

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**ENERJİ YOĞUNLUĞUNU ETKİLEYEN FAKTÖRLERİN REGRESYON
ANALİZİ VE MAKİNE ÖĞRENMESİ YAKLAŞIMLARIYLA
MODELLENMESİ**

Hülya EREN

İSTATİSTİK ANABİLİM DALI

**ANKARA
2025**

Her hakkı saklıdır

ÖZET

Yüksek Lisans Tezi

ENERJİ YOĞUNLUĞUNU ETKİLEYEN FAKTÖRLERİN REGRESYON ANALİZİ VE MAKİNE ÖĞRENMESİ YAKLAŞIMLARIYLA MODELLENMESİ

Hülya EREN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Doç. Dr. Sibel AÇIK KEMALOĞLU

Enerji yoğunluğu, bir birim ekonomik çıktı (GSYH) üretmek için tüketilen enerji miktarını ifade eden temel bir göstergedir. Bir ekonomide enerji kaynaklarının ne kadar verimli kullanıldığını yansıtan bu ölçüt, aynı zamanda ekonomik yapının karakteristiklerini ve çevresel sürdürülebilirlik düzeyini değerlendirmek açısından da kritik bir rol oynamaktadır. Enerji yoğunluğunun düşük olması, daha az enerji kullanılarak daha yüksek ekonomik üretim gerçekleştirildiğini gösterir. Bu nedenle, enerji yoğunluğunun azaltılması; sürdürülebilir kalkınma hedeflerine ulaşılması, sera gazı emisyonlarının azaltılması ve doğal kaynakların etkin yönetimi açısından büyük önem taşımaktadır. Bu tez çalışmasında, ülkelerin enerji yoğunluğunu belirleyen faktörlerin ortaya konulması amaçlanmıştır. Bağımlı değişken olarak enerji yoğunluğu ele alınırken; doğalgaz, elektrik, kömür ve petrol tüketimleri ile yenilenebilir enerji kaynakları ve nüfus yoğunluğu bağımsız değişkenler olarak kullanılmıştır. Analizlerde ülkeler, Dünya Bankası sınıflandırmasına göre alt-orta, orta-yüksek ve yüksek gelirli olmak üzere üç gruba ayrılmıştır. Enerji yoğunluğunun $[0,1]$ aralığında sınırlı olması nedeniyle Beta Regresyon yöntemi tercih edilmiş; çoklu bağlantı sorununu gidermek için Beta Ridge Regresyon uygulanmıştır. Karşılaştırmalı analiz amacıyla ayrıca Destek Vektör Regresyonu, Rastgele Orman ve Karar Ağacı algoritmaları kullanılmış, modeller MSE, RMSE, MAE ve R^2 ölçütleriyle performans açısından değerlendirilmiştir. Ek olarak, Rastgele Orman modeli üzerinden değişkenlerin enerji yoğunluğu üzerindeki göreceli etkileri incelenmiştir. Elde edilen bulgular, enerji yoğunluğunu etkileyen faktörlerin gelir gruplarına göre farklılık gösterdiğini ortaya koymuş ve politika yapıcılar için gruba özgü enerji verimliliği stratejilerinin geliştirilmesi gerektiğini vurgulamıştır.

Eylül 2025, 78 sayfa

Anahtar Kelimeler: Enerji yoğunluğu, Beta Regresyon, Makine Öğrenmesi, Destek Vektör Regresyonu, Rastgele Orman, Karar Ağacı, Gelir Grupları, Yenilenebilir Enerji Kaynakları

ABSTRACT

Master Thesis

MODELING THE DETERMINANTS OF ENERGY INTENSITY USING REGRESSION ANALYSIS AND MACHINE LEARNING APPROACHES

Hülya EREN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Assoc. Prof. Sibel AÇIK KEMALOĞLU

Energy intensity is a fundamental indicator that expresses the amount of energy consumed to produce one unit of economic output (GDP). This measure, reflecting how efficiently energy resources are utilized within an economy, also plays a critical role in assessing the characteristics of the economic structure and the level of environmental sustainability. A lower level of energy intensity indicates that higher economic output is achieved with less energy consumption. Therefore, reducing energy intensity is of great importance for achieving sustainable development goals, mitigating greenhouse gas emissions, and ensuring the efficient management of natural resources. In this thesis, the aim is to identify the factors that determine countries' energy intensity. While energy intensity is considered as the dependent variable, natural gas, electricity, coal, and oil consumption, along with renewable energy sources and population density, are used as independent variables. In the analyses, countries are grouped into three categories according to the World Bank classification: lower-middle, upper-middle, and high income. Since energy intensity is bounded within the $[0,1]$ interval, the Beta Ridge Regression method has been employed; Beta Ridge Regression is applied to address the multicollinearity problem. For comparative analysis, Support Vector Regression, Random Forest, and Decision Tree algorithms are also used, and the models are evaluated in terms of performance using MSE, RMSE, MAE, and R^2 metrics. In addition, the relative impacts of the variables on energy intensity are examined through the Random Forest model. The findings reveal that the factors influencing energy intensity differ across income groups and highlight the need for policymakers to develop group-specific energy efficiency strategies.

September 2025, 78 pages

Key Words: Energy Intensity, Beta Regression, Machine Learning, Support Vector Regression, Random Forest, Decision Tree, Income Levels, Renewable Energy Resources

ÖNSÖZ ve TEŞEKKÜR

Bu çalışmanın ortaya çıkmasında bana rehberlik eden, destek olan ve katkı sağlayan, değerli danışman hocam Doç. Dr. Sibel AÇIK KEMALOĞLU'na teşekkürlerimi sunarım.

Çalışmanın yazım sürecinde desteğini esirgemeyen aileme ve sevgili eşime teşekkürü bir borç bilirim.

Hülya EREN
Ankara, Eylül 2025

İÇİNDEKİLER

TEZ ONAY SAYFASI

ETİK.....	i
ÖZET.....	ii
ABSTRACT	iii
ÖNSÖZ ve TEŞEKKÜR.....	iv
KISALTMALAR DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	ix
1. GİRİŞ	1
2. KURAMSAL TEMELLER ve KAYNAK ÖZETLERİ	5
3. BETA REGRESYON ANALİZİ	10
3.1 Bilgi Kriterleri	13
3.2 Beta Regresyon Analizinde Kullanılan Bağlantı Fonksiyonları	14
4. BETA RİDGE REGRESYON ANALİZİ	16
5. MAKİNE ÖĞRENMESİ YÖNTEMLERİ	17
5.1 Denetimsiz Öğrenme	18
5.2 Pekiştirmeli Öğrenme	19
5.3 Denetimli Öğrenme	19
5.3.1 Karar ağaçları.....	19
5.3.2 Rastgele orman regresyonu	23
5.3.3 Destek vektör regresyonu.....	24
5.4 Model Performans Ölçüm Kriterleri	26
6. UYGULAMA.....	29
6.1 Enerji Yoğunluğu Grafikleri.....	31
6.2 Beta Regresyon Yöntemi İle Ülkelerin Enerji Yoğunluğunu Etkileyen Faktörlerin Belirlenmesi	33
6.2.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi.....	34
6.2.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi.....	39
6.2.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi.....	45
6.3 Karar Ağacı Yöntemi İle Enerji Yoğunluğunu Etkileyen Faktörlerin Belirlenmesi	51
6.3.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi.....	51
6.3.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi	52

6.3.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi.....	52
6.4 Rastgele orman regresyonu yöntemi ile enerji yoğunluğunu etkileyen faktörlerin belirlenmesi.....	53
6.4.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi	54
6.4.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi	54
6.4.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi	55
6.5 Destek vektör regresyonu yöntemi ile enerji yoğunluğunu etkileyen faktörlerin belirlenmesi.....	56
6.5.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi	57
6.5.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi.....	57
6.5.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi	58
6.6 Farklı Gelir Gruplarındaki Ülkeler için Makine Öğrenmesi Yöntemlerinin Kıyaslanması	59
6.6.1 Alt-orta gelirli ülkeler için makine öğrenmesi yöntemlerinin karşılaştırılması	59
6.6.2 Orta-yüksek gelirli ülkeler için makine öğrenmesi yöntemlerinin karşılaştırılması.....	62
6.6.3 Yüksek gelirli ülkeler için beta regresyon ve makine öğrenmesi yöntemlerinin karşılaştırılması	65
7. SONUÇ VE DEĞERLENDİRME	69
KAYNAKLAR	73
ÖZGEÇMİŞ.....	78

KISALTMALAR DİZİNİ

AIC	Akaike Bilgi Kriteri
BIC	Bayes Bilgi Kriteri
MAE	Ortalama Mutlak Hata
MSE	Hata Kareler Ortalaması
RMSE	Karekök Hata Kareler Ortalaması
R^2	Belirlilik Katsayısı
VIF	Varyans Şişirme Faktörü

ŞEKİLLER DİZİNİ

Şekil 3.1 Beta dağılımının parametreleri $\mu = 0.10, 0.25, 0.50, 0.75, 0.90$, $\phi = 5$ ve $\phi = 100$ iken yoğunluk fonksiyonu grafiği	12
Şekil 5.1 SVR modelinde tahmin doğrusu, $\pm\epsilon$ sınırları ve destek vektörlerinin gösterimi	25
Şekil 6.1 Seçilmiş ülkelerin enerji yoğunluğu grafikleri	32

ÇİZELGELER DİZİNİ

Çizelge 3.1 Çeşitli bağ fonksiyonları, formülleri ve açıklamaları	15
Çizelge 5.1 Çeşitli model performans ölçüm kriterleri	28
Çizelge 6.1 Değişkenlerin tanımları ve kaynakları	30
Çizelge 6.2 Alt-orta gelirli ülkelerin bağ fonksiyonları için bilgi kriteri değerleri.....	34
Çizelge 6.3 Alt-orta gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği.....	35
Çizelge 6.4 Alt-orta ülkelerin log-log bağlantılı ortalama model için katsayıları	36
Çizelge 6.5 Alt-orta gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri	37
Çizelge 6.6 Alt-orta gelirli ülkelerin VIF değerleri	37
Çizelge 6.7 Alt-orta ülkelerin log-log bağlantılı beta ridge ortalama model için katsayıları	38
Çizelge 6.8 Orta-yüksek gelirli ülkelerin bağ fonksiyonları için bilgi kriteri değerleri..	39
Çizelge 6.9 Orta-yüksek gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği.....	40
Çizelge 6.10 Orta-yüksek gelirli ülkelerin cauchit bağlantılı ortalama model için katsayılar.....	41
Çizelge 6.11 Orta-yüksek gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri	42
Çizelge 6.12 Orta-yüksek gelirli ülkelerin VIF değerleri	42
Çizelge 6.13 Orta-yüksek gelirli ülkelerin cauchit bağlantılı beta ridge ortalama model için katsayılar.....	43
Çizelge 6.14 Yüksek gelirli ülkelerin bağ fonksiyonları için bilgi kriterleri değeri	45
Çizelge 6.15 Yüksek gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği.....	46
Çizelge 6.16 Yüksek gelirli ülkelerin cauchit bağlantılı ortalama model için katsayılar	47
Çizelge 6.17 Yüksek gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri	48
Çizelge 6.18 Yüksek gelirli ülkelerin VIF değerleri.....	48
Çizelge 6.19 Yüksek gelirli ülkelerin cauchit bağlantılı beta ridge ortalama model için katsayılar.....	49
Çizelge 6.20 Alt-orta gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri ..	51
Çizelge 6.21 Orta-yüksek gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri	52

Çizelge 6.22 Yüksek gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri...	53
Çizelge 6.23 Alt-orta gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri	54
Çizelge 6.24 Orta-yüksek gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri	55
Çizelge 6.25 Yüksek gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri	55
Çizelge 6.26 Alt-orta gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri	57
Çizelge 6.27 Orta-yüksek gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri	57
Çizelge 6.28 Yüksek gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri	58
Çizelge 6.29 Alt-orta gelirli ülke grubu için uygulanan modellerin performans ölçüm .	59
Çizelge 6.30 Alt-orta gelirli ülke grubu için değişken önem düzeyi değerleri	60
Çizelge 6.31 Orta-yüksek gelirli ülke grubu için uygulanan modellerin performans ölçüm	62
Çizelge 6.32 Orta-yüksek gelirli ülke grubu için değişken önem düzeyi değerleri	63
Çizelge 6.33 Yüksek gelirli ülke grubu için uygulanan modellerin performans ölçüm .	65
Çizelge 6.34 Yüksek gelirli ülke grubu için değişken önem düzeyi değerleri.....	66

1. GİRİŞ

Enerji, günümüz toplumlarının ekonomik, sosyal ve teknolojik kalkınmasının temel dinamiklerinden biri olarak ön plana çıkmaktadır. Tarihsel süreçte özellikle Sanayi Devrimi ile enerjiye olan talebin hızlı bir artış göstermesi, üretim biçimlerinde ve kalkınma modellerinde köklü değişimlerin yaşanmasına neden olmuştur. 21. yüzyıla gelindiğinde ise söz konusu dönüşüm, yalnızca üretim odaklı değil; çevresel sürdürülebilirlik, kaynak yönetimi ve enerji verimliliği boyutlarında çok daha karmaşık ve çok katmanlı bir hâl almıştır. Günümüzde artan dünya nüfusu, hızlı kentleşme süreci, sanayileşmenin yaygınlaşması ve dijital teknolojilerin enerji talebini artırıcı etkisi, enerji tüketiminin hem miktar hem de kaynak türleri açısından büyük bir çeşitlilik kazanmasına yol açmıştır (Çapık vd. 2012). Özellikle gelişmekte olan ülkelerde ekonomik büyüme ile paralel olarak enerji tüketiminde gözlemlenen hızlı artış, enerji kaynaklarının etkin ve verimli kullanımını stratejik bir gereklilik hâline getirmiştir.

Bununla birlikte, söz konusu gelişmelerin çevresel sürdürülebilirlik üzerindeki olumsuz etkileri de ulusal ve uluslararası düzeyde ciddi bir tehdit oluşturmaktadır. Küresel ısınma, sera gazı emisyonlarının artışı, fosil yakıt rezervlerinin tükenme tehlikesi ve iklim değişikliği gibi çevresel problemler, enerji politikalarının yalnızca üretim ve tüketim dengesine odaklanmasının yetersiz olduğunu açıkça ortaya koymaktadır. Bu bağlamda, enerji politikalarının yeniden tasarlanarak çevre dostu, sürdürülebilir ve sosyal refahı gözetilen bir çerçevede ele alınması zorunlu hâle gelmiştir. Özellikle yenilenebilir enerji kaynaklarının daha fazla kullanılması, enerji verimliliğinin artırılması ve enerji yoğunluğunun düşürülmesi gibi politikalar uluslararası gündemde öncelikli bir yer tutmaktadır.

Fosil yakıtların çevre üzerindeki olumsuz etkilerinin giderek daha görünür hâle gelmesi ve bu kaynakların sınırlılığı, ülkeleri alternatif ve sürdürülebilir enerji kaynaklarına yönelmeye zorlamaktadır. Ancak bu dönüşüm süreci, yüksek maliyetler, altyapı eksiklikleri ve teknolojik adaptasyon zorlukları nedeniyle zaman alıcı ve karmaşık bir süreç olarak karşımıza çıkmaktadır. Bu noktada enerji verimliliği ve enerji yoğunluğu

kavramları, sürdürülebilir kalkınma stratejilerinin merkezinde kritik öneme sahiptir. Enerji yoğunluğu, bir ülkenin ekonomik üretim kapasitesi ile enerji tüketimi arasındaki ilişkiyi ortaya koyan ve genellikle Gayri Safi Yurtiçi Hasıla (GSYH) başına tüketilen enerji miktarı ile ölçülen temel bir göstergedir (Karabat ve Aydın, 2018). Enerji yoğunluğunun yüksek olması, enerji kullanımında verimlilik sorununa işaret ederken, düşük enerji yoğunluğu ise daha az enerji harcayarak daha yüksek ekonomik çıktı üretilbildiğini göstermektedir.

Buna karşılık, enerji verimliliği kavramı, mevcut enerji kaynaklarının minimum kayıpla ve maksimum çıktı sağlayacak şekilde kullanılmasını ifade eder ve dolaylı olarak enerji yoğunluğu ile ters orantılıdır (Patterson, 1996; Ediger, 2009). Bu iki kavram, enerji politikalarının değerlendirilmesinde yalnızca teknik boyutuyla değil; aynı zamanda ekonomik kalkınma, çevre koruma ve sosyal refah ekseninde bütüncül bir yaklaşımın temel taşları olarak ele alınmalıdır.

Enerji yoğunluğu, sadece enerji kullanım etkinliğini yansıtan teknik bir gösterge olmakla kalmayıp; aynı zamanda ülkelerin sanayi yapısını, enerji kaynakları kompozisyonunu, ekonomik gelişmişlik düzeylerini ve çevresel politikalarını doğrudan etkileyen çok boyutlu bir kavramdır. Bu nedenle enerji yoğunluğunun azaltılması, özellikle gelişmekte olan ülkelerde çevresel sürdürülebilirlik hedeflerine ulaşılmasında ve enerji bağımlılığının azaltılmasında önemli bir rol oynamaktadır. Avrupa Birliği tarafından belirlenen “20-20-20” hedefleri kapsamında, sera gazı salınımlarının düşürülmesi, enerji verimliliğinin yükseltilmesi ve yenilenebilir enerji kullanımını yaygınlaştırma politikaları, bu kapsamda küresel ölçekte öncü uygulamalar arasında yer almaktadır (Paoli, 2014).

Bu çerçevede hazırlanan tez çalışmasının temel amacı, ülkelerin gelir düzeyleri bağlamında enerji yoğunluğunu etkileyen faktörleri ayrıntılı biçimde incelemek ve bu faktörlerin enerji yoğunluğu üzerindeki etkisini hem klasik istatistiksel yöntemler hem de modern makine öğrenmesi yaklaşımlarıyla karşılaştırmalı olarak analiz etmektir. Çalışmada bağımlı değişken olarak enerji yoğunluğu; bağımsız değişkenler olarak ise

kömür, doğalgaz, petrol, elektrik tüketimi, yenilenebilir enerji kullanımı ve nüfus yoğunluğu ele alınmıştır.

Enerji yoğunluğunun (0,1) aralığında sürekli değerler alması nedeniyle, ilk aşamada Beta regresyon modeli tercih edilmiştir. Beta regresyon, oransal veri yapısına sahip değişkenlerin analizinde klasik doğrusal regresyon yöntemlerine göre daha uygun ve esnek bir modeldir. Analizlerde logit, probit ve log-log gibi çeşitli bağlantı fonksiyonları kullanılmış; model seçiminde Akaike Bilgi Kriteri (AIC), Bayes Bilgi Kriteri (BIC) ve hata ölçütleri (MSE, MAE, RMSE) dikkate alınmıştır. Bağımsız değişkenler arasındaki yüksek korelasyonun (multicollinearity) katsayı tahminlerinde sapmaya yol açabileceği göz önünde bulundurularak, Beta regresyonun ardından Beta ridge regresyon yöntemi de uygulanmıştır. Beta ridge regresyon, katsayılara ceza terimi ekleyerek tahminlerdeki oynaklığı azaltmakta ve değişkenlerin model üzerindeki etkilerini daha dengeli hale getirmektedir. Bu sayede enerji yoğunluğunu belirleyen faktörler daha güvenilir biçimde analiz edilmiştir.

Çalışmanın üçüncü aşamasında ise, yüksek boyutlu ve karmaşık veri setlerinde güçlü tahmin performansı ile öne çıkan makine öğrenmesi yöntemleri uygulanmıştır. Bu kapsamda Rastgele Orman Regresyonu (Random Forest Regression), Destek Vektör Regresyonu (Support Vector Regression) ve Karar Ağaçları (Decision Trees) olacak şekilde üç farklı makine öğrenmesi algoritması seçilmiştir. Çalışmada kullanılan bu makine öğrenmesi tekniklerinin tercih edilmesinin temel nedeni, bahsi geçen algoritmaların yüksek boyutlu ve karmaşık veri setlerindeki doğrusal olmayan ilişkileri güçlü bir tahmin yeteneğiyle yakalayabilmesidir. Karar Ağaçları, verideki önemli özelliklerin ve bunların bağımlı değişken üzerindeki etkilerinin kolayca yorumlanabilmesini sağlarken, Rastgele Orman algoritması birden fazla karar ağacını bir araya getirerek aşırı öğrenmeyi (overfitting) azaltır ve genelleme yeteneğini önemli ölçüde artırır. Destek Vektör Regresyonu ise verideki aykırı değerlere karşı daha dayanıklı bir yapı sunar ve karmaşık veri örüntülerini yakalamak için farklı matematiksel yaklaşımlar kullanır.

Analizlerde kullanılan veri seti, 1990–2021 yılları arasında 8 farklı ülkenin yıllık verilerinden oluşmaktadır. Ülkeler, Dünya Bankası'nın gelir sınıflandırması temel alınarak alt-orta, orta-yüksek ve yüksek gelir grupları olmak üzere üç kategoriye ayrılmıştır. Bu sınıflandırma sayesinde enerji yoğunluğunu etkileyen faktörlerin yalnızca genel düzeyde değil, aynı zamanda gelir düzeyine bağlı olarak nasıl farklılaştığı da ortaya konulmuştur. Analiz sürecinde, Beta regresyon, Beta ridge regresyon ve Makine Öğrenmesi yöntemleri her bir gelir grubu için ayrı ayrı uygulanmış ve elde edilen bulgular karşılaştırmalı olarak değerlendirilmiştir.

Tüm analiz süreçlerinde R ve Python programlama dillerinin sunduğu istatistiksel ve makine öğrenmesi kütüphanelerinden etkin şekilde yararlanılmış; veri temizleme, görselleştirme, hata analizi ve model performans değerlendirmeleri bu yazılımlar aracılığıyla gerçekleştirilmiştir.

Sonuç olarak, bu tezde benimsenen kapsamlı ve çok yönlü modelleme yaklaşımı sayesinde enerji yoğunluğunu etkileyen faktörler hem doğrusal hem de doğrusal olmayan yapılar çerçevesinde analiz edilmiştir. Bu kapsamlı analizlerin, enerji verimliliği odaklı politika geliştirmek isteyen karar vericilere yol gösterici olması; özellikle sürdürülebilir kalkınma, çevresel koruma ve enerji güvenliği alanlarında daha etkili stratejilerin hayata geçirilmesine katkı sağlaması beklenmektedir.

2. KURAMSAL TEMELLER ve KAYNAK ÖZETLERİ

Enerji yoğunluğu, bir ülkenin ekonomik faaliyetlerini sürdürmesi için ihtiyaç duyduğu enerji tüketimini ortaya koyan kritik bir verimlilik göstergesidir. Bu bakımdan, fosil yakıt tüketimi (kömür, petrol, doğalgaz), elektrik tüketimi, nüfus yoğunluğu, yenilenebilir enerji kullanımı ve buna benzer faktörlerin enerji yoğunluğu üzerindeki etkisi sürekli araştırma konusu olmuştur. Literatürde, bu değişkenlerin enerji verimliliği, çevresel sürdürülebilirlik ve makroekonomik göstergelerle olan ilişkisi farklı ülke grupları ve dönemler için çeşitli yöntemlerle ele alınmaktadır.

Enerji politikaları üzerine yaptığı çalışmada Pamir (2005), enerji fiyatları ve teknolojik gelişmelerin enerji yoğunluğunu azalttığını belirtmiştir. Ancak fosil kaynak rezervlerinin keşfiyle birlikte bu kaynakların kullanımının artacağı ve buna bağlı olarak enerji yoğunluğunun yükselebileceği öngörülmüştür. Ayrıca hidroelektrik ve yenilenebilir enerji kaynaklarının artması beklense de bu kaynakların toplam enerji tüketimindeki payının kayda değer biçimde artmayacağı ifade edilmiştir.

Bildirici ve Bakırtaş (2013), BRİCTS¹ ülkeleri üzerine yaptıkları çalışmada, 1980–2011 yılları arasında kömür, doğalgaz ve petrol tüketimi ile ekonomik büyüme arasındaki nedensellik bağına incelemişlerdir. Araştırmaları, özellikle Brezilya, Rusya ve Türkiye'de doğalgaz kullanımının enerji yoğunluğunu düşürdüğünü ve enerji verimliliğini yükselttiğini göstermiştir.

¹ Brezilya, Rusya, Hindistan, Çin, Güney Afrika Cumhuriyeti, Mısır, Etiyopya, Endonezya, İran ve Birleşik Arap Emirlikleri

Gelişmekte olan 76 ülkeyi kapsayan araştırmasında Sadorsky (2013), gelir, kentleşme ve sanayileşmenin enerji yoğunluğu üzerindeki etkilerini heterojen panel regresyon yöntemleriyle analiz etmişlerdir. Çalışmanın bulguları, gelirdeki %1’lik artışın enerji yoğunluğunu %0,35 ile %0,45 oranında azalttığı ifade edilmiştir.

Türkiye’nin enerji verimliliği potansiyelini değerlendiren çalışmasında Doğan ve Yılankırkan (2015), enerji tüketim miktarları, sektörel dağılım, verimlilik teknolojileri gibi unsurları ele alarak sanayi, konut ve ulaşım sektörlerinde enerjinin verimli kullanımının çevresel koruma ve ekonomik katkı sağlayabileceğini belirtmiştir.

Avrupa Birliği ülkeleri üzerine yaptığı araştırmasında Balitsky vd. (2016), ekonomik büyümenin doğalgaz tüketimini artırdığını; ancak doğalgaz tüketiminin büyüme üzerinde olumsuz etki oluşturduğunu saptamıştır. Bu bulgulara ek olarak doğalgaz tüketimi ile ekonomik büyüme arasında uzun dönemde iki yönlü bir nedensellik ilişkisi ortaya konulmuştur.

Osorio vd. (2017), kentleşme biçimi ve enerji tüketimi arasındaki ilişkiyi incelemiş ve yüksek nüfus yoğunluğuna sahip bölgelerde kişi başına enerji tüketiminin daha düşük olduğunu, böylece enerji verimliliğinin arttığını ortaya koymuştur.

Asya ülkelerini inceleyen Bilgili vd. (2017), 1990–2014 döneminde kentleşmenin enerji yoğunluğu üzerindeki etkisini panel veri yöntemiyle analiz etmiş ve kişi başına GSYH, kentleşme ile kırsallaşma arasındaki uzun süreli bir ilişkinin varlığı gösterilmiştir. Kentleşmenin hem kısa hem uzun vadede enerji yoğunluğunu azalttığı bulunmuştur.

Enerji yoğunluğundaki iyileşmelerin ve yenilenebilir enerjiye geçişin ekonomik büyüme üzerindeki etkilerini analiz eden çalışmasında Diaz vd. (2019), 1960–2010 dönemine ait 134 ülkenin verilerini kullanmıştır. Elde edilen bulgular, enerji yoğunluğunun azalmasının büyümeyi teşvik ettiğini ve fosil yakıtlardan rüzgâr, güneş ve jeotermal gibi yenilenebilir kaynaklara geçişin ekonomik büyümeyle pozitif ilişkili olduğunu göstermektedir.

Deicman vd. (2019) tarafından yayımlanan çalışmada, 137 ülkeyi kapsayan analizinde, kişi başına gelir arttıkça enerji yoğunluğunun azaldığını; ancak 5 bin dolar gelir eşiğinden sonra bu azalışın %30'dan fazla yavaşladığını belirtmiştir. Yapısal değişimin tüm ülkelerde önemli bir faktör olduğu; verimlilik artışının ise özellikle yüksek gelirli ülkelerde belirleyici olduğu vurgulanmıştır.

Ünüvar ve Keskinkılıç (2020) tarafından yapılan yenilenebilir enerji ile ekonomik büyüme ilişkisi üzerine olan çalışmada, yenilenebilir enerji üretiminin ekonomik büyümeyi olumlu yönde etkilediği sonucuna varılmıştır.

Fosil yakıt desteklerinin enerji yoğunluğu üzerindeki etkilerini ele alan çalışmasında Solarin (2020), bu desteklerin enerji yoğunluğunu artırdığını belirtmiştir. Ayrıca kentleşme, sanayileşme ve doğrudan yabancı yatırımın enerji yoğunluğunu artırıcı; kişi başına GSYH ve insan sermayesi gelişiminin ise azaltıcı etkiye sahip olduğu ifade edilmiştir.

Kenya özelinde çevresel Kuznets eğrisi hipotezini test eden çalışmasında Sarkodie ve Öztürk (2020), 1971–2013 dönemi verileriyle yaptığı analizlerde, enerji tüketiminin uzun vadede CO₂ emisyonlarını artırdığını; yenilenebilir kaynaklardan elde edilen elektriğin ise bu emisyonları azalttığını ortaya koymuştur. Enerji ithalatının uzun vadeli bir etkisi bulunmazken, enerji tasarrufu bilincinin artırılması gerektiği vurgulanmıştır.

G7 ülkeleri üzerine yaptığı çalışmada Chu ve Le (2021), çevresel kalite üzerinde etkili olan ekonomik karmaşıklık, politika belirsizliği, yenilenebilir enerji ve enerji yoğunluğu gibi faktörleri panel veri analiziyle test etmiştir. Çalışmada, yüksek enerji yoğunluğunun çevresel bozulmayı artırdığı, buna karşılık yenilenebilir enerjinin çevresel kaliteyi iyileştirdiği belirlenmiştir.

Yüksek emisyon üreten beş ülkeyi kapsayan çalışmasında Kartal (2022), kömür, petrol ve doğalgaz tüketiminin karbon salımı üzerinde en güçlü etkiye sahip faktörler olduğunu ortaya koymuştur. Ayrıca, değişkenlerin etkisinin ülkelere göre farklılık gösterdiği ve tek tip çevre politikalarının yetersiz kalabileceği belirtilmiştir.

Gelişmekte olan 23 ülkeyi inceleyen Eyyüboğlu vd. (2022), enerji yoğunluğu ve kentleşme arasında uzun dönemde herhangi bir ilişkinin varlığını tespit edememişlerdir. Bununla birlikte, kısa dönemde bu iki değişken arasında karşılıklı bir nedensellik ilişkisi gözlemlenmiştir.

Macaristan'da enerji verimliliğini etkileyen faktörler Naimoğlu vd. (2022) tarafından incelenmiştir. Çalışmada, kişi başına düşen GSYH ve sanayileşmenin enerji verimliliğinde artışa yol açtığı, buna karşılık fosil yakıt tüketiminin ise verimliliği düşürdüğünü ortaya koymuştur.

Enerji ithalatçısı ülkelerde enerji yoğunluğunu etkileyen faktörleri analiz eden Özbek (2023), uzun dönemde fosil yakıt kullanımının enerji yoğunluğunu artırdığını; sanayileşme, yenilenebilir enerji tüketimi, enerji fiyatı ve kentleşmenin ise enerji yoğunluğunun azaltıcı etkisini ortaya koymuştur.

Enerji kaynaklarının enerji yoğunluğu üzerindeki etkilerini inceleyen bir başka çalışmada Naimoğlu ve Özel (2022), 16 yükselen ekonomi için panel veri analizi uygulamış ve hidroelektrik ile elektrik tüketiminin enerji yoğunluğunu azalttığını; kömür, petrol ve doğalgaz kullanımının ise artırdığını göstermiştir.

Enerji yoğunluğu, ekonomik büyüme ve yönetim ilişkisini 182 ülke kapsamında ele alan Ernawati vd. (2022), yüksek gelirli ülkelerde enerji kullanımının daha verimli olduğunu ve yönetim kalitesinin enerji yoğunluğunu azalttığını saptamıştır. Ayrıca, enerji yoğunluğu ile ekonomik büyüme arasında pozitif; yönetim ile negatif korelasyon tespit edilmiştir.

Yenilenebilir enerji tüketiminin CO₂ emisyonları üzerindeki etkisini inceleyen çalışmasında Mirziyoyeva ve Salahodjaev (2022), yüksek karbon yoğunluklu ülkelerde yenilenebilir enerjideki %1’lik artışın emisyonları %0,98 oranında azalttığını ortaya koymuştur. Bunun yanında, kişi başına gelir ve kadın temsilindeki artışların çevresel kaliteyi artırdığı belirlenmiştir.

Enerji yoğunluğunu etkileyen faktörleri ele alan Gyamfi vd. (2022), temiz enerji ve gelirin enerji yoğunluğunu azalttığını; buna karşılık ticaret, teknolojik yenilik ve yenilenemeyen enerji tüketiminin artırıcı etki yaptığını ortaya koymuştur.

Yenilenebilir enerjinin çevresel kalite üzerindeki etkilerini analiz eden çalışmasında Gao vd. (2023), bu enerji türünün ulaşım ve üretim kaynaklı çevresel kaliteyi hem kısa hem de uzun vadede olumlu etkilediğini; yenilenemeyen enerji kaynaklarının ise uzun vadede her iki alanda da çevresel kaliteyi olumsuz etkilediğini belirtmiştir. Kısa vadede ise bu etkinin yalnızca üretim kaynaklı olduğu saptanmıştır.

Literatür taraması, enerji yoğunluğunu etkileyen faktörlerin ülkeden ülkeye ve zaman içerisinde değişkenlik gösterdiğini ortaya koymaktadır. Bununla birlikte, fosil yakıt tüketiminin genel olarak enerji yoğunluğunu artırdığı; yenilenebilir enerji kullanımının, sanayileşme düzeyinin ve kişi başına gelirin ise azaltıcı yönde etkide bulunduğu yönünde bulgular yaygındır. Bu çalışma, mevcut literatürü temel alarak kömür, petrol, doğalgaz, elektrik ve yenilenebilir enerji kaynaklarının yanı sıra nüfus yoğunluğunun ülkelerin enerji yoğunluğu üzerindeki etkilerini analiz etmeyi amaçlamaktadır.

3. BETA REGRESYON ANALİZİ

Günümüzde sosyal bilimlerden ekonomiye, sağlık araştırmalarından mühendislik uygulamalarına kadar pek çok disiplinde, bağımlı değişkenin 0 ile 1 aralığında sürekli değerler aldığı veri yapılarıyla sıkça karşılaşilmektedir. Ancak bu tür veri yapılarında, klasik doğrusal regresyon modelleri bağımlı değişkenin sınır kısıtlamalarını ve değişken varyans özelliklerini göz ardı ettiği için yetersiz kalmakta ve yanıltıcı sonuçlara yol açabilmektedir. Bu gereksinimden hareketle geliştirilen beta regresyon modeli, bağımlı değişkenin (0,1) aralığında tanımlı olduğu ve varyansın ortalamaya bağlı olarak değişkenlik gösterdiği durumlar için etkili bir modelleme yaklaşımı sunmaktadır (Ferrari ve Cribari-Neto 2004). Beta dağılımı temelinde yapılandırılan bu model hem bağımlı değişkenin ortalamasını hem de varyans yapısını esnek bir biçimde modelleyebilmesi sayesinde, çeşitli araştırma alanlarında geniş bir uygulama alanı bulmuştur (Espinheira vd. 2008).

Beta regresyon analizinin tercih edilmesinin temel nedeni, bağımlı değişkenin doğası gereği (0,1) aralığında sürekli değerler alması ve bu yapıdaki verilerin beta dağılımı ile uyumlu özellikler göstermesidir. Beta dağılımı, şekil parametreleri aracılığıyla esnek bir şekilde farklı dağılım biçimlerini modelleyebilmesi sayesinde, özellikle oranlar, olasılıklar ve başarı oranı gibi değişkenlerin analizinde etkin bir araç sunmaktadır. Beta dağılımının olasılık yoğunluk fonksiyonu aşağıdaki şekilde ifade edilir:

$$\pi(y; p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1}, \quad 0 < y < 1. \quad (2)$$

Burada $p > 0$, $q > 0$ ve $\Gamma(\cdot)$ gama fonksiyonudur. Fonksiyonun beklenen değeri ve varyansı sırasıyla aşağıdaki gibidir:

$$E(Y) = \frac{p}{p+q} \quad (3)$$

ve

$$V(Y) = \frac{pq}{(p+q)^2(p+q+1)} \quad (4)$$

Burada amaç beta dağılmış rastgele değişkenlerden bir regresyon modeli oluşturmaktır. Beta dağılımının olasılık yoğunluk fonksiyonu (2) de verilmiş ve dağılımın beklenen değeri ve varyansı p ve q parametreleri ile gösterilmiştir. Fakat oluşturulacak regresyon modelini beklenen değeri kullanarak bir parametre dönüşümü uygulanırsa daha açıklayıcı olacağından,

$$\mu = \frac{p}{p+q}$$

ve

$$\phi = p + q$$

dönüşümleri uygulanır ve aşağıda verilen eşitlikler elde edilir:

$$p = \mu\phi \quad \text{ve} \quad q = (1 - \mu)\phi$$

Buradan;

$$E(Y) = \frac{\mu\phi}{\mu\phi + (1 - \mu)\phi} = \frac{\mu\phi}{\phi(\mu + 1 - \mu)} = \mu$$

ve

$$Var(Y) = \frac{(\mu\phi)(1 - \mu)\phi}{(\mu\phi + (1 - \mu)\phi)^2(\mu\phi + (1 - \mu)\phi + 1)}$$

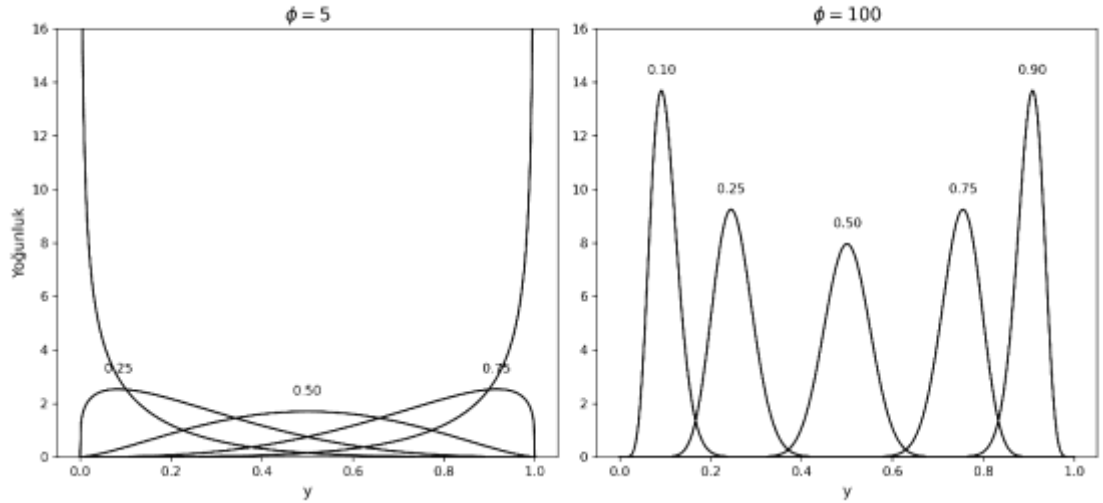
$$= \frac{(\mu\phi)(1-\mu)\phi}{((\mu\phi) + (\phi - \mu\phi))^2(\mu\phi + \phi - \mu\phi + 1)} = \frac{(\mu\phi)(\phi - \mu)}{\phi^2(\phi + 1)}$$

$$\Rightarrow Var(Y) = \frac{\mu\phi(1-\mu)}{\phi(\phi + 1)} = \frac{\mu(1-\mu)}{1 + \phi}$$

olur ve $E(Y)$ ve $Var(Y)$ eşitlik (2)'de verilen olasılık yoğunluk fonksiyonunda yerine koyulursa yeni olasılık yoğunluk fonksiyonu

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad 0 < y < 1 \quad (5)$$

şeklinde gösterilir, $0 < \mu < 1$ ve $\phi > 0$ 'dır. Beta regresyon modelinde, aynı ortalama değerine sahip gözlemler varyans yapılarının (hassasiyet parametresi ϕ) farklılığı nedeniyle oldukça farklı olasılık dağılımları gösterebildiğinden, yoğunluk değerleri birbirine yakın olsa dahi model tahminlerinde belirgin farklılıklar oluşabilmektedir.



Şekil 3.1 Beta dağılımının parametreleri $\mu = 0.10, 0.25, 0.50, 0.75, 0.90$, $\phi = 5$ ve $\phi = 100$ iken yoğunluk fonksiyonu grafiği

Özellikle $\mu = 1/2$ iken simetrik ve $\mu \neq 1/2$ iken asimetriktir. Ayrıca μ sabitken, ϕ arttıkça dağılımın sivrililiğinin arttığını Şekil 3.1’ de gözlemlenebilir. $\mu = 1/2$ ve $\phi = 2$ iken dağılım düzgün dağılım olur.

$y_1, y_2, \dots, y_n \sim B(\mu, \phi)$ [y, B(μ, ϕ)] beta yoğunluğunda tesadüfi örneklem olsun. O zaman beta regresyonu;

$$g(\mu_i) = \beta_0 + x_{i1}\beta_1 + \dots + x_{ik}\beta_k = \eta_i \quad i = 1, 2, \dots, n \quad (6)$$

$x_{i1}, \dots, x_{ik}$ ’ler bağımsız (açıklayıcı) değişkenler, $\beta_0, \beta_1, \dots, \beta_k$ ise her bir bağımsız (açıklayıcı) değişkene karşılık gelen tahmin edilen kesişim ve katsayılarıdır. n örneklem genişliği ve η_i ’ de gözlemin doğrusal kestiricisidir. Burada g(.) doğrusal kestirici ve bağımlı değişkeni bağlayan bağıntı fonksiyonudur (Yellareddy vd. 2016).

Eşitlik (5) temel alınarak, beta regresyon modeli, beta dağılımı ile kurulan log-olabilirlik fonksiyonu aracılığıyla oluşturulur. Bu model için elde edilen log-olabilirlik fonksiyonu ise şöyledir:

$$L(\mu, \delta) = \log \Gamma(\delta) - \log \Gamma(\mu\delta) - \log \Gamma((1 - \mu)\delta) + (\mu\delta - 1) \log y + ((1 - \mu)\delta - 1) \log(1 - y) \quad (7)$$

Regresyon katsayıları, log-olabilirlik fonksiyonunun maksimum değerini verecek en çok olabilirlik tahminlerinin hesaplanmasıyla elde edilir (Zhao vd. 2014).

3.1 Bilgi Kriterleri

İstatistiksel model seçiminde birçok bilgi kriteri kullanılmaktadır. Bu çalışmada, model seçimi için Akaike Bilgi Kriteri (AIC) ve Bayes Bilgi Kriteri (BIC) esas alınacaktır. Karşılaştırmalı değerlendirmelerde, AIC ve BIC değerleri daha düşük olan modeller, veriyle daha iyi uyum sağladıkları için tercih edilmektedir. Dolayısıyla, çalışmada en iyi

performansı gösteren model, bilgi kriteri deęerleri en dūřuk olan model olarak belirlenecektir.

Model seiminde kullanılacak bilgi kriterlerinin matematiksel ifadeleri ařaęıda sunulmaktadır:

$$AIC = -2 \ln(L) + 2d$$
$$BIC = -2 \ln(L) + d \ln(n)$$

Burada

d: Toplam parametre sayısı

n: rneklem byklę

L: En ok olabilirlik fonksiyonunu gstermektedir.

3.2 Beta Regresyon Analizinde Kullanılan Baęlantı Fonksiyonları

Beta regresyon analizinde, baęımlı deęiřkenin sınırlı bir aralıkta deęerler alması, verinin doęru biimde modellenebilmesi iin uygun baę fonksiyonunun seilmesini kritik hle getirmektedir. Baęlantı fonksiyonu, baęımlı deęiřkenin beklenen deęeri ile doęrusal tahminleyici arasındaki iliřkiyi kurar. Doęru fonksiyonun seilmesi, modelin hem uyum bařarısını hem de tahmin performansını doęrudan etkiler.

Bu tez alıřmasında, kullanılacak baę fonksiyonunun belirlenmesinde Akaike Bilgi Kriteri (AIC) ve Bayes Bilgi Kriteri (BIC) esas alınacaktır. Farklı baęlantı fonksiyonlarıyla kurulan modellerin performansları bu kriterler aracılıęıyla karřılařtırılacak ve en dūřuk AIC ve BIC deęerlerine sahip fonksiyon tercih edilen model olarak belirlenecektir.

Beta regresyon analizinde literatrde yaygın olarak kullanılan bařlıca baę fonksiyonları ve bunlara ait matematiksel ifadeler izelge 3.1’de sunulmaktadır.

Çizelge 3.1 Çeşitli bağ fonksiyonları, formülleri ve açıklamaları

Bağ Fonksiyonu	Formül	Açıklama
Logit	$\log\left(\frac{\mu}{1-\mu}\right)$	Olasılık oranının (odds) logaritmasını alarak (0,1) aralığındaki değerleri tüm reel sayılar kümesine dönüştürür. Simetrik yapısı nedeniyle sık tercih edilir.
Log-log	$-\log(-\log(\mu))$	Olasılığın logaritmasının negatif logaritmasını alarak dönüşüm yapar. Özellikle sağa çarpık dağılımlar için uygundur.
Tamamlayıcı log-log	$\log(-\log(1-\mu))$	Olasılığın tamamlayıcısının logaritmasının negatif logaritmasını alır. Asimetrik yapıda olup üst sınıra yaklaşan olasılıkların modellenmesinde avantaj sağlar.
Probit	$\Phi^{-1}(\mu)$	Standart normal dağılımın ters kümülatif dağılım fonksiyonunu kullanır. Simetrik yapısı logit fonksiyonuna benzer, ancak uç değerlerde daha yumuşak geçiş sağlar.
Cauchit	$\tan(\pi(\mu - 0.5))$	Cauchy dağılımına dayalı dönüşüm sağlar. Aykırı değerlere karşı dayanıklıdır ve uç olasılıkların modellenmesinde kullanılabilir.
Log	$\log(\mu)$	Olasılığın logaritmasını alır ve çarpık dağılımlar için kullanılabilir.

4. BETA RIDGE REGRESYON ANALİZİ

Ağırlıklı hata kareler toplamını minimize etmek amacıyla $\widehat{\beta}$, β vektörünün bir tahmin edicisi olduğunu varsayalım. Bu durumda hata kareler toplamı şu şekilde yazılabilir:

$$\begin{aligned}\theta &= (y - \widehat{\beta})'(y - \widehat{\beta}) \\ &= (y - X\widehat{\beta}_{MLE})'(y - X\widehat{\beta}_{MLE}) + (\widehat{\beta} - \widehat{\beta}_{MLE})'X'\widehat{W}X\widehat{\beta} - \widehat{\beta}_{MLE}) \\ &= \theta_{min} + \theta(\theta)\end{aligned}$$

Burada $\theta(\widehat{\beta}) > 0$ ise $\widehat{\beta}_{MLE}$ yerine $\widehat{\beta}$ konulduğunda ağırlıklı hata kareleri toplamındaki artışı ifade eder. Beta ridge tahmin edicisi, Hoerl ve Kennard (1970a) tarafından önerilen bir kısıtlama altında elde edilir. Yani, $(\widehat{\beta} - \widehat{\beta}_{MLE})'X'\widehat{W}X(\widehat{\beta} - \widehat{\beta}_{MLE}) = \theta_0$ kısıtlamasına bağlı olarak $\widehat{\beta}'\widehat{\beta}$ 'nin minimize edilmesiyle elde edilir (Gedikli 2024).

Toplam hatayı temsil eden minimize yapılan Q fonksiyonu tanımlanırsa;

$$\begin{aligned}\text{Minimize } (Q) &= (y - \widehat{\beta})'(y - \widehat{\beta}) \\ &= (y - X\widehat{\beta}_{MLE})'(y - X\widehat{\beta}_{MLE}) + (\widehat{\beta} - \widehat{\beta}_{MLE})'X'\widehat{W}X\widehat{\beta} - \widehat{\beta}_{MLE}) \\ Q &= \widehat{\beta}'\widehat{\beta} + \frac{1}{k}((\widehat{\beta} - \widehat{\beta}_{MLE})'X'\widehat{W}X(\widehat{\beta} - \widehat{\beta}_{MLE}) - \theta_0)\end{aligned}\quad (8)$$

Burada $\frac{1}{k}$ Lagrange çarpanı olarak adlandırılır. Denklem (8)'in yerine $\widehat{\beta}$, 'ye göre türevi alınarak 0'a eşitlenirse beta ridge tahmin edicisi aşağıdaki şekilde elde edilir:

$$\widehat{\beta}_{RE} = (X'\widehat{W}X + kI)^{-1}X'\widehat{W}X\widehat{\beta}_{MLE}$$

X : Bağımsız değişken matrisi

$\widehat{\beta}_{MLE}$ = En çok olabilirlik tahmin edicisi ile elde edilen beta parametre vektörü

$\widehat{W}$ = Ağırlık matrisinin tahmini

I : Birim matris

k : Yanlılık parametresi

5. MAKİNE ÖĞRENMESİ YÖNTEMLERİ

Makine öğrenmesinin temelinde öğrenme kavramı yer almaktadır. Öğrenme kavramı farklı şekillerde tanımlanmakla birlikte, en çok kullanılan tanımı Simon tarafından yapılmıştır. Simon, öğrenmeyi “zaman içinde yeni bilgilerin keşfedilmesi yoluyla davranışların iyileştirilmesi süreci” olarak tanımlamıştır (Öztemel 2006). Bu öğrenme sürecinin bilgisayarlar aracılığıyla yapılmasını amaçlayan makine öğrenmesi, bilginin örnekler üzerinden kazanımına dayanır. Bir bilgisayar, ilk olarak öğrenmesi gereken bilgiden bir örnek alır ve buradan bazı veriler öğrenir; sonrasında ikinci bir örnek seçerek bu yeni örnekten ilave bilgiler edinir. Tüm örnekleri tekrar tekrar gözden geçirerek olayla ilgili bir genelleme yapar ve bu süreç sonunda bir makine öğrenmesi sistemi oluşturur.

Makine öğrenmesi, bilgisayar bilimlerinde en hızlı büyüyen alanlardan biridir. Bunun nedeni yalnızca verilerin sürekli olarak "daha büyük" hale gelmesi değil, aynı zamanda bu verileri işleyerek bilgiye dönüştürecek teorilerin de gelişmesidir. Birçok bilim dalında ve günlük hayatta dijital teknolojinin etkisinin giderek artması ve dijital ayak izimizin derinleşmesiyle birlikte, sürekli olarak daha fazla veri üretilmekte ve toplanmaktadır (Alpaydın 2014).

Makine öğrenmesi uygulamaları beş temel aşamada gerçekleştirilmektedir. İlk aşama, veri toplama aşamasıdır. Bu aşamada, ilgilenilen konu ile alakalı veriler excel dosyaları, metin dosyaları veya Uygulama Programlama Arayüzü (API) gibi çeşitli ortamlardan ham veri olarak toplanır. Verilerin toplanması, öğrenme aşamasında bilgisayarın başarılı olabilmesi için oldukça önemlidir.

İkinci aşama, verilerin hazırlanmasıdır. Bu, en fazla insan müdahalesi gerektiren ve en çok zaman harcanan aşamadır. Verilerin incelenecek konudan sapmaması, eksiksiz olması ve gereksiz verilerin temizlenmesi bu aşamada gerçekleştirilir.

Üçüncü aşama, modelin eğitilmesidir. Veri hazırlama sürecinde, bilgisayarın öğrenmesi gereken bilgiler hakkında yeterli bilgi edinilir. Bu aşamada uygun bir algoritma seçilerek

veriler analiz edilebilir bir modele dönüştürülür. Denetimli öğrenme algoritmaları kullanılıyorsa, veriler eğitim ve test setleri olarak ikiye ayrılır ve öğrenme bu şekilde sağlanır.

Dördüncü aşama, modelin değerlendirilmesidir. Bu aşamada, modelin başarısı değerlendirilir ve algoritmanın öğrenme sürecinden ne kadar iyi sonuç çıkardığı ölçülür. Kullanılan modele bağlı olarak, modelin doğruluğu test veri setiyle ölçülebilir veya farklı performans metrikleri kullanılarak analiz edilebilir. Böylece hem modelin doğruluğu hem de seçilen algoritmanın uygunluğu belirlenir.

Beşinci ve son aşama, performans artırmadır. Eğer model değerlendirildikten sonra performansın iyileştirilmesi gerekiyorsa, modele farklı değişkenler ilave edilebilir. Ayrıca mevcut model yetersiz kalıyorsa, tamamen farklı bir model türüne geçmek ve süreci en baştan başlatmak da bir seçenek olabilir (Selvi 2019).

Makine öğrenmesi yöntemleri, öğrenme biçimlerine göre denetimli, denetimsiz ve pekiştirmeli öğrenme olmak üzere üç ana başlık altında sınıflandırılır.

5.1 Denetimsiz Öğrenme

Denetimsiz öğrenme, sisteme yalnızca girdi değerlerinin sunulduğu bir yaklaşımdır. Bu yöntemde, girilen örneklemdaki parametreler arasındaki ilişkilerin sistem tarafından bağımsız olarak öğrenilmesi beklenir. Kümeleme ve boyut indirgeme denetimsiz öğrenme problemine örneklerdir. Kümeleme, verilerin doğal gruplar halinde ayrılmasını sağlar. Boyut indirgeme ise yüksek boyutlu verilerde önemli bilgileri koruyarak veri boyutunu azaltır (Mohri 2018).

5.2 Pekiştirmeli Öğrenme

Dinamik programlama ile denetimli öğrenme yaklaşımlarını bir araya getiren bir yöntemdir. Bu öğrenme türünde, öğreniciye bir hedef belirlenir ve öğrenici, hedefe ulaşmanın yollarını deneme-yanılma yöntemiyle kendi başına keşfeder.

Pekiştirmeli öğrenme teknikleri arasında, bagging, boosting ve stacking bulunmaktadır (Islam 2020).

5.3 Denetimli Öğrenme

Makine öğrenmesi çeşitleri arasında en sık rastlanan ve en çok tercih edilen yöntemdir. Öğrenilecek olaya dair örnekler, girdi-çıkı setleri şeklinde sisteme sunulur. Bu yöntemde sistemin hedefi, belirtilen girdileri ilgili çıktılarla eşleştirerek, olayın girdi ve çıktıları arasındaki bağıntıları kavramaktır (Şahin 2022).

Denetimli öğrenme algoritmaları Sınıflandırma ve Regresyon olarak ikiye ayrılır. Sınıflandırma algoritma çeşitleri İkili Sınıflandırma, Çoklu Sınıf Sınıflandırma ve Çoklu Etiket Sınıflandırmadır. Regresyon algoritmalarına verilebilecek örnekler ise; En Yakın Komşuluk, Destek Vektör Mekanizmaları, Yapay Sinir Ağları, Rastgele Orman, Karar Ağaçları, Doğrusal Regresyon ve Lojistik Regresyon verilebilir.

Bu tez çalışmasında denetimli öğrenme algoritmalarından karar ağaçları, rastgele orman ve destek vektör regresyonu yöntemleri kullanılacaktır.

5.3.1 Karar ağaçları

Karar ağaçları, denetimli öğrenme yöntemlerinden biridir. Sık kullanımlarının temel nedeni, genellikle doğrudan uygulanabilir olmalarıdır. Karar ağaçları az sayıda parametreye dayanır ve veri dönüşümlerinden etkilenmezler. Bu özellik, eğitim sırasında

veri özelliklerini derinlemesine analiz etme ihtiyacını azaltarak modeli pratik hale getirir. Ayrıca, tek bir karar ağacı, yorumlanması kolay olan az sayıdaki makine öğrenmesi modellerinden biridir.

Bununla birlikte, birden fazla ağacın birleşimiyle oluşturulan topluluk modellerinde (ensemble methods) modelin iç işleyişi ve karar verme süreçleri daha karmaşık hale gelir. Karar ağaçları, genel veri yapısından çıkarımlarda bulunarak verileri birbiriyle ilişkili birçok bölgeye ayırır ve böylece karmaşık regresyon eğrileri oluşturabilir.

Bir karar ağacı modelinde, kök düğümde giriş değişkenleri bulunur. Ardından, bağımlı değişkenleri en iyi şekilde ayırt etmek amacıyla bu değişkenlerin değerlerine dayalı kurallar oluşturulur. Bağımsız değişkenler (X) ile karşılık gelen bağımlı değişkenler (Y) verildiğinde, eğitilmiş regresyon ağacı modeli bu veri çiftlerini belirli kurallara göre gruplara ayırır. Yaprak düğümler bu (X, Y) çiftlerini çakışmayan alt gruplara böler; yani her bir (X, Y) çifti yalnızca tek bir yaprağa ait olur. Böylece, model benzer özelliklere sahip veri noktalarını aynı yaprakta toplayarak daha doğru tahminler yapmayı amaçlar.

Bir karar ağacı oluşturulurken, veriyi bölen her adımın ne kadar uygun olduğunu değerlendirmek için bir hata ölçütü kullanılır. Sınıflandırma ağaçları için en yaygın kullanılan ölçütler Gini safsızlığı ve entropi iken (Saad, 2022), bu tez çalışmasında regresyon ağaçları kullanılacağı için ortalama kare hata (MSE) ölçütü tercih edilmiştir. Regresyon ağaçlarında, tahmin edilen değer ile gerçek değer arasındaki farkın karesi alınarak her dalda hata hesaplaması gerçekleştirilir. Amacımız, tüm veri noktaları için hata karelerinin toplamını minimize eden bir model oluşturmaktır.

Matematiksel olarak, m düğümü için X_m , X 'in m düğümüne ulaşan alt kümesi olarak tanımlanır. Yani kökten m düğümüne kadar olan yoldaki karar düğümlerinde tüm koşulları sağlayan tüm $x \in X$ elemanları X_m kümesine dahildir. Buna göre:

$$b_m(x) = \begin{cases} 1, & x \in X_m \\ 0 & \text{diğer durumlarda} \end{cases}$$

Burada g_m , m düğümündeki tahmini değeri ifade eder ve m düğümüne ulaşan örnekler için ortalama (veya çok fazla gürültü varsa medyan) alınarak hesaplanır:

$$g_m = \frac{\sum_t b_m(x^t) r^t}{\sum_t b_m(x^t)}$$

Düğüm m için hata, aşağıdaki gibi ifade edilir:

$$E_m = \frac{1}{N_m} \sum_t (r^t - g_m)^2 b_m(x^t) \quad (9)$$

Burada;

$$N_m = |X_m| = \sum_t b_m(x^t)$$

olarak tanımlanır. Eğer bir düğümdeki hata kabul edilebilir seviyenin (θ_r) altında ise, bu düğüm bir yaprak düğüm olarak belirlenir ve g_m değeri saklanır. Eğer hata yüksekse, düğüm daha fazla alt kümeye bölünerek hata azaltılmaya çalışılır.

Bir düğüm m 'den yapılan dallanmalarda, her bir dal j için alt kümeler X_{mj} olarak tanımlanır ve tüm alt kümelerin birleşimi orijinal kümeyi oluşturur:

$$\sum_{j=1}^n X_{mj} = X_m$$

Buna karşılık, her alt küme için:

$$b_{mj}(x) = \begin{cases} 1, & x \in X_{mj} \\ 0, & \text{diğer durumlarda} \end{cases}$$

Her alt dal için tahmini değer g_{mj} şu şekilde hesaplanır:

$$g_{mj} = \frac{\sum_t b_{mj}(x^t)(r^t)}{\sum_t b_{mj}(x^t)}$$

Bölme işlemi sonrasında toplam hata aşağıdaki gibi ifade edilir:

$$E'_m = \frac{1}{N_m} \sum_j \sum_t (r^t - g_{mj})^2 b_{mj}(x^t) \quad (10)$$

Bölme işleminin etkinliğini değerlendirmek için başlangıçtaki hata (eşitlik 9) ile bölme sonrası hata (eşitlik 10) arasındaki fark kullanılır. Maksimum hata farkını bulmak veya eşdeğer olarak minimum toplam hatayı aramak amaçlanır. Ayrıca, belirli bir hata eşiğinin aşılmaması da garanti altına alınır:

$$E_m = \max_j \max_t |r^t - g_{mj}| b_{mj}(x^t)$$

Daha karmaşık modellere ulaşmak amacıyla yaprak düğümlerinde basit doğrusal modeller de kullanılabilir:

$$g_m(x) = w_m^T x + w_{m0}$$

Bu tür bir yaklaşım daha küçük ağaçlar oluşturulmasına olanak tanırken, yaprak düğümlerinde ek hesaplama maliyetleri doğurur (Alpaydın, 2014).

Karar ağacı teknikleri arasında Sınıflandırma ve Regresyon Ağaçları (CART), Çok Değişkenli ve Adaptif Regresyon Spline (MARS), Ağaç Toplulukları ve Rastgele Orman gibi yöntemler yer almaktadır.

5.3.2 Rastgele orman regresyonu

Rastgele Orman algoritması, istatistiksel olarak bağımsız ve aynı dağılıma sahip rastgele değişkenlere dayalı bir topluluk tahmin modeli olarak tanımlanır. Rastgele ormanlar, her bir ağacın, ormandaki tüm ağaçlarla aynı dağılıma sahip ve bağımsız olarak seçilmiş rastgele vektörlere dayanan bir dizi ağaç tahmincisinden oluşur.

Rastgele Orman yönteminin bu kadar popüler olmasının nedeni farklı tahmin problemlerine kolayca uygulanabilmeleri ve yalnızca birkaç önemli parametreyi ayarlayarak hızla optimize edilebilmeleridir. Kullanımı kolay olmasının yanı sıra, genellikle yüksek doğruluk içeren; küçük örneklem boyutlarına, karmaşık ve yüksek boyutla sahip veri yapılarına karşı gösterdiği dayanıklılık ile tanınır (Scornet vd. 2015).

Bir Rastgele Orman Regresyonu aşağıdaki gibi modellenenir:

$\theta_1, \dots, \theta_m$ bağımsız ve aynı dağılıma sahip rastgele değişkenlerdir. Bu değişkenler örneklemden bağımsız olarak rastgele seçilir. Bir rastgele orman, $t_1(x, \theta_1, D_n), \dots, t_m(x, \theta_m, D_n)$ şeklinde m tane ağacı birleştirerek regresyon tahminini oluşturur.

$$r_n(x) = \frac{1}{m} \sum_{j=1}^m t_j(x, \theta_j, D_n)$$

θ rastgeleleştirme değişkeni, ağaç yapısı oluşturulurken hangi özelliklerin (koordinatların) seçileceğini ve bu özellikler üzerinde nasıl bölme yapılacağını belirleyen bir faktördür. Her bir bireysel rastgeleleştirilmiş ağaç $t_j(x, \theta_j, D_n)$ genellikle budama yapılmadan inşa edilir, yani ağaç oluşturma süreci, her düğümde en fazla k veri noktası kalana kadar devam eder. Burada k, önceden belirlenmiş pozitif bir tam sayıdır (Biau ve Devroye 2008).

5.3.3 Destek vektör regresyonu

Destek vektör regresyonu, sınıflandırma ve tahmin (regresyon) problemlerini çözmek amacıyla geliştirilen ve örnek verilerden öğrenmeye dayalı bir yaklaşımdır. Bu yöntem, öğrenme sürecini istatistiksel kurallara dayandırır ve modelin yalnızca mevcut verilere değil, gelecekte karşılaşılabilecek durumlara da uyum sağlamasını hedefler (Vapnik, 1992).

Eğitim veri kümesinin oluşturduğu problem durumu dikkate alındığında, bu veri kümesi üzerinden yapılacak tahmin süreci belirli kurallara göre modellenenir.

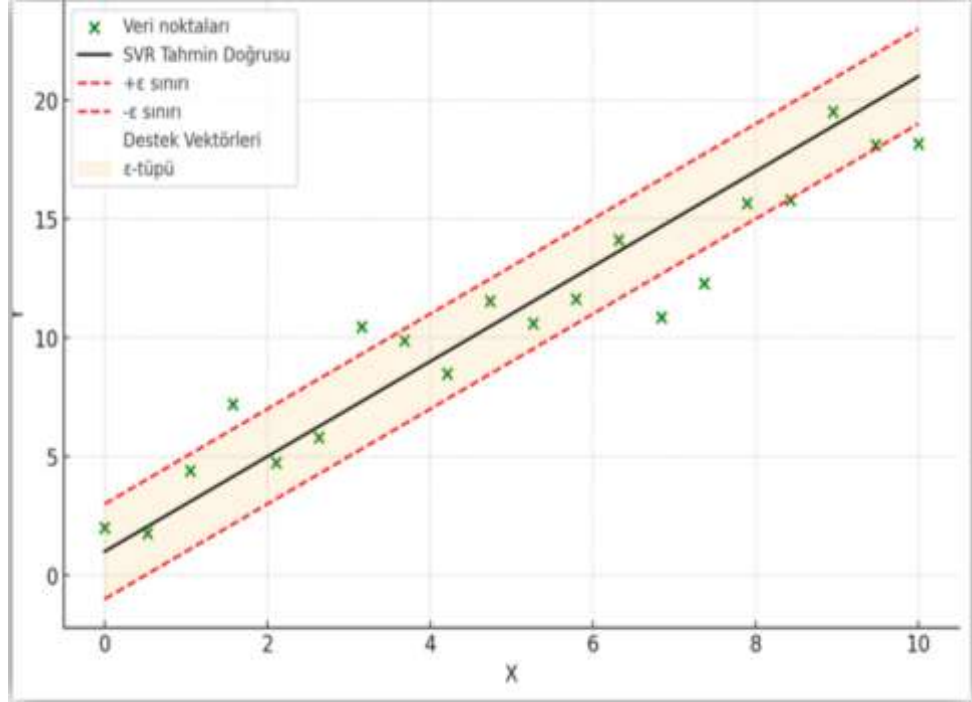
$$D = \{(x^1, y^1), (x^2, y^2), \dots, (x^m, y^m)\}$$

Burada $x^i \in R_m$ giriş değişkenleri vektörü $y^i \in R_m$ ise hedef değişkeni temsil etmektedir.

Destek vektör, doğrusal bir tahmin fonksiyonu oluşturarak veri kümesindeki örnekleri modellemeye çalışır:

$$y = \langle w, x \rangle + b$$

Burada ağırlık vektörü w ile gösterilirken, b sabit terimi ifade etmektedir.



Şekil 5.1 SVR modelinde tahmin doğrusu, $\pm\epsilon$ sınırları ve destek vektörlerinin gösterimi

Şekil 5.1’de, Destek Vektör Regresyonu (SVR) modeline ilişkin temel bileşenler görsel olarak sunulmaktadır. Grafikte yeşil çarpılar, modelin eğitildiği veri noktalarını temsil ederken; bu noktalar üzerinden geçen siyah çizgi, SVR modelinin oluşturduğu tahmin doğrusunu göstermektedir. Tahmin doğrusunun etrafındaki kırmızı kesikli çizgiler, model tarafından belirlenen $\pm\epsilon$ sınırlarını ifade etmektedir. Bu sınırlar arasında kalan turuncu bant, ϵ -tüpü olarak adlandırılmakta ve modelin belirli bir hata toleransı dâhilinde kabul ettiği bölgeyi göstermektedir. Bu sınırları aşan ve modelin öğrenme sürecine doğrudan katkıda bulunan noktalar ise destek vektörleri olarak adlandırılmakta ve SVR modelinin karar sınırlarını belirlemede kritik rol oynamaktadır.

SVR modelinin temel bileşenleri olan tahmin doğrusu, $\pm\epsilon$ sınırları ve destek vektörleri yapısı, yalnızca ϵ sınırını aşan örneklerin (destek vektörlerinin) modele katkı sağlaması prensibine dayanır. Bu sezgisel yaklaşım, aşağıda yer alan amaç fonksiyonu ve kısıtlar ile matematiksel olarak ifade edilen optimizasyon problemine temel oluşturmaktadır.

Amaç fonksiyonu (Minimizasyon):

$$\frac{1}{2} \|w\|^2 + c \sum (\xi_i + \xi_i^*)$$

Kısıtlar:

$$y_i - (wx_i) - b \leq \varepsilon + \xi_i$$

$$wx_i + b - y_i \leq \varepsilon + \xi_i^*$$

$$\xi_i, \xi_i^* \geq 0$$

ε tolerans parametresi, ξ_i, ξ_i^* gevşeme (sapma) değişkenleri ve $C > 0$ ise ceza parametresi olup, sapmaları dengeleyen bir katsayıdır (Gunn, 1998).

Bu kısıtlı optimizasyon problemi, her bir kısıt için tanımlanan Lagrange çarpanları a_i, a_i^* kullanılarak ikili (dual) problem biçimine dönüştürülür. Lagrange çarpanları, kuadratik programlama yöntemiyle çözülerek belirlenir.

Optimum çözüm elde edildikten sonra, tahmin fonksiyonu aşağıdaki şekilde ifade edilir (Shokry vd., 2015).

$$y = \sum_i^m (a_i^* - a_i)(x_i - x) + b$$

5.4 Model Performans Ölçüm Kriterleri

Model performans ölçüm kriterleri (performans metrikleri veya hata metrikleri), hem regresyon analizi hem de makine öğrenmesi tabanlı tahmin modellerinin en kritik bileşenlerinden birini oluşturmaktadır. Bu metrikler, geliştirilen modelin öngördüğü değerler ile gerçek gözlem değerleri arasındaki farkı nicel olarak ölçmeyi amaçlayan,

mantıksal ve matematiksel temellere dayalı yapılar olarak tanımlanabilir. Başka bir ifadeyle, bir performans metriği, tahmin edilen sonucun gerçeğe ne kadar yakın olduğunu değerlendiren bir ölçüttür.

Literatürde, model türüne ve analiz amacına göre çok sayıda performans metriği tanımlanmış ve önerilmiştir (Botchkarev, 2018). Bu metrikler genel olarak Sınıflandırma, Regresyon, Kümeleme ve Çok Sınıflı Sınıflandırma olmak üzere dört ana kategori altında toplanmaktadır. Her bir kategori, kendine özgü hata ölçütleri ve başarı göstergelerine sahiptir.

Regresyon analizlerinde yaygın olarak kullanılan performans metrikleri arasında Hata Kareler Ortalaması (Mean Squared Error – MSE), Hata Mutlak Ortalama (Mean Absolute Error – MAE), Karekök Hata Kareler Ortalaması (Root Mean Squared Error – RMSE) ve Belirlilik Katsayısı (R^2) yer almaktadır. Bu metriklerden MSE, MAE ve RMSE değerlerinin düşük olması, modelin tahminlerinin gerçek değerlere yakınlığını gösterirken; R^2 katsayısının yüksek olması ise modelin bağımlı değişken varyansını açıklama gücünün yüksek olduğuna işaret eder.

Bu tez çalışmasında, kullanılan regresyon ve makine öğrenmesi modellerinin başarısı, söz konusu performans metrikleri üzerinden değerlendirilecektir. Analiz sürecinde, modellerin performans metriklerinden elde edilen sonuçlar karşılaştırmalı olarak değerlendirilecek ve bu kriterler doğrultusunda en iyi performansı gösteren model tercih edilecektir. Bu kapsamda ilgili metriklerin isimleri, matematiksel formülleri ve kısa açıklamaları Çizelge 5.1’de özetlenmektedir.

Çizelge 5.1 Çeşitli model performans ölçüm kriterleri

Model Performans Ölçüm Kriterleri	Formülleri	Açıklama
Hata Kareler Ortalaması (MSE)	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$	Gerçek değerler ile tahmin edilen değerler arasındaki farkların karelerinin ortalaması olarak hesaplanır. Hataları karesel biçimde değerlendirdiğinden, büyük sapmalara daha fazla ağırlık verir ve bu nedenle uç değerlere karşı duyarlıdır.
Ortalama Mutlak Hata (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y}_i $	Gerçek değerler ile tahmin edilen değerler arasındaki farkların mutlak değerlerinin ortalaması olarak tanımlanır. MAE, tahminlerin ne kadar hata yaptığını, yönünden (pozitif /negatif) bağımsız olarak, ortalama hata büyüklüğü şeklinde ifade eder. Hedef değişkenin birimi ile aynıdır, bu nedenle yorumlaması kolaydır.
Karekök Hata Kareler Ortalaması (RMSE)	$RMS = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$	MSE'nin karekökü olup hatayı hedef değişkenin birimi cinsinden ifade eder. Büyük hatalara duyarlıdır ve yorumlanması kolaydır.
Belirlilik Katsayısı (R^2)	$R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	Modelin bağımlı değişkendeki toplam varyansı ne ölçüde açıkladığını gösterir. 1'e yakın değerler yüksek açıklama gücünü, 0'a yakın değerler ise düşük açıklama gücünü ifade eder.

Burada n : Toplam gözlem sayısı

y_i : Gerçek değer

$\hat{y}_i$: Tahmin edilen değeri göstermektedir.

6. UYGULAMA

Uygulama aşamasında, enerji yoğunluğunu etkileyen faktörlerin farklı gelir grupları bağlamında incelenmesi amaçlanmıştır. Bu doğrultuda, Dünya Bankası gelir sınıflandırmasına göre üç farklı gelir grubundan toplam sekiz ülke çalışmaya dâhil edilmiştir. Bu ülkeler: alt-orta gelir grubundan Mısır ve Fas; orta-yüksek gelir grubundan Arjantin, Meksika ve Tayland ile yüksek gelir grubundan Danimarka, Hollanda ve İngiltere'dir. Bu gruplandırma, enerji yoğunluğunu etkileyen etmenlerin yalnızca genel düzeyde değil, ülkelerin ekonomik gelişmişlik seviyesine göre nasıl farklılaştığının da ayrıntılı olarak analiz edilmesini sağlamıştır.

Çalışmada bağımlı değişken olarak enerji yoğunluğu ele alınmış; doğalgaz tüketimi, elektrik tüketimi, kömür tüketimi, petrol tüketimi, yenilenebilir enerji kaynaklarının kullanımı ve nüfus yoğunluğu ise bağımsız değişkenler olarak çalışmaya dâhil edilmiştir. Analizlerde 1990–2021 dönemine ait ülke bazlı yıllık veriler kullanılmıştır. Analizin ilk aşamasında, bağımlı değişkenin 0 ile 1 aralığında sürekli değerler alması nedeniyle, oransal veri yapısına uygunluk gösteren Beta Regresyon modeli uygulanmıştır. Bu model, enerji yoğunluğu ile açıklayıcı değişkenler arasındaki ilişkileri parametrik bir çerçevede detaylı biçimde ortaya koymuştur. Daha sonraki aşamada, bağımsız değişkenler arasında gözlenen çoklu bağlantı (multicollinearity) sorunu dikkate alınmış ve bu problemi azaltmak amacıyla Beta ridge Regresyon yöntemi uygulanmıştır. Beta ridge yaklaşımı, katsayı tahminlerindeki oynaklığı azaltarak modelin istikrarını güçlendirmiş ve sonuçların güvenilirliğini artırmıştır. Bu kapsamda elde edilen katsayı tahminleri değerlendirilmiş, enerji yoğunluğunu belirleyen değişkenlerin görece etkileri analiz edilmiştir.

Son aşamada ise, model performansının artırılması ve doğrusal olmayan karmaşık ilişkilerin yakalanabilmesi amacıyla makine öğrenmesi tabanlı algoritmalara başvurulmuştur. Karar Ağacı, Rastgele Orman ve Destek Vektör Regresyonu yöntemleri train değeri %80 test değeri ise %20 alınarak ayrı ayrı uygulanmış; elde edilen sonuçlar karşılaştırmalı olarak değerlendirilmiştir. Böylece, klasik istatistiksel yaklaşımların yanı

sıra modern yöntemlerin de enerji yoğunluğu üzerindeki belirleyicileri nasıl ortaya koyduğu incelenmiştir.

Modelde kullanılacak olan bağımlı ve bağımsız değişkenlerin tanımları ile veri kaynakları Çizelge 6.1’de sunulmaktadır. Ayrıca, bağımsız değişkenlere z-skoru standartlaştırması uygulanarak değişkenler arasındaki ölçek farklılıklarının giderilmesi sağlanmıştır. Bu adım, modellerin daha sağlıklı ve karşılaştırılabilir sonuçlar üretmesine katkı sağlamıştır.

Tüm analizler R ve Python programlama dillerinin sunduğu ileri düzey istatistiksel ve makine öğrenmesi kütüphaneleri kullanılarak gerçekleştirilmiştir. Bu yazılımlar; veri ön işleme, modelleme, hata analizi ve model karşılaştırmalarının sistematik bir şekilde yürütülmesine olanak sağlamıştır.

R ortamında Beta Regresyon için betareg (Cribari-Neto ve Zeileis 2010), Varyans Şişme Faktörü için car (Fox ve Weisberg 2019), model performans ölçüm kriterleri için Metrics (Hamner ve Frasco 2018) kütüphaneleri kullanılmıştır.

Python ortamında Karar Ağacı, Rasgele Orman Regresyonu ve Destek Vektör Regresyonu analizleri scikit-learn kütüphanesi (Pedregosa vd. 2011) kullanılarak uygulanmıştır.

Çizelge 6.1 Değişkenlerin tanımları ve kaynakları

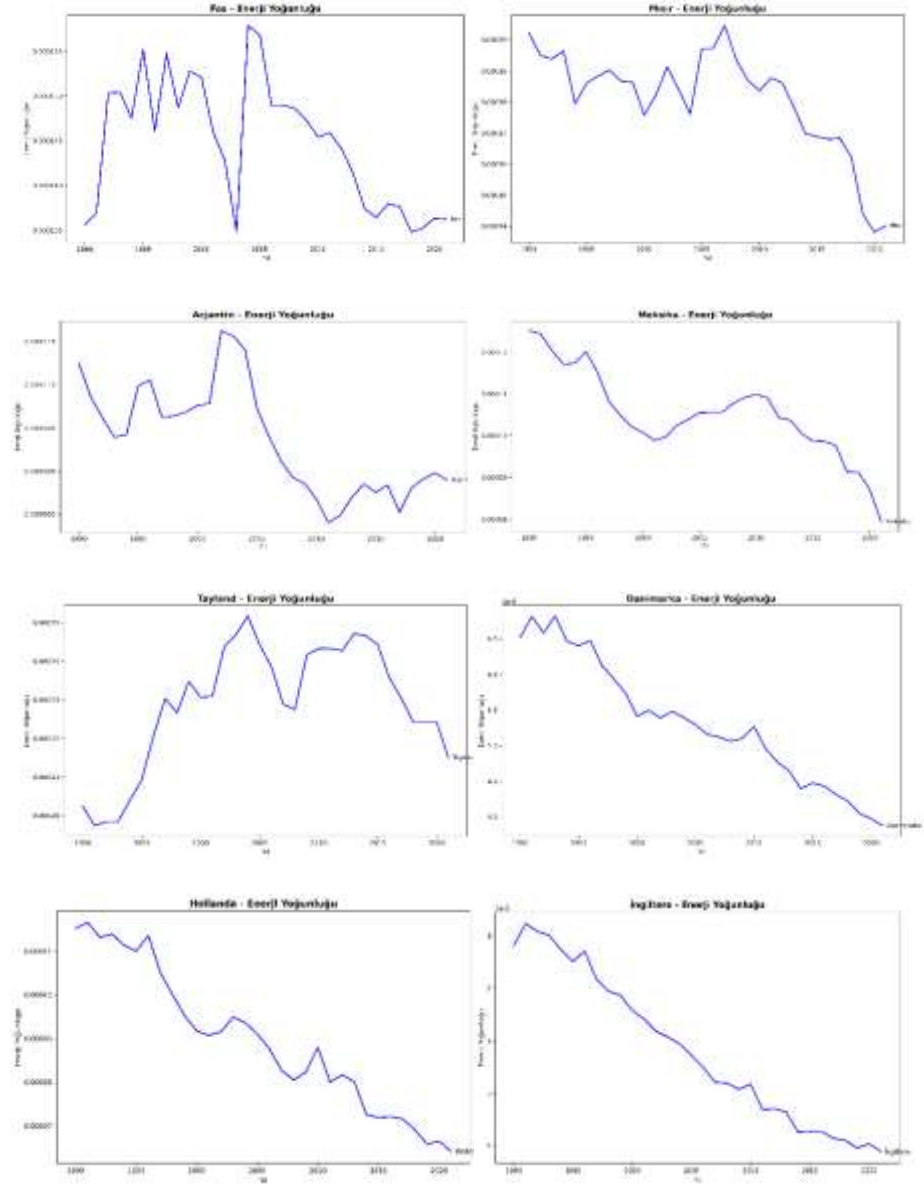
Değişken	Tanımı	Kaynak
Enerji Yoğunluğu	$\frac{\text{Toplam Enerji Kullanımı}(\text{tep})}{\text{GSYH (2015 temel yılı \$)}}$	Toplam Enerji Kullanımı: Uluslararası Enerji Ajansı (IEA) GSYH: Dünya Bankası (WB)
Doğalgaz Tüketimi	Doğalgaz tüketiminin z-skoru ile standartlaştırılmış değeri	IEA
Elektrik Tüketimi	Elektrik tüketiminin z-skoru ile standartlaştırılmış değeri	IEA

Çizelge 6.1 Değişkenlerin tanımları ve kaynakları (devam)

Değişken	Tanımı	Kaynak
Kömür Tüketimi	Kömür tüketiminin z-skoru ile standartlaştırılmış değeri	IEA
Petrol Tüketimi	Petrol tüketiminin z-skoru ile standartlaştırılmış değeri	IEA
Yenilenebilir Enerji Kaynakları	Yenilenebilir enerji kaynaklarının z-skoru ile standartlaştırılmış değeri	IEA
Nüfus Yoğunluğu	Nüfus yoğunluğunun z-skoru ile standartlaştırılmış değeri	WB

6.1 Enerji Yoğunluğu Grafikleri

Bu çalışmada bağımlı değişken olarak kullanılan enerji yoğunluğu, bir ülkenin veya bölgenin bir birim gayrisafi yurt içi hasıla (GSYH) üretebilmek için tükettiği enerji miktarını gösteren bir enerji verimliliği göstergesidir (Anonim, 2024). Enerji yoğunluğu ile enerji verimliliği kavramları sıklıkla birbirine karıştırılsa da bu iki kavram ters orantılı bir ilişkiye sahiptir. Başka bir ifadeyle enerji yoğunluğundaki azalma, enerji verimliliğinde artış anlamına gelmektedir. Şekil 6.1’de çalışmaya dâhil edilen ve farklı gelir gruplarında yer alan toplam sekiz ülkenin (iki alt-orta gelirli, üç orta-yüksek gelirli ve üç yüksek gelirli) yıllar itibarıyla enerji yoğunluklarındaki değişim görsel olarak sunulmaktadır.



Şekil 6.1 Seçilmiş ülkelerin enerji yoğunluğu grafikleri

Bu çalışmada analiz edilen sekiz ülkenin enerji yoğunluğu eğilimleri, 1990 sonrası dönemde genel olarak azalma yönünde bir seyir izlemektedir. Bu durum hem gelişmiş hem de gelişmekte olan ülkelerde enerji verimliliğine yönelik politika çabalarının etkili olmaya başladığını göstermektedir.

Danimarka, Hollanda ve İngiltere gibi yüksek gelirli ve gelişmiş ülkelerde enerji yoğunluğunun istikrarlı bir biçimde gerilediği görülmektedir. Bu ülkelerde izlenen düşüş

eğilimi; yenilenebilir enerji kaynaklarının yaygınlaştırılması, enerji verimliliği odaklı yatırımların artması ve düşük enerji yoğunluklu sektörlerin ekonomideki payının büyümesi ile ilişkilendirilebilir.

Benzer şekilde, Fas ve Arjantin gibi gelişmekte olan ülkelerde uzun vadede enerji yoğunluğunda düşüş gözlenmekle birlikte, bu gerileme daha dalgalı bir yapı sergilemektedir. Bu durum, söz konusu ülkelerde politika reformlarının belirli dönemlerde etkili olduğunu ancak uzun vadeli istikrarın sağlanmasında zorluklar yaşandığını düşündürmektedir.

Öte yandan, Mısır, Tayland ve Meksika gibi ülkelerde enerji yoğunluğu zaman içerisinde dalgalı bir seyir izlemiştir. Bu dalgalanmalar; ekonomik büyümedeki değişkenlik, enerji yoğun sektörlerdeki üretim yapısındaki oynamalar, sübvansiyon politikalarındaki değişiklikler veya dışsal şoklar ile açıklanabilir. Bu ülkelerde enerji politikalarının tutarlı şekilde uygulanmaması ya da sektörler arası dönüşümün yavaş ilerlemesi, enerji yoğunluğundaki oynaklığın temel nedenleri arasında değerlendirilebilir.

Sonuç olarak, enerji yoğunluğundaki farklılaşmalar; ülkelerin ekonomik yapıları, enerji politikalarının kapsamı, sektör kompozisyonları ve dışsal koşullara verdikleri tepkiler doğrultusunda şekillenmektedir. Enerji yoğunluğu, ülkelerin enerji verimliliği politikalarının başarısını ve ekonomik dönüşüm süreçlerini izlemek açısından stratejik öneme sahip bir göstergedir.

6.2 Beta Regresyon Yöntemi ile Ülkelerin Enerji Yoğunluğunu Etkileyen Faktörlerin Belirlenmesi

Bu bölümde, ülkelerin enerji yoğunluğunu etkileyen faktörlerin belirlenmesi amacıyla Beta Regresyon analizleri gerçekleştirilmiştir. Analiz kapsamında ülkeler, Dünya Bankası gelir sınıflandırmasına göre üç farklı gelir grubuna ayrılmış ve her grup için ayrı modelleme yapılmıştır. Enerji yoğunluğu değişkeninin 0 ile 1 aralığında sürekli değerler

alması ve oransal veri yapısına sahip olması nedeniyle, Beta Regresyon yöntemi tercih edilmiştir (Ferrari ve Cribari-Neto 2004).

Modelleme sürecinde, farklı bağlantı fonksiyonları denenmiş ve her model için Akaike Bilgi Kriteri (AIC) ile Bayes Bilgi Kriteri (BIC) değerleri karşılaştırılarak en uygun model belirlenmiştir. Modelin tahmin performansı ise Hata Kareler Ortalaması (MSE), Karekök Hata Kareler Ortalaması (RMSE), Ortalama Mutlak Hata (MAE) ve Belirlilik Katsayısı (R^2) ölçütleri kullanılarak değerlendirilmiştir.

6.2.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi

Bu bölümde, alt-orta gelir grubunda yer alan Mısır ve Fas için enerji yoğunluğunu etkileyen faktörler Beta Regresyon yöntemi ile analiz edilmiştir. İnceleme, söz konusu faktörlerin enerji yoğunluğu üzerindeki etkilerinin yönünü (pozitif/negatif) ve büyüklüğünü, aynı zamanda istatistiksel anlamlılık düzeylerini değerlendirmeyi amaçlamaktadır.

Çizelge 6.2 Alt-orta gelirli ülkelerin bağ fonksiyonları için bilgi kriteri değerleri

Bağ Fonksiyonu	df	AIC	BIC
Logit	8	-1371.808	-1354.537
Log-log	8	-1371.998	-1354.727
Log	8	-1371.807	-1354.536
C-Loglog	8	-1371.808	-1354.537
Probit	8	-1371.916	-1354.645
Cauchit	8	-1369.494	-1352.223

Çizelge 6.2’de verilen AIC ve BIC değerleri incelendiğinde, altı farklı bağlantı fonksiyonu arasında en iyi modelin, her iki ölçütte de en düşük değerlere sahip olan model olduğu görülmektedir. Karşılaştırma sonuçlarına göre, log-log bağlantı fonksiyonu AIC (-1371.998) ve BIC (-1354.727) değerleri bakımından diğer tüm bağlantı fonksiyonlarından daha iyi performans göstermiştir. Bu bulgu, log-log bağlantı fonksiyonunun veri ile en yüksek uyumu sağladığını ve model karmaşıklığını minimum düzeyde tuttuğunu göstermektedir. Bu nedenle, analizlerin devamında log-log bağlantı fonksiyonu kullanılmış; bu modele ilişkin artıkların betimsel istatistikleri ve regresyon katsayıları sırasıyla Çizelge 6.3 ve Çizelge 6.4’te sunulmuştur.

Çizelge 6.3 Alt-orta gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği

En Küçük Değer	1. Çeyreklik	Medyan	3. Çeyreklik	En Büyük Değer
-3.4365	-0.4745	0.0137	0.6275	1.9141

Alt-orta gelirli ülkeler için standartlaştırılmış ağırlıklı artıkların medyanının sıfıra oldukça yakın olması, modelde sistematik bir sapma bulunmadığını göstermektedir. Artık değerlerin büyük bölümünün dar bir aralıkta yoğunlaşması, tahminlerin dengeli olduğunu ortaya koymaktadır. Bununla birlikte, negatif uça gözlenen yüksek değerler bazı ülkelerde model öngörülerinden sapmalar yaşandığını göstermektedir. Nitekim Chang ve arkadaşları (2024) da enerji yoğunluğu dinamiklerinin genel olarak benzer eğilimler sergilemesine rağmen ülkeye özgü farklılıklar içerdiğini ortaya koymaktadır.

Alt-orta gelir grubunda yer alan ülkelerde enerji yoğunluğunu etkileyen değişkenlerin katsayı tahminleri ve anlamlılık düzeyleri, log-log bağlantı fonksiyonu kullanılarak elde edilmiş ve sonuçlar Çizelge 6.4’te sunulmuştur.

Çizelge 6.4 Alt-orta ülkelerin log-log bağlantılı ortalama model için katsayıları

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-2.1668290	0.0004203	-5155.804	< 2e-16
Doğalgaz Tüketimi	0.0095306	0.0018269	5.217	1.82e-07
Elektrik Tüketimi	-0.0136287	0.0032321	-4.217	2.48e-05
Kömür Tüketimi	-0.0002333	0.0005951	-0.392	0.695
Petrol Tüketimi	0.0284015	0.0017671	16.073	< 2e-16
Yenilenebilir Enerji Kaynakları	0.0023672	0.0005018	4.717	2.39e-06
Nüfus Yoğunluğu	-0.0190022	0.0012628	-15.048	< 2e-16

Çizelge 6.4’te yer alan katsayılar, petrol ve doğalgaz tüketiminin enerji yoğunluğunu artırdığını; elektrik tüketiminin ise azaltıcı ve istatistiksel olarak anlamlı bir etkiye sahip olduğunu göstermektedir. Yenilenebilir enerji kaynaklarının pozitif işaretli bulunması, alt-orta gelir grubunda bu yatırımların kısa vadede verimlilik kazanımlarına tam olarak dönüşemediğini düşündürmektedir. Bu sonuç, geçiş maliyetleri, altyapı eksiklikleri ve teknolojik uyum süreçleri ile açıklanabilir. Ayrıca nüfus yoğunluğunun negatif ve güçlü biçimde anlamlı etkisi, yerleşim yoğunluğunun özellikle ulaştırma ve ısınma gibi alanlarda enerji kullanımında verimlilik sağladığına işaret etmektedir.

Elde edilen bulgular, literatürdeki benzer çalışmalarla da büyük ölçüde tutarlılık göstermektedir. Elektrifikasyonun enerji verimliliğini desteklediğine ilişkin sonuç, enerji-büyüme-teknoloji etkileşimini inceleyen Csereklyei (2016) ile örtüşmektedir. Sadorsky (2013), gelişmekte olan ülkelerde gelir artışının yenilenebilir enerji tüketimini uzun vadede artırdığını, ancak kısa vadede bu etkinin sınırlı kaldığını ortaya koymaktadır. Bu durum, alt-orta gelir grubunda yenilenebilir enerji talebindeki hızlı artışın kısa vadede enerji yoğunluğunu düşürmek yerine artırıcı yönde etkiler yaratabileceğini göstermektedir. Ayrıca, nüfus ve kentleşme dinamiklerinin enerji yoğunluğu üzerindeki

etkilerinin ülke gelir düzeyine göre farklılık gösterebileceğine dair bulgular, Poumanyong ve Kaneko (2010) çalışmasıyla paralellik arz etmektedir.

Çizelge 6.5 Alt-orta gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri

MSE	RMSE	MAE	Psuedo R^2
0.000000000022	0.0000047	0.0000034	0.95

Çizelge 6.5’te sunulan performans ölçüm kriterleri, alt-orta gelir grubuna ait Beta regresyon modelinin yüksek bir açıklama gücüne sahip olduğunu göstermektedir. Özellikle R^2 değerinin 0.95 seviyesinde olması, enerji yoğunluğunun bağımsız değişkenler yardımıyla büyük ölçüde açıklandığını ortaya koymaktadır. Bununla birlikte, bağımlı değişken olan enerji yoğunluğunun 0–1 aralığında tanımlı olması nedeniyle hata ölçütlerinin (MSE, RMSE ve MAE) düşük çıkması beklenen bir durumdur. Bu nedenle, model değerlendirmesinde esas olarak R^2 değeri ile katsayıların istatistiksel anlamlılığı dikkate alınmıştır.

Modelin performans ölçütleri güçlü olmakla birlikte, katsayıların güvenilir bir şekilde yorumlanabilmesi için bağımsız değişkenler arasında çoklu doğrusal bağlantı riskinin de değerlendirilmesi gerekmektedir. Çünkü yüksek R^2 ve düşük hata değerleri, değişkenler arasındaki yüksek korelasyon varlığında yanıltıcı olabilir ve katsayı tahminlerinde sapmalara yol açabilir. Bu nedenle, modelin sağlamlığını artırmak amacıyla Varyans Şişirme Faktörü (Variance Inflation Factor, VIF) değerlerinin incelenmesi önem taşımaktadır. Literatürde genellikle VIF değerinin 10’un üzerinde olması ciddi düzeyde çoklu bağlantıya işaret etmektedir (Gujarati & Porter, 2009). Alt-orta gelirli ülkeler için VIF değerleri hesaplanmış ve sonuçlar Çizelge 6.6’de sunulmuştur.

Çizelge 6.6 Alt-orta gelirli ülkelerin VIF değerleri

Doğalgaz	Elektrik	Kömür	Petrol	Yenilenebilir	Nüfus
Tüketimi	Tüketimi	Tüketimi	Tüketimi	Enerji Kaynakları	Yoğunluğu
20.757439	60.218834	2.059658	17.847317	1.349406	8.523798

Çizelge 6.6’da verilen VIF sonuçları incelendiğinde, özellikle elektrik tüketimi (60.21) ve doğalgaz tüketimi (20.76) için oldukça yüksek değerler elde edilmiştir. Bunun yanında petrol tüketimi değişkeni de 17.85 değeri ile kritik eşik değerinin üzerinde yer almaktadır. Bu bulgular, söz konusu değişkenler arasında güçlü doğrusal ilişkiler bulunduğunu ve modelde belirgin çoklu bağlantı sorunu yaşandığını göstermektedir. Buna karşılık, kömür tüketimi (2.06), yenilenebilir enerji kullanımı (1.35) ve nüfus yoğunluğu (8.52) nispeten düşük değerlere sahiptir ve bu değişkenler açısından ciddi bir çoklu doğrusal bağlantı problemi bulunmamaktadır.

VIF değerlerinde gözlenen yükseklikler, özellikle elektrik, doğalgaz ve petrol tüketimi değişkenlerinin katsayı tahminlerinde güvenilirliğin azalmasına yol açabilmektedir. Bu nedenle, çoklu bağlantının etkilerini azaltmak ve daha kararlı katsayı tahminleri elde etmek amacıyla Beta ridge regresyon analizi uygulanmıştır. Beta ridge regresyon, katsayılara ceza terimi ekleyerek parametre tahminlerindeki varyansı düşürmekte ve modelin istikrarını artırmaktadır.

Çizelge 6.7 Alt-orta ülkelerin log-log bağlantılı beta ridge ortalama model için katsayıları

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-2.16683051	0.00042031	-5155.34996	0
Doğalgaz Tüketimi	0.00953254	0.00183478	5.19546	2.0422e-07
Elektrik Tüketimi	-0.01362755	0.00324332	-4.20173	2.6489e-05
Kömür Tüketimi	-0.00024145	0.00059410	-0.40642	6.8444e-01
Petrol Tüketimi	0.02840206	0.00176242	16.11542	1.9881e-58
Yenilenebilir Enerji Kaynakları	0.00236399	0.00050133	4.71541	2.4123e-06
Nüfus Yoğunluğu	-0.01899871	0.00126184	-15.05637	3.1355e-51

Çizelge 6.7’de sunulan Beta ridge regresyon sonuçları, alt-orta gelir grubuna ilişkin enerji yoğunluğu modeline dair önemli bulgular ortaya koymaktadır. Katsayılar incelendiğinde, doğalgaz (0.0095) ve petrol tüketiminin (0.0284) pozitif ve istatistiksel olarak anlamlı etkiler gösterdiği görülmektedir. Buna karşılık, elektrik tüketimi (-0.0136) ve nüfus yoğunluğu (-0.0190) negatif ve anlamlı katsayılarla sahiptir. Bu sonuçlar, doğalgaz ve petrol tüketimindeki artışın enerji yoğunluğunu yükselttiğini, elektrik tüketimi ile nüfus yoğunluğunun ise enerji verimliliğini destekleyici yönde etkiler yarattığını göstermektedir. Kömür tüketimi ise modelde istatistiksel olarak anlamlı bulunmamıştır.

Karşılaştırmalı olarak değerlendirildiğinde, Beta ridge regresyon sonuçları beta regresyon ile büyük ölçüde örtüşmektedir. Katsayıların işaretleri aynı kalmış, anlamlılık düzeylerinde herhangi bir değişim olmamış ve yalnızca küçük sayısal farklılıklar ortaya çıkmıştır. Bu durum, Beta ridge yönteminin yüksek VIF değerlerinden kaynaklanan çoklu doğrusal bağlantı sorununa karşı katsayı tahminlerini istikrarlı hale getirdiğini ve beta regresyon bulgularını güvenilir biçimde desteklediğini göstermektedir.

6.2.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi

Bu bölümde, Arjantin, Meksika ve Tayland gibi orta-yüksek gelir grubunda yer alan ülkeler için enerji yoğunluğunu belirleyen temel değişkenler analiz edilmektedir.

Çizelge 6.8 Orta-yüksek gelirli ülkelerin bağ fonksiyonları için bilgi kriteri değerleri

Bağ Fonksiyonu	df	AIC	BIC
Logit	8	-2005.185	-1984.670
Log-log	8	-2004.323	-1983.808
Log	8	-2005.186	-1984.671
C-Loglog	8	-2005.185	-1984.670
Probit	8	-2004.703	-1984.189
Cauchit	8	-2014.843	-1994.329

Çeşitli bağ fonksiyonları için istatistiksel model seçiminde kullanılan AIC ve BIC değerleri Çizelge 6.8’de sunulmuştur. Model seçim kriterleri açısından değerlendirildiğinde, en küçük AIC ve BIC değerlerini veren Cauchit bağ fonksiyonunun en uygun seçenek olduğu görülmektedir. Bu nedenle, analizlerde Cauchit bağ fonksiyonunun kullanılması tercih edilmiştir.

Bu kapsamda Cauchit bağ fonksiyonu için artıklarının betimsel istatistiği ve cauchit bağ fonksiyon modeli için katsayılar Çizelge 6.9 ve 6.10’da yer almaktadır.

Çizelge 6.9 Orta-yüksek gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği

En Küçük Değer	1. Çeyreklik	Medyan	3. Çeyreklik	En Büyük Değer
-2.6890	-0.7647	0.0890	0.7206	2.3041

Orta-yüksek gelirli ülkeler için standartlaştırılmış ağırlıklı artıkların betimsel istatistikleri, modelin genel olarak dengeli tahminler ürettiğini göstermektedir. Artıkların ortanca değeri sıfıra yakın çıkmış, bu da modelin sistematik bir yanlılık taşımadığını ortaya koymaktadır. Artıkların büyük bölümü dar bir aralıkta yoğunlaşmış olup, tahminlerin çoğu gözlem için başarılı olduğu söylenebilir. Bununla birlikte, uç değerler bazı ülkelerde modelin öngörülerden sapabildiğine işaret etmektedir. Bu durum, orta-yüksek gelir grubunda enerji yoğunluğunu belirleyen faktörlerin ülkeler arasında farklılık gösterebildiğini ortaya koymaktadır.

Orta-yüksek gelir grubunda yer alan ülkelerde enerji yoğunluğunu etkileyen değişkenlerin katsayı tahminleri ve anlamlılık düzeyleri, Cauchit bağlantı fonksiyonu kullanılarak elde edilmiş ve sonuçlar Çizelge 6.10’da sunulmuştur.

Çizelge 6.10 Orta-yüksek gelirli ülkelerin cauchit bağlantılı ortalama model için katsayılar

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-2507.67	13.68	-183.365	< 2e-16
Doğalgaz Tüketimi	256.23	36.49	7.023	2.18e-12
Elektrik Tüketimi	-482.06	36.96	-13.044	< 2e-16
Kömür Tüketimi	113.23	19.52	5.801	6.60e-09
Petrol Tüketimi	79.29	30.96	2.561	0.0104
Yenilenebilir Enerji Kaynakları	-72.56	28.46	-2.549	0.0108
Nüfus Yoğunluğu	987.66	55.89	17.673	< 2e-16

Çizelge 6.10'daki sonuçlar incelendiğinde, orta-yüksek gelirli ülkelerde enerji yoğunluğunu artıran en önemli değişkenlerin doğalgaz, kömür ve petrol tüketimi olduğu görülmektedir. Bu değişkenler pozitif ve anlamlı katsayılara sahiptir, yani bu ülkelerde fosil yakıt kullanımındaki artış enerji yoğunluğunu yükseltmektedir. Elektrik tüketimi ise negatif katsayıya sahip olup enerji yoğunluğunu azaltıcı yönde ve güçlü biçimde anlamlı bulunmuştur. Bu durum, elektrik kullanımının enerji verimliliği açısından önemli bir unsur olduğunu göstermektedir.

Yenilenebilir enerji kaynaklarının da negatif ve anlamlı bir etkiye sahip olması, bu enerji türünün enerji yoğunluğunu azaltıcı rol oynadığını ortaya koymaktadır. Nüfus yoğunluğu ise pozitif ve yüksek düzeyde anlamlı bulunmuş, bu da nüfus artışının ve kentleşmenin enerji talebini artırarak enerji yoğunluğunu yükselttiğine işaret etmektedir. Buna karşılık, kömür dışındaki fosil yakıtlar güçlü şekilde anlamlı bulunurken, kömür tüketiminin modeldeki etkisi daha sınırlı olmakla birlikte yine pozitif yönde gerçekleşmiştir.

Genel olarak sonuçlar, fosil yakıtların enerji yoğunluğunu artırıcı; elektrik ve yenilenebilir enerji kullanımının ise azaltıcı etkiye sahip olduğunu, nüfus yoğunluğunun ise enerji talebini artırarak enerji yoğunluğunu yükselttiğini göstermektedir.

Orta-yüksek gelirli ülkeler için elde edilen beta regresyon modelinin tahmin performansına ilişkin bulgular Çizelge 6.11’de yer almaktadır.

Çizelge 6.11 Orta-yüksek gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri

MSE	RMSE	MAE	Pseudo R^2
0.000000000042	0.0000065	0.000005	0.97

Çizelge 6.11’de sunulan performans ölçüm kriterleri, orta-yüksek gelirli ülkeler için kurulan Beta regresyon modelinin enerji yoğunluğunu açıklamada yüksek düzeyde bir başarıya sahip olduğunu ortaya koymaktadır. Modelin hata oranlarının düşük, R^2 değerinin ise yüksek olması, bağımlı değişken ile açıklayıcı değişkenler arasındaki ilişkinin güçlü ve istatistiksel olarak anlamlı olduğunu göstermektedir. Bununla birlikte, bağımlı değişkenin 0–1 aralığında tanımlanmış olması ve değişkenlerin z-skoru ile standartlaştırılması nedeniyle MSE, RMSE ve MAE gibi hata ölçütlerinin doğal olarak düşük seviyelerde çıkabileceği dikkate alınmalıdır. Bu nedenle modelin değerlendirilmesinde yalnızca hata ölçütleri değil, aynı zamanda R^2 değerinin büyüklüğü ve katsayıların istatistiksel anlamlılık düzeyleri birlikte göz önünde bulundurulmuştur.

Modelin performans ölçütleri güçlü olmakla birlikte, katsayıların güvenilir bir şekilde yorumlanabilmesi için bağımsız değişkenler arasında çoklu doğrusal bağlantı riskinin de değerlendirilmesi gerekmektedir. Çünkü yüksek R^2 ve düşük hata değerleri, değişkenler arasındaki yüksek korelasyon varlığında yanıltıcı olabilir ve katsayı tahminlerinde sapmalara yol açabilir. Bu nedenle, modelin sağlamlığını artırmak amacıyla VIF değerlerinin incelenmesi önem taşımaktadır. Bu amaçla Orta-yüksek gelirli ülkeler için VIF değerleri hesaplanmış ve sonuçlar Çizelge 6.12’de sunulmuştur.

Çizelge 6.12 Orta-yüksek gelirli ülkelerin VIF değerleri

Doğalgaz Tüketimi	Elektrik Tüketimi	Kömür Tüketimi	Petrol Tüketimi	Yenilenebilir Enerji Kaynakları	Nüfus Yoğunluğu
9.887	5.801	6.964	7.308	18.611	42.652

Çizelge 6.12’de verilen VIF değerleri incelendiğinde, özellikle nüfus yoğunluğu (42.652) değişkeni çok yüksek, yenilenebilir enerji kaynakları (18.61) değişkeninde ise orta düzeyde çoklu doğrusal bağlantı (multicollinearity) olduğu görülmektedir. Genel kabul gören eşik değere göre VIF’in 10’un üzerinde olması ciddi çoklu doğrusal bağlantı riskine işaret etmektedir. Buna karşın, doğalgaz (9.88), kömür (6.96), petrol (7.30) ve elektrik tüketimi (5.80) tüketimi değişkenleri eşik değerin sınırında olmakla birlikte görece kabul edilebilir düzeydedir.

Bağımsız değişkenlerde gözlenen bu yüksek korelasyon yapısı, katsayı tahminlerinin güvenilirliğini azaltabileceğinden, Beta ridge regresyon analizi uygulanarak katsayıların daha istikrarlı hale getirilmesi hedeflenmiştir. Beta ridge regresyonu ile elde edilen katsayı tahminleri ve ilgili istatistikler Çizelge 6.13’te sunulmaktadır.

Çizelge 6.13 Orta-yüksek gelirli ülkelerin cauchit bağlantılı beta ridge ortalama model için katsayılar

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-2491.622	15.92817	-156.4287	0
Doğalgaz Tüketimi	49.847	38.93748	1.2802	2.0048e-01
Elektrik Tüketimi	-334.772	37.23667	-8.9904	2.4637e-19
Kömür Tüketimi	148.774	21.21214	7.0136	2.3223e-12
Petrol Tüketimi	-22.746	32.48514	-0.7002	4.8380e-01
Yenilenebilir Enerji Kaynakları	54.253	31.90763	1.7003	8.9073e-02
Nüfus Yoğunluğu	661.028	60.51525	10.9233	8.9161e-28

Çizelge 6.13’te sunulan Beta ridge regresyon bulguları, çoklu doğrusal bağlantı probleminin azaltılmasıyla birlikte klasik beta regresyona kıyasla daha istikrarlı katsayı tahminleri ortaya koymuştur. Analiz sonuçlarına göre, elektrik tüketimi enerji yoğunluğunu negatif ve güçlü bir şekilde etkilemekte, bu bulgu elektrifikasyonun enerji verimliliğini desteklediğini gösteren literatürle uyumludur (Csereklyei vd., 2016). Petrol

tüketimi negatif fakat anlamsız çıkmış, bu nedenle genel eğilimle zıt yönde olsa da istatistiksel olarak desteklenmemiştir.

Doğalgaz tüketimi pozitif katsayıya sahip olmasına rağmen istatistiksel olarak anlamlı bulunmamıştır. Bu durum, literatürde genellikle vurgulanan fosil yakıtların artırıcı etkisiyle (Sadorsky, 2013; Apergis ve Payne, 2010) tam olarak örtüşmemekle birlikte, Beta ridge regresyonun çoklu bağlantı sorununu azaltarak doğalgazın yüksek gelirli ekonomilerde sanayi yapısı, enerji dönüşüm verimliliği veya alternatif enerji kaynaklarıyla ikame edilme düzeyi gibi faktörlerden dolayı farklı bir etki gösterebileceğini işaret etmektedir. Kömür tüketimi ise pozitif ve anlamlı etkiye sahip olup, enerji yoğunluğunu artırıcı yöndeki etkisine işaret eden çalışmaları doğrulamaktadır (Sadorsky, 2013; Bilgili vd., 2016).

Yenilenebilir enerji kaynakları ise pozitif katsayılara sahip olmakla birlikte yalnızca sınırlı düzeyde anlamlı bulunmuştur. Yenilenebilir enerjinin bu şekilde pozitif işaretli çıkması, kısa vadede geçiş maliyetleri, teknoloji adaptasyon süreçleri veya altyapı yetersizlikleri nedeniyle enerji verimliliği üzerindeki etkilerin tam olarak yansımaması ile açıklanabilir (Bilgili et al., 2016). Bununla birlikte, uzun vadede yenilenebilir enerji yatırımlarının enerji yoğunluğunu azaltıcı etkisi daha belirgin olabilecektir.

Nüfus yoğunluğu ise her iki modelde de (beta ve beta ridge) güçlü ve anlamlı etkisini korumuş, bu durum Poumanyong ve Kaneko (2010)'nun nüfus ve kentleşmenin enerji tüketimi üzerinde belirleyici bir faktör olduğu yönündeki bulgularıyla uyumlu bulunmuştur.

Sonuç olarak, Ridge regresyon modeli, yüksek VIF değerlerinin neden olduğu sapmaları azaltarak daha güvenilir katsayı tahminleri sağlamış ve özellikle elektrik tüketimi, kömür tüketimi ve nüfus yoğunluğunu enerji yoğunluğunu belirleyen temel faktörler olarak öne çıkarmıştır. Bu bulgular, çoklu doğrusal bağlantının göz ardı edilmemesi gerektiğini ve Ridge regresyonun enerji yoğunluğu analizlerinde istatistiksel olarak daha kararlı bir alternatif sunduğunu göstermektedir.

6.2.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin beta regresyon ile belirlenmesi

Yüksek gelirli ülkeler grubunda yer alan Danimarka, Hollanda ve İngiltere için enerji yoğunluğu ile bu yoğunluğu etkileyen çeşitli ekonomik ve çevresel faktörler arasındaki ilişki, Beta regresyon yöntemiyle incelenecektir. Bu analiz ile söz konusu ülkelerde enerji verimliliğini belirleyen temel değişkenlerin ortaya konulması amaçlanmaktadır.

Çeşitli bağ fonksiyonları için istatistiksel model seçiminde kullanılan AIC ve BIC değerleri Çizelge 6.14'te sunulmuştur.

Çizelge 6.14 Yüksek gelirli ülkelerin bağ fonksiyonları için bilgi kriterleri değeri

Bağ Fonksiyonu	df	AIC	BIC
Logit	8	-2057.774	-2037.259
Log-log	8	-2054.032	-2033.517
Log	8	-2057.776	-2037.261
C-Loglog	8	-2057.775	-2037.260
Probit	8	-2055.712	-2035.197
Cauchit	8	-2083.529	-2063.014

Model seçim kriterleri incelendiğinde, en düşük AIC ve BIC değerlerini veren bağ fonksiyonunun cauchit bağ fonksiyonu olduğu görülmektedir. Bu nedenle, yukarıdaki çizelgeye dayanarak modellemede Cauchit olabilirlik fonksiyonunun tercih edilmesi gerektiği sonucuna ulaşılmıştır.

Yüksek gelirli ülkeler için kurulan Beta regresyon modelinin tahminlerinin kalitesini ve modelin genel uyumunu değerlendirmek amacıyla, standartlaştırılmış ağırlıklı artıkların betimsel istatistikleri incelenmiştir (Çizelge 6.15). Bu analiz, modelin sistematik bir

yanlılık taşıyıp taşımadığını ve tahminlerin gözlemlere ne kadar yakın olduğunu ortaya koymaktadır.

Çizelge 6.15 Yüksek gelirli ülkelerin standartlaştırılmış ağırlıklı artıkların betimsel istatistiği

En Küçük Değer	1. Çeyreklik	Medyan	3. Çeyreklik	En Büyük Değer
-3.2231	-0.4328	-0.0703	0.4775	2.5893

Çizelge 6.15'teki bulgulara göre, artıkların medyan değeri -0.0703 olup, bu değer in sıfıra oldukça yakın olması, modelin tahminlerinde genel olarak sistematik bir sapma olmadığını göstermektedir. Ayrıca, artıkların büyük bir bölümünün, -0.4328 ile 0.4775 gibi dar bir aralıkta yoğunlaşması, modelin gözlemlerin büyük bir kısmı için oldukça isabetli ve dengeli tahminler ürettiğini kanıtlamaktadır. Bu durum, kullanılan bağımsız değişkenlerin, yüksek gelirli ülkelerin çoğunda enerji yoğunluğunu açıklama gücünün yüksek olduğunu göstermektedir. Bununla birlikte, artık dağılımında en küçük değer in -3.2231 ve en büyük değerin 2.5893 olarak gözlemlenmesi, modelin bazı ülkelerdeki enerji yoğunluğu seviyelerini öngörmeye önemli sapmalar yaşadığına işaret etmektedir. Bu uç değerler, modele dâhil edilmeyen ve sadece belirli ülkelere özgü olan yapısal veya politik faktörlerin, enerji yoğunluğu üzerinde beklenmedik etkiler yaratabileceğini düşündürmektedir. Bu bulgu, analiz edilen gelir grubunda dahi, enerji yoğunluğu dinamiklerinin tüm ülkeler için homojen olmadığını ve özel durumların ayrıntılı bir şekilde incelenmesi gerektiğini ortaya koymaktadır.

Yüksek gelir grubunda yer alan ülkelerde enerji yoğunluğunu etkileyen değişkenlerin katsayı tahminleri ve anlamlılık düzeyleri, Cauchit bağlantı fonksiyonu kullanılarak elde edilmiş ve sonuçlar Çizelge 6.16'da sunulmuştur.

Çizelge 6.16 Yüksek gelirli ülkelerin cauchit bağlantılı ortalama model için katsayılar

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-5192.91	40.49	-128.260	< 2e-16
Doğalgaz Tüketimi	1229.15	182.22	6.745	1.53e-11
Elektrik Tüketimi	-3264.37	498.66	-6.546	5.90e-11
Kömür Tüketimi	-23.52	109.00	-0.216	0.829153
Petrol Tüketimi	2075.37	583.39	3.557	0.000375
Yenilenebilir Enerji Kaynakları	-732.18	73.98	-9.896	< 2e-16
Nüfus Yoğunluğu	478.93	63.64	7.525	5.26e-14

Çizelge 6.16'daki sonuçlar, yüksek gelirli ülkelerde enerji yoğunluğunu etkileyen değişkenlerin çoğunun istatistiksel olarak anlamlı olduğunu ortaya koymaktadır. Doğalgaz ve petrol tüketimi pozitif ve anlamlı katsayılarla enerji yoğunluğunu artırmakta, yani fosil yakıt kullanımının enerji verimliliği üzerinde olumsuz bir etki yarattığını göstermektedir. Buna karşılık, elektrik tüketimi ile yenilenebilir enerji kullanımının negatif ve anlamlı katsayılarla sahip olması, bu kaynakların enerji yoğunluğunu azaltarak verimliliği artırıcı bir rol oynadığını göstermektedir. Nüfus yoğunluğunun pozitif ve güçlü biçimde anlamlı çıkması, yoğun kentleşmenin enerji talebini artırarak enerji yoğunluğunu yükselttiğini ortaya koymaktadır. Öte yandan kömür tüketiminin katsayısı anlamsız bulunmuş, bu da yüksek gelirli ülkelerde kömürün enerji portföyündeki sınırlı rolünü yansıtmaktadır.

Genel olarak, sonuçlar fosil yakıt bağımlılığının azaltılması ve yenilenebilir enerji ile elektrifikasyon politikalarının yaygınlaştırılmasının enerji verimliliğini artırmada kritik öneme sahip olduğunu göstermektedir.

Yüksek gelirli ülkeler için elde edilen beta regresyon modelinin tahmin performansına ilişkin bulgular Çizelge 6.17'de yer almaktadır.

Çizelge 6.17 Yüksek gelirli ülkelerin beta regresyon performans ölçüm kriteri değerleri

MSE	RMSE	MAE	Pseudo R^2
0.000000000015	0.0000039	0.0000029	0.90

Çizelge 6.17’de yer alan performans ölçüm kriterleri, yüksek gelirli ülkeler için kurulan beta regresyon modelinin güçlü bir tahmin kapasitesine sahip olduğunu göstermektedir. R^2 değerinin %90 seviyesinde olması, bağımlı değişken olan enerji yoğunluğunun açıklanmasında modelin yüksek bir açıklayıcılık gücüne sahip olduğunu ortaya koymaktadır. Bununla birlikte, MSE, RMSE ve MAE değerlerinin oldukça düşük çıkması, bağımlı değişkenin 0 – 1 aralığında tanımlı olmasından kaynaklanabilmektedir. Bu nedenle, hata ölçütleri tek başına değerlendirilmemiş; modelin güvenilirliği, R^2 ’nin yüksekliği ve katsayıların istatistiksel anlamlılık düzeyleri ile birlikte ele alınmıştır. Bu sonuçlar, modelin yüksek gelirli ülkelerde enerji yoğunluğunu etkileyen faktörleri belirlemede güvenilir ve sağlam bir araç olduğunu teyit etmektedir.

Modelin performans ölçütleri güçlü olmakla birlikte, katsayıların güvenilir bir şekilde yorumlanabilmesi için bağımsız değişkenler arasında çoklu doğrusal bağlantı riskinin de değerlendirilmesi gerekmektedir. Bu nedenle yüksek gelirli ülkeler için VIF değerleri hesaplanmış ve Çizelge 6.18’de sunulmuştur.

Çizelge 6.18 Yüksek gelirli ülkelerin VIF değerleri

Doğalgaz Tüketimi	Elektrik Tüketimi	Kömür Tüketimi	Petrol Tüketimi	Yenilenebilir Enerji Kaynakları	Nüfus Yoğunluğu
53.821	172.079	10.579	231.122	2.474	4.591

Çizelge 6.18’de sunulan VIF değerleri incelendiğinde, özellikle petrol (231.12), elektrik (172.07), doğalgaz (53.82), ve kömür (10.57) değişkenlerinde çok yüksek düzeyde çoklu doğrusal bağlantı sorunu bulunduğu görülmektedir. Genel olarak VIF değerinin 10’un üzerinde olması ciddi bir çoklu bağlantı problemine işaret etmektedir. Bu sonuç, yüksek

gelirli ülkelerde özellikle fosil yakıt tüketimi ile demografik göstergelerin birbirleriyle yüksek korelasyon içinde olduğunu göstermekte ve katsayı tahminlerinin güvenilirliğini zayıflatmaktadır. Buna karşılık, nüfus yoğunluğu (4.59) ve yenilenebilir enerji kullanımı (2.47) değişkenleri görece düşük VIF değerlerine sahip olup önemli bir çoklu doğrusal bağlantı riski taşımamaktadır. Bu nedenle, katsayı tahminlerinin güvenilirliğini artırmak ve çoklu bağlantı etkisini azaltmak amacıyla Beta ridge regresyon uygulanmıştır. Beta ridge regresyon sonuçları Çizelge 6.19’da sunulmaktadır.

Çizelge 6.19 Yüksek gelirli ülkelerin cauchit bağlantılı beta ridge ortalama model için katsayılar

Katsayı	Tahmin	Standart Hata	z değeri	p
Sabit Terim	-4898.555	72.81673	-67.27238	0
Doğalgaz Tüketimi	209.035	86.31999	2.42162	1.5451e-02
Elektrik Tüketimi	-229.223	88.49518	-2.59023	9.5911e-03
Kömür Tüketimi	167.287	69.07315	2.42188	1.5440e-02
Petrol Tüketimi	-59.715	86.37014	-0.69138	4.8933e-01
Yenilenebilir Enerji Kaynakları	-697.145	86.12853	-8.09424	5.7625e-16
Nüfus Yoğunluğu	667.420	62.72337	10.64069	1.9269e-26

Beta ridge regresyon bulguları incelendiğinde, doğalgaz tüketiminin enerji yoğunluğunu pozitif ve anlamlı biçimde etkilediği görülmektedir. Bu sonuç, fosil yakıtların genel olarak enerji yoğunluğunu artırıcı etkisine işaret eden literatürle uyumlu olup (Sadorsky, 2013; Apergis ve Payne, 2010), aynı zamanda yüksek gelirli ülkelerde doğalgazın sanayi yapısı ve enerji dönüşüm süreçlerindeki rolünü göstermektedir.

Elektrik tüketiminin Beta ridge modelinde negatif ve anlamlı çıkması, elektrifikasyonun enerji yoğunluğunu azaltıcı bir unsur olduğunu ve enerji verimliliğini desteklediğini göstermektedir. Bu bulgu, elektrifikasyonun enerji verimliliği ile ilişkisini ortaya koyan

alıřmaları (Csereklyei vd., 2016) destekler niteliktedir. Bununla birlikte, bu etkinin srdrlebilir biimde gerekleēebilmesi iin elektrifikasyon politikalarının fosil temelli deėil, yenilenebilir odaklı tasarlanması gerektiėi; Avrupa Yeēil Mutabakatı gibi karbon ntr hedeflerin bu dnēimde yol gsterici rol oynayabileceėi deėerlendirilmektedir.

Kmr tketimi pozitif ve anlamlı katsayılarla enerji yoėunluėunu artırıcı bir etki gstermektedir. Bu bulgular, Sadorsky (2013) ve Bilgili vd. (2016)'nın alıřmalarında raporlanan, fosil yakıt baėımlılıėının enerji yoėunluėunu artırıcı etkisiyle uyumludur. Kmr tketiminin Beta ridge regresyonda anlamlı hale gelmesi, kmrn yksek gelirli lkelerde enerji portfynde tamamen dıřlanmamıř olduėunu, bazı sektrlerde (rneėin aėır sanayi) hlen kritik bir rol oynadıėını ortaya koymaktadır. Bu sonu, temiz kmr teknolojilerinin geliētirilmesi ya da kmrden ıkıř stratejilerinin hızlandırılması gerektiėini gstermektedir.

Petrol tketimi ise negatif ancak istatistiksel olarak anlamsız bulunmuřtur. Bu durum, nceki literatrde raporlanan artırıcı etkiyle tam olarak rtřmemekte, yksek gelirli lkelerde petroln enerji portfyndeki grel neminin azalmıř olabileceėine iřaret etmektedir.

Yenilenebilir enerji kullanımı ise negatif ve anlamlı bulunmuř, bu da yenilenebilir kaynakların enerji verimliliėi ve evresel srdrlebilirlik aısından kritik bir ara olduėunu teyit etmiřtir. Literatrde de (Sadorsky, 2009; Bilgili vd., 2016) yenilenebilir enerji kullanımının uzun vadede enerji yoėunluėunu azaltıcı etkiler yarattıėı vurgulanmaktadır.

Nfus yoėunluėunun pozitif ve anlamlı etkisi ise, kentleřmenin enerji talebini artırarak enerji yoėunluėunu ykselttiėini gstermektedir. Bu durum, yksek gelirli lkelerde srdrlebilir řehirleřme politikalarının (rneėin enerji verimli ulařım altyapısı, akıllı řehir uygulamaları ve kompakt kent planlaması) enerji yoėunluėunu azaltmada kilit bir rol oynayabileceėini dřndrmektedir. Poumanyong ve Kaneko (2010)'nun gelir dzeyine baėlı farklılařmaları ortaya koyan bulguları da bu yorumu desteklemektedir.

Sonuç olarak, Beta ridge regresyon yöntemi, yüksek gelirli ülkelerde enerji yoğunluğunun belirleyicilerinin daha sağlam ve güvenilir şekilde tespit edilmesine olanak tanımış; fosil yakıt bağımlılığının olumsuz, yenilenebilir enerjinin olumlu etkisinin yanı sıra doğalgazın koşullara bağlı olarak verimlilik avantajı sunabileceğini göstermiştir. Analiz, enerji politikasında teknoloji, altyapı ve kentsel dönüşüm odaklı stratejilerin önemine işaret etmektedir.

6.3 Karar Ağacı Yöntemi İle Enerji Yoğunluğunu Etkileyen Faktörlerin Belirlenmesi

Karar Ağacı Regresyonu, veriyi ardışık bölümlere ayırarak dallar üzerinden tahmin yapan, yüksek yorum kapasitesine sahip ve uygulaması nispeten kolay bir makine öğrenmesi tekniğidir. Bu yöntem, değişkenler arasındaki ilişkileri hiyerarşik bir yapı içinde ele alarak enerji yoğunluğunun tahmininde esnek ve anlaşılır bir modelleme seçeneği sunmaktadır.

Bu bölümde sunulan çalışmada farklı gelir grupları için ayrı ayrı karar ağacı modelleri geliştirilecek ve bu modellerin tahmin performansı Hata Kareler Ortalaması (MSE), Karekök Hata Kareler Ortalaması (RMSE), Ortalama Mutlak Hata (MAE) ve Belirlilik Katsayısı (R^2) gibi hata metrikleriyle değerlendirilecektir.

6.3.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi

Mısır ve Fas gibi alt-orta gelir grubundan seçilen ülkelerde enerji yoğunluğunu tahmin etmek üzere Karar Ağacı Regresyonu yöntemi kullanılacaktır. Model karşılaştırmasında faydalanılan MSE, RMSE, MAE ve R^2 gibi performans metrikleri de belirlenmiştir.

Çizelge 6.20 Alt-orta gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.000000000004	0.0000063	0.0000046	0.93

Çizelge 6.20’de sunulan karar ağacı modeli sonuçları, alt-orta gelirli ülkeler için modelin yüksek bir tahmin doğruluğuna sahip olduğunu göstermektedir. Hata metriklerinin son derece düşük olması, modelin enerji yoğunluğu değerlerini gerçek değerlere oldukça yakın biçimde tahmin ettiğini ortaya koymaktadır. Bu durum, karar ağacı yönteminin söz konusu ülke grubunda enerji yoğunluğunu etkileyen faktörler arasındaki ilişkileri başarılı bir şekilde yakalayabildiğini ve karmaşık yapıları etkin biçimde modelleyebildiğini düşündürmektedir.

6.3.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi

Orta-yüksek gelir grubundaki ülkelerden seçilen Arjantin, Meksika ve Tayland ülkeleri için de Karar Ağacı Regresyonu yöntemi ayrı olarak uygulanacaktır. Model karşılaştırılmasında kullanılan performans metriklerinden MSE, RMSE, MAE ve R^2 değerleri hesaplanarak model seçiminde kullanılacaktır.

Çizelge 6.21 Orta-yüksek gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.00000000007	0.0000084	0.000006	0.98

İlgili gelir grubu için Çizelge 6.21’de ifade edilen performans metrikleri sonuçları modelin enerji yoğunluğunu yüksek bir doğrulukla tahmin ettiğini göstermektedir. R^2 değerinin %98 seviyesinde olması, karar ağacının bu gelir grubundaki ülkelerde değişkenler arasındaki ilişkiyi başarıyla yansıttığını ortaya koymaktadır.

6.3.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin karar ağacı yöntemi ile belirlenmesi

Yüksek gelir grubuna ait ülkelerde enerji yoğunluğu tahmininde Karar Ağacı Regresyonu yöntemi kullanılacaktır. Elde edilen MSE, RMSE, MAE ve R^2 değerleri doğrultusunda

modelin doğruluk düzeyi değerlendirilecek ve genel karşılaştırma için temel oluşturacaktır.

Çizelge 6.22 Yüksek gelirli ülkelerin karar ağacı performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R²
0.0000000000004	0.000002	0.0000016	0.98

Yüksek gelirli ülke grubuna ait model için elde edilen performans ölçüm kriteri değerleri Çizelge 6.22’de yer almaktadır. İlgili ülke grubuna ait model son derece düşük hata metrikleri ile dikkat çekmektedir. Özellikle MSE, RMSE ve MAE değerlerinin çok küçük olması, modelin enerji yoğunluğu tahminlerinde yüksek bir doğruluk ve istikrar sağladığını göstermektedir. Bu bulgular, karar ağacı yönteminin yüksek gelirli ülkelerde enerji yoğunluğunu etkileyen faktörler arasındaki ilişkileri başarıyla yakalayabildiğini ve bu ülke grubuna özgü veri desenlerini etkili biçimde temsil edebildiğini ortaya koymaktadır. Modelin tahmin gücünün bu denli yüksek olması, karar ağacı algoritmasının bu grup özelinde güvenilir ve güçlü bir tahmin aracı olarak değerlendirilebileceğini göstermektedir.

6.4 Rastgele orman regresyonu yöntemi ile enerji yoğunluğunu etkileyen faktörlerin belirlenmesi

Bu bölümde, enerji yoğunluğunu etkileyen faktörleri belirlemek amacıyla Rastgele Orman Regresyonu (Random Forest Regression) yönteminden faydalanılmıştır. Rastgele Orman, birden çok karar ağacının birleşimiyle oluşturulan ve bağımsız değişkenlerle bağımlı değişken arasındaki karmaşık ilişkileri başarıyla modelleyebilen güçlü bir öngörü modelidir.

Çalışmamızda, Dünya Bankası'nın gelir sınıflamasına uygun olarak ülkeler alt-orta, orta-yüksek ve yüksek gelir gruplarına ayrılarak her bir grup için özelleştirilmiş Rastgele Orman modelleri geliştirilmiştir. Bu modeller sayesinde enerji yoğunluğunu etkileyen temel etmenler belirlenmiştir. Modellerin tahmin başarısı ise Hata Kareler Ortalaması

(MSE), Karekök Hata Kareler Ortalaması (RMSE), Ortalama Mutlak Hata (MAE) ve Belirlilik Katsayısı (R^2) gibi performans ölçütleri kullanılarak değerlendirilmiştir.

6.4.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi

Mısır ve Fas gibi alt-orta gelir grubundan seçilen ülkelerde enerji yoğunluğunu etkileyen faktörleri tespit etmek amacıyla Rastgele Orman Regresyonu yöntemi uygulanmıştır. Bu uygulama neticesinde, modelin performansını gösteren MSE, RMSE, MAE ve R^2 değerleri ortaya konulmuştur.

Çizelge 6.23 Alt-orta gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.00000000000002	0.000000044	0.000000033	0.99

Çizelge 6.23’de sunulmuş olan performans metrikleri incelendiğinde, modelin düşük hata oranlarına sahip olduğu ve enerji yoğunluğu üzerindeki yapısal ilişkiyi başarılı bir şekilde yansıttığı görülmektedir. MSE ve RMSE değerlerinin çok düşük olması, modelin tahmin hatalarının sınırlı düzeyde gerçekleştiğini göstermektedir.

Rastgele Orman yöntemi ile oluşturulan modelin yüksek açıklayıcılık gücüne sahip olması, enerji yoğunluğunu etkileyen faktörlerin doğru bir şekilde belirlenmesine imkân sağlamış ve modelin analiz açısından güvenilir bir araç olduğunu ortaya koymuştur.

6.4.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi

Orta-yüksek gelir grubunda yer alan Arjantin, Meksika ve Tayland için gerçekleştirilen Rastgele Orman Regresyonu analizi, enerji yoğunluğunu etkileyen temel faktörlerin belirlenmesine yönelik olarak uygulanmıştır. Modelin performans metriklerine ilişkin sonuçlar elde edilmiştir.

Çizelge 6.24 Orta-yüksek gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.00000000000007	0.000000083	0.000000044	0.99

Çizelge 6.24'te yer verilen performans kriterleri değerlendirildiğinde, modelin enerji yoğunluğunu tahmin etmede yüksek bir doğruluk sergilediği görülmektedir. Düşük MSE ve RMSE değerleri, modelin hata oranlarının sınırlı olduğunu ve tahmin kapasitesinin güçlü olduğunu göstermektedir.

Elde edilen sonuçlar, orta-yüksek gelir grubunda enerji yoğunluğunu etkileyen faktörlerin belirlenmesinde Rastgele Orman yönteminin etkili bir analiz aracı olduğunu ve bu yöntemin politika yapıcılar için önemli çıkarımlar sunabileceğini ortaya koymaktadır.

6.4.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin rastgele orman regresyonu yöntemi ile belirlenmesi

Yüksek gelir grubunda yer alan Danimarka, Hollanda ve İngiltere için gerçekleştirilen Rastgele Orman Regresyonu analizi, enerji yoğunluğunu etkileyen değişkenlerin belirlenmesi amacıyla uygulanmıştır. Modelin performans ölçüm kriterlerine ilişkin sonuçlar Çizelge 6.25'de yer almaktadır.

Çizelge 6.25 Yüksek gelirli ülkelerin rastgele orman performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.00000000000004	0.000000068	0.000000045	0.99

Yüksek gelirli ülkeler grubunda uygulanan Rastgele Orman Regresyonu modelinin enerji yoğunluğunu açıklama konusunda yüksek bir başarıya ulaştığını göstermektedir. Düşük MSE, RMSE ve MAE değerleri, modelin hata oranlarının son derece düşük olduğunu ortaya koymaktadır.

Elde edilen sonuçlar, yüksek gelirli ülkelerde enerji yoğunluğu ile ekonomik ve demografik faktörler arasındaki ilişkinin düzenli ve öngörülebilir bir yapıya sahip olduğunu göstermekte ve Rastgele Orman yönteminin bu grup özelinde güvenilir tahminler sunduğunu teyit etmektedir.

Rastgele Orman yöntemi, üç gelir grubunda da en yüksek tahmin doğruluğunu sağlamıştır (alt-orta: $R^2 = 0.99$; orta-yüksek: $R^2 = 0.99$; yüksek: $R^2 = 0.99$). Literatürde de vurgulandığı üzere, topluluk (bagging) temelli orman modelleri tekil ağaçlara kıyasla daha düşük varyans ve daha güçlü genelleme kapasitesi sunmaktadır; bu çalışmada elde edilen üstün performans da yöntem literatürüyle uyumludur (Breiman, 2001). Enerji talebi ve enerji yoğunluğu gibi karmaşık sistemlerde Rastgele Orman'ın öne çıkan başarısı, doğrusal olmayan örüntüleri yakalama kapasitesiyle açıklanabilir.

6.5 Destek Vektör Regresyonu Yöntemi İle Enerji Yoğunluğunu Etkileyen Faktörlerin Belirlenmesi

Enerji yoğunluğunu etkileyen etmenlerin belirlenmesine yönelik olarak bu kısımda Destek Vektör Regresyonu (Support Vector Regression) yönteminden yararlanılmıştır. SVR, hem doğrusal hem de doğrusal olmayan veri setlerinde yüksek performans sergileyen, dayanıklı bir regresyon algoritmasıdır. Bu algoritma, veri noktalarını en uygun şekilde ayırarak karmaşık örüntüleri modelleyebilir.

SVR modeli, Radyal Temelli Fonksiyon (RBF) çekirdeği kullanılarak ve varsayılan hiperparametreler ($C=1.0$, $\epsilon=0.1$, $\gamma=\text{'scale'}$) altında uygulanmıştır.

Çalışmada, ülkeler Dünya Bankası'nın gelir sınıflamasına uygun biçimde alt-orta, orta-yüksek ve yüksek gelir düzeylerine sahip gruplara ayrılmıştır. Söz konusu her bir ülke grubu özelinde bağımsız SVR modelleri kurulmuş ve enerji yoğunluğunun ana belirleyicileri tespit edilmiştir. Modellerin tahmin gücü; MSE, RMSE, MAE ve R^2 gibi yaygın performans ölçütleri dikkate alınarak analiz edilmiştir.

6.5.1 Alt-orta gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi

Enerji yoğunluğunu etkileyen değişkenlerin analizi, alt-orta gelir grubundaki Mısır ve Fas'ta Destek Vektör Regresyonu yöntemiyle gerçekleştirilmiştir. Bu sürecin sonunda, modelin tahmin başarısına ilişkin MSE, RMSE, MAE ve R^2 değerleri elde edilmiştir.

Çizelge 6.26 Alt-orta gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.000000000004	0.0000062	0.00000444	0.93

Çizelge 6.26'da yer alan performans metrikleri incelendiğinde, modelin enerji yoğunluğunu yüksek doğrulukla tahmin ettiği görülmektedir. Düşük hata oranları, modelin tahmin gücünü desteklemekte; R^2 değerinin yüksekliği ise değişkenler ile enerji yoğunluğu arasındaki ilişkinin başarılı şekilde yakalandığını göstermektedir. Elde edilen bulgular, alt-orta gelir grubunda SVR yönteminin etkili bir analiz aracı olduğunu ortaya koymaktadır.

6.5.2 Orta-yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi

Arjantin, Meksika ve Tayland'ın yer aldığı orta-yüksek gelir grubunda, enerji yoğunluğunu etkileyen faktörler Destek Vektör Regresyonu modeli aracılığıyla incelenmiştir. Çalışma sonucunda modelin performans ölçüm kriterleri tespit edilmiştir.

Çizelge 6.27 Orta-yüksek gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.000000000004	0.0000065	0.0000052	0.98

Enerji yoğunluğu tahmininde modelin güçlü bir performans gösterdiği tespit edilmiştir. Hata metriklerinin düşük, R^2 değerinin ise yüksek olması, enerji yoğunluğu üzerindeki bağımsız değişken etkilerinin başarılı bir şekilde modellendiğine işaret etmektedir. Çizelge 6.27’de de gösterildiği üzere, SVR yönteminin ilgili ülkeler için güvenilir ve etkili bir tahmin aracı olarak değerlendirilebileceği söylenebilir.

6.5.3 Yüksek gelirli ülkelerin enerji yoğunluğunu etkileyen faktörlerin destek vektör regresyonu yöntemi ile belirlenmesi

Yüksek gelir grubunda yer alan Danimarka, Hollanda ve İngiltere için Destek Vektör Regresyonu yöntemi kullanılarak enerji yoğunluğunu etkileyen faktörler incelenmiştir. Çalışma kapsamında, modelin başarısını değerlendirmek üzere kritik hata metriklerine ait değerler belirlenmiştir. Bu sonuçlar, modelin enerji yoğunluğu tahminindeki doğruluğu ve güvenilirliği hakkında önemli bilgiler sunmaktadır.

Çizelge 6.28 Yüksek gelirli ülkelerin destek vektör performans ölçüm kriteri değerleri

MSE	RMSE	MAE	R^2
0.00000000002	0.0000044	0.0000032	0.94

Modelin performans göstergeleri, enerji yoğunluğunun yüksek doğrulukla tahmin edildiğini ve değişkenler arasındaki ilişkinin başarılı bir şekilde yansıtıldığını ortaya koymaktadır. Çizelge 6.28’deki performans ölçüm kriteri değerlerinde görüldüğü üzere, düşük hata oranları ve yüksek açıklayıcılık gücü, Destek Vektör yönteminin yüksek gelirli ülkelerde enerji yoğunluğu analizinde etkin bir araç olduğunu göstermektedir.

Destek Vektör Regresyonu (SVR), üç gelir grubunda da yüksek doğruluk sağlamıştır (alt-orta: $R^2 = 0.93$; orta-yüksek: $R^2 = 0.98$; yüksek: $R^2 = 0.94$). Çekirdek fonksiyonları sayesinde doğrusal olmayan ilişkilerin etkin biçimde modellenbildiği görülmektedir. Bununla birlikte, SVR’nin yorumlanabilirliği Rastgele Orman ve Karar Ağaçlarına kıyasla daha sınırlıdır. Literatürde de vurgulandığı üzere, SVR yüksek boyutlu ve

doğrusal olmayan problemler için güçlü bir yaklaşım olarak kabul edilmektedir (Smola & Schölkopf, 2004; Drucker vd., 1997).

6.6 Farklı Gelir Gruplarındaki Ülkeler için Makine Öğrenmesi Yöntemlerinin Kıyaslanması

Bu bölümde, enerji yoğunluğunu etkileyen faktörlerin modellenmesine yönelik olarak kullanılan makine öğrenmesi yöntemlerinin performansları karşılaştırılmıştır. Her bir modelin performansı; Hata Kareler Ortalaması (MSE), Karekök Hata Kareler Ortalaması (RMSE), Ortalama Mutlak Hata (MAE) ve Belirlilik Katsayısı (R^2) gibi temel tahmin doğruluğu metrikleri üzerinden değerlendirilmiştir. Bu kapsamda yapılan analizler, her yöntemin farklı gelir düzeylerine sahip ülkelerin veri yapısına ne derece uyum sağladığını ve hangi modellerin daha yüksek tahmin başarısı sunduğunu ortaya koymaktadır.

6.6.1 Alt-orta gelirli ülkeler için makine öğrenmesi yöntemlerinin karşılaştırılması

Çizelge 6.29, alt-orta gelir grubundaki Mısır ve Fas'a uygulanan Rastgele Orman Regresyonu, Destek Vektör Regresyonu ve Karar Ağacı gibi üç farklı regresyon yönteminin model performanslarını karşılaştırmak amacıyla hesaplanan hata metriklerini içermektedir.

Çizelge 6.29 Alt-orta gelirli ülke grubu için uygulanan modellerin performans ölçüm kriteri değerleri

Model	MSE	RMSE	MAE	R^2
Rastgele Orman	0.0000000000002	0.000000044	0.000000033	0.99
Destek Vektör	0.000000000004	0.00000062	0.000000444	0.93
Karar Ağacı	0.000000000004	0.00000063	0.000000460	0.93

Çizelge 6.29'daki performans ölçütleri incelendiğinde, alt-orta gelirli ülkeler için kurulan üç farklı makine öğrenmesi modelinin tahmin başarılarının belirgin şekilde farklılaştığı görülmektedir. Rastgele Orman modeli, son derece düşük MSE, RMSE ve MAE

değerlerinin yanı sıra %99 düzeyindeki R^2 ile açık ara en yüksek tahmin doğruluğunu sağlamıştır. Destek Vektör Regresyonu ve Karar Ağacı yöntemleri ise benzer performans sergilemiş, her ikisi de $R^2 = 0.93$ düzeyinde açıklayıcılık sunmakla birlikte hata metriklerinin görece yüksek olması, bu modellerin tahmin gücünü sınırlamıştır. Genel bir değerlendirme yapıldığında, Rastgele Orman modeli alt-orta gelir grubundaki enerji yoğunluğunun tahmininde en güvenilir yöntem olarak öne çıkarken, Destek Vektör ve Karar Ağacı modelleri kabul edilebilir ancak görece daha düşük doğruluk düzeyine sahiptir. Bu nedenle, Rastgele Orman modeliyle elde edilen değişken önem sonuçları, alt-orta gelir grubunda enerji yoğunluğunu belirleyen temel faktörlerin karşılaştırmalı analizine zemin oluşturmak amacıyla Çizelge 6.30'da sunulmuştur.

Çizelge 6.30 Alt-orta gelirli ülke grubu için değişken önem düzeyi değerleri

Değişken	Değişken Önem Düzeyi Değeri
Petrol Tüketimi	0.31
Elektrik Tüketimi	0.27
Doğalgaz Tüketimi	0.23
Nüfus Yoğunluğu	0.1
Yenilenebilir Enerji Kaynakları	0.05
Kömür Kullanımı	0.02

Alt-orta gelir grubuna ilişkin Rastgele Orman modeliyle elde edilen değişken önem analizine göre, enerji yoğunluğunu etkileyen en önemli değişkenin petrol tüketimi olduğu belirlenmiştir. Bu değişken, modelde en yüksek önem skorunu (0.31) alarak tahmin gücüne en fazla katkı sağlayan unsur olmuştur. Ardından elektrik (0.27) ve doğalgaz tüketimi (0.23) modelin kararlarında etkili olmuştur. Bu sonuç, söz konusu gelir grubunda enerji yoğunluğunun büyük ölçüde fosil yakıt tüketimine bağımlı olduğunu ve enerji verimliliğinin artırılması için fosil enerji tüketiminde dönüşümün kritik bir rol oynayabileceğini ortaya koymaktadır. Buna karşılık, nüfus yoğunluğu (0.10), yenilenebilir enerji kaynakları (0.05) ve kömür kullanımı (0.02) görece daha düşük önem düzeylerine sahiptir. Özellikle yenilenebilir enerji ve kömürün sınırlı katkısı, bu gelir

grubunda hem enerji arz yapısının hâlâ fosil temelli olduğunu hem de alternatif enerji kaynaklarının enerji yoğunluğu üzerindeki etkisinin henüz zayıf kaldığını göstermektedir. Dolayısıyla, alt-orta gelirli ülkelerde enerji yoğunluğunu düşürmeye yönelik politikaların, özellikle petrol ve elektrik kullanımına odaklanarak daha verimli ve sürdürülebilir enerji stratejilerinin geliştirilmesine yönelmesi gerektiği söylenebilir.

Çizelge 6.7’de verilen Beta ridge regresyon sonuçları ile Çizelge 6.30’daki Rastgele Orman modelinde değişken önem düzeyleri birlikte değerlendirildiğinde, modeller arasında genel bir tutarlılık olduğu görülmektedir. Beta ridge regresyon modelinde doğalgaz ve petrol tüketimi pozitif ve istatistiksel olarak anlamlı bulunmuş, benzer şekilde Rastgele Orman modelinde de petrol ve elektrik en yüksek önem düzeylerine sahip değişkenler olarak öne çıkmıştır. Elektrik tüketimi Beta ridge modelinde negatif ve anlamlı çıkarken, Rastgele Orman modelinde ikinci sırada yüksek bir önem değeriyle yer almıştır. Bu durum, elektrik tüketiminin enerji yoğunluğu üzerindeki rolünü güçlü biçimde teyit etmektedir; ancak Beta ridge regresyon, bu ilişkinin enerji verimliliğini artırıcı (negatif) yönde olduğunu göstermektedir.

Nüfus yoğunluğu Beta ridge modelinde negatif ve oldukça anlamlı bir etki göstermesine karşın, Rastgele Orman analizinde değişken önem değeri yalnızca %10 düzeyinde kalmıştır. Bu farklılık, nüfus yoğunluğunun doğrusal modelde güçlü bir etki sergilerken, orman tabanlı modelde görece daha sınırlı bir katkı sağladığını ortaya koymaktadır. Yenilenebilir enerji kullanımı Beta ridge regresyonda pozitif ve anlamlı çıkmış; ancak Rastgele Orman modelinde düşük bir önem düzeyine (%5) sahip olmuştur. Bu da yenilenebilir kaynakların alt-orta gelirli ülkelerde enerji yoğunluğu üzerinde istatistiksel olarak anlamlı olsa da göreceli olarak zayıf bir belirleyici olduğuna işaret etmektedir. Kömür ise her iki yöntemde de düşük veya anlamsız bulunmuştur (Beta ridge: anlamsız katsayı; orman: %2 önem düzeyi).

Sonuç olarak, her iki yöntemin bulguları birlikte değerlendirildiğinde alt-orta gelir grubunda enerji yoğunluğunun en temel belirleyicilerinin fosil yakıt tüketimi, özellikle

petrol ve doğalgaz olduğu; elektrik tüketimi ve nüfus yoğunluğunun ise modele bağlı olarak farklı yönlerde fakat anlamlı katkılar sağladığı görülmektedir.

6.6.2 Orta-yüksek gelirli ülkeler için makine öğrenmesi yöntemlerinin karşılaştırılması

Orta-yüksek gelir grubunda yer alan Arjantin, Meksika ve Tayland ülkeleri için Çizelge 6.31’de gerçekleştirilen üç farklı modelin performans ölçüm kriteri değerleri detaylandırılmıştır.

Çizelge 6.31 Orta-yüksek gelirli ülke grubu için uygulanan modellerin performans ölçüm kriteri değerleri

Model	MSE	RMSE	MAE	R^2
Rastgele Orman	0.00000000000007	0.000000083	0.000000044	0.99
Destek Vektör	0.000000000004	0.00000065	0.00000052	0.98
Karar Ağacı	0.000000000007	0.00000084	0.0000006	0.98

Orta-yüksek gelir grubundaki ülkeler için uygulanan üç farklı modelinin performansları karşılaştırıldığında, en başarılı sonuçları Rastgele Orman modeli vermiştir. Bu model, son derece düşük hata oranları (MSE, RMSE, MAE) ve yüksek açıklayıcılık düzeyi ile öne çıkmakta; enerji yoğunluğunu en isabetli şekilde tahmin eden yöntem olarak dikkat çekmektedir. Destek Vektör Regresyonu ve Karar Ağacı modelleri de oldukça yüksek doğruluk düzeylerine ulaşarak başarılı performanslar ortaya koymuştur. Özellikle Karar Ağacı modelinin, alt-orta gelir grubuna kıyasla bu gelir grubunda daha güçlü bir performans sergilemesi dikkat çekicidir. Bu sonuçlar, orta-yüksek gelirli ülkelerde enerji yoğunluğu gibi karmaşık yapıdaki değişkenlerin modellenmesinde, doğrusal olmayan yapıları yakalayabilen ve veri esnekliği yüksek makine öğrenmesi algoritmalarının etkin olarak kullanılabileceğini ortaya koymaktadır.

Orta-yüksek gelir grubuna ait ülkeler için de alt orta gelir grubunda olduğu gibi çeşitli modelleme yöntemleri uygulanmış ve sonuçlar karşılaştırılmıştır. Bu grup özelinde de

Rastgele Orman regresyonu, diğer yöntemlere kıyasla daha başarılı sonuçlar vererek en yüksek tahmin doğruluğuna ulaşmıştır. Bu nedenle, modelin kararlarında hangi değişkenlerin daha belirleyici olduğunu ortaya koymak amacıyla değişken önem analizine başvurulmuştur. Elde edilen önem skorları, orta yüksek gelir grubunda enerji yoğunluğunu etkileyen temel faktörleri ortaya koymak üzere aşağıdaki çizelge 6.32’de sunulmaktadır.

Çizelge 6.32 Orta-yüksek gelirli ülke grubu için değişken önem düzeyi değerleri

Değişken	Değişken Önem Düzeyi Değeri
Petrol Tüketimi	0.006
Elektrik Tüketimi	0.06
Doğalgaz Tüketimi	0.44
Nüfus Yoğunluğu	0.17
Yenilenebilir Enerji Kaynakları	0.35
Kömür Kullanımı	0.002

Model sonuçlarına göre, orta-yüksek gelir grubundaki ülkelerde enerji yoğunluğunu etkileyen en belirleyici faktörün doğalgaz tüketimi (0.44) olduğu görülmektedir. Bunu sırasıyla yenilenebilir enerji kaynakları (0.35) ve nüfus yoğunluğu (0.17) izlemektedir. Özellikle yenilenebilir enerjinin bu gelir grubunda görece daha yüksek bir etki düzeyine sahip olması, enerji dönüşüm süreçlerinin bu ülkelerde daha etkili biçimde uygulanmakta olduğunu göstermektedir. Buna karşılık kömür ve petrol tüketiminin oldukça düşük önem düzeylerine sahip olması, fosil yakıt kullanımının bu grupta enerji yoğunluğu üzerindeki etkisinin sınırlı kaldığını ortaya koymaktadır.

Çizelge 6.13’te verilen Beta ridge regresyon sonuçları ile Çizelge 6.32’deki Rastgele Orman modelinde değişken önem düzeyleri birlikte değerlendirildiğinde, modeller arasında hem örtüşen hem de ayrışan noktaların bulunduğu görülmektedir. Beta ridge regresyon modelinde doğalgaz tüketimi pozitif ve anlamsız çıkarken, Rastgele Orman modelinde %44 ile en yüksek önem düzeyine sahip değişken olarak belirlenmiştir. Bu

bulgu, her iki yöntemin de doğalgazın enerji yoğunluğu üzerindeki kritik rolünü teyit ettiğini göstermektedir; ancak Beta ridge modeli bu etkinin artırıcı yönde olduğunu ortaya koymaktadır. Bu sonuç, doğalgazın sanayideki kullanım yapısı, enerji dönüşüm verimliliği veya yenilenebilir kaynaklarla ikame edilme düzeyi gibi faktörlerden dolayı, enerji yoğunluğunu farklı yönlerde etkileyebileceğini düşündürmektedir.

Elektrik tüketimi Beta ridge modelinde negatif ve yüksek düzeyde anlamlı çıkarken, Rastgele Orman modelinde %6'lık bir önem değeriyle orta düzeyde katkı sağlamıştır. Bu farklılık, elektrik tüketiminin etkisinin doğrusal olmayan yapılar altında daha görünür hale geldiğini göstermektedir. Kömür tüketimi Beta ridge regresyonda pozitif anlamlı çıkarken, Rastgele Orman modelinde %0,2 ile ihmal edilebilir düzeydedir. Dolayısıyla kömürün bu gelir grubunda enerji yoğunluğu üzerinde kayda değer bir belirleyici olmadığı görülmektedir.

Petrol tüketimi Beta ridge modelinde negatif ve anlamsız, Rastgele Orman modelinde ise %0,6'lık düşük bir önem düzeyi göstermiştir. Bu durum, petrolün orta-yüksek gelir grubunda enerji yoğunluğu üzerindeki rolünün zayıfladığını ve diğer faktörler karşısında ikincil konumda kaldığını göstermektedir. Yenilenebilir enerji kaynakları Beta ridge regresyonda pozitif ve sınırlı düzeyde anlamlı bulunurken, Rastgele Orman modelinde %35'lik yüksek bir önem değeriyle öne çıkmıştır. Bu farklılık, yenilenebilir enerji yatırımlarının kısa vadede Beta ridge modelinde net verimlilik kazançlarını yansıtmamasına rağmen, orman tabanlı modelde enerji yoğunluğunu açıklayan güçlü bir değişken olduğunu göstermektedir.

Nüfus yoğunluğu ise her iki yöntemde de tutarlı biçimde güçlü bir etki göstermiştir. Beta ridge modelinde pozitif ve yüksek düzeyde anlamlı çıkan nüfus yoğunluğu, Rastgele Orman'da da %17'lik bir önem değeriyle önemli bir belirleyici olmuştur. Bu bulgu, nüfus baskısının ve kentleşmenin orta-yüksek gelir grubunda enerji yoğunluğunu artıran en istikrarlı faktörlerden biri olduğunu teyit etmektedir.

Genel olarak değerlendirildiğinde, orta-yüksek gelir grubunda enerji yoğunluğunun en önemli belirleyicileri doğalgaz tüketimi ve nüfus yoğunluğu olurken yenilenebilir enerji ve elektrik tüketimi de yöntemlere bağlı olarak dikkate değer katkılar sağlamaktadır. Politika yapıcılar açısından bu bulgular, fosil yakıtlara bağımlılığın azaltılması, enerji dönüşüm verimliliğinin artırılması ve yenilenebilir enerji altyapısına yönelik yatırımların desteklenmesi gerektiğine işaret etmektedir. Ayrıca, nüfus yoğunluğu ve kentleşmenin enerji talebini artırıcı etkileri dikkate alınarak, enerji verimliliğini önceleyen şehir planlaması, ulaşım sistemleri ve altyapı politikalarının geliştirilmesi de bu gelir grubunda enerji yoğunluğunu düşürmeye yönelik kritik adımlar arasında değerlendirilebilir.

6.6.3 Yüksek gelirli ülkeler için beta regresyon ve makine öğrenmesi yöntemlerinin karşılaştırılması

Danimarka, Hollanda ve İngiltere gibi yüksek gelirli ülkelere ait modellerin performans ölçütleri Çizelge 6.33'te gösterilmiştir; bu değerler, modellerin etkinliğini ortaya koyan önemli verilerdir.

Çizelge 6.33 Yüksek gelirli ülke grubu için uygulanan modellerin performans ölçüm kriteri değerleri

Model	MSE	RMSE	MAE	R ²
Rastgele Orman	0.00000000000004	0.000000068	0.000000045	0.99
Destek Vektör	0.00000000000002	0.00000044	0.00000032	0.94
Karar Ağacı	0.00000000000004	0.0000002	0.00000016	0.98

Yüksek gelirli ülkeler için uygulanan modeller karşılaştırıldığında, Rastgele Orman yönteminin bu grup için de en yüksek tahmin doğruluğunu sağladığını ortaya koymaktadır. Modelin hem hata oranları son derece düşük hem de açıklayıcılığı en yüksek düzeydedir. Karar Ağacı ve Destek Vektör Regresyonu modelleri de başarılı performanslarıyla öne çıkmıştır. Özellikle Karar Ağacı modeli, bu gelir grubunda önceki gruplara kıyasla daha güçlü bir tahmin yeteneği sergileyerek dikkat çekmiştir. Destek Vektör Regresyonu da hata oranlarının düşüklüğü ve yüksek açıklama gücüyle öne

çıkılmaktadır. Bu bulgular, yüksek gelirli ülkelerde enerji yoğunluğu gibi çok değişkenli ve karmaşık yapıya sahip göstergelerin tahmininde, makine öğrenmesi tabanlı algoritmaların daha doğru, güçlü ve güvenilir sonuçlar sunduğunu göstermektedir.

Yüksek gelir grubuna ait ülkeler için de aynı yöntemsel çerçevede model karşılaştırmaları gerçekleştirilmiş ve Rastgele Orman regresyonu bu grupta da en yüksek tahmin doğruluğunu sağlayan yöntem olmuştur. Bu nedenle, enerji yoğunluğunu etkileyen değişkenlerin model üzerindeki göreceli etkilerini değerlendirmek amacıyla değişken önem analizi yapılmıştır. Aşağıdaki 6.34'deki çizelgede, yüksek gelir grubuna yönelik olarak elde edilen değişken önem düzeyleri yer almakta olup, bu bulgular grup içi öncelikli belirleyicilerin ortaya konmasına imkân tanımaktadır.

Çizelge 6.34 Yüksek gelirli ülke grubu için değişken önem düzeyi değerleri

Değişken	Değişken Önem Düzeyi Değeri
Petrol Tüketimi	0.01
Elektrik Tüketimi	0.02
Doğalgaz Tüketimi	0.03
Nüfus Yoğunluğu	0.52
Yenilenebilir Enerji Kaynakları	0.33
Kömür Kullanımı	0.07

Yüksek gelir grubuna ait ülkelerde, Rastgele Orman modeline göre enerji yoğunluğunu etkileyen en önemli değişkenin nüfus yoğunluğu (0.52) olduğu belirlenmiştir. Bunu, dikkat çekici şekilde yüksek bir oranla yenilenebilir enerji kaynakları (0.33) kullanımı izlemiştir. Bu sonuç, söz konusu ülkelerde enerji verimliliği politikalarının daha çok nüfus yapısı ve yenilenebilir enerji yatırımları üzerinden şekillendiğini göstermektedir. Öte yandan, fosil yakıt tüketimine ilişkin değişkenlerin (petrol, doğalgaz, kömür) modelde görece düşük önem düzeyine sahip olması, bu grubun enerji yoğunluğu açısından fosil yakıt bağımlılığının azaldığına işaret etmektedir.

Çizelge 6.19'daki Beta ridge regresyon bulguları ile Çizelge 6.34'te verilen Rastgele Orman değişken önem düzeyleri birlikte değerlendirildiğinde, yüksek gelirli ülkelerde enerji yoğunluğunu belirleyen faktörlerin bazı noktalarda örtüştüğü, bazı noktalarda ise ayrıştığı görülmektedir. Her iki model de nüfus yoğunluğunu en güçlü belirleyici olarak ortaya koymuştur. Beta ridge modelinde nüfus yoğunluğu pozitif ve son derece anlamlı bir katsayıya sahipken, Rastgele Orman modelinde %52 gibi en yüksek önem değeriyle ilk sırada yer almıştır. Bu tutarlılık, kentleşmenin ve yoğun nüfus baskısının yüksek gelirli ülkelerde enerji talebini artırarak enerji yoğunluğunu yükselten en kritik faktörlerden biri olduğunu teyit etmektedir.

Yenilenebilir enerji kaynakları Beta ridge modelinde negatif ve anlamlı katsayı ile enerji yoğunluğunu azaltıcı bir unsur olarak bulunmuştur. Rastgele Orman modelinde de %33 gibi yüksek bir önem derecesiyle ikinci sırada yer almıştır. İki yöntem arasındaki bu uyum, yenilenebilir enerjinin uzun vadede verimlilik artışı ve karbon emisyonlarının azaltılması açısından temel politika aracı olduğunu güçlü biçimde desteklemektedir.

Fosil yakıtlar açısından ise daha farklı bir tablo ortaya çıkmaktadır. Beta ridge modelinde kömür ve doğalgaz tüketimi pozitif ve anlamlı katsayılarla enerji yoğunluğunu artırıcı bulunmuş; bu da literatürdeki fosil bağımlılığının olumsuz etkilerine dair bulgularla örtüşmüştür (Sadorsky, 2013; Bilgili vd., 2016). Rastgele Orman modelinde ise kömür %7, doğalgaz %3 gibi oldukça düşük önem derecelerine sahip çıkmıştır. Bu durum, kömür ve doğalgaz tüketimi değişkenlerinin yüksek gelirli ülkelerin enerji yoğunluğu üzerindeki etkilerinin belirli sektörlerle (örneğin ağır sanayi, ulaştırma) özgü kaldığını, dolayısıyla orman tabanlı modelde görece daha sınırlı bir rol oynadığını düşündürmektedir.

Petrol tüketimi, Beta ridge modelinde negatif ve anlamsız bulunmuş Rastgele Orman modelinde ise yalnızca %3'lük düşük bir önem düzeyi göstermiştir. Bu farklılık, Beta ridge modelinin doğalgazın verimlilik avantajlarını yakalayabildiğini, Rastgele Orman modelinin ise enerji yoğunluğunu açıklarken doğalgazın rolünü görece daha sınırlı değerlendirdiğini göstermektedir. Bu sonuç, yüksek gelirli ülkelerde doğalgazın elektrik

üretimi ve sanayi süreçlerinde yenilenebilir enerjiyle entegrasyonunun, enerji verimliliğinde belirgin kazanımlar yarattığını ortaya koymaktadır.

Elektrik tüketimi Beta ridge regresyonda negatif ve anlamlı bulunurken, Rastgele Orman modelinde %2’lik oldukça düşük bir önem derecesine sahip olmuştur. Bu paralellik, elektrifikasyonun tek başına enerji yoğunluğunu açıklamada sınırlı olduğunu, ancak yenilenebilir kaynaklarla bütünleştğinde anlamlı katkı sağlayabileceğini göstermektedir.

Genel olarak değerlendirildiğinde, yüksek gelirli ülkelerde enerji yoğunluğunun azaltılmasına yönelik stratejilerin üç temel eksen üzerine odaklanması gerektiği söylenebilir: (i) kentleşme ve nüfus yoğunluğunun enerji talebine baskısını azaltacak sürdürülebilir şehirleşme politikaları, (ii) yenilenebilir enerji yatırımlarının artırılması, (iii) kömür ve petrol bağımlılığının azaltılması ve doğalgazın geçiş yakıtı olarak daha verimli kullanılması.

7. SONUÇ VE DEĞERLENDİRME

Enerji yoğunluğu ve enerji verimliliği, günümüz dünyasında yalnızca ekonomik performansı değil, aynı zamanda çevresel sürdürülebilirliği ve sosyal refah düzeyini de doğrudan etkileyen stratejik göstergeler arasında yer almaktadır. Enerji yoğunluğu, bir ekonominin üretim kapasitesine karşılık ne kadar enerji tükettiğini gösteren bir ölçüttür ve enerji verimliliğinin tersini yansıtır. Bu göstergenin zaman içindeki azalma eğilimi, ekonomik büyümenin daha az enerji harcanarak sürdürülebilir biçimde gerçekleştirilebildiğini ortaya koymaktadır. Dolayısıyla enerji yoğunluğundaki iyileşmeler hem üretkenliğin artışı hem de kaynakların daha etkin kullanımı anlamına gelmektedir.

Enerji verimliliğinin artırılması, yalnızca enerji kaynaklarının tasarruflu kullanılmasına değil, aynı zamanda sera gazı emisyonlarının azaltılmasına, enerji arz güvenliğinin sağlanmasına ve dışa bağımlılığın düşürülmesine de hizmet etmektedir. Özellikle fosil yakıtlara olan bağımlılığın azaltılması, çevresel zararın azaltılması açısından kritik önemdedir. Bu nedenle, enerji verimliliği uygulamaları ve yenilenebilir enerji yatırımları, Paris İklim Anlaşması, Birleşmiş Milletler Sürdürülebilir Kalkınma Hedefleri ve Avrupa Yeşil Mutabakatı gibi uluslararası politikaların merkezinde yer almaktadır.

Günümüzde ülkeler yalnızca ekonomik büyümeyi değil, aynı zamanda bu büyümenin çevreye olan etkilerini de dikkate alarak politika geliştirme zorunluluğuyla karşı karşıyadır. Enerji yoğunluğu bu bağlamda, üretim yapısındaki dönüşümün, enerji teknolojilerindeki ilerlemenin ve çevresel politikaların başarısının ölçülebilir bir çıktısı olarak değerlendirilmektedir. Bu nedenle, enerji yoğunluğunu etkileyen faktörlerin istatistiksel ve analitik yöntemlerle incelenmesi, sadece akademik değil aynı zamanda pratik ve yönetsel anlamda da önemli bir katkı sunmaktadır.

Bu tezde enerji yoğunluğunu etkileyen faktörler, Dünya Bankası sınıflandırmasına göre alt-orta, orta-yüksek ve yüksek gelirli ülke grupları için ayrı ayrı incelenmiştir. Çalışmada öncelikle Beta regresyon modelleri uygulanmış ve her gelir grubu için en uygun bağlantı fonksiyonları belirlenmiştir. Alt-orta gelirli ülkelerde Log-log, orta-yüksek ve yüksek gelirli ülkelerde ise Cauchit bağlantısı en iyi uyum ölçütlerini sağlamıştır. Bu aşamada elde edilen sonuçlar, enerji yoğunluğunu etkileyen faktörlerin ülkelerin gelir düzeyine göre farklılık gösterebildiğini ortaya koymuştur.

Beta regresyon bulguları, alt-orta gelirli ülkelerde petrol ve doğalgaz tüketiminin enerji yoğunluğunu artırıcı, elektrik tüketiminin ise azaltıcı etkisini vurgulamıştır. Orta-yüksek gelirli ülkelerde fosil yakıtların artırıcı etkisi belirginleşirken, elektrik tüketimi ve yenilenebilir enerji kullanımı azaltıcı yönde öne çıkmıştır. Yüksek gelirli ülkelerde ise fosil yakıtların olumsuz, yenilenebilir enerji kullanımının ise olumlu etkileri teyit edilmiş; ayrıca nüfus yoğunluğunun güçlü ve pozitif katkısı dikkat çekmiştir.

Bununla birlikte, yüksek R^2 değerleri ve düşük hata ölçütlerine rağmen, Beta regresyon katsayılarının güvenilirliği açısından çoklu doğrusal bağlantı sorununun dikkate alınması gerekmiştir. Bu amaçla Varyans Şişirme Faktörü (VIF) değerleri hesaplanmış ve özellikle petrol, doğalgaz, elektrik ve yenilenebilir enerji değişkenlerinde yüksek değerler elde edilmiştir. Çoklu bağlantının yaratabileceği sapmaları azaltmak amacıyla Beta ridge regresyon analizi uygulanmıştır. Beta ridge modelleri, daha istikrarlı katsayı tahminleri sunmuş ve Beta regresyon bulgularını büyük ölçüde desteklemiştir. Örneğin alt-orta gelirli ülkelerde petrol ve doğalgaz tüketimi pozitif ve anlamlı etkiler gösterirken, elektrik tüketimi ile nüfus yoğunluğu negatif ve anlamlı bulunmuştur. Orta-yüksek gelirli ülkelerde kömür ve doğalgaz tüketimi ile nüfus yoğunluğu değişkenleri pozitif ve anlamlı çıkmış; yenilenebilir enerjinin kısa vadede pozitif işaret alması geçiş maliyetleri ile açıklanmıştır. Yüksek gelirli ülkelerde ise kömür ve doğalgaz tüketimi artırıcı, yenilenebilir enerji azaltıcı ve anlamlı bulunmuş; elektriğin negatif işareti ise gelişmiş ülkelerdeki görece verimli kullanımına işaret etmiştir.

Tezin üçüncü aşamasında, değişkenler arasındaki doğrusal olmayan ilişkileri yakalayabilmek amacıyla makine öğrenmesi yöntemleri uygulanmıştır. Karar Ağaçları, Destek Vektör Regresyonu ve Rastgele Orman modelleri karşılaştırılmış; en yüksek tahmin başarısı Rastgele Orman yöntemiyle elde edilmiştir. Model performans ölçütleri (MSE, RMSE, MAE, R^2) bu yöntemin enerji yoğunluğunu en düşük hata ve en yüksek açıklama gücüyle tahmin ettiğini göstermiştir. Bu nedenle değişken önem düzeyleri Rastgele Orman sonuçları üzerinden değerlendirilmiştir. Bulgular, alt-orta gelirli ülkelerde petrol ve elektrik tüketiminin; orta-yüksek gelirli ülkelerde doğalgaz ve yenilenebilir enerjinin; yüksek gelirli ülkelerde ise nüfus yoğunluğu ve yenilenebilir enerjinin enerji yoğunluğunu belirleyen temel değişkenler olduğunu ortaya koymuştur.

Sonuçlar bir bütün olarak değerlendirildiğinde, her üç gelir grubunda da fosil yakıt tüketiminin enerji yoğunluğunu artırıcı, yenilenebilir enerji kaynaklarının ise azaltıcı etkiler sunduğu genel eğilimin teyit edildiği görülmektedir. Bununla birlikte, ülkelerin gelir düzeyine göre farklılıklar öne çıkmaktadır:

Alt-orta gelir grubunda enerji yoğunluğu daha çok petrol ve doğalgaz tüketimi tarafından şekillendirilmekte, elektrik ve nüfus yoğunluğu verimlilik sağlayıcı rol üstlenmektedir.

Orta-yüksek gelir grubunda doğalgaz ve nüfus yoğunluğu en güçlü belirleyiciler olurken, yenilenebilir enerji kaynaklarının önemi giderek artmaktadır.

Yüksek gelir grubunda ise nüfus yoğunluğu ve yenilenebilir enerji belirleyici rol oynamakta, kömür ve petrolün olumsuz etkileri devam etmektedir.

Elde edilen bulgular, politika düzeyinde üç temel yönelimi ortaya koymaktadır. İlk olarak, fosil yakıt bağımlılığının özellikle petrol, doğalgaz ve kömür tüketiminde azaltılması gerekmektedir. Alt-orta gelirli ülkelerde fosil yakıt kullanımının hâlen enerji yoğunluğu üzerinde baskın bir rol oynadığı, dolayısıyla bu ülkelerin enerji dönüşümünü hızlandırmaları gerektiği anlaşılmaktadır. İkinci olarak, yenilenebilir enerji yatırımlarının güçlendirilmesi tüm gelir grupları için kritik bir öneme sahiptir. Orta-yüksek gelirli

lkelerde altyapı ve teknoloji yatırımları, yksek gelirli lkelerde ise karbon ntr hedefleri doęrultusunda yenilenebilir kullanımının yaygınlaştırılması n plana çıkmaktadır. nc olarak, nfus yoęunluęunun enerji talebi zerindeki baskısını azaltacak srdrlebilir řehirleřme politikalarının geliřtirilmesi gerekmektedir. zellikle yksek gelirli lkelerde kentleřmenin enerji yoęunluęunu artırıcı etkisi gl biimde gzlenmiř, bu nedenle enerji verimli ulařım, akıllı řehir uygulamaları ve kompakt kent planlaması gibi stratejilerin kritik rol oynayacaęı ortaya çıkmıřtır.

Sonuç olarak, bu alıřma enerji yoęunluęunu etkileyen faktrlerin gelir grupları arasında farklılık gsterdięini, ancak ortak hedefin enerji verimlilięinin artırılması ve dřk karbonlu kalkınma yoluna geiř olduęunu ortaya koymuřtur. Beta ridge regresyon ve makine ęrenmesi yntemlerinin birlikte kullanımı, enerji yoęunluęunun daha saęlam ve gvenilir biimde analiz edilmesine imkn tanımıř; elde edilen bulgular politika yapıcılar iin enerji verimlilięi politikalarının hem gelir gruplarına zg hem de kresel hedeflerle uyumlu biimde řekillendirilmesi gerektięini aıka ortaya koymuřtur. Bu genel ereve, enerji yoęunluęunun azaltılmasının yalnızca ekonomik byme aısından deęil, aynı zamanda evresel srdrlebilirlik ve toplumsal refahın artırılması bakımından da stratejik bir neme sahip olduęunu gstermekte; dolayısıyla fosil yakıt baęımlılıęının azaltılması, yenilenebilir enerji yatırımlarının glendirilmesi ve srdrlebilir řehirleřme politikalarının hayata geirilmesi gereklilięini vurgulamaktadır.

KAYNAKLAR

- Alpaydın, E. 2014. Introduction to Machine Learning. The MIT Press, 613, London.
- Anonim. 2024. Web Sitesi: <https://enerji.gov.tr/bilgi-merkezi-enerji-verimlilik-olcme-degerlendirme>, Erişim Tarihi: 02.05.2025.
- Apergis, N., ve Payne, J. E. 2010. Renewable energy consumption and economic growth: Evidence from a panel of OECD countries. *Energy Policy*, 38(1), 656–660.
- Balitsky, S., Bilan, Y., Strielkowski, W. ve Štreimikienė, D. 2016. Energy efficiency and natural gas consumption in the context of economic development in the European Union. *Renewable and Sustainable Energy Reviews*, 55, 156–168.
- Biau, G. ve Devroye, L. 2008. On the layered nearest neighbour estimate, the bagged nearest neighbour estimate and the random forest method in regression and classification. *Journal of Multivariate Analysis*, 101, 2499–2518.
- Bildirici, M.E. ve Bakırtaş, T. 2014. The relationship among oil, natural gas and coal consumption and economic growth in BRICS countries. *Energy*, 65, 134–144.
- Bilgili, F., Koçak, E., & Bulut, Ü. 2016. The dynamic impact of renewable energy consumption on CO₂ emissions: A revisited Environmental Kuznets Curve approach. *Renewable and Sustainable Energy Reviews*, 54, 838–845.
- Bilgili, F., Koçak, E., Bulut, Ü. ve Kuloğlu, A. 2017. The impact of urbanization on energy intensity: Panel data evidence considering cross-sectional dependence and heterogeneity. *Energy*, 133, 242–256.
- Breiman, L. 2001. Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A. ve Stone, C. J. 1984. *Classification and Regression Trees*. Wadsworth.
- Botchkarev, A. 2018. Web Sitesi: <https://ssrn.com/abstract=3177507>. Erişim Tarihi: 24.04.2025.
- Chang, Y., Choi, Y., Kim, C. S., Miller, J. I. ve Park, J. Y. 2024. Common trends and country-specific heterogeneities in long-run world energy consumption. *Unpublished manuscript*.
- Chu, L.K. ve Le, N.T.M. 2022. Environmental quality and the role of economic policy uncertainty, economic complexity, renewable energy, and energy intensity: The case of G7 countries. *Environmental Science and Pollution Research*, 29, 2866–2882.

- Cribari-Neto F. ve Zeileis A. 2010. Beta Regression in R. *Journal of Statistical Software*, 34(2), 1-24.
- Csereklyei, Z., Rubio-Varas, M. M. ve Stern, D. I. 2016. Energy and economic growth: The stylized facts. *The Energy Journal*, 37(2), 223–255.
- Çapık, M., Yılmaz, A.O. ve Çavuşoğlu, İ. 2012. Present situation and potential role of renewable energy in Turkey. *Renewable Energy*, 46, 1-13.
- Diaz, A., Marrero, G.A., Puch, L.A. ve Rodríguez, J. 2019. Economic growth, energy intensity and the energy mix. *Energy Economics*, 81, 1056–1077.
- Doğan, H. ve Yılankırkan, N. 2015. Türkiye'nin enerji verimliliği potansiyeli ve projeksiyonu. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 3(1), 375–383.
- Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. J. ve Vapnik, V. 1997. Support vector regression machines. In *Advances in Neural Information Processing Systems*, 155–161.
- Ediger, V. Ş. 2009. Doğalgaz Piyasası 2011 Yılı Sektör Raporu. Ankara.
- Espinheira, P. L., Ferrari, S. L. P. ve Cribari-Neto, F. 2008. On beta regression residuals. *Journal of Applied Statistics*, 35(4), 407–419.
- Ernawati, E., Syarif, M., Suriadi, L.O., Rosnawintang, R. ve Madi, R.A. 2024. Does energy intensity correlate with economic growth and government governance? *IOP Conference Series: Earth and Environmental Science*, 1324(1), 012098.
- Eyyüboğlu, K., Akdağ, S., Yıldırım, H. ve Alola, A. A. 2022. The causal trend of energy intensity and urbanization in emerging countries. *Letters in Spatial and Resource Sciences*, 15, 653–663.
- Ferrari, S. ve Cribari-Neto, F. 2004. Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*, 31(7), 799-815.
- Fox J. ve Weisberg S. 2019. *An R Companion to Applied Regression*. Sage Publications, 160, Thousand Oaks CA.
- Gao, J., Hassan, M.S., Kalim, R., Sharif, A., Alkhateeb, T.T.Y. ve Mahmood, H. 2023. The role of clean and unclean energy resources in inspecting N-shaped impact of industrial production on environmental quality: A case of high polluting economies. *Resources Policy*, 80, 103217.
- Gedikli, P. 2024. Beta regresyon modelinde tahmin yöntemlerinin karşılaştırılması Yüksek lisans tezi, Necmettin Erbakan Üniversitesi, Fen Bilimleri Enstitüsü, Matematik Anabilim Dalı, 79, Konya.

- Gunn, S. R.1998.Support vector machines for Classification and regression. ISIS technical report, 14(1), 5-16.
- Gyamfi, B.A., Kwakwa, P.A. ve Adebayo, T.S. 2023. Energy intensity among European Union countries: The role of renewable energy, income and trade. International Journal of Energy Sector Management, 17(4), 801–819.
- Hamner B., Frasco M. 2018. Web Sitesi: <https://CRAN.R-project.org/package=Metrics>. Erişim Tarihi: 09.09.2025
- Hoerl, A. E. ve Kennard, R. W. 1970a. *Ridge regression: Biased estimation for nonorthogonal problems*. Technometrics, 12(1), 55–67.
- Islam, S. 2020. Data Science: The Impact of Machine Learning. International Journal of Scientific & Engineering Research Volume, 11(5), 1046-1049.
- Karabat, S. ve Aydın, B. 2018. İyi Tarım Uygulamalarının Mandarin Üretiminde Enerji Kullanım Etkinliği ve Ekonomik Analiz Üzerine Etkisi: İzmir İli Örneği. Toprak Su Dergisi, 7 (1), 1-10.
- Kartal, M.T. 2022. The role of consumption of energy, fossil sources, nuclear energy, and renewable energy on environmental degradation in top-five carbon producing countries. Renewable Energy, 184, 871–880.
- Li, Y., Zou, C., Berecibar, M., Nanini-Maury, E., Jonathan, C., Chan, W., Bossche, P., Mierlo, J. ve Omar, N. 2018. Random forest regression for online capacity estimation of lithium-ion batteries. Applied Energy, 232, 197-210.
- Mirziyoyeva, Z. ve Salahodjaev, R. 2022. Renewable energy and CO2 emissions intensity in the top carbon intense countries. Renewable Energy, 192, 507–512.
- Mohri, M., Rostamizadeh, A. ve Talwalker, A. 2018. Foundations of Machine Learning. The MIT Press, 475, London.
- Naimoğlu, M. ve Özel, B. 2022. Enerji Kaynaklarının Enerji Yoğunluğu Üzerindeki Etkileri: Enerji İthalatçısı Yükselen Ekonomilerden Kanıtlar. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 47, 1–15.
- Naimoğlu, M., Sahabi, A.M. ve Özbek, S. 2022. Macaristan ekonomisinde enerji verimliliğini etkileyen faktörlerin Fourier ADL eşbütünleşme yaklaşımıyla belirlenmesi. Sosyoekonomi, 30(53), 487–507.
- Osorio, C. J., Steinberger, J. K. ve Roberts, J. T. 2017. Urban form and energy consumption: A review of the evidence and policy implications. Urban Studies, 54(12), 2658–2679.
- Özbek, S. 2023. The Determinants of Energy Intensity in Energy Importing Developing Countries. Journal of Economics and Administrative Sciences, 24(1), 114–126.

- Öztemel, E. 2006. Yapay Sinir Ağları. Papatya Yayıncılık, 118, İstanbul.
- Pamir, N. 2005. Enerji politikaları ve küresel gelişmeler. ASAM Yayınları. Web Sitesi: <http://www.turksam.org>, Erişim Tarihi: 08.05.2025.
- Paoli, L. D. 2014. The EU climate and energy policy: what about the new targets for 2030?. *Economics And Policy Of Energy And The Environment*, 2014(1), 27-51.
- Patterson, M.G. 1996. What is energy efficiency? Concepts, indicators and methodological issues”, *Energy Policy*, 24(5), 377-390.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. ve Duchesnay, É. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Poumanyong, P. ve Kaneko, S. 2010. Does urbanization lead to less energy use and lower CO₂ emissions? A cross-country analysis. *Ecological Economics*, 70(2), 434–444.
- Saad, D. ve Rykov, Y. P. 2022. Principled Machine Learning. *Ieee Journal Of Selected Topics In Quantum Electronics*, 28 (4).
- Sadorsky, P. 2009. Renewable energy consumption, CO₂ emissions and oil prices in the G7 countries. *Energy Economics*, 31(3), 456–462.
- Sadorsky, P. 2013. Do urbanization and industrialization affect energy intensity in developing countries? *Energy Economics*, 37, 52–59.
- Sarkodie, S.A. ve Ozturk, I. 2020. Investigating the Environmental Kuznets Curve hypothesis in Kenya: A multivariate analysis. *Renewable and Sustainable Energy Reviews*, 117, 109481.
- Schwarz, G. 1978. Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.
- Scornet, E., Biau, G. ve Vert, J. P. G. 2015. Consistency of Random Forests. *The Annals of Statistics*, 43(4), 1716-1741.
- Shokry, A., Audino, F., Vicente, P., Escudero, G., Moya, M. P., Graells, M. ve Espuña, A. 2015. Modeling and simulation of complex nonlinear dynamic processes using data-based models: Application to photo-Fenton process. In *Computer Aided Chemical Engineering*, 37, 191-196

- Selvi, E. 2019. Kentsel Atıksu Arıtma Tesisi Biyolojik Oksijen İhtiyacının (Boi5) Makina Öğrenmesi Yöntemleri ile Tahmin Edilmesi. Yüksek Lisans Tezi. Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya.
- Smola, A. J. ve Schölkopf, B. 2004. A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222.
- Solarin, S.A. 2020. Towards sustainable development in developing countries: Aggregate and disaggregate analysis of energy intensity and the role of fossil fuel subsidies. *Sustainable Production and Consumption*, 24, 254–265.
- Şahin, T. 2022. Dizel Makinanın Makina Öğrenmesi Yöntemi Kullanılarak Modellenmesi ve Karar-Destek Mekanizması Oluşturulması. İstanbul Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, İstanbul.
- Ünüvar, İ. ve Keskinç, S. 2020. Yenilenebilir Enerji ve Ekonomik Büyüme İlişkisi: G20 Ülkeleri Örneği (2000–2016). *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 16(2), 251–266
- Yellareddygar, S. KR., Pasche J.S., Taylor R.J., Hua S. ve Gudmestad N.C. 2016. Beta regression model for predicting the development of pink rot in potato tubers during storage. *The American Phytopathological Society*, 100(6), 1118-1124.
- Vapnik, V. 1992. Principles of risk minimization for learning theory. Paper presented at the Advances in Neural Information Processing Systems.x
- Zhao, W., Zhang, R., Lv, Y. ve Liu, J. 2014. Variable Selection for Varying Dispersion Beta Regression Model. *Journal of Applied Statistics*. 41(1):95-108