

Yayın Geliş Tarihi (Submitted): 26/11/2020

Yayın Kabul Tarihi (Accepted): 28/01/2021

Makele Türü (Paper Type): Araştırma Makalesi – Research Paper

Please Cite As/Atıf için:

Demirezen S. ve Çetin M. (2021), Rassal orman regresyonu ve destek vektör regresyonu ile piyasa takas fiyatının tahmini, *Nicel Bilimler Dergisi*, 3(1), 1-15. doi: 10.51541/nicel.832164

RASSAL ORMAN REGRESYONU VE DESTEK VEKTÖR REGRESYONU İLE PİYASA TAKAS FİYATININ TAHMİNİ

Sinan Demirezen¹ ve Meral Çetin²

ÖZET

Antik çağdan beri varlığı bilinen statik elektrik 1880’de New York’ta üretilen elektrik ile farklı bir anlam kazanarak insan hayatının vazgeçilmez bir unsuru olmuştur. Günümüzde, temel ihtiyaç alanına girmiş olan elektriğin üretiminden dağıtımına kadar önceleri devlet tekeliyle gerçekleştirilse de özellikle 1980’li yıllardan itibaren elektrik piyasası serbestleştirilmeye başlanarak rekabetçi bir yapıya dönüşmesi amaçlanmıştır. Serbestleşme adımları başta Şili olmak üzere, İngiltere, Avustralya, Yeni Zelanda ve Baltık ülkelerinde gerçekleşmiş ve günümüzde de bu dönüşüm süreci devam etmektedir. Ülkemizde ise elektrik piyasasındaki serbestleşme çalışmaları tam olarak 2000’li yıllarda gerçekleşmeye başlamıştır. 2015 yılında EPDK’dan aldığı piyasa işletim lisansı ile Enerji Piyasaları İşletme Anonim Şirketi (EPIAŞ) faaliyete geçerek elektrik piyasasının serbestleştirilmesinde önemli bir adım atılmıştır. Bu çalışmada, EPIAŞ tarafından işletilmekte olan Gün Öncesi Piyasası’nda belirlenen saatlik Piyasa Takas Fiyatının (PTF) tahmin edilmesi amaçlanmıştır. PTF’nin geçmiş değerlerinin ve gün öncesi piyasasında oluşan işlem hacminin PTF tahminindeki başarısı araştırılmıştır. Tahmin yöntemi olarak, makine öğrenmesi yöntemlerinden rassal orman regresyonu ve destek vektör regresyonu kullanılmıştır. Analiz sonucunda, makine öğrenmesi yöntemlerinin tahmin performanslarının karşılaştırılmasında literatürde sıklıkla kullanılan RMSE, MAE ve MAPE kriterlerine göre rassal orman regresyon yöntemi ile gerçekleştirilen ve işlem hacminin de dahil olduğu değişken grubu PTF’yi en iyi tahmin eden model (RFR-2.grup) olmuştur. Bu çalışma

¹Sorumlu yazar, Araş. Gör., Ekonometri Bölümü, İktisat Fakültesi, İstanbul Üniversitesi, İstanbul, Türkiye, ORCID ID: <https://orcid.org/0000-0002-5009-6421>

²Prof. Dr., İstatistik Bölümü, Fen Fakültesi, Hacettepe Üniversitesi, Ankara, Türkiye, ORCID ID: <https://orcid.org/0000-0003-0247-7120>

ile işlem hacminin PTF için önemli bir değişken olduğu belirlenmiş olup PTF tahmin çalışmalarında diğer yöntemlere göre görece daha az kullanılan rassal orman regresyonunda bu yöntemler kadar önemli olduğu görülmüştür.

Anahtar Kelimeler: Destek Vektör Regresyonu, Gün Öncesi Piyasası, Makine Öğrenmesi, Piyasa Takas Fiyatı, Rassal Orman Regresyonu

FORECASTING OF MARKET CLEARING PRICE WITH RANDOM FOREST REGRESSION AND SUPPORT VECTOR REGRESSION

ABSTRACT

Electricity known as static electricity until nineteenth century took on a new meaning with it was generated in New York in 1880. Electricity have been an indispensable instrument for human life. Initially, governments undertook the electricity process from generating to distribution. However, in 1980's, also in Turkey at the beginning of the 2000's, countries such as Chile, England, Australia started to liberalize their electricity markets for competition. This paper aims to predict hourly Market Clearing Price (MCP) announcing in Day Ahead Market being operated by Turkish Energy Exchange (EXIST) with market operating licence dated from 2015. It is researched that lagged values of MCP and trade value of day ahead market how prediction success on MCP. As prediction method, random forest and support vector machine of machine learning methods was used. Analysis period involve 1 Jan 2019 00:00 and 10 Mar 2020 23:00 and consist of 10440 data divided into two subset as training set (%84) and test set (%16). K-fold Cross-validation method is used for describing best parameter. Analysis was implemented in R program with caret and e1071 package. As a result of the analysis, according to the RMSE, MAPE, MAE using frequently in literature for comparing forecast performance, the best method and the best variable group which predict MCP is respectively random forest regression and the group including trade value. Therefore, this paper demonstrated that trade value is important variable for MCP and random forest is important method just as other methods used prediction of MCP.

Keywords: Day Ahead Market, Machine Learning, Market Clearing Price, Random Forest Regression, Support Vector Regression

1. GİRİŞ

Elektriğin günümüzdeki anlamıyla kullanılmaya başlanmasından itibaren üretiminden dağıtımına kadar devlet tekelinde olan elektrik endüstrisi dikey bütünleşik bir yapıda iken petrol krizinin küresel etkisiyle birlikte dönüşüm sinyalleri vererek rekabetçi bir yapının ortaya çıkmasına zemin hazırlamıştır. Elektrik endüstrisinin serbestleştirilmesi dünya genelinde çok önemli bir konu haline gelmiş olup serbestleştirilmedeki temel amaç elektrik üretiminde verimi artırmak, elektrik tüketimini en üst seviyeye çıkarmak ve elektrik fiyatlarını düşürmektir (Sahay ve Tripathi, 2014). Bu nedenle 1980’li yıllardan itibaren elektrik piyasasının serbestleştirilmesi ve özel yatırımların piyasaya çekilmesi amaçlanmıştır. İlk olarak 1982 yılında Şili’de elektrik piyasası serbestleştirilmeye başlanmış daha sonra bu düşünce İngiltere, Arjantin, Avustralya, İspanya, Norveç gibi ülkelere yayılmıştır. Türkiye’de ise serbestleşme çalışmaları her ne kadar 1980’li yıllarda temelleri atılsa da tam anlamıyla 2000’li yılların başında yaşanmaya başlamıştır.

Serbestleşen piyasada, piyasa katılımcıları, verecekleri tekliflerde en az maliyet ve en çok kâr hedefini gerçekleştirebilmek için piyasa yapısı hakkında bilgi sahibi olma çabası içerisinde olacaklardır. Bu bakımdan elektrik fiyatları, piyasa yapısı hakkında önemli bilgi vermekte olup katılımcılar açısından önem arz etmektedir. Bu açıdan, katılımcılar hedefleri doğrultusunda piyasaya verecekleri tekliflerde elektrik fiyatlarının seyrini önceden bilmek isteyeceklerdir. Bu yönüyle, elektrik fiyatlarının tahmini katılımcılara önemli bir gösterge olacaktır. Elektrik piyasası doğası gereği diğer piyasalardan daha oynak bir yapıya sahip olduğundan fiyat tahmin çalışmalarında elektrik fiyatlarını etkileyen değişkenler ve tahmin yöntemleri önem kazanmaktadır (Nargele ve Patil, 2016). Bu anlamda, literatürde elektrik fiyatını tahmin eden çalışmalar bu bölüm içerisinde özetlenmeye çalışılmıştır.

Catalão ve diğerlerinin (2005) çalışmasında, Backpropagation algoritması ile eğitilen üç katmanlı ileri beslemeli yapay sinir ağlarını kullanarak saatlik elektrik fiyat tahmini yapmışlardır. Conejo ve diğerlerinin (2005) çalışmasında İspanya gün öncesi piyasasında ARIMA, Wavelet-ARIMA ve Naïve yöntemi kullanarak elektrik fiyatını tahmin etmişlerdir. Sun ve Zang’ın (2008) çalışmasında, elektrik fiyatlarını saatlik olarak yapay sinir ağları ve genetik algoritmalara dayanan destek vektör makineleri (GA-SVM) yöntemi ile tahmin etmişlerdir. Tang ve Gu’nun (2009) çalışmasında, önerdikleri Backpropagation algoritması ile elektrik fiyatlarını tahmin etmişlerdir. Saini ve diğerlerinin (2016) çalışmasında, parçacık sürü optimizasyonuna dayalı destek vektör makineleri ve doğrusal regresyonu kullanarak elektrik fiyat tahmini gerçekleştirmişlerdir. Yan ve Chowdhury’nin (2013) çalışmasında, piyasa takas

fiyatını, en küçük kareler SVM (LSSVM) ve ARMAX'ın birleşiminden oluşan hibrit model ve LSSVM ile tahmin etmişlerdir. Anamika ve Kumar'ın (2016) çalışmasında, Hindistan elektrik piyasasında piyasa takas fiyatını yapay sinir ağları ve regresyon ile tahmin etmişlerdir. Ding ve Ge'nin (2018) çalışmasında, PJM elektrik piyasasında uyarlamalı kalman filtresi yöntemiyle elektrik fiyatlarını tahmin ederlerken Tat'ın (2018) çalışmasında, Litvanya gün öncesi elektrik piyasasında Monte Carlo simülasyonu yöntemi ile elektrik fiyatlarını öngörmüştür. McGlynn ve diğerleri (2019) ise, Büyük Britanya elektrik piyasasında elektrik fiyatlarını genelleştirilmiş regresyon sinir ağları ve AdaBoosted algoritması ile tahmin etmişlerdir.

Literatür incelendiğinde, elektrik fiyat tahmininde birçok yöntem kullanılmıştır. Bu yöntemler arasında yapay sinir ağları, destek vektör makineleri gibi makine öğrenmesi yöntemleri; ARIMA, ARMAX, GARCH, Wavelet gibi zaman serileri yöntemleri; genetik algoritmalar, bulanık yöntemler ve hibrit yöntemler yer almaktadır. Bu çalışmada ise, Piyasa Takas Fiyatı (PTF) tahmin çalışmalarında diğer yöntemlere göre görece daha az kullanılan rassal orman regresyonu (RFR) ve yaygın olarak kullanılan destek vektör regresyonu (DVR) ile PTF tahmin edilmeye çalışılmış ve kullanılan bu iki yöntemin tahmin performanslarının karşılaştırılması amaçlanmıştır. PTF ise Gün Öncesi Piyasasına (GÖP) sunulan tekliflerin eşleştirilmesiyle belirlenen saatlik elektrik enerjisi alıŖ- satıŖ fiyatı olarak tanımlanmaktadır (EPIAŖ Ŗeffaflık Platformu, 2020)

Bu çalışmada GiriŖ bölümüne ek olarak ikinci bölümde gün öncesi piyasası ve uygulamada kullanılan yöntemler hakkında bilgiler, üçüncü bölümde uygulama ve dördüncü bölümde elde edilen bulgular ve beşinci bölümde ise sonuç ve öneriler bölümü yer almaktadır.

2. GENEL BİLGİLER

2.1. Gün Öncesi Elektrik Piyasası

Türkiye'de Gün öncesi elektrik piyasası, bir gün sonrası teslim edilecek uzlaŖtırma dönemi temelinde elektrik enerjisi alıŖ-satıŖ işlemleri için kurulan ve 2015'ten beri piyasa işleme lisansına sahip EPIAŖ kurumu tarafından işletilen organize toptan elektrik piyasası olarak tanımlanmaktadır. GÖP'te işlemler saatlik düzeyde gerçekleştirilmektedir (Elektrik Piyasası Dengeleme ve UzlaŖtırma Yönetmeliđi, 2009). Piyasada bir gün saat 00:00'dan başlayıp ertesi gün 00:00'a kadar devam eden zaman dilimlerinden oluşmaktadır. GÖP'ün uzlaŖtırmasında uygulanan fiyatlar ve miktarlar günlük olarak ve her bir saat için belirlenmektedir. GÖP'e katılan piyasa katılımcıları her gün 12:30'a kadar bir sonraki güne

ilişkin tekliflerini EPIAŞ'a bildirmektedirler. Verilen her bir teklif, 12:30-13:00 arasında EPIAŞ tarafından değerlendirilerek doğrulanmaktadır. Doğrulananan teklifler, 13:00-13:30 arasında optimizasyon aracı ile değerlendirilip bir sonraki günün her bir saatine ait PTF ve piyasa takas miktarı belirlenerek piyasa katılımcılarına bildirilmektedir. 13:30-13:50 arasında piyasa katılımcılarına itiraz hakkı tanınmaktadır. Yapılan itirazlar saat 14:00'e kadar değerlendirilerek bir sonraki güne ait saatlik fiyatlar ve eşleşmeler nihai olarak ilan edilmektedir (EPIAŞ, 2020; GÖP kullanıcı kılavuzu, 2016).

2.2. Metodoloji

Bu bölüm içerisinde sırasıyla RF ve SVR yöntemlerine ait bilgiler verilmiştir.

2.2.1. Rassal Orman Regresyonu

Denetimli makine öğrenmesi yöntemlerinden birisi olan rassal orman (RF) yöntemi, rastgele değişkenler topluluğuna bağlı ağaçlardan oluşan bir ağaç tabanlı topluluktur. Amit ve Geman'ın 1997 çalışmasından esinlenen Leo Breiman tarafından tanıtılmıştır. RF, Breiman'ın Bagging fikrinin bir uzantısı olup boosting yöntemine alternatif olması için geliştirilmiştir (Breiman, 2001; Cutler vd., 2012). RF hem regresyon hem de sınıflandırma problemlerine uygulanması, diğer yöntemlere göreli olarak daha hızlı eğitilmesi ve tahmin hızının daha yüksek olması, düzenleme parametre sayısının daha az olması, çok boyutlu problemlere doğrudan uygulanabilmesi gibi özelliklerinden dolayı dikkat çekmektedir (Hastie vd., 2017).

RF, karar ağaçlarının en önemli dezavantajı olarak görülen eğitim verilerini aşırı öğrenme eğiliminde olması problemine bir çözüm yolu olma özelliği taşımaktadır. Bağımsız karar ağaçlarını hem gözlemleri hem de eğitim verilerindeki değişkenleri örneklemeye dayanmaktadır. RF, sınıflandırma problemleri için kullanıldığında her bir ağaçtan bir sınıf oyu olarak oy çoğunluğuna göre sınıflandırma yapmaktadır. Regresyon problemlerinde (RFR) ise tahminlerin ortalaması alınmaktadır (Dangeti, 2017). RF, temelde iki düzenleme parametresine sahiptir. Bunlar, her düğümde rassal seçilen tahmin edicilerin sayısı (m) ve ormandaki ağaç sayısı (J) parametreleridir. Sınıflandırmada, varsayılan m değeri $m = \sqrt{p}$ olup burada p , toplam tahmin edici sayısını göstermektedir. Öte yandan, RFR'de ise bu değer $p/3$ olarak temel alınmaktadır (Cutler vd., 2012).

2.2.2. Destek Vektör Regresyonu

Destek vektör makineleri (DVM), ilk kez Vapnik ve arkadaşları tarafından tanıtılmış olup Vapnik-Chervonenkis (VC) teorisine dayanan denetimli bir makine öğrenmesi yöntemidir.

VC teorisi, test verilerini iyi bir şekilde tahmin etmesini sağlayan makine öğrenmesinin özelliklerini belirlemektedir. DVM, hem sınıflandırma hem de regresyon problemlerine uygulanabilmektedir (Schölkopf ve Smola, 2004).

Destek vektör regresyonunun (DVR) mantığı, doğrusal olmayan bir fonksiyon üzerinden girdi verilerinin eşlendiği daha yüksek boyutlu girdi uzayında (feature space) bir doğrusal regresyon fonksiyonunun hesaplanmasına dayanmaktadır. Böylelikle, düşük boyutlu girdi uzayındaki doğrusal olmayan bir regresyon probleminden yüksek boyutlu girdi uzayındaki doğrusal bir regresyon problemi elde edilmekte olup çözüm bu uzayda gerçekleştirilmektedir (Mohandes, 2002).

$$f(x) = \sum_{i=1}^D w_i \phi_i(x) + b \quad (1)$$

Burada $\{\phi_i(x)\}_{i=1}^D$ girdi (features) olarak adlandırılmaktadır. b ve $\{w_i\}_{i=1}^D$ ise tahmin edilecek katsayılar olup y çıktısından en fazla ε değeri kadar bir sapmaya sahip bir f fonksiyonunda optimal değerlerinin bulunması amaçlanmaktadır. $\{w_i\}_{i=1}^D$ katsayıları aşağıdaki verilen denklemin minimize edilmesiyle tahmin edilmektedir (Mohandes, 2002).

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (2)$$

$$w^T \phi(x_i) + b - y_i \leq \varepsilon$$

$$y_i - w^T \phi(x_i) - b \leq \varepsilon$$

Burada, ε ' dan daha yüksek sapmalar kabul edilmemekte olup ε ' dan daha düşük sapmalar ise dikkate alınmamaktadır. Bu problemin çözümü için, optimizasyon probleminde kısıtlara ξ_i , ξ_i^* notasyonu ile ifade edilen ve gevşeklik adı verilen parametreler eklenerek aşağıdaki denklemler elde edilmektedir (Schölkopf ve Smola, 2004; Davò vd., 2016).

$$\min_{w,b,\xi_i,\xi_i^*} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^D (\xi_i + \xi_i^*) \quad (3)$$

$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i$$

$$y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i^*$$

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

Denklemdaki C sabiti, sıfırdan büyük bir değer alıp model karmaşıklığı ile ε ' dan büyük sapmaların göz ardı edildiği miktar arasındaki mübadeleyi belirlemektedir. Bu problem doğrusal DVR' de, (3) duali alınarak aşağıda verilen şekliyle,

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) \langle x_i, x \rangle + b \quad (4)$$

Doğrusal olmayan DVR' de ise kernel fonksiyonu vasıtasıyla

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) k(x_i, x) + b \quad (5)$$

Şeklinde çözülmektedir (Schölkopf ve Smola, 2004; Davò vd., 2016).

Burada, a_i, a_i^* 3'teki problemin kısıtlarının dual değişkenleri olup $[0, C]$ aralığında değerler almaktadır. $k(x_i, x)$ ise Kernel fonksiyonu olarak adlandırılmakta olup yüksek boyutlu girdi uzayındaki X_i ve X_j şeklinde iki vektörün iç çarpımlarına eşittir (Cao ve Tay, 2001). Böylelikle, doğrusal olmayan DVR, girdi uzayından yüksek boyutlu girdi uzayına veriler eşlenerek doğrusal bir şekilde çözümlenmesi sağlanır. En çok bilinen Kernel fonksiyonları, doğrusal, polinom, üstel, Gauss'dur (Santamaría vd., 2013).

3. UYGULAMA

Bu çalışmada, Piyasa takas fiyatını makine öğrenmesi yöntemlerinden rassal orman regresyonu ve destek vektör regresyonu ile tahmin edilmesi amaçlanmıştır. Çalışmada analiz dönemi, 1 Ocak 2019 00:00 ve 10 Mart 2020 23:00 dönemini kapsamakta olup 10440 veriden oluşmaktadır. Veri, makine öğrenmesi yöntemlerinin bir gereği olarak eğitim kümesi ve test kümesi olmak üzere iki alt kümeye ayrılmıştır. Eğitim kümesi 1 Ocak 2019 00:00 ile 31 Aralık 2019 23:00 dönemini kapsamakta olup 8760 veriden ve test kümesi ise 1 Ocak 2020 00:00 ile 10 Mart 2020 00:00 dönemini kapsamakta olup 1680 veriden oluşmaktadır. Verinin %84' lük bir kısmının eğitim kümesine ayrılma nedeni PTF'nin bir yıl içerisinde tüm mevsimlere ait seyrinin eğitim kümesine dahil edilerek model eğitiminde kullanılmak istenmesidir. Böylece, eğitilen modelin tahmin performansının artırılması amaçlanmıştır. Uygulamada, eğitim kümesinin sadece bir dönem aralığında değerlendirilecek olmasından (verilerin sonbahar mevsimine denk gelmesi gibi) veya verilerin rassal olarak dağıtılacak olması zaman serisi verisi niteliğinde olan bu veriye uygun olmamasından dolayı geçerlilik kümesi (validation set) oluşturulmamıştır. Bunun yerine, modelde aşırı öğrenme problemini engellemek ve en iyi parametreleri belirlemek için k- katlı çapraz geçerlilik (k-fold Cross-Validation) yönteminden yararlanılmıştır. k değeri, literatürde 2, sıklıkla 5 ve 10 değerleri alınarak analizler gerçekleştirildiği tespit edilmiştir. Bu çalışmada, hem sapma (bias) ve varyans arasındaki dengeyi sağlamak hem de modelin aşırı öğrenme eğilimini engellemek adına k değeri 5 olarak alınmıştır. Uygulama R programında gerçekleştirilmiştir. RFR için "caret", DVR için "e1071" paketi kullanılmıştır. caret paketi kapsamında RFR için parametreler mtry (m) ve ntree (J); e1071 paketi kapsamında ise doğrusal, radyal, polinom ve sigmoid kernel tipine göre cost,

epsilon ve gamma parametreleri yer almakta olup birçok deneme gerçekleştirilerek tahmin açısından en iyi düzenleme parametreleri belirlenmeye çalışılmıştır. En iyi parametrelerin seçiminde yukarıda da bahsedildiği üzere 5-katlı çapraz geçerlilik yönteminden yararlanılmıştır. Öte yandan, iyi bir tahmin yapabilmek için analiz öncesi değişkenlere Min-Max normalleştirilmesi uygulanarak [0,1] aralığında veriler ölçeklendirilmiştir. Tahmin grafiği çizdirilirken veriler tekrar denormalize edilerek gerçek PTF değerleriyle karşılaştırılması sağlanmıştır. Analizde gerek literatürden elde edilen bilgiler gerekse de PTF'yi etkilediği düşünülen değişkenler kullanılarak iki farklı grup oluşturulmuş ve bu yöntemler her iki gruba da uygulanmıştır. Tablo 1'de uygulamada kullanılan değişkenlere ait açıklamalar ve gruplarda yer alan değişkenler belirtilmiştir.

Tablo 1. Uygulamada kullanılan değişkenler ve oluşturulan gruplar

Değişkenler	Açıklama	1. Grup	2. Grup
DPTF	Bir gün önceki PTF (\$)	•	•
WPTF	Bir hafta önceki PTF (\$)	•	•
MPTF	Bir ay önceki PTF (\$)	•	•
Hafta	Hafta içi günler 0; hafta sonu günler 1	•	•
Tatil	Tatil ve bayram günleri 1; diğer günler 0	•	•
H1PTF	Bir saat önceki PTF (\$)	•	•
H2PTF	İki saat önceki PTF (\$)	•	•
H3PTF	Üç saat önceki PTF (\$)	•	•
İşlem hacmi	GÖP'te eşleşen alış tekliflerinin saatlik toplam mali değeridir (₺)	○	•

Tablo 1 incelendiğinde PTF' nin tahmininde bir gün önceki, bir hafta önceki ve bir ay önceki PTF değerleri, haftanın günleri (hafta içi günlere 0, hafta sonu günlere 1 değeri atanmıştır), tatil ve bayram günleri (tatil ve bayram günlerine 1, diğer günlere 0 değeri atanmıştır), üç saat öncesine kadar PTF' nin gecikmeli değerleri ve işlem hacmine yer verilmiştir. Bu çalışmanın bir diğer amacı da işlem hacminin PTF tahminindeki etkisinin araştırılması olarak belirlenmiştir. Bu amaçla iki grup oluşturulmuş ve ikinci grupta birinci gruptaki değişkenlerden farklı olarak işlem hacmi değişkenine yer verilmiştir. Ayrıca, PTF

değerleri Dolar cinsinden alınarak döviz kurundaki değişmelerin tahmin performansını düşürmemesi amaçlanmıştır.

Makine öğrenmesi yöntemlerinin tahmin performanslarının karşılaştırılmasında birçok alternatif kriter mevcuttur. Bu çalışmada, en çok kullanılan kriterler arasında yer alan MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error) kriterleri göz önünde bulundurularak tahmin performansı karşılaştırılmıştır.

$$MAE = \frac{1}{N} \sum_{t=1}^N |\hat{y}_t - y_t| \quad (6)$$

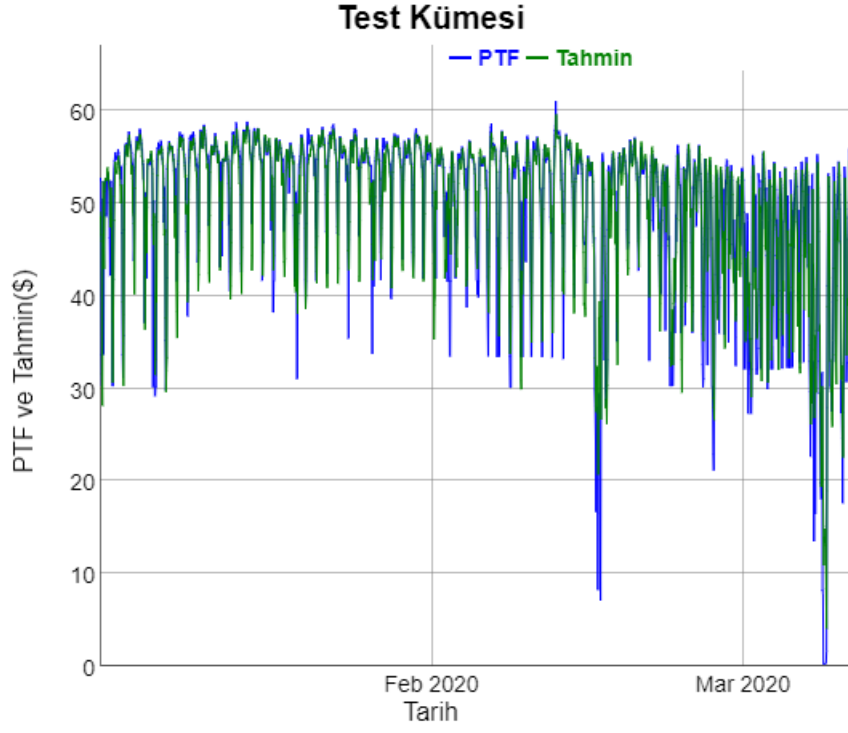
$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{\hat{y}_t - y_t}{y_t} \right| \cdot 100 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N |\hat{y}_t - y_t|^2} \quad (8)$$

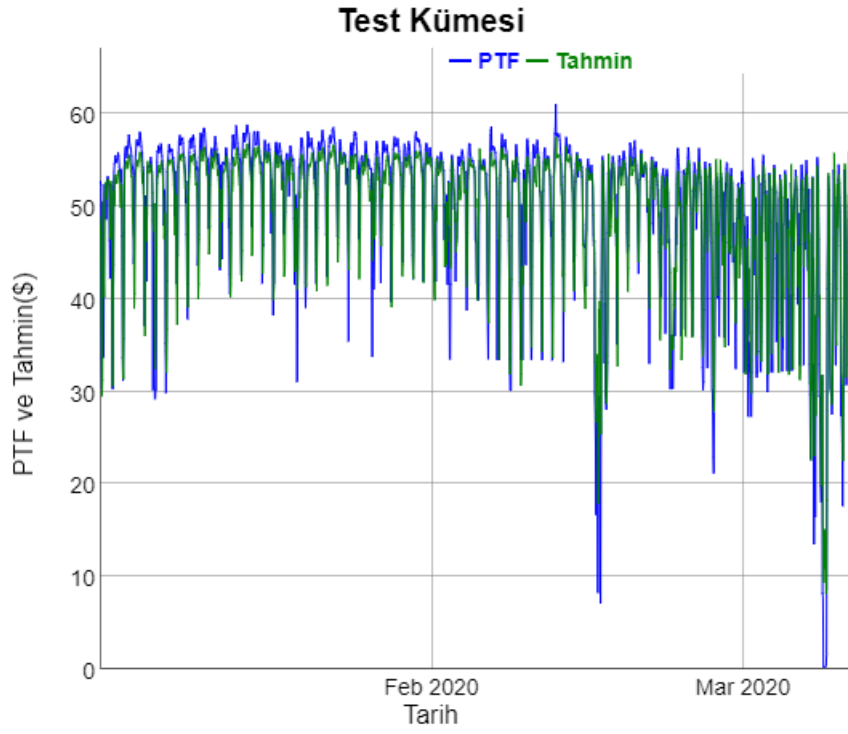
Burada, y_t 'ler gerçek değeri, \hat{y}_t 'lar tahmin değerlerini, N gözlem sayısını ve t ise verinin zaman serisi verisi niteliğinde olduğunu ifade etmektedir (Bello vd., 2010).

4. BULGULAR

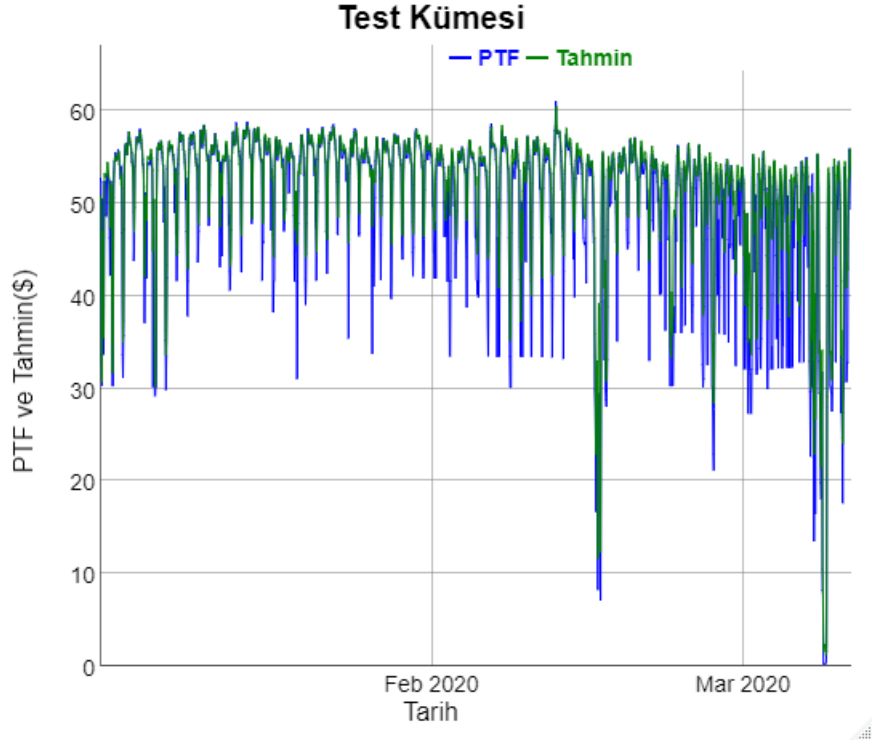
Bu bölümde, RFR ve DVR ile gerçekleştirilen analizin sonuçları özetlenmiştir. Tahmin performansını en az hata ile gerçekleştirmek için RFR' de en iyi parametre değerleri belirlenirken mtry parametresi 2 ve 3 değerleri (Değişken sayısı 1. Grup için 8, 2. Grup için 9 olduğundan), ntree parametresi için ise 300 ile 1000 arasında değerler alınarak Grid Search yöntemiyle birçok deneme gerçekleştirilmiştir. RMSE kriterine göre 1. Grup için en iyi parametre değerleri sırasıyla 2 ve 600, 2.Grup için ise 3 ve 900 olarak belirlenmiştir. Öte yandan, SVR' de cost parametresi 10^{-2} ile 10 arasında, epsilon parametresi 2^{-4} ile 10^{-1} arasında ve gamma parametresi ise 10^{-3} ile 1 arasında Random Search yöntemiyle en iyi kernel tipi ve parametreler belirlenmeye çalışılmıştır. Yapılan denemelerin ardından en az hata miktarına sahip kernel tipi 1. Grup için radial kerneli olup cost parametresi 1, epsilon parametresi 0.25, gamma parametresi ise 0.1 olarak elde edilmiştir. 2.Grupta ise kernel tipi radial kerneli olup cost parametresi 10, epsilon parametresi 0.1 ve gamma parametresi ise 0.1 olarak elde edilmiştir. Böylelikle, bu parametreler ile model eğitilerek eğitim ve test kümesi tahmin edilmiştir. Şekil 1'de ve Şekil 2'de 1.Grubun test kümesine ait sırasıyla RFR ve DVR yöntemiyle gerçekleştirilen tahmin grafiği, Şekil 3'te ve Şekil 4'te ise 2.Grubun test kümesine ait sırasıyla RFR ve DVR yöntemiyle gerçekleştirilen tahmin grafiği verilmiştir.



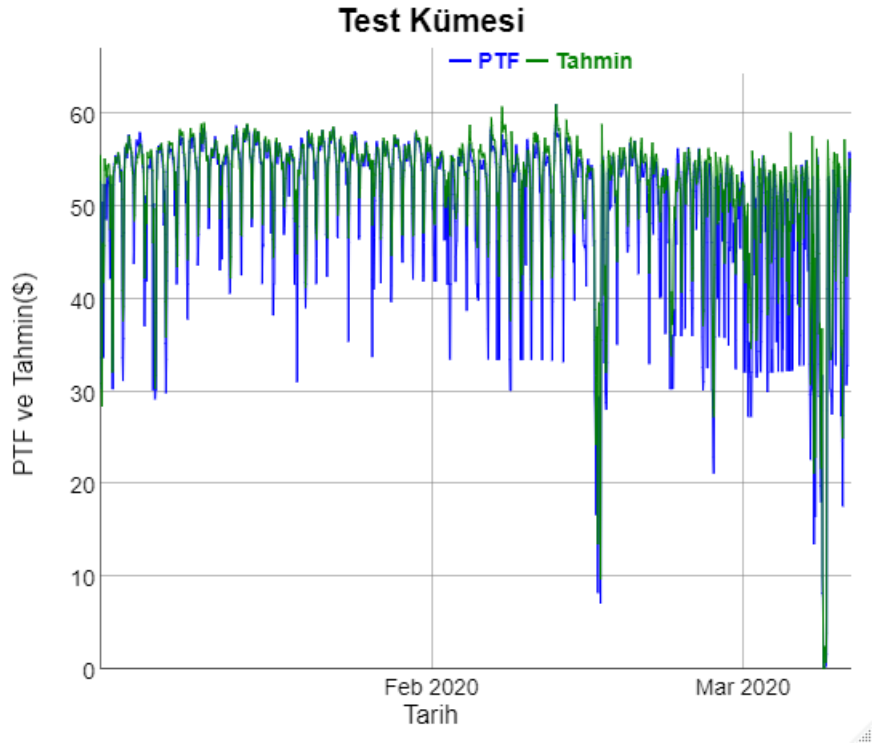
Şekil 1. 1.Grubun test kümesine ait RFR ile gerçekleştirilen tahmin grafiği



Şekil 2. 1.Grubun test kümesine ait DVR ile gerçekleştirilen tahmin grafiği



Şekil 3. 2.Grubun test kümesine ait RFR ile gerçekleştirilen tahmin grafiği



Şekil 4. 2.Grubun test kümesine ait DVR ile gerçekleştirilen tahmin grafiği

Şekiller incelendiğinde, mavi renk ile PTF değerleri ve yeşil renk ile PTF'nin tahmin değerleri gösterilmiştir. Verilen şekillerden hareketle, PTF'yi en iyi açıklayan yöntem RFR ve değişken grubunun ise işlem hacminin yer aldığı 2.Grubun olduğu görülmektedir. Ancak sonuçların güvenilir olabilmesi için Tablo 2'de ve Tablo 3'te verilen sonuçların incelenmesi gerekmektedir.

Tablo 2. Eğitim kümesine ait tahmin performansları

Kriter/Model	Eğitim Kümesi			
	RFR-1.Grup	RFR-2.Grup	DVR-1.Grup	DVR-2.Grup
MAE	0.02285	0.00973	0.04799	0.02110
MAPE	0.08277	0.04054	0.15770	0.09586
RMSE	0.03843	0.01607	0.07757	0.03343

Tablo 2'de eğitim kümesine ait tahmin performansları özetlenmiştir. Bu sonuçlardan hareketle en düşük hata performans kriterine sahip yöntem RFR olup 2. Grubu oluşturan değişkenler PTF'yi daha iyi açıklamıştır. Tablo 3'te ise test kümesine ait tahmin performansları özetlenmiştir.

Tablo 3. Test kümesine ait tahmin performansları

Kriter/Model	Test Kümesi			
	RFR-1.Grup	RFR-2.Grup	DVR-1.Grup	DVR-2.Grup
MAE	0.02914	0.02604	0.03224	0.02922
MAPE	0.06185	0.05238	0.06589	0.05890
RMSE	0.04882	0.04479	0.04906	0.04671

Tablo 3 incelendiğinde ise Test kümesinde en düşük hata performansına sahip yöntem RFR iken 2.Grubu oluşturan değişkenler PTF'yi daha iyi açıklamıştır. Bu sonuçlar, yukarıda yapılan grafik yorumunu desteklemektedir.

5. SONUÇ

Bu çalışmada, Gün öncesi piyasasında belirlenen Piyasa Takas Fiyatının tahmini gerçekleştirilmiştir. Analizde, PTF'yi etkilediği düşünülen ve literatürden elde edilen bilgiler ışığında iki grup değişken oluşturularak işlem hacminin PTF açısından önemli olup olmadığı araştırılmıştır. Öte yandan, rassal orman regresyonu ve destek vektör regresyonu ile PTF tahmin

edilmiş olup hangi yöntemin veri seti dönemi kapsamında PTF açısından daha önemli olduğu belirlenmeye çalışılmıştır. Bu çalışma kapsamında sunulan gerek şekiller gerekse tablolardan elde edilen bulgulardan hareketle ilgili veri seti döneminde işlem hacminin PTF açısından önemli bir değişken olduğu ve RFR yönteminin DVR yöntemine göre PTF'yi daha iyi açıkladığı sonucuna ulaşılmıştır. Böylelikle, bu çalışma, işlem hacmi değişkenin PTF tahmini açısından önemli olduğunu göstermiş olup bundan sonra yapılacak ilgili tahmin çalışmalarında yol gösterici olması beklenmektedir. Öte yandan, PTF'yi etkilediği düşünülen diğer değişkenler çalışmaya dahil edilerek ve regresyon çalışmasına olanak sağlayan diğer makine öğrenmesi yöntemleri kullanılarak çalışmanın kapsamı genişletilebilir.

ETİK BEYAN

“Rassal Orman Regresyonu ve Destek Vektör Regresyonu ile Piyasa Takas Fiyatının Tahmini” başlıklı çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamış ve bu çalışma herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiştir.

KAYNAKÇA

Breiman, L. (2001), Random forests. *Machine learning*, 45(1), 5-32.

Catalão, J., Mariano, S., Mendes, V. and Ferreira, L. (2005), An artificial neural network approach for day-ahead electricity prices forecasting, *WSEAS Transactions on Systems*, 4(4), 451-454.

Conejo, A. J., Plazas, M. A., Espinola, R. and Molina, A. B. (2005), Day-ahead electricity price forecasting using the wavelet transform and ARIMA models, *IEEE transactions on power systems*, 20(2), 1035-1042.

Cutler, A., Cutler, D. R. and Stevens, J. R. (2012), Random forests, *Ensemble machine learning Methods and Applications*, Springer, Boston, MA.

Dangeti, P. (2017), *Statistics for machine learning*, Packt Publishing Ltd., Birmingham, UK.

Davò, F., Vespucci, M. T., Gelmini, A., Grisi, P. and Ronzio, D. (2016, October). Forecasting Italian electricity market prices using a Neural Network and a Support Vector Regression. In *2016 AEIT International Annual Conference (AEIT)*, 1-6.

- de Marcos, R. A., Bello, A. and Reneses, J. (2019), Electricity price forecasting in the short term hybridising fundamental and econometric modelling, *Electric Power Systems Research*, 167, 240-251.
- Ding, L. and Ge, Q. (2018), Electricity market clearing price forecast based on adaptive Kalman filter. In *2018 International Conference on Control, Automation and Information Sciences (ICCAIS)*, 417-421.
- Elektrik Piyasası Dengeleme ve Uzlaştırma Yönetmeliği, <https://www.epias.com.tr/mevzuat/dengeleme-ve-uzlastirma-yonetmeliği/>, Erişim Tarihi:11.10.2020.
- EPIAŞ Şeffaflık Platformu, <https://seffaflik.epias.com.tr/transparency/piyasalar/gop/ptf.xhtml>, Erişim Tarihi:11.10.2020.
- EPIAŞ, <https://www.epias.com.tr/gun-oncesi-piyasasi/surecler/>, Erişim Tarihi:11.10.2020.
- GÖP kullanıcı kılavuzu, <https://www.epias.com.tr/gun-oncesi-piyasasi/gop-kullanici-kilavuzu/>, Erişim Tarihi:12.10.2020.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Kumar, N. (2016). Market clearing price prediction using ANN in indian electricity markets. In *2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS)*, 454-458.
- McGlynn, D., Coleman, S., Kerr, D. and McHugh, C. (2018), Day-Ahead Price Forecasting in Great Britain's BETTA Electricity Market. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2112-2116.
- Mohandes, M. (2002), Support vector machines for short-term electrical load forecasting. *International Journal of Energy Research*, 26(4), 335-345.
- Nargale, K. K. and Patil, S. B. (2016), Day ahead price forecasting in deregulated electricity market using Artificial Neural Network, In *2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS)*, 527-532.
- Sahay, K. B. and Tripathi, M. M. (2014), Day ahead hourly load forecast of PJM electricity market and ISO New England market by using artificial neural network, In *ISGT 2014*, 1-5.

- Saini, D., Saxena, A., and Bansal, R. C. (2016). Electricity price forecasting by linear regression and SVM, In *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, 1-7.
- Santamaría-Bonfil, G., Frausto-Solís, J. and Vázquez-Rodarte, I. (2015), Volatility forecasting using support vector regression and a hybrid genetic algorithm, *Computational Economics*, 45(1), 111-133.
- Smola, A. J. and Schölkopf, B. (2004), A tutorial on support vector regression. *Statistics and computing*, 14(3), 199-222.
- Sun, W. and Zhang, J. (2008), Forecasting day ahead spot electricity prices based on GASVM, In *2008 International Conference on Internet Computing in Science and Engineering*, 73-78.
- Tang, Q. and Gu, D. (2009), Day-ahead electricity prices forecasting using artificial neural networks, In *2009 International Conference on Artificial Intelligence and Computational Intelligence*, 2, 511-514.
- Tat, A. N. (2018), Electricity Price Forecasting Using Monte Carlo Simulation: The Case of Lithuania, *Ekonomika (Economics)*, 97(1), 76-86.
- Tay, F. E. and Cao, L. (2001), Application of support vector machines in financial time series forecasting, *Omega*, 29(4), 309-317.
- Yan, X. and Chowdhury, N. A. (2013), Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach. *International Journal of Electrical Power & Energy Systems*, 53, 20-26.