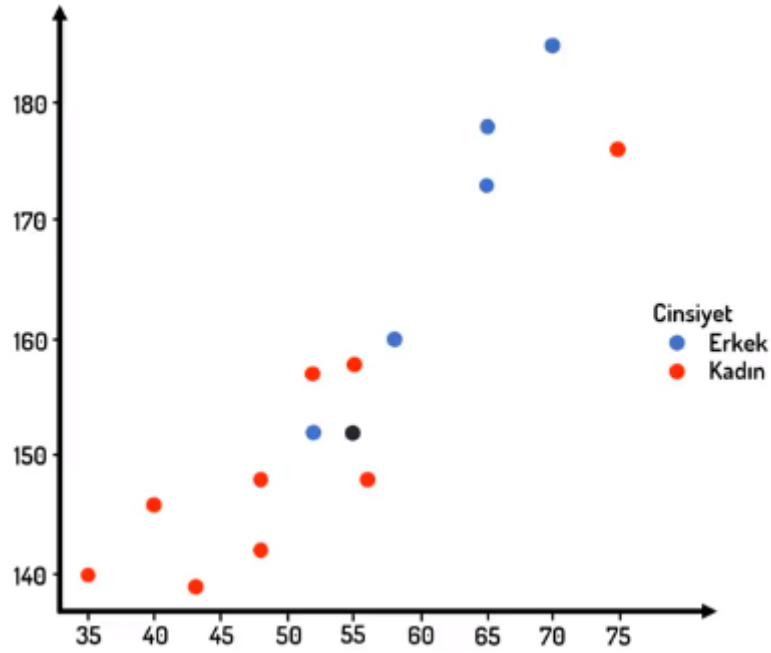


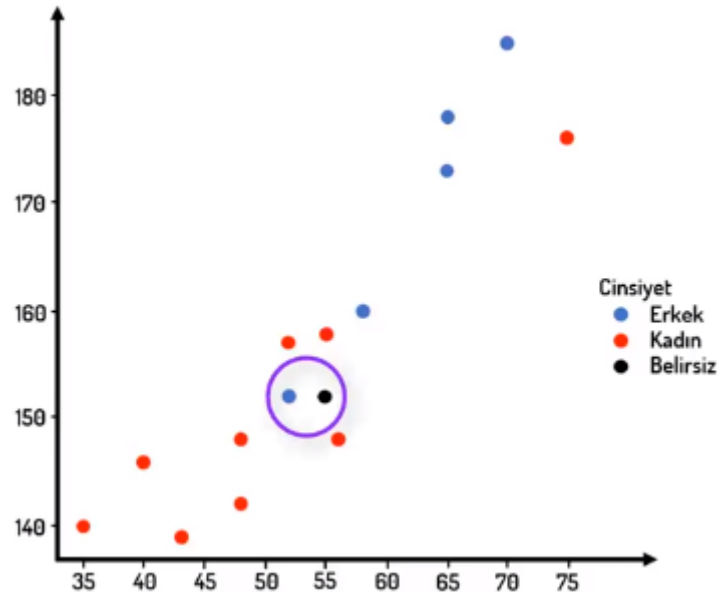
KNN ALGORİTMASI

KNN algoritmasının adı aslında İngilizcedeki K-Nearest Neighbors kelimelerinin baş harflerinden oluşmaktadır. K'ıncı En Yakın Komşu olarak çevrilebilir.

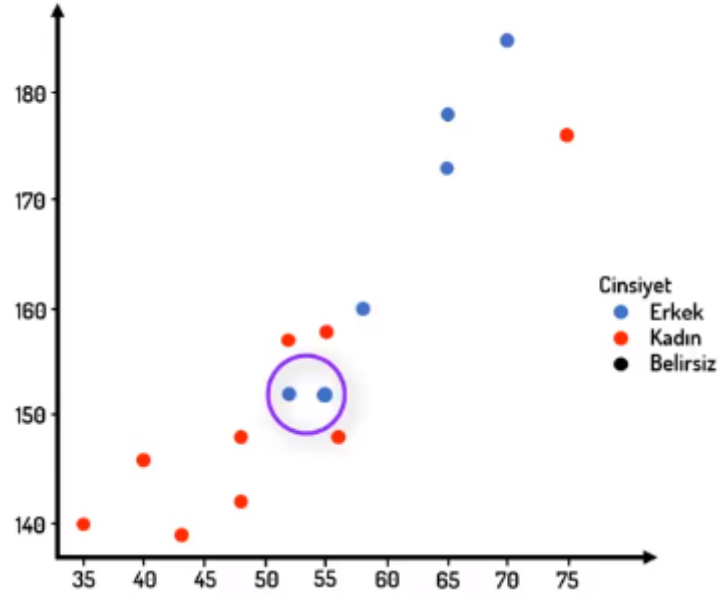
Örnek: **Mavi**-> erkek öğrenci, **Kırmızı**->kız öğrenci, **Siyah** sınıfa yeni katılan öğrenci ama cinsiyetini bilmiyoruz sadece boy ve kilosunu biliyoruz.



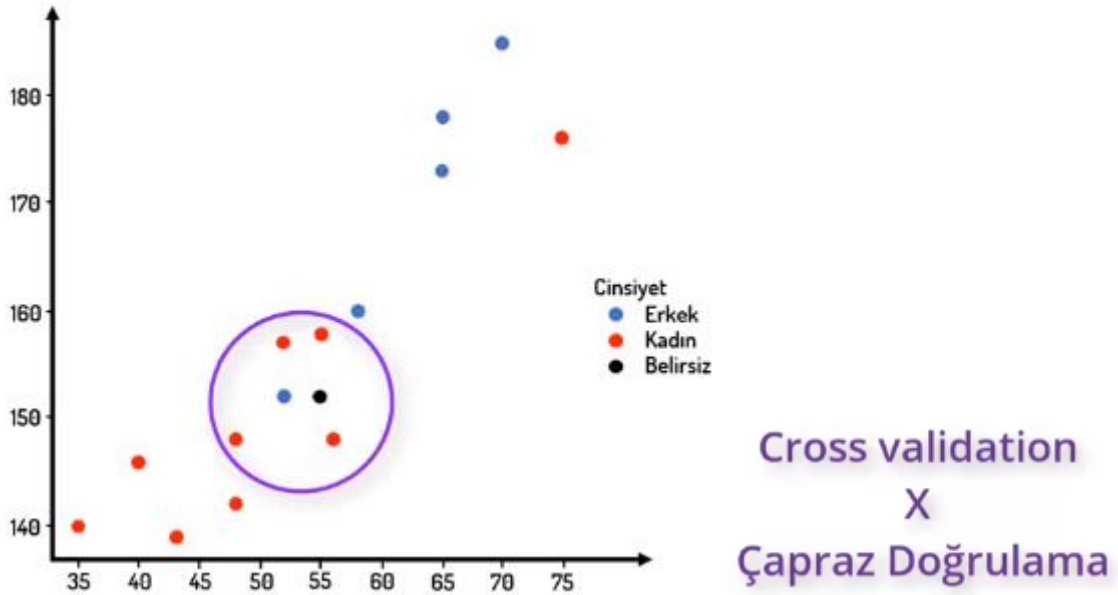
Algoritmanın ismindeki K harfi, kaç tane en yakın komşuya bakmamız gerektiği anlamına gelir.



Bu örnekte k 'yı 1 aldığımızda siyah öğrenci, erkek olarak belirlenmektedir.

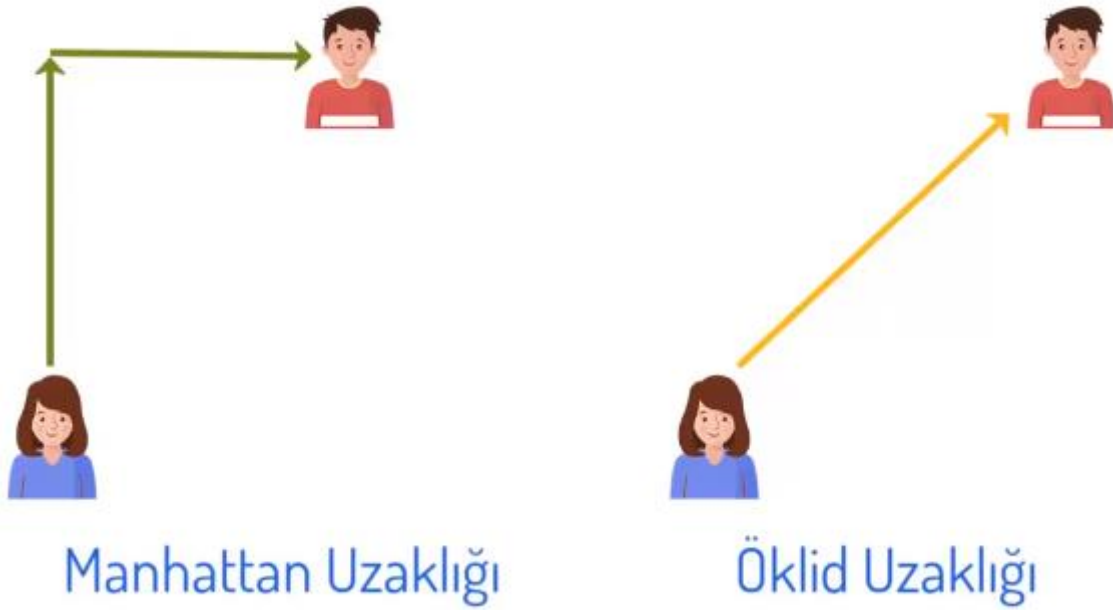


Fakat k değeri değiştirildiğinde farklı sonuçlar elde edilebilir. Örneğin k 'yı 5 aldığımız da bu bilinmeyen öğrenci, kız öğrenci olarak nitelendirilmektedir.



k değişkenini cross validation dediğimiz çapraz doğrulama yöntemi ile belirleyerek daha doğru bir k değeri seçmeyi sağlayabiliriz.

Örneklemelerin birbirine olan uzaklıklarını hesaplamak için de farklı uzaklık metrikleri kullanıyoruz. Bazıları:



KNN ALGORİTMA UYGULAMASI

Kütüphanemizi import edip liste oluşturup DataFrame dönüştürme:

```
[1] import pandas as pd
```

```
[2] boy = [185,178,176,173,160,158,157,148,148,146,140,139,142,152]  
kilo = [70,65,75,65,58,55,52,56,48,40,35,43,48,52]  
cinsiyet = ['E','E','K','E','E','K','K','K','K','K','K','K','K','E']
```

```
df = pd.DataFrame(list(zip(boy, kilo, cinsiyet)),  
                  columns=['Boy', 'Kilo', 'Cinsiyet'])  
df
```

	Boy	Kilo	Cinsiyet
0	185	70	E
1	178	65	E
2	176	75	K
3	173	65	E
4	160	58	E
5	158	55	K
6	157	52	K
7	148	56	K
8	148	48	K
9	146	40	K
10	140	35	K
11	139	43	K
12	142	48	K
13	152	52	E

Fonksiyon oluşturma ve fonksiyon için gerekli kütüphaneler:

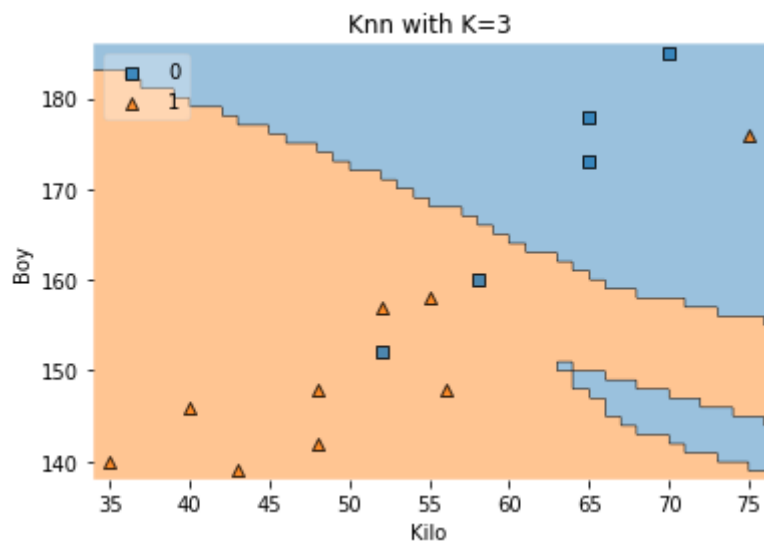
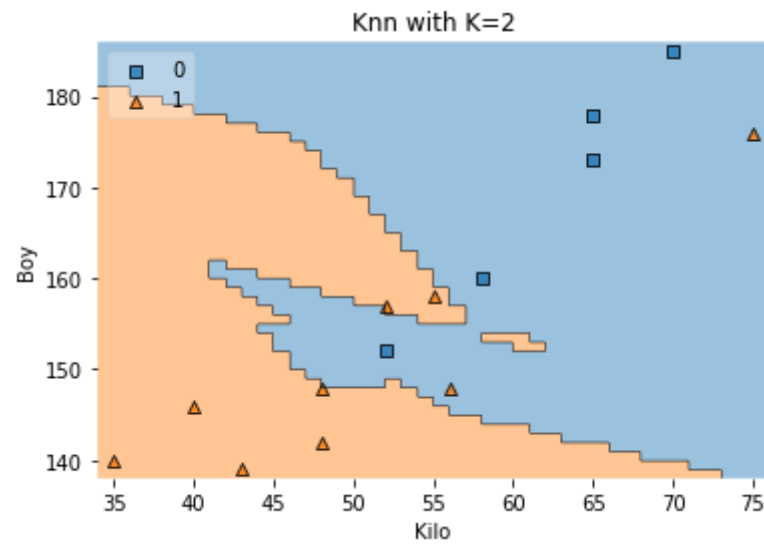
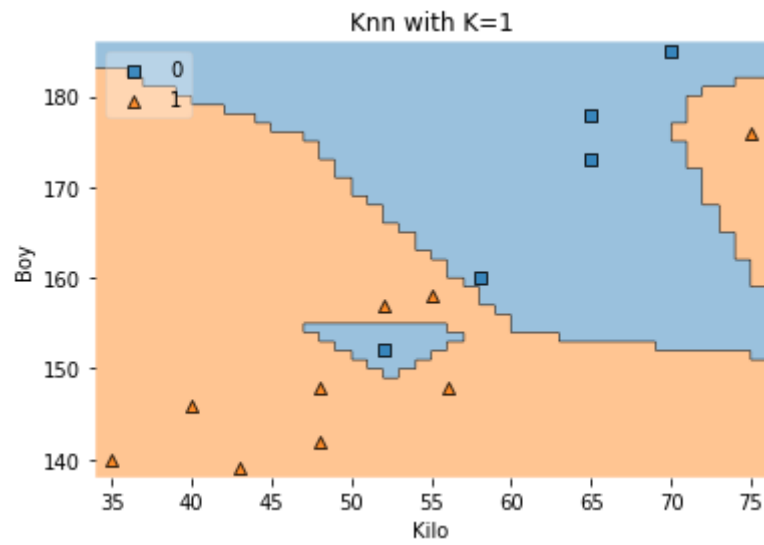
```
[4] import matplotlib.pyplot as plt
import pandas as pd
from sklearn import datasets, neighbors
from mlxtend.plotting import plot_decision_regions

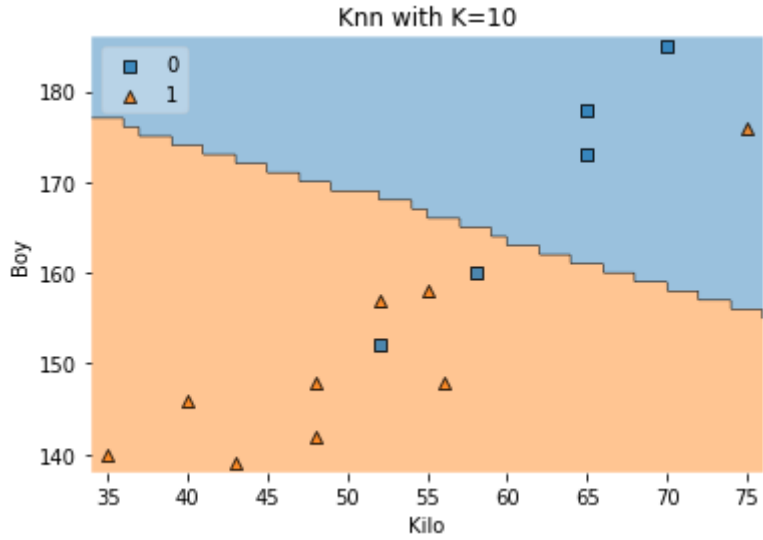
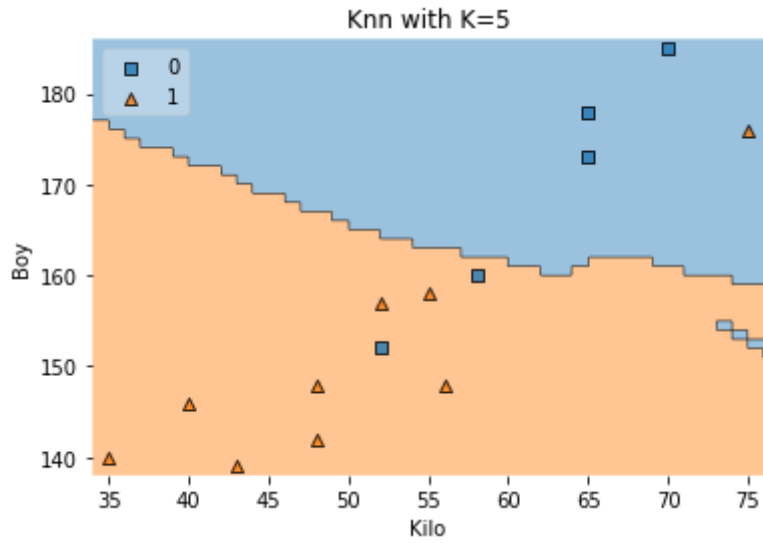
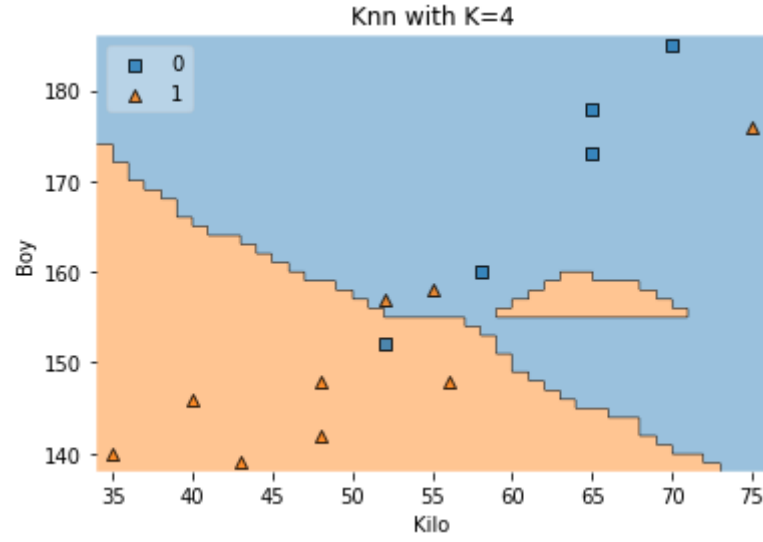
def knn_comparison(data, k, m="euclidean"):
    x = data[['Kilo', 'Boy']].values
    y = data['Cinsiyet'].astype('category').cat.codes.to_numpy()
    clf = neighbors.KNeighborsClassifier(n_neighbors=k, metric = m)
    clf.fit(x, y)
    plot_decision_regions(x, y, clf=clf, legend=2)
    plt.xlabel('Kilo')
    plt.ylabel('Boy')
    plt.title('Knn with K='+ str(k))
    plt.show()
```

1, 2, 3, 4, 5, 10 değerlerini kullanma:



```
for i in [1,2,3,4,5,10]:  
    knn_comparison(df, i)
```



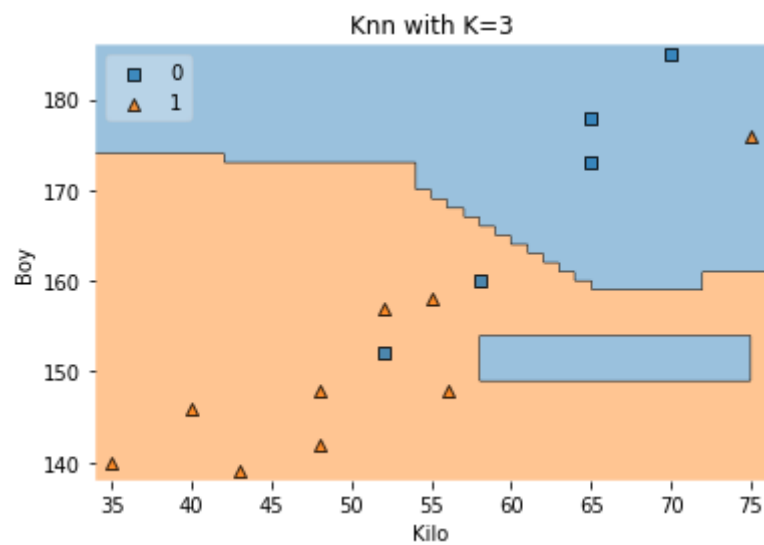
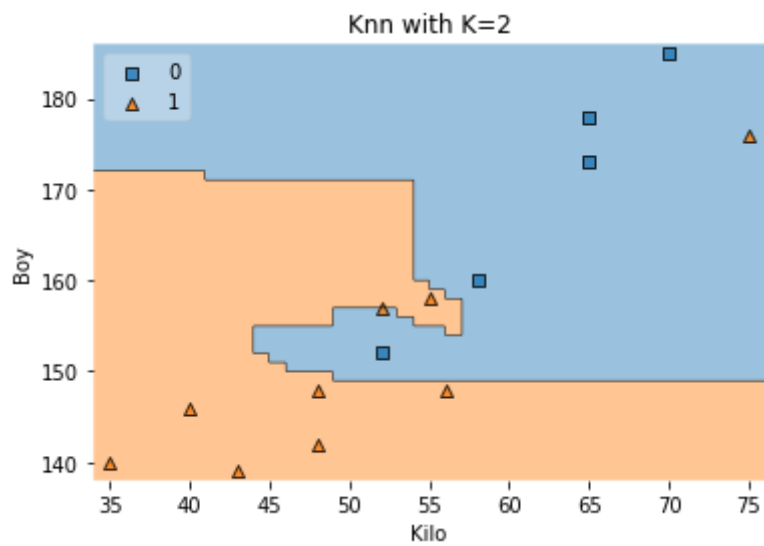
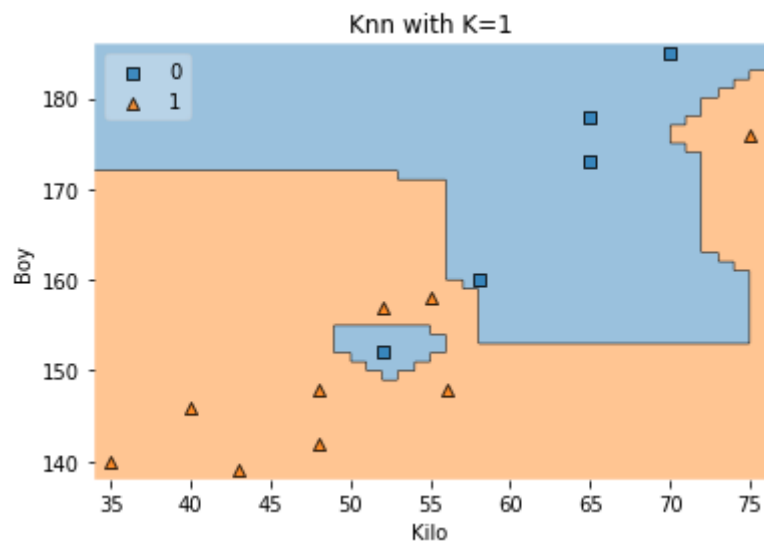


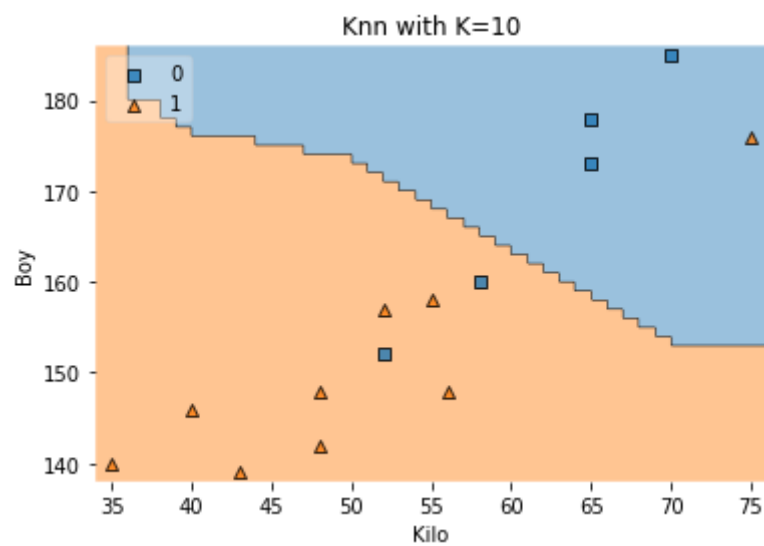
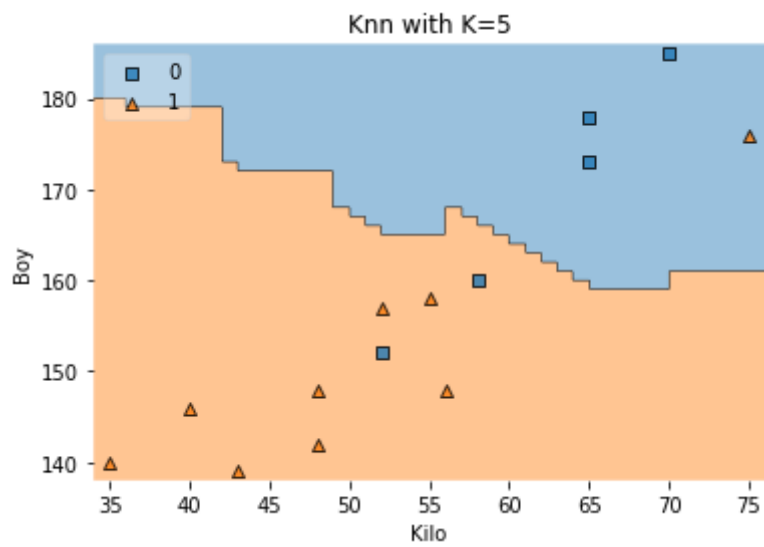
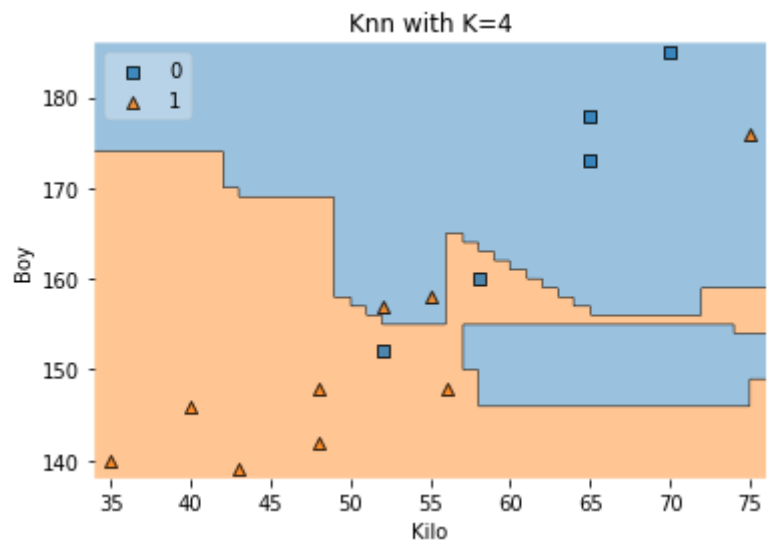
1 komşu olduğunda, 2 komşu olduğunda, 3, 4, 5, 10 komşular olduğunda benim alanım ne şekilde bölünüyor? Ne değiştiriyor kaç komşu olması? Bunları gösteriyor. Öklid mesafe ile yapıldı.

Manhattan kullanımı:



```
for i in [1,2,3,4,5,10]:  
    knn_comparison(df, i, "manhattan")
```





KAYNAKÇA

Bilgeiř “Herkes iin Yapay Zekâ” eđitimi.

