

YILDIZ TEKNİK ÜNİVERSİTESİ
ELEKTRİK-ELEKTRONİK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ



Yapay Zeka Dersi Dönem Proje Ödevi

Burak Başol

24011037

Öğretim Görevlisi

Prof. Dr. Mehmet Fatih AMASYALI

İstanbul

2025

Contents

Giriş	3
Ödev Açıklaması	3
Veri Oluşturma	4
Kullanılan Model Yapısı	4
Hiperparametre Optimizasyonu	4
Problem A.....	5
Problem B	6
Problem C	7
Problem D	8
Problem E.....	9
Sonuç ve Yorumlarım	9

Giriş

Yapay zeka dersi dönem projesine ait kod dosyası, video linki ve raporu teslim edilmesi istenmiştir.

Video Linki: <https://youtu.be/R3kkIRw370M>

Ödev Açıklaması

Ödevde, 25x25 binary matrisler kullanılarak proje tanımında istenilen problemler için yapay sinir ağları kullanılarak tahmin modelleri geliştirilmiştir.

Veri Oluřturma

Projede istenildiđi üzere her problem için 800 train 200 test olmak üzere 1000 adet veri oluşturulmuřtur. Bu veriler ile yapay sinir ađı modelleri eğitilmiřtir.

Veriler numpy array formatında (.npy) saklanmıřtır.

Kullanılan Model Yapısı

Projede Convolutional Neural Network (CNN) kullanılmıřtır. Katmanlar řu řekildedir:

- Conv2D
- MaxPooling2D
- Conv2D
- MaxPooling2D
- Conv2D
- MaxPooling2D
- Flatten
- Dense
- Dropout
- Dense

Hiperparametre Optimizasyonu

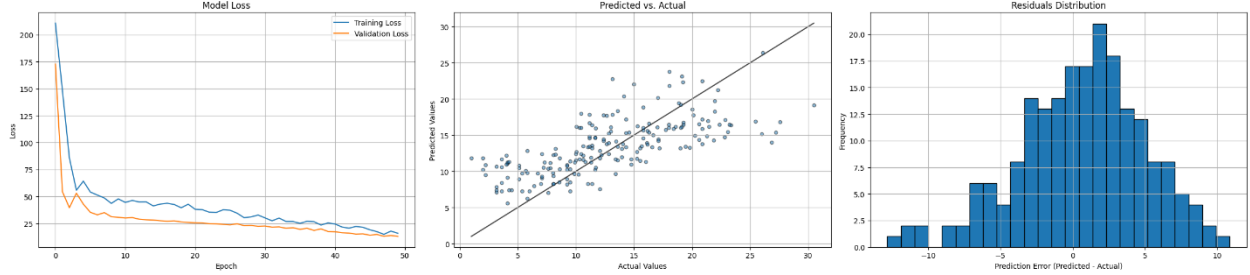
Hiperparametre optimizasyonu olarak, model için optimizasyon yöntemi “Adam” kullanılmıřtır ve lr değeri 0.001 olarak belirlenmiřtir.

Ayrıca erken durdurma, batch size ve epoch değeri de modele ve veri setine uygun řekilde belirlenmiřtir.

Problem A

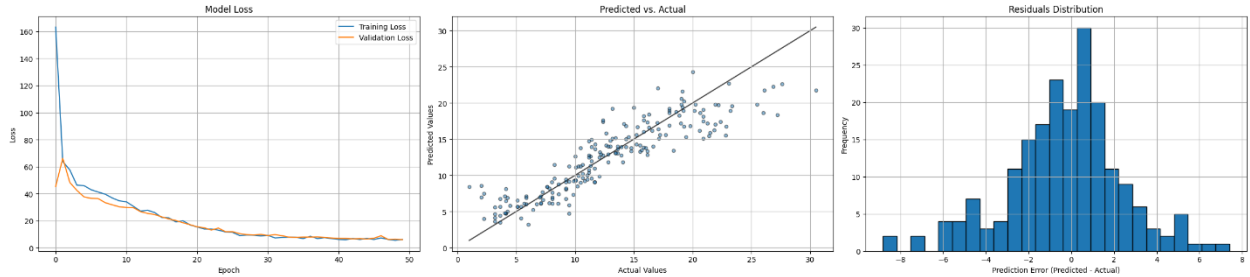
Matris üzerinde rastgel konumlarda iki adet nokta bulunmaktadır ve aralarındaki mesafenin tahmin edilmesi istenmiştir.

Problem A - CNN Performance (quarter Training Data)



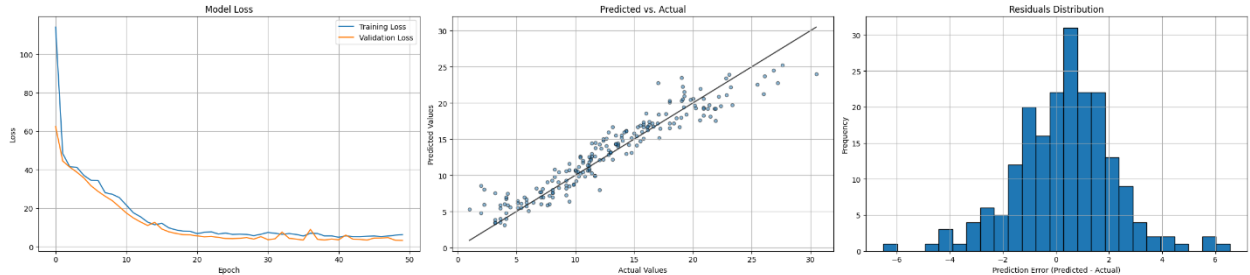
- Veri setinin çeyreği kullanıldığı zaman “loss” değeri 20-25 arasına kadar düşmüştür ve orada takılmıştır. Veri yetersiz olduğu için yeterince verimli çalışmamıştır. Dağılım grafiğinde ve nokta grafiğinde de değerler yayılmıştır.

Problem A - CNN Performance (half Training Data)



- Veri setinin yarısı kullanıldığı zaman “loss” değeri yaklaşık 10’a kadar düşmüştür ve orada kalmıştır. Veri setinin çeyreğinin kullanıldığı senaryoya göre daha az bir dağınıklık ve hata görülmektedir.

Problem A - CNN Performance (full Training Data)

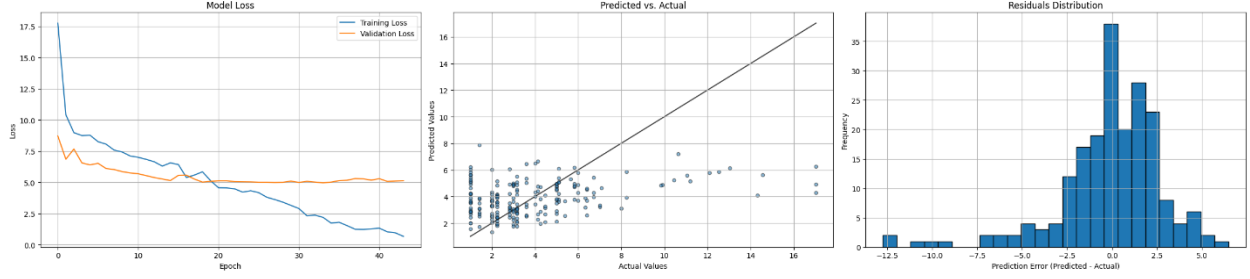


- Tüm veri seti kullanıldığında en optimum sonuç elde edilmiştir. Hata oranı diğer sonuçlara göre çok daha düşüktür.

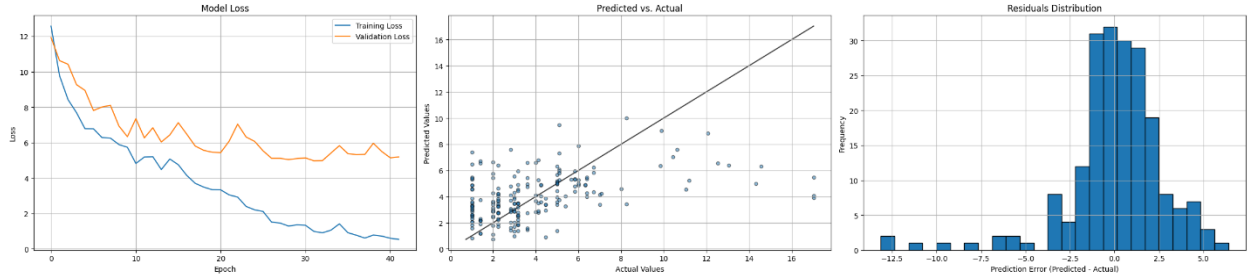
Problem B

Matriste 3 ile 10 adet arası nokta bulunmaktadır ve model bu noktalardan en yakın iki tanesi arasındaki mesafeyi tahmin etmektedir.

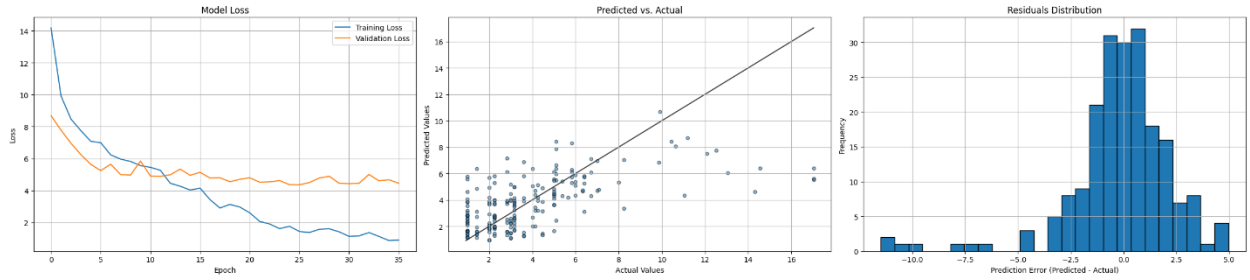
Problem B - CNN Performance (quarter Training Data)



Problem B - CNN Performance (half Training Data)



Problem B - CNN Performance (full Training Data)

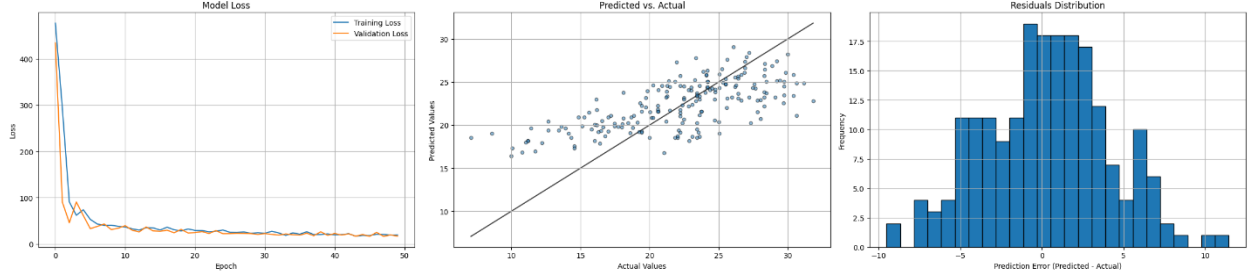


- Üç durum için de alınan çıktıda küçük değişiklikler gerçekleşmiştir. Önemli değişiklik dağılım yoğunluğunda gerçekleşmiştir.

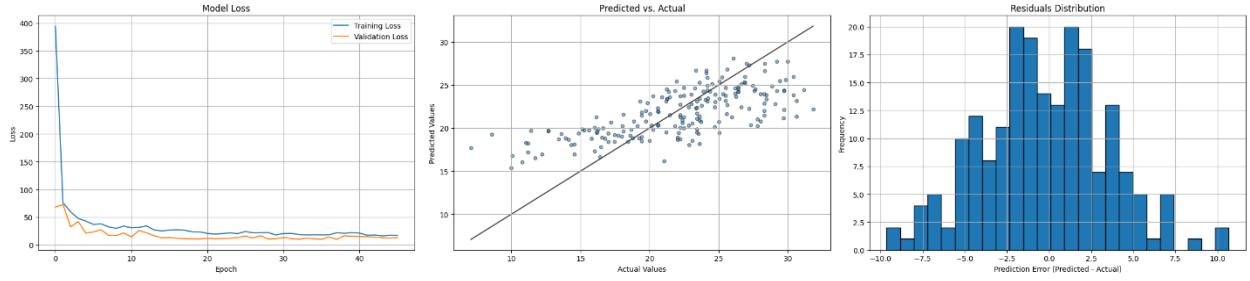
Problem C

Bir önceki problemdeki koşullar kullanılmıştır. En yakın yerine en uzak iki nokta arasındaki mesafe tahmin edilmektedir.

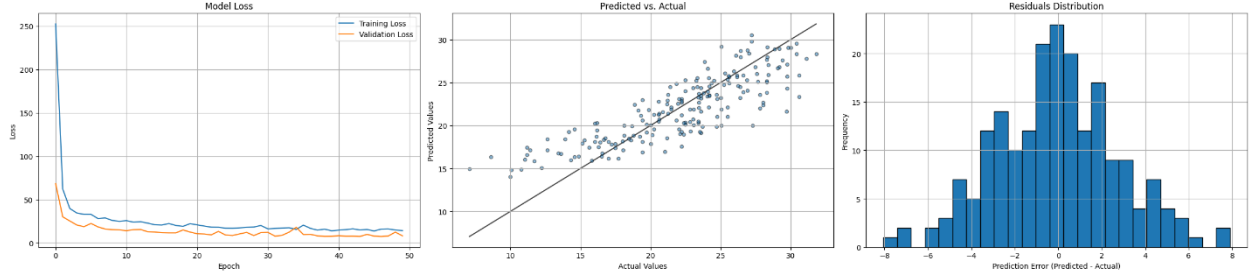
Problem C - CNN Performance (quarter Training Data)



Problem C - CNN Performance (half Training Data)



Problem C - CNN Performance (full Training Data)

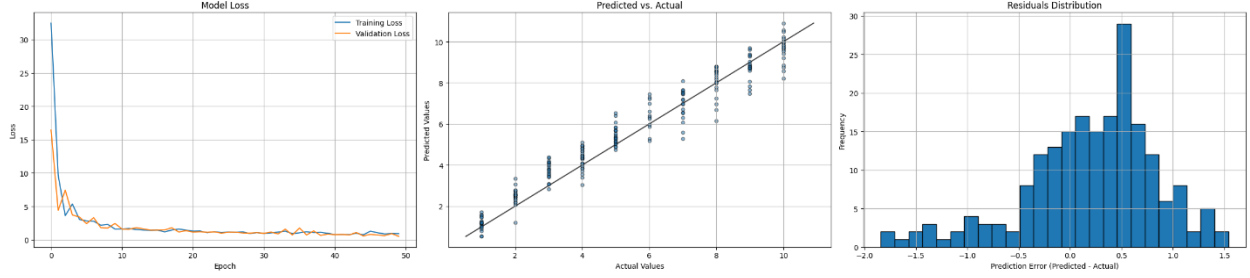


- Sonuçlar bu soru için de yaklaşık olarak aynı kalmıştır. Dağılım aralıklarında ve “loss” grafiklerinde küçük değişiklikler mevcuttur.

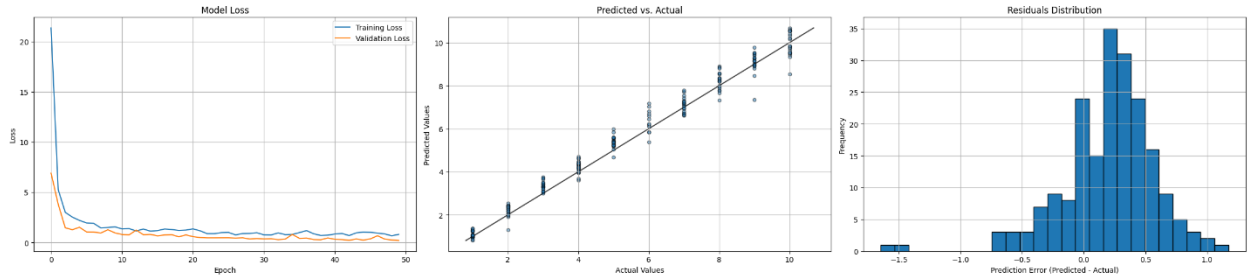
Problem D

Matriste 3 ile 10 adet arası nokta bulunmaktadır ve model nokta sayısını tahmin etmektedir.

Problem D - CNN Performance (quarter Training Data)

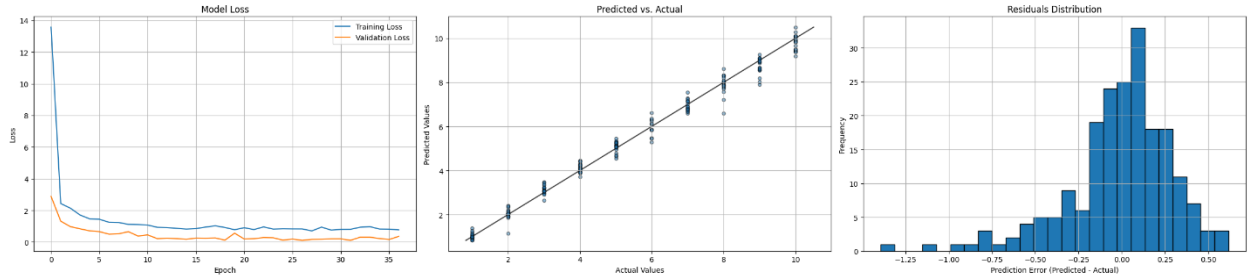


Problem D - CNN Performance (half Training Data)



- Veri setinin çeyreği ve yarısı kullanıldığı durumlarda “loss” grafiğinde fark minimal kalmıştır. Dağılım grafiğinde ise gözle görülebilir bir fark vardır. Ayrıca veri setinin çeyreği ile çalışıldığı durumda hata oranı 0.5 olduğu durumlar çok fazladır. Veri setinin yarısı kullanıldığı durumda ise bu oran 0.2-0.5 arasında gerçekleşmektedir.

Problem D - CNN Performance (full Training Data)

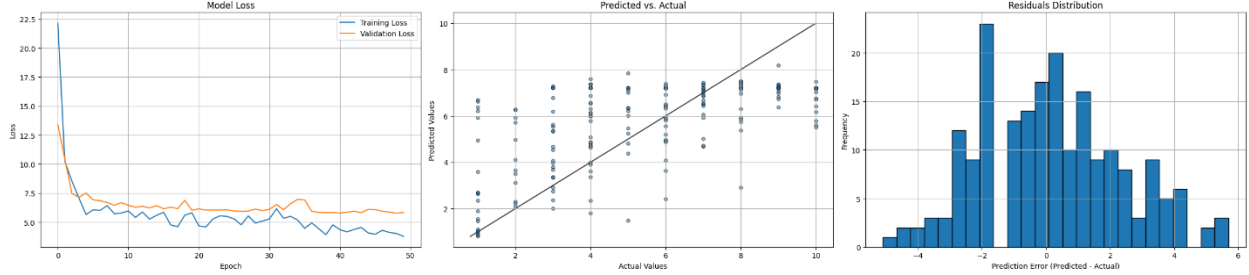


- Tüm veri seti kullanıldığında hata oranı minimize edilmiştir. Dağılım grafiğinde de görüleceği üzere hata oranı 0'a çok yakındır.

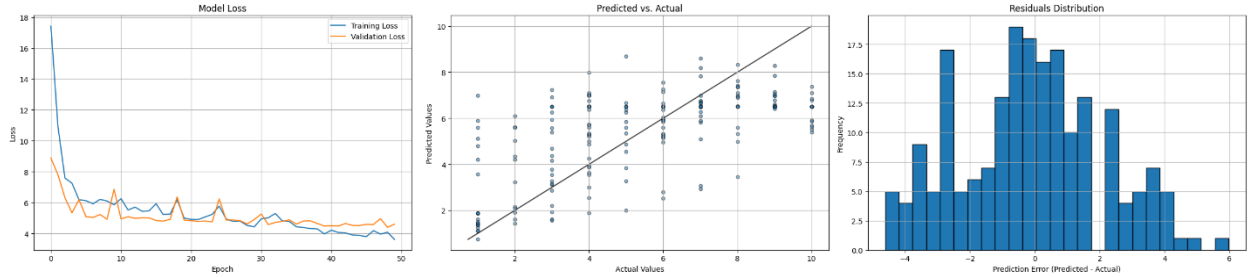
Problem E

Matriste 1 ile 10 adet arası deęiřken büyüklükte kareler bulunmaktadır. Model kare sayısını tahmin etmektedir.

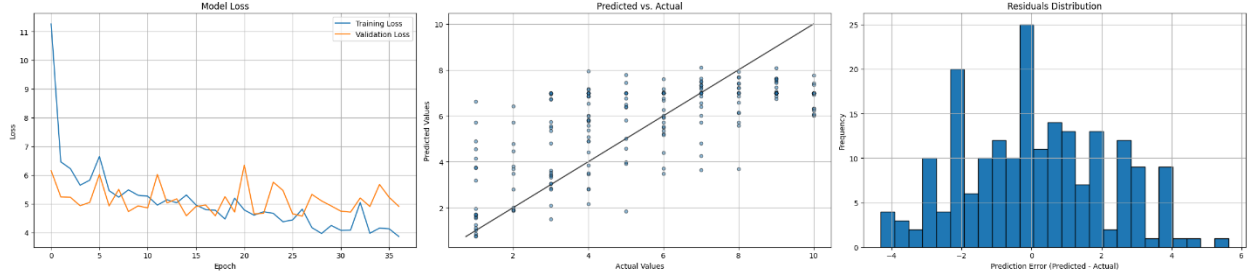
Problem E - CNN Performance (quarter Training Data)



Problem E - CNN Performance (half Training Data)



Problem E - CNN Performance (full Training Data)



- Model düzgün çıktı verse de bence yetersiz kalmıştır. Hata oranı düşüktür fakat tahminindeki hata dağılımı düzgün ilerliyor. (Şans eseri doğru cevabı vermiş gibi)

Sonuç ve Yorumlarım

CNN modeli ödevde istenilen durumlarda verimli olarak çalışmakta ve düşük hata payı ile çıktı üretmektedir.

Veri seti daha büyük olduğu zaman (çeyrek ve yarım veri setli eğitimlerde bu durum tam tersidir) daha düzgün sonuçlar üreterek daha düşük hata payı ile çıktı üretmiştir.

Ayrıca basit bir CNN modeli son problem için tam olarak verimli çalışmamıştır.