

SOM TİPİNDE YAPAY SINİR AĞLARINI KULLANARAK TÜRKİYE’NİN İHRACAT YAPTIĞI ÜLKELERİN KÜMELENMESİ ÜZERİNE BİR ÇALIŞMA

Hüdaverdi BİRCAN^(*)
Metin ZONTUL^(**)
Ahmet Gürkan YÜKSEK^(***)

Özet: Bu çalışmada, SOM tipinde yapay sinir ağları kullanılarak Türkiye'nin ihracat yaptığı ülkeler kümelenmiştir. SOM sinir ağları ile kurulan modeller, veriler için herhangi bir dağılım ve korelasyon varsayımı içermemektedir. Ayrıca, bu modeller verilerdeki satır sayısı ve değişken sayısı arttıkça daha iyi sonuç vermektedir. Üstelik, bu modeller, elde edilen kümelerin topolojik komşuluğunu da vermektedir. Bu sebeple, klasik istatistiksel kümeleme yöntemleri yerine SOM sinir ağları tercih edilmiştir.

Anahtar Kelimeler: Yapay Sinir Ağları, SOM Sinir Ağları, Kümeleme Analizi

Abstract: In this study, the countries to which Turkey exports, are clustered using SOM type neural network. The models established by SOM neural networks don't contain any distribution and correlation assumptions for data. Moreover, these models provide better results when there is an increase in number of rows and variables in data. Also, these models give topologic neighbourhood of the gained clusters. Therefore, SOM neural networks are preferred to classic statistical clustering methods.

Keywords: Neural Networks, SOM Neural Networks, Cluster Analysis

I.Giriş

1980'lerin ortalarından itibaren istatistik alanında çalışan bilim adamları yapay sinir ağlarıyla ilgilenmeye başlamışlardır. White (1989) yapay sinir ağlarında kullanılan öğrenme algoritmaları ile klasik istatistiksel yöntemlerdeki matematiksel modellerin birbirine benzerlik gösterdiğini ispatlamıştır. Cheng ve Titterington (1994) bir çalışmalarında, istatistiksel açıdan yapay sinir ağlarını incelemişlerdir. Bu araştırmacılar, yapay sinir ağları ile istatistiksel yöntemlerin birbirine alternatif gibi gözükmemesine rağmen aslında aynı temaların farklı isimler altında işlendiğini, bir nörobilimcinin ne kadar istatistiksel bilgiye ihtiyacı varsa bir istatistikçinin de o ölçüde yapay sinir ağları bilgisine sahip olması gerektiği belirtilmiştir. Sarle (1994) bir çok yapay sinir ağı modelinin istatistiksel yöntemlere uygunluk gösterdiğini belirtmiştir.

İstatistiksel yöntemlerin çoğunda verilerle ilgili dağılım varsayımı ve değişkenlerle ilgili varyasyon ve korelasyon varsayımları vardır (Neter, Wasserman, Kutner 1989). Sinir ağları yaklaşımının klasik istatistiksel

^(*)Yrd.Doç.Dr.Cumhuriyet Üniversitesi İİBF Fakültesi İşletme Böl

^(**)Cumhuriyet Üniversitesi Anadolu Bil Meslek Yüksek Okulu

^(***)Cumhuriyet Üniversitesi SMYO Bilgisayar Programı

yöntemlere göre avantajı, verilerin dağılım varsayımları ile değişkenlerle ilgili varsayımlara gereksinim duymamasıdır. Yapay sinir ağları, bazı değişkenlere ait eksik verileri de tolere etme özelliğine sahiptir. Ancak, sinir ağlarının eğitimi için çok sayıda veri gerekmesi ve ağ modelinin kurulmasındaki güçlükler, dezavantaj olarak görülebilir (Öztemel 2003).

Yapay sinir ağlarının istatistiksel uygulamalarına bakıldığında, genellikle “Geri Yayılım (Back-Propagation)” ve “SOM (Self-Organizing Maps)” öğrenme algoritmalarının kullanıldığı görülür. Bunlardan ilki, ileri beslemeli çok katmanlı ağ modellerinde, sonraki ise ileri beslemeli tek katmanlı ağ modellerinde kullanılmaktadır. SOM algoritması, kullanıldığı yapay sinir ağlarıyla özdeşleşmiştir. Bu ağlar Kohonen SOM sinir ağları olarak adlandırılır. Teuvo Kohonen 1982 yılında bu ağ modelini öne sürmüştür. SOM ağları tek katmanlı bir ağıdır. SOM algoritması bir denetimsiz öğrenme (Unsupervised Learning) algoritmasıdır. Bu ağın eğitiminde kullanılacak veriler bağımlı değişken içermez. Çoğunlukla bu değişkenler, özellikler (Features) olarak anılır (Kohonen 2001:159). Bu sebeple, SOM tipindeki ağlar kümeleme analizi ilgili problemlerin çözümü için idealdir. Ancak, elde edilen sonuçların doğruluğunu denetlemek için konuyla ilgili uzman görüşüne başvurmak gerekebilir.

SOM ağları, hem verilerin kümelenmesi hem de görselleştirilmesi açısından tercih edilmektedir. Bu ağlar çok boyutlu bir veriyi iki boyutlu bir haritaya indirgemektedir. Her bir küme için oluşturulan referans vektörleri bir araya geldiğinde bir haritayı meydana getirmektedir. Bu harita, üzerindeki topolojik komşuluk kümeler arasındaki ilişkiyi göstermektedir. SOM ağlarının geliştirilmiş bir çok çeşidi vardır. Kangas, Kohonen ve Laaksonen (1990) SOM algoritması üzerinde yapılabilecek çeşitli değişiklikler üzerinde durmuşlardır. Referans vektörlerine ilk değer atanması, dinamik topolojik komşuluk ve oluşan referans vektörlerinin LVQ algoritmasıyla iyileştirilmesi konusunda önerilerde bulunmuşlardır. Pal, Bezdek ve Tsao (1993) denetimsiz SOM algoritması ile denetimli LVQ algoritmasını karşılaştıran bir çalışma yapmışlar ve LVQ yerine GLVQ algoritmasını önermişlerdir. LVQ, SOM algoritmasına benzemekle beraber denetimli olması ve topolojik komşuluk içermemesi açısından SOM’dan farklıdır. LVQ algoritmasında sadece kazanan nöronun katsayıları güncellenirken, GLVQ algoritmasında diğer nöronların da katsayıları güncellenmektedir. Martinetz, Berkovich ve Schulten (1993) SOM algoritmasına alternatif olarak Neural-Gas algoritmasını öne sürmüşlerdir. Neural-Gas, SOM algoritması ile K-ortalama yöntemini birlikte kullanır. Bu araştırmacılar, Neural-Gas algoritmasının “zaman serilerinin tahmini” konusunda bir uygulamasını gerçekleştirmişlerdir. Rauber (1998) bir çok iki ve üç boyutlu SOM modeli üzerinde çalışma yapmıştır. Bütün modellerde üç ayrı veri setinin uygulaması yapılmış ve sonuçlar karşılaştırılmıştır. Weijters, Bosh ve Herik (1997) bir çalışmalarında geri yayılım ağları ile SOM ağlarını birlikte kullanmışlardır. Böylelikle, daha az sayıda gizli nöron kullanmanın mümkün olduğunu iddia etmektedirler.

Son yıllarda, SOM sinir ağları doküman kümeleme (Text Clustering) çalışmalarında yoğun olarak kullanılmaktadır. Honkela, Kaski ve Lagus, Kohonen (1996;1997) tam metin (full-text) dokümanları WEBSOM metoduyla organize etmişlerdir. Bu çalışmalarda, dokümanlar önce kelimelere parçalanmış, daha sonra anlamca birbirine yakın kelimeler gruplanıp bu kelimelerin frekans tabloları oluşturulmuştur. Böylelikle, istenilen konulara göre WEB üzerinden ilgili dokümanlar bir araya getirilebilmektedir. Lagus, Kaski ve Kohonen (2003) bir diğer çalışmalarında, WEBSOM metoduyla Britannica ansiklopedisinin doküman haritasını çıkartmışlardır. Choudhary ve Bhattacharyya (2002) “Evrensel Ağ Dili” (Universal Network Language=UNL) ile dokümanlardaki cümleleri grafiklerle temsil etmişlerdir. Daha sonra, bu grafiklere göre evrensel kelime (Universal Word) frekanslarını içeren giriş vektörleri oluşturulmuştur. Kelimenin frekansı, grafikte bu kelime için kullanılan bağlantı sayısı ile ölçülmektedir. Bu vektörler SOM sinir ağıyla modellenmiş ve böylelikle dokümanlar kümelenebilir.

Hsu ve Halgamuge (2003) hem genişleyen bir harita yapısı hem de hiyerarşik bir SOM modeli önermektedir. Böylelikle hem optimal harita büyüklüğü dinamik olarak tesbit edilebilmekte hem de büyük kümeleri alt kümelere parçalamak mümkün olabilmektedir. Jin, Shum, Leung ve Wong (2003) yeni bir topolojik komşuluk öneren ESOM algoritmasını sunmakta ve bu algoritmayı benzetim (simülasyon) verileri üzerinde SOM algoritmasıyla karşılaştırmaktadır. ESOM, SOM’dan farklı olarak öğrenme kuralına bir de genişleme katsayısı faktörünü eklemektedir.

İstatistik bilimi dışında yapay sinir ağlarıyla yapılan çalışmaların çoğu model tanıma, sınıflama ve bağımlı değişkenlerin tahmini gibi problemlerin çözümü ile ilgilidir. Dolayısıyla, doğrudan belirtilmese de yapılan çalışmalar istatistikselidir. Gerek fen gerekse sosyal bilimlerde yapılan pek çok çalışmada klasik istatistik yöntemleri yerine yapay sinir ağı modelleri kullanılmış ve çok başarılı sonuçlar elde edilmiştir. Creedy ve Martin (1997) borsa endeksi ve döviz kurlarının tahmini için geri yayılım ağlarını önermektedir. Larkin (1999) portföy yönetimini geri yayılım ağı ile gerçekleştirmiştir. Kaski ve Kohonen (1995) bir çalışmada SOM tipinde bir yapay sinir ağı kullanarak ülkeleri refah ve zenginlik seviyelerine göre kümelemişlerdir. Bu çalışmada ülkelerin ekonomik, eğitim ve refah düzeylerini gösteren 39 değişken kullanılmıştır. Markey, Lo, Tourassi ve Floyd (2003) yaptıkları bir çalışmada bir göğüs kanseri veritabanını kullanmışlar ve bu veriler üzerinde SOM ile kümeleme analizi çalışması gerçekleştirmişlerdir. Plaehn ve Laundahl (2003) bir SOM ağıyla müşterileri homojen gruplara ayırmışlardır. Bu araştırmacılar, SOM modeli ile yapılan kümeleme çalışmasının diğer klasik istatistiksel yöntemlere göre daha iyi sonuç verdiğini göstermişlerdir.

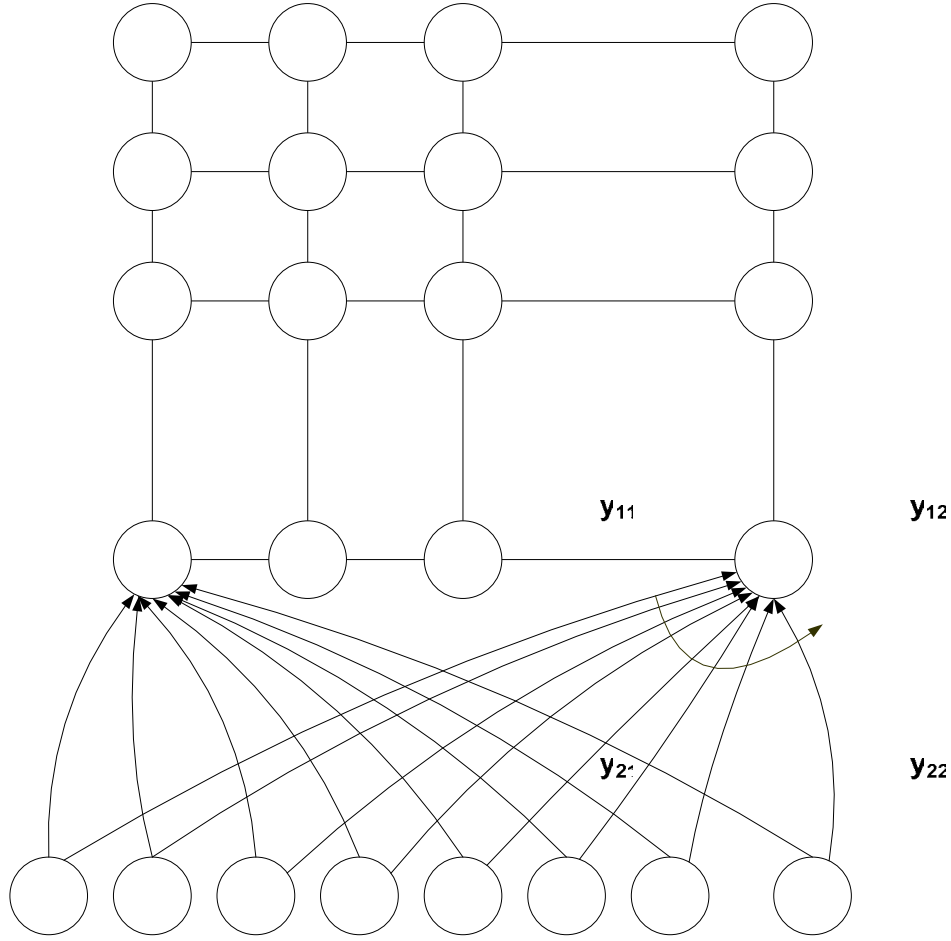
Bu çalışmanın amacı, yapay sinir ağları ile bir kümeleme çalışması gerçekleştirmektir. Bu çalışmada, Devlet İstatistik Enstitüsü Dış Ticaret Şubesi’nden alınan Türkiye’nin 2002 yılına ait ihracat verileri kullanılmıştır. Bu

veriler, geliştirilen bir yazılım ile öncelikle bir veritabanında toplanmış ve sonra da SOM tipinde bir yapay sinir ağı ile modellenmiştir. Bu SOM sinir ağı modeliyle, Türkiye'nin ihracat yaptığı ülkeler harmonize sistemdeki fasıl gruplarında (bölümler) yapılan ticaret tutarlarına göre kümelenmiştir. Ülkelerin kümelenmesinde etkili olan faktörler belirtilmiştir.

II.SOM Sinir Ağları

Kümeleme çalışmalarında, klasik istatistiksel yöntemler yerine yapay sinir ağları kullanılabilir. Yapay sinir ağları, veriler için dağılım varsayımlarına gerek duymaz. Bir veri setindeki eleman ve değişken sayısının çok fazla olması sinir ağları için bir zorluk yaratmaz. Kümeleme çalışmalarında en çok kullanılan yapay sinir ağları SOM (Self-Organizing Maps) sinir ağlarıdır (Kohonen 2001:105). Bu ağlar, 1982 yılında Teuvo Kohonen tarafından geliştirilmiştir. Bu sebeple, Kohonen SOM ağları olarak da bilinir. SOM ağları, klasik istatistikteki K-Ortalama ile çok boyutlu ölçekleme yöntemlerinin her ikisinin de işlevlerini yerine getirebilir. Yani, veri setindeki elemanları hem kümelendirir hem de haritalandırır. Bu sebeple, bu ağlar son yıllarda oldukça popüler olmuştur.

SOM ağları, tek katmanlı bir ağ olup giriş ve çıkış nöronlarından oluşur. Giriş nöronlarının sayısını veri setindeki değişken sayısı belirler. Çıkış nöronlarının her biri bir kümeyi temsil eder. Şekil 2.1'de bir SOM ağı görülmektedir. Diğer yapay sinir ağlarından farklı olarak, çıkış katmanındaki nöronların dizilimi çok önemlidir. Bu dizilim doğrusal, dikdörtgensel, altıgen veya küp şeklinde olabilir. En çok dikdörtgensel ve altıgen şeklindeki dizilimler tercih edilmektedir. Pratikte, çoğu kez dikdörtgensel dizilim karesel dizilim olarak uygulanır. Buradaki dizilim topolojik komşuluk açısından önemlidir. Aslında, çıkış nöronları arasında doğrudan bir bağlantı yoktur. Giriş nöronları ile her bir çıkış nöronu arasındaki bağlantıyı referans vektörleri (code-book vectors) gösterir. Bu vektörler bir katsayılar matrisinin sütunları olarak da düşünülebilir. SOM sinir ağları eğitilirken, bu topolojik komşuluk referans vektörlerinin yenilenmesinde kullanılır.



Şekil 2.1: Kohonen SOM Sinir Ağı

III. SOM Öğrenme Algoritması

Kohonen ağlarında kullanılan öğrenme algoritması bu ağlara ismini de veren, SOM (Self Organizing Maps) algoritmasıdır. Bu ağlarda kullanılan öğrenme algoritması denetimsizdir. Yani, ağ eğitilirken bağımlı değişken kullanılmaz. Veri setindeki giriş vektörleri ağı girildikçe, ağ kendi kendini düzenler ve referans vektörleri oluşur. Bu algoritma aşağıda verilmiştir (Fauset 1994:170).

 y_{m1} y_{m2} y_{m3}

Bu algoritmada kullanılan semboller:

w_{ij} : i . satır j . sütundaki çıkış nöronuna ait referans vektörü,

x : Giriş vektörü.

$D(i,j)$: x vektörünün (i,j) koordinatındaki çıkış nöronuna olan öklid uzaklığının karesi.

I, J : x vektörünün en yakın olduğu çıkış nöronun koordinatları.

α : öğrenme katsayısı.

Algoritma:

0. w_{ij} katsayılarına ilk değer ata.

Topolojik komşuluk parametrelerini belirle.

Öğrenme katsayı parametrelerini ayarla.

1. Bitiş şartı yanlışken adım 2-8 i takip et.

2. Her bir x giriş vektörü için adım 3-5 i takip et.

3. Her bir i, j için $D(i,j) = \sum_{i,j} (w_{ij} - x)^2$ oklid uzaklık değerlerini

hesapla.

4. $D(I,J)$ 'nin minimum olduğu I, J değerini bul.

5. I, J 'nin belirtilen komşuluğundaki bütün çıkış nöronları için

$w_{ij}(\text{yeni}) = w_{ij}(\text{eski}) + \alpha (x - w_{ij}(\text{eski}))$

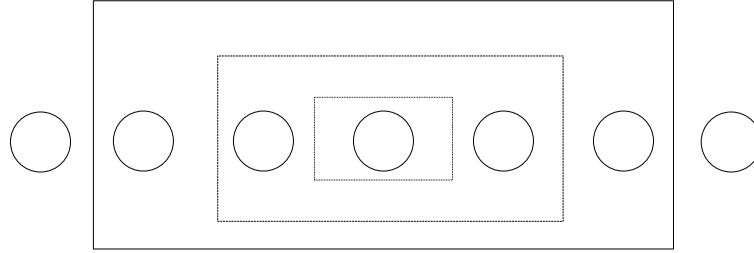
6. Öğrenme katsayısını güncelle.

7. Belirtilen zamanlarda topolojik komşuluk parametresini azalt.

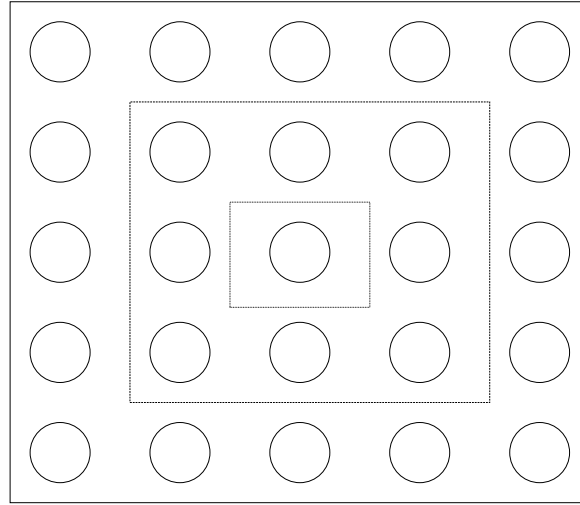
8. Bitiş şartını kontrol et.

Yukarıdaki algoritmadan da anlaşılacağı üzere, ilk önce referans vektörlerine bir ilk değer verilir. Döngüye başlamadan önce öğrenme katsayısı (α) ve komşuluk değişkenine (R) yüksek bir değer atanır. α 'ya 0 ile 1 arasında bir değer atanır. Bu değer 1'e yakın olması tercih edilir. R değişkeni ise çıkış katmanındaki dizilimin boyu veya eninden büyük olanının değeri ile başlar. Algoritma için bir döngü veri setindeki tüm satırların birer kere SOM ağına girdi olarak sunulmasıdır. Veri setinin bir satırı x vektörüdür. x vektörünün çıkış katmanındaki her bir nörona olan öklid uzaklığının karesi bulunur. Çıkış katmanındaki her bir nöronu bir referans vektörü (w_{ij}) temsil eder. Dolayısıyla, bu uzaklık x vektörü ile w_{ij} arasındaki uzaklıktır. Hesaplanan uzaklıklardan en küçüğü bulunur. Bu uzaklık hangi çıkış nöronuna aitse, o nöron kazanan (Winner neuron) nörondur. Yani, SOM ağları "yarışmacı" bir ağıdır. Kazanan nöron ve komşu nöronların referans vektörleri yeniden hesaplanır. Şekil 2.2'de kazanan nöronun doğrusal komşuluğu ve Şekil 2.3'de kazanan nöronun dikdörtgensel komşuluğu görülmektedir. Bu şekillerden de görüldüğü gibi, dikdörtgensel komşulukta kazanan nöronun etrafında daha fazla komşu nöron bulunmaktadır. Bu hesaplamada $w_{ij}(\text{yeni}) = w_{ij}(\text{eski}) + \alpha (x - w_{ij}(\text{eski}))$ eşitliği kullanılır. Burada, x vektörü ile w_{ij} referans vektörü arasındaki fark, öğrenme katsayısı α ile çarpılır ve w_{ij} referans vektörüne ilave edilir. Bu sebeple, w_{ij} referans vektörlerine ilk değer olarak çok küçük değerler verilmişse, α değeri

1'e yakın bir değer alınmalıdır. Böylece, referans vektörleri kendilerini oluşturma şansına sahip olurlar. Böylece, veri setindeki tüm satırlar için bu işlemler tekrarlandığında bir döngü tamamlanmış olur. Döngüler devam ettikçe referans vektörleri değişmeye devam eder. Döngünün belirli periyodlarında α ve R değerleri azaltılır. Kaç döngüde bir bu değişkenlerin azaltılacağı kesin kurallara bağlanmamıştır. Bu konuda değişik görüşler vardır. Çoğu zaman bu değişkenlerin doğrusal bir fonksiyonla azaltılması yeterli olur. Referans vektörlerindeki değişim sona erdiğinde döngü de sona ermiş olur.



Şekil 2.2: Kazanan Nöronun (#) Doğrusal Komşuluğu
(İçten Dışa Doğru Sırasıyla $R=0$, $R=1$ ve $R=2$).



Şekil 2.3: Kazanan nöronun (#) Dikdörtgensel Komşuluğu
(İçten Dışa Doğru Sırasıyla $R=0$, $R=1$ ve $R=2$).

Ağın eğitimi tamamlanıp referans vektörleri oluştuktan sonra veri setindeki elemanlar kümelere ayrılır. Veri setindeki tüm satırlar art arda ağa girilir. Giriş vektörü çıkış nöronlarının referans vektörleriyle çarpılır. Hangi çarpım değeri daha büyükse, eleman o kümeye aittir. Bu işlem sonucunda,

elemanlar hem kümelenmiş hem de iki boyutlu bir harita üzerine yerleştirilmiş olur. Bu haritadan, birbirine yakın ve uzak olan kümeleri veya elemanları görmek mümkündür. Çıkış katmanındaki nöron dizilimi 3 boyutlu ise harita da 3 boyutlu olacaktır. Bu haritalar, kümelerin özelliklerine göre değişik şekillerde renklendirilip gölgelendirilebilir. Böylece, daha görsel bir harita elde edilebilir.

IV. SOM Modelinde Kümelemeyi Etkileyen Faktörler

Ağ yapısı ve öğrenme algoritması yukarıda açıklanan bir SOM modelinde başarılı bir kümeleme çalışması gerçekleştirebilmek için bazı faktörlere dikkat etmek gerekir. Ancak, bunlar kesin kurallara bağlı olmayıp sadece bir çerçeve çizmek için verilecektir. Çoğu kez deneme yanılma yoluyla bu faktörler için en iyi değerler bulunur. Giriş vektörüyle referans vektörleri arasındaki fark hata olarak kabul edilirse, en küçük mutlak hata ortalamasına sahip model en iyidir denilebilir (Kohonen 2001:105-176). Bu faktörler şunlardır:

- 1) Çıkış katmanındaki nöron sayısı:
Çıkış katmanındaki nöron sayısı, elde edilebilecek maksimum küme sayısını belirtir. Genellikle, veri setindeki eleman sayısının %10'u civarında çıkış nöronu tercih edilir.
- 2) Verilerin normalleştirilmesi:
Veri setindeki değişken değerleri arasında büyük farklar varsa değerler normalleştirilmelidir. Böylece, tüm değişkenlere eşit şans verilmiş olur. Sütun bazında yapılan normalleştirme üç şekilde yapılabilir: a) Her bir sütundaki değerler sütun standart sapmasına bölünür. b) Her bir sütundaki değerlerden sütun ortalaması çıkartılıp çıkan sonuç sütun standart sapmasına bölünür. Böylece, ortalaması 0 standart sapması 1 olan standart dağılım elde edilir. c) Her bir sütundaki değer, sütun vektör uzunluğuna bölünür (Graupe 1997).
- 3) Referans vektörlerine ilk değer atanması:
Referans vektörlerine ilk değer atanması, SOM modelinde çok kritik bir yere sahiptir. Bu ilk değerler atanırken veri setindeki giriş vektör değerleri göz önünde bulundurulmalıdır. Pratikte, referans vektörlerine rasgele değerler atamak bazen sakıncalı olabilir. Tüm vektörlere 0'dan çok az büyük bir değer atanırsa öğrenme katsayısı 1'e yakın bir değerle başlatılmalı ve belli bir döngü boyunca (1000 döngü gibi) azaltılmamalıdır. Ayrıca, komşuluk değişkeni büyük bir değerle başlatılmalı ve öğrenme katsayısının değişmediği periyotta sabit kalmalıdır. Böylece, referans vektörleri giriş vektörlerine uygun bir forma kavuşurlar. Referans vektörlerine, giriş vektörlerinin dağılımına uygun bir ilk atama yapıldığında öğrenme katsayısı ve komşuluk değişkeni daha küçük bir değerle başlatılabilir. Bu da algoritmanın öğrenme hızını artırır.

- 4) Uzaklık ölçüsü:
SOM algoritmasında giriş vektörleriyle referans vektörleri arasındaki uzaklık, öklid uzaklığının karesi ile ifade edilmektedir. Ancak, öklid uzaklığı yerine başka uzaklık ölçüleri de kullanılabilir: Minkowski ölçüsü, city-block uzaklığı gibi (Johnson, Wichern 1988:545). Bazı çalışmalarda uzaklık ölçüsü yerine vektör çarpımı kullanılmıştır (Taner 1997). Vektör çarpımı kullanıldığında, en büyük çarpım değerine sahip olan nöron, kazanan nörondur. Bununla ilgili SOM algoritmasında değişiklik yapmak gerekir.
- 5) Öğrenme katsayısı ve Komşuluk değişkeni:
4.maddede, bu değişkenlerin referans vektörlerine ilk değer atanması ile olan ilişkisinden bahsedilmişti. Öğrenme katsayısı 0 ile 1 arasında bir değerle başlamalı ve döngü arttıkça 0'a yaklaşmalıdır. Komşuluk değişkeni, çıkış katmanındaki dizilime uygun büyük bir değerle başlamalı ve döngü ilerledikçe azaltılmalıdır. Bu değişkenler genellikle lineer olarak azalan fonksiyonlarla temsil edilir.

$$\alpha(t) = \alpha(t) - \frac{A}{t + B}$$

gibi bir fonksiyon kullanılabilir. Burada t döngü sayısını simgelemektedir (Kohonen 2001). Bu fonksiyonda kullanılan A ve B değerleri ile bu değişkenler, istenilen döngü periyotlarında azaltılabilir.

V.Uygulama

Bu çalışmada, Devlet İstatistik Enstitüsünden elde edilen dış ticaret verileri kullanılmıştır. Dış ticarete ticareti yapılan 12000 civarında mal vardır. Ancak, kolay istatistik tutmak için bu mallar 98 fasıl (chapter) ve bu fasıllar da 22 bölüme (section) ayrılmıştır. Her bir bölüm çeşitli fasılların gruplanmış halidir. Dış ticaretteki bu sisteme harmonize sistem denilmektedir. Bu veriler, Türkiye'nin 2002 yılında çeşitli ülkelere yaptığı USD cinsinden ihracat tutarlarını içermektedir. Verilerin her bir satırında, ülke numarası, ülke adı ve 98 fasıl için ihracat tutarları yer almaktadır. Uygulamayı kolaylaştırmak için verilerdeki 98 fasıl, 22 bölüm şeklinde yeniden düzenlenmiştir. Verilerin düzenlenmesi ve modelin kurulması için Delphi programlama diliyle özel bir yazılım geliştirilmiştir. Tablo 3.1 de ihracat yapılan 22 bölüm görülmektedir.

2002 yılına ait Türkiye'nin ihracat verileri bölüm bazında düzenlendikten sonra normalleştirilmiştir. Çünkü, ülkelerden çeşitli mal bölümleri bazında ihracat tutarlarının değişim aralıkları birbirinden çok farklıdır. Kümeleme işleminde her bir mal bölümünün bu farklılıklardan dolayı baskın çıkması bu şekilde önlenmiştir. Normalleştirme yapılırken, önce her bir

değişkene karşı gelen sütunların aritmetik ortalaması ve standart sapması bulunmuştur. Sonra da,

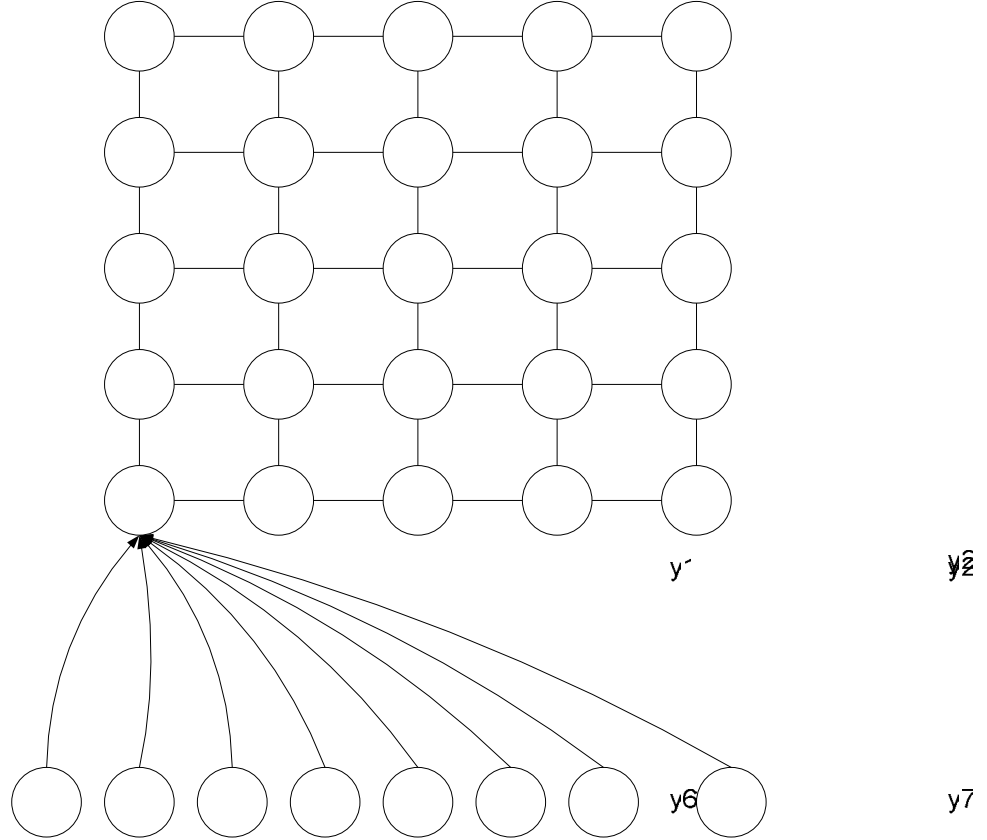
$$\frac{x - \mu}{\sigma}$$
 formülüyle veriler normalleştirilmiştir.

Tablo 3.1: *Harmonize Sistemde Fasıl Grupları (Bölümler)*
(Devlet İstatistik Enstitüsü)

BÖLÜM I:	Canlı Hayvanlar; Hayvan Ürünleri
BÖLÜM II:	Bitkisel Ürünler
BÖLÜM III:	Hayvansal/Bitkisel Katı/Sıvı Yağlar, Bitkisel Mumlar vb. Müstahsalları
BÖLÜM IV:	Hazır Yiyecekler; Meşrubat, Alkollü İçkiler Ve Sirke; Tütün Ve Tütün Yerine Geçen İşlenmiş Maddeler
BÖLÜM V:	Mineral Ürünler
BÖLÜM VI:	Kimyasal Veya Birleşik Endüstri Ürünleri
BÖLÜM VII:	Muhtelif Kimyasal Maddeler; Plastik Ve Plastikten Mamul Eşya
BÖLÜM VIII:	Ham Postlar, Deriler ve Köseleler, Postlar, Kürkler, Taklit Kürkler ve Mamulleri; Deri Eşya, Saraciye Eşyası, Seyahat Eşyası, Bağırıskattan Eşyası
BÖLÜM IX:	Ağaç ve Ağaçtan Mamul Eşya; Odun Kömürü, Mantar ve Mantardan Eşya; Hasır, Saz vb. Örülebilir Maddelerden Mamuller
BÖLÜM X:	Odun Hamuru; Lifli Selülozik Maddelerin Hamurları, Hurdalar; Kağıt ve Karton; Kağıt Hamurundan Kağıt ve Kartondan Eşya; Basılı Kitap, Gazete, Resim vb Baskı Sanayi Mamulü, El Yazmaları
BÖLÜM XI:	Mensucat ve Mensucat Eşyaları
BÖLÜM XII:	Ayakkabı, Başlıklar, Şemsiye, Baston, Kamçı, Kırbaç vb. Aksanı, Hazır Kuş Tüyü ve İnsan Saçı Mamulleri, Yapma Çiçekler
BÖLÜM XIII:	Taş, Alçı, Çimento, Amyant, Mika vb. Maddelerden Eşya; Seramik Mamulleri; Cam ve Cam Eşya
BÖLÜM XIV:	İnci, Kıymetli Taş ve Metal Mamulleri, Madeni Para
BÖLÜM XV:	Adi Metaller ve Adi Metal Eşyaları
BÖLÜM XVI:	Makine, Mekanik Cihaz ve Aletler; Elektrikli Makine ve Cihazlar, Aksam; Parçaları
BÖLÜM XVII:	Motorlu Kara Taşıtları, Hava Taşıtları, Gemiler ve İlişkili Taşıma Gereçleri
BÖLÜM XVIII:	Optik, Fotoğraf, Sinema, Ölçü, Kontrol, Ayar Cihazları, Tıbbi Alet; Saatler; Müzik Aletleri; vb. Aksam, Parça ve Aksesuarı
BÖLÜM XIX:	Silahlar Ve Mühimmat, Bunların Aksam, Parça ve Aksesuarları
BÖLÜM XX:	Çeşitli Eşyalar
BÖLÜM XXI:	Sanat Eseri, Koleksiyon Eşyası, Antikalar
BÖLÜM XXII:	Diğer Eşyalar

Veriler normalleştirildikten sonra ülkelerin kümelenmesi için bir SOM sinir ağı modeli kurulmuştur. Bu modelde, 22 giriş nöronu ve 25 çıkış nöronu kullanılmıştır. Burada, 22 giriş nöronu 22 bölüm değişkenine ve 25 çıkış nöronu olası 25 kümeye denk düşmektedir. 250 civarındaki ülkenin %10'u oranında

küme öngörülmüştür. Ancak, bütün kümelerin dolu olması gerekmemektedir. 25 çıkış nöronu 5x5'lik bir kare şeklinde dizilmiştir. Şekil 3.1'de oluşturulan SOM modeli görülmektedir.



Şekil 3.1: Türkiye'nin İhracat Yaptığı Ülkeleri Kümelemek İçin Oluşturulan SOM Sinir Ağı Modeli.

Her bir sütunu bir referans vektöründen oluşan katsayılar matrisi, giriş nöronlarıyla çıkış nöronları arasındaki ilişki gösteren bir matristir. Bu matrisin değerleri SOM öğrenme algoritması ile oluşmaktadır. Ancak, algoritmaya başlamadan önce bu matrise ilk değer atanması gerekir. Bu değer 0 dan farklı olmalıdır. Normalleştirilmiş ihracat verileri -0.4209 ile 14.5913 arasında değerler almaktadır. Bu değerlerin değişim aralığı küçük olduğu için referans vektörlerine ilk değer olarak rasgele değerler atamak uygun görülmemiştir. Bu sebeple, katsayılar matrisine "0"dan biraz büyük 0.001 gibi bir ilk değer atanmıştır. Yani, referans vektörlerinin iyi bir başlangıç değeri alması algoritmaya bırakılmıştır. Bundan dolayı, döngü sayısı olarak 10000 gibi yüksek bir değer öngörülmüştür.

SOM algoritmasında referans vektörlerinin öğrenilmesini etkileyen bir diğer önemli faktör de öğrenme katsayısıdır (α). Algoritmada yeni katsayılar aşağıdaki formüle göre hesaplanmaktadır:

$$w_{ij}(\text{yeni}) = w_{ij}(\text{eski}) + \alpha(x_i - w_{ij}(\text{eski}))$$

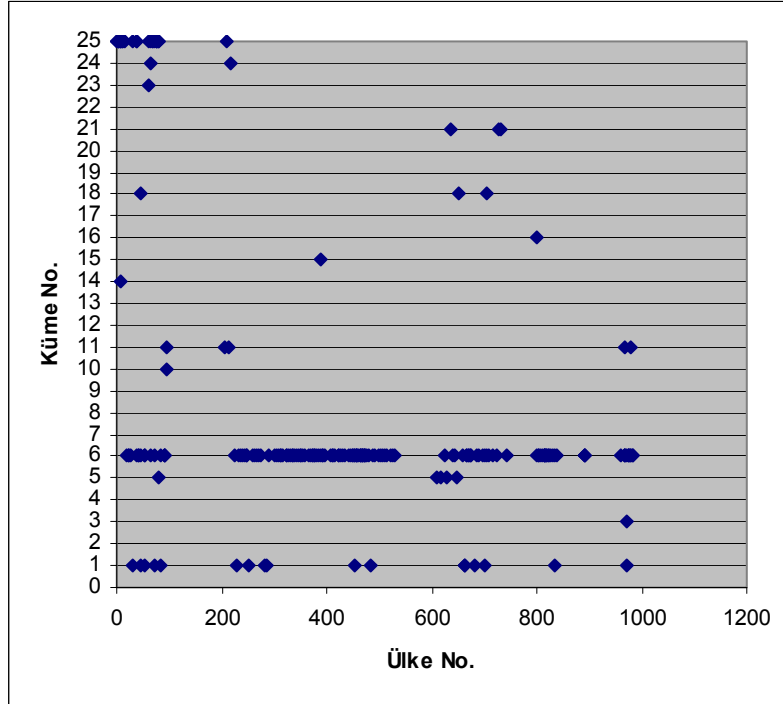
Bu formülden de anlaşılacağı gibi yeni katsayılar giriş değerleri ile eski katsayılar arasındaki farkın öğrenme katsayısı ile çarpılıp eski katsayıya eklenmesi ile bulunmaktadır. Öğrenme katsayısı 0 ile 1 arasında bir değerdir. Döngü sayısı arttıkça, bu değer belli bir oranda azaltılmalıdır. Aksi takdirde, katsayılardaki değişim sürekli devam edecektir. Bu da sonsuz döngüye neden olur. Bazı çalışmalarda, öğrenme katsayısı yerine sürekli azalan ve 0 ile 1 arasında değer alan bir fonksiyon kullanılmıştır (Rauber 1998). Referans vektörlerine ilk değer olarak 0.001 atandığı için öğrenme katsayısı 1'e yakın bir değer alınmıştır ($\alpha=0.8$). Öğrenme katsayısının yüksek olması algoritmanın yavaşlatılması içindir. İlk 1000 döngü boyunca bu değer değiştirilmemiştir. Böylece, referans vektörlerinin kendi ilk değerlerini giriş vektörlerinden oluşturması sağlanmıştır. Daha sonra her bin döngüde, α 1/10 oranında lineer olarak azaltılmıştır. 10000 döngü sonunda α , 0'a yakın bir değere çekilmiştir. Bu da, öğrenmenin sona ermesi demektir.

SOM modelinde referans vektörlerinin veya bir diğer deyişle katsayılar matrisinin değişimini etkileyen bir diğer faktör de kazanan nöronun komşuluğudur. Modelde çıkış nöronlarının dizilimi topolojik komşuluk açısından önemlidir. Çıkış nöronları doğrusal dizildiğinde, komşuluk tek bir eksen üzerinde olmaktadır. Komşuluğun hem yatay hem de dikey eksen üzerinde olabilmesi için bu çalışmada çıkış nöronları karesel olarak dizilmiştir. R'nin ilk değerine, çıkış nöronlarından oluşan karenin bir kenarındaki nöron sayısı verilmiştir. Bu çalışmada kullanılan 5x5'lik dizilime göre R=5 olacaktır. Döngü sayısı arttıkça R değeri de belli aralıklarla azaltılmıştır. Böylece algoritmanın bitiminde R=0 olacak ve yalnızca kazanan nöronun referans vektörü güncellenecektir. Bu çalışmada her bir 2000 döngüde, R değeri 1 azaltılmaktadır. Ancak, ilk 2000 döngü boyunca R değerinin sabit kalması sağlanmıştır. Böylece, ilk değerleri 0.001 verilen referans vektörlerinin iyi bir değer alması sağlanmıştır.

Bütün x vektörleri belli bir döngü sayısınca SOM modeline girildikten sonra katsayılar matrisi veya referans vektörleri kararlı bir hale gelir. Bu modelde döngü sayısı 10000 alınmıştır. 2002 yılı ihracat verileri 10000 defa modele girilmiş ve katsayıların artık değişmediği görüldüğünden yapay sinir ağının eğitiminin tamamlandığına karar verilmiştir. Yani modelin oluşumu tamamlanmıştır. Bundan sonra yapılacak işlem, hangi ülkenin hangi kümeye ait olduğunu belirlemektir. Modeldeki referans vektörleri giriş nöronları ile çıkış nöronları arasındaki ilişkiyi göstermektedir. Giriş nöronları her bir ülkenin 22 değişkenini ve çıkış nöronları da olası 25 kümeyi göstermektedir. Ancak, 25 kümenin de dolması gerekmemektedir. O halde oluşacak küme sayısı 25'den daha aşağı da olabilecektir. Kümelerin belirlenebilmesi için her bir ülkenin

değişken değerleri, modele yalnız bir kere yerleştirilir. O ülkenin x vektörü çıkış katmanındaki her bir nöronun referans vektörü ile çarpılır. Hangi çarpım değeri daha yüksek ise ülke o kümeye aittir. $m=5$ (5×5 'lik dizilim), $\alpha=0.8$, $R=5$ ve 10000 döngü sonucu Türkiye'nin ihracat yaptığı ülkeler Tablo 3.2'de kümelenmiştir.

Türkiye'nin ihracat yaptığı ülkeler ve ait oldukları kümeler Tablo 3.2'de görülmekle beraber, bu tablonun grafiksel gösterimleri ülkelerin kümelenme eğilimlerini daha net gösterecektir. Şekil 3.2'de bu tablo X-Y dağılım grafiği şeklinde görülmektedir. Bu grafikten de anlaşıldığı üzere ülkeler 6 nolu kümede yoğunlaşmaktadır. Ancak, yoğunlaşmanın çok olduğu bu küme ihracatın çok az olduğu ülkeleri kapsamaktadır. Bu tür grafikler her ne kadar yoğunlaşmayı gösterse de kümelerin topolojik komşuluğunu gösteremez. Bir diğer önemli husus da kümelerde hangi değişkenlerin daha etkin olduğudur. Kümeler üzerindeki değişken



Şekil 3.2: Türkiye ile Ticareti Olan Ülkelerin Kümelenmesini Gösteren X-Y Dağılım Grafiği

Tablo 3.2: 2002 yılına Türkiye'nin İhracat Verilerine Göre Ülkeler ve Ait Oldukları Kümeler.

Ülke Adı	Küme No.	Ülke Adı	Küme No.	Ülke Adı	Küme No.	Ülke Adı	Küme No.	Ülke Adı	Küme No.	Ülke Adı	Küme No.
FİNLANDIYA	1	SUDAN	6	ZİMBABVE	6	PERU	6	PALAU	6	HOLLANDA	25
GİNE	1	MALİ	6	MALAVİ	6	BREZİLYA	6	AMERİKAN SAMOA	6	ALMANYA	25
SİERRA LEONE	1	BURKİNA FASO	6	NAMİBYA	6	ŞİLİ	6	NORFOLK AD.	6	İTALYA	25
BEYAZ RUSYA	1	NİJER	6	BOTSVANA	6	BOLIVYA	6	COOK ADALARI	6	İNGİLTERE	25
TACİKİSTAN	1	CAD	6	SVAZİLİND	6	PARAGUAY	6	ANTARTİKA	6	DANİMARKA	25
MORİTANYA	1	CAPE VERDE	6	LESOTHO	6	URUGUAY	6	BOUVET ADALARI	6	YUNANİSTAN	25
GAMBIYA	1	SENEGAL	6	GRÖNLAND	6	ARJANTİN	6	S.GEOR.VE SAND.	6	PORTEKİZ	25
TOGO	1	GİNE	6	MEKSİKA	6	FİLİSTİN	6	ANTALYA SERB.BLG	6	İSPANYA	25
BENİN	1	SİERRA LEONE	6	BERMUDA	6	BAHREYN	6	TRAKYA S.BÖLGE	6	BELÇİKA	25
BAHAMALAR	1	LİBERYA	6	GUATEMALA	6	KATAR	6	TRABZON S.BÖLG	6	İSVEÇ	25
VENEZUELLA	1	FİLDİŞİ KIYISI	6	BELİZE	6	AFGANİSTAN	6	MARDİN S.BÖLGE	6	AVUSTURYA	25
PAKİSTAN	1	GANA	6	HONDURAS	6	BANGLADEŞ	6	YUMURTALIK S.BÖL	6	İSVİÇRE	25
HİNDİSTAN	1	NİJERYA	6	EL SALVADOR	6	MALDİVLER	6	SAMSUN S.BÖLGE	6	POLONYA	25
TAYLAND	1	KAMERUN	6	NİKARAGUA	6	SRİ LANKA	6	KAYSERİ S.BÖLGE	6	ROMANYA	25
MALEZYA	1	O. AFRIKA CUM.	6	KOSTARİKA	6	BUTAN	6	AVRUPA S..BÖLGE	6	BULGARİSTAN	25
CHRİSTMAS DALARI	1	EKVATOR GİNESİ	6	PANAMA	6	BURMA	6	G.ANTEP S..BÖLGE	6	ARNAVUTLUK	25
RİZE SERB.BÖLGE	1	SAO T. VE PRİN.	6	ANGUİLLA	6	LAOS	6	TUBİTAK TECH.S.B.	6	UKRAYNA	25
MENEMEN DERİ S.B.	3	GABON	6	KÜBA	6	VİETNAM	6	KOCAELİ S..BÖLGE	6	RUSYA	25
TÜRKMENİSTAN	5	KONGO	6	ST.KITTS VE NEV.	6	KAMBOÇYA	6	DENİZLİ S.BÖLGESİ	6	GÜRCİSTAN	25
SURİYE	5	ZAİRE	6	HAITİ	6	ENDONEZYA	6	MAKEDONYA	10	AZERBEYCAN	25
İRAN	5	RUANDA	6	DOMİNİK CUM.	6	BRUNEİ	6	YUGOSLAVYA	11	KAZAKİSTAN	25
ÜRDÜN	5	BURUNDİ	6	ABD VİRJİN AD.	6	FİLİPİNLER	6	FAS	11	ÖZBEKİSTAN	25
UMMAN	5	ANGOLA	6	ANTİGUA ve BAR.	6	MOĞOLİSTAN	6	TUNUS	11	CEZAYİR	25
LÜKSEMBURG	6	ETİOPYA	6	DOMİNİK	6	KUZEY KORE	6	MERSİN S. BÖLGE	11	MİSİR	25
CEUTA VE MELLİLA	6	ERİTRE	6	CAYMAN ADA.	6	MAKAO	6	BURSA S.BÖLGE	11	A.B.D.	25
İZLANDA	6	CİBUTİ	6	JAMAİKA	6	PA. YENİ GİNE	6	İRLANDA	14	KANADA	25
NORVEÇ	6	SOMALİ	6	ST LUCIA	6	AVUSTRALYA OK.	6	GÜNEY AFRIKA	15	K.K.TÜRK.CUM.	25
LİCHTENSTEİN	6	KENYA	6	ST VİNCENT	6	YENİ ZELANDA	6	AVUSTRALYA	16	LÜBNAN	25
FAROE AD.	6	UGANDA	6	İNGİLİZ VİRJİN A.	6	Y. KALEDONYA	6	MALTA	18	İSRAİL	25
ANDORRA	6	TANZANYA	6	BARBADOS	6	AMERİKAN OK.	6	YEMEN	18	SUUDİ ARABİSTAN	25
SAN MARİNO	6	ŞEYSELLER	6	TRİNİDAD VE TOBAGO	6	Y. ZELENDA OKY.	6	SİNGAPUR	18	B.A.E.	25
ESTONYA	6	İNGİL.OK.AD.	6	GRENADA	6	FİJİ	6	KUVEYT	21	ÇİN	25
SLOVAKYA	6	MOZAMBİK	6	ARUBA	6	VANUATU	6	GÜNEY KORE	21	TAYVAN	25
MOLDOVA	6	MADAGASKAR	6	HOLLANDA ANTİL.	6	TONGA	6	JAPONYA	21	HONG KONG	25
KIRGIZİSTAN	6	MARİTUS	6	KOLOMBİYA	6	SAMOA	6	ÇEK CUM.	23	KES.VE BİL.ÜLK.	25
SLOVENYA	6	KOMOROLAR	6	GUYANA	6	KUZEY MARİANA AD.	6	MACARİSTAN	24	AHL SERB.BÖLGE	25
HİRVATİSTAN	6	MAYOTTE	6	SURİNAM	6	FRANSIZ POLİNEZY.	6	LİBYA	24	İST.DERİ SERB.BÖL.	25
BOSNA-HERSEK	6	ZAMBİA	6	EKVATOR	6	MARSHALL ADALARI	6	FRANSA	25	EGE SERBEST BOL	25"

ağırlıklarını SOM modelinde elde ettiğimiz referans vektörlerinden bulabiliriz. Tablo 3.3’de ihracat kümelerine ait referans vektörleri görülmektedir. Bu tabloda satırlar değişkenleri sütunlar ise kümeleri göstermektedir. Tabloya dikkat edilirse 22 satır ve 25 sütun mevcuttur. Her bir sütun, ilgili kümenin referans vektörüdür.

Her bir satır, harmonize sistemdeki bölüm değişkenlerini temsil etmektedir. Her bir sütundaki değerler değişkenlerin küme üzerindeki etkilerini göstermektedir. Sütunlardaki pozitif değerler olumlu yönde, negatif değerler ise ters yönde kümelere etki etmektedir. Referans tablosundaki yüksek değerler ilgili mal grubundan o kümenin yoğun olarak ticareti olduğunu gösterir. Düşük değerler de ilgili mal grubundan o kümenin ticaretinin az olduğunu gösterir.

SOM modelindeki topolojik yapıyı ve kümelerin komşuluğunu görebilmek için kümeleri dikdörtgensel bir grid şeklinde dizmek gerekir. Şekil 3.3’de ihracat kümeleri topolojiye uygun olarak dizilmişlerdir. Her bir küme yoğunluklarına göre gölgelendirilmiştir. Hafif gölgeli bölgeler yoğunluğun az olduğunu koyu gölgeli bölgeler ise yoğunluğun çok olduğunu göstermektedir.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Şekil 3.3: *İhracat Kümelerinin Topolojik Gösterimi. (Koyu Alanlar Yoğun Açık Alanlar Seyrek Kümeleri Gösteriyor).*

Tablo 3.3: İhracat Kümelerine Ait Referans Vektörleri (Sütunlar Kümeleri, Satırlar Değişkenleri Gösteriyor).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	-0.24	-0.23	-0.13	-0.03	-0.01	-0.24	-0.24	-0.19	-0.02	0.13	-0.24	-0.24	-0.15	0.03	0.16	-0.13	-0.17	-0.07	0.24	0.84	2.22	2.09	0.42	1.42	1.85
2	-0.28	-0.28	-0.22	-0.05	0.22	-0.27	-0.27	-0.23	-0.05	0.30	-0.27	-0.27	-0.22	-0.14	0.26	-0.23	-0.23	-0.15	0.17	1.37	0.02	0.26	1.18	2.66	3.76
3	-0.24	-0.24	-0.19	0.73	3.32	-0.25	-0.25	-0.22	0.69	3.03	-0.25	-0.25	-0.21	-0.09	0.56	-0.24	-0.25	-0.22	-0.01	1.25	0.18	0.20	0.29	1.28	2.55
4	-0.30	-0.29	-0.18	0.19	0.76	-0.29	-0.29	-0.21	0.18	0.77	-0.29	-0.29	-0.23	-0.17	0.17	-0.22	-0.24	-0.18	0.05	1.58	0.70	0.79	0.55	2.33	4.10
5	-0.24	-0.24	-0.20	-0.15	0.34	-0.25	-0.25	-0.22	-0.12	0.67	-0.25	-0.25	-0.17	0.04	0.89	-0.22	-0.22	0.06	0.57	2.59	0.14	0.12	0.08	1.74	3.63
6	-0.34	-0.31	0.04	0.28	1.38	-0.40	-0.40	-0.18	0.28	1.41	-0.40	-0.40	-0.22	0.19	1.28	-0.36	-0.36	-0.16	0.40	2.57	-0.09	-0.03	0.73	2.16	3.82
7	-0.33	-0.28	0.20	0.61	0.56	-0.23	-0.23	0.22	0.61	0.58	-0.23	-0.22	0.32	0.54	0.68	-0.02	-0.02	0.09	0.12	1.75	-0.10	-0.04	0.52	1.85	3.80
8	0.11	0.15	1.09	0.35	-0.08	-0.17	-0.16	0.43	0.36	-0.04	-0.19	-0.18	0.08	0.40	0.24	-0.18	-0.18	-0.15	-0.03	1.61	-0.10	-0.11	0.53	1.70	4.13
9	-0.27	-0.26	-0.17	0.91	4.28	-0.27	-0.27	-0.21	0.87	3.88	-0.27	-0.27	-0.15	0.24	1.63	-0.25	-0.25	-0.12	0.32	1.13	-0.02	0.09	0.38	1.07	1.58
10	-0.35	-0.30	0.16	0.61	1.45	-0.34	-0.34	-0.09	0.61	1.99	-0.34	-0.33	-0.04	0.55	2.20	-0.29	-0.29	-0.14	0.43	3.01	-0.18	-0.19	0.04	0.74	1.84
11	-0.19	-0.17	-0.03	0.06	-0.07	-0.18	-0.18	-0.09	0.06	-0.04	-0.18	-0.18	-0.12	-0.09	0.01	-0.13	-0.13	-0.14	-0.05	1.39	-0.08	-0.07	0.15	1.81	4.26
12	-0.29	-0.29	-0.27	0.07	0.80	-0.29	-0.29	-0.28	0.08	1.25	-0.29	-0.29	-0.28	-0.17	0.69	-0.28	-0.28	-0.12	0.63	1.88	0.08	0.66	2.75	3.91	2.46
13	-0.29	-0.25	0.13	0.47	0.30	-0.28	-0.28	-0.07	0.48	0.61	-0.28	-0.27	0.03	0.42	0.62	-0.25	-0.25	-0.16	0.21	2.76	0.11	0.21	0.62	2.84	5.25
14	-0.16	-0.16	-0.11	-0.06	-0.02	-0.16	-0.16	-0.14	-0.06	0.22	-0.16	-0.16	-0.13	-0.03	0.41	-0.14	-0.15	-0.15	0.05	2.60	0.62	0.59	0.12	2.20	5.04
15	-0.36	-0.35	-0.25	0.14	1.11	-0.34	-0.33	-0.23	0.29	1.63	-0.33	-0.33	-0.04	0.63	1.57	-0.24	-0.20	0.62	1.67	3.39	0.80	0.85	1.05	2.59	5.20
16	-0.25	-0.21	0.11	0.51	0.72	-0.21	-0.21	0.02	0.51	0.67	-0.21	-0.20	0.08	0.27	0.14	-0.12	-0.13	-0.10	-0.06	0.97	-0.13	-0.10	0.17	1.26	3.33
17	-0.27	-0.22	0.20	0.50	0.27	-0.27	-0.26	-0.04	0.51	0.41	-0.26	-0.25	0.08	0.51	0.48	-0.24	-0.25	-0.07	0.06	1.19	-0.17	-0.15	0.11	1.02	2.86
18	-0.28	-0.23	0.23	0.81	1.24	-0.24	-0.23	0.10	0.80	1.11	-0.23	-0.22	0.16	0.48	0.49	-0.11	-0.11	-0.08	0.00	1.40	-0.06	0.03	0.28	1.93	4.17
19	-0.15	-0.15	-0.15	-0.10	0.03	-0.15	-0.15	-0.15	-0.10	0.03	-0.15	-0.15	-0.15	-0.14	-0.06	-0.15	-0.15	-0.15	-0.01	2.21	0.02	-0.01	-0.01	2.40	5.49
20	-0.28	-0.25	0.00	0.27	0.27	-0.27	-0.26	-0.10	0.26	0.58	-0.27	-0.26	-0.13	0.03	0.53	-0.22	-0.22	-0.19	0.08	1.70	0.04	0.14	0.52	1.75	2.95
21	-0.10	-0.10	-0.11	-0.10	-0.08	-0.11	-0.11	-0.11	-0.10	-0.08	-0.11	-0.11	-0.10	-0.10	-0.08	-0.10	-0.11	-0.10	-0.07	0.50	0.10	0.03	-0.05	0.68	1.92
22	-0.18	-0.15	0.13	0.36	0.15	-0.18	-0.18	-0.03	0.36	0.14	-0.18	-0.17	0.03	0.27	-0.16	-0.18	-0.18	-0.17	-0.13	0.87	-0.13	-0.15	-0.15	0.97	2.66

Tablo 3.3'de 1 ve 6 nolu kümenin referans vektör değerleri çok düşüktür. Bu kümedeki ülkeler Türkiye'nin en az ihracat yaptığı ülkelerdir. Aynı tabloda, 24 ve 25 nolu kümelerin referans vektör değerleri çok yüksektir. Bu kümedeki ülkeler de Türkiye'nin en çok ihracat yaptığı ülkelerdir. Şekil 3.3'e bakıldığında 1 ile 25 nolu kümeler birbirine en uzak konumda yer almaktadır. Kümelerin referans vektör değerleri ile topolojik komşulukları birbirini desteklemektedir. Bu da, bu çalışmada kurulan SOM modellerinin başarılı olduğunun kanıtıdır.

VI.Sonuç ve Öneriler

Bu çalışmada, ihracat modeli için 25 adet referans vektörü elde edilmiştir. Her referans vektörü olası bir kümeye karşılık gelmektedir. Ancak, uygulama sonucunda bazı kümelerin boş kaldığı görülmektedir. İhracat modelinde 14 küme elde edilmiştir. Bu kümelerin referans vektörleri incelediğinde, vektör değerleri yüksek olan kümelerin Türkiye ile ticareti yoğun olan ülkelere oluştuğu görülmektedir. Vektör değerleri düşük olan kümelerin ise Türkiye ile ticareti az olan ülkelere oluştuğu gözlemlenmiştir. Buradan, referans vektör değerleri ile o kümedeki ülkelere ait veriler arasında bir doğru orantı olduğu sonucuna varılmıştır.

Uygulama sonucu elde edilen küme haritaları da kayda değer bilgiler içermektedir. 5x5'lik bu haritalarda her bir kare bir kümeyi temsil etmektedir. İhracat modeli için bir harita oluşturulmuş ve bu harita üzerindeki kümeler yoğunluklarına göre gölgelendirilmiştir. Böylece, küme yoğunlukları ve kümelerin birbirine olan topolojik komşulukları harita üzerinde gösterilmiştir. Bu harita üzerinde, Türkiye ile ticareti en yoğun olan ülke kümeleri ile Türkiye ile ticareti en az olan ülke kümelerinin birbirine en uzak konumda olduğu belirlenmiştir. Ticaret eğilimleri birbirine yakın olan kümeler ise haritalar üzerinde birbirine en yakın konumda bulunmaktadır. Buradan, modellerdeki öğrenme süreçlerinin başarılı olduğu sonucuna varılmıştır.

İhracat verilerine göre yapılan kümeleme çalışmasında, ülkelerin bazı kümelerde yoğunlaştığı görülmüştür. Bazı ülkelerin ise tek başına küme oluşturduğu görülmektedir. Bu da belli başlı ülkelerle ticaretin yoğun olduğunu ve bir çok ülke ile de bazı mal grupları dışında önemli bir ticaretin olmadığını göstermektedir. Buradan Türkiye'nin dış ticaretinin dünya bazında homojen olmadığı ve kırılabilir bir yapıya sahip olduğu anlaşılmaktadır.

İleriki çalışmalarda 2002 yılından günümüze kadar olan tüm verileri temel alan bir çalışma yapılabilir. Her bir yılda yaklaşık 250 ülke verisi olduğu düşünülürse, toplam 1000 satırlık bir veri seti elde edilecektir. Bu da, daha büyük bir SOM haritasının kullanılabilmesine imkan verecektir. Oluşturulan bu harita üzerinde, aynı ülke verilerinin kümelere dağılımı incelenebilir. Böylece, ülkelerin yıllara göre küme değişimleri gözlemlenebilir. Ayrıca, denetimsiz SOM modelinin sonuçları, denetimli bir ağ modeli ile yeniden değerlendirilebilir.

Kaynaklar

- Cheng, B. and D. M. Titterington (1994), "Neural Networks: A Review from a Statistical Perspective.", *Statistical Science*, Vol. 9, No,1,2-54.
- Choudhary, B. and P. Bhattacharyya (2002), "Text Clustering Using Semantics.", *The Eleventh International World Wide Web Conference*.
- Creedy, J. and Vance L. Martin (1997), *Nonlinear Economic Models*, Cheltenham, UK: Edward Elgar Publishing Ltd .
- Fausett, Laurene (1994), *Fundamentals of Neural Networks*, NJ:Prentice Hall.
- Graupe, Daniel (1997), *Principles of Artificial Neural Networks*, Singapore: World Scientific Publishing.
- Honkela, T., S. Kaski, K. Lagus and T. Kohonen (1997), "WEBSOM-Self-Organizing Maps of Document Collections.", *Proceedings of WSOM'97, Workshop on Self-Organizing Maps*, Espoo, Finland, June 4-6.
- Hsu, Arthur L. and S. K. Halgamuge (2003), "Enhancement of Topology Preservation and Hierarchical Dynamic Self- Organizing Maps for Data Visualisation.", *International Journal of Approximate Reasoning*.
- Jin, H., Wing-Ho Shum, Kwong-Sak Leung and Man-Leung WONG (2003), "Expanding Self-Organizing Map For Data Visualization And Cluster Analysis.", *Information Sciences*.
- Johnson, R. A. and D. W. Wichern (1988), *Applied Multivariate Statistical Analysis*, New Jersey: Prentice Hall.
- Kangas, Jari A., T. Kohonen and Jorma T. Laaksonen (1990), "Variants of Self-Organizing Maps.", *IEEE Transactions on Neural Networks* Vol. 1.
- Kaski, S. and T. Kohonen (1995), "Exploratory Data Analysis By the Self-Organizing Map: Structures of Welfare and Poverty in the World.", *Proceedings of the Third International Conference on Neural Networks in the Capital Markets*, London.
- Kohonen, Teuvo (2001), *Self-Organizing Maps*, Berlin:Springer.
- Lagus, K., S. Kaski and T. Kohonen (2003), "Mining Massive Document Collections by the WEBSOM Method.", *Information Sciences*.
- Lagus, K., T.Honkela, S. Kaski and T. Kohonen (1996), "Self-Organizing Maps of Document Collections: A New Approach to Interactive Exploration.", *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, California: AAAI Press, Menlo Park, pages 238-243.
- Larkin, George Richard (1999), " Policy Subsystem Portfolio Management:A Neural Network Model of the Gulf of Mexico Program." (Yayınlanmamış Doktora Tezi), Virginia: Faculty of the Virginia Polytechnic Institute and State University.

- Markey, Mia K., Joseph Y. Lo, Georgia D. Tourassi, Carey E. FLOYD (2003), "Self-Organizing Map for Cluster Analysis of a Breast Cancer Database.", *Artificial Intelligence in Medicine*, Vol. 27, pp. 113–127.
- Martinetz, Thomas M., Stanislav G. Berkovich, Klaus J. Schulten (1993), "Neural Gas Network for Vector Quantization and its Application to Time-Series Prediction", *IEEE Transactions on Neural Networks*, Vol. 4, NO. 4.
- Neter, J., W. Wasserman and M. H. Kutner (1989), *Applied Linear Regression Models*, Boston:IRWIN.
- Öztemel, Ercan (2003), *Yapay Sinir Ağları*, İstanbul:Papatya Yayınevi.
- Pal, Nikhil R., James C. Bezdek and Eric C.-K. TSAO (1993), "Generalized Clustering Networks and Kohonen's Self Organizing Scheme.", *IEEE Transactions on Neural Networks*, Vol. 4, NO. 4.
- Plaehn, D., T. Fiez and D. Lundahl (2003), "An Advancement in Consumer Segmentation Using an Unsupervised Self-Organizing Neural Network.", "<http://www.camo.com/rt/news/infodec/neural>", Web-Seminar.
- Rauber, Anderas (1998), "Cluster Visualization in Unsupervised Neural Networks.", http://www.ifs.tuwien.ac.at/ifs/research/pub_html/rau_masterth96/thesis.html
- Sarle, W. S. (1994), "Neural Networks and Statistical Models", *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute.
- Taner, M. Turhan (1997), "Kohonen's Self Organizing Networks with Conscience.", "www.rocksolidimages.com/pdf/kohonen.pdf".
- Weijters, T., A. V. Bosch and H. J. Herik (1997), "Intelligible Neural Networks with BP-SOM.", *Proceedings of NAIC-97, the Ninth Dutch Conference on Artificial Intelligence*, Helmond, Netherlands: University of Antwerp.
- White, Halbert (1989), "Learning in Artificial Neural Networks: A Statistical Perspective.", *Neural Computation*, Vol. 1, No. 4, pp. 425—464.