



T.C.
OSMANIYE KORKUT ATA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANA BİLİM DALI

İNSANİ GELİŞİMİŞLİK ENDEKSİNİN
VERİ MADENCİLİĞİ TEKNİKLERİNDEN OLAN
C5.0 VE GINI ALGORİTMALARI KULLANARAK
MODELENMESİ

YÜKSEK LİSANS TEZİ

ALİ KORKMAZ

OSMANIYE / 2019

T.C.
OSMANIYE KORKUT ATA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANA BİLİM DALI

**İNSANİ GELİŞMİŞLİK ENDEKSİNİN VERİ MADENCİLİĞİ
TEKNİKLERİNDEN OLAN C5.0 VE GINI ALGORİTMALARI
KULLANARAK MODELLENMESİ**

YÜKSEK LİSANS TEZİ

ALİ KORKMAZ

Danışman: Dr. Öğretim Üyesi Emre YAKUT


Jüri Üyesi: Prof. Dr. Bülent ÖZ

Jüri Üyesi: Dr. Öğretim Üyesi Ömer Faruk RENÇBER

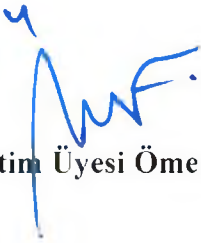
OSMANIYE / 2019

Osmaniye Korkut Ata Üniversitesi Sosyal Bilimler Enstitüsü Müdürlüğüne;

Bu çalışma, jürimiz tarafından İşletme Ana Bilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.


Başkan: Prof. Dr. Bülent ÖZ


Üye: Dr. Öğretim Üyesi Emre YAKUT


Üye: Dr. Öğretim Üyesi Ömer Faruk RENÇBER

ONAY

Yukarıdaki imzaların, adı geçen öğretim elemanlarına ait olduklarını onaylarım.

...../...../2019

Doç. Dr. Müjdat AVCI
Enstitü Müdürü

NOT: Bu tezde kullanılan ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunu'ndaki hükümlere tabidir.

ETİK BEYANI

Osmaniye Korkut Ata Üniversitesi Sosyal Bilimler Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde ve ortaya çıkan sonuçlarda herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim. 27/09/2019


Ali KORKMAZ

ÖZET

İNSANI GELİŞİMİŞLİK ENDEKSİNİN VERİ MADENCİLİĞİ TEKNİKLERİNDEN OLAN C5.0 VE GINI ALGORİTMALARI KULLANARAK MODELLENMESİ

ALİ KORKMAZ

Yüksek Lisans, İşletme Ana Bilim Dalı

Danışman: Dr. Öğretim Üyesi Emre YAKUT

Ağustos 2019, 72 sayfa

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 ülkeler arası karşılaştırma yapmak için başvuru 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 düzeyde 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 ve gini algoritması.

ABSTRACT**MODELING OF HUMAN DEVELOPMENT INDEX (HDI) USING C5.0, A
DATA MINING TECHNIQUE, AND GINI ALGORITHMS****ALİ KORKMAZ****Osmaniye Korkut Ata University, Institute of Social Sciences****Master's Thesis, Department of Business Administration****Advisor: Assistant Professor Emre YAKUT****August 2019, 72 pages**

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 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 on 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.

ÖN SÖZ

Tez çalışmasının konusunun seçiminde, planlanmasında ve bu araştırmanın her safhasında engin tecrübeleriyle desteğini esirgemeyerek çalışmayı bilimsel temeller ışığında şekillendiren danışmanım Sayın Dr. Öğretim Üyesi Emre YAKUT'a, kendisinden çok şey öğrendiğim Prof. Dr. Mustafa AKAR'a, baskı öncesinde tezi okuyarak, hata ve eksikliklerin düzeltilmesinde önemli katkısı bulunan Soner ÇERÇİ'ye, maddi ve manevi desteğini esirgemeyip daima yardımcı olan aileme teşekkür ederim.

Göstermiş oldukları sonsuz hoşgörü ve sabırlarından dolayı sevgili eşim Aysun'a, kızlarım Özge Yağmur, Gökçe Bahar ve Gülce Pınar ile oğlum Rüstem Ertuğrul'a gönülden teşekkür ediyor, şükranlarımı sunuyorum.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
ÖN SÖZ.....	vi
KISALTMALAR	xi
TABLolar LİSTESİ.....	xii
ŞEKİLLER LİSTESİ.....	xiii
GİRİŞ	1

BÖLÜM I

VERİ MADENCİLİĞİ

1.1. Veri Madenciliğine Giriş	3
1.2. Veri Madenciliğinin Diğer Disiplinlerle Olan İlişkisi	4
1.3. Veri Madenciliği Tekniklerinin Uygulandığı Alanlar	5
1.4. Veri Madenciliği Uygulamasına Yönelik Süreç	6
1.4.1. Problemin Belirlenmesi	7
1.4.2. Verilerin Hazır Duruma Getirilmesi	7
1.4.2.1. Veri Entegrasyonu.....	8
1.4.2.2. Veri Temizleme.....	8
1.4.2.3. Veri Dönüştürme.....	8
1.4.2.3.1. Veri Düzeltme.....	9
1.4.2.3.2. Veri Birleştirme	9
1.4.2.3.3. Veri Genelleştirme.....	9
1.4.2.3.4. Veri Normalleştirme	9
1.4.2.3.4.1. Ondalık Ölçekleme	10

1.4.2.3.4.2. Z-skor Normalleştirme.....	10
1.4.2.3.4.3. Min-Max Normalleştirme.....	11
1.4.2.3.4.4. Logaritmik Normalleştirme.....	11
1.4.2.3.4.5. Değişken Oluşturma.....	12
1.4.3. Veri İndirgeme.....	12
1.4.4. Verinin Seçimi.....	13
1.4.5. Modelin Oluşturulması.....	14
1.4.6. Modelin Değerlendirilmesi.....	14
1.4.7. Modelin Uygulamaya Konulması.....	14
1.4.8. Modelin İzlenmesi.....	14
1.5. Veri Madenciliğine Ait Bazı Yöntemleri.....	15
1.5.1. Sınıflandırma Yöntemi.....	16
1.5.1.1. Karar Ağaçları.....	16
1.5.1.1.1. Gini Algoritması.....	17
1.5.1.1.2. C5.0 Algoritması.....	20
1.5.2. Kümeleme Yöntemi.....	21
1.5.3. Birliktelik Kuralları Yöntemi.....	22

BÖLÜM II

GELİŞMİŞLİK TANIMI VE İNSANİ GELİŞME ENDEKSİ

2.1. Gelişmişliği Esas Alan Tanımlar.....	23
2.1.1. İktisadi Büyüme Tanımı.....	23
2.1.2. Gelişmişlik ve Kalkınmışlığa İlişkin Tanımlar.....	24
2.1.3. Gelişmişliğin Ölçümüne Yönelik Tarihsel Boyutu.....	25
2.1.4. Kalkınmışlığın Ölçümüne Yönelik Temel Göstergeler.....	26
2.1.4.1. Doğrudan Kalkınmışlığa Yönelik Göstergeler.....	26

2.1.4.2. Dolaylı Olarak Kalkınmışlığa Yönelik Göstergeler.....	27
2.1.5. Gelişmişliğin Ölçümüne Yönelik Temel Göstergeler	27
2.2. İnsani Gelişme Endeksi Kavramı	27
2.3. İnsani Gelişme Endeksinin Hesaplanmasında Kullanılan Göstergeler.....	28
2.3.1. Sağlık Faktörü: Doğumda Yaşam Beklentisi Göstergesi	28
2.3.2. Eğitim: Beklenen Okullaşma ve Ortalama Okullaşma Oranı	29
2.3.3. Gelir: Satıl Alma Gücü Paritesine İlişkin Kişi Başına Gelir Hesaplaması ...	30
2.4. Toplumsal Cinsiyet Eşitsizliği Endeksi Tanımı.....	33

BÖLÜM III

İNSANİ GELİŞME ENDEKSİ İLE İLGİLİ LİTERATÜR İNCELEMESİ

3.1. İGE Konulu Çalışmalar	34
----------------------------------	----

BÖLÜM IV

ARAŞTIRMA BULGULARI VE DEĞERLENDİRME

4.1. Metodoloji.....	42
4.2. Çalışmanın Amacı.....	42
4.3. Çalışmanın Önemi	42
4.4. Çalışmanın Kapsamı	43
4.5. Veri Setine Ait Değişkenlerin Seçimi.....	47
4.6. Verilerin Analiz Yöntemi	48
4.7. Spearman Korelasyon Analizine İlişkin Sonuçlar	49
4.8. C5.0 Algoritması ile Karar Ağacı Modellemesi	50
4.9. Gini Algoritması ile Karar Ağacı Modellemesi.....	53
4.10. C5.0 ve Gini Algoritmalarının Performanslarının Karşılaştırılması.....	56

BÖLÜM V

SONUÇ VE ÖNERİLER.....	60
KAYNAKÇA.....	62
EKLER.....	71
ÖZ GEÇMİŞ.....	72

KISALTMALAR

AB	: Avrupa Birliđi
AID	: Automatic Interaction Detection
ANFIS	: Bulanık Yapay Sinir Ađı
BOY	: Beklenen Okullařma Yılı
CART	: Classification and Regression Trees
ÇLR	: Çoklu Lojistik Regresyon
DA	: Diskriminant Analizi
DYB	: Doğumda Yařam Beklentisi
GSMH	: Gayri Safi Milli Hasıla
GSYİH	: Gayri Safi Yurtiçi Hasıla
HDI	: Human Development Index
HPI	: Mutlu Gezegen Endeksi
İGE	: İnsani Geliřme Endeksi
İGR	: İnsani Geliřme Raporu
İİT	: İslam İşbirliđi Teřkilatı
LR	: Lojistik Regresyon
NEF	: Yeni Ekonomi Kurumu
OECD	: Ekonomik İşbirliđi ve Kalkınma Örgütü
OYY	: Ortalama Okullařma Yılı
PISA	: Uluslararası Öğrenci Deđerlendirme Programı
TCEE	: Toplumsal Cinsiyet Eřsitsizliđi Endeksi
TÜFE	: Tüketici Fiyat Endeksi
TÜİK	: Türkiye İstatistik Kurumu
UNDP	: Birleřmiş Milletler Kalkınma Programı
VM	: Veri Madenciliđi
VZA	: Veri Zarflama Analizi
YSA	: Yapay Sinir Ađları

TABLULAR LİSTESİ

	Sayfa
Tablo 1. İnsani Gelişme Endeks Değerine Göre Yapılan Sınıflandırma	31
Tablo 2. 2018 İGE Hesaplamasında Kullanılan Minimum ve Maksimum Değerler....	32
Tablo 3. Gelişme Düzeylerine Göre Çalışma Kapsamındaki Ülkeler	44
Tablo 4. Gelişme Düzeylerine Göre Çalışma Kapsamındaki Ülkelerin Sayısı ve Yüzdesi	45
Tablo 5. Kıtalara Göre Çalışma Kapsamındaki Ülkelerin Sayısı ve Yüzdesi.....	46
Tablo 6. Gelişmişlik Düzeylerine ve Kıtalara Göre Ülke Sayıları ve Yüzdesi.....	46
Tablo 7. İGE, Alt Endeksler ve Faktörlere İlişkin Değişkenler	48
Tablo 8. Değişkenlere ait Spearman Korelasyon Analizi Sonuçları	50
Tablo 9. Karar Ağacı Modeli için C5.0 Algoritmasından Elde Edilen Kurallar.....	52
Tablo 10. Karar Ağacı Modeli için Gini Algoritmasından Elde Edilen Kurallar	55
Tablo 11. Benzer Çalışmalarla Sonuçların Karşılaştırılması	61

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1. Veri madenciliğinin diğer disiplinlerle olan ilişkisi	5
Şekil 2. Veri indirgeme yöntemleri	12
Şekil 3. Veri madenciliği yöntemleri.....	15
Şekil 4. İnsani gelişme endeksi boyut, bileşenler ve oluşumu	32
Şekil 5. Araştırmanın kavramsal yapısı	49
Şekil 6. C5.0 Algoritması ile İGE'nin Kategorik olarak Sınıflandırılmasına ait Karar Ağacı Modeli	51
Şekil 7. Gini Algoritması ile İGE'nin Kategorik olarak Sınıflandırılmasına ait Karar Ağacı Modeli	54
Şekil 8. C5.0 ve Gini Algoritmalarına İlişkin Ülkelerin İGE Sınıflandırma Performans Sonuçları	56
Şekil 9. C5.0 ve Gini Algoritmalarına ait Eğitim ve Test Sınıflandırma Performans Sonuçları	57
Şekil 10. C5.0 ve Gini Algoritmalarının Duyarlılık İstatistiklerinin Performans Sonuçları	58
Şekil 11. C5.0 ve Gini Algoritmalarının Belirleyicilik İstatistiklerinin Performans Sonuçları	59

GİRİŞ

Enformasyon teknolojileri dünyasındaki hızlı gelişmeler, birçok alanda verinin anlık olarak veri ambarlarında kolayca depolanmasına imkân tanımaktadır. Ekonomi, eğitim, sağlık, pazarlama, ticaret, iletişim gibi birden fazla alanda depolanan veriler zaman içinde artarak birer veri yığını haline gelmektedir. Veri ambarlarında bulunan gizli, anlaşılmayan bu veri yığınları analiz edilmediği müddetçe herhangi bir anlam ifade etmemektedir. Birden fazla veriyi bağlantılarıyla birlikte aynı anda çözümleyebilmek ve anlamlandırabilmek amacıyla veri madenciliği kavramı ortaya çıkmış ve zamanla çeşitli veri madenciliği yöntemleri geliştirilmiştir.

Veri madenciliği birden fazla disiplinin katkılarıyla elde edilmiş birikimli bir yapıdır. Bu yapının ana birimleri: yapay zekâ, veri tabanı yönetim sistemleri, makine öğrenmesi ve çok değişkenli istatistik analiz teknikleridir (Akpınar, 2014, s. XII).

Veri madenciliği, büyük ölçekli veri ambarı içerisinde dağınık, karmaşık ve kaotik bir yapıda bulunan verilerden, yararlı ve mantıksal bilgilerin ortaya çıkarılması amacı ile birçok analiz yönteminin kullanımına imkân tanımaktadır. Diğer bir ifade ile veriler arasındaki gizlenmiş olan bağlantı ve eğilimleri çeşitli analiz teknikleriyle ortaya çıkarmaya yarayan bir yöntem olan veri madenciliği, akademik, finans, tıp, iş çevreleri gibi birçok alanda yoğun ilgi gören bir veri analiz tekniğidir (Albayrak, 2017, s. 752; Gürsakal, 2014, s. 35).

Çok boyutlu bir kavram olan insani gelişme, genel anlamıyla insanların içinde bulunduğu ekonomilerin zenginliğinden ziyade, insan hayatının sosyal ve ekonomik olmak üzere tüm alanlarda zenginleşmesi ile ilgili bir yaklaşımdır. Bu yönüyle insani gelişme, temeline insanı koyarak insanların kendilerine ve sahip oldukları fırsatlara odaklanmaktadır. İnsani gelişmenin ölçülmesinde UNDP tarafından sahiplenilen İGE, kapsayıcılık ve bütüncül perspektif yönüyle öne çıkmaktadır.

İGE, sosyal ve ekonomik boyutlardaki genel başarıyı ölçmek için kullanılan istatistiksel bir araçtır. Bir ülkenin sosyal ve ekonomik boyutları; insanların sağlığına, eğitim kazanımlarına ve yaşam standartlarına dayanmaktadır. İGE, 1990 yılında Pakistanlı ekonomist Mahbubu ul Haq tarafından Birleşmiş Milletler Kalkınma Programını daha da ölçeklendirmek ve ileri safhaya taşımak için geliştirilmiştir. İGE, sağlık için tahmin edilen yaşam süresi, beklenen eğitim süresi, ortalama eğitim süresi ve

yaşam standardı için kişi başına düşen gayri safi ulusal gelir olmak üzere dört ana göstereyi hesaplamaktadır. İGE, ekonomik kalkınmadan sorumlu tüm önemli sosyal ve ekonomik göstergeleri bir araya getirmesi nedeniyle bir ülkenin gelişmişlik düzeyini takip etmek için kullanılan en iyi araçlardan birisidir. UNDP, her yıl yayımladığı İGE raporuna göre ülkeleri sıralamaktadır.

Bu tez çalışması beş bölüme ayrılmaya çalışılmıştır. Birincisinde, veri madenciliği yöntemlerine ilişkin kavramlar ve algoritmalar bahsedilmiştir. İkincisinde, gelişmişlik tanımı ve insani gelişme endeksi'nin kavramları açıklanarak hesaplama yöntemlerine değinilmiştir. Üçüncüsünde, insani gelişme endeksi ile ilgili daha önce yapılan çalışmalar incelenmiştir. Dördüncüsünde, araştırmanın uygulama kısmından bahsedilmiştir. Beşincisinde ise sonuç ve öneriler yer almaktadır.

BÖLÜM I

VERİ MADENCİLİĞİ

1.1. Veri Madenciliğine Giriş

Teknoloji ve bilginin son derece yoğun olarak kullanıldığı çalışma hayatımızda bilgisayar teknolojilerinin gelişmesine bağlı olarak çok sayıda veri birikimi ortaya çıkmış, ve bu çok sayıda veri kümeleri içerisinde arasından anlamlı ilişkilerin belirlenmesine yönelik, eğilimlerin ortaya çıkarılması ve dolayısıyla ileriye dönük doğru tahminlerin yapılması önem arz etmiştir. Bu açıdan bakıldığında verileri işleyecek, istatistiksel analiz ve modellemelerle birlikte yapay zekâ tekniklerini dikkate alan veri madenciliği teknikleri uygulanmaya başlanmıştır (Gürsoy, 2012, s. 3).

Veri madenciliği (VM), yaklaşık yarım asırdır teknolojiye paralel olarak gelişimini sürdüren ve kullanım alanı her geçen gün yayılmakta olan bir konudur. VM'nin kullanıldığı yer ve zamana göre çeşitli tanımlamaları yapılmakta olsa bile dijital dünyada meydana gelen hızlı gelişmeler karşısında yapılan tüm bu tanımlamalar yetersiz kalmaktadır (Silahtaroglu, 2013, s. 12). Bu kapsamda araştırmacılar tarafından VM ile ilgili yapılmış birbirinden farklı tanımlamalar bulunmaktadır.

VM, gelişen bilgisayar teknolojilerinin büyük miktardaki veriyi saklayabilmesi sonucunda depolanan veriden; istatistik, yapay zekâ, makine öğrenimi ve veri görselleştirme gibi analiz teknikleri vasıtasıyla anlaşılır olmayan fakat kullanışlı ve yararlı bilgilerin yorumlanması işlemini yerine getirir. Diğer bir ifade ile veri ambarı içerisinde bulunan saklı bağlantıları, kullanıcıların gelecekte yararlanmaları ve öngörülerde bulunmalarını sağlamak için ortaya çıkartma veya aydınlatma işlemidir (Armutlu, 2018, s. 5).

VM, daha önceden bahsedilmeyen uygulamaya yönelik verilerin veri tabanlarından temin edilmesi ve temin edilen verilerin işletmenin geleceğini planlama sürecinde karar verilirken kullanılmasıdır (Silahtaroglu, 2004, s. 4).

VM, sadece bir problemin çözümüne yönelik değil aynı zamanda çözüme ulaşmaya yönelik karar alma sürecinde gerekli aydınlatıcı bilgilerin belirlenmesini sağlayan bir tekniktir (Baykal, 2006, s. 96).

Geçmişten günümüze karar vermede önemli bir rol üstlenen VM, büyük ölçekli veri ambarından değerli olan bir bilgileri elde etme faaliyetidir (Özkan, 2008, s. 38).

Yukarıda belirtilen tanımlardan da yararlanarak VM, modern teknolojilerin sağlanmış olduğu süratli veri işleme ve sınırsız depolama imkânlarının ve çeşitli disiplinlerin (veri ambarı, istatistik, yapay zekâ, makine öğrenmesi ...) katkılarıyla elde edilen çok büyük boyutlardaki verilerden, karar vericinin faydalı bilgiye dayalı karar verme sürecinde faydalanabilmesine yönelik; eskiden ön görülmeyen, saklı, kapalı, standart metotlarla gün yüzüne çıkarılması zor, faydalı, anlaşılabilir verilerin ortaya çıkarılması olarak tanımlanabilir (Şentürk, 2006, s. 3).

Veri tabanı sistemlerinin gelişmesi sonucunda VD'nin; sağlık, bankacılık, güvenlik, eğitim, ekonomi gibi pek çok alandaki uygulama sahasının artarak genişlediği görülmektedir. VM uygulama yöntemlerinin her alanda kullanılmasının nedeni; ekonomik boyut düşünülerek maliyetleri azaltması, satışları fazlalaştırılması ve ileriye dönük tahminleme yaparak oluşacak olan riskleri azaltmasıdır (Gürsoy, 2012, s. 3).

Genel olarak yapılan tanımları kısaca bahsetmek gerekirse literatürde bahsedilen farklı tanımlar birbirleriyle aynı kavramlara değinmektedir. İstatistik gibi veri madenciliği teknikleri geniş hacimli ve karmaşık veriden hareketle anlamlı faydalı bilgiler üretmeye yarayan bir bilim alanıdır. VM, çok büyük boyutlu veri tabanlarındaki karmaşık ve anlaşılması güç verilerin içindeki faydalı bilgilerin ortaya çıkartılmasını, analizler için önemli olabilecek bilgilerin özetlenmesini, elde edilen bilgiler ışığında anlaşılabilir bir şekilde dataların ortaya konmasını sağlar ve bununla birlikte geleceğe yönelik tahminler yapılmasını da mümkün sağlamaktadır (Emre & Selçukcan Erol, 2017, s. 163).

1.2. Veri Madenciliğinin Diğer Disiplinlerle Olan İlişkisi

VM, birden çok disiplinle ilişkiler içerisinde olan ve aynı zamanda farklı bilim dallarından etkilenecek yöntemler geliştiren disiplinler arası bir tekniktir. Şekil 1'de de görüldüğü üzere veri madenciliğinin işbirliği halinde olduğu disiplinler; makine öğrenme, görselleşme, istatistik, veri tabanı teknolojisi, bilgi bilimi ve diğer bilim dallarıdır (Alagöz, Öge, & Ortakarpuz, 2014, s. 4). Kısacası VM, belirtilen disiplinler arasında köprü görevini icra eden bir daldır (Demiral, Soba, & Armutlu, 2017, s. 243).



Şekil 1. Veri madenciliğinin diğer disiplinlerle olan ilişkisi

Kaynak: Gürsoy, 2009, s. 28

1.3. Veri Madenciliği Tekniklerinin Uygulandığı Alanlar

Verinin bulunduğu her alanda veri madenciliğini kullanma olanağı bulunmaktadır. Çağımızda karar verme aşamasında gereksinim duyulan pek çok sahada veri madenciliği yöntemleri yoğun bir şekilde uygulanmaktadır. Literatür taraması neticesinde veri madenciliğinin sık kullanıldığı alanların tıp, biyoloji ve genetik olduğu gözlemlenmektedir (Savaş, Topaloğlu, & Yılmaz, 2012, s. 5).

Bu sık kullanılan alanların dışında veri madenciliğinin uygulamaya yönelik oldukça geniş alanı mevcuttur: Elektronik ticaret, sağlık alanında, pazarlama sektöründe, endüstri, sigortacılık faaliyetlerinde, bankacılık işlemlerinde, eğitim gibi birçok sektör alanında veri madenciliğinin uygulama alanını oluşturmaktadır (Ayık, Özdemir, & Yavuz, 2007, s. 448).

1.4. Veri Madenciliği Uygulamasına Yönelik Süreç

VM, verilere kolay bir şekilde uygulanan ve belirli varsayımları olmadığından literatürlerde tekrarlı bir süreç olarak açıklanmıştır. Veri ambarı içerisinde özetlenmiş faydalı çıkarımlar yaparak veriyi ortaya çıkarmanın yanında, birbirleriyle ilişkili örüntüleri ortaya çıkararak süzmek ve daha sonraki aşamada verileri hazır hale getirmek veri madenciliği sürecinin bir aşamasıdır (Savaş, vd., 2012, s. 19).

VM bir adım olarak değil birbirini takip eden bir süreç olarak karşımıza çıkmaktadır. Birbirini takip eden bu sürecin temelinde süreci uygulayan gerçekleştirci bulunmaktadır. Süreçte var olan adımlar tam ve doğru bir şekilde tamamlanmadığı müddetçe sonuca ulaşılmasının imkânı bulunmamaktadır.

VM sürecinde çalışmanın amacı yönünde izlenmesi gereken adımlar aşağıda açıklanmıştır (Terzi, Küçüksille, Ergin, & İlker, 2011, s. 29-37).

- ❖ Problemin Kavranması
- ❖ Verilerin Hazır Duruma Getirilmesi
 - Veri Entegrasyonu
 - Veri Temizleme
 - Veri Dönüştürme
 - Ver Düzeltme
 - Veri Birleştirme
 - Veri Genelleştirme
 - Veri Normalleştirme
 - ◆ Ondalık Ölçekleme
 - ◆ Z-skor Normalleştirme
 - ◆ Min-max Normalleştirme
 - ◆ Logaritmik Normalleştirme
 - ◆ Değişken Oluşturma
 - Veri İndirgeme
- ❖ Verinin Seçimi
- ❖ Modelin Kurulması Aşaması
- ❖ Modelin Değerlendirilmesi Aşaması
- ❖ Modelin Uygulanmaya Konulması
- ❖ Modelin İzlenmesi Aşaması

1.4.1. Problemin Belirlenmesi

VM sürecinin en önemli evresidir. VM çalışmalarında başarılı olmanın ilk şartı problemin kavranarak açık bir dille tanımlanmasının yapılmasıdır (Gürsoy, 2009, s. 31-32).

VM'nin anlaşılır bir dille tanımlanması, veri hakkında ilk izlenimlerin edinilmesine, veri dizilerinde saklı bulunan enformasyonun gün yüzüne çıkarılması için varsayımların geliştirilmesine, problemin üzerine odaklanılmasına, verinin kavranmasına ve diğer faaliyetlerin gerçekleştirilmesine olanak sağlar (Akpınar, 2014, s. 74; Tüzüntürk, 2010, s. 68).

1.4.2. Verilerin Hazır Duruma Getirilmesi

Veri madenciliğinde, veri kalitesi büyük önem taşımaktadır. Veri madenciliğinde güvenilirliğin üst seviyelere çıkarılabilmesi için veri ön işleme ihtiyacı duyulmaktadır (Oğuzlar, 2003, s. 67-76).

Verinin hazırlanması veri madenciliğinin diğer bir önemli evresidir. Veriyi hazırlama safhasında elde edilen veriler işlem görek veri madenciliği sürecinde kullanılabilir hale getirilir. Bu evre, veri madenciliği sürecinde en fazla zaman harcanan ve iş gücü gerektiren aşama olarak görülür (Altıntaş, 2018).

Verinin hazırlanması aşaması sağlam bir temel üzerine inşa edilmezse, modelin kurulması aşamasında sorunların çıkmasına neden olacağından sonrasında tekrar tekrar bu aşamaya dönülmesine neden olacaktır (Gürsoy, 2009, s. 32). Bunun neticesinde verinin hazırlanarak modelin kurulması evreleri, veri sürecinin tamamı içerisinde enerji ve zamanın % 50 - % 85 oranında harcanmasına sebebiyet verecektir (Akpınar, 2014, s. 77).

Veri hazırlama evresi kendi bünyesinde; veri entegrasyonu işlemi, veriyi temizleme aşaması, veri dönüştürme süreci ve veri indirgeme kısmından meydana gelmektedir (Akpınar, 2014, s. 78). Bu aşamada temin edilen veriler birtakım işlemlerden sonra veri madenciliği uygulamalarına hazır hale getirilmektedir. Böylece, veriler üzerindeki aşamalar ile birlikte veri dönüştürme işlemi yapılmaktadır.

1.4.2.1. Veri Entegrasyonu

Tanımlanan probleme yönelik yapılması gereken veri dizilerinin ve bu veri dizilerinin bir araya getirileceği veri kaynaklarının tespit edilmesi aşamasıdır. Ne tür verilerden faydalanılacağı önemli bir husustur. Olması gerekenden daha az bir veri temini veri madenciliği analiz çalışmasını yetersiz kılabileceği gibi, olması gereken daha fazla veri temini sürecin uzamasına yönelik veri kirliliğine yol açacaktır (Terzi, vd., 2011, s. 3).

1.4.2.2. Veri Temizleme

Veri dizilerinin çeşitli kaynaklardan temin edilmesi veri uyumsuzluklarına zemin hazırlamaktadır. Bu uyumsuzlukların temel nedenlerinin bazıları: farklı zaman dilimlerinden elde edilmesi, kodlamaya yönelik farklılıklar (örnek olarak bir veri tabanında yer alan eğitim özelliklerinin ön lisans/lisans/ yüksek lisans/doktora eğitimi, farklı bir veri tabanında 0/1/2/3 olarak kodlama işleminin yapılması) ve çeşitli ölçü birimleri örnek verilebilir. Bunların dışında veri dizilerinin nerede, nasıl ve hangi şartlar altında bir araya getirildiği de önem arz etmektedir.

Baştan savma ve basit yöntemlerle yapılacak veri temizleme aşamasının, sonraki süreçlerde daha büyük sorunlara neden olacağı göz ardı edilmemelidir. Kaliteli sonuç alınacak modeller sadece iyi veriler üzerine inşa edilebileceğinden, toplanan veri dizilerinin hangi oranda ölçülü oldukları bu süreçte incelenerek değerlendirilmelidir (Oğuzlar, 2003, s. 5; Akpınar, 2014, s. 78).

1.4.2.3. Veri Dönüştürme

Veri ambarlarında veya veritabanında karmaşık veya birbirleriyle ilişkili bulunan veriler, daha anlamlı bir yapı oluşturabilmek için bir takım dönüşüm işlemine tabi tutulurlar (Ayık, Özdemir, & Yavuz, 2007, s. 443).

Diğer bir ifade ile tercih edilen algoritmalara göre verinin kendi içinde uyumlaştırılmasıdır. Veri dönüştürme aşamasında veriler, bir takım işlemlerle veri madenciliği uygulamalarına hazır hale getirilir. Bu amacı gerçekleştirmek için veri dönüştürme aşması; veri düzeltme, birleştirme, genelleştirme ve normalleştirme gibi çeşitli aşamalardır (Akpınar, 2014, s. 78; Yalçın L., 2019, s. 18).

1.4.2.3.1. Veri Düzeltme

Verilerden gürültülü verilerin ayıklanmasını ve azaltılması işlemini sağlamaktadır. Bu aşamada, bölümlene kısmı, kümeleme yöntemi ve regresyon analizi gibi teknikler uygulanmaktadır (Şentürk, 2006, s. 15). Veri özelliklerinde ya da sınıflarında bulunan hataların toplu olarak adlandırılmasına gürültü denilmektedir (Oğuzlar, 2003, s. 68; Akgöbek & Çakır, 2009, s. 804).

1.4.2.3.2. Veri Birleştirme

Çeşitli veri tabanlarından veya veri kaynaklarından temin edilen verilerin bir bütün olarak değerlendirmeye tabi tutulması için çeşitli türdeki verilerin özetlenmesi veya birleştirilmesi bu aşamada gerçekleştirilir (Özkan, 2008, s. 40-41).

1.4.2.3.3. Veri Genelleştirme

Ham veya işlenmemiş verileri daha üst seviyedeki değişkenlere tanımlayabilecek şekle dönüştürme aşamasıdır (Şentürk, 2006, s. 15).

1.4.2.3.4. Veri Normalleştirme

Veri dönüştürme işlemleri içerisinde en çok kullanılan dönüşümlerden birisi olan veri normalleştirme, veri dizisinde bulunan değerlerin $[-1, +1]$ veya $[0, +1]$ aralığında dönüştürülme işlemidir (Akpınar, 2014, s. 133). Verilerin veri bütünlüğüne aykırı olan durumlarda başvurulmaktadır..

Veri normalleştirme yöntemi olarak sıklıkla; ondalık ölçekleme, z-skor normalleştirme, min-max normalleştirme, logaritmik normalleştirme ve değişken oluşturma yöntemleri kullanılmaktadır.

1.4.2.3.4.1. Ondalık Ölçekleme

Ondalık ölçeklendirme ile normalleştirme, değerlendirmeye alınacak değişken değerlerinin ondalık kısmı atılarak normalleştirme işlemi yapılır. Bahsedilen ölçekleme işlemi, değişkenin alacağı değerlerin -1 ile +1 arasında olacak şekilde dönüştürülme işlemi gerçekleştirilir. Değişkenin alacağı maksimum mutlak değeri, atılacak ondalık sayısını belirlemektedir (Tunç & Ülger, t.y.).

Eşitliğin her bir değeri 10'nun kuvvetine bölünerek değerler daha geniş bir alandan daha dar ve küçük bir aralığa toplanmaktadır (Şentürk, 2006, s. 16). Örneğin 1000 maksimum değer ise, $n=4$ olduğundan 0,1 olarak normalleştirilir. Bu işlem aşağıdaki 1.1'deki denklem yardımıyla hesaplanır.

$$V^1(i) = V(i)/10^k \quad (1.1)$$

1.4.2.3.4.2. Z-skor Normalleştirme

Sıfır-ortalama yöntemi, adından da anlaşılacağı üzere 0 ortalama etrafında normalleştirme olarak adlandırılmaktadır. Burada ortalama ve standart sapma değerleri kullanılır (Silahtaroglu, 2013, s. 35). Z-skor normalleştirme yöntemi eşitlik 1.2'deki denklem yardımıyla hesaplanır.

$$x'_i = \frac{x_i - \bar{x}}{\sigma_x} \quad (1.2)$$

Burada x'_i verinin dönüştürülmüş değerleri, x_i gözlem değerlerini, \bar{x} veri tabanındaki ilgili alanın ortalaması, σ_x ise o alanın standart sapmasını ifade etmektedir (Özkan, 2008, s. 43).

1.4.2.3.4.3. Min-Max Normalleştirilmesi

Normalize işlemi olarak bilinen min-max normalleştirilmesi verileri doğrusal olarak 1-0 aralığında bir değer verir. **Min**, bir verinin alabileceği en küçük değeri ifade ederken, **max** verinin alabileceği en büyük değeri açıklamaktadır. Veriyi 1-0 aralığına indirmek için min-max normalleştirme uygulaması eşitlik 1.3'deki denklem yardımıyla hesaplanır (Akpınar, 2014, s. 134).

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1.3)$$

Eğer normalizasyon 1-0 aralığından farklı bir aralıksa, bu durumda söz konusu dönüştürme eşitlik 1.4'deki denklem yardımıyla hesaplanır..

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} (\text{yeni_max} - \text{yeni_min}) + \text{yeni_min} \quad (1.4)$$

Burada max ve min değişkenin bilinen orijinal maksimum ve minimum değerlerini ifade etmektedir. Yeni_max ve yeni_min değişkenleri ise dönüştürülmek istenen değişkenin maksimum ve minimum değerlerini göstermektedir (Şentürk, 2006, s. 16).

1.4.2.3.4.4. Logaritmik Normalleştirme

Veri dizisini normalleştirmek için faydalanılabilecek bir diğer teknik logaritma kullanımıdır (Akpınar, 2014, s. 135). Belirli bir logaritma tabanına göre değişkenin değerlerini logaritma olarak küçültme işlemidir. Bu dönüşüm neticesinde orijinal verilerin değerleri küçülürken herhangi bir veri kaybı söz konusu değildir.

Geniş bir aralıktan dar bir aralığa etkin bir ölçekleme ile dönüştürme işlemi yapılmak isteniyorsa bunu sağlamak için logaritmik normalleştirme kullanılmaktadır (Şentürk, 2006, s. 16).

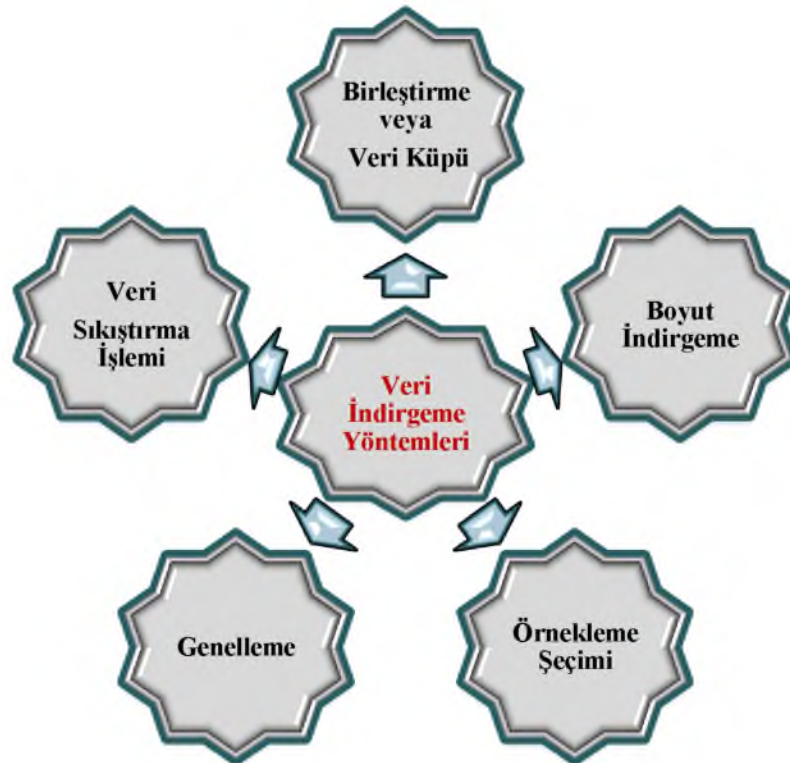
1.4.2.3.4.5. Değişken Oluşturma

Bu yöntemde ise orijinal değişkenlerden faydalanılarak veri madenciliği sürecine katkı sağlayacak yeni değişkenlerin oluşturulması hedeflenmektedir (Şentürk, 2006, s. 16).

1.4.3. Veri İndirgeme

Veri indirgeme süreci, daha küçük boyutlu olacak şekilde bir veri kümesinin indirgenmiş bir yapısını oluşturmak için uygulanmaktadır. Böylece oluşan bir indirgenmiş veri kümesine, veri madenciliği analizleri uygulanarak daha etkin ve verimli sonuçlar elde edilebilecektir (Oğuzlar, 2003, s. 73).

Veri madenciliği problemine ilişkin çözümlerinde ortaya çıkacak sonucun değişmeyeceği düşünülüyorsa değişken ya da veri sayısında azaltma işlemine gidilebilir. Veri indirgeme aşamaları; veri sıkıştırma işlemi, örnekleme seçimi, genelleme, birleştirme veya veri küpü ve boyut indirgemedir (Coşlu, 2013, s. 616).



Şekil 2. Veri indirgeme yöntemleri
Kaynak: Özkan, 2008, s. 41

Boyut İndirgeme: Bu aşamada verileri, çok boyutlu olacak şekilde veri küpleri haline dönüştürme söz konusudur. Bunun sonucunda yapılacak çözümler belirlenecek boyutlara göre yapılacaktır.

Birleştirme veya Veri Küpü: Veriler içerisinde bir eleme işlemi yapılması sonucunda, lüzumsuz verilerin veritabanından çıkarılarak boyut azaltılmasının sağlanması aşamasıdır.

Veri Sıkılaştırma: Bu aşamada, büyük veri kümeleri sıkılaştırılmak suretiyle daha az yer kaplamaları sağlanır.

Örnekleme: Veri anakütlesi yerine onu temsil edebilecek daha küçük verili kümelerinin oluşturulmasıdır.

Genelleme: Verilerin tek başına değil genel kavramlarla açıklanmasıdır (Özkan, 2008, s. 41-42).

1.4.4. Verinin Seçimi

Bu aşamada ilk önce kurulacak modele ilişkin olarak veri seçimi gerçekleştirilir. Tahminleyici bir modele yönelik olarak, bağımlı ve bağımsız değişkenler ile modelde kullanılmak için yer alacak bir veri kümesinin belirlenmesi aşamasıdır (Gürsoy, 2009, s. 35). Başka bir ifade ile veri seçimi, veri kaynaklarından verilerin seçilerek veri dosyasını hazır hale getirme aşamasıdır (Yıldırım, Uludağ & Görür, 2008, s. 430).

Veri madenciliği uygulamasında, evreni temsil edecek şekilde örnekleme yapılabilmesi için kullanılan modelde veri alanının çok geniş tabanlı tutulması gerekmektedir. Kullanılmakta olan hazır paket programların performansı ne kadar iyi olursa olsun, çok büyük veri tabanlarında birden fazla modelin uygulanması sorun teşkil etmektedir.

Bundan dolayı tüm veri tabanını kullanarak çok sayıda model uygulamak yerine, bir veri tabanında tesadüfî örnekleme yöntemiyle model seçilerek içerisinde en güvenilir, anakütleyi en iyi temsil eden güçlü modelin tercih edilmesi daha uygun olabilecektir. Başka bir ifade ile modellerin çıktılarını seçilmiş uygun bir karar yöntemi ile denemelidir (Terzi, vd., 2011, s. 32).

1.4.5. Modelin Oluřturulması

Veri madencilięinde bilgi bankalarından en üst seviyede faydanın saęlanabilmesi için, modelin kurulması adımı son derece önemlidir. Probleme iliřkin bir model, analizler için elde edilecek sonuçların performansını etkileyebilecektir (Gürsoy, 2009, s. 35).

Probleme iliřkin en uygun model seçimi, birden fazla modelin uygulanması ile mümkündür. Buna baęlı olarak veri hazırlama ile model kurma aşamaları, en çok fayda saęlayan modeli ulařıncaya kadar tekrarlama işlemidir (Ayık, vd., 2007, s. 443). Eęer model yanlış kurulursa veri seti içerisinde yer alan birbirleriyle iliřkili bilgiler ortaya çıkartılamaz. Bundan dolayı modelden başarılı sonuç elde etmek mümkün olmayabilir (Gürsoy, 2009, s. 35).

1.4.6. Modelin Deęerlendirilmesi

Tanımlanan problem için uygulanan modelin, verimlilięinin ve kalitesinin ölçüldüęü, veri madencilięi süreçlerinin tekrar gözden geçirilerek belirlenen kriterlere uyumlu olup olmadıęının tespiti ve bunların neticesinde sonuçların kullanımını için karar verilmesi süreçlerini içermektedir (Fakı & Öztayři, (t.y.), s. 3).

1.4.7. Modelin Uygulamaya Konulması

Veri madencilięi evresinin bu aşamasında, model uygulanmaya konulur. Uygulanmaya konu olan model en uygun model de olabilir ya da farklı bir modelin alt yapısını da oluşturabilir. (Gürsoy, 2009, s. 37).

1.4.8. Modelin İzlenmesi

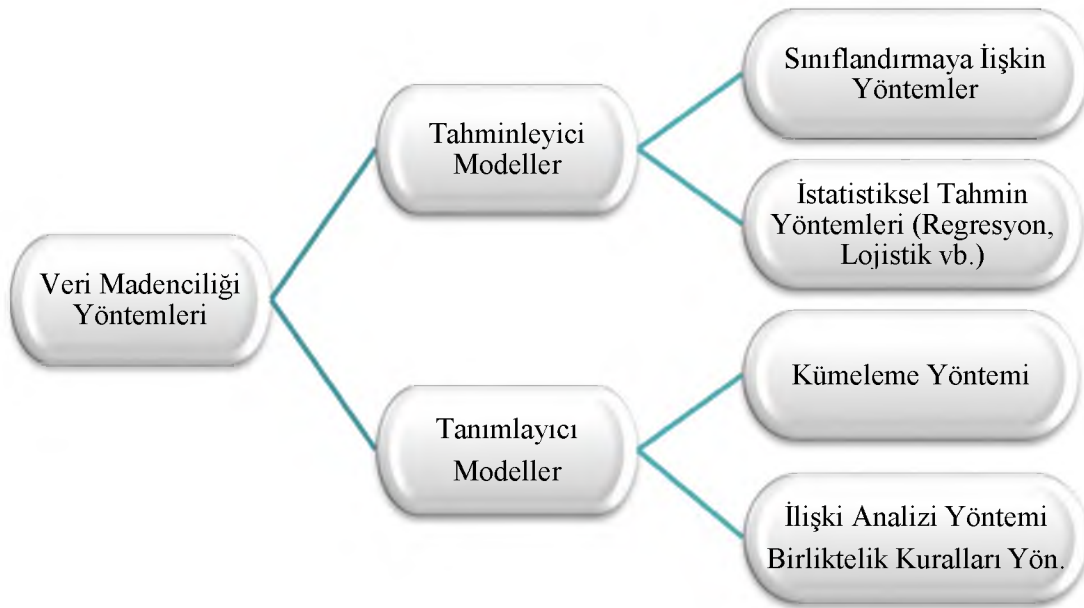
Veri madencilięi evresinin son aşamasıdır. Uygulamaya konu olan modelin, zaman içerisinde meydana gelebilecek deęişikliklere baęlı olarak güncellenmesi ve yeniden dizayn edilmesidir (Savaş, vd., s. 8).

1.5. Veri Madenciliğine Ait Bazı Yöntemler

Veri madenciliği uygulamalarında çok sayıda yöntem ve algoritmalar gerçekleştirilmiştir. Bu algoritmaları tahminleyici ve tanımlayıcı modeller olarak ikiye ayırmak mümkündür.

Tahminleyici modellerde amaç, değişkenlerin aldığı değerlerden faydalanılarak bir model oluşturulması ve oluşturulan bu modeller yardımıyla hedef değerlerinin tahmin edilmesidir.

Tanımlayıcı modellerdeki amaç ise, karar vermede kullanılabilecek mevcut verilerdeki ilişkilerin tanımlanmasını sağlamaktır (Özkes, 2003, s. 67). Bir takım veri madenciliği modelleri, Şekil 3'te gösterilmiştir.



Şekil 3. Veri madenciliği yöntemleri

Şekil 3'te de görüldüğü üzere veri madenciliği alanında geliştirilmiş olan çok sayıda yöntem bulunmaktadır. Bu tez çalışması kapsamında ilgili yöntemlerden sınıflandırma, kümeleme ve birliktelik kuralları ele alınmıştır.

- ❖ Sınıflandırma Yöntemi
 - Karar Ağaçları
 - ◆ Gini Algoritması
 - ◆ C5.0 Algoritması
- ❖ Kümeleme yöntemi
- ❖ Birliktelik kuralları yöntemi

1.5.1. Sınıflandırma Yöntemi

Veri madenciliğinde sıklıkla başvurulan bir yöntem olan sınıflandırma, büyük veri tabanları içerisindeki saklı kestirimsel bilgilerin örüntülerini meydana çıkarmakta kullanılır (Özkan, 2008, s. 45). Sınıflandırma yöntemi iki aşamalı bir süreçtir. Birinci aşamada tahmin amacıyla kullanılacak bir model oluşturulurken ikincisinde ise oluşturulan bu modellerin sınıfları bilinmeyen veriler üzerinde uygulanması sonucunda verilerin sınıfları tahmin edilmektedir (Yakut, 2012, s. 18).

1.5.1.1. Karar Ağaçları

Karar ağaçları; mümkün tüm faaliyetlerin yönlerini, faaliyetlerin etkisi olabilecek tüm olası faktörleri ve tüm bu faktörlere ilişkin her bir olası sonucu, verilere ilişkin olarak değerlendiren çeşitli geometrik semboller kullanımı yoluyla bir veri setinden, kullanıcıların basit bir şekilde anlayabileceği kuralları yapısında düğüm, dal, yaprak bulunan bir ağaç olarak düzenleme biçiminde tanımlanabilir (Gürsoy, 2009, s. 42).

Tahmin edici ve açıklayıcı özellikleri içinde barındıran karar ağaçları; ortaya çıkarılmasında/getirilmesinde düşük maliyetli olması, analiz edilerek yorumlanmasının kolay ve açıklayıcı olması, veri tabanı sistemleri ile kolay entegre sağlayarak, güvenilirliklerinin üst seviyede olması gibi sebeplerden, sınıflandırma problemlerinde en çok aranan ve alanında en fazla yazılımların bulunduğu algoritmalarındandır (Baysakaloğlu, 2005, s. 2-5).

1.5.1.1.1. Gini Algoritması

CART veya C&RT (Classification and Regression Trees) ağacı olarak da tanımlanan GİNİ algoritması, karar ağacı belirlenmesinde kullanılan bir algoritma türüdür (Adak & Yurtay, 2013, s. 2).

Klasik karar ağacı algoritmaları içerisinde önemli bir yere sahip olan CART, Morgan ve Sonquist tarafından AID (Automatic Interaction Detection) olarak adlandırılan karar ağacı algoritmalarından olan; 1984 yılında Breiman, Friedman, Olshen ve Stone tarafından geliştirilmiş ikili (binary) ağaç olarak oluşturulan bir algoritma çeşididir (Akcınar Sezer, Bozkır, Yağız, & Gökçeoğlu, 2010, s. 2).

CART karar ağacı, her bir karar düğümünden dallanarak iki alt kümeye ayrılması ilkesine bağlıdır. Buradaki amaç bir sonraki adımda oluşacak olan alt kümede daha homojen hale getirilmiş düğümlerin elde edilmesidir (Şengür & Tekin, 2013, s. 9). Hem nicel hem de nitel veri çeşitlerini, girdi değişkeni olarak alan CART, iki işlevi birden yerine getirerek sınıflandırma ve regresyon problemlerin çözümünde kullanılmaktadır (Yakut, 2012, s. 27).

Sürekli ve kategorik değerlerle çalışabilen CART, sürekli değişkenlerinin bölen noktasının bulunmasında en küçük kareler sapması; kategorik değerler için bölen değişkeninin bulunmasında ise gini, twoing, ordered twoing kriterlerini kullanır (Akcınar, 2014, s. 225).

CART algoritmasında her bir düğümde olasılıkları dikkate alarak bölünme işlemi gerçekleştirilmektedir. CART algoritması, bu işlemi gerçekleştirirken hangi düğümden sonra bölünme işleminin gerçekleşeceğine karar vermektedir. Bölünme işlemi düğümlerin ikiye ayrılması ile gerçekleşmektedir. Diğer bir ifade ile CART algoritması; karar ağacına ait dalları ayırma işleminde uygun değişkeni belirler, değişkenin aldığı değerlere bağlı olarak grupların nasıl ayrılacağını tespit eder. Bu işlemi gerçekleştirirken bağımsız değişken olarak ifade edilen tüm niteliklerin aldığı değerler dikkate alınarak, iki bölünmeler belirli işlemlere göre gerçekleştirilir ve seçme işlemi uygulanarak karar ağacının dalları oluşturulur. Tüm bu işlemler aşağıdaki parametreler kullanılarak gerçekleştirilir (Akcınar, 2014, s. 82; Yakut, 2012, s. 27).

Herhangi bir t düğümündeki s dallara ayrılma kriteri $\Psi(s/t)$ olarak gösterilirse; hesaplanan $\Psi(s/t)$ değerleri içinden en yüksek değer, düğüm olarak belirlenir ve işlem tüm yapraklar oluşturuluncaya kadar devam ettirilir (Aksoy, 2018, s. 44). Bu işlem eşitlik 1.5, 1.6, 1.7, 1.8 ve 1.9'daki denklemler yardımıyla hesaplanır.

$$\Psi(s|t) = 2P_L P_R \sum_{j=1}^{\pi} |P(C_j|t_L) - P(C_j|t_R)| \quad (1.5)$$

t : Dallanmanın yapılacağı düğüm

c : Kriter

L : Ağacın sol tarafı

R : Ağacın sağ tarafı

P_L, P_R : Öğrenim kümesindeki bir kaydın sağda veya solda olma olasılığı

$P(C_j|t_L)$ ve $P(C_j|t_R)$: C_j sınıfındaki bir kaydın sağda veya solda olma olasılığı

$$P_L = \frac{t_L \text{ deki kayıtların sayısı}}{\text{Eğitim setindeki kayıtların sayısı}} \quad (1.6)$$

$$P_R = \frac{t_r \text{ deki herbir nitelik değerinin ilgili nitelik sütunundaki tekrar sayısı}}{\text{Eğitim setindeki örneklerin sayısı}} \quad (1.7)$$

$$P(C_j|t_L) = \frac{t_R \text{ deki } C_j \text{ sınıfının sayısı}}{t \text{ deki kayıtların sayısı}} \quad (1.8)$$

$$P(C_j|t_R) = \frac{t_L \text{ deki } C_j \text{ sınıfının sayısı}}{t \text{ deki kayıtların sayısı}} \quad (1.9)$$

GİNİ algoritmasının hesaplanmasında ise bir bağımsız değişken niteliğinin GİNİ değeri hesaplanmadan önce niteliğin her bir nitelikle ilgili sol ve sağ taraftaki $Gini_{sol}$ ve $Gini_{sağ}$ değerleri hesaplanmalıdır (Adak & Yurtay, 2013, s. 2). Bu işlem eşitlik 2.10 ve 2.11'deki denklemler yardımıyla hesaplanır.

$$Gini_{sol} = 1 - \sum_{i=1}^k \left[\frac{L_i}{|T_{sol}|} \right]^2 \quad (1.10)$$

$$Gini_{sağ} = 1 - \sum_{i=1}^k \left[\frac{R_i}{|T_l|} \right]^2 \quad (1.11)$$

Denklem $Gini_{sol}$ ve $Gini_{sağ}$ değerlerinin sembollerinin açıklamaları;

- k : Sınıfların sayısı
- T : Bir düğümdeki örnek sayısı
- T_{sol} : Sol koldaki örneklerin sayısı
- $T_{sağ}$: Sağ koldaki örneklerin sayısı
- L_i : Sol koldaki i kategorisindeki örneklerin sayısı
- R_i : Sağ koldaki i kategorisindeki örneklerin sayısı

Hesaplanan bu sol ve sağ değerler değişken niteliğinin Gini değerinin hesaplanmasında kullanılır. Bir değişken niteliğinin Gini değeri eşitlik 1.12'deki $Gini_j$ denklemi kullanılarak hesaplanır (Adak & Yurtay, 2013, s. 2).

$$Gini_j = \frac{1}{n} (|T_{sol}| Gini_{sol} + |T_{sağ}| Gini_{sağ}) \quad (1.12)$$

Her bir deęişken nitelięi için hesaplanan Gini deęerleri arasından minimize olanı alınır ve bölünme bu nitelik üzerinden gerçekleştirilmektedir. Geriye kalan veri seti üzerinde en baştaki adımlara tekrar dönülür, aynı işlemlere devam edilerek dięer bölünmeye ilişkin GİNİ deęerleri hesaplanır (Özkan, 2008, s. 106).

1.5.1.1.2. C5.0 Algoritması

C5.0 algoritmasını Quinlan, 1994 yılında Unix/Linux için geliştirmiş ve ticari olarak piyasaya sunmuştur (Akpınar, 2014, s. 222). C5.0, C4.5'e göre süratli olması, daha etkin amacıyla bellek kullanımı, daha küçük karar ağaçları oluşturma, boosting desteęi ve sonuca herhangi bir katkısı olmayacak özneliklerin elenmesini sağlayan winnoving gibi özellikleri nedeniyle sınıfı içerisindeki algoritmalara göre daha üstün özelliklere sahiptir (Terlemez, 2008, s. 40).

C5.0'ın denetimli bir sınıflandırma algoritma olması nedeniyle, C5.0'da ilk olarak hedef parametreleri bir veri kümesi ile karar ağacı eğitilerek bir model oluşturulur ve ardından oluşturulan bu modelle baęımlı deęişkenin daha önce belirtilmedięi yeni bir veri kümesi üzerinde uygulanarak modelin performansı ölçülür (Seyrek & Ata, 2010, s. 78).

Daha küçük karar ağaçları oluşturabilen, verinin ayrıştırılması ve bulanık verilerin iyileştirilmesinde etkili olan C5.0 algoritması, optimal olarak düğümleri ayırmada bilgi kazancı ve entropi deęerlerinin maksimize edilmesi esasına dayanmaktadır (Diler, 2016, s. 41; Yakut & Gemici , 2017, s. 465).

X deęişkeni (attribute) için k adet olasılıklar sırasıyla $P_1, P_2, P_3, \dots, P_k$ olarak adlandırılmaktadır. X deęişkeni veya nitelięi için entropi hesaplama denklemi eşitlik 1.13'te verilmiştir (Yakut & Gemici , 2017, s. 465).

$$Entropi = H(X) = - \sum_{j=1}^k p_j \log_2 (P_j) \quad (1.13)$$

X nitelięine baęlı olarak hedef nitelięini gösteren T alt kümelerinin $P_1, P_2, P_3, \dots, P_k$ olarak alt bölmelere ayrıldığı varsayılmaktadır. Her bir T hedef nitelięinin sınıfını belirlemede kullanılan bilgilerin aęırlıklı ortalaması ile entropilerin aęırlıklandırılmış toplamları hesaplanır. Bu işlem eşitlik 1.14'deki denklem yardımıyla hesaplanmaktadır.

$$H_S(T) = \sum_{i=1}^k p_i H_S(T_i) \quad (1.14)$$

İlgili hesaplama işlemi yapıldıktan sonra ayırma işlemini gerçekleştirmek için bilgi kazancı hesaplanmaktadır. C5.0 algoritması, böylelikle her bir karar verme düğümünden elde edilen maksimize bilgi kazancına sahip olan ayırma ölçütünü hesaplayarak optimal ayırma işlemini gerçekleştirmektedir. Information Gain değerinin hesaplama denklemi eşitlik 1.15'te verilmiştir (Yakut & Gemici, 2017, s. 466).

$$\text{Information Gain}(S) = H(T) - H_S(T) \quad (1.15)$$

1.5.2. Kümeleme Yöntemi

Çok değişkenli istatistik tekniklerinden birisi olan ve çok sayıda veri içeren bir alandaki verileri gruplara ayırarak, oluşan grupları karşılaştırma esasıyla kullanılan bir veri madenciliği yöntemidir. Kullanımının kolay ve sonuçlarının anlaşılabilir olması kümeleme analizinin avantajları olarak görülebilir (Tekin, 2015, s. 392; Kayri, 2008, s. 106).

Tüm bu yönleriyle kümeleme analizi; tıp, pazarlama, biyoloji, arkeoloji, ekonomi, psikoloji, antropoloji gibi birçok ve birbirinden çok farklı bilim dallarında yararlanılan bir yöntemdir (Aşan, 2007, s. 3; Silahtaroglu, 2013, s. 155).

Kümeleme yönteminin amacı; birbirinden farklı olan veya aralarında yüksek oranda doğal benzerlik bulunan veri ya da değişkenleri uzaklık ve benzerliklerine göre gruplara ayırmaktır (Çelik & Kahyaoğlu, 2007). Diğer bir ifade ile uzaklık matrisi ya da belirsizlik matrisinden yararlanan kümeleme yönteminde birbirine en çok benzeyen birimler aynı kümelerde toplanır (Aşan, 2007, s. 3).

Kümeleme yöntemiyle oluşturulan bu tür kümeler, kendi içlerinde yüksek derecede homojen, diğer kümeler arasında ise yüksek derecede heterojen durumundadırlar. (Taşkın & Emel, 2010, s. 397)

1.5.3. Birliktelik Kuralları Yöntemi

Yaygın olarak kullanılan birliktelik kuralları yöntemi, veri madenciliğinin en önemli konularından biridir (Birant, ve diğerleri, 2010, s. 215). Birliktelik kuralları, bazı veri madenciliği yöntemleri tarafından, veritabanı içindeki kayıtların birbirleriyle olan ilişkilerinin incelenmesi sonucu hangi olayların eş zamanlı olarak birlikte gerçekleşebileceklerinin belirlenmesiyle, elde edilen verideki potansiyel ilişkileri tanımlar (Özkan, 2008, s. 48).

Diğer bir ifade ile olayların birlikte gerçekleşme olasılıklarını dikkate alan veri madenciliği tekniklerinden olan birliktelik kuralları denilmektedir (Alan, 2012, s. 167). Özellikle pazarlama alanında uygulama alanı bulan birliktelik kuralları, müşterilerin satın alma eğilimlerinin tanımlanmasını sağlayan pazarlama amaçlı olarak kullanılan “Pazar Sepet Analizi” adı altında kullanılan bir analizdir (Özkan, 2008, s. 49; Erpolat, 2012, s. 138).

Kümeleme ve birliktelik süreçlerinden oluşan Pazar Sepeti analizinde, kümeleme süreci ile birbirine benzer aynı kategoride incelenmesi gereken müşterilerin kümeler halinde gruplanması sağlanırken, birliktelik sürecindeyse kümelenmiş ve birbirine benzer özellikteki müşterilerin benzer satın alma davranışı eğiliminde bulunacağı varsayılır. Belirlenen bir küme üzerinden müşterilerin satın alma davranışlarındaki ortak noktalar tespit edilerek kümeler oluşturulur (Kırtay, ve diğerleri, 2015, s. 173).

Birliktelik kuralları analizi ile müşterilerin alışveriş alışkanlıkları, veritabanındaki bilgiler aracılığıyla ortaya çıkarılmaya çalışılır. Birliktelik kuralları müşteri tarafından bir ürün satın alındığında, o ürünle birlikte başka hangi ürünleri satın aldığı araştırılır (Gürsoy, 2009, s. 50).

BÖLÜM II

GELİŞMİŞLİK TANIMI VE İNSANİ GELİŞME ENDEKSİ

2.1. Gelişmişliği Esas Alan Tanımlar

Çalışmanın bu bölümünde büyüme, gelişme ve kalkınma kavramlarının tanımına, gelişmişlik ölçümündeki tarihsel sürece ve kalkınmışlık ile gelişmişlik ölçümünde kullanılan temel göstergeler açıklanmıştır.

2.1.1. İktisadi Büyüme Tanımı

Büyüme, iktisadi hayatın temel verilerinde (fiziki sermaye, istihdam, ,doğal kaynaklar, teknolojik gelişme), kişi başına bir yıldan diğerine daha fazla bir reel gelir temin edecek şekilde devamlı artışlara denir (Ülgener, 1991, s. 409). Diğer bir ifade ile büyüme, bir ülkedeki ekonomide meydana gelen artışı ve kişi başına gelir artışını açıklamaktadır. Kişi başına reel gelir veya hasılda oluşan artışın büyüme olarak ifade edilebilmesi için bu artışın devamlı olması gerekmektedir. Bu nedenle iktisadi büyüme kısa dönemli statik değil, uzun dönemli dinamik bir yapıdır (Taban, 2016, s. 1).

Bir ülkenin belirli bir dönem içinde (genellikle bir yıl içinde) üretim kapasitesi veya GSYİH’de görülen ve sayısal olarak hesaplanabilen reel artışlara iktisadi büyüme denir (Üzümcü, 2012, s. 3). İktisadi büyümenin ölçülmesinde birçok kriter bulunmakla birlikte genel olarak GSYH’de meydana gelen nicel artış ile ölçüm yapılmaktadır. David Ricardo, Adam Smith, Karl Marx ve Robert Malthus’un öncülüğünü yaptığı büyüme teorilerinin temelini klasik büyüme teorisi oluşturur (Rençber, 2018, s. 5).

2.1.2. Gelişmişlik ve Kalkınmışlığa İlişkin Tanımlar

Kalkınmışlık ve gelişmişlik ile ilgili çalışmalar genellikle II. Dünya Savaşı sonrası ivme kazanmıştır. 1950-1970 döneminde ülkelerin kalkınmışlık göstergeleri daha çok milli gelirdeki artışla ifade edildiğinden o dönemdeki kalkınmanın temel amacı üretim ve istihdam yapısını, tarım yerine gelir getirici sanayi ve hizmet sektörleri için artırmak için kullanılmıştır. Bir ülkenin refah seviyesini ölçmenin temel göstergesi olarak kişi başına düşen milli gelir ölçütünün kullanılması bu alanlara geçilmesinde etkin rol oynamıştır. 1970’den itibaren gelişmişliği, insani, sosyal, kültürel, mekânsal ve çevresel boyutlarıyla tanımlamak amacıyla farklı yaklaşımlar ortaya çıkmıştır. Bu dönemde yoksulluk, bölgesel eşitsizlik ve gelir dağılımı, işsizlik gibi tanımlar iktisadi büyüme kavramı yanında kalkınma tanımlarının içinde bulunmaya başlamıştır (Başar, Eren, & Eren, 2015, s. 846).

Kalkınmışlık, sürekli ekonomik büyümeyi devam ettirebilmek için kıt kaynakların etkin ve verimli bir şekilde kullanarak ekonomik, siyasi ve kurumsal olanakların insanların yaşam kalitesini iyileştirilmesi yönünde kullanılmasını ifade eder. Kalkınmışlık uzun dönemli ekonomik büyümenin yanı sıra sosyal ve kültürel yapının da devamlı olarak iyileştirilmesi olarak açıklanır. Büyüme ve kalkınmışlık arasındaki farka bakılacak olursa büyüme nicelikteki artışı dikkate alırken, kalkınmışlık ise niteliksel iyileştirilmeye göre açıklanmaktadır (Üzümcü, 2012, s. 5; Rençber, 2018, s. 6).

Anlam itibarıyla gelişmişlik ile aynı anlama gelen kalkınmışlık, bir ülkede gelir ve üretim artışlarıyla beraber ekonomik, siyasi ve kültürel alanlarda yaşanan yapısal değişim süreci olarak tanımlanmaktadır. Buna bağlı olarak kalkınma bir ülkede yaşanan niceliksel artışlarla birlikte niteliksel değişme yolundaki her şeyi ifade etmektedir (Erbay & Özden, 2013, s. 4). Tanımlardan da anlaşılacağı üzere kalkınmadan söz edilebilmesi için; üretim ve tüketim olgularının yapısal değişime uğraması, büyümenin durağan değil sürdürülebilir olması, teknolojik ilerleme, siyasi, sosyal ve kurumsal yenileşme ile birlikte yaşam kalitesinde de geniş çaplı bir iyileşmenin ortaya çıkması gerekmektedir (Günsoy, 2005, s. 36).

Gelişme, bir ülkenin sadece üretiminin ve fert başına gelirin artırılması olmayıp bunu yanında toplumun ekonomik, sosyal, kültürel ve siyasi yapısında iyileşmelerin olmasıdır. Kavram açıklamasından da anlaşılacağı üzere temelinde insan faktörü olan gelişme, bir ülkenin ekonomik bakımdan gelişmesiyle birlikte fertlerin uzun, sağlıklı ve mutlu bir yaşama kavuşmasını amaç edinmiştir (Berber, 2017, s. 239-240).

2.1.3. Gelişmişliğin Ölçümüne Yönelik Tarihsel Boyutu

Uzun bir süre ekonomik kalkınma ile özdeşleştirilen gelişmişlik, 1970'lere kadar kişi başına düşen gelir ile hesaplanmaktaydı ve sanayileşmeye dayalı ekonomik büyüme olarak bilinmekteydi. Buna bağlı olarak GSYH, kişi başına düşen milli gelir ve diğer iktisadi göstergeler ile kalkınmışlığı en çok açıklayan nicel veriler olarak ifade edilmekteydi. Ekonomik büyümenin yoksulluğu azaltmaması ve ortaya çıkan çeşitli toplumsal sorunlara çözüm sunmadığı gerekçesiyle gelişmişlik kavramını ekonomik büyüme ile özdeşleştiren bu yaklaşıma yönelik eleştiriler olmuştur (Tunç & Ertuna, 2015, s. 133).

20. yüzyılda ekonomik büyümede hız kazana bazı ülkelerde siyasi istikrarsızlık, artan işsizlik oranları, gelir dağılımdaki uçurumlar gibi sorunlarla karşı karşıya kalındığında ekonomik büyümenin gelişme için tek başına yeterli olmadığı görüldü. Gelişmenin yalnız ekonomik büyüme ile ölçülmesinin yetersizliğinin anlaşılması üzerine, ülkelerin en alt katmanlarının bile menfaatlerini göz önünde bulunduracak alternatif politikaların uygulanabilirliği gündeme gelerek yeni ölçme tekniklerinin gelişmesine zemin hazırlanmıştır. Bu aşamada UNDP tarafından, 1990 yılında Pakistanlı maliye bakanı ve iktisatçı olan Mahbub ul Haq öncülüğündeki bir ekip tarafından hazırlanan ilk İGE raporu yayımlandı. Bu raporun amacı toplumsal kalkınma politikalarının merkezinde insanın bulunması gerektiğini belirterek ve milletlerin ilgisini fertlerin hayat kalitesine ve mutluluğuna çekmekti (Kaynak, 2017, s. 42; Tunç & Ertuna, 2015, s. 134).

Temelde Amartya Sen'in yapabilirlik yaklaşıma dayanan insani gelişim kavramı ile Amartya Sen tarafından yapabilirlik yaklaşımı ile yoksulluk ve kalkınma anlayışlarına farklı ve yeni bir görüş açısı getirilmiştir. Sen, gelişmenin temel hedefinin insan için iyi ve onurlu bir yaşam sağlamak olması gerektiğini açıklayarak gelişme aşamasında kişinin ne olabildiği ve neleri başarabildiği konusunu temel kalkınma ölçütü olarak açıklamıştır (Şanlısoy, 2019, s. 1581).

UNDP tarafından Amartya Sen'in yapabilirlik yaklaşımdan yola çıkılarak geliştirilmiş olan İGE, insani gelişmeyi nicel olarak açıklamıştır. Süregelen ölçüm tekniklerinden farklı olarak İGE, bir ülkeyi gelişmişlik bakımından açıklarken; hesaplamada eğitim ve gelir boyutları ile yaşam beklentisini kullanır.

2.1.4. Kalkınmışlığın Ölçümüne Yönelik Temel Göstergeler

Kalkınmışlığın ölçümünde her ne kadar çeşitli göstergelerden yararlanılıyor olsa da genelde ülkelerin kalkınmışlığını değerlendirmek için doğrudan ve dolaylı kalkınma olmak üzere iki temel gösterge kullanılmaktadır. Bunlardan doğrudan kalkınma göstergelerini, ulusların kalkınma seviyelerini direk etkileyen ve niceliksel olarak değerlendirilebilen göstergeler olarak tanımlamak mümkünken, dolaylı kalkınma göstergeleri ise ulusların kalkınma düzeylerine dolaylı olarak etki eden nitel göstergelerle tanımlanabilmektedir (Çiftçi, 2008, s. 53).

2.1.4.1. Doğrudan Kalkınmışlığa Yönelik Göstergeler

Nicel anlamda bir ülkenin kalkınmışlık seviyesinin belirlenmesinde yararlanılan doğrudan kalkınma göstergelerinde, ülkenin ekonomik görünümünü açıklamada makro-ekonomik, ülkede yaşayan halkın ekonomik durumunu açıklamada ise sosyo-ekonomik göstergelerden faydalanılmaktadır. Makro-ekonomik göstergeler, GSMH, Tüketici Fiyat Endeksi (TÜFE), perakende satış, iş gücü piyasa verileri, kur değişim oranları ve faizlerin etkisi, sanayi üretimi, işsizlik oranı gibi değişkenlerle ölçülmektedir. GSMH, bir ülke ekonomisini malların ve hizmetlerin söz konusu yıl içerisindeki değerlerini ve pazardaki yerini gösteren en geniş indikatördür. Kalkınmışlık kavramı ile birlikte ekonomik gelişmişliğin yanında sosyal gelişmişlik göstergeler de hesaplamalarda kullanılmaya başlanmıştır (Rençber, 2018, s. 10).

Sosyo-ekonomik göstergeler, kişi başına milli gelirin artmasının yanısıra yapısal ve insani gelişmeyi de içinde barındıran ve ölçülebilen bütün sosyal değişkenleri kapsamaktadır. Bu bağlamda sosyo-ekonomik göstergeler; istihdam, sağlık, eğitim, demografik, altyapı ile diğer refah göstergelerinden oluşmaktadır (Erol, 2013, s. 199).

2.1.4.2. Dolaylı Olarak Kalkınmışlığa Yönelik Göstergeler

Bir ülkenin kalkınmışlığını dolaylı yollardan etkileyen ve nitel kavramlar olan hukuki ve siyasi göstergelerden oluşan doğrudan kalkınma göstergeleri; siyasi göstergelerin ölçümünde, meclisteki kadın milletvekili sayısı, ülkedeki seçme ve seçilme yaşı, seçime katılım yüzdesi, yolsuzluk yüzdesi gibi göstergeler aracılığıyla ölçülmektedir. Hukuki göstergelerin ölçümünde ise yasal düzenlemeler, hane halkının her türlü yaşam standartlarını düzenleyen kanunlar gibi göstergeler kullanılmaktadır (Taban, 2016, s. 39).

2.1.5. Gelişmişliğin Ölçümüne Yönelik Temel Göstergeler

Çok boyutlu bir kavram olan insani gelişmişliğin ölçümünde, Birleşmiş Milletlerin İGE’de kullandığı üç temel değişken kullanılmaktadır. Bu üç temel değişken; uzun, sağlıklı ve yaratıcı bir hayat yaşamak, bilgi ve eğitim alabilme imkânına sahip olmak ve insana yaraşır bir hayat için gerekli kaynaklara erişilebilmektir. Diğer bir ifade ile insani gelişme için gelir, sağlık ve eğitim göstergelerinden oluşan bu üç temel unsur sağlandığında, yaşamın farklı alanlarında da ilerleme ve gelişme fırsatları da artacaktır (Erdoğan, 2010, s. 326).

2.2. İnsani Gelişme Endeksi Kavramı

İnsani gelişme kavramı, II. Dünya Savaşı öncesinde siyasi otoriteler, iktisatçılar ve kurumlar tarafından sadece ülkelerin ekonomik kalkınmışlıklarının göstergesi olarak görülmekteydi. Bir ülkede ekonomik büyüme gerçekleşmişse o ülke aynı oranda gelişmiş ve kalkınmış sayılmaktaydı (Zor, 2019, s. 25). İlerleyen yıllarda ekonomik bakımdan gelişmiş pek çok ülkede sosyal sorunların bir çözüme ulaşamadığının görülmesi, ekonomik ilerleme ile insani gelişme arasındaki ilişkinin daha iyi kurulması gerekliliğini ortaya çıkarmıştır (Doğan & Tatlı, 2014, s. 101).

Ülkelerin kalkınmışlık göstergelerinin boyutlarının değişerek, milletlerarası sosyo-ekonomik gelişmişlik seviyelerinin de dikkate alındığı insani gelişme kavramı 1990 yılından itibaren kalkınma literatüründe çok sık kullanılmaya başlanmıştır (Günsoy, 2005, s. 37). İnsani gelişme kavramı ilk defa Birleşmiş Milletler Kalkınma Programı (UNDP) tarafından kullanılmıştır. İnsani gelişme; İlk defa UNDP tarafından 1990 yılında

maliye bakanı olan Pakistanlı ekonomist Mahbul ul Haq liderliğinde hazırlanan İnsani Gelişme Raporunda (İGR) bireylerin seçeneklerini fazlalaştırma süreci şeklinde tanımlanmıştır (UNDP, Human Development Report, 1995).

Bu seçenekler sonsuz ve değişken olmakla birlikte Amartya Sen'in yapabilirlik yaklaşımı dikkate alınarak UNDP tarafından ortaya konulan İGE'de, gelişmişlik düzeyi değerlendirilirken üç tane eşit öneme sahip temel değişken olan sağlık, eğitim ve gelir baz alınmıştır (Günsoy, 2005, s. 36).

İnsani gelişme için öncelikli olarak bu üç temel faktör elde edildiğinde, yaşamın diğer alanlarında da ilerleme ve gelişme fırsatlarında artış sağlanacaktır (Ünal, 2008, s. 90).

UNDP tarafından kullanılan bu üç ana faktörü değerlendirmek amacıyla kullanılan değişkenler, 1990 yılı başta olmak üzere düzenli olarak 2010 yılına kadar aynı olmakla birlikte hesaplama yönteminde zamanla değişiklikler meydana gelmiştir (Kocatepe Cebeci & Dilber, 2014, s. 125). İGE hesaplamasında kullanılan formül değişiklikleri sebebiyle, ülkelere ait verilerin yıllar itibarıyla karşılaştırılabilirliğinde farklılıklar olduğundan, bu farklılıkları gidermek için en son üzerinde karar verilen formül kullanılarak daha önceki dönem verileri beş senelik periyotlarla tekrar hesaplanıp yayımlanmaktadır (Demir Şeker, 2011, s. 4).

2.3. İnsani Gelişme Endeksinin Hesaplanmasında Kullanılan Göstergeler

İGE'de bir ülke için temel insani ölçütleri veren ve elde edilen verilerle bir ülkenin gelişmişlik düzeyi değerlendirilen temel faktörler, aşağıda hesaplama yöntemleriyle birlikte açıklanmıştır.

2.3.1. Sağlık Faktörü: Doğumda Yaşam Beklentisi Göstergesi

Uzun, sağlıklı ve yaratıcı bir hayatı kapsayan doğumda yaşam beklentisini (DYB) ifade etmektedir (Gülel, vd., 2017, s. 210). Doğumda yaşam beklentisi, bireyin dünyaya geldiği andan itibaren yaşaması beklenen ortalama süre olarak adlandırılmaktadır. Sağlıklı bir hayatın olduğu toplumlarda ortalama ömür artarken buna bağlı olarak eğitim seviyesi de yükselmektedir. Diğer bir ifade ile sağlık göstergelerinin yüksek olması, sağlıklı bir toplumda hem sağlık hizmetlerinin hem de yaşam kalitesinin iyi olduğu anlamına gelmektedir (Sakarya & İbişoğlu, 2015, s. 222).

Uzun ve sağlıklı bir yaşamı ifade eden endeks bireylerin yaşam beklentisini kullanarak eşitlik 2.1'deki denklem yardımıyla hesaplanır (Şahin & Gökdemir, 2016, s. 8).

$$I_H = \frac{L - L_{min}}{L_{max} - L_{min}} \quad (2.1)$$

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

2.3.2. Eğitim: Beklenen Okullaşma ve Ortalama Okullaşma Oranı

Toplumların bilgi ortaya koymalarını ve bilgiye ulaşma imkânlarını etkileyen ana faktörün, bireylerin eğitim seviyesi olduğu bilinmektedir. İGE hesaplamasında kullanılan eğitim endeksini hesaplamak için, ülkedeki beklenen okullaşma yılı (BOY) ve ortalama okullaşma yılını ifade eden (OYY) kullanılmaktadır. BOY, ülkedeki okula başlama yaşındaki çocukların göreceği beklenen eğitim süresini veya yaşa özgü okullaşma oranlarının, çocuğun yaşamı boyunca aynı kalacağı varsayımıyla çocuğun eğitim almayı bekleyebileceği toplam süre ile hesaplanmaktadır. OYY'un hesaplanmasında ise yetişkin nüfusun ortalama eğitim süresiyle veya diğer bir ifadeyle 25 yaş ve üzerindeki kişilerin hayatı süresince gördükleri eğitim süresi ile ölçülmektedir (Zor, 2019, s. 30; Ünal, 2008, s. 278).

Bilgiye erişimi ifade eden eğitim endeksi ortalama okullaşma yılları ve beklenen okullaşma yılları üzerinden eşitlik 2.2'deki denklem yardımıyla hesaplanır (Şahin & Gökdemir, 2016, s. 8).

$$I_E = \left(\frac{MYS - MYS_{min}}{MYS_{max} - MYS_{min}} + \frac{EYS - EYS_{min}}{EYS_{max} - EYS_{min}} \right) / 2 \quad (2.2)$$

Formüldeki MYS ortalama okullaşma yılını, EYS ise beklenen okullaşma yılını gösteren sembollerdir. Hesaplama MYS ve EYS 'nin eşit ağırlıklandırılmış ortalamaları alınır.

2.3.3. Gelir: Satıl Alma Gücü Paritesine İlişkin Kişi Başına Gelir Hesaplaması

Gelir, insana yaraşır bir hayat için gerekli olan tüm kaynaklara ulaşabilme imkânına sahip olmayı ifade eder (Tıraş, 2019, s. 18). Kaynaklara ulaşabilme imkânı hakkında net verilerin elde edilmemesi insani gelişmenin ölçülmesinde kişi başı gelirin kullanımını zorunlu kılmaktadır. Bireylerin belirli bir dönemde (genellikle 1 yıl) yaptığı ekonomik faaliyetler neticesinde tüm mal ve hizmetlerinin parasal değerlerinin karşılığına Gayri Safi Milli Hasıla (GSMH) denilmektedir.

Kaliteli bir yaşam standardını ifade eden endeks, GSMH üzerinden eşitlik 2.3'deki denklem yardımıyla hesaplanmaktadır (Şahin & Gökdemir, 2016, s. 18).

$$I_I = \frac{\ln(GNI) - \ln(GNI)_{min}}{\ln(GNI)_{max} - \ln(GNI)_{min}} \quad (2.3)$$

Formüldeki *GNI* logaritması alınmış satın alma gücü paritesine göre uyarlanan kişi başına GSMH'yi, $\ln(GNI)_{min}$ 100 değerine eşitliği temsil eden semboldür.

İGE, ölçüm parametreleri boyutlarının her biri için normalize edilmiş bu üç temel endeksin geometrik ortalaması alınarak hangi yıl hesaplanıyorsa o yıla ait insani gelişme endeksi değerine ulaşılır.

İGE sonucunu hesaplamak amacıyla endeksler üzerinde eşit ağırlıklandırma yapılır ve edinilen sonuçlar bir araya getirilir. Bu işlem eşitlik 2.4'deki denklem yardımıyla hesaplanır.

$$HDI = \sqrt[3]{I_H \times I_E \times I_I} = D(0 \text{ ile } 1 \text{ arasında}) \quad (2.4)$$

İGE'de ülkeler değerlendirilirken, insani gelişme alanındaki gösterdikleri performansa göre yukarıda belirtilen hesaplama yönteminden yararlanarak ülkelerin endeks değerleri 0 ile 1 arasında sıralanma yapılmakta; 0 minimum, 1 ise maksimum insani gelişme değerini ifade etmektedir. Her bir gösterge değeri için maksimum performansı gösteren ülkeye 1, minimum performansı gösteren ülkeye 0 değeri verilmekte, diğer ülkelerin değerleri ise bu iki sınır arasında açıklanmaktadır (Mıhçı & Mıhçı, 2003, s. 32).

Her yıl UNDP tarafından hesaplanarak yayımlanmakta olan İGE'de, ülkeler endeks değerine göre çok yüksek insani gelişim, yüksek insani gelişim, orta insani gelişim ve düşük insani gelişim olmak üzere dört kategoriye ayrılmaktadır (Kaya, 2017, s. 171).

UNDP, 2018 yılı İGE ve göstergelerinde yapılan değişiklikler sonucunda ülke gruplarının aldıkları puanlara göre sınıflandırılması Tablo 1'de gösterilmektedir. (Doğan & Tatlı, 2014, s. 105; Tıraş, 2019, s. 20).

Tablo 1

İnsani Gelişme Endeks Değerine Göre Yapılan Sınıflandırma

Endeks Değeri	Kategori Tanımı
0.800 ≥	Çok Yüksek İnsani Gelişim
0.700 – 0.799	Yüksek İnsani Gelişim
0.550 – 0.699	Orta İnsani Gelişim
<0,550	Düşük İnsani Gelişim

Kaynak: UNDP, 2018b, s. 3

İGE değerlendirilirken boyut endekslerinin oluşturulmasında birtakım minimum ve maksimum değerler belirlenmiştir. En son 2018 yılı İGE hesaplamasında kullanılan minimum ve maksimum değerler Tablo 2'de gösterilmektedir (Tıraş, 2019, s. 18).

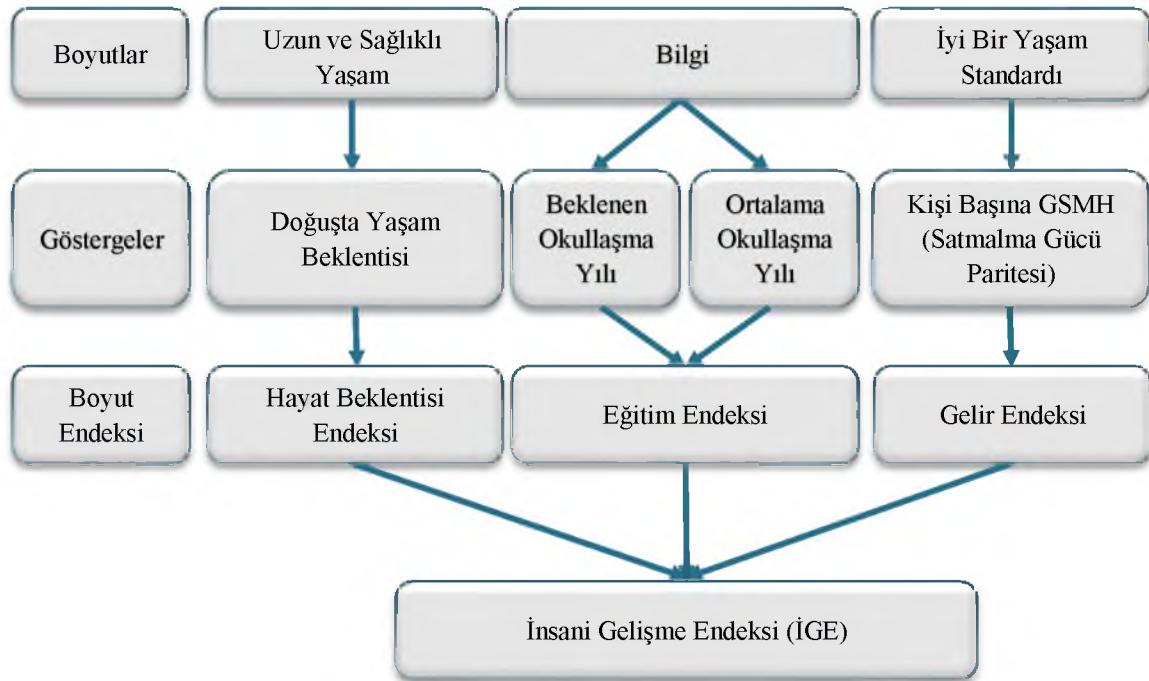
Tablo 2

2018 İGE Hesaplamasında Kullanılan Minimum ve Maksimum Değerler

Boyutlar	Göstergeler	Minimum	Maksimum
Sağlık	Yaşam Süresi (Beklenen)	20	85
	Okullaşma Yılı (Beklenen)	0	18
Eğitim	Okullaşma Yılı (Ortalama)	0	15
	Satılma Gücü Paritesi Kişi Başı GSMH (\$)	100	75.000

Kaynak: UNDP, 2018b, s. 2

İGE, üç boyutlu dört gösterge kullanılarak hesaplanmaktadır. Boyutlar hesaplanmasında sağlık (uzun ve sağlıklı yaşam), eğitim (bilgi) ve yaşam standardı (iyi bir yaşam standardı) dikkate alınmaktadır. İGE'yi oluşturan boyutlar, göstergeler ve boyut endeksinin oluşumu ayrıntılı olmak üzere Şekil 4'te gösterilmiştir.



Şekil 4. İnsani gelişme endeksi boyut, bileşenler ve oluşumu

Kaynak: UNDP B. , 2018, s. 2

Yayımlandığı yıllarda ortaya çıkan hadiseler dikkate alınarak bu hadiselerin insani gelişme üstündeki açılımına değinilen ve her yıl gelişmelere göre farklı bir tema ile çıkarılan İGR'de, 1990'da "İnsani Gelişim Kavramı ve Ölçülmesi" ana başlığı altında yayımlanmaya başlanılan raporlar yıllara ve temalarına göre değişiklik göstermektedir.

2.4. Toplumsal Cinsiyet Eşitsizliği Endeksi Tanımı

Toplumsal Cinsiyet Eşitsizliği Endeksi (TCEE) ilk defa 2010 yılında hesaplanmış ve cinsiyet eşitsizliğini; üreme sağlığı, kadının güçlendirilmesi ve ekonomik faaliyetler olmak üzere üç temel boyut üzerinde açıklanmaya çalışılmıştır.

Bu üç temel boyut: **üreme sağlığı**; anne ölüm ve ergen doğurganlık oranları, **kadının güçlendirilmesi**; bir ülke parlamentosundaki kadın milletvekili sayısı ile kadın ve erkeklerin orta ve yükseköğrenime devam etme oranları, **ekonomik faaliyetler**; bir ülkedeki kadın ve erkelerin iş gücüne katılımları ile değerlendirilmektedir.

Hesaplama Seth (2009), ortaya koyduğu yöntem ve geometrik ortalama metodu kullanılmaktadır. Endeks hesaplanmasında başlangıçta üç boyutun cinsiyet grubuna göre geometrik ortalamaları hesaplanır, sonrasında elde edilen verilerin (kadın ve erkekler ayrı ayrı olmak üzere) harmonik ortalamaları alınmak suretiyle endeks değeri elde edilmiş olur (Fırat & Aydın, 2015, s. 67-68; Rençber, 2018, s. 22-23).

BÖLÜM III

İNSANİ GELİŞME ENDEKSİ İLE İLGİLİ LİTERATÜR İNCELEMESİ

3.1. İGE Konulu Çalışmalar

İGE'nin, farklı sınıflandırma teknikleri ve algoritmaları yöntemleriyle modellenmesi, araştırmacılar tarafından kullanılan bir alan olarak karşımıza çıkmaktadır. Literatürde İGE ile ilgili, yerli ve yabancı araştırmalar tarafından yapılmış birçok çalışmanın yer aldığı görülmektedir.

Erdem & Çelik (2019), gerçekleştirmiş oldukları çalışmada 33 Afrika ülkesinin 1995-2014 yılları arasındaki verilerinden yararlanılarak insani gelişme ile ekonomik büyüme arasındaki ilişkiyi analiz etmişlerdir. Kişi başına sağlık harcamaları ile GSYH bağımsız değişken, İGE ise bağımlı değişken olarak değerlendirilmiştir. Çalışma sonucunda Afrika ülkelerinin, ekonomik büyümeyi sağlamak amacıyla yalnız geliri değil bunun yanında sağlık ve eğitim yatırımlarına da önem vermelerinin gereğine vurgulamışlardır. Siyasetçilere, ülkenin ekonomik büyüme ve kalkınma planlarını yaparken kısa dönem için gelir değişkenini, uzun vade de ise eğitim değişkenini baz alan politikaları devreye almaları konusunda önerilerde bulunmuşlardır.

Zanbak & Özeş Özgür (2019), yapmış oldukları çalışmada 1990-2017 yıllarında Türkiye'nin de içinde bulunduğu AB'ye aday ülkeler ile AB üyesi ülkelerin (28), Birleşmiş Milletler değerleriyle hesaplamış, İGE değerlerini karşılaştırmışlar ve AB üyeliğinin olası etkilerini değerlendirmişlerdir. Değerlendirme neticesinde kurucu üyelerin insani gelişme açısından günümüzde ileride olduğu; İngiltere, İrlanda, İsveç, Finlandiya ve Danimarka'nın da aynı seviyeye ulaştığının, AB üyeliğine kabul edilmelerine rağmen özellikle Romanya, Hırvatistan ve Bulgaristan'ın ise fazla bir gelişme göstermediği sonucunu vurgulamışlardır. Bunun yanında aday statüsünü devam ettiren Makedonya'nın tüm ülkeler arasında en düşük performansa sahip ülke olduğunu, Türkiye'nin ise geride olmasına rağmen en süratli değişimi gerçekleştiren ülkelerin başında geldiği değerlendirmesini yapmışlardır.

Çağlar & Keten (2018), çalışmalarında Türkiye'deki illerin 2013 yılındaki İGE'sini, Veri Zarflama Analizi (VZA) yöntemiyle hesaplanmasını öneren Mahlberg ve Obersteiner (2001)'in çalışmalarını baz alarak belirlemişlerdir. Söz konusu çalışmada veriler, UNDP'nin İGE göstergelerini temel alınarak 2013 yılındaki Türkiye'nin il düzeyinde İGE'sini hesaplayan Gülel vd. (2017) çalışmalarından temin etmişlerdir. İki yaklaşımın kullandıkları çalışmada birinci olarak VZA modellerinden çık yönlü CCR modeli, ikincisinde ise başka bir yaklaşımla İGE hesaplaması yapmışlardır. Değerlendirme neticesinde Türkiye'nin doğusunda kalan illerin her iki yaklaşımda da son sırada bulunduğu sonucunu açıklamaya çalışmışlardır.

Çoban & Yayar (2018), gerçekleştirmiş oldukları çalışmada AB ülkelerinde demokrasi göstergelerinin önemli bir sosyal refah göstergesi olarak kabul edilen insani gelişmişlik üzerine olan etkisini 2007-2014 yıllarını baz alarak, The Economist Endeksi, Uluslararası Mülkiyet Hakları Endeksi ve insani gelişmişlik göstergesi olarak da İGE'yi tercih etmişlerdir. Açıklayıcı değişkenler olarak ise tercihlerini ticari açlık oranı, işsizlik oranı, kamu sağlık harcamaları ve kamu eğitim harcamalarından yana kullanmışlardır. Dengeli panel veri analizi yöntemini kullanarak AB ülkelerinin demokrasinin göstergelerinin insani gelişmişlik üzerindeki etkisini test etmişlerdir. Yapılan çalışma sonucunda AB ülkelerinde demokrasinin göstergeleri ile insani gelişmişlik arasında istatistiksel olarak anlamlı bir ilişkinin olduğunu açıklamışlardır.

Meydan & Sarı (2018), yapmış oldukları çalışmada 1990 yılından itibaren UNDP tarafından yayımlanan İGE'nin, 2014 yılı verilerini temel alınarak Türkiye'nin iller bazında İGE ve alt endekslerin arasındaki ilişkiyi analiz etmişlerdir. Değerlendirme neticesinde İGE'nin Türkiye'nin doğusunda batısına göre daha düşük değer aldığını tespit etmişlerdir. Bunun yanında İGE ile eğitim ve gelir alt endeksleri arasında istatistiki olarak anlamlı bir ilişki olduğu sonucuna varmışlardır.

Tıraş & Ağır (2018), gerçekleştirmiş oldukları çalışmada 2016 yılı İGE değerleri ve alt grup göstergeleri baz alınarak Türkiye'nin de üyesi olduğu 57 İslam İşbirliği Teşkilatı ülkelerinin değerlendirmesini yapmışlardır. İGE açısından 188 ülke içerisinde 0,767 puanla 71. sırada bulunan Türkiye'nin, İslam İşbirliği Teşkilatı ülkeleri içerisinde ise 11. sırada yer aldığı vurgulamışlardır. Değerlendirme sonucunda Türkiye'nin eğitim ve sağlık alanlarındaki yatırımlarını artırmasıyla insani gelişim bakımından daha üst basamaklara yerleşeceği yönünde önerilerde bulunmuşlardır.

Kaya (2017), gerçekleştirmiş olduğu çalışmada finansal piyasalardaki gelişmelerin insani gelişmişlik üzerine etkisini belirlemek amacıyla, insani gelişmişlik düzeyini temsilen İGE'nin 1986-2015 dönemini kapsayan yıllık verilerini ve finansal piyasalardaki gelişmeleri temsilen Borsa İstanbul 100 Endeksi (BİST) büyüme oranını kullanarak zaman serisi analizini yapmıştır. Nedenselliğin belirlenmesi için Granger Nedensellik testi, değişkenler arasındaki uzun süreli ilişkilerin varlığını tespit etmek amacıyla Johansen-Juselius Eşbütünleşme testini uygulamıştır. Yapılan değerlendirme sonucunda değişkenlerin uzun süreli birlikte hareket ettikleri ve İGE endeksinin BİST endeksinin bir nedeni olduğu sonucunu açıklamıştır.

Rençber (2017), çalışmasında Çoklu Lojistik Regresyon, bulanık mantık yönteminin birleşimi olan ve hibrid öğrenme tekniğine dayanan ANFİS tekniklerinin sınıflandırma performanslarının karşılaştırmasını yapmıştır. Bu kapsamda Birleşmiş Milletler Dünya Gelişmişlik Göstergeleri ve ilgili sınıflandırma yöntemleri kullanılarak İGE'ye göre ülkeler sınıflandırmış ve elde edilen sonuçlar İGE ile karşılaştırmıştır. Yapılan değerlendirme sonucunda Adaptif Ağ Tabanlı Bulanık Çıkarım Sistemi yönteminin diğer sınıflandırma yöntemlerinden daha başarılı olduğu tespit etmiştir.

Öztürk (2016), İGE, ekonomik gelişmişlik ve beslenme ilişkisini değerlendirmiştir. Çalışmasında Türkiye, Norveç, Kore, İtalya, Yunanistan, Bulgaristan, İran ve Amerika'ya ait verileri karşılaştırmıştır. Karşılaştırmalar sonucunda Türkiye'nin insani gelişme ve beslenme yetersizliği açısından, İran ile benzer özellikler gösterdiğini ancak Avrupa ülkeleri, Kore ve Amerika'dan ayrıştığı tespit etmiştir. Ülkelerin karşılaştırılmasında ekonomik kalkınmışlık seviyesinin değil İGE çok daha doğru sonuçlar verdiği yargısına ulaşmıştır.

Yalçın & Çakmak (2016), yapmış oldukları çalışmada regresyon analizi yöntemiyle kamu sağlık harcamalarının insani gelişim düzeyi üzerine olan etkisini araştırmışlardır. Bu kapsamda 1991-2013 yıllarına ait yataklı sağlık kurumu, kamu sağlık harcamasının Gayri Safi Yurtiçi Hasılaya (GSYİH) oranı ve hekim başına düşen birey sayısı verilerini kullanılmışlardır. Çalışmada kamu sağlık harcamaları değişkeni ile insani gelişme arasında istatistiksel olarak olumlu yönde anlamlı bir ilişkinin olduğu tespit etmişlerdir.

Fırat & Aydın (2015), iki bölümden oluşan çalışmalarının birinci bölümünde kalkınma kavramını göstergeler ve kalkınmayı ölçme yönündeki endekslerle ele alarak teorik olarak tarihsel perspektifle açıklamışlardır. İkinci bölümünde ise İGE'nin alt endeksi olan Eğitim Endeksi'nin Türkiye göstergelerini, zaman dizisi kapsamında analiz

etmişlerdir. Eğitim alanındaki gelişmişlik seviyesinin tespiti amacıyla Türkiye'nin eğitim endeksi göstergeleri OECD ülkeleri ile karşılaştırmışlardır. Değerlendirme neticesinde Türkiye'nin, eğitim göstergelerinin OECD ülkelerine oranla düşük olduğu ve insani gelişme seviyesinin yeterli düzeyde olmadığı sonucuna ulaşmışlardır.

Murat & Gürsakal (2015), gerçekleştirmiş oldukları çalışmada ülkelerin gelişmişlik seviyelerinin belirlenmesinde sadece belirlenmiş makro-ekonomik göstergelerin değerlendirilmesinin yeterli olmadığı vurgulamışlardır. Yeni Ekonomik Kurumu (NEF) tarafından ilk defa 2006 yılında hesaplanan Mutlu Gezegen Endeksi (HPI) ile 150 ülke için hesaplanan İGE ve HPI arasındaki ilişkiler çok değişkenli istatistiksel yöntemlerden kanonik korelasyon analizi ile araştırılmışlardır. Çalışma neticesinde elde edilen ampirik bulgulardan İGE ve HPI arasında oldukça anlamlı ve güçlü bir kanonik ilişki olduğu tespit edilmiştir.

Paksoy (2015), gerçekleştirmiş olduğu çalışmada Avrupa ülkelerinin ve Türkiye'nin performanslarını ve gelişmişlik düzeylerini değerlendirmek amacıyla farkı kurumlar tarafından geliştirilen bileşik göstergelerin Vikor yöntemiyle birlikte analiz ederek değerlendirmiştir. Söz konusu çalışmanın genel olarak değerlendirilmesi sonucunda İsveç, Danimarka ve Hollanda gibi gelişmiş ülkelerin gelişmişliklerini pek çok alanda sağladıklarına vurgu yapılarak bunun sonucu olarak göstergeler ile bileşik göstergelerin bir arada değerlendirilmesinde de ilgili ülkelerin yüksek performansa sahip oldukları tespiti yapılmıştır.

Yakut, Gündüz & Demirci (2015), gerçekleştirmiş oldukları çalışmada UNDP İnsani Gelişme Raporunda yer alan 81 ülkenin, 2010-2012 yılları arasındaki İGE değerlerini kullanarak söz konusu ülkeleri çok yüksek, yüksek ve orta insani kalkınmışlık sınıfı olarak sınıflandırarak İGE sınıflandırma başarılarını lojistik regresyon ve yapay sinir ağları kullanarak karşılaştırılmasını yapmışlardır. Değerlendirme neticesinde sıralı lojistik regresyon modeli sonuçları incelendiğinde; sağlık harcamaları, ithalat ve ihracat, bebek ölüm oranı ile internet kullanıcı sayısı değişkenlerinin istatistik olarak anlamlı olduğunu tespit etmişlerdir. Sıralı Lojistik Regresyon analizi yönteminin % 88,1'lik, çok katmanlı Yapay Sinir Ağları analizi modelinin ise % 97,1'lik doğru sınıflandırma başarısı gösterdiği ilgili analiz yöntemlerinin karşılaştırmışlardır.

Akar (2014), Daha İyi Yaşam Endeks değerlerini kullanarak OECD ülkelerini karşılaştırmış ve OECD ülkeleri arasında Türkiye'nin daha iyi bir yaşam açısından son sırada bulunduğunu belirtmiştir. Gerçekleştirmiş olduğu çalışmada insani gelişmişliklerin yalnız GSYH ile değerlendirilmeyeceğini bireylerin sosyal (gelir, iş, sağlık, güvenlik vb.)

hayatlarının da deęişkenlerde yer alması gerektięini ifade etmiştir.

Doęan & Tatlı (2014), yapmış oldukları çalışmada insani gelişme ve insani yoksulluk kavramlarını kullanarak insani gelişme ve yoksulluk endekslerinin temel göstergelerinin nelerden oluştuęunu ve nasıl hesaplandığını inceleyerek İGE ve dięer endeksler açısından Türkiye'nin performansını karşılaştırmayı amaçlamışlardır. Yapılan deęerlendirme neticesinde Türkiye'nin yıllar itibariyle insani gelişmesini artırmış olmasına karşılık pek çok gelişmiş ülkenin gerisinde kaldığı sonucunu açıklayıp Türkiye için öneride bulunmuşlardır.

Kocaefe Cebeci & Dilber (2014), gerçekleştirmiş oldukları çalışmada büyüme ve kalkınma arasındaki farkı kavramsal olarak irdeleyerek insani gelişme kavramının meydana çıkışına deęinmişler ve İGE'yi oluşturan alt bileşenlerin 2011 ve 2012 yılı deęerlerini karşılaştırarak endeks hesaplaması yapmışlardır. Buna göre 0,699 İGE ile 2011 yılında 187 ülke içerisinde 92'nci sırada olan Türkiye'nin, 2012 yılında doğumda beklenen yaşam süresinde, birey başına düşen gelirden ve beklenen okullaşma yaşındaki artışlar sonucunda 0,722 İGE deęeri ile 90'ıncı sıralamada yer aldığını tespit etmişlerdir.

Çemrek & Bayraç (2013), gerçekleştirmiş oldukları çalışmada Temel Bileşenler Analizi yöntemi ile sürdürülebilir kalkınma göstergeleri kullanılarak Bağımsız Devletler Topluluęu, Baltık Cumhuriyetleri ve Rusya için sürdürülebilir kalkınma skoru ile İGE deęerleri arasındaki ilişkiyi karşılaştırmışlardır. Yapılan deęerlendirme sonucunda sürdürülebilir kalkınma skoru ile İGE arasındaki korelasyon katsayısı 0,872 olarak hesaplanmış ve ilgili deęişkenler arasındaki ilişki %1 önem düzeyinde tespit edilmesi sonucunda istatistiksel olarak anlamlı bulmuşlardır.

Kazar & Kazar (2013), gerçekleştirmiş oldukları çalışmada UNDP tarafından oluşturulan İGE yeni ölçüm sistemi ve eşitsizliklerinde deęerlendirildięi eşitsizlikle uyumlandırılmış İGE ve alt gösterge deęerlerinin gelişmekte ve gelişmiş bulunan ülkelerle karşılaştırılarak Türkiye'nin performansının deęerlendirilmesini amaçlamışlardır. Deęerlendirme neticesinde Türkiye'nin 1980-2011 döneminde endeks deęerlerini iyileştirip dünya ülkeleri ortalamasına ulaşmasına rağmen göstermiş olduęu performansın gelişmiş ülkelere göre yetersiz kaldığı sonucuna ulaşmışlardır. Türkiye'nin gelir ve eğitim eşitsizliği nedeniyle potansiyel endeks deęeri ile performansı arasında meydana gelen kaybın bu yetersizlikte önemli rol oynadığını belirtmişlerdir. Sonuç olarak Türkiye'nin gelişmiş ülkelerin potansiyeline ulaşabilmesi için öncelikle eğitim, saęlık ve gelir dağılımındaki eşitsizliklerin giderilmesi gerektięini vurgulamışlardır.

Korkmaz & Şahin (2013), gerçekleştirmiş oldukları çalışmada betimsel olarak doküman analizi modeli kullanılarak insani gelişme endeks değerine göre yapılan sınıflandırmaya göre ülkeleri ve insani gelişim seviyeleri arasında nasıl bir farkın olduğunu incelemişlerdir. Türkiye'nin, başarı durumlarına göre genel kapsamında PISA sınavından başarı gösteren ülkelerin, genel ve insani gelişim düzeylerinin üst seviyede olduğu, Türkiye'nin ise başarılı ülkeler seviyesinin gerisinde kaldığını açıklamışlardır. Verilerin incelenmesi sonucunda başarı ile insani gelişmişlik arasında olumlu bir ilişkinin olduğu tespit etmişlerdir.

Nartgün, Kösterelioğlu & Sipahioğlu (2013), çalışmalarında Avrupa Birliği'ne (AB) üye ülkeler ile adaylık süreçleri devam eden ülkelerin İGE, eğitim indeksi ve GSYİH verilerine göre, AB'ye aday olan Türkiye'nin kalkınma düzeyinin bu ülkeler içerisindeki durumunu incelemişlerdir. Yapılan çalışma neticesinde Türkiye'nin hem İGE değerinin hem de eğitim indeksi değerinin bu ülkelerin değerlerinden düşük olduğu açıklamışlardır.

Tireli, Coşkun & Kunduracı (2013), çalışmada İslam İşbirliği Teşkilatının (İİT) bir alt kuruluşu olan SESRIC tarafından oluşturulan üye ülkelerin eğitim, sağlık ve istihdam verileri ile İGE verileri arasındaki bağlantı incelenmişlerdir. İncelenme neticesinde yoksullukla; sağlık, eğitim ve istihdam arasındaki ilişki ve etkileşim seviyesi, İİT ve UNDP kuruluşlarının verileri temel alınarak değerlendirilmiş ve İİT üyesi ülkelerin toplumsal ve beşeri kalkınmasını sağlayabilmesi için yoksullukla mücadelede; sağlık, istihdam ve eğitim alanında daha fazla çalışması gerektiği sonucuna varmışlardır.

Arpacıoğlu & Yıldırım (2011), gerçekleştirmiş oldukları çalışmada ülkelerin İGE seviyelerine göre sınıflandırılmasında ekonomik ve sosyal değişkenlerin her ikisinin de birlikte değerlendirilmesi sonucunda ülkelerdeki yoksulluk kavramının dikkate alınmadığını açıklamışlardır. Çalışmada ülkelerin İGE'ye göre buldukları sınıflar ile yoksulluk düzeyleri karşılaştırılarak Türkiye için tavsiyelerde bulunmuşlardır. Öneri olarak yoksullukla mücadele edilmesini ve bununla ilgili destek projeleri yapılması gerektiği vurgulanmıştır.

Erol (2011), gerçekleştirmiş olduğu çalışmada Temel Bileşenler Analizi ile gelişmekte olan bazı ülkelerle Türkiye'nin insani gelişme kriterlerini incelemiştir. Çalışma kapsamında değerlendirilmeye alınan 23 gelişmekte olan ülkenin analiz sonuçlarıyla UNDP raporları arasında herhangi bir farklılık bulunmadığını tespit etmiştir. Sonuç olarak ilgili ülkelerin ve özellikle Türkiye'nin, gelişme stratejisi politikalarını benimseyerek uygulaması neticesinde ilerleyen yıllarda bu ülkelerinde refah seviyesinde

ve İGE değerlerinde ivme kazanarak yükselme olacağı değerlendirilmiştir.

Öngel, Sözen & Çelik (2011), Orta Asya Ülkeleri olan Kazakistan, Türkmenistan, Özbekistan, Kırgızistan ve Tacikistan'ın İGE sonuçlarının 2010 yılındaki bulunduğu konumu ile söz konusu ülkelerin bağımsız olduktan sonraki dönemde ortaya çıkan değişimleri yıllar itibari ile karşılaştırmışlardır. Sonuç olarak bağımsızlık sonrası 20 yıllık süre içerisinde GSYİH'de meydana gelen yüksek oranlı artışların İGE'ye yansımadağı ifade etmişlerdir. Orta Asya ülkeleri içerisindeki diğer ülkelere oranla Kazakistan'ın İGE değerinin daha çok arttığı yargısına ulaşmışlardır.

Çalışkan (2010), çalışmasında Türkiye'de beşeri sermayenin temel iki bileşeni olan sağlık ve eğitim amacıyla yapılan harcamaların analizini yapmıştır. Ülkelerin sağlık ve eğitim harcamalarının GSYH'ye oranını İGE'ye göre hesaplamıştır. Yapılan değerlendirme sonucunda Türkiye'nin eğitim ve sağlık harcamalarının kişi başına miktarının GSYH'ye oranı gelişmiş ülkelerle karşılaştırıldığında yetersiz kaldığı ve bütçeden ayrılan harcamaların azlığı nedeniyle, insani gelişmişliğin özellikle eğitim ve sağlık boyutuna ilişkin göstergelerinde sorunların bulunduğu tespit etmiştir. Kamu ve özel sektörün eğitim ve sağlık alanlarına bütçeden daha fazla kaynak aktarmaları durumunda Türkiye'nin gelişmişlik seviyesinin yükseltilmesinde önemli katkılar sağlayacağı önerisinde bulunmuştur.

Gürses (2009), gerçekleştirmiş olduğu çalışmada, Türkiye'nin insani gelişme bakımından milletlerarası konumu ile yıllar itibarıyla ortaya koyduğu performansı değerlendirmeyi amaçlamıştır. Yapılan değerlendirme neticesinde; Türkiye'nin, insani gelişmişlik raporu verileri neticesinde düzenli fakat dengesiz bir gelişme eğilimi içerisinde olduğu, aynı kategorideki veya daha aşağı gelir grubundaki ülkelere kıyasla eğitim (eğitim ve okullaşma) endeks değerlerinin düşük olduğu sonucuna varılmıştır. Bununla birlikte Türkiye'nin gelir göstergesinin sağlam lakin bilgi göstergesinin zayıf olduğu tespit etmiştir.

Karabulut, Kaya & Gürsoy (2009), gerçekleştirmiş oldukları çalışmada 2006 yılı İGE değerlerine göre OECD ülkelerinin insani gelişmişlik düzeylerini analiz etmişlerdir. Yapılan değerlendirme sonucunda, Türkiye dışındaki OECD ülkelerinin tamamının yüksek insani gelişmişlik seviyesine sahip ülkeler kategorisinde bulunduğu Türkiye'nin ise eğitim endeksi dışındaki diğer endeks verilerine göre orta insani gelişmişlik seviyesine sahip ülkeler kategorisinde bulunduğu tespit etmişlerdir.

Ünal (2008), yapmış olduğu çalışmada Türkiye İstatistik Kurumu (TÜİK) tarafından oluşturulan ulusal ve uluslararası seçilmiş ekonomik göstergeler yardımıyla

Türkiye'nin, hem gelişimi hem de AB ülkeleri ve AB'ye aday ülkeler arasındaki konumunun belirlenmesini amaçlamıştır. Değerlendirme neticesinde Türkiye'nin, İGE'de önemli gelişmeler kaydetmiş olmasına rağmen birçok endeksin değerlendirilmesi sonucunda sadece AB'ye üye ülkeler değil üyeliğe aday ülkelerin de gerisinde kaldığı sonucuna ulaşmıştır.

Erkekoğlu (2007), çalışmasında kümeleme yöntemi ile Türkiye ile AB ülkelerini sosyo-ekonomik değişkenler açısından karşılaştırmıştır. AB ülkeleri, Romanya ve Bulgaristan'ın da içinde bulunduğu yirmi sekiz ülke içerisinde oluşturulan kümelerden Türkiye'nin hangisinde yer aldığı tespit edilmesi amaçlamıştır. Kümeleme analizi neticesinde beş gruba ayrılan ülkelere aynı gelişmişlik düzeyine sahip Türkiye, Letonya, Bulgaristan, Litvanya, Polonya, Romanya ve Bulgaristan'ın aynı kümede bulunduğu sonucuna varmıştır. Oluşturulan tüm grup ortalamaları ile Türkiye göstergeleri karşılaştırıldığında en büyük farklılığın kişi başına düşen GSMH olduğunu açıklamıştır.

Demir (2006), gerçekleştirmiş olduğu çalışmada Türkiye'nin İGE değerlerindeki gelişmeler ile İGE'nin yıllar itibarıyla geçirdiği değişimlerin nedenleri ve ileride Türkiye İGE değerine yönelik tahminleme yapmıştır. Düzenlenen raporda Türkiye'nin, İGE değerinde yıllar itibarıyla artış meydana gelmesine rağmen İGE sıralamasında düzenli bir yükseliş gözlenmediği belirtilmiştir. Türkiye'nin 2003 yılı verilerine göre İGE değerindeki artışın nedeni olarak kişi başı GSYİH endeksi gösterilmiş, diğer alt endekslere de gereken önemin verilmesi neticesinde daha sonraki yıllarda da İGE değerinde artışın olabileceği değerlendirilmesini yapmıştır.

Tüylüoğlu & Karalı (2006), yapmış oldukları çalışmada İGE'yi sadece Türkiye açısından değerlendirerek bağımsız olarak hesaplamışlar ve UNDP'nin hesaplamalarına oranla daha düzenli olduğunu tespit etmelerine rağmen Türkiye'nin orta insani kalkınma düzeyinde yer aldığını belirtmişlerdir. Türkiye'de iller düzeyinde UNDP standartlarında resmi verilerin oluşturulması sonucunda, hesaplamaların daha sağlıklı bir şekilde yapılabileceği ve sonuç odaklı başarılı bir insani gelişme politikasının uygulanabileceği değerlendirmesinde bulunmuşlardır.

Günsoy (2005), gerçekleştirmiş olduğu çalışmada Türkiye ve aynı kategorideki 14 ülkeyi; gelişmişlik düzeyi, sağlık harcamaları ve bireylerin sağlık durumları itibarıyla karşılaştırmıştır. Yapılan değerlendirme sonucunda, sağlıkla ilgili göstergeleri nedeniyle Türkiye'nin, benzer ülkelere kıyasla insani gelişme açısından dezavantajlı durumda olduğu ve sağlık politikalarının tekrar gözden geçirilmesi gerektiği ifade etmiştir.

BÖLÜM IV

ARAŞTIRMA BULGULARI VE DEĞERLENDİRME

4.1. Metodoloji

Bu çalışmada; UNDP tarafından hazırlanan, her yıl yayımlanan insani gelişme raporunda yer alan İGE ve alt göstergelerinden yararlanılarak, veri madenciliğinin karar ağacı tekniklerinden olan C5.0 ve GİNİ algoritmaları ile karar ağaçları modellenerek insani gelişme endeksine etki eden faktörler belirlenmiş ve sınıflandırma tahmini gerçekleştirilmiştir. İGE ve alt endeks verilerinin analizinde yöntem olarak VM tekniklerinden olan C5.0 ve GİNİ algoritmaları kullanılmıştır. Ülkeler gelişmişlik sınıflarına göre çok yüksek, yüksek, orta ve düşük gelişmiş ülkeler olmak üzere 1-4 arasında değerlere atanarak İGE'ye ait kategorik değişkenler oluşturulmuştur. Verilerin gerekli düzenleme işlemleri tamamlandıktan sonra excel formatında kaydedilerek analiz için hazır hale getirilmiştir.

4.2. Çalışmanın Amacı

Bu tez çalışmasının amacı, UNDP tarafından her yıl yayımlanan insani gelişme raporunda bulunan kategorik değişken olan insani gelişme endeksini etki eden faktörlerin, veri madenciliğinin karar ağacı tekniklerinden olan C5.0 ve GİNİ algoritmalarını kullanarak modellenmesinin yapılmasıdır.

4.3. Çalışmanın Önemi

İnsani gelişme endeksi ülkelerin gelişmişliklerini göz önünde bulundurarak insanların mutluluğunu, sağlıklı bir yaşam ile birlikte başarılı bir hayat sürmelerini dikkate alan bir kalkınma endeksidir. Ekonomik kalkınmanın aksine insani gelişme endeksi 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 ülkeler arası karşılaştırma yapmak için başvuru bir

gösterge değeri olmuştur. Çalışmada veri madenciliğinin karar ağaçları analizleriyle İGE literatürüne katkı sağlayacağı düşünülmektedir.

4.4. Çalışmanın Kapsamı

Çalışma, UNDP tarafından hazırlanan her yıl yayımlanan insani gelişme raporunda yer alan 189 ülkenin eksik verileri çıkarılarak, geriye kalan 79 ülke çalışma kapsamına dâhil edilmiştir. Veri setine ilişkin olarak yıllar itibari ile farklı değişkenlerin oluşması ve bazı değişkenlerin bütün ülkelerde açıklanmaması nedeniyle 110 ülke çalışma kapsamının dışına alınmıştır. Böylece ülkeler çok gelişmiş, gelişmiş, orta ve düşük düzey gelişmiş ülkeler olmak üzere kategorileştirilmiştir. Tablo 3'te çalışma kapsamındaki ülkelerin 2010-2017 dönemine ait gelişmişlik düzeyleri gösterilmiştir

Tablo 3

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

Çalışma kapsamındaki ülkelerin gelişme düzeylerine göre 2010 ve 2017 yıllarındaki durumlarının belirlendiği Tablo 3'e göre, 2010 yılı verilerinde düşük düzey gelişmiş ülkeler kategorisinde bulunan Bangladeş, Kenya, Zambiya ve Pakistan'ın 2017 yılında bir üst kategori olan orta düzey gelişmiş ülkeler düzeyinde yer aldığı görülmektedir. 2010 yılında orta düzey gelişmiş ülke seviyesinde bulunan Moğolistan, Maldivler, Belize, Paraguay ve Moldova'nın 2017 yılında gelişmiş ülkeler kategorisine yükseldiği gözlemlenmektedir. 2010 yılında gelişmiş ülke seviyesinde bulunan Rusya, Karadağ, Bulgaristan, Romanya, Belarus, Uruguay ve Malezya'nın 2017 yılında çok gelişmiş ülkeler kategorisine yükseldiği, 2010 yılında gelişmiş ülke düzeyinde bulunan Türkiye'nin ise 2017 yılında da yerini koruyarak yine gelişmiş ülke düzeyinde yer aldığı görülmektedir. Tablo 4'te ülkelerin 2010-2017 dönemine ait gelişmişlik düzeylerine ait sayılar ve yüzdeler dilimler gösterilmiştir.

Tablo 4

Gelişme Düzeylerine Göre Çalışma Kapsamındaki Ülkelerin Sayısı ve Yüzdesi

Sınıf Kodu	Gelişmişlik Düzeyi	Ülke Sayısı		Yüzde (%)	
		2010	2017	2010	2017
1	Çok Yüksek İnsani Gelişim	31	38	39,24	48,10
2	Yüksek İnsani Gelişim	24	22	30,38	27,85
3	Orta İnsani Gelişim	16	15	20,25	18,99
4	Düşük İnsani Gelişim	8	4	10,13	5,06

Gelişme düzeylerine göre çalışma kapsamındaki ülkelerin 2010 ve 2017 yıllarının karşılaştırıldığı Tablo 4'e göre, 2017 yılındaki çok yüksek insani gelişim sayısında 2010 yılına göre artış olduğu söylenebilir. Yüksek insani gelişim, orta insani gelişim ve düşük insani gelişim düzeylerinde ise olumlu olarak bazı ülkelerin çok yüksek insani gelişim düzeyine geçmesi nedeniyle 2017 yılındaki ülke sayısında 2010 yılına göre azalma olduğu görülmektedir. Tablo 5'te ülkelerin kıtalara göre sayı ve yüzde değerleri verilmiştir.

Tablo 5

Kıtalara Göre Çalışma Kapsamındaki Ülkelerin Sayısı ve Yüzdesi

Kıta	Ülke Sayısı	Yüzde (%)
Avrupa	36	45,57
Asya	19	24,05
Kuzey Amerika	8	10,13
Afrika	8	10,13
Güney Amerika	7	8,86
Okyanusya	1	1,27
Toplam	79	100

Çalışma kapsamındaki ülkelerin kıtalara göre ayrıştırıldığı Tablo 5 incelendiğinde ülkelerin büyük çoğunluğunun Avrupa ve Asya kıtasında bulunduğu gözlemlenirken diğer kıtalardaki ülkelerin daha az sayıda bulunduğu görülmektedir. Kuzey ve Güney Amerika, Afrika ve Okyanusya kıtalarındaki ülke sayılarının fazla olmamasında bu kıtalarda bulunan ülke sayılarının az olmasının etkili olduğu söylenebilir. Tablo 6'da ülkelerin gelişmişlik düzeyleri ve kıtalara göre sayı ve yüzde değerleri verilmiştir.

Tablo 6

Gelişmişlik Düzeylerine ve Kıtalara Göre Ülke Sayıları ve Yüzdesi

Gelişmişlik Düzeyi	Kıta	Ülke Sayısı		Yüzde (%)	
		2010	2017	2010	2017
Çok Yüksek İnsani Gelişim	Avrupa	27	32	87,10	84,21
	Asya	2	3	6,45	7,89
	Kuzey Amerika	1	1	3,23	2,63
	Güney Amerika	0	1	0,00	2,63
	Okyanusya	1	1	3,23	2,63
	Toplam		31	38	100,00
Yüksek İnsani Gelişim	Avrupa	8	4	33,33	18,18
	Asya	8	9	33,33	40,91
	Kuzey Amerika	3	4	12,50	18,18
	Güney Amerika	5	5	20,83	22,73
	Toplam		24	22	100,00
Orta İnsani Gelişim	Avrupa	1	0	6,25	0,00
	Asya	7	7	43,75	46,67
	Kuzey Amerika	4	3	25,00	20,00
	Afrika	2	4	12,50	26,67
	Güney Amerika	2	1	12,50	6,67
	Toplam		16	15	100,00
Düşük İnsani Gelişim	Asya	2	0	25,00	0,00
	Afrika	6	4	75,00	100,00
	Toplam		8	4	100,00

Gelişmişlik düzeylerine ve kıtalara göre ülke sayılarının ve yüzdelerinin verildiği Tablo 6'ya göre çok yüksek insani gelişim dâhilindeki kıtalar içerisinde ülkelerin büyük çoğunluğunun Avrupa kıtasında bulunduğu ve bunun 2010 ve 2017 yıllarında değişmediği görülmektedir. Yüksek insani gelişim kategorisinde çoğunluk ülkelerin Avrupa ve Asya kıtalarında bulunduğu, orta insani gelişim düzeyinde bulunan ülkelerin ise çoğunluk olarak Asya ve Kuzey Amerika kıtasında olduğu söylenebilir. Düşük insani gelişim kategorisinde sadece Asya ve Afrika kıtalarının olduğu gözlemlenirken bu durum kıtalar arasında gelişmişlik kapsamında çok büyük farklılıkların olduğunun da göstergesidir. Okyanusya kıtasının sadece çok yüksek insani gelişim kategorisinde bulunmasında bu kıtada bulunan ülke sayısının az olması etkili olmuştur.

4.5. Veri Setine Ait Değişkenlerin Seçimi

Bu çalışmada, UNDP'nin İnsani Kalkınma Programının internet sitesinden 79 ülkenin 2010-2017 yıllarını kapsayan sekiz yılına ait veriler kullanılarak çalışmanın veri seti oluşturulmuştur. Kategorik değişken olarak insani gelişme endeksine etki eden faktörlerin belirlenmesi ve karar ağaçlarının oluşturulması için eğitim, sağlık, ekonomi, istihdam ve iletişim faktörlerinden toplam 15 bağımsız değişken belirlenmiştir. Tablo 7'de analizlerde kullanılacak olan değişkenler verilmiştir.

Tablo 7

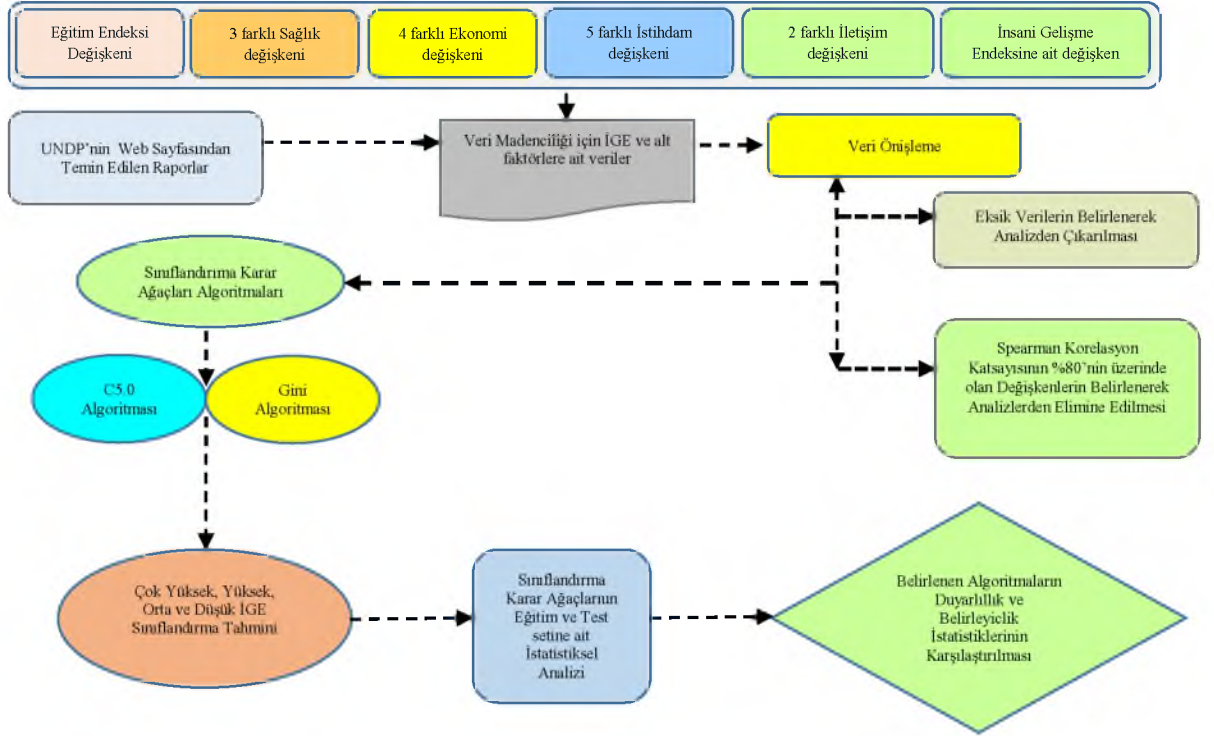
İGE, Alt Endeksler ve Faktörlere İlişkin Değişkenler

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'nın %'si)
X8	Foreign direct investment, net inflows (% of GDP)	Doğrudan yabancı yatırımlar, net girişler (GSYİH'nın %'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)

4.6. Verilerin Analiz Yöntemi

Çalışmanın bu kısmında Şekil 5'te görüldüğü üzere karar ağaçlarının analizinde kullanılmak üzere veri ön işleme gerçekleştirilmiş, eksik olan veriler analiz kapsamından çıkarılmıştır. Şekil 5'te görüldüğü gibi araştırmanın kavramsal yapısına ilişkin olarak 189 ülkenin 110'unun eksik verileri ve yıllar itibari ile bazı değişkenlerin bütün ülkelerde açıklanmaması sebebiyle analiz kapsamına alınmamıştır. Geri kalan 79 ülkenin değişkenleri arasında spearman korelasyon analizi gerçekleştirilerek değişken sayısı azaltılmıştır. Böylece veri ön işleme gerçekleştirildikten sonra 79 ülkenin toplam 632 veri kaydı karar ağaçları algoritmalarında kullanılmak üzere hazır hale getirilmiştir. C5.0 algoritması ve Gini algoritması yardımıyla, hazır hale getirilen veriler ile karar ağacı

modelleri oluşturularak kategorik değişken olan İGE'ye ait kurallar belirlenmiştir. Verilerin %75'i ile modeller eğitilmiş, geriye kalan %25'lik veriler ile test işlemi gerçekleştirilmiştir. Bu aşamadan sonra sınıflandırma algoritmalarının belirleyicilik "specificity" ve duyarlılık "sensitivity" istatistikleri karşılaştırılmak üzere incelenmiştir. Araştırmanın kapsamına ilişkin kavramsal yapı Şekil 5'te gösterilmiştir.



Şekil 5. Araştırmanın kavramsal yapısı

4.7. Spearman Korelasyon Analizine İlişkin Sonuçlar

Spearman korelasyon analizi değişkenlerin normal dağılmadığı durumlara ilişkin olarak kullanılmaktadır. Spearman korelasyon katsayısı Pearson'ın sıralı verilerle kullanılmak üzere belirlenmiş parametrik olmayan bir yöntemidir. Spearman korelasyon katsayısı -1 ile +1 arasında değer almakta olup, +1 olduğunda değişkenler arasında pozitif yönlü çok kuvvetli bir ilişkinin, -1 ise değişkenler arasında negatif yönlü çok kuvvetli bir doğrusal ilişkinin olduğunu açıklamaktadır. Spearman korelasyon katsayısının 0 olması durumunda, değişkenler arasında doğrusal bir ilişkin olmadığını ifade etmektedir (Kalaycı, 2008, s. 116).

Tablo 8

Değişkenlere ait Spearman Korelasyon Analizi Sonuçları

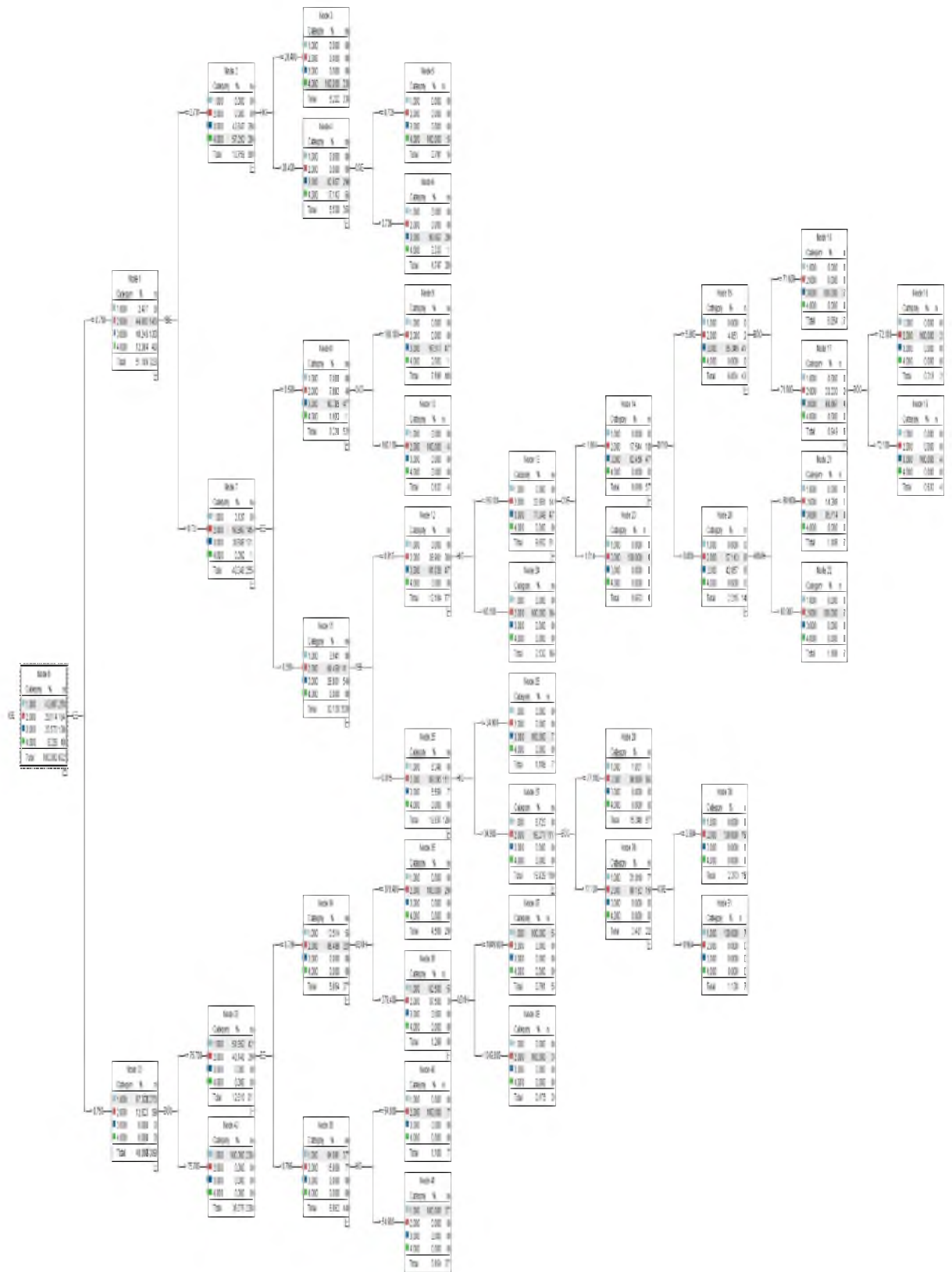
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
X1	1.000	.720**	.244*	.720**	.860**	-.890**	.429**	.522**	-.780**	.167**	.866**	.380**	.704**	.920**	-.038
X2	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.343
X3	.720**	1.000	.357**	1.000**	.858**	-.843**	.248**	.257**	-.743**	.101*	.811**	.350**	.730**	.805**	-.009
X4	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.011	.000	.000	.000	.000	.824
X5	.244*	.357**	1.000	.357**	.443*	-.279**	-.330**	.069	-.332**	-.178*	.331**	.079	.244*	.353**	-.318*
X6	.000	.000	.000	.000	1.000	.000	.000	.081	.000	.000	.000	.185	.000	.000	.000
X7	.720**	1.000**	.357**	1.000**	.858**	-.843**	.247**	.257**	-.744**	.102*	.812**	.350**	.731**	.805**	-.009
X8	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.011	.000	.000	.000	.000	.819
X9	.860**	.858**	.443*	.858**	1.000	-.887**	.302**	.433**	-.885**	.075	.900**	.405**	.828**	.892**	-.091*
X10	.000	.000	.000	.000	.000	.000	.000	.000	.000	.060	.000	.000	.000	.000	.022
X11	-.890**	-.843**	-.279**	-.843**	-.887**	1.000	-.401**	-.471**	.761**	-.115*	-.865**	-.389**	-.710**	-.874**	.045
X12	.000	.000	.000	.000	.000	.000	.000	.000	.000	.004	.000	.000	.000	.000	.260
X13	.429**	.248**	-.330**	.247**	.302**	-.401**	1.000	.343**	-.339**	.064	.366**	.375**	.267**	.286**	.327**
X14	.000	.000	.000	.000	.000	.000	.000	.000	.110	.000	.000	.000	.000	.000	.000
X15	.522**	.257**	.069	.257**	.433**	-.471**	.343**	1.000	-.417**	.124**	.439**	.401**	.383**	.499**	.023
X1	.000	.000	.081	.000	.000	.000	.000	.000	.000	.002	.000	.000	.000	.000	.563
X2	-.780**	-.743**	-.332**	-.744**	-.885**	.761**	-.359**	-.417**	1.000	-.120**	-.810**	-.430**	-.921**	-.788**	.073
X3	.000	.000	.000	.000	.000	.000	.000	.000	.000	.003	.000	.000	.000	.000	.068
X4	.167**	.101*	-.178*	.102*	.075	-.115*	.064	.124**	-.120**	1.000	.106**	.105**	.170**	.179**	.072
X5	.000	.011	.000	.011	.060	.004	.110	.002	.003	.000	.008	.008	.000	.000	.072
X6	.866**	.811**	.331**	.812**	.900**	-.865**	.366**	.459**	-.810**	.106**	1.000	.390**	.768**	.841**	-.051
X7	.000	.000	.000	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.204
X8	.380**	.350**	.053	.350**	.405**	-.389**	.375**	.401**	-.430**	.105**	.390**	1.000	.387**	.363**	.082*
X9	.000	.000	.185	.000	.000	.000	.000	.000	.000	.008	.000	.000	.000	.000	.040
X10	.704**	.730**	.244*	.731**	.828**	-.710**	.267**	.383**	-.921**	.170**	.768**	.387**	1.000	.749**	-.014
X11	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.718
X12	.920**	.805**	.353**	.805**	.892**	-.874**	.286**	.499**	-.788**	.179**	.841**	.363**	.749**	1.000	-.059
X13	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.141
X14	-.038	-.009	-.318**	-.009	-.091*	.045	.327**	.023	.073	.072	-.051	.082*	-.014	-.059	1.000
X15	.343	.824	.000	.819	.022	.260	.000	.563	.068	.072	.204	.040	.718	.141	

** . Korelasyon katsayısı 0.01 önem seviyesinde anlamlıdır. * . Korelasyon katsayısı 0.05 önem seviyesinde anlamlıdır.

Tablo 8’de değişkenlere ait Spearman korelasyon analiz sonuçlarına yer verilmiştir. C5.0 ve Gini algoritmalarında kullanılacak olan bağımsız değişkenlerin normallik varsayımı sağlanmadığından spearman korelasyon analizi gerçekleştirilerek, %80’in üzerinde korelasyon katsayısına sahip olan değişkenler karar ağacı analizlerine dahil edilmeyecektir (Alpar, 2013). Sırasıyla 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”na ait değişkenlerdir.

4.8. C5.0 Algoritması ile Karar Ağacı Modellemesi

İGE’ye ait kategorik değişkene ilişkin karar ağacının modellenmesinde entropi tabanlı C5.0 algoritması kullanılmıştır. Şekil 6’da C5.0 algoritmasına ait karar ağacı gösterilmiştir. Karar ağacı modelinde 21 kural oluşturulmuştur. Oluşturulan kurallar Tablo 9’da gösterilmiştir. Şekil 6’da görüldüğü üzere karar ağacının modelinde İGE’nin sınıflandırılmasına etki eden eğitim endeksi %43,1’lik oran ile en kuvvetli etkiye sahip, ardından %29,5’lik oran ile ikinci en kuvvetli etkiye sahip olan değişkenin hizmet sektöründeki istihdam oranı değişkeni olduğu tespit edilmiştir.



Şekil 6. C5.0 Algoritması ile İGE'nin Kategorik olarak Sınıflandırılmasına ait Karar Ağacı Modeli

Tablo 9

Karar Ağacı Modeli için C5.0 Algoritmasından Elde Edilen Kurallar

Kural No	Eğer Kuralı	%	Ülke Sayısı	Kategorik Değişkenler
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

Tablo 9’da C5.0 algoritması ile elde edilen karar ağacı modeline ilişkin toplam 21 kural verilmiştir. Oluşturulan kurallar içerisinde en önemlileri çok yüksek İGE için 8, yüksek İGE için 14, orta İGE için 4, düşük İGE için 1 no’lu kuraldır.

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 İGE %100 oran ile düşük İGE’dir.

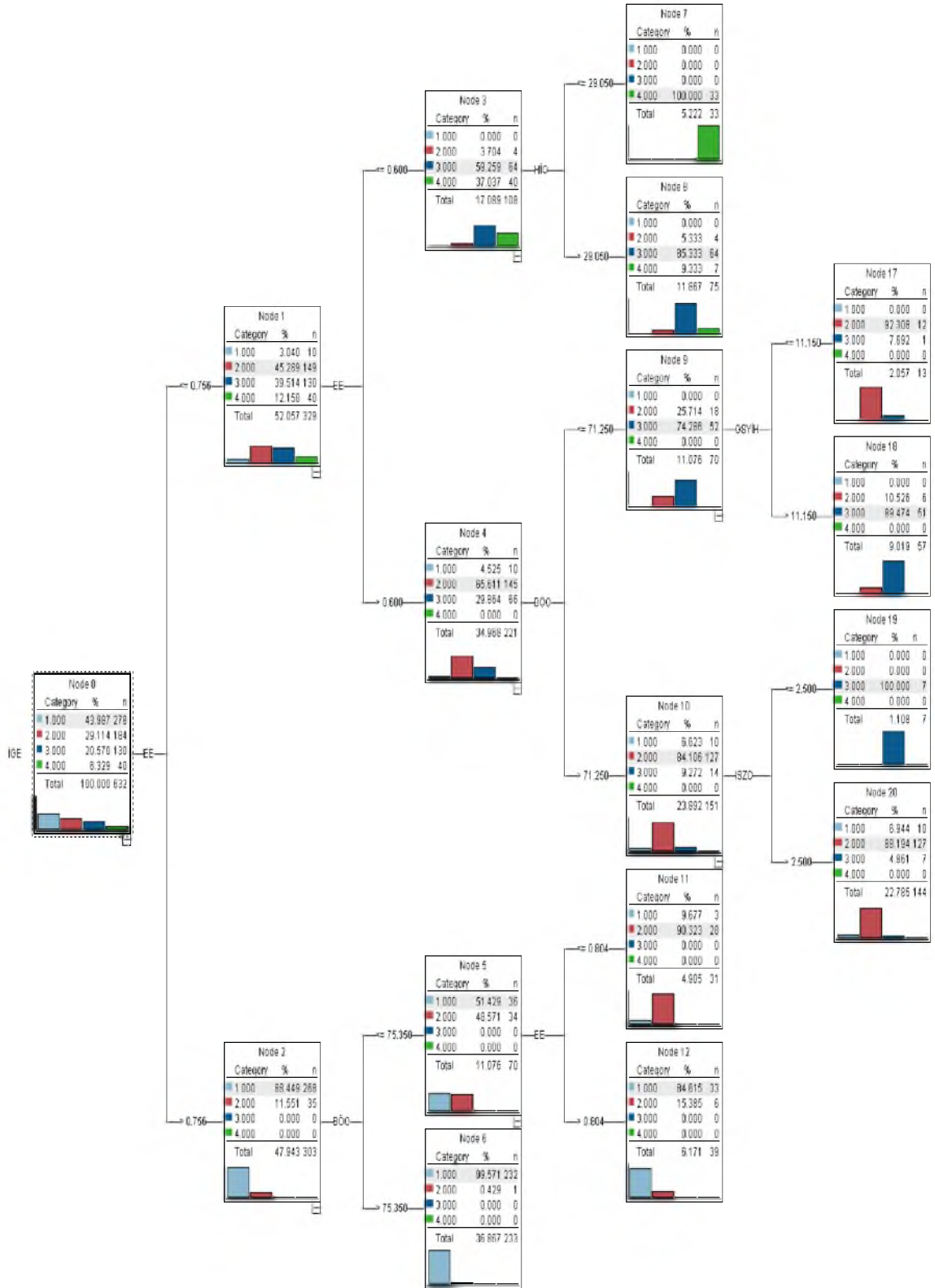
Kural 4: Eđer İGE'nin X1 deęişkeni "eđitim endeksi" 0.599'a eđit ya da kúçük, X3 deęişkeni "yaşam beklentisi endeksi" 0.731'den büyük ve X15 deęişkeni "cep telefonu abonelik oranı" 160,1'e eđit ya da kúçük ise İGE %97.9 oran ile orta İGE'dir.

Kural 8: Eđer İGE'nin X1 deęişkeni "eđitim endeksi" 0.796'dan büyük, X2 deęişkeni "dođumdan sonra beklenen yaşam süresi" 75.7'den büyük ve X13 deęişkeni "hizmet sektöründeki istihdam oranı" 54.9'dan büyük ise İGE %100 oran ile çok yüksek İGE'dir.

Kural 14: Eđer İGE'nin X1 deęişkeni "eđitim endeksi" 0.599'dan büyük ve 0.750'ye eđit ya da kúçük, X3 deęişkeni "yaşam beklentisi endeksi" 0.815'ten büyük, X13 deęişkeni "hizmet sektöründeki istihdam oranı" 34.9'dan büyük ve X2 deęişkeni "dođumdan sonra beklenen yaşam süresi" 77.1'e eđit ya da kúçük ise İGE %98,97 oran ile yüksek İGE'dir.

4.9. Gini Algoritması ile Karar Ađacı Modellemesi

Karar ađacının modellenmesinde kullanılan ikinci bir algoritma Gini algoritmasıdır. Gini algoritması ile karar ađacı modelinde 9 kural türetilmiřtir. İGE'nin sınıflandırılmasına iliřkin karar ađacı modelinde %61,9'luk oran ile en kuvvetli etkiye sahip olan deęişken eđitim endeksi, daha sonra %25,9'luk oran ile hizmet sektöründeki istihdam oranı deęişkeni ikinci sırada gelmektedir. řekil 7'de Gini algoritmasına ait karar ađacı verilmiřtir.



Şekil 7. Gini Algoritması ile İGE'nin Kategorik olarak Sınıflandırılmasına ait Karar Ağacı Modeli

Tablo 10

Karar Ağacı Modeli için Gini Algoritmasından Elde Edilen Kurallar

Kural No	Eğer Kuralları	%	Ülke Sayısı	Kategorik Değişkenler
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 10’da Gini algoritması ile oluşturulan karar ağacı modeline ilişkin toplam 9 kural oluşturulmuştur. Belirlenen kurallar içerisinde en önemlileri çok yüksek İGE için 3, yüksek İGE için 9, orta İGE için 2, düşük İGE için 1 no’lu kuraldır.

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 İGE %100 oran ile düşük İGE’dir.

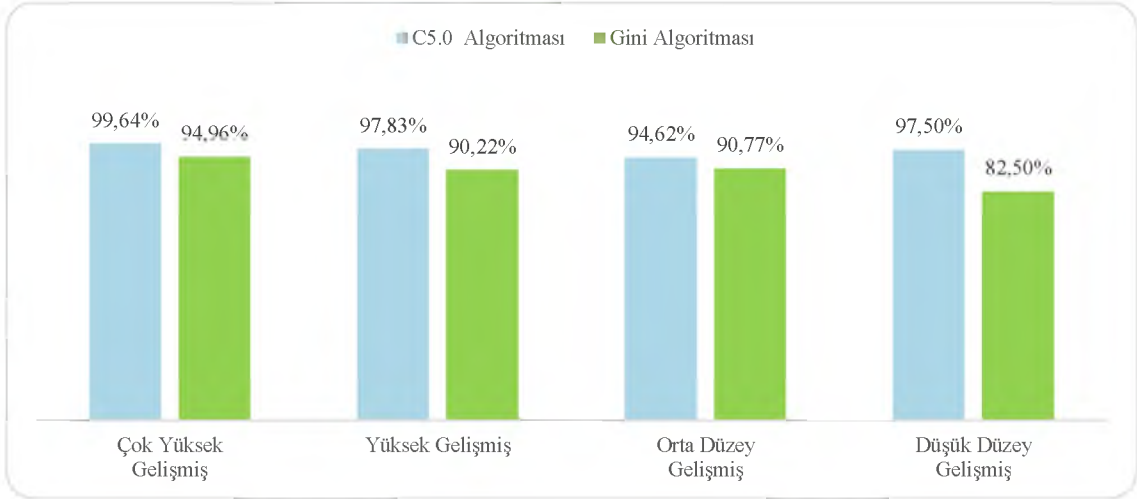
Kural 2: Eğer İGE’nin X1 değişkeni “eğitim endeksi” 0.60’a eşit ya da küçük ve X13 değişkeni “hizmet sektöründeki istihdam oranı” 29.05’ten büyük ise İGE %85,33 oran ile orta İGE’dir.

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 İGE %99,57 oran ile çok yüksek İGE’dir.

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 İGE %88,14 oran ile yüksek İGE’dir.

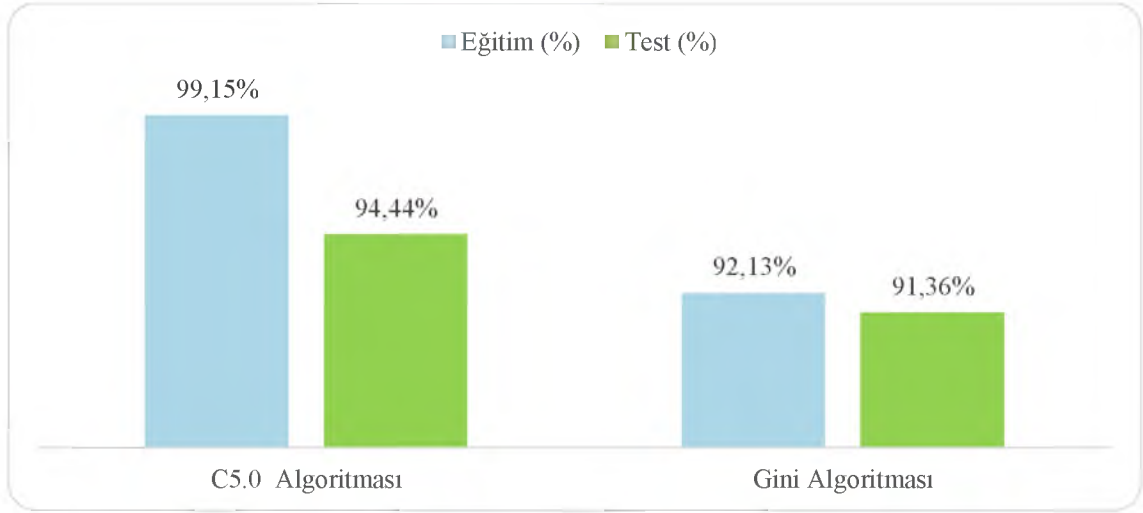
4.10. C5.0 ve Gini Algoritmalarının Performanslarının Karşılaştırılması

Bu aşamada C5.0 ve Gini 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 8’de ülkelerin İGE sınıflandırma performans sonuçları verilmiştir.



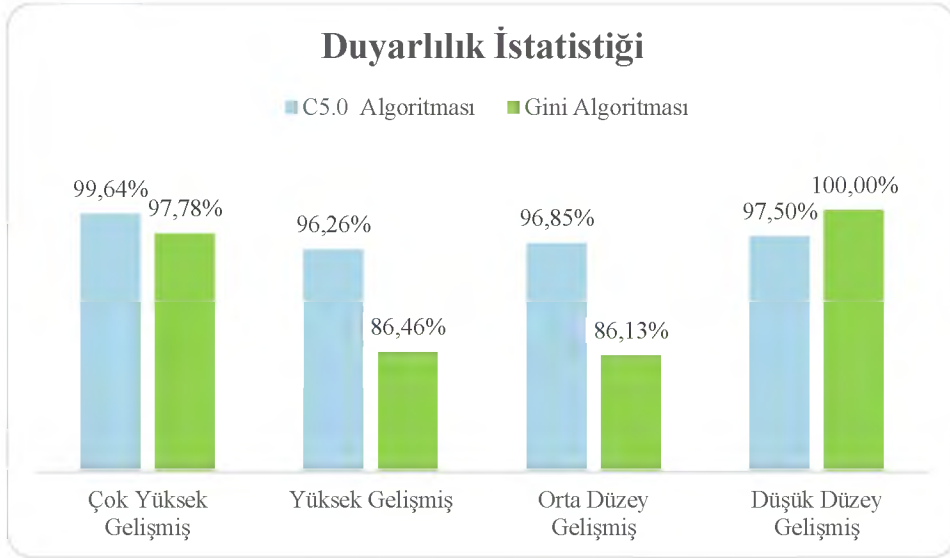
Şekil 8. C5.0 ve Gini Algoritmalarına İlişkin Ülkelerin İGE Sınıflandırma Performans Sonuçları

Şekil 8’e göre C5.0 algoritması analizi ile 278 çok yüksek gelişmiş ülkeden 277’sini, 184 yüksek gelişmiş ülkeden 180’ni, 130 orta düzey gelişmiş ülkeden 123’ünü ve 40 düşük düzey gelişmiş ülkeden 39’unu doğru tahmin ederek; çok yüksek gelişmiş ülkeler için %99,64’lük, yüksek gelişmiş ülkeler için %97,83’lük, orta düzey gelişmiş ülkeler için %94,62’lik ve düşük düzey ülkeler için %97,5’lik doğru sınıflandırma tahmini elde edilmiştir. Gini algoritması analizi kullanıldığında doğru sınıflandırma başarısı çok yüksek gelişmiş ülkeler için %94,96, yüksek gelişmiş ülkeler için %90,22, orta düzey gelişmiş ülkeler için %90,77 ve düşük düzey gelişmiş ülkeler için %82,5’tir. C5.0 ve Gini algoritmalarının toplam doğru sınıflandırma başarıları sırasıyla %97,94 ile %91,93’tür.



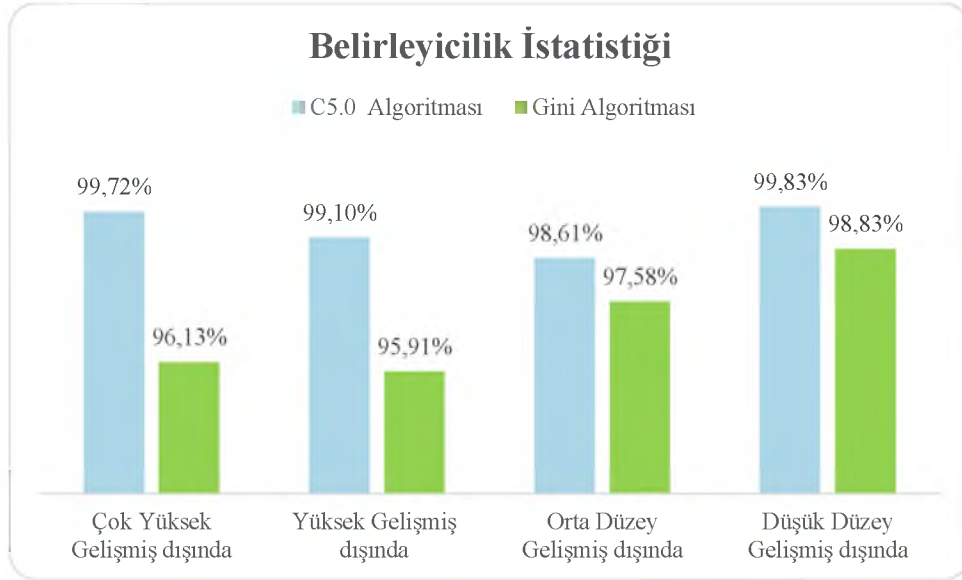
Şekil 9. C5.0 ve Gini Algoritmalarına ait Eğitim ve Test Sınıflandırma Performans Sonuçları

Şekil 9’da C5.0 ve Gini algoritmalarının eğitim ve test veri setlerine ilişkin performans sonuçları gösterilmiştir. 470 ülkenin yer aldığı eğitim seti için sınıflandırma performansı C5.0 için %99,15 ve Gini için %92,13’tür. 162 ülkenin bulunduğu test seti için C5.0 algoritmasında %92,13 ve gini algoritmasında %91,36’lık performans düzeyi tespit edilmiştir. Bazı durumlarda toplam doğru sınıflandırma performansına ilişkin analizler açısından test performanslarının karşılaştırılması hatalı sonuçlar verebilir. Bundan dolayı test performanslarının karşılaştırılmasında belirleyicilik ve duyarlılık istatistiğinin hesaplanması daha uygundur. Duyarlılık istatistiği (gerçek pozitif oran), İGE’si sınıflandırılan ülkeler açısından yüzde kaçının gerçekten doğru İGE olduğunu belirten orandır. Belirleyicilik istatistiği (gerçek negatif oran), İGE’si sınıflandırılan ülkeler açısından yüzde kaçının gerçekten yanlış İGE olduğunu açıklayan orandır (Öğüt, Doğanay, & Aktaş, 2009; Han & Kamber, 2006). Şekil 10 ve Şekil 11’de seçilen algoritmalara ilişkin belirleyicilik ve duyarlılık istatistiği gösterilmiştir.



Şekil 10. C5.0 ve Gini Algoritmalarının Duyarlılık İstatistiklerinin Performans Sonuçları

Şekil 10’da hesaplanan duyarlılık istatistiğine göre C5.0 algoritması analizi ile çok yüksek gelişmiş ülkeler için %99,64’lük, yüksek gelişmiş ülkeler için %96,26’lık, orta düzey gelişmiş ülkeler için %96,85’lik ve düşük düzey ülkeler için %97,50’lik; Gini algoritması analizi ile çok yüksek gelişmiş ülkeler için %97,78’lik, yüksek gelişmiş ülkeler için %86,46’lık, orta düzey gelişmiş ülkeler için %86,13’lük ve düşük düzey ülkeler için %100’lük doğru sınıflandırma düzeyi tespit edilmiştir.



Şekil 11. C5.0 ve Gini Algoritmalarının Belirleyicilik İstatistiklerinin Performans Sonuçları

Şekil 11’de görüldüğü üzere belirleyicilik istatistiğine göre C5.0 algoritması analizi ile çok yüksek gelişmiş dışındaki ülkeler için %99,72’lik, yüksek gelişmiş dışındaki ülkeler için %99,10’luk, orta düzey gelişmiş dışındaki ülkeler için %98,61’lik ve düşük düzey gelişmiş dışındaki ülkeler için %99,83’lük; Gini algoritması analizi ile çok yüksek gelişmiş dışındaki ülkeler için %96,13’lük, yüksek gelişmiş dışındaki ülkeler için %95,91’lik, orta düzey gelişmiş dışındaki ülkeler için %97,58’lik ve düşük düzey dışındaki ülkeler için %98,83’lük başarı düzeyi elde edilmiştir.

BÖLÜM V

SONUÇ VE ÖNERİLER

Bu çalışmada UNDP'nin 2010-2017 yıllarını kapsayan 79 ülkenin kategorik değişken olan insani gelişme endeksine etki eden faktörlerin, veri madenciliği karar ağacı teknik tekniklerinden olan C5.0 ve Gini algoritmalarını kullanarak modellenmesi amaçlanmıştır. 79 ülkenin 632 veri kaydı için 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 değişkenlerin normallik varsayımı sağlanmadığından spearman korelasyon analizi gerçekleştirilerek 5 değişken analizden çıkarılmıştır. Karar ağaçlarının oluşturulmasında C5.0 algoritması ile %97,94 ve Gini algoritmasıyla %91,93'lük 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 tüm yıllar boyunca yüksek düzey İGE doğru olarak sınıflandırılmıştır. 2010-2017 yılları arasında C5.0 algoritması ile 13 ülke 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 karar ağacı analizi ile modellenmesinde 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 faktörünün olduğu tespit edilmiştir. Seçilen algoritmaların karşılaştırılmasında sınıflandırma, eğitim ve test, belirleyicilik ve duyarlılık istatistik performans kriterleri olmak üzere dört farklı kriter kullanılmıştır. İstatistikî performans açısından karşılaştırılmasında C5.0 algoritmasının, Gini algoritmasına göre daha güçlü performans gösterdiği gözlenmiştir. Tablo 11'de İGE'nin sınıflandırılmasına yönelik benzer çalışmalar, önerdiğimiz modellerle karşılaştırılmıştır.

Tablo 11

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)	-	%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	-	%92,5- %100
Bolat & Çılan, (t.y.)	38 Ülke	3 değişken	DA	2004	%73,7
Önerdiğimiz Model	BM, 79	15 değişken	C5.0 ve Gini Algoritması	2010-2017	%97,94-%91,93

Sonuç itibari ile bu araştırma insan hayatının zenginliği açısından ülkeleri İGE göre karşılaştırma yapma imkânı sunmaktadır. Seçilen algoritmalar tarafından karar ağacı kurallarına göre bir ülkenin İGE göre hangi sınıfta yer alabileceği tahmin edilebilecektir. Gelecekteki yapılacak çalışmalarda daha geniş veri seti ve farklı analiz yöntemleri ile bu çalışmanın sonuçları karşılaştırılabilir.

KAYNAKÇA

- Adak, M., & Yurtay, N. (2013). Gini Algoritması Kullanarak Karar Ağacı Oluşturmayı Sağlayan Bir Yazılımın Geliştirilmesi. *Bilişim Teknolojileri Dergisi*, 1-6.
- Akar, S. (2014). Türkiye'de Daha İyi Yaşam Endeksi: OECD Ülkeleri ile Karşılaştırma. *Journal of Life Economics*, 1(1), 1-12.
- Akcapınar Sezer, E., Bozkır, A., Yağız, S., & Gökçeoğlu, C. (2010). Karar Ağacı Derinliğinin CART Algoritmasında Kestirim Kapasitesine Etkisi: Bir Tünel Açma Makinesinin İlerleme Hızı Üzerine Uygulama. *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*. Kayseri.
- Akgöbek, Ö., & Çakır, F. (2009). Veri Madenciliğinde Bir Uzman Sistem Tasarımı. *XI. Akademik Bilişim Konferansı Bildirimi* (s. 801-806). Şanlıurfa: Harran Üniversitesi.
- Akpınar, H. (2014). *Data Veri Madenciliği Veri Analizi* (1 b.). İstanbul: Papatya Yayıncılık Eğitim.
- Aksoy, B. (2018, Temmuz). İşletmelerde Finansal Başarısızlık Tahmininde Veri Madenciliği Yöntemlerinin Karşılaştırılması: BİST'te Bir Uygulama. *Doktora Tezi*. Kayseri: Erciyes Üniversitesi Sosyal Bilimler Enstitüsü.
- Alagöz, A., Öge, S., & Ortakarpuz, M. (2014). Bir Kurumsal Zekâ Teknolojisi Olarak Veri Madenciliği İle Muhasebe Bilgi Sistemi İlişkisi. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 1-21.
- Alan, M. A. (2012, Ağustos). Veri Madenciliği ve Lisansüstü Öğrenci Verileri Üzerine Bir Uygulama. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*(33), 165-174.
- Albayrak, M. (2017). Bilimsel Araştırmada Veri Madenciliği Kullanımı. *International Journal of Social Sciences and Education Research*, 3(2), 751-760.
- Alpar, R. (2013). *Çok Değişkenli İstatistik Yöntemler* (4. b.). Ankara: Detay Yayıncılık.
- Altıntaş, E. (2018). Borsa İstanbul İşletmelerinin (BİST) Veri Madenciliği ile Kümelenmesi. *Yüksek Lisans Tezi*, 6. Isparta: Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü.
- Armutlu, Ş. (2018, Kasım). Veri Madenciliği ile Kütüphane Kullanımı ve Ders Başarısı Arasındaki İlişkinin İncelenmesi. *Yüksek Lisans Tezi*. Uşak: Uşak Üniversitesi Sosyal Bilimler Enstitüsü.

- Arpacioğlu, Ö., & Yıldırım, M. (2011). Dünyada ve Türkiye'de Yoksulluğun Analizi. *Nigde Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 4(2), 60-76.
- Aşan, Z. (2007). Kredi Kartı Kullanan Müşterilerin Sosyo Ekonomik Özelliklerinin Kümeleme Analiziyle İncelenmesi. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 1-12.
- Ayık, Y., Özdemir, A., & Yavuz, U. (2007). Lise Türü ve Lise Mezuniyet Başarısının, Kazanılan Fakülte ile İlişkisinin Veri Madenciliği Tekniği ile Analizi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(2), 448.
- Ayık, Y., Özdemir, A., & Yavuz, U. (2007). Lise Türü ve Lise Mezuniyet Başarısının, Kazanılan Fakülte ile İlişkisinin Veri Madenciliği Tekniği ile Analizi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 441-454.
- Başar, S., Eren, C. M., & Eren, M. (2015). Ülkelerin İnsani Gelişmişlik Endeksi Değişkenlerine Göre Etkinliklerinin İncelenmesi. *International Conferanse on Eurasian Economies*, 846-854.
- Baykal, A. (2006). Veri Madenciliği Uygulama Alanları. *Düzce Üniversitesi Ziya Gökalp Fakültesi Dergisi*, 96.
- Baysakaloğlu, A. (2005). Veri Madenciliği ve Çimento Sektöründe Bir Uygulama. *Gaziantep Üniversitesi Endüstri Mühendisliği Bölümü*, 2-5.
- Berber, M. (2017). *İktisadi Büyüme ve Kalkınma* (6. b.). Trabzon: Celepler Matbaacılık Yayın ve Dağıtım.
- Birant, D., Kut, A., Venture, M., Altınok, H., Altınok, B., Altınok, E., & Ihlamur, M. (2010). İş Zekası Çözümleri İçin Çok Boyutlu Birliktelik Kuralları Analizi. *XII. Akademik Bilişim Konferansı Bildirileri* (s. 215-222). Muğla: Muğla Üniversitesi.
- Bolat, B., & Çılan, Ç. (t.y.). İnsani Gelişme Endeksi Bileşenleri Açısından Gelişmekte Olan Ülkelerin Diskriminant Analizi ile Karşılaştırılması. 223-234.
- Burmaoğlu, S., & Oktay, E. (2015). A Statistical Classification Study Countries' Human Development Level By Discriminat Analysis. *Ataturk University Journal of Economics & Administrative Sciences*, 29(4), 579-599.
- Burmaoğlu, S., Oktay, E., & Özen, Ü. (2009). Birleşmiş Milletler Kalkınma Programı Beşeri Kalkınma Endeksi Verilerini Kullanarak Diskriminat Analizi ve Lojistik Regresyon Analizinin Sınıflandırma Performanslarının Karşılaştırılması.
- Coşlu, E. (2013). Veri Madenciliği. *Akademik Bilişim 2013-XV. Akademik Konferansı Bildirileri* (s. 573-585). Antalya: Akdeniz Üniversitesi.

- Çağlar, A., & Keten, N. (2018). İllerin İnsani Gelişme Endeksinin Veri Zarflama Analizi İle Ölçülmesi. *Ege Akademik Bakış Dergisi*, 565-578.
- Çalışkan, Ş. (2010). Türkiye'de Beşeri Sermaye Harcamaları ve İnsani Gelişmişlik. *İş. Güç Endüstri İlişkileri ve İnsan Kaynakları Dergisi*, 12(1), 7-28.
- Çelik, H. C., & Kahyaoğlu, M. (2007). İlköğretim Öğretmen Adaylarının Teknolojiye Yönelik Tutumlarının Kümeleme Analizi. *Türk Eğitim Bilimleri Dergisi*, 5(4), 571-586.
- Çemrek, F., & Bayraç, H. (2013). Sürdürülebilir Kalkınma Skorunun Hesaplanması. *Eskişehir Kalkınma Skorunun Hesaplanması*, 14(2), 131-152.
- Çiftçi, M. (2008). Kalkınma Göstergesi olarak Ortalama Yaşam Beklentisine Göre Türkiye'nin AB İçindeki Konumu: Kritikler ve Çok Değişkenli İstatistik Uygulamaları. *İstanbul Üniversitesi İktisat Fakültesi Ekonometri ve İstatistik Dergisi*(7), 51-87.
- Çoban, N., & Yayar, R. (2018). Demokrasinin Göstergelerinin İnsani Gelişmişlik Üzerine Etkisi: AB Ülkeleri Üzerine Bir Panel Veri Analizi. *Journal of Politics Economy and Management*, 1(2), 55-61.
- Demir Şeker, S. (2011). *Türkiye'nin İnsani Gelişme Endeksi ve Endeks Sıralamasının Analizi*. Ankara: T.C. Kalkınma Bakanlığı Sosyal Sektörler ve Koordinasyon Genel Müdürlüğü.
- Demir, S. (2006). *Birleşmiş Milletler Kalkınma Programı İnsani Gelişme Endeksi ve Türkiye Açısından Değerlendirme*. Ankara: Devlet Planlama Teşkilatı.
- Demiral, G., Soba, M., & Armutlu, Ş. (2017). Kütüphane Veri Tabanında Veri Madenciliği: Uşak Üniversitesi Örneği. *Bartın Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 241-264.
- Diler, S. (2016). Veri Madenciliği Süreçleri ve Karar Ağaçları Algoritmaları ile Bir Uygulama. *Yüksek Lisans Tezi*. Van: Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü.
- Doğan, E., & Tatlı, H. (2014). İnsani Gelişme ve İnsani Yoksulluk Bağlamında Türkiye'nin Dünyadaki Yeri. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 28(1), 99-124.
- Emre, İ. E., & Selçukcan Erol, Ç. (2017). Veri Analizinde İstatistik mi Veri Madenciliği mi? *Bilişim Teknolojileri Dergisi*, 10(2), 161-167.
- Erbay, E. R., & Özden, M. (2013). İktisadi Kalkınma Kuramlarına Eleştirel Yaklaşım. *Sosyal Bilimler Metinleri*, 1, 1-27.

- Erdem, E., & Çelik, B. (2019). İnsani Gelişme ve Ekonomik Büyüme İlişkisi: Bazı Afrika Ülkeleri Üzerine Bir Uygulama. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(17), 13-36.
- Erdoğan, M. (2010). Beşeri Gelişme Perspektifinden Türkiye'de Kamu Eğitim ve Sağlık Harcamaları. 324-336.
- Erkekoğlu, H. (2007). AB'ye Tam Üyelik Sürecinde Türkiye'nin Üye Ülkeler Karşısındaki Görelî Gelişme Düzeyi: Çok Değişkenli İstatistiksel Bir Analiz. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 28-50.
- Erol, E. (2011). İnsani Gelişme Yaklaşımı Doğrultusunda Beşeri Kalkınmanın Boyutları: Gelişmekte Olan Ülkeler. *Sosyal ve Beşeri Bilimler Dergisi*, 3(2), 1309-8012.
- Erol, E. (2013). Türkiye ve Avrupa Birliği Üyesi Ülkelerin Sosyo-Ekonomik Gelişmişlik Düzeylerinin Karşılaştırmalı Analizi. *Sosyal ve Beşeri Bilimler Dergisi*, 5(1), 198-208.
- Erpolat, S. (2012). Otomobil Yetkili Servislerinde Birliktelik Kurallarının Belirlenmesinde Apriori ve FB-Growth Algoritmalarının Karşılaştırılması. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 137-146.
- Fakı, B., & Öztayşı, B. (tarih yok). Veri Madenciliği Yöntemlerini Kullanarak Anemi Sınıflandırılmasına Yönelik Bir Uygulama.
- Fırat, E., & Aydın, A. (2015). İnsani Kalkınma Endeksine Göre Türkiye'nin Eğitim Endeks Göstergelerinin OECD Ülkeleri İle Karşılaştırılması. *Selçuk Üniversitesi İktisadi ve İdari Bilimler Fakültesi Sosyal Ekonomik Araştırmalar Dergisi*(29), 62-87.
- Günel, F., Çağlar, A., Kangallı Uyar, S., Karadeniz, O., & Yeşilyurt, M. (2017). Türkiye'de İllere Göre İnsani Gelişme Endeksi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 208-216.
- Gündüz, M., & Soyer, M. (2019). Küresel Rekabetçilik ile Ekonomik Özgürlük Arasındaki İlişkinin Binary Lojistik Regresyon Analizi ile İncelenmesi. *İnsan ve Toplum Bilimleri Araştırma Dergisi*, 8(3), 1906-1930.
- Günsoy, G. (2005). İnsani Gelişme Kavramı ve Sağlıklı Yaşam Hakkı. *ZKÜ Sosyal Bilimler Dergisi*, 35-52.
- Gürsakar, N. (2014). *Büyük Veri*. Bursa: DORA Basım-Yayın Dağıtım Ltd. Şti.
- Gürses, D. (2009, Haziran). İnsani Gelişme ve Türkiye. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(21), 339-350.

- Gürsoy, U. (2009). *Veri Madenciliği ve Bilgi Keşfi*. İstanbul: Pegem Akademi.
- Gürsoy, U. (2012). *Uygulamalı Veri Madenciliği Sektörel Analizler*. Ankara: Pegem Akademi.
- Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques. *The Morgan Kaufmann*.
- İnsani Gelişme Vakfı*. (2019, 8 15). ingev: <http://ingev.org/hakkimizda/insani-gelisme-nedir/> adresinden alındı
- Kalaycı, Ş. (2008). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri* (3 b.). Ankara: Asil Yayın.
- Karabulut, T., Kaya, N., & Gürsoy, Z. (2009). Ekonomik Kalkınma ve İşbirliği Örgütü'ne Üye Ülkelerin 2016 Yılı İnsani Gelişmişlik Düzeylerinin Analizi. *Niğde Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 2(2), 1-18.
- Kaya, A. (2017). Finansal Piyasalardaki Gelişmelerin İnsani Gelişmişlik Üzerine Etkisi: Türkiye Örneği. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 169-180.
- Kaynak, S. (2017). AB'ye Tam Üyelik Sürecinde Aday Ülkelerin İnsani Gelişme, Eğitim ve Gelir İndekslerinin Çok Amaçlı Doğrusal Programlama ile Çoklu Periyotlu Olarak Değerlendirilmesi. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 41-49.
- Kayri, M. (2008, Haziran). Elektronik Portfolyo Değerlendirmeleri İçin Veri Madenciliği Yaklaşımı. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 5(1), 98-110.
- Kazar, G., & Kazar, A. (2013). Eşitsizlikle Uyumlaştırılmış İnsani Gelişme Endeksinin Türkiye İçin Değerlendirilmesi. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 22(2), 195-208.
- Kırtay, S. H., Ekmekçi, N., Halıcı, T., Ketenci, U., Aktaş, M. S., & Kalıpsız, O. (2015, Eylül 09-11). Pazar Sepeti Analizi İçin Örneklem Oluşturulması ve Birliktelik Kurallarının Çıkarılması. *9. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS'15)*, 172-178. İzmir.
- Kocatepe Cebeci, A., & Dilber, İ. (2014). Türkiye'nin İnsani Gelişme Endeks Değerinin 2011-2012 Karşılaştırılması. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 24(2), 123-131.
- Korkmaz, C., & Şahin, M. (2013). 2009 PISA Başarılarına Göre Ülkelerin Genel ve İnsani Gelişmişlik Düzeyleri Arasındaki İlişki. *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 İller İçin Ölçülmesi. *İdealkent Dergisi*, 9(24), 387-429.

- Mihçı, H., & Mihçı, S. (2003). Türkiye'nin Yakın Dönemdeki İnsani Gelişme Eğilimleri. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21(2), 21-47.
- Murat, D., & Gürsakal, S. (2015). Mutluluk ve İnsani Gelişmişlik Arasındaki İlişkinin Belirlenmesi: Çok Değişkenli İstatistiksel Yaklaşım. *Alphanumeric Journal*, 3(1), 67-79.
- Nartgün, Ş. S., Kösterelioğlu, M., & Sipahioğlu, M. (2013). İnsani Gelişim İndeksi Göstergeleri Açısından AB Üyesi ve AB Üyeliğine Aday Ülkelerin Karşılaştırılması. *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 3(1), 80-89.
- Oğuzlar, A. (2003). Veri Ön İşleme. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 67-76.
- 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.
- Öngel, V., Sözen, İ., & Çelik, A. (2011). Orta Asya Ülkelerinin İnsani Gelişmişlik Endeksleri Açısından Değerlendirilmesi. *International Conferance On Eurasian Economies* (s. 434-439). Tashkent: Eurasian Economists Association.
- Özekes, S. (2003). Veri Madenciliği Modelleri ve Uygulama Alanları. *İstanbul Ticaret Üniversitesi Dergisi*, 65-82.
- Özkan, Y. (2008). *Veri Madenciliği Yöntemleri*. İstanbul: Papatya Yayıncılık Eğitim A.Ş.
- Öztürk, H. M. (2016). Human Development Index of Turkey, Economic Development, Relationship Nutrition and Comparison with Selected Countries Türkiye'nin İnsani Gelişme Endeksi, Ekonomik Gelişmişlik, Beslenme İlişkisi ve Farklı Ülkelerle Karşılaştırılması. *Journal of Human Sciences*, 13(2), 3402-3412.
- Paksoy, S. (2015). Ülke Göstergelerinin Vikor Yöntemi İle Değerlendirilmesi. *Ekonomik ve Sosyal Araştırmalar Dergisi*, 11(2), 153-169.
- Rençber, Ö. F. (2017). Bulanık ve Yapay Sinir Ağları İle Çoklu Lojistik Regresyon Yöntemlerinin Sınıflandırma Performanslarının Karşılaştırılması: Ülkelerin Gelişmişlik Düzeylerinin Sınıflandırılması Üzerine Bir Uygulama. *Doktora Tezi*. Aksaray: Aksaray Üniversitesi Sosyal Bilimler Enstitüsü.
- Rençber, Ö. F. (2018). *Sınıflandırma Problemlerinde Çoklu Lojistik Regresyon, Yapay Sinir Ağı ve ANFIS Yöntemlerinin Karşılaştırılması: İnsani Gelişmişlik Endeksi Üzerine Uygulama*. Ankara: Gazi Kitabevi Tic. Ltd. Şti.

- Rençber, Ö. F., & Mete, S. (2018). Ülkelerin İnsani Gelişmişlik Endeksine Göre Yeniden Sınıflandırılması: Yapay Sinir Ağı ve ANFIS Yöntemleri ile Bir Uygulama. *Business & Management Studies: an International Journal*, 6(3), 228-252.
- Sakarya, A., & İbişoğlu, Ç. (2015). Türkiye'de İllerin Sosyo-Ekonomik Gelişmişlik Endeksinin Coğrafi Ağırlıklı Regresyon Modeli İle Analizi. *Marmara Coğrafya Dergisi*, 211-238.
- Savaş, S., Topaloğlu, N., & Yılmaz, M. (2012). Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 1-23.
- Seyrek, İ., & Ata, H. (2010). Veri Zarflama Analizi ve Veri Madenciliği ile Mevduat Bankalarında etkinlik Ölçümü. *BDDK Bankacılık ve Finansal Piyasalar Dergisi*, 4(2), 67-84.
- Silahtaroglu, G. (2004). Veri Madenciliğinde Kümeleme Analizi ve Öğretim Başarısının Değerlemesine İlişkin Bir Uygulama. *Doktora Tezi*. İstanbul: İstanbul Üniversitesi Sosyal Bilimler Enstitüsü.
- Silahtaroglu, G. (2013). *Veri Madenciliği*. İstanbul: Papatya Yayıncılık Eğitim.
- Şahin, G., & Gökdemir, L. (2016). İnsani Gelişme Endeksi Bileşenlerinin Türkiye Ölçeğinde ARDL Sınır Testi İle Sınanması. *Gazi İktisat ve İşletme Dergisi*, 1-24.
- Şanlısoy, S. (2019). İnsani Gelişmişlik Endeksine Türk Cumhuriyetleri Örneğinde Eleştirel Bir Yaklaşım. *Uluslararası Sosyal Araştırmalar Dergisi*, 12(62), 1307-9581.
- Şengür, D., & Tekin, A. (2013, Eylül). Öğrencilerin Mezuniyet Notlarının Veri Madenciliği Metotları İle Tahmini. *Bilişim Teknolojileri Dergisi*, 6(3), 7-16.
- Şentürk, A. (2006). *Veri Madenciliği Kavram ve Teknikler*. Bursa: Ekin Basım Yayın Dağıtım.
- Taban, S. (2016). *İktisadi Büyüme Kavram ve Modeller*. Bursa: EKİN Basım Yayın Dağıtım.
- Taşkın, Ç., & Emel, G. (2010). Veri Madenciliğinde Kümeleme Yaklaşımları ve Kohonen Ağları İle Perakendecilik Sektöründe Bir Uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 15(3), 395-409.
- Tekin, B. (2015). Temel Sağlık Göstergeleri Açısından Türkiye'deki İllerin Gruplandırılması: Bir Kümeleme Analizi Uygulaması. *Çankırı Karatekin Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 5(2), 389-419.

- Terlemez, L. (2008, Şubat). Eş İşlem Stratejisi Yöntemiyle İMKB'de Portföy Oluşturmada Veri Madenciliği Uygulamaları. *(Doktora Tezi, basılmamış)*. Eskişehir: Atatürk Üniversitesi Fen Bilimleri Enstitüsü.
- Terzi, Ö., Küçüksille, E., Ergin, G., & İlker, A. (2011). Veri Madenciliği Süreci Kullanılarak Güneş Işınımı Tahmini. *Süleyman Demirel Üniversitesi Uluslararası Teknolojik Bilimler Dergisi*, 3(2), 29-37.
- Tıraş, H. (2019). Türkiye İçin İnsani Gelişmişlik Göstergeleri. *Bilgi Ekonomisi ve Yönetimi Dergisi*, 14(1), 15-31.
- Tıraş, H., & Ağır, H. (2018). İnsani Gelişmişlik Göstergeleri Açısından İslam İşbirliği Teşkilatı Üye Ülkelerinin Değerlendirilmesi. *Sosyal Ekonomik Araştırmalar Dergisi*, 18(35), 20-40.
- Tireli, M., Çoşkun, S., & Kunduracı, N. F. (2013). İİT Ülkeleri ve Türkiye'ye Ait Sosyal Göstergeler İle İnsani Gelişim Göstergeleri Arasındaki İlişkinin Karşılaştırmalı Analizi. *Sosyal Politikalar Çalışmaları Dergisi*, 7(30), 61-87.
- Tunç, A., & Ülger, İ. (t.y.). Veri Madenciliği Uygulamalarında Özellik Seçimi İçin Finansal Değerlere Binning ve Five Number Summary Metotları ile Normalizasyon İşleminin Uygulanması. *XVIII. Akademik Bilişim Konferansı*. Adnan Menderes Üniversitesi.
- 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.
- Tüylüoğlu, Ş., & Karalı, B. (2006). İnsani Kalkınma Endeksi ve Türkiye İçin Değerlendirilmesi. *Selçuk Üniversitesi İktisadi ve İdari Bilimler Fakültesi Sosyal ve Ekonomik Araştırmalar Dergisi*, 6(12), 53-88.
- Tüzüntürk, S. (2010). Veri Madenciliği ve İstatistik. *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*(1), 65-90.
- UNDP. (1995). *Human Development Report 1995*. New York: Oxford University Press.
- UNDP. (2016). *Human Development Report 2016*. New York: Gerry Quinn and Human Development Report Office.
- UNDP. (2018). *Human Development Indices and Indicators 2018 Statistical Update, Technical Notes*. UNDP.
- UNDP. (2018). *Human Development Report*. New York: Gerry Quinn and Human Development Report Office.

- Ülgener, S. F. (1991). *Milli Gelir, İstihdam ve İktisadi Büyüme*. İstanbul: DER Yayınlar.
- Ünal, Ç. (2008). İnsani Gelişmişlik Endeksine Göre Türkiye'nin Bölgesel Farklılıkları. *Coğrafi Bilimler Dergisi*, 89-113.
- Üzümçü, A. (2012). *İktisadi Büyüme (Teori, Model ve Türkiye Üzerine Gözlemler)*. İstanbul: Beta Basım Yayım Dağıtım A.Ş.
- Yakut, E. (2012, 12). Veri Madenciliği Tekniklerinden C5.0 Algoritması ile Destek Vektör Makineleri ile Yapay Sinir Ağlarının Sınıflandırma Başarılarının Karşılaştırılması: İmalat Sektöründe Bir Uygulama. *Doktora Tezi*. Erzurum: Atatürk Üniversitesi Sosyal Bilimler Enstitüsü.
- Yakut, E., & Gemici, E. (2017, Ekim). LR, C5.0, CART, DVM Yöntemlerini Kullanarak Hisse Senedi Getiri Sınıflandırma Tahmini Yapılması ve Kullanılan Yöntemlerin Karşılaştırılması: Türkiye'de BİST'te Bir Uygulama. *Ege Akademik Bakış*, 17(4), 461-479.
- 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*, 172-199.
- Yalçın, A. Z., & Çakmak, F. (2016). Türkiye'de Kamu Sağlık Harcamalarının İnsani Gelişim Üzerindeki Etkisi. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 30(4), 705-724.
- Yalçın, L. (2019, Ocak). Sağlık Sektöründe Veri Madenciliği. *Yüksek Lisans Tezi*. Milli Savunma Üniversitesi Hezarfen Havacılık ve Uzay Teknolojileri Enstitüsü.
- Yıldırım, P., Uludağ, M., & Görür, A. (2008). Hastane Bilgi Sisteminde Veri Madenciliği. *Çanakkale Onsekiz Mart Üniversitesi Akademik Bilişim*, 429-434.
- Zanbak, M., & Özeş Özgür, R. (2019). İnsani Gelişme Endeksi Bağlamında Avrupa Birliği'ne Üye ve Aday Ülkelerin Karşılaştırmalı Analizi. *Yönetim ve Ekonomi Araştırmaları Dergisi*, 17(2), 175-192.
- Zor, E. G. (2019). Türkiye'de 1994 ve Sonrasında Yaşanan Krizlerin İnsani Gelişme Endeksine Etkisi. *Yüksek Lisans Tezi*. Manisa: Celal Bayar Üniversitesi Sosyal Bilimler Enstitüsü.

EKLER

EK 1:

Analiz Kapsamında İGE ve Alt Endeks Verileri Değerlendirilen Ülkeler

S.No	Ülke	S.No	Ülke
1	Norveç	40	Türkiye
2	İsviçre	41	Sırbistan
3	Avustralya	42	Arnavutluk
4	İrlanda	43	Gürcistan
5	Almanya	44	Meksika
6	İzlanda	45	Sri Lanka
7	İsveç	46	Brezilya
8	Singapur	47	Azerbeycan
9	Hollanda	48	Ermenistan
10	Danimarka	49	Çin
11	Kanada	50	Ukrayna
12	Birleşik Krallık	51	Peru
13	Finlandiya	52	Kolombiya
14	Belçika	53	Moğolistan
15	Avusturya	54	Dominik Cumhuriyeti
16	Kore Cumhuriyeti)	55	Ürdün
17	Fransa	56	Surinam
18	Slovenya	57	Maldivler
19	İspanya	58	Belize
20	Çekya	59	Paraguay
21	İtalya	60	Moldova Cumhuriyeti
22	Estonya	61	Filipinler
23	Yunanistan	62	Güney Afrika
24	Kıbrıs	63	Endonezya
25	Polonya	64	Vietnam
26	Litvanya	65	Bolivya
27	Slovakya	66	El Salvador
28	Letonya	67	Kırgızistan
29	Portekiz	68	Fas
30	Macaristan	69	Nikaragua
31	Hırvatistan	70	Guatemala
32	Rusya	71	Tacikistan
33	Karadağ	72	Bangladeş
34	Bulgaristan	73	Kenya
35	Romanya	74	Zambiya
36	Belarus	75	Pakistan
37	Uruguay	76	Ruanda
38	Malezya	77	Uganda
39	Kosta Rika	78	Sudan
		79	Malavi

ÖZ GEÇMİŞ

KİŞİSEL BİLGİLERİ

Adı Soyadı : Ali KORKMAZ
Doğum Yeri : Kahramanmaraş
Adres : Osmaniye Korkut Ata Üniversitesi
Strateji Geliştirme Daire Başkanlığı
Fakültesi/OSMANIYE
E-Posta : alikorkmaz@osmaniye.edu.tr

EĞİTİM BİLGİLERİ

Lisans : Anadolu Üniversitesi
İşletme Fakültesi / 2010
Ön Lisans : Kahramanmaraş Sütçü İmam Üniversitesi
Kahramanmaraş Meslek Yüksekokulu
Turizm ve Otel İşletmeciliği / 2007

İŞ DENEYİMİ

2009-2012 : Osmaniye Korkut Ata Üniversitesi
Personel Daire Başkanlığı
Memur/Bilgisayar İşletmeni
2013-2018 : Osmaniye Korkut Ata Üniversitesi
Strateji Geliştirme Daire Başkanlığı
Şef
2018-Devam ediyor : Osmaniye Korkut Ata Üniversitesi
Strateji Geliştirme Daire Başkanlığı
Şube Müdürü

OSMANİYE KORKUT ATA ÜNİVERSİTESİ SOSYAL BİLİMLER ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

OSMANİYE KORKUT ATA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 27/09/2019

Tez Başlığı / Konusu: İNSANİ GELİŞİMİŞLİK ENDEKSİNİN VERİ MADENCİLİĞİ TEKNİKLERİNDEN OLAN C5.0 ve GINI ALGORİTMALARI KULLANARAK MODELLENMESİ

Yukarıda başlığı/konusu belirlenen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 72 sayfalık kısmına ilişkin, 27/09/2019 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtreleme tiplerinden biri uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 26'dır.

Filtreleme Tip 1 (maksimum %30)

- 1- Kabul/Onay ve Bildirim sayfaları hariç,
- 2- Kaynakça dâhil,
- 3- Alıntılar dâhil.

Filtreleme Tip 2 (maksimum %10)

- 1- Kabul/Onay ve Bildirim sayfaları hariç,
- 2- Kaynakça hariç,
- 3- Alıntılar dâhil,
- 4- 5 Kelimedenden daha az örtüşme içeren metin kısımları hariç.

Osmaniye Korkut Ata Üniversitesi Sosyal Bilimler Enstitüsü Tez Çalışması Orjinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Adı Soyadı: Ali KORKMAZ

.Öğrenci No: 15YLIS1115

Anabilim Dalı: İŞLETME

Programı: İŞLETME

Statüsü: Y.Lisans Doktora

27.09.2019

DANISMAN ONAYI

UYGUNDUR.

Dr. Öğretim Üyesi Emre YAKUT

ENSTİTÜ ONAYI

UYGUNDUR.

(Unvan, Ad Soyad, İmza)
Arş. Gör. Ali Korkmaz