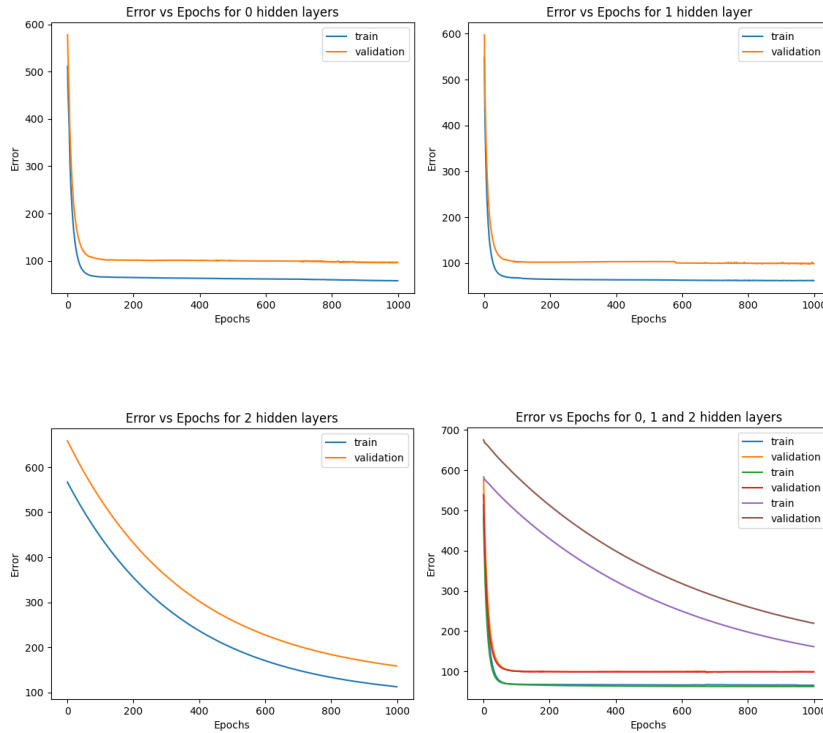


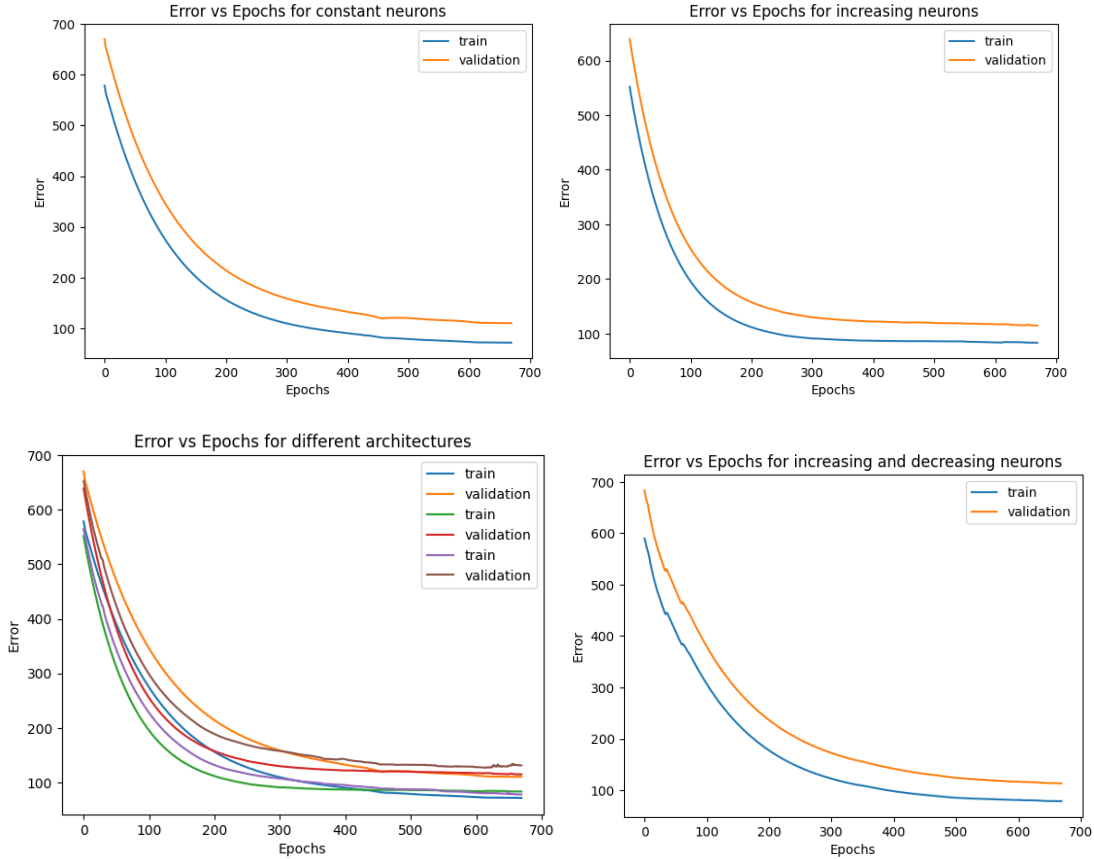
# YAP470 ÖDEV2 RAPORU

**1)** Gizli katman sayısının sırasıyla 0, 1 ve 2 olduğu 3 tane aynı eğitim verisi üzerinde yapay sinir ağı modeli eğitilmiştir. Modellerde oluşturulan gizli katmanlardaki nöron sayıları (64) ve öğrenme kat sayısı (0.000001) aynıdır. Mümkün olduğunca değişen tek durumun katman sayısı olması sağlanmıştır. Modellerin eğitim 1000 adım ile sağlanmıştır. Aşağıdaki grafiklerde eğitim sırasında değişen eğitim ve doğrulama hatası gözlemlenmiştir.

Sonuç Tartışma: Katman sayısı arttıkça model daha karmaşık hale gelmiştir dolayısıyla aynı öğrenme kat sayısı ile daha fazla katmanda model yeteri kadar öğrenememiştir. 1000 adım sonucu modellerin eğitim, doğrulama ve test hataları sırasıyla şu şekildedir: 0 gizli katman sayılı model hataları (61.82,99.96,48.03), 1 gizli katman sayılı model hataları (61.31,98.54,51.03), 2 gizli katman sayılı model hataları (112.74,158.71,89.42). Son model için öğrenme kat sayısı artırılıp model tekrar eğitildiğinde hata değerleri sırasıyla (71.20, 104.53, 53.62) şeklinde değişmiştir. Bu şartlarda ve bu eğitim verisi üzerinde modelimize ekstra gizli katman eklemek bir katkı sağlamamıştır. Bu durumda daha basit olan model kullanılmalıdır. Fakat gizli katmanlardaki nöron sayıları üzerinde değişiklik sağlanabilirse daha fazla katman olan modellerde gelişme sağlanabilir. (Bunun yapılp yapılamayacağı ödev dokümanında belirtilmemiştir.)



**2)** Modeller 2 gizli katmanlı şekilde nöron sayılarını belirlemede farklı yöntemler denenerek eğitilmiştir. 3 farklı model eğitilmiştir bu modellerde sırasıyla bu yöntemler uygulanmıştır: Gizli katmanlarda sabit nöron sayısı ((13,64)(64,64)(64,1)), Gizli katmanlarda azalan miktarda nöron sayısı ((13,128)(128,64)(64,32)(32,1)), Gizli katmanlarda önce artan sonra azalan şekilde nöron sayısı ((13,64),(64,128)(128,64)(64,1)). Adım sayısı 670 olarak seçilmiştir.



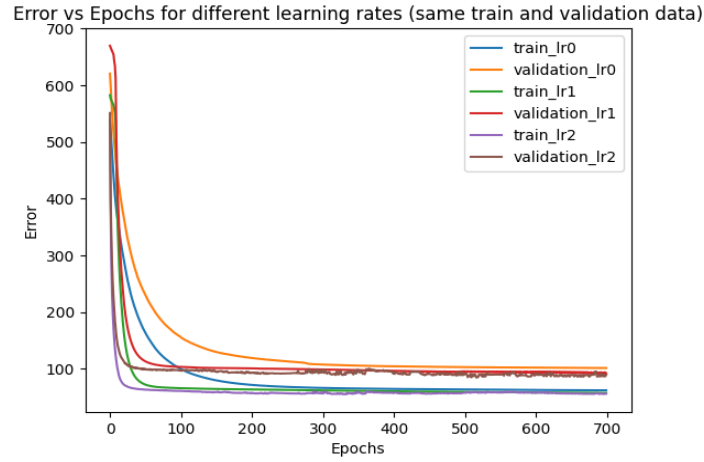
Sonuç ve Tartışma: Eğitilen ilk model için eğitim, doğrulama ve test hataları sırasıyla (72.40 ,110.05, 53.24) eğitilen ikinci model için hatalar sırasıyla (78.51, 105.44, 61.00) ve eğitilen üçüncü model için hatalar sırasıyla ( 79.33, 115.22, 59.05) şeklinde gözlemlenmiştir. Önceki duruma benzer şekilde basit olan model tüm hata değerleri üzerinde daha iyi bir öğrenme sağlamıştır. Bu şartlarda ve bu veri üzerinde daha nöron sayısının sabit olması daha farklı bir deyişle daha basit bir modelin seçilmesi daha uygun olduğu gözlemlenmiştir.

**3)** Farklı öğrenme parametreleri altında yine 3 model geliştirilmiştir. Öğrenme parametreleri sırasıyla 0.0000001, 0.0000005, 0.000001 seçilmiştir. Normal yapay sinir ağı kütüphanelerinde bu değerler çok daha yüksek olarak kullanılmaktadır. Genel olarak modellerimizde bu aralığa göre daha küçük oranlar tercih etmemizin sebebi az veri olmasından kaynaklı batch boyutunu 1 olarak belirlememizden kaynaklanıyor olabilir.

## Sonuç ve Tartışma:

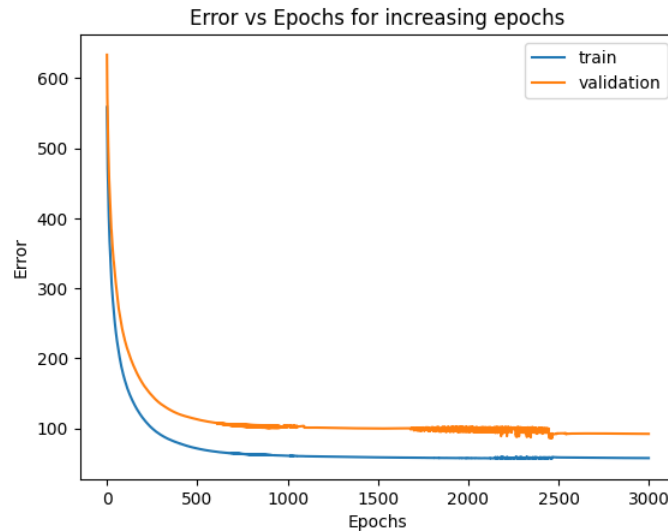
Bu farklı öğrenme katsayılarında modeller görüldüğü üzere hepsi convex mse yapısındaki minimum noktasına yönelip ulaşmıştır ya da küçük bir hata payı ile çevresindedir.

Öğrenme katsayı parametresi bu minimum noktasına ulaşmadaki adım sayısı üzerinde etkili olmuştur. Eğitim süreleri arasında inanılmaz bir fark olmadığı için öğrenme katsayısı küçük olan daha garanti bir sonuç çıkaracağı için seçilmesinin uygun olduğu gözlemlenmiştir.



**4)** Adım sayısı ve hata oranları arasındaki ilişkiyi göstermek için daha kompleks bir model geliştirilmiştir. Kompleks bir model geliştirilmesinin amaç fazla adım sayısındaki overfit durumunu gözlemlemektir. Bu model 4 gizli katmana sahip olup nöron sayıları ((13,64)(64,128)(128,256)(256,128)(128,64)(64,1)) şeklindedir.

Sonuç ve Tartışma: En başta bekleneneği üzere doğrulama ve eğitim verisi üzerinde ilk 500 adımda azalma gerçekleşmiştir. Çünkü model git gide öğrenmekte ve daha iyi bir şekilde tahmin etmektedir. Fakat amaçlanmış olan overfit durumu 3000 adım eğitilmesi sonucunda rastlanmamıştır. Bu durum grafikte açık bir şekilde görülmektedir. Overfit olması durumunda gözlemlenecek grafik doğrulama hatasının artarken eğitim hatasının düşme durumudur.



---

**5)**Hyperparametreler modelde farklı yöntemler ile nöron sayısı değişimi, farklı sayıda katmanlar, katmanlar arası farklı aktivasyon fonksiyonlarının seçilmesi, adım sayısı, öğrenme katsayısı olarak saptanmıştır. Bu parametreler arasında en iyisi bulunurken nöronların değişimi artan, azalan, önce artıp sonra azalan ve sabit olarak denenmiştir. Bu nöron sayıları belirlenirken 2'nin katları seçilmiştir. Katman sayısı olarak 0,1,2,3,4 denenmiştir. Ayrıca aktivasyon fonksiyonları olarak tanh, relu, sigmoid, softmax seçilmiştir. Modeller 500, 1000, 2000 ve 3000 adım sayısında eğitilmiştir. Modellerde denen en öğrenme katsayısı sırasıyla 0.00001, 0.000005, 0.000001, 0.0000001 şeklindedir. Tüm bunların değerlendirilmesi sonucu çok fazla grafik oluşacağı için çizdirilip rapora eklenmemiştir.

Sonuç ve Tartışma En İyi Model:

Bütün modeller ortalama 1000 adım sayısı sonunda modelin mse minimum noktasına yönelip küçük bir hata payı ile ulaşmıştır. En iyi aktivasyon fonksiyonları nöron sayısı ve katman sayısına göre değişiklik göstermektedir. En iyi nöron yapısı azalan nöron sayısı olarak gözlemlenmiştir. Çoğu modelde azalan nöron sayısı daha iyi şekilde sonuç vermektedir.

- Gizli katman Sayısı: 1
- Nöron Sayısı: ((13,128)(128,32)(32,1))
- Kullanılan aktivasyon fonksiyonu: ara katmanda sigmoid, son katmanında relu
- Ağırlık ilk değerlerinin tanımlanma biçimi: önceki ve sonraki nöron sayısına göre +/-0.5 arası eşit dağılım
- Epoch sayısı: 1000
- Normalizasyon: Kullanılmadı
- Batch Boyutu:1
- Eğitim hatası: 36.212223
- Test hatası: 34.012111