

Makine Öğrenmesi

7. hafta

- Çok Katmanlı Algılayıcı (MLP) Ağı
- Geri-yayılım Algoritması
- Eğitimi Durdurma Kriteri

Yrd. Doç. Dr. Umut ORHAN

1

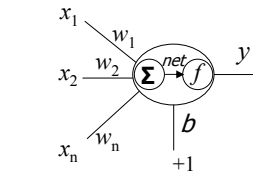
Çok Katmanlı Algılayıcı Ağı

Algılayıcı ve Adaline yöntemleri doğrusal olmayan çözümler üretemediği için hem mimari hem de eğitim algoritması açısından iyileştirilmiş Çok Katmanlı Algılayıcı (MLP) ağı önerilmiştir. Mimari açıdan doğrusal olmayan aktivasyon fonksiyonuna sahip birçok nöronun birbirine hiyerarşik olarak bağlandığı bir yapıya sahip olan MLP, Algılayıcı ve Adaline yöntemlerinin avantajları yanı sıra geri-yayılım adındaki öğrenme sistemini kullanmaktadır.

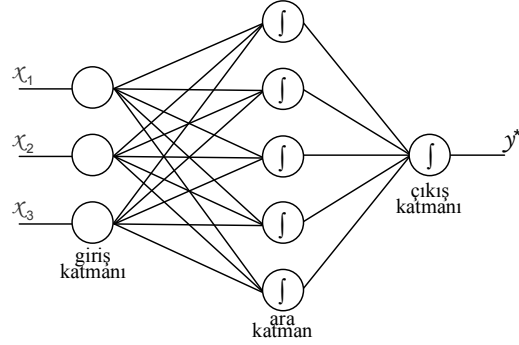
Yrd. Doç. Dr. Umut ORHAN

2

MLP Mimarisi



MLP ağlarında kullanılan bir nöronun genel yapısı.



Bir MLP ağının genel mimarisi.

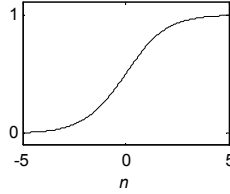
Aktivasyon Fonksiyonu

Genel kavram olarak aktivasyon fonksiyonu, bir değişkeni farklı bir boyuta taşıyan doğrusal veya doğrusal olmayan bir fonksiyondur. Türevinin kolay alınabilmesi eğitim hızını artırır. Sıklıkla kullanılan üç aktivasyon fonksiyonu vardır: Sigmoid, Hiperbolik Tanjant ve Adım Basamak.

Aktivasyon Fonksiyonu

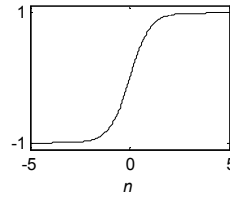
Sigmoid

$$y = \frac{1}{1 + e^{-net}}$$



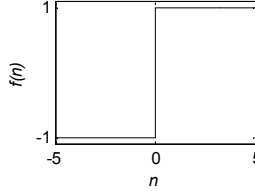
Hiperbolik tanjant

$$y = \frac{1 - e^{-2net}}{1 + e^{2net}}$$



Adım basamak

$$y = \begin{cases} 1 & net \geq 0 \\ -1 & net < 0 \end{cases}$$



Yrd. Doç. Dr. Umut ORHAN

5

Geri-yayılım (Backpropagation)

Tüm ağırlık değerleri, eğim azaltma (gradient descent) denilen bir yöntemin hata fonksiyonunu minimize etmesiyle bulunur. Hesaplanan hata, nöronun çıkışından girişine doğru giderken aktivasyon fonksiyonunun türevine maruz kalır. Fonksiyonun türevine uygulanan hata değeri tüm ağırlıklara, giriş değerleri oranında dağıtılır.

Yrd. Doç. Dr. Umut ORHAN

6

Geri-yayılım (Backpropagation)

$$E = \frac{1}{2} \sum_j e_j^2 = \frac{1}{2} \sum_j (d_j - y_j)^2$$

$$\begin{aligned} \Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \left[\frac{\partial E}{\partial y_j} \right] \left[\frac{\partial y_j}{\partial net_i} \right] \left[\frac{\partial net_i}{\partial w_{ij}} \right] \\ &= -\eta [-e_j] \left[\frac{\partial y_j}{\partial net_i} \right] [x_i] \end{aligned}$$

Aktivasyon fonksiyonu olarak sigmoid kullanılırsa;

$$\Delta w_{ij} = -\eta [-e_j] [(1 - y_j) y_j] [x_i] = \eta e_j (1 - y_j) y_j x_i$$

Yrd. Doç. Dr. Umut ORHAN

7

Geri-yayılım (Backpropagation)

Aşağıdaki ifadenin doğruluğunu $y_j = \frac{1}{1 + e^{-net}}$ için ispatlayınız.

$$-\eta \frac{\partial E}{\partial w_{ij}} = \eta e_j (1 - y_j) y_j x_i$$

Yrd. Doç. Dr. Umut ORHAN

8

Geri-yayılım (Backpropagation)

Eğim azaltma ile sistem adım adım hedefe yaklaşırken bazen yerel minimum değerleri yüzünden adım miktarı sıfıra gider ve mutlak hedefe ulaşamaz. Bunun için eğim azaltma yönteminin önceki adımlardaki değişim değerlerinin belirli oranda (α) katkısını sağlayan momentum etkisi kullanılır.

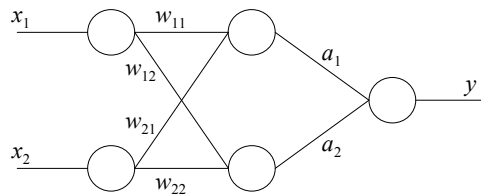
$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t)$$

Yrd. Doç. Dr. Umut ORHAN

9

Örnek

Tek ara (gizli) katmanında 2 nöronu olan bir MLP ağ için iki giriş tek çıkışlı XOR problemini çözelim.



$$\eta = 0.5$$

$$\alpha = 0.1$$

$$w_{11} = 1$$

$$w_{12} = -2$$

$$w_{21} = -1$$

$$w_{22} = 2$$

$$a_1 = 3$$

$$a_2 = -3$$

Yrd. Doç. Dr. Umut ORHAN

10

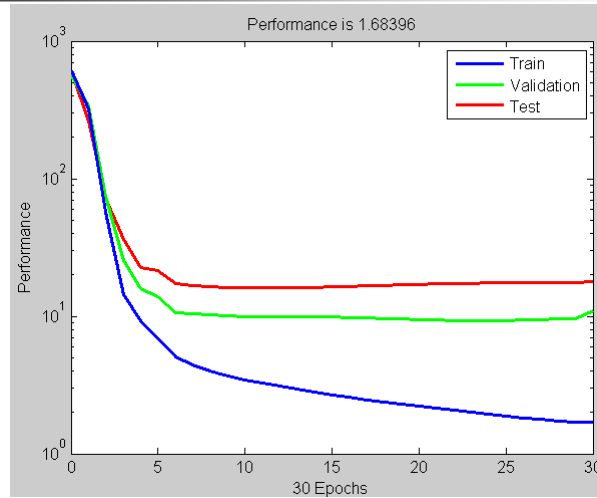
Eğitimi Durdurma Kriteri

MLP ağının eğitiminin ne zaman durması gerektiğini belirleyecek bir kriter olmazsa aşırı eğitim (overfittig) kaçınılmazdır. Çapraz geçerliğe göre bölünen verinin bir altkütmesi olan eğitim kümesi bir daha ikiye bölünür. Bu parçalardan birisi eğitimde ağırlıkların güncellenmesi için kullanılırken diğer parça (validation set) üzerinde eğitimin başarısı hesaplanır. Başarının keskin düşüşü gözlenirse eğitim durdurulur.

Yrd. Doç. Dr. Umut ORHAN

11

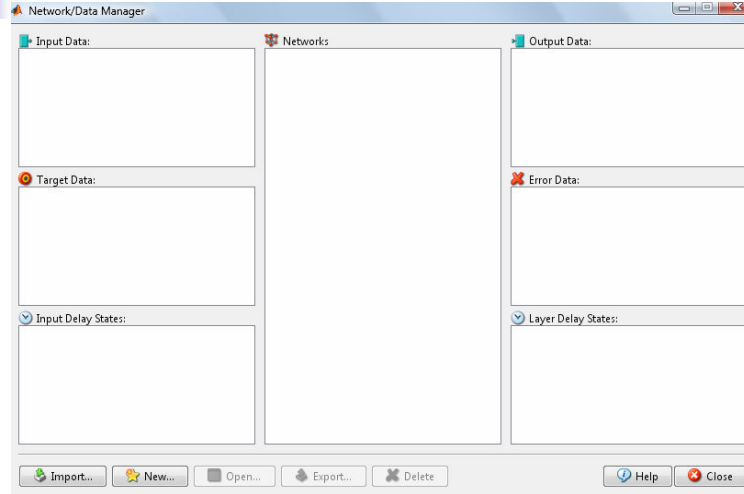
Eğitimi Durdurma Kriteri



Yrd. Doç. Dr. Umut ORHAN

12

Matlab Uygulaması - NNTool



Yrd. Doç. Dr. Umut ORHAN

13

MATLAB Uygulaması

```
>edit MLP_ornek.m
```

Hazırlanmış olan farklı datasetler yüklenerek sınıflandırma ve tahminleme için Çok Katmanlı Algılayıcı ağı kullanılmaktadır. Matlab `newff` ve `train` komutlarının ve eğitilen ağın testinin nasıl yapıldığı irdelenmelidir.

Yrd. Doç. Dr. Umut ORHAN

14