





ENERJİSA

# ENERJİSA ENERJİ VERİ MARATONU

FİNAL SUNUMU

Anıl Öztürk



# PROBLEM

**Başkent bölgesinde yer alan lisanssız güneş enerjisi santrallerinin 01.12.2021 – 31.12.2021 arasındaki üretim toplamları saatlik bazda ne olacaktır?**

- Metrik: RMSE
- Dönemlere ait hava durumu bilgileri verilmiş
- Test dönemine ait paylaşılan hava durumu verisi tahmini

# LİTERATÜR

- <https://www.solar.com/learn/solar-panel-production-factors/>
- <https://www.sciencedirect.com/science/article/abs/pii/S1364032117311620>
- <https://www.mdpi.com/2071-1050/12/6/2247/pdf>
- <https://www.emerald.com/insight/content/doi/10.1016/j.aci.2019.11.002/full/html>

## BULGULAR

### 1. PROBLEME DAİR EN ETKİLİ FAKTÖRLER

- Panellerin eğimi
- Panellerin konumu
- İklim ve hava durumu
- Enlem
- Güneş radyasyonu (Solar irradiance)

### 2. ÇÖZÜMLERDE KULLANILAN ÖZNİTELİKLER

- Anlık hava durumu bilgileri (bulut oranı, rüzgar hızı, sıcaklık, nem)
- Panellerin anlık eğimi ve rotasyonu (Hedefe kitlenmeli olanlar için)
- Günün saati

### 3. TAHMİN MODELİ MİMARİLERİ

- Gelişmiş karar ağacı algoritmaları
- Ardışıl yapay sinir ağları (RNN, GRU, LSTM)
- Zaman serisi özel tahmin modelleri (ARIMA, Prophet vb.)



# ÇÖZÜM AŞAMALARI

## VERİ ANALİZİ VE ZENGİNLEŞTİRME

Verilerin birleştirilmesi, ek verilerin dahil edilmesi.  
Verilerin hedefle ilişkisinin analizi.

## ÖZELLİK ÇIKARIMI

Yeni sentez verilerin oluşturulması

## ÇAPRAZ DOĞRULAMA

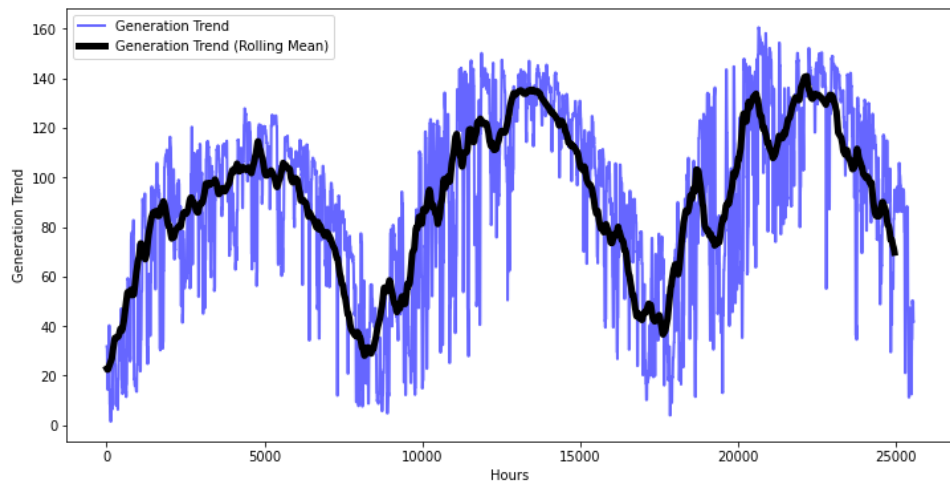
Modelin güvenilirliğinin test edilebilmesi adına  
değişken eğitim-test setleri üzerindeki  
performanslarının değerlendirilme stratejisine  
karar verilmesi

## TAHMİN MODELİ SEÇİMİ

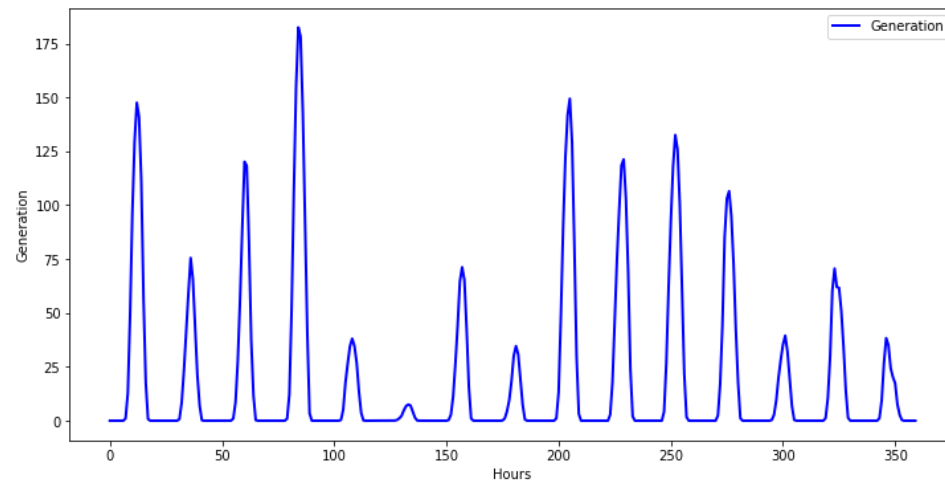
Model mimarisine, hyper-parameter setine ve  
eğitim sürecine karar verilmesi

# VERİ ANALİZİ

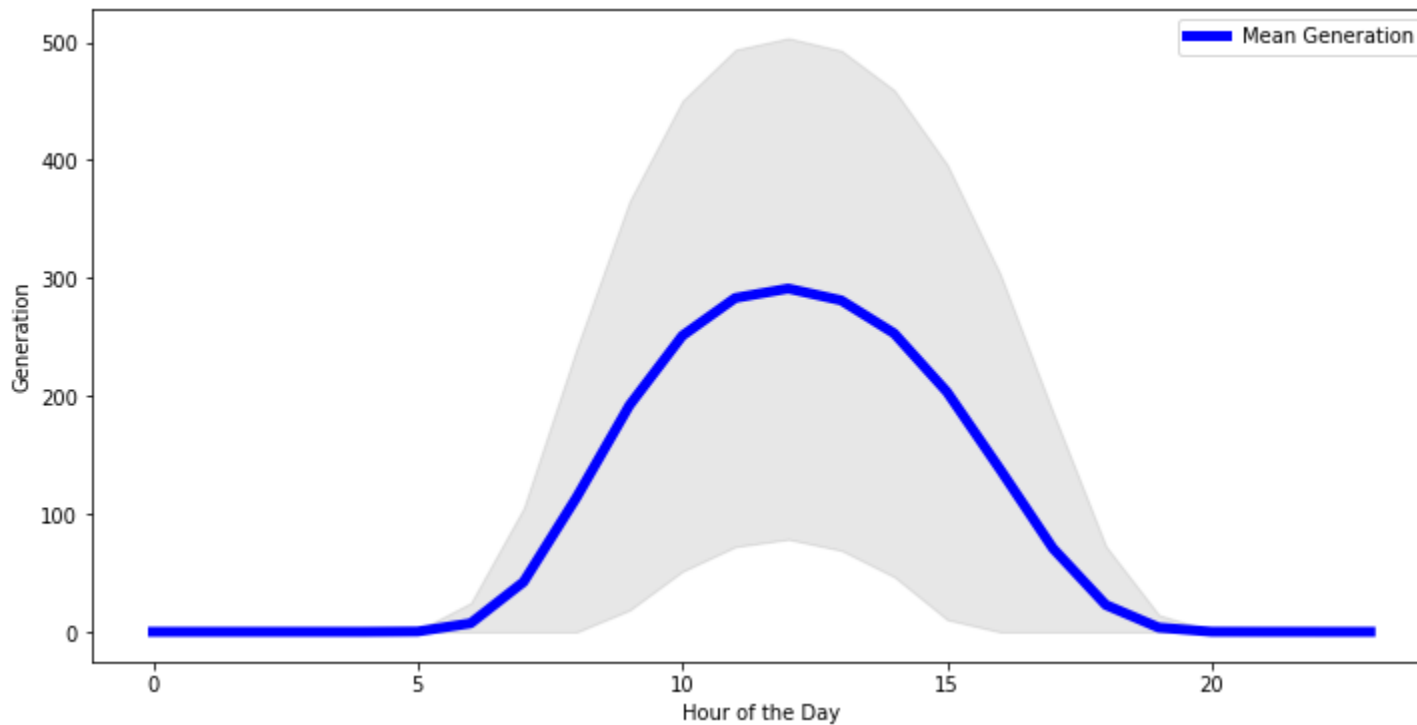
Yıl Ölçeği



Gün Ölçeği



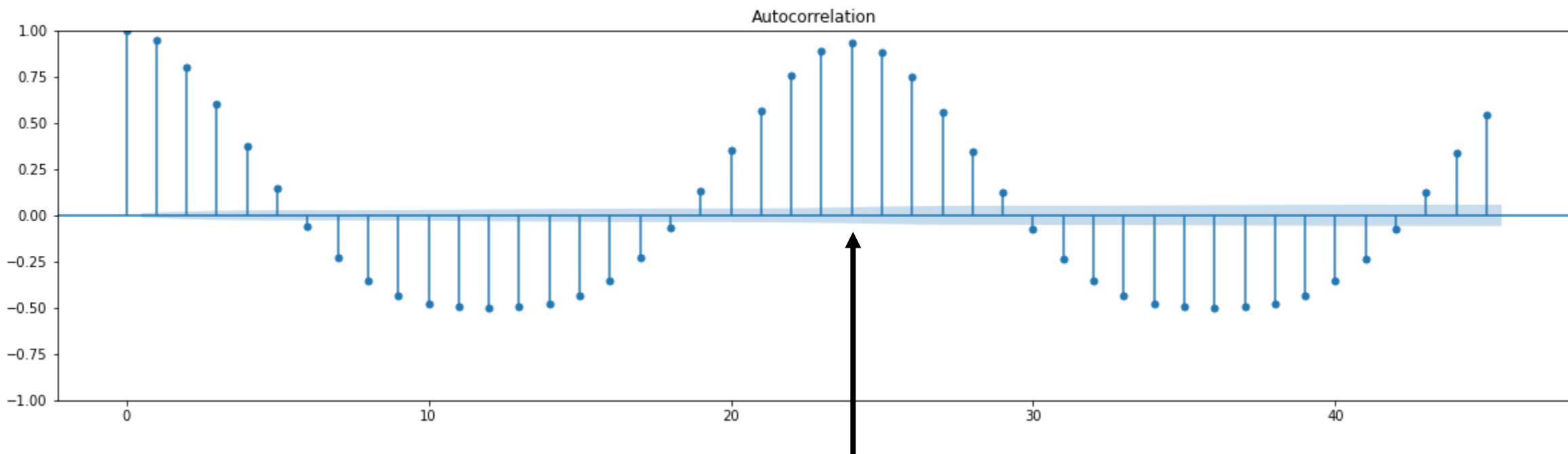
# VERİ ANALİZİ



Günün saatlerine göre ortalama üretim miktarı

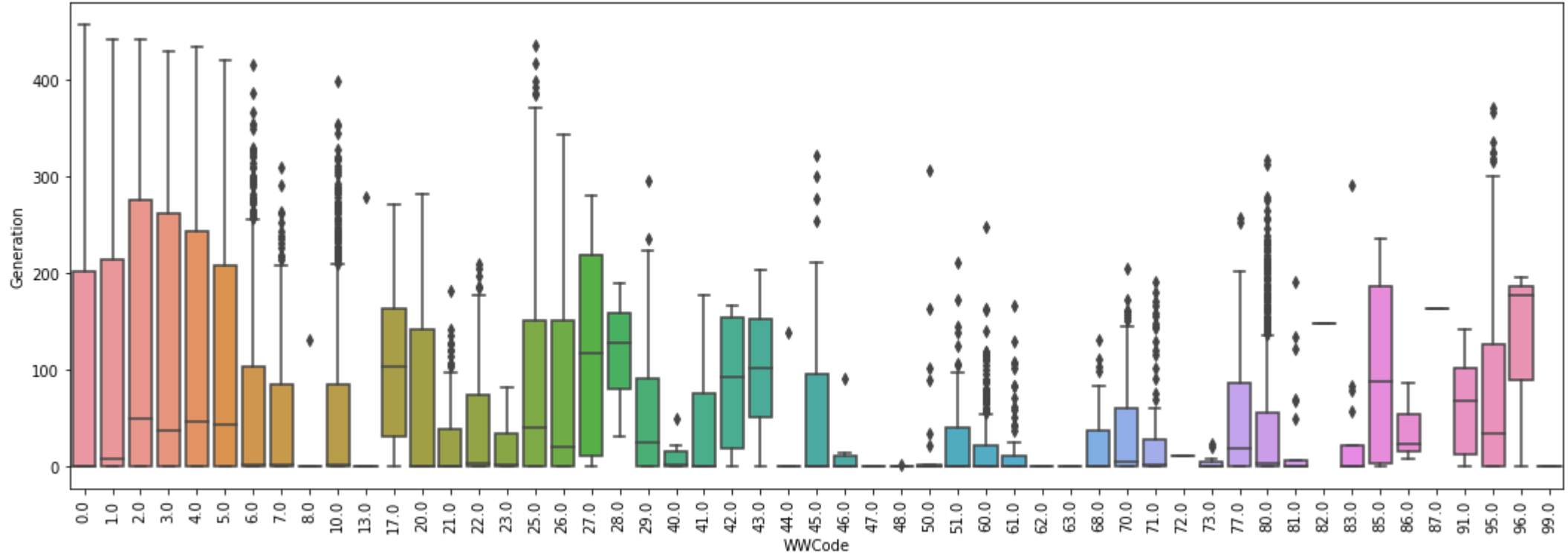
# VERİ ANALİZİ

Üretim miktarının saat bazında otokorelasyon analizi



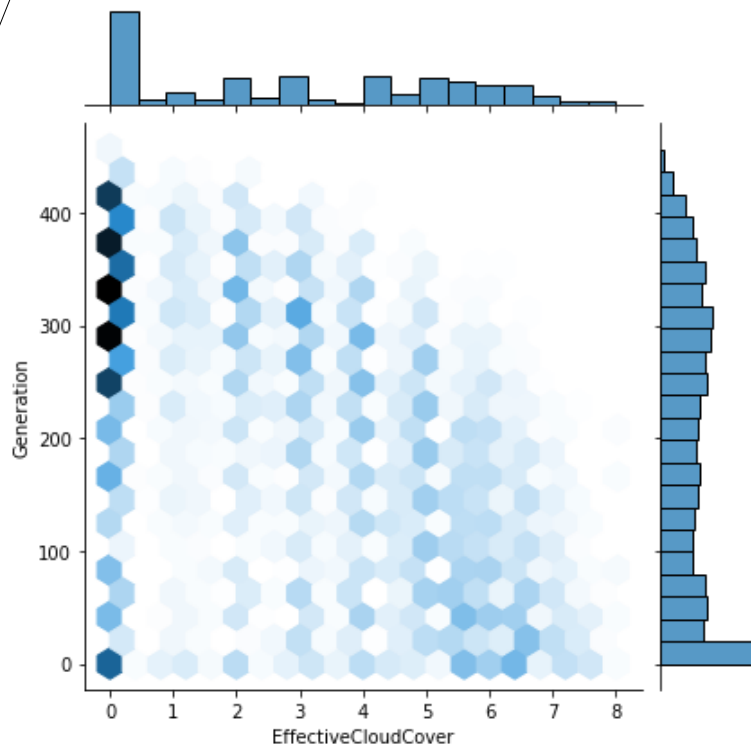


# VERİ ANALİZİ

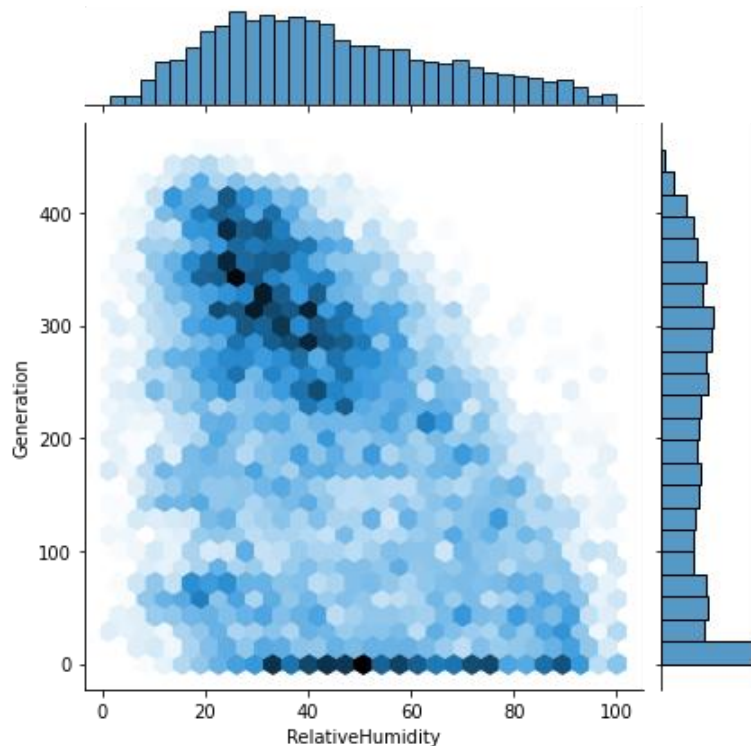


Hava durumu sınıflandırmasına göre ortalama üretim miktarları

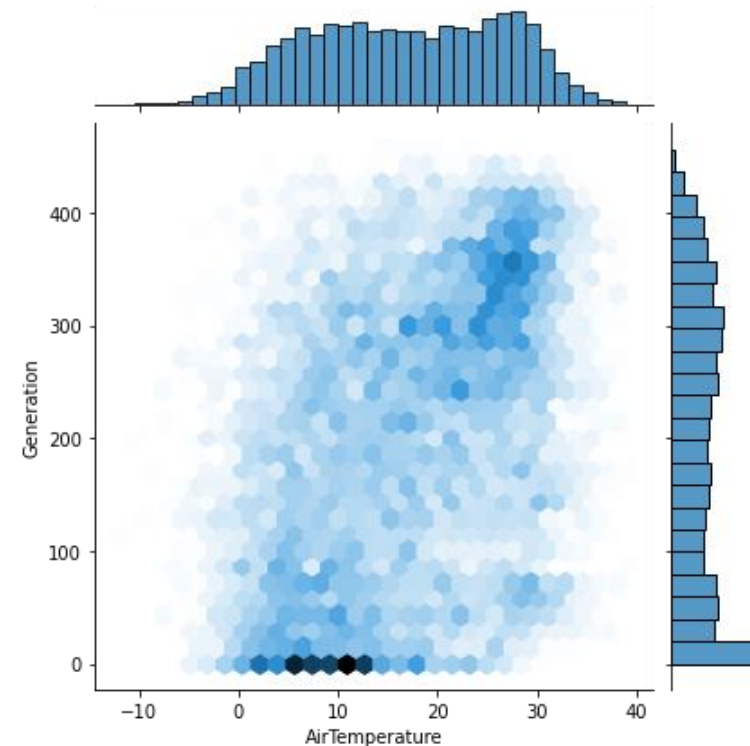
# VERİ ANALİZİ



Bulut Oranı – Üretim ilişkisi

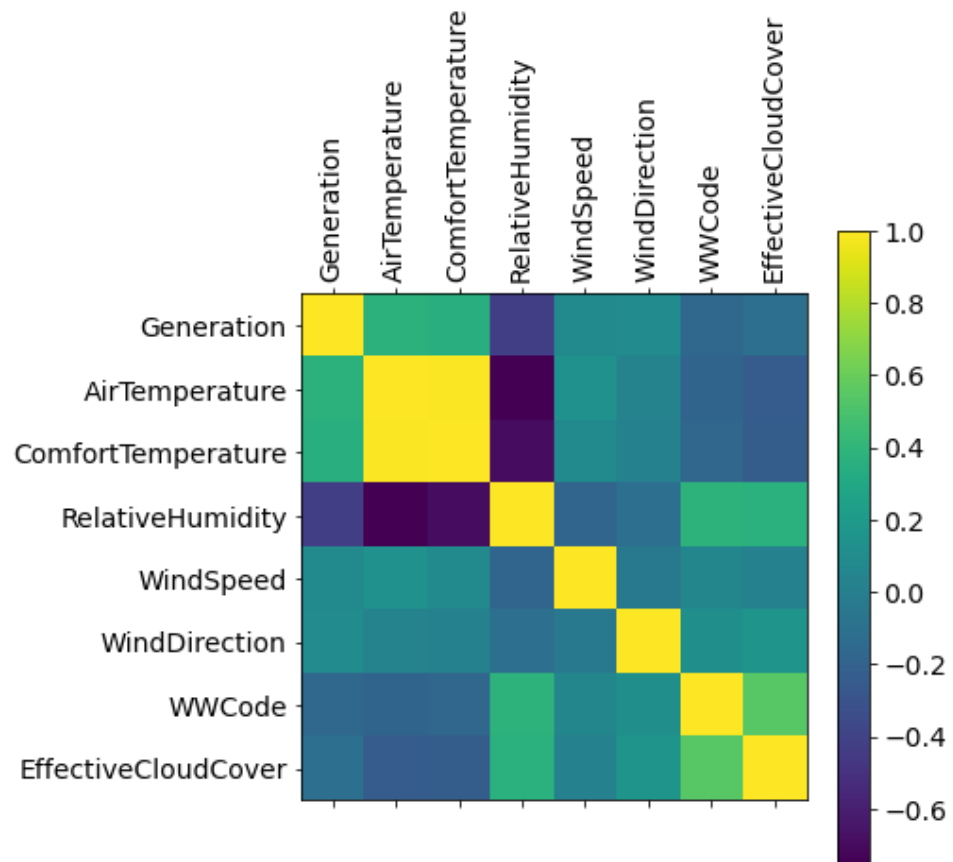


Bağıl Nem – Üretim ilişkisi



Hava Sıcaklığı – Üretim ilişkisi

# VERİ ANALİZİ



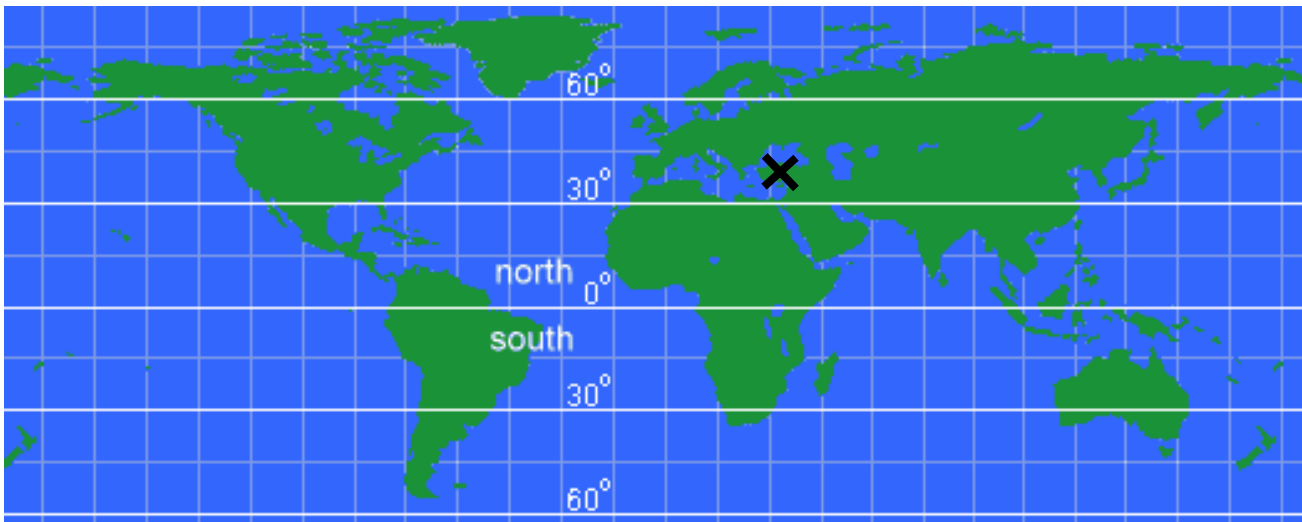
Hava durumu özniteliklerinin ve üretim miktarının korelasyon analizi

# VERİ ZENGİNLEŞTİRME

## Solar Enerji Parametreleri

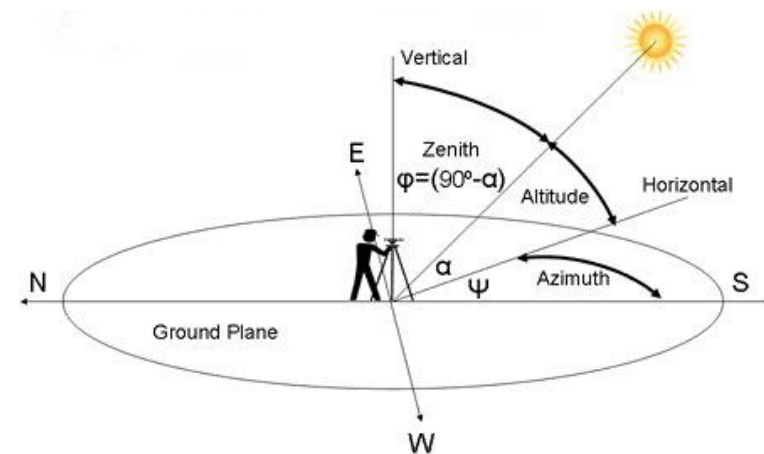
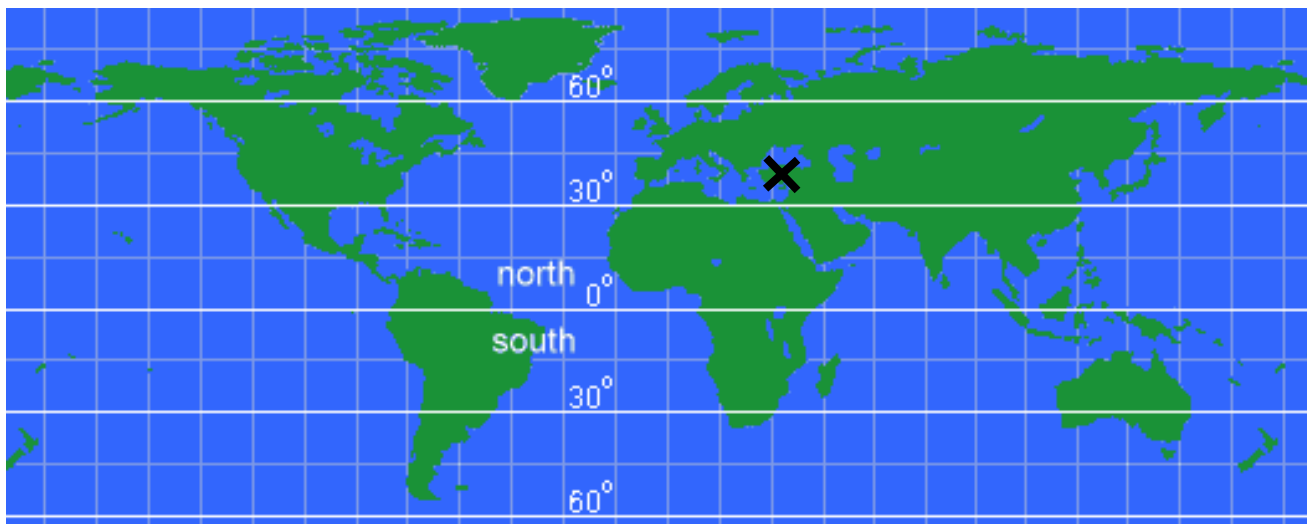
- ~~Panellerin eğimi~~
  - ~~Panellerin konumu~~
  - ~~Güneş radyasyonu (Solar Irradiance)~~
  - İklim ve hava durumu
  - Enlem
  - Günün saati
- Tekil panel bilgisi elimizde yok
  - Tekil panel bilgisi elimizde yok
  - Ücretsiz bir şekilde sunulmuyor. Ayrıca tahmin veriseti de yok.
  - **EnerjiSA tarafından sağlanan bir bilgi.**
  - **Biliniyor**
  - **Biliniyor**

# VERİ ZENGİNLEŐTİRME

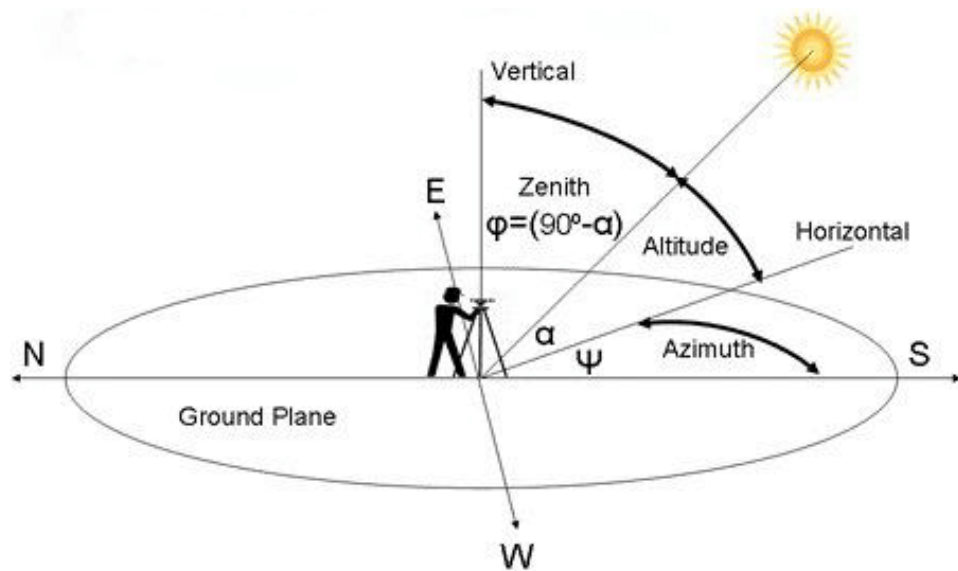


Enlem
39.925533
39.925533
39.925533
...

# VERİ ZENGİNLEŞTİRME



# VERİ ZENGİNLEŞTİRME



## Güneş Parametreleri

- Altitude Angle
- ~~Zenith Angle~~
- Azimuth Angle
- Sun Distance

# VERİ ZENGİNLEŞTİRME

## Güneş Parametreleri

- Altitude Angle
- Azimuth Angle
- Sun Distance

**Güneşin yukarıdan yaptığı açı**

**Güneşin yan ekseninde yaptığı açı**

**Mevsimsel etkiler**



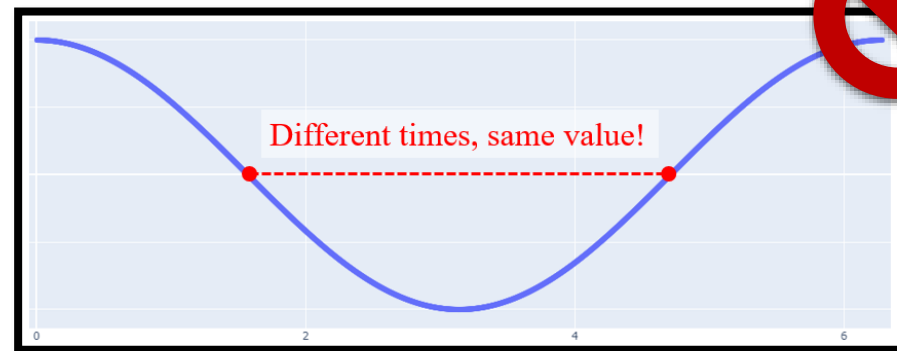
**Dönemsel değişen saat farklılığına karşı daha kararlı**



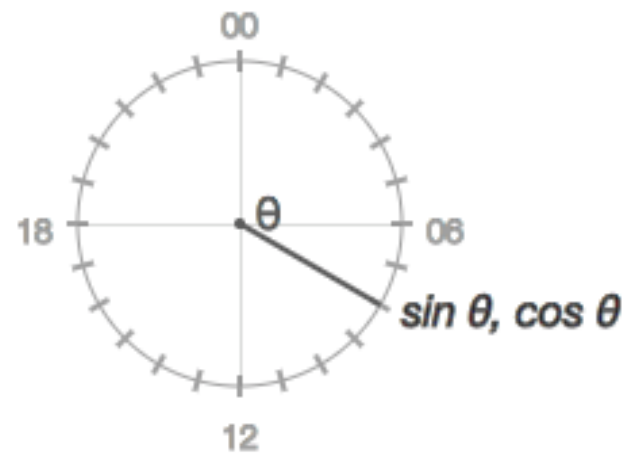
# ÖZELLİK ÇIKARIMI

## Döngüsel özellikler

- Saat
- Yılın günü
- Hafta numarası
- Derece cinsinden yön

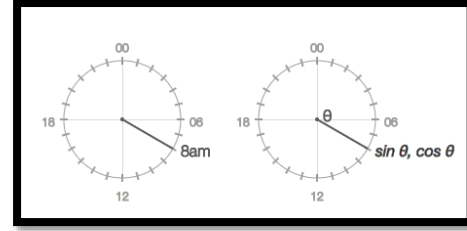


# ÖZELLİK ÇIKARIMI



# ÖZELLİK ÇIKARIMI

Rüzgar hızı  
Rüzgar yönü



Yataydaki rüzgar hızı  
Dikeydeki rüzgar hızı

# ÖZELLİK ÇIKARIMI

## Geçmiş (Lag) Temelli Ortalamalar

- Son **3** saatin **ortalama** bulut oranı
- Son **3** saatin **ortalama** bağıl nemi
- Son **3** saatin **ortalama** hava sıcaklığı

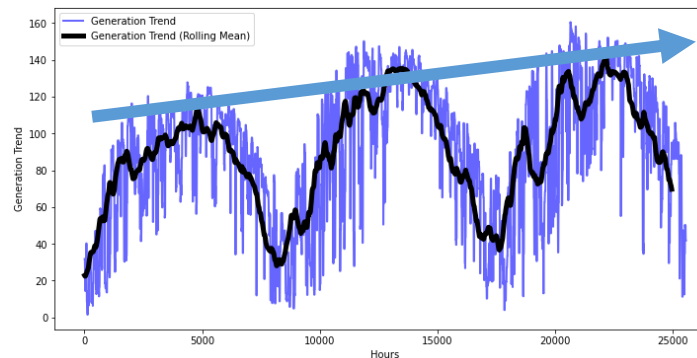
# ÖZELLİK ÇIKARIMI

## Gelecek (Lead) Sensör Tahmin Verileri

- Gelecek **8** saat için bulut oranları
- Gelecek **8** saat için bağıl nem miktarı
- Gelecek **8** saat için hissedilen sıcaklık
- Gelecek **8** saat için hava sıcaklığı
- Gelecek **8** saat için hava durumu kodu

# ÖZELLİK ÇIKARIMI

İl	Sayı	Toplam Güç(kW)
ANKARA	664	382804.941
BARTIN	2	520
KARABÜK	25	14561.2
KASTAMONU	25	9616.6
KIRIKKALE	61	55336.1
ZONGULDAK	5	2830
ÇANKIRI	66	56336.16



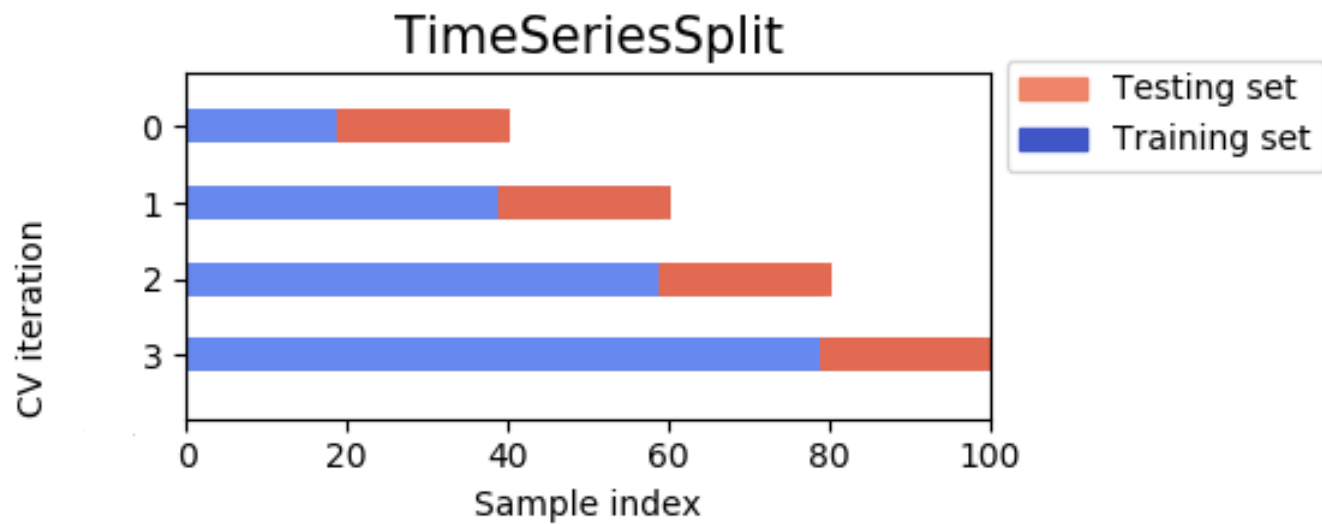
## Dolaylı özellikler

- Yıl
- Hafta No

# SON VERSİYONDA KULLANILMAYAN ÖZELLİKLER

- Dew temperature
- Air mass
- Wind Chill
- Cloud Base
- Basınç – Nem (WMO Dataset)
- Hava durumu kodu temelli mean encoding, LOO encoding, CatBoost encoding
- Sıcaklık, nem ve bulut oranı için lag temelli değerler
- PCA çıktıları
- K-Means çıktıları

# ÇAPRAZ DOĞRULAMA



- ~30 günlük test seti
- **CV3** (son 3 ay)



# TAHMİN MODELİ

- SVR
  - Linear Regression Ailesi (LinReg, Lasso, Ridge, ElasticNet)
  - kNN Regression
- ~70 RMSE
- ~35 RMSE
- LGBM
  - XGBoost
  - CatBoost
- ~18 RMSE
- Keras TabNet
- ~23 RMSE
- Keras LSTM
  - Keras GRU
- ~35 RMSE
- Prophet
  - SARIMAX
- ~50 RMSE

# TAHMİN MODELİ

- SVR
- Linear Regression Ailesi (LinReg, Lasso, Ridge, ElasticNet)
- kNN Regression

~70 RMSE  
~35 RMSE

- LGBM
- XGBoost
- CatBoost

~18 RMSE

## Stacked-Ensemble

- LR Family Head
- GBDT-Type Head

~18 RMSE

- Keras TabNet

~23 RMSE

- Keras LSTM
- Keras GRU

~35 RMSE

- Prophet
- SARIMAX

~50 RMSE

# TAHMİN MODELİ

## Hyper-Parameter Tuning

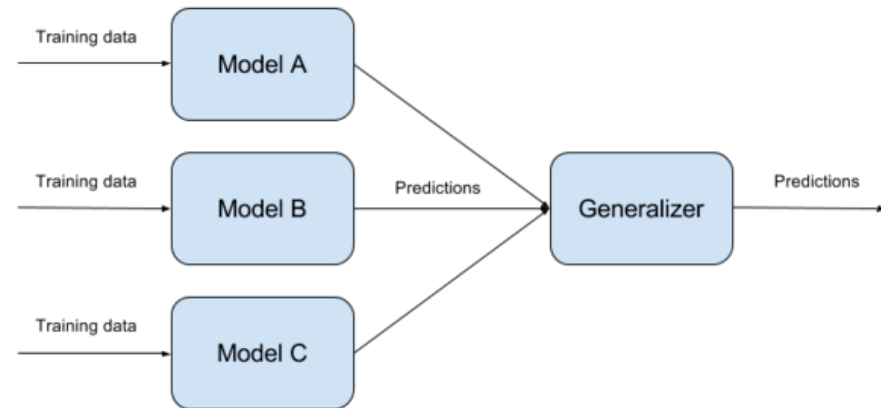
- **RMSE minimizasyonu**
- Depth
- Learning Rate
- Estimators
- Max Bin
- Max Data in Leaf
- Leaf Regression Coefficient



OPTUNA

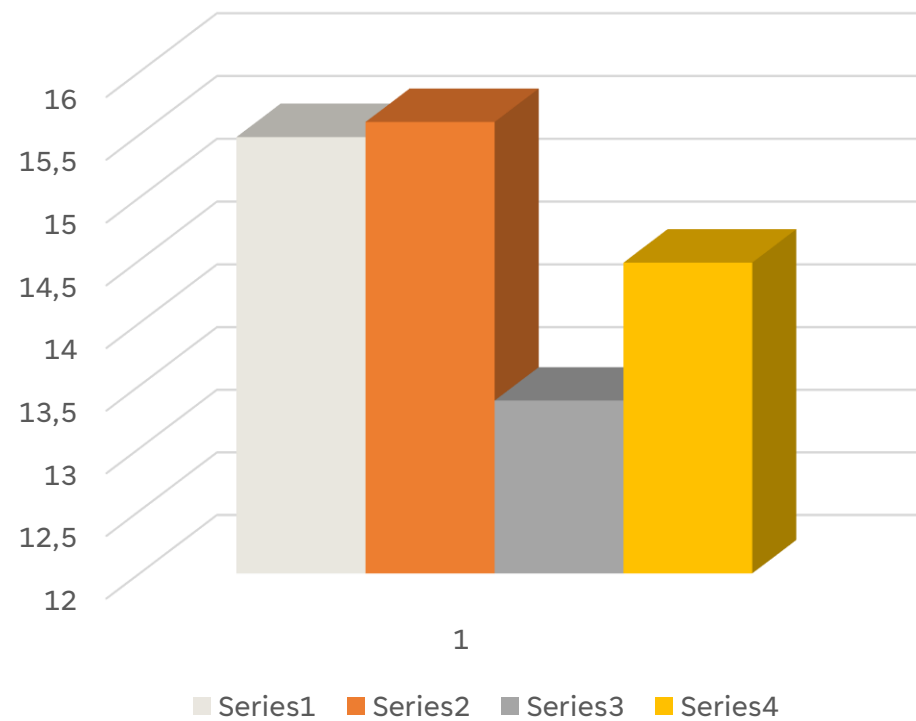
# TAHMİN AŞAMASI

- Farklı seedler ile model eğitimi
- Her model ile tahmin
- Modellerin tahminlerin ortalamasının alınması



# SONUÇLAR

Başarı Kıyası (RMSE)





# TEŞEKKÜRLER

Anıl Öztürk

[anilozturk96@gmail.com](mailto:anilozturk96@gmail.com)



# SORU & CEVAP