



Türkiye'de Limanlarda Yük Taşımacılığı Veri Analizi

Hasan Şen, Arzum Yaşayanlar, Berrin Denizhan

Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Sakarya, Türkiye, (ORCID: 0000-0003-0785-2778), hasansen387@gmail.com
Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Sakarya, Türkiye (ORCID: 0000-0001-6771-6626), arzumyasayanlar@hotmail.com
Sakarya Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, Sakarya, Türkiye (ORCID: 0000-0002-0212-0087), denizhan@sakarya.edu.tr

(İlk Geliş Tarihi Aralık 2020 ve Kabul Tarihi Ocak 2021)

(DOI: 10.31590/ejosat.851537)

ATIF/REFERENCE: Şen, H., Yaşayanlar, A. & Denizhan, B. (2021). Türkiye'de Limanlarda Yük Taşımacılığı Veri Analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (22), 436-444.

Öz

Gelişen bilişim teknolojisi ile birlikte veriler içerisinden anlamlı bilgileri ortaya çıkarmak ve stratejik kararlarda bir adım öne geçebilmek veri analitiğini her alanda gerekli hale getirmiştir. Bu alanlardan biri de ülkemizin ticaret yükünü taşıyan deniz limanlarına ait yük taşımacılığı verileridir. Limanlar ticaretin en önemli aracıdır. Rekabet gücünü artırmak için ülkemizin ihracat hedefini artırma önceliğinin olduğu günümüzde, lojistik öngörü, politika ve analizler, geçmiş taşımacılık verilerimizin analizini de içermelidir. Bu çalışmada, T.C. Ulaştırma ve Altyapı Bakanlığı'ndan alınan Türk limanlarındaki son 15 yıla ait yük istatistik verileri üzerine veri analizi uygulanmıştır. Çalışma aynı zamanda limanlardaki yük taşımacılığına ait veriler üzerine yapılmış ilk uygulama niteliğindedir. Lineer regresyon ve yapay sinir ağları yöntemleri, sırası ile Weka üzerinde ve Phyton dilinde yazılarak uygulanmış, sonuçlar karşılaştırılmıştır. Geçmiş veriler üzerinde yapılan tahminlerin doğruluk oranları kıyaslanmıştır. Elde edilen sonuçlara göre yapay sinir ağları yönteminin gerçeğe daha yakın sonuçlar verdiği görülmüştür. Çalışma sonucunda limanlarımız ile yapılan ticarete geçmiş yük taşımacılığına ait verilerin kullanılması, yapay sinir ağları gibi tekniklerle gelecekte yapılacak ihracat ve ithalat karar ve politikaları için önemli bir yöntem ve bilgi kaynağını literatüre sunulmaktadır. Bu bilgileri doğru şekilde kullanmak, ülkeleri rekabette daha avantajlı konuma getirebilecektir. Aynı zamanda yük taşımacılığında veri madenciliği uygulamalarının yaygınlaşması, yapılacak ticari faaliyetler için analitik öngörülerde bulunma fırsatı sunabilecektir.

Anahtar Kelimeler: Yük İstatistik Verileri, Veri Madenciliği, Yapay Sinir Ağları, Lineer Regresyon, Zaman Serisi Analizi.

Data Analysis of The Port of Freight Transportation In Turkey

Abstract

Revealing meaningful information from the data and being able to take a step forward in strategic decisions made data analytics necessary in every field of developing information technology. One of these areas is the freight transport data of the seaports that carry the commercial cargo of our country. Ports are the most important tool of the trade. In today's world where our country has priority to increase its export target to increase its competitiveness, logistics foresight, policy and analysis should also include the analysis of our past transportation data. In this study, T.C. Data analysis has been applied to the freight statistics data of the last 15 years in Turkish ports taken from the Ministry of Transport and Infrastructure. The study is also the first application made on the data of freight transportation in ports. Linear regression and artificial neural network methods were applied on Weka and written in Phyton language, respectively, and the results were compared. The accuracy rates of the predictions made on the historical data were compared. According to the results obtained, it has been seen that the artificial neural network method gives more realistic results. As a result of the study, an important method and source of information are presented to the literature for future export and import decisions and policies with techniques such as the use of data of past freight transportation in trade with our ports and artificial neural networks. Using this information correctly will bring countries to a more advantageous position in the competition. At the same time, the widespread use of data mining applications in freight transportation will provide an opportunity to make analytical predictions for commercial activities.

Keywords: Load Statistics Data, Data Mining, Artificial Neural Networks, Linear Regression, Time Series Analysis.

* Sorumlu Yazar: denizhan@sakarya.edu.tr

1. Giriş

Taşımacılık, lojistik kavramı içerisinde çok önemli bir yer tutmaktadır. Bunun sebebi taşımacılık maliyetlerinin lojistik maliyetleri içerisinde yüksek orana sahip olmasıyla beraber taşımacılık kavramının lojistikle özdeşleşmiş olmasıdır. Taşımacılık kavramı, üretimden tüketime kadar, belirli yer ve zamanlarda gerekli olan mal veya hizmetlerin, talebe uygun olarak, belirli bir yerden hedeflenen başka bir alana gönderilmesi işlemidir. (Demir, 2005 ; Kokotos ve Linardatos ,2010). Taşınan mal veya hizmetler çeşitlilik göstermekle birlikte temelde hammadde, yarı mamul veya bitmiş ürün olarak genel bir sınıflandırılma yapılabilir. Sanayi devrimi ile, üretilen çok miktardaki ürünü başka ülkelerin tüketicileriyle buluşturma isteği ve rekabet ortamı, işletmelerin taşımacılık faaliyetlerine eğilmelerine sebep olmuştur (Kılıcı, 2017 ; Jingjing ve diğerleri, 2018). Günümüzde ise rekabet ortamında bir adım öne geçmenin en temel faktörlerinden birisi, veri madenciliği kavramını lojistik sektörüne entegre edebilmektir.

Uluslararası rekabet ortamında, taşınmanın daha hızlı ve daha ekonomik olması zorunlu hale gelmiştir (Nataraj vds, 2020). Bu noktada taşımacılık kavramının temel hedefi, taşınması planlanan yükleri veya gerçekleştirilecek hizmeti, en uygun taşıma sistemiyle birlikte belirlenen başlangıç noktasından varış noktasına kadar, minimum maliyet, kısa zaman ve güvenli şekilde ulaştırmak olmuştur (Lambert ve James, 1999). Taşıyıcı türüne göre başlıca ulaştırma sistemleri; denizyolu, karayolu, demiryolu, havayolu ve diğer taşıma türleri olarak sınıflandırılmaktadır. Bu çalışmada, Türkiye’de denizyolu taşımacılığı üzerine bir veri analizi ele alınmıştır. Denizyolu aracılığı ile taşınan malzemelerin, gemiler vasıtası ile bir limandan diğer limana taşınma işlemine deniz yolu taşımacılığı denir. Günümüzde uluslararası ticarete, denizyolu ülkelere, büyük miktarlarda yüklerin en düşük maliyetle taşınması denizyolu taşımacılığı ile gerçekleşmektedir. Bundan dolayıdır ki dünya genelinde yaklaşık olarak ticaretin %80’i deniz taşımacılığı ile sağlanmaktadır. (Kalaycı, 2014).

Büyük miktarlardaki yüklerin tek seferde taşınabilmesine imkan tanınması, taşıma maliyetlerinin hava yoluyla yapılan taşımalara göre 14 kat, karayolu taşımacılığına göre 7 kat, demiryolu ile taşımaya göre ise 3.5 kat ucuz olması, daha çevreci olması, harcanılan enerjinin taşınan yüklere oranla maliyetinin daha az olması ve oluşacak kaza riskinin en alt seviyede olması gibi sebeplerden denizyolu taşımacılığı, taşıma türleri arasında ilk olarak tercih edilen taşıma yöntemi olmaktadır. (Arslan, 2011). Denizyolu ile yapılacak olan taşımacılık da, gidilen limanları ve bu limanlara taşınan yükleri ifade etmektedir. (Aksoy, 2019).

WorldBank’ın 2018 yılında yayımladığı Lojistik Performans Index verilerinde, ulaştırma ve ticaret ile ilgili alt yapının kalitesi gözönüne alınarak yapılan sıralamaya göre, Türkiye 3.21 puan ile 160 ülke arasından 33. sırada yer almaktadır. LPI oran ile birlikte yapılan sıralama sonucunda ise, Türkiye 47. sırada yer almıştır (www.worldbank.org).

WorldBank’ın yayımlanmış olduğu sonuçlara göre, Türkiye’nin, ulaştırma ve ticaret ile ilgili alt yapının kalitesi ile ilgili yapılan sıralamada 33. sırada yer almış olması, Türkiye’nin bulunduğu bölgede daha güçlü yani işlem ve ihracat hacmi daha büyük bir lojistik merkez olabilmesi için daha da gelişmesi gerektiğini göstermektedir (Gün, 2013). Bu gelişmeler; altyapı, dijitalleşme, gümrük mevzuatına yönelik düzenlemeler, operasyonel etkinliği artıracak makine ve ekipmanlara yatırımlar

ve nitelikli bilimsel araştırmalar ve iyileştirmelerin artırılması şeklinde olacaktır. Bunlarla birlikte, Türkiye gelecekte lojistik sektöründe hayat bulabilecek ve var olabilecek 10 ülke arasında kendisine 4. Sıraya çıkabilecek yetkinliğe sahiptir.

T.C. Ulaştırma ve Altyapı Bakanlığı’nın verilerine göre Türkiye’de toplam 71 adet liman başkanlıkları mevcuttur. 2004 yılına ait veriler incelendiğinde, 267 farklı ürün kalemi ihracat ve ithalat sürecinde yer almaktadır. 2019 yılı sonunda yayımlanan veriler tekrar incelendiğinde liman başkanlıklarından yapılan ihracat ve ithalat çalışmalarında, 560 farklı ürün kalemi vardır. Bu demek oluyor ki ihracat ve ithalatta 15 yıllık zaman diliminde 293 farklı ürün kalemi ortaya çıkmıştır.

1.1. Deniz Taşımacılığı ve Veri Madenciliği

Teknoloji kullanımının yaygınlaşması, veri ve bilgi akışlarının büyük boyutlara ulaşmasını ve hatta büyük verilerin ortaya çıkmasını sağlamıştır (Sousa ve Campos, 2018). Kuruluşların dünya üzerinde artan bu veri yığınlarından kendilerine faydalı olacak bilgileri ve ilişkileri ortaya çıkarma ihtiyaçları küresel rekabette kendileri için de önemli bir avantaj durumuna gelmiştir. Buna ek olarak bilginin güvenilirliği, doğru yöntemler kullanılarak doğru bilginin ortaya çıkması günümüzde oldukça önem kazanmıştır (Akçetin ve diğerleri, 2013).

Kurumlar etkili kararlar verirken, mevcut kullanılmamış veri yığınlarını en uygun şekilde işleyerek, yeni bilgi kaynağını ortaya çıkarıp bunlardan stratejik olarak faydalanmayı amaçlarlar (Jose ve diğerleri 2017). Bu gibi amaçları gerçekleştirebilmek ve hayata geçirebilmek için yapay zekanın alt dalı olan veri madenciliği yöntemine başvururlar. Veri madenciliği yönteminin kullanılmasıyla, istatistiksel analizler ve çeşitli hesaplamalar yapılabilir (Cornejo-Bueno ve diğerleri, 2018). Kurumların; kuruluş yeri seçiminde, çizelgeleme faaliyetlerinde, geleceğe yönelik tahminlerinde, araç ve rotalama problemlerinde veri madenciliği kullanılmaktadır.

Limanlar, deniz taşımacılığı ve lojistik üzerinde de veri madenciliği uygulamaları literatürde yer almaktadır. Wang (2012), veri madenciliği yöntemini kullanarak limandan limana yük hareketi için Amerika Birleşik Devletleri ve diğer ülkeler arasında bir veritabanı oluşturmuştur. Alnıpak, v.d. (2007) , Otomatik Tanıma Sistemleri olarak bilinen RFID teknolojisinin konteyner limanlarında kullanılması ve büyük veri analizi ile ilgili bir çalışma yapmıştır. Çalışma sonucunda departmanlar arasındaki veri akışının otomatikleşmesi ve veriminin artması, israfa azalma, süreç içindeki bütün insan , araç ve yüklerin etkin bir şekilde takip edilebilmesi ve limanların ve konteynerların taşıma sürecindeki güvenliğindeki artış, bekleme sürelerinin azalması , insan hatalarının çok az miktarlara indirilmesi , trafikte azalış, gümrük sürelerinde kısalma, konteyner depo yönetiminde verimlilik artışı, kayıp veya çalıntı malzeme azalışı, işlem sürelerinde azalmalar hedeflenmiştir.

Birant (2006), deniz verilerinin modellenmesi ve analizi üzerine yaptığı çalışma kapsamında konumsal-zamansal verilerin kümelenmesi ve konumsal-zamansal sıra dışı verilerin tespiti gibi iki farklı veri madenciliği algoritması tanıtılmıştır. Çözüm aracı olarak Mardex (Marine Data Explorer) programı kullanılmıştır. Çalışmada ayrıca web servis tabanlı paralel kümeleme adlı yeni bir yaklaşım önerilmiştir.

Kokotos ve Linardatos (2010), ise 1995-2006 yılları arasında Yunanistan’da gerçekleşen gemi kazalarını araştırmak için yapılan bir çalışma gerçekleştirmiştir. CHAID algoritmasını

kullanarak , gerçekleşen kazaların insan kaynaklı kaza veya insan kaynaklı olmayan kaza olarak tahmin edilmesi sağlanmıştır. Kiremitçi (2005), lojistik alanında firmaların veri depoladıkları veriden firmanın karar verme süreçlerinde kullanabilecekleri bilgiyi ortaya çıkarmak için veri madenciliği çalışması yapmıştır. Kümeleme Analizi ve Karar Ağaçları kullanılarak gerçekleştirilen çalışmada, kurulan modelin sonucunda ortaya çıkacak bilginin anlaşılabilirliği dikkate alınmıştır. Akçetin ve diğerleri (2013), lojistik sektörü açısından veri madenciliği kavramının öneminden ve birçok alanda uygulamasından söz etmiştir. Türkiye için deniz lojistiğinde yapılacak veri madenciliği çalışmalarının önemi belirtilmiştir.

Dünya çapında kurumların gerçekleştirdiği ticarete, lojistik faaliyetleri büyük önem arz etmektedir. Global ticareti katlanabilir kılan kritik lojistik faaliyeti ise tartışmasız taşımacılıktır. Deniz yolu, belirtilen taşımacılık türleri arasında düşük maliyet ve güvenlik açısından çok önemli bir yere sahiptir (Şenyiğit ve Arsav, 2019). Gün geçtikçe değeri artan veri madenciliği uygulamaları her sektörde uygulanabildiği gibi denizyolu taşımacılığında da kendine yer bulmuştur. Her sektöre farklı tür getiriler sunan veri madenciliği uygulamaları, denizyolu taşımacılığı açısından değerlendirilecek olursa;

- Limanlarda verimlilik analizleri yapılabilir.
- Daha doğru lojistik politikaları belirlenebilir.
- Müşteriye göre tespitler yapılabilir.
- Tedarikçiye göre tespitler yapılabilir.
- Riskin azaltılması ve minimuma indirilmesi için risk analizleri yapılabilir.
- Yük akışının hızlandırılması için yük akış hızının analizi yapılabilir.
- Limanlardaki yük verilerinin yıllara göre tahmin ve analizi yapılabilir (Akçetin ve diğerleri., 2013).

Bu çalışmada T.C Ulaştırma ve Altyapı Bakanlığı Denizcilik İstatistikleri'ne ait 16 yıllık (2004-2019) yük istatistik verileri kullanılmıştır. Veriler, aylar bazında Türkiye limanlarında gerçekleştirilen toplam elleçleme (ton bazında) istatistiklerini içermektedir. (Ulaştırma Bakanlığı, 2019)

Elleçleme; temelde iki ana faaliyetten oluşmaktadır. Bunlar; yükleme (ihracat) ve boşaltma (ithalat)dır. İhracat ve ithalat verileri kendi içerisinde Türk Bayraklı, Kendi Ülke Bayraklı ve Diğer Ülke Bayraklı olmak üzere sınıflandırılmıştır. Yapılacak çalışmada, 2019 yılı içerisindeki ihracat ve ithalat miktarlarının aylar bazında veri madenciliği yöntemleri ile tahmini yapılacaktır. Tahmin için Wekaforecast de yer alan Lineer Regresyon ve Yapay Sinir Ağları algoritmaları kullanılacaktır. Çalışma sonunda kullanılan yöntemlerden hangisinin daha iyi sonuç vereceği belirlenecektir..

2. Materyal ve Metot

Problemin çözümünde Weka ve Python dili ile gelecek değer tahmini için ayrı ayrı hesaplamalar yapılmıştır. Bu yöntemlerin problemin çözümünde takip edilecek aşamalar aşağıda gösterilmiştir;

1. Weka lineer regresyon yöntemi ve yapay sinir ağları yöntemi ile tahmin yapılması,

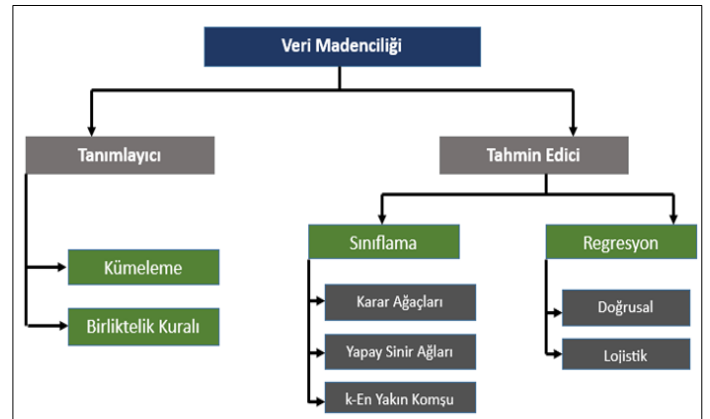
2. Weka lineer regresyon yöntemi ve yapay sinir ağları yönteminin sonuçlarının karşılaştırılması,
3. Python ile lineer regresyon yöntemi ve yapay sinir ağları uygulanarak tahmin yapılması,
4. Python lineer regresyon yöntemi ve yapay sinir ağları yönteminin sonuçlarının karşılaştırılması,
5. Elde edilen verilerinin istatistiksel analizi,
6. Weka ve Python kullanılarak 2020 Yılı İhracat ve İthalat Tahmini.

Veri Madenciliği, büyük miktarlardaki veriler içinden 'anamlı' bir bilgi etme işlemidir. Veri madenciliği sayesinde veriler arasındaki ilişkileri ortaya çıkarmak ve bu veriler ile ileriye dönük tahminlerde bulunarak yorumlayabilmek mümkün hale gelmektedir. Veri madenciliği uygulamaları, işletme, finans, sağlık, pazarlama, bankacılık, sigortacılık, borsa ve internet gibi birçok alanda kendisine yer bulmuştur (Özkan, 2016), (Khair ve diğerleri 2017).

Veri madenciliği birtakım süreçlerden oluşmaktadır. Bu süreçler veriyi;

- Temizleme
- Bütünleştirme
- İndirgeme
- Dönüştürme

Son olarak da veri madenciliği algoritmasını uygulama ve sonuçları sunum ve değerlendirme olarak belirtilebilir (Özkan, 2016). Veri madenciliğine ait yöntemlerin sınıflandırılması Şekil 1.'de gösterilmiştir.



Şekil 1. Veri madenciliğinde kullanılan model ve yöntemler (Arabacı, 2007)

3. Araştırma Sonuçları ve Tartışma

3.1. Weka Lineer Regresyon Yöntemi İle Tahmin Uygulaması

T.C Ulaştırma ve Altyapı Bakanlığı'ndan alınan Denizcilik istatistikleri'ne ait 16 yıllık (2004-2019) yıllık veriler Weka programı için belirli düzenlemelere tabi tutulmuştur. Aylar bazında düzenlenen yük istatistik verilerinden alınan bir kısım örnek aşağıdaki Tablo 1.'de yer almaktadır.

Problem çözüm aşamasında ilk olarak, ayrı ayrı yıllara göre bulunan veriler excel dosyasında birleştirilerek 166 yıllık veri setinin tek bir dosyada toplanması sağlanmıştır. Ardından verilerin weka programına uygun hale getirilmesi sağlanmıştır.

Weka programı ile analiz için uygun hale getirilen veriler Weka'nın forecast özelliği kullanılarak tahmin edilmiştir.

Wekaforecast, Weka programının en gelişmiş özelliklerinden biridir.

Weka programında lineer regresyon yöntemi ve yapay sinir ağları yöntemi ile 2019 yılına ait verilerin tahmini gerçekleştirilecek ve gerçek sonuçları ile karşılaştırılacaktır. Weka’da, lineer regresyon fonksiyonu gerçek değerinin, tahmin edilmek istenen değişkenin değerine göre yapılan bir işlemdir. Lineer regresyonun arka planında ya arkadan ekleme kullanılarak ya da tüm niteliklerden tam bir model oluşturularak ve bir durdurma

ölçütü elde edilene kadar standartlaştırılmış katsayıların sırası düşünülerek tahmin hesaplamaları yapılır (Aydemir, 2019).

Tahminlemeye uygun hale getirilen veriler üzerinde %95 eğitim, %5 test oranında verilerin eğitimi ve testi gerçekleştirilmiştir. 2004-2018 yıllarına ait veriler gelecek tahminleri için eğitim verisi olarak kullanılmış, 2019 yılına ait veri ise test verisi olarak bir sonraki yılı belirlemek üzere kullanılmıştır.

Tablo 1. 2010 Yılı Yük İstatistik Veril

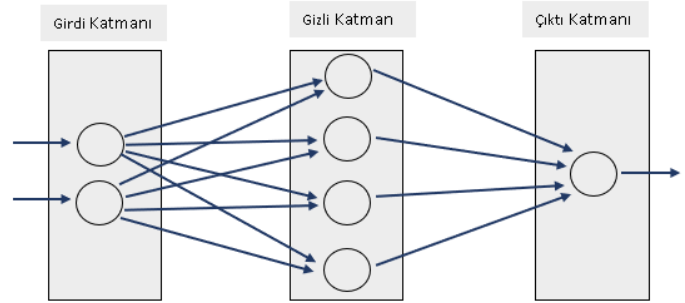
Dönem	İhracat Türk Bayraklı	İhracat Kendi Bayraklı	İhracat Diğer Ülke Bayraklı	İthalat Türk Bayraklı	İthalat Kendi Bayraklı	İthalat Diğer Ülke Bayraklı
28-01-2010	863.553	325.348	4.583.653	2.186.211	678.061	8.698.487
28-02-2010	825.541	214.151	4.881.417	1.723.264	654.582	8.028.930
28-03-2010	978.869	460.396	5.680.063	2.503.276	848.090	10.264.192
28-04-2010	979.920	290.895	6.009.028	2.453.381	792.318	9.654.404
28-05-2010	998.656	540.611	6.576.840	2.638.125	756.710	10.924.954
28-06-2010	994.732	305.703	5.732.826	2.726.211	741.234	10.048.995
28-07-2010	996.267	372.731	6.390.617	2.721.131	830.950	11.841.482
28-08-2010	882.450	404.101	5.378.268	2.274.118	834.286	12.109.388
28-09-2010	1.012.933	234.003	5.703.279	2.498.819	675.708	9.665.718
28-10-2010	957.219	239.426	5.941.399	2.461.653	564.552	11.949.475
28-11-2010	1.026.063	327.251	5.225.908	2.506.592	717.866	10.389.037
28-12-2010	1.098.715	269.661	6.230.895	2.179.438	954.018	11.093.072

3.2. Weka Yapay Sinir Ağları Yöntemi İle Tahmin Uygulaması

Yapay sinir ağları, insan beyninin özelliklerinden çalışma ve öğrenme ilkelerinden esinlenerek yeni bilgiler türetebilme, yeni bilgiler ortaya çıkarabilme ve keşfedebilme gibi yetenekleri yardımsız bir şekilde bilgisayarlar üzerinde taklit ederek uygulayan bir yöntemdir.

Yapay sinir ağları, sınıflandırma, olasılıksal fonksiyon kestirimi, örüntü işleme, örüntü tanıma, optimizasyon, zaman serisi analizleri, tahmin ve sinyal işleme gibi farklı alanlarda kullanılmaktadır. Yapay sinir ağlarının asıl görevi, bilgisayarların öğrenmesini sağlamaktır. Bu çalışmayla beraber bilgisayar öğrenme yetisi kazanmış olur ve karar verme süreçlerinde olayları yorumlayabilme olanakları kazanır.

Sınıflandırma, tanıma, tanımlama ve sinyal işleme gibi birçok problemin çözümünde İleri Beslemeli Yapay Sinir Ağları modeli kullanılmaktadır. Bu modelde bir katmandaki sinirin çıktıları kendisinden bir sonraki katmana ağırlıklar üzerinden girdi olarak gönderilmektedir. Modelin yapısı aşağıda yer alan Şekil 2.’de gösterilmiştir (Aydemir, 2019).



Şekil 2. İleri Beslemeli Yapay Sinir Ağlar

Yapay Sinir Ağları yönteminde tahminlemeye uygun hale getirilen veriler üzerinde %95 eğitim, %5 test oranında verilerin eğitimi ve testi gerçekleştirilmiştir. Bu kez 2019 yılı tahminleri Yapay Sinir Ağları ile çözümlenmiştir.

Problemin 2019 yılı veri tahminleri WekaForecast üzerinde hesaplanmıştır. Elde edilen sonuçlar Tablo2.’de karşılaştırılmıştır. Altı farklı kategoride metodların verdiği doğruluk oranları hangi yöntemin daha gerçeğe yakın tahmin edildiğini açıkça göstermektedir. Lineer Regresyonda ortalama doğruluk oranı %81,75 iken Yapay Sinir Ağlarında ise bu oran %69,91 çıkmıştır. Dolayısıyla problemimizde yer alan veri seti için Weka Programında Lineer Regresyon yöntemi ile tahmin yapmak, Yapay Sinir Ağları yöntemine göre daha uygundur.

Tablo 2. Weka Linear Regresyon Yöntemi ile Yapay Sinir Ağları Yöntemi Karşılaştırılması

İthalat-İhracat Oranları	Lineer Regresyon Yöntemi İle Doğruluk Oranı	Yapay Sinir Ağları Yöntemi İle Doğruluk Oranı
İhracat Türk Bayraklı	0,8206	0,8136
İhracat Kendi Ülke Bayraklı	0,8893	0,6745
İhracat Diğer Ülke Bayraklı	0,8415	0,7870
İthalat Türk Bayraklı	0,6311	0,5003
İthalat Kendi Ülke Bayraklı	0,7826	0,6347
İthalat Diğer Ülke Bayraklı	0,9396	0,7846

3.3. Python Linear Regresyon ve Yapay Sinir Ağları Yöntemi İle Tahmin Uygulaması

T.C Ulaştırma ve Altyapı Bakanlığı'ndan alınan Denizcilik İstatistikleri'ne ait, Weka programı için altı başlıkta kullanılan 16 yıllık (2004-2019) yıllık veriler python programı için ihracat ve ithalat verileri olmak üzere iki başlık altında toplanmıştır. Tablo 3.'de 180 adet verinin, ilk 24 verilik parçası verilmiştir

Tablo 3. Bir Yıllık İhracat ve İthalat Verileri

Dönem(Ay)	İHRACAT(ton)	İTHALAT(ton)
1	3.869.790	9.333.088
2	3.676.000	9.639.313
3	5.220.926	10.609.523
4	5.409.012	9.412.340
5	5.555.230	10.297.391
6	4.822.120	10.229.015
7	4.770.913	10.362.404
8	4.845.585	10.411.391
9	4.706.893	9.908.589
10	5.051.356	10.617.950
11	4.616.468	9.613.938
12	4.800.149	10.281.634

Python programı için tahminlemeye uygun hale getirilen veriler üzerinde %75 eğitim, %25 test oranında verilerin eğitimi ve testi gerçekleştirilmiştir. Test verisininin %25 seçilmesi sebebiyle, Python programında 16 yıllık verilerden 2016, 2017, 2018 ve 2019 yılına ait verileri test verisi olarak kullanılacaktır.

Girilen gerçek değerlerin eğitilmesi ve test edilmesiyle yıllara ait verilerin ait tahmini Python dili ile programlanarak hesaplanmıştır. Hesaplamalar; Weka programında ana başlıklar olarak Türk bayraklı ihracat, kendi ülke bayraklı ihracat, diğer

ülke bayraklı ihracat, türk bayraklı ithalat, kendi ülke bayraklı ithalat ve diğer ülke bayraklı ithalat olmak üzere ayrı ayrı hesaplanması yerine, İhracat ve İthalat başlıkları altında gerçekleştirilmiştir. Ardından Weka'da olduğu gibi Python'da da gerçek değerle tahmin edilen değerler arasındaki ilişki tutarlılığının ölçülmesi için mutlak hata oranları belirlenmiştir.

Yapay Sinir Ağları yönteminde tahminlemeye uygun hale getirilen veriler üzerinde %75 eğitim, %25 test oranında verilerin eğitimi ve testi gerçekleştirilmiştir.

Tablo 4. Python Linear Regresyon Yöntemi ve Yapay Sinir Ağları Yöntemi Karşılaştırılması

İthalat-İhracat Oranları	Lineer Regresyon Yöntemi İle Doğruluk Oranı	Yapay Sinir Ağları Yöntemi İle Doğruluk Oranı
İhracat	0,6181	0,9094
İthalat	0,8424	0,9157

2019 yılı veri tahminleri Python programı üzerinde gerçekleştirilmiştir. Hesaplama ile elde edilen değerler Tablo 4.'de karşılaştırılmıştır. İki farklı kategoride İhracat ve İthalat başlıkları adı altında karşılaştırmalar yapılmıştır.

Lineer Regresyonda ortalama doğruluk oranı %73,025 iken Yapay Sinir Ağlarında ise bu oran %91,25 çıkmıştır. Dolayısıyla problemimizde yer alan veri seti için Python programında Yapay Sinir Ağları ile tahmin yapmak, Lineer Regresyon yöntemine göre daha uygundur.

3.4. İstatistiksel Analiz ve Karşılaştırma

Veri setinde kullanılan 2019 yılı gerçek değerleri ile 2019 yılı tahmin değerleri arasında ortalama mutlak hata (mae) analizi ve ortalama kare hata (mse) analizi yapılacaktır.

Ortalama mutlak hata kavramı, sürekli değişken olan iki ifadenin aralarındaki farklarını belirtir. ortalama mutlak hata'nın aldığı değer, 0'dan başlayarak ∞ değere kadar değişkenlik gösterebilir. Ortalama mutlak hatası daha az değere sahip olan tahminleyiciler, yüksek değerli olan ortalama mutlak hataya göre daha yakın sonuç verir.

Tablo 5. Weka Verileri ile Ortalama Mutlak Hata

Weka	Ortalama Mutlak Hata
Lineer Regresyon İhracat Türk Bayraklı	195376,4231
Lineer Regresyon İhracat Kendi Ülke Bayraklı	42200,8118
Lineer Regresyon İhracat Diğer Ülke Bayraklı	1535011,9589
Lineer Regresyon İthalat Türk Bayraklı	390684,0969
Lineer Regresyon İthalat Kendi Ülke Bayraklı	244359,9989
Lineer Regresyon İthalat Diğer Ülke Bayraklı	968020,9375
Yapay Sinir Ağları İhracat Türk Bayraklı	202266,6453
Yapay Sinir Ağları İhracat Kendi Ülke Bayraklı	119367,8189
Yapay Sinir Ağları İhracat Diğer Ülke Bayraklı	1996420,3605
Yapay Sinir Ağları İthalat Türk Bayraklı	545227,6489
Yapay Sinir Ağları İthalat Kendi Ülke Bayraklı	350634,2087
Yapay Sinir Ağları İthalat Diğer Ülke Bayraklı	3546366,9875

Weka programı üzerinde hesaplanan değerler ışığında, ortalama mutlak hata fonksiyonu için değerler Tablo 5.'de hesaplanmıştır.

Lineer Regresyon yöntemi ve Yapay Sinir Ağları yöntemi, Ortalama Mutlak Hata verileri üzerinde karşılaştırılmıştır. Ortalama Mutlak Hata fonksiyonunda elde edilen sonuçlarda Lineer Regresyon Yöntemi verileri Yapay Sinir Ağları verilerinden daha küçük olduğu için, WekaForecast ile Lineer Regresyon yöntemi daha iyi sonuç vermiştir.

Tablo 6. Python Verileri ile Ortalama Mutlak Hata

Python	Ortalama Mutlak Hata
Lineer Regresyon İhracat	4253384,4886
Lineer Regresyon İthalat	2984447,0235
Yapay Sinir Ağları İhracat	1018326,25
Yapay Sinir Ağları İthalat	1551705,1666

Python programı üzerinde hesaplanan değerler ışığında, Ortalama Mutlak Hata fonksiyonu için Ortalama Mutlak Hata sonuçları Tablo 6'de hesaplanmıştır.

Lineer Regresyon yöntemi ve Yapay Sinir Ağları yöntemi, Ortalama Mutlak Hata verileri üzerinde karşılaştırılmıştır.

Ortalama Mutlak Hata fonksiyonunda elde edilen sonuçlarda Yapay Sinir Ağları verileri Lineer Regresyon Yöntemi verilerinden daha küçük olduğu için, Python Yapay Sinir Ağları yönteminin daha iyi sonuç vermiştir.

Ortalama Kare Hata, bir makine öğrenmesi modelinin, tahminleyicinin göstermiş olduğu performansı değerlendirir, pozitif değerlerden oluşur ve Orta Kare Hata değeri sıfıra yakın olan tahminleyicinin göstermiş olduğu performans daha başarılıdır.

Tablo 7. Weka Verileri ile Ortalama Kare Hata

Weka	Ortalama Kare Hata
Lineer Regresyon İhracat Türk Bayraklı	52408508713,2152
Lineer Regresyon İhracat Kendi Ülke Bayraklı	2669961793,6117
Lineer Regresyon İhracat Diğer Ülke Bayraklı	2887613581745,527
Lineer Regresyon İthalat Türk Bayraklı	178811354370,0350
Lineer Regresyon İthalat Kendi Ülke Bayraklı	105372308420,1585
Lineer Regresyon İthalat Diğer Ülke Bayraklı	1345585061325,3809
Yapay Sinir Ağları İhracat Türk Bayraklı	56261521128,659294
Yapay Sinir Ağları İhracat Kendi Ülke Bayraklı	18724516651,622463
Yapay Sinir Ağları İhracat Diğer Ülke Bayraklı	6224761310927,28
Yapay Sinir Ağları İthalat Türk Bayraklı	480100059275,1989
Yapay Sinir Ağları İthalat Kendi Ülke Bayraklı	180205538525,51065
Yapay Sinir Ağları İthalat Diğer Ülke Bayraklı	17300167872300,781

Weka programı üzerinde hesaplanan değerler ışığında, Ortalama Kare Hata fonksiyonu için Ortalama Kare Hata değerleri Tablo 7'de hesaplanmıştır.

Lineer Regresyon yöntemi ve Yapay Sinir Ağları yöntemi, ortalama kare hata verileri üzerinde karşılaştırılmıştır. Ortalama kare hata fonksiyonunda elde edilen sonuçlarda lineer regresyon yöntemi verileri yapay sinir ağları verilerinden daha küçük olduğu için, wekaforecast uygulandığında lineer regresyon yöntemi daha iyi sonuç vermiştir.

Tablo 8. Python Verileri ile Ortalama Kare Hata

Python	Ortalama Kare Hata
Lineer Regresyon İhracat	21850917895059,457
Lineer Regresyon İthalat	18336940622799,812
Yapay Sinir Ağları İhracat	1617305870253,0964
Yapay Sinir Ağları İthalat	3383154514427,4966

Python programı üzerinde hesaplanan değerler ışığında, Ortalama Kare Hata fonksiyonu için Ortalama Kare Hata değerleri Tablo 8’de hesaplanmıştır.

Lineer Regresyon yöntemi ve Yapay Sinir Ağları yöntemi, ortalama kare hata verileri üzerinde karşılaştırılmıştır. ortalama kare hata fonksiyonunda elde edilen sonuçlarda yapay sinir ağları verileri lineer regresyon yöntemi verilerinden daha küçük değerlere sahip olduğu için, python dilinde yapay sinir ağları yönteminin daha gerçeğe yakın sonuç vermiştir.

3.5. 2020 Yılı İhracat ve İthalat Tahmini ve Gerçekleşen Değerler ile Karşılaştırma

Elde edilen veriler sonucunda, Weka programı üzerinde Lineer Regresyon yöntemi, Python programında Yapay Sinir Ağları yöntemi daha gerçeğe yakın sonuç vermiştir. Bu bilgiler ışığında, Weka programında Lineer regresyon yöntemi ile, Python programında Yapay Sinir Ağları ile 2020 yılı tahmini gerçekleştirilmiştir.

Tablo 9. Weka Programı2020 yılı İhracat ve İthalat Tahmini

Dönem(Ay)	İHRACAT(ton)	İTHALAT(ton)
1	10157323	19179764
2	10222337	18073946
3	11776189	20180525
4	11173987	19282839
5	11098633	19772010
6	10858882	19219970
7	10344919	18348420
8	11197428	18765711
9	11548752	18517378
10	12195989	19229276
11	11828890	19321409
12	11377269	19520156

2020 Yılı ile ilgili Weka programı üzerinde yapılan tahminler Tablo 9.’de verilmiştir. İhracat ve İthalat tahminleri yapılırken kullanılan yöntem lineer regresyon yöntemidir.

Tablo 10. Python Program 2020 yılı İhracat ve İthalat Tahmini

Dönem(Ay)	İHRACAT(ton)	İTHALAT(ton)
1	10284938	20507074
2	9951037	20261772
3	9502345	20195968
4	9242197	19732638
5	8780814	190955114
6	8241287	18330078
7	8111201	17854850
8	8013423	17105300
9	7905110	16742764
10	7551840	15964641
11	7288175	16102641
12	7297451	16161033

2020 Yılı ile ilgili Python programı üzerinde yapılan tahminler Tablo 10.’da tabloda verilmiştir. İhracat ve İthalat tahminleri yapılırken yapay sinir ağları yöntemi kullanılmıştır.

Çalışma sonuçları Ocak 2021’de açıklanan ithalat ve ihracat rakamları ile karşılaştırılmıştır. Beklenen tahmin değerleri ve gerçekleşen değerler Tablo 11. ‘de gösterilmektedir.

Tablo 11. 2020 yılı tahmin ve gerçekleşen ithalat ve ihracat rakamları

Dönem (Ay)	İHRACAT (ton) Tahmini	İTHALAT (ton) Tahmini	İHRACAT (ton) Gerçek	İTHALAT (ton) Gerçek
1	10157323	19179764	11529896	32953498
2	10222337	18073946	10876501	29507308
3	11776189	20180525	12092311	32159263
4	11173987	19282839	10557637	28843534
5	11098633	19772010	10219872	28028476
6	10858882	19219970	10953497	28147557
7	10344919	18348420	11625919	30253863
8	11197428	18765711	10733845	29500707
9	11548752	18517378	12154193	31177030
10	12195989	19229276	13017990	31988333
11	11828890	19321409	12421908	30121378
12	11377269	19520156	12719254	32761349

Bu sonuçlara göre 2020 yılında pandemi etkisi ile ihracatın beklenin altında, ithalatın ise beklenen değerlerin üzerinde gerçekleştiği görülmektedir. Yapılan veri analizi ve tahmin açısından ise Tablo 12’ de, 2020 yılına ait İhracat ve İthalat verilerinin gerçekleşme oranlarını göstermektedir. 2020 yılına ait verilerin Weka ile yapılan analizde İhracat ve İthalat verileri 2020

yılına ait gerçek verilerle oranlanması ile, İhracat ve İthalat başlıkları adı altında yüzdelik sonuçlar elde edilmiştir.

Tablo 12. 2020 İhracat ve İthalat Verileri Tahmin Gerçekleşme Oranları

Weka	
İhracat Tahmini değer/ Gerçek değer	İthalat Tahmini değer/ Gerçek değer
0,881	0,582
0,940	0,613
0,974	0,628
0,942	0,669
0,914	0,705
0,991	0,683
0,890	0,606
0,957	0,636
0,950	0,594
0,937	0,601
0,952	0,641
0,894	0,596

Bu sonuçlara göre 2020 yılı ihracat rakamları tahmin açısından çok yakın değerlere sahipken, ithalat oranlarının beklenenin üzerinde çıktığı tahmin hatası ile de görülmektedir. İthalat rakamları yük istatistik rakamlarına göre beklenenin üzerinde bir artış göstermiştir. Bu artışın, Covid 19-Pandemisi ile ilişkisi olduğu, kimi ithal malzemelere talep ve ihtiyacın arttığı şeklinde değerlendirilebilir.

4. Sonuç

Ülkemizde, lojistik açıdan bulunduğu konuma istinaden, deniz taşımacılığı ve limancılık gün geçtikçe ilerlemektedir. Deniz ticareti üzerine yapılabilecek veri analizi çalışmaları ise gelecek stratejik ve operasyonel hedefler ve faaliyetler oluşturulmasında etkili olacaktır. Bu çalışma kapsamında, Weka ve Python ile Lineer Regresyon Yöntemi ve Yapay Sinir Ağları yöntemleri kullanılarak deniz ticaretinde ithalat ve ihracat tahmini yapılmıştır. Veri seti olarak, T.C Ulaştırma ve Altyapı Bakanlığı'na ait 16 yıllık (2004-2019) yük istatistik verileri kullanılmıştır. Gerçekleştirilen veri madenciliği uygulaması ile, kullanılan hangi yöntemin gerçek değere daha yakın sonuçlar vereceği belirlenmiş ve elde edilen sonuçlar ile de 2020 yılı için tahminleme işlemi gerçekleştirilmiştir.

Yapılan bu çalışmada Weka programı kapsamında ihracat ve ithalat verileri kendi içerisinde Türk Bayraklı, Kendi Ülke Bayraklı ve Diğer Ülke Bayraklı olarak sınıflandırılarak 6 ana başlık altında analiz edilmiştir. Çalışma sonucunda Weka programı üzerinde oluşan tahmin ve gerçek değerler arasında ortalama mutlak hatanın ve ortalama kare hatanın sonuçları hesaplanıp yorumlanmıştır. En iyi sonucu veren yöntem altı başlık altında da lineer regresyon yöntemi seçilmiştir. Python programında yapılan çalışmada, weka programında kullanılan veri setinin ihracat ve ithalat olarak iki ana başlıkta analizi gerçekleştirilmiştir. Hesaplanan ortalama mutlak hata'nın ve ortalama kare hata'nın sonuçları doğrultusunda, en iyi sonucu veren yöntem python programında yapay sinir ağları olarak seçilmiştir.

Weka programında incelenen veriler üzerinde lineer regresyon yöntemi daha doğru, başarılı sonuçlar vermiştir. aynı şekilde Python programında ise Yapay Sinir Ağları yöntemi doğruluk oranı yüksek sonuçlar vermiştir. Weka programında daha başarılı sonuç veren model olan lineer regresyon yöntemi ortalama olarak %80,43 oranında 2019 yılı verilerini doğru tahmin etmiştir. Python programında daha başarılı sonuç veren model olan yapay sinir ağları yöntemi ise ortalama olarak %91,25 oranında 2019 yılı verilerini doğru tahmin etmiştir. Elde edilen sonuçlar ışığında, 2020 yılına ait ihracat ve ithalat verilerinin tahminleme işlemi, Weka programında Lineer Regresyon yöntemi ile, Python dili ile Yapay Sinir Ağları yöntemi uygulanarak gerçekleştirilmiştir. 2020 yılı gerçekleşen rakamlar ile kıyaslandığında ise, ihracatın beklenen oranda sonuçlandığı bununla birlikte Covid-19 pandemisi etkisi ile ithalatta beklenenin üzerinde artış olduğu gözlemlenmiştir.

Çalışmada weka programında test veri oranında değişiklik yapılarak doğruluk oranı daha yüksek sonuçlar elde edilebilir. Mevsimsel bazda çalışmaların yapılması ile daha tutarlı sonuçlar elde edilebilir. 2020 yılı gerçekleşen verileri ile tahmin sonuçlarını kıyasladığımızda ithalat alanındaki tahmin hatasının çoklu başka etkenlere bağlı olduğu düşünülürse veri boyutu genişletilerek tahmin doğruluğu artırılabilir. Yapılan tahmin ve liman/yük düzeyinde yapılacak çalışmalar; hedeflenen ihracat seviyesinde yük dağılımının planlanması, limanların elleçleme kapasitelerinin planlanması, kapasitelerin artırılması için limanlarda ne tip yatırımlara öncelik verilebileceği gibi konulara yol gösterici olacaktır.

Kaynakça

- Akçetin, E., Çelik U., Takçı H., (2013). 'Lojistik ve Denizcilik Sektörü Açısından Veri Madenciliği Uygulamalarının Önemi', Journal of ETA Maritime Science Vol. 1, No. 2, 73-80.
- Aksoy, M. B. (2019). 'Türkiye'de Taşımacılık Alt Yapılarının Ekonomik Büyüme Üzerine Etkileri.', Yüksek Lisans Tezi, Gaziantep Üniversitesi.
- Alnıpak, S., Alkan, F., Günay, G. (2017). 'Veri Madenciliği, Büyük Veri Analizi Ve Rfid Teknolojisi Tabanlı Konteyner Limanları', III. Ulusal Liman Kongresi.
- Arslan, A. (2011). 'Avrupa Birliği Uyum Sürecinde, Türkiye'de Denizyolu Konteyner Taşımacılığı ve Limanların Ekonomideki Yeri. Yüksek Lisans Tezi', Marmara Üniversitesi, İstanbul.
- Aydemir, E. (2019). 'Weka ile Yapay Zekâ, Seçkin Yayınları', ISBN: 978-975-02-5536-6, Ankara, Türkiye.
- Birant, D. (2006). 'Modeling and Analyzing Marine Data Using Data Mining Techniques. Yüksek Lisans Tezi', Dokuz Eylül Üniversitesi, İzmir.
- Cornejo-Bueno, L., Garrido-Merchán, E.C., Hernández-Lobato, D., Salcedo-Sanz, S. (2018). 'Bayesian optimization of a hybrid system for robust ocean wave features prediction', Neurocomputing, 818-828.
- Demir, Ş. (2005) Uluslararası Taşımacılık/Lojistik KDV İstisnası ve İadesi, Gelirler 74 Kontrolörleri Derneği Yayını, İstanbul, s. 1.

Gün, D., (2013). 'Değişim Çağında Sürdürülebilir Lojistik Süreç ve Stratejilerinin Yönetimsel Bakış Açısıyla Değerlendirilmesi ve Küresel Lojistik Üs Vizyonu', Rize Kalkınma Sempozyumu.

<https://lpi.worldbank.org/international/global?sort=asc&order=LPI%20Rank#datatabl>

Jingjing Y., Guolei T., Xiangqun S., Xuhui Y., Yue Q., Da L., Yong Z., (2018). 'Ship Arrival Prediction and Its Value on Daily Container Terminal Operation', *Ocean Engineering* 157 (2018) 73-86.

Jose A.S., Juan A.G., Agapito L., (2017). 'Forecasting Sea Level Changes Applying Data Mining Techniques to the Cristobal Bay Time Series Panama', IWA Publishing.

Kalaycı, İ. (2014): 'Deniz Ticareti ve Küresel Mali Kriz: İpek Yolu'nda Türkiye için Yeni Stratejiler'.

Khair A.F., Awang M.K., Zakarala Z.A., Mazlan M., (2017). 'Daily Streamflow Prediction On Time Series Forecasting', *Journal of Theoretical and Applied Information Technology*.

Kılıcı, Ö. G. H., (2017). 'Taşımacılık Faaliyetlerinin Rekabet Üstünlüğü Oluşturmada İşletmecilikte Yeri ve Önemi Türkiye Örneği ve Türkiye'de Taşımacılığın Gelişimi', *Balkan Sosyal Bilimler Dergisi*.

Kiremitçi, B., (2005). 'Veri Ambarlarında Veri Madenciliği ve Ulaştırma-Lojistik Sektöründe Bir Uygulama. Yüksek Lisans Tezi', İstanbul Üniversitesi, İstanbul.

Kokotos D.X., Linardatos D.S., (2010). 'An application of data mining tools for the study of shipping safety in restricted waters', *Safety Science* 49, 192-197.

Lambert, D.M.- James R.S. (1999). 'Strategic Logistics Management. Irwin\McGraw-Hill.' 3. Baskı. Boston.

Özkan, Y., (2016). 'Veri Madenciliği Yöntemleri, Papatya Yayıncılık Eğitimi', ISBN:978-975-6797-82-2, İstanbul, Türkiye, Eylül.

S. Nataraj, C. Alvarez, L. Sada, A. A. Juan, J. Panadero, C. Bayliss, (2019). 'Applying Statistical Learning Methods for Forecasting Prices and Enhancing the Probability of Success in Logistics Tenders', *Transportation Research Procedia* 47 (2020) 529-536.

Sousa N., Campos A. (2018). 'Data mining for anomaly detection in maritime traffic data', *Escola Naval*.

Şenyiğit, E., Arsav, M. S. (2019). 'The Outbound Containers' Storage Location Assignment Problem for a Maritime Terminal. *European Journal of Science and Technology*', 349-355.

Ulaştırma ve Altyapı Bakanlığı, (t.y.), İstatistiklerle Ulaştırma Denizcilik Ve Haberleşme 2004-2019.
<http://www.uab.gov.tr/images/istatistik/2004.2019.pdf>

Wang, J., Simin Pulat, P. and Shen, G., (2012). 'Data mining for the development of a global port-to-port freight movement database', *Int. J. Shipping and Transport Logistics*, Vol. 4, No. 2, pp.137-156.