İSTANBUL TEKNİK ÜNİVERSİTESİ
ELEKTRİK-ELEKTRONİK FAKÜLTESİ

EHB 420 - Yapay Sinir Ağları

2021-2022 Güz Dönemi

(CRN:10834)

Öğretim Üyesi: Prof. Dr. Neslihan Serap Şengör

Dersin Yardımcısı: Araş. Gör. Mehmet Onur Demirtürk

Ödev No: 1

<u>Ödev Grubu</u>

Ali İbrahim Yılmaz - 040170010

Ahmet Özler Aktaş - 040170104

1)

a) Altı boyutlu düzlemde 50 noktadan oluşan bir noktalar kümesi belirleyiniz. Bu noktalar kümesini lineer ayrıştırılabilir iki gruba ayırınız. Bu ayrıştırmayı nasıl yaptığınızı açıklayınız.

Lineer ayrıştırılabilir küme: İki küme, lineer tek bir düzlem ile ayrılabiliyorsa bu kümelere lineer ayrıştırılabilir küme denir.

n boyutlu iki kümenin elemanlarının, doğrusal fonksiyon ya da fonksiyonlar ile ayrıştırılması aşağıdaki gibi gerçekleştirilir.

$g_a = a_1x_1 + a_2x_2 + \dots + a_nx_n + c_a$: g_a doğrusal fonksiyonu

$g_b = b_1x_1 + b_2x_2 + \dots + b_nx_n + c_b$: g_b doğrusal fonksiyonu

a kümesinin her elemanı iki fonksiyonda da yerine yazıldığında : $g_a > g_b$,

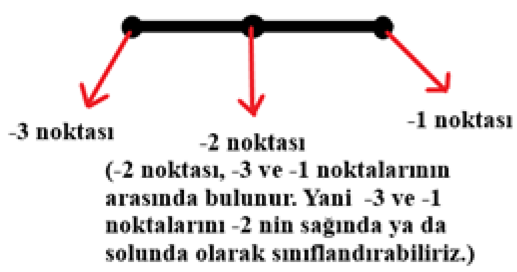
b kümesinin her elemanı iki fonksiyonda da yerine yazıldığında : $g_b > g_a$

koşulları sağlanıyorsa, bu iki küme lineer ayrıştırılabilirdir. Bu koşulu sağlayabilecek birden çok g_a ve g_b fonksiyonları bulunabilir.

Ancak bu fonksiyonların el ile hesaplanması, boyut sayısı arttıkça zorlaşır. Bu soruda lineer ayrıştırılabilir iki küme oluşturmak için temelde aynı mantığa sahip, farklı bir yol izlenmiştir.

İzlenilen Yöntem:

Bir doğru üzerindeki iki noktayı ele alalım. Bu noktaları ayırmak için ikisi arasında bir nokta koyulmalıdır. (Tek boyutlu düzlem için geçerli)



Yandaki şekilde de görüldüğü gibi -2 noktası -3 ve -1 noktaları için ayırt edici bir noktadır.

1 boyutlu uzaydaki kümeler bu şekilde ayırt edilebiliyor.

6 boyutlu uzay için nasıl bir yöntem izleyebiliriz ?

İlk olarak 3 boyutlu uzay üzerinden düşünelim. x,y ve z koordinatları 3 boyutu ifade eder. 4. boyut olarak renk uzayını düşünebiliriz. Şu anda elimizde x, y ve z düzleminde renkli noktalar mevcut. 5. ve 6. boyutu birçok farklı yoldan ifade edebiliriz. 5. boyut için noktaları farklı geometrik şekle dönüştürebilir, 6. boyut için her noktayı bir harf ile isimlendirebiliriz. Bu durumda 6 boyutlu uzayımızı [x,y,z,r,g,h] olarak belirtebiliriz.

Kısaca özetlersek, uzay sayısı arttıkça noktalara her defasında yeni bir sıfat eklemiş olur, yeni gelen matris sütununda da bu sıfatı sayısallaştırmış oluruz.

İki farklı kümenin lineer olarak ayrıştırılabilmesi için en az bir uzayın belirleyici özelliğe sahip olması yeterlidir. (Sınıftaki bütün öğrencileri uzun ya da kısa olarak sınıflandırabiliriz. Bu durumda boy özelliği dışındaki ayırt edici özellikler dikkate alınmamış olur. Diğer bir örnek olarak nesnelerin sadece renklerine ya da hem renklerine hem de şekillerine göre ayırt edilebildiğini düşünebiliriz.)

x, y ve z uzaylarının değer aralığı [-3, 3] aralığında belirlendi.

Renk, geometrik şekil ve harf değerleri de [-3, 3] aralığında belirlendi.

İki kümeyi [0, 3, -1, 1, 0, -1] düzlemi ile ayırmak istediğimizi varsayalım:

[x,y,z,r,g,h] düzlemindeki değerleri kendi başlarına noktalar olarak düşünürsek bu noktaların iki ayrı kümedeki noktaları ayırt edici bir göreve sahip olduklarını söyleyebiliriz. Ancak y değeri [-3,3] aralığının sınır değeri olduğu için iki kümeyi y değerlerine bakarak ayırt edemeyiz.

Bununla birlikte z değerini ayırt edici olarak düşünmez isek ;

Örneğin 15'er elemana sahip iki ayrı kümemiz var ve bu kümeler [0, 3, -1, 1, 0, -1] karar düzlemi ile ayrıştırılacak. 1. kümedeki 14 noktanın z değerleri -2 ya da -3 ve 15. nokta 1 değerinde olsun. Bu durumda z değeri (-1) de y değeri (3) gibi ayırt edici olmaktan çıkar. Karar düzlemindeki z değerinin ayırt edici olmasını istiyorsak, 1. kümedeki bütün elemanların z değerleri ya -2 veya -3 ya da 0,1,2 veya 3 (elimizdeki kümenin değer aralığı [-3, 3] olduğu için) olmalıdırki kümedeki noktaları z = -1 noktasının sağında ya da solunda olarak sınıflandırabilelim. Yani 1. kümedeki z'ler 0,1,2 veya 3 değerlerinde olursa diğer kümedeki elemanların z değerleri -2 ya da -3 değerlerinde olmalıdır.

[0, 3, -1, 1, 0, -1] karar düzlemi referans alınarak 1. ve 2. kümeler oluşturulmuştur. Ancak yukarıdaki açıklamalardan farklı olarak şunu da belirtmek gerekir:

$g_a = a_1x_1 + a_2x_2 + \dots + a_nx_n + c_a$ denkleminde c_a sabiti de mevcuttur. Bu sabit, karar düzleminin kaydırılabileceğini belirtir. Yani karar düzlemi, kümedeki elemanların değer aralığının dışında olabilir. Böyle durumlarda karar düzleminin, c sabiti ile kaydırılarak yine ayırt edici bir hale gelmesi mümkündür. Ancak bu soruda $g_a = a_1x_1 + a_2x_2 + \dots + a_nx_n$ kısmı öncelenmiş, c sabitinin etkisi çok fazla önemsenmemiştir.

Belirlediğiniz kümedeki 30 noktayı eğitim kümeniz, 20 noktayı test kümeniz olarak ayırınız. Oluşturduğunuz bu sınıflama problemini, genlikte ayrık algılayıcıyı (Perceptron) ile çözmeniz isteniyor.

Eğitim Kümeleri:

	0	1	2	3	4	5
0	-1	0	3	3	-3	0
1	-1	-2	0	3	-3	3
2	-1	0	3	3	-2	0
3	-2	2	1	2	-3	2
4	-1	1	0	2	-3	2
5	-1	0	3	3	-2	3
6	-2	0	1	2	-2	1
7	-1	-1	3	2	-2	2
8	-3	1	2	3	-2	3
9	-1	-1	3	2	-3	1
10	-3	-2	3	2	-2	3
11	-3	2	1	2	-2	3
12	-1	-1	2	3	-2	3
13	-2	1	1	3	-3	3
14	-1	-3	0	3	-2	0

Tablo-1. 15 elemanlı 1. küme

	0	1	2	3	4	5
0	3	0	-2	-3	-1	-1
1	2	1	-1	0	1	-2
2	1	-2	-2	-2	-1	-3
3	3	3	-3	-3	-1	-1
4	3	3	-2	0	3	-2
5	2	-1	-1	0	1	-1
6	1	1	-1	0	0	-1
7	3	-1	-3	-2	-1	-2
8	1	-2	-3	0	0	-1
9	3	3	-2	-2	2	-2
10	1	2	-2	-3	1	-3
11	1	-2	-2	-3	-1	-2
12	3	3	-2	-1	-1	-2
13	3	1	-1	0	3	-1
14	3	-2	-1	0	2	-2

Tablo-2. 15 elemanlı 2. küme

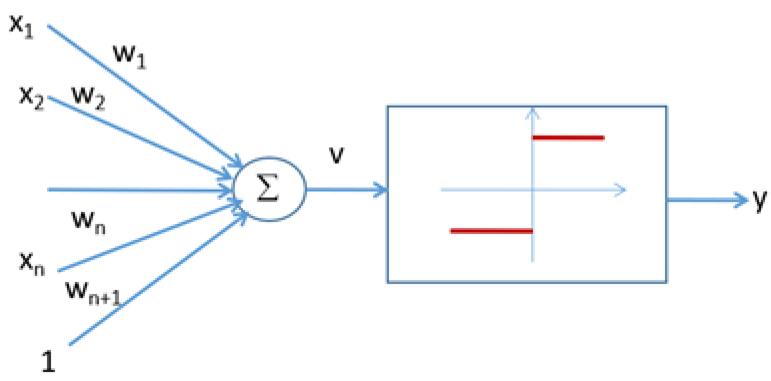
Test Kümeleri:

	0	1	2	3	4	5
0	-3	-3	3	3	-2	2
1	-3	3	1	2	-3	1
2	-2	-1	3	2	-2	3
3	-3	3	3	2	-2	2
4	-1	3	0	3	-2	1
5	-2	-1	1	2	-3	3
6	-1	0	2	3	-3	1
7	-1	1	3	2	-3	0
8	-1	-1	0	2	-2	0
9	-2	2	2	2	-2	2

Tablo-3. 10 elemanlı 1. küme

	0	1	2	3	4	5
0	1	-3	-2	-2	3	-1
1	2	3	-1	-2	1	-1
2	1	3	-3	0	1	-1
3	3	1	-2	-2	1	-2
4	1	3	-1	0	1	-1
5	2	0	-1	-3	2	-1
6	1	-3	-1	-1	-1	-1
7	1	2	-2	-1	-1	-1
8	3	0	-3	-3	2	-1
9	3	2	-3	-3	-1	-1

Tablo-4. 10 elemanlı 2. küme



Genlikte Ayrık Algılayıcının (Perceptron) Yapısı

$$\mathbf{v = w^T * x \quad , \quad y = f(v) \{ 1, v>=0 \ ; \ -1, v<0\}}$$

$$\mathbf{\Delta w_i = c*0.5*(y_d - y)*x_i \quad , c \in R}$$

i) Ağırlıkların ilk koşullarının, öğrenme hızının, eğitim kümesinin farklı sıralanmasının eğitim sürecine etkisini tartışınız.

Ağırlıkların ilk koşulları:

[-3,3] aralığındaki değerlere sahip kümeler için ağırlıkların ilk koşulunu, yine [-3,3] aralığında seçmek daha iyi sonuç verecektir. Daha önce belirtildiği gibi, noktalar kümesini ayırt etmek için onları ikiye bölebilen bir düzleme ihtiyacımız vardır.

[-3,3] aralığından farklı aralıkta verilen ilk koşullar, sonuca ulaşmayı zorlaştırır. Böyle durumlarda istenilen sonuca ulaşmak için iterasyon sayısının arttırılması gerekebilir.

Öğrenme hızı:

Δw_i= c*0.5*[y_d-y]*x_i denkleminde, c öğrenme hızını gösterir. c'nin hangi değerinin daha iyi sonuç vereceğini deneme yanılma yolu ile bulmak süre ve enerji açısından avantaj sağlar.

Yine de veri kümesini inceleyerek c'yi hangi değerlerde seçmemiz konusunda bir fikir edinebiliriz.

Eğitim için oluşturduğumuz küme [-3,3] aralığındaki rakamlardan oluştuğu için c=1 seçmek daha iyi ve hızlı sonuç almamızı sağlar. c=0.001 gibi bir değer seçersek, çok daha fazla sayıda iterasyon yapmamız gerekir. Eğitilen model de test kümelerinde hatalı sınıflandırma yapabilir.

Eğitim kümesinin farklı sıralanması:

Oluşturduğumuz kümeler lineer olarak çok rahat ayrıştırılabildiği için farklı sıralama yapmamız sonucu çok fazla etkilememektedir. Yine de eğitim kümesindeki sırayı rastgele karıştırmanın, aynı kümedeki elemanların arka arkaya dizilmiş halinden daha çabuk sonuç vereceği öngörülebilir.

ii) Elde ettiğiniz sonuçları eğitim ve test kümesi için yorumlayınız.

Öğrenme hızı c=0.001 iken eğitim kümesini %100 doğru bilirken, test kümesinde %85 sonuç veriyor.

c=1 iken 1 iterasyonda %100 doğru sonuç veriyor. Ve test kümesinde de %100 doğru sonuç veriyor.

Sonuçlar, ağırlık ilk koşullarından ve eğitim kümesinin farklı sıralanmasından çok fazla etkilenmiyor.

Hem eğitim hem de test kümeleri doğrusal ayrıştırılabilir olduğu için Perceptron yapısı ile kolayca ayrıştırılabiliyor.

b) Altı boyutlu düzlemde yine 50 noktadan oluşan ama lineer ayrıştırılabilir olmayan iki grup oluşturun, eğitim ve test kümelerini belirleyin. Bu iki grubu sınıflandırmak için genlikte ayrık algılayıcıyı kullanın. Elde ettiğiniz sonucu yorumlayın.

Rastgele değerlerle oluşturduğumuz iki küme büyük olasılıkla lineer ayrıştırılabilir olmayacaktır. Her ihtimale karşı kümeyi oluşturup test edebiliriz. Eğer kümeler lineer ayrıştırılabilir ise a şıkkındaki gibi az bir adım sayısı ile %100 sonuç vermelidir.

Bu yolları kullanarak lineer ayrıştırılamaz küme elde edemiyorsak, lineer ayrıştırılabilirliği engelleyecek bir noktayı kümelere ekleyebiliriz.

Eğitim Kümeleri:

	0	1	2	3	4	5
0	-1	3	2	-2	-3	-3
1	-1	-2	-1	1	2	3
2	-2	-3	1	-1	1	3
3	3	-1	1	0	-3	2
4	2	-2	3	2	3	-3
5	-1	3	2	3	-3	-2
6	-1	-1	-3	-3	0	-1
7	2	-2	-3	3	2	1
8	0	-1	0	-1	-2	-1
9	-2	-3	-2	-2	-2	0
10	-3	1	-1	3	3	0
11	3	3	0	-3	-1	3
12	2	-1	-2	-1	2	0
13	-3	2	3	1	0	0
14	0	0	-3	-2	2	-3

Tablo-5. 15 elemanlı 1. küme

	0	1	2	3	4	5
0	0	3	2	3	3	-3
1	-2	-3	1	3	0	-3
2	3	-3	1	-2	2	-3
3	0	1	2	0	-2	3
4	1	3	2	-1	3	-2
5	-2	-1	-2	-2	3	-2
6	-1	-3	3	2	-1	-3
7	-3	1	3	1	-2	0
8	0	-1	1	3	-2	0
9	0	-1	-2	2	1	1
10	2	3	3	0	0	0
11	1	-2	0	0	3	0
12	2	-2	-2	2	-3	3
13	3	-1	-2	3	-3	2
14	-1	2	0	-3	2	3

Tablo-6. 15 elemanlı 2. küme

Test Kümeleri:

	0	1	2	3	4	5
0	2	-2	3	-1	0	-3
1	-2	-2	-1	-2	-1	2
2	-3	1	2	-1	-3	3
3	0	1	-1	2	1	-3
4	-1	0	-1	1	-3	0
5	1	3	-1	2	3	-2
6	3	-1	2	2	-3	0
7	-1	-1	2	-3	-2	1
8	-3	-3	3	-1	3	1
9	0	-1	-3	-3	-2	3

Tablo-7. 10 elemanlı 1. küme

	0	1	2	3	4	5
0	1	3	0	-1	-1	0
1	3	0	-2	2	-1	1
2	-3	-2	-2	1	3	-1
3	1	0	1	2	-1	1
4	0	3	-2	-3	2	-3
5	-3	0	3	0	3	-1
6	-3	0	2	-1	-1	3
7	-2	-3	3	1	-3	-2
8	-3	1	-2	3	2	-2
9	3	-2	-3	-2	1	-3

Tablo-8. 10 elemanlı 2. küme

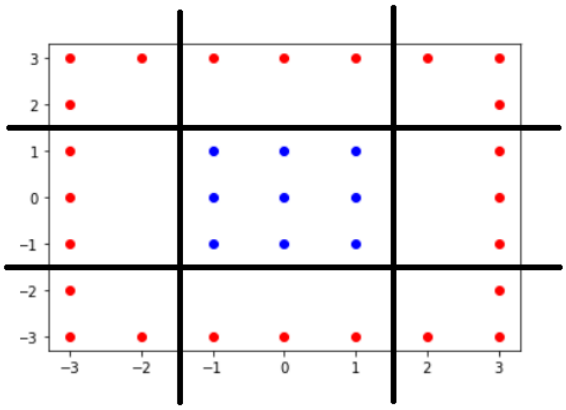
Genlikte ayrık algılayıcı, lineer ayrıştırılabilir olmayan iki kümeyi ayrıştıramaz. İterasyon sayısının arttırılması bu problemi çözmeyecektir. Bu problem, nöron sayısı arttırılarak çözülebilir. Ancak nöron sayısını arttırmak karmaşık problemler için çok masraflı olmaktadır.

2)

$$\left\{\left(\begin{bmatrix}0\\-1\end{bmatrix},1\right),\left(\begin{bmatrix}0\\0\end{bmatrix},1\right),\left(\begin{bmatrix}0\\1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\-1\end{bmatrix},1\right),\left(\begin{bmatrix}1\\0\end{bmatrix},1\right),\left(\begin{bmatrix}1\\1\end{bmatrix},1\right),\left(\begin{bmatrix}-1\\-1\end{bmatrix},1\right),\left(\begin{bmatrix}-1\\0\end{bmatrix},1\right),\left(\begin{bmatrix}-1\\1\end{bmatrix},1\right),\right.$$
$$\left(\begin{bmatrix}-3\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\0\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\-1\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\-3\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}-1\\-3\end{bmatrix},-1\right),$$
$$\left(\begin{bmatrix}0\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}0\\-3\end{bmatrix},-1\right),\left(\begin{bmatrix}1\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}1\\-3\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\1\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\0\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\-1\end{bmatrix},-1\right),$$
$$\left.\left(\begin{bmatrix}3\\-3\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\2\end{bmatrix},-1\right),\left(\begin{bmatrix}-3\\-2\end{bmatrix},-1\right),\left(\begin{bmatrix}-2\\-3\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\3\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\2\end{bmatrix},-1\right),\left(\begin{bmatrix}3\\-2\end{bmatrix},-1\right),\left(\begin{bmatrix}2\\-3\end{bmatrix},-1\right)\right\}$$

i) Yukarıda verilen küme lineer ayrıştırılabilir midir? Değilse ayrıştırmak için nasıl bir yöntem önerirsiniz ?

Bu kümeye genel olarak bakarsak, 1'ler kümesi $-1<=x_1<=1$ ve $-1<=x_2<=1$ düzleminin içerisinde bulunmaktadır. -1' ler kümesinin elemanları ise bu düzlemin sınırları dışında bulunmaktadır. Bu küme aşağıda gösterilmektedir.



Şekilden de anlaşılacağı gibi bu şekildeki iki küme 4 farklı doğru kullanılarak ayrıştırılabilir.

Bu kümeler, m boyut sayısı arttırılarak ya da P örüntü sayısı azaltılarak ayrıştırılabilir.

Kümedeki elemanlar için bir fi($x_{1,}$ x_2) fonksiyonu hesaplanmalıdır.

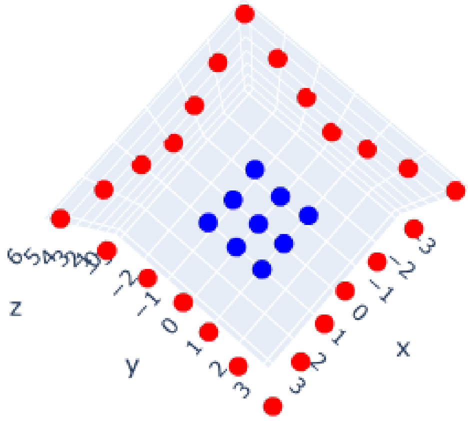
Bu fonksiyon sayesinde boyut arttırma ya da örüntü azaltma durumu gerçeklenebilir.

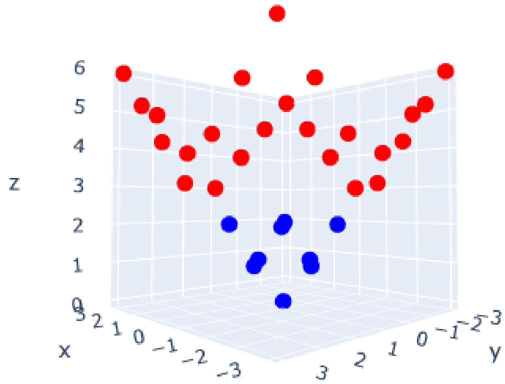
ii) Verilen kümeyi ara katmanında üç birim olan Rosenblatt'ın genlikte ayrık algılayıcısını kullanarak sınıflandırınız. Üç birimi nasıl seçtiğinizi açıklamayı unutmayınız.

Yukarıda bahsettiğimiz işlemi, Rosenblatt'ın genlikte ayrık algılayıcısı için yapalım. Elimizdeki kümeleri lineer ayrıştırmak için m boyut sayısını arttıracak bir fi fonksiyonu belirleyelim:

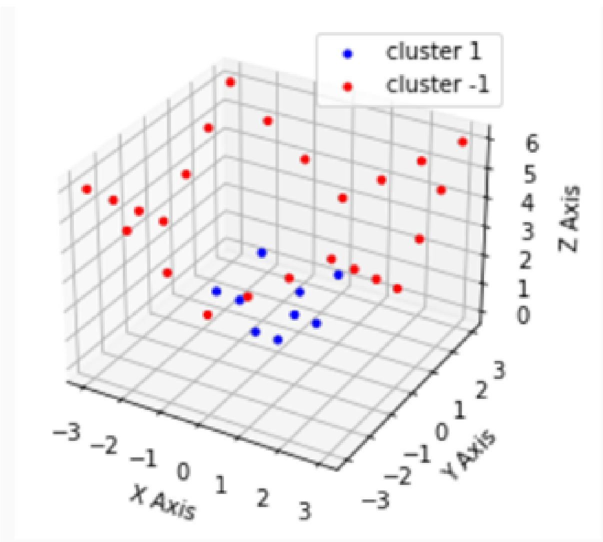
$$\text{fi}(x_{1,}\ x_2) = [x_1, x_2, |\ x_1|+|\ x_2|]$$

Bu fonksiyon x_1 ve x_2 değerlerini aynen verecektir. Eklediğimiz boyutu kullanarak ayrıştırma işlemini gerçekleştireceğiz. $|\ x_1|+|\ x_2|$ sonucu 1'ler kümesi için maksimum 2, -1'ler kümesi için minimum 3 değerini verecektir. z = 2.5 düzlemi ile bu iki kümeyi lineer ayrıştırabiliriz.





Yukarıdaki resimler, kümenin elemanlarının fi(x_1, x_2) fonksiyonu yardımıyla 3 boyutlu düzleme aktarılmış halini gösterir . Çizimler plotly kütüphanesi kullanılarak yapılmıştır. Kırmızı noktalar -1 kümesini, mavi noktalar 1 kümesini temsil etmektedir.



Yandaki gösterim matplotlib kütüphanesi kullanılarak oluşturulmuştur.

Python koduna plotly kütüphanesi ve çizim kodları eklenmemiştir. Ancak matplotlib kütüphanesi ve çizim kodları mevcuttur .

3) Genlikte Sürekli Algılayıcı (ADALİNE) ile aşağıda verilen fonksiyonu yaklaşık olarak ifade etmeniz isteniyor.

$$f(x_1, x_2, x_3) = 0.5x_1x_2 + x_2^2 e^{-x_3}, \qquad x_i \in [0,1]$$

a) Eğitim ve test kümesini oluşturun.

	0	1	2
0	0.153054	0.157686	0.985491
1	0.26771	0.334051	0.644592
2	0.579798	0.618811	0.367772
3	0.55252	0.825631	0.50345
4	0.642304	0.385004	0.0801039
5	0.050829	0.594252	0.968262
6	0.689881	0.623238	0.599231
7	0.460011	0.547737	0.956497
8	0.813136	0.606818	0.635172
9	0.304222	0.759696	0.399904
10	0.441419	0.59567	0.917362
11	0.932431	0.690009	0.958979
12	0.0425459	0.968084	0.636336
13	0.127159	0.663864	0.893559
14	0.477736	0.336287	0.0651107
15	0.288667	0.107237	0.559471
16	0.172229	0.548736	0.597238
17	0.415999	0.162344	0.488767
18	0.898565	0.291583	0.0923233
19	0.845628	0.036555	0.978494
20	0.907464	0.611651	0.707241
21	0.383796	0.527138	0.720337
22	0.187752	0.732938	0.426658
23	0.477233	0.548645	0.0199017
24	0.601894	0.570507	0.793237
25	0.742022	0.635109	0.803838
26	0.175812	0.15445	0.778982
27	0.334897	0.296169	0.376699
28	0.491313	0.247514	0.663451
29	0.387473	0.232813	0.105946

	0	1	2
30	0.255768	0.265827	0.505062
31	0.554885	0.113528	0.163011
32	0.798941	0.523326	0.988521
33	0.272834	0.821832	0.32244
34	0.822798	0.611875	0.0675916
35	0.102452	9.88397e-05	0.484767
36	0.42032	0.246303	0.730217
37	0.386739	0.787918	0.289563
38	0.695709	0.41642	0.682168
39	0.846509	0.0953269	0.0935469
40	0.597098	0.519554	0.016984
41	0.518483	0.867566	0.112602
42	0.396897	0.985131	0.8001
43	0.0758497	0.254863	0.814763
44	0.464465	0.232016	0.117043
45	0.332131	0.867206	0.865283
46	0.851618	0.581225	0.658151
47	0.479219	0.814408	0.324222
48	0.557343	0.459419	0.461137
49	0.340717	0.770498	0.539822
50	0.185424	0.190881	0.664729
51	0.340851	0.651341	0.951972
52	0.68027	0.920148	0.0480063
53	0.743107	0.406184	0.20769
54	0.295611	0.575725	0.78003
55	0.995784	0.68409	0.347476
56	0.733876	0.923568	0.584726
57	0.273481	0.287562	0.554783
58	0.0752848	0.263387	0.275332
59	0.444357	0.297788	0.569906

Tablo-9. 60 elemanlı eğitim kümesi

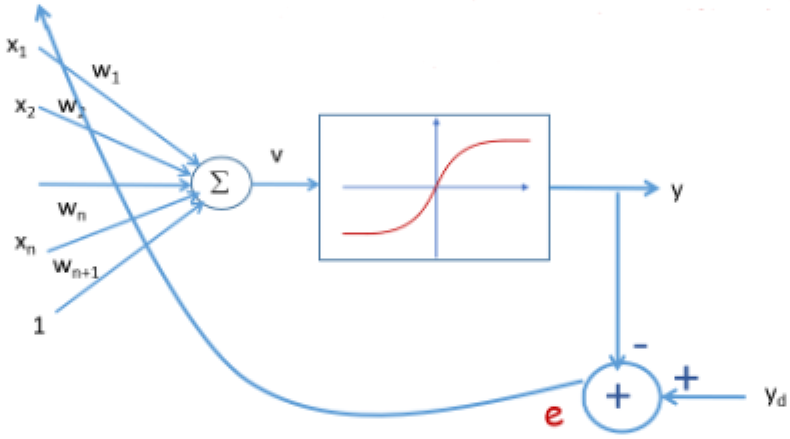
	0	1	2
0	0.375485	0.394296	0.059864
1	0.879765	0.697219	0.263728
2	0.458021	0.34603	0.142651
3	0.973884	0.889787	0.649085
4	0.347174	0.274756	0.017983
5	0.950464	0.708705	0.743288
6	0.215545	0.458756	0.588486
7	0.0371865	0.766326	0.685168
8	0.107116	0.534422	0.548962
9	0.872191	0.13827	0.707156
10	0.861169	0.376637	0.255175
11	0.00826578	0.936771	0.227829
12	0.409447	0.155636	0.213013
13	0.743621	0.612233	0.488384
14	0.86753	0.791406	0.913295
15	0.896264	0.278942	0.614128
16	0.269993	0.974131	0.579353
17	0.593347	0.593151	0.654676
18	0.937657	0.0800461	0.388802
19	0.450118	0.0888931	0.366714

	0	1	2
20	0.940858	0.571149	0.245928
21	0.272425	0.26034	0.137767
22	0.906702	0.761574	0.941156
23	0.360072	0.567752	0.773927
24	0.0999441	0.401954	0.577411
25	0.400661	0.902748	0.0427358
26	0.294176	0.971073	0.54468
27	0.804872	0.560414	0.111671
28	0.793605	0.691298	0.236596
29	0.21915	0.418664	0.468845
30	0.813989	0.238985	0.708845
31	0.333562	0.00692437	0.263032
32	0.0540255	0.693347	0.978597
33	0.703696	0.207215	0.905119
34	0.245933	0.185332	0.428378
35	0.750374	0.603628	0.404558
36	0.185108	0.606904	0.0458724
37	0.495709	0.248386	0.234815
38	0.789566	0.50824	0.969603
39	0.769493	0.4585	0.299588

Tablo-10. 40 elemanlı test kümesi

b) Durdurma kriterinizi nasıl belirlediğinizi ve test kümesinde elde ettiğiniz sonuçlara dayanarak kriterin sonucu nasıl etkilediğini tartışınız.

Genlikte sürekli algılayıcı (Adaline) yapısı aşağıda gösterilmektedir.



Hatayı belirlemek için hata fonksiyonunu yazalım :

$E = 0.5 * e^T * e = 0.5 * e^2$

$e = y_d - y = y_d - \boldsymbol{\sigma} (w_1x_1 + w_2x_2 + \ \ldots + w_nx_n + w_{n+1})$

$v = w^T * x$, $y = \boldsymbol{\sigma} (v)$

$w_{k+1} = w_k - \eta * \nabla E(w) \quad ; \ \eta \in [0,1]$

$$\frac{\partial E}{\partial wi} = \frac{\partial E}{\partial e} \ \frac{\partial e}{\partial y} \ \frac{\partial y}{\partial v} \ \frac{\partial v}{\partial wi} = e * (-1) * \boldsymbol{\sigma}'(v) * x_i$$

$w_{k+1} = w_k + \eta * e * \boldsymbol{\sigma}'(v) * x_i$

Eğitimi durdurmak için izlenebilecek iki yöntem:

1- Kümedeki elemanlardan, maksimum hataya sahip olan elemanın hatası belirli bir sınırın altına düşünce eğitim durdurulur.

2- Hataların ortalaması belirli bir sınırın altına inince eğitim durdurulur.

Bunların dışında başka bir yol kullanarak da eğitim belirtilen iterasyona gelmeden bitirilebilir.

Durdurma kriteri olarak ikinci yol seçildi.

$E = 0.5 * e^T * e = 0.5 * e^2$

Bir iterasyondaki hataların toplamı : $s = \sum_{i=1}^{k} E_i$, k : kümedeki eleman sayısı

Hataların ortalaması : $m = s/k$

Hataların ortalaması, bir süre sonra belirli bir sınırdan daha aşağı bir değere inemiyor. Durdurma kriterinin çalışması için sınırın biraz daha yukarıya çekilmesi gerekir. Ancak bu işlem, iterasyon sayısının yüksek olmadığı eğitim modellerinde, eğitilen modelin hata oranını arttıracağı için istenilen bir yöntem değildir.

Adım sayısı (η), 0.49 olarak belirlendi. İterasyon sayısı artmasına rağmen eğitim esnasında hataların ortalaması 0.00104 değerinden daha aşağı bir değere düşmedi. Durdurma değeri olarak 0.00100 ve altında kullanmak durdurma kriterinin işlevsiz hale gelmesine sebep olacaktır. 0.001 değerinden daha yüksek bir değer durdurma kriteri olarak belirlenebilir.

Durdurma kriteri istenilen hassasiyete göre ayarlanabilir. Daha hassas sonuç vermesi istenen modellerde, durdurma kriteri için hataların ortalamasının çok küçük bir değere gelmesi beklenebilir.

4)

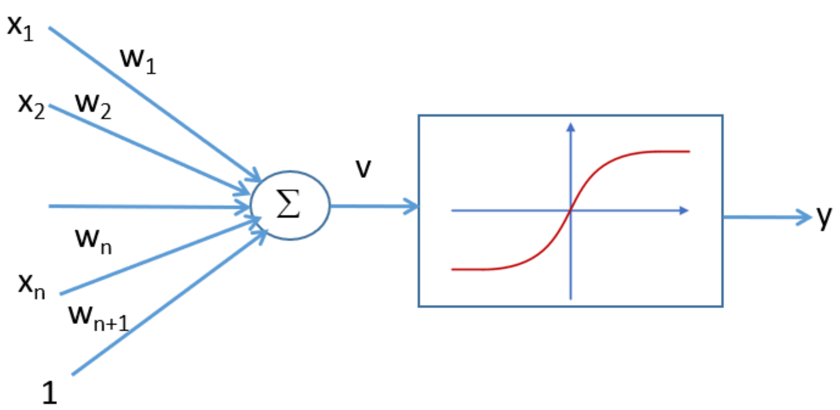
$x_1, x_2, \ldots, x_n$ değişkenlerine değeri bilinmeyen c_i, c_{ij} katsayıları ile

$$f(x_1, x_2, \ldots, x_n) = \sum_i c_i x_i + \sum_i \sum_j c_{ij} x_i x_j$$

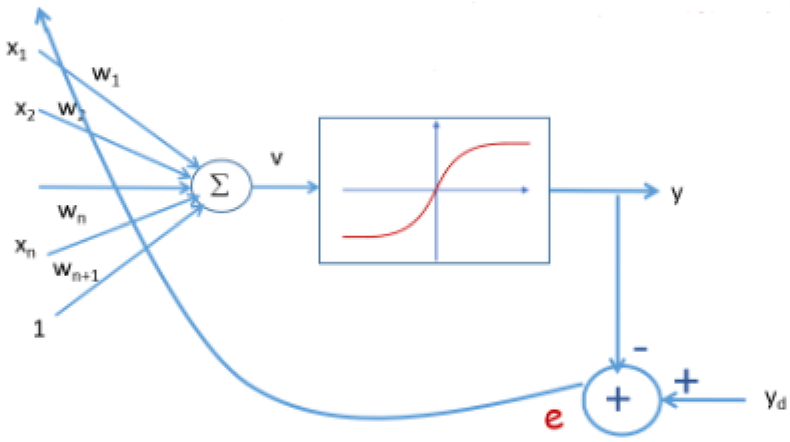
şeklinde bağlı olan $f(x_1, x_2, \ldots, x_n)$ fonksiyonu genlikte sürekli algılayıcı ile belirlenmek isteniyor. Genlikte sürekli algılayıcıda nasıl bir değişiklik yaparsınız? Önerdiğiniz değişiklik ile elde ettiğiniz genlikte sürekli algılayıcı yapısını kullanarak (3)'deki problemi çözünüz ve sonuçları karşılaştırınız.

Yukarıda belirtilen $f(x_1, x_2, \ldots, x_n)$ fonksiyonundaki c_i, c_{ij} katsayıları bilinmiyor. Bu fonksiyon genlikte sürekli algılayıcı ile belirlenmek isteniyor. $c_i, c_{ij} < \infty$ kabul edelim.

Genlikte sürekli algılayıcı, girdilerin ağırlık matrisi ile çarpımının tanh ya da lojistik fonksiyon gibi bir sigmoid fonksiyonundan geçirilmesi ile sonuç veren bir yapıdır.



y çıktısının, elimizdeki f fonksiyonunun sonucu olan y_d çıktısı ile farkı alınır ve epsilon hatası elde edilmiş olur. Buradaki sigmoid fonksiyonunun çıktısı, tanh fonksiyonu için [-1,1] aralığında, lojistik fonksiyon için [0,1] aralığında değerler alabilir.



Ancak f fonksiyonunun c_i, c_{ij} katsayıları bilinmemektedir. Bu durumda için epsilon hatası c_i, c_{ij} katsayılarının etkisiyle çok büyük değerler çıkabilir. Böyle bir durumda eğitim gerçekleştirilemez. Bunun önüne geçmek için f fonksiyonunun çıktısı y_d'ye normalizasyon işlemi uygulamamız gereklidir.

Normalizasyon işlemi, bir kümedeki değerleri belirli bir aralığa sıkıştırma işlemidir.

3. soruda $x_i \in [0,1]$ için f fonksiyonu [0, 1.5] değer aralığında değerler alabilmektedir. Bu fonksiyonun [0,1] aralığına getirmek için normalizasyon işlemi uygulamamız gerekli.

f fonksiyonunun alabileceği maksimum değer 1.5 olduğu için sonuç kümesindeki bütün elemanları 1.5'ye bölmemiz durumunda sonuç kümesi [0,1] aralığına sıkıştırılmış olacaktır.

f fonksiyonunun çıktıları normalize edildikten sonra test kümesi için hataların ortalaması, 0.00068' den 0.00022' ye düştü. Buradan, normalizasyon yapıldıktan sonra eğitimde daha iyi bir sonuç elde edildiği sonucu çıkartılabilir .