

Graident Descent, hataları minimum noktaya getirmeye çalışır.

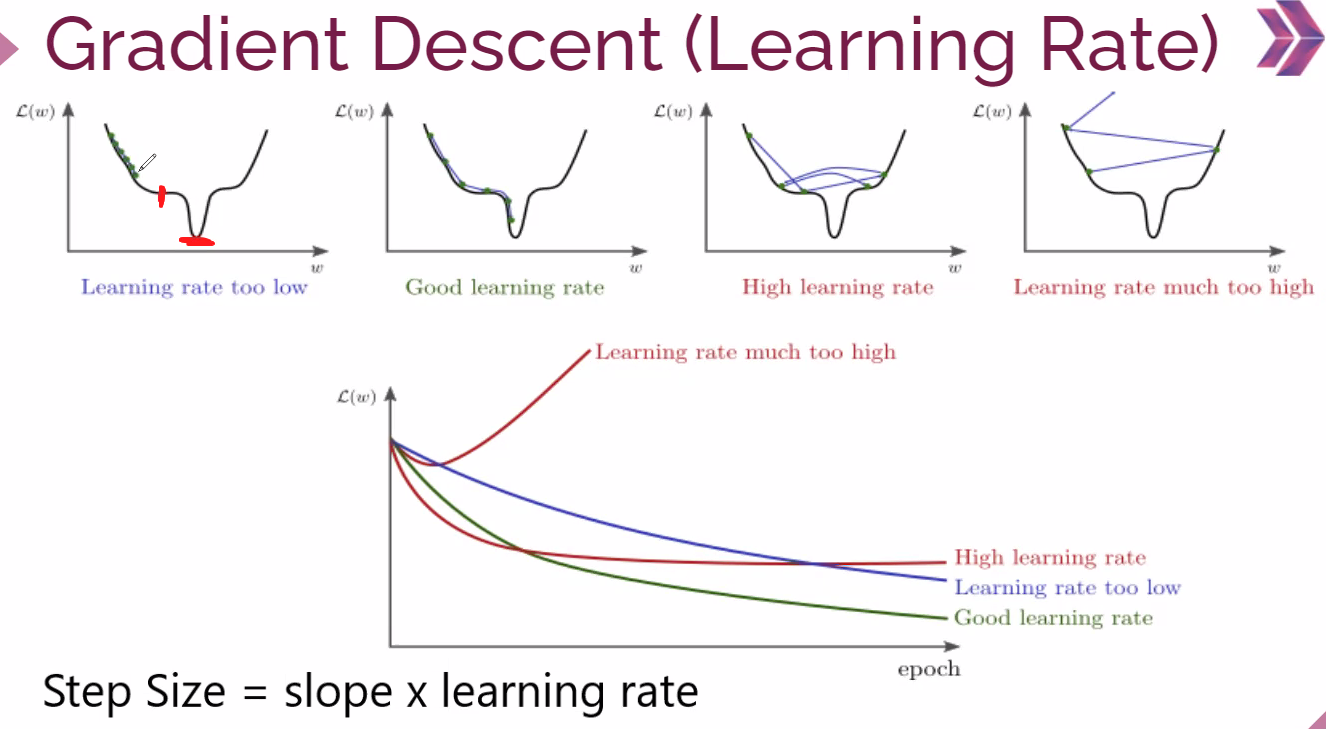
Gradient Descent işlemini yapan optimizer parametresidir.

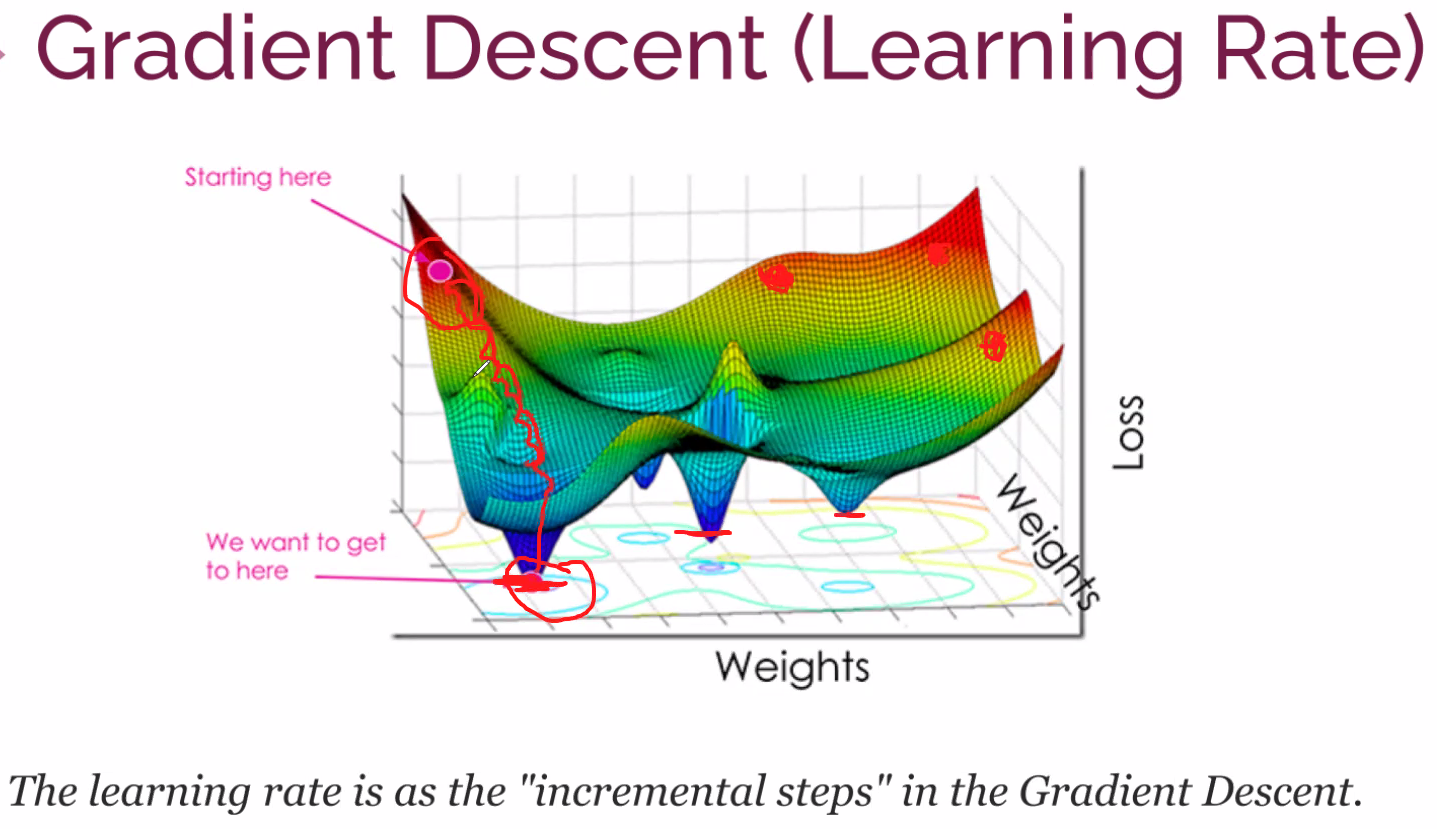
Learning Rate her adımda ne kadar öğrenmesi gerektiğini söylüyoruz. Yani minimum noktaya ulaşırken adımının büyüklüğüne etki eden şey learning rate dir. Her bir noktada fonksiyonun türevi alınıyor. Türev bize eğimi veriyor. Bu eğim modelin bize bir sonraki stepte kontrol edeceği yeri veriyor. eğim (slop) ne kadar fazla olursa adım o kadar büyük, ne kadar az ise adım o kadar küçük oluyor.

Learning rate i çok küçük seçersek çok minik adımlarla gideriz. Adımlarımız çok küçük olursa eğitim çok uzun sürebilir. Aynı zamanda global mean’i de bulamayabiliz. Başka minimun değerlerde takılır kalır (durur) sistem.

Learning rate büyük olursa 3. Şekildeki gibi dip nokta üzerinden devamlı atlayarak onu bir türlü bulamayabilir.

İdeal durum 2. Şekildeki gibi optimum learning rate değeridir. (adımları optimum seviyede tutacak değer)





BAtch (Vanilla) Gradient Descent: Tüm train setini modele sokuyor. Aldığı hata değerlerinin ortalamasına göre. Minimum noktayı bulmaya çalışıyor. Bu noktalar için bulduğu ağırlık değerlerini güncelleyipi modelini eğitmeye çalışıyor. Hepsi için hata değerleri buluyor. Gradient descent yapıyor optimizasyon çalıştıyor.

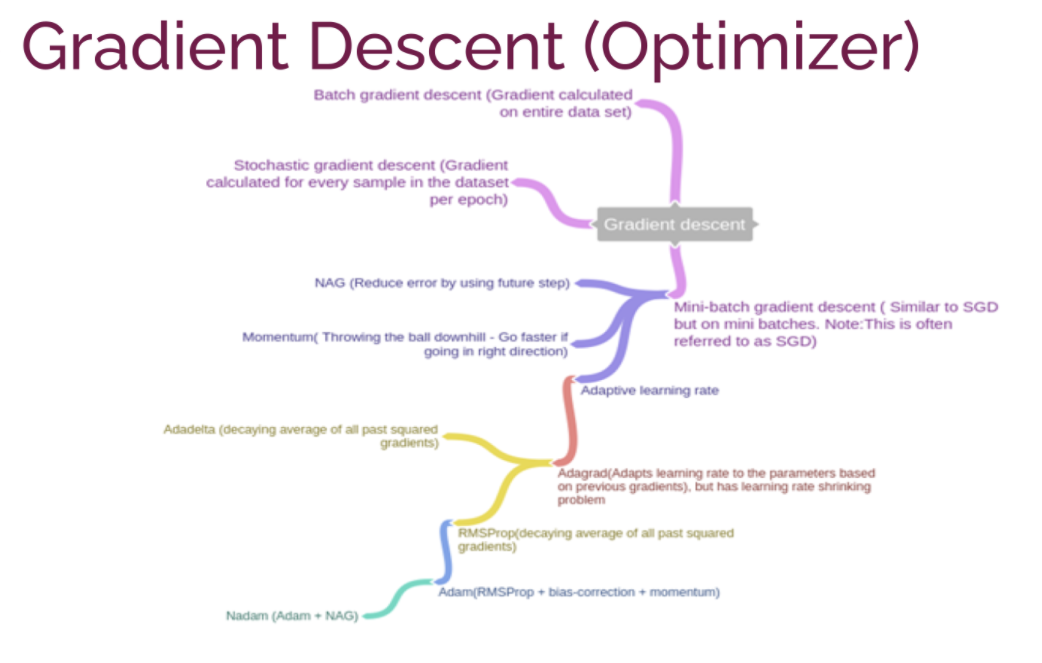
Stochastic Gradient: yukardaki işlemi her örnek için ayrı ayrı yapıyor. Yani her adımda rastgele alınan bir veri üzerinde tüm işlemleri yeniden yapar.

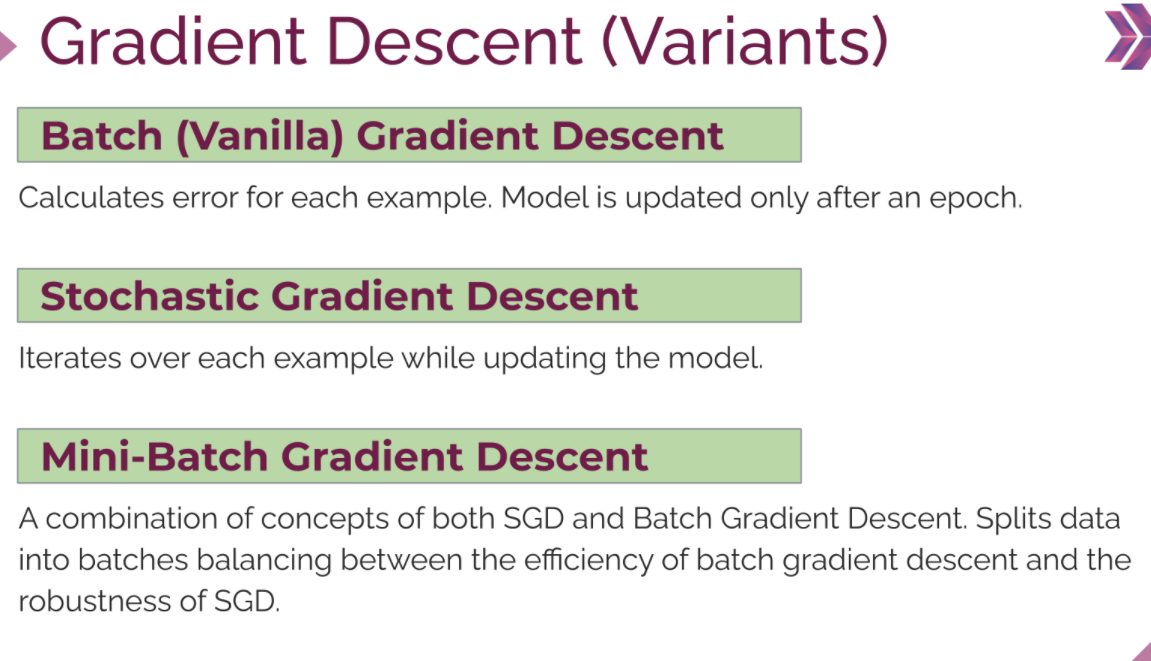
Mini-Batch Gradient Descent : Bazı yerlerde stokastik gradient olarak da geçiyor. Çünkü stokastic in gelişmiş halidir. 3’erli 5’erli 10’arlı örnek grupları halinde gradient descent işlemlerini uyguluyor.

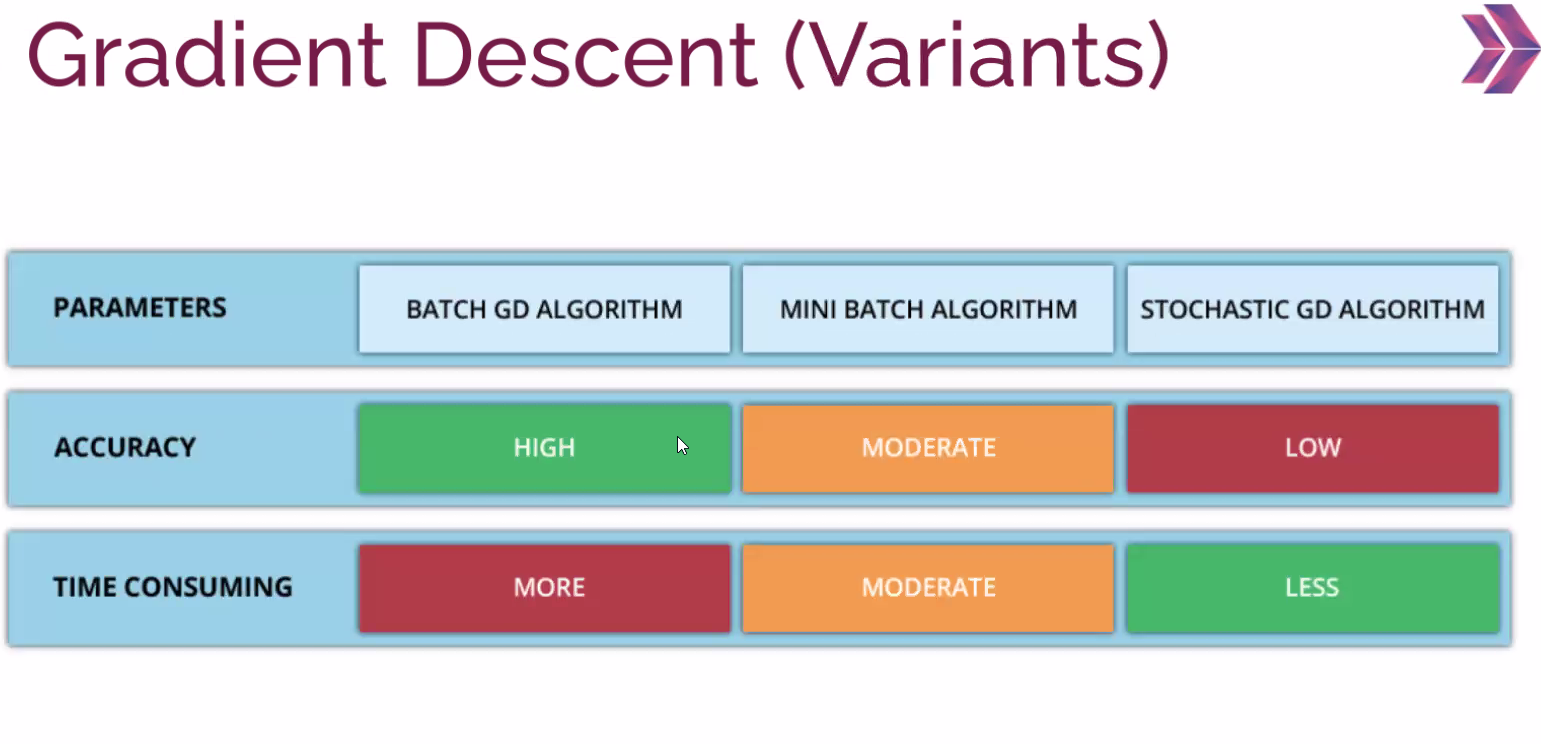
Stochastic, her bir gözlem için yukardaki haritayı tekrar tekrar çıkartıyor. İşlem çok uzun sürüyor.

Batch te ise hızlı oluyor ama hatayı stochastic kadar başarılı aşağıya çekemiyor.

O yüzden genelde mini-batch tercih ediliyor

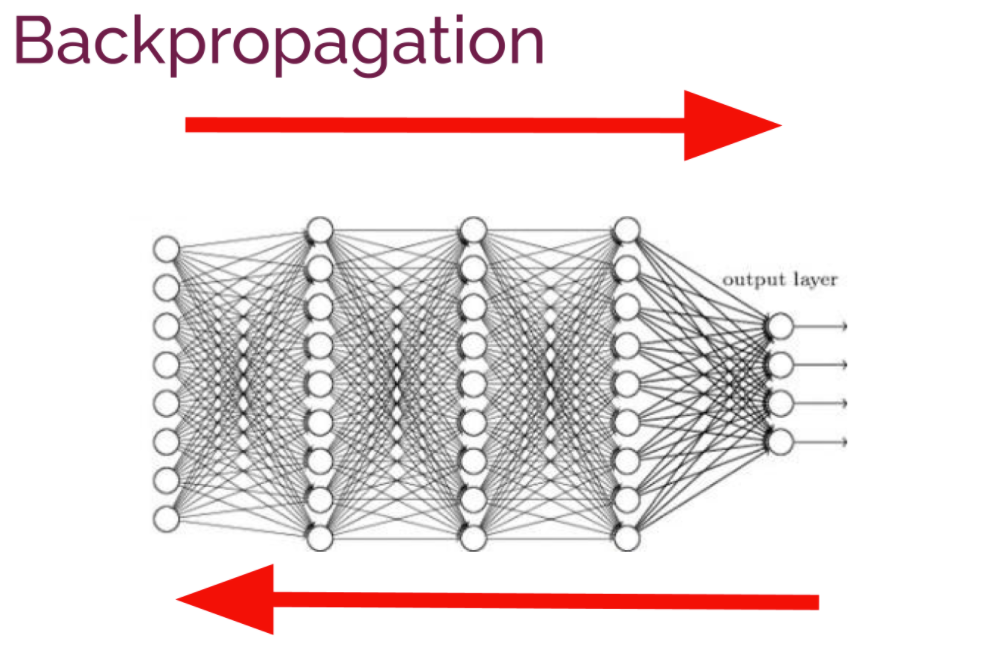




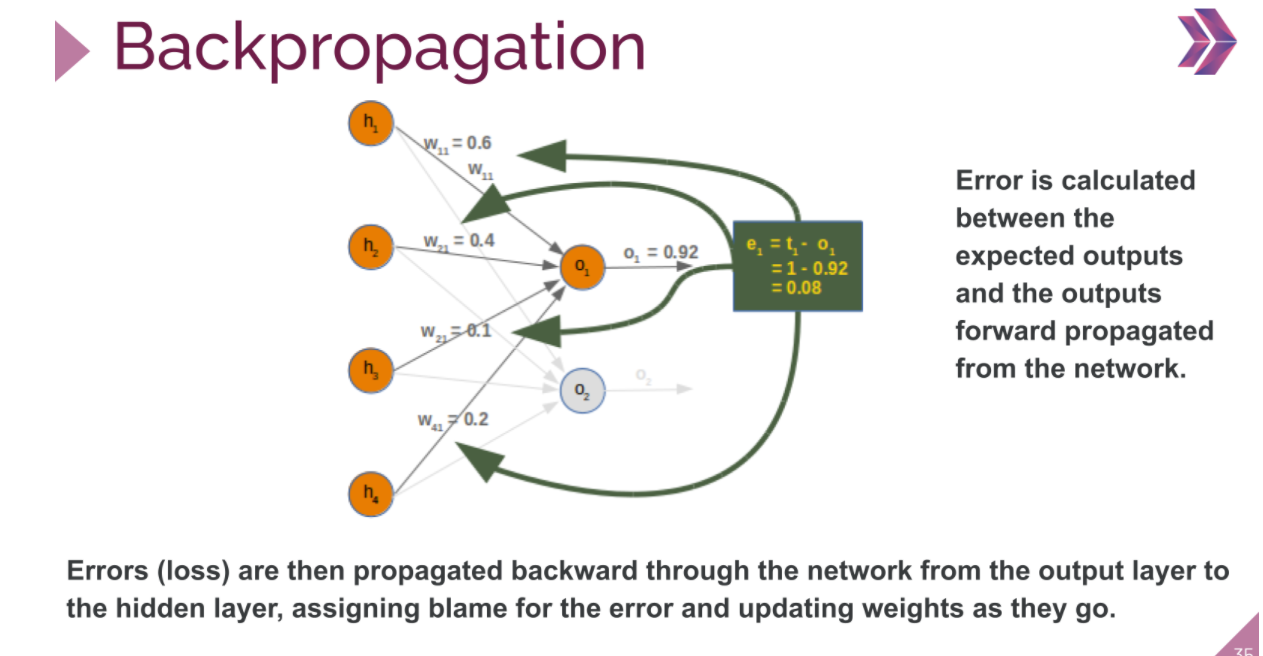


**BACKPROPAGATION**

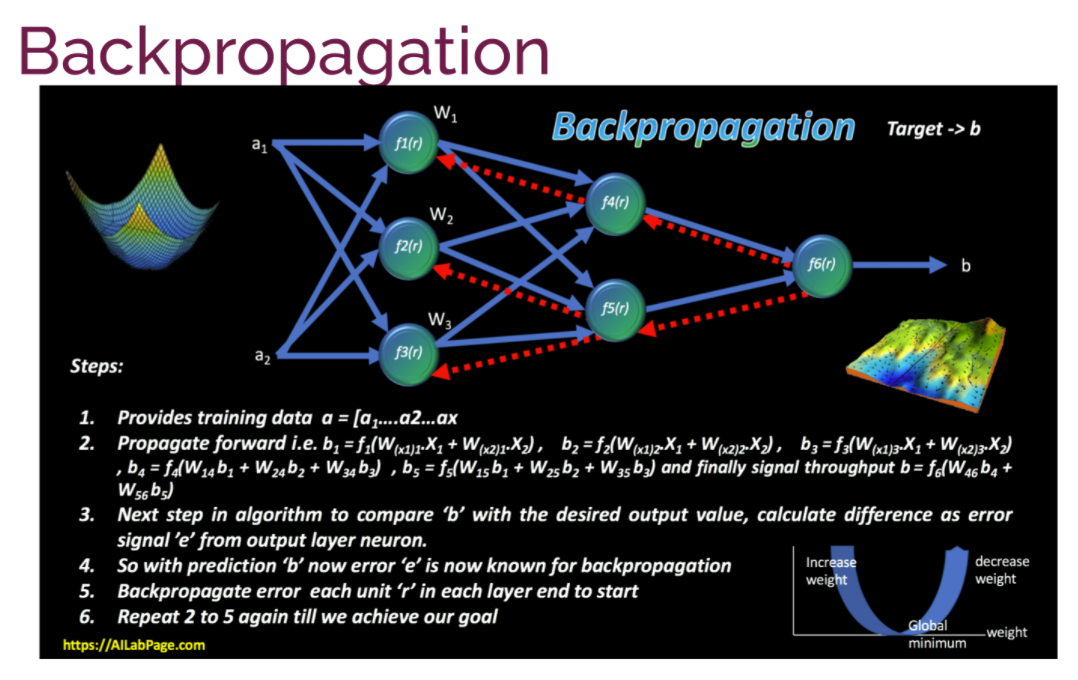
Sistemi ileri doğru besliyoruz. Output alıyoruz. Outputlardan hataları tespit ediyoruz. Bu hatalara göre ağırlıkları geriye doğru güncelliyoruz.

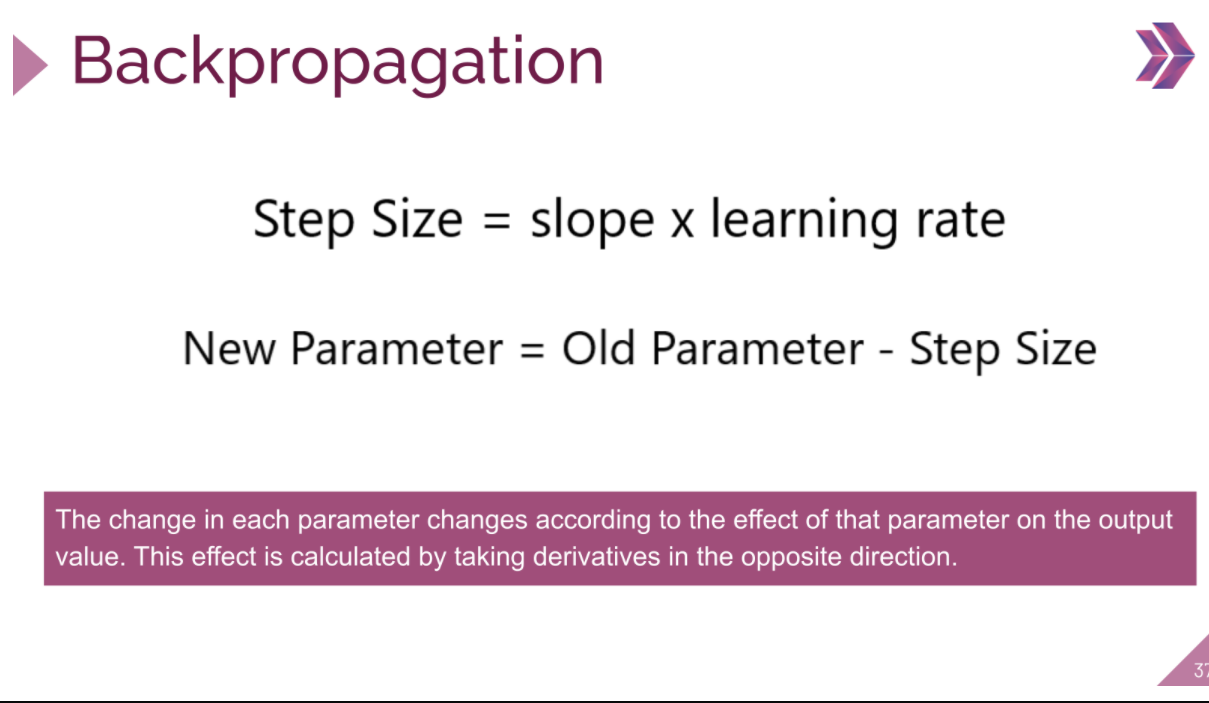
****

1 olması beklenirken 0.92 çıkmış. 1 - 0.92 = 0.08



1. Trainin datası model için sağlanıyor
2. Model ileri doğru besleniyor.
3. Sonuç alındı. Gerçek değerle karşılaştırıldı.
4. Artık elimizde bir hata değer var.
5. O hata değerine göre sondan başa doğru tüm parametrelerimiz ve ağırlık değerlerimiz güncellendi.
6. Tekrar başa dönülerek ileri doğru beseleme başladı.





Herhangi bir nöronun çıkışı Y olsun..

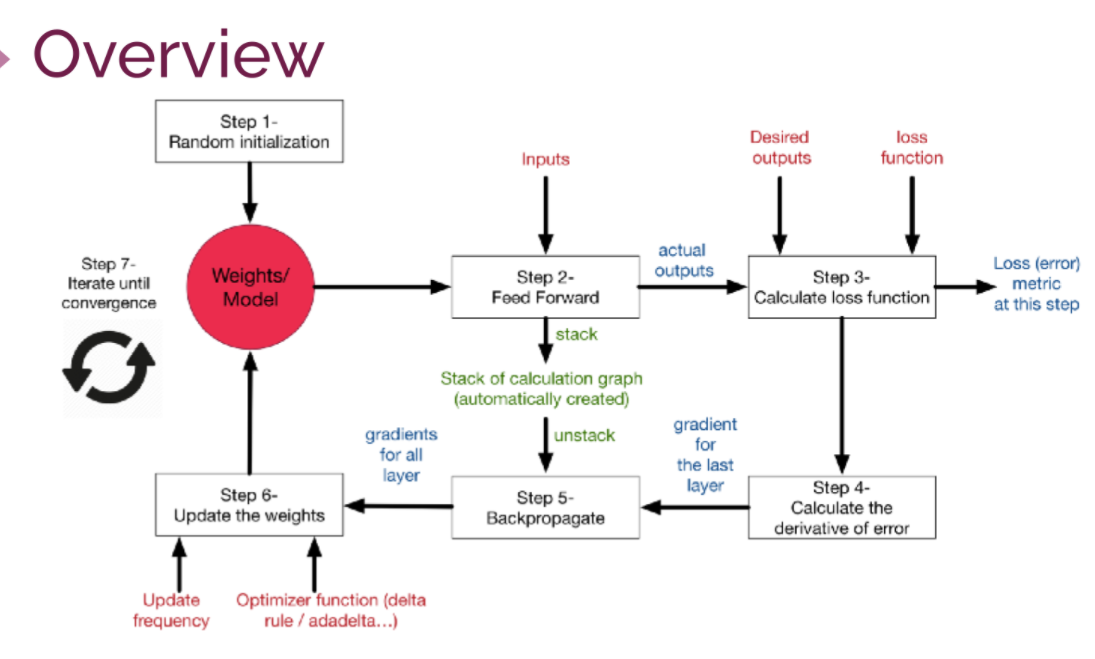
Bu Y’ye etki edenler: Bias (b), ağırlık (w) ve X

Eski parametreyi güncellerken, her birindeki 1 birimlik değişimin (nöronun) Y çıkışını ne kadar etkilediğine bakıyor ve ne kadar etkiliyorsa o oranda bu üç değerde güncelleme yapıyor.

**epochs** : eğitim setimizin kaç sefer modele fit leneceğini belirliyor.

**batch\_size**: her seferinde modelin içerisine kaç data atacağını söylüyor. (örneğin train datası 12800 veriden oluşsun. Batch\_size eğer 128 olarak belirlenmiş ise, tüm datayı 100 iterasyonda bitiriyor.)

batch\_size tüm datayı bitirdiğinde epochs un 1 iterasyonu mu tamamlanmış oluyor



**TÜREV:**

**Türev**, bir şeyin bir diğer şeye göre değişim miktardır. Yani türev, "değişim"i ölçmek için kullanılır. Genellikle türevi bir şeyin zaman geçtikçe ne kadar değiştiğini hesaplamak veya ifade etmek için kullanırız. Bunu az sonra örneklendireceğiz.

**İntegral**ise, belli bir aralıktaki toplam değişimi, ya da "biriken değişim miktarını" ifade etmek için kullanılır.

Türev ve integrali anlamak için, integrali çözme yöntemleri bir kenara bırakılarak, hayattan örneklere bakılabilir.  
Örneğin tavanınız akıtıyorsa ve etrafı su götürmemesi için akıtan noktanın hizasına büyük bir kova koyduysanız, kova içerisindeki su damla damla birikecektir. Birim zamanda (örneğin saatte 1 veya günde 1) kovadaki suyun hacmindeki değişim miktarı türev ile hesaplanır. Çok basit tabiriyle, hacim miktarındaki değişimin, zamandaki değişime oranı türevdir! Zamanı saatle ölçersiniz, dersiniz ki 8 saat geçmiş, buna 8 birim zaman diyelim. Kovaya bakarsınız, boşken 8 litre dolmuş. Kovadaki bu hacim değişiminin, zamandaki değişime oranı türevdir!