



12.05.2023

M.FURKAN PORTAKAL

18360859046

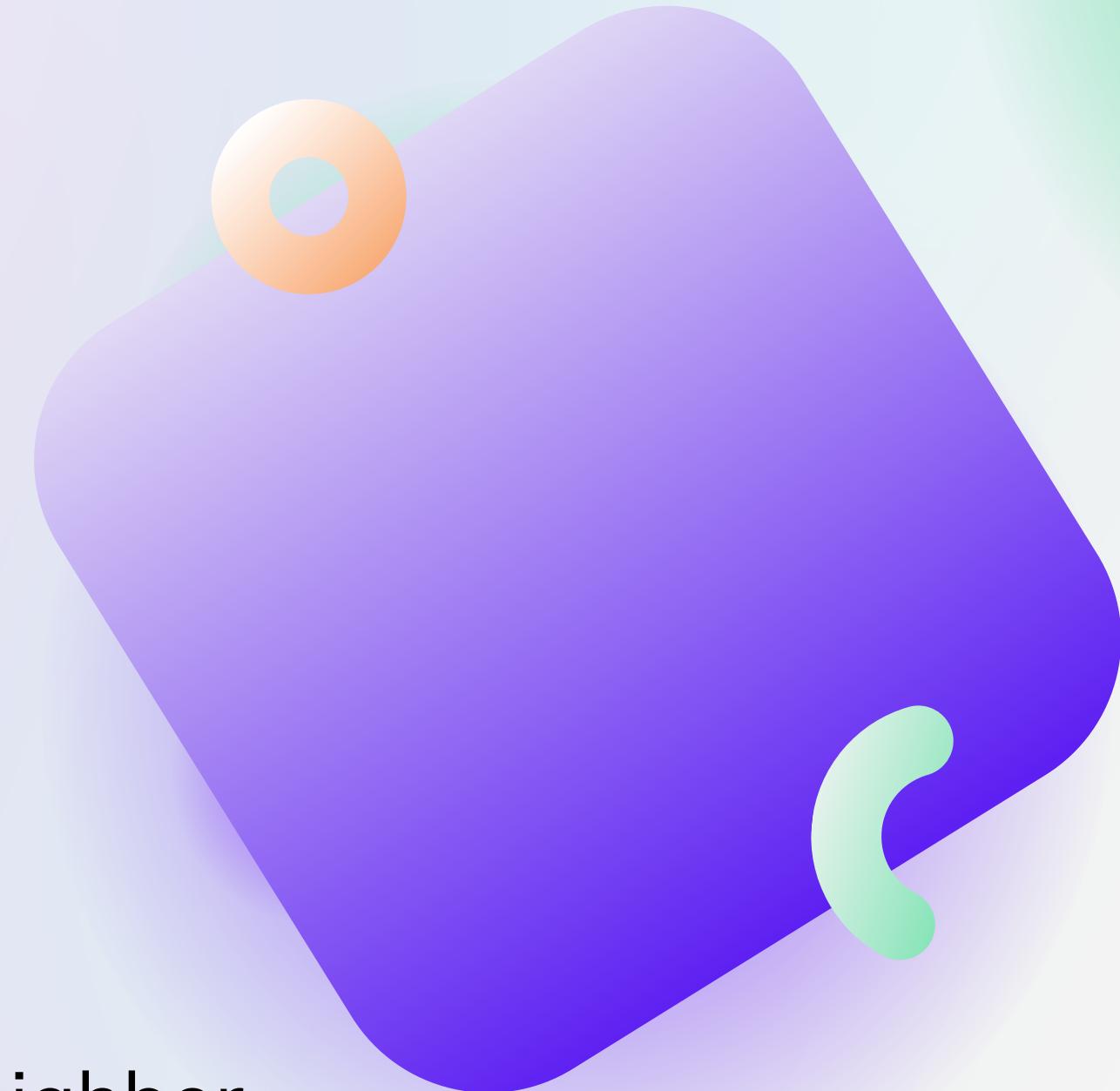
# BLM463

# Veri

# Madenciliğine

# Giriş

Bank Marketing Veri Seti Üzerinde K-Nearest Neighbor  
Sınıflandırıcısı Kullanımı ve Sonuçlar

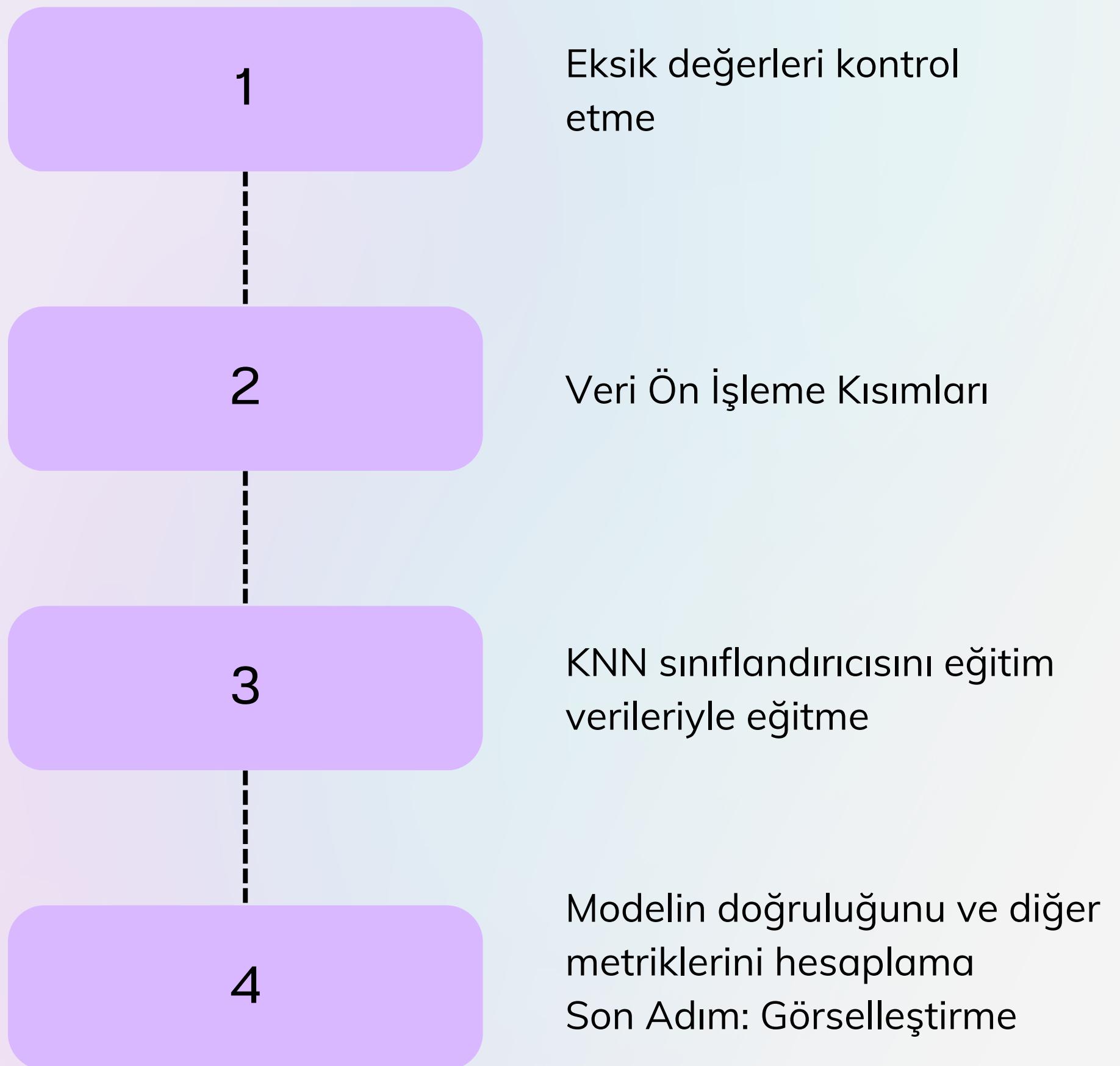


# Giriş

Merhaba herkese, bugün sizlere UCI Machine Learning Repository'den alınan Bank Marketing veri seti üzerinde K-Nearest Neighbor benimseyerek gerçekleştirdiğim sınıflandırma çalışmasını sunacağım.

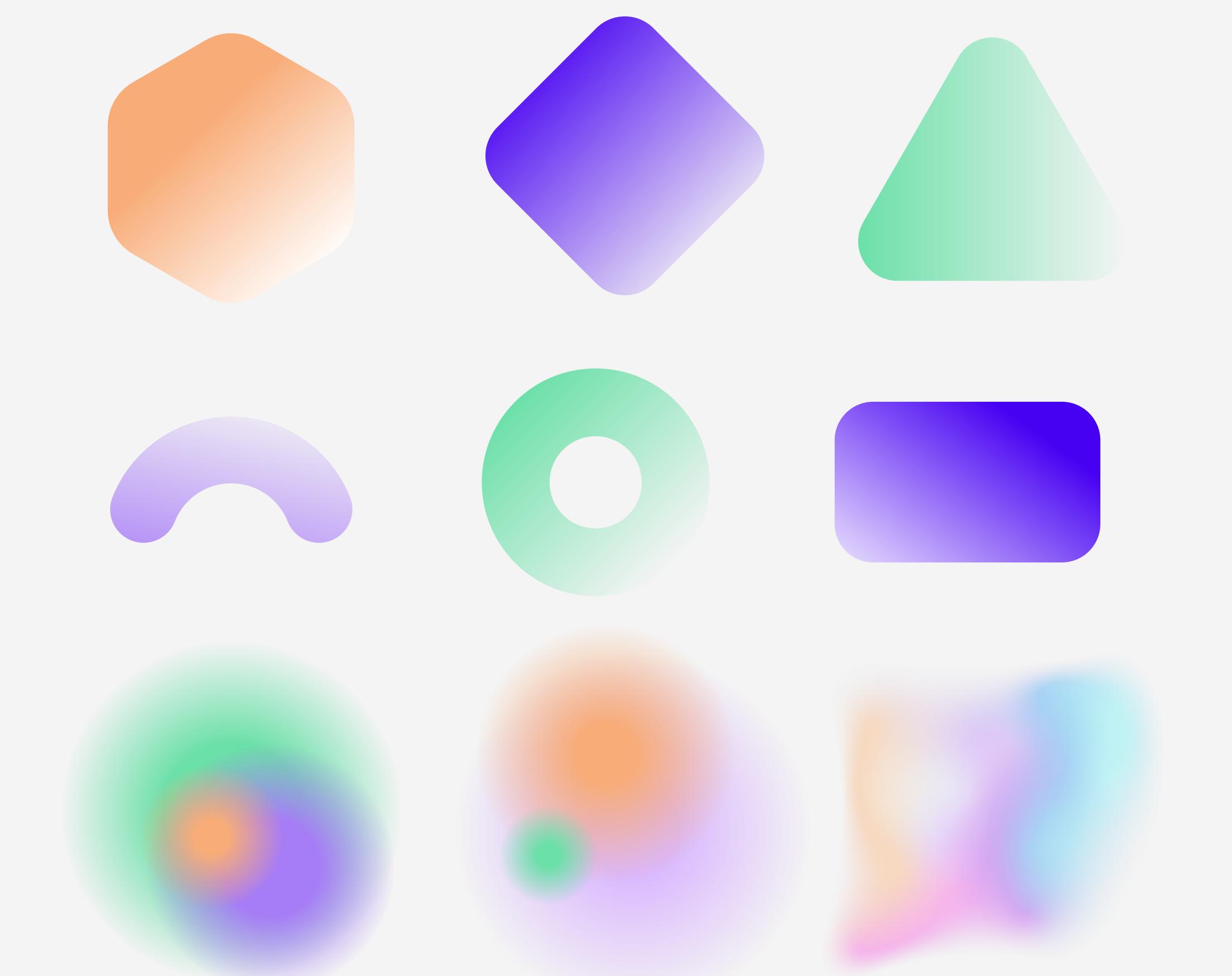
# Adımlar

Projedeki Adımlar



# Veri Seti Hakkında

Bu veri seti, bir Portekiz bankasının müşterilerine yaptığı telefon pazarlama kampanyası sonucu toplanmış verileri içermektedir. Toplamda 45.211 gözlem ve 17 özellik bulunmaktadır.

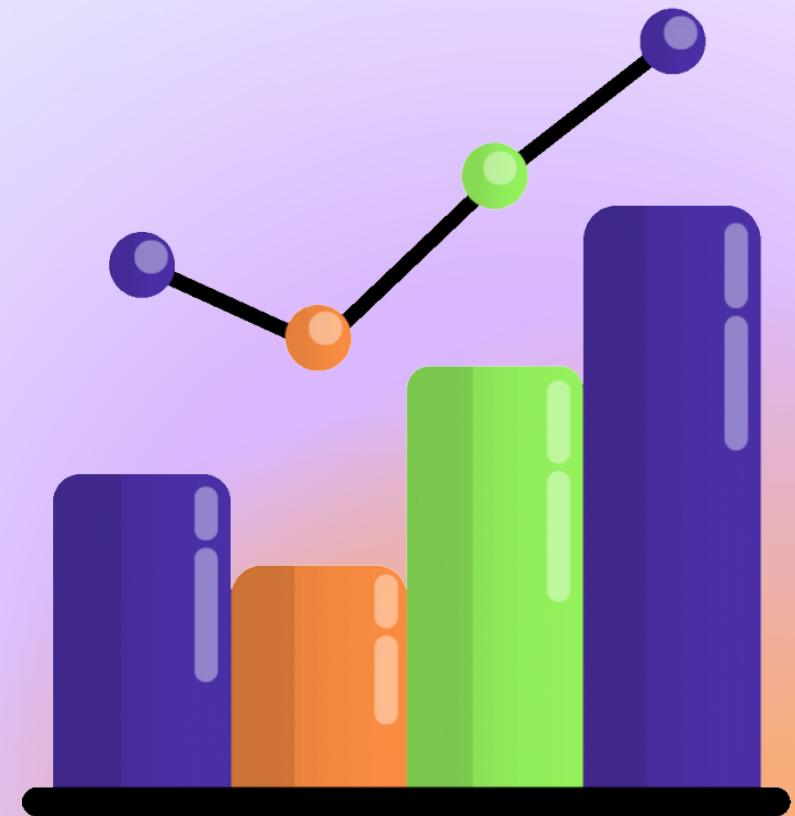


# K-Nearest Neighbor Nedir?

K-Nearest Neighbor (KNN), sınıflandırma ve regresyon problemlerinde kullanılan bir makine öğrenmesi algoritmasıdır. Sınıflandırma problemlerinde, yeni bir gözlem noktası komşularının etiketlerine göre sınıflandırılır.

# Veri Ön İşleme

Veri seti UCI Machine Learning Repository'den alındı ve banka pazarlama kampanyalarının sonuçlarını içeriyor. Ön işlem adımlarında, eksik veri kontrolü gerçekleştirdik ve hiç eksik veri olmadığını tespit ettik. Ardından, kategorik özellikler için One-Hot Encoding ve sayısal özellikler için StandardScaler kullanılarak veri ölçeklendirildi.



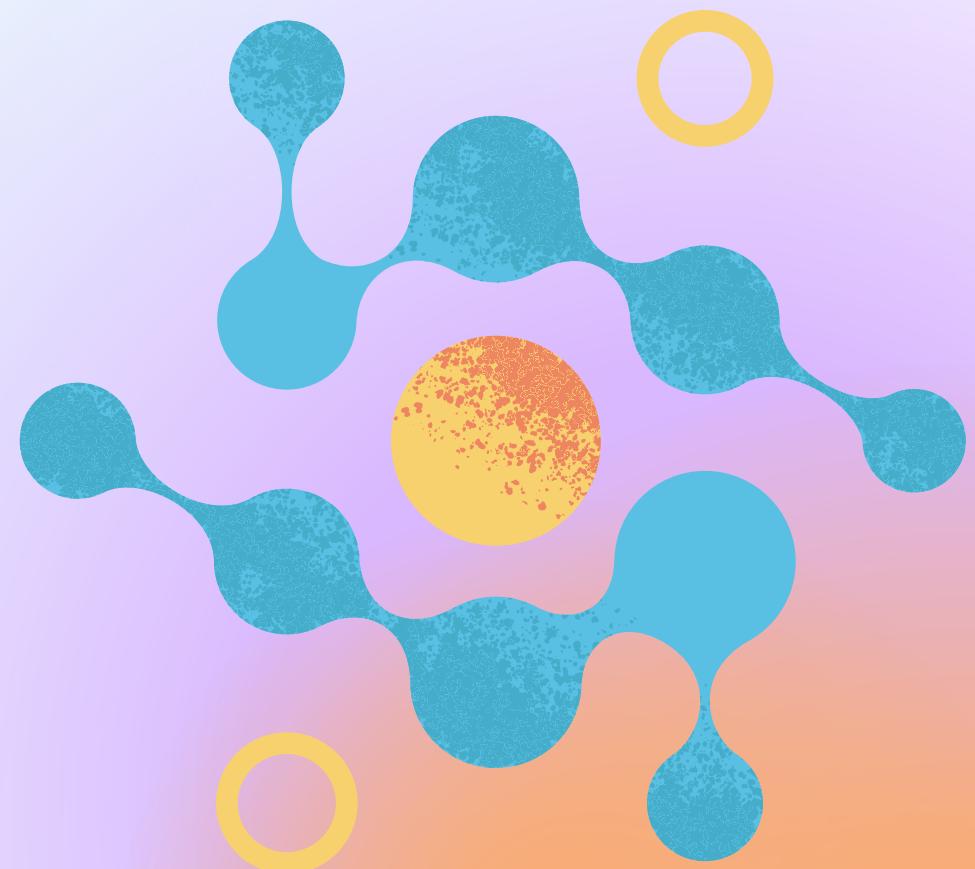
# Model Eğitimi

Veri kümesi eğitim ve test seti olarak ayrıldı ve K-Nearest Neighbor sınıflandırıcısını eğitim verileriyle eğitme gerçekleştirdik.



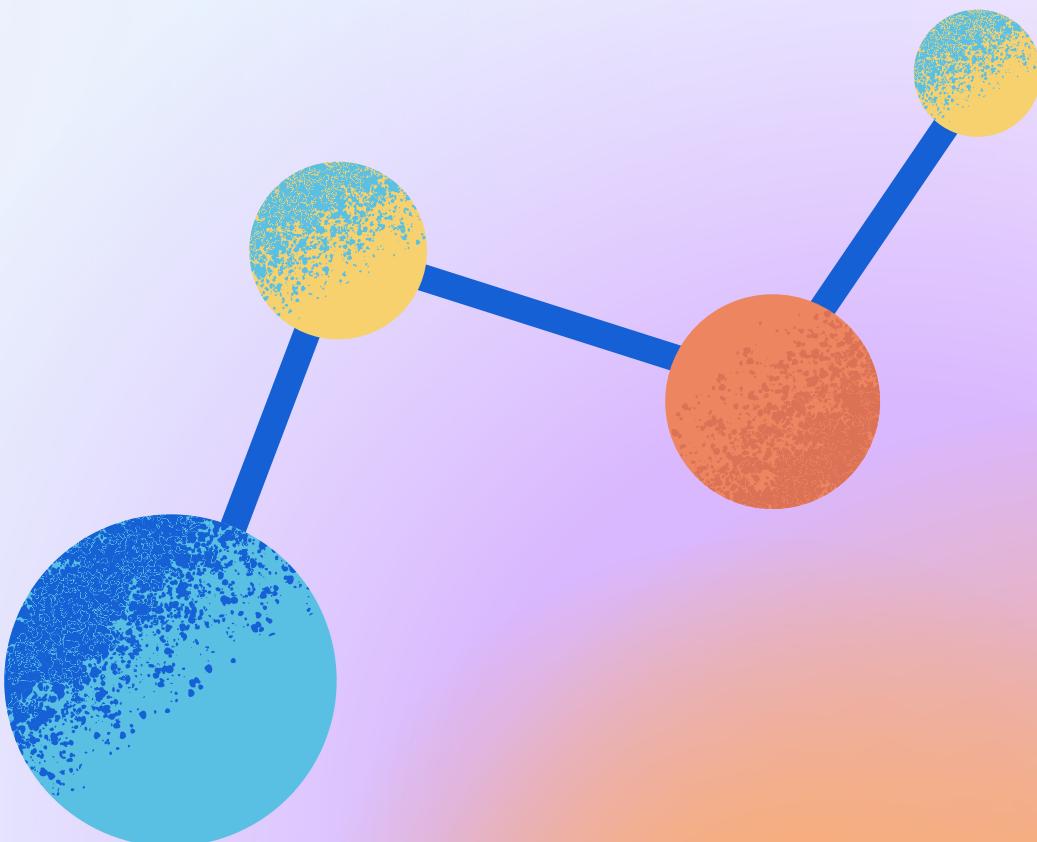
# Model Performansı

Modelin genel doğruluk oranı hesaplandı ve ayrıca sınıflandırma raporu ve karmaşıklık matrisi kullanılarak daha ayrıntılı bir değerlendirme yapıldı.



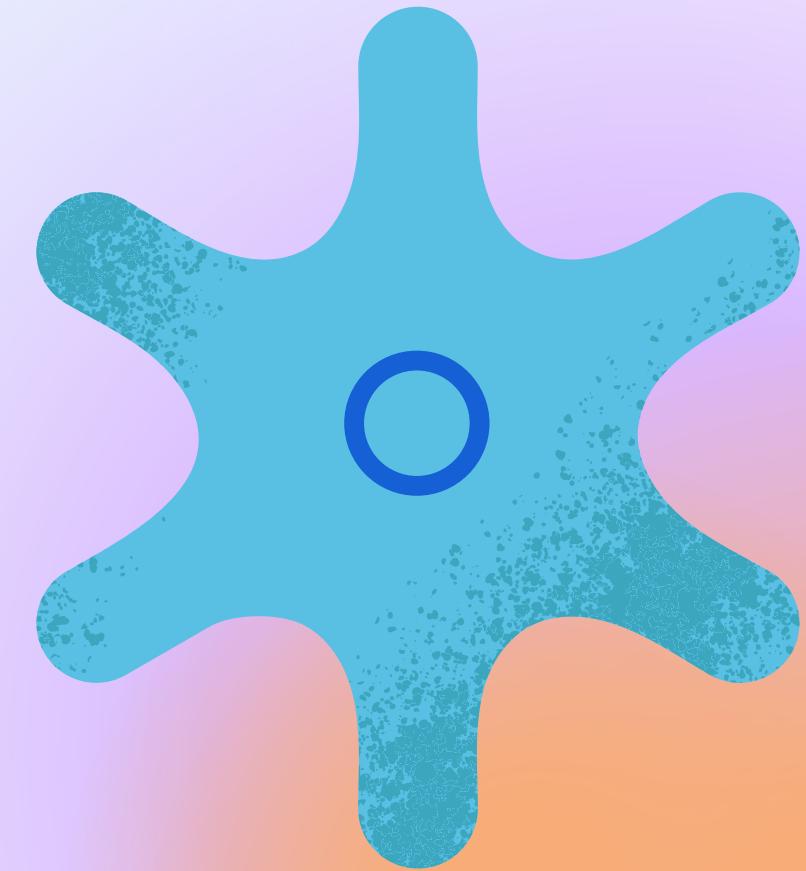
# Görselleştirme

Modelin performansını daha net anlamak adına, karmaşıklık matrisini ve ROC eğrisini görselleştirdik. Bu, modelin genel performansını ve sınıflandırma yeteneğini daha iyi anlamamıza yardımcı oldu.



# Karşılaştırma ve Tartışma

Önceden yapılmış bir çalışmayı bulduk ve sonuçlarını kendi sonuçlarımıza karşılaştırdık. Bu, modelimizin performansını daha geniş bir perspektifte değerlendirmemize yardımcı oldu.



# Sonuç

Sonuç olarak, K-Nearest Neighbor sınıflandırıcı, Bank Marketing veri setinde oldukça iyi bir performans sergiledi. Modelin genel doğruluk oranı yükseltti ve ROC eğrisi de modelin iyi bir sınıflandırma yeteneğine sahip olduğunu gösterdi.

