**PROJE RAPORU**

**İnovasyon Mühendislik Ltd. Şti.**

**Proje Adı:** Lithium-Ion Batarya SoH Analizi

**Danışman**

Okan Ulu

**Hazırlayan**

Hasan Taşkın

29.07.2024

Eskişehir

İçindekiler

**1. Giriş** 3

**2. Veri Seti ve Ön İşleme** 3

**2.1 Veri Seti** 3

**2.2 Veri Yükleme ve İnceleme** 3

**2.3 Eksik Veri Kontrolü** 5

**2.4 SoH (State of Health) Değerinin Hesaplanması** 6

**2.5 Gürültülü Verilerin Temizlenmesi (IQR Yöntemi)**7

**2.6 SoH Değerlerinin Görselleştirilmesi**8

**2.7 Veri Normalizasyonu**9

**3. Keşifsel Veri Analizi (EDA)** 10

**3.1 Korelasyon Analizi ve Heatmap Oluşturma** 10

**3.2 Özellik Önemini Belirleme**11

**4. Model Eğitim ve Değerlendirme** 12

**4.1 Veri Setlerinin Eğitim ve Test Setlerine Ayrılması**12

**4.2 Model Eğitimi ve Değerlendirme**13

**4.2.1 K-Nearest Neighbors (KNN)**13

**4.2.2 Gradient Boosting Regressor** 14

**4.2.3 Support Vector Regression (SVR)**14

**4.2.4 Neural Network (MLPRegressor)**15

**4.2.5 Random Forest Regressor**15

**4.3 Modellerin Karşılaştırılması ve Performans Grafiği**16

**4.4 Modeli Pickle ile Kaydetme**17

**5. Sonuç ve Değerlendirme**18

**1. Giriş**

Bu proje, lityum-iyon bataryaların sağlık durumu (SoH) tahmin modelinin geliştirilmesini ve performanslarının değerlendirilmesini hedeflemektedir. Bataryaların etkili ve verimli bir şekilde kullanılmasını destekleyerek enerji depolama teknolojilerinin gelişimine katkıda bulunmayı amaçlamaktadır. Proje süreci, veri ön işleme, keşifsel veri analizi (EDA), model geliştirme ve performans değerlendirme aşamalarını içermektedir.

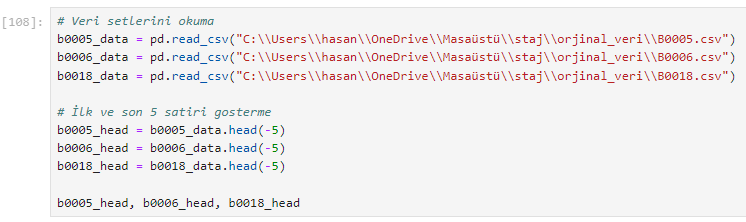
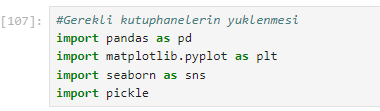
**2. Veri Seti ve Ön İşleme**

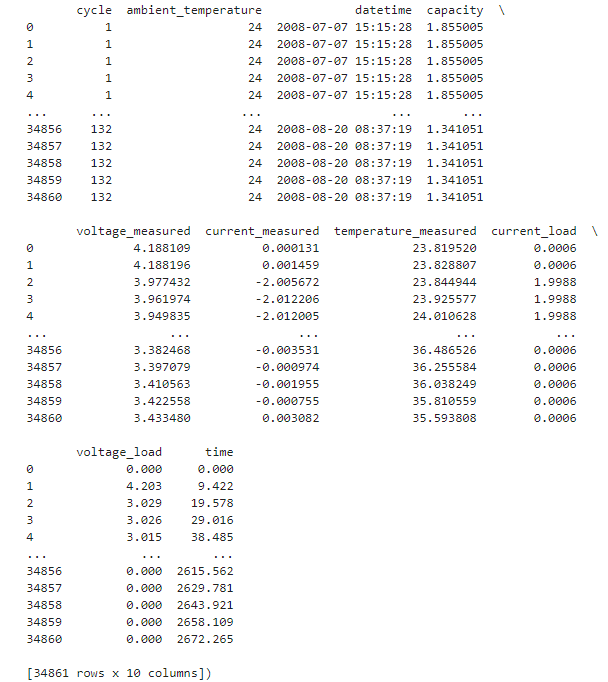
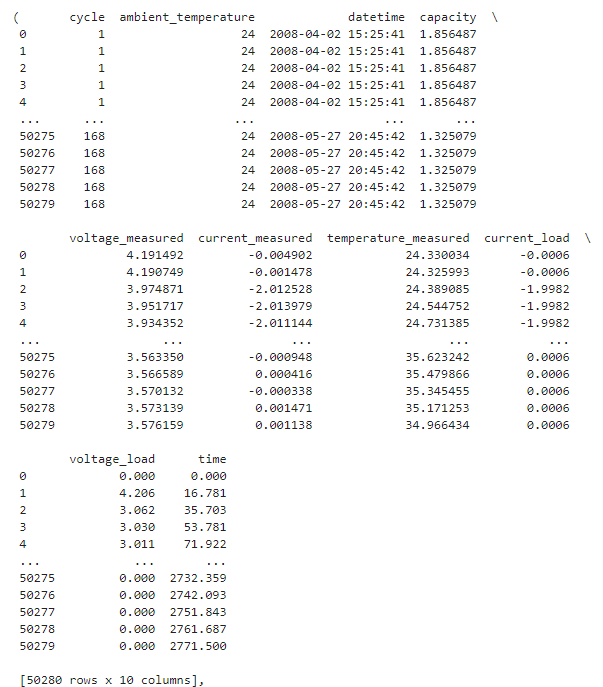
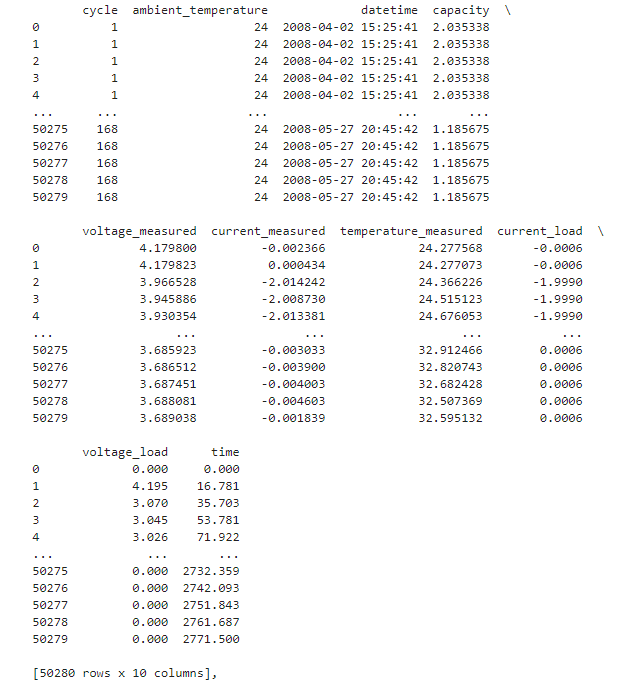
**2.1 Veri Seti**

Bu projede kullanılan veri setleri, üç farklı lithium-ion bataryaya ait performans verilerini içermektedir: B0005, B0006, ve B0018. Her bir veri seti, bataryaların kapasite, döngü sayısı, akım, voltaj ve diğer ilgili parametrelerini içermektedir.

**2.2 Veri Yükleme ve İnceleme**

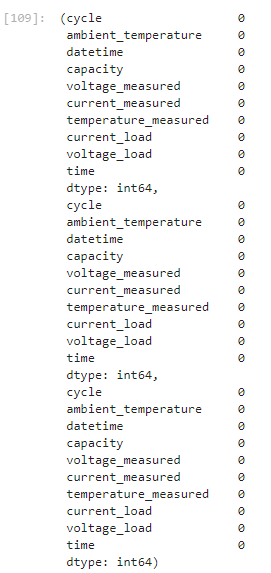
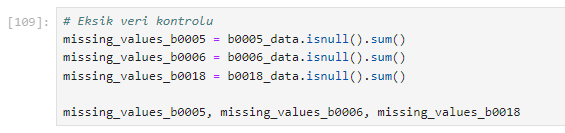
Veri setleri, pandas kütüphanesi kullanılarak CSV dosyalarından okunur ve incelenir. Bu adımda, veri setlerinin ilk ve son birkaç satırı gösterilerek verilerin yapısı hakkında bilgi edinilir.





**2.3 Eksik Veri Kontrolü**

Veri setlerinde eksik veri olup olmadığı kontrol edilir. Eksik veri analizi, verinin kalitesini ve modelleme sürecine olan etkisini anlamak için gereklidir. Eksik veri kontrolü, isnull().sum() fonksiyonu kullanılarak yapılır.



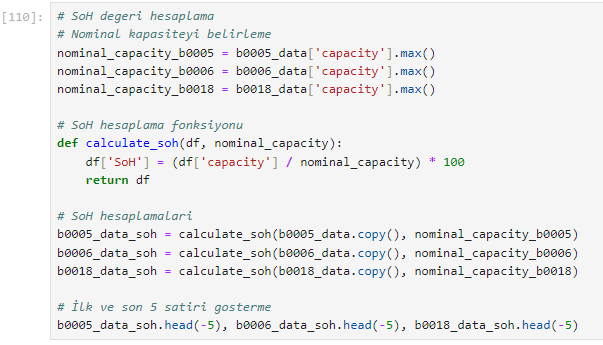
**2.4 SoH (State of Health) Değerinin Hesaplanması**

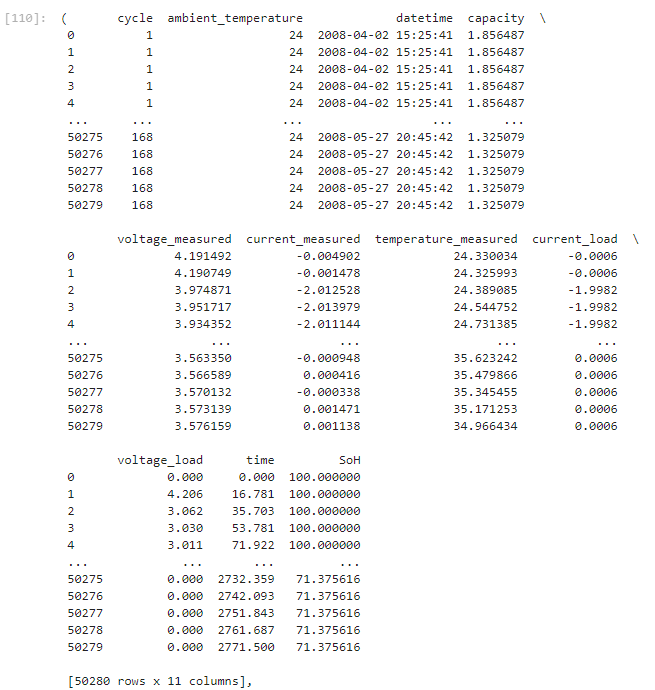
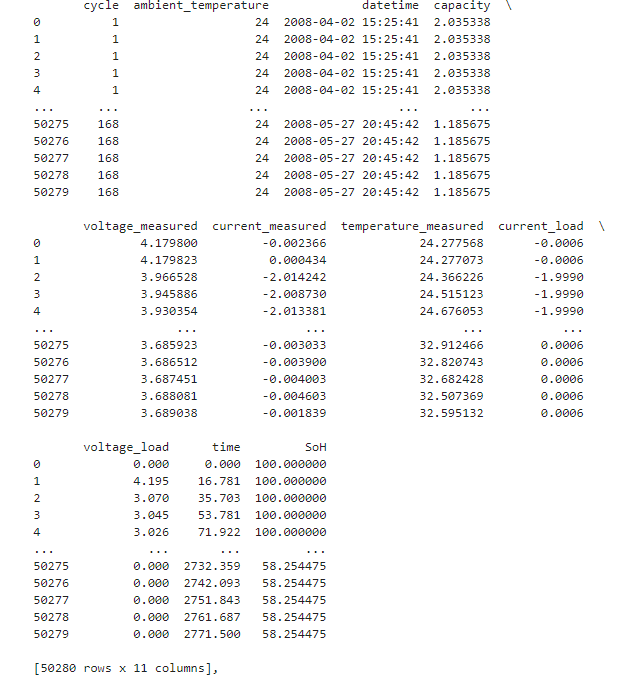
SoH (State of Health), bataryanın sağlık durumunu yüzdelik olarak gösterir. Bu değerin hesaplanması için öncelikle her bataryanın nominal kapasitesi belirlenir. Nominal kapasite, bir bataryanın tasarım gereği maksimum kapasitesini temsil eder ve bataryanın tamamen şarj edildiğinde sağlayabileceği enerji miktarını ifade eder. Genellikle ampere-saat (Ah) cinsinden ölçülür.

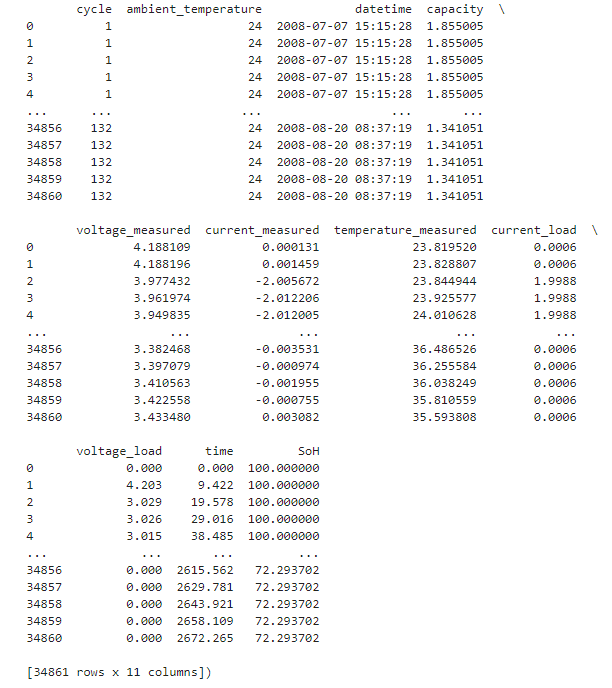
Bataryanın kullanım ömrü boyunca kapasitesi azalabilir, bu nedenle SoH hesaplamalarında nominal kapasite referans olarak kullanılır. SoH, bataryanın mevcut kapasitesinin nominal kapasiteye oranı olarak hesaplanır ve yüzde olarak ifade edilir. Örneğin, bir bataryanın nominal kapasitesi 2 Ah ise ve mevcut kapasitesi 1.5 Ah ise, SoH %75 olur.

Veri setinizde nominal kapasiteyi belirlemek için, veri setindeki maksimum kapasite değerini kullanırız çünkü bu değer, genellikle bataryanın yeni durumdayken sahip olduğu kapasiteyi temsil eder.

Aşağıda bataryaların SoH hesaplaması yapılmış ve verilerin son hali incelenmiştir.

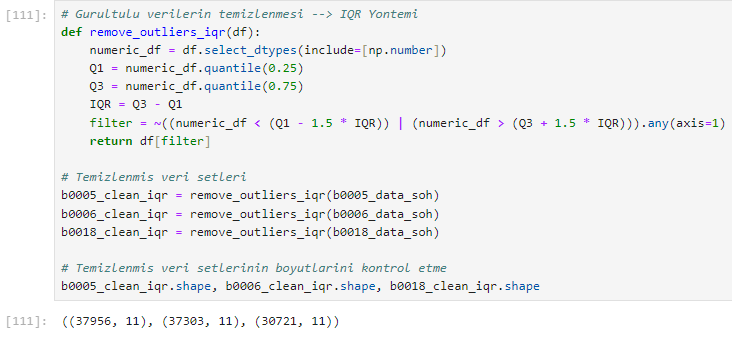






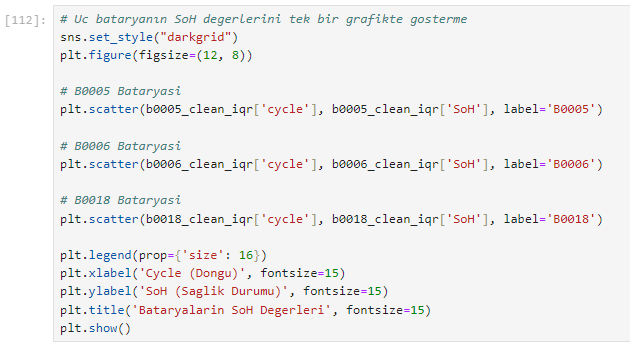
**2.5 Gürültülü Verilerin Temizlenmesi (IQR Yöntemi)**

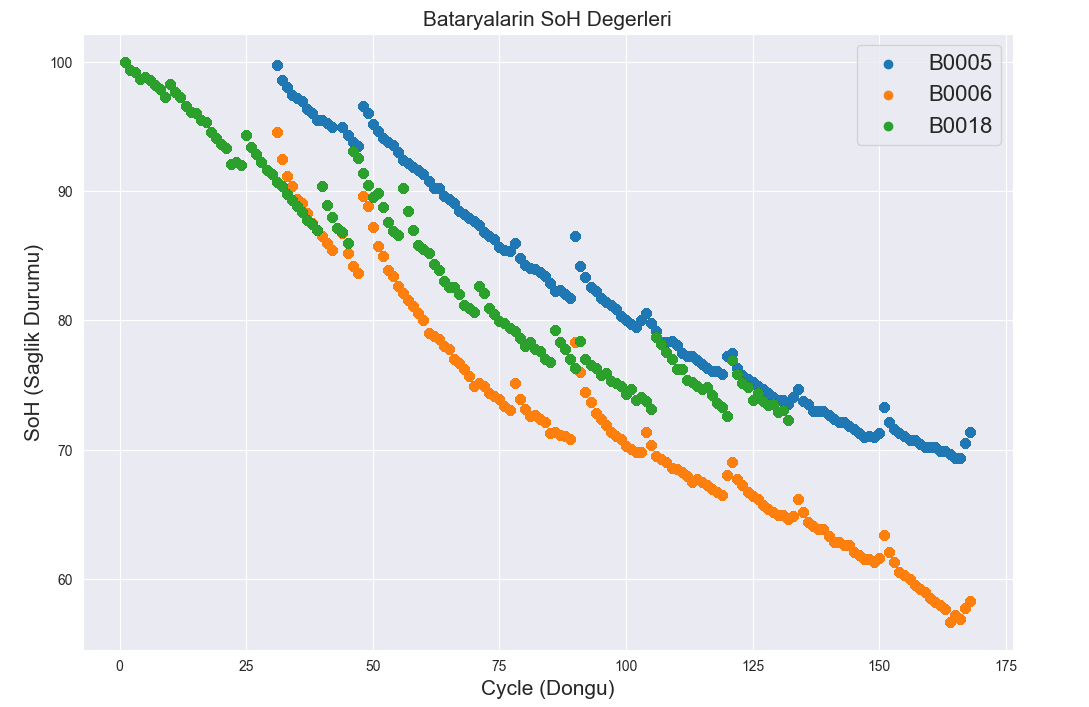
Gürültülü veriler, veri setlerinde normal dağılımdan sapmış anormal değerlerdir. Bu tür veriler modelin performansını olumsuz etkileyebilir. IQR yöntemi, bu tür verilerin temizlenmesinde kullanılan yaygın bir yöntemdir. Bu yöntemde, veri setinin birinci ve üçüncü çeyrek değerleri (Q1 ve Q3) hesaplanır ve aykırı değerler bu aralığın dışındaki değerler olarak tanımlanır.



**2.6 SoH Değerlerinin Görselleştirilmesi**

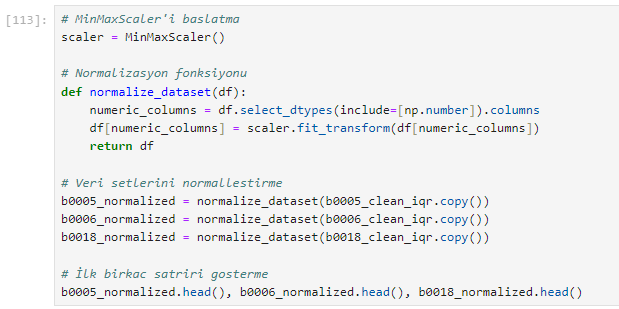
Veri görselleştirme, veri analizinin önemli bir parçasıdır. Bu adımda, her üç bataryanın SoH değerleri bir scatter plot kullanılarak görselleştirilir. Bu grafik, bataryaların döngü sayısına göre SoH değişimini gösterir.

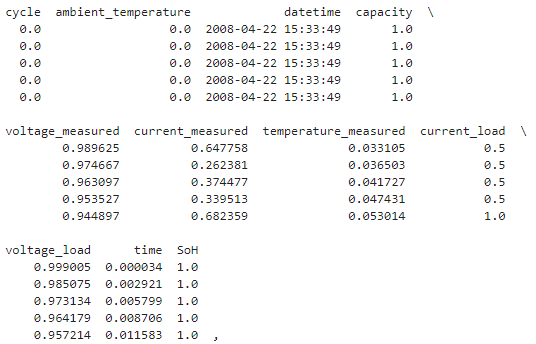
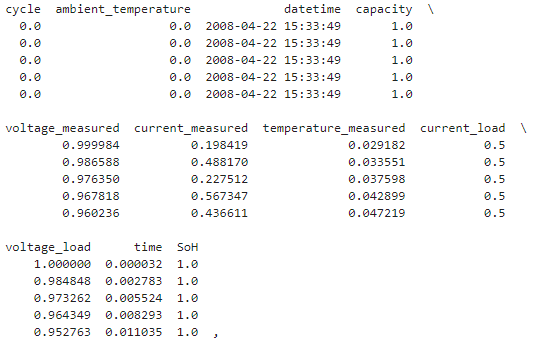




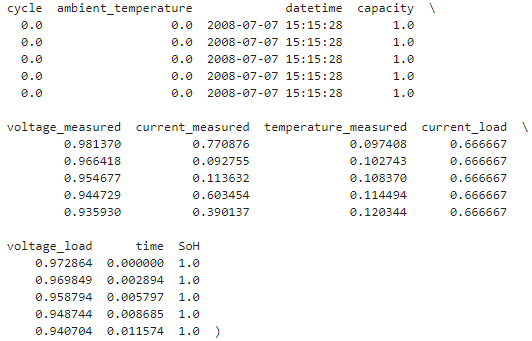
**2.7 Veri Normalizasyonu**

Veri setlerindeki sayısal değerler MinMaxScaler kullanılarak 0 ile 1 arasında ölçeklendirilir. Normalizasyon, verilerin belirli bir ölçeğe indirgenmesi işlemi olup, model eğitiminde verilerin daha iyi performans göstermesini sağlar.





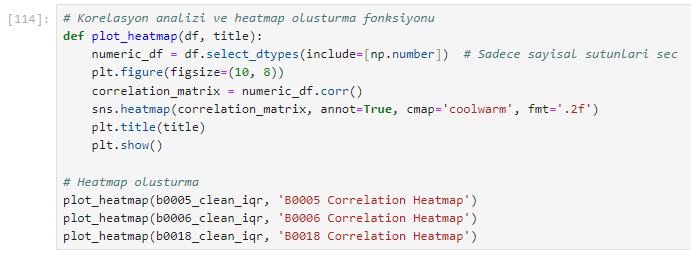
B0005 B0006

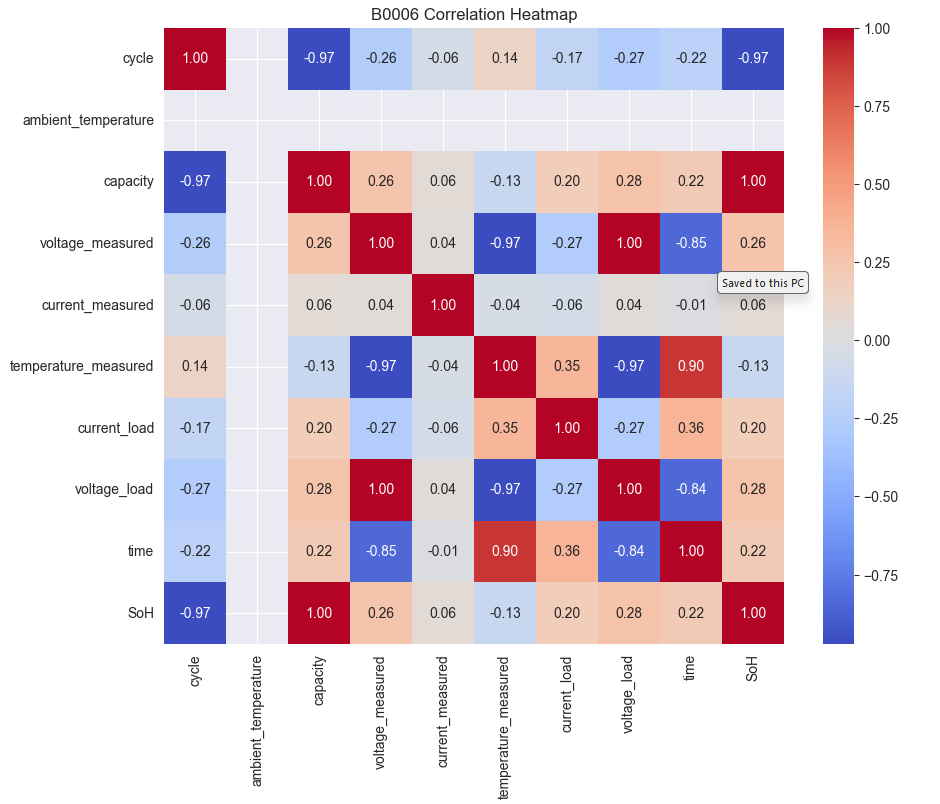
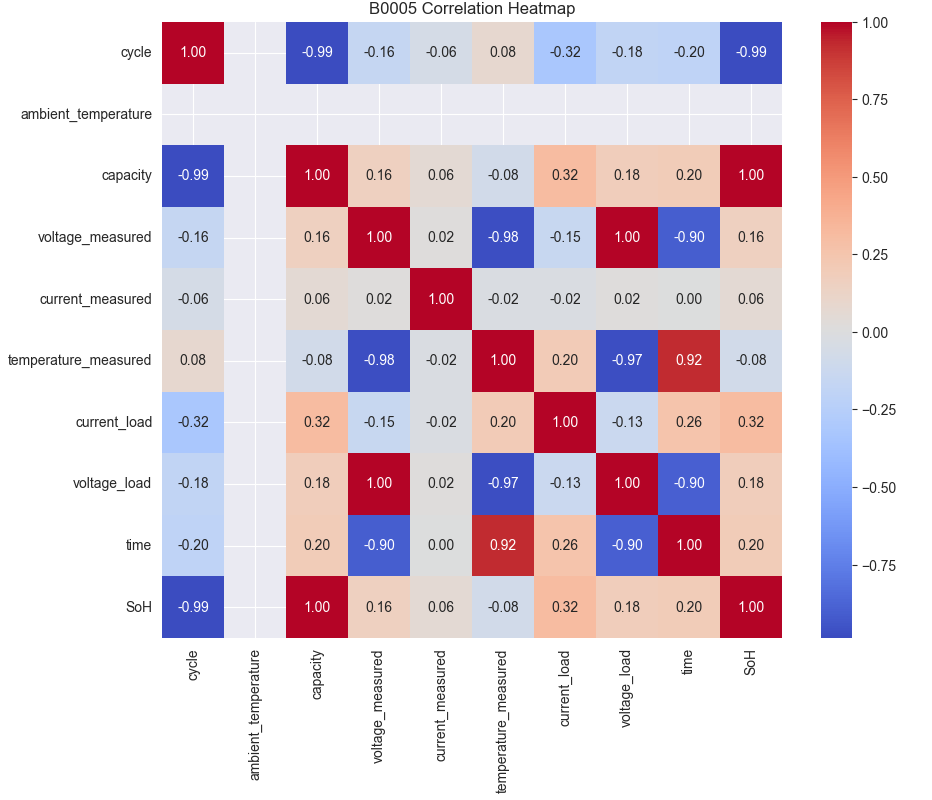


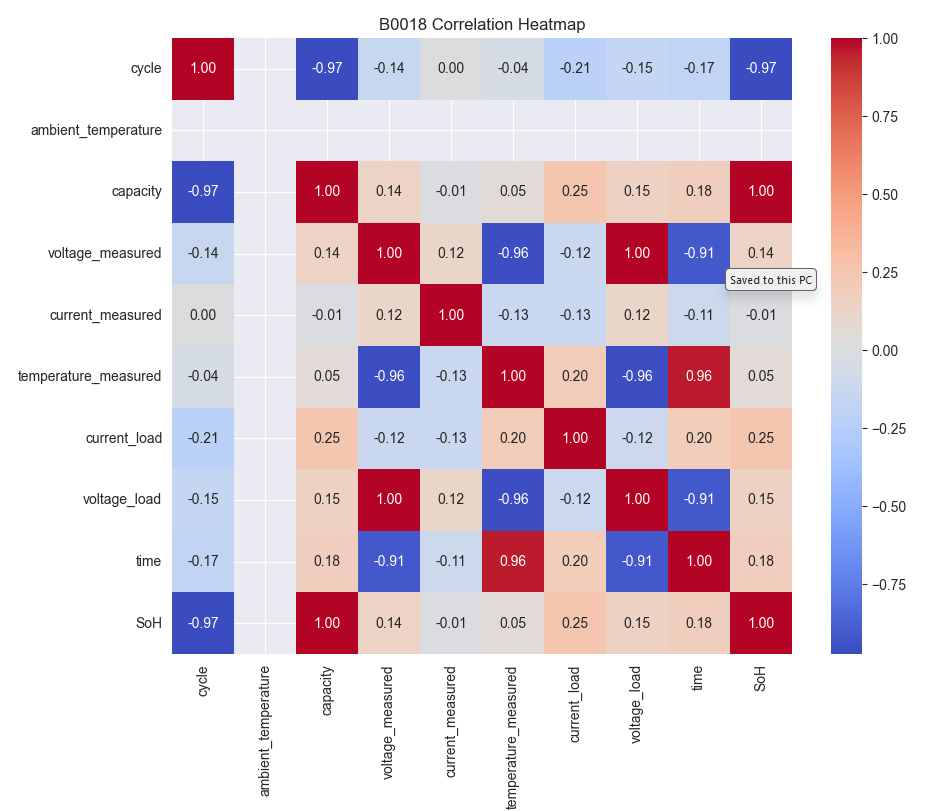
B0018

**3. Keşifsel Veri Analizi (EDA)**

**3.1 Korelasyon Analizi ve Heatmap Oluşturma**

Korelasyon analizi, veri setindeki özellikler arasındaki ilişkileri belirlemeye yardımcı olur. Korelasyon matrisi, bu ilişkilerin derecesini gösterir ve heatmap, bu matrisi görselleştirir. Bu adımda, veri setlerindeki özellikler arasındaki ilişkiler incelenir ve korelasyon matrisi heatmap ile görselleştirilir.



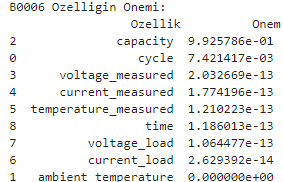
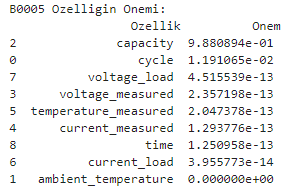


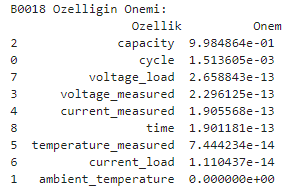
Bu adımda, plot\_heatmap fonksiyonu kullanılarak her bataryanın veri seti için korelasyon matrisi ve heatmap oluşturulur. Korelasyon matrisi, veri setindeki sayısal sütunlar arasındaki korelasyonu hesaplar ve heatmap, bu matrisi görselleştirir. Bu analiz, hangi özelliklerin birbiriyle ilişkili olduğunu ve modelleme sürecinde hangi özelliklerin daha önemli olduğunu belirlemek için kullanılır.

**3.2 Özellik Önemini Belirleme**

Özellik önemi analizi, makine öğrenmesi modellerinin, özellikle karar ağaçları ve ensemble modelleri gibi belirli türlerinin, her bir özelliğin hedef değişken üzerindeki etkisini değerlendirmek için kullanılan bir tekniktir. Bu analiz, hangi özelliklerin model performansı üzerinde en büyük etkiye sahip olduğunu belirlememize yardımcı olur.







Bu adımda, feature\_importance fonksiyonu kullanılarak her bataryanın veri seti için özellik önemleri hesaplanır. Karar ağacı modeli kullanılarak, her bir özelliğin SoH üzerindeki etkisi belirlenir ve bu etkiler önem derecesine göre sıralanır. Bu analiz, hangi özelliklerin SoH tahminlerinde daha önemli olduğunu belirlemek için kullanılır.

Özellik önemini belirleme analizinin sonuçları aşağıdaki gibidir:

**B0005 Bataryası için Özellik Önemi:**

* **capacity**: 89.89%
* **cycle**: 10.10%
* Diğer özellikler çok düşük önemlere sahiptir.

**B0006 Bataryası için Özellik Önemi:**

* **capacity**: 99.20%
* **cycle**: 0.79%
* Diğer özellikler çok düşük önemlere sahiptir.

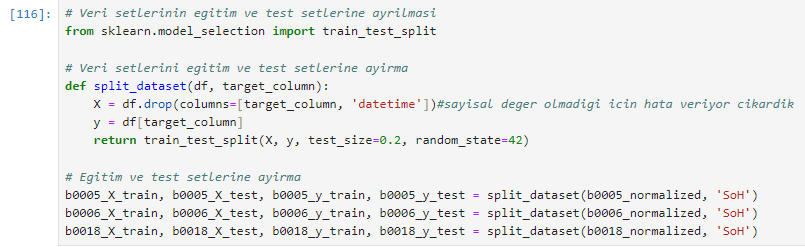
**B0018 Bataryası için Özellik Önemi:**

* **capacity**: 99.67%
* **cycle**: 0.32%
* Diğer özellikler çok düşük önemlere sahiptir.

**4. Model Eğitim ve Değerlendirme**

**4.1 Veri Setlerinin Eğitim ve Test Setlerine Ayrılması**

Veri setleri, eğitim ve test setlerine ayrılır. Bu adımda, veri seti %80 eğitim ve %20 test olacak şekilde bölünür.



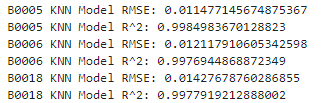
Bu adımda, split\_dataset fonksiyonu kullanılarak her bataryanın veri seti için eğitim ve test setleri oluşturulur.

**4.2 Model Eğitimi ve Değerlendirme**

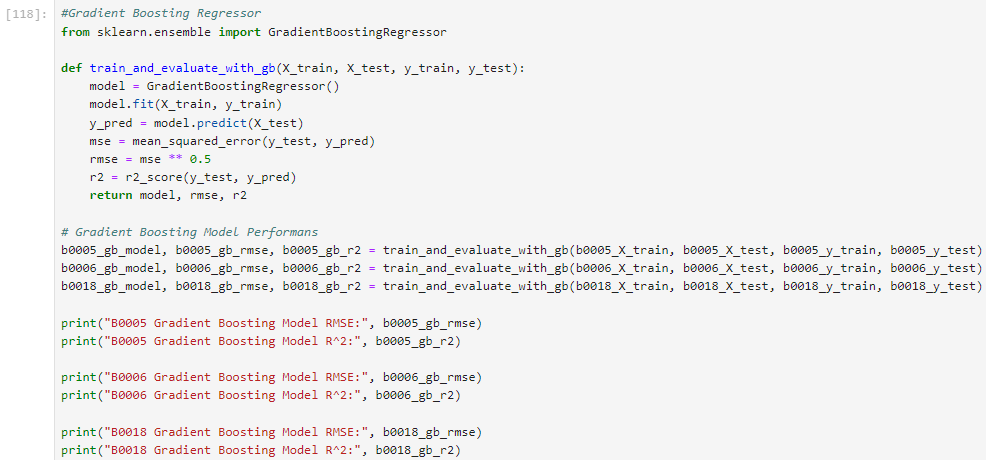
Farklı makine öğrenmesi modelleri kullanılarak SoH tahmin modelleri eğitilir ve performansları değerlendirilir.

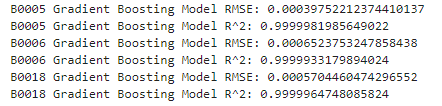
**4.2.1 K-Nearest Neighbors (KNN)**

****

****

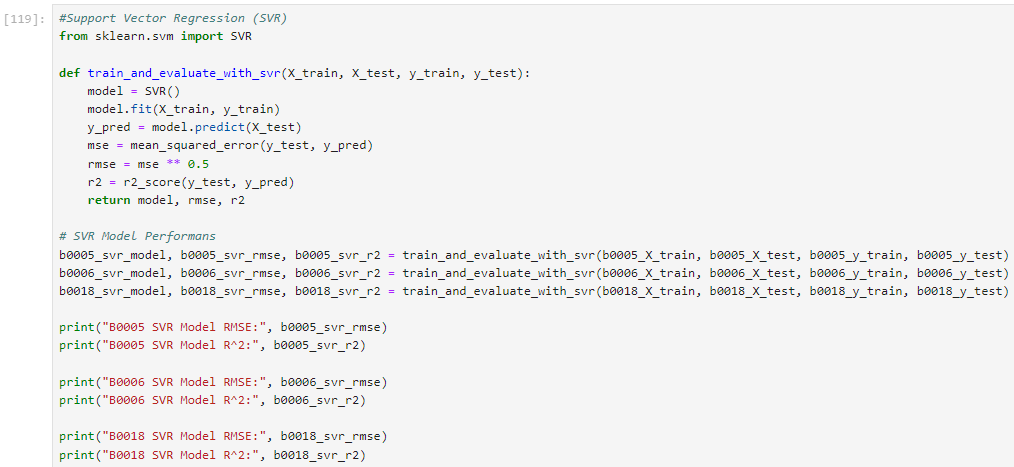
Bu adımda, KNeighborsRegressor modeli kullanılarak KNN modeli eğitilir ve performansı değerlendirilir. Modelin performansı RMSE ve R^2 metrikleri ile ölçülür.

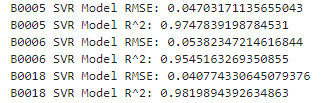
**4.2.2 Gradient Boosting Regressor**

****

Bu adımda, GradientBoostingRegressor modeli kullanılarak Gradient Boosting modeli eğitilir ve performansı değerlendirilir. Modelin performansı RMSE ve R^2 metrikleri ile ölçülür.

**4.2.3 Support Vector Regression (SVR)**

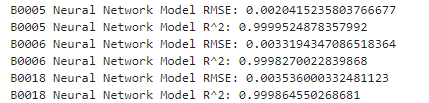
****

****

Bu adımda, SVR modeli kullanılarak Support Vector Regression modeli eğitilir ve performansı değerlendirilir. Modelin performansı RMSE ve R^2 metrikleri ile ölçülür.

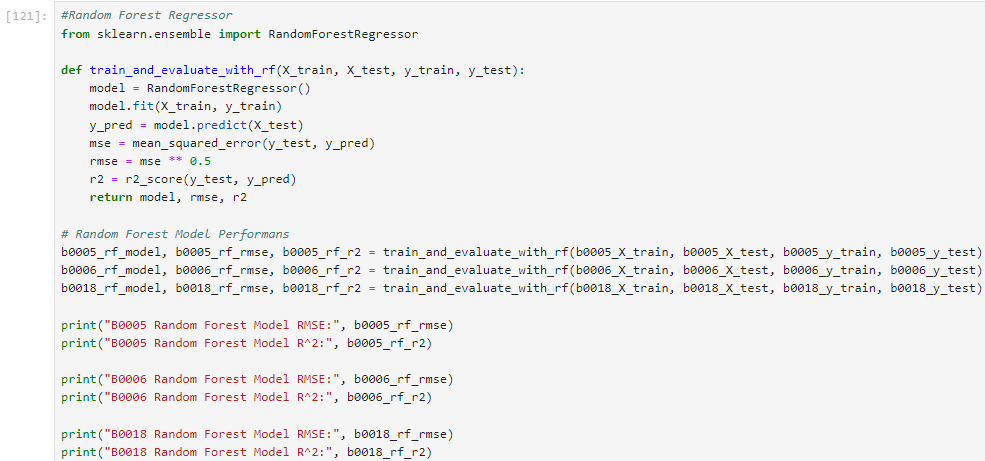
**4.2.4 Neural Network (MLPRegressor)**

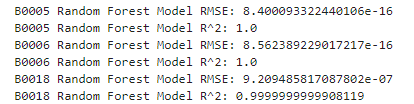




Bu adımda, MLPRegressor modeli kullanılarak Neural Network modeli eğitilir ve performansı değerlendirilir. Modelin performansı RMSE ve R^2 metrikleri ile ölçülür.

**4.2.5 Random Forest Regressor**

****

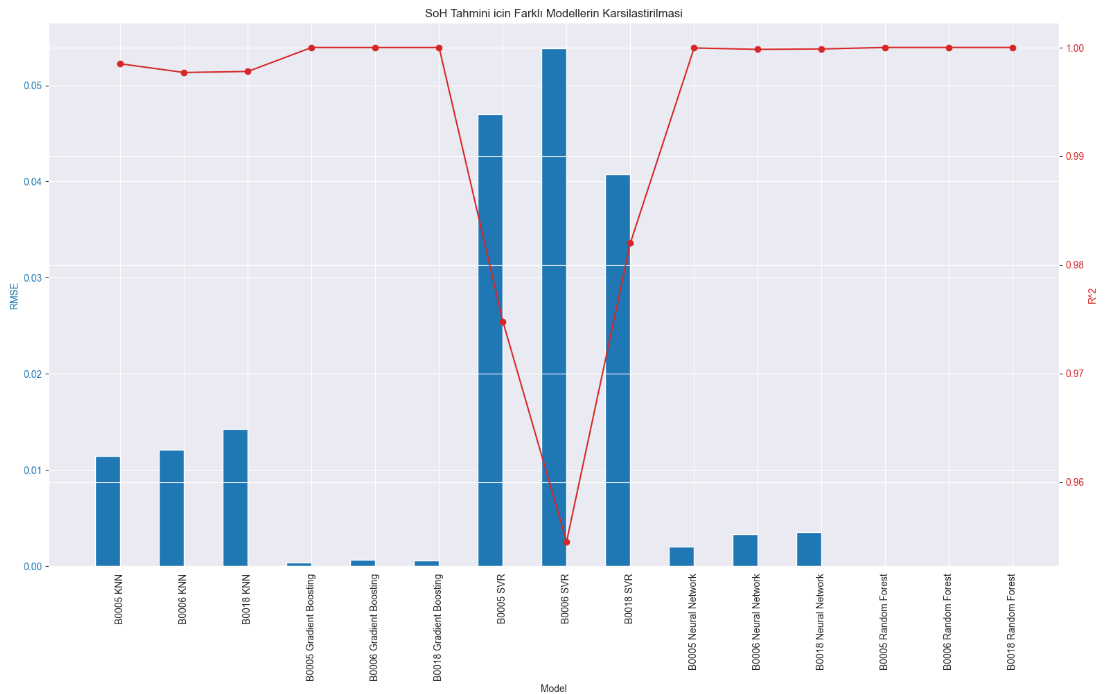
****

Bu adımda, RandomForestRegressor modeli kullanılarak Random Forest modeli eğitilir ve performansı değerlendirilir. Modelin performansı RMSE ve R^2 metrikleri ile ölçülür.

**4.3 Modellerin Karşılaştırılması ve Performans Grafiği**

Farklı modellerin performansları karşılaştırılır ve bir performans grafiği oluşturulur.

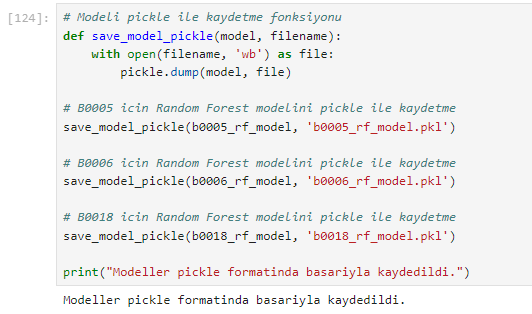




Bu adımda, farklı modellerin performanslarının RMSE ve R^2 değerleriyle karşılaştırıldığı bir grafik oluşturulur. Bu grafik, modellerin tahmin performansını görsel olarak kıyaslamayı sağlar.

**4.4 Modeli Pickle ile Kaydetme**

En iyi performans gösteren Random Forest modeli pickle kullanarak kaydedilmiştir.



**5 Sonuç ve Değerlendirme**

Bu proje kapsamında, üç farklı lithium-ion bataryanın (B0005, B0006, B0018) SoH (State of Health) değerlerini tahmin etmek için çeşitli makine öğrenmesi modelleri kullanılmış ve değerlendirilmiştir. Modellerin performansı, RMSE ve R^2 metrikleri ile ölçülmüş ve en iyi performans gösteren model olarak Random Forest belirlenmiştir.

Random Forest modeli, diğer modellere kıyasla en düşük RMSE ve en yüksek R^2 değerlerine sahip olarak en iyi performansı göstermiştir. Bu modelin üstünlüğü, bataryaların SoH tahmininde güvenilir ve doğru sonuçlar verdiğini göstermektedir.

Bu proje sonucunda elde edilen modeller ve analizler, lithium-ion bataryaların sağlık durumlarını daha doğru bir şekilde tahmin etmeye yardımcı olacak ve batarya yönetim sistemlerinin geliştirilmesine katkıda bulunacaktır.