

**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**



## ÇOX XƏTTİ REQRESSİYA

Çoxlu reqressiya xətti reqressiya kimidir, lakin birdən çox müstəqil dəyərə malikdir, yəni biz iki və ya daha çox dəyişənə əsaslanan dəyəri proqnozlaşdırmağa çalışırıq. Aşağıdakı verilənlər bazasına nəzər salın, orada avtomobillər haqqında bəzi məlumatlar var.



**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**

Biz avtomobilin CO2 emissiyalarını mühərrikin ölçüsünə əsasən proqnozlaşdırma biləri, lakin çoxlu reqressiya ilə proqnozu daha dəqiq etmək üçün avtomobilin çəkisi kimi daha çox dəyişən əlavə edə bilərik. Python-da bizim üçün işi görəcək modullarımız var.

Pandas modulunu import etməklə başlayır.

İmport pandas

Pandas modulu bizə csv fayllarını oxumağa və DataFrame obyektini qaytarmağa imkan verir.

```
df = pandas.read_csv("maşınlar.csv")
```

Sonra müstəqil dəyərlərin siyahısı hazırlanır və bu dəyişən X adlanır. Eynilə, asılı dəyərlər y adlı dəyişənə qoyulur.

```
X = df[['Weight', 'Volume']]
```

```
y = df['CO2']
```

İpucu: Müstəqil dəyərlərin siyahısını böyük X hərfi ilə və asılı dəyərlər siyahısını y hərfi ilə adlandırmaq adi haldır. Biz Sklearn modulundan bəzi üsullardan istifadə edəcəyik, ona görə də həmin modulu da import etməliyik.

```
from sklearn import linear_model
```

Sklearn modulundan xətti reqressiya obyektini yaratmaq üçün LinearRegression() metodundan istifadə edəcəyik. Bu obyektə müstəqil və asılı dəyərləri parametr kimi qəbul edən və reqressiya obyektini əlaqəni təsvir edən datalarla dolduran fit() adlı metod var.

```
regr = linear_model.LinearRegression()
```

```
regr.fit(X,y)
```

İndi avtomobilin çəkisi və həcminə əsaslanaraq CO2 dəyərlərini proqnozlaşdırmağa hazır bir reqressiya obyektimiz var.

*Çəkisi 2300 kq və həcmi 1300 sm3 olan avtomobilin CO2 emissiyasını qiymətləndirək.*

```
predictedCO2 = regr.predict([[2300,1300]])
```

**MÜƏLLİF:** İBRAHİM İSMAYİL

**MƏQALƏ** COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.

**TƏTBİQ** PYTHON

NÜMUNƏNİ TAM OLARAQ PYTHON-DA TƏTBİQ EDƏK:

```
import pandas
from sklearn import linear_model

# CSV faylından məlumatları oxunması
df = pandas.read_csv("maşınlar.csv")

# Xüsusiyyətləri (müstəqil dəyişənlər) və hədəf dəyişəni (asılı dəyişən) müəyyənləşdirək.
X = df[['Weight', 'Volume']]
y = df['CO2']

# Xəttili reqressiya modeli yaradaq.
regr = linear_model.LinearRegression()

# Xüsusiyyətlərdən və hədəf dəyişəndən istifadə edərək modeli uyğunlaşdıraraq.
regr.fit(X, y)

# Çəkisi 2300 kq, həcmi isə 1300 sm3 olan avtomobilin CO2 emissiyasını proqnozlaşdıraraq.
predictedCO2 = regr.predict([[2300, 1300]])
print(predictedCO2)

[107.2087328]
```

1.3 litr mühərriki və çəkisi 2300 kq olan avtomobilin getdiyi hər kilometrə təxminən 107 qram CO2 buraxacağı təxmin edilmişdir.

Əmsal:

Əmsal naməlum dəyişənlə əlaqəni təsvir edən amildir.

Misal: Əgər  $x$  dəyişəndirsə,  $2x$  iki dəfə  $x$ -dir.  $x$  məchul dəyişən,  $2$  rəqəmi isə əmsaldır.

Bu halda, CO2-yə qarşı çəkinin əmsal dəyərini və CO2-yə qarşı həcmnin dəyərini istəyə bilərik. Aldığımız cavablar müstəqil dəyərlərdən birini artırısaq və ya azaltsaq nə baş verəcəyini bildirir.

**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**

Misal: Regressiya obyektinin əmsal dəyərlərinin tapılması.

```
import pandas
from sklearn.linear_model import LinearRegression

df = pandas.read_csv("maşınlar.csv")
X = df[['Weight', 'Volume']]
y = df['CO2']

regr = LinearRegression()
regr.fit(X, y)

print(regr.coef_)

[0.00755095 0.00780526]
```

Nəticə Təsviri:

Yaranan massiv çəki və həcmə əmsal dəyərlərini təmsil edir.

Çəki: 0.00755095

Həcmi: 0.00780526

Bu dəyərlər bizə deyir ki, çəki 1 kq artarsa, CO2 emissiyaları 0,00755095 qram artır.

Mühərrikin həcmi (Həcmi) 1 sm<sup>3</sup> artarsa, CO2 emissiyası 0,00780526 qram artır.

Düşünürəm ki, bu, ədalətli bir qiymətləndirmədir, amma gəlin bunu sınaqdan keçirək!

Biz artıq təxmin etmişik ki, 1300 sm<sup>3</sup> mühərriki olan avtomobilin çəkisi 2300 kq olarsa, onun CO2 emissiyası təxminən 107 qram olacaq.

Çəkisi 1000 kq artırısaq nə olar?

Nümunə: Əvvəlki nümunəni kopyalayaq, lakin çəkisini 2300-dən 3300-ə dəyişdirək.

```
import pandas
from sklearn import linear_model
df = pandas.read_csv("maşınlar.csv")
X = df[['Weight', 'Volume']]
y = df['CO2']
regr = linear_model.LinearRegression()
regr.fit(X, y)
predictedCO2 = regr.predict([[3300, 1300]])
print(predictedCO2)

[114.75968007]
```

**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**

1.3 litrlik mühərriki və çəkisi 3300 kq olan avtomobilin getdiyi hər kilometr üçün təxminən 115 qram CO2 buraxacağını təxmin etdik. Bu 0.00755095 əmsalının düzgün olduğunu göstərir:

$$107.2087328 + (1000 * 0.00755095) = 114.75968$$

### **Scaling Features:**

Datalarınız fərqli dəyərlərə və ya hətta fərqli ölçü vahidlərinə malik olduqda, onları müqayisə etmək çətin ola bilər. Bir kiloqram bir metrle müqayisədə nədir? Yoxsa zamanla hündürlük? Bu sualın cavabı miqyasdır. Məlumatları müqayisə etmək daha asan olan yeni dəyərlərə miqyaslaya bilərik. Aşağıdakı cədvələ baxın, eyni data seti, lakin bu dəfə həcm sütununda sm3 əvəzinə litrlə dəyərlər var. Misal üçün (1000 əvəzinə 1.0).

1.0-ın həcmi 790-ın çəkisi ilə müqayisə etmək çətin ola bilər, lakin hər ikisini müqayisə edilə bilən dəyərlərə çevirsək, bir dəyərin digəri ilə necə müqayisə olunduğunu asanlıqla görə bilərik. Verilənləri miqyaslaşdırmaq üçün müxtəlif üsullar var, biz bu standartlaşdırma adlı metoddan istifadə edəcəyik. Standartlaşdırma metodu düsturdan istifadə edirik:

$$z = (x - u) / s$$

burada z yeni qiymət, x ilkin dəyər, u orta, s standart kənarlaşmadır. Yuxarıdakı verilənlər toplusundan çəki sütununu götürsəniz, ilkin dəyər 790, miqyaslı dəyər isə belə olacaq:  $(790 - 1292.23) / 238.74 = -2.1$

Əgər yuxarıdakı verilənlər toplusundan həcm sütununu götürsəniz, ilkin dəyər 1.0, miqyaslı dəyər isə belə olacaq:  $(1.0 - 1.61) / 0.38 = -1.59$

İndi 790-ı 1.0 ilə müqayisə etmək əvəzinə -2.1-i -1.59 ilə müqayisə edə bilərsiniz. Bunu əl ilə etməyə ehtiyac yoxdur, Python sklearn modulunda verilənlər toplusunu dəyişdirmək üsulları ilə Scaler obyektini qaytaran StandardScaler() adlı metod var.

Misal: Çəki və Həcmi sütunlarındakı bütün dəyərləri ölçün.

```
import pandas
from sklearn import linear_model
from sklearn.preprocessing import StandardScaler
scale = StandardScaler()
df = pandas.read_csv("maşınlar.csv")
X = df[['Weight', 'Volume']]
scaledX = scale.fit_transform(X)
print(scaledX)
```

**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**

```
[[-2.10389253 - 1.59336644]
[-0.55407235 - 1.07190106]
[-1.52166278 - 1.59336644]
[-1.78973979 - 1.85409913]
[-0.63784641 - 0.28970299]
[-1.52166278 - 1.59336644]
[-0.76769621 - 0.55043568]
[ 0.3046118 - 0.28970299]
[-0.7551301 - 0.28970299]
[-0.59595938 - 0.0289703 ]
[-1.30803892 - 1.33263375]
[-1.26615189 - 0.81116837]
[-0.7551301 - 1.59336644]
[-0.16871166 - 0.0289703 ]
[ 0.14125238 - 0.0289703 ]
[ 0.15800719 - 0.0289703 ]
[ 0.3046118 - 0.0289703 ]
[-0.05142797 1.53542584]
[-0.72580918 - 0.0289703 ]
[ 0.14962979 1.01396046]
[ 1.2219378 - 0.0289703 ]
[ 0.5685001 1.01396046]
[ 0.3046118 1.27469315]
[ 0.51404696 - 0.0289703 ]
[ 0.51404696 1.01396046]
[ 0.72348212 - 0.28970299]
[ 0.8281997 1.01396046]
[ 1.81254495 1.01396046]
[ 0.96642691 - 0.0289703 ]
[ 1.72877089 1.01396046]
[ 1.30990057 1.27469315]
[ 1.90050772 1.01396046]
[-0.23991961 - 0.0289703 ]
[ 0.40932938 - 0.0289703 ]
[ 0.47215993 - 0.0289703 ]
[ 0.4302729 2.31762392]]
```

**MÜƏLLİF: İBRAHİM İSMAYİL**

**MƏQALƏ COP 29-LA BAĞLI MAARİFLƏNDİRMƏ MƏQSƏDİ İLƏ YAZILMIŞDIR.**

**TƏTBİQ PYTHON**

### **CO2 (Karbon Qazı) Dəyərlərinin Qiymətləndirilməsi:**

Tapşırıq yalnız çəkisini və həcmi bildiyiniz bir avtomobildən CO2 emissiyalarını qiymətləndirmək idi. Data seti miqyaslandıqda, siz dəyərləri qiymətləndirərkən miqyasdan istifadə etməlisiniz.

Misal: 2300 kiloqram ağırlığında 1.3 litrlik avtomobilin CO2 emissiyasını qiymətləndirək.

```
import pandas
from sklearn import linear_model
from sklearn.preprocessing import StandardScaler

scale = StandardScaler()
df = pandas.read_csv("maşınlar.csv")
X = df[['Weight', 'Volume']]
y = df['CO2']
scaledX = scale.fit_transform(X)

regr = linear_model.LinearRegression()
regr.fit(scaledX, y)

scaled = scale.transform([[2300, 1.3]])
predictedCO2 = regr.predict([scaled[0]])

print(predictedCO2)
```

[97.07204485]

### **Modelin Qiymətləndirilməsi:**

Maşın Öyrənməsində modellər müəyyən hadisələrin nəticəsini proqnozlaşdırmaq üçün qurulur, məsələn, avtomobilin çəkisini və mühərrik ölçüsünü bildiyimiz zaman onun CO2 emissiyalarını proqnozlaşdırdığımız olduğu kimi olacaq. Modelin kifayət qədər yaxşı olub-olmadığını ölçmək üçün Train/Test adlı metoddan istifadə edilir.