**R2 degeri :** bu tahmini yapmak yerine direk ortalamalarini alinsaydi tahmin icin bu sizin kurdugunuz modelden daha mi iyi olur daha mi kotu olur. – ise cok kotudur, + cok iyidir. Tek bir feature ile kullanilir ve bu yetersiz kalabilir o yuzden adjusted r sequare kullanilir regresyonda.

**Adjusted r sequare :** birden fazla feature varsa kullanilabilir.

Bunlarin yerine hata metriklerini kullanmak cok daha mantiklidir.

**Hata Merikleri :**

* R-Squared (R2)
* Mean squared error : hatalarin karalerinin ortalamasi
  + Mse yi min yapacak egim ve intercept noktalarinin turevleri 0’a esitlenir.
* Mean absolute percentage error : veriden tahminlerimizi cikarip, gercek veriye boleriz. Ardindan bunlarin mutlak degerlerini toplayip 100’le carpip veri sayisina boleriz. Outlier veri coksa kullanilir, ama veriden sifirlarin temizlenmesi gerek, yoksa paydada 0 olur.

**K-Fold Cross Validation :** modelin tum verilerini test etmesini sagliyoruz.

**Stratified K-Fold Cross Validation :** siniflandirma problemlerinde kullaniriz. Her fold icin esit sayida sinif degerleri bulunmaktadir. Siniflarin esit dagilmasini saglamaktadir.

**Shuffle Split Cross Validation :** Rastgele sectigi bir foldu kaldiriyor, siliyor.

**Veri Sizintisi :**

**Grid Search Cross Validation :** random forest, kesikli degerlerin kombinasyonlarini aliyor

**Randomized search cross validation :** olasilik dagilimi veriliyor ve o degerler uzerinde kombinasyonlari aliyor. Arama yaptiginiz aralik daha genisse ve bu aralikta daha cok deper varsa kullanilir. Iterasyon sayisini 1000 yaparsak bir parametre icin 1000’den fazla deger deneyecektir.

**Uniform distribution :** a ve b degeri arasindaki esit olasiliklarla aralik seicyor. A loc olarak veriliyor, b ise scale oluyor.