* **Bias**, modelin gerçek değerden ne kadar sapma gösterdiğini ifade eder.**Yüksek bias, modelin aşırı basit olduğu ve verilerin karmaşıklığını yeterince yansıtamadığı anlamına gelir**.**Buna karşılık, düşük bias, modelin verileri iyi yansıttığı anlamına gelir.**
* Overfitting: modelin eğitim verilerine **aşırı uyduğu** ve **farklı veri kümesinde kötü performans gösterebileceği**anlamına geliyor.
* **Varyans**, modelin aynı veri kümesindeki farklı örnekler üzerinde nasıl performans gösterdiğini ifade eder.Yüksek varyans, **modelin eğitim verilerine aşırı uyduğu ve farklı veri kümesinde kötü performans gösterebileceği anlamına gelir.**Düşük varyans, **modelin farklı veri kümesinde iyi performans gösterebileceği anlamına gelir**
* **Bias-variance trade off**(*Bias-Varyans Dengelemesi)*, bu iki faktör arasındaki dengeyi bulmayı amaçlıyor.Modelin yüksek bias ve düşük varyansı olması,**aşırı basit**bir modeli ifade ederken,düşük bias ve yüksek varyans**, aşırı karmaşık**bir modeli ifade eder.
* İdeal durumda, modelin hem düşük bias hem de düşük varyansı olmalıdır. Fakat gerçek hayattaki datalarda genellikle bu mümkün olamayacağından modeller arasında bir seçim yapmak gerekebilir. Bu nedenle, model seçiminde **bias-variance trade off**udikkate alıyoruz.
* total error = bias + variance + irreducible error  
  formulasyonunda ie değiştirilemezdir, totral erorde sabit olduğundan bias artarsa varyans düşer, bias azalırsa varyans artar
* matematiksel linearlik ile modellemdeki linearlik ayrı aman dikkat. Bir feature bir katsayı ile ifade ediliyorsa lineer, birden fazla katsayı ile ifade ediliyorsa non-lineer regresyonla ifade edilir.
* Aslında tüm mesele, modelin train datası ile eğitiminde öyle bir ayrıntı seviyesinde kalmalı ki eğitim sonunda dışardan gelecek her türlü gerçek datayı da kabul edilebilir optimum bir hata ile tanımlayabilmeli.
* Train datasında 30 değeri olmayabilir. model farklı independent feature değerleriyle (x-trainlerle) eğitildi ve o değerlerden datayı tanımlayan bir fonksiyon buldu. İşte o fonksiyona test datasından 30 değerini koyduğumuzda bir predict yapıyor. ve bu predictionın hatası en düşük bir prediction olmasını bekliyoruz. neden? çünkü datayı tanımlayan en doğru polinomial fonksiyonu bulmuştu.

çizelge içeren bir resim

Açıklama otomatik olarak oluşturuldu

Train datası ezberlendiği için r2 değeri test datasında negatif gelir.