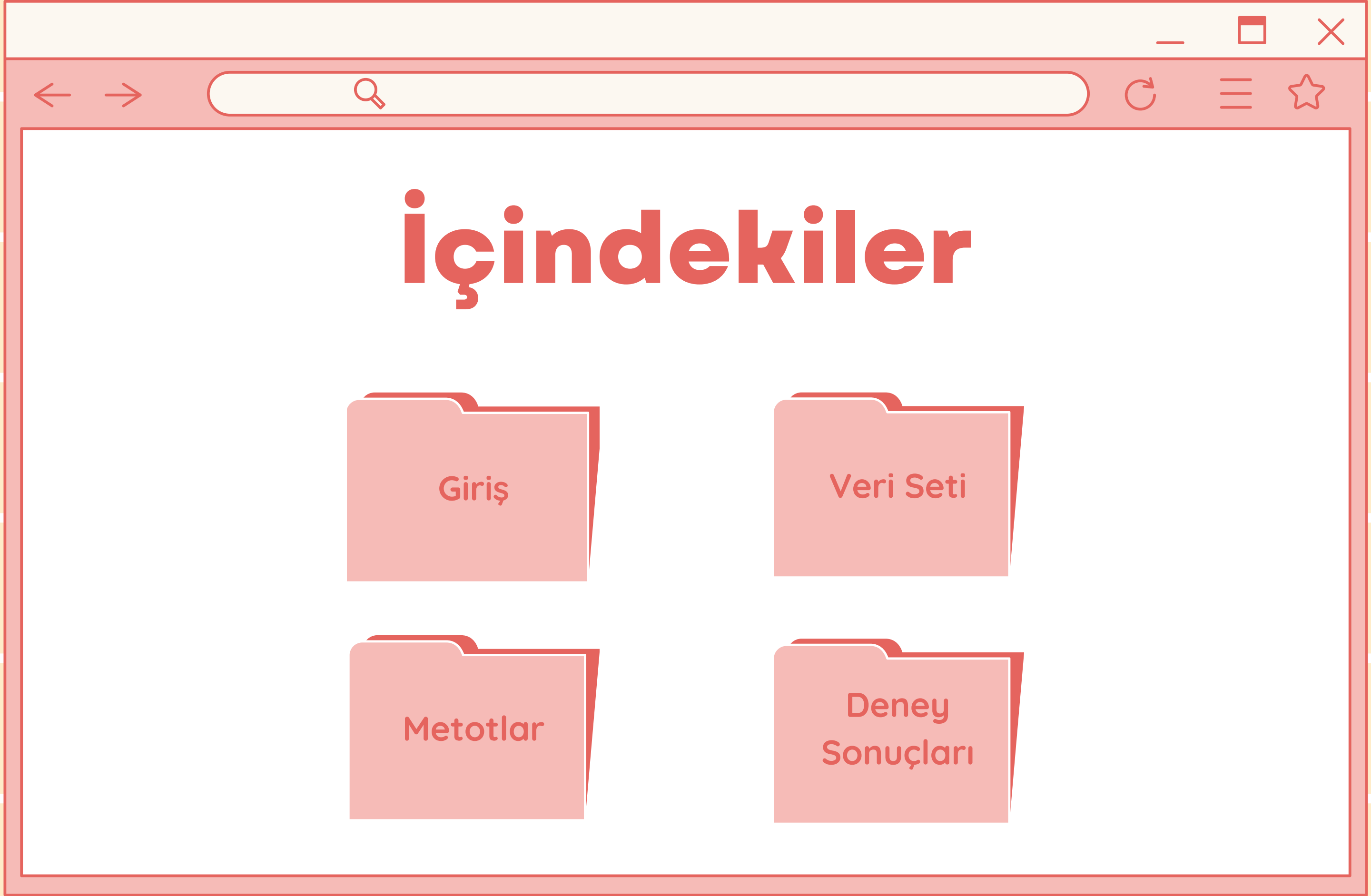




Digit Recognition

Alperen Topak



İçindekiler

Giriş

Veri Seti

Metotlar

Deney
Sonuçları



Giriş

Projenin Amacı: Bu projenin amacı, çeşitli makine öğrenimi modellerini kullanarak, rakamlardan oluşan bir veri setindeki rakamları tanımayı sağlamaktır.



Veri Seti

Kullanılan Veri Seti, Kaggle üzerinden seçilmiş olup içinde rakamların bulunduğu bir görüntü veri setidir.

<https://www.kaggle.com/datasets/pintowar/numerical-images/data>

Kullanılan Modeller

Bu projede iki farklı model kullanıldı.

- 1.) KNN Modeli
- 2.) Yapay Sinir Ağları

KNN

KNN modeli, bir sınıflandırma algoritmasıdır. Bu algoritma, herhangi bir örneği sınıflandırmak için ona en yakın komşularını kullanır. Yani, bir örneği sınıflandırmak için, bu örneğe en yakın olan eğitim verisindeki örneklerin etiketlerine bakarız. En yakın komşuların etiketlerine göre, yeni örneği ait olduğu sınıfa atarız.

Yapay Sinir Ağları

Yapay Sinir Ağları, insan beyninin çalışma prensibinden esinlenerek tasarlanmış matematiksel modellerdir. Bilgiyi işlemek ve desenleri tanımak için kullanılır. Birbirine bağlı olan yapay sinir hücrelerinden (nöronlar) oluşur. Her bir nöron, girdileri alır, bu girdileri işler ve bir çıktı üretir. YSA'lar, genellikle katmanlar halinde düzenlenir. Her bir katman, belirli bir görevi yerine getirir ve bir sonraki katmana çıktı sağlar.

Deney Sonuçları- KNN

KNN sonuçları:

MAE: 0.26427165354330706

MSE: 1.5585629921259843

RMSE: 1.2484242035966717

R2 Score: 0.8129864508007312

Deney Sonuçları - KNN

MAE (Mean Absolute Error):

- MAE'nin 0.264 olması, modelin tahminlerinin gerçek değerlerden ortalama olarak yaklaşık 0.26 birim uzaklıkta olduğunu gösterir.
- Bu, ortalama olarak makul bir hata miktarıdır ve modelin tahminlerinin kabul edilebilir olduğunu düşündürebilir.

MSE (Mean Squared Error):

- MSE'nin 1.558 olduğu göz önüne alındığında, modelin tahminlerinin gerçek değerlerden ortalama olarak yaklaşık 1.56 birim uzaklıkta olduğu görülür.
- MSE genellikle büyük hataları vurguladığından, bu değer modelin bazı tahminlerde büyük hatalar yaptığını gösterebilir.

Deney Sonuçları - KNN

RMSE (Root Mean Squared Error):

- RMSE'nin 1.248 olduğu göz önüne alındığında, modelin tahminlerinin gerçek değerlerden ortalama olarak yaklaşık 1.25 birim uzaklıkta olduğu görülür.
- RMSE, MSE'nin orijinal veri birimine dönüştürülmüş halidir. Bu, modelin tahminlerinin hala belirli bir sapma içerdiğini gösterir.

R2 Score (R-squared):

- R2 skorunun 0.813 olduğu göz önüne alındığında, modelin bağımsız değişkenlerin bağımlı değişkeni açıklama oranının %81.3 olduğu görülür.
- Bu, modelin veriye göre iyi bir uyum sağladığını gösterir, ancak belki de daha iyisi mümkün olabilir.

Deney Sonuçları - KNN

- MAE değeri kabul edilebilir bir hata miktarını gösterirken,
- MSE ve RMSE değerleri modelin bazı büyük hatalar yaptığını işaret edebilir.
- R2 skoru, modelin veriye iyi bir uyum sağladığını gösterir, ancak hala iyileştirilebileceğini gösterir.

Deney Sonuçları - YSA

Yapay Sinir Ağları sonuçları:

Test Accuracy: 93.6%

Test Loss: 24.5%

Test Precision: 94.5%

Test Recall: 93.0%

Deney Sonuçları - YSA

Accuracy -> Doğru sınıflandırma oranı

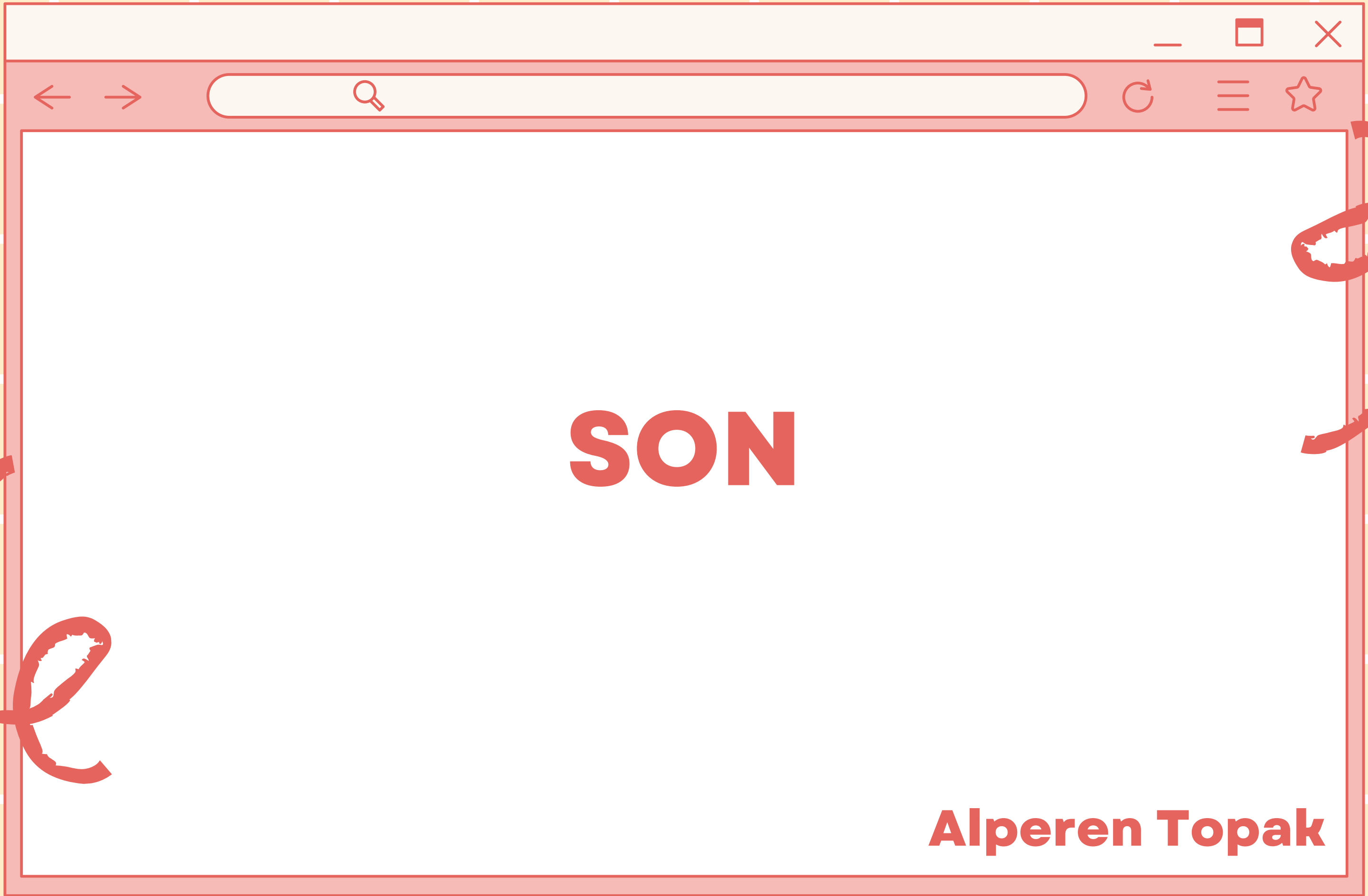
Precision -> Örneğin 5 olarak tahmin ettiğimiz sınıfların ne kadarının gerçekten 5 olduğunu ifade eder

Recall -> Önce gerçek değerlere odaklanır ve bu modelle bunların kaç tanesini doğru tahmin ettiğimizi ifade eder

Precision ve Recall birbirine yakın çıktığı için sınıflar arası bir dengesizlik yok denilebilir.

Deney Sonuçları - YSA

- Yapay Sinir Ağı modelinin yüksek doğruluk ve hassasiyet oranları, genel olarak kabul edilebilir bir performans sergilediğini gösterir.
- Ancak, yüksek test kaybı ve belirli durumlarda düşük geri çağırma oranı, modelin hala iyileştirilebileceğini veya bazı senaryolarda daha iyi performans gösterebileceğini işaret edebilir.



SON

Alperen Topak