

SOM TİPİNDE YAPAY SİNİR AĞLARINI KULLANARAK TÜRKİYE’NİN İTHALAT YAPTIĞI ÜLKELERİN KÜMELENMESİ ÜZERİNE BİR ÇALIŞMA

Metin ZONTUL*, Oğuz KAYNAR** ve Hüdaverdi BİRCAN***

ÖZET

Bu çalışmada, SOM tipinde yapay sinir ağları kullanılarak Türkiye'nin ithalat yaptığı ülkeler kümelendi. 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. Uygulama verisi olarak Türkiye'nin 2002 yılına ait ithalat verileri kullanılmıştır. Bu verilerin düzenlenmesi ve SOM sinir ağı modelinin kurulması için Delphi programlama diliyle bir yazılım geliştirilmiştir. Elde edilen sonuçlar tablo ve grafiklerle gösterilmiştir.

Anahtar kelimeler: Yapay Sinir Ağları, SOM Sinir Ağları, Kümeleme Analizi, Dış Ticaret.

ABSTRACT

By Using SOM Type Neural Network, a Study on Clustering of the Countries Turkey Imports From.

In this study, the countries Turkey imports from have been clustered by using SOM type neural network. The models established by SOM neural networks have no distribution assumption or correlation assumption for data. Also, this models gives better results when the number of rows and variables in data increase. Moreover, they give the topological neighbourhood of the clusters obtained, too. For this reason, SOM neural networks have been preferred to classical statistical clustering methods. As application data, Turkey's importing data belonging to 2002 year has been used. A software has been written by using Delphi programming language to organize the raw data and to establish the SOM model. The results obtained have been illustrated by means of tables and figures.

Key Words: Neural Networks, SOM Neural Networks, Cluster Analysis, Foreign Trade.

* Öğr. Gör. Dr. Cumhuriyet Üniversitesi, Sivas M.Y.O. Bilgisayar Prog.

** Öğr. Gör. Dr. Cumhuriyet Üniversitesi, Sivas M.Y.O. Bilgisayar Prog.

*** Yrd. Doç. Dr. Cumhuriyet Üniversitesi, İİBF, İşletme Bölümü

1. GİRİŞ

İnsan zekasını bilgisayar ortamında taklit eden bilim dalı yapay zeka (Artificial Intelligence) olarak adlandırılmaktadır. Yapay zekanın dallarından biri yapay sinir ağlarıdır (Neural Networks). Yapay sinir ağları, insan beynindeki sinir hücrelerinin (nöron) bilgisayar ortamında matematiksel ve grafiksel bir modellemesidir. Yapay sinir ağlarında kullanılan öğrenme algoritmaları da klasik bilgisayar algoritmalarından farklıdır. Bu algoritmalar insan beyninin sezgisel gücünü içinde taşırlar. Bu sebeple, bir çok bilim dalı yapay sinir ağları ile ilgilenmektedir. Genellikle, yapay sinir ağları ile oluşturulan modeller istatistiksel veri analizi veya sistem optimizasyonu için kullanılmaktadır.

1980'lerin ortalarından itibaren istatistik 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. Amaçları istatistikçileri bu konuda bilgilendirmektir. Bu araştırmacılar, yapay sinir ağları ile istatistiksel yöntemlerin birbirine alternatif gibi gözükmesine rağmen aslında aynı temaların farklı isimler altında işlendiğini, bir nörobilimcinin ne kadar istatistiksel bilgiye ihtiyacı varsa o ölçüde de bir istatistikçinin yapay sinir ağları bilgisine ihtiyacı olduğunu belirtmektedir. Sarle (1994) bir çok yapay sinir ağı modelinin istatistiksel yöntemlere uygunluk gösterdiğini belirtmiştir. Gizli katmansız ileri beslemeli ağlar lineer modellere, bir gizli katmanlı ileri beslemeli ağlar lineer olmayan regresyon modellerine, Hebbian öğrenme temel bileşenler analizine ve Kohonen SOM ağları kümeleme analizine karşılık gelmektedir. Kay ve Titterington (1999) istatistik ve yapay sinir ağlarının kesişiminde yapılan bir çok çalışmayı bir kitap haline getirmişlerdir. Bu çalışmalarda, çok değişkenli istatistikte kullanılan bir çok metot yerine yapay sinir ağları kullanılmıştır.

İstatistiksel metotların çoğunda verilerle ilgili dağılım varsayımı ve değişkenlerle ilgili varyasyon ve korelasyon varsayımları vardır (Neter, Wasserman, Kutner 1989). Sinir ağları yaklaşımının klasik istatistiksel yöntemlere göre avantajı, verilerin dağılım varsayımları ile değişkenlerle ilgili varsayımlara gereksinim duymamasıdır. Yapay sinir ağları, bazı değişkenlere ait eksik verileri de tolare 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. Geri yayılım algoritmasını kullanan ağlar genellikle doğrusal

olmayan regresyon problemleri için kullanılmaktadır. Diskriminant analizi gibi bazı sınıflama yöntemleri de regresyon modeli ile ifade edilebilir. Geri yayılım algoritması bir denetimli öğrenme (Supervised Learning) algoritmasıdır. Ağın eğitilmesinde kullanılacak veriler bağımlı ve bağımsız değişkenleri içermelidir. Verilerin bir kısmı ağın eğitimi için bir kısmı da eğitilen ağın testi için kullanılır. Test işleminde bağımsız değişkenlerden oluşan veriler sinir ağına girildiğinde elde edilen değerler eldeki bağımlı değişkenlerle yaklaşık aynı değeri veriyorsa ağ doğru eğitilmiş demektir (Patterson 1996). 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 metodunu 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 çalışmada geri yayılım ağları ile SOM ağlarını birlikte kullanmışlardır. Böylelikle, daha az sayıda gizli nöron kullanmanın mümkün olduğunu iddia etmektedirler.

Son yıllarda, SOM sinir ağıları 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.

Walter ve Schulten (1993) SOM ağıları ile bir robot motorunu kontrol etmişlerdir. Robotun hareketleri çevreden gelen etkilere göre belirlenmektedir. Bu etkilere karşı ne tür tepkiler vereceğine kendi kendine öğrenen bir sinir ağı modeli ile karar verilmektedir. Moshou, Hostens, Papaioannou ve Ramon (2000) SOM ağılarını EMG sinyallerinin analizinde kullanmışlardır. EMG sinyalleri kas aktivitelerine bağlı olarak üretilen sinyallerdir. Bu sinyaller sinir ağı modeline girdi olarak verilmiş ve kas aktiviteleri sınıflandırılmıştır.

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 simülasyon verileri üzerinde SOM algoritmasıyla karşılaştırmaktadır. ESOM, SOM’dan farklı olarak öğrenme kuralına bir de genişleme katsayısı faktörünü eklemektedir.

İstatistik bilimi dışında yapay sinir ağılarıyla yapılan çalışmaların çoğu model tanıma, sınıflama ve bağımlı değişkenlerin tahmini gibi problemlerin çözümü ile ilgilidir. Dolayısıyla, direkt olarak belirtilmese de yapılan çalışmalar istatistikseldir. 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 ağılarını 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 ekonomi, 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

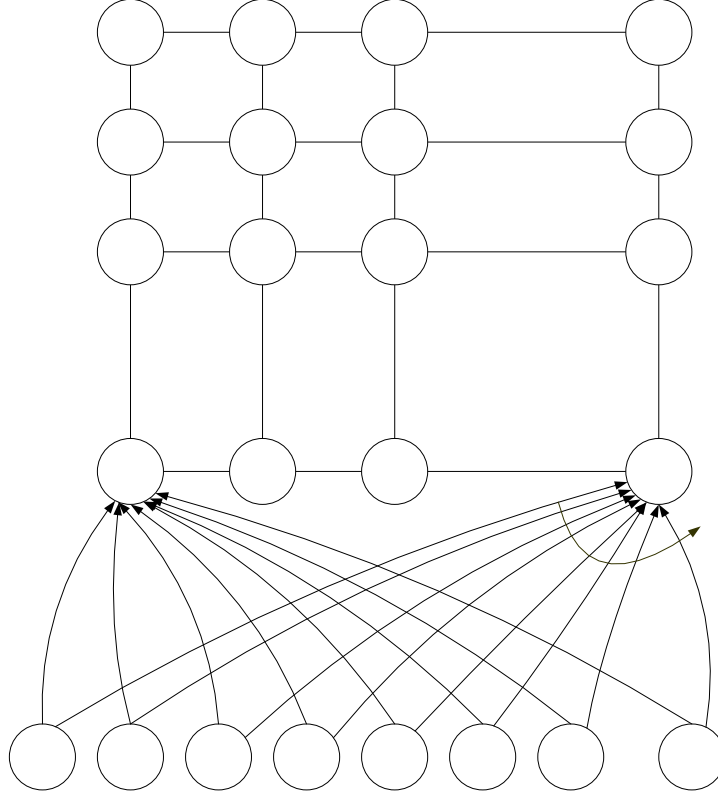
üzerinde SOM ile kümeleme analizi çalışması gerçekleştirmişlerdir. Plaehn ve Laundahl (2003) bir SOM ağıyla müşterileri homojen gruplara ayırmışlardır. Bu araştırmacılar, SOM modeli ile yapılan kümeleme çalışmasının diğer klasik istatistiksel yöntemlere göre daha iyi sonuç verdiğini göstermişlerdir.

Bu çalışmanın amacı, yapay sinir ağları ile bir kümeleme çalışması gerçekleştirmek ve bu konuda bir yazılım geliştirmektir. Bu çalışmada, Devlet İstatistik Enstitüsü Dış Ticaret Şubesi'nden alınan Türkiye'nin 2002 yılına ait ithalat verileri kullanılmıştır. Bu veriler, geliştirilen bir yazılım ile öncelikle bir veritabanında toplanmış ve sonrada SOM tipinde bir yapay sinir ağı ile modellenmiştir. Bu SOM sinir ağı modeliyle, Türkiye'nin ithalat yaptığı ülkeler harmonize sistemdeki fasıl gruplarında (bölümler) yapılan ticaret tutarlarına göre kümelendiği. Ülkelerin kümeleneğinde etkili olan faktörler belirtilmiştir.

2.SOM SİNİR AĞLARI

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, bu ağlar 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.

2.1 SOM ÖĞRENME ALGORİTMASI

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 ağa girildikçe ağ kendi kendini düzenler ve referans vektörleri oluşur. Bu algoritma aşağıda verilmiştir: (Fauset 1994:170):

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.

Y11

Y21

Y61

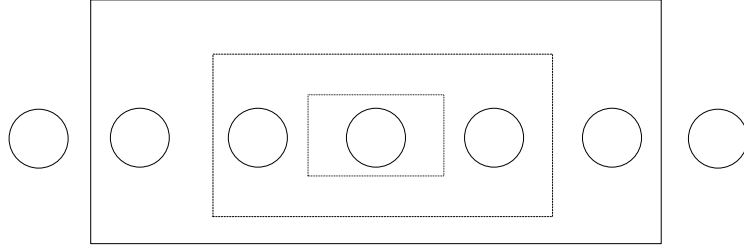
I, J : x vektörünün en yakın olduğu çıkış nöronun koordinatları.
 α : öğrenme katsayısı.

Algoritma:

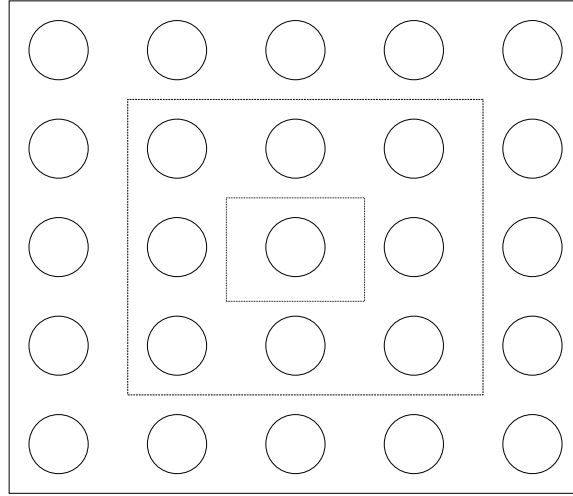
0. w_{ij} katsayılarına ilk değer ata.
 Topolojik komşuluk parametrelerini belirle.
 Öğrenme katsayı parametrelerini ayarla.
1. Bitiş şartı yanlışken adım 2-8 i takip et.
 2. Her bir x giriş vektörü için adım 3-5 i takip et.
 3. Her bir i,j için $D(i,j) = \sum_{i,j} (w_{ij} - x)^2$
 oklid uzaklık değerlerini 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 algorithmandan 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 oklid 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) nöronudur. 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 periyotlarında 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 (#) dikdörtgenel 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 kümelenecek 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.

2.2 SOM MODELİNDE KÜMELEMİYİ ETKİLEYEN 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.

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

3. 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 denmektedir. Bu veriler, Türkiye'nin 2002 yılında çeşitli ülkelerden yaptığı USD cinsinden ithalat tutarlarını içermektedir. Verilerin her bir satırında, ülke numarası, ülke adı ve 98 fasıl için ithalat 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 ithalatı yapılan 22 bölüm görülmektedir.

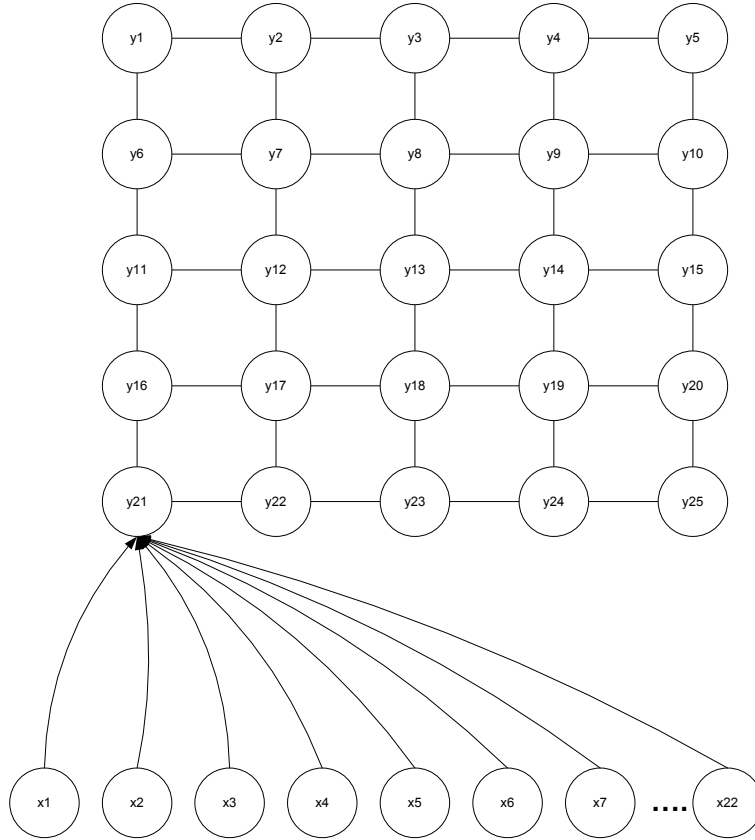
Tablo 3.1 Harmonize sistemde fasıl grupları (Bölümler) (Devlet İstatistik Enstitüsü).

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, Bağırsaktan 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 Mamulu, El Yazmaları
BÖLÜM XI:	Mensucat Ve Mensucat Eşyaları
BÖLÜM XII:	Ayakkabı, Başlıklar, Şemsiye, Baston, Kamçı, Kırbaç Vb. Aksamı, 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 Makina 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 Aksesuarları
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

2002 yılına ait Türkiye'nin ithalat verileri bölüm bazında düzenlendikten sonra normalleştirilmiştir. Çünkü, ülkelerden çeşitli mal bölümleri bazında ithalat 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.

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'lık bir dikdörtgen şeklinde dizilmiştir. Şekil 3.1'de oluşturulan SOM modeli görülmektedir.



Şekil 3.1 Türkiye'nin ithalat yaptığı ülkeleri kümelemek için oluşturulan SOM sinir ağı modeli.

Normalleştirilmiş ithalat verileri -0.3200 ile 14.5435 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.

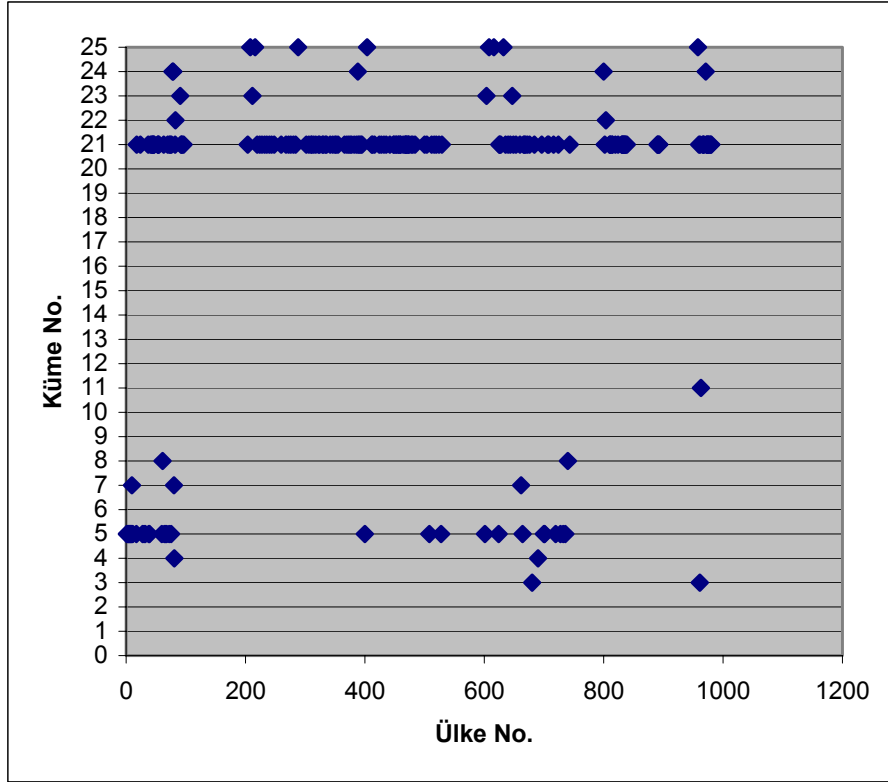
Referans vektörlerine ilk değer olarak 0.001 atandığı için öğrenme katsayısı 1'e yakın bir değer alınmıştır ($\alpha=0.8$). Öğrenme katsayısının yüksek olması algoritmanın yavaşlatılması içindir. İlk 1000 döngü boyunca bu değer değiştirilmemiştir. Böylece, referans vektörlerinin kendi ilk değerlerini 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.

Komşuluğun hem yatay hem de dikey ekseninde 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 ithalat 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. $m=5$ (5x5'lik dizilim), $\alpha=0.8$, R=5 ve 10000 döngü sonucu Türkiye'nin ithalat yaptığı ülkeler Tablo 3.2'de kümelenebilir.

Tablo 3.2 2002 yılına Türkiye'nin ithalat 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.
TAYLAND	3	UKRAYNA	5	FAROE AD.	21	MALİ	21	SOMALİ	21	PANAMA	21
AHL SERB.BÖLGE	3	RUSYA	5	ANDORRA	21	BURKİNA FASO	21	KENYA	21	KÜBA	21
ÖZBEKİSTAN	4	A.B.D.	5	CEBELİTARİK	21	NİJER	21	UGANDA	21	ST.KITTS VE NEVIS	21
VIETNAM	4	BREZİLYA	5	VATİKAN	21	CAD	21	TANZANYA	21	KUVEYT	21
FRANSA	5	ARJANTİN	5	MALTA	21	SENEGAL	21	ŞEYSELLER	21	KIRGIZİSTAN	22
HOLLANDA	5	K.K.TÜRK.CUM.	5	SAN MARİNO	21	GİNE	21	MOZAMBİK	21	YENİ ZELANDA	22
ALMANYA	5	İSRAİL	5	ESTONYA	21	LİBERYA	21	MADAGASKAR	21	SLOVENYA	23
İTALYA	5	HİNDİSTAN	5	LETONYA	21	FİLDİŞİ KIYISI	21	MARİTUS	21	TUNUS	23
İNGİLTERE	5	ENDONEZYA	5	LİTVANYA	21	GANA	21	KOMOROLAR	21	LÜBNAN	23
İRLANDA	5	MALEZYA	5	SLOVAKYA	21	TOGO	21	ZAMBİA	21	B.A.E.	23
DANİMARKA	5	ÇİN	5	ARNAVUTLUK	21	BENİN	21	ZİMBABVE	21	AZERBEYCAN	24
YUNANİSTAN	5	GÜNEY KORE	5	BEYAZ RUSYA	21	KAMERUN	21	MALAVİ	21	KAZAKİSTAN	24
İSPANYA	5	JAPONYA	5	MOLDOVA	21	ORTA AFRIKA	21	NAMİBYA	21	GÜNEY AFRIKA	24
BELÇİKA	5	TAYVAN	5	GÜRCİSTAN	21	EKVATOR GİN.	21	BOTSVANA	21	AVUSTRALYA	24
NORVEÇ	5	PORTEKİZ	7	TACİKİSTAN	21	SAO TOME	21	SVAZİLİLAND	21	MENEMEN S.BÖLGE	24
İSVEÇ	5	TÜRKMENİSTAN	7	HIRVATİSTAN	21	GABON	21	LESOTHO	21	CEZAYİR	25
FİNLANDİYA	5	PAKİSTAN	7	BOSNA-HERSEK	21	KONGO	21	MEKSİKA	21	LİBYA	25
AVUSTURYA	5	ÇEK CUM.	8	YUGOSLAVYA	21	ZAİRE	21	BERMUDA	21	NİJERYA	25
İSVİÇRE	5	HONG KONG	8	MAKEDONYA	21	RUANDA	21	GUATEMALA	21	KANADA	25
POLONYA	5	EGE SERBEST BÖL.	11	FAS	21	ST HELENE	21	HONDURAS	21	SURİYE	25
MACARİSTAN	5	LÜKSEMBURG	21	MISIR	21	ANGOLA	21	EL SALVADOR	21	İRAN	25
ROMANYA	5	İZLANDA	21	SUDAN	21	ETİOPYA	21	NİKARAGUA	21	SUUDİ ARABİSTAN	25
BULGARİSTAN	5	LİECHTENSTEİN	21	MORİTANYA	21	ERİTRE	21	KOSTARİKA	21	KES.VE BİL.ÜLK.	25



Şekil 3.2 Türkiye ile ticareti olan ülkelerin kümelenmesini gösteren X-Y dağılım grafiği.

Türkiye'nin ithalat 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 21 nolu kümede yoğunlaşmaktadır. Ancak, yoğunlaşmanın çok olduğu bu küme ithalatın çok az olduğu ülkeleri kapsamaktadır. Bu tür grafikler her ne kadar yoğunlaşmayı gösterse de kümelerin topolojik komşuluğunu gösteremez. Bir diğer önemli hususta kümelerde hangi değişkenlerin daha etkin olduğudur. Kümeler üzerindeki değişken ağırlıklarını SOM modelinde elde ettiğimiz referans vektörlerinden bulabiliriz. Tablo 3.3'de ithalat 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.

Table 3.3 İthalat 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	-0.27	-0.25	0.41	2.03	2.52	-0.29	-0.29	0.01	0.52	0.98	-0.30	-0.30	-0.29	-0.19	0.06	-0.31	-0.31	-0.14	0.03	-0.17	-0.31	-0.31	-0.06	0.04	-0.22
2	-0.13	-0.14	0.00	1.53	2.42	-0.19	-0.19	-0.12	0.19	0.72	-0.21	-0.21	-0.19	-0.08	0.05	-0.21	-0.21	-0.06	0.19	0.06	-0.21	-0.21	0.00	0.27	0.19
3	-0.06	-0.12	-0.14	1.04	1.66	-0.15	0.21	0.30	0.98	1.31	-0.15	-0.14	0.16	1.30	1.31	-0.16	-0.16	0.16	0.52	1.13	-0.16	-0.16	-0.09	-0.05	0.07
4	-0.19	-0.17	-0.10	1.46	2.38	-0.21	-0.20	-0.19	0.27	0.80	-0.22	-0.22	-0.21	0.13	0.34	-0.23	-0.23	-0.21	-0.03	0.16	-0.23	-0.23	-0.15	-0.03	0.17
5	0.01	-0.05	0.27	0.02	0.04	-0.19	-0.19	-0.15	-0.11	-0.07	-0.19	-0.19	-0.19	-0.16	-0.09	-0.19	-0.19	-0.15	0.89	1.71	-0.19	-0.19	-0.15	1.46	2.71
6	-0.17	-0.18	0.26	1.53	2.59	-0.22	-0.21	0.03	0.65	1.10	-0.25	-0.25	-0.17	0.42	0.52	-0.27	-0.27	-0.20	-0.03	0.04	-0.27	-0.27	-0.23	-0.11	-0.04
7	-0.03	0.03	0.37	1.37	2.15	-0.10	-0.01	0.15	0.89	1.22	-0.20	-0.20	-0.07	0.76	0.75	-0.25	-0.25	-0.16	0.05	0.34	-0.25	-0.25	-0.20	-0.14	-0.02
8	-0.22	-0.22	0.08	1.00	1.81	-0.18	-0.18	-0.04	0.16	0.73	-0.23	-0.23	-0.20	-0.16	0.09	-0.25	-0.17	1.65	1.96	-0.10	-0.25	-0.17	1.67	2.00	0.01
9	-0.01	0.00	0.29	0.63	0.88	-0.24	-0.17	-0.03	0.22	0.35	-0.27	-0.27	-0.20	0.06	0.03	-0.27	-0.27	-0.22	-0.15	-0.01	-0.27	-0.27	-0.26	-0.24	-0.22
10	-0.15	-0.17	0.03	0.92	1.79	-0.21	-0.20	-0.16	0.01	0.49	-0.24	-0.24	-0.22	-0.06	0.11	-0.25	-0.25	-0.23	-0.20	-0.09	-0.25	-0.25	-0.24	-0.20	-0.12
11	1.22	1.27	1.80	2.06	2.94	1.13	1.28	1.64	1.86	2.16	0.22	0.22	0.56	1.40	1.22	-0.28	-0.28	0.00	0.49	0.96	-0.28	-0.28	-0.10	-0.17	-0.16
12	0.62	0.64	1.38	2.52	3.05	0.44	0.52	0.99	1.62	1.71	-0.14	-0.14	0.01	0.27	0.12	-0.16	-0.16	-0.06	0.01	0.01	-0.16	-0.16	-0.16	-0.15	-0.13
13	-0.10	-0.10	0.31	1.36	2.25	-0.15	-0.12	0.13	0.60	1.03	-0.07	-0.07	0.06	0.26	0.37	-0.22	-0.22	-0.03	-0.08	0.08	-0.22	-0.22	-0.07	-0.17	-0.11
14	-0.08	-0.08	-0.05	0.05	0.12	-0.08	-0.07	-0.06	0.11	-0.02	-0.09	-0.09	-0.08	0.04	-0.06	-0.09	-0.09	-0.09	-0.08	-0.07	-0.09	-0.09	-0.09	-0.08	-0.08
15	0.00	-0.06	0.46	1.00	1.56	-0.22	-0.22	0.00	0.48	0.73	-0.25	-0.25	-0.17	0.34	0.40	-0.26	-0.26	-0.24	-0.18	-0.12	-0.26	-0.26	-0.24	-0.20	-0.14
16	0.04	0.06	0.80	2.25	3.23	0.01	0.04	0.49	1.47	1.87	-0.15	-0.15	0.06	1.12	1.16	-0.23	-0.23	-0.13	-0.06	0.09	-0.23	-0.23	-0.16	-0.22	-0.20
17	-0.16	-0.15	0.29	1.47	2.26	-0.17	-0.16	0.09	0.98	1.44	-0.17	-0.17	-0.05	0.91	1.07	-0.18	-0.18	-0.17	-0.01	0.21	-0.18	-0.18	-0.17	-0.16	-0.15
18	-0.12	-0.10	0.89	3.24	4.33	-0.01	0.02	0.65	1.93	2.43	-0.17	-0.17	0.08	1.36	1.41	-0.21	-0.21	-0.15	-0.10	-0.03	-0.21	-0.21	-0.19	-0.19	-0.19
19	-0.18	-0.15	0.08	1.06	1.56	-0.18	-0.18	-0.16	0.07	0.62	-0.18	-0.18	-0.18	0.01	0.43	-0.18	-0.18	-0.18	-0.10	0.18	-0.18	-0.18	-0.17	-0.18	-0.17
20	0.54	0.55	1.94	3.49	4.36	0.54	0.56	1.49	2.25	2.60	-0.05	-0.05	0.18	0.80	0.87	-0.21	-0.21	-0.15	-0.10	0.00	-0.21	-0.21	-0.19	-0.20	-0.19
21	-0.16	-0.16	0.61	1.86	2.64	0.09	0.09	0.53	1.10	1.42	-0.14	-0.14	0.07	0.82	0.94	-0.16	-0.16	-0.09	-0.14	-0.11	-0.16	-0.16	-0.08	-0.16	-0.15
22	-0.15	-0.12	-0.06	1.28	2.10	-0.15	-0.15	-0.14	1.63	2.30	-0.06	-0.06	-0.02	1.49	1.80	-0.15	-0.15	-0.06	0.63	1.64	-0.15	-0.15	-0.06	-0.15	-0.15

SOM modelindeki topolojik yapıyı ve kümelerin komşuluğunu görebilmek için kümeleri dikdörtgensel bir grid şeklinde dizmek gerekir. Şekil 3.3’de ithalat kümeleri topolojiye uygun olarak dizilmiştir. 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 İthalat kümelerinin topolojik gösterimi.

Tablo 3.3’de 5 nolu kümenin referans vektör değerleri gayet yüksektir. Bu kümedeki ülkeler Türkiye’nin yoğun olarak ithalat yaptığı ülkelerdir. Aynı tabloda, 21 nolu kümenin referans vektör değerleri çok düşüktür. Bu kümedeki ülkeler de Türkiye’nin en az ithalat yaptığı ülkelerdir. Şekil 3.3’e bakıldığında 5 ile 21 nolu kümeler birbirine en uzak konumda yer almaktadır. Kümelerin referans vektör değerleri ile topolojik komşulukları birbirini desteklemektedir. Bu da, bu çalışmada kurulan SOM modellerinin oldukça başarılı olduğunu kanıtlar.

4. SONUÇ VE ÖNERİLER

Bu çalışmada, ithalat modeli için 25 adet referans vektörü elde edilmiştir. Her referans vektörü olası bir kümeye karşılık gelmektedir. Ancak, uygulama sonucunda bazı kümelerin boş kaldığı görülmektedir. İthalat modelinde 11 küme elde edilmiştir. Bu kümelerin referans vektörleri incelendiğinde, 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. İthalat modeli için bir harita oluşturulmuş ve bu harita üzerindeki kümeler yoğunluklarına göre gölgelendirilmiştir. Böylece, küme yoğunlukları ve kümelerin birbirine olan topolojik komşulukları bir 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.

İthalat verilerine göre yapılan kümeleme çalışmasında, ülkelerin bazı kümelerde yoğunlaştığı görülmüştür. Bazı ülkelerin ise tek başına küme oluşturduğu görülmektedir. Bu da belli başlı ülkelere 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 kırılmalı bir yapıya sahip olduğu anlaşılmaktadır.

İleriki çalışmalarda 1993 ile 2002 arasındaki tüm verileri temel alan bir çalışma yapılabilir. Elde edilecek harita üzerinde, aynı ülke verilerinin kümelere dağılımları incelenebilir. Böylece, ülkelerin yıllara göre küme değişimleri gözlenebilir. Ayrıca, denetimsiz SOM modelinin sonuçları denetimli bir ağ modeli ile yeniden değerlendirilebilir. Oluşacak bu model, SOM sınır ağları ile geri yayılım ağlarının birlikte kullanıldığı karma bir model olabilir.

Bu çalışmada geliştirilen yazılım benzer ve başka uygulamalar için de kullanılacak şekilde genel yazılmıştır. Bu yazılımda, veri setindeki değişkenler seçilerek yeni veri setleri oluşturulabilmektedir. Geliştirilen yazılım, simülasyon çalışmalarında da kullanılabilir. Çeşitli kesikli ve sürekli dağılımlara göre rasgele sayı üreten bir modül yazılıma eklendiğinde veri setleri istenilen büyüklükte üretilebilir.

KAYNAKÇA

- 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 Hierchical Dynamic Self- Organizing Maps for Data Visulisation.", *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.
- Kay, J. W. and D.M. Titterington (1999), *Statistics and Neural Networks*, New York:Oxford University Pres.
- 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.
- Moshou, D., I. Hostens, G. Papaioannou, and Herman RAMON (2000), “Wavelets and Self-Organizing Maps in Electromyogram (EMG) Analysis.”, *European Symposium on Intelligent Techniques*, Aachen, Germany.
- 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.
- Patterson, Dan W (1996), *Artificial Neural Networks*, New York:Prentice Hall.
- 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”.
- Walter, Jorg A. and Klaus J. Schulten (1993), “Implementation of Self-Organizing Neural Networks for Visuo-Motor Control of an Industrial Robot.”, *IEEE Transactions On Neural Networks*, Vol. 4, NO. 1.

- 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.