

MAKİNE  
ÖĞRENİMİ  
PROJE

WATER QUALITY

FATMA SENA YÜKSEL

# WATER QUALITY

- Su kalitesi, insan sağığı ve çevresel denge açısından kritik bir öneme sahiptir. Bu projede, mevcut su kaynaklarının çeşitli parametreler üzerinden analizi yapılarak suyun temizliğı, içilebilirliğı ve çevresel etkileri değerlendirilmiştir. Kullanılan veri seti, kimyasal bileşenler, mikrobiyolojik özellikler, pH, çözünmüş oksijen seviyeleri gibi çeşitli faktörleri içermekte ve bu verilerin analiziyle su kalitesinin detaylı değerlendirilmesi yapılmıştır.

# WATER QUALITY PARAMETERS

Bu veri seti toplam 21 parametreden oluşmaktadır

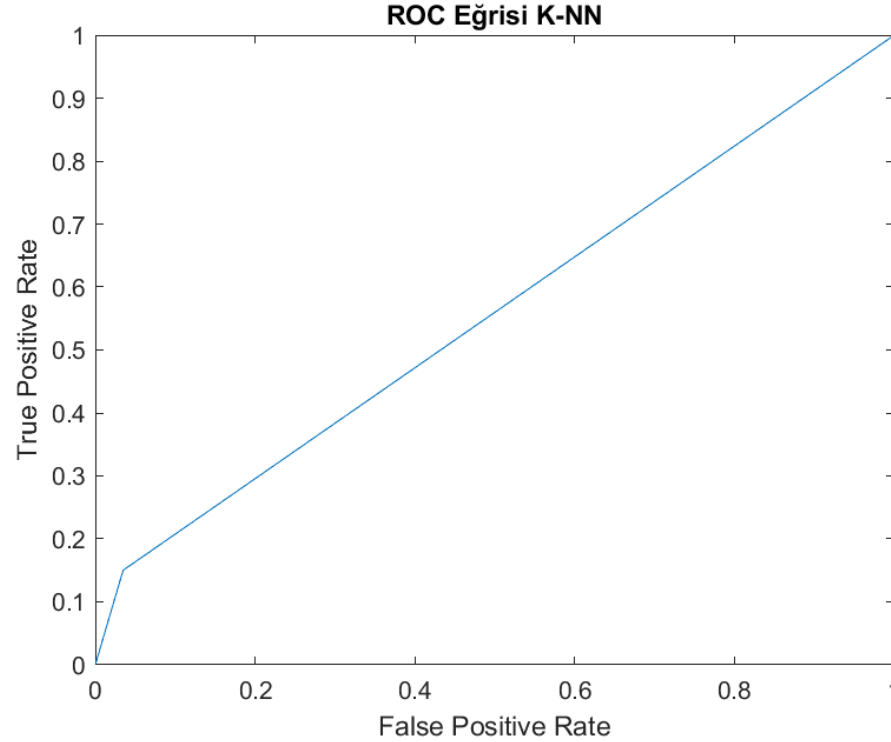
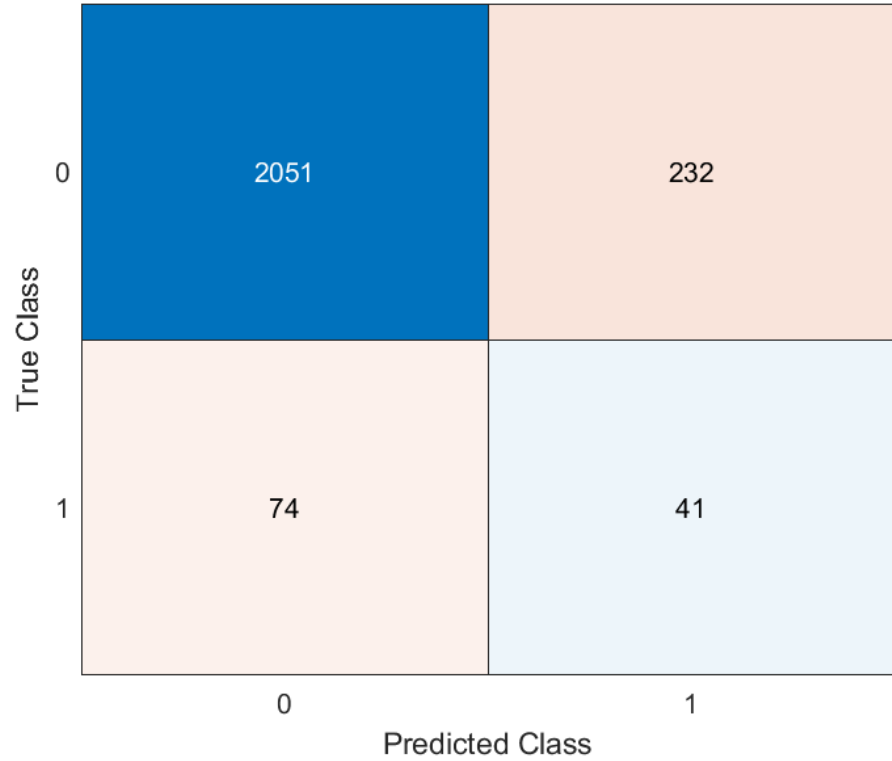
- alüminyum - 2,8'den büyükse tehlikeli
- amonyak - 32,5'tan büyükse tehlikeli
- arsenik - 0,01'den büyükse tehlikeli
- baryum - 2'den büyükse tehlikeli
- kadmiyum - 0,005'ten büyükse tehlikeli
- kloramin - 4'ten büyükse tehlikeli
- krom - 0,1'den büyükse tehlikeli
- bakır - 1,3'ten büyükse tehlikeli
- florür - 1,5'tan büyükse tehlikeli
- bakteri - 0'dan büyükse tehlikeli
- virüsler - 0'dan büyükse tehlikelidir
- kurşun - 0,015'ten büyükse tehlikeli
- nitratlar - 10'dan büyükse tehlikeli
- nitritler - 1'den büyükse tehlikeli
- cıva - 0,002'den büyükse tehlikeli
- perklorat - 56'dan büyükse tehlikeli
- radyum - 5'ten büyükse tehlikeli
- selenyum - 0,5'ten büyükse tehlikeli
- gümüş - 0,1'den büyükse tehlikeli
- uranyum - 0,3'ten büyükse tehlikeli
- is\_safe - sınıf özelliği {0 - güvenli değil, 1 - güvenli}

# UYGULANAN YÖNTEMLER

- K-NN
- SVM
- Decision Tree
- YSA
- Naive Bayes

# K-NN

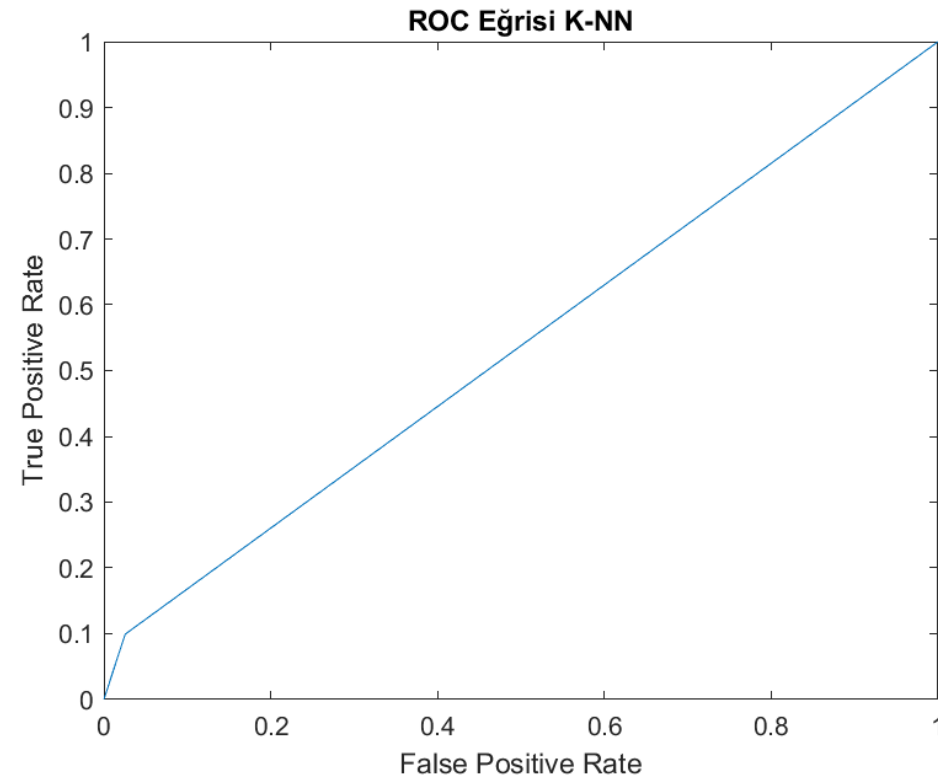
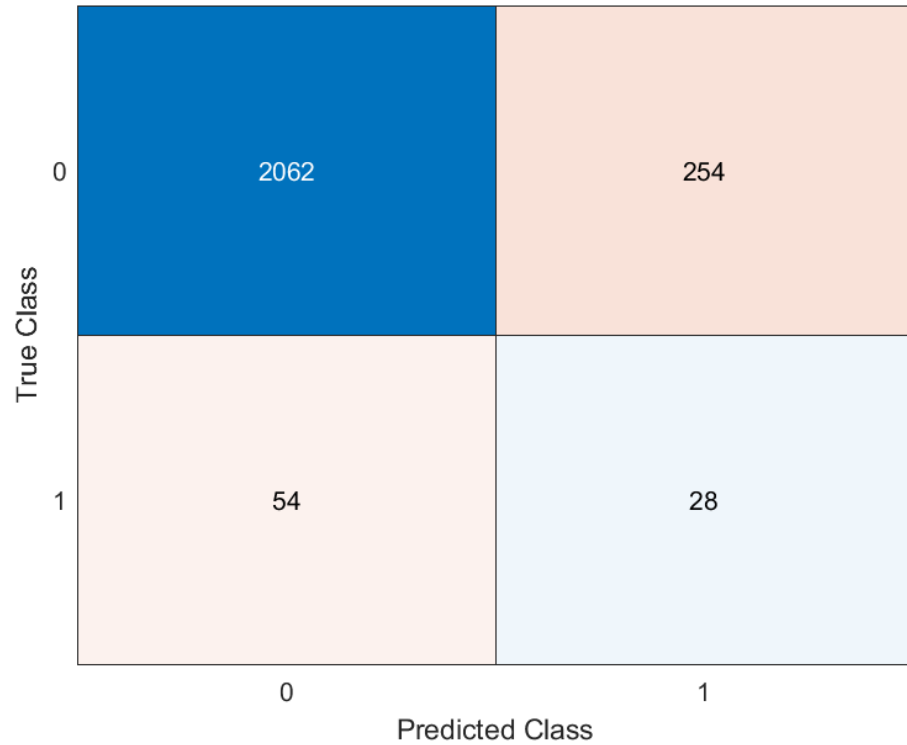
0.3 holdout 3 neighbors



Accuracy: 87.2394  
AUC:0.55768  
Kappa Değeri: 0.1543  
Duyarlılık: 0.8984  
Özgüllük: 0.3565  
Precision: 0.9652  
Recall: 0.898379  
F-Ölçümü: 0.930581

# K-NN

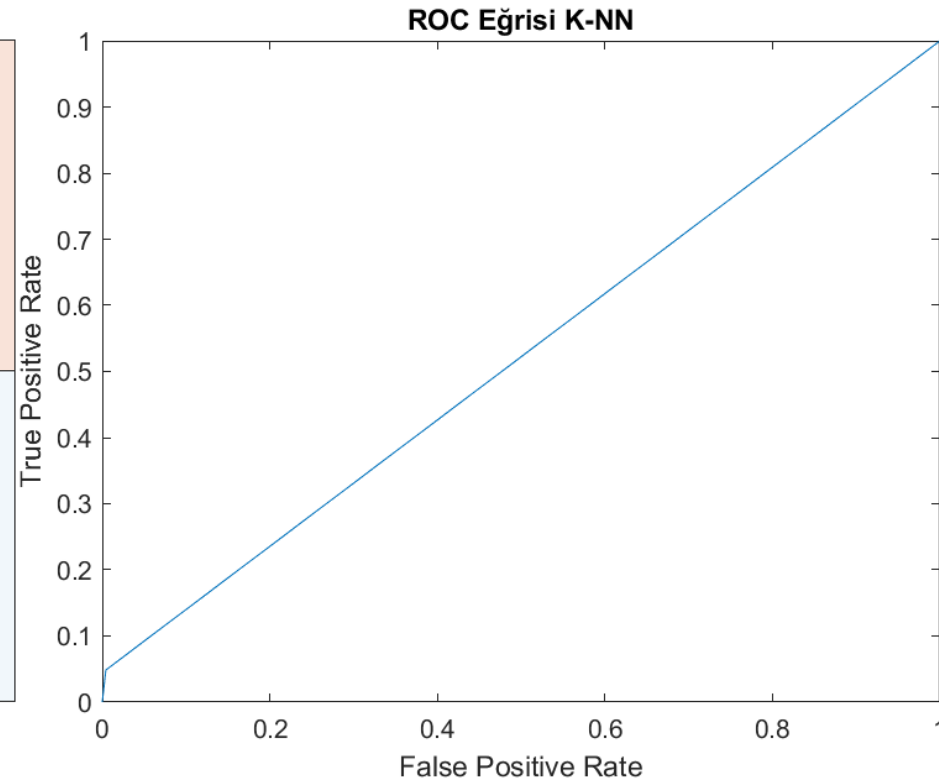
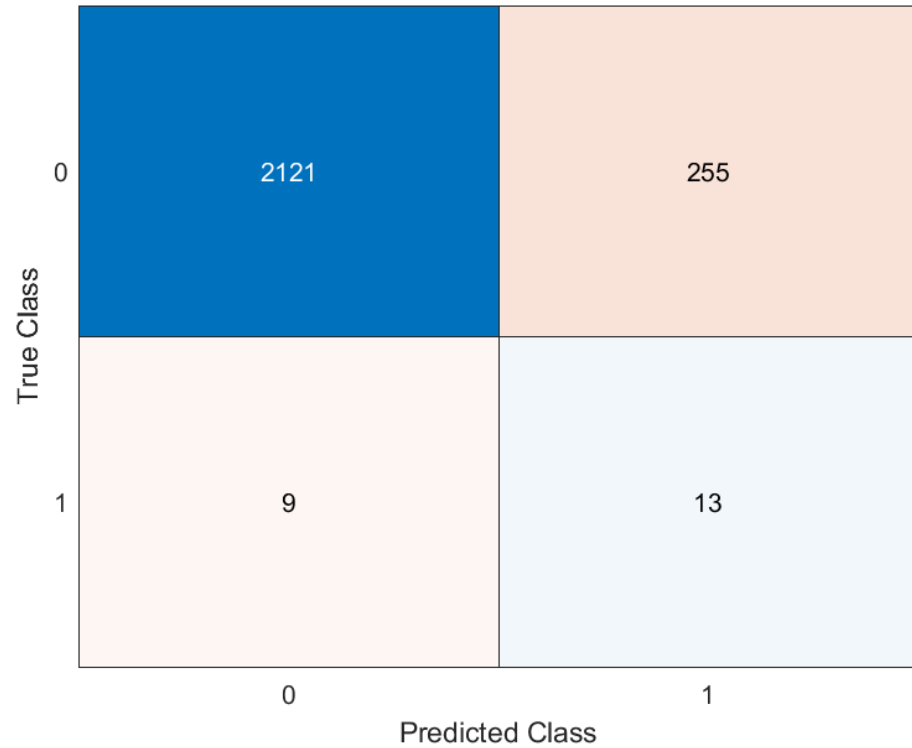
0.3 holdout 5 neighbors



Accuracy: 87.1560  
AUC:0.53689  
Kappa Değeri: 0.1065  
Duyarlılık: 0.8903  
Özgüllük: 0.3415  
Precision: 0.9745  
Recall: 0.890328  
F-Ölçümü: 0.930505

# K-NN

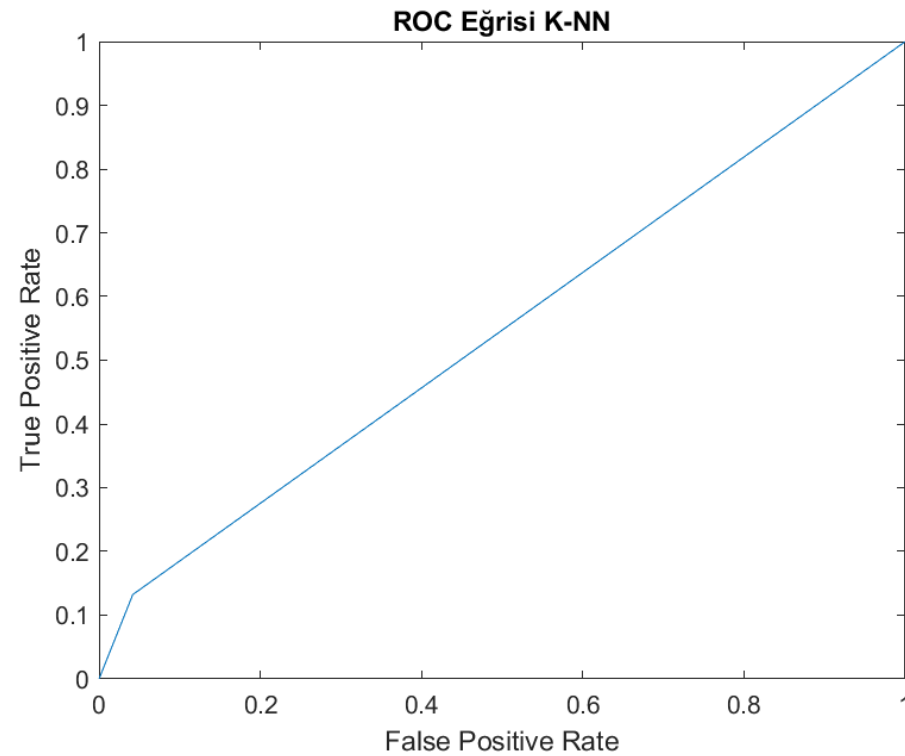
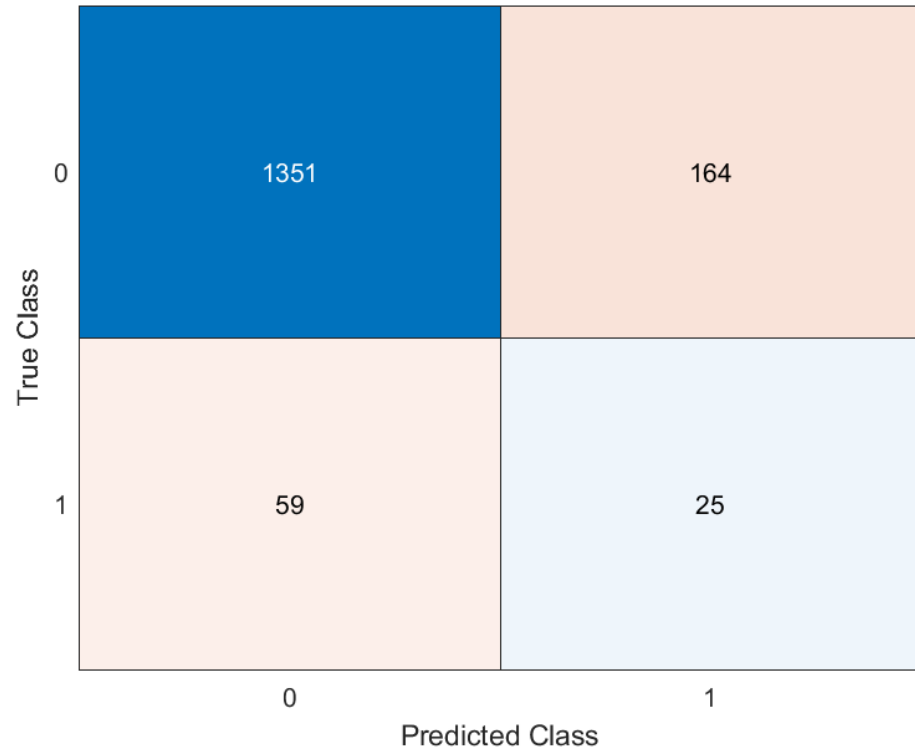
0.3 holdout 10 neighbors



Accuracy: 88.9908  
AUC:0.52214  
Kappa Değeri: 0.0740  
Duyarlılık: 0.8927  
Özgüllük: 0.5909  
Precision: 0.9958  
Recall: 0.892677  
F-Ölçümü: 0.941411

# K-NN

0.2 holdout 3 neighbors

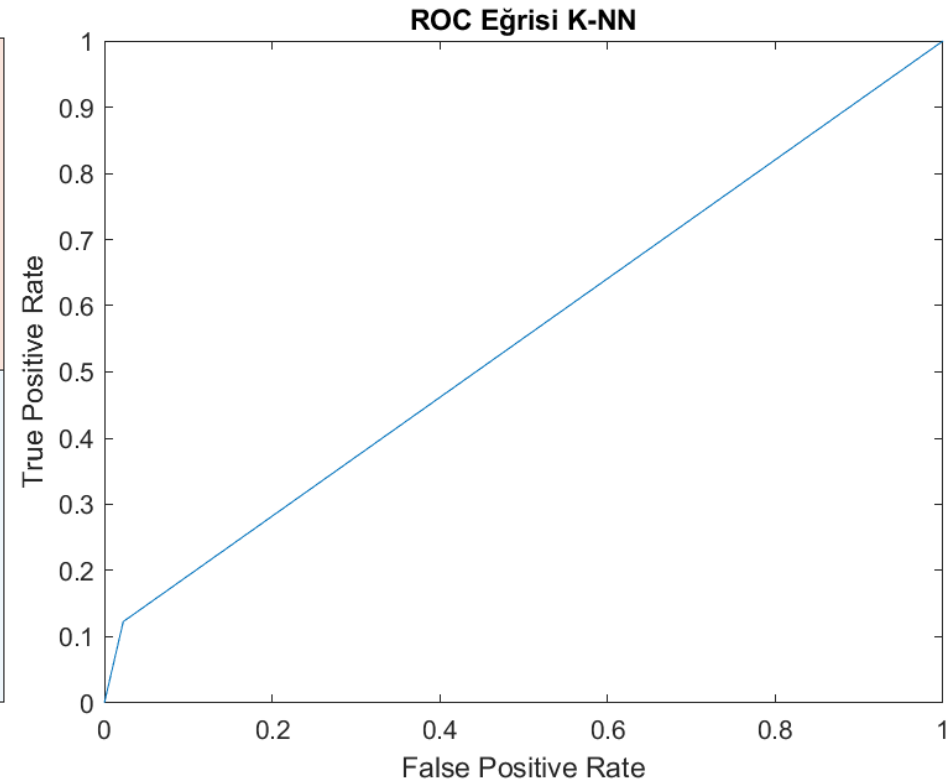
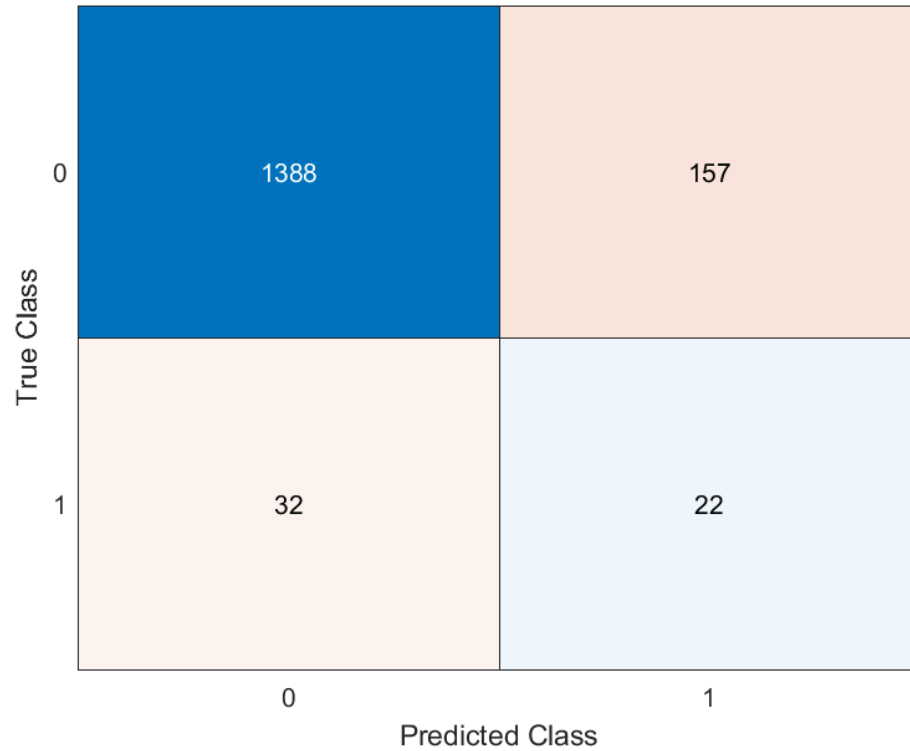


Accuracy: 86.0538  
AUC:0.54522  
Kappa Değeri: 0.1191  
Duyarlılık: 0.8917  
Özgüllük: 0.2976  
Precision: 0.9582  
Recall: 0.891749  
F-Ölçümü: 0.923761



# K-NN

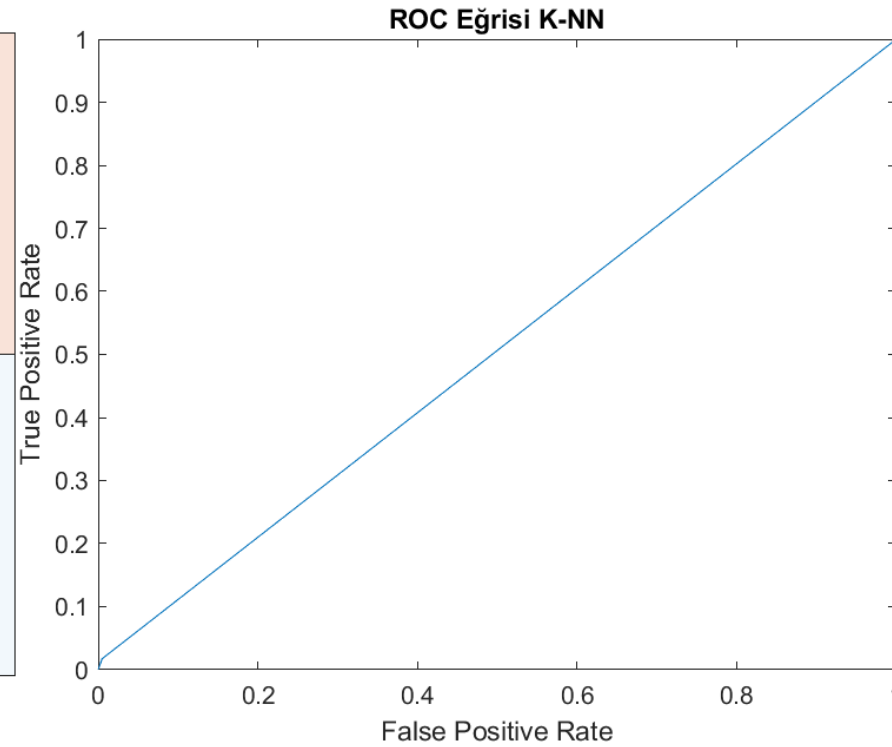
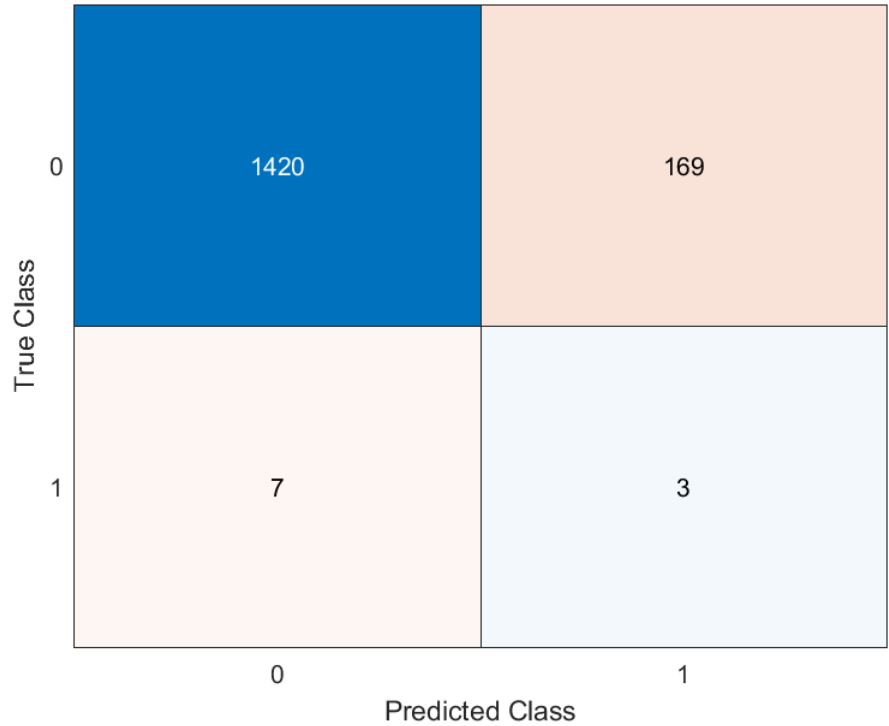
0.2 holdout 5 neighbors



Accuracy: 88.1801  
AUC:0.55018  
Kappa Değeri: 0.1444  
Duyarlılık: 0.8984  
Özgüllük: 0.4074  
Precision: 0.9775  
Recall: 0.898382  
F-Ölçümü: 0.936256

# K-NN

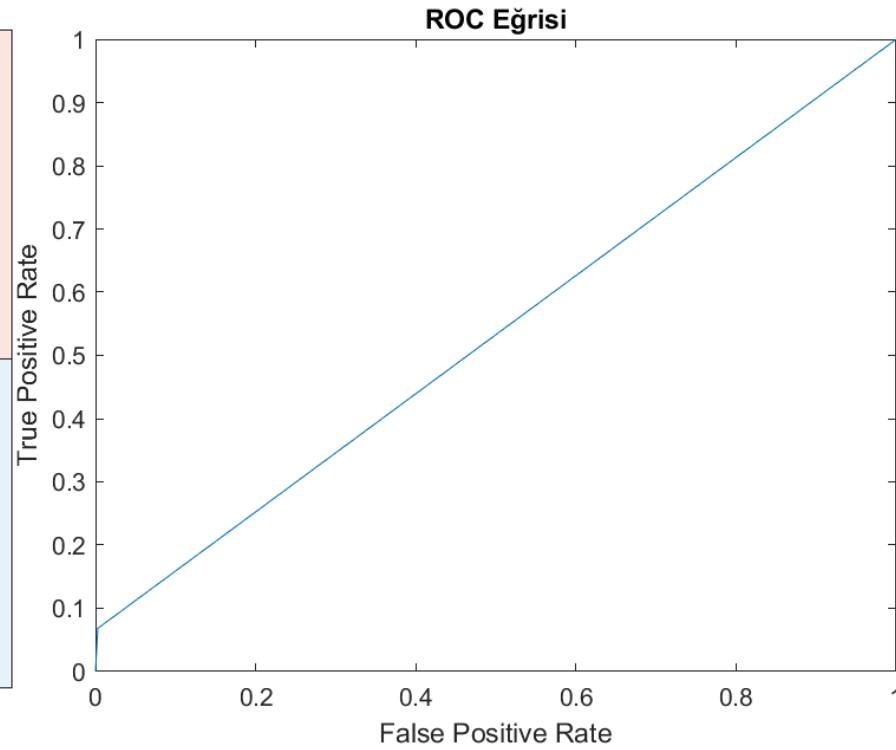
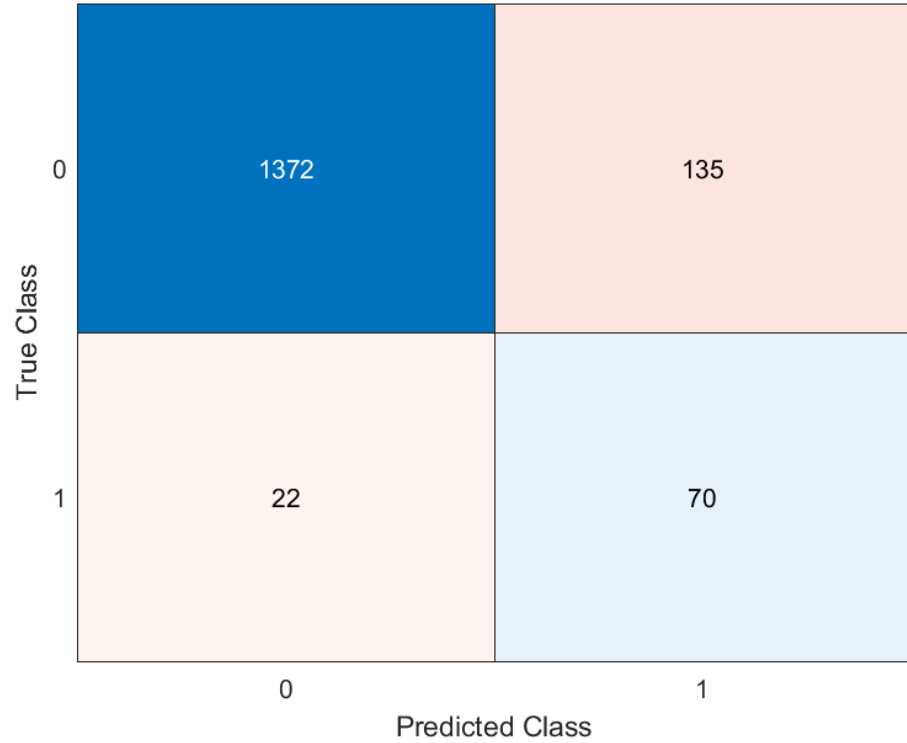
0.2 holdout 10 neighbors



Accuracy: 88.9931  
AUC:0.50627  
Kappa Değeri: 0.0214  
Duyarlılık: 0.8936  
Özgüllük: 0.3000  
Precision: 0.9951  
Recall: 0.893644  
F-Ölçümü: 0.941645

# SVM

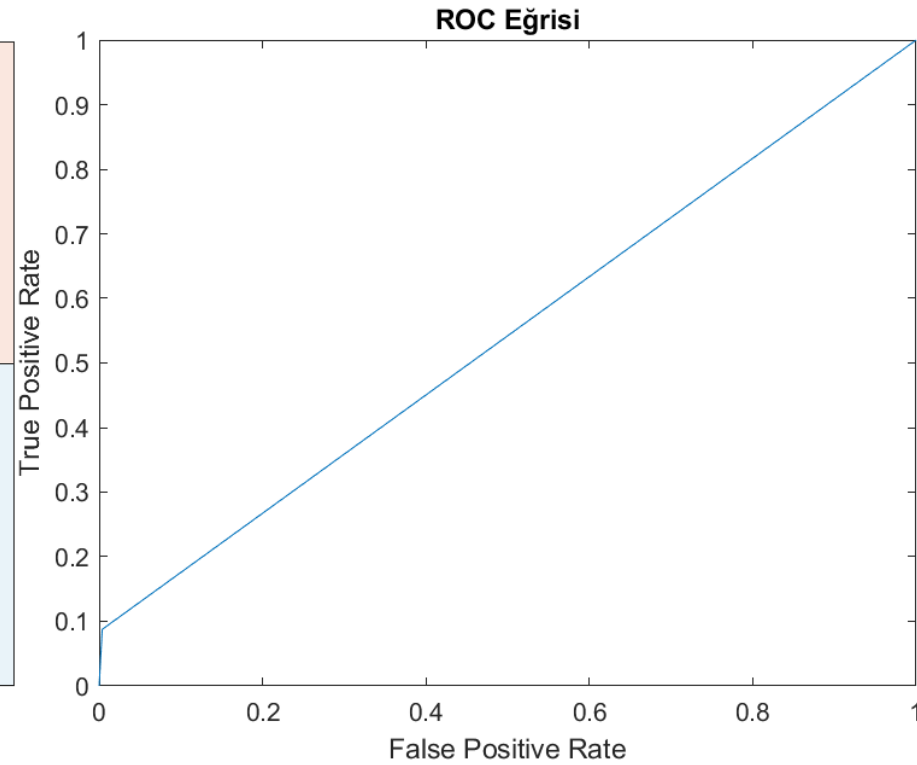
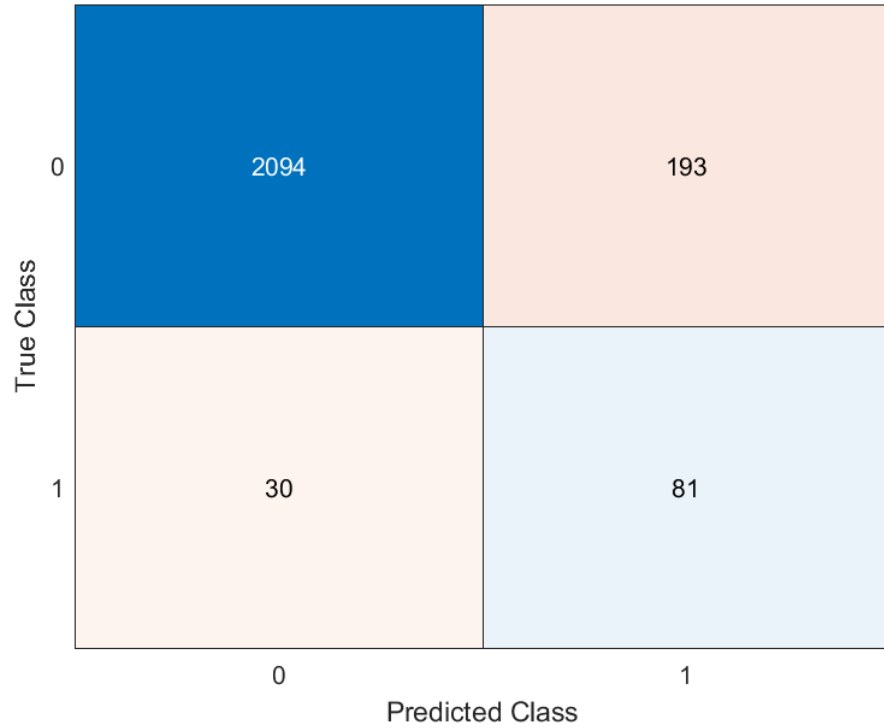
## 0.2 holdout



Accuracy: 90.1814  
AUC:0.53271  
Kappa Değeri (SVM ile): 0.4258  
Duyarlılık: 0.9104  
Özgüllük: 0.7609  
Precision: 0.9842  
Recall: 0.910418  
F-Ölçümü: 0.945881

# SVM

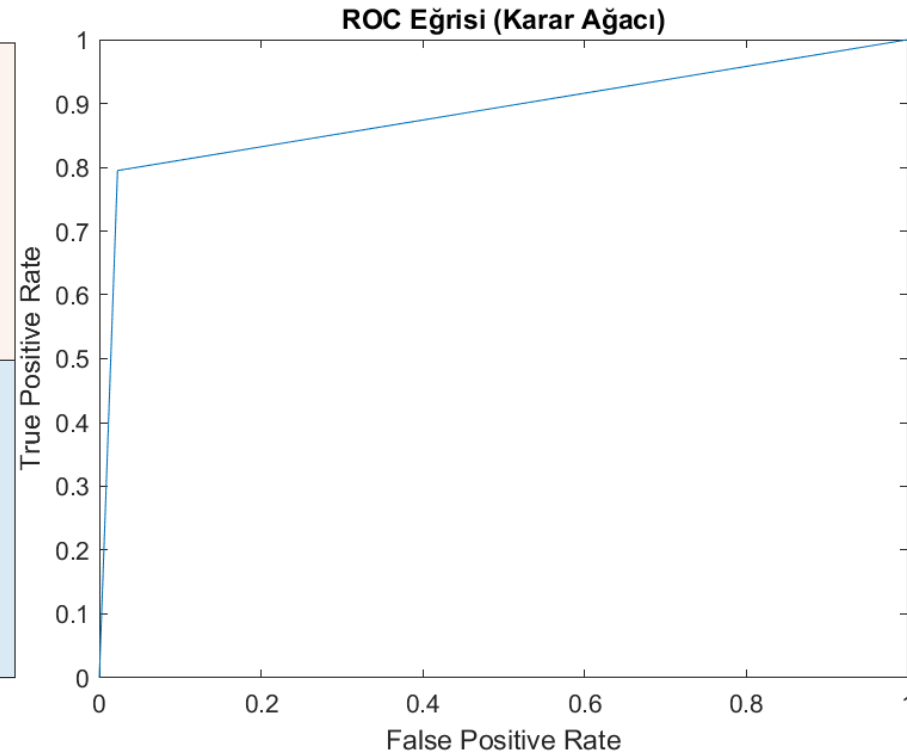
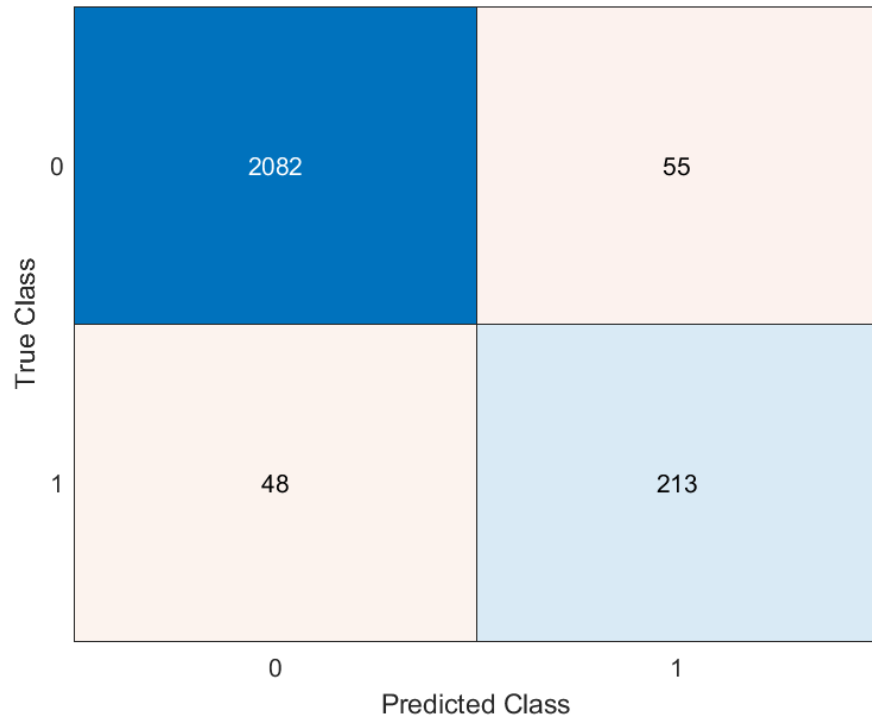
## 0.3 holdout



Accuracy: 90.7006  
AUC:0.54191  
Kappa Değeri (SVM ile): 0.3799  
Duyarlılık: 0.9156  
Özgüllük: 0.7297  
Precision: 0.9859  
Recall: 0.915610  
F-Ölçümü: 0.949445

# Decision Tree

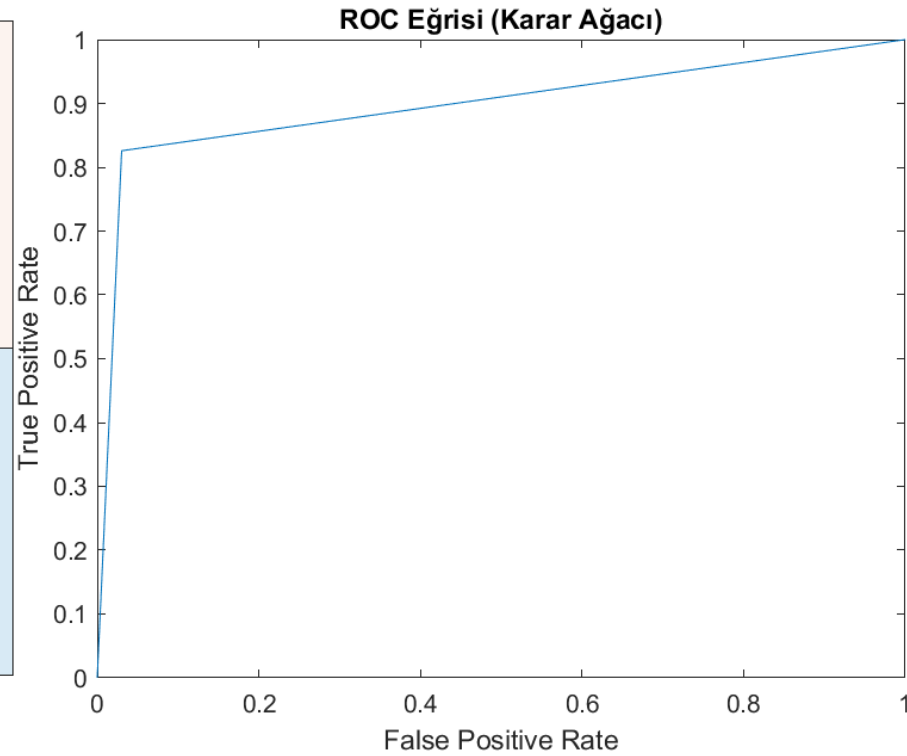
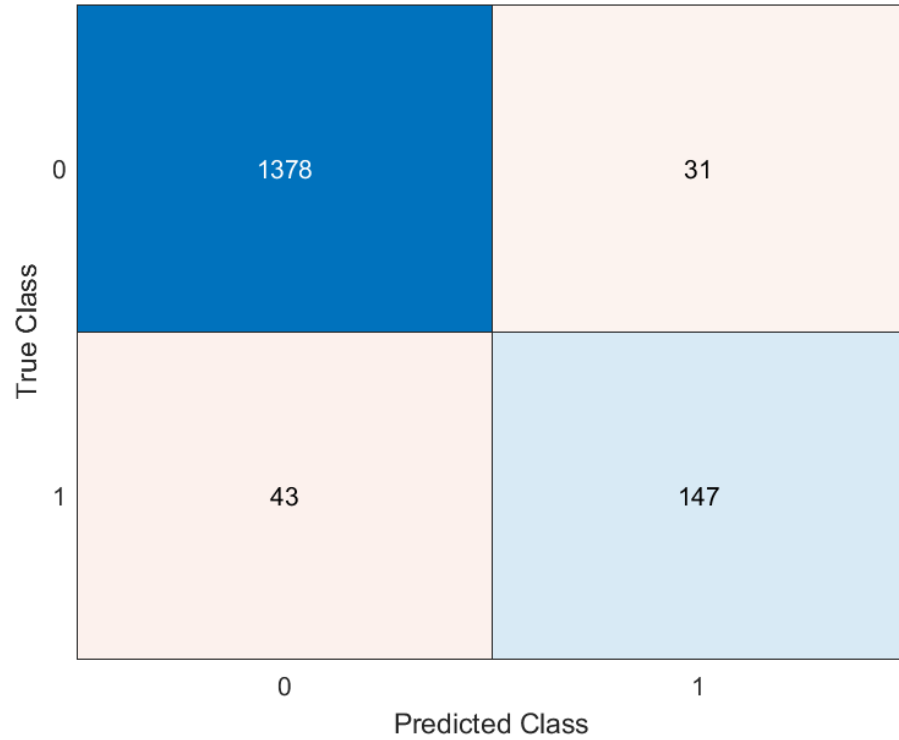
## 0.3 holdout



Accuracy: 95.7048  
AUC (Karar Ağacı): 0.8861  
Kappa Değeri (Karar Ağacı ile): 0.7812  
Duyarlılık: 0.9743  
Özgüllük: 0.8161  
Precision: 0.9775  
Recall: 0.974263  
F-Ölçümü: 0.975861

# Decision Tree

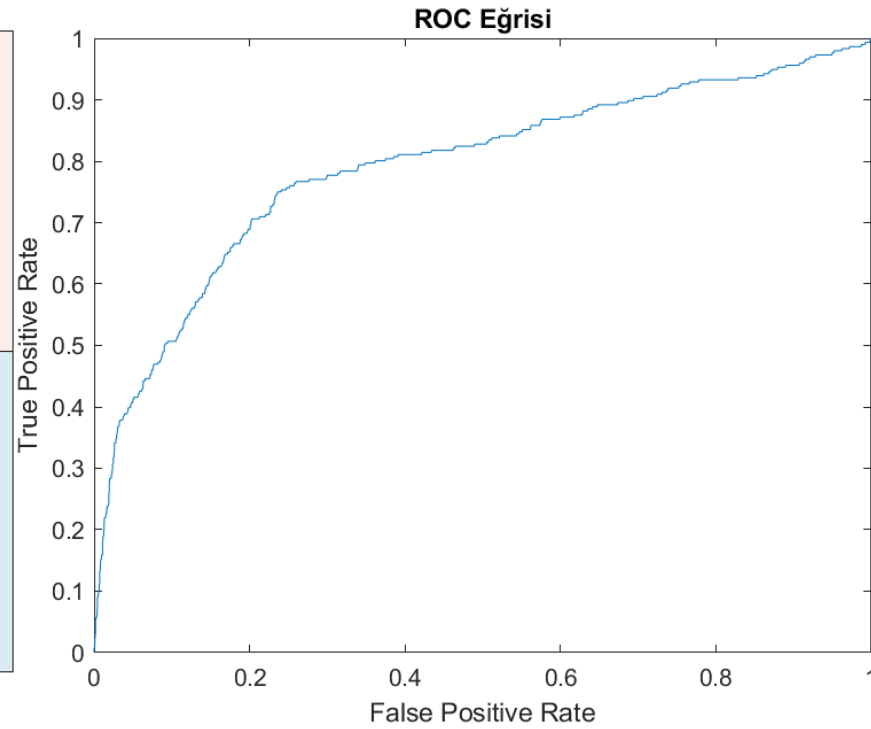
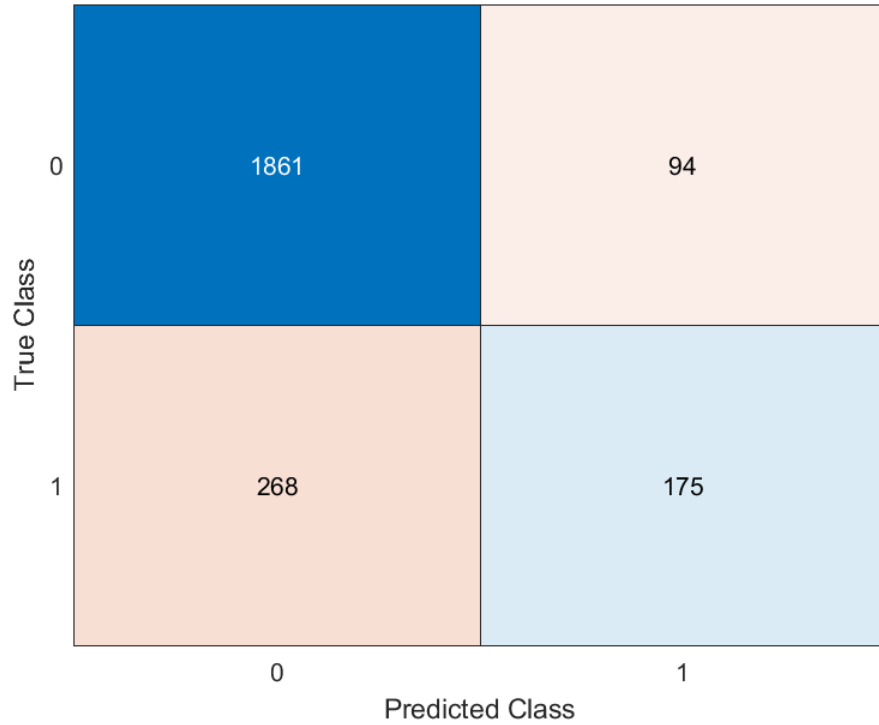
## 0.2 holdout



Accuracy: 95.3721  
AUC (Karar Ağacı): 0.8978  
Kappa Değeri (Karar Ağacı ile): 0.7728  
Duyarlılık: 0.9780  
Özgüllük: 0.7737  
Precision: 0.9697  
Recall: 0.977999  
F-Ölçümü: 0.973852

# Naive Bayes

## 0.3 holdout



Accuracy: 83.5279  
AUC:0.79052  
Naive Bayes için Kappa Değeri: 0.3635  
Duyarlılık: 0.9339  
Özgüllük: 0.3852  
Precision: 0.8739  
Recall: 0.933910  
F-Ölçümü: 0.902925