ISSN: 1309 - 0712
İşletme Araştırmaları Dergisi

Journal Of
Business Research
Turk
www.isarder.org

# İnsani Kalkınmışlık Düzeyinin Sınıflandırma Başarılarının Karşılaştırılmasında Sıralı Lojistik Regresyon Analizi ve Yapay Sinir Ağları Yöntemlerinin Kullanılması[1]

*Comparison Of Classification Success Of Human Development Index By Using Ordered Logistic Regression Analysis And Artificial Neural Network Methods*

**Emre YAKUT**
Osmaniye Korkut Ata Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, Osmaniye, Türkiye
emreyakut@osmaniye.edu.tr

**Murat GÜNDÜZ**
Uşak Üniversitesi İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, Uşak Türkiye
murat.gunduz@usak.edu.tr

**Ayhan DEMİRCİ**
Türk Silahlı Kuvvetleri
kho1993@hotmail.com.tr

**Özet**

İktisadi kalkınma ve büyüme, çoğu ülke açısından en önemli hedefler arasında ön sırada yer almaktadır. Bu hedefin gerçekleştirilmesinde sadece ekonomik kalkınma değil, insanların yaşam becerilerinin artırılması ve daha kaliteli bir hale getirilmesi süreci olarak ifade edilen insani kalkınma da önemli bir rol oynamaktadır. Bu nedenle ülkeler arası karşılaştırma ve sınıflandırmaya olanak sağlaması sebebiyle insani kalkınma endeksi tercih edilen ve başvurulan sayısal bir gösterge olmuştur. Bu araştırmanın amacı sıralı lojistik regresyon ve yapay sinir ağları kullanarak İnsani Kalkınma Endeksi'nin sınıflandırma başarılarının karşılaştırılmasıdır. Uygulamada Birleşmiş Milletler Kalkınma Programı İnsani Kalkınma Endeksi'ne sahip olan 81 ülkenin 2010-2012 yılları arasındaki verileri kullanılmıştır. Sınıflandırma işlemi ülkeleri, çok yüksek, yüksek ve orta insani kalkınmışlık sınıfı olarak sınıflandırılmıştır. Yapılan analizler neticesinde sıralı lojistik regresyon modeli sonuçları, bebek ölüm oranı, sağlık harcamaları, internet kullanıcı sayısı, ithalat ve ihracat değişkenlere ait belirleyicilerinin istatistiksel olarak anlamlı olduğunu göstermektedir. Analizler olarak karşılaştırma yapıldığında Sıralı Lojistik Regresyon Analizinde % 88,1'lik başarı gerçekleştirirken, buna karşın Çok katmanlı Yapay Sinir Ağları Analizi Modeli % 97,1'lik doğru sınıflandırma başarısı göstermiştir.

**Anahtar Kelimeler:** İnsani Kalkınma Endeksi, Sıralı Lojistik Regresyon, Yapay Sinir Ağları

---

[1] Bu çalışma "Birleşmiş Milletler Kalkınma Programı Verileriyle Sıralı Lojistik Regresyon ve Yapay Sinir Ağları Yöntemlerini Kullanarak İnsani Kalkınma Endeksinin Sınıflandırma Başarılarının Karşılaştırılması" başlıklı çalışma ile, 15th International Symposium on Econometrics, Operations Research and Statistics, 22-25 Mayıs 2014, Isparta'da bildiri olarak sunulmuştur.

***Abstract***

*Economic development and growth are among the most important objectives for many countries. Not only economic development but also human development, which means enhancing and improving people's quality of life, plays an important role for reaching this objective. For this reason, human development index has become a widely preferred and recognized numerical indicator for comparison and classification of countries. The purpose of this research is to compare the classification success of Human Development Index by using ordered logistic regression and artificial neural network. The data of 81 countries, which has United Nations Development Program's Human Development Index, between the years of 2010-2012 were used in this study. Countries are classified for having very high, high and moderate levels of human development. The results of the ordered lojistic regression model indicate that determinants including infant mortality rate, health expenses, number of internet users, import and export were observed as statistically significant. As a comparison of the analysis, Ordered Logistic Regression Analysis proved 88,1 % success in classification while Multilayer Neural Networks Model showed 97,1 % success.*

***Keywords:*** *Human Development Index, Ordered Logistic Regression, Artificial Neural Network*

## GİRİŞ

İnsani kalkınma, insanların yaşam becerilerinin artırılması ve daha kaliteli bir hale getirilmesi sürecidir. Bu süreçte amaçlanan elde edilen becerilerin ve sağlanan kapasitenin ekonomik, sosyal, siyasal ve kültürel alanlarda insana ve onun yaşam standardına olumlu bir katkı sunmaktır (UNDP 1990, s.1).

Birleşmiş Milletler Kalkınma Programı (UNDP) 1990 yılında geliştirdiği *İnsani Kalkınma Endeksi*'nde (İKE), kalkınmanın yalnızca ekonomik boyutunu ele almamış, böylece sadece ekonomik büyümeye odaklanmış dar bir bakış açısından sıyrılarak kalkınmanın eğitim, sağlık ve refah boyutlarını kapsama birlikte dahil ederek daha kompozit ve insan odaklı bir yaklaşım sergilemiştir (Lind, 1992, s.89).

2010 yılına kadar endeks kapsamında ekonomik boyut olarak satın alma gücü paritesine göre hesaplanan kişi başına düşen GSYİH, sağlık boyutu olarak doğumdan itibaren ortalama yaşam süresi ve eğitim boyutu olarak da okuryazarlık oranı ve okullaşma oranı hesaplamaya dahil edilmiştir. İKE (HDI) bu üç boyutun aritmetik ortalaması olarak hesaplanmıştır. Ekonomi ve sağlık boyutları birer gösterge ile belirtilirken eğitim boyutu iki alt göstergeden meydana gelmiş ve bu iki alt göstergeden okuryazarlık oranı 2/3, okullaşma oranı ise 1/3 oranında katkı sağlayarak eğitim boyutunu oluşturmuştur (Ivanova, vd., 1999, ss.159-160). 2010 yılından itibaren endeks hesaplamasında bazı büyük değişikliklere gidilmiştir. Bu kapsamda genel olarak endeksin hesabında aritmetik ortalama yerine geometrik ortalama kullanılmıştır. Eğitim boyutunu hesaplarken ise yetişkin okuryazarlık oranı artık hesaplamaya dahil edilmemiş; eğitim boyutu okullaşma oranı ve tahmini okullaşma oranının ortalaması alınarak hesaplanmıştır (Morse 2014, s.249).

İKE, 0 ile 1 arasında değerler alır. Bu aralıkta değerler 1'e doğru yaklaştıkça daha yüksek bir insani kalkınmışlık seviyesini göstermektedir. 2014 yılı insani kalkınma raporuna göre insani kalkınmışlık düzeyi çok yüksek, yüksek, orta ve düşük olarak 4

düzeyde belirlenmiştir. İnsani kalkınma endeks değeri 0.550'den küçük olan ülkeler düşük, 0.550–0.699 arasında olan ülkeler orta, 0.700–0.799 arasında olanlar yüksek ve 0.800'den büyük olan ülkeler ise çok yüksek insani kalkınmışlık sınıfında yer almışlardır (UNDP 2014, s.156).

Bu araştırmanın amacı sıralı lojistik regresyon analizini ve yapay sinir ağları yöntemlerinden Elman YSA, çok katmanlı YSA ve LVQ ağını kullanarak insani kalkınmışlık düzeyinin çoklu sınıflandırma başarılarını karşılaştırmak ve buna etki eden faktörleri tespit etmektir.

Çalışma üç bölümden oluşmaktadır. Birinci bölümde sıralı lojistik regresyon analizi tanıtılmış, ikinci bölümde yapay sinir ağları, Elman YSA, LVQ ağından bahsedilmiştir. Üçüncü bölümde uygulama sonuçları karşılaştırılmıştır.

## 1. Sıralı Lojistik Regresyon Modeli

Lojistik regresyon modelleri, çeşitli kategoriler halinde ölçülmüş bağımlı değişken ile kategorik veya sürekli olarak ölçülmüş bağımsız değişkenler arasındaki ilişkinin modellenmesi amacıyla kullanılan regresyon modelleridir. Sıralı lojistik regresyon (SLOGREG) modeli ise bağımlı değişkenin en az üç kategoriden oluştuğu ve sıralı ölçekle ölçüldüğü durumlarda kullanılan bir yöntemdir (Demirtas v.d., 2009, s.869).

Sıralı lojistik regresyon modelinin temel özellikleri aşağıdaki gibi sayılabilir (Chen and Hughes, 2004, s.4):

- ✓ Kategorik ve sıralı ölçülmüş çıktı değişkeni, gözlenmemiş sürekli gizli (latent) bir değişkenden tekrar düzenlenebilir bir değişkendir, fakat bu sıralı çıktı değişkeninin kategorileri arasındaki mesafenin eşit olup olmadığı belirgin değildir.
- ✓ Sıralı lojistik regresyon analizi, bağımsız değişkenlerin sıralı ve kategorik çıktı değişkeni üzerindeki etkilerini açıklamak için bir bağlantı fonksiyonu kullanılır. Bu modelde normallik ve sabit varyans varsayımının sağlanmasına gerek yoktur.
- ✓ Regresyon katsayısının değeri kategorik çıktı değişkeninin kategorilerine bağlı olmadığından sıralı lojistik regresyon modeli açıklayıcı değişkenler ile sıralı kategorik çıktı değişkeni arasındaki ilişkinin kategorilerden bağımsız olduğunu varsayar.

Sıralı lojistik regresyon modeli, kategorik bağımlı Y değişkeninin altında sürekli ve aslında gözlenemeyen rastgele bir $Y^*$ gizli değişkenin varlığına dayanarak temellendirilir. Bu değişkendeki kategoriler kesme noktası veya eşik değeri olarak adlandırılan sürekli bir düzlemdeki ardışık aralıklar olarak öngörülmektedir (McCullagh, 1980, s.109).

Bu gizli $Y^*$ değişkeni;

$\theta_{s-1} < Y^* < \theta_s$ , s=1,…, j aralığında ve $\theta_0=-\infty$ ve $\theta_j=+\infty$ olmak üzere (Anderson, 1984, s.3) eşitlik (1)'deki şekilde ifade edilir.

$$Y^* = \sum_{k=1}^{K} \beta_k x_k + \varepsilon \qquad (1)$$

Burada θ eşik değerini, $x_k$ bağımsız değişkenler vektörünü, $\beta_k$ parametre vektörünü, $\varepsilon$ ise hata terimini göstermektedir.

Gözlenen Y değişkeni ile gözlenemeyen $Y^*$ arasında aşağıdaki ilişki eşitlik (2)'de gösterilmiştir (Liao, 1994, ss.37-38):

$$y = \begin{cases} 1 & \text{eğer} \quad y_i^* \leq \theta_1 \\ 2 & \text{eğer} \quad \theta_1 < y_i^* \leq \theta_2 \\ 3 & \text{eğer} \quad \theta_2 < y_i^* \leq \theta_3 \\ \vdots & \vdots \\ j & \text{eğer} \quad \theta_{j-1} < y_i^* \end{cases} \tag{2}$$

Burada θ'lar bağımlı değişkenin kategorilerini birbirinden ayıran eşik değerlerini göstermektedir. F lojistik dağıldığı varsayılan hata teriminin dağılım fonksiyonu olmak üzere, verilen bağımsız değişkenler için gözlenen bağımlı değişkenin k. kategoriye düşmesinin genel olasılığı eşitlik (3)'de gösterilmiştir:

$$Prob(y = j\backslash x) = F\left[\theta_j - \sum_{k=1}^{K} \beta_k x_k\right] - F\left[\theta_{j-1} - \sum_{k=1}^{K} \beta_k x_k\right] \tag{3}$$

Sıralı lojistik modelin tahmin edilmesi için kümülatif olasılıkların bir dönüşümü olan birçok bağlantı fonksiyonu vardır. Bu fonksiyonlar Tablo 1'de gösterilmiştir. (Elamir ve Sadeq, 2010, s.652):

**Tablo 1. Bağlantı Fonksiyonları ve tipik uygulamaları**

| Fonksiyon | Biçimi | Uygulama Alanı |
|---|---|---|
| **Logit** | $\log\left(\frac{x}{1-x}\right)$ | Kategoriler eşit olarak dağılır |
| **Tamalayıcı Log-Log** | $\log\left(-\log(1-x)\right)$ | Yüksek kategoriler daha olasıdır |
| **Negatif Log-Log** | $-\log\left(-\log\left(x\right)\right.$ | Düşük kategoriler daha olasıdır |
| **Probit** | $F^{-1}(x)$ | Değişken normal dağılır |
| **Couchit** | $\tan\left(\pi(x-0.5)\right)$ | Değişken aşırı değerlere sahiptir |

Sıralı logit modelde L lojit dağılım fonksiyonunu göstermek üzere, gözlemlerin bağımlı değişkenin kategorilerine düşme olasılıkları eşitlik (4)'de verilmiştir (Akkus v.d., 2010, s.323):

$$Prob(y = 1) = L\left(-\sum_{k=1}^{K} x_k \beta_k\right)$$

$$Prob(y = 2) = L\left(\theta_2 - \sum_{k=1}^{K} x_k\, \beta_k\right) - L\left(-\sum_{k=1}^{K} x_k \beta_k\right)$$

$$Prob(y = 3) = L\left(\theta_3 - \sum_{k=1}^{K} x_k\, \beta_k\right) - L\left(\theta_2 - \sum_{k=1}^{K} x_k \beta_k\right)$$

$$Prob(y = j) = 1 - L\left(\theta_{j-1} - \sum_{k=1}^{K} x_k \beta_k\right) \tag{4}$$

Sıralı lojistik regresyon modelinde en önemli varsayım paralel eğriler varsayımıdır. Bu varsayıma göre modelde elde edilen regresyon parametreleri bağımlı değişkenin tüm kategorilerinde eşittir. Diğer bir ifadeyle, bağımsız değişkenler ile bağımlı değişken arasındaki ilişki bağımlı değişkenin kategorilerine göre değişiklik göstermez ve parametre tahminleri birbirinden farklı eşik değerlerine göre değişmez. Dolayısıyla eğer J kategorili bağımlı değişken bulunuyorsa, "$\beta_k$"parametreleri bir

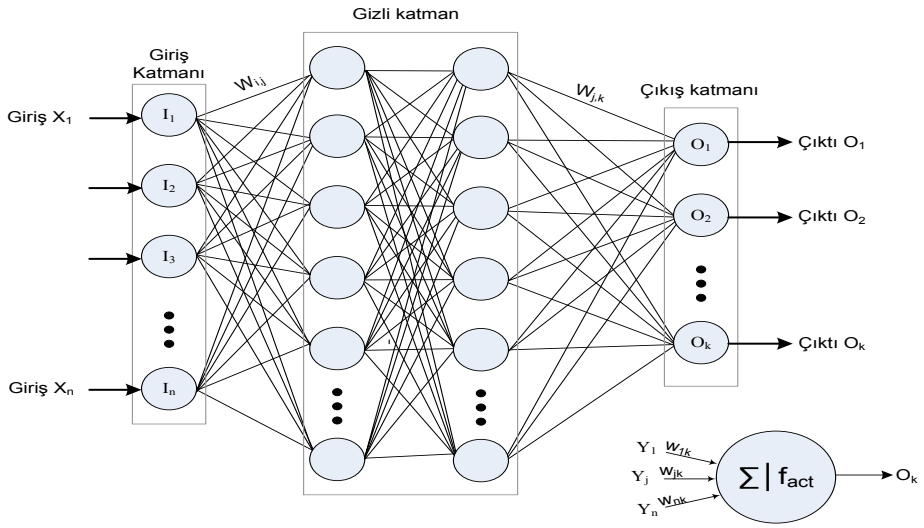
tanedir. Bununla birlikte J-1 lojit karşılaştırması için $\theta_{j-1}$ kesme noktası (eşik değeri) yer almaktadır (Akın ve Şentürk, 2012, s.185).

Sıralı lojistik regresyonda parametrelerin yorumlanması zorlu bir işlemdir. Parametrelerin yorumunda standartlaştırılmış katsayıları hesaplama, tahmin edilen olasılıkları hesaplama, tahmin edilen olasılıklardaki faktör değişmeyi hesaplama ve tahmin edilen olasılıklarda yüzde değişmeyi hesaplama yöntemleri kullanılmaktadır. Parametrelerin yorumunda fark oranı (odds ratio)'da kullanılabilir. Kukla değişkende diğer tüm değişkenler sabit olduğunda exp($\beta_k$) fark oranını vermektedir. Fark oranlarını standardize etmek için $s_k$: standart sapmayı göstermek üzere diğer tüm değişkenler sabit iken $\exp(\beta_k * s_k)$ hesaplanır. Sürekli değişkenlerde ise; $[\exp(\beta - 1) * 100]$ ile yüzde değişim bulunur (Üçdoğruk vd., 2001).

## 2. Yapay Sinir Ağları

Yapay sinir ağları bilgiyi dış dünyadan alan, gerektiğinde depolayan ve daha sonra kullanabilen hücresel sistemli yapılardır. YSA'lar nöronların birbirlerine bağlanmasıyla bir araya gelerek katmaları oluşturan sistemlerdir. YSA'lar içerisinde en çok kullanılan türü çok katmanlı yapay sinir ağlarıdır (Lippman 1987, s.15). Yapay sinir ağları, dış dünyadan gelen bilgileri girdi katmalarına alan belirli bir işlemlerden geçirerek ara katmanlara veya çıktı katmanına aktaran, istenilen çıktı seviyesine erişebilmek amacıyla tekrar girdi katmanına gönderen bilgisayar sistemleridir (Pissarenko, 2002, s.35).YSA'lar basit biyolojik sinir sistemine benzer bir çalışma mantığı içerisinde işlev görmektedir. YSA'lar insanoğlunun düşünebilme ve genelleyebilme yeteneklerine benzer bir şekilde problemleri bilgisayarlar tarafından çözebilme imkanı sağlamaktadır. YSA'lar insan beynin çalışmasına benzer bir şekilde öğrenmeyi gerçekleştirerek mevcut bilgilerden farklı ve yeni bilgiler elde edilebilme yeteneklerine sahip ve bu işlemi farklı öğrenme metotları ile gerçekleştirebilen bilgisayar sistemleridir (Öztemel, 2003, s.29). YSA eldeki verilerden haraketle öğrenme, ilişki kurma, sınıflandırma, genelleme ve optimizasyon amaçlarını gerçekleştirmede kullanılabilemektedir (Şen, 2004, s.13).

Şekil 1'de bir yapay sinir ağının genel yapısı gösterilmektedir.




**Şekil 1. Bir Yapay Sinir Ağının Genel Yapısı**

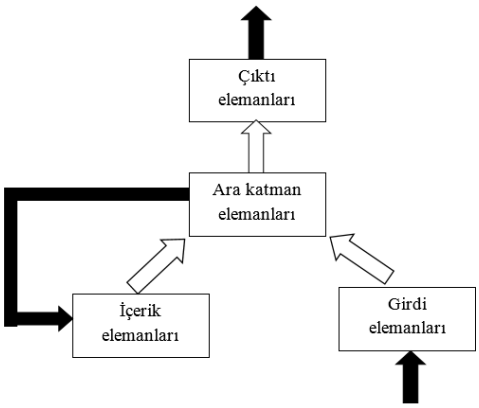
**Kaynak:** M. Bilgehan and P. Turgut, (2010), The Use of Neural Networks In Concrete Compressive Strength Estimation, Computers and Concrete, 7(3), s. 275.

YSA'nın doğrusal ya da doğrusal olmayan bir probleme ait verilerinden yola çıkarak öğrenme teknikleri kullanmak suretiyle daha önce gösterilmemiş problemin verileri verilerek çözümler gerçekleştirmesi kullanıcıların ilgisini çekmiştir. YSA'daki en büyük karmaşıklık uygun yapay sinir ağı mimarisinin belirlenmesidir (Kohonen, 1987, ss. 1-79). YSA, bir takım işlemlerden meydana gelen, verileri kullanarak deneyimsel bilgilerin kullanılmasını imkan sağlayan bilgisayar sistemleridir (Haykin, 1999, s.2). YSA'da nöronlar biraraya katmaları oluşturur (Yıldız, 2001, ss.51-67). Giriş katmanından çıktı katmanına doğru bilgi işleme söz konusudur. Bu bilgi işleme olayı katmanlardaki nöronların birbirlerine bağlanmasıyla gerçekleşir.

### 2.1. Elman Ağı

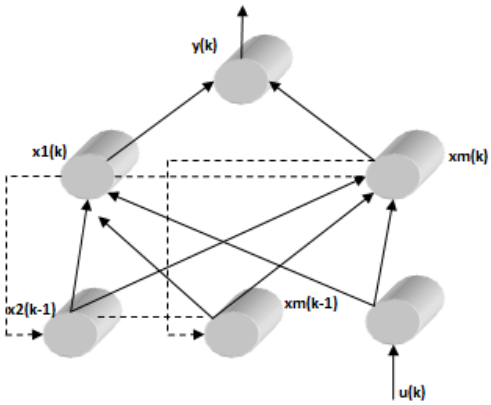
Çok tabakalı YSA yapısının tümüne sahip olan ilave olarak ara tabaka çıktılarını paralel bir girdi tabakası olarak yapısında bulunduran YSA türüne Elman ağı denir (Şen, 2004, s.144). Elman ağında girdi elemanları içerisinde bilgi işleme özellikleri söz konusu değildir (Öztemel, 2003, s.166). Ağın girdilerinin belirlenmesinin ardından ağ artık ileri beslemeli birçok katmanlı algılayıcıya dönüşür. Bu girdilerin kullanılmasıyla ileri doğru ağın çıktıları belirlenir (Elman, 1990, s.182). Elman ağı, Jordan ağına oldukça benzer olmakla beraber aralarında iki önemli farklılık bulunmaktadır. Bu farklılıklardan ilki, geri besleme yaptıkları aktivasyon değerlerini çıktı katmanından almak yerine ara katmandan almaları, ikincisi ise içerik elemanlarının kendilerine bağlantı durumlarının olmamasıdır.

Şekil 2'de Elman Ağ yapısı gösterilmektedir (Küçükönder, 2011, s.78).




**Şekil 2.** Elman Ağ Yapısı

Elman ağında öğrenme, genelleştirilmiş delta öğrenme kuralına göre iki adımda gerçekleşmektedir. İşlem adımlarından ilki, ara katmanda yer alan işlemci elemanlara gelen net girdi değerinin, girdi katmanında yer alan eleman değerleriyle ağırlıklarının çarpılıp toplanmasıdır. İkinci adım ise içerik elemanlarından gelen bu bağlantı değerlerinin ara katmanlarda oluşan bir önceki aktivasyon değerleri ile çarpılıp toplanmasıdır. Elman ağının bu ayrıntılarının gösterildiği yapı Şekil 3'de gösterilmiştir (Öztemel, 2003, s.167).




**Şekil 3. Elman ağının ayrıntılı gösterimi**

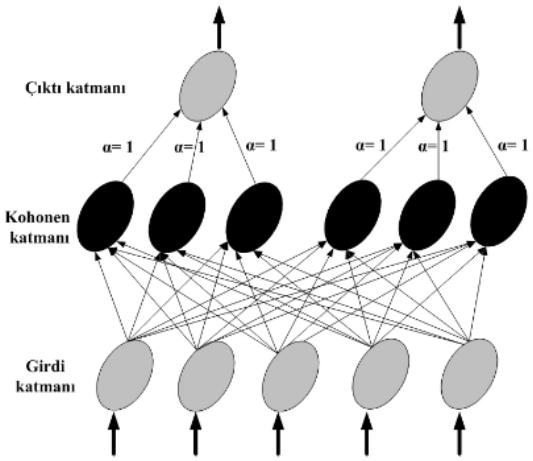
### 2.2. LVQ Modeli (Learning Vector Quantization)

Kohonen (1984) tarafından geliştirilen LVQ modeli destekleyici öğrenme modelini kullanan bir ağ yapısıdır. LVQ ağları daha çok sınıflandırma problemlerini analiz etmek için kullanılır.

Öğrenme, girdi vektörünün hangi vektör seti (referans vektör) tarafından temsil edilmesi gerektiğinin belirlenmesidir. LVQ ağının işlevi öğrenme yolu ile vektör seti ile ifade edilen girdi vektörlerinin üyesine karşılık gelecek vektör sınıfını oluşturmaktır. Kohonen öğrenme kuralına göre öğrenir. Çıktı katmanındanki değerlerden sadece birisi 1, diğerleri ise 0 değerini alarak hangi çıktının hangi sınıfa ait olduğu belirlenir. LVQ ağı istatistiksel bir sınıflandırma ve ayırt etme metodu olduğu için amacı girdi verilerini sınıflandırmada kullanmaktır (Kohonen, 2001, s.245). LVQ ağında, bir giriş, bir çıkış ve kohenon katmanı mevcut olup, giriş katmanındaki tüm nöronlar ara katmanlardaki tüm nöronlarla bağlantılı bir yapı sergilemektedir. Bu ağda asıl amaç, n boyutlu bir vektörün, vektörler seti halinde haritalanmasıdır (Küçükönder, 2011, s.70).

Eğitim sırasında girdi vektörünün sınıflara ayrılması olayı en yakın k komşu algoritmasına göre belirlenmektedir. Girdi vektörü ile referans vektörleri arasındaki en kısa mesafe belirlenmekte, girdi vektörüne en kısa mesafede yer alan vektör katmandaki ağın ağırlıkları değiştirilerek girdileri doğru sınıflara atayacak referans vektörlerinin belirlenmesi amaçlanmaktadır. Kullanılan öğrenme stratejisi destekleyici (reinforcement learning) öğrenme yapısıdır. Çıktı değerinin belirlenmesi aşamasında "kazanan her şey alır" stratejisi ele alınmaktadır. Ağ eğitilirken her bir iterasyonun sonunda ağın meydana getirdiği çıktı değerinin yerine sınıfının doğru olup olmadığı söylenir. Böylece girdi vektörüne en yakın olan vektörün kazanan vektör olduğu, değerlerinin ise ağın bu vektöre ait ağırlıklarının değiştirilmesiyle değişmektedir (Öztemel, 2003: s.115:117).

LVQ ağının bu ayrıntılarının gösterildiği mimari yapı Şekil 4'de gösterilmiştir (Öztemel, 2003, s.116).




**Şekil 4. LVQ Ağ yapısı**

LVQ ağı üç katmandan oluşur (Öztemel, 2003, s.116):

Girdi katmanı: Bu katmanda bilgi işleme olmaz, gelen bilgiler girdi vektörünü oluşturur. Bu katmandaki her proses elemanı kohonen katmanında yer alan her proses elemanıyla ilişkilidir. Girdi katmanı ve kohonen katmanındaki ağırlıklar değiştirilerek öğrenme sağlanır. Kohonen katmanı: Bu katmanda girdi setine en yakın olan ağırlık vektörü belirlenir. Kohonen katmandaki her bir proses elemanı bir referans vektörünü göstermektedir. Girdi değerleri girdi katmanındaki proses elamanlarının Kohonen katmanındaki proses elemanlarına bağlayan bağlantıların ağırlık değerlerinden oluşmaktadır. Referans vektöründeki eleman sayısı ise girdi katmanındaki eleman sayısına eşittir. Çıktı katmanı: Girdinin ait olduğu sınıf bu katmanda belirlenir. Kohonen katmanındaki proses elemanlarının hepsi çıktı katmanındaki bir tek proses elemanına bağlıdır. Kohonen katmanı ile çıktı katmanılarındaki proses elemanlarının ikili binary (değerler) olup sadece bir proses elemanı 1 diğeleri ise 0 değerini alır.

## 3. METODOLOJİ

İktisadi kalkınma ve büyüme her zaman tüm devletlerin en önemli hedefleri arasında yer almıştır. Bu nedenle ülkeler arası karşılaştırma ve sınıflandırmaya olanak sağlaması sebebiyle insani kalkınma endeksi tercih edilen ve başvurulan sayısal gösterge olmuştur. Bu araştırmanın amacı sıralı lojistik regresyon analizini ve yapay sinir ağları yöntemlerinden Elman YSA, çok katmanlı YSA ve LVQ ağını kullanarak insani kalkınmışlık düzeyinin çoklu sınıflandırma başarılarını karşılaştırmak ve etki eden faktörleri tespit etmektir. Bu çalışmada, her yıl yayınlanmakta olan Birleşmiş Milletler Kalkınma Programı (UNDP) İnsani Kalkınma Rapor Ofisinin belirlediği çok gelişmiş, gelişmiş, orta ve düşük düzeyde ülkelere ait veriler ele alınmıştır. Bu veriler Birleşmiş Milletler Kalkınma Programının internet sitesinden Beşeri Kalkınma Endeksinin 2010-2012 yılları arasındaki üç seneye ait verilerdir. Düşük düzeydeki ülkelere ait veri bulma noktasında zorluk yaşandığından analizde düşük insani kalkınmışlık düzeyindeki ülkeler dahil edilmemiştir. Araştırmada 81 ülkenin üç yıla ait verileri ile sınıflandırma işlemi gerçekleştirilmiştir. 2012 yılında bir ülkenin daha ilave edilmesiyle toplamda 244 ülkeye ait veri, çalışmanın ana kütle büyüklüğünü oluşturmuştur. İnsani kalkınma endeksinin hesaplanmasında, gelir, eğitim ve yaşam beklentisi olmazsa olmaz değişkenler olarak kullanılmaktadır. Çalışmada insani kalkınmışlık düzeyi çoklu sınıflandırma işlemine tabi tutulacağı için bu üç değişkenin

yanı sıra, yöntemlerin daha ayrıntılı olarak incelenebilmesine olanak sağlamak amacıyla kalkınma ile ilgili olarak analize 11 bağımsız değişken daha eklenerek toplam 14 bağımsız değişken inceleyeme alınmıştır. Analizde kullanılan değişkenler sırasıyla:

X1: BÖO "bebek ölüm oranı",
X2: GSMH "gayri safi milli hasıla",
X3: LKO "lise kayıt oranı",
X4: BY "büyüme",
X5: DYY "doğrudan yabancı yatırım",
X6: ET "enerji tüketimi",
X7: EÜ "enerji üretimi",
X8: E "enflasyon",
X9: IH "ihracat",
X10: IKS "internet kullanıcı sayısı",
X11:ISZ "işsizlik",
X12: ITH "ithalat",
X13: MTAS "mobil telefon abone sayısı",
X14: SH "sağlık harcamaları"dır.

Sıralı Lojistik Regresyon analizinde kullanılmak üzere cevap değişkenleri sırasıyla;

0: Orta İnsani Kalkınmışlık Düzeyi,
1: Gelişmiş İnsani Kalkınmışlık Düzeyi,
2: Çok Gelişmiş İnsani Kalkınmışlık Düzeyi olarak kodlanmıştır.

### 3.1. Verilerin Analiz Yöntemi

Ülkelere ait İnsani Kalkınmışlık Düzeyinin sınıflandırma işlemini gerçekleştirmek için Stata 11.2 paket programı kullanılarak Sıralı Lojistik Regresyon ve Matlab 2012 yazılımından kullanılarak Elman, Çok Katmanlı Yapay Sinir Ağları ve LVQ Ağı yöntemleri ile ülkelerin insani kalkınmışlık düzeyinin sınıflandırılması gerçekleştirilmiştir. Yapay sinir ağları yöntemini kullanırken bir ülke çok gelişmiş kalkınmışlık düzeyinde yer alıyorsa 1 diğer kategorilerin kalkınmışlık düzeyinde yer alıyor ise 0 olmak üzere çıktı değişkenleri oluşturulmuştur.

### 3.2. Sıralı Lojistik Regresyon Analizi

Ülkelere ait İnsani Kalkınma Endeksi'nin sınıflandırma işlemini analiz etmek için Stata 11.2 istatistiksel analiz programından yararlanılmıştır. Tablo 2'de modelin uygunluk değerleri verilmiştir.

**Tablo 2: Kurulan Modelin Uygunluk Değerleri**

| | | | |
|---|---|---|---|
| Log-Lik Intercept Only: | -256,865 | Log-Lik Full Model: | -98,392 |
| D(228): | 196,784 | LR(12): | 316,947 |
| | | Prob > LR: | 0,000 |
| McFadden's $R^2$: | 0,617 | McFadden's Adj $R^2$: | 0,555 |
| ML (Cox-Snell) $R^2$: | 0,727 | Cragg-Uhler(Nagelkerke) $R^2$: | 0,828 |
| McKelvey & Zavoina's $R^2$: | 0,886 | | |
| Variance of y*: | 28,850 | Variance of error: | 3,290 |
| Count $R^2$: | 0,881 | Adj Count $R^2$: | 0,779 |
| AIC: | 0,938 | AIC*n: | 228,784 |
| BIC: | -1.056.570 | BIC': | -250,981 |
| BIC used by Stata: | 273,745 | AIC used by Stata: | 224,784 |

Tablo 2'de araştırmada kullanılan modelin uyum iyiliği testi incelenmiştir. Akaika bilgi kriteri (AIC) 228,784 ve Bayes bilgi kriteri (BIC) -250,981 olarak hesaplanmıştır. Düşük değerli AIC ve negatif değerli BIC değeri logit modelin uyum iyiliğinin tatmin edici düzeyde olduğunu göstermektedir. Sıralı logit modelin olabilirlik oran değeri (likelihood ratio) 316,947 ve p<0,05 değeri anlamlı olup tüm bağımsız değişkenlerin yer aldığı model istatistiksel açıdan anlamlı olduğu sonucuna ulaşılmıştır. Sıralı logit modelin Pseudo- $R^2$ değeri, 0,617 olarak tespit edilen McFadden's $R^2$ değerine eşittir (Long ve Freese, 2001, s.148). Bu değer bağımsız değişkenlerin bağımlı değişken üzerindeki ilgili değişikliklerin %61,7'sini açıklayabileceği anlamına gelmektedir.

Tablo 3'de sıralı lojistik regresyon modeli paralel eğriler varsayımı için test sonuçları aşağıda verilmiştir.

**Tablo 3: Parelel Eğriler Varsayımı Testi**

| Değişken | Ki-kare | P>$\chi^2$ | sd |
|---|---|---|---|
| Tümü | 14.12 | 0.118 | 14 |

Tablo 3'e göre paralel eğriler varsayımının test edilmesinde sıfır hipotez ($H_0$) ve alternatif hipotez ($H_a$) aşağıdaki gibidir.

"$H_0$: İlişkili regresyon katsayıları, bağımlı değişkenin tüm kategorilerinde aynıdır

$H_a$: İlişkili regresyon katsayıları, bağımlı değişkenin tüm düzeylerinde farklıdır"

Hipotezleri kurulmuştur. Parelel Eğriler varsayımı Ki-kare testi ile test edilmiştir. P=0,118>0,05 olduğu için $H_0$ hipotezi rededilemez. Böylelikle regresyon katsayılarının bağımlı değişkenin her bir kategorisinde aynı olduğu ve ülkelerin kalkınmışlık kategorilerinin birbirine parelel olduğu sonucuna varılmıştır.

Tablo 4'te kurulan modelin uyum iyiliği testine yer verilmiştir.

**Tablo 4: Kurulan Modelin Uyum İyiliği Testi**

| Sıralı Logit Modelin Uyum İyiliği Testi | |
|---|---|
| Pearson $\chi^2$ test istatistiği | 10.24 |
| sd | 14 |
| P>$\chi^2$ | 0.332 |

Tablo 4'e göre kurulan modelin uyum iyiliğinin test edilmesinde sıfır hipotez ($H_0$) ve alternatif hipotez ($H_a$) aşağıdaki gibidir.

"$H_0$: Parametreler belirleyicilik açısından model-veri uyumu yeterli düzeydedir

$H_a$: Parametreler belirleyicilik açısından model-veri uyumu yeterli düzeyde değildir"

şeklindedir. Olabilirlik oran testi (Likelihood ratio) ki-kare uyum iyiliği testi olarak anılan test, sıralı lojistik regresyon modelini bir bütün olarak değerlendirmektedir. Modelde P=0,332>0.05 olduğu için $H_0$ hipotezi kabul edilmiştir. Sıralı lojistik regresyon modelinin veri uyumunun yeterli düzeyde olduğunu göstermiştir.

Bağımsız değişken ve bağımlı değişkenler için belirlenen sıralı lojistik regresyon analizine ait sonuçlar Tablo 5'te gösterilmiştir.

**Tablo 5. İnsani Kalkınışlık Düzeyini Etkileyen Değişkenlerin Sıralı Lojistik Regresyon Analizi Sonuçları**

| | | | | | | Gözlem Sayısı | | 244 |
|---|---|---|---|---|---|---|---|---|
| | | | | | | **LR Ki-kare (12)** | | **316,95** |
| | | | | | | **P > $\chi^2$** | | **0.0000** |
| Log likelihood | | -110,706 | | | | **Pseudo $R^2$** | | **0,617** |
| HDI | **Katsayı** | **Std. Hata** | **Wald** | **Z** | **P>Z** | **Odds Oranı** | **[95% Güven Aralığı]** | |
| BÖO | -0,132 | 0,035 | 13,985 | -3,740 | 0,000 | 0,876 | 0,817 | 0,939 |
| GSMH | 0,000 | 0,000 | 1,840 | 1,360 | 0,174 | 1,000 | 1,000 | 1,000 |
| LKO | 0,017 | 0,013 | 1,749 | 1,320 | 0,186 | 1,017 | 0,992 | 1,043 |
| BY | -0,076 | 0,063 | 1,482 | -1,220 | 0,223 | 0,926 | 0,819 | 1,048 |
| DYY | 0,000 | 0,000 | 0,175 | -0,420 | 0,674 | 1,000 | 1,000 | 1,000 |
| ET | 0,000 | 0,000 | 0,603 | -0,780 | 0,438 | 1,000 | 1,000 | 1,000 |
| EÜ | 0,000 | 0,000 | 0,011 | -0,110 | 0,916 | 1,000 | 1,000 | 1,000 |
| E | 0,051 | 0,037 | 1,837 | 1,360 | 0,175 | 1,052 | 0,978 | 1,131 |
| IH | 0,058 | 0,023 | 6,475 | 2,540 | 0,011 | 1,059 | 1,013 | 1,108 |
| IKS | 0,104 | 0,020 | 26,893 | 5,190 | 0,000 | 1,109 | 1,067 | 1,154 |
| ISZ | -0,005 | 0,044 | 0,012 | -0,110 | 0,914 | 0,995 | 0,914 | 1,084 |
| ITH | -0,062 | 0,024 | 6,788 | -2,610 | 0,009 | 0,940 | 0,897 | 0,985 |
| MTAS | -0,006 | 0,009 | 0,506 | -0,710 | 0,477 | 0,994 | 0,977 | 1,011 |
| SH | 0,237 | 0,110 | 4,616 | 2,150 | 0,032 | 1,267 | 1,021 | 1,573 |
| /cut1 | 1,563 | 1,539 | | | | | -1,453 | 4,580 |
| /cut2 | 6,487 | 1,614 | | | | | 3,323 | 9,651 |

Tablo 5'te görüldüğü gibi modelde ele alınan gözlem sayısı 244 olup modele ait $\chi^2$ değeri istatistiksel olarak anlamlı olup (p<0,01), modele ait log olabilirlik değeri -110,71 olarak hesaplanmıştır. Tablo 5'in ilk sütunu sıralı lojistik regresyon analizinin $\beta$ katsayılarını ifade etmektedir. Ülkelerin İnsani kalkınmışlık düzeylerinin belirlenmesinde yer alan bağımsız değişkenlerin katsayılarının ($\beta$) anlamlılığının test edilmesinde Wald testi kullanılmıştır. Wald testine göre BÖO "bebek ölüm oranı", İKS "internet kullanıcı sayısı" ve ITH "ithalat" 0,01 önem seviyesinde; IH "ihracat" ve SH "sağlık harcamaları" 0,05 önem seviyesinde istatistiksel olarak anlamlı olduğu gözlenmiştir. İstatistiksel olarak bağımlı değişken üzerindeki BÖO ve ITH değişkenleri negatif, İKS, İH ve SH değişkenlerin tahmin edilen değeri pozitif işaretlidir.

Odds oranı bağısız değişkenlerin logit üzerindeki etkisini göstermek için doğrudan yorumlanabilir. Odds'daki yüzdelik değişim (100 × [OR-1]) formülü kullanılarak hesaplanabilir. Odds oranı (OR) 1'den büyük ise bağımsız değişkenin lojit'te artırıcı bir etkisi olduğunu, OR değerinin 1'den küçük olması bağımsız değişkenin lojit'te azaltıcı bir etkisi olduğunu ve OR değerinin 1'e eşit olması bağımsız değişkenin lojit'te değişime neden olmadığı şeklinde değerlendirilebilir (O'Connell, 2006, s.16; Alpar, 2013, s.643). Odds oranları için, BÖO değişkenindeki bir birimlik artışın Exp($\beta$)diğer tüm bağımsız değişkenler sabit kalmak şartıyla orta ve gelişmiş insani kalkınmışlık düzeyine karşın, çok gelişmiş insani kalkınmışlık düzeyinin odds'unda negatif etkiye sahip olduğunu göstermiştir. BÖO değişkeni değerinin üstel lojistik değeri $e^{-0,132} = 0,876$ olarak bulunmuştur. Bu değerin çok gelişmiş insani kalkınmışlık düzeyinin odds'unda orta ve gelişmiş insani kalkınmışlık düzeyine karşın

%12,4'lük (1-0,876=0,124*100) bir oranda azalışa neden olmuştur. Benzer şekilde sırasıyla ITH, IH, IKS ve SH değişkenlerinin üstel lojistik değerleri $e^{-0,062} = 0,94$; $e^{0,058} = 1,059$; $e^{0,104} = 1,109$ ve $e^{0,237} = 1,267$ olarak gerçekleşmiştir. Böylelikle değişkenlerin bir birimlik artışı çok gelişmiş insani kalkınmışlık düzeyinin odds'unda ITH değişkeni %6'lık (1-0,94=0,06*100) azalışa; IH değişkeni %5,9'luk (1,059-1=0,059*100), IKS değişkeni %10,9'luk (1,109-1=0,109*100) ve SH değişkeni %26,7'lik artışa neden olmuştur. Buradan hareketle insani kalkınmışlık düzeyine pozitif yönde etki eden en önemli değişkenin SH "sağlık harcamaları" değişkeni ve ikinci sıradaki değişkenin ise IKS "internet kullanıcı sayısı" değişkeni, üçüncü sıradaki değişkenin ise IH "ihracat" değişkeni; negatif yönde etki eden en önemli değişkenin ise BÖO "bebek ölüm oranı" değişkeni olduğu gözlenmiştir.

Tablo 5'te olasılık hesaplarında kullanılmak üzere 3 kategorili bağımlı değişken için oluşturulan sıralı lojistik regresyon modeli aşağıda verilmiştir. Kategorik değişken sayısının M-1=2 kadar cut off (kesme) değeri elde edilir.

$Z=\sum_{k=1}^{K} \beta_k X_k = -0,132BÖO+1,37*10^{-12}GSMH+0,017LKO- 0,076BY-1*10^{-11}DYY - 5,38*10^{-6}ET-3,42*10^{-7}EÜ+0,051E+0,058IH+0,104IKS-0,005ISZ-0,062ITH-0,006MTAS+0,237SH$

Yukarıda verilen sıralı lojistik regresyon modeline herhangi bir ülkenin istenen yılına ait değişken değerleri yazılarak Z değeri bulunur. Herhangi bir ülke için hesaplanan Z değeri aşağıda verilen formüllere yazılarak o ülkenin orta, gelişmiş ve çok gelişmiş insani kalkınmışlık düzeyinde yer alma olasılıkları hesaplanır. Hesaplanan en yüksek olasılık değeri o ülkenin hangi insani kalkınmışlık düzeyinde yer aldığını gösterir.

Orta düzeyde kalkınmışlık düzeyine ait olasılık değeri:

$$P(Y = 0) = 1 - \frac{\exp(Z_i - cut_1)}{1+\exp(Z_i - cut_1)}$$

Gelişmiş düzeyde kalkınmışlık düzeyine ait olasılık değeri:

$$P(Y = 1) = \frac{\exp(Z_i - cut_2)}{1+\exp(Z_i - cut_2)} - \frac{exp(Z_i - cut_1)}{1+\exp(Z_i - cut_1)}$$

Çok gelişmiş kalkınmışlık düzeyine ait olasılık değeri:

$$P(Y = 2) = \frac{\exp(Z_i - cut_2)}{1+\exp(Z_i - cut_2)}$$

Buna göre 2010 yılı Türkiye'nin İnsani Kalkınmışlık Düzeyine ait olasılık değerleri aşağıdaki gibidir:

$Z_1(x_i) = -0,132BÖO(16)+1,37*10^{-12}GSMH(7,31E+11)+0,017LKO(56)-0,076BY(9,29)-1*10^{11}DYY(9,04E+09)-5,38*10^{-6}ET(105133,1-3,42*10^{-7}EÜ(32225)+0,051E(5,7)+0,058IH(21)+0,104IKS(39,8)-0,005ISZ(11,9)-0,062ITH(27)-0,006MTAS(85)+0,237SH(6,7)=3,441$

$$Z_i - cut_1 = 3,441 - 1,563 = 1,878$$
$$Z_i - cut_2 = 3,441 - 6,487 = -3,046$$

Bu hesaplara göre,

Orta düzeyde gelişmiş ülke olma olasılığı: P(Y=0)= 0,133,

Gelişmiş düzeyde ülke olma olasılığı: P(Y=1)=0,822,

Çok gelişmiş düzeyde ülke olma olasılığı: P(Y=2)=0,045,

2010 yılında gelişmiş insani kalkınmışlık düzeyinde yer alan Türkiye'nin, sıralı lojistik regresyon analizi sonucu 0,822 ile en yüksek olasılığa sahip olan gelişmiş insani kalkınmışlık düzeyinde yer almasıyla doğru bir şekilde sınıflandırıldığı tespit edilmiştir.

2011 yılı Türkiye'nin İnsani Kalkınmışlık Düzeyine ait olasılık değerleri aşağıdaki gibidir:

$Z_1(x_i) = -0{,}132$BÖO(15)+1,37*$10^{-12}$GSMH(7,75E+11)+0,017LKO(61)–0,076BY(8,8)-1*$10^{11}$DYY(1,6E+10)–5,38*$10^{-6}$ET(112458,7)–3,42*$10^{-7}$EÜ(32064)+0,051E(8,6)+0,058IH(24)+0,104IKS(43,1)-0,005ISZ(9,8)-0,062ITH(33)-0,006MTAS(89)+0,237SH(6,7)=3,916

$Z_i - cut_1 = 3{,}916 - 1{,}563 = 2{,}353$
$Z_i - cut_2 = 3{,}916 - 6{,}487 = -2{,}571$

Bu hesaplara göre,

Orta düzeyde gelişmiş ülke olma olasılığı: P(Y=0)= 0,087,

Gelişmiş düzeyde ülke olma olasılığı: P(Y=1)=0,842,

Çok gelişmiş düzeyde ülke olma olasılığı: P(Y=2)=0,071,

2011 yılında da gelişmiş insani kalkınmışlık düzeyinde yer alan Türkiye'nin, sıralı lojistik regresyon analizi sonucu 0,842 ile en yüksek olasılığa sahip olan gelişmiş insani kalkınmışlık düzeyinde yer almasıyla doğru bir şekilde sınıflandırıldığı gözlenmiştir.

2012 yılı Türkiye'nin İnsani Kalkınmışlık Düzeyine ait olasılık değerleri aşağıdaki gibidir:

$Z_1(x_i) = -0{,}132$BÖO(14)+1,37*$10^{-12}$GSMH(7,89E+11)+0,017LKO(62)–0,076BY(2,2)-1*$10^{11}$DYY(1,26E+10)–5,38*$10^{-6}$ET(115701,2)–3,42*$10^{-7}$EÜ(31117)+0,051E(6,8)+0,058IH(26)+0,104IKS(45,1)
-0,005ISZ(9,3)-0,062ITH(32)-0,006MTAS(91)+0,237SH(6,7)=4,891

$Z_i - cut_1 = 4{,}891 - 1{,}563 = 3{,}328$
$Z_i - cut_2 = 4{,}891 - 6{,}487 = -1{,}596$

Bu hesaplara göre,

Orta düzeyde gelişmiş ülke olma olasılığı: P(Y=0)= 0,035,

Gelişmiş düzeyde ülke olma olasılığı: P(Y=1)=0,797,

Çok gelişmiş düzeyde ülke olma olasılığı: P(Y=2)=0,168,

2012 yılında da gelişmiş insani kalkınmışlık düzeyinde yer alan Türkiye'nin, sıralı lojistik regresyon analizi sonucu 0,797 ile en yüksek olasılığa sahip olan gelişmiş insani kalkınmışlık düzeyinde yer almasıyla doğru bir şekilde sınıflandırıldığı

gözlenmiştir. Böylelikle bu üç yıl içinde Türkiye'nin insani kalkınmışlık düzeyi doğru olarak sınıflandırılmıştır.

2012 yılı baz alınarak Türkiye için önerilen SH, IKS ve IH değişkenlerinde %20'lik bir artış ve BÖO değişkeninde %10'luk bir azalışla beraber Türkiye'nin çok gelişmiş insani kalkınmışlık düzeyinde yer alabilmesi muhtemeldir.

2012 yılı için önerilen değişkenlerdeki değişikliğe bağlı olarak Türkiye'nin İnsani Kalkınmışlık Düzeyine ait olasılık değerleri aşağıdaki gibidir:

$Z_1(x_i) = -0{,}132$BÖO(12,6)+1,37*$10^{-12}$GSMH(7,89E+11)+0,017LKO(62)–0,076BY(2,2)-1*$10^{11}$DYY(1,26E+10)–5,38*$10^{-6}$ET(115701,2)–3,42*$10^{-7}$EÜ(31117)+0,051E(6,8)+0,058IH(31,2)+0,104IKS(54,12) -0,005ISZ(9,3)-0,062ITH(32)-0,006MTAS(91)+0,237SH(8,04)=6,629

$Z_i - cut_1 = 6{,}629 - 1{,}563 = 5{,}066$
$Z_i - cut_2 = 6{,}629 - 6{,}487 = 0{,}142$

Bu hesaplara göre,

Orta düzeyde gelişmiş ülke olma olasılığı: P(Y=0)=0,006,

Gelişmiş düzeyde ülke olma olasılığı: P(Y=1)=0,458,

Çok gelişmiş düzeyde ülke olma olasılığı: P(Y=2)=0,536,

2012 yılı için önerilen değişkenlerdeki değişme ile beraber Türkiye'nin sıra lojistik regresyon analizi sonucu 0,536 ile en yüksek olasılığa sahip çok gelişmiş insani kalkınmışlık düzeyinde yer alması olasıdır.

Tablo 7'de sıralı lojistik regresyon analizi kullanılmasıyla 2010-2012 yılları arasındaki 81 ülkenin üçer yıllık toplamda 244 ülkenin insani kalkınmışlık düzeyi sınıflandırma sonuçları gösterilmiştir.

**Tablo 7: Sıralı Lojistik Regresyon Analizi Sınıflandırma Başarısı Sonucu**

| Sıralı Lojistik Regresyon Analizi | | Tahmin Edilen Grup | | | | Doğruluk Yüzdesi |
|---|---|---|---|---|---|---|
| | | **Orta** | **Gelişmiş** | **Çok Gelişmiş** | **Toplam** | |
| Gözlenen Grup | **Orta** | **44** | 9 | 0 | 53 | 83 |
| | **Gelişmiş** | 7 | **66** | 5 | 78 | 84,6 |
| | **Çok Gelişmiş** | 2 | 6 | **105** | 113 | 92,9 |
| | **Toplam** | 53 | 81 | 110 | **244** | 88,1 |

Tablo 7'ye göre kalkınmışlık sınıflandırma değişkenlerine ait veriler sıralı lojistik regresyon analizinde kullanılarak 53 orta düzey gelişmiş ülkeden 44'ünü; 78 gelişmiş ülkeden 66'sını ve 113 çok gelişmiş ülkeden 105'ini doğru tahmin ederek, orta düzey ülkeler için %83'lük, gelişmiş ülkeler için %84,6'lık ve çok gelişmiş ülkeler için %92,9'luk doğru sınıflandırma başarısı gerçekleştirdiği gözlenmiştir. Tüm ülkeler için toplam sınıflandırma başarısı %88,1 olarak tespit edilmiştir.

### 3.2. Yapay Sinir Ağları Analizi

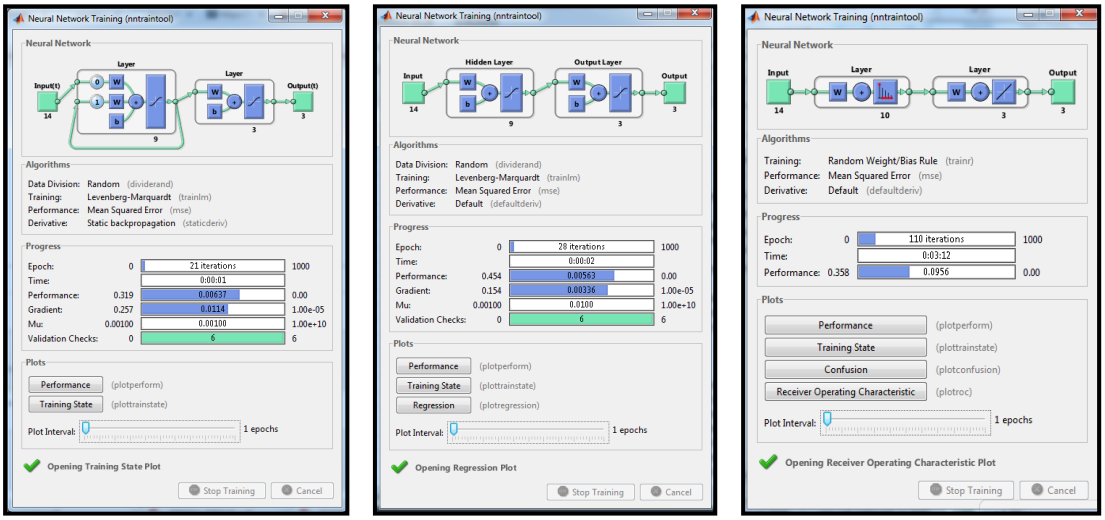
YSA modellerin oluşturulmasında, Matlab R2012a bilgisayar yazılımından faydalanılmıştır. YSA yönteminin tercih edilmesinde çok defa deneme yanılma yönteminden yararlanılmış ve çok sayıda test işlemi gerçleştirilmiştir. Bu doğrultuda gizli katman sayısı, gizli katmanlardaki düğüm sayısı, momentum terimi, aktivasyon fonksiyonu, devir sayısı gibi parametrelerin çeşitli kombinasyonları deneme yanılma yöntemi ile gerek eğitim seti gerekse test seti üzerinde en iyi performansı gösteren ağ belirlenerek sınıflnadırma işlemleri gerçekleştirilmiştir. Ele alınan YSA türleri ise geri dönüşümlü elman yapay sinir ağı, çok katmanlı yapay sinir ağı ve LVQ ağlarıdır.

Tablo 8'de sınıflandırma tahmini için ele alınan YSA'ya ait belirlenen en uygun katman ve düğüm sayılarının yer aldığı modellerin mimari yapısı verilmiştir.

**Tablo 8: Elman YSA, Çok Katmanlı ve LVQ Ağ Parametreleri**

| Ağın türü | Elman | Çok Katmanlı | LVQ |
|---|---|---|---|
| | **YSA** | **YSA** | **YSA** |
| Öğrenme Algoritması | Levenberg-Marquardt Optimization (Denetimli Öğrenme) | Levenberg-Marquardt Optimization (Denetimli Öğrenme) | Learnk (Destekleyici Öğrenme) |
| Öğrenme Kuralı | Gradient descent rule | Gradient descent rule | Kohonen rule |
| Giriş Katmanındaki Düğüm Sayısı | 14 | 14 | 14 |
| Gizli Katman Sayısı | 1 | 1 | 1 |
| Gizli Katmandaki Düğüm Sayısı | 9 | 9 | 10 |
| Çıkış katmanındaki Düğüm Sayısı | 3 | 3 | 3 |
| Öğrenme Oranı | 0,01 | 0,01 | 0,01 |
| Devir Sayısı | 21 | 28 | 110 |
| Öğrenme zamanı (sn) | 1 | 2 | 192 |
| Gizli Katmanlar için Transfer Fonksiyonu | Tansig | Tansig | Learnlv1 (LVQ1 weight learning function) |
| Çıktı Katman İçin Transfer Fonksiyonu | Purelin | Purelin | Purelin |
| Geri Yayılım Ağının Eğitim Fonksiyonu | Trainlm | Trainlm | Geri yayılım yok |

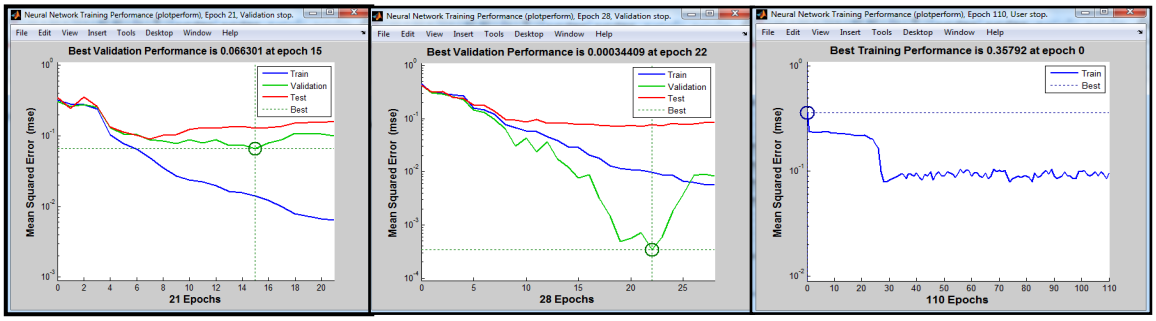
Tablo 8'e göre ülkelere ait insani kalkınmışlık düzeyi sınıflandırma işlemini analiz etmek için Elman YSA, çok katmanlı YSA ve LVQ modelleri kullanılmıştır. Modellerin her üçünde de giriş katmanında 14 düğüm ele alınmınmıştır ve bu durum insani kalkınmışlık düzeyini sınıflandırmak için kullanılan 14 bağımsız değişkenin normalleştirme sonucundaki değerleridir. Çok katmanlı ve Elman YSA türlerinde gizli katman sayısı 1 belirlenmiş her katmanda 9 düğüm oluşturulmuştur, LVQ türünde ise gizli katmanında 10 düğüm belirlenmiştir. Elman YSA, çok katmanlı YSA ve LVQ ağı çeşitlerinde 3 kategorili sınıflandırma işlemini analiz etmek için çıktı katmanında 3'er düğüm oluşturulmuştur. Matlab 2012 yazılımıyla oluşturulan YSA Modellerine yönelik mimari gösterim Şekil 5'de yer verilmiştir.




**Şekil 5. Ülkelerin İnsani Kalkınmışlık Düzeyini Sınıflandırma İşlemi için Matlab Yazılımıyla Oluşturulan YSA Modelleri**

Oluşturulan YSA modellerine ait eğitim işlemleri gerçekleştirildikten sonra, çok sayıda test işlemi yapılmıştır. Şekil 5'de görüldüğü üzere Elman YSA ve çok katmanlı YSA modellerinde ülkelere ait insani kalkınma düzeyinin en iyi sınıflandırma işlemini veren öğrenme algoritması Levenberg-Marquardt Optimizationu'dur. Bu kapsamda gizli katmanlardaki ele alınan fonksiyon çeşidi sigmoid transfer fonskiyonu "tansig"'dur. Çıktı katmanında ele alınan transfer fonksiyonu "purelin" olup geri yayılım ağının eğitimi için ise "trainlm" fonksiyonları denenmiştir. LVQ ağları için kullanılan öğrenme algoritması destekleyici öğrenme algoritması, gizli katmanlarında LVQ1 weight learning fonksiyon türü, çıktı katman için "purelin" fonksiyonları kullanılmıştır. Elman YSA, Çok katmanlı YSA ve LVQ ağı için 14-1-3 network mimarileri en uygun yapıdaki YSA modeli olarak belirlenmiştir.

Şekil 6'da YSA modelleri için devir performansları gösterilmektedir.




**Şekil 6. YSA Modelleri Devir Performansları**

Şekil 6'ya göre oluşturulan yapay sinir ağlarının eğitimi için başlangıçta 1000 iterasyon atanmış , Elman YSA için 21 iterasyonu 1 saniyede, çok katmanlı YSA için 28 iterasyonu 2 saniyede ve LVQ ağı için 110 iterasyonu 192 saniyede tamamlayarak öğrenme işlemlerini gerçekleştirmişlerdir.

Tablo 9'da Elman YSA analizi kullanılmasıyla 2010-2012 yılları arasındaki 81 ülkenin üçer yıllık toplamda 244 ülkenin insani kalkınmışlık düzeyi sınıflandırma sonuçları gösterilmiştir.

**Tablo 9: Elman YSA Analizi Sınıflandırma Başarısı Sonucu**

| Elman YSA | | Tahmin Edilen Grup | | | | Doğruluk Yüzdesi |
|---|---|---|---|---|---|---|
| | | **Orta** | **Gelişmiş** | **Çok Gelişmiş** | **Toplam** | |
| Gözlenen Grup | **Orta** | **48** | 5 | 0 | 53 | 90,6 |
| | **Gelişmiş** | 0 | **74** | 4 | 78 | 94,9 |
| | **Çok Gelişmiş** | 2 | 3 | **108** | 113 | 95,6 |
| | **Toplam** | 50 | 82 | 112 | **244** | 94,3 |

Tablo 9'a göre kalkınmışlık sınıflandırma değişkenlerine ait veriler Elman YSA analizinde kullanılarak 53 orta düzey gelişmiş ülkeden 48'ini; 78 gelişmiş ülkeden 74'ünü ve 113 çok gelişmiş ülkeden 108'ini doğru tahmin ederek, orta düzey ülkeler için %90,6'lık, gelişmiş ülkeler için %94,9'luk ve çok gelişmiş ülkeler için %95,6'lık doğru sınıflandırma başarısı gözlenmiştir. Tüm ülkeler için toplam sınıflandırma başarı düzeyi %94,3'tür.

Tablo 10'da Çok katmanlı YSA analizi kullanılmasıyla 2010-2012 yılları arasındaki 81 ülkenin üçer yıllık toplamda 244 ülkenin insani kalkınmışlık düzeyi sınıflandırma sonuçları gösterilmiştir.

**Tablo 10: Çok Katmanlı YSA Analizi Sınıflandırma Başarısı Sonucu**

| Çok Katmanlı YSA | | Tahmin Edilen Grup | | | | Doğruluk Yüzdesi |
|---|---|---|---|---|---|---|
| | | **Orta** | **Gelişmiş** | **Çok Gelişmiş** | **Toplam** | |
| Gözlenen Grup | **Orta** | **53** | 0 | 0 | 53 | 100 |
| | **Gelişmiş** | 2 | **75** | 1 | 78 | 96,2 |
| | **Çok Gelişmiş** | 2 | 2 | **109** | 113 | 96,5 |
| | **Toplam** | 57 | 77 | 110 | **244** | 97,1 |

Tablo 10'a göre çok katmanlı YSA analizinde kullanılarak 53 orta düzey gelişmiş ülkeden 53'ünü (tamamını); 78 gelişmiş ülkeden 75'ini ve 113 çok gelişmiş ülkeden 109'unu doğru tahmin ederek, orta düzey ülkeler için %100'lük, gelişmiş ülkeler için %96,2'lik ve çok gelişmiş ülkeler için %96,5'lik doğru sınıflandırma başarısı gözlenmiştir. Tüm ülkeler için sınıflandırma başarı düzeyi %97,1'dir.

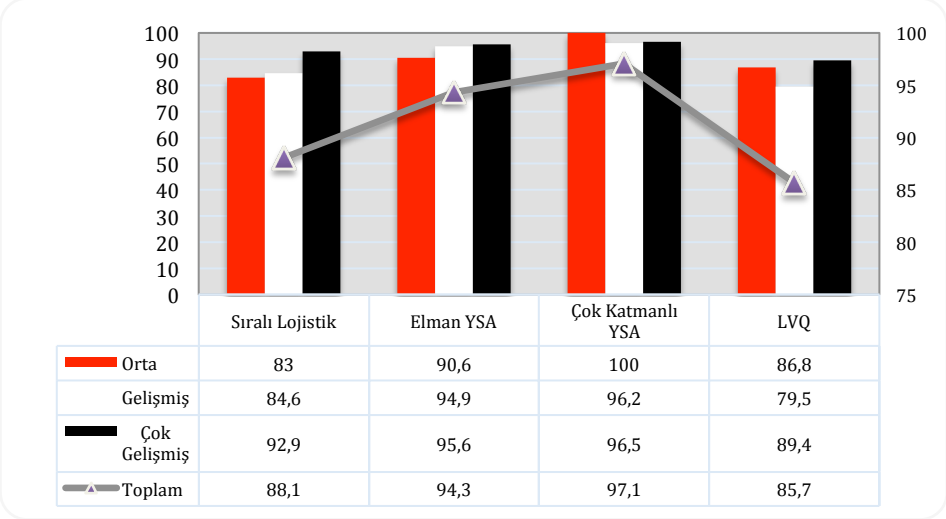
Tablo 11'de LVQ ağı analizi kullanılmasıyla 2010-2012 yılları arasındaki 81 ülkenin üçer yıllık toplamda 244 ülkenin insani kalkınmışlık düzeyi sınıflandırma sonuçları gösterilmiştir.

**Tablo 11: LVQ Ağı Analizi Sınıflandırma Başarısı Sonucu**

| LVQ Ağı | | Tahmin Edilen Grup | | | | Doğruluk Yüzdesi |
|---|---|---|---|---|---|---|
| | | **Orta** | **Gelişmiş** | **Çok Gelişmiş** | **Toplam** | |
| Gözlenen Grup | **Orta** | **46** | 7 | 0 | 53 | 86,8 |
| | **Gelişmiş** | 13 | **62** | 3 | 78 | 79,5 |
| | **Çok Gelişmiş** | 4 | 8 | **101** | 113 | 89,4 |
| | **Toplam** | 63 | 77 | 104 | **244** | 85,7 |

Tablo 11'e göre LVQ ağı analizinde kullanılarak 53 orta düzey gelişmiş ülkeden 46'sını; 78 gelişmiş ülkeden 62'sini ve 113 çok gelişmiş ülkeden 101'ini doğru tahmin ederek, orta düzey ülkeler için %86,8'lik, gelişmiş ülkeler için %79,5'lik ve çok gelişmiş ülkeler için %89,4'lük doğru sınıflandırma başarısı gözlenmiştir. Tüm ülkeler için toplam sınıflandırma başarı düzeyi %85,7'dir.

Grafik 1.'de 2010-2012 yılları arasındaki 81 ülkenin üçer yıllık toplamda 244 ülkenin dört analiz yöntemine göre orta, gelişmiş, çok gelişmiş insani kalkınmışlık düzeyi ve toplam sınıflandırma doğruluğuna sahip % değerler gösterilmiştir.




**Grafik 1. Dört Analiz Yöntemine İlişkin Performans Değerleri**

Grafik 1'e göre en iyi sınıflanırma sonucu çok katmanlı YSA olarak tespit edilmiştir. Orta insani kalkınmışlık düzeyini %100, gelişmiş insani kalkınmışlık düzeyini %96,2, çok gelişmiş insani kalkınmışlık düzeyini %96,5 ve toplam sınıflandırma başarısını %97,1 olarak gerçekleştirmiştir. Analizler arası karşılaştırma gerçekleştirildiğinde çok katmanlı YSA, Elman YSA'dan; sıralı lojistik regresyon analizi ise LVQ ağına göre daha yüksek sınıflandırma yüzdesiyle sonuçlar verdiği tespit edilmiştir. Dört analiz yönteminde de çok katmanlı YSA yönteminin orta, gelişmiş, çok gelişmiş ve toplam sınıflandırma sonuçlarıyla diğer üç yönteme göre performansının daha iyi olduğu gözlenmiştir.

## SONUÇ

Bu çalışmada ülkelerin insani kalkınmışlık düzeyini sıralı lojistik regresyon, Elman YSA, çok katmanlı YSA ve LVQ ağı kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Bağımlı değişken olarak orta düzeyde gelişmiş, gelişmiş ve çok gelişmiş ülkeler olmak üzere 3 kategorili sıralı değişken ile 14 bağımsız değişken ile arasındaki ilişkiyi belirlemek için çoklu sınıflandırma yöntemleri ele alınmıştır. Uygulama verisi için Birleşmiş Milletler Kalkınma Programı İnsani Kalkınma Endeksi kapsamında 81 ülkenin 2010-2012 yılları arasındaki verileri ele alınmıştır.

Sıralı lojistik regresyon analizinde 14 bağımsız değişkenden BÖO "bebek ölüm oranı", ITH "ithalat" negatif ve İKS "internet kullanıcı sayısı", IH "ihracat" ve SH "sağlık harcamaları" pozitif işaretli olarak istatistiksel açıdan anlamlı olduğu gözlenmiştir. İnsani kalkınma düzeyine pozitif yönde etki eden en önemli değişkenin SH "sağlık harcamaları" değişkeni ve ikinci sıradaki değişkenin ise IKS "internet kullanıcı sayısı" değişkeni, negatif yönde etki eden en önemli değişkenin ise BÖO "bebek ölüm oranı" değişkeni olduğu tespit edilmiştir. 2010, 2011 ve 2012 yıllarında gelişmiş insani kalkınmışlık düzeyinde yer alan Türkiye'nin, sıralı lojistik regresyon analizi ile sınıflandırma işlemi doğru şekilde tahmin edilmiştir. Analizler arası karşılaştırmada çok katmanlı YSA, Elman YSA'dan; sıralı lojistik regresyon analizi ise LVQ ağına göre daha yüksek sınıflandırma yüzdesiyle sonuçlar gerçekleştirmiştir.

2012 yılı Türkiye için önerilen SH, IKS ve IH değişkenlerinde %20'lik bir artış ve BÖO değişkeninde %10'luk bir azalışla beraber Türkiye'nin sıra lojistik regresyon analizi sonucu 0,536 ile en yüksek olasılığa sahip çok gelişmiş insani kalkınmışlık düzeyinde yer alabilmesi muhtemeldir.

Analiz sonuçları literatürde yer alan çalışılmalarla karşılaştırıldığında, Burmaoğlu, vd.(2009), binary lojistik regresyon analizinden yararlanarak birleşmiş milletler kalkınma programı beşeri kalkınma endeksinin sınıflandırılmasında en fazla etkiye sahip olan değişkenlerin sağlık harcamaları ve doğumda yaşam beklentisi olduğunu tespit etmişlerdir. Ayrıca yapılan başka bir çalışmada Eren, vd., (2014), doğumda yaşam beklentisi (bebek ölüm oranı) ve okullaşma oranı değişkenlerinin insani kalkınmışlık düzeyine etki eden en önemli değişkenler olduklarını belirtmişlerdir.

Araştırma sonucunda elde edilen sonuçlar dikkate alındığında, bundan sonraki çalışmalarda bağımsız değişkenlerin sayısı artırılarak ve geniş bir dönemi kapsayacak şekilde değerlendirilmelidir. Gelecekte yapılacak çalışmalarda veri madenciliği teknikleri kullanılarak insani kalkınmışlık endeksine etki eden değişkenler yardımıyla karar ağaçları ile analiz gerçekleştirilebilir. Daha sonraki yapılacak çalışmalarda insani kalkınmışlık endeksinin sınıflandırılması hem araştırmacılara hem de literatüre katkı sağlayacağı düşünülmektedir.

## KAYNAKÇA

Akın, H.B., ve Şentürk, E., (2012). "Bireylerin Mutluluk Düzeylerinin Ordinal Lojistik Regresyon Analizi İle İncelenmesi", Öneri, C.10, S.37, 183-193.

Akkuş, Ö, Türkan, S., Tatlıdil, H., (2010). "Sıralı Bağımlı Değişken Modeli Ve Diskriminant Analizi'nin Ticari Bankaların Mali Performanslarına Göre Sınıflandırılmasında Kullanımı", Suleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences, Vol.15, No.2, 319-332.

Alpar, R., (2013). Çok Değişkenli İstatistiksel Yöntemler, 4. Baskı, Detay Yayıncılık, Ankara,

Bilgehan, M., and Turgut, P., (2010). "The Use of Neural Networks In Concrete Compressive Strength Estimation", Computers and Concrete, 7(3), 2010, 271-283.

Burmaoğlu, S., Oktay, E. ve Özen, Ü. (2009). "Birleşmiş Milletler Kalkınma Programı Beşeri Kalkınma Endeksi Verilerini Kullanarak Diskriminant Analizi ve Lojistik Regresyon Analizinin Sınıflandırma Performanslarının Karşılaştırılması", KHO Savben Dergisi, Cilt:8, Sayı:2, 23-49.

Chen, C-K., and Hughes, J. Jr., (2004). "Using Ordinal Regression Model to Analyze Student Satisfaction Questionnaires", IR Applications, Association for Institutional Research, Vol. 1,1-13.

Demirtas E. A., Anagun A.S., Koksal G., (2009). "Determination Of Optimal Product Styles By Ordinal Logistic Regression Versus Conjoint Analysis For Kitchen Faucets". International Journal of Industrial Ergonomics, 39, 866–875.

Elamir, E., ve Sadeq, H., (2010). "Ordınal Regression to Analyze Employees Attitudes Towards The Application of Total Quality Management", Journal of Applied Quantitative, Vol.5, No.4, 647-658.

Elman, J.L., (1990). "Finding structure in time", Cognitive Science, 14(2), 179-211.

Eren, M., Çelik, A.K., Kubat, A., (2014). "Determinants of the Levels of Development Based on the Human Development Index: A Comparison of Regression Models for Limited Dependent Variables", Review of European Studies, Vol. 6, No. 1, 10-22.

Haykin, S., (1999). Neural Networks: A Comprehensive Foundation, Prentice Hall Inc, New Jersey.

Ivanova, I., Arcelus F.J. and Srinivasan G., (1999). "An assessment of The Measurement Properties of The Human Development Index", Social Indicators Research, 46, 157-179.

Kohonen, T. (1987). "State of the Art in Neural Computing", Int. Conf. on AI,1-79.

Kohonen, T., (2001). Self-Organizing Maps, Springer Series in Information Sciences, (3rd Edition), New York.

Küçükönder, H., (2011). "Yapay Sinir Ağları ve Tarımda Bir Uygulama", (Yayımlanmamış Doktora Tezi), Kahranmanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü, Kahramanmaraş.

Liao, T. F., (1994). Interpreting Probability Models, Logit, Probit, and Other Generalized Linear Model, Sage Publications, In Sage University Papers Series. Thousand Oaks. California

Lind Niels C., (1992). "Some Thoughts on The Human Development Index", Social Indicators Research, 27, 89-101.

Lippman, R. (1987)., "An introduction to computing with neural nets", IEEE ASSP Mag., 4, 4-22.

Long, J. Scott, ve Freese, J., (2001). Regressıon Models For Categorıcal Dependent Varıables Using Stata, A Stata Press Publication, Texas.

McCullagh, P., (1980). "Regression Models for Ordinal Data", Journal of the Royal Statistical Society. Series B (Methodological), Vol. 42, No. 2, 109-142.

Morse, S., (2014). "Stirring The Pot. Influence of Changes in Methodology of The Human Development Index on Reporting By The Pres", Ecological Indicators 45: 245–254.

O'Connell, A.A., (2006). Logistic Regression Models for Ordinal Response Variables. Thousand Oaks, Sage Publications, London.

Öztemel, E., (2003), Yapay Sinir Ağları", Papatya Yayıncılık, İstanbul.

Pissarenko, D., (2002). Neural Networks For Financial Time Series Prediction: Overwiev Over Recent Research, http://members.inode.at/d.pissarenko/fyp/Pissarenko2002.pdf

Şen, Z., (2004). Yapay Sinir Ağlarının İlkeleri, Özener Matbacılık, İstanbul.

UNDP, Human Development Report, 1990.

UNDP, Human Development Report, 2010.

UNDP, Human Development Report, 2011.

UNDP, Human Development Report, 2012.

UNDP, Human Development Report, 2014.

Üçdoğruk, Ş., Akın, F., ve Emeç, H., (2001). "Hanehalkı Harcamaların Olasılıklarını Sıralı Regresyon Modeli ile Tahmin Etme", V. Ulusal Ekonometri ve İstatistik Sempozyumu, Adana. (http://idari.cu.edu.tr/sempozyum/bil13.htm).

Yıldız, B., (2001). "Finansal Başarısızlığın Öngörülmesinde Yapay Sinir Ağı Kullanımı ve Halka Açık Şirketlerde Ampirik Bir Uygulama", İMKB Dergisi, Sayı:17, 51-67.

**EK-1 Ülkelere ait İnsani Kalkınmışlık Düzeyi**

| Ülke | Kalkınmışlık Düzeyi | Ülke | Kalkınmışlık Düzeyi | Ülke | Kalkınmışlık Düzeyi | Ülke | Kalkınmışlık Düzeyi |
|---|---|---|---|---|---|---|---|
| Almanya | Çok Gelişmiş | **Guatemala** | Orta | **Latvia** | Gelişmiş | **Sırbistan** | Gelişmiş |
| Amerika | Çok Gelişmiş | **Gürcistan** | Gelişmiş | **Litvanya** | Gelişmiş | **Slovenya** | Çok Gelişmiş |
| Avustralya | Çok Gelişmiş | **Hırvatistan** | Gelişmiş | **Lüksemburg** | Çok Gelişmiş | **Solovakya** | Çok Gelişmiş |
| Avusturya | Çok Gelişmiş | **Hollanda** | Çok Gelişmiş | **Macaristan** | Çok Gelişmiş | **Srilanka** | Orta |
| Azerbaycan | Gelişmiş | **Honduras** | Orta | **Makedonya** | Gelişmiş | **Suriye** | Orta |
| Belçika | Çok Gelişmiş | **İngiltere** | Çok Gelişmiş | **Malezya** | Gelişmiş | **Suudi Arabistan** | Gelişmiş |
| Bolivya | Orta | **İrlanda** | Çok Gelişmiş | **Malta** | Çok Gelişmiş | **Şili** | Gelişmiş |
| Bosna Hersek | Gelişmiş | **İspanya** | Çok Gelişmiş | **Meksika** | Gelişmiş | **Tayland** | Orta |
| Brezilya | Gelişmiş | **İsrail** | Çok Gelişmiş | **Mısır** | Orta | **Trinidad ve Tobago** | Gelişmiş |
| Bulgaristan | Gelişmiş | **İsveç** | Çok Gelişmiş | **Moldovya** | Orta | **Tunus** | Gelişmiş |
| Cezayir | Gelişmiş | **İsviçre** | Çok Gelişmiş | **Nikaragua** | Orta | **Türkiye** | Gelişmiş |
| Çek Cumhuriyeti | Çok Gelişmiş | **İtalya** | Çok Gelişmiş | **Norveç** | Çok Gelişmiş | **Ukrayna** | Gelişmiş |
| Çin | Orta | **İzlanda** | Çok Gelişmiş | **Pakistan** | Orta | **Uruguay** | Gelişmiş |
| Danimarka | Çok Gelişmiş | **Jamaika** | Gelişmiş | **Panama** | Gelişmiş | **Ürdün** | Gelişmiş |
| Dominik Cumhuriyeti | Orta | **Japonya** | Çok Gelişmiş | **Paraguay** | Orta | **Venezülla** | Gelişmiş |
| Ekvador | Orta | **Kanada** | Çok Gelişmiş | **Peru** | Gelişmiş | | |
| El Salvador | Orta | **Kazakistan** | Gelişmiş | **Polonya** | Çok Gelişmiş | | |
| Estonya | Çok Gelişmiş | **Kıbrıs** | Çok Gelişmiş | **Portekiz** | Çok Gelişmiş | | |
| Fas | Orta | **Kırgızistan** | Orta | **Romanya** | Gelişmiş | | |
| Filipinler | Orta | **Kolombiya** | Gelişmiş | **Rusya** | Gelişmiş | | |
| Finlandiya | Çok Gelişmiş | **Kore** | Çok Gelişmiş | **Yeni Zelanda** | Çok Gelişmiş | | |
| Fransa | Çok Gelişmiş | **Kosta Rika** | Çok Gelişmiş | **Yunanistan** | Çok Gelişmiş | | |

# Comparison Of Classification Success Of Human Development Index By Using Ordered Logistic Regression Analysis And Artificial Neural Network Methods

**Emre YAKUT**
Osmaniye Korkut Ata University, Faculty of Economics and Administrative Sciences, Osmaniye, Turkey
emreyakut@osmaniye.edu.tr

**Murat GÜNDÜZ**
Uşak University
Faculty of Economics and Administrative Sciences, Uşak, Turkey
murat.gunduz@usak.edu.tr

**Ayhan DEMİRCİ**
Turkish Armed Forces
kho1993@hotmail.com.tr

## Extensive Summary

### Introduction

Human development is the process of enhancing and improving people's life skills. This process aims to make a positive contribution to human and their living standards by equipping them with skills and capacity (UNDP 1990, p.1). By its *Human Development Index* (HDI) developed in 1990, United Nations Development Program (UNDP) takes a more composite and human oriented perspective by taking education, health and welfare dimensions of development into consideration by widening the perspective which is focused narrowly on economic growth only (Lind, 1992, p.89). Until 2010, GDP calculated per person based on purchasing power parity was taken into consideration for the economic dimension while life expectancy since birth was used for the health dimension and literacy and schooling were used for the education dimension. HDI calculates the arithmetic mean. Both economy and health dimension has one indicator while education dimension has two being literacy (2/3) and schooling (1/3) (Ivanova et al, 1999, pp.159-160). In 2010, index calculation was significantly changed. In this context, index calculation was based on arithmetic average instead of geometric average. With regards to education dimension, literacy among adults was excluded and the average of schooling rate and estimated schooling rate was considered (Morse 2014, p.249). HDI is scored between 0 and 1. 1 shows the highest human development status. The human development report in 2014 stated 4 levels of human development as very high, high, moderate and low. Countries with HDI value lower than 0,550 was classified as low, 0,550–0,699 as moderate, 0,700–0,799 as high and higher than 0,800 as very high (UNDP 2014, p.156). The purpose of this study is to compare the success of multiple classifications and to determine the effective factors by using logistic regression analysis and Elman ANN, multi-layer ANN and LVQ network. This study is comprised of 3 parts. In the first part, logistic regression analysis is introduced while the second part focuses on Elman ANN and LVQ network. In the third part, application results are compared.

### 1. Ordered Logistic Regression Model

Logistic regression models are used for modelling the relation between dependent variables measured in different categories and independent variables of categorical or continuous measurement. Ordered logistic regression (OLOGREG) is

used when dependent variables consists of at least three categories and measured by ordinal scale (Demirtas v.d., 2009, p.869).

The main features of ordered logistic regression model are as follows (Chen and Hughes, 2004, p.4):

- ✓ Outcome variable of categorical and ordinal measurement is a variable, which can be rearranged multiple times from an unobserved continuous latent variable, however it's not clear whether the space between the categories of this ordinal outcome variable is equal.
- ✓ Ordered logistic regression analysis, uses a correlation function to explain the effects of independent variables on ordered and categorical outcome variable. This model does not require normality and constant variance assumption.
- ✓ Since regression coefficient value is not dependent on the categories of categorical output variable, ordered logistic regression model assumes that the relation between explanatory variables and ordered categorical output variable is independent from categories.

Ordered logistic regression model is actually based on the existence of an continuous and unobserved random $Y^*$ latent variable under a categorical dependent Y variable. The categories of this variable are estimated as sequential intervals on a continuous plane named as cut-off point or threshold value (McCullagh, 1980, p.109).

The most important assumption in ordered logistic regression model is the assumption of parallel curves. According to this assumption, regression parameters obtained in the model is the same in all categories of the dependent variable. In other words, the relation between independent variables and dependent variable does not change according to the categories of dependent variable, and parameter estimations do not change according to different threshold values. Thus, if there's a dependent variable of J category, "$\beta_k$ " parameter is only one. On the other hand, there is $\theta_{j-1}$cut-off point (threshold value) for J-1 logit comparisons (Akin and Senturk, 2012, p.185).

It's challenging to interpret parameters in ordered logistic regression. Methods of calculation of standardizes coefficients, calculation of estimated probabilities, calculation of factor change in estimated probabilities and percentage change in estimated probabilities are used for interpreting parameters. Odds ratio can also be used for interpreting parameters. In the event that all other variables are held constant, $\exp(\beta_k)$ is odd ratio for dummy variable. To standardize odds ratios, $s_k$: showing standard deviation, $\exp(\beta_k * s_k)$ is calculated provided that all other variables are held constant. For continuous variables; the percentage is found by $[\exp(\beta - 1) * 100]$(Ucdogruk vd., 2001).

## 2. Artificial Neural Networks

ANNs are cellular systems that can receive, store and use information. ANNs are parallel systems, which are formed by connecting many connecting elements with links of variable weights. Multi-layer artificial neural network is the most popular one among many artificial neural networks (Lippman 1987, p.15). ANN is a system based on simple neural networks, which can receive interconnected information as input, process them and submits to other units, and which can even use the outputs as inputs again (Pissarenko, 2002, p.35). ANN simulates the operation of a simple biological neural

system. ANNs provide solutions to problems, which normally requires a person's natural ability to think and observe.Artificial neural networks are computer systems which are developed to perform some characteristics of a human brain automatically without getting any help such as getting new information through learning, creating new information and discovering (Oztemel, 2003, p.29). Artificial neural networks are used to achieve one or more processes including learning through using available data, associating, classification, generalization and optimization (Sen, 2004, p.13).

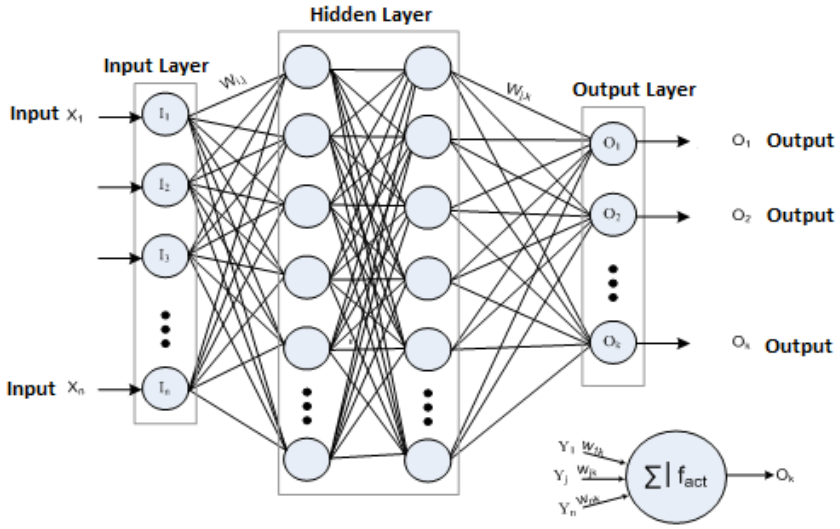



**Figure 1. General Structure of an Artificial Neural Network** (Bilghan and Turgut, 2010, p.275).

Use of artificial neural networks in such problems whose algorithmic solution could not be found has increased due to the fact that artificial neural networks can find solutions to new occurrences by way of examining former instances and learning the relationship between inputs and outputs of the said occurrence, regardless of whether the relationship is linear or not, from the current instances in hand. The biggest problem in artificial neural networks is that there is a need for such artificial neural networks that contain either very large neurons or multi-layered and a great amount of neurons in order to solve complicated problems (Kohonen,1987, pp. 1-79). An artificial neural network is an intensively parallel-distributed processor which is comprised of simple processing units, has a natural tendency to collecting experiential information, and enabling them to be used (Haykin, 1999, p.2). In a general artificial neural network system, neurons gather on the same direction to form layers (Yildiz, 2001, pp.51-67). There is parallel flow of information from the input layer to the exit in an architectural structure. Such flow is possible with parallel placed cells.

## 3. Method

Economic development and growth are among the most important objectives for all countries. For this reason, human development index has become a widely preferred and recognized numerical indicator for comparison and classification of countries. The purpose of this research is to compare the classification success of Human Development Index and determine the by using ordered logistic regression as well as Elman ANN, multi layer ANN and LVQ network as artificial neural networks.

In this study, data of highly developed, developed, moderately developed and less developed countries produced by Human Development Report Office of United Nations Development Program (UNDP), which is published annually. This data covers a period of three years and is obtained from 2010-2012 Human Development Index published at United Nations Development web site.Since it was difficult to find data about less developed countries, they are excluded from the analysis. The study classifies 81 countries based on data of three years. With the addition of one more country in 2012, the universe of the study is total of 244 countries.

Income, education and life expectancy are indispensable components in calculation of human development index. Since human development level would be subject to multiple classification process, 11 independent variables were added to these three variables, thus 14 independent variables were studied in order to enable a more detailed examination of methods. The following variables were used in the analysis respectively:

X1: BÖO "infant mortality rate",
X2: GSMH "gross national product",
X3: LKO "high school enrolment",
X4: BY "growth",
X5: DYY "direct foreign investment",
X6: ET "energy consumption",
X7: EÜ "energy production",
X8: E "inflation",
X9: IH "export",
X10: IKS "number of internet users",
X11:ISZ "unemployment",
X12: ITH "import",
X13: MTAS "number of mobile phone subscribers",
X14: SH "health expenses"

Answer variables for Human Development are coded as follows for Ordered Logistic Regression Analysis:

0: Moderately Developed,
1: Developed,
2: Highly Developed.

### 3.1. Data Analysis Method

Countries are classified for their Human Development Level by using Stata 11.2 package, Ordered Logistic Regression and Matlab 2012 software with Elman, Multi Layer Neural Networks and LVQ Network methods. In artificial neural network method, if a country is highly developed in human development, it is valued as 1 while the others' output values are stated as 0.

### 3.2. Ordered Logistic Regression Analysis

Stata 11.2 statistical analysis program is used to classify Human Development Index of countries.

**Results of Ordered Logistic Regression Analysis of Variables Effecting Human Development Level**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Number of obs | | 244 |
| | | | | | | **LRchi2(12)** | | **316,95** |
| | | | | | | **Prob> chi2** | | **0.0000** |
| Log likelihood | | -110,706 | | | | **Pseudo R2** | | **0,617** |
| HDII | **Coef.** | **Std. Err.** | **Wald** | **z** | **P>z** | **Odds Ratio** | **[95% Conf. Interval]** | |
| BÖO | -0,132 | 0,035 | 13,985 | -3,740 | 0,000 | 0,876 | 0,817 | 0,939 |
| GSMH | 0,000 | 0,000 | 1,840 | 1,360 | 0,174 | 1,000 | 1,000 | 1,000 |
| LKO | 0,017 | 0,013 | 1,749 | 1,320 | 0,186 | 1,017 | 0,992 | 1,043 |
| BY | -0,076 | 0,063 | 1,482 | -1,220 | 0,223 | 0,926 | 0,819 | 1,048 |
| DYY | 0,000 | 0,000 | 0,175 | -0,420 | 0,674 | 1,000 | 1,000 | 1,000 |
| ET | 0,000 | 0,000 | 0,603 | -0,780 | 0,438 | 1,000 | 1,000 | 1,000 |
| EÜ | 0,000 | 0,000 | 0,011 | -0,110 | 0,916 | 1,000 | 1,000 | 1,000 |
| E | 0,051 | 0,037 | 1,837 | 1,360 | 0,175 | 1,052 | 0,978 | 1,131 |
| IH | 0,058 | 0,023 | 6,475 | 2,540 | 0,011 | 1,059 | 1,013 | 1,108 |
| IKS | 0,104 | 0,020 | 26,893 | 5,190 | 0,000 | 1,109 | 1,067 | 1,154 |
| ISZ | -0,005 | 0,044 | 0,012 | -0,110 | 0,914 | 0,995 | 0,914 | 1,084 |
| ITH | -0,062 | 0,024 | 6,788 | -2,610 | 0,009 | 0,940 | 0,897 | 0,985 |
| MTAS | -0,006 | 0,009 | 0,506 | -0,710 | 0,477 | 0,994 | 0,977 | 1,011 |
| SH | 0,237 | 0,110 | 4,616 | 2,150 | 0,032 | 1,267 | 1,021 | 1,573 |
| /cut1 | 1,563 | 1,539 | | | | | -1,453 | 4,580 |
| /cut2 | 6,487 | 1,614 | | | | | 3,323 | 9,651 |

Results of the ordered logistic regression analysis of dependent and independent variables using mlogit command in Stata 11.2 package program are shown in Table 5. As seen in Table 5, the number of observations is 244 in the model and $\chi^2$value is statistically significant (p <0.01). Log likelihood value of the model was found to be -110.71. The first column of Table 5 shows β coefficients of ordered logistic regression analysis. BÖO "infant mortality rate", IKS "number of internet users" and ITH "import" with 0.01 significance level and IH "export" and SH "health expenses" with 0.05 significance level were observed as statistically significant. BÖO and ITH variables above the statistically dependent variable is marked negative while estimated value of İKS, İH and SH variables are marked positive. In addition, marginal effects will be calculated for the change in probabilities of dependent variable pursuant to change in β coefficients.

In odds ratio, one unit increase in BÖO variable decreases odds of high level human development rate by 12.4% while one unit increase in ITH variable decreases it by 6% provided that all other independent variables are held constant against moderate and low level of human development rate. One unit increase in IH variable increases odds of high level human development rate by 5.9% while one unit increase in IKS variable increases it by 10.9% and one unit increase in SH variable increases it by 26.7% against moderate and low level of human development rate. Thus, the most important variable that has a positive effect on human development level is SH "health expenses" variable, the second one is IKS "number of internet users" variable and the third one is IH "export" variable while the most important variable with negative effect is BÖO "infant mortality rate". Table 7 shows the results of 244 countries' human development level classification results by using ordered logistic regression analysis.

### 3.3. Artificial Neural Networks Analysis

Matlab R2012a computer software was used to establish artificial neural network models. In order to determine the appropriate artificial neural network method,

trial-and-error method is commonly used and many tests are performed. In this context, different combinations of parameters including number of hidden layers, number of nodes in hidden layers, momentum term, activation function, number of cycles were tested on both the training set and the test set to find the best performing network. Elman artificial neural network, multi layer artificial neural network and LVQ network were used as artificial neural networks in this study.

### 3.4. Findings and Conclusion

In this study, human development level of countries were classified by using ordered logistic regression, Elman ANN, multi layer ANN and LVQ network. Multiple classification methods were used to determine the relation between moderately developed, developed and highly developed countries as the dependent variable with ordered variable of 3 categories and 14 independent variables. The data of 81 countries, which has United Nations Development Program's Human Development Index, between the years of 2010-2012 were used in this study.

In ordered logistic regression analysis among 14 independent variables, it was observed that BÖO "infant mortality rate" and ITH "import" had a statistically significant negative effect while IKS "number of internet users", IH "export" and SH "health expenses" a statistically significant positive effect. Thus, the most important variable that has a positive effect on human development level was SH "health expenses" variable and the second one was IKS "number of internet users" varible while the most important variable with negative effect was BÖO "infant mortality rate". The classification of Turkey, which was classified as a developed country with regards to human development in 2010, 2011 and 2012, was successfully estimated by using ordered logistic regression analysis. Chart 1 shows percentage values of all four methods of analysis for moderate, development and high development levels and the overall classification accuracy.

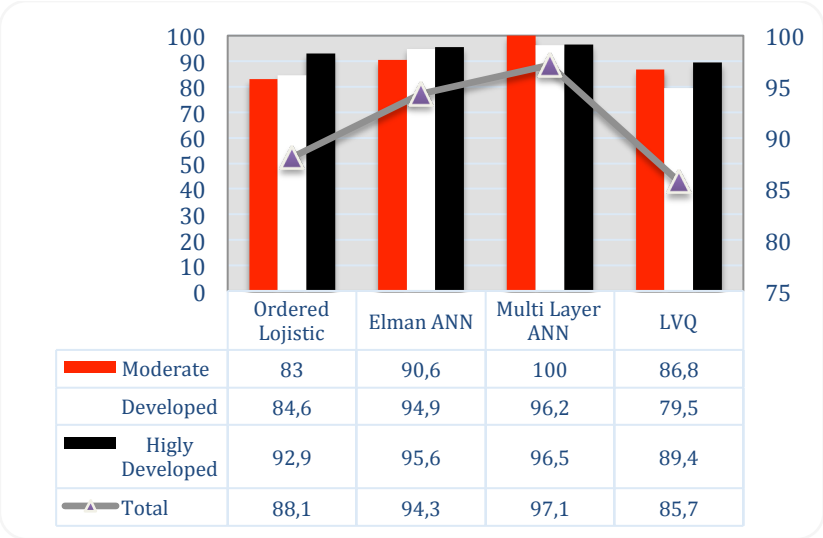



**Chart 1. Performance Criterion for All Four Methods of Analysis**

Multi Layer ANN analysis had the best performance among all. Its classification success was 100% for moderately developed countries, 96.2% for developed countries and 96.5% for highly developed countries. As a result of comparison of analyses, it's seen that Muli Layer ANN provides results with a higher accuracy percentage compared

to Elman ANN while ordered logistic regression analysis provides results with a higher accuracy compared to LVQ network. In all four methods of analysis, it was proved that Multi Layer ANN had a better performance compared to the other three methods with regards to total classification results of moderately, developed and highly developed countries.