



Proje Konu Adı	Karma Döngü Enerji Santrali
Öğrenci No Adı ve Soyadı	

1. Proje Tanımı ve Açıklaması

Karma Döngü Enerji Santrali için elektrik enerjisi çıkışı tahmin etmek için Lineer Regresyon, Polinom Regresyon ve *Rprop MLP Learner*(Yapay Sinir Ağları) modellerinin kullanılmıştır. Bu çalışmada kullanılan makine öğrenme yöntemleri akıllı şehirlerde kullanılabilecek bir tesis olup, kullanılan algoritmanın performansları değerlendirilmiştir.

2. Veri Kümesi Bilgileri

Veri seti, santral tam yükte çalışmaya ayarlandığında 6 yıl boyunca (2006-2011) bir Kombine Çevrim Santralinden toplanan 9568 veri noktasını içerir. Özellikler, tesisin net saatlik elektrik enerjisi çıkışı (PE) tahmin etmek için saatlik ortalama ortam değişkenlerinden Sıcaklık (T), Ortam Basıncı (AP), Bağıl Nem (RH) ve Egzoz Vakumundan (V) oluşur. Elektrik, tek çevrimde birleştirilen gaz ve buhar türbinleri tarafından üretilir. Üç ortam değişkeni gaz türbininin performansını etkiler ve egzoz vakumu buhar türbininin performansını etkiler. Veriler, her saniye ortam değişkenlerini kaydeden tesis çevresinde bulunan çeşitli sensörlerden alınan saatlik ortalamalardan oluşur.

3. Özellik Bilgileri (Statistic ile bulunmuştur.)

İsim	Tabloda Gösterimi	Değer Aralığı (Değişkenler normalizasyon yapılmadan verilmiştir.)
Temperature (Sıcaklık)	(AT)	1.81°C ile 37.11°C aralığında sıcaklık
Ambient Pressure (AP) (Ortam Basıncı)	(AP)	992.89-1033.30 milibar aralığında Ortam Basıncı
Relative Humidity (RH) (Bağıl Nem)	(RH)	%25,56 ile %100,16 aralığında Bağıl Nem
Exhaust Vacuum (V) (Egzoz Vakum)	(V)	Egzoz Vakum 25,36-81,56 cm Hg aralığında
Net hourly electrical energy output (PE) (ÇIKTI)	(PE)	Net saatlik elektrik enerjisi çıkışı 420,26-495,76 MW

Özellikler, saatlik ortalama ortam değişkenlerinden oluşur Ortalamalar, ortam değişkenlerini her saniye kaydeden tesis çevresinde bulunan çeşitli sensörlerden alınır. Değişkenler normalizasyon yapılmadan verilmiştir.

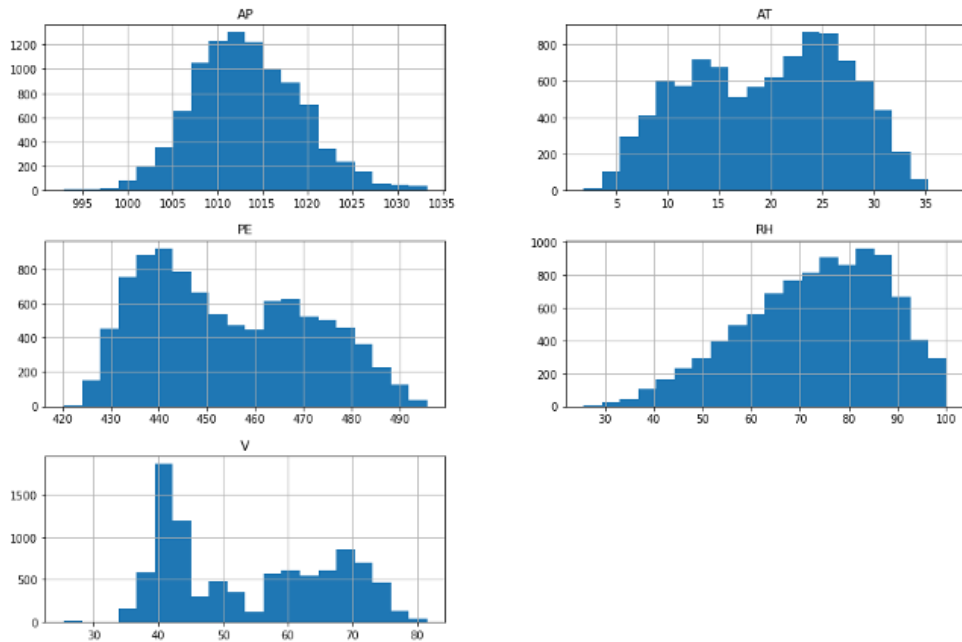
Proje Aşamaları

- A) Veri Keşfi
- B) Modelleme

A. Veri Keşfi

Verimizde eksik değer yok (Missing Value) ve tümü sayısal niteliklerdir. Hemen verimizi keşfetmeye başlayalım. Öncelikle Histogram grafiğini kullanalım.

a) Histogram

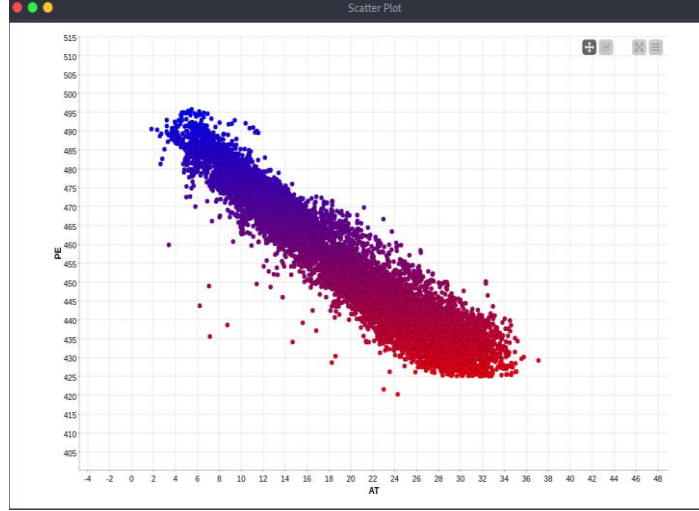


Gözlemler:

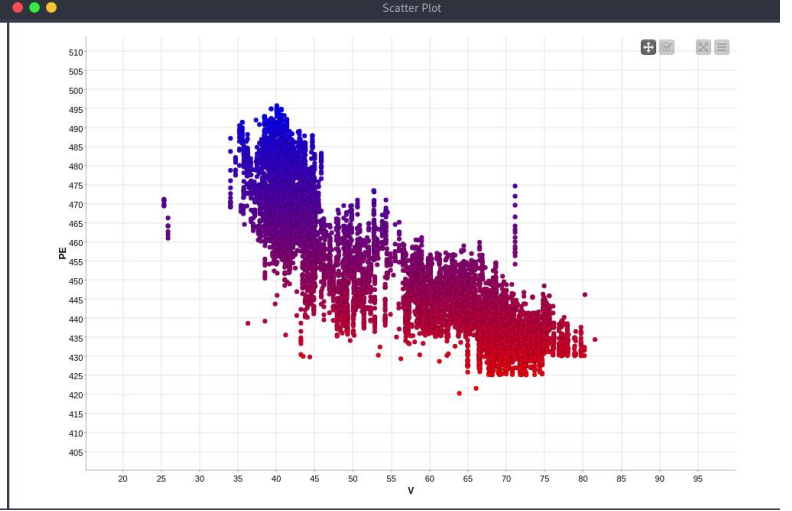
- Sıcaklık ve egzoz vakum özelliklerinin yanıt değişkeni ile önemli bir doğrusal ilişkiye sahip olduğu dağılım grafiklerinden çıkarabiliriz.
- Ortam Basıncı (AP) Normal olarak dağıtılır .
- Bağıl Nem (RH) sağa Eğik .
- Nitelik ölçekleri farklıdır.

b) Girişler ve çıkış PE arasındaki Doğrusal Korelasyon

AT (Ortam Basıncı)



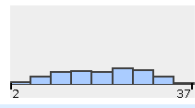
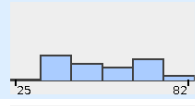
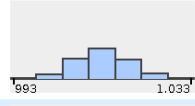
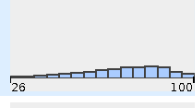
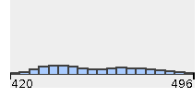
V(Egzoz Vakum)



Gözlemler:

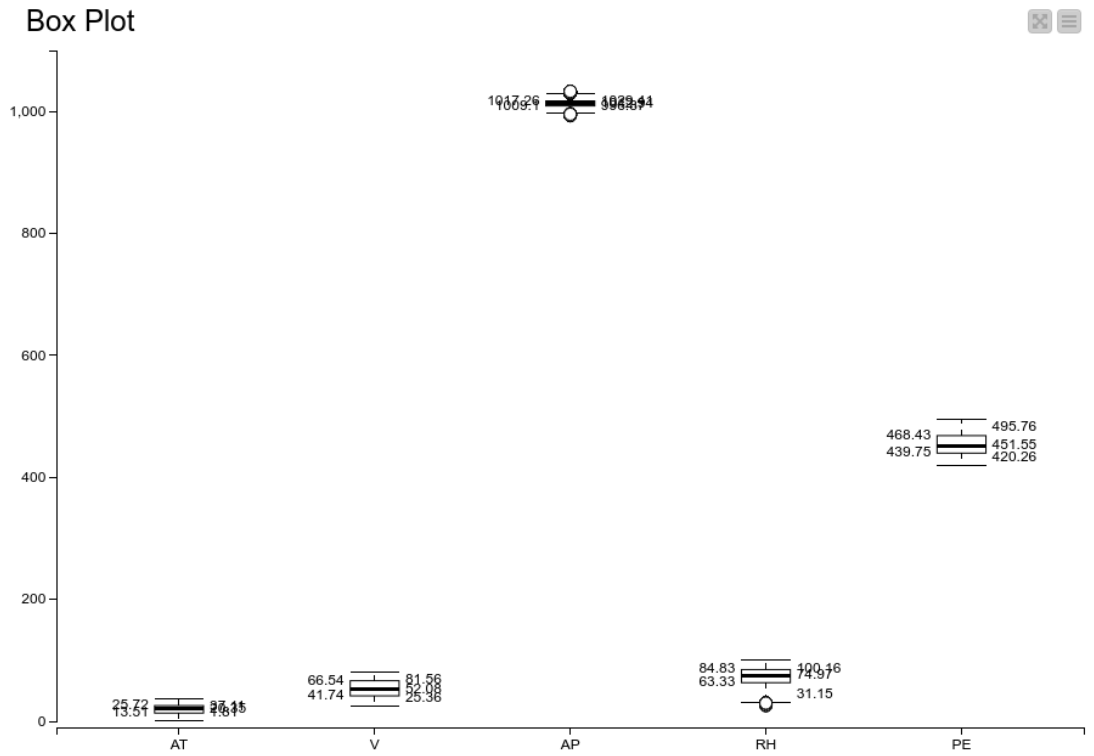
- Grafiği incelediğimizde ortam basıncının önemli bir özellik olduğu sonucuna varılabilir.
- Ortam basıncı ve egzoz vakum tahminleri ile yanıt arasında istatistiksel olarak anlamlı bir ilişki olduğunu göstermektedir.

c) Statistic

Column	Min	Mean	Median	Max	Std. Dev.	Skewness	Kurtosis	No. Missing	No. +∞	No. -∞	Histogram
AT	1.81	19.6512	?	37.11	7.4525	-0.1364	-1.0375	0	0	0	
V	25.36	54.3058	?	81.56	12.7079	0.1985	-1.4443	0	0	0	
AP	992.89	1.013.2591	?	1.033.3	5.9388	0.2654	0.0942	0	0	0	
RH	25.56	73.309	?	100.16	14.8003	-0.4318	-0.4445	0	0	0	
PE	420.26	454.365	?	495.76	17.067	0.3065	-1.0485	0	0	0	

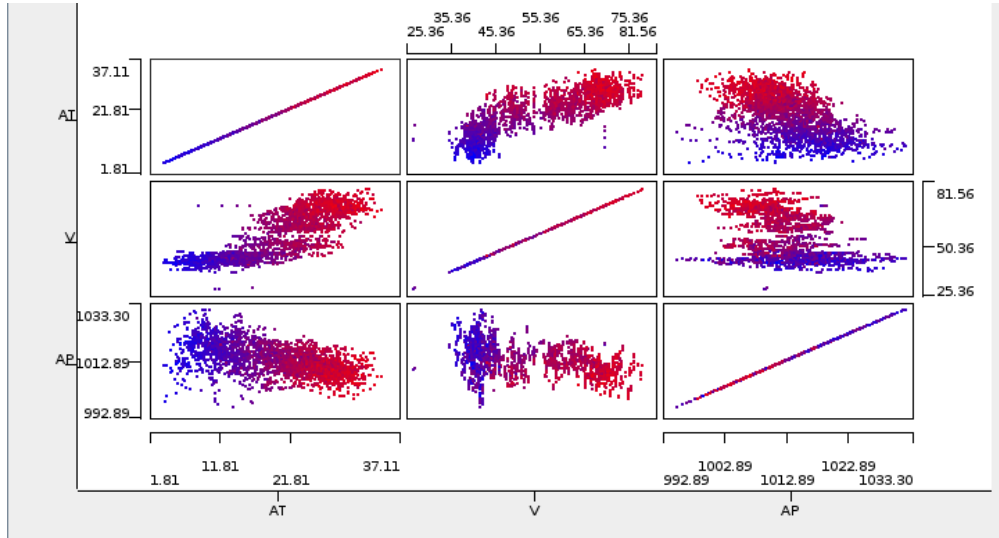
Sonuçlar:

- Verimizin hangi aralıkta ne kadar olduğunu görebiliriz.
- 1.81°C ile 37.11°C aralığında sıcaklık (T)
- 992.89-1033.30 milibar aralığında Ortam Basıncı (AP) Normal dağılmıştır.
- %25,56 ila %100,16 aralığında Bağıl Nem (RH)
- Egzoz Vakum (V) 25,36-81,56 cm Hg aralığında
- Net saatlik elektrik enerjisi çıkışı (EP) 420,495

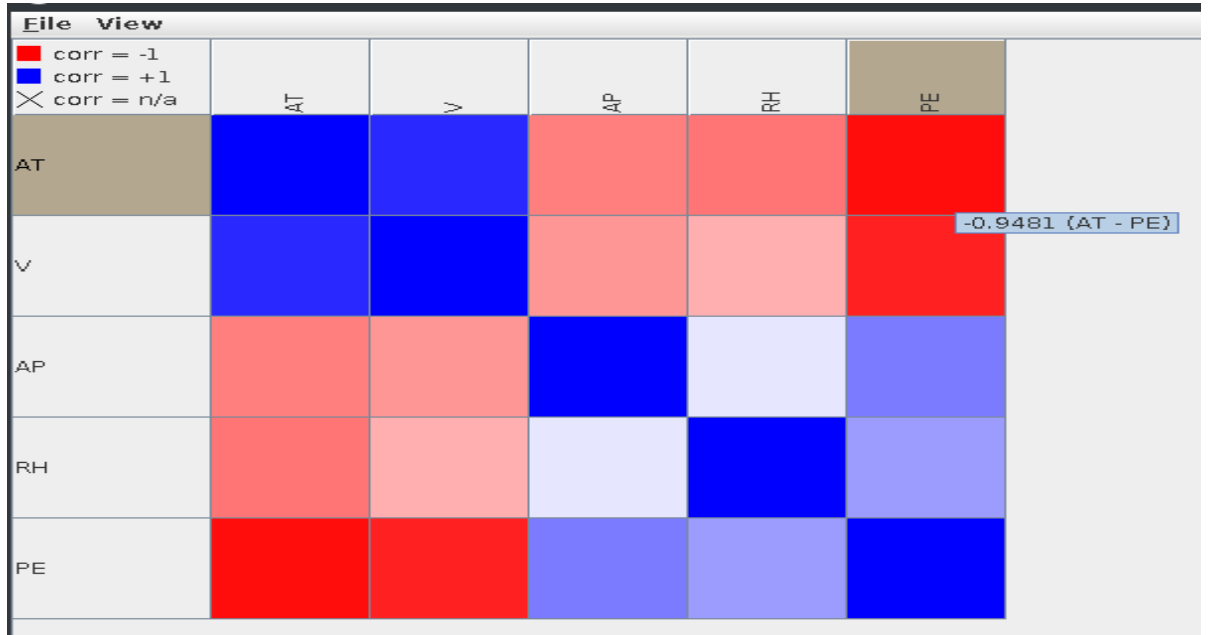
d) Box Plot**Gözlemler:**

- AP ve RH da uç noktalar bulunmaktadır fakat uç nokta ile arasında büyük mesafe olmadığı için aşırı değer yoktur.

E) Scatter Matrix



f) Korelasyon Matrisi



Gözlemler:

- Grafik, Ortalama Sıcaklık tahmincisi ile yanıt arasında istatistiksel olarak anlamlı bir ilişki olduğunu göstermektedir.
- Ortalama Sıcaklık (AT), Net saatlik elektrik enerjisi çıktısı yani PE ile güçlü -ve korelasyona sahiptir ve -0.948'e eşittir.
- Egzoz Vakum(V) ayrıca PE ile korelasyona sahiptir ve -0.869'a eşittir.

B) Modelleme

a) Linear Regression

Linear Regression Result View - 3:9 - Linear Regression Learner (ogrenme dugumu)				
File				
Statistics on Linear Regression				
Variable	Coeff.	Std. Err.	t-value	P> t
AT	-0.9312	0.0084	-110.3356	0.0
V	-0.1717	0.0064	-26.7616	0.0
AP	0.0381	0.006	6.3666	2.06E-10
RH	-0.1553	0.0049	-31.6564	0.0
Intercept	1.0901	0.0069	158.2296	0.0
R-Squared: 0.931				
Adjusted R-Squared: 0.931				

Gözlemler:

- Net saatlik elektrik enerjisi çıkışını (PE) olumlu etkileyen ortam basıncı (AP) oldu.
- R-Squared: 0.931
- Adjusted R-Squared: 0.93
- R-Squared: 0.931 değeri 1 e yakın olduğu için modelimiz başarılı bir şekilde öğrenmiştir diyebiliriz.

Eğitim ve test karşılaştırma:

R ² :	0.931
Mean absolute error:	0.047
Mean squared error:	0.004
Root mean squared error:	0.06
Mean signed difference:	-0
Mean absolute percentage error:	NaN
Adjusted R ² :	0.931

R ² :	0.923
Mean absolute error:	0.049
Mean squared error:	0.004
Root mean squared error:	0.062
Mean signed difference:	-0.003
Mean absolute percentage error:	0.154
Adjusted R ² :	0.923

Sonuç:

- Her iki sonuç birbirine yakın.
- Model ne AŞIRI UYUMLU ne de UYUMLU DEĞİLDİR.



b) Polinom Regresyon

File

Statistics on Polynomial Regression

Variable	Coeff.	Std. Err.	t-value	P> t
AT	-1.3202	0.0191	-68.9956	0.0
V	-0.29	0.0253	-11.4647	0.0
AP	0.2968	0.0248	11.9757	0.0
RH	-0.0498	0.0201	-2.4779	0.0132
AT^2	0.4696	0.0192	24.4627	0.0
V^2	0.0644	0.0219	2.9438	0.0033
AP^2	-0.2304	0.0232	-9.9476	0.0
RH^2	-0.0516	0.016	-3.2265	0.0013
Intercept	1.075	0.0102	105.542	0.0

R-Squared: 0.9386

Adjusted R-Squared: 0.9385

Eğitim ve test karşılaştırma:

R²: 0.939
Mean absolute error: 0.045
Mean squared error: 0.003
Root mean squared error: 0.056
Mean signed difference: -0
Mean absolute percentage error: NaN
Adjusted R²: 0.939

R²: 0.932
Mean absolute error: 0.046
Mean squared error: 0.003
Root mean squared error: 0.058
Mean signed difference: -0.003
Mean absolute percentage error: 0.144
Adjusted R²: 0.932

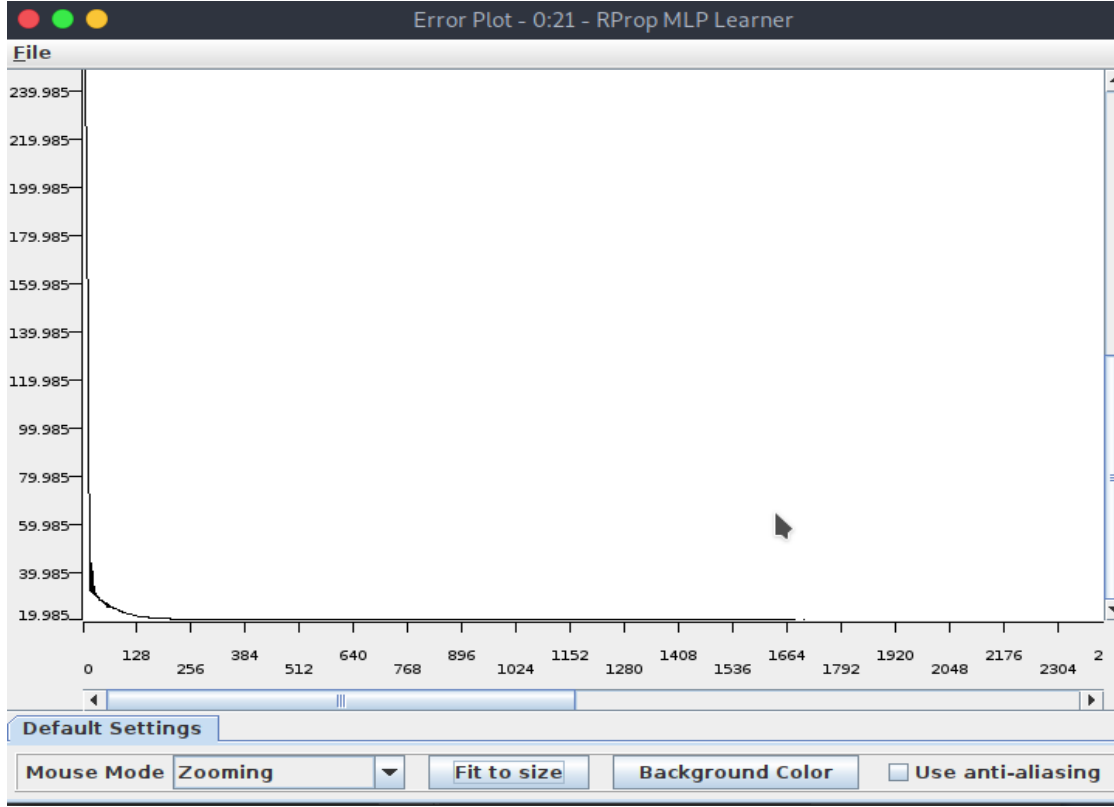
Sonuc:

- Modelimiz eğitimde ve test linear regresyona göre daha başarılı bir şekilde öğrenmiştir.
- Karmaşıklığı artırdığımız zaman modelimiz ezberleme yapmamaktadır.
- Karmaşıklığı artırdığımız zaman modelimiz daha iyi öğrenmeye başlar ve bu yüzden Polinom regresyonda istediğimiz başarıyı sağlamaya biraz daha yaklaştık.

c) Rprop MLP Learner

Kullanılan itterasyon sayısı: 5,000.

0 a ne kadar yakın ise model o kadar iyi öğrenmiştir.

**Gözlem:**

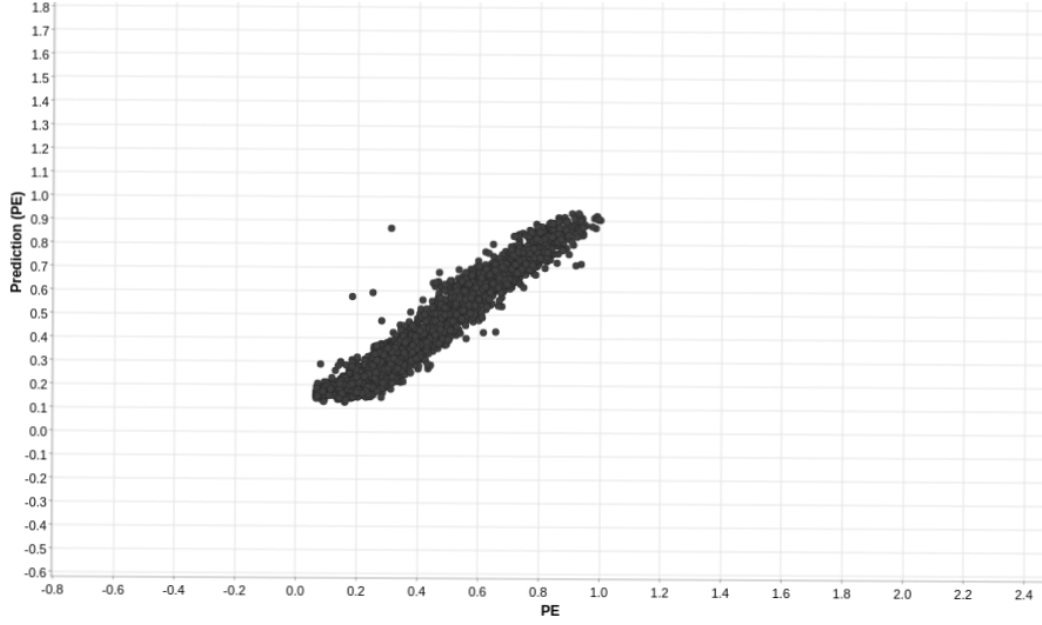
- İtterasyon değerini yükseltip tekrar inceleyelim.

Eğitim ve test karşılaştırması:

R ² :	0.941
Mean absolute error:	0.043
Mean squared error:	0.003
Root mean squared error:	0.055
Mean signed difference:	-0
Mean absolute percentage error:	0.142
Adjusted R ² :	0.941

R ² :	0.941
Mean absolute error:	0.042
Mean squared error:	0.003
Root mean squared error:	0.055
Mean signed difference:	-0
Mean absolute percentage error:	NaN
Adjusted R ² :	0.941

GRAFİK

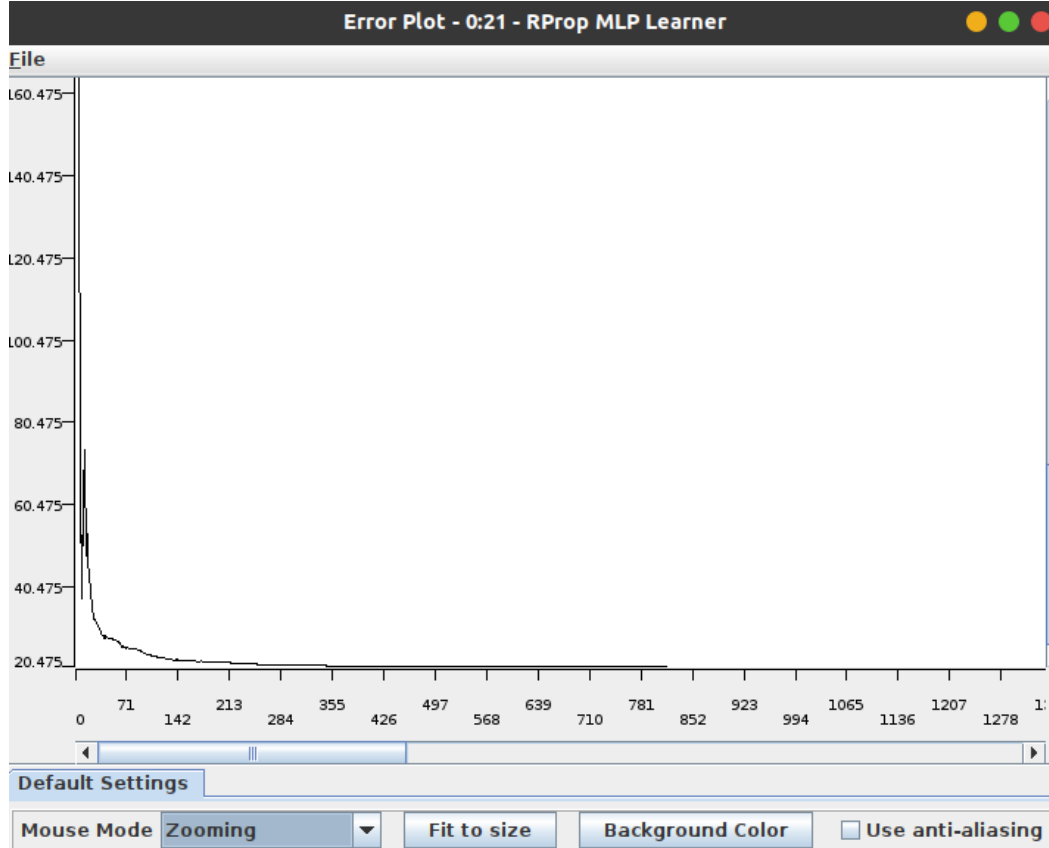


Sonuç:

- *Hata kareler 0 a yakın olduğu için modelimiz başarılı bir şekilde öğreniyor diyebiliriz.*
- *Şu ana kadar en yüksek test ve eğitim başarısını yakaladık.*
- *Projemize yapay sinir ağı ile ilerlersek daha iyi başarı sağlayabiliriz.*

Deneme 2

Kullanılan iterasyon sayısı: 6,000.



Gözlem:

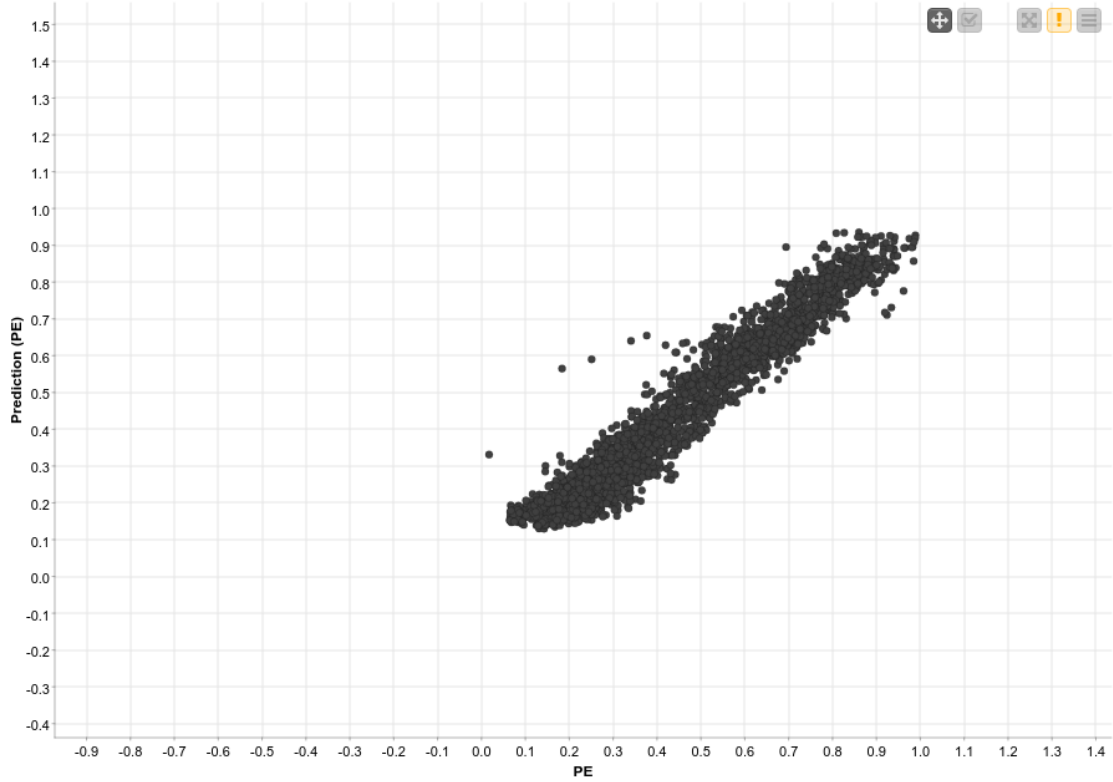
- İterasyon değerini yükselttiğimiz zaman modelimiz daha başarılı öğrenmiştir.

Eğitim ve test karşılaştırması:

R ² :	0.941
Mean absolute error:	0.042
Mean squared error:	0.003
Root mean squared error:	0.055
Mean signed difference:	-0
Mean absolute percentage error:	NaN
Adjusted R ² :	0.941

R ² :	0.941
Mean absolute error:	0.043
Mean squared error:	0.003
Root mean squared error:	0.055
Mean signed difference:	-0.003
Mean absolute percentage error:	0.142
Adjusted R ² :	0.941

GRAFİK



Scatter Plot grafiğinde gerçek değerler ile tahmin değerleri bir grafiğe aktardığımız zaman modelimiz başarılı bir şekilde çalıştığını gözlemleyebiliriz. (Noktalar birbirlerine ne kadar yakınsa model o kadar başarılıdır.)



Genel Sonuç:

Tüm modelleme tekniklerini kullandıktan sonra modelimiz en iyi Yapay Sinir Ağında (**Rprop MLP Learner**) başarılı olduğunu söyleyebiliriz. Bu yüzden Yapay Sinir Ağı ile devam etmeliyiz.

Yapay Sinir Ağı yöntemiyle temsil edilen akıllı şehirlerde kullanılabilir. Yapay Sinir Ağı yönteminin performansı değerlendirildi ve ortalama mutlak hata (0.042) ve bir kök ile en iyi sonuçları verdi. (0.055) ortalama kare hatası.

Referanslar

1. <https://archive.ics.uci.edu/ml/datasets/Combined+Cycle+Power+Plant?ref=datanews.io>
2. Pınar Tüfekci, Çorlu Faculty of Engineering, Namık Kemal University, TR-59860 Çorlu, Tekirdağ, Turkey
Email: ptufekci '@' nku.edu.tr
3. Heysem Kaya, Department of Computer Engineering, Boğaziçi University, TR-34342, Beşiktaş, İstanbul, Turkey
Email: heysem '@' boun.edu.tr

Projenin tamamı github hesabımda yer almaktadır:
<https://github.com/alii76tt>