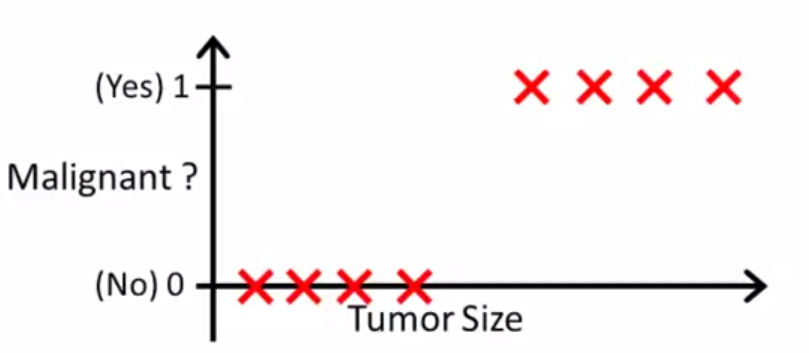
Min-Max Algoritması

Veriler arasında farklılıkların fazla olduğu durumlarda verilerimizi ölçeklememiz gerekebilir. Ölçekleme işlemi için kullanılan yöntemlerden bir tanesi Min-Max yöntemidir.

En büyük x değerimiz 1 değerine ve en küçük x değerimiz 0 değerine eşlenir.Geriye kalanla ise 0-1 değer aralığı arasına eşlenir.

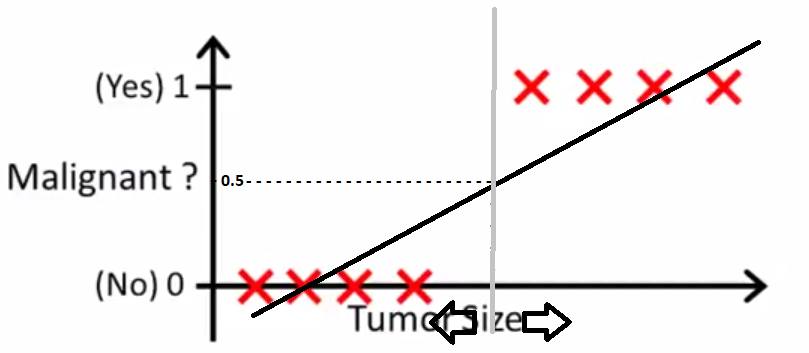
**LOGISTIC REGRESYON İLE SINIFLANDIRMA**

Sınıflandırma işlemi regresyon promblemine benzer ancak çıkış değerimiz (y) sadece ayrık değerler alır(0 veya 1). Linear regresyon kullanarak sınıflandırma yapabiliriz ancak verilerimizin düzenli ve grupsal bir şekilde dağılması gerekir.Genel olarak sınıflandırma problemlerinde linear regresyon kullanılmaz.



Şekil 1: Örnek bir sınıflar üzerinden veri dağılımı gösterimi

Yukarıdaki şekilde görüldüğü üzere verilerimiz düzenli olarak birbirinden ayrı olarak dağılım göstermiştir. Bu problem için linear regresyon kullanılabilir.



Şekil 2:Linear regresyon eğrisinin uydurulması.

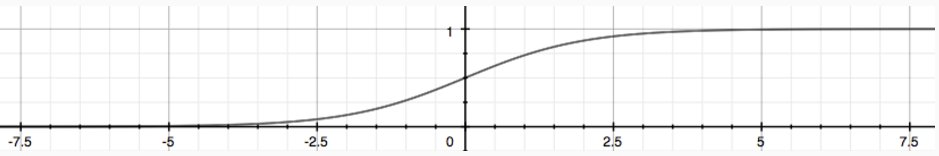
Şekil 2’de görüldüğü üzere linear doğrumuzu uydurduk. Ok işareti ile gösterilen kısım sınıflarımızın karar verme sınırıdır. Ok işaretinin sol tarafında kalan kısım çıkış olarak 0 değeri alan sınıftır.Aynı şekilde ok işaretinin sağ tarafında kalan kısım ise çıkış olarak 1 değeri alan sınıftır. Ancak çoğu zaman verilerimiz bu kadar düzenli değildir. Ayrık bir değer olduğu zaman hipotezimiz değişeceğinden karar verme sınırlarıda değişir ve kötü sonuçlar elde edebiliriz.

Buradan çıkarılacak sonuç ile hipotez değerlerimizin 1’den büyük veya 0’dan küçük değerler alması bizim için kötü sonuçlar doğurmaktadır. Bu yüzden hipotezimizi () sınırlandırmalıyız. Bunun için “Sigmoid Fonksiyonu” kullanılır.

=

g() =

Aşagıdaki şekilde sigmoid fonksiyonu gösterilmiştir.



Şekil 3: Sigmoid fonksiyonu

denklemi bize çıkış değerinin 1 olması olasılığını verir.

Örneğin = 0.7 ise %70 oranında çıkış değeri 1 olur anlamına gelir. Olasılık toplamlarının 1 olması gerektiğinden çıkış değerinin %30 olasılıkla 0 değeri olması anlamına gelir.

Hipotez fonskiyonumuzun () değeri 0.5 değerinden büyük veya eşit ise çıkış değerimiz 1 olur. Eğer 0.5 değerinden küçük ise çıkış değerimiz 0 olur.

Bu durumda giriş değeri (z) 0’dan büyük veya eşit ise çıkış değeri 0.5’den büyük veya eşittir.

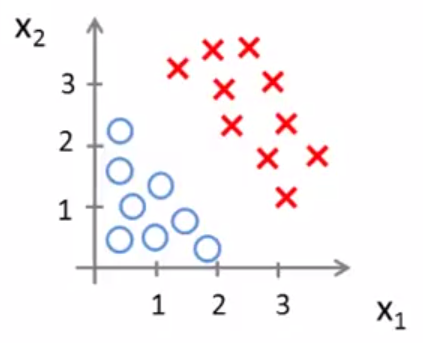
z 0 ise

Sigmoid fonksiyonumuzda giriş olarak 0 değerini aldığımızda g() = formulüne göre oldugu için çıkış değeri g() = olur. Giriş değerimiz değerine yaklaşırsa değerine yaklaştığı için çıkış değerimizi 1 olur. Son olarak giriş değerimiz -değerine yaklaşırsa değerine yaklaştığı için çıkış değerimiz 0 olur.

=

Bu iki eşitliği göz önüne alarak hakkında yorum yapabiliriz.

Aşağıda verilen örnek ile yukarıdaki aşamalar daha iyi anlaşılacaktır.



Şekil 4: Örnek soru veri dağılımları.

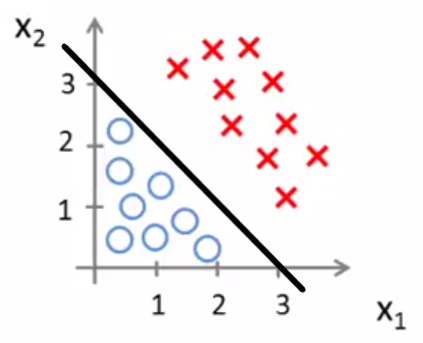
Yukarıdaki şekile göre matrisimiz ve hipotezimiz aşağıda verilmiştir.

denklemimizi yukarıda verilen bilgilere göre oluşturursak

y = 1 durumunda

= -3

eşitliğini buluruz. Elde ettiğimiz bu denklemi şekil üzerinde çizersek karar verme sınırımızı oluştururuz.



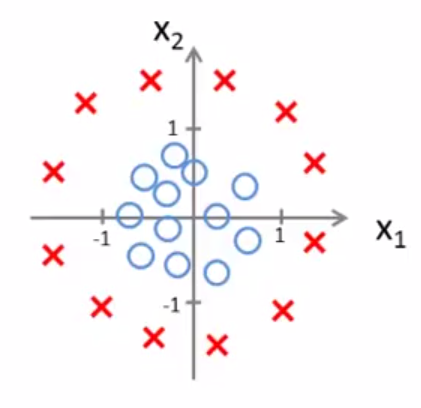
Şekil 5: Karar verme çizgisi ile oluşturulan son durum

Oluşturulan çizginin sağ tarafında kalan kısım, y = 1 çıkışı veren sınıfı temsil eder. Diğer kısım ise y = 0 çıkışı veren sınıfı temsil eder.

**Doğrusal Olmayan Karar Verme Sınırı Problemi**

Bir önceki kısımda doğrusal denklemler üzerinden sonuca ulaşmıştık. Bu kısımda doğrusal olmayan denklemler üzerinden karar verme sınırı oluşturmayı öğreneceğiz.

Aşağıda örnek bir problem verilmiştir şekilde görüldüğü üzere doğrusal bir çizgi ile karar verme sınırı oluşturmamız imkansızdır.



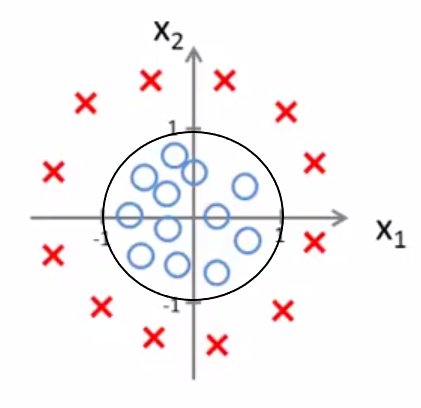
Şekil 6: Doğrusal olmayan veri dağılımları.

matrisimiz ve hipotezimiz aşağıda verilmiştir.

Verilen değerleri hipotezimize yerleştirildiğinde aşağıdaki denklem elde edilir.

y = 1 durumunda

Yukarıdaki eşitliği düzlem üzerinde gösterecek olursak aşağıdaki şekli elde ederiz.

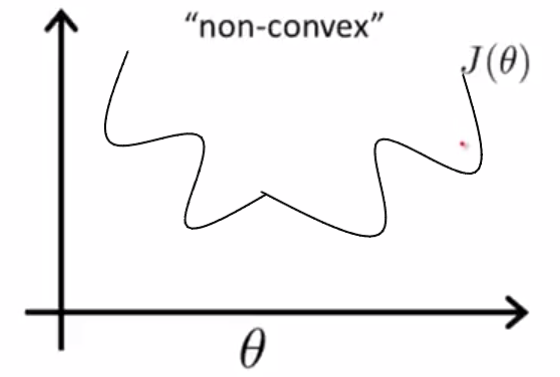


Şekil 7:Doğrusal olmayan eğrinin uydurulması.

Şekil 7’de görülen daire şeklindeki siyah çizgi karar verme sınırımızdır. Kırmızı çarpıların olduğu grubun çıkış değeri 1’dir. Mavi yuvarlakların olduğu grubun ise çıkış değeri 0’dır.

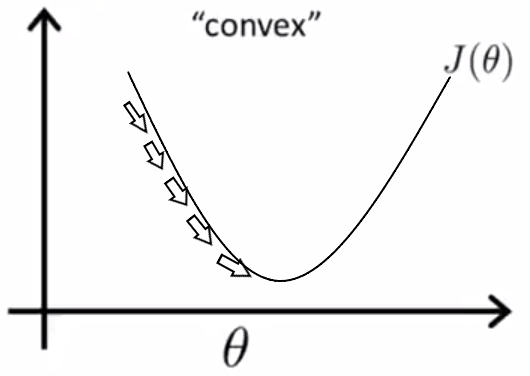
Logistic Regresyon Cost Function

Linear regresyon için kullandığımız cost fonksiyonunu logistic regresyon için kullanamayız.Bunun sebebi ise çok sayıda local optimum değerinin olması başka bir ifade ile convex fonksiyonun elde edilememesidir.



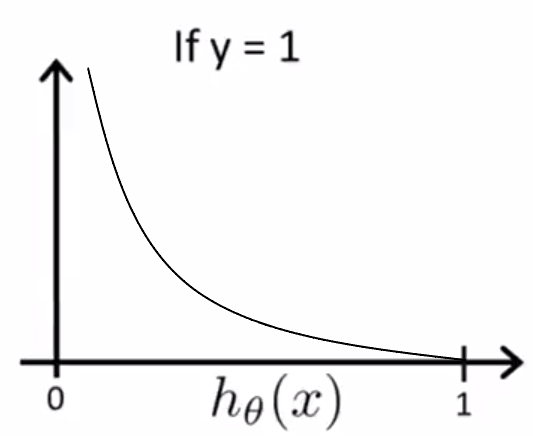
Şekil 8:Convex olmayan fonksiyonun örneksel gösterimi.

Yukarıdaki şekilde görüldüğü üzere çok sayıda local minimum değeri bulunmaktadır ve bu yüzden optimal olan değere yakınlaşmak oldukça zordur.



Şekil 9:Convex olan fonksiyonun örneksel gösterimi

Yukarıdaki şekilde ise tek bir tane local minimum bulunur. Bu şekilde kolayca optimum değere yakınlaşır.



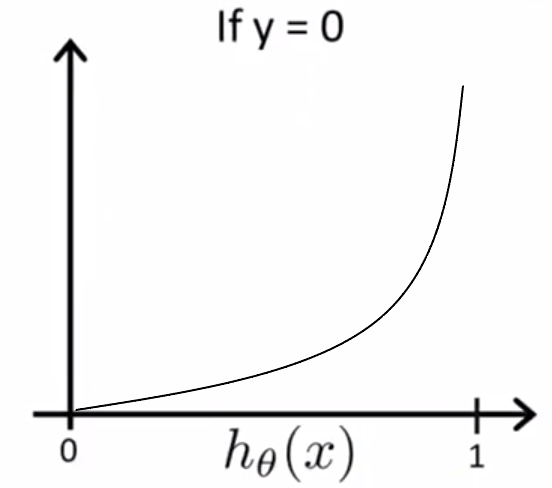
Şekil 10: y=1 durumunda cost fonksiyonunun karakteristiği

y=1 durumunda ;

= 1 olduğu zaman cost fonksiyonun değeri 0 olur.

0 değerine yaklaştığı zaman cost fonksiyonu değerine yaklaşır.

Buradan çıkarıcağımız sonuç ise y=1 durumunda 0 değerine yakınlaştığı zaman cost fonksiyonun çıkış değeri gittikçe artar yani öğrenme algoritması daha çok cezalandırılır.



Şekil 11: y = 0 durumunda cost fonksiyonunun karakteristiği

y = 0 durumunda ise;

değerine yaklaştığı zaman cost fonksiyonun değeri değerine yaklaşır.

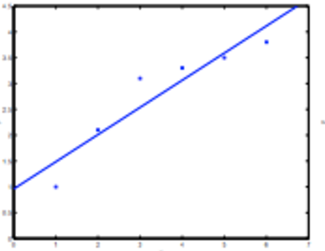
= 0 olduğu zaman ise cost fonksiyonunun değeri 0 olur.

Buradan çıkarıcağımız sonuç 1 değerine yaklaştığı zaman cost fonksiyonun çıkış değeri gittikçe artar yani öğrenme algoritması daha çok cezalandırılır.

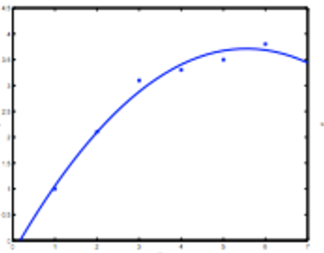
**Regularization**

Öncelikle kullanıcağımız bağımsız değişkenlerin sayısı bizim için önemlidir. 3 farklı durum ile karşılaşabiliriz. Bunlar;

* Underfitting: Bağımsız değişken sayısı azdır.
* Fitting: Veriler tam anlamı ile hipotez fonksiyonuna uyum sağlaması.
* Overfitting: Bağımsız değişken sayısı fazladır. Bu nedenle hipotez fonksiyonumuz mevcut veriye oldukça yüksek bir şekilde uyum sağlar ancak mevcut veriden hariç yeni bir çıkış değeri tahmin etmeye çalıştığımız zaman kötü sonuçlar çıkar bu durumun üstesinden gelmek için Regularization yöntemi kullanılır.



Şekil 12: Underfit durumu



Şekil 13: Fitting durumu

Ayrıca overfitting durumundan bağımsız değişken sayısını azaltarakda kurtulabiliriz.

Elimizide şu şekilde bir denklem olsun;

ve değerlerinin etkisini azaltmak istiyoruz. Bu değerlerden kurtulmak yerine aşağıdaki şekilde bu değerlerin denklem üzerindeki etkisini azaltabiliriz.

ve değerlerinin maliyetini yükseltmek için iki tane ekstra değer ile çarptık. Maliyet fonksiyonu değerinin sıfır değerine yakın olması için ve değerlerinin sıfıra yakın olması gerekir ve bu şekilde overfitting durumundan kurtulabiliriz.

Bir önceki denklemi bu şekilde sade halde yazabiliriz. Burada değerinin seçimi oldukça önemlidir.Eğer çok yüksek bir değer seçilirse underfitting durumu meydana gelebilir.