**Kod**

metin içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Kullanılan Kütüphaneler**

***Numpy:*** NumPy, dizilerle çalışmak için kullanılan bir Python kütüphanesidir. Ayrıca doğrusal cebir, fourier dönüşümü ve matrisler alanında çalışmak için de gerekli işlevlere sahiptir. Python’da dizilerin amacına hizmet eden listelerimiz var, ancak işlenmesi yavaştır. NumPy, geleneksel Python listelerinden 50 kata kadar daha hızlı bir dizi nesnesi sağlamayı amaçlamaktadır. NumPy’deki dizi nesnesi (array) ndarray olarak adlandırılır ve ndarray, çalışmayı çok kolaylaştıran birçok destekleyici işlev sağlar. Diziler, hız ve kaynakların çok önemli olduğu veri biliminde çok sık kullanılır.

***Pandas:*** Pandas ile bir veri bilimi projesindeki veri okuma, veri ön işleme ve veri temizleme aşamaları yapılır. Verinin çok büyük hacimde olması ve farklı kaynaklardan gelmesi ile parazitli, eksik ve tutarsız veriler ortaya çıktı. Amerika’da 2010 yılında yapılan bir araştırmada veri kalitesi sorunları nedeni ile yılda 1 milyon dolardan fazla para kaybedildi. Veriyi topladıktan sonra veriyi analize uygun hale getirmek, veri analizi için oldukça önemlidir. Veri bilimcilerinin en fazla zamanlarını alan aşama veri ön işleme ve veri temizlemedir. Pandas kütüphanesi ile veri organize edilerek analizler daha kolay ve hızlı yapılır. Pandas Python’ın en önemli kütüphanelerinden biridir.

***Scikit-learn:*** Scikit-learn veya Sklearn makine öğrenmesi modelleri oluşturmak için kullanılan Python temelli bir kütüphanedir. Regresyon, kümeleme ve sınıflandırma için kullanılan pek çok öğrenme algoritmasına sahiptir. Sklearn NumPy ve SciPy ile uyumludur. Yani farklı Python kütüphaneleri ile kolaylıkla çalışabilir.

***Matplotlib:*** Günümüzün altın cevheri olarak gösterilen veri, popüleritesini artırmakta. Bu artış veri analizi, veri temizleme ve veri görselleştirme gibi alt dalları da gittikçe yaygın hale getirmekte. Veri bilimciliğinin son zamanlarda yaşadığı artış aynı şekilde programlama dillerinin ve kütüphanelerinin kullanım sıklığını da etkiledi. Bu kütüphanelerin başını Pandas, NumPy, Scikit-learn ve Matplotlib gibi Python kütüphaneleri çekmektedir. Bir veri yığınını analiz etmek için doğru soruları sorması gereken veri bilimci elindeki sonuçları da en iyi şekilde analiz edip raporlamalıdır. Bu analiz için en büyük silahı ise “veri görselleştirmedir“. Matplotlib bize verilerin görselleştirilmesini sağlar.

**Pandas:**

***.read Metodu:*** Dataseti okumaya yarar. Bir datasetin csv dosyasından işlenebilir hale gelmesini sağlar.

**Sklearn:**

***train\_test\_split***: train\_test\_split, dizilerimizi veya matrislerimizi rastgele subsetlere ve test datalarına ayırır. Sklearn ile test datası oluşturmak bu kadar kolay.

***.fit Metodu:*** .fit metodu modelimizin eğitimini gerçekleştiren ve bunun sonucunda bize fonksiyon parametrelerini en optimum şekilde çıkaran metoddur. Lineer Regresyonun formülünde olan bias ve weight değerlerini en optimum şekilde bulmaya yarayan fonksiyondur.

***Predict Fonksiyonu:*** Regresyon, sınıflandırma, kümeleme gibi yöntemler kullanarak yapacağımız çalışmalarda tahmin edilen etiket bilgisini predict fonksiyonuyla elde edebiliriz.

