیره مدلهای نهایی و آپلود آنها در سایت مکتبخونه به همراه کد نهایی در قالب یکJupyter Notebook

اختیاری) قدرت الگوریتم خود را بر روی دادههای test.csv اعمال نموده و جواب را گزارش دهید.

نمونه خروجی نهایی در فایل sampleSubmission.csv موجود می باشد.