

ELM472 – Makine Öğrenmesinin Temelleri

Ödev 5

Sigmoid Fonksiyonu ile Gradyan Azaltma Yöntemi

Son teslim tarihi: 03.12.2022 – 17:00

Alican Bayındır

a.bayindir2020@gtu.edu.tr

Elektronik Mühendisliği Bölümü, GTÜ, Kocaeli, Türkiye

I. GİRİŞ

Makine öğreniminin amacı, bir maliyet işlevini optimize ederek bir hedef işlevi maksimize etmek veya azaltmaktır. Kayıp veya maliyet işlevi bunun için bilinen addır. Genellikle, bu işlevi azaltmak isteriz. Model parametrelerinin ve modelinizden veriler geçirildiğinde üretilen hataların bir fonksiyonu olarak, C maliyet fonksiyonu bir ceza atar.

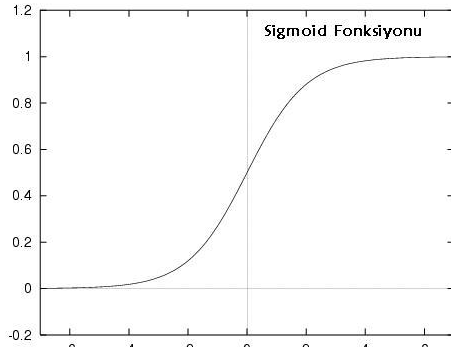
Bir resmin köpek mi yoksa kedi mi içerdiğini belirlemeye çalıştığımız durumu inceleyelim. İdeal modelimiz olsaydı, ona bir fotoğraf gösterebilirdik ve bize kedi mi köpek mi olduğunu söylerdi. Hiçbir model kusursuz değildir ve kusurlar kaçınılmaz olarak ortaya çıkacaktır. Bu gibi ikili karşılaştırmalarda sigmoid fonksiyonları kullanabiliriz. [1]

Sigmoid eğri olarak da bilinen, ayırt edici bir "S"-şekilli eğriye sahip bir matematiksel fonksiyona sigmoid fonksiyon denir. Eksi sonsuz ile artı sonsuz arasındaki herhangi bir değer, 0 ile 1 arasında bir sayıya dönüştürülür. [1]

II. UYGULAMA

Sigmoid işlevi, herhangi bir gerçek sayıyı 0 ile 1 arasında bir değere dönüştürebildiğinden, veri analitiğinde ve diğer birçok alanda kullanışlıdır. Örnekler şunları içerir: [2]

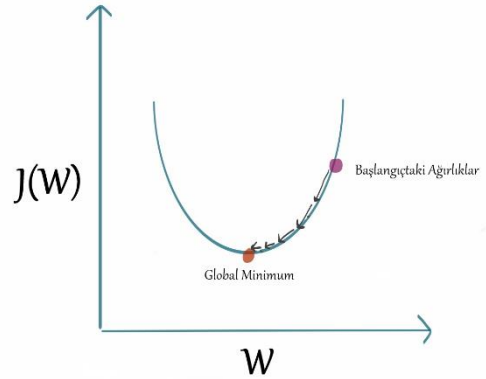
- 1- Yapay sinir ağları, nöronlar içinde doğrusal olmayan bir aktivasyon işlevi kullanan derin öğrenmeyi kullanarak veriler arasındaki doğrusal olmayan korelasyonları keşfedebilir.
- 2- Sigmoid işlevi, bir ikili değişkenin olasılığını tahmin etmek için lojistik regresyon olarak da bilinen ikili sınıflandırmada kullanılır.



Şekil 1 Sigmoid Fonksiyonunun görselleştirilmesi.

Aslında, makine öğreniminin çoğu, bir maliyet işlevini en aza indirerek bir dağılımı belirleyebilen bir grup çerçeveden oluşur. Bir fonksiyonun nasıl küçültülebileceği olası bir sorudur.[2]

Bunu iyi bilinen bir benzetmeyle açıklayalım. Bir 2B küçültme sorunu düşünün. Bu, uzak bir dağ yürüyüşüne çıkmakla karşılaştırılabilir. En alçak noktada bulunan yerleşime geri dönmek istiyorsunuz ve bunu biliyorsunuz. köyün ana yönlerine aşına olmasanız bile. En zor yolda yokuş aşağı gitmeye devam ederseniz, nihayetinde yerleşim yerine ulaşacaksınız. Buna göre eğimin ne kadar dik olduğuna bağlı olarak yüzeyden aşağıya doğru hareket edeceğiz. Belirli bir noktadan sonra eğim yaklaşık olarak belirli bir değerde kalmaya başlayacak. O zaman ise varmak istediğimiz en ideal noktaya ulaştığımızı varsayacağız. Bu yöntem azalan gradyan yöntemi olarak

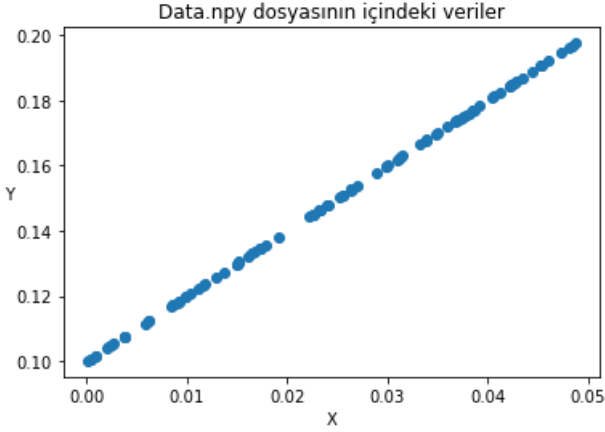


Şekil 2 Azalan gradyan yöntemi görselleştirilmesi.

bilinmektedir. Anlatılan örneğin görselleştirilmiş hali aşağıdaki şekil olarak tanımlanabilir. Başlangıçta eğimi yüksek bir noktadayken sürekli aşağı inerek eğimin artık yükselmediği bir nokta bulduğumuz nokta şekilde global minimum olarak gösterilmiştir. [3]

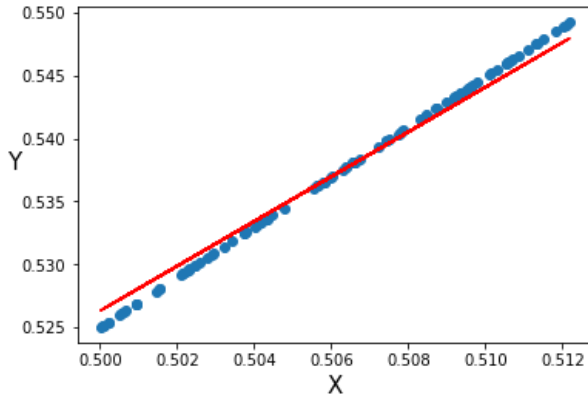
A. Python kodunun yazılması

Gerekli araştırma ve konu incelenmesi yapıldıktan sonra Python kodu yazılmış ve ekte iletilmiştir. Formülün koda işlenmesi sırasında ilk olarak izin verilen kütüphane ödevde bize teslim edilen veri dosyası ile içeri aktarılmıştır. [4]



Şekil 3 data.npy dosyasının içindeki veriler.

Veriler oldukça doğrusal gözüktüğü için kontrol amaçlı lineer regresyon yapıldı ve katsayıları yazdırıldı. Daha sonra sigmoid fonksiyonu için bir tanımlama yapıldı. Sonrasında bu verilerin hepsi sigmoid fonksiyonuna göre tekrar hesaplandı. Sonrasında azalan gradyanlar metodu uygulandı ve aşağıdaki grafikler elde edildi.



Şekil 4 Azalan gradyanlar metodu uygulandıktan sonra elde edilen grafik.

III. SONUÇ

Ödevde bize teslim edilen veriler başarılı bir şekilde analiz edilmiş olup, gradyan azalan yöntemi ile verilen verilerin analizi yapılmıştır. Bu aşamalar gerçekleştirilirken azalan gradyan yöntemi yanlış uygulandığında hatalı bir sonuç elde edilmiştir bu problem referans 4'ün incelenmesiyle çözülmüştür. Bu probleme bağlı olarak bir sonraki aşamada iterasyon ve eta

(öğrenme katsayısı) değerinin ne olacağı konusunda bir kararsızlık yaşanmıştır. Bu problem ise dersin ana kitabı 2. Referansta belirtilen “Introduction to Machine Learning” ile çözülmüştür. Başlangıçta gelen theta değerleri rastgele seçildiğinden dolayı bazen verilen iterasyon sonucuna kadar gittiğinde doğru noktaya denk gelinmediği için Şekil 4'te çizilen doğru her zaman veriler ile uyuşur biçimde çıkmamıştır. Bu problemin çözümü olarak ise herhangi bir işlem yapılmamıştır. Elde edilen en doğru görüntü Şekil 4'te verilmiştir. Ödev sırasında yazılan kod ise ödev raporu ile birlikte teslim edilmiştir.

IV. KAYNAKÇA

- [1] Ahmet Güneş, Dr. Öğr. Üyesi, ELM 472 - Makine Öğrenmesi Temelleri dersi, Gebze Teknik Üniversitesi.
- [2] E. Alpaydin, Introduction to Machine Learning, 3. bs. Cambridge, MA, USA: MIT Press, 2014.
- [3] JAHKNOWS. (2020, MAR 28). Role derivative of sigmoid function in neural networks. <https://datascience.stackexchange.com/questions/30676/role-derivative-of-sigmoid-function-in-neural-networks>
- [4] Nadim Kawwa, Gradient Descent, Show me the Math!, <https://towardsdatascience.com/gradient-descent-show-me-the-math-7ba7d1caef09>