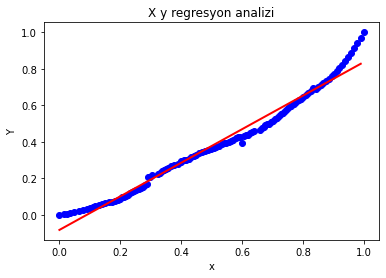
**Regresyon Analizi**



Bu doğrusal regresyon analizde y bağımlı değişken olarak kış covid-19 enfekte verileri, x bağımsız değişken olarak yaz covid-19 enfekte verileri kullanılmıştır. Veriler modelde kullanmadan önce minimum-maksimum normalizasyonu kullanılarak verileri 0-1 aralığına ölçeklendirilmiştir. Elde edilen regresyon modelinde denklemi şu şekildedir

Y=[-0.08135807]+X[0.91811684]

**Modelin Performans Ölçümü**

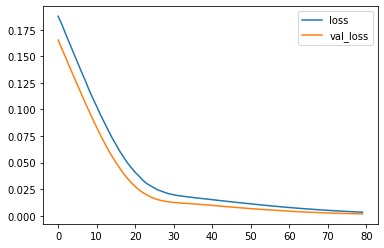
Determinasyon katsayısı ile bulunan değerde, model 0.9758 oranında verilere uyum sağladı. (1’e yakın daha performanslı)

Ortalama karesel hata yöntemi ile regresyon eğrisinin bir dizi noktaya 0.0016 ornanında yakın olduğu bulundu.(0’a yakın daha performanslı)

Ortalama mutlak hata ile bir dizi tahmindeki hataların ortalama büyüklüğü 0.0283 olarak bulundu.(0’a yakın daha performanslı)

**Derin Öğrenme**

**Öğrenme Eğrisi**

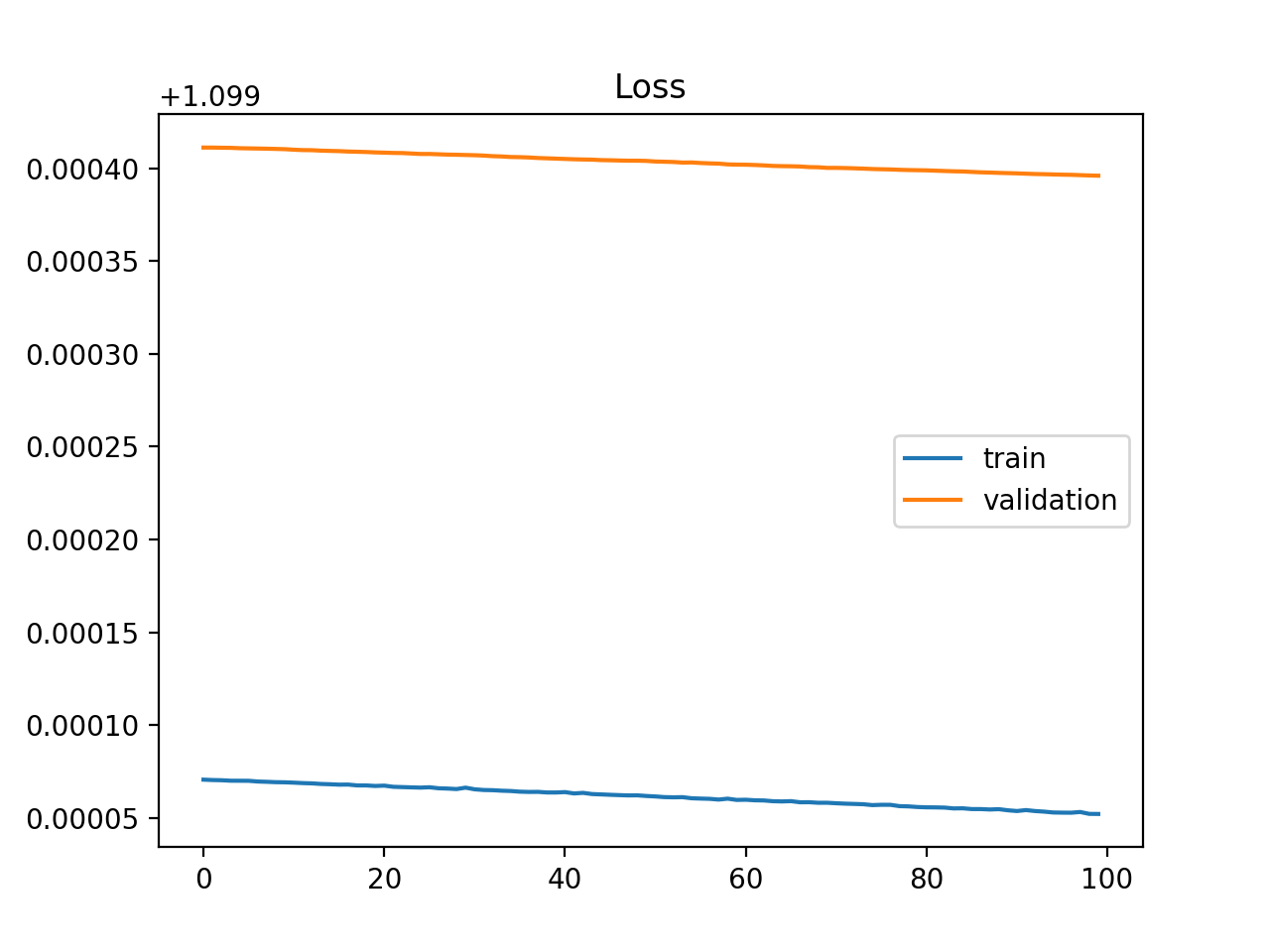


Bu derin öğrenme yönteminde y bağımlı değişken olarak kış covid-19 enfekte verileri, x bağımsız değişken olarak yaz covid-19 enfekte verileri kullanılmıştır. Kış verileri modelde kullanmadan önce minimum-maksimum normalizasyonu kullanılarak verileri 0-1 aralığına ölçeklendirilmiştir. Bu modelde 84 tane eğitim verisi 36 tane test verisi olmak üzere toplam 120 veri vardır. Öğrenme için 4 tane katman eklenmiş ve bu katmanlarda doğrusal olmayan “Rectified Linear Unit (ReLU)” aktivasyon fonksiyonu kullanılmıştır. Ardından hesaplama aşamasında hesaplama açısından verimli ve küçük hafıza gereksinimi olduğu için “Adam” optimizasyon algoritması, kayıp değerler içinse “Mean Squared Error (MSE)“ metriği kullanılmıştır.

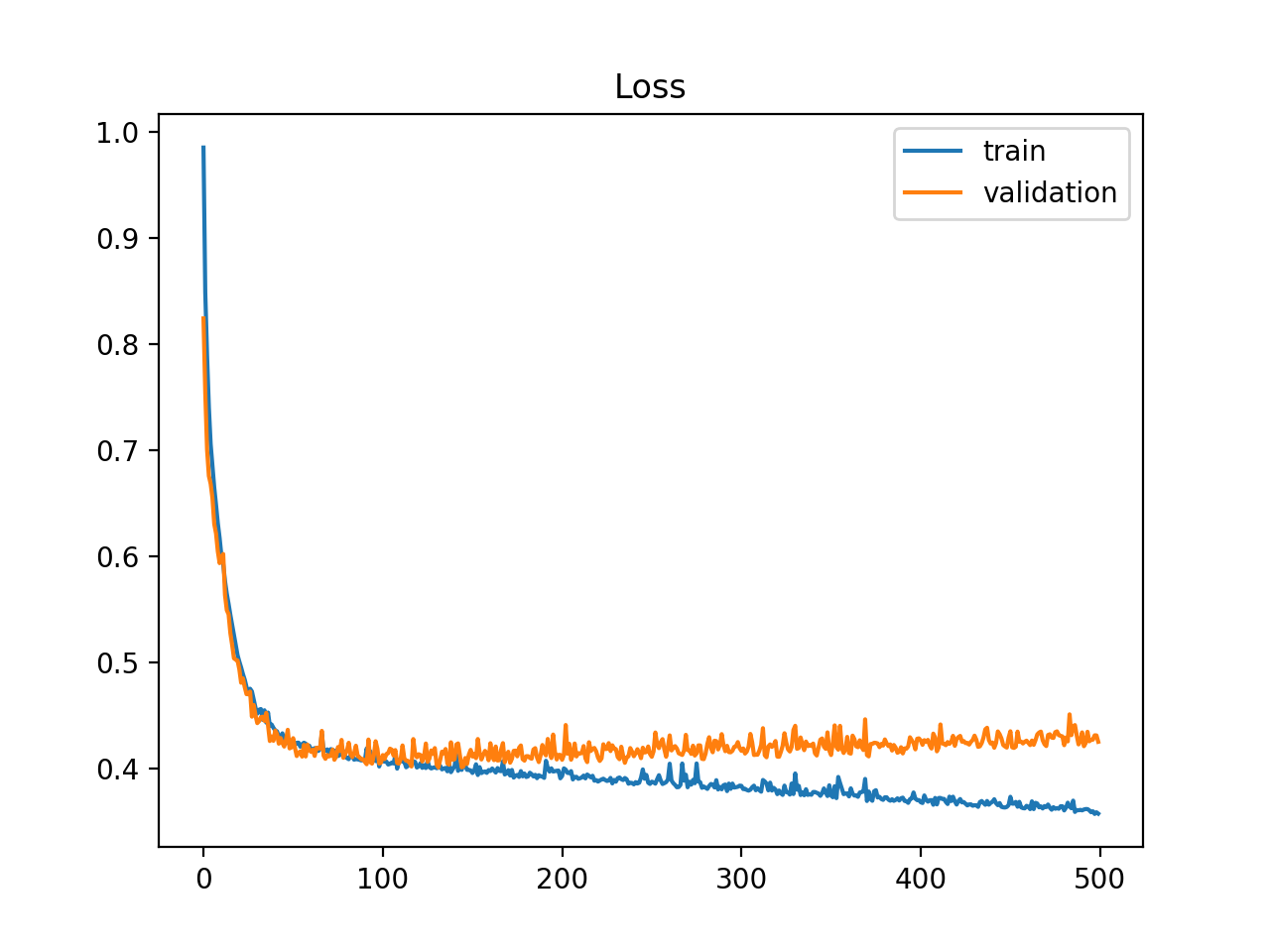
**Modelin Performansı**

Bir öğrenme eğrisinin şekli ve dinamiği derin öğrenme modelinin performansı hakkında bilgi verir. Y eksenindeki düşük değerler görece daha iyi bir model performansına işaret eder. Üç ortak dinamik şu şekildedir:[[1]](#footnote-2)

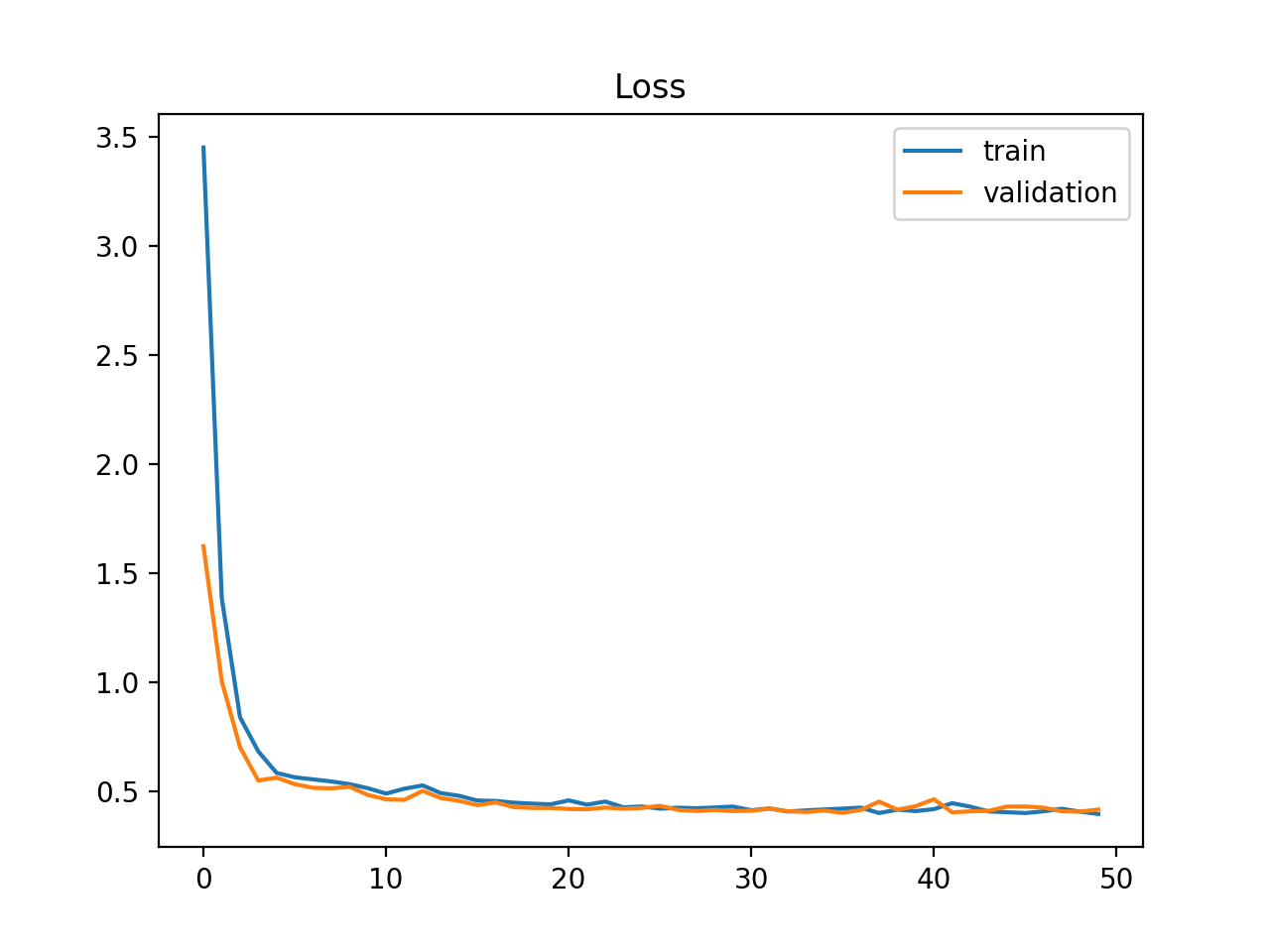
1. Uyumsuz(Underfit)



2.Aşırı Uyum (Overfiting)



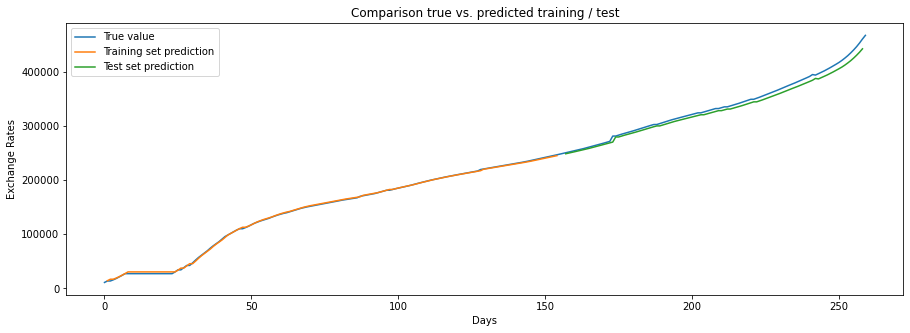
3. İyi Uyum(Good Fit)



**Değerlendirme**

Görüldüğü üzere eğitilen model 2. modele (overfiting) daha yalın durmaktadır. Bu model eğitim setini çok iyi öğrenmiş bir modele işaret eder. Problemi, veri setine aşırı bağlı olabilir bu da veri setini genellerken sorun yaratabilir. Model uzun süre eğitilirsede ortaya çıkabilir.[[2]](#footnote-3)

**Yapay Sinir Ağları**



Bu yapay sinir ağı yönteminde -LSTM- eğitim verisi rastgele olarak 154 tane 2020 yılının Covid-19 vaka sayısı verileri, rastgele test verisi olarak 102 tane 2020 yılının Covid-19 vaka sayısı verileri kullanılmıştır. Tüm veriler modelde kullanmadan önce minimum-maksimum normalizasyonu kullanılarak verileri 0-1 aralığına ölçeklendirilmiştir. Öğrenme için 1 tane katman eklenmiş ve bu katmanda LSTM kullanılmıştır. Ardından hesaplama aşamasında hesaplama açısından verimli ve küçük hafıza gereksinimi olduğu için “Adam” optimizasyon algoritması, kayıp değerler içinse “Mean Squared Error (MSE)“ metriği kullanılmıştır. Devir (epoch) sayısı 80 olarak belirlenmiştir.

**Modelin Performans Değerleri**

Birçok kez eğitildi:

Training MAPE: 147,481287 (0‘a yakın daha iyi)

Training data score: 0,114019 RMSE(0’a yakın daha iyi)

Test MAPE: 47,96095358 (0‘a yakın daha iyi)

Test data score: 0,444447 RMSE (0‘a yakın daha iyi)

İlkkez eğitildi:

Training MAPE: 147,543607 (0‘a yakın daha iyi)

Training data score: 0.122103 RMSE (0‘a yakın daha iyi)

Test MAPE: 47.63319095 (0‘a yakın daha iyi)

Test data score: 0.688683 RMSE (0‘a yakın daha iyi)

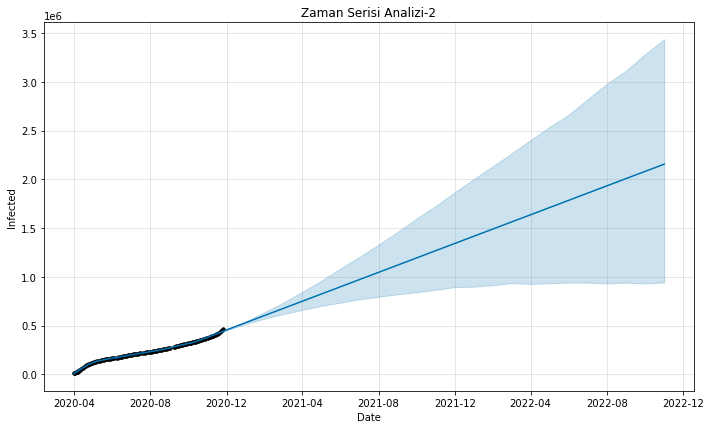
Test data Bias: -0.32669184 (0‘a yakın daha iyi)

Training data Bias: -0.14301209 (0‘a yakın daha iyi)

**Performans Değerlendirmesi**

Eğitim verisinin MAPE metriğinin değeri dışındaki veriler bu model için kabul edilebilir performans değerleridir. Eğitim versinin MAPE değeri ise daha fazla değer aldığı için daha yüksek bir sonuç çıkmıştır.

**Zaman Serisi Analizi**



Yukarıdaki zaman serisi analizinde -veya tahminlemesinde- Facebook Prophet kütüphanesi kullanılmıştır. Covid-19 vaka sayısını yaz sezonu ve diğerleri olarak ayrıldı. Test verisi yaz aylarındaki vaka sayıları olarak alındı, eğitim verisi ise diğer kalan günlerdeki vaka sayıları olarak alındı. Test verisi sayısı 90, eğitim verisi sayısı 169 olmak üzere toplam 259 veri vardır. Prophet forecasting yöntemi ile tahminleme yapılmış ve üsteki grafik elde edilmiştir.

**Modelin Performans Verisi**

Bu modelde zaman serisi tahminlemeleri için uygun olan Forecast Bias yöntemi kullanılmıştır. 0’ a yakın değer daha iyidir.

Forecast Bias: -0.376644

**Modelin Performans Değerlendirmesi**

Kullanılan eğitim ve test verilerine göre modelin performansı ortalamadır. Grafikte vaka sayısı -herhangi bir bağımsız değişkenden etkilenmezse- doğrusal olarak ilerlemektedir.

1. Kaynak:https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/ [↑](#footnote-ref-2)
2. Kaynak: https://machinelearningmastery.com/learning-curves-for-diagnosing-machine-learning-model-performance/ [↑](#footnote-ref-3)