

Destek Vektör Makineleri ve Türkiye'deki Enerji Santrallerinde Doğal Gaz Tüketimi Üzerine Bir Uygulama

Gizem MERAL¹, Sinan SARAÇLI^{*2}

^{1,2}Afyon Kocatepe Üniversitesi, Fen-Edebiyat Fakültesi, İstatistik Bölümü, 03200, Afyonkarahisar, Türkiye

(Alınış / Received: 03.02.2020, Kabul / Accepted: 11.05.2020, Online Yayınlanma / Published Online: 20.08.2020)

Anahtar Kelimeler

Destek vektör makineleri,
Destek vektör regresyon,
Çekirdek fonksiyonları

Özet: Bu çalışmanın amacı, destek vektör makinelerinde regresyon yöntemini kullanarak Türkiye'deki enerji santrallerinin doğalgaz tüketimi üzerine ön kestirim yapmaktır. Bu amaçla, veri seti 2013-2018 yılları arasında Türkiye Enerji Piyasası Düzenleme Kurumu ve Enerji İşleri Genel Müdürlüğünden elde edilmiştir. Bu çalışmada ilk olarak, Türkiye'de doğalgazın enerji piyasasındaki yeri, birincil enerji kaynakları içindeki payı, üretim, tüketim, ithalat ve ihracat değerleri incelenmiştir. Bu değerlerin ölçümündeki farklılıklardan dolayı, ilgili veri seti istatistiksel analizden önce standartlaştırılmıştır. Enerji santralleri tüketimi bağımlı değişken iken; sanayi tüketimi, şehir tüketimi, üretim, ithalat ve ihracat değerleri bağımsız değişken olarak belirlenmiştir. Destek Vektör Regresyonda kullanılan tüm kernel fonksiyonları (Doğrusal, Polinomial, Radyal Tabanlı Fonksiyon (RTF) ve Sigmoid) test edilmiştir. En küçük Hata Kareler Ortalaması (HKO)'na sahip olan RTF tahmin kernel fonksiyonu olarak seçilmiştir. Daha sonra, destek vektörler, ağırlıklar ve karar sabiti belirlenmiştir. Ağırlıklar ve destek vektörler çarpılıp sapma eklenerek, son model elde edilmiştir. Son model yardımıyla da Mayıs-Aralık 2018 için Türkiye'deki enerji santrallerinin doğalgaz tüketimlerine ait önkestirimler yapılmıştır. Sonuçlar ilgili tablolarda gösterilmiştir.

Support Vector Machines and an Application on Natural Gas Consumptions of Power Plants in Turkey

Keywords

Support vector machines,
Support vector regression,
Kernel functions

Abstract: The aim of this study is to forecast the natural gas consumptions of power plants in Turkey via support vector machines regression method. With this aim the data set is obtained from Turkey's Energy Market Regulatory Authority and Energy Affairs General Directorate between the years 2013-2018. In this study, first of all, the place in Turkish market, ratio within the primary energy sources, production, consumption, import and export values of natural gas, as a power supply is examined. Because of the differences in measurements of these values, the related data set is standardized before the statistical analysis. While, the consumption in energy plants is considered as a dependent variable, industrial consumption, city consumption, production, import and export values are considered as independent variables. All types of Kernel functions (Linear, Polynomial, Radial Basis Function (RBF) and Sigmoid) in Support Vector Regression are tested. RBF is chosen as the forecasting Kernel function because of having the minimum Mean Square Error (MSE). Then, support vectors, weights and decision constants are determined. By multiplying weights with support vectors and adding the bias, the final model is obtained. By the help of final model, forecasts of natural gas consumption of power plants in Turkey, for May-December 2018 are obtained. The results are given in related tables and figures.

1. Giriş

İlk olarak ısınma amaçlı ülkemizde kullanımına başlanılan doğal gaz, bugün başta elektrik üretimi

olmak üzere konut, sanayi ve diğer kullanımlarda etkinliğini artırmıştır [1]. Ancak Türkiye'deki doğal gaz piyasası ithalata dayalıdır. Doğal gaz ithalatında kaynak çeşitliliğinin ve arz güvenliği ile sürekliliğinin

*İlgili yazar: ssaracli@aku.edu.tr

sağlanması ülkemiz için çok önemlidir [2]. Özellikle kış döneminde ortaya çıkan doğal gaz arz-talep dengesizliklerini gidermek için tüketim miktarlarının doğru tahmin edilmesi gereklidir [3]. Doğal gaz tüketim tahminlerinde genellikle klasik zaman serileri kullanılmaktadır. Geçmiş dönemlere ait değerler yardımıyla ön kestirim yapmayı amaçlayan zaman serileri; veriler arasında doğrusal bir ilişki olduğu varsayımı ile durağan ya da çeşitli yöntemlerle durağan hale getirilmiş serilere uygulanmaktadır. Veriler arasında doğrusal olmayan ilişki söz konusu olduğu zaman da yapay sinir ağları tercih edilir ve tahminlerde kullanılmaktadır [4].

Bu çalışma da ise zaman serisi analizi ve yapay sinir ağları ile gelecek dönem tahmini dışında Destek Vektör Makineleri (DVM) kullanılarak Türkiye'deki enerji santrallerinin doğal gaz tüketimi üzerine ön kestirim yapmak amaçlanmıştır.

Bu amaçla öncelikle makine öğrenmesi konusuna kısaca değinilmiş ve DVM'nin alt yapısını oluşturan istatistiksel öğrenme ile aralarındaki farklara yer verilmiştir. Daha sonra ise DVM'nin hangi problemlerde nasıl kullanıldığı araştırılmış ve problemlerin çözümünde kullanılan çekirdek fonksiyonları kavramı üzerinde durulmuştur.

1.1. Makine Öğrenmesi

Zeka, insanın düşünme, akıl yürütme, objektif gerçekleri algılama, yargılama ve sonuç çıkarma yeteneklerinin tamamı olarak tanımlanmaktadır [5]. Yapay zeka ise insana özgü bu özelliklerin analiz edilip makinelerle kazandırılmasıdır [6]. Makine öğrenmesi algoritmaları kendilerine verilen örnek olayları inceleyerek bu olayların meydana geliş biçimlerinden eğitilir ve bu olaylar üzerinden genelleme yapma yeteneği kazanır. Bu genelleme yeteneğine ağırlık öğrenmesi denir [7].

1.2. İstatistiksel Öğrenme Teorisi

İstatistiksel Öğrenme Teorisi günümüzün birçok makine öğrenmesi algoritmalarına teorik alt yapı oluşturur ve yapay zeka konusunun en iyi şekilde geliştirilmiş dallarından biridir. 1960'larda Rusya'da geliştirilmeye başlayan teorem DVM'nin tanınmasıyla birlikte büyük bir popüleriteye kavuşmuştur. Teorem yeni geliştirilen öğrenme algoritmalarına temel teşkil etmekle birlikte deneysel veriden gerekli bilgileri çekmek ve geçerli sonuçları elde etmek için de kullanılmaktadır [8].

Makine öğrenmesinde öğrenme işi istatistik ile yapılır çünkü amaç örnek veri üzerinden genellenebilir yapılar ortaya çıkarmaktır. Daha sonra ise öğrenme yaklaşımlarının optimizasyonu ve performansı ile ilgilenilir [9]. Örneğin, bir kedi resmine kedi demek için açıklanabilir bir modele ihtiyaç yoktur, bu makine öğrenmesidir. Fakat; bir

müşterinin sizi terk etmesi örneğinde neden-sonuç bağlamında model kurmanız gerekir. Bu ise istatistiksel öğrenmedir [10].

1.3. Makine Öğrenmesi ve İstatistiksel Öğrenme Arasındaki Farklar

İstatistiksel öğrenme;

- İstatistiğin bir alt alanıdır,
- Model kurmak için kullanılan bağımsız değişkenlerde meydana gelen değişimlerin bağımlı değişken üzerindeki etkisi yorumlanabilir,
- Bu öğrenme modellerinde eksik gözlem, aykırı gözlem gibi durumların çok fazla önemi vardır. Veri kalitesi ve veri ön işleme oldukça önemlidir [10].

Matematiksel öğrenme ise;

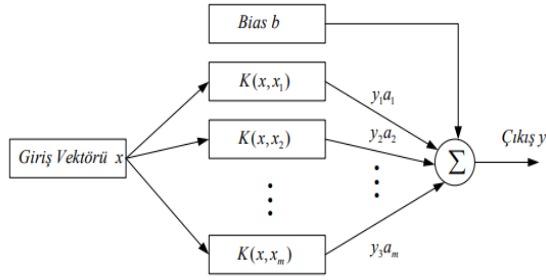
- Yapay zekanın bir alt alanıdır,
- Ağırlıklı olarak büyük ölçekli uygulamalarda kullanılır ve tahmin yakınlığı ile ilgilenir.
- Sonuç odaklıdır.
- Aykırı gözlem, eksik gözlem gibi durumlardan daha az etkilenir [10].

1.4. VapnikChervonenkis (VC) Boyutu

VC-Boyut, bir sınıflandırıcının karmaşıklığının ölçüsüdür. Yani, bu ölçüm bir ikili sınıflandırıcının iki sınıf arasındaki sınırı ne kadar iyi ölçmesi olarak bilinir. Eğer VC-Boyut çok büyük ise sınıflandırıcı daha karmaşık bir hâl alır ve böylece ilgili sınıflara ait olan gözlemleri daha iyi bölebilir. N nokta içeren bir veri kümesi olsun. Bu N örnek 2^N farklı biçimde artı ve eksi olarak etiketlenebilir. Yani, N örnekle 2^N tane farklı öğrenme problemi tanımlanabilir. Bu problemlerin herhangi biri için pozitif örnekleri negatif örneklerden ayıran $h \in H$ olacak şekilde bir hipotez tanımlanabilir. O zaman H, N adet noktaları parçalara ayırır denir. Yani, N adet örnek yolu ile tanımlanabilen herhangi bir öğrenme problemi H'den çekilen bir hipotez yolu ile hatasız öğrenilebilir. H tarafından parçalara ayrılabilen maksimum noktaların sayısı, H'nin VC-Boyutu olarak adlandırılır ve VC(H) olarak gösterilir [9].

1.5. Destek Vektör Makineleri (DVM)

DVM, sınıflandırma ve regresyon analizinde kullanılan, verileri analiz eden ve örneklerden öğrenen denetimli bir öğrenme yöntemidir. DVM, ilk kez Vapnik vd. tarafından 1992 yılında Computational Learning Theory (COLT) konferansında önerilmiş olup günümüzde kullanılan şeklini Cortes ve Vapnik tarafından 1995 yılında almıştır. 1997 yılında ise DVM algoritması yine Vapnik vd. tarafından regresyon uygulamalarını da kapsayacak şekilde genişletilmiş ve bu tarihten sonra son on yılda gelişen bilgisayar teknolojisi sayesinde artarak kullanılmaya başlanmıştır [11]. Şekil 1'de DVM'nin ağ yapısı gösterilmiştir.

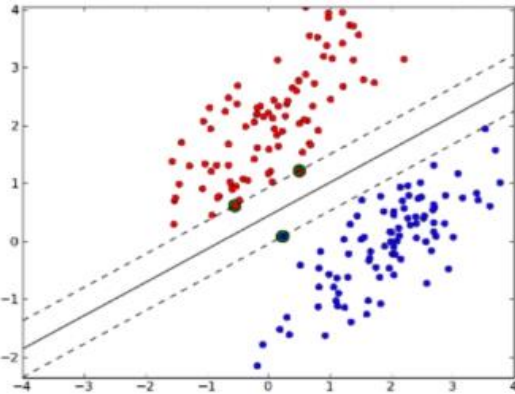


Şekil 1. DVM'nin Ağ Yapısı [12].

1.5.1. Destek Vektör Sınıflandırma

DVM'de sınıflandırma problemlerinin birkaç türü vardır. Bunlardan en temel olanı eğitim hatasının bulunmadığı doğrusal ayrılabilirlik durumudur. Bir başka durum ise, eğitim hatasının bulunmadığı doğrusal olmayan ayrılabilirlik durumudur [13].

Eğitim Hatasının Bulunmadığı Doğrusal Ayrılabilirlik Durumunda Sınıflandırma: Doğrusal olarak ayrılabilen iki sınıflı öğrenme için, destek vektör sınıflandırıcının hedefi verilen örneklerin iki sınıfın maksimal bir marj ile bölebileceği bir aşırı düzlem tanımlamaktır [14]. Doğrusal olarak ayrılabilen bir durumda aşırı düzlem Şekil 2'de gösterilmiştir. Sınırdaki kalan noktalar destek vektörleri olarak tanımlanabilir.



Şekil 2. Doğrusal Olarak Ayrılabilen Bir Durumda Aşırı Düzlem [15].

Eğitim Hatasının Bulunmadığı Doğrusal Olmayan Ayrılabilirlik Durumu: DVM, bir çekirdek fonksiyonu yardımıyla doğrusal olmayan dönüşümler yapabilmekte ve bu şekilde verilerin yüksek boyutta doğrusal olarak ayrılmasına imkan sağlamaktadır. Sonuç olarak, çekirdek fonksiyonu kullanarak doğrusal olarak ayrılamayan iki sınıflı bir problemin çözümü ile ilgili karar kuralı Eşitlik 1'de gösterilmiştir [16].

$$f(x) = \text{sign}\left(\sum_i a_i y_i \varphi(x) \cdot \varphi(x_j) + b\right) \quad (1)$$

Burada dikkat edilmesi gereken nokta, orijinal girdi uzayını daha yüksek boyutlu uzayına haritalama yaparken, yeni uzayın boyutu çok fazla olduğundan hesaplamaları (çarpımları) yapmanın zorlaşacağıdır.

Bu yüzden işlemi daha da basit hale getirebilmek için çekirdek fonksiyonları kullanılır [17].

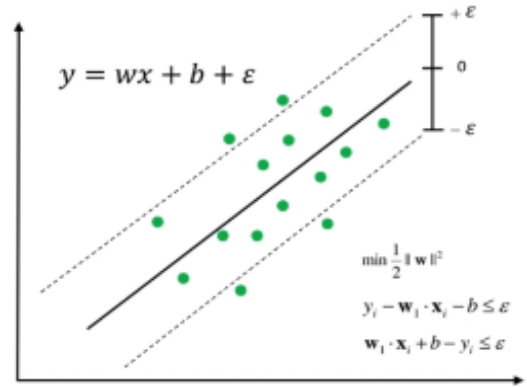
DVM'de kullanılan bazı çekirdek fonksiyonları kısaca Tablo 1'de gösterildiği gibidir [18].

Tablo 1. DVM'de Kullanılan Çekirdek Fonksiyonları

Çekirdek Tipleri	Fonksiyon
Doğrusal Çekirdek	$K(x_i, x_j) = x_i \cdot x_j$
Polinomial Çekirdek	$K(x_i, x_j) = (y(x_i \cdot x_j) + c)^d$
Sigmoid Çekirdek	$K(x_i, x_j) = \tanh(y(x_i \cdot x_j) + c)$
Radyal Tabanlı Çekirdek	$K(x_i, x_j) = \exp(-\gamma x_i - x_j ^2)$

1.5.2. Destek Vektör Regresyon (DVR)

Burada amaç, bir marjın aralığına maksimum noktayı en küçük hata ile alabilecek bir doğru (fonksiyon) belirlemektir. Şekil 3'te DVR'ningrafiği gösterilmiştir [15].



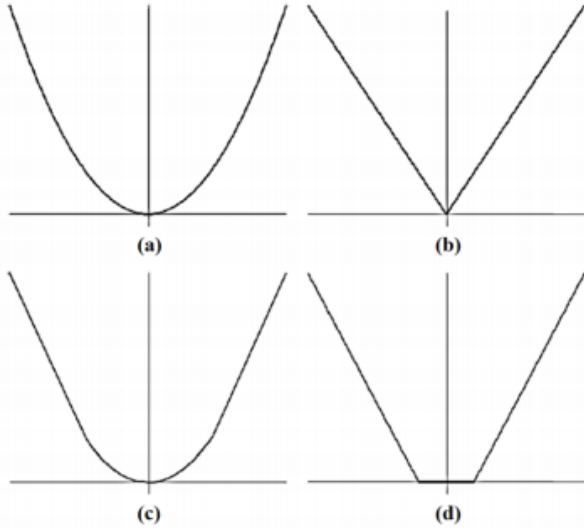
Şekil 3. DVR [15].

DVR'nin Özellikleri;

- Destek Vektör Regresyon robusttur,
- Aykırı gözlemlerin regresyon denklemi üzerindeki etkilerini minimize eder,
- Eşik değeri ile gözlemlerin regresyon eğrisine olan katkıları ayarlanabilmektedir,
- Doğrusal ve doğrusal olmayan formları vardır [15].

DVR'nin ana prensibi, birkaç küçük farkla sınıflandırma ile aynıdır. Temel olarak DVM, alternatif bir kayıp fonksiyonunun içe aktarılmasıyla regresyon problemlerine uygulanabilir. Şekil 4'te farklı olası kayıp fonksiyonları görülmektedir [19].

Şekil 4'te (a) en küçük hata kriterleri ile ilgili olan ikinci dereceden kayıp fonksiyonudur, (b) ikinci dereceden kayıp fonksiyonundan ziyade sapmalara daha az duyarlı olan Laplacian tip kayıp fonksiyonu, (c) Veri dağılımı bilinmediğinde en uygun özellikleri içeren kayıp fonksiyonu, (d) ise regresyon problemleri için bağımsız değişkende meydana gelen kötü performansla sahip olan aykırı verilerin var olması ile birlikte tasarlanan model içinde ufak farklılıklara karşın ε duyarsızlık kayıp fonksiyonu tanımlanmıştır [19, 21].



Şekil 4. DVR için Kayıp Fonksiyonları [20]. (a) Quadratic, (b) Laplace, (c) Huber, (d) Duyarsızlık.

1.6. Model Belirleme Kriterleri

Bir modele eklenen her açıklayıcı değişken uyumu iyileştirmez. İstenmeyen değişkenler eklemek, parametre tahminlerinin kesinliğini azaltır. Yaygın olarak kullanılan model belirleme kriterleri şöyle sıralanabilir [22]:

R^2 Kriteri, Düzeltilmiş R^2 , Akaike Bilgi Kriteri (AIC), Bayesçi Bilgi Kriteri (BIC), Hannan-Quinn Kriteri'dir.

Çalışmada model belirleme kriteri olarak AIC ve BIC kullanılmıştır. Her iki bilgi kriteri de bir modelin uyum iyiliğini, bir ceza terimiyle dengeler [23]. Eşitlik 2 ve Eşitlik 3 sırasıyla AIC ve BIC'nin nasıl hesaplandığını göstermektedir.

$$AIC = -2\log L(\hat{\theta}) + 2k \quad (2)$$

Burada:

$\log L(\hat{\theta}) = \text{model tarafından elde edilebilecek}$
 θ 'nın en yüksek olasılık tahmini
 $k = \text{modeldeki tahmin parametrelerinin sayısı}$

$$BIC = -2\log L(\hat{\theta}) + k \ln(N) \quad (3)$$

Burada:

$N = \text{uyumda kullanılan veri noktalarının sayısı}$
[24].

Bu iki kriterin uygulamaları genellikle ulaşılan sonuçlarda geniş bir uzlaşma göstermiştir, ancak modellerin detaylı sıralamasında ara sıra farklar ortaya çıkmaktadır [25].

1.7. Performans Fonksiyonları

Kurulan modellere ait performansların değerlendirilmesinde kullanılan literatürde birçok fonksiyon bulunmaktadır. Bu fonksiyonlar Tablo 2.'de verilmiştir.

Tablo 2. Performans Fonksiyonları [26]

OMH	$\frac{1}{N} \sum_{j=1}^n e_j $
HKO	$\frac{1}{N} \sum_{k=1}^N (y_{gerçek} - y_{tahmin})^2$
OHKK	$\sqrt{\frac{1}{N} \sum_{k=1}^N (y_{gerçek} - y_{tahmin})^2}$
HKT	$\sum_{k=1}^N (y_{gerçek} - y_{tahmin})^2$

OMH: Ortalama Mutlak Hata

HKO: Hata Kareler Ortalaması

OHKK: Ortalama Hata Karelerin Karekökü

HKT: Hata Kareler Toplamı

2. Materyal ve Metot

Bu çalışmada Türkiye'nin doğal gaz piyasasındaki yeri, sanayi tüketimi, şehir tüketimi, üretim, ithalat ve ihracat değerleri incelenmiş, Türkiye'deki enerji santrallerinin doğal gaz tüketimi üzerine ön kestirim yapmak amaçlanmıştır. Bu amaçla ilgili veri seti Enerji İşleri Genel Müdürlüğü'nün ve Türkiye Enerji Piyasası Düzenleme Kurumunun doğal gaz piyasası sektör raporlarından derlenerek elde edilmiştir ve veriler 2013 Ocak - 2018 Nisan arasındaki aylık değerleri kapsamaktadır.

Son yıllarda literatürde gelecek dönem doğal gaz tüketim tahmini için zaman serileri ve yapay sinir ağları oldukça kullanılmıştır. Yine bu çalışmalarda girdi olarak birçok farklı bağımsız değişken kullanılmıştır. Doğal gazın enerji santralleri tüketiminde en çok kullanılan değişkenlerin arasında sanayi tüketimi, şehir tüketimi, depolama, üretim, ithalat ve ihracat verileri gösterilebilir. Bu değişkenlerden bazıları, elde edilebilecek tahmin üzerinde büyük etkiye sahipken bazılarının daha zayıf etkisi olacaktır. Bu yüzden öncelikle analizde hangi bağımsız değişkenlerin kullanılacağı seçilerek destek vektör makineleri yöntemiyle tahminler elde edilmiştir. Daha sonra ise veriler üzerinde trend analizi yapılarak ortaya çıkan tahmin değerleri ile destek vektör makineleri ile elde edilen tahmin değerleri karşılaştırılmıştır.

Çalışmada kullanılacak olan çekirdek fonksiyonunun, destek vektörlerin ve ağırlıkların belirlenmesi işlemleri için Statistica programı, model katsayılarının hesaplanması için ise MATLAB programı kullanılmıştır. Ayrıca veriler üzerinde trend analizi uygulanabilmesi için ise Minitab programı kullanılmıştır.

3. Bulgular

Çalışmada kullanılacak bağımsız değişken seçimi için 4 farklı model kurulmuştur. Bu modeller Tablo 3'te gösterilmiştir. Tabloda bulunan kısaltmalar ST:

Sanayi Tüketimi, ŞT: Şehir Tüketimi, Ü: Üretim, İT: İthalat ve İH: İhracat anlamlarına gelmektedir. Bu kurulan modellere ilişkin AIC ve BIC değerleri ise Tablo 4'te yer almaktadır.

Tablo 3. Kurulan Modeller

Modeller	X ₁	X ₂	X ₃	X ₄	X ₅	SD
Model1	ST	ŞT	Ü	İT	İH	5
Model2	ŞT	Ü	İT	İH		4
Model3	ŞT	Ü	İH			3
Model4	ŞT	İT				2

Tablo 4. Modellere İlişkin AIC ve BIC Değerleri

Modeller	AIC	BIC
Model1	99,869	112,823
Model2	167,504	180,953
Model3	169,660	186,614
Model4	167,684	186,638

Çalışmada model belirlendikten sonra veriler %75 eğitim ve %25 test verisi olmak üzere ikiye ayrılmıştır.

Çalışmada kullanılan bağımsız değişkenler; Sanayi tüketimi, Şehir tüketimi, Üretim, İthalat ve İhracat olarak belirlendikten sonra DVM'de kullanılacak olan çekirdek (kernel) fonksiyonunun seçimi için Tablo 1'de bahsedilen tüm çekirdek fonksiyonları denenmiş ve performanslarına göre en iyi çekirdek fonksiyonu en düşük HKO değerine sahip olan RTF çekirdek fonksiyonu olarak belirlenmiştir. Tüm çekirdek fonksiyonlarına ait performans fonksiyonları Tablo 5'te gösterilmiştir.

Tablo 5. Çekirdek Fonksiyonlarına Ait Performans Fonksiyonları

Fonksiyon Çeşiti	Performans Fonks.	Eğitim	Test
Doğrusal	MAE	0,374	0,522
	HKO	0,254	0,431
	HKT	16,256	27,584
Polinomial	MAE	0,533	0,433
	HKO	0,461	0,290
	HKT	29,504	18,560
RBF	MAE	0,333	0,440
	MSE	0,249	0,276
	SSE	15,936	17,664
Sigmoid	MAE	0,412	0,454
	MSE	0,353	0,328
	SSE	22,592	20,992

Çalışmanın devamında MATLAB programına geçilerek RTF çekirdek fonksiyonu kullanılarak oluşturulan modelin kat sayıları hesaplanmıştır. Burada da Şekil 1'de gösterilen DVM'nin yapısı gereği öncelikle belirlenen destek vektörlerle ağırlıklar çarpılmış, üzerine yan(bias) eklenmiş ve Eşitlik 4'teki

model ortaya çıkmıştır. C parametresi 8 ve bias terimi 0,1 iken hesaplanan model kat sayıları;

$$\hat{Y} = -0,349 - 0,072 * X_1 - 4,449 * X_2 + 0,459 * X_3 + 4,890 * X_4 + 0,895 * X_5 \quad (4)$$

Burada -0,349 sabit kat sayı ve

X1: Sanayi Tüketimi

X2: Şehir Tüketimi

X3: Üretim

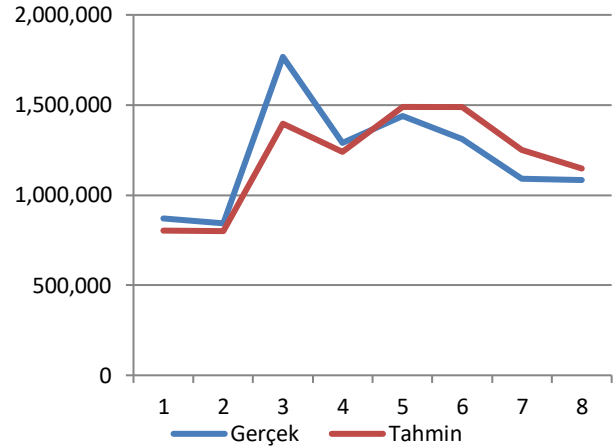
X4: İthalat

X5: İhracat

bağımsız değişkenlerdir. Burada 2018 yılının son sekiz aylık dönemi için tahminler yapılmış ve gerçek değerler ile karşılaştırılmıştır çıkan sonuçlar Tablo 6'da, sonuçlara ait grafik ise Şekil 5'te gösterilmiştir.

Tablo 6. 2018'in Son Sekiz Aylık Dönemi İçin DVM ile Yapılan Tahminler ve Gerçek Değer Karşılaştırması

Aylar	Tahmin	Gerçek
Mayıs	873 048	869 451
Haziran	822 499	843 344
Temmuz	1 881 988	1 767 067
Ağustos	1 306 750	1 289 299
Eylül	1 401 496	1 438 252
Ekim	1 316 965	1 311 603
Kasım	1 155 373	1 089 548
Aralık	1 070 842	1 082 018



Şekil 5. DVM ile Yapılan Tahminlerin Gerçek Değerleri Arasındaki Farkı Gösteren Grafik.

Ayrıca veriler üzerinde trend analizi yapmak için doğrusal, üstel ve ikinci dereceden denklemler denenmiş ve performanslarına göre en iyi denklem en düşük Ortalama Kareli Sapma (OKS) değerine sahip olan ikinci dereceden denklem olarak seçilmiştir. Tüm denklemlere ait OKS değerleri Tablo 7'de gösterilmiştir.

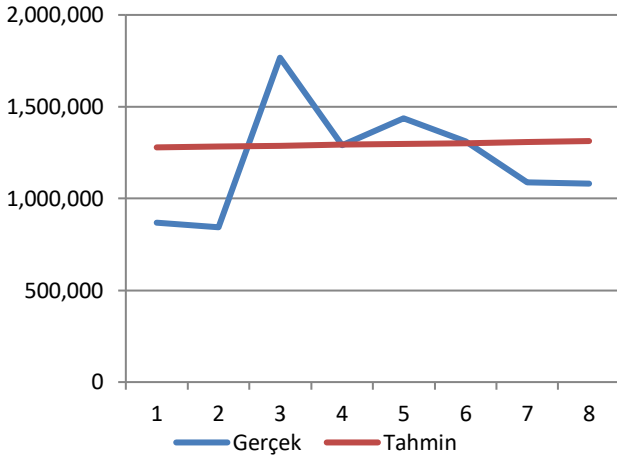
Tablo 7. Denklemlere Ait MSD değerleri

Denklemler	OKS
Doğrusal	0,410
Üstel	0,413
İkinci Dereceden	0,404

Daha sonra ise ikinci dereceden denklemle ait tahmin değerleri ve gerçek değerler ile karşılaştırılması Tablo 8'de gösterilmiştir.

Tablo 8. Trend Analizi Elde Edilen Tahmin Değerleri ve Gerçek Değerlerin Karşılaştırılması

Aylar	Gerçek	Tahmin
Mayıs	869 451	1 278 501
Haziran	843 344	1 282 944
Temmuz	1 767 067	1 287 547
Ağustos	1 289 299	1 292 308
Eylül	1 438 252	1 297 227
Ekim	1 311 603	1 302 306
Kasım	1 089 548	1 307 543
Aralık	1 082 018	1 312 940



Şekil 6. Trend Analizi ile Yapılan Tahminlerin Gerçek Değerleri Arasındaki Farkı Gösteren Grafik.

Şekil 5'e göre DVM ile yapılan tahminlerin gerçek değerlerle uyumunun, Şekil 6'ya göre trend analiziyle yapılan tahminlerin gerçek değerlerle uyumdan daha iyi olduğu görülmektedir.

Bu sonuçlardan hareketle doğal gaz tüketim tahminlerinde yapay sinir ağları ve zaman serisi analizi yanında destek vektör makinelerinin de kullanılabilirliği söylenebilir.

4. Tartışma ve Sonuç

Avrupa birliği ilk ulusal belirgin bir katkı olarak, Paris anlaşması kapsamında 2030 yılı itibarı ile sera gazı emisyonunu en az %40 azaltmayı taahhüt etmiştir. Daha temiz enerji kullanımı için, tüm Avrupa Birliği ülkelerin son yasama eylemi paketinde yenilenebilir enerji için %32, etkin enerji kullanımının geliştirilmesi için ise %32,5 kolektif hedef olarak belirlendi [27].

Türkiye'nin de önüne benzer hedefleri koyabilmesi bakımından mevcut doğalgaz tüketimine ilişkin verileri kullanarak gelecek dönem tahminlerini daha net bir şekilde öngörmesi önemlidir. Bu bağlamda Destek Vektör Makinelerinde regresyon yöntemi kullanılarak Türkiye'deki enerji santrallerinde doğal gaz tüketimi tahmininin bulunması amaçlanan bu

çalışmada; veriler üzerinde trend analizi yapılmış en uygun HKO değerine sahip denklem olan ikinci dereceden denklem seçilerek tahmin değerleri elde edilmiştir. Daha önce destek vektör makineleri kullanılarak yapılan çalışmalar incelendiğinde farklı alanlarda birçok çalışma yapılmıştır [8, 11, 12, 28]. Ancak daha önceki doğal gaz tahminine yönelik yapılan çalışmalarda Akgül (2013), yapay sinir ağları yöntemi ve klasik zaman serileri yöntemi kullanılarak doğal gaz tüketim tahminlerini yıllık olarak elde etmiş sonucunda doğrusal olmayan problemlerde iyi sonuçlar veren yapay sinir ağlarının yine zaman serilerine göre daha iyi sonuçlar verdiğine karar vermiştir.

Topçu (2013) ise, doğal gaz tüketim tahminlerini üretirken sadece zaman serileri yöntemini kullanmış ve uzun dönem yıllık doğal gaz tüketim tahminlerini elde etmiştir.

Bu çalışmada elde edilen sonuçlar kısaca açıklanacak olursa ilk olarak, değişkenlerdeki ölçü birimlerinin farklılığından dolayı veriler standart normal dağılıma uygun hale getirilmiştir. Hemen ardından bağımsız değişkenlerin seçimi için serbestlik dereceleri sırasıyla 5, 4, 3, 2 olan modeller kurulmuş ve bunların AIC ve BIC değerlerine bakılmıştır. Serbestlik derecesi 5 ve en düşük AIC değerine sahip (99,869) olan model en iyi olarak seçilmiş ve çalışmaya bu modeldeki bağımsız değişkenler üzerinden devam edilmiştir. Daha sonra doğrusal olmayan destek vektör regresyonda kullanılan çekirdek fonksiyonlarının hepsi denenmiş ve performansları karşılaştırılmıştır en iyi performansın RTF (Eğitim HKO= 0,249, Test HKO=0,276 olmak üzere) tarafından verildiği görülmüştür. Bunun üzerine ortaya çıkan destek vektörler ile ağırlıklar çarpılmış ve üzerine yan eklenerek katsayılar belirlenmiş böylece5 numaralı denklem elde edilmiştir.

$$\hat{Y} = -0,349 - 0,072 * X_1 - 4,449 * X_2 + 0,459 * X_3 + 4,890 * X_4 + 0,895 * X_5 \quad (5)$$

Bu denkleme bağlı olarak tahminler elde edilmiş ve daha sonra verilere trend analizi uygulanmış burada da en düşük HKO değerine sahip olan ikinci dereceden denkleme göre tahminler bulunmuştur. Sonuçlar karşılaştırıldığında destek vektör makinelerinin daha iyi sonuçlar verdiği gözlemlenmiştir.

Bu çalışmada ise diğer çalışmalarda yapılan klasik zaman serileri ve yapay sinir ağları ile doğal gaz tahminlerinin dışında DVR metodunu kullanarak gelecek dönem (aylık) doğal gaz tahmini yapılmış ve gerçeğe uygunluğuna bakılmıştır. Daha sonra veriler üzerinde trend analizi de uygulanmış, tahminler elde edilmiş ve gerçeğe uygunluğuna bakılmıştır. Yapılan karşılaştırmalara göre DVR metodu kullanılarak elde edilen tahminlerin daha iyi sonuçlar verdiği gözlemlenmiştir.

Dolayısıyla, bu çalışmada karmaşık model problemlerinde daha çok tercih edilen DVR'nin doğal gaz tüketimi tahmininde kullanılabileceği gösterilmiş ve bu yönde literatüre katkı sağlayacağı öngörülmektedir.

Çalışmadan elde edilen bulgular doğrultusunda, doğal gaz tüketim tahminine yönelik gelecek çalışmada kullanılan bağımsız değişkenler dışında farklı değişkenler (Örneğin: Türkiye'nin nüfusu, doğal gaz fiyatı, ekonomik koşullar vb.) kullanılarak DVR, yapay sinir ağları, zaman serileri analizi gibi yöntemler yardımı ile gelecek dönemler için öngörülerde bulunulabilir.

Teşekkür

Bu çalışma Gizem Meral'ın Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsünde, Sinan Saraçlı danışmanlığındaki YL tezinden üretilmiş olup, Afyon Kocatepe Üniversitesi Bilimsel Araştırma ve Projeler Koordinasyonunca Desteklenmiştir. Proje No: 18.FEN.BİL.15. Desteklerinden dolayı Afyon Kocatepe Üniversitesi BAPK'ne teşekkür ederiz.

Kaynakça

- [1] Anonim, 2012. Doğal Gaz Piyasası Sektör Raporu. EPDK, Ankara. <https://www.epdk.org.tr/Detay/Icerik/3-0-94-1007/dogal-gazyillik-sektor-raporu> (Erişim Tarihi: 05.04.2019).
- [2] Kaynar, O., Taştan, S., Demirkoparan, F. 2011. Yapay Sinir Ağları ile Doğal Gaz Tüketim Tahmini. Atatürk Üniversitesi İİBF Dergisi, 10. Ekonometri ve İstatistik Sempozyumu Özel Sayısı, 463-474.
- [3] Aras, H., Aras, N. 2002. Konutsal Doğal Gaz Tüketimi Tahmini. Mühendis ve Makine Dergisi, 46 (540), 20-27.
- [4] Akgül, S. 2013. Doğal Gaz Tüketim Tahmini. Sosyal ve Beşeri Bilimler Dergisi, 5 (1), 440-452.
- [5] TDK. http://tdk.gov.tr/index.php?option=com_gts&arama=gts&kelime=zek%C3%A2&uid=58140&guid=TDK.GTS.5901a6b63382e6.22318161 (Erişim Tarihi: 08.04.2019).
- [6] Yiğit, P. 2011. Yapay Sinir Ağları ve Kredi Taleplerinin Değerlendirilmesi Üzerine Bir Uygulama. İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, 134s, İstanbul.
- [7] Kaytez, F. 2012. En Küçük Kareler Destek Vektör Makineleri İle Türkiye'nin Uzun Dönem Elektrik Tüketim Tahmini ve Modellenmesi. Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 129s, Ankara.
- [8] Arat, M. 2014. Destek Vektör Makineleri Üzerine Bir Çalışma. Hacettepe Üniversitesi, Fen

- Bilimleri Enstitüsü, Yüksek Lisans Tezi, 98s, Ankara.
- [9] Alpaydın, E. 2010. Yapay Öğrenme. 2.Baskı, Boğaziçi Üniversitesi Yayınevi. İstanbul, 486s.
- [10] İstatistiksel Öğrenme ve Makine Öğrenme Arasındaki Fark.<https://www.veribilimi.co/makine-ogrenmesi-ile-istatistiksel-ogrenme-arasinda-ne-fark-var/> (Erişim Tarihi:22.06.2019).
- [11] Bilişik, M.T. 2011. Destek Vektör Makinesi, Çoklu Regresyon ve Doğrusal Olmayan Programlama ile Perakendecilik Sektöründe Gelir Yönetimi İçin Dinamik Fiyatlandırma. İstanbul Kültür Üniversitesi, XI. Üretim Araştırmaları Sempozyumu, 23-24 Haziran, İstanbul, 785-799.
- [12] Uçak, K. 2012. Destek Vektör Regresyonu ile Pid Kontrolör Tasarımı. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 89s, İstanbul.
- [13] Çomak, E. 2008. Destek Vektör Makinelerinin Etki Eğitimi İçin Yeni Yaklaşımlar. Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 118s, Konya.
- [14] Kuzey, C. 2012. Veri Madenciliğinde Destek Vektör Makineleri ve Karar Ağaçları Yöntemlerini Kullanarak Bilgi Çalışanlarının Kurum Performansı Üzerine Etkisinin Ölçülmesi ve Bir Uygulama. İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi, 318s, İstanbul.
- [15] UDEMY. <https://www.udemy.com/veri-bilimi-ve-makine-ogrenmesi-egitimi/learn/v4/t/lecture/13064922?start=75> (Erişim Tarihi: 23.04.2019).
- [16] Osuna, E.E, Freund, R.,Girosi, F. 1997. SupportVectorMachines: Training and Applications. Massachusetts Institute of TechnologyandArtificialIntelligenceLaboratory and Center ForBiologicalandComputationalLearningDepartment of Brain andCognitiveSciences, 1605 (144), 2-31.
- [17] Bican, B. 2014. Zaman Serilerinin Öngörüsü İçin GKA Tabanlı DVR Metodları. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 79s, İstanbul.
- [18] Yu, P.S.,Chen, S.T., Chang, F. 2016. SupportVectorRegressionFor Real-Time FloodStageForecasting. Journal of Hydrology, 328 (3-4), 704-716.
- [19] Haykin, S. 2001. Neural Networks: A Compherensive Foundation. Second Edition, PearsonEducation Publishing. India, 818s.
- [20] Gunn, S.R. 1998. Support Vector Machines for Classification and Regression. Technial Report

- <http://ce.sharif.ir/courses/85-86/2/ce725/resources/root/LECTURES/SVM.pdf> (Erişim Tarihi: 15.04.2019).
- [21] Kor, K. 2015. Penetration Rate Optimization with Support Vector Regression Method. Istanbul Technical University, Graduate School of Science Engineering and Technology, M.Sc. Thesis, 209s, İstanbul.
- [22] Model Belirleme Kriterleri. http://www.acikders.org.tr/pluginfile.php/2616/mod_resource/content/2/ekonometri2-tuba (Erişim Tarihi: 26.03.2019).
- [23] Fabozzi, F.J., Focardi, S.M., Rachev, S.T. Arshanpalli, B.G. 2014. The Basics of Financial Econometrics: Tools, Concepts and Asset Management Applications. First Published, John Wiley & Sons, Inc. Published. New York, 448s.
- [24] Liddle, A.R. 2007. Information Criteria For Astrophysical Model Selection. Monthly Notices of the Royal Astronomical Society: Letters, 377 (1), 71-76.
- [25] Burnham, K., Anderson, D. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second Edition, Springer-Verlag. New York, 485s.
- [26] Performans Fonksiyonları. <https://veribilimcisi.com/2017/07/14/mse-rmse-mae-mape-metrikleri-nedir/> (Erişim Tarihi: 28.03.2019).
- [27] Cātuşi, M., Egenhofer, C., and Elkerbout M. 2019. The future of gas in Europe: Review of recent studies on the future of gas. Research Report No. 2019/03, https://www.ceps.eu/wp-content/uploads/2019/08/RR2019-03_Future-of-gas-in-Europe.pdf (Erişim Tarihi: 29.03.2020).
- [28] Yalçın, C. 2016. Destek Vektör Regresyon İle Uçaklarda Anlık Yakıt Tüketiminin Analizi. Mimar Sinan Güzel Sanatlar Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 141s, İstanbul.
- [29] Topçu, G.Y. 2013. Türkiye Doğalgaz Tahmini. Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi, Ankara.