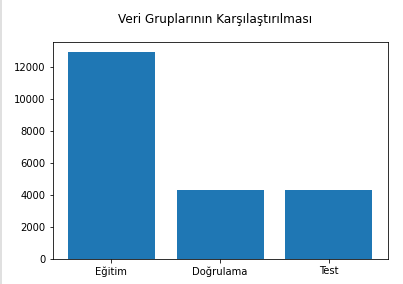
**YMH418 - Yazılım Mühendisliği Güncel Konular Dersi Aşama 5 Raporu**

1. **Giriş**

Veri seti karıştırıldı ve gruplara ayırıldı. Geçen hafta oluşturulan model eğitildi ve test edildi.

1. **Verilerin Karıştırılması ve Gruplara Ayırılması**

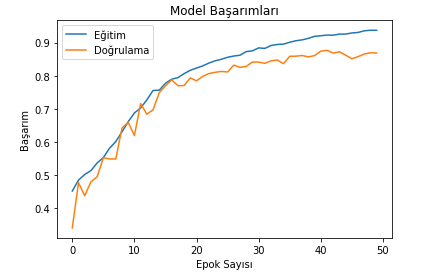
Toplanılan veri yıl sırasına göre dizilmiş haldeydi. Sıralanmış verinin eğitimde modelin veriyi ezberlemesine neden olabileceği için, Python programlama dilin kütüphanelerinden olan sklearn kütüphanesinin shuffle modulü kullanılarak veri kendi içerisinde karıştırılmıştır. Daha sonra karıştırılan bu veriler eğitim, doğrulama ve test olarak 3 gruba ayırıldı.



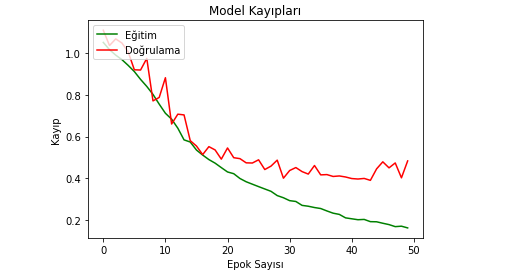
Şekil 1. Veri gruplarının büyüklüklerinin karşılaştırılması

1. **Modelin Eğitilmesi ve Test Edilmesi**

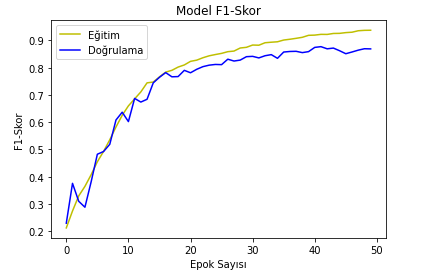
Projenin amacı yeni gelen hava kirliliği verilerinin istasyon tipini tahmin eden bir model eğitmektir. Bu amaç için oluşturulan LSTM modeli eğitim veri grubuyla eğitildi ve doğrulama veri grubu ile eğitim sonuçları doğrulandı. Modelin doğrulama başarımı %75 civarındadır kayıp ise %58 civarındadır bu durum bazı durumlarda sorun olmazken bazı durumlarda ezberlemenin işareti olmaktadır. Eğitim ve doğrulama çıktı değerleri aşağıdaki grafiklerde görüldüğü gibidir.



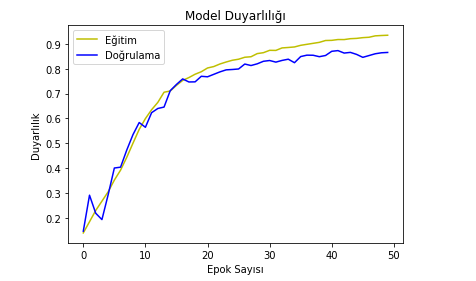
Şekil 2.1. Başarım Grafiği



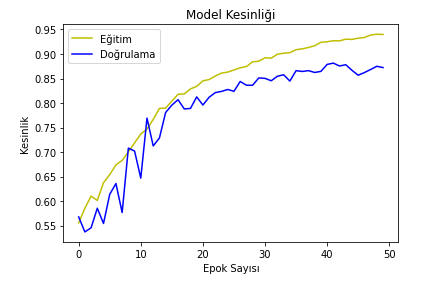
Şekil 2.2. Kayıp Grafiği



Şekil 2.3. F1-Skor Grafiği

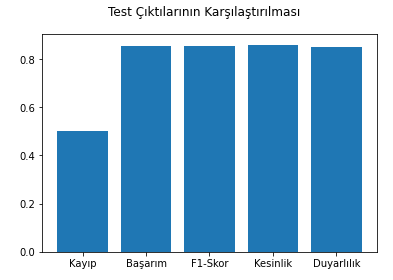


Şekil 2.4. Model Duyarlılığı



Şekil 2.5. Model Kesinliği

Test sonuçlarında %85 başarım vardır, başarım değeri eğitim ve doğrulamanın üzerine çıkmışken kayıp değeri hemen hemen aynı kalmıştır. Model bu problemi çözmekte başarılı olmuştur. Test sonucunda elde edilen veriler aşağıda görüldüğü gibidir.



Şekil 2.6. Test Çıktılarının Karşılaştırılması

1. **Sonuç**

Model eğitimde ve test verilerinde başarılı olmuştur. Verileri artırarak tekrar test edilmesi ve gerekirse eğitilmesi kayıp değerlerini daha düşük seviyelere çekebilir.