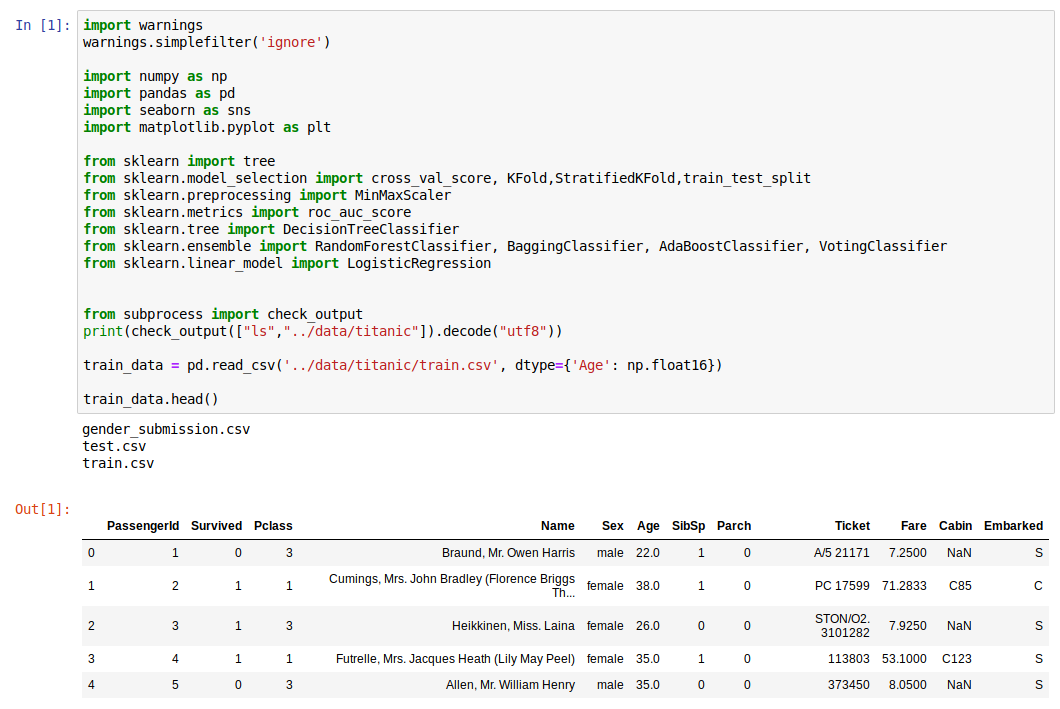
MÜCAHİT TAHA AZ

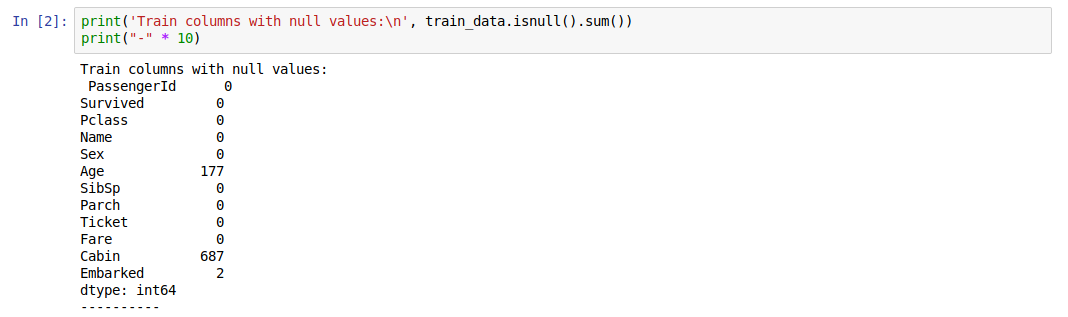
ENM - 593

Endüstriyel Veri Madenciliği Uygulamaları

**RAPOR – 3**

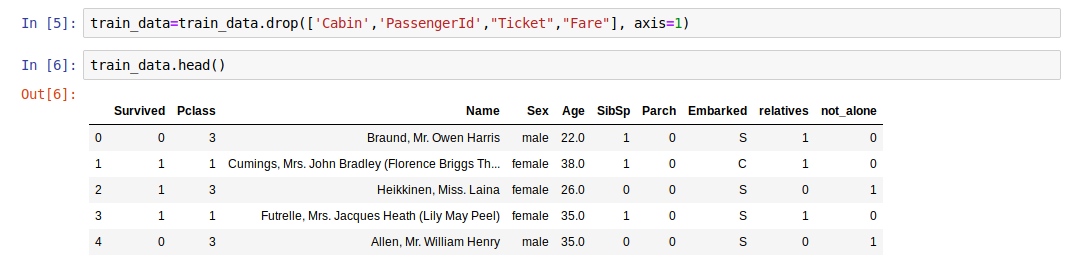
Gerekli kütüphaneler yüklenir. Veriseti yüklenir ve incelenir.





Gerekli düzenlemeler yapılır. (Veri temizleme ,eksikleri doldurma)

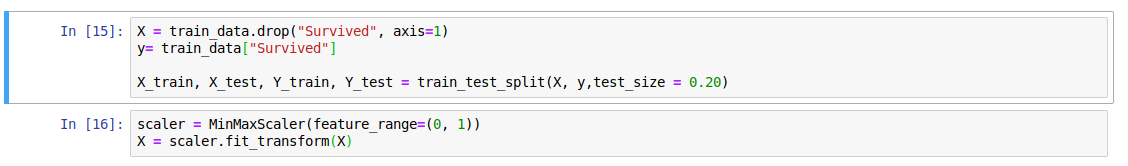






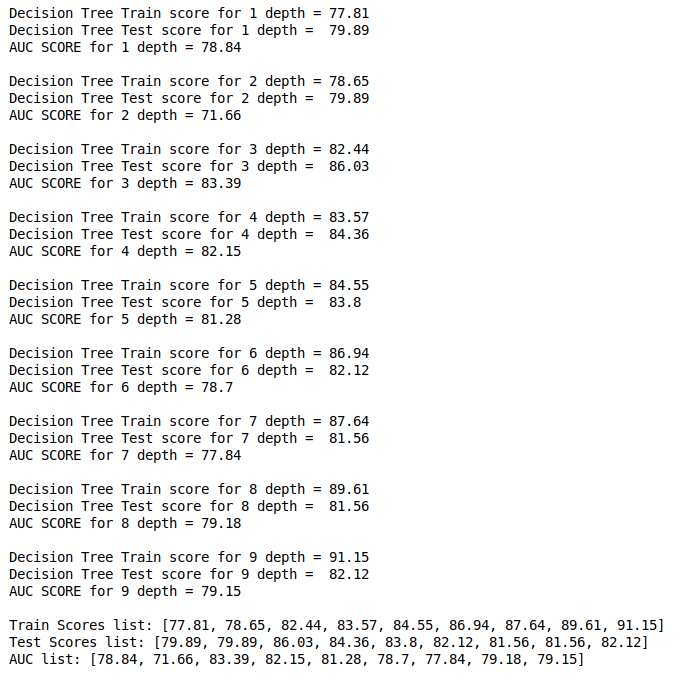


Veriseti test ve train olarak ayrıştırılır.



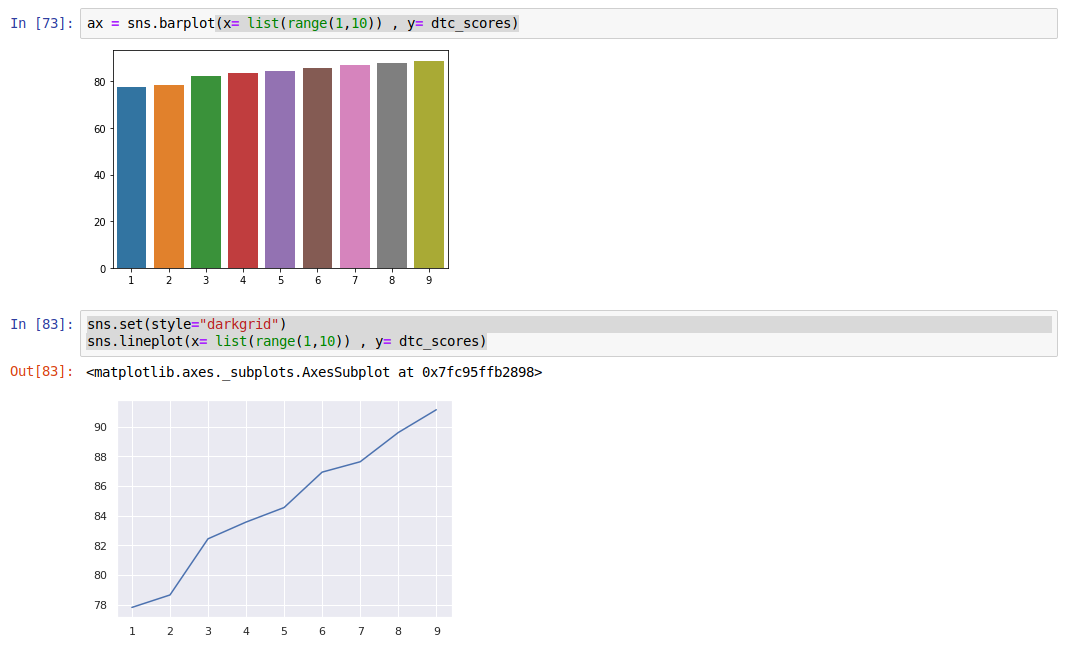
Tek bir karar ağacı ve farklı maksimum derinliklere göre çözüm yaptırılır.



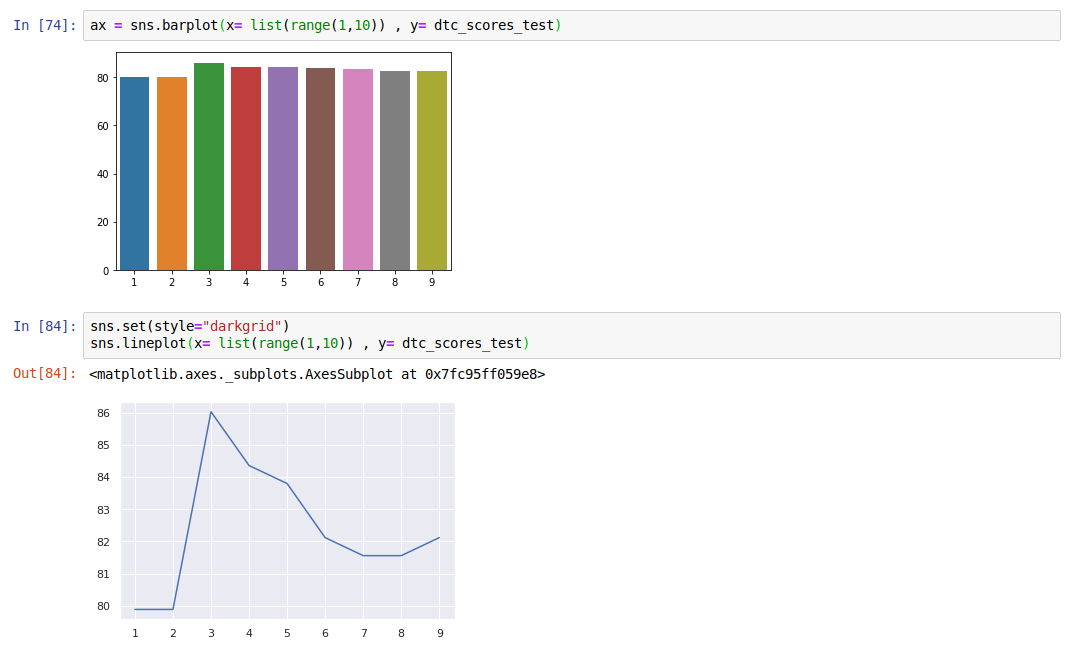


**Bu durumda max depth 3 olduğunda en yüksek skorları elde ettiğimiz görülür. Ayrıca overfitting yaşanmamıştır.**

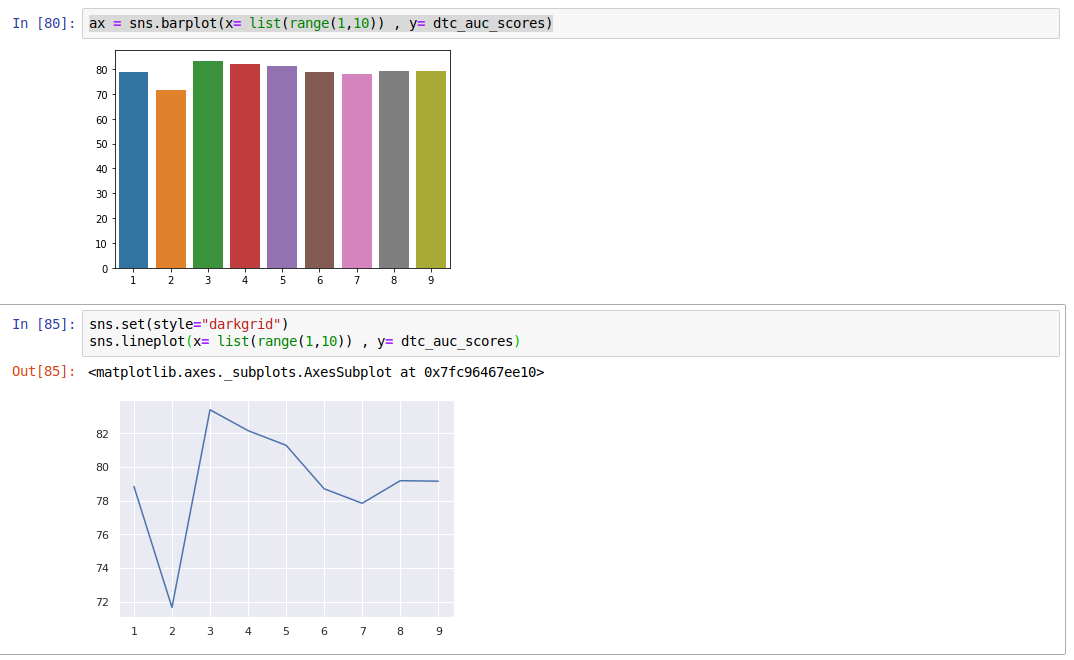
Train için grafikler



Test İçin Grafikler



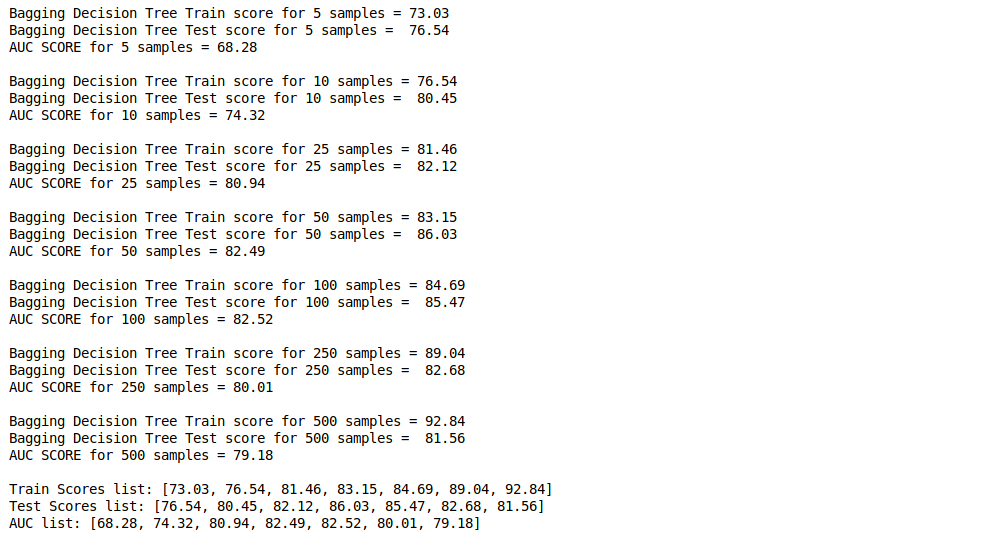
AUC Score İçin Grafikler



Bundan sonra ise Bagging kullanılarak Karar Ağacı çözümü yapılır.

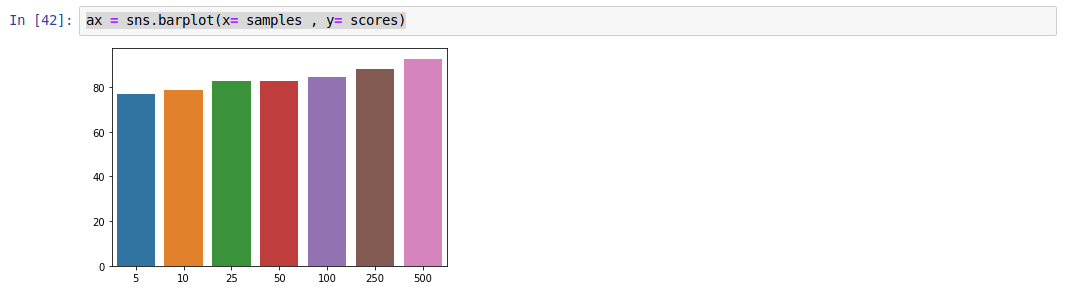


Bu çözümde farklı örneklem miktarları karşılaştırılmıştır.

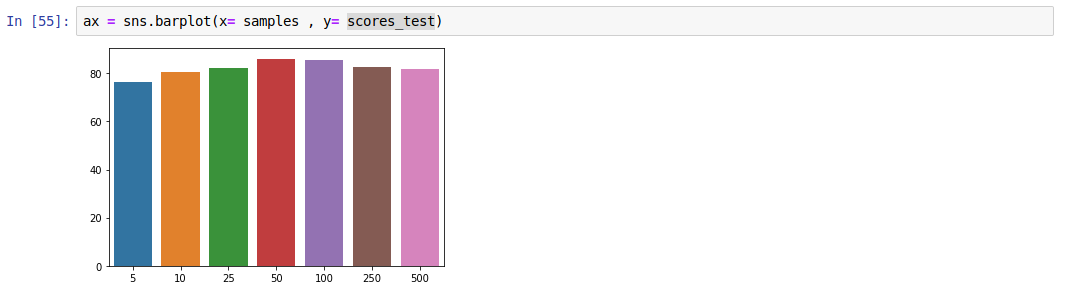


**En yüksek AUC Skoru 50 örneklemli karar ağacında bulunmuştur.**

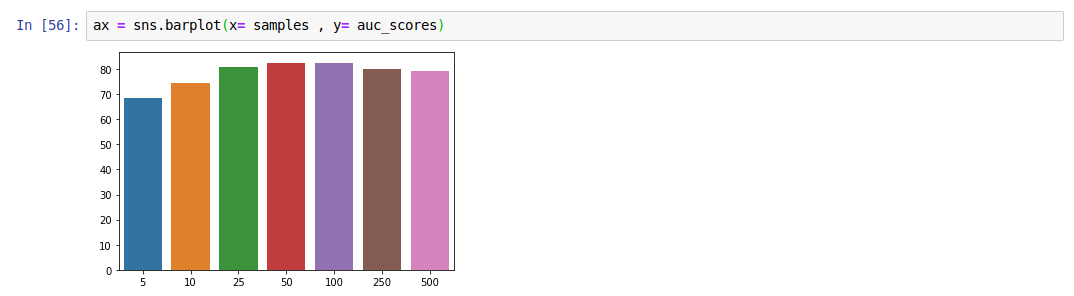
Train data için grafik



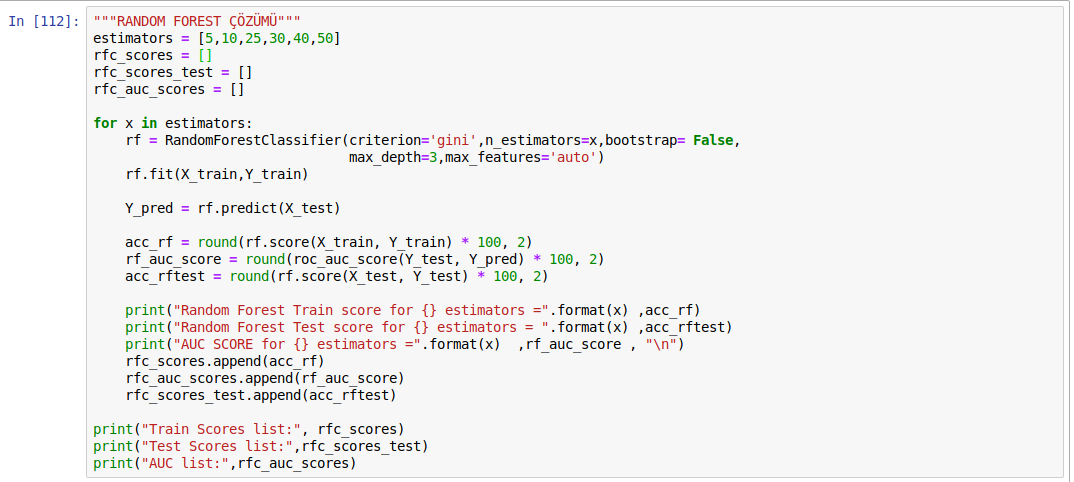
Test data için grafik

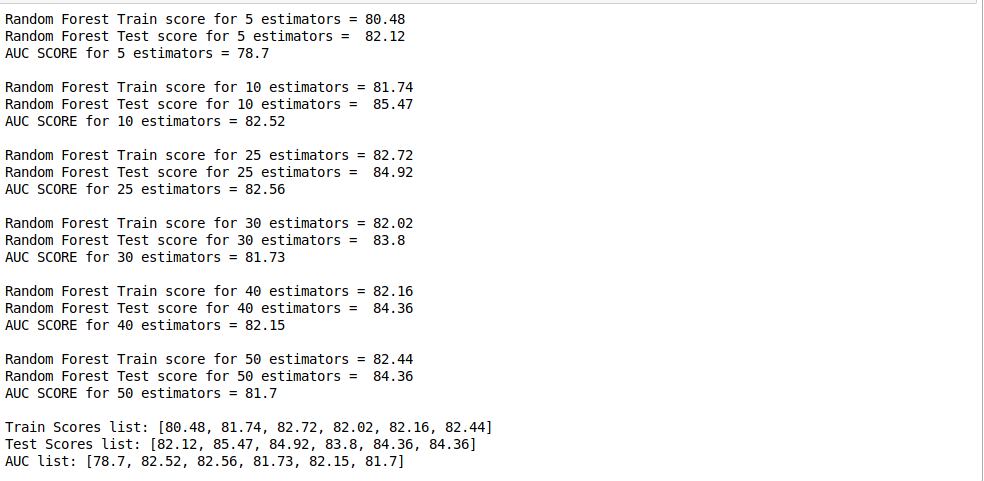


AUC Score için grafik



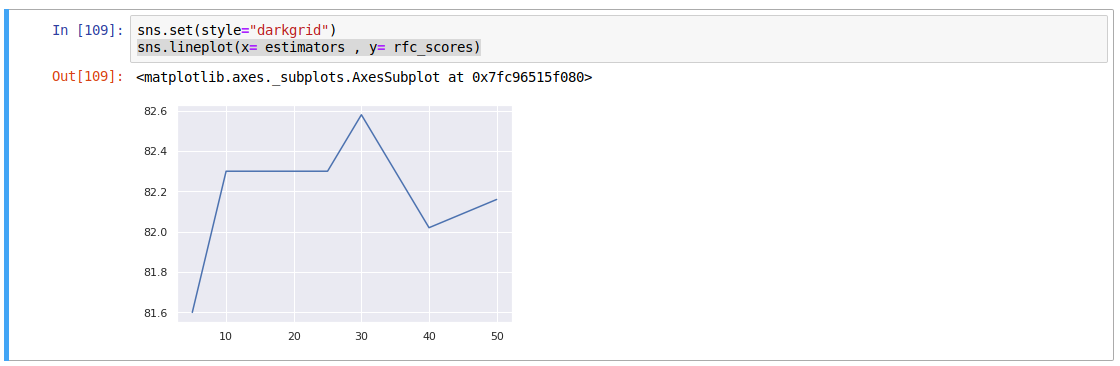
Rassal Orman çözümü yaptırılır. Max depth 3 olduğunda karar ağacının en yüksek AUC skoru görüldüğü için bu çözüm yapılırken max depth 3 olarak alınmıştır. Maksimum öznitelik sayısı otomatik olarak ayarlanmıştır. Boostsrap kapatılmıştır. Çözüm yapılırken ormandaki ağaç sayıları farklı olarak verilmiş olup bunlara göre değişim gözlenmiştir.



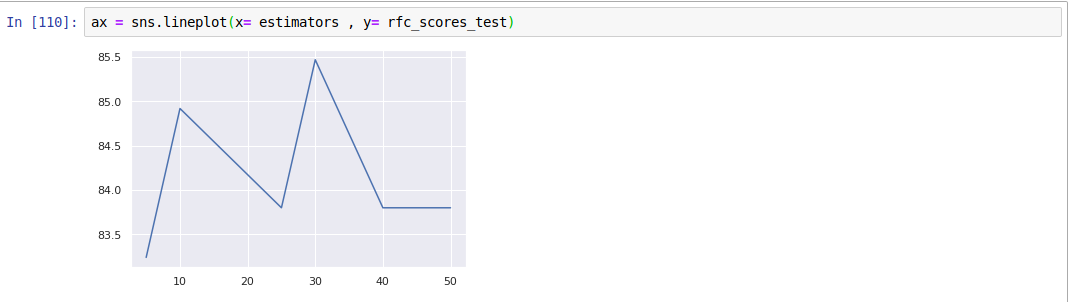


**En yüksek AUC skoru 25 ağaçlı rassal ormanda bulunmuştur.**

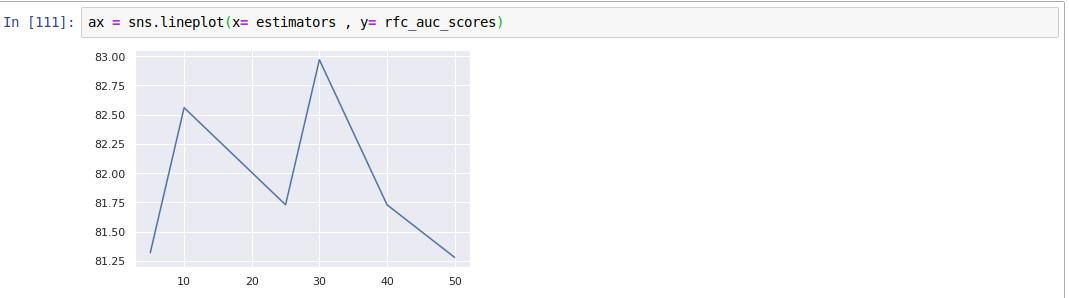
Train verisi için grafik



Test verisi için grafik

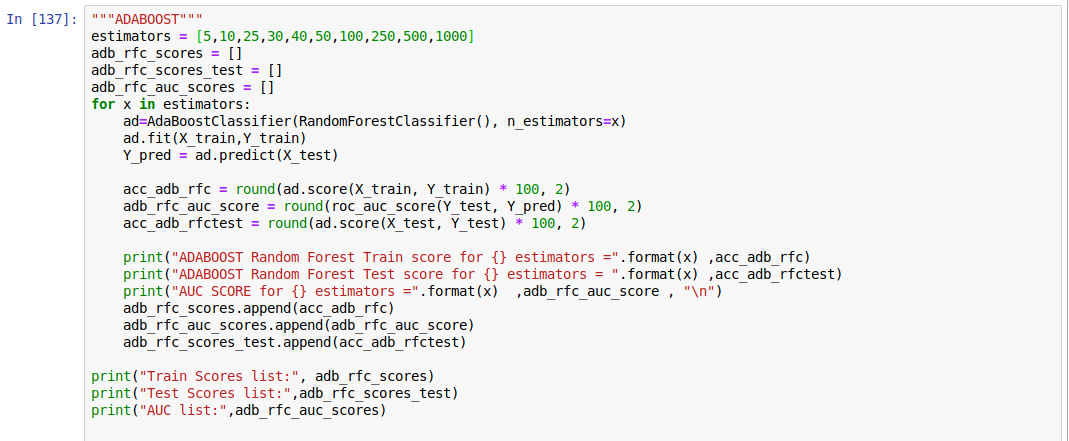


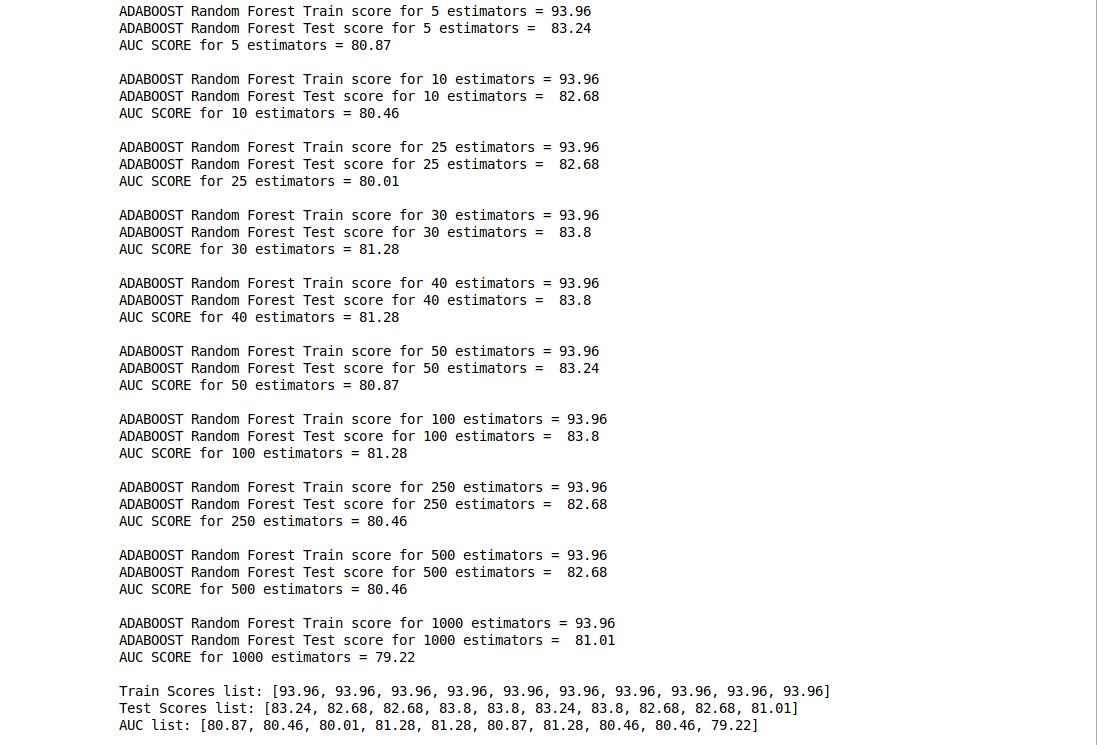
AUC skoru için grafik



Adaboost ile Rassal Orman çözümü

Bu çözümde de ormandaki ağaç sayısına göre değişim gözlenmiştir.

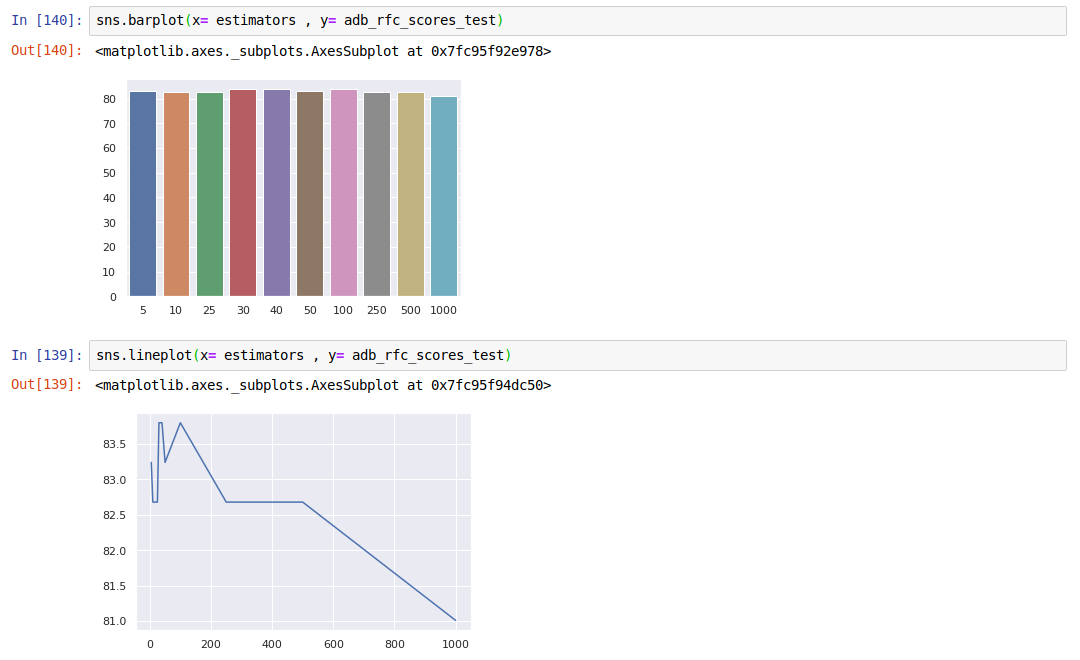




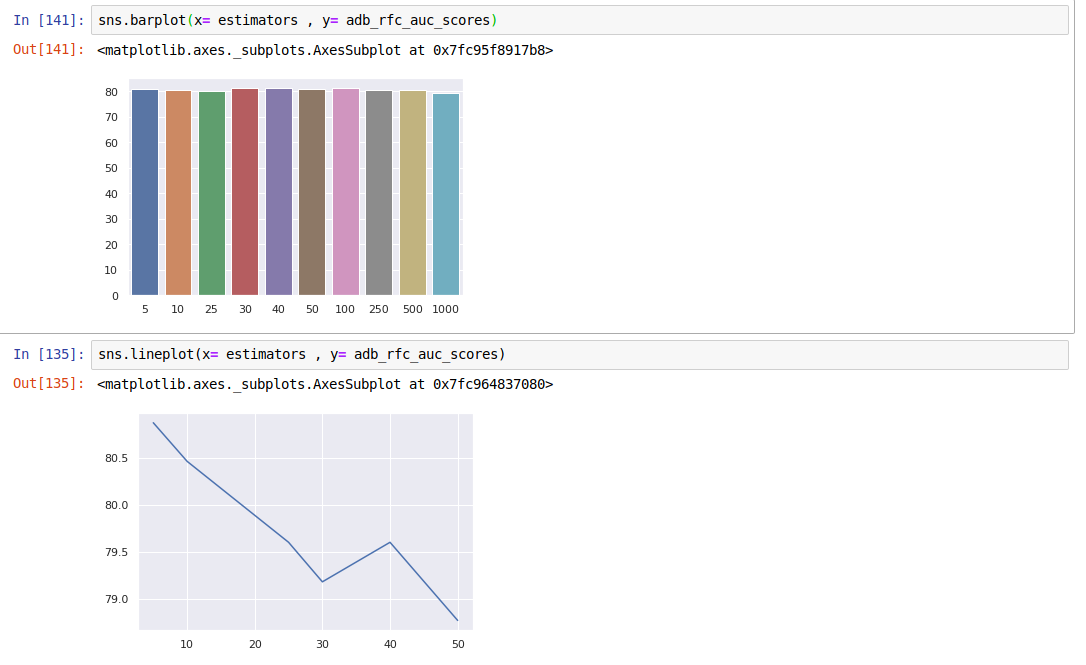
**Bu çözümde 30 ve 40 ağaçlı orman çözümleri en yüksek ve aynı sonucu vermiştir.**

Çözümde train verilerinin skoru her denemede aynı kalmıştır.

Test verisi için grafikler

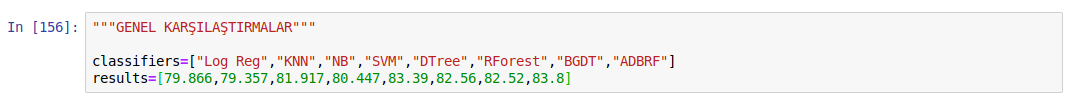


AUC Skoru için grafikler

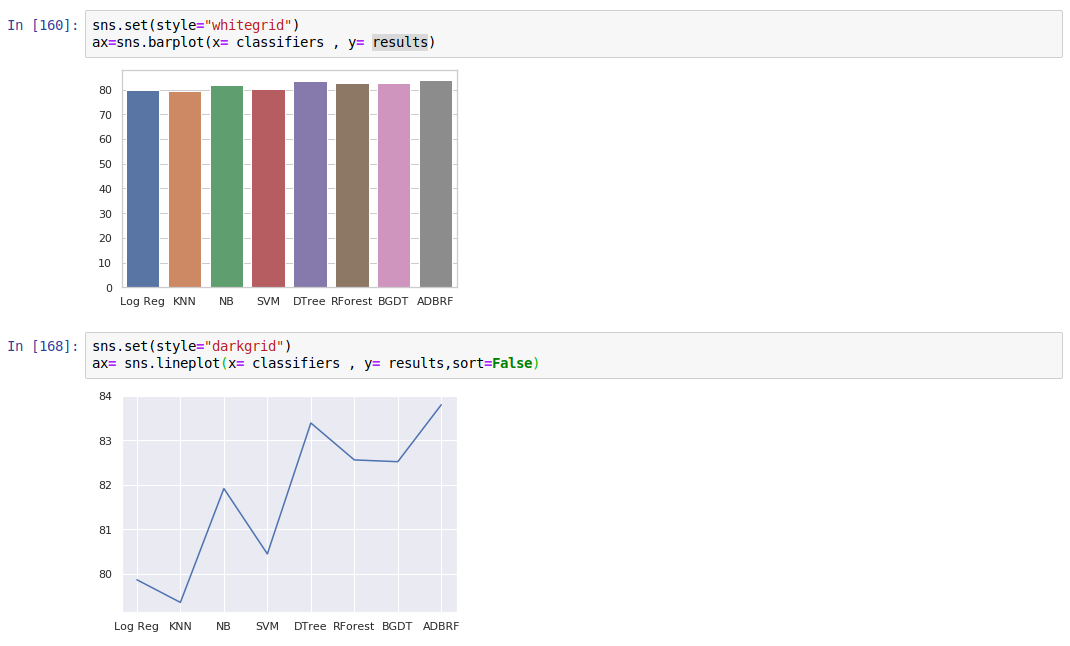


**GENEL KARŞILAŞTIRMALAR**

Bu kısımda şuana kadar eğitilmiş olan tüm algoritmaların en yüksek AUC skoru veren sonuçları seçilmiştir. Eski raporlardaki sonuçlarda eklenmiş olup genel bir karşılaştırma yapılmıştır.



Grafikler



Sonuçlar incelendiğinde Adaboostlu Rassal Orman en yüksek AUC skoruna sahiptir ve en başarılı algoritmadır.