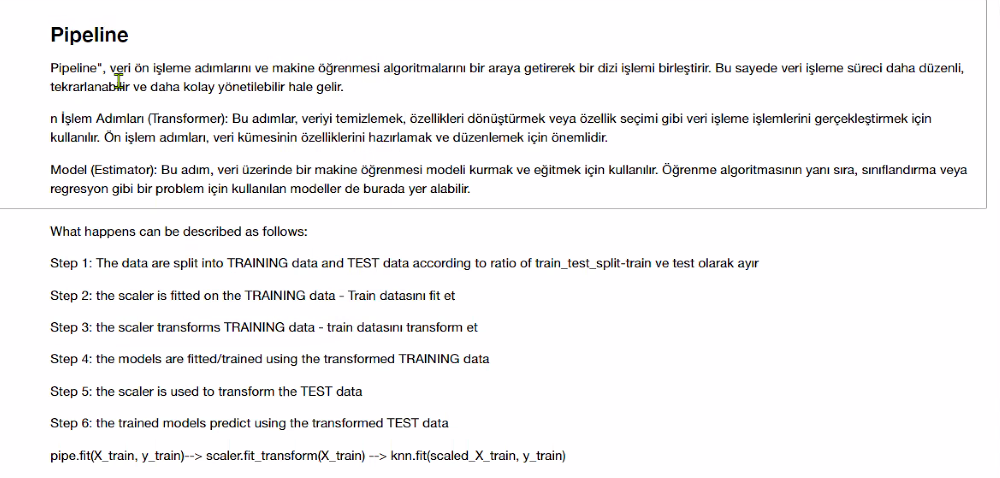
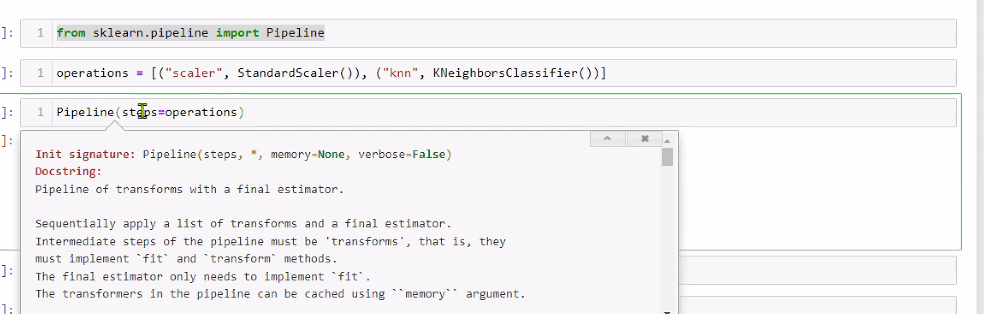
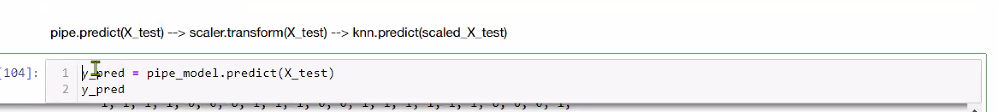
ML\_pipeline\_27\_07\_2023



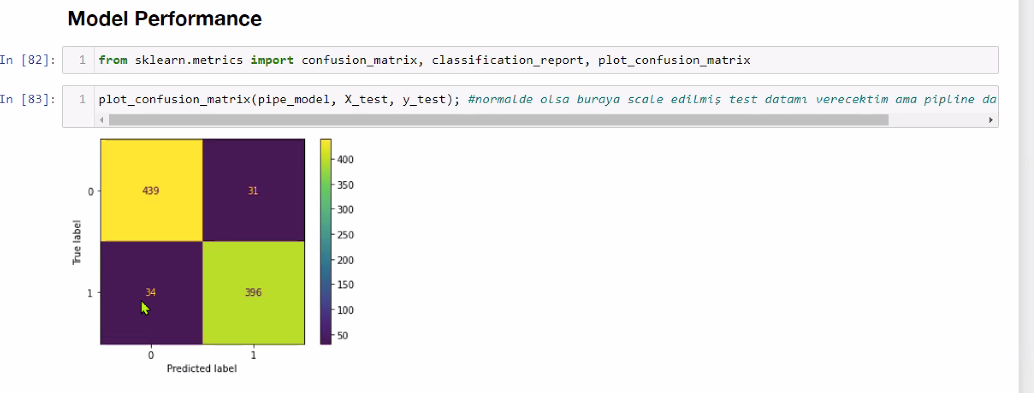


Pipeline", veri ön işleme adımlarını ve makine öğrenmesi algoritmalarını bir araya getirerek bir dizi işlemi birleştirir. Bu sayede veri işleme süreci daha düzenli, tekrarlanabilir ve daha kolay yönetilebilir hale gelir.  
n İşlem Adımları (Transformer): Bu adımlar, veriyi temizlemek, özellikleri dönüştürmek veya özellik seçimi gibi veri işleme işlemlerini gerçekleştirmek için kullanılır. Ön işlem adımları, veri kümesinin özelliklerini hazırlamak ve düzenlemek için önemlidir.  
Model (Estimator): Bu adım, veri üzerinde bir makine öğrenmesi modeli kurmak ve eğitmek için kullanılır. Öğrenme algoritmasının yanı sıra, sınıflandırma veya regresyon gibi bir problem için kullanılan modeller de burada yer alabilir.

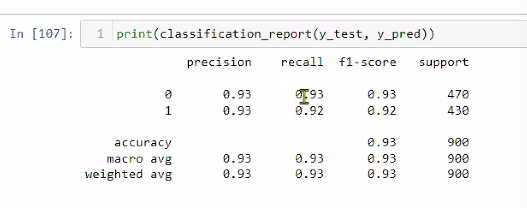




Scale islemi fiziken yapmadim. Pipeline benim icin yapiyor. Scale edilmis x testi ppredict yapiyor boylece otomatik olarak.

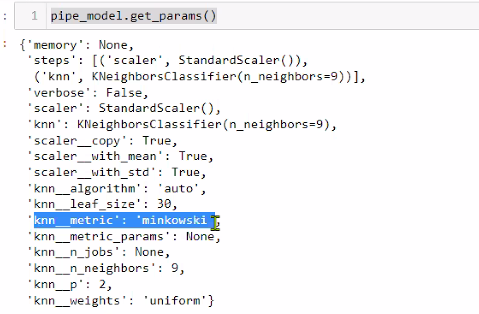


Toplam 65 yanlisim var.



Pipeline; Data leakage önlediği gibi tek bir kod satırıyla birden fazla adımın tek tek yapılmasını ve hata riskinin azaltılmasını sağlar. Ayrıca, modele uygulanacak değişikliklerin tek bir yerden yapılmasını ve kodun daha okunabilir ve anlaşılır olmasını sağlar.





get\_params sonrası gelen isimlendirmeler her modele göre farklı gelecek. pipeline ı her algoritmada kullanabiliriz.

O modelin icerisnde nasil vermem gerekiyor bunlari bana gosteriyor.

Sadece scaler yapmasi icin ve sadece modeli calistirmasi icin yapilacaklar:



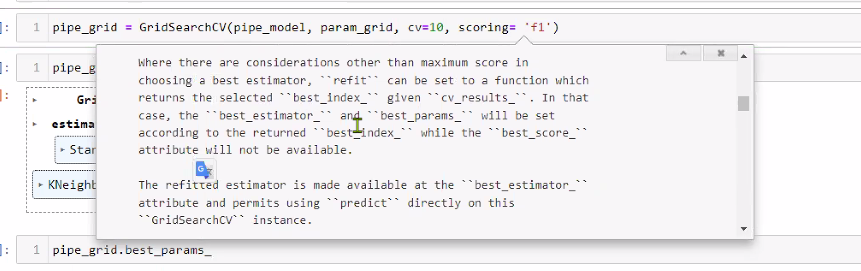


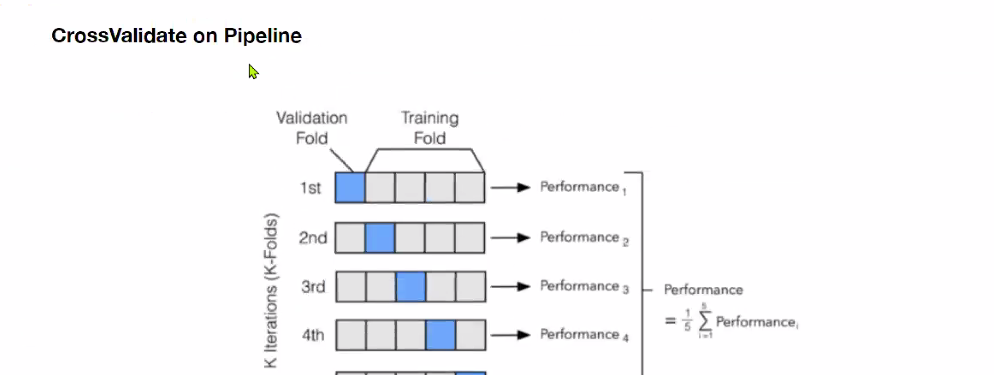
Param\_grid ile sadece komsu sayisini degerlendirttiriyorum.

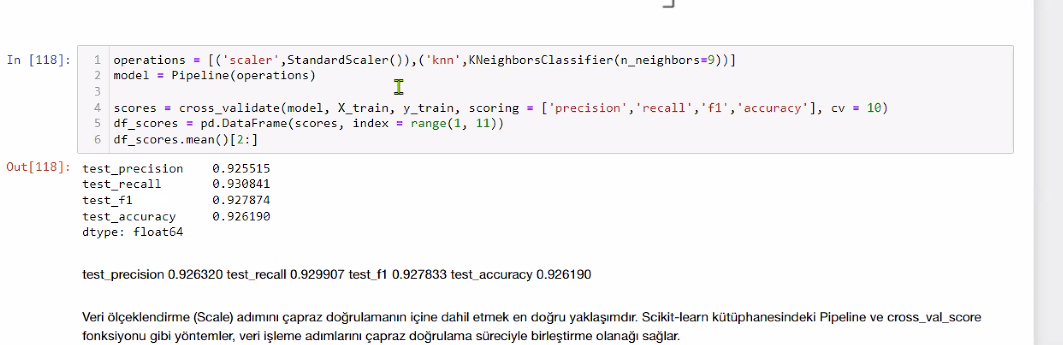
Operation olarak da sadece modelim olarak knn ve scale islemi olarak da standard scale

Scoring de f1 diyorum. 1’den fazla score verebiliyoruz icine aslinda ama **refit ile. Ama yaygin degil**

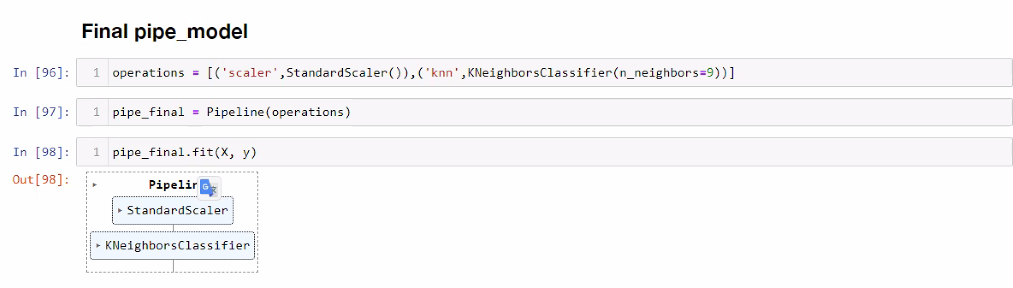
**Unbalanced data da f1, balanced datada accuracy’e bak**





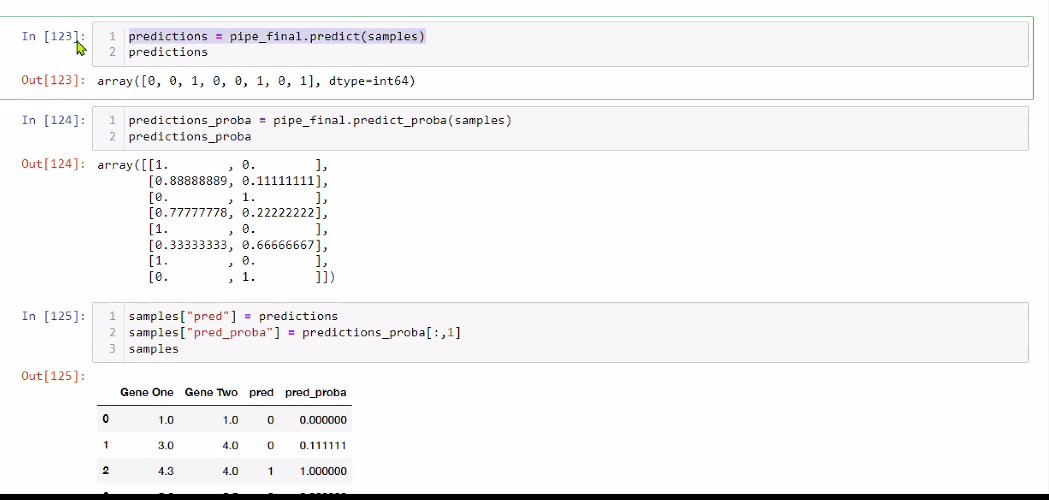


9 ile ilerledik.



Modelimi final olarak, orda aldigim degerlerle kuruyorum.

Veri ölçeklendirme (Scale) adımını çapraz doğrulamanın içine dahil etmek en doğru yaklaşımdır. Scikit-learn kütüphanesindeki Pipeline ve cross\_val\_score fonksiyonu gibi yöntemler, veri işleme adımlarını çapraz doğrulama süreciyle birleştirme olanağı sağlar.



Scale yapmiyoruz sample icin de pipe modelde.

