



İnsani Gelişmişlik Endeksinin Karar Ağacı Algoritmaları ile Modellenmesi: BM'de Bir Uygulama 2010-2017 Dönemi

Modeling of Human Development Index with Decision Tree Algorithms:
An Application at the UN 2010-2017 Period

Dr. Öğr. Üyesi Emre Yakut¹ - Ali Korkmaz²

Başvuru Tarihi: 24.12.2019

Kabul Tarihi: 12.06.2020

Makale Türü: Araştırma Makalesi

Öz

İnsani Gelişme Endeksi (İGE), ülkelerin gelişmişliklerini gözönünde bulundurarak insanların mutluluğunu, sağlıklı bir yaşam ile birlikte başarılı bir hayat sürmelerini dikkate alan bir kalkınma endeksidir. Ülkelerin milli gelirlerini karşılaştırarak o ülkenin daha gelişmiş olduğunu açıklamak yeterli değildir. Birleşmiş Milletler Kalkınma Programı (UNDP) tarafından yayımlanan İGE, insan hayatının daha kaliteli bir hayat sürmesi açısından ülkelerin milli geliri, eğitim ve sağlık göstergelerine dayanarak hesaplanmaktadır. Dolayısıyla İGE, insan hayatının zenginliği açısından ülkelerarası karşılaştırma yapmak için başvurulan bir gösterge değer olmuştur. Bu çalışmada UNDP'nin 2010-2017 yıllarını kapsayan 79 ülkenin verileri kullanılarak veri madenciliğinin karar ağacı tekniklerinden C5.0 ve Gini algoritmaları ile karar ağaçları oluşturulmuştur. Karar ağaçları ile birlikte İGE'ye etki eden faktörler belirlenmiş ve ülkeler çok yüksek, yüksek, orta ve düşük düzey gelişmiş ülkeler olarak sınıflandırılarak kurallar elde edilmiştir. Yapılan analizler sonucunda C5.0 algoritması ile %97,94 ve Gini algoritması ile %91,93'lük doğru sınıflandırma başarısı elde edilmiştir. Bunun dışında duyarlılık ve belirleyicilik istatistikleri de hesaplanmıştır. İGE'ye en fazla etki eden değişkenlerin eğitim, istihdam ve sağlık göstergelerindeki değişkenler olduğu tespit edilmiştir.

Anahtar Kelimeler: *İnsani Gelişme Endeksi, Veri Madenciliği, C5.0, Gini Algoritması*

Abstract

Human Development Index (HDI) is an indicator that ranks nations by their development in health, welfare and success of their citizens. Using national income alone is not sufficient to determine the level of development of a nation. HDI, published by United Nations Development Plan (UNDP), is calculated based on nation's income, education and health indicators. Hence, HDI became a metric to rank the quality of human life between countries. This study contains

¹ Osmaniye Korkut Ata Üniversitesi, İİBF Fakültesi, YBS Bölümü, emreyakut@osmaniye.edu.tr, ORCID: 0000-0002-1978-0217

² Osmaniye Korkut Ata Üniversitesi, SGDB Şube Müdürü, alikorkmaz@osmaniye.edu.tr, ORCID: 0000-0001-8581-7615

decision trees that are generated using C5.0 and GINI Algorithms based on UNDP data for 79 countries dating from 2010 till 2017 In addition to the decision trees, the factors that affect HDI also were determined and rules were produced by grouping countries as highly developed, developed, developing and under developed countries. Based the analyses, C5.0 and GINI algorithms produced 97.94% and 91.93% ranking accuracy respectively. Additionally, sensitivity and determinability statistics were also calculated, and the variables in education, employment and health were determined to be the most significant.

Keywords: *Human Development Index, Data Mining, C5.0, Gini Algorithm*

Giriş

Ülkelerin gelişmişliğinin hangi düzeyde olduğu, daha gelişmiş bir ülke sınıfına atlamak için hangi faktörlerin dikkate alınması gerektiği insanların yaşam kalitesini karşılaştırma açısından önem arz etmektedir. Bu bağlamda ülkeleri karşılaştırmak için büyüme ve gelişme kavramlarının açıklanması gerekmektedir. Büyüme, bir ülkenin belirli bir dönem içerisinde üretim kapasitesinde ve GSYH’de meydana gelecek artışın devamlı olmasını ifade eder ve iktisadi anlamda sayısal olarak nicel artışı ya da azalışı gösterir (Taban, 2016; Üzümcü, 2012; Ülgener, 1991). Gelişme, kalkınma ile aynı anlamı ifade etmekte olup, ülke ekonomisinin istikrarının devamı için ülkenin kıt kaynaklarını etkin ve verimli bir şekilde kullanarak insanların yaşam kalitesinin iyileştirilmesi amacıyla kullanılmaktadır (Berber, 2017; Erbay ve Özden, 2013; Üzümcü, 2012). İnsani gelişmişlik göstergesi, UNDP tarafından iktisatçı Mahbub ul Hap’ın ekibinin toplumsal kalkınmanın merkezinde insanların hayat kalitesine ve mutluluğuna dikkat çekerek belirledikleri insani gelişme endeksiyle (İGE) ölçülmektedir (Eren ve Kaynak, 2017; Tunç ve Ertuna, 2015). İGE kavramı, bir ülkenin gelişmişlik düzeyini açıklarken, aynı zamanda ülkenin ekonomik, sosyal ve politik gösterge değerlerinin insanların yaşam kalitesine ne düzeyde katkıda bulunduğunu göstermektedir (Zanbak ve Özgür, 2019; Rençber, 2018). İGE, gelişmişlik düzeyinin belirlenmesinde sağlık, eğitim ve gelir olmak üzere eşit önem derecesine sahip bu üç faktörle değerlendirilmektedir (Günsoy, 2012). İnsani gelişme için öncelikle bahsedilen faktörler iyileştirildiğinde insan yaşamının diğer faaliyetlerinde de gelişme imkânı sağlanacaktır (Ünal, 2008).

İGE, ülkelerin gelişmişliklerini göz önünde bulundurarak insanların mutluluğunu, sağlıklı bir yaşam ile birlikte başarılı bir hayat sürmelerini dikkate alan bir kalkınma endeksidir. Ekonomik kalkınmanın aksine İGE GSMH’ye alternatif olarak insan hayatını konu alan bir hesaplama ve değerlendirme yöntemidir. Bu hususlar dikkate alındığında insani gelişme endeksi, insan hayatının daha kaliteli bir yaşam sürmesi açısından çok önemli bir rol oynamaktadır. Dolayısıyla İGE, insan hayatının zenginliği açısından ülkelerarası karşılaştırma yapmak için başvurulan bir gösterge değer olmuştur.

İnsani Gelişme Endeksinin Hesaplanmasında Kullanılan Göstergeler

İGE’de bir ülkenin gelişmişlik düzeyi için değerlendirilen temel faktörler, aşağıda hesaplama yöntemleriyle birlikte açıklanmıştır.

Sağlık Endeksi: Uzun, sağlıklı ve yaratıcı bir hayatı kapsayan doğumda yaşam beklentisini (DYB) açıklamaktadır (Gülel, vd., 2017, s. 210). Sağlık endeksinin yüksek olması, sağlıklı bir toplumda hem sağlık hizmetlerinin hem de yaşam kalitesinin iyi olduğu anlamına gelmektedir (Sakarya ve İbişoğlu, 2015, s.222). Uzun ve sağlıklı bir yaşamı ifade eden sağlık endeksi eşitlik 1'deki denklem ile hesaplanmaktadır (Şahin ve Gökdemir, 2016, s.8).

$$I_S = \frac{L - L_{min}}{L_{max} - L_{min}} \quad (1)$$

Formüldeki L doğumda yaşam beklentisini, L_{min} belirlenen alt yaş sınırını, L_{max} ise belirlenen üst yaş sınırını göstermektedir.

Eğitim Endeksi: Bilgiye ulaşımı ifade eden eğitim endeksi, ortalama okullaşma yılları (MYS) ve beklenen okullaşma yıllarına (EYS) ait değerleri kullanarak eşitlik 2'deki denklemle hesaplanır (Şahin ve Gökdemir, 2016, s.8).

$$I_E = \left(\frac{MYS - MYS_{min}}{MYS_{max} - MYS_{min}} + \frac{EYS - EYS_{min}}{EYS_{max} - EYS_{min}} \right) / 2 \quad (2)$$

Gelir Endeksi: Kaliteli bir hayat standardını açıklayan endeks, GSMH değerlerini kullanarak eşitlik 3'teki denklem ile ölçülmektedir (Gülel, vd., 2017, s. 210; Şahin ve Gökdemir, 2016, s.8). $\ln(GSMH)$, bir ülkenin GSMH değerinin doğal logaritmasını, $\ln(GSMH)_{min}$ ve $\ln(GSMH)_{mak}$, gelir serisinde en düşük ve en yüksek gelirlere sahip ülkelerin GSMH değerlerinin doğal logaritmasını göstermektedir.

$$I_G = \frac{\ln(GSMH) - \ln(GSMH)_{min}}{\ln(GSMH)_{max} - \ln(GSMH)_{min}} \quad (3)$$

İGE, eşit ağırlıklandırma yapılan sağlık, eğitim ve gelir endekslerinin geometrik ortalaması alınarak eşitlik 4'teki denklem yardımıyla ölçülmektedir.

$$İGE = \sqrt[3]{I_S \times I_E \times I_G} \quad (4)$$

Ülkelerin İGE değerleri 0 ile 1 arasında değişkenlik göstermekte; 0 değeri en düşük ve 1 değeri en yüksek insani gelişme değerine sahip ülkeyi ifade etmektedir (Mıhçı ve Mıhçı, 2003, s. 32). UNDP tarafından hesaplanarak her sene yayımlanan İGE raporunda, ülkeler endeks değerlerine göre çok yüksek, yüksek, orta ve düşük insani gelişme gösteren ülkeler olarak dört kategoriye ayrılmaktadır (Kaya, 2017, s. 171). UNDP'nin 2018 yılı raporundaki İGE göstergelerinde yapılan değişikliklere bağlı olarak ülkelerin endeks değerlerine göre sınıflandırılması Tablo 1'de gösterilmektedir (Doğan ve Tatlı, 2014, s. 105; Tıraş, 2019, s. 20).

Tablo 1. İnsani Gelişme Endeks Değerine Göre Yapılan Sınıflandırma (UNDP, 2018b, s. 3)

Endeks Değeri	Kategori Tanımı
0.800 ≥	Çok Yüksek İnsani Gelişim
0.700 – 0.799	Yüksek İnsani Gelişim
0.550 – 0.699	Orta İnsani Gelişim
<0,549	Düşük İnsani Gelişim

Literatür Taraması

İGE ile ilgili literatürde yer alan çalışmaların bazıları aşağıda özetlenmiştir.

Erdem ve Çelik (2019), 33 Afrika ülkesinin 1995-2014 yılları arasındaki verilerinden yararlanarak İGE ile ekonomik büyüme verileri arasındaki ilişkiyi analiz etmişlerdir. Çalışma sonucunda İGE ile gelir değişkeni arasında istatistiksel açıdan pozitif ve anlamlı, eğitim değişkeni ile negatif ve anlamlı buna karşın sağlık değişkeni ile pozitif fakat istatistiksel açıdan anlamsız bir ilişkinin olduğunu tespit etmişlerdir.

Zanbak ve Özeş Özgür (2019), 1990-2017 yıllarına ait Türkiye'nin de içinde bulunduğu AB'ye aday ile AB üyesi ülkelerinin İGE değerlerini karşılaştırmışlar ve AB üyeliğinin endeks üzerindeki olası etkilerini değerlendirmişlerdir. Çalışma sonucunda kurucu üyelerin insani gelişme açısından günümüzde ileride, aday statüsünü devam ettiren Makedonya'nın tüm ülkeler arasında en düşük performansa sahip ülke olduğunu ve Türkiye'nin ise geride olmasına rağmen en süratli değişimi gerçekleştiren ülkelerin başında geldiği değerlendirmesini yapmışlardır.

Çağlar ve Keten (2018), Türkiye'deki illerin 2013 yılındaki İGE'sini, iki farklı VZA yöntemiyle hesaplamışlardır. Değerlendirme neticesinde Türkiye'nin doğusundaki illerin her iki yaklaşımda da son sırada bulunduğunu açıklamışlardır.

Çoban ve Yayar (2018), 2007-2014 yılları arasındaki AB ülkelerinin demokrasi göstergelerinin İGE üzerine etkisini panel veri analizi ile incelemeye çalışmışlardır. Çalışma sonucunda ülkelerin demokrasi göstergeleri yükseldikçe, mülkiyet haklarının korunması ve güçlendirilmesi arttıkça ülkelerin insani gelişmişlik düzeyinin arttığını açıklamışlardır.

Meydan ve Sarı (2018), 2014 yılı verilerini kullanarak Türkiye'deki illerin İGE ile sağlık, eğitim ve gelir endeksleri arasında bir ilişkinin olup olmadığını basit regresyon analizi ile araştırmışlardır. Çalışma neticesinde İGE'nin Türkiye'nin doğu illerinin batı illerine göre daha düşük değerler aldıklarını tespit etmişler, ayrıca eğitim ve gelir endeksinin sağlık endeksine göre İGE'yi açıklama gücünün istatistiksel açıdan anlamlı ve daha yüksek olduğunu belirtmişlerdir. Tıraş ve Ağır (2018), 57 İslam İşbirliği Teşkilatı ülkelerinin 2016 yılı verilerini kullanarak İGE'yi, eğitim, gelir ve sağlık endeksleri açısından değerlendirmişlerdir. Araştırma sonucunda, Türkiye'nin daha yüksek İGE seviyesine ulaşabilmesi için eğitim ve sağlık alanlarında daha fazla yatırım yapması gerektiğini tavsiye etmişlerdir.

Kaya (2017), 1986-2015 dönemi arasındaki İGE'nin BIST100 endeksi üzerinde bir etkisinin olup olmadığını zaman serileri analizi ile incelemiştir. Çalışma sonucunda Granger Nedensellik testine göre değişkenler arasında uzun süreli ilişkilerin varlığını ve İGE'nin BIST endeksinin bir nedeni olduğunu açıklamıştır.

Rençber (2017), 2015 yılı BM'nin resmi internet sitesinde yer alan 185 ülkenin verilerini kullanarak ÇLR, YSA ve ANFIS yöntemleriyle ülkelerin İGE değerlerini yeniden hesaplamış, gerçek grup üyelikleri ile karşılaştırmıştır. Yapılan değerlendirme sonucunda ANFIS yönteminin diğer iki yöntemle göre sınıflandırma performansının daha başarılı olduğunu tespit etmiştir.

Öztürk (2016), 1990-2015 yılları arasındaki 8 ülkenin İGE değerlerini ülkelerin ekonomik gelişmişlik göstergesi değerleri ve beslenme ilişkileri ile karşılaştırmıştır. Çalışmanın sonucunda Türkiye'nin İGE değeri ve beslenme yetersizliği dikkate alındığında, İran ile benzer özellikler gösterdiğini buna karşın Avrupa ülkeleri, Kore ve Amerika'dan ayrıştığını tespit etmiştir.

Yalçın ve Çakmak (2016), 1991-2013 yıllarına ait Türkiye'nin kamu sağlık harcamalarının İGE üzerine etkisini logaritmik regresyon analizi ile araştırmışlardır. Çalışma sonucunda kamu sağlık harcamaları ile İGE arasında istatistiksel açıdan pozitif anlamlı bir ilişkinin olduğunu tespit etmişlerdir.

Fırat ve Aydın (2015), 2014 yılına ait OECD ülkelerinin İGE ve eğitim endeksi değerlerini karşılaştırmışlardır. Değerlendirme neticesinde Türkiye'nin eğitim endeksinin OECD ülkelerine göre düşük olduğunu ve İGE düzeyinin yeterli düzeyde olmadığını vurgulamışlardır. Murat ve Gürsakal (2015), 2012 yılına ait UNDP'nin resmi internet sitesinde yer alan 150 ülkenin İGE ve Mutlu Gezegen Endeksi (HPI) arasındaki ilişkileri kanonik korelasyon analizi ile araştırmışlardır. Araştırma sonucunda İGE ve HPI arasında istatistiksel açıdan anlamlı ve güçlü bir kanonik ilişkinin olduğunu tespit etmişlerdir.

Paksoy (2015), 2014 yılı Avrupa ülkelerinin ve Türkiye'nin performanslarını ve gelişmişlik düzeylerini sağlık, eğitim ve gelir değişkenlerinin değerlerini kullanarak VIKOR yöntemiyle sıralamıştır. Çalışma sonucunda İsveç, Danimarka ve Hollanda gibi gelişmiş ülkelerin performanslarının daha yüksek olduklarını belirtmiştir.

Yakut, vd., (2015), 2010-2012 dönemine ait UNDP raporunda yer alan 81 ülkenin İGE değerlerini sıralı lojistik regresyon analizi ve YSA yöntemleri ile sınıflandırarak karşılaştırmışlardır. Çalışma sonucunda sıralı lojistik regresyon analizine göre sağlık harcamaları, ithalat ve ihracat, bebek ölüm oranı ile internet kullanıcı sayısı değişkenlerinin istatistik açıdan anlamlı olduğunu ve YSA'nın SLR analizine göre daha başarılı performans gösterdiğini tespit etmişlerdir.

Doğan ve Tatlı (2014), 1990-2012 yılları arasındaki UNDP raporunda yer alan ülkelerin İGE ve yoksulluk endekslerini hesaplayarak, Türkiye'nin performansını diğer ülkelere göre karşılaştırmışlardır. Yapılan değerlendirme neticesinde Türkiye'nin İGE performansının yükselmesi için eğitim endeksinin daha fazla iyileştirilmesi gerektiğini açıklamışlardır.

Kocatepe ve Dilber (2014), 2011-2012 yıllarına ilişkin Türkiye'nin eğitim, gelir ve sağlık endekslerine ait verileri kullanarak İGE değerini hesaplamışlardır. Çalışma sonucunda Türkiye'nin İGE değerine bağlı olarak diğer ülkeler içerisindeki sıralama sonucunu yorumlamışlardır.

Çemrek ve Bayraç (2013), 2010 yılına ait Bağımsız Devletler Topluluğu'ndan 12, Baltık Cumhuriyeti Devletleri'nden 3 ve Rusya dahil olmak üzere toplam 15 ülkenin ekonomik, sosyal ve çevresel nitelikli 10 farklı değişkenini kullanarak sürdürülebilir kalkınma skorları ile İGE değerleri arasındaki ilişkiyi incelemiştir. Analiz sonucunda SDK skorları ile İGE değerleri arasında istatistiksel açıdan anlamlı bir ilişkinin olduğunu tespit etmişlerdir.

Korkmaz ve Şahin (2013), OECD'ye üye olan ülkelerin 2009 yılı PISA sınav sonuçları ile İGE değerleri arasındaki ilişkisini araştırmışlardır. Araştırma sonucunda PISA sınav sonuçları ile İGE değerleri arasında olumlu bir ilişkinin olduğunu, Türkiye'nin PISA sınav sonucundan daha başarılı olması için İGE düzeyinin çok gelişmiş ülkeler kategorisinde yer alması gerektiğini ifade etmişlerdir.

Nartgün, vd., (2013), 2011 yılı UNDP raporuna göre AB'ye üye, aday ve aday aday ülkeler arasında İGE ve eğitim endeksi açısından bir farklılık olup olmadığını araştırmışlardır. Çalışma sonucunda Türkiye'nin hem İGE hem de eğitim endeksinin aday ülkelerin değerlerine göre düşük olduğunu açıklamışlardır.

Tireli, vd., (2013), 2007-2010 yıllarına ilişkin İslam İşbirliği Teşkilatının bir alt kuruluşu olan SESRIC tarafından oluşturulan 57 üye ülkenin eğitim, sağlık, istihdam verileri ile İGE değerleri arasında anlamlı bir ilişkinin olup olmadığını araştırmışlardır. Çalışma sonucunda İGE'nin eğitim verileri arasında anlamlı buna karşın sağlık harcamaları ve istihdam verileri arasında anlamlı olmayan bir ilişkinin olduğunu ortaya koymuşlardır.

Metodoloji

Veri madenciliği geniş hacimli olmak üzere karmaşık verileri kullanarak anlamlı ve faydalı bilgiler oluşturmaya çalışan bir bilim dalıdır. Başka bir ifadeyle veri madenciliği büyük boyutlu veri tabanlarından anlaşılması zor verileri çeşitli algoritmalarla işleyerek analizciler için özetlenmiş bilgiler ve kurallar elde edilmesine, geleceğe yönelik tahminler yapılmasına imkan sağlar. Bu çalışmada karar ağaçlarının sınıflandırma yöntemleri içerisinde yer alan C5.0 ve Gini algoritmaları anlatılacaktır.

C5.0 Algoritması: Quinlan, C5.0 algoritmasını Unix/Linux için geliştirmiş ve ticari olarak piyasaya sunmuştur (Akpınar, 2014, s. 222). C5.0 algoritması entropiyi kullanarak karar

ağaçlarını oluşturmaktadır. Tesadüfî bir değişkendirki belirsizliğin ölçüsü olan Entropi, 0 ile 1 arasında bir değer almakta, iki temelli logaritma ile hesaplanarak elde edilmektedir. Veri seti içerisinde bütün değerler tek bir sınıfa ait olduklarında entropi sıfır, eşit olarak birden fazla sınıfa ait olduklarında entropi 1 olacaktır.

C5.0 algoritmasıyla sınıflandırmada karar düğümlerini ayırmak için niteliklerin bilgi kazancı ve entropi değerleri maksimize edilerek karar ağacı oluşturulmaktadır. C5.0 algoritmasının karar ağacına ilişkin adımlar aşağıda gösterilmiştir (Akpınar,2014,s.208;Alpaydın,2013,s.156; Silahtaroglu, 2013, s.76; Özkan, 2008, s.58):

Adım 1: X değişkeni (nitelik) için k adet olasılıklar sırasıyla $P_1, P_2, P_3, \dots, P_k$ olarak isimlendirilmekte ve entropi değeri eşitlik 5 yardımıyla hesaplanmaktadır.

$$Entropi = H(T) = - \sum_{i=1}^k p_i \log_2 (p_i) \quad (5)$$

Adım 2: X değişkeninin hedef niteliği olan T'nin alt kümeleri $T_1, T_2, T_3, \dots, T_k$ olarak adlandırılmakta ve her bir T hedef niteliğinin sınıfını belirlemede kullanılan bilgilerin ağırlıklı ortalamaları ile entropilerin ağırlıklandırılmış toplamları eşitlik 6 ile belirlenmektedir.

$$H_S(T) = \sum_{i=1}^n p_i H_S(T_i) \quad (6)$$

Adım 3: Karar verme düğümünden elde edilen maksimum bilgi kazancının ayırma ölçütü eşitlik 7 yardımı ile hesaplanarak gerçekleştirilmektedir.

$$Bilgi Kazancı (S) = H(T) - H_S(T) \quad (7)$$

Gini Algoritması: Gini algoritması, sınıflandırma ve regresyon ağaçları algoritmalarından (CART) birisidir. Breiman (1984) tarafından Gini algoritması geliştirilmiştir. Gini algoritması, hedef nitelik değerlerinin olasılık dağılımları arasındaki farkı ölçen gini endeksiyle sınıflandırma işlemini gerçekleştirmektedir (Rokach ve Maimon, 2009, s. 63). Nitelik değerlerinin sol ve sağ tarafta olmak üzere iki sütuna bölünmesi ile $gini_{sol}$ ve $gini_{sağ}$ değerleri hesaplanarak minimum gini değerine karşılık gelen niteliklerle karar ağacı oluşturulur. Gini algoritmasıyla karar ağacı oluşturulurken aşağıdaki adımlar uygulanır (Akpınar, 2014, s. 212, Adak ve Yurtay, 2013, s. 2; Özkan, 2008, s. 106):

Adım 1: Gini değeri hesaplanmadan önce her bir nitelik için sol ve sağ taraftaki $Gini_{sol}$ ve $Gini_{sağ}$ değerleri eşitlik 8 ve eşitlik 9'daki denklemler yardımıyla hesaplanır.

$$Gini_{sol} = 1 - \sum_{i=1}^k \left[\frac{L_i}{|T_{sol}|} \right]^2 \quad (8)$$

$$Gini_{sağ} = 1 - \sum_{i=1}^k \left[\frac{R_i}{|T_l|} \right]^2 \quad (9)$$

$Gini_{sol}$ ve $Gini_{sağ}$ değerlerinin hesaplanmasında kullanılan sembollerin açıklamaları aşağıda verilmiştir.

k : Sınıfların sayısı,

T : bir düğümdeki örnek sayısı,

T_{sol} : sol sütundaki örneklerin sayısı,

$T_{sağ}$: sağ sütundaki örneklerin sayısı,

L_i : sol sütundaki i kategorisindeki örneklerin sayısı,

R_i : sağ sütundaki i kategorisindeki örneklerin sayısı.

Adım 2: Bir niteliğin Gini değeri eşitlik 10'daki $Gini_j$ denklemiyle hesaplanır.

$$Gini_j = \frac{1}{n} (|T_{sol}|Gini_{sol} + |T_{sağ}|Gini_{sağ}) \quad (10)$$

Adım 3: Her bir nitelik için hesaplanan Gini değerleri arasından minimum olanı seçilerek bölünme bu nitelik üzerinden gerçekleştirilmektedir. Daha sonra geri kalan veri seti için en baştaki adımlara tekrar dönülerek, diğer bölünelere ait gini değerleri ile karar ağacı oluşturulur.

Uygulama

Bu çalışmada, kategorik değişken olan İGE'nin veri madenciliğinin karar ağacı tekniklerinden Gini ve C5.0 algoritmalarıyla modellenmesi amaçlanmıştır. Çalışmanın ilk aşamasında spearman korelasyon analizi sonucunda birbirleriyle yüksek korelasyona sahip değişkenler belirlenerek analizden çıkarılmış, gini ve C5.0 algoritmalarıyla karar ağaçları oluşturularak kurallar belirlenmiş ve yöntemlerin sınıflandırma performanslarını karşılaştırmak için duyarlılık ve belirleyicilik istatistikleri hesaplanmıştır.

Veri Seti ve Değişkenler: BM tarafından her yıl yayımlanan insani gelişme raporunun 2010-2017 dönemine ait 189 ülkenin eksik verileri çıkarılarak geriye kalan 79 ülkenin gösterge değerleri çalışmanın kapsamını oluşturmuştur. Böylece ülkeler İGE'ye göre çok gelişmiş, gelişmiş, orta ve düşük düzey gelişmiş ülkeler olarak sınıflandırılmıştır. Tablo 2'de çalışma kapsamındaki ülkelerin 2010-2017 dönemine ait gelişmişlik düzeyleri gösterilmiştir.

Tablo 2. Gelişme Düzeylerine Göre Çalışma Kapsamındaki Ülkeler

Düşük Düzey Gelişmiş Ülkeler		Orta Düzey Gelişmiş Ülkeler		Gelişmiş Ülkeler		Çok Gelişmiş Ülkeler	
2010	2017	2010	2017	2010	2017	2010	2017
Bangladeş	Ruanda	Moğolistan	Filipinler	Rusya	Kosta Rika	Norveç	Norveç
Kenya	Uganda	Maldivler	G. Afrika	Karadağ	Türkiye	İsviçre	İsviçre
Zambiya	Sudan	Belize	Endonezya	Bulgaristan	Sırbistan	Avustralya	Avustralya
Pakistan	Malavi	Paraguay	Vietnam	Romanya	Arnavutluk	İrlanda	İrlanda
Ruanda		Moldova	Bolivya	Belarus	Gürcistan	Almanya	Almanya
Uganda		Filipinler	El Salvador	Uruguay	Meksika	İzlanda	İzlanda
Sudan		G. Afrika	Kırgızistan	Malezya	Sri Lanka	İsveç	İsveç
Malavi		Endonezya	Fas	Kosta Rika	Brezilya	Singapur	Singapur
		Vietnam	Nikaragua	Türkiye	Azerbaycan	Hollanda	Hollanda
		Bolivya	Guatemala	Sırbistan	Ermenistan	Danimarka	Danimarka
		El Salvador	Tacikistan	Arnavutluk	Çin	Kanada	Kanada
		Kırgızistan	Bangladeş	Gürcistan	Ukrayna	Birl. Krallık	Birl. Krallık
		Fas	Kenya	Meksika	Peru	Finlandiya	Finlandiya
		Nikaragua	Zambiya	Sri Lanka	Kolombiya	Belçika	Belçika
		Guatemala	Pakistan	Brezilya	Dominik C.	Avusturya	Avusturya
		Tacikistan		Azerbaycan	Ürdün	Kore Cum.	Kore Cum.
				Ermenistan	Surinam	Fransa	Fransa
				Çin	Moğolistan	Slovenya	Slovenya
				Ukrayna	Maldivler	İspanya	İspanya
				Peru	Belize	Çekya	Çekya
				Kolombiya	Paraguay	İtalya	İtalya
				Dominik C.	Moldova	Estonya	Estonya
				Ürdün		Yunanistan	Yunanistan
				Surinam		Kıbrıs	Kıbrıs
						Polonya	Polonya
						Litvanya	Litvanya
						Slovakya	Slovakya
						Letonya	Letonya
						Portekiz	Portekiz
						Macaristan	Macaristan
						Hırvatistan	Hırvatistan
						Rusya	Rusya
						Karadağ	Karadağ
						Bulgaristan	Bulgaristan
						Romanya	Romanya
						Belarus	Belarus
						Uruguay	Uruguay
						Malezya	Malezya

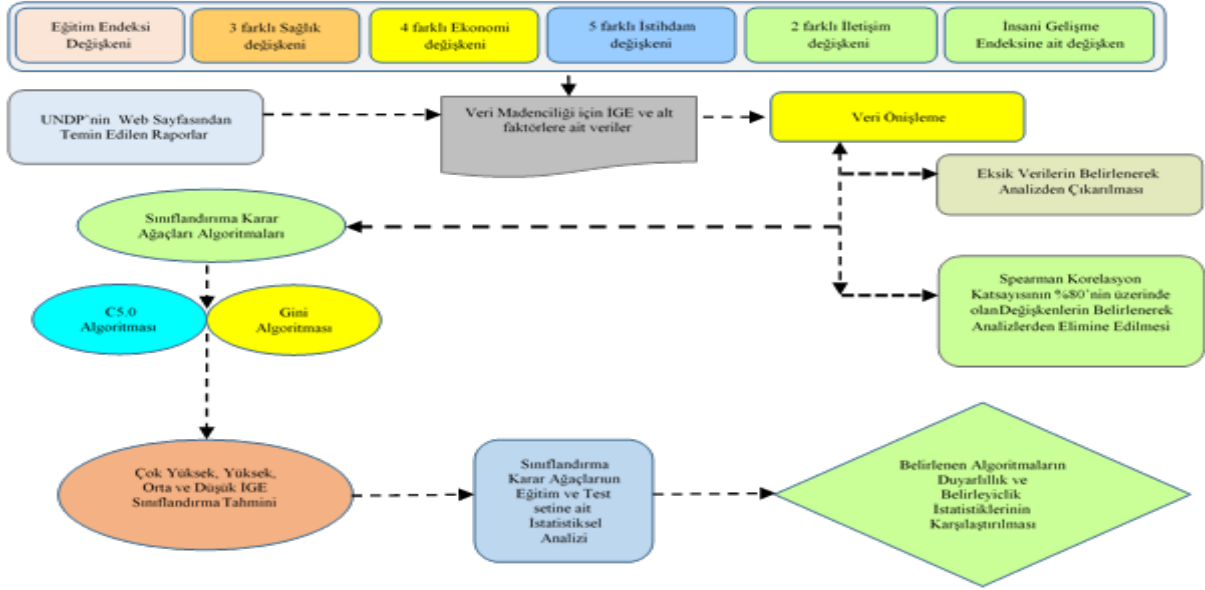
Kategorileştirilmiş İGE'ye etki eden faktörlerin belirlenmesi ve karar ağaçlarının oluşturulması için eğitim, sağlık, gelir, istihdam ve iletişim faktörlerinden toplam 15 bağımsız değişken belirlenmiştir. Tablo 3'te analizlerde kullanılacak değişkenler verilmiştir.

Tablo 3. İGE'nin Faktörlerine İlişkin Değişkenler

Kod	İGE ve Alt Endekslerin İngilizce ve Türkçe Karşılıkları	Faktörler
Y	Human development index (HDI)	İnsani gelişme endeksi
X1	Education Index	Eğitim endeksi
X2	Life expectancy at birth (years)	Doğumdan sonra beklenen yaşam süresi (yıl)
X3	Life expectancy index	Yaşam beklentisi endeksi
X4	Expected years of schooling (years)	Beklenen eğitim süresi (yıl)
X5	Gross domestic product (GDP), total (2011 PPP \$ billions)	Gayri Safi yurtiçi hâsıla (GSYİH), toplam (2011 PPP \$ milyar)
X6	Income index	Gelir endeksi
X7	Exports and imports (% of GDP)	İhracat ve ithalat'ın GSYİH payı (GSYİH'nin % 'si)
X8	Foreign direct investment, net inflows (% of GDP)	Doğrudan yabancı yatırımlar, net girişler (GSYİH'nin% 'si)
X9	Gender inequality index (GII)	Cinsiyet eşitsizliği endeksi
X10	Gender development index (GDI)	Cinsiyet gelişme endeksi
X11	Employment in agriculture (% of total employment)	Tarım sektöründeki istihdam oranı (toplam istihdamın % 'si)
X12	Unemployment, total (% of labour force)	İşsizlik oranı
X13	Employment in services (% of total employment)	Hizmet sektöründeki istihdam oranı
X14	Internet users, total (% of population)	İnternet kullanım oranı
X15	Mobile phone subscriptions (per 100 people)	Cep telefonu abonelik oranı (100 kişi başına)

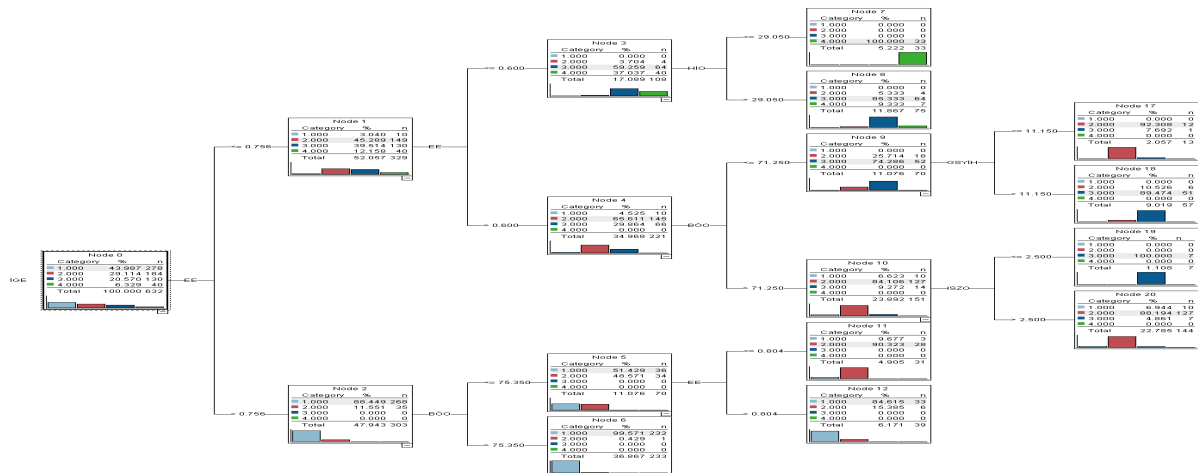
Bulgular

Araştırmanın bu kapsamında gerekli düzenlemelerin yapıldığı modelin kavramsal yapısına ve bulgulara ilişkin bilgiler verilmiştir. Araştırma modelinin kavramsal yapısı içerisinde yer alan veri ön işleminin birinci aşamasında eksik verilere sahip olan ülkeler çalışmanın kapsamına dahil edilmemiştir. İkinci aşamasında değişkenler için normallik varsayımı sağlanmadığından, spearman korelasyon analizi gerçekleştirilerek %80'in üzerinde yüksek korelasyon ilişkisine sahip olan değişkenler karar ağacı analizlerinden çıkarılmıştır (Alpar, 2013). Analizlerden çıkarılan bu değişkenler: X4 değişkeni “beklenen eğitim süresi (yıl)”, X6 değişkeni “gelir endeksi”, X9 değişkeni “cinsiyet eşitsizliği endeksi”, X11 değişkeni “tarım sektöründeki istihdam oranı”, X14 değişkeni “internet kullanım oranı”dır. Böylece veri ön işleme tamamlandıktan sonra 79 ülkenin toplam 632 veri kaydı karar ağaçları algoritmalarında kullanılmak üzere hazır hale getirilmiştir. Verilerin %75'i ile modeller eğitilmiş, geriye kalan %25'lik kısmı ile test işlemi gerçekleştirilmiştir. Bu aşamadan sonra sınıflandırma algoritmalarının belirleyicilik “specificity” ve duyarlılık “sensitivity” istatistikleri karşılaştırılmak üzere hesaplanmıştır. Araştırmanın kapsamına ilişkin kavramsal yapı Şekil 1'de gösterilmiştir.



Şekil 1. Karar Ağaçları için Araştırmanın Kavramsal Yapısı

Gini Algoritması ile Karar Ağacı Modeli: Gini algoritması ile karar ağacı modelinde 9 kural türetilmiştir. İGE'ye ait karar ağacı modelinin sınıflandırılmasında en kuvvetli etkiye sahip olan değişkenin %61,9 ile eğitim endeksi, ikinci sırada %25,9 ile hizmet sektöründeki istihdam oranı değişkeni olduğu tespit edilmiştir. Şekil 2'de Gini algoritmasına ait karar ağacı modeli verilmiştir.



Şekil 2. Gini Algoritmasına ait Karar Ağacı Modeli

Tablo 4. Gini Algoritmasının Karar Ağacı Modelinden Türetilen Kurallar

Kural No	Eğer Kuralı	%	Ulke Sayısı	Kategorik İGE Düzeyi
1	$X1 \leq 0.60$ & $X13 \leq 29.05$	100%	33	Düşük İGE
2	$X1 \leq 0.60$ & $X13 > 29.05$	85,33%	64	Orta İGE
3	$X1 > 0.756$ & $X2 > 73.35$	99,57%	222	Çok Yüksek İGE
4	$0.756 < X1 \leq 0.804$ & $X2 \leq 73.35$	90,32%	28	Yüksek İGE
5	$X1 > 0.804$ & $X2 \leq 73.35$	84,62%	33	Çok Yüksek İGE
6	$0.60 < X1 \leq 0.756$ & $X2 \leq 71.25$ & $X5 \leq 1150$	92,31%	12	Yüksek İGE
7	$0.60 < X1 \leq 0.756$ & $X2 \leq 71.25$ & $X5 > 1150$	89,47%	51	Orta İGE
8	$0.60 < X1 \leq 0.756$ & $X2 > 71.25$ & $X11 \leq 2.5$	100,00%	7	Orta İGE
9	$0.60 < X1 \leq 0.756$ & $X2 > 71.25$ & $X11 > 2.5$	88,14%	127	Yüksek İGE

Tablo 4'te Gini algoritmasının karar ağacı modelinden türetilen 9 kural verilmiştir. Belirlenen kurallar içerisinde çok yüksek İGE düzeyi için 3, yüksek İGE düzeyi için 9, orta İGE düzeyi için 2, düşük İGE düzeyi için 1 no'lu kural en önemli kurallardır.

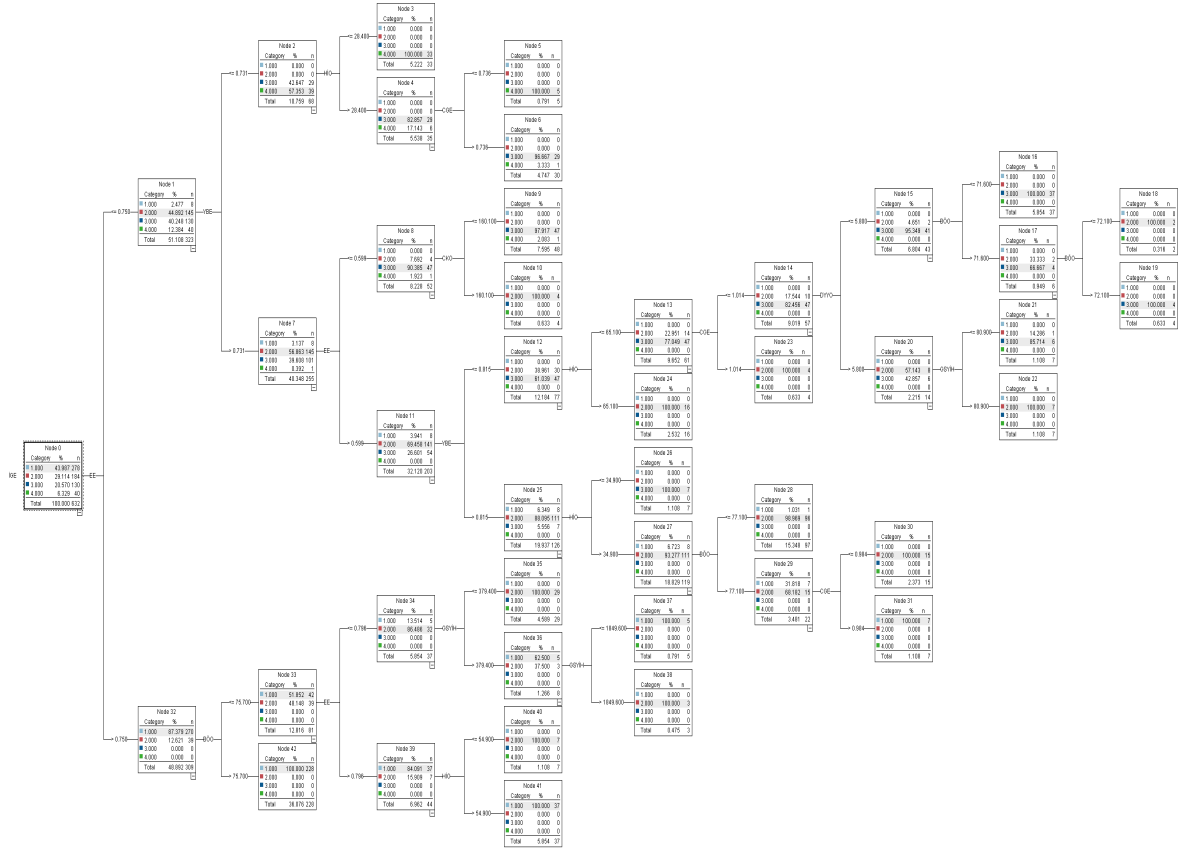
Kural 1: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.60'a eşit ya da küçük ve X13 değişkeni "hizmet sektöründeki istihdam oranı" 29.05'e eşit ya da küçük ise o ülkenin İGE %100 oran ile düşük düzeydedir.

Kural 2: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.60'a eşit ya da küçük ve X13 değişkeni "hizmet sektöründeki istihdam oranı" 29.05'ten büyük ise o ülkenin İGE %85,33 oran ile orta düzeydedir.

Kural 3: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.756'dan büyük ve X2 değişkeni "doğumdan sonra beklenen yaşam süresi" 73,35'ten büyük ise o ülkenin İGE %99,57 oran ile çok yüksek düzeydedir.

Kural 9: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.60'tan büyük ve 0.756'ya eşit ya da küçük, X2 değişkeni "doğumdan sonra beklenen yaşam süresi" 71,25'ten büyük ve X12 değişkeni "işsizlik oranı değişkeni" 2,5'ten büyük ise o ülkenin İGE %88,14 oran ile yüksek düzeydedir.

C5.0 Algoritması ile Karar Ağacı Modeli: Entropi tabanlı C5.0 algoritmasının karar ağacı Şekil 3'te gösterilmiştir. Karar ağacı modelinden 21 kural oluşturulmuştur. İGE düzeyinin tespitinde en kuvvetli etkiye sahip olan değişkenin %43,1 ile eğitim endeksi, %29,5 ile hizmet sektöründeki istihdam oranı değişkeni olduğu belirlenmiştir.



Şekil 3. C5.0 Algoritmasına ait Karar Ağacı Modeli

C5.0 algoritmasının karar ağacı modelinden elde edilen 21 kural Tablo 5'te gösterilmiştir. Çok yüksek İGE düzeyi için 8, yüksek İGE düzeyi için 14, orta İGE düzeyi için 4, düşük İGE düzeyi için 1 no'lu kural tespit edilenler içerisinde en önemli kurallardır.

Kural 1: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.750'den küçük, X3 değişkeni "yaşam beklentisi endeksi" 0.731'e eşit ya da küçük ve X13 değişkeni "hizmet sektöründeki istihdam oranı" 28.4'e eşit ya da küçük ise o ülkenin İGE %100 oran ile düşük düzeydedir.

Kural 4: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.599'a eşit ya da küçük, X3 değişkeni "yaşam beklentisi endeksi" 0.731'den büyük ve X15 değişkeni "cep telefonu abonelik oranı" 160,1'e eşit ya da küçük ise o ülkenin İGE %97.9 oran ile orta düzeydedir.

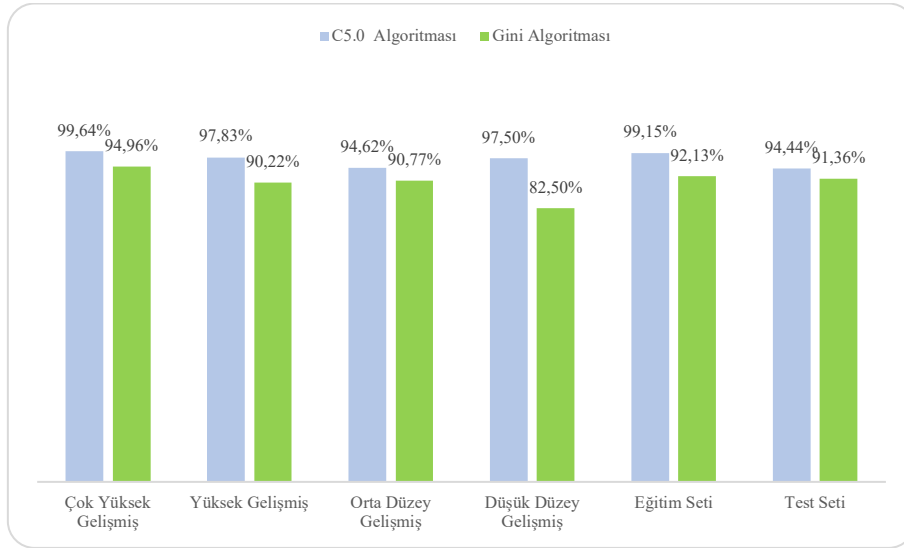
Kural 8: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.796'dan büyük, X2 değişkeni "doğumdan sonra beklenen yaşam süresi" 75.7'den büyük ve X13 değişkeni "hizmet sektöründeki istihdam oranı" 54.9'dan büyük ise o ülkenin İGE %100 oran ile çok yüksek düzeydedir.

Kural 14: Eğer İGE'nin X1 değişkeni "eğitim endeksi" 0.599'dan büyük ve 0.750'ye eşit ya da küçük, X3 değişkeni "yaşam beklentisi endeksi" 0.815'ten büyük, X13 değişkeni "hizmet sektöründeki istihdam oranı" 34.9'dan büyük ve X2 değişkeni "doğumdan sonra beklenen yaşam süresi" 77.1'e eşit ya da küçük ise o ülkenin İGE %98,97 oran ile yüksek düzeydedir.

Tablo 5. Karar Ağacı Modeli için C5.0 Algoritmasından Elde Edilen Kurallar

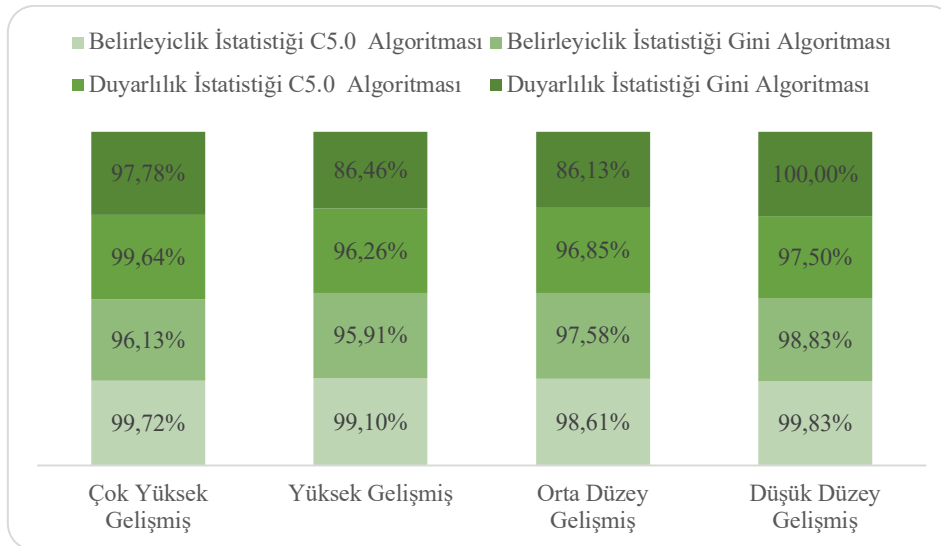
Kural No	Eğer Kuralı	%	Ülke Sayısı	Kategorik İGE Düzeyi
1	$X1 \leq 0.750 \& X3 \leq 0.731 \& X13 \leq 28.4$	100%	33	Düşük İGE
2	$X1 \leq 0.750 \& X3 \leq 0.731 \& X13 > 28.4 \& X10 \leq 0.736$	100%	5	Düşük İGE
3	$X1 \leq 0.750 \& X3 \leq 0.731 \& X13 > 28.4 \& X10 > 0.736$	96,60%	29	Orta İGE
4	$X1 \leq 0.599 \& X3 > 0.731 \& X15 \leq 160,1$	97,90%	47	Orta İGE
5	$X1 \leq 0.599 \& X3 > 0.731 \& X15 > 160,1$	100%	4	Yüksek İGE
6	$0.750 < X1 \leq 0.796 \& X2 \leq 75.7 \& X5 < 379.4$	100%	29	Yüksek İGE
7	$X1 > 0.796 \& X2 > 75.7 \& X13 \leq 54.9$	100%	7	Yüksek İGE
8	$X1 > 0.796 \& X2 > 75.7 \& X13 > 54.9$	100%	37	Çok Yüksek İGE
9	$0.599 < X1 \leq 0.750 \& 0.731 < X3 \leq 0.815 \& X13 > 85.1$	100%	16	Yüksek İGE
10	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X13 < 34.9$	100%	7	Orta İGE
11	$0.750 < X1 \leq 0.796 \& X2 \leq 75.7 \& 379.4 < X5 \leq 1849$	100%	5	Çok Yüksek İGE
12	$0.750 < X1 \leq 0.796 \& X2 \leq 75.7 \& X5 > 1849$	100%	3	Yüksek İGE
13	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 > 1.014$	100%	4	Yüksek İGE
14	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X13 > 34.9 \& X2 \leq 77.1$	98,97%	96	Yüksek İGE
15	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X13 > 34.9 \& X2 > 77.1 \& X10 < 0.984$	100%	15	Yüksek İGE
16	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X13 > 34.9 \& X2 > 77.1 \& X10 > 0.984$	100%	7	Çok Yüksek İGE
17	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 \leq 1.014 \& X8 \leq 5.8 \& X2 \leq 71.8$	100%	37	Orta İGE
18	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 \leq 1.014 \& X8 > 5.8 \& X5 \leq 80.9$	85,70%	6	Orta İGE
19	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 \leq 1.014 \& X8 > 5.8 \& X5 > 80.9$	100%	7	Yüksek İGE
20	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 \leq 1.014 \& X8 \leq 5.8 \& 71.8 < X2 \leq 72.1$	100%	2	Yüksek İGE
21	$0.599 < X1 \leq 0.750 \& X3 > 0.815 \& X10 \leq 1.014 \& X8 \leq 5.8 \& X2 > 72.1$	100%	4	Orta İGE

Algoritmaların Performanslarının Karşılaştırılması: Bu aşamada Gini ve C5.0 algoritmalarıyla oluşturulan karar ağacı modellerinin performanslarını karşılaştırmak için dört farklı kriter kullanılmıştır. Bu kriterler sırasıyla sınıflandırma, eğitim ve test, belirleyicilik ve duyarlılık istatistik performans kriterleridir. Şekil 4'te algoritmaların İGE, eğitim ve test performanslarının sonuçları gösterilmiştir.



Şekil 4. C5.0 ve Gini Algoritmalarının İGE, Eğitim ve Test Performanslarının Sonuçları

Şekil 4'e göre C5.0 ve Gini algoritmalarının İGE sınıflandırma performans sonuçları sırasıyla çok yüksek gelişmiş ülkeler için %99,64 ile %94,96, yüksek gelişmiş ülkeler için %97,83 ile %90,22, orta düzey ülkeler için %94,62 ile %90,77 ve düşük düzey gelişmiş ülkeler için %97,50 ile %82,50 ve toplam doğru sınıflandırma başarı düzeyleri %97,94 ile %91,93'tür. Eğitim setinin sınıflandırma performansları C5.0 için %99,15 ve Gini için %92,13; test setinin sınıflandırma performansları C5.0 için %92,13 ve Gini için %91,36'dır. Algoritmaların ayrıca duyarlılık ve belirleyicilik istatistikleri de hesaplanmıştır. Duyarlılık istatistiği (gerçek pozitif oran), İGE'de sınıflandırılan ülkelerlerden yüzde kaçının gerçekten doğru İGE olduğunu; Belirleyicilik istatistiği (gerçek negatif oran), İGE sınıflandırılan ülkelerden yüzde kaçının gerçekten yanlış İGE olduğunu belirten orandır (Öğüt, vd., 2009; Han ve Kamber, 2006). Şekil 5'te seçilen algoritmalara ilişkin belirleyicilik ve duyarlılık istatistikleri gösterilmiştir.



Şekil 5. C5.0 ve Gini Algoritmalarının Duyarlılık ve Belirleyicilik İstatistiklerinin Performans Sonuçları

Şekil 5'te görüldüğü üzere C5.0 ve Gini algoritmalarının duyarlılık istatistik sonuçları sırasıyla çok yüksek gelişmiş ülkeler için %99,64 ile %97,78, yüksek gelişmiş ülkeler için %96,26 ile %86,46, orta düzey gelişmiş ülkeler için %96,85 ile %86,13 ve düşük düzey ülkeler için %97,50 ile %100; belirleyicilik istatistik sonuçları sırasıyla çok yüksek gelişmiş ülkeler için %99,72 ile %96,13, yüksek gelişmiş ülkeler için %99,10 ile %95,91, orta düzey gelişmiş ülkeler için %98,61 ile %97,58 ve düşük düzey ülkeler için %99,83 ile %98,83'tür.

Sonuç ve Öneriler

Bu çalışmada, BM tarafından her yıl yayımlanan insani gelişme raporunun 2010-2017 dönemine ait 79 ülkenin insani gelişme endeksine etki eden faktörlerin, veri madenciliğinin karar ağacı teknik tekniklerinden C5.0 ve Gini algoritmalarıyla modellenmesi amaçlanmıştır. Ülkelerin İGE düzeyleri çok yüksek, yüksek, orta düzey ve düşük düzey olmak üzere kategorileştirilmiştir. İGE'ye etki eden faktörlerin belirlenmesinde eğitim, sağlık, ekonomi, istihdam ve iletişim faktörlerinden toplam 15 bağımsız değişken kullanılmıştır.

Bağımsız değişkenler için spearman korelasyon analizi gerçekleştirilerek 5 değişken analizden çıkarılmıştır. Karar ağaçlarının oluşturulmasında C5.0 algoritması ile %97,94 ve Gini algoritmasıyla %91,93'lük toplam doğru sınıflandırma başarı düzeyi gerçekleştirilmiştir. C5.0 ve Gini algoritmalarının karar ağacı modellerinde sırasıyla 21 ve 9 kural türetilmiştir. Her iki algoritma ile Türkiye'nin çalışma dönemi boyunca yüksek düzey olan İGE'sini doğru olarak sınıflandırmıştır. 2010-2017 dönemine ait C5.0 algoritması ile 13 ve Gini algoritması ile 51 ülke yanlış olarak sınıflandırılmıştır. C5.0 algoritması ile yanlış sınıflandırılan ülkeler: Bangladeş, Belize, El Salvador "üç kez", Malezya, Maldivler, Moldova, Moğolistan, Nikaragua "iki kez", Pakistan ve Rusya'dır. Gini algoritması ile yanlış sınıflandırılan ülkeler: Azerbaycan "üç kez", Bangladeş, Belize, Kenya, Malezya, Maldivler "dört kez", Moğolistan "beş kez", Pakistan "beş kez", Paraguay "dört kez", Portekiz "dört kez", Romanya "beş kez", Rusya "1", Ukrayna "beş kez", Uruguay "dört kez", Vietnam "yedi kez" dir.

İGE'ye etki eden faktörlerin tespitinde en fazla etki eden faktörlerin sırasıyla: C5.0 algoritmasında %43,1 ile eğitim, %29,5 ile istihdam ve %12, 1 ile iletişim faktörü; Gini algoritmasında ise %61,9 ile eğitim, %25,9 ile istihdam, %6,7 ile sağlık faktörü olduğu tespit edilmiştir. Algoritmaların karşılaştırılmasında sınıflandırma, eğitim ve test, belirleyicilik ve duyarlılık istatistik performans kriterleri olmak üzere dört farklı kriter kullanılmıştır. İstatistiki performans açısından karşılaştırılmasında C5.0 algoritmasının, Gini algoritmasına göre daha güçlü performans gösterdiği gözlenmiştir. Tablo 6'da İGE'nin sınıflandırılmasına ilişkin benzer çalışmalar, önerdiğimiz model ile karşılaştırılmıştır.

Tablo 6. Benzer Çalışmalarla Sonuçların Karşılaştırılması

Yazar/Yıl	Veri Seti	Girdi Verileri	Analiz Yöntemleri	Verinin Alındığı Tarih Aralığı	En İyi Sınıflandırma Doğruluğu
Gündüz & Soyer, (2019)	138 Ülke	12 Farklı Ekonomik Özgürlük Endeksi	Binary Lojistik Regresyon Analizi	2018	%85,5
Rençber & Mete, (2018)	BM, 185 Ülke	27 değişken	YSA, ANFİS	2015	%87,5-91,36
Rençber, (2018)	BM, 185 Ülke	118 değişken	ÇLR, YSA, ANFİS	2015	%81,6- %85,4-%91,36
Burmaoğlu & Oktay, (2015)	BM, 120 Ülke	16 değişken	Diskriminant Analizi (DA)	2007-2008	%92,5
Yakut, Gündüz, & Demirci, (2015)	BM, 81 Ülke	14 değişken	ÇLR, YSA	2010-2012	%88,1-%97,1
Burmaoğlu, Oktay, & Özen, (2009)	BM. 120 Ülke	16 değişken	DA, LR	2007-2008	%92,5- %100
Önerdiğimiz Model	BM, 79	15 değişken	C5.0 ve Gini Algoritması	2010-2017	%97,94-%91,93

Tablo 6’da görüldüğü üzere önerdiğimiz modelin sınıflandırma sonuçları İGE’ye ilişkin diğer çalışmaların sınıflandırma sonuçlarına göre daha iyi performans göstermiştir. Sonuç itibari ile bu araştırma insan hayatının zenginliği açısından ülkeleri İGE’ye göre karşılaştırma yapma imkânı sunmaktadır. Araştırma sonucunda algoritmaların karar ağacı kurallarına göre bir ülkenin İGE’nin hangi sınıfta yer aldığı tahmin edilebilmektedir. Gelecekteki yapılacak çalışmalarda daha geniş veri seti ve diğer analiz yöntemleri ile bu çalışmanın sonuçları karşılaştırılabilir.

Kaynakça

- Adak, M. F., & Yurtay, N. (2013). Gini algoritmasını kullanarak karar ağacı oluşturmayı sağlayan bir yazılımın geliştirilmesi. *International Journal of Informatics Technologies*, 6(3), 1-6.
- Alpar, R. (2013). *Çok Değişkenli İstatistik Yöntemler*, Ankara: Detay.
- Alpaydın, E. (2013). *Yapay öğrenme*, İstanbul: Boğaziçi Üniversitesi.
- Berber, M. (2017). *İktisadi büyüme ve kalkınma*, Trabzon: Celepler.
- Burmaoğlu, S., & Oktay, E. (2015). A statistical classification study of countries'human development level by discriminant analysis, *Ataturk University Journal of Economics & Administrative Sciences*, 29(4), 579-598.
- Burmaoğlu, S., Oktay, E., & Üstün, Ö. (2009). Birleşmiş milletler kalkınma programı beşeri kalkınma endeksi verilerini kullanarak diskriminant analizi ve lojistik regresyon analizinin sınıflandırma performanslarının karşılaştırılması, *Savunma Bilimleri Dergisi*, 8(2), 23-49.
- Çağlar, A., & Keten, N. D. (2018). İllerin insani gelişme endeksinin veri zarflama analizi ile ölçülmesi, *Ege Akademik Bakış Dergisi*, 18(4), 565-578.

- Çemrek, F., & Naci Bayraç, H. (2013). Sürdürülebilir kalkınma skorunun hesaplanması. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 14(2), 131-152.
- Çoban, N., & Yayar, R. (2018). Demokrasinin göstergelerinin insani gelişmişlik üzerine etkisi: ab ülkeleri üzerine bir panel veri analizi, *Journal of Politics Economy and Management*, 1(2), 55-61.
- Doğan, E., & Tatlı, H. (2014). İnsani gelişme ve insani yoksulluk bağlamında Türkiye'nin dünyadaki yeri, *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 28(1), 99-124.
- Erbay, E. R., & Özden, M. (2013). İktisadi kalkınma kuramlarına eleştirel yaklaşım. *Sosyal Bilimler Metinleri*, 3(1), 1-27.
- Erdem, E., & Çelik, B. (2019). İnsani gelişme ve ekonomik büyüme ilişkisi: Bazı Afrika ülkeleri üzerine bir uygulama, *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(17), 13-36.
- Fırat, E., & Aydın, A. (2015). İnsani kalkınma endeksine göre Türkiye'nin eğitim endeks göstergelerinin OECD ülkeleri ile karşılaştırılması, *Sosyal Ekonomik Araştırmalar Dergisi*, 15(29), 62-87.
- Günel, F., Çağlar, A., Kangallı Uyar, S., Karadeniz, O., & Yeşilyurt, M. (2017). Türkiye'de illere göre insani gelişme endeksi, *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 27, 208-216.
- Gündüz, M., & Soyer, M. (2019). Küresel rekabetçilik ile ekonomik özgürlük arasındaki ilişkinin Binary Lojistik Regresyon Analizi ile incelenmesi. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 8(3), 1906-1930.
- Günsoy, G. (2012). İnsani gelişme kavramı ve sağlıklı yaşam hakkı, *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 1(2), 35-52.
- Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques*, San Francisco: The Morgan Kaufmann.
- Kaya, A. (2017). Finansal piyasalardaki gelişmelerin insani gelişmişlik üzerine etkisi: Türkiye örneği, *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 20, 169-180.
- Eren, M., & Kaynak, (2017). AB'ye tam üyelik sürecinde aday ülkelerin insani gelişme, eğitim ve gelir indekslerinin çok amaçlı doğrusal programlama ile çoklu periyotlu olarak değerlendirilmesi, *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 18(2), 41-50.
- Kocatepe, C.A., & Dilber, İ. (2014). Türkiye'nin insani gelişme endeksi değerinin 2011-2012 karşılaştırılması. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 24(2), 123-131.
- Korkmaz, C., & Şahin, M. (2013). 2009 PİSA başarılarına göre ülkelerin genel ve insani gelişmişlik düzeyleri arasındaki ilişki, *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(22), 225-247.
- Meydan, M., & Sarı, V. İ. (2018). İnsani gelişme endeksi ve alt endekslerinin Türkiye'deki iller için ölçülmesi, *İdealkent Dergisi*, 9(24), 387-429.
- Mıhçı, H., & Mıhçı, S. (2003). Türkiye'nin yakın dönemdeki insani gelişme eğilimleri, *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21(2), 21-47.

- Murat, D., & Gürsakal, S. (2015). Mutluluk ve insani gelişmişlik arasındaki ilişkinin belirlenmesi: çok değişkenli istatistiksel yaklaşım, *Alphanumeric Journal*, 3(1), 67-79.
- Nartgün, Ş. S., Kösterelioğlu, M., & Sipahioğlu, M. (2013). İnsani gelişim indeksi göstergeleri açısından AB üyesi ve AB üyeliğine aday ülkelerin karşılaştırılması, *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 3(1), 80-89.
- Öğüt, H., Doğanay, M., & Aktaş, R. (2009). Detecting stock-price manipulation in an emerging market: The case of Turkey, *Expert Systems With Applications*, 36(9), 11944-11949.
- Özkan, Y. (2008). *Veri madenciliği yöntemleri*, İstanbul: Papatya Yayıncılık Eğitim.
- Öztürk, H. M. (2016). Türkiye'nin insani gelişme endeksi, ekonomik gelişmişlik, beslenme ilişkisi ve farklı ülkelerle karşılaştırılması, *Journal of Human Sciences*, 13(2), 3402-3412.
- Paksoy, S. (2015). Ülke göstergelerinin VİKOR yöntemi ile değerlendirilmesi, *Ekonomik ve Sosyal Araştırmalar Dergisi*, 11(2), 153-169.
- Rençber, Ö. F. (2017). *Bulanık ve Yapay Sinir Ağları İle Çoklu Lojistik Regrasyon Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması: Ülkelerin Gelişmişlik Düzeylerinin Sınıflandırılması Üzerine Bir Uygulama* (Yayınlanmamış Doktora Tezi), Aksaray Üniversitesi Sosyal Bilimler Enstitüsü, Aksaray.
- Rençber, Ö. F. (2018). *Sınıflandırma problemlerinde çoklu lojistik regresyon, yapay sinir ağı ve ANFİS yöntemlerinin karşılaştırılması: İnsani gelişmişlik endeksi üzerine uygulama*, Ankara: Gazi.
- Rençber, Ö. F., & Sinan, M. (2018). Reclassification of countries according to human development index: An application with ann and anfis methods, *Business & Management Studies: An International Journal*, 6(3), 228-252.
- Rokach, L., & Maimon, O. Z. (2008). *Data mining with decision trees: theory and applications*, Singapore: World scientific.
- Sakarya, A., & İbişoğlu, Ç. (2015). Türkiye'de illerin sosyo-ekonomik gelişmişlik endeksinin coğrafi ağırlıklı regresyon modeli ile analizi. *Marmara Coğrafya Dergisi*, 32, 211-238.
- Silahtaroglu, G. (2013). *Veri madenciliği*. İstanbul: Papatya Yayıncılık Eğitim.
- Şahin, G., & Gökdemir, L. (2016). İnsani gelişme endeksi bileşenlerinin Türkiye ölçeğinde ARDL sınır testi ile sınanması, *Gazi İktisat ve İşletme Dergisi*, 2(1), 1-24.
- Taban, S. (2016). *İktisadi büyüme kavram ve modeller*, Bursa: Ekin.
- Tıraş, H. (2019). Türkiye için insani gelişmişlik göstergeleri, *Bilgi Ekonomisi ve Yönetimi Dergisi*, 14(1), 15-31.
- Tireli, M., Coşkun, S., & Kunduracı, N. F. (2013). İİT ülkeleri ve Türkiye'ye ait sosyal göstergeler ile insani gelişim göstergeleri arasındaki ilişkinin karşılaştırmalı analizi, *Sosyal Politika Çalışmaları Dergisi*, 7(30), 61-87.

- Tunç, O., & Ertuna, Ö. (2015). Human development index of Turkey simulation and comparison of selected countries, *Journal of Management Marketing and Logistics*, 2(2), 132-157.
- UNDP (2018). *Human development report 2018*, <http://hdr.undp.org/en/2018-update>
- Ülgener, S. F. (1991). *Milli gelir, istihdam ve iktisadi büyüme*. İstanbul: Der.
- Çiğdem, Ü. (2008). İnsani gelişmişlik endeksine göre Türkiye'nin bölgesel farklılıkları, *Coğrafi Bilimler Dergisi*, 6(2), 89-113.
- Üzümçü, A. (2012). İktisadi büyüme (Teori, model ve Türkiye üzerine gözlemler), İstanbul: Beta.
- Yakut, E., Gündüz, M., & Demirci, A. (2015). İnsani kalkınmışlık düzeyinin sınıflandırma başarılarının karşılaştırılmasında sıralı lojistik regresyon analizi ve yapay sinir ağları yöntemlerinin kullanılması, *İşletme Araştırmaları Dergisi*, 7(4), 172-199.
- Yalçın, A. Z., & Çakmak, F. (2016). Türkiye'de kamu sağlık harcamalarının insani gelişim üzerindeki etkisi, *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 30(4), 705-724.
- Zanbak, M., & Özeş Ö. R. (2019). İnsani gelişme endeksi bağlamında Avrupa Birliği'ne üye ve aday ülkelerin karşılaştırmalı analizi, *Yönetim ve Ekonomi Araştırmaları Dergisi*, 17(2), 175-192.