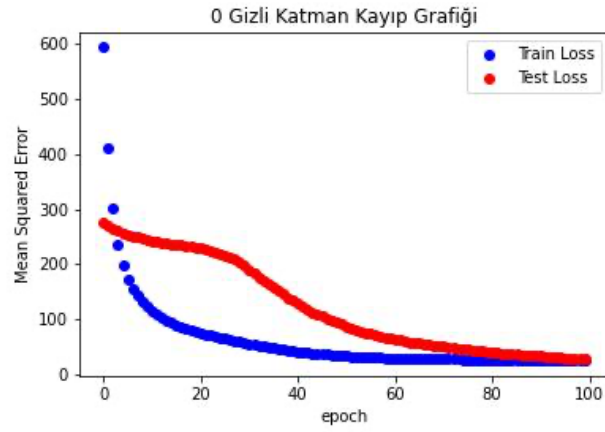




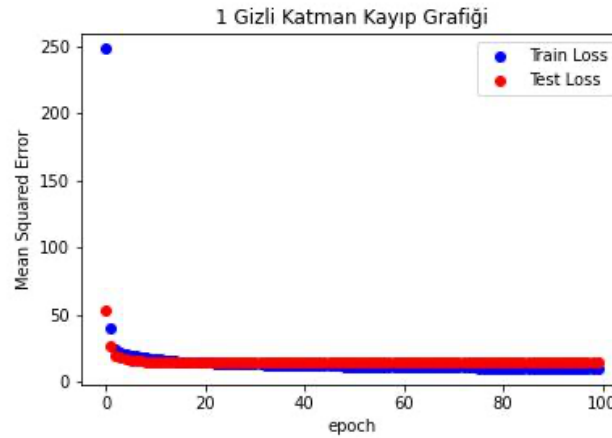
TOBB University of Economics and
Technology
Faculty of Engineering
YAP470 – Machine Learning
Homework 2

Name	Ecem Deniz Babaoğlu
ID	181201071
Date	26/11/2022

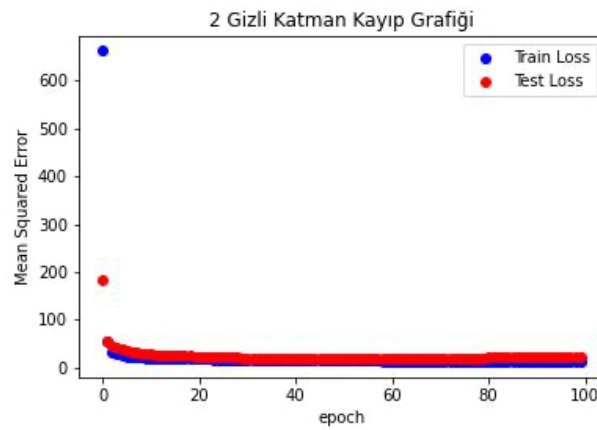
Boston Housing veriseti üzerinde eğitilen her model için learning rate=0.001, epoch sayısı=100 seçilerek her epoch'ta ulaşılan eğitim kaybı ve test kaybı, mean squared error fonksiyonu kullanılarak elde edilmiş ve çizdirilmiştir. Epoch sayısı görselleştirebilmek adına 100 alınmıştır. Learning rate seçimi sebebine ise ileride değinilecektir.



Figür 1



Figür 2



Figür 3

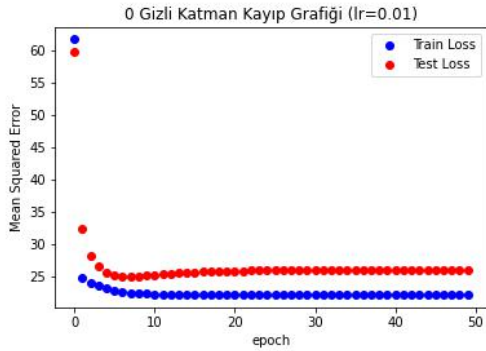
Grafiklerde de görülebileceği üzere, katman sayısı arttıkça kayıp daha hızlı azalmakta. Fakat, 0 katmanlı model daha büyük learning rate (bkz:0.01, epochs=50) ile de yeterince başarılı bir performans sergilerken, 1 ve 2 katmanlı modellerde bu geçerli değil.

1 ve 2 katmanlı modellerde learning rate veya epoch sayısı arttıkça “Dying ReLU” problemi gözlemlendi. Bu problemin üzerinden gelebilmek için learning rate veya epoch sayısı düşürüldüğünde ise model hiç öğrenmedi. Bu sebeple parametreler daha dikkatli seçildi. Bu grafiklerin çizimi için seçilen learning rate ise bu iki modele uyum sağlayabilmesi için seçildi.

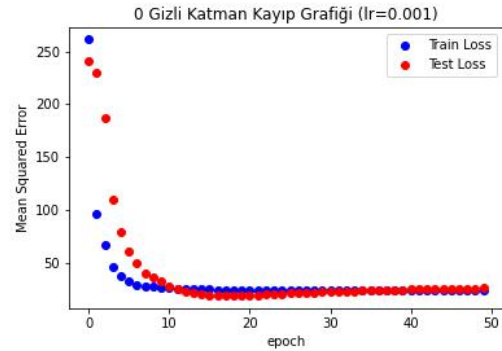
Tek katmanlı model, learning rate=0.0001, epochs=100 ve nöron sayısı 5, 7, 10, 12, 13 alınarak gözlem yapıldı. Yüksek nöron sayılarında daha iyi performans gösterdiği gözlemlendi. En iyi performans ise 12 nöronda alındı.

İki katmanlı model, learning rate=0.0001, epochs=100 ve nöron sayıları 12-10, 10-7, 7-5, 5-3, 12-12, 10-10, 7-7, 5-5, 3-3, 10,5 ve 10-3 için gözlemlendi. 10-10 nöronlu modelde “Dying ReLU” problemi gözlemlendi. En iyi sonuçlar ise 12-12 ve 7-7 nöron sayılı modeller için gözlemlendi. Bu sonuçlara göre birden fazla katmanlı modellerde üst üste aynı ve yüksek nöron sayılarının daha başarılı sonuçlar verebileceği gözlemlendi.

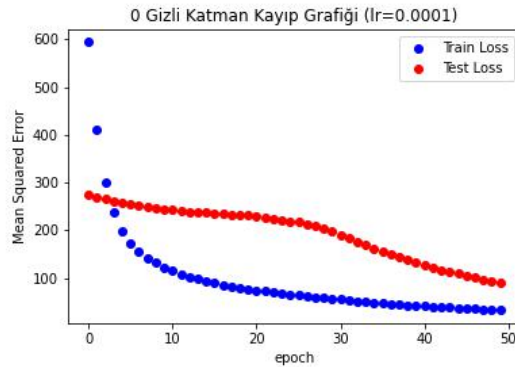
0 katmanlı model $lr=0.01$, 0.001 , 0.0001 için gözlemlendi. Bu modellerin kayıp grafikleri sırasıyla Figür 4, Figür 5 ve Figür 6’da gösterilmiştir



Figür 4



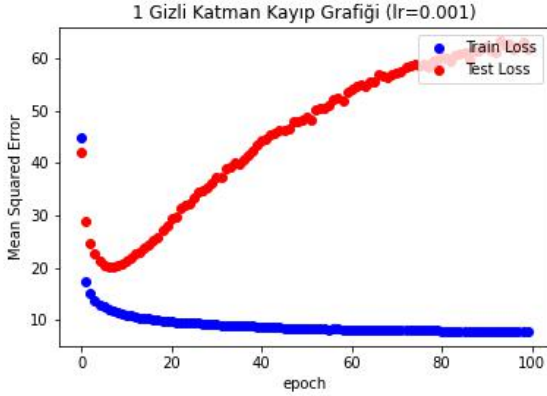
Figür 5



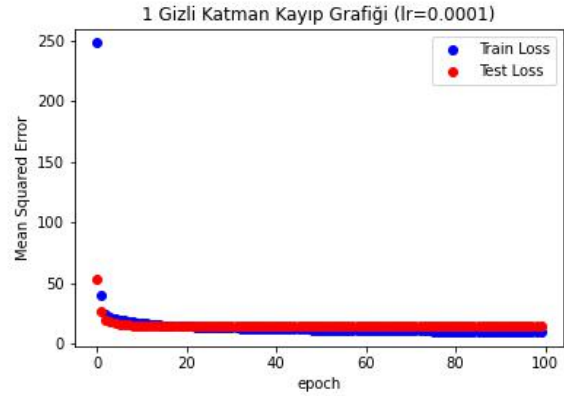
Figür 6

Bu modelin learning rate yükseldikçe daha iyi sonuç verdiği gözlemlenmiştir

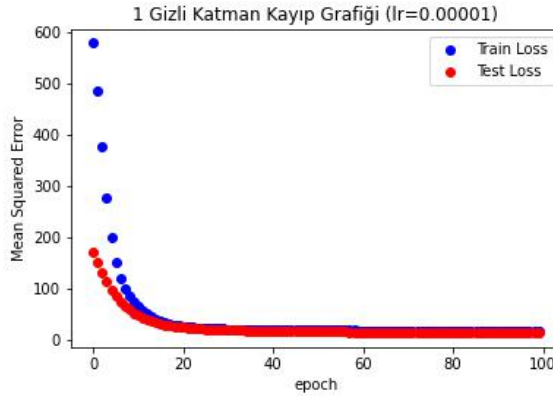
Tek gizli katmanlı model $lr=0.00001$, 0.0001 , 0.001 için gözlemlendi. Bu modellerin kayıp grafikleri sırasıyla Figür 7, Figür 8 ve Figür 9’de gösterilmiştir.



Figür 7



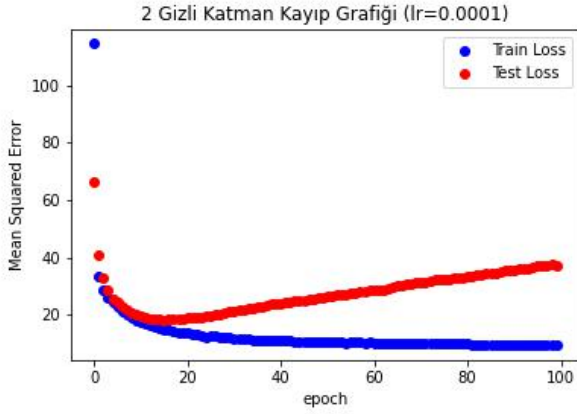
Figür 8



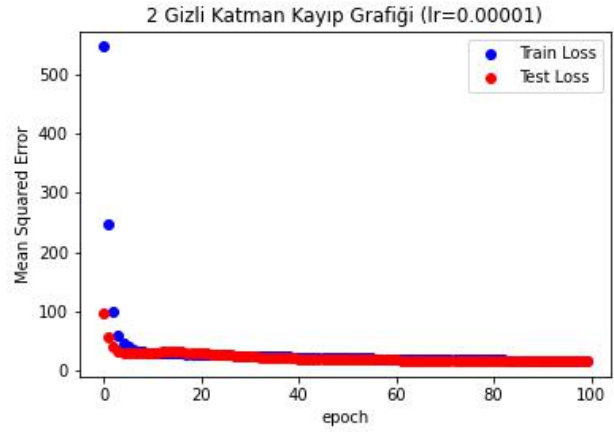
Figür 9

Bu modelde en iyi sonuç $lr=0.0001$ için elde edildi. Model daha yüksek bir learning rate ile doğru tahminlere yaklaşıyor. Daha küçük bir learning rate için ise daha yavaş şekilde doğru tahminlere yaklaşıyor.

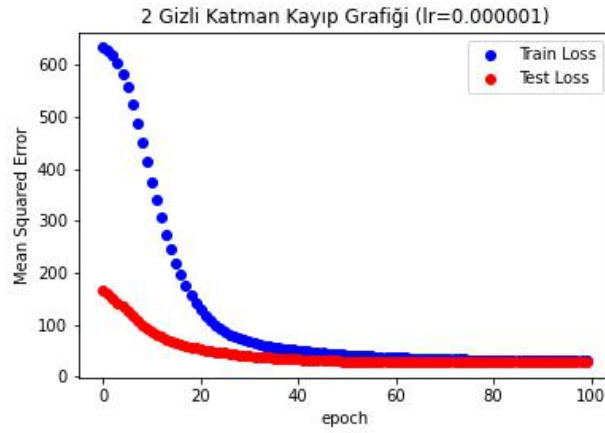
İki gizli katmanlı model ise $lr=0.0001$, 0.00001 , 0.000001 için gözlemlendi. Bu modellerin kayıp grafikleri sırasıyla Figür 10, Figür 11 ve Figür 12’de gösterilmiştir.



Figür 10



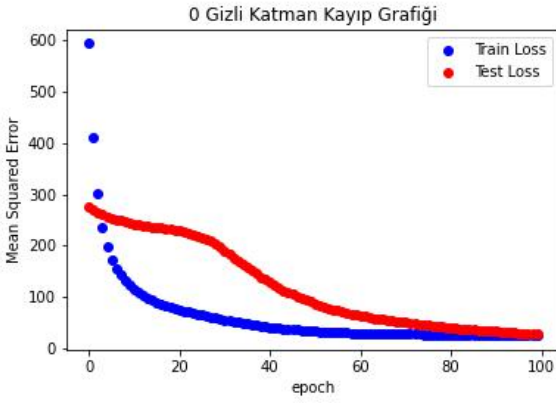
Figür 11



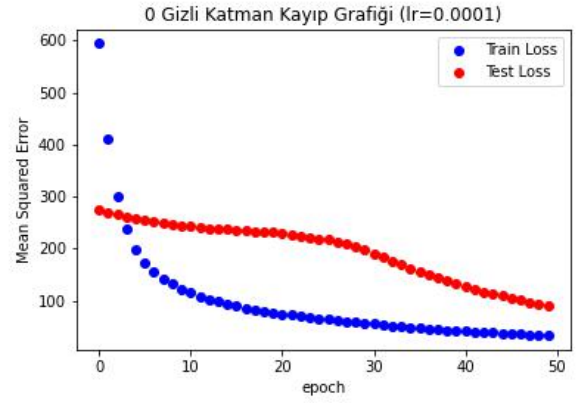
Figür 12

Bu modelde ise en iyi sonuç $lr=0.00001$ için elde edildi. Tek katmanlıdaki gözlemlendiği üzere iki gizli katmanlı modelde daha yüksek bir learning rate ile doğru tahminlere yaklaşılamazken daha küçük bir learning rate için ise daha yavaş şekilde doğru tahminlere yaklaşıyor.

0 katmanlı model, Figür 1'de 100 epoch boyunca ve Figür 6'da 50 epoch boyunca $lr=0.0001$ için eğitilmiştir. Bu figürler tekrar incelendiğinde, Figür 1'in test ve train kaybı azalarak birbirine yaklaşıırken, Figür 6'da kayıpların daha yüksek ve birbirlerinden uzak olduğu gözlemlenmektedir. Bu figürler karşılaştırılma yapılabilmesi için tekrar sırasıyla Figür 1.1 ve Figür 6.1'de verilmiştir.



Figür 1.1



Figür 6.1

Doğru eğitim miktarı hesaplanırken, modelin katmanları, nöron sayıları, öğrenim hızı (learning rate), kullanılan aktivasyon fonksiyonları, verisetinin boyutu gibi bir çok parametre göz önünde bulundurulmalıdır. Eğitim miktarı ile ilgili bazı yorumlara raporun önceki sayfalarında da yer verilmiştir.

Üzerinde çalışılan Boston Housing veriseti için,

- 0 gizli katmanlı model için en iyi parametreler: $lr=0.01$, epoch=50, train loss= [23.24738137], test loss= [26.74012985]
- Tek gizli katmanlı model için en iyi parametreler: $lr=0.0001$, epoch=100, nöron sayısı=12, train loss= [15.37723945], test loss= [15.60485321]
- İki gizli katmanlı model için en iyi parametreler: $lr=0.00001$, epoch=400, nöron sayıları=12-12, train loss= [17.21537575], test loss= [18.32953712]
- Tüm modeller normalizasyon kullanılmadan, min-max normalizasyon ile ve standart normalizasyon ile denendi. En iyi sonuç standart normalizasyon ile alındı.
- Tüm modeller arasından en iyi performansın elde edildiği modelin özellikleri:
 1. Tek katman
 2. 12 nörona sahip
 3. ReLU aktivasyon fonksiyonu kullanıldı
 4. Ağırlıklar ve biasın ilk değeri $2 \cdot (\text{random}) - 1$ şeklinde tanımlandı
 5. 100 epoch boyunca eğitildi
 6. $lr=0.0001$
 7. Standart normalizasyon kullanıldı
 8. Eğitim hatası: [15.37723945]
 9. Test hatası: [15.60485321]