

MAKİNE ÖĞRENMESİNİN TEMELLERİ

1. Veri Seti

- Tanım:** Makine öğrenmesi modellerinin öğrenmesi için kullanılan veri.
- Örnek:** CSV dosyasında saklanan iris çiçekleri hakkında bilgi içeren tablo.

2. Özellikler (Features) ve Hedef Değişken (Target Variable)

- Özellikler:** Modelin öğrenmesi için kullanılan bağımsız değişkenler.
- Hedef Değişken:** Modelin tahmin etmeye çalıştığı bağımlı değişken.
- Örnek:** Ev fiyatı tahmini için özellikler: oda sayısı, evin büyüklüğü; hedef değişken: evin fiyatı.

3. Eğitim ve Test Veri Seti (Training and Test Set)

- Eğitim Seti:** Modelin öğrenmesi için kullanılan veri.
- Test Seti:** Modelin performansını değerlendirmek için kullanılan veri.
- Örnek:** %80 eğitim, %20 test olarak veri setinin bölünmesi.

4. Model

- Tanım:** Veriden öğrenen ve tahmin yapabilen matematiksel yapı.
- Örnek:** Doğrusal regresyon modeli, karar ağacı modeli.

5. Aşırı Uydurma (Overfitting) ve Yetersiz Uydurma (Underfitting)

- Aşırı Uydurma:** Modelin eğitim verisine çok iyi uyum sağlayıp, test verisinde kötü performans göstermesi.
- Yetersiz Uydurma:** Modelin ne eğitim verisine ne de test verisine iyi uyum sağlayamaması.
- Önlem:** Düzenleştirme, çapraz doğrulama, daha fazla veri kullanma.

6. Doğrulama ve Çapraz Doğrulama (Validation and Cross-Validation)

- **Doğrulama Seti:** Modelin hiperparametrelerini ayarlamak için kullanılan veri seti.
- **Çapraz Doğrulama:** Veriyi farklı alt kümelerine ayırarak modelin performansını daha sağlam değerlendirme yöntemi.
- **Örnek:** K-Katlı Çapraz Doğrulama (K-Fold Cross Validation)

7. Hata Metrikleri (Error Metrics)

- **Regresyon:** Ortalama Kare Hatası (MSE), Ortalama Mutlak Hata (MAE), R^2 Skoru
- **Sınıflandırma:** Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Skoru, ROC-AUC

8. Model Seçimi ve Hiperparametre Ayarlama (Model Selection and Hyperparameter Tuning)

- **Model Seçimi:** Farklı algoritmaları deneyerek en iyi performansı gösteren modeli seçmek.
- **Hiperparametre Ayarlama:** Modelin performansını optimize etmek için hiperparametreleri ayarlamak (örneğin, ağacın derinliği, öğrenme oranı).

9. Özellik Mühendisliği (Feature Engineering)

- **Tanım:** Ham veriden anlamlı özellikler oluşturma süreci.
- **Örnek:** Zaman damgası verilerinden gün, ay, yıl bilgisi çıkarma.

10. Özellik Ölçekleme (Feature Scaling)

- **Tanım:** Verinin belirli bir aralığa ölçeklenmesi.
- **Örnek:** Standartlaştırma (Standardization), Min-Max Ölçekleme (Normalization).

11. Düzenleştirme (Regularization)

- **Tanım:** Modelin karmaşıklığını azaltarak aşırı uydurmayı önleme yöntemi.
- **Örnek:** Lasso (L1) ve Ridge (L2) regresyon.

12. Algoritma Türleri

- **Denetimli Öğrenme (Supervised Learning):** Etiketli veri ile model eğitimi (örneğin, sınıflandırma, regresyon).
- **Denetimsiz Öğrenme (Unsupervised Learning):** Etiketlenmemiş veri ile model eğitimi (örneğin, kümeleme, boyut indirgeme).
- **Pekiştirmeli Öğrenme (Reinforcement Learning):** Bir ajan (agent) ödül (reward) veya ceza (punishment) alarak öğrenir.

13. Sınıflandırma ve Regresyon

- **Sınıflandırma:** Kategorik hedef değişkenleri tahmin etme (örneğin, spam/ham).
- **Regresyon:** Sürekli hedef değişkenleri tahmin etme (örneğin, ev fiyatı tahmini).

14. Boyut İndirgeme (Dimensionality Reduction)

- **Tanım:** Verideki özellik sayısını azaltarak veri setini daha yönetilebilir hale getirme.
- **Örnek:** Ana Bileşen Analizi (PCA), T-SNE.

15. Doğal Dil İşleme (Natural Language Processing, NLP)

- **Tanım:** Metin verisiyle çalışmak ve bu veriden anlamlı bilgiler çıkarmak.
- **Örnek:** Sentiment analizi, metin sınıflandırma, özetleme.