

Destek Vektör Makineleri ile Sanayi Üretim Endeks Değerinin Tahmin Edilmesi

Muhammed Fatih Yürük*

Öz

Üretim faktörü ülkelerin refah içinde yaşamaları için önemli bir etkidir. Ülkelerin makroekonomik durumlarındaki değişim veya küresel krizler ile birlikte üretim düzeyleri değişime uğramaktadır. Bu çalışmanın amacı ülkemiz için son derece önemli olan sanayi üretim endeks değerinin tahmin edilmesi üzerinedir. Çalışmada sanayi üretim endeks değerini tahmin etmek için makine öğrenmesi metodlarından destek vektör makineleri kullanılmıştır. BİST 100 endeksi, BİST sanayi endeksi, USD, EUR, tüketici fiyat endeksi, yurt içi üretici fiyat endeksi, ticari kredi faiz oranları ve imalat sanayi kapasite kullanım oranlarının Ocak 2007- Ağustos 2021 tarihleri arası aylık verileri kullanılmıştır. Çıktı değerleri MSE, MAE, MAPE, RMSE, CORR ölçütleri kullanılarak performans değerlendirilmesi yapılmıştır. Eğitim ve test seti için MAPE değerleri %7 değerlerini alarak başarılı bir performans göstermiştir.

Anahtar Kelimeler: Yapay Zekâ, Destek Vektör Makineleri, BİST, İmalat Sanayi, Makine Öğrenmesi

* Dr. Öğretim Üyesi, Dicle Üniversitesi, Sivil Havacılık Yüksekokulu,
E-mail: mfyuruk@hotmail.com,
Orcid: <http://orcid.org/0000-0001-7429-2278>
Geliş Tarihi: 15.12.2021 ▪ Kabul Tarihi: 30.12.2021

Using Support Vector Machines to Forecast the Industrial Production Index Values

Muhammed Fatih Yürük*

Abstract

The factor of production is an important factor for countries to live in prosperity. Production levels change with the change in the macroeconomic conditions of the countries or with the global crises. The aim of this study is to estimate the industrial production index value, which is extremely important for our country. In the study, support vector machines from machine learning methods were used to estimate the industrial production index value. Monthly data of BIST 100 index, BIST industry index, USD, EUR, consumer price index, domestic producer price index, commercial credit interest rates and capacity utilization rate of the manufacturing industry between January 2007 and August 2021 are used. Performance evaluation was made using output values MSE, MAE, MAPE, RMSE, CORR criteria. The MAPE values for the training and test set showed a successful performance by taking the values of 7%.

Keywords: Artificial Intelligence, Support Vector Machines, BIST, Manufacturing Industry, Machine Learning

* Assoc. Prof., Dicle University, School of Civil Aviation,
E-mail: mfyuruk@hotmail.com,
Orcid: <http://orcid.org/0000-0001-7429-2278>
Received Date: 15.12.2021 ▪ Accepted Date: 30.12.2021

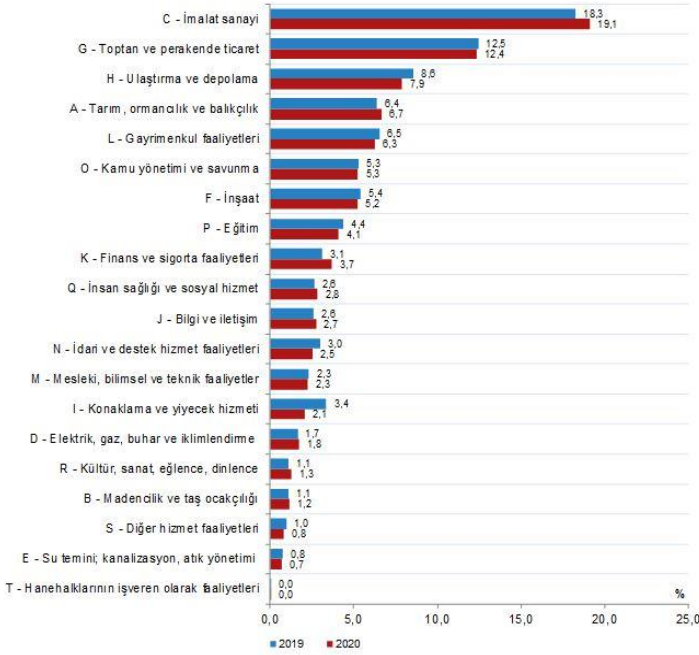
Giriş: Türkiye’de Sanayi Üretim Endeksi

Ekonominin sanayi kesiminde meydana gelen gelişmelerin ve uygulanan ekonomik politikaların, kısa dönemde olumlu veya olumsuz etkilerinin ölçülebilmesi için “Sanayi Üretim Endeksi” hesaplanmaktadır (TÜİK, 2021). Türkiye İstatistik Kurumu (TÜİK) tarafından aylık olarak yayınlanan sanayi üretim endeksi, sanayi sektöründe yer alan işletmelerin üretimlerindeki faaliyetlerin durumunu geçmiş yılları baz alarak karşılaştırmalı olarak takibini sağlayan bir göstergedir.

Sanayi üretim endeksi, sektörün durumunu, sanayi üretimindeki artış ya da azalışın seneler baz alınarak takibini sağlayan bir gösterge olup, Türkiye İstatistik Kurumu (TÜİK) tarafından aylık olarak yayımlanmaktadır. Sanayi üretim endeksi, işyerlerine yapılan anketlerden faydalanarak hesaplanmaktadır. Sanayi üretim endeksi hesaplamasına eklenen sanayi sektöründeki kuruluşlar üç alt sektörde sınıflandırılmış olup, bu alt sektörler ve bunların toplam üretim endeksi içindeki ağırlıkları; 2010 yılı referans endeks değerine göre, imalat sanayi sektörü (%81,51), elektrik, gaz, buhar ve iklimlendirme üretim ve dağıtım alt sektörü (%12,44), madencilik ve taş ocakçılığı sektörü (%6,05) şeklindedir. Belirlenen iş yerlerinde yapılan anket sonuçları ile yukarı belirtilen sektörel ağırlıklar dikkate alınarak aylık olarak sanayi üretim endeksi hesaplanmaktadır (Koç, vd. 2016).

Ülkelerin refah içinde yaşamasının önemli bir etkenei ürettiği mal ve hizmetlerdir. Toplumun ürettiği mal ve hizmeti gösteren bir ölçü gayri safi milli hasıladır (GSMH). GSMH ekonomik kalkınmanın bir ölçüsüdür ve büyüyen GSMH o ülkenin ekonomik istikrarını ve büyüyen pazar potansiyelinin bir göstergesidir (Kar ve Tatlısöz, 2008). GSMH’nin artması ülkelerin ellerindeki kaynaklarını etkili ve verimli kullanması, üretim faktörlerinin işlevlik kazanması ile mümkündür. Üretim faktörlerinin en önemli ayağı ise sanayidir.

Destek Vektör Makineleri ile Sanayi Üretim Endeks Değerinin Tahmin Edilmesi



Şekil 1. Üretim Yöntemine göre GSYH hesabında sektörlerin payları, cari fiyatlarla, 2019, 2020 (TUİK,2021)

Gayri safi yurt içi hasıla (GSYH), 2020 yılında bir önceki yıla göre %16,9 artarak 5 trilyon 46 milyar 883 milyon TL olmuştur. Şekil 1 'de görüldüğü üzere gayrisafi yurt içi hasılda en yüksek payı 2020 yılında %19,1 ile imalat sanayi almıştır. İmalatı, %12,4 ile toptan ve perakende ticaret, motorlu kara taşıtlarının ve motosikletlerin onarımı ve %7,9 ile ulaştırma ve depolama sektörü izlemiştir (TUİK, 2021). İmalat sanayinin bir önceki yılda olduğu gibi 2020 yılında da GSYH'de en yüksek payı aldığı görülmektedir. Tablo 1'de 2020'nin 2. Çeyreğinde salgının etkisi ile bir yavaşlama görülse de 3. ve 4. çeyreklerde yükselişe geçmiştir. Toplamda 5.046.883.307 TL GSYH'nin 1.149.175.754 TL sanayi sektöründen gelmektedir. Sanayi sektörü dolayısıyla üretim milli gelirin artmasında en büyük etkidir.

Tablo 1. Gayri Safi Yurtiçi Hasıla (TL)

Sektörler	2020				Toplam
	I	II	III	IV	
1. Tarım, ormancılık ve balıkçılık	29.864.863	54.732.487	165.546.880	87.015.820	337.160.051
2. Sanayi	250.947.913	222.888.289	304.141.045	371.198.507	1.149.175.754
İmalat Sanayi	210.554.567	184.474.526	256.305.669	314.005.764	965.340.526

İmalat Hariç Sanayi	40.393.345	38.413.763	47.835.377	57.192.742	183.835.228
3. Hizmetler	684.619.454	648.837.608	782.007.648	884.855.098	3.000.319.809
İnşaat	56.654.511	66.553.873	65.783.912	75.906.486	264.898.782
Ticaret, Ulaştırma ve Konaklama	247.605.046	210.533.303	311.636.351	358.123.813	1.127.898.513
Bilgi ve iletişim	27.724.217	30.198.143	33.672.030	47.038.533	138.632.924
Finans ve sigorta faaliyetleri	39.926.476	53.113.201	56.487.397	37.789.964	187.317.038
Gayrimenkul faaliyetleri	76.540.793	75.839.316	79.919.592	83.866.298	316.165.999
Mesleki, İdari ve Destek Hizmet Faaliyetleri	53.675.151	49.330.822	59.469.446	78.43.984	241.019.401
Kamu Yönetimi, Eğitim, İnsan Sağlığı ve Sosyal Hizmet Faaliyetleri	153.417.575	144.262.004	150.367.386	167.406.702	615.453.668
Diğer hizmet faaliyetleri	29.075.685	19.006.946	24.671.536	36.179.317	108.933.483
4. Sektörler Toplamı (1+2+3)	965.432.230	926.458.384	1.251.695.574	1.343.069.425	4.486.655.613
5. Vergi-Sübvansiyonlar	105.791.094	111.233.938	162.187.087	181.015.575	560.227.694
6. Alıcı Fiyatlarıyla GSYH (4+5)	1.071.223.324	1.037.692.322	1.413.882.660	1.52.085.000	5.046.883.307

Kaynak: (SBB, 2021)

Avrupa Topluluğundaki Ekonomik Faaliyetlerin İstatistiksel Sınıflandırması (NACE)'ye göre sanayi sektörü; madencilik ve taş ocakçılığı, imalat ve elektrik, gaz, buhar ve iklimlendirme üretimi ve dağıtım ve su temini; kanalizasyon, atık yönetimi ve iyileştirme faaliyetleri sektörlerinden oluşmaktadır (TUİK, 2021b). Özellikle imalat sanayi, sanayi sektörü içinde önemli bir paya sahiptir. Tablo 1'de görüldüğü üzere GSYH da sanayi sektörünün 250.947.913 TL hasılanın 210.554.567 TL'si imalat sektörüne aittir.

Tablo 2. İmalat sanayinde büyüklük grubu ve teknoloji düzeyine göre temel göstergeler

Büyük- lük grubu	Girişim sayısı	Ciro	Üretim değeri	Faktör maliyetiyle değer	Çalışanlar katmasıyısı değer	
Türkiye	Toplam	409.476	3.051.570.23.048	2.837.643.952.074	644.137.309.822	4.308.302
	Yüksek teknoloji/	2.535	103.299.227.284	102.876.856.955	39.930.645.317	108.408
	Orta-ileri teknoloji/	45.170	812.382.291.818	751.663.087.815	188.949.594.078	797.140
	Orta-düşük teknoloji/	129.840	980.591.902.098	915.102.199.524	191.700.113.626	1.167.946
	Düşük teknoloji/	231.931	1.155.296.809.848	1.068.001.807.780	223.556.956.801	2.234.808
Mikr	Toplam	346.452	127.582.285.721	116.640.478.026	20.363.452.340	784.436
	Yüksek teknoloji/	1.783	1.035.881.359	937.693.752	226.722.713	3.789
	Orta-ileri teknoloji/	33.978	20.857.314.877	19.242.797.164	3.351.377.547	80.754

Destek Vektör Makineleri ile Sanayi Üretim Endeks Değerinin Tahmin Edilmesi

Küçük	Orta-düşük teknoloji/	111.374	44.671.091.762	40.623.146.091	7.158.626.519	242.356
	Düşük teknoloji	199.317	61.017.997.723	55.836.841.019	9.626.725.561	457.537
	Toplam	47.634	356.715.514.467	317.373.639.152	60.569.111.162	830.928
	Yüksek teknoloji	460	4.217.286.931	3.841.569.289	1.213.242.905	8.265
	Orta-ileri teknoloji	8.265	78.640.032.531	70.566.255.789	15.119.779.905	143.140
	Orta-düşük teknoloji	14.146	112.752.213.575	100.162.118.569	19.730.828.949	24.019
	Düşük teknoloji	24.763	161.105.981.430	142.803.69.505	24.505.259.403	438.504
Orta	Toplam	12.223	562.008.789.798	511.124.563.397	117.265.005.757	1.049.598
	Yüksek teknoloji/	199	10.522.031.679	9.900.109.417	3.363.741.700	17.832
	Orta-ileri teknoloji	2.240	117.816.972.844	109.083.726.649	30.676.298.273	182.927
	Orta-düşük teknoloji	3.440	166.413.657.303	150.730.210.401	35.805.697.270	288.204
	Düşük teknoloji	6.344	267.256.127.972	241.410.516.930	47.419.268.514	560.635
KOBİ	Toplam	406.309	1.046.306.589.986	945.138.680.575	198.197.569.259	2.664.962
	Yüksek teknoloji	2.442	15.775.199.969	14.679.372.458	4.803.707.318	29.886
	Orta-ileri teknoloji	44.483	217.314.320.252	198.892.779.602	49.147.455.725	406.821
	Orta-düşük teknoloji	128.960	323.836.962.640	291.515.475.061	62.695.152.738	771.579
	Düşük teknoloji	230.424	489.380.107.125	440.051.053.454	81.551.253.478	1.456.676
Büyük	Toplam	3.167	2.005.263.641.062	1.892.505.271.499	445.939.740.563	1.643.340
	Yüksek teknoloji/	93	87.524.027.315	88.197.484.497	35.126.937.999	78.522
	Orta-ileri teknoloji	687	595.067.971.566	552.770.308.213	139.802.138.353	390.319
	Orta-düşük teknoloji	880	656.754.939.458	623.586.724.463	129.004.960.888	396.367
	Düşük teknoloji	1.507	665.916.702.723	627.950.754.326	142.005.703.323	778.132

Kaynak: (TÜİK, 2021a)

Türkiye’de Küçük ve Orta Ölçekli işletmelerin tanımında değişiklik yapılmasına dair yönetmelik 24.06.2018 tarihinde Resmî Gazete ‘de yayınlanarak yürürlüğe girmiştir. Buna göre, KOBİ, "250 kişiden az yıllık çalışan istihdam eden ve yıllık net satış hasılatı ya da mali bilançosu 125 milyon lirayı aşmayan ve yönetmelikte mikro işletme, küçük işletme ve orta büyüklükteki işletme olarak sınıflandırılan ekonomik birimler" olarak tanımlanmıştır (T.C. Resmî Gazete, 24 Haziran 2018, Numara: 11828). Tablo 2’de belirtildiği gibi, Türkiye’de imalat sanayinde KOBİ ve büyük işletmelerin toplam sayısı 409.476’dir. Toplam 4.308.302 çalışanı ve 3.051.570.231.048 TL cirosu ile önemli bir yer teşkil etmektedir.

Konuyla İlgili Yurt İçinde ve Yurt Dışında Yapılan Çalışmalar

Destek Vektör Makinesinin (DVM) finans alanında kullanımı konusunda yurt içinde ve yurt dışında çeşitli çalışmaların yapıldığı literatürde görülmektedir. Yurt içindeki çalışmalara yakından bakacak olursak; Özdemir vd. (2011) lojistik regresyon ve destek vektör makineleri yöntemleri ile borsa endeksinin getirisinin yönünü tahmin etmişlerdir. Yaptıkları çalışma sonucunda modelleme kümesi ve tahmin

kümesi için doğru sınıflandırma oranlarını lojistik regresyon için yaklaşık %75, destek vektör makineleri için %86 olarak bulmuşlardır. Başka bir çalışmada Karagül (2014) tarafından borsada işlem gören firmaların sınıflandırılması üzerine yapılmıştır. Yazar çalışmada destek vektör makinelerini kullanmıştır. Borsaya kote olan 42 şirketi, 10 finansal oran ile kümelemeye çalışmıştır. Kullanılan oranlar etkisizden en etkili olan faktöre doğru sıra ile modelden elenerek, bu dört faktörden hangilerinin analize alınmasının en güçlü performansı veren DVM modeli elde edilebileceği araştırılmıştır. En güçlü performans gösteren modelin ilk 3 faktöre bağlı olan model olduğu belirlenmiştir. Bu yeni modelde sınıflandırma başarı oranı %97,61 ve destek vektör sayısı 12 olarak belirlenmiştir.

Borsada işlem gören firmalar üzerine makine öğrenmesi metodunu kullanan Yakut vd. (2014) yapay sinir ağları ve destek vektör makinelerini beraber kullanarak BIST endeksi üzerinde etkin bir tahminin yapılıp yapılmayacağını ortaya koymak için analiz yapmışlardır. Analizlerinin sonucunda yapay sinir ağları ve Destek Vektör makineleri yöntemlerinin borsa endeksinin tahmin edilmesinde kullanılabileceğini belirtmişlerdir. Literatürde finans alanında makine öğrenmesini kullanarak borsa endeksi üzerine çalışmaların fazla olduğu görülmektedir. Pabuççu (2019), Yakut vd. yaptıkları çalışmadan farklı olarak üç farklı makine öğrenme algoritması olan yapay sinir ağları, destek vektör makineleri ve naive bayes sınıflandırıcı algoritması kullanılmış ve performansları karşılaştırılmıştır. Çalışmanın sonucunda üç modelin de borsa endeks hareketlerini yakalamada kullanılabilir olduğunu, yapay sinir ağı algoritmasının ise daha iyi bir sınıflandırıcı olduğunu belirtmiştir.

Destek vektör makinelerinin finans alanında işletmelerin mali başarısızlığını tespit eden çalışmaların yapıldığı görülmektedir. Yürük ve Ekşi (2019) yaptıkları çalışmada yapay sinir ağları ve destek vektör makinelerini (DVM) kullanarak BIST imalat sanayi işletmelerinin finansal başarısızlığını tespit etmeye çalışmışlardır. Çalışma sonucunda iki modelinde başarılı bir performans gösterdiği belirtilmiştir. Benzer bir çalışmada Kuzu ve Yakut (2020) tarafından Borsa İstanbul imalat sanayi firmalarının DVM kullanarak başarısızlık tahmini yapmaya çalışmışlardır. Son yıllarda ilgi odağı olan kripto para birimlerinin piyasa değerleri kullanılarak BIST 100 endeksinin hareket yönleri tahmin etmişlerdir. Çalışma sonucunda %52 doğruluk başarısı elde etmişlerdir.

Yurt dışında yapılan çalışmalara yakından bakacak olursak, Pai ve Lin (2005) ARIMA ve DVM modellerini Hibrit bir model oluşturarak beraber kullanmışlardır. Yazarlar iki modelin birleşiminin daha iyi bir performans göstereceğini amaçlayarak çalışmayı yapmışlardır. Çalışma sonucunda hibrit modelin beklediği kadar performans göstermediğini belirtmişlerdir. Hansen vd. (2006) yaptıkları çalışmada her biri farklı zaman serisi özellikleri sergileyen dokuz problem alanı üzerinde za-

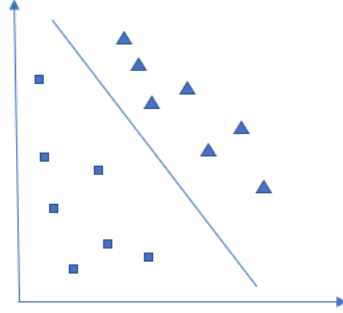
man serisi analizinin karşılaştırmalı sonuçlarının bir çalışmasını sunmaktadır. Yazarlar ARIMA ve DVM metotlarını kullanıp dokuz zaman serisinin performansını ölçmüşlerdir. Çalışmada Mean Absolute Prediction Error (MAPE), Mean Absolute Deviation (MAD) ve Root-Mean-Squared Error (RMSE) performans ölçütleri ile karşılaştırma yapılmış ve DVM çok yakın farkla ARIMA modeline üstünlük sağlamıştır. Ullrich (2009) destek vektör makinesi (DVM) modellerini kullanarak günlük EUR / GBP, EUR / JPY ve EUR / USD döviz kuru dönüş yönlerini tahmin etme ve işlem yapma genel yeteneğini incelemiş ve analiz yapmıştır. Hiperbolik SVM'lerin sürekli olarak iyi performans gösterdiği bulunmuştur. Ayrıca, p-Gauss SVM'leri EUR/GBP ve EUR/USD dönüş yönlerini tahmin etmede oldukça iyi performans gösterdiğini belirtmiştir.

Papadimitriou vd. (2014) DVM tabanlı tahmin modellemesini elektrik fiyatlarının ertesi gün yönlü değişimi için kullanmışlardır. Yazarlar 200 günlük bir dönem analiz etmişlerdir ve analiz sonucunda %76,12 oranında bir tahmin başarı yüzdeliği elde etmişlerdir. Plakandaras vd. (2020) kredi derecelendirme kuruluşlarının belirlediği şekilde, AB bankacılık kurumlarının kredi notlarını tahmin etmeye çalışmışlardır. Çalışmalarında bankaların derecelendirmelerinin %91.07'sini doğru bir şekilde tanımlayan doğrusal olmayan RBF çekirdeği ile birleştirilmiş bir SVM modeli olduğunu belirtmişlerdir. Literatürde özellikle ARIMA modeli ile makine öğrenmesi modellerini karşılaştırarak yapılan çalışmalar görülmektedir. Bunlardan biride Syriopoulos vd. (2021) yaptıkları çalışmadır. Yazarlar gemi fiyatlarının tahmini için ARIMA ve DVM kullanmışlardır. Çalışmalarında performans ölçütü olarak RMSE ve MAPE ölçütlerini kullanmış ve karşılaştırma yapmışlardır. DVM modelinin ARIMA modeline göre çok daha iyi performans gösterdiğini yazmışlardır.

Tasarım, Yöntem ve Uygulama

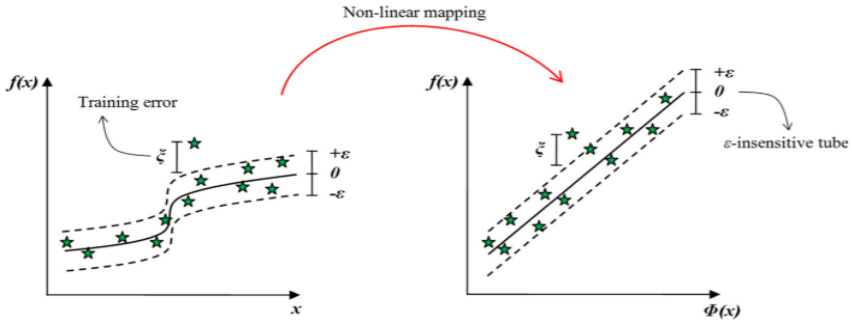
Destek Vektör Makineleri (DVM), veri sınıflandırmasında ve regresyon analizinde kullanılan denetimli bir makine öğrenme yöntemidir (Plakandaras vd., 2020). İlk defa Cortes ve Vapnik (1995) tarafından geliştirilmiştir. Önceleri sınıflandırma problemlerinde kullanılan DVM Smola ve Schölkopf (2004) tarafından regresyon problemlerinde kullanılmak üzere geliştirilmiştir. Destek Vektör Makinelerinde veriler lineer olarak ayrılacakları bir yapıda olabileceği gibi lineer olarak ayrılamayan yapıda olabilir. Lineer ayrılabilme durumunda; $\theta = \{x_i, y_i\}$, $i=1,2,3,4,\dots,N$ elemandan oluşan veri olduğunu varsayalım. Burada $y_i \in \{-1,1\}$ etiket değerleri ve $x_i \in R^d$ özellikler vektörüdür. İki değerli veriler lineer ayrılabilme durumunda direkt olarak aşırı düzlem ile ayrılabilir. DVM 'nin

amacı aşırı düzlemin iki farklı sınıfta bulunan örnek grubuna eş uzaklıkta olmasını sağlamaktır (Yakut vd., 2014).



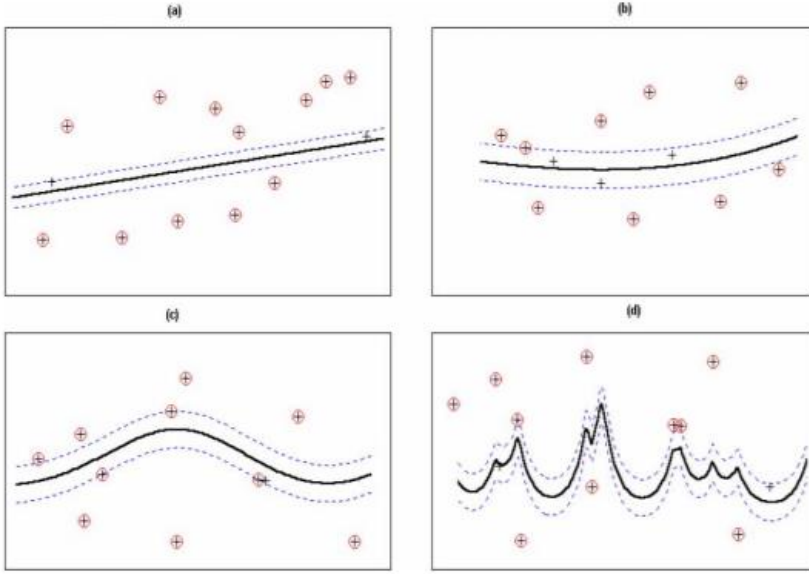
Şekil 2. Lineer ayrılabilen veriler

Verilerin doğrusal olarak ayrılamadığı durumlarda şekil 3’de görüldüğü gibi doğrusal olmayan veriler çekirdek fonksiyonları kullanılarak daha yüksek boyutlara eşlenir (Chanklan vd., 2018). Doğrusal, sigmoid, radyal tabanlı (RBF) ve polinomial DVM’de kullanılan çekirdek fonksiyonlarından bazılarıdır (Şekil 4).



Şekil 3. Doğrusal Olmayan Destek Vektör Regresyonlarını (SVR) Nitelik Uzayına Dönüştürme Kaynak: (Mahdevari vd., 2014)

Şekil 3’de görüldüğü üzere doğrusal olmayan SVR çekirdek fonksiyonları kullanılarak nitelik uzayına dönüştürülmüştür.



Şekil 4. Destek Vektör Regresyon Çekirdek fonksiyonları Kaynak: (Gani vd., 2010)

Şekil 4’de (a) şekli doğrusal, (b) şekli polinomial, (c) şekli gauss RBF, (d) exponential RBF çekirdek fonksiyonlarının kullanımı örnekleri görülmektedir.

Bu çalışmada sanayi üretim endeksi DVM yöntemi ile tahmin edilmeye çalışılmıştır. Ocak 2007- Ağustos 2021 arasındaki aylık veriler kullanılmıştır. Sanayi üretim endeksi tahmini için tablo 3’te görüldüğü üzere BİST 100 endeksi, BİST sanayi endeksi, USD, EUR, tüketici fiyat endeksi, yurt içi üretici fiyat endeksi, ticari kredi faiz oranları ve İmalat sanayi kapasite kullanım oranları olmak üzere 8 bağımsız değişken kullanılmıştır.

Tablo 3. Çalışmada Kullanılan Değişkenler

Sıra	Değişkenler	Veri Türü	Kaynak
1	BİST 100	Aylık	TCMB
2	BİST Sanayi Endeksi	Aylık	TCMB
3	USD	Aylık	TCMB
4	EUR	Aylık	TCMB
5	Tüketici Fiyat Endeksi	Aylık	TCMB
6	Yurt İçi Üretici Fiyat Endeksi	Aylık	TCMB
7	Ticari Kredi Faiz Oranları	Aylık	TCMB
8	İmalat Sanayi Kapasite Kullanım Oranları	Aylık	TCMB

Tahmin modelinin oluşturulmasında Statistica 12 versiyon paket program kullanılmıştır. Verilerin %70'i eğitim %30'ü ise test için ayrılmıştır. Veriler program tarafından rastgele dağıtılmıştır. DVM çekirdek tipi olarak radyal tabanlı çekirdek (RBF) kullanılmıştır. RBF parametreleri cost (C) ve gamma (γ)'nın optimizasyonu için çapraz geçerlilik (cross-validation) yöntemi uygulanmıştır.

DVM'de, doğrusal (linear), polynominal, Radial Basis Function (RBF) ve sigmoid gibi çeşitli çekirdek fonksiyonları mevcuttur. Bu çalışmada kernel fonksiyonu olarak Radial Basis Function (RBF) kullanılmıştır. RBF kernel fonksiyonu doğruluğu önemli derecede etkileyen iki parametreye sahiptir. Bunlar cost (C) ve gamma (γ)'dır. (Yürük ve Ekşi, 2019) Literatürde, bir öğrenme yönteminin hata miktarını tespit etmek için, 10 kere çapraz doğrulama (10-fold cross validation) yöntemi yaygın olarak kullanılmaktadır. Bu çalışmada RBF parametrelerinin tespit edilmesinde çapraz geçerlilik (cross-validation) yöntemi uygulanmıştır. Bu yöntem aynı zamanda aşırı öğrenmenin önüne geçmektedir. Oluşturulan modele ait bilgiler tablo 4'te verilmiştir.

Tablo 4. SVM Modeline Ait Bilgiler

SVM	C: 10,000 epsilon:0,200
Çekirdek türü	(RBF) Radial Basis Function ($\gamma=0,125$)
Destek vektörlerinin sayısı	24

Tabloda görüldüğü üzere tahmin modelinde C: 10 epsilon: 0,20 ve gamma: 0,125 değerleri atanmıştır. Modelin başarı durumuna göre belirtilen parametreler değiştirilerek en iyi performansı veren model seçilecektir. Bu çalışmada en iyi performansı gösteren model parametreleri tablo 4'de belirtildiği gibidir.

Oluşturulan Modelin Genel Performans Değerlendirilmesi

Literatürde; oluşturulan modelin tahmin sonuçlarının performansını değerlendirmek için farklı ölçütler kullanıldığı görülmektedir. Bu çalışmada, en çok kullanılan ölçütlerden Hata Kareleri Ortalaması (MSE), Ortalama Mutlak Yüzde Hata (MAPE), Korelasyon Katsayısı (CORR) ve Determinasyon Katsayısı (R^2) kullanılmıştır. y_i = Gerçek değerleri belirtirken, \hat{y}_i = Tahmin değerlerini belirtmektedir.

Ortalama Hata Kare (MSE):

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} = \frac{\sum \text{tahmin hataları}^2}{n} \quad (1)$$

Ortalama Mutlak Yüzde Hata (MAPE):

$$MAPE = 100 \frac{\sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}}{n} \quad (2)$$

Hata kareler ortalamasının karekökü (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

Tablo 5. Modelin Performans Değerleri

Performans Ölçütü	Eğitim Seti	Test Seti	Tüm Seti
Ortalama Hata Kareleri (MSE)	59,788	75,52	64,52
Hatanın Mutlak Ortalaması (MAE)	6,3806	6,974	6,559
Ortalama Mutlak Yüzde Hata (MAPE)	%7,00	%7,65	%7,71
Standart Sapma Oranı	0,345	0,380	0,356
Hata Kareler Ortalamasının Karekökü (RMSE)	7,732	8,690	5,959
Korelasyon Katsayısı (CORR)	0,9351	0,925	0,935

Destek vektör makinesi modelinin performansını ölçmek için Tablo 5’de verildiği gibi MSE, MAE, MAPE, RMSE, CORR ölçütleri kullanılmıştır. Performans ölçütleri incelendiğinde korelasyon katsayısı eğitim seti 0,93, test seti 0,92 değerini alırken tüm set için 0,93 bulunmuştur. Literatürde tahmin performanslarının ölçümünde ya da karşılaştırılmasında en çok kullanılan ölçütün MAPE olduğu görülmektedir. Literatürde MAPE değerinin %10’un altında olmasının modelin performansının çok iyi olduğunu göstermektedir. Tablo 5’te görüldüğü üzere MAPE değerleri eğitim seti için; (%7), test seti için; (%7,65) değerleri bulunmuştur. Bu sonuçlar modelin performansının çok iyi olduğunu göstermektedir.

Sonuç ve Değerlendirme

Sanayi bir ülkenin kalkınması ve refah seviyesinin yükselmesi için son derece önem arz etmektedir. Konu hakkında geleceğe yönelik tahminler yapmak konunun muhataplarının elini güçlendirecektir. Özellikle sanayi gibi önemli bir konuda tahminlerde bulunmak son derece önemlidir. Karar alıcılar geliştirilen tahmin modelleri ile hedeflerini daha sağlıklı belirleyebileceklerdir.

Makine öğrenmesinin bir uygulaması olan destek vektör makineleri birçok alanda kullanıldığı gibi literatürde finans alanında da artık sık kullanılmaktadır. Bu

çalışmada da destek vektör makineleri kullanılmıştır. Yapay zekanın (YZ) gündemde olduğu bu dönemlerde tahmin modellemeleri artık geleneksel metotlardan ziyade YZ kullanılarak yapılabilmektedir. Makine öğrenmelerinde sisteme girilen verilerin eğitim aşamasında matematiksel bağlantılar kurarak analitik bir öğrenme ile geri kalan verilerin tahmin edilmesi mantığı ile çalışmaktadır. Burada sisteme tanıtılan değişkenler ve verilerin sayısı önemli olmaktadır. Bu çalışmada da 8 bağımsız değişken ile 1 bağımlı değişkenin değerleri tahmin edilmeye çalışılmıştır. Kullanılan değişkenler sanayi üretim endeksinin değişimi ile yakın ilişkili olduğu düşünülen değişkenlerdir. Modelde BİST 100 endeksi, BİST sanayi endeks, USD, EUR, tüketici fiyat endeksi, yurt içi üretici fiyat endeksi, ticari kredi faiz oranları ve İmalat sanayi kapasite kullanım oranlarının Ocak 2007- Ağustos 2021 tarihleri arasındaki 176 aylık veriler kullanılmıştır. Modelin çalışması sonucu üretilen çıktılar çeşitli performans ölçütleri ile değerlendirilmiştir.

Literatürde sanayi üretim endeksinin içeren çalışmaların yapıldığı görülmekte fakat makine öğrenmesi ile sanayi üretim endeksinin tahmini üzerine bir çalışma yapılmadığı görülmektedir. Sanayi üretim endeksinin tahmini konusunda Ozel ve Gunay (2015) basit bir istatistikî yöntem kullanmışlardır. Çalışmalarında kullanılan yöntemin tahmin başarısı, CNBC-e tarafından ekonomistlerle yapılan anketlerden elde edilen piyasa beklentileri ile gerçek zamanlı olarak karşılaştırmışlardır. Yani yazarlar anket sonucu ortaya çıkan piyasa beklenti değeri ile kendi buldukları değerleri karşılaştırmışlardır. Tahmin hatalarının değerlendirilmesinde hata kareler ortalamasının karekökü (Root Mean Squared Error) değerleri karşılaştırmışlardır. Karşılaştırma sonucunda beklenti değerleri ile önerilen yöntemin yakın değerde tahmin hataları yaptığı görülmüştür.

Bu çalışmada ise farklı bağımsız değişkenler kullanılarak gerçek zamanlı açıklanan sanayi üretim endeks değerinin tahmin edilmesine çalışılmıştır. Ozel ve Gunay (2015) gibi RMSE performans ölçütleri yanında MSE, MAE, MAPE ölçütleri de kullanılmıştır. Bu ölçütler incelendiğinde performansların başarılı olduğu ve metodun tahmin yöntemlerinde kullanılabileceği yönündedir. Özellikle literatürde sık kullanılan MAPE ölçütü değerinin %10 değerinin altında çıkması modelin çok iyi olduğunu göstermektedir (Lewis, 1982). Bu çalışmada da MAPE değerleri %10'un altında sonuç üretmiştir. Çalışma endeks değerinin tahmin edilmesi üzerine odaklansa da çalışmada kullanılan bağımsız değişkenlerin endeks tahmini üzerindeki etkisi de dolaylı yönden araştırılmıştır. Çıkan performans sonuçları kullanılan bağımsız değişkenlerin tahminde etkili olduğunu ortaya çıkarmıştır.

Makine öğrenmesi kullanılarak borsa endeks yönü tahmini, hisse senedi değeri tahmini üzerine çalışmaların çoğunlukla yapıldığı görülmektedir. Bu sebeplerden dolayı çalışma literatürdeki açığı kapatabilecektir. Daha sonraki çalışmalarda farklı değişkenler ve farklı YZ metotları ile tahmin çalışmaları yapılabilir.

Kaynakça

- Chanklan, R., Kaoungku, N., Suksut, K., Kerdprasop, K. ve Kerdprasop, N. (2018). Runoff prediction with a combined artificial neural network and support vector regression. *International Journal of Machine Learning and Computing*, 8, 39-43. doi: 10.18178/ijmlc.2018.8.1.660
- Cortes, C. ve Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. doi: 10.1007/BF00994018
- Gani, W., Taleb, H. ve Limam, M. (2010). Support vector regression based residual control charts. *Journal of Applied Statistics*, 37, 309-324. doi: 10.1080/02664760903002667
- Hansen, J. V., McDonald, J. ve Nelson, R. D. (2006). Some evidence on forecasting time-series with support vector machines. *Journal of the Operational Research Society*, 57(9), 1053-1063.
- Kar, M. ve Tatlısöz, F. (2008). Türkiye'de doğrudan yabancı sermaye hareketlerini belirleyen faktörlerin ekonometrik analizi. *Karamanoğlu Mehmetbey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi*, 2008(1), 436-458.
- Koç, E., Kaya, K. ve Şenel, M. C. (2016). Türkiye'de sanayi sektörü ve temel sanayi göstergeleri-sanayi üretim endeksi. *Mühendis ve Makina*, 57(682), 42-53.
- Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann.
- Mahdevari, S., Shahriar, K., Yagiz, S. ve Akbarpour Shirazi, M. (2014). A support vector regression model for predicting tunnel boring machine penetration rates. *International Journal of Rock Mechanics and Mining Sciences*, 72, 214-229. doi: 10.1016/j.ijrmms.2014.09.012
- Ozel, O. ve Gunay, M. (2015). Sanayi üretim endeksi yıllık değişiminin tahminine istatistiksel yaklaşım. Research and Monetary Policy Department, Central Bank of the Republic of Turkey.
- Pai, P.-F. ve Lin, C.-S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33(6), 497-505.
- Papadimitriou, T., Gogas, P. ve Stathakis, E. (2014). Forecasting energy markets using support vector machines. *Energy Economics*, 44, 135-142. doi: <https://doi.org/10.1016/j.eneco.2014.03.017>
- Plakandaras, V., Gogas, P., Papadimitriou, T., Doumpa, E. ve Stefanidou, M. (2020). Forecasting credit ratings of EU banks. *IJFS*, 8(3), 1-15.
- SBB. (2021). SBB (Türkiye Cumhuriyeti Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı). 12 Ekim 2021 tarihinde <https://www.sbb.gov.tr/temel-ekonomik-gostergerler/#1594716589132-d3a64e97-2238> adresinden erişilmiştir.
- Smola, A. J. ve Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, 14(3), 199-222.
- Syriopoulos, T., Tsatsaronis, M. ve Karamanos, I. (2021). Support Vector Machine Algorithms: An Application to Ship Price Forecasting. *Computational Economics*, 57(1), 55-87.

- TUİK. (2021a). İmalat sanayinde büyüklük grubu ve teknoloji düzeyine göre temel göstergeler. <https://data.tuik.gov.tr/Bulten/Index?p=Kucuk-ve-Orta-Buyuklukteki-Girisim-Istatistikleri-2020-41129>. (01.10.2021)
- TUİK. (2021b). Sanayi İstatistikleri. <https://www.resmiistatistik.gov.tr/detail/subject/sanayi-istatistikleri/>. (11.10.2021)
- TUİK, 2021. (2021). <https://data.tuik.gov.tr/Bulten/Index?p=Yillik-Gayrisafi-Yurt-Ici-Hasila-2020-37184>. (10.10.2021)
- Ullrich, C. (2009). Exchange rate forecasting with support vector machines (ss. 183-184), Springer.
- Yakut, Y. B. E. T. Y., Yakut, E. ve Yavuz, S. (2014). Yapay sinir ağları ve destek vektör makineleri yöntemleriyle borsa endeksi tahmini. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 19(1), 139-157.
- Yürük, M. F., ve Ekşi, İ. H. (2019). Yapay zeka yöntemleri ile işletmelerin finansal başarısızlığının tahmin edilmesi: bist imalat sektörü uygulaması. *Mukaddime*, 10(1), 393-422.