

# Yapay Zeka

## Bitcoin Fiyat Tahmini

Mustafa Katipoğlu

16011084

# Bitcoin Fiyat Tahmini

Bitcoin fiyat tahmini, elimizdeki tarihsel bitcoin verilerini kullanarak gelecek bitcoin fiyatlarının saptanmasıdır. Örneğin 2013-2017 arası bitcoin fiyatlarının elimizde olduğu durumda 2017 yılı ocak ayında bitcoin fiyatları nasıl bir hal alacak gibi soruları cevaplamaya çalışacağız.

Veriseti: Tarihsel bitcoin fiyatları  
coinmarketcap.com ve blockchain.com  
sitelerindeki en güncel verilerinden  
oluşmaktadır.

# Time Series Prediction with Timebin Based Nearest Neighbourhood

## Time Series Tahmini:

Zamansal olarak ard arda gelen noktalardan oluşmuş verilere **time series** denir ve bu ard ard gelen veriler arasında ilişkiler olduğu düşünülür. Bitcoin tarihsel değer değişim verileride tam bir time series verisi oluşturduğu için bitcoin fiyat tahminimiz aslında elimizdeki time series verilerinin gelecek noktalarını tahmin etmekten oluşacak.

Elimizdeki time series verilerini tahmin ederken, hazır makine öğrenmesi algoritmaları kullanmak yerine kendim KNN algoritmasına bir türev olabilcek bir algoritmayı sıfırdan geliştirmeye çalıştım. Algoritma time series verisini farklı parçalara bölüp( bu parçalara bin diyoruz ) sonrasında her bin ' i incelediği ve bitcoin fiyat tahminleride bu bin ' lere benzer olan komşu binler üzerinden yaptığı için, tasarladığım metoda Timebin Based Nearest Neighbourhood adını verdim.

## **Timebin Based Nearest Neighbourhood :**

Timebin based nearest neighbourhood ile bitcoin fiyat tahmini bitcoin fiyatlarına etki eden faktörlerin, farklı zaman dilimlerinde benzer şekilde davranabiliceğine dayanan bir yöntem.

Bu yöntem elimizdeki tarihsel verilerden, fiyat tahmini yapmak istediğimiz zamana en yakın olan bölgeden bir dizi örneği alır ve bir timebin oluşturur. Timebin'ler aslında bir dizi örneğin fiyat değişimlerinden oluşan veri kümeleridir.

Amacımız tüm tarihsel veriler arasında bu timebin'e en benzer şekile davranmış veri kümelerini yada timebin'leri bulup, elde ettiğimiz komşu timebinler yardımıyla gelecek fiyat tahmini yapmaktan oluşuyor. Elde edilen komşu timebin'ler ile bitcoin fiyatını tahmin etmek istediğimiz zamana en yakın tarihsel verilere çok benzer şekilde davrandığı bulunduğu, bu verilerin bir sonraki adımlarda sahip olucakları değer değişimlerinde bitcoinin tahmin etmek istediğimiz bölgedeki değer değişimleriyle benzer olacağı düşüncesiyle bu komşu timebinler üzerinden geleceği tahmin edeceğiz.

## Timebin Based Nearest Neighbourhood Algoritması:

1. Bitcoin fiyatları tahmin edilmesi istenen tarihi seç.
2. Bu tarihe en yakın verileri timebin olarak al.
3. Seçilen timebin ile verisetinden oluşabilecek tüm olası timebin'ler arası benzerlikleri bul.
4. Benzer timebinler yardımıyla bitcoin fiyat değişimlerini tahmin et.
5. Fiyat değişimlerini kullanarak bitcoin fiyatını tahmin et.

Algoritma'nın uygulanmasında aşağıdaki gibi küçük ayrıntılar da öne çıkmakta.

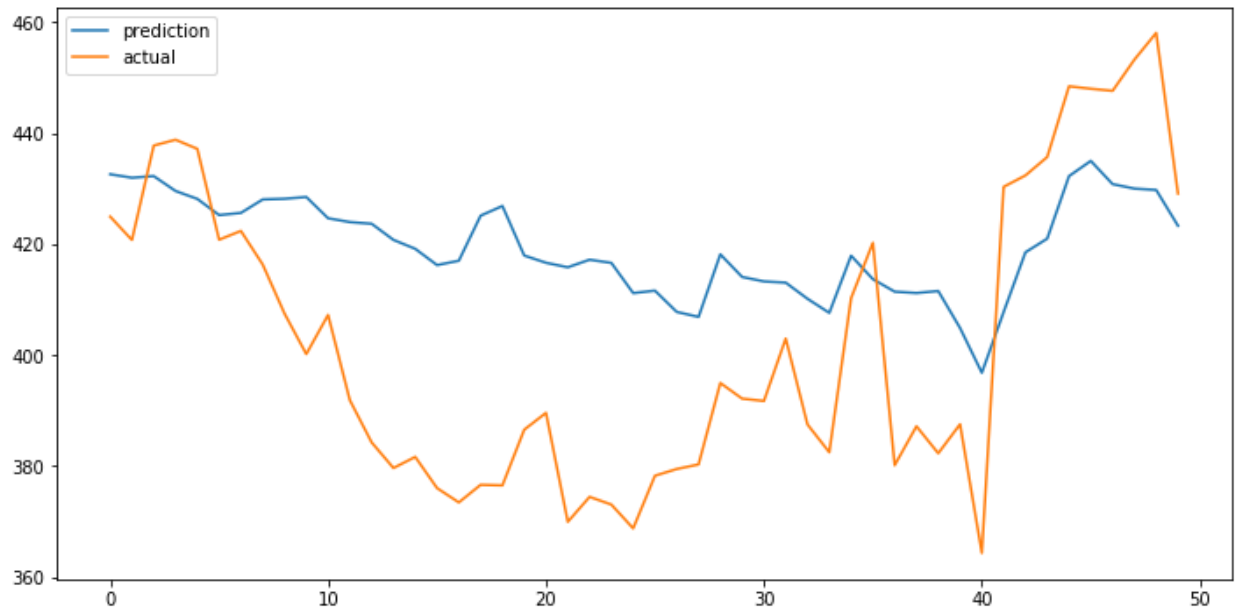
- Bitcoin fiyat tahmin edilmesi istenen tarihe en yakın kaç adet veriyi almalıyız?
  - Deneyimlerimize göre 50 adet veriyi almak daha optimize sonuçlar ortaya çıkarttı.
- Timebinler arası benzerliği nasıl buluruz?
  1. Timebin içindeki her fiyat verinin timebin'in ortalama fiyatına göre nasıl bir yüzde değişim içerdiğini kaydet. Örneğin ilk veri yüzde5 düşük, ikinci yüzde 1 yüksek...
  2. Timebin içindeki bu değişim yüzdelerinden oluşan veri dizisini bir vektör olarak kabul et.
  3. Farklı timebin'lerin değişim vektörleri arası uzaklığı öklid yardımıyla bul.
  4.  $(1 / \text{uzaklık})$  bize iki timebin arasındaki benzerliği verir.

- Fiyat değişimi tahmini nasıl yapılır?
  - Benzerlikleri bulunan komşuların timebinlerinden hemen sonraki fiyat verilerini timebin olarak al.
  - Timebin içindeki verilerin, timebin ortalamasına göre değişim yüzdelerini kaydet.
  - Benzerliği bulunan her timebin'in gelecek verilerini benzerlik oranlarıyla ağırlıklı ortalama alarak gelecek tahminimizi hesapla.
    - Her benzer timebin için,  
(gelecek timebin verisi \* benzerlik oranı) şeklinde hesaplıyıp tüm bu hesapları topla ve toplam benzerlik oranına bölerek tahmin ettiğimiz gelecek timebin verisi elde et.

# Çalıştırma Örnekleri

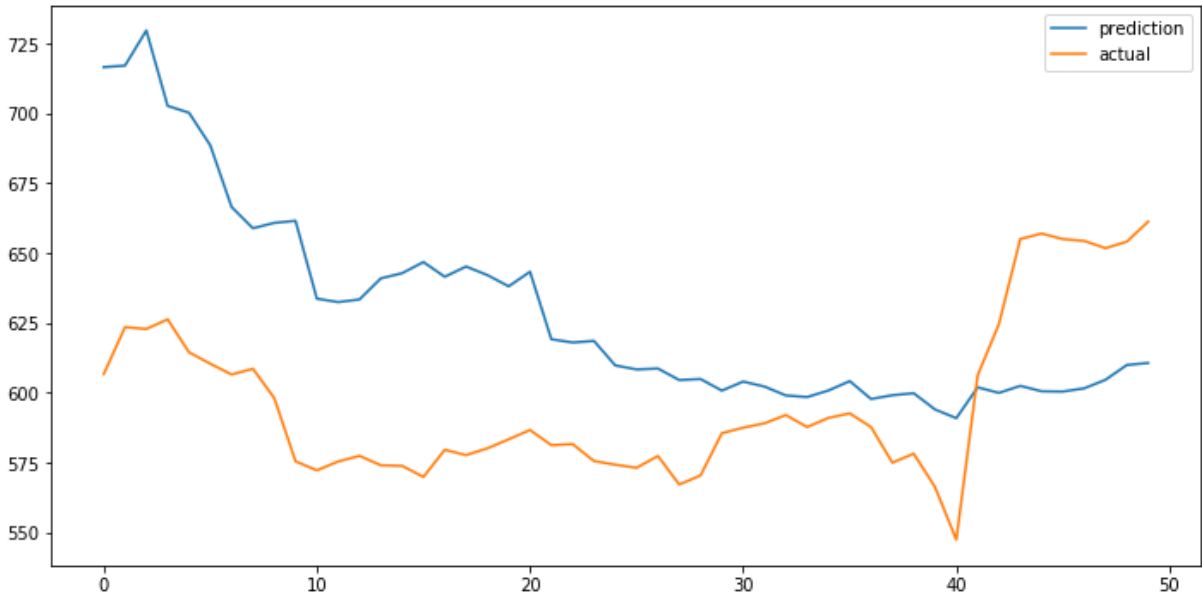


## Örnek 1:



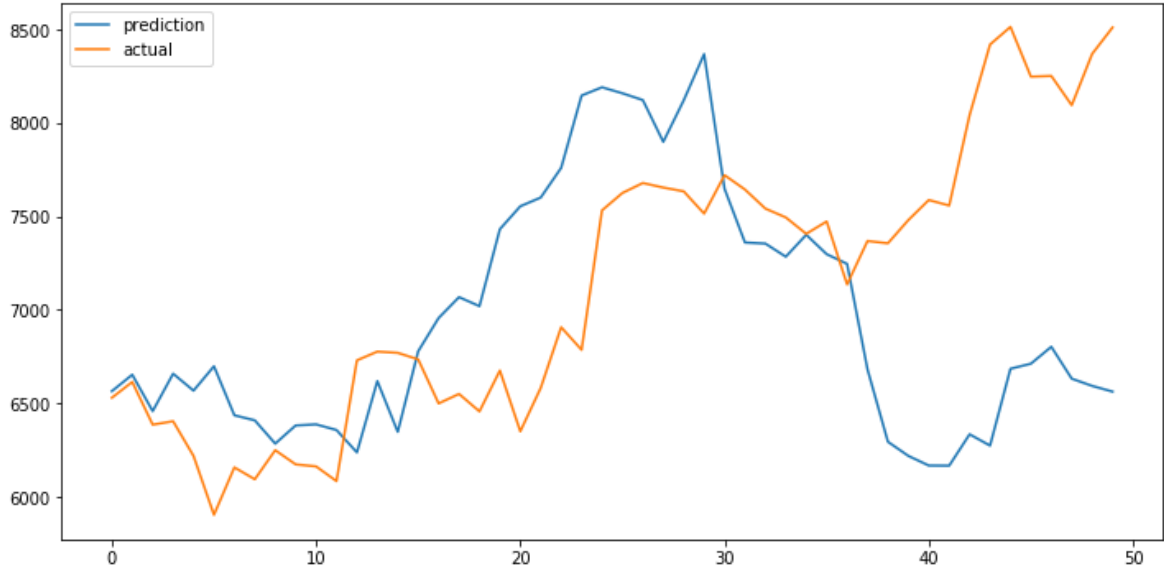
Bu örnekte, algoritmamızın fiyat düşüş ve yükselişlerini oldukça doğru tahmin ettiğini görmekle beraber, fiyatların ne kadar yükselip düşüceği konusunda çok başarılı olamadığını gözlemliyoruz.

## Örnek 2:



Yine örnek 1’de olduğu gibi, fiyatların yükselip düşme konusu oldukça başarılı olmasına karşık bu yükselişim miktarını yakalama konusunda çok başarılı değil.

### Örnek 3:



Bu örneğin büyük bir kısmında yükseliş-düşüş miktarı konusunda oldukça başarılı bir tahmin yaptığını gözlemlemekler beraber, son kısımlarda çok başarılı olmadığını gözlemliyoruz.

# Geliřtirme Sürecinde Yařananlar

## **1- Veriseti'nin küçük olması.**

Bitcoin yeni bir para birimi olduğundan tarihsel verisi oldukça az, bu veri setindeki benzerlikleri yakalama konusunda sorunlar oluşturdu.

## **2- Bitcoin Tarihsel Verilerinin Tutarsızlığı**

Bitcoin fiyatları oldukça tutarsız yükselme ve düşüşler yaşıyor. Özellikle verisetinin küçük olması sebebiyle bu tutarsızlıklara benzer tutarsız veri bulmakta zorluk yaşadığımızdan bir çok noktada başarı oranının istediğimiz iyi seviyelere gelememesiyle sonuçlandı.

## **3- Timebin Boyutunun Seçimi**

Timebin boyutları hem benzerlik hesaplarında hemde fiyat tahmini yaptığımız aşamada önemli bir kriter. Bu hiperparametrenin değerini farklı değerler deniyerek bulmaya çalıştık.

## 4- Timebinler Arası Benzerliđi Bulunması

Timebin Based Neighbourhood algoritmasının en kritik noktası benzerlik hesabı. Bu yüzden, benzerlik hesabını çok doğru yapmak için farklı seçenekler denenik. İlk önce pearson correlation ile başlayıp sonrasında elimizdeki benzerliğini bulmak istediđimiz verileri vektör olarak kabul edip aralarındaki uzaklıklarla ters orantılı olarak benzerliđi bulma yoluna gittik. Ancak ve ancak, elimizdeki 4 farklı parametrenin hepsini aynı katkı katsayısıyla vektörün içersine koyduğumuzdan, aslında bakarsanız iyi bir benzerlik fonksiyonu oluşturmuş olmadık. Daha detaylı analizler yardımıyla benzerlik fonksiyonu geliştirilip, sistemin başarı oranının arttırılması mümkün olduđu kanısındaım.

# Sistemin Başarısının Değerlendirilmesi

Sistemin başarısının değerlendirilmesinde  $R^2$  ve RMSE(Root Mean Square Error) kullanıldı.

Sistemin başlangıçtaki başarı oranlarına göre farklı benzerlik fonksiyonları kullanılması yardımıyla 9-10 katlık bir iyileştirme sağlandı.

# Kaynakça

- [1] Predicting Bitcoin Prices (<https://towardsdatascience.com/predicting-prices-of-bitcoin-with-machine-learning-3e83bb4dd35f>)
- [2] Time Series ([https://en.wikipedia.org/wiki/Time\\_series](https://en.wikipedia.org/wiki/Time_series))
- [3] Model Performance Measures (<https://pbiecek.github.io/ema/modelPerformance.html>)