

**Büyük Veri Proje Ödevi**  
**2022-2023**  
**Güz Dönemi**

**Grup üyeleri:**

Barış Yılmaz 2-A G191210303 xbarisyilmaz@gmail.com  
Cemal Aslan 2-A G191210387 cemas\_1999@outlook.com

**Grup üyelerinin projeye katkısı:**

Barış Yılmaz -> Hands-On Big Data Modeling kitabında bulunan Modeling Bitcoin Data Points With Python bölümünü gerekli uygulamaları kurarak gerçekleştirimi ve yapılan uygulamanın analizi, kod incelemesi ve işlemlerin araştırılması.

Cemal Aslan -> Yapılan uygulamadaki işlemlerin araştırılması, eğitim için açıklanmayarak yazılan kodların araştırılması.

**Projenin adı:** Modeling Bitcoin Data Points With Python

**Projenin amacı:**

Bitcoin verilerinden oluşturulabilecek farklı model türleri hakkında bilgi edinmek ve gelecekteki Bitcoin fiyatını tahmin etmek.

**Verinin açıklanması:**

Bitcoin'in 10 senelik tarihsel verilerini barındırmaktadır. İçerisinde bulunan sütunlar ise şunlardır;

- Timestamp (in Unix time)
- Open: Bitcoin price in currency units at the opening of the time period
- High: Highest Bitcoin price in currency units during the time period
- Low: Lowest Bitcoin price in currency units during the time period
- Close: Bitcoin price in currency units at the closing of the time period
- Volume\_(BTC): Volume of BTC transacted in the time period
- Volume\_(Currency): Volume of currency transacted in the time period
- Volume-weighted average price (VWAP)

**Kullanılan büyük veri platformları:** Jupyter Notebook

## **Projenin açıklanması:**

Açık kaynaklı araç ve kütüphanelerin bulunduğu yazılım olan Anaconda-Navigator üzerinden Jupyter Notebook kullanarak Python dilinde projeyi geliştirdik. Projede öncelikle bitcoin veri setimizi import ettik ve içerisinde bulunan nan satırları drop ettik. Ardından price değerlerini kullanarak önce bir istatistiğini görüntüledik. Sonrasında price'ın zamanda nasıl değiştiğinin ve bu setten önce 100 bin, ardından 50 bininin örneğini alarak bunları da makine öğrenmesine tabi tutup test verilerini ve grafiklerini yazdırdık. Sonrasında bu tahmin işlemini derin öğrenme yapısı olan RNN yapısını kullanarak yaptık. Bu yapıda veri setini import ettikten sonra öncelikle tarih Unix olarak tutulduğundan ve dakikaya göre değil de tarihe göre sıralama yapacağımız için tarihi düzenledik. Sonrasında tarihe göre gruplandırma yaparak bunların ortalamasını aldık. Son 1 aylık tahminde bulunmak için ve makine öğrenmesi için dataya normalleştirme, yeniden şekillendirme ve ölçeklendirme yaptık. Sonrasında RNN modeli oluşturduk, oluştururken bu modeli tanımladık gerekli katmanları ekleyip derledik. Son olarak ise RNN'i eğitim setine uydurduk ve ilk 100 girdisini ekrana yazdırdık. Ardından oluşturmuş olduğumuz bu RNN modelini kullanarak tahminde bulunmak için bugünün fiyatını kullanarak ertesi günün fiyatını tahmin ettik. Son işlem olarak ise gerçek bitcoin fiyatı ile tahmin edilen bitcoin fiyatını karşılaştırmak için bunların grafiğini çıkarttık.

## **Projenin çıktıları:**

10 sene gibi bir datayı barındıran Bitcoin verilerinin neye benzediğini ve çeşitli modeller oluşturmak için nasıl kullanılabileceğini ve buna ek olarak, iPython Notebook'taki RNN gibi çeşitli algoritmaları, bir zaman damgasına dayalı fiyatı tahmin etmek için bir Bitcoin veri kümesi üzerinde kullanma şansımız oldu.