Zaman Serilerinde Anlık Tahmin (Nowcasting): Klasik Ekonometrik Yöntemler ile LSTM Modelinin Karşılaştırmalı Analizi

G. Merve GÖKÇİN[1]

Giriş

Zaman serilerinde sıklıkla karşımıza çıkan "Nowcasting", "Şimdi/ An" ve "Tahmin"in birleşiminden oluşan bir kavramdır ve türkçeye anlık / şimdiki tahmin olarak çevrilmektedir. Anlık tahminde amaç, mevcut olan bilgileri kullanarak bir hedef değişkenin mevcut veya ona yakın zamanda ileri veya geri durumunun daha zamanlı bir şekilde tahminidir. Keith Browning, bu terimi ilk defa, 1981'de (WMO, 2017) çok yakın gelecekteki hava durumunu, mevcut durumuna dayanarak tahmin etmeyi tanımlamak için kullanmıştır. Makroekonomik duruma ilişkin "gerçek zamanlı tahminler" kavramı ise Mariano ve Murasawa (2003) tarafından ilk defa ele alınmıştır.

Gayrisafi Yurt İçi Hasılanın (GSYH) ekonominin büyüklüğünü ve performansını ölçmedeki önemi hafife alınamaz. Hükümet yetkilileri, politika yapıcılar ve hatta genel halk için ekonomik sağlık ve dolayısıyla sosyal refah gibi sayısız ilgili ölçüt hakkında fikir edinmek için başvurulan bir ölçüttür (Dynan vd., 2018; IMF, 2020; Kapoor ve Debroy, 2019). Ancak bu rakama ne kadar bağımlı olunsa da, çoğu zaman önemli bir gecikmeyle yayınlanmaktadır. Bu nedenle 2010'lu yıllarda GSYH, ticaret, ekonomik faaliyetlere ilişkin değişkenlerin "anlık / şimdiki tahmin"i üzerine birçok çalışma yapılmıştır. Bu çalışmalarda klasik zaman serisi yöntemlerinden yararlanılmaktadır. Bu yöntemler genellikle tek değişkenli tahminler sağlar ve mevsimsellik ile diğer parametreleri modellemek için elle ayarlanmış özellikler gerektirir. Ancak ulusal istatistik ofislerinin mevcut veri kaynaklarının alternatif veri kaynakları ile güçlendirmesi ile ekonomik veriler daha ayrıntılı ve daha kısa zaman diliminde incelenebilmektedir. Ele alınan

[1] Dr. Öğr. Gör., Çukurova Üniversitesi, gmgokcin@cu.edu.tr, ORCID: 0000-0002-2644-1942

ekonometrik modeldeki değişken sayısı arttıkça ve zaman frekansları farklılık gösterdikçe, model kurulumuna ilişkin çoklu doğrusallık, eksik gözlemler, karma frekanslı verilerin varlığı (bağımlı ve bağımsız değişkenlerin farklı periyotlarda olması) ve değişen yayın tarihleri gibi sorunlar ortaya çıkmaktadır. Bu sorunların çözümüne ilişkin çalışmaların artmasıyla ekonomik verilerin anlık tahmininde Yapay Sinir Ağ (YSA) modellerinin kullanımında son on yılda metodolojik ilerlemeler kaydedilmiştir.

Geleneksel ileri beslemeli YSA'ların zaman serisi tahmininde uzun bir kullanım geçmişi vardır (Kohzadi, vd., 1996). Literatürde bir çok çalışmada YSA model tahminleri ile klasik zaman serileri model sonuçları karşılaştırılmaktadır. Bu çalışmalarda en çok dikkat çeken ekonomik verilerin yorumlanmasında YSA modellerinin elverişli olmamasıdır. Ayrıca bu modellerin açık bir zamansal yönü yoktur. Zaman boyutunun mevcut YSA modellerine dahil edilmesi ile tekrarlayan sinir ağları (RNN) ortaya çıkmıştır (Amidi vd. 2019). İleri beslemeli ağlarda girdiler ve çıktılar arasındaki tek yönlü ilişkinin aksine, RNN'ler, katman çıktılarının ağa geri beslenebildiği bir geri besleme döngüsü sunar (Stratos, 2020), (Dematos vd., 1996). Bu mimari, RNN'leri doğal dil işleme veya konuşma işleme gibi zamansal yönü veya akışı olan uygulamalara çok uygun hale getirir. Bununla birlikte, kaybolan gradyanlar nedeniyle, RNN'ler çok "kısa" bir belleğe sahip olma eğilimindedir ve bu da onların şimdi tahmin uygulamasındaki kullanışlılığını sınırlamaktadır (Grosse, 2017; Hoop, 2021). Uzun Kısa Süreli Bellek (LSTM) (Hochreiter ve Schmidhuber, 1997) gibi RNN'ye dayalı zaman serisi modelleri, uçtan uca eğitimleri, dışsal yardımcı değişkenleri dahil etme kolaylığı ve derin öğrenmenin ayırt edici özellikleri olan otomatik özellik çıkarma yetenekleri nedeniyle popüler yöntemler haline gelmiştir (Rasul vd., 2020).

Bu çalışma zaman serilerinin anlık tahmininde sıklıkla kullanılan LSTM modelleri ile örnek olarak seçilen klasik ekonometrik yöntemleri karşılaştırırken, LSTM modellerine ilişkin detaylı bilgi sağlamayı amaçlamaktadır. Bu kapsamda ikinci bölümde konu ile ilgili yapılmış ulusal ve uluslararası çalışmalara yer verilmiştir. Üçüncü bölümde Yapay Sinir Ağlarının, Tekrarlayan Sinir Ağlarının (RNN) ve Uzun Kısa Süreli Bellek (LSTM) ağlarının yapısı teorik olarak ele alınmıştır. Uygulama bölümünde ise farklı frekanslarda yayınlanan iki seri ile öngörü modelleri için deneysel bir çalışma yapılmıştır. Türkiye Elektrik Tüketimi için saatlik frekansta, GSYH büyümesi için aylık frekansta TRAMO /SEATS, ARIMA ve LSTM modeli tahminleri elde edilmiş, sonuçlar yorumlanmıştır. Bu bölümde bilimsel katkılardan dolayı ayrıca Çukurova Üniversitesi Bilgisayar Mühendisliği

Anabilim Dalı Dr. Öğr. Üyesi Mehmet Sarıgül'e ve Yüksek Frekanslı Zaman Serileri üzerine çalışmaları olan İlayda Tosun'a teşekkür ederim. Çalışmanın RNN&LSTM modellerinin teorik yapısını incelemesi ve ekonomik serilerde kullanımına örnek temsil etmesi nedeniyle alan çalışmalarına katkıda bulunacağı düşünülmektedir.

Literatür Taraması

Politika yapıcılar ve piyasa katılımcıları, politika kararlarını ve yatırım stratejilerini belirlemek, ekonominin mevcut durumunu değerlendirmek için gerçek makro ekonomik göstergeleri takip ederler. Gerçekte ekonominin mevcut durumuyla ilgili temel istatistikler önemli bir gecikmeyle elde edilmektedir ve bu gecikmelerden dolayı anlık tahmin ekonomide oldukça önemlidir. Özellikle üç aylık bazda toplanan Gayri Safi Yurt İçi Hasıla (GSYH) bunun önemli bir örneğidir. Örneğin, Amerika Birleşik Devletleri'nde (ABD) veya Birleşik Krallık'ta GSYH'nin ilk resmi tahmini referans çeyreğinin bitiminden yaklaşık bir ay sonra yayınlanır. Avro bölgesinde buna karşılık gelen yayın gecikmesi 2-3 hafta daha uzundur. Ülkemizde de GSYH referans çeyreğine göre bir gecikmeyle yayınlanır. Bu gecikmelerden dolayı "Anlık tahminin" temel ilkesi de resmi rakam mevcut olmadan önce bir 'erken tahmin' elde etmek için, ilgi duyulan hedef değişkenden daha erken ve muhtemelen daha yüksek frekanslarda yayınlanan bilgilerin kullanılması olarak belirtilmektedir (Banbura vd., 2013). Nowcasting kavramının 2000'li yıllarda ekonomi literatürüne girmesiyle birlikte zaman serisi analizi ve şimdi tahmin konularında yapılan çalışmalar giderek artmıştır. Bu dönemde yapılan çalışmalara bakıldığında, ülkelerin ekonomik konjonktürel hareketlerine ilişkin GSYH, ticaret, işgücü piyasası vb. anlık gerçek zamanlı tahminlere ihtiyaç duyulmuş ve çeşitli ekonometrik analiz yöntemleri geliştirilmiştir. Örneğin, Stock ve Watson (2002) çalışmasında konjonktürel dalgaların tespiti için şimdiki zaman tahmininde dinamik faktör modellerinden (DFM) yararlanmış ve mevcut veri sorunlarını değerlendirmiştir. Daha sonra Mariano ve Murasawa (2003), Aruoba vd. (2009) çalışmalarında gerçek ekonomik faaliyet verilerini ölçmek için, yüksek ve karışık frekansta gözlemlenen çeşitli stok ve akım verilerine yönelik filtrelemeye izin veren dinamik faktör modellerini geliştirmişlerdir.

Anlık tahminde özellikle GSYH'nın büyüme oranının tahmininde ARMA modellerinden sıklıkla yararlanılmaktadır. Bu modelde ARMA modeli yalnızca GSYH büyüme oranının kendi geçmişini girdi değişkeni olarak içermektedir. Örnek olarak Cimadomo vd. (2020)'nin çalışmaları ile

Yeni Zelanda Hükümeti Ulaştırma Bakanlığının uygulamaları verilebilir (Hoop, 2022). Baffigi, Golinelli ve Parigi (2004) de çalışmalarında çeyreklik GSYH büyümesinin tahmini için "bridge regression" olarak adlandırdıkları regresyon modelini kullanmışlardır. Bu modelde bağımlı değişken için belirlenen kısıtlı çeyreklik veri olduğundan aylık ekonomik veriler (örneğin sanayi üretim verileri) ARIMA gibi yardımcı modellerle tahmin edilmiştir. Giannone, Reichlin ve Small (2008) ve Evans (2005) ise çalışmalarında şimdiki zaman tahmin sürecinde istatistiksel bir çerçeve oluşturmaya çalışmış ve durum uzay (state space) modellerinden yararlanmışlardır.

2010'lı yıllara gelindiğinde ise anlık tahmin konusunda birden fazla bağımlı değişkenin ele alındığı ve dolayısıyla metodolojinin ve modellemelerin farklılaştığı görülmektedir. Bu çalışmalarda en çok ele alınan bağımlı değişkenler GSYH ve ticaret verileri olup en çok ele alınan metodolojiler ise Dinamik Faktör Modelleri (DFM) (Guichard ve Rusticelli, 2011), (Antolin-Diaz vd, 2020), Karma Frekanslı Vektör Otoregresif Model (MF-VAR) (Kuzin vd, 2009), (Marcellino and Schumacher, 2010), Bayesyen Karma Frekanslı Vektör Otoregresif Model (Bayesyen VAR) ve karma veri toplama (MIDAS-mixed data sampling) yöntemleridir (Hoop, 2021). Bu modellerden MF-VAR'da genel yaklaşım, daha yüksek frekanslı göstergeleri en düşük frekanslı göstergenin frekansına toplamaktır. Örneğin, aylık değişkenler kullanılarak üç aylık bir büyüme oranı tahmininde, aylık değişkenler üç ayın büyüme oranlarının ortalaması alınarak dönüştürülmektedir. Bayesyen VAR modelinde ise model parametrelerinin tahmin edildiği ve sabit değerler olarak alındığı standart VAR'ların aksine, değişkenler parametreleri atanmış bir ön olasılığa sahip rastgele değişkenler olarak ele alınır. Bayesyen VAR'ın karmaşık dinamik sistemleri modelleme, parametrelendirme ve doğrusallığı ele almadaki gücü, onları şimdiki zaman tahmini alanında popüler bir seçim haline getirmiştir (Banbura vd., 2010), (Cimadomo vd., 2020), (Hoop, 2022). MIDAS modelinde ise karma frekanslı VAR yöntemine benzemekle birlikte daha yüksek frekanslı değişkenlerin gecikmeleri de modele dahil edilmektedir.

Klasik ekonometrik yöntemlerin yanı sıra son yıllarda makine öğrenmesi ve derin öğrenme yöntemleri de zaman serileri analizi ve şimdiki tahminde tercih edilen yöntemler haline gelmiştir. Kitabın bu bölümünde ele alınacak olan derin öğrenme yöntemleri ve özellikle RNN&LSTM üzerine yapılan ilk çalışmalardan biri Hochreiter vd. (1997)'ne aittir.

Derin öğrenme tabanlı tahmin yöntemleri ekonomik ve finans verileri haricinde ayrıca büyük ölçekli endüstriyel tahmin uygulamalarında, satış hacmi, elektrik tüketimi, trafik yoğunluğu, güneş enerjisi, meteorolojik verilerin tahmini gibi alanlarda yaygın olarak kullanılmaktadır. Bu pratik başarı, derin tahmin yöntemlerini anlama ve iyileştirme konusundaki akademik ilgiyi daha da artırmıştır. Benidis vd. (2022), bu bakış açısıyla makalelerinde tahmin modelleri için derin öğrenme yöntemlerini incelemiş ve yapılan tüm çalışmalara ilişkin bir literatür kaynağı oluşturmuşlardır. Makalelerinde derin tahmin yöntemlerinden RNN, CNN, MLP, SVM, LSTM, State Space gibi çeşitli yöntemlerin kullanıldığı makaleleri, kullanılan metrikler, veri tipleri ve kayıp fonksiyonları bakımından özetlemişlerdir. Salinas vd. (2019) da çalışmalarında çok sayıda ilişkili zaman serisinde otoregresif tekrarlayan sinir ağı modelini eğitmeye, doğru olasılıksal tahminler üretmeye dayalı DeepAR algoritmasını önermektedir. Makalede derin öğrenme tekniklerinin zaman serisi tahminine yönelik kullanımı ile klasik yaklaşımlarda karşılaşılan zorlukların üstesinden gelinebileceği vurgulanmıştır.

Son yıllarda LSTM ile ilgili en çok atıf alan akademik çalışmalardan biri de Houdt vd. (2020) tarafından ele alınmıştır. Değerlendirme makalesinde LSTM ile ilgili yapılan çalışmalara, modelin nasıl eğitilmesi gerektiğine, zaman serisi tahmini ve doğal dil işleme modellerine yer verilmektedir.

Yüksek varyanslı serilerde (örneğin tatiller) doğru zaman serisi tahmini, anormallik tespiti, optimum kaynak tahsisi, bütçe planlaması ve diğer ilgili görevler için kritik öneme sahiptir. Laptev vd. (2017), ulaşım sektörü örneği olan Uber'de özel etkinlikler sırasında tamamlanan yolculuklar için doğru tahmin modellemesi üzerine çalışmışlardır. Çalışmanın amacı daha verimli bir sürücü tahsisi elde edilirken, yolcular için daha az bekleme süresinin sağlanmasıdır. Bu amaç için genellikle tek değişkenli tahmin modelleri (örneğin Holt-Winters) ve makine öğrenimi yöntemlerinin (örneğin rastgele orman) bir kombinasyonu ele alınmaktadır. Ancak böyle bir sistemin ayarlanması, ölçeklenmesi ve dışsal değişkenler eklenmesi zordur. Bu nedenle çalışmalarında Uber verilerinde mevcut en son olay tahmin yöntemlerinden daha iyi performans gösteren ve zaman serisi tahmin yarışmaları için kullanılan LSTM yapısına ek olarak yeni bir uçtan uca yinelemeli sinir ağı mimarisi önermişlerdir.

Uydu görüntüleri ve meteorolojik ölçüm verilerinden elde edilen parametrelerle anlık yağış tahmini için çeşitli yapay sinir ağ modellerinin kullanımı son zamanlarda literatürde yer almaktadır. Rahimi vd. (2023) anlık

yağış tahminlerini elde edebilmek için LSTM modellerini kullanmışlardır. Çalışmaları, dört saatlik bir ön hazırlık süresiyle neredeyse her otuz dakikada bir küresel yağışın anlık tahmini için derin öğrenme mimarisini sunar. Mimari, bir U-Net ve bir evrişimli LSTM sinir ağını birleştirir ve GPM için Entegre Çoklu Uydu Almaları (IMERG) ve Küresel Tahmin Sistemi'nden (GFS) gelen birkaç önemli yağış sürücüsünden gelen veriler kullanılarak eğitilir. Bir diğer çalışma Patel vd. (2018) tarafından ele alınmıştır. Hava tahmini, büyük miktarda verideki temel yapıyı belirlemeyi içerdiğinden, çalışmalarında derin öğrenmeye dayalı yağış tahminlerinin geleneksel doğrusal ekstrapolasyon yöntemlerinden sezgisel olarak daha iyi performans gösterdiği sonucuna varmışlardır.

Zaman serisi tahminlerindeki belirsizlikleri modelleme sorununu göze alan Rasul vd. (2020), çalışmalarında, zaman serilerinin çok değişkenli zamansal dinamiklerini, veri dağılımının koşullu bir normalleştirme akışıyla temsil edildiği otoregresif bir derin öğrenme modeli tasarımı üzerine çalışmışlardır. Bu model Maskeli Otoregresif Akış (Papamakarios, vd., 2017) veya Gerçek NVP (Dinh vd., 2017) gibi bir normalleştirme akışı ile modelleyen, olasılıksal tahmin için eğitilebilir otoregresif derin öğrenme mimarileri içermektedir.

Zaman serilerinde zamansal bağımlılıkları otomatik olarak öğrenme kapasiteleri göz önüne alındığında derin öğrenme yöntemlerinin etkili bir çözüm oldukları kanıtlanmıştır. Ancak, en uygun derin sinir ağı türüne ve parametrelerine karar vermek, önemli uzmanlık gerektiren karmaşık bir süreçtir. Bu nedenle, farklı tahmin görevleri için tüm mevcut mimarilerin uygunluğu hakkında daha derin çalışmalara ihtiyaç vardır. Bu nedenle Benitez vd. (2021), zaman serilerinin öngörüsü için derin öğrenme yapıları üzerine bir makale hazırlamıştır. Bu çalışmada, öncelikle zaman serisi tahmini için derin öğrenmeyi kullanan en son çalışmalar kapsamlı bir şekilde incelenmiş ve en popüler mimarilerin performansını karşılaştıran deneysel bir çalışma hazırlanmıştır. Karşılaştırma, doğruluk ve verimlilik açısından yedi tür derin öğrenme modelinin kapsamlı bir analizini içermektedir.

Farahani vd. (2025), çalışmalarında zaman serisi tahmin yöntemlerini ele alırken, Yapay Zeka (YZ) ve Makine Öğrenimi (ML) metodlarını incelemişlerdir. İlgili çalışma, YZ ve ML algoritmalarının akıllı üretim ortamlarında uygulanabilirliğine odaklanarak on üç üretimle ilgili veri kümesinde zaman serisi tahmini için geliştirilen algoritmalarının deneysel değerlendirmesini ele aldığı için dikkat çekicidir.

Literatürde diğer dikkat çeken bir çalışma Hoop (2021) tarafından ele alınmıştır. Hoop, 2020 yılındaki pandemi ile gelişen olaylar nedeniyle, politika kararlarını bilgilendirmek için makroekonomik serilerin zamanında ve doğru tahminlerinin değerinin önemini vurgulamıştır. Bu amaçla LSTM'lerin daha güçlü bir şekilde değerlendirilmesi için çalışmasında çeşitli metriklerle modellerin tahmin performanslarını hesaplayarak, gelecekteki araştırmaları kolaylaştırmak için bir Python kütüphanesi de paylaşmaktadır. Çalışmasında mal ihracatına ait değer ve hacim verileri ile hizmet ihracatı verilerini anlık olarak tahmin etmeyi amaçlamıştır. Belirlediği üç ayrı hedef serisinin (üç aylık frekansta) her birini tahmin etmek için, 116 adet karışık frekanslı aylık ve üç aylık bağımsız seriden oluşan geniş bir veri havuzu kullanılmıştır ve çalışmasında LSTM modellerinin tahmininin DFM modellerinden daha üstün sonuçlar ürettiği sonucuna varmıştır.

Özdemir ve Yozgatlıgil (2022) de çalışmalarında yüksek ve düşük frekanslı zaman serilerde geleneksel yöntemler ile makine öğrenmesi yöntemlerinin tahmin performanslarını karşılaştırmışlardır. Geleneksel tahmin yöntemlerine (Naive ve Mevsimsel Naive Yöntemler, S/ARIMA, Üstel Düzeltme, TBATS, Trend Değişiklikleriyle Bayes Üstel Düzeltme Modelleri ve STL Ayrıştırma) ek olarak, tahminler ayrıca yedi farklı makine öğrenme yöntemi kullanılarak elde edilmiştir, bunlar Rastgele Orman, Destek Vektör Regresyonu, XGBoosting, BNN, RNN, LSTM ve FFNN ve hem istatistiksel zaman serisi hem de makine öğrenme yöntemlerinin hibrid yaklaşımla ele alınmasıdır. Veri seti olarak M4 Yarışması veri setindeki çeşitli zaman alanlarından orantılı olarak seriler ele alınmıştır. İlgili çalışma hangi yöntemin farklı zaman serisi frekanslarında daha iyi performans gösterdiğini belirtmektedir.

Metodoloji

YSA modelleri

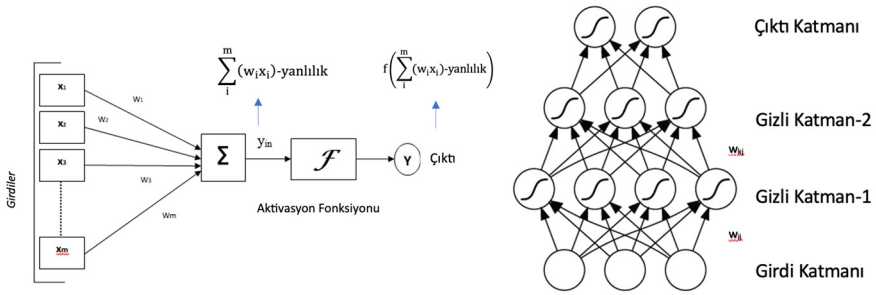
Makine öğrenmesi yöntemlerinden Yapay sinir ağları (YSA/ANN), biyolojik beyinlerin bilgi işleme yeteneklerine ilişkin matematiksel olarak oluşturulan modellerdir (McCulloch ve Pitts, 1988; Rosenblatt, 1963; Rumelhart, vd., 1986). İlk yapay sinir ağı modeli 1943 yılında, bir nörolog olan Warren McCulloch ile bir matematikçi olan Walter Pitts tarafından gerçekleştirilmiştir. Bu model insan beyninin hesaplama yeteneğinden esinlenerek geliştirilen elektrik devrelerine uygulanan basit bir sinir ağı modelidir. Ancak nörolog olan Donald Hebb (1949) bugünün sinir ağı

teorisinin öncüsü olarak adlandırılmaktadır ve beynin nasıl öğrendiği ile ilgili çalışmalar yapmıştır. Çalışmalarına beynin en temel birimi olan sinir hücresini ele alarak başlamıştır. Bilgisayarlar da insan beyninin çalışma sistemini kopyalayan makinelerdir. 1990'lı yıllardan beri bilgisayarların öğrenmesini sağlayan yapay sinir ağları teknolojisinde oldukça hızlı bir gelişme görülmektedir.

Bir YSA'nın temel yapısı ağırlıklı bağlantılarla birbirine bağlanan küçük işlem birimlerinden ya da düğümlerinden oluşan bir ağdır. İnsan beyninin yani biyolojik sinir sistemine bakıldığında; öğrenme nöronlar arasındaki sinaptik bağlantıların ayarlanması ve bir sinir hücresinin işlediği bilgileri aksonlar yolu ile diğer hücrelere göndermesi ile gerçekleşmektedir. Öğrenme süreci hayat boyunca devam etmektedir ve bu süreç içinde beyin sürekli olarak bir gelişim gösterir. YSA modelleri için de bu süreç incelenmiştir ve biyolojik sinir sistemi taklit edilmektedir. Öğrenme, girdi/çıktı verilerinin işlenmesiyle, eğitme algoritmasının bu verileri kullanarak bağlantı ağırlıklarını bir yakınsama sağlanana kadar ayarlamasıyla gerçekleşmektedir (Gökçin, 2019).

YSA'ları anlamak için onların temel taşı olan tek nöron'un yapısı incelenebilir. Şekil 1 (a)'da m girdisi (x_1,x_2,x_3,…, x_m) ve tek çıktısı (fonksiyon) olan bir nöron yapısı bulunmaktadır. Literatürde bu yapıya perseptron (perceptron) ismi de verilmektedir. Bir yapay sinir hücresine dışarıdan ya da başka hücrelerden gelen bilgilere "girdi" denilmektedir. Girdiler, ağın öğrenmesi istenen örnekler tarafından belirlenmektedir. Bu yapıda bir yapay sinir hücresine gelen bilginin önemini ve hücre üzerindeki etkisini "ağırlıklar" göstermektedir. Toplama fonksiyonu ise bir hücreye gelen net girdiyi hesaplamaktadır ve en yaygın olanı ağırlıklı toplamı bulmakla elde edilmektedir. Ağırlıklı toplam her bir girdinin kendi ağırlığı ile çarpılarak toplanmasından oluşmaktadır. Aktivasyon fonksiyonunda ise hücreye gelen net girdi işlenerek hücrenin bu girdiye karşılık üreteceği çıktı belirlenmektedir (Öztemel, 2012).

Temel olarak bir YSA modelinde üç temel katman bulunmaktadır, bunlar girdi, gizli (ara) ve çıktı katmanlarıdır. Giriş desenleri girdi katmanına sunulur, ardından gizli katmana yayılır. Gizli katmanlar da desenleri çıktı katmanına aktarır. Örnek olarak Şekil 1 (b) de 1 girdi, 2 gizli ve 1 çıktı katmanlı bir yapay sinir ağı modeli bulunmaktadır. Bu modelde gizli ve çıktı katmanlarındaki her bir daire bir nöronu (node), nöronların içindeki S şekli, sigmoid fonksiyonunu belirtmektedir (Graves, 2012).



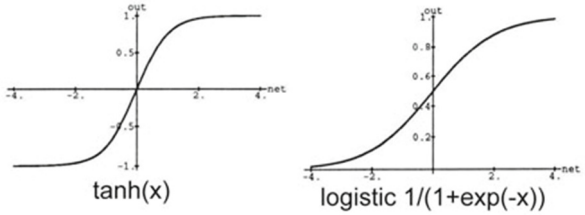
(a) Bir nöron ya da node (b) MLP

Şekil 1. (a) Bir nöron ya da node. (b) Çok katmanlı ileri beslemeli bir YSA modeli (MLP)

Biyolojik nöronların elektriksel aktivitesi tipik olarak bir dizi keskin hareketi takip eder, bunlara 'ani artışlar' denilebilir ve bir YSA düğümünün aktivasyonu başlangıçta, bu sivri uçların ortalama atış hızının modellemesini amaçlamıştır. Yıllar içinde teknoloji geliştikçe çok çeşitli yapay sinir ağları (YSA) modelleri ortaya çıkmıştır. Bu modeller bağlantıları döngüsel olanlar ve olmayanlar olmak üzere temel olarak ikiye ayrılmaktadır. Döngülere sahip YSA'lar geri beslemeli, özyinelemeli veya yinelenen sinir ağları olarak adlandırılır. Döngüsü olmayan YSA'lara ileri besleme sinir ağları (FNN'ler) denilmektedir. FNN'lerin iyi bilinen örnekleri arasında, algılayıcılar (perception) (Rosenblatt, 1958), radyal tabanlı fonksiyon ağları (Broomhead ve Low, 1988), Kohonen haritaları (Kohonen, 1989) ve Hopfield ağları (Hopfield, 1982) bulunur (Graves, 2012).

İleri besleme sinir ağlarının (FNN'ler) en yaygın kullanılan biçimi çok katmanlı algılayıcılardır (MLP). Çok katmanlı bir algılayıcıdaki (MLP) birimler, bağlantıların bir katmandan diğerine iletilmesiyle katmanlar halinde düzenlenir. Şekil 1(b)' deki gizli katmanlar desenleri çıktı katmanına aktarır ve bir döngü modeli bulunmadığı için MLP modeli olarak adlandırılmaktadır. Bu yapıda girdiler, girdi katmanını oluşturan birimleri aynı anda besler. Bu birimlerin ağırlıklı çıktıları (w_{ij}), nöronlara benzeyen "gizli katman" olarak bilinen ikinci katmanı eş zamanlı olarak besler. Gizli katmanların ağırlıklı çıktıları, başka bir gizli katmanın girdisi olabilir. Son gizli katmanın ağırlıklı çıktıları (w_{kj}), çıktı katmanını oluşturmada girdi olarak kullanılır. Gizli ve çıktı katmanlarındaki birimler, bazen nöronlar bazen de çıktı birimleri olarak adlandırılırlar (Han ve Kamber, 2001).

Yapay sinir ağları karmaşık, doğrusal olmayan değerleri tahmin gücü yüksek modeller olarak tanımlanmaktadır. Bu tahmin gücünü doğrusal olmayan bir aktivasyon fonksiyonu ile gerçekleştirir. Çok katmanlı bir algılayıcıda eğer belirli bir ağırlık değerleri kümesi mevcutsa, girdiden çıktı vektörlerine bir fonksiyon tanımlanır. Bu durumda ağırlıklar değiştirilerek, tek bir MLP ile birçok farklı fonksiyon sonucu elde edilebilmektedir. Ağlar arasındaki geçişleri belirten aktivasyon fonksiyonunun pek çok farklı türü olmasına rağmen yaygın olarak kullanılanlar Tanh ve Sigmoid fonksiyonudur. Bu fonksiyonlara ait grafiksel gösterim ve eşitlikler aşağıdaki gibidir. Bu fonksiyonların en önemli özelliği doğrusal olmamalarıdır.



(a) tanh(x) fonksiyonu (b) sigmoid fonksiyonu

Şekil 2. (a) tanh(x) fonksiyonu, (b) sigmoid fonksiyonu

Eş. 1 ve 2' de a_h gizli birim h için elde edilecek ağırlıklı toplam birimini, θ_h aktivasyon fonksiyonunu, b_h birimin son aktivasyonunu ve w_{ij} i. birimden j. birime olan etkiyi yani ağırlığı belirtmektedir.

$$a_h = \sum_{i=1}^{I} w_{ih} x_i \tag{1}$$

$$b_h = \theta_h(a_h) \tag{2}$$

$$tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{3}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

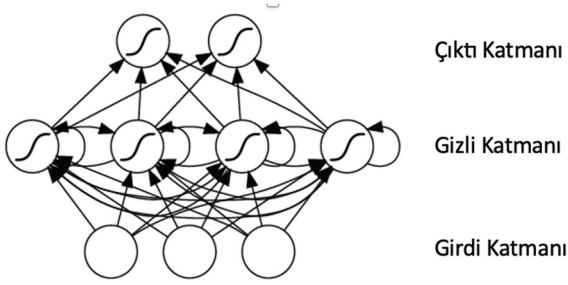
Bu fonksiyonların diğer en önemli özelliği ise her iki fonksiyonun da türevlenebilir olmasıdır, bu da ağın gradyan azaltma ile eğitilmesine olanak

tanımasıdır. Diğer bir aktivasyon fonksiyonu ise ReLU (Düzeltilmiş Doğrusal Birimler) ve Leakly ReLu fonksiyonudur.

ReLu Fonksiyonunda; bir girdi değeri sıfırdan büyük veya sıfıra eşitse değişmeden aynı kalır ve girdi değeri sıfırdan küçükse sıfır çıktısı alınır.

$$f(x) = \max(0, x) \tag{5}$$

Döngüler oluşturmayan ileri beslemeli sinir ağlarında, döngüsel bağlantılara da izin verilirse yinelemeli sinir ağları (RNN'ler) elde edilmektedir. Yinelemeli sinir ağlarına örnek olarak Elman ağları (Elman, 1990), Jordan ağları (Jordan, 1990), zaman gecikmeli sinir ağları (Lang, vd., 1990) ve yankı durumu ağları (Jaeger, 2001) verilebilir. Şekil 3'de örnek olarak kendi kendine bağlı tek bir gizli katman içeren basit bir RNN modeli bulunmaktadır (Graves, 2012).



Şekil 3. RNN modeli (Graves, 2012)

Çok katmanlı bir algılayıcı ile bir RNN arasındaki temel fark, dizi öğrenimi için elde edilen çıkarımlardadır. Bir MLP yalnızca girdiden çıktı vektörlerine eşleme yapabilirken, bir RNN önceki girdilerin tüm geçmişinden her çıktıya eşleme yapabilir. RNN modelindeki en önemli nokta, tekrarlayan bağlantıların önceki girdilerin bir 'hafızasının' ağın iç durumunda kalmasına izin vermesi ve böylece ağ çıktısını etkilemesidir.

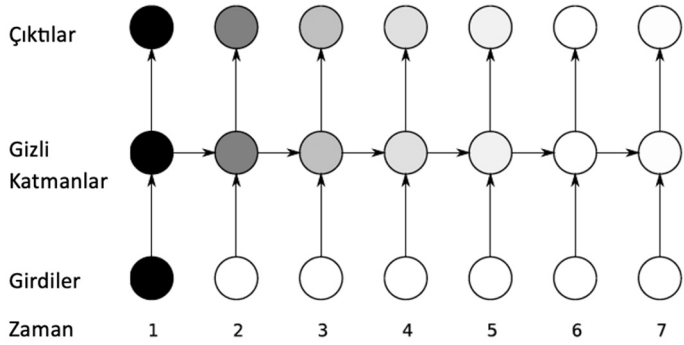
Bir RNN'nin ileri geçişi, tek gizli katmana sahip çok katmanlı bir algılayıcınınkiyle aynıdır, ancak aktivasyonlar gizli katmana, hem mevcut harici girdiden hem de önceki zaman adımından (t-1) gelen gizli katman aktivasyonlarından gelir. Bu durum Eş. (6-7) ile gösterilmektedir. Burada, I Girdi, H gizli, K çıktı birim sayısını, x_i^t t zamanındaki i girdisinin değerini, b_h^t t zamanında j birimine ağ girişini ve j biriminin t zamanındaki aktivasyonunu göstermektedir.

$$a_h^t = \sum_{i=1}^{I} w_{ih} x_i^t + \sum_{h'=1}^{H} w_{h'h} b_{h'}^{t-1} \tag{6}$$

$$b_h^t = \theta_h(a_h^t) \tag{7}$$

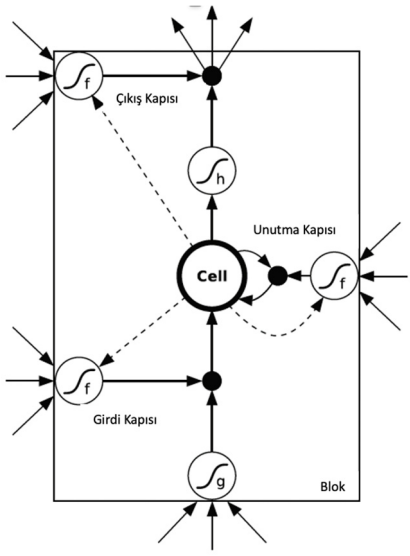
RNN & LSTM

Tekrarlayan sinir ağlarının önemli bir faydası, girdi ve çıktı katmanları arasında eşleme yaparken bağlamsal bilgileri kullanma yetenekleridir ancak standart RNN yapıları için, pratikte erişilebilen bağlam aralığı oldukça sınırlıdır. Sorun, belirli bir girdinin gizli olan üzerindeki etkisinin sınırlı olmasıdır. Ağın yinelenen bağlantılarında duyarlılık giderek azalır ve yeni girişler gizli katmanın ve ağın aktivasyonlarının üzerine yazıldığından zaman ilk girişler unutulur, bu duruma literatürde RNN'ler için kaybolan gradyan problemi denilmektedir (Graves, 2012). Kaybolan gradyan problemi Şekil 4'de şematik olarak gösterilmiştir. Açılmış ağdaki düğümlerin gölgelenmesi, birinci zamandaki girdilere olan duyarlılıklarını gösterir (gölge ne kadar koyuysa duyarlılık o kadar fazladır). Duyarlılık, yeni girdiler gizli katmanın aktivasyonlarını üzerine yazdıkça zamanla azalır ve ağ ilk girdileri 'unutur'. 1990'lı yıllarda RNN'ler için kaybolan gradyanlar sorununu çözmek için çok sayıda çalışmalar yapılmıştır ve bunlardan bir tanesi de 1997'de Hochreiter ve Schmidhuber tarafından önerilen Uzun kısa süreli bellek ağları (LSTM) modeli'dir.



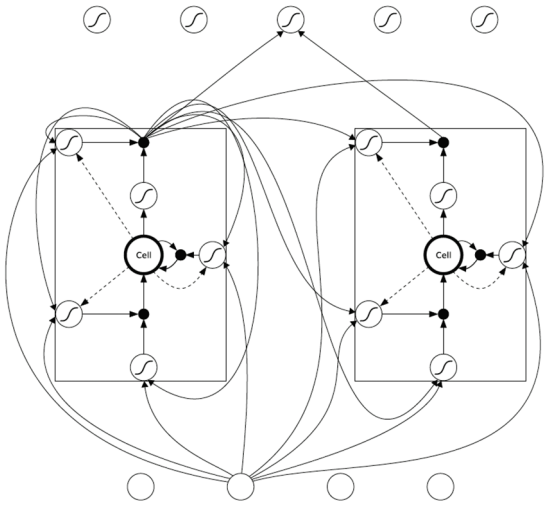
Şekil 4. RNN'ler için kaybolan gradyan problemi (Graves, 2012)

LSTM yapısı, tekrar tekrar bağlanan "bellek blokları (memory blocks)" olarak bilinen bir dizi alt ağdan oluşur . Bu modeller her biri, bir veya daha fazla tekrarlayan bağlı bellek hücreleri ve bu hücreler için sürekli yazma, okuma ve sıfırlama işlemlerini sağlayan girdi, çıktı ve unutma (forget) kapıları olan üç çarpımsal birim içerir. Ağ, hücrelerle ancak kapılar aracılığıyla etkileşime girebilir (Rothman, 2018). Şekil 5, tek hücreli bir LSTM bellek bloğunun bir örneğini göstermektedir.



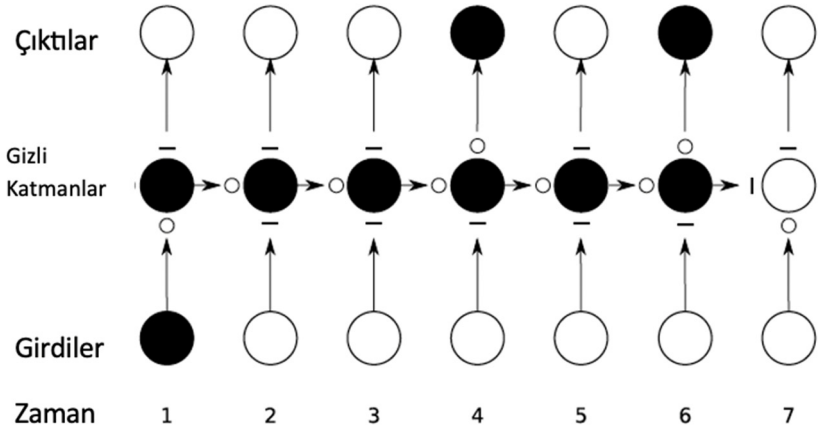
Şekil 5. Tek hücreli bir LSTM bellek bloğu (Graves, 2012)

Bir LSTM ağı, gizli katmandaki toplama birimlerinin Şekil 6'da gösterildiği gibi bellek bloklarıyla değiştirilmesi dışında standart bir RNN ile aynıdır. LSTM ağları için standart RNN'ler için olduğu gibi aynı çıktı katmanları kullanılabilir. Çarpımsal kapılar, LSTM bellek hücrelerinin uzun süreler boyunca bilgileri depolamasına ve erişmesine olanak tanır ve böylece kaybolan gradyan sorununu çözmüş olur. Binlerce zaman adımı sürebilen RNN için kısa süreli bir bellek sağlamayı amaçladığı için ağın ismi "Uzun kısa süreli bellek" ağları (LSTM)'dır. (Hochreiter, vd., 1997). LSTM tarafından gradyan bilgisinin zaman içinde korunması Şekil 7'de gösterilmiştir.



Şekil 6. LSTM ağı (Graves, 2012)

Örnek olarak verilen LSTM ağı, dört girdi biriminden, bir gizli katmanlı iki tek hücreli LSTM bellek bloğundan ve beş çıktı biriminden oluşmaktadır. Şekilde her bloğun dört girişi ve bir çıkışı olduğu görülmektedir ancak tüm bağlantılar gösterilmemiştir (Graves, 2012).



Şekil 7. LSTM ile gradyan bilgisinin korunması (Graves, 2012)

LSTM ağına veri girişi, (gözlem sayısı) x (zaman adımı sayısı) x (özellik sayısı) şeklindedir. Zaman adımları boyutunun eklenmesi, modelin yalnızca eşzamanlı gözlemler yerine her değişkenin birden fazla gecikmesine göre eğitilmesine olanak tanır (Hoop, 2021). Ayrıca LSTM modellerinde bir

diziden bir gözlem ağa girdiğinde, bu gözlemin daha önce ağ yapısında gördüğü hangi gözlemlerle alakalı olduğunu, bunların bir tahmin ile ilişkisini hatırlayabilir. Tüm bu nedenlerle LSTM modelleri, verilerden zamansal bağımlılığı öğrenebilir ve kullanabilir. Son yıllarda zaman serileri analizinde özellikle öngörü modellerinde LSTM modellerinin tercih edilmesinin sebebi bu bağımlılığı kolaylıkla öğrenebilmesidir.

Zaman serilerinde "nowcasting" yani "anlık" tahmin için LSTM modellerinin tercih edilmesinin sebebi ekonomik verilerin yayınlanma frekanslarının farklılık göstermesidir. LSTM modelleri yapısı gereği, doğrusal olmayan aktivasyon fonksiyonları ile çok farklı frekanslara (dakikalık-saatlik-günlük- haftalık-aylık-üç aylık) sahip değişkenlerin karmaşık, doğrusal olmayan ilişkilerini kolaylıkla öğrenebilmektedir.

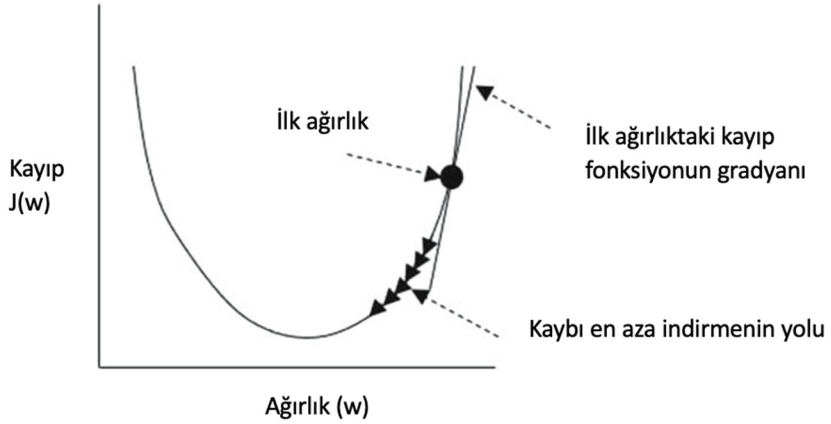
Geri Yayılım Algoritması ve Gradyan Azaltma

Yapay sinir ağları insan zekası gibi örneklerle eğitilirler ve öğrenme türüne göre de denetimli ve denetimsiz olmak üzere ikiye ayrılmaktadır. Denetimli öğrenme yapısında ağa ait hem girdi değerleri, hem de çıktı değerleri sunulur ve her iki iterasyonda örneğe ait çıktı değerleriyle ağ çıktı değeri karşılaştırılarak ağın hatası bulunur. Bu hata kontrol edilebilir dereceye indirilene kadar, yapay sinir ağı nöronlar arasındaki ağırlıkları değiştirerek iterasyona devam eder. Denetimsiz öğrenmede ise ağa sadece girdi veri grubu sunulur ve ağdan bu veri grubuna uyumlu bir çıkış değeri üretecek şekilde kendisini uygun ağırlıklarla düzenlemesi istenir (Özdemir, 2013).

Yapay Sinir Ağlarının elde ettiği çıktı yani tahmin değerlerinin doğruluğunu artırmak için ağın nasıl daha iyi öğrenebileceği konusunda çalışmalar yapılmıştır. İlk aşamada yani ağın eğitimi kısmında tüm ağırlıklar rasgele başlatılır ve ilk tahminler elde edilir. Aynı veriler ile başlangıç değerleri ve başlangıç ağırlıkları farklı olduğunda, birden fazla YSA modeli elde edilebilir ancak bu tahminlerin performansları düşük kalitede olacaktır çünkü ağırlıklar rasgele seçilmiştir. Modelin öğrenmesi, yani ağırlıklarını güncellemesi ve tahminlerinin kalitesini artırması için iki adım daha gereklidir. Bu adımlar geri yayılım ve gradyan azaltmadır (Hoop, 2022).

Geri yayılım, bir yapay sinir ağı öğrenme algoritmasıdır. Yapay sinir ağları aslında girdi ve çıktı birimlerinin bağlantılı bir kümesidir ve her bir bağlantı girdi ve çıktılarla ilgili olan bir ağırlığa sahiptir. Öğrenme algoritmasında sinir ağı, girdileri doğru sınıflandırmak için ağırlıkları her aşamada ayarlar. Bu nedenle sinir ağları, birimler arası bağlantıya bağlı olarak "bağlantılı öğrenme" olarak da adlandırılır (Han ve Kamber, 2001). Geri

yayılım algoritması, çok katmanlı ileri-besleme sinir ağ yapılarında öğrenme ile oluşur. Ayrıca bir ağ, hatayı veya kaybı en aza indirmek için eğitilmeden önce, bu hatayı belirlemek için bir kayıp fonksiyonu seçilmelidir. Seçilen kayıp fonksiyonunun türü uygulamaya bağlıdır. Örneğin, regresyon problemleri için yaygın kayıp fonksiyonları arasında ortalama mutlak hata (MAE) ve ortalama karesel hata (MSE) bulunmaktadır. Geri yayılım algoritması da, belirlenen kayıp fonksiyonunun eğimlerini veya kısmi türevlerini, ağdaki her katmanın ağırlıklarına göre zincir kuralını kullanarak hesaplayan bir algoritmadır.



Şekil 8. Tek bir ağırlıkla geri yayılım ve gradyan azalma grafiksel gösterimi, Hoop (2022)

Şekil 8, tek bir ağırlığa sahip bir ağdaki kayıp fonksiyonunun grafiğini göstermektedir. Amaç kaybı en aza indirmektir. Bu basit durumda, türevimizi sıfıra eşitleyebilir ve ağırlık için çözebiliriz. Ancak, daha karmaşık ağlarda, kayıp fonksiyonunun kapalı form türevi yoktur, bu nedenle kaybı en aza indirmek için bunun yerine gradyan azaltma kullanılır. Başlangıç ağırlığımızda, gradyan geri yayılım algoritması aracılığıyla hesaplanır, bu bilgi daha sonra kaybı azaltmak için ağırlığın hangi yöne hareket ettirileceğini belirlemek için kullanılır, Şekil 8'deki fonksiyon minimumuna işaret eden ok kümesiyle gösterilir (Hoop, 2022). İşlem ağırlıkların yenilenmesi ile daha sonra tekrarlanır ve ağırlıklar birbirine yakınsamaya başlayınca son değerlerde sabitlenir ve ağ eğitilmiş olarak kabul edilir. Daha sonra yeni veriler beslenerek yeni çıktılar veya tahminler elde edilebilir.

Uygulama

Bu bölümde zaman serilerinde nowcasting amaçlı kullanılan klasik ekonometrik yöntemler ile makine öğrenmesi yöntemlerinin kullanımına yönelik iki farklı örnek uygulama gerçekleştirilmiştir. Birinci uygulama yüksek frekanslı bir seri olan Türkiye'nin elektrik tüketimi miktarının (kWh) saatlik olarak tahmin edilmesine ilişkindir. Bu uygulamada Box-Jenkins ARIMA yöntemi ile LSTM modeli tercih edilmiştir. İkinci uygulama ise çeyreklik dönemlerde yayınlanan mevsimsel düzeltilmiş GSYH büyüme oranının aylık olarak tahmin edilmesine yönelik bir çalışmadır. Çalışmada TRAMO/ SEATS ve LSTM modelleri sonuçları karşılaştırılmıştır. Uygulama kısmında E-views 15, TSW programları ile Pyhton matplotlib, numpy, pandas ve torch kütüphanelerinden yararlanılmıştır[2].

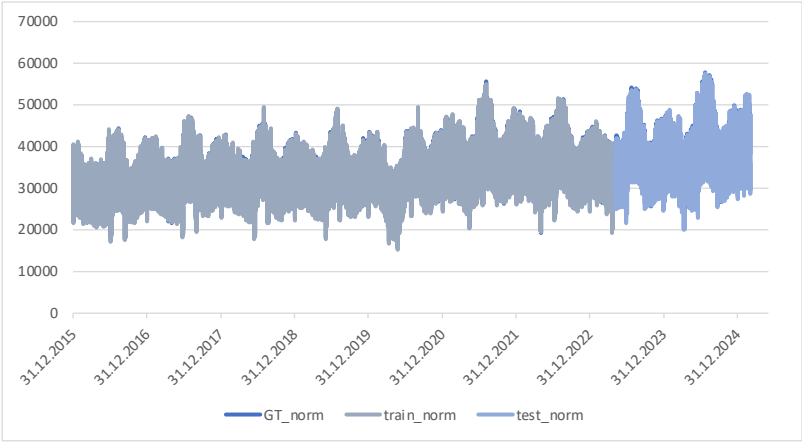
Saatlik Elektrik Tüketim Tahmini

Elektrik enerjisi, günümüzde vazgeçilmez bir enerji kaynağı olup her geçen gün bu enerjiye duyulan ihtiyaç artmaktadır. Sürdürülebilir Kalkınma Göstergelerinden 7. Maddede herkes için karşılanabilir, güvenilir, sürdürülebilir ve modern enerjiye erişimi sağlamak amaçlanmıştır. 2030'a kadar uygun fiyatlı, güvenilir ve modern enerji hizmetlerine evrensel erişimin sağlanması birinci hedeftir. Bu kapsamda üretilen elektriğin makul fiyatlarla tüketiciye sunulabilmesi için ne kadar elektrik tüketileceğinin daha önceden tahmin edilmesi gerekmektedir. Literatürde elektrik tüketim tahminine ilişkin birçok yayın bulunmakla birlikte genellikle tercih edilen modeller Yapay Sinir Ağları, Otoregresif Hareketli Ortalamalar ve Basit/Çok Değişkenli Regresyon Modelidir.

Bu çalışmada ise saatlik elektrik tüketiminin Box- Jenkins ARIMA ve LSTM modeli ile tahmin edilmesi amaçlanmaktadır. Enerji piyasalarının şeffaf, güvenilir, adil ve öngörülebilir bir şekilde işletilmesi için gerekli verileri sağlayan EPİAŞ Şeffaflık Platformundan yararlanılmıştır. 31.12.2015 04:00 - 04.03.2025 12:00 saatleri arasındaki Türkiye geneli elektrik tüketimi verisi %80 eğitim ve %20 test verisi olacak şekilde iki kısma ayrılmıştır. Bu kapsamda 31.12.2015 04:00 – 04.05.2023 04:00 arası 64542 kayıt eğitim veri setini, 04.05.2023 05:00 - 04.03.2025 12:00 arası 16137 kayıt test verisidir. Veri 50000'e bölünerek normalizasyon işlemi yapılmış, model sonuçları elde edildikten sonra ters işlem gerçekleştirilmiştir. Elektrik tüketimi için tek girdi

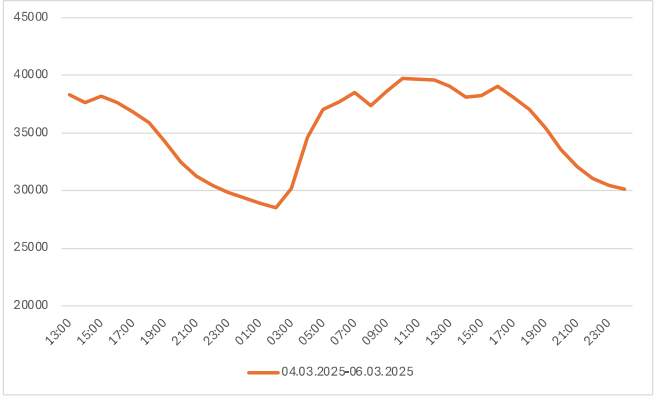
[2] Değerli katkılarından dolayı Çukurova Üniversitesi Bilgisayar Mühendisliği Bölümünden Dr. Öğr. Üyesi Mehmet Sarıgül'e teşekkür ederim.

olarak kendi serisi ele alınmıştır. İlk çalışma olarak ilgili zaman serisinin 24 saat gecikmeli değerlerini hesaba alan LSTM modeli için lockback katsayısı 24 seçilmiştir. 100 gizli, 1 girdi , 1 çıktı katmanından oluşan 2 tabakalı LSTM modelinde aktivasyon fonksiyonu olarak ReLu seçilmiştir. Kayıp fonksiyonu için MSE değerleri hesaplanarak model en küçük MSE değerine erişen aşamada eğitimi durdurmaktadır. Başlangıç olarak epoch deneme sayısı 100 olarak belirtilmiştir. Eğitim modeline ilişkin MSE 0,000167562; test modeline ilişkin MSE 0,000395777 olarak hesaplanmıştır.



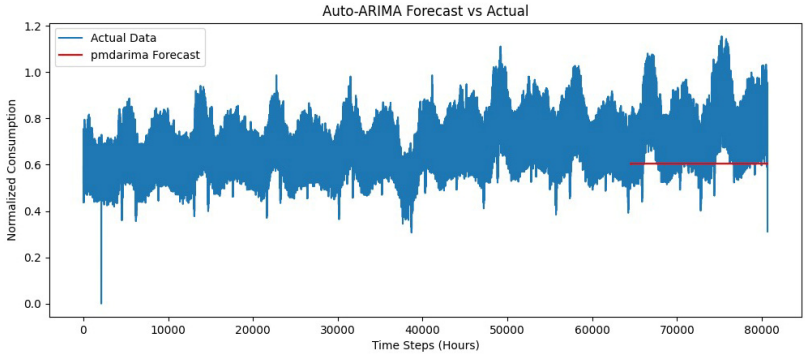
Şekil 9. Elektrik Tüketimine ilişkin LSTM eğitim ve test sonuçları

04.03.2025 13:00 - 06.03.2025 00:00 tarihleri arasında saatlik elektrik tüketimi LSTM kısa dönemli şimdiki / anlık tahmin sonuçları Şekil 10' da belirtilmiştir. Beklendiği gibi saatlik elektrik tüketim dalgalanmalarının gece saatlerine düşmesi ve sonrasında artması ile döngüsel beklenti karşılanmıştır.

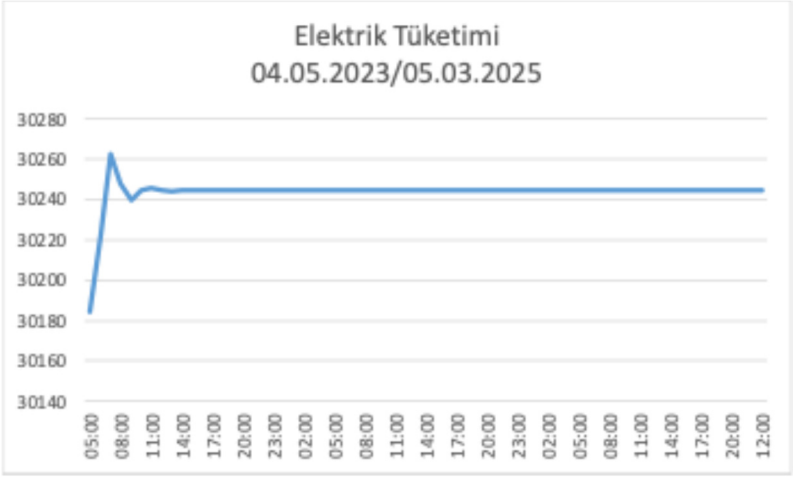


Şekil 10. Elektrik Tüketimine ilişkin LSTM kısa dönemli anlık tahmin sonuçları

İkinci olarak en uygun ARIMA modelinin seçimi için Python pmdarima kütüphanesinden yararlanılmış ve veri seti LSTM modelinde olduğu gibi iki kısma ayrılmıştır. Elde edilen model denemeleri sonucunda en düşük AIC değeri (AIC:-319276.991) ile ARIMA(2,1,3)(0,0,0) modeli seçilmiştir. İlgili modele ait tahmin sonuçları grafiği Şekil 11 ve 12'de görülmektedir. Model sonuçları incelendiğinde tahmin sonuçlarının 04.05.2023 saat 16:00 dan itibaren aynı olduğu dikkat çekicidir. ARIMA modelinin saatlik düzeyde ölçülen büyük veri kapsamında sayılan bu tip bir örnek uygulama için tahmin performansının düşük kaldığı grafiklere bakarak söylenebilir. ARIMA modelleri kısa dönemli tahminleri için tutarlı sonuç verebilmektedir ancak serinin sonraki hareketlerini tahmin etmekte geri kalmaktadır. Bu nedenle yüksek frekanslı serilerdeki hareketi yakalamakta etkili olamamaktadır.



Şekil 11. Elektrik Tüketimine ilişkin ARIMA model tahmin sonuçları



Şekil 12. Elektrik Tüketimine ilişkin LSTM kısa dönemli anlık tahmin sonuçları

Aylık GSYH büyüme oranı tahmini

GSYH büyümesi, GSYH rakamları ekonomik aktivite hakkında kapsamlı bilgi sağladığından en önemli ekonomik göstergelerden biridir. GSYH verileri önemli bir gecikmeyle yayınlanır, bu nedenle GSYH büyümesinin erken tahminleri değerli olabilir. Bu amaçla literatürde çok sayıda GSYH büyüme oranının tahminine ilişkin bir çok çalışma bulunmaktadır. Merkez Bankası çalışma notları kapsamında Akkoyun ve Günay (2012) tarafından hazırlanan çalışmada, Stock ve Watson eş zamanlı gösterge modelinin genişletilmiş bir versiyonu kullanılarak Türkiye için GSYH büyümesi tahmin edilmiştir. Bu modelin seçilmesinin amacı karışık frekans (çeyreklik ve aylık değişkenler gibi), düzensiz uçlar (bazı göstergelerin diğerlerinden önce yayınlanması) ve eksik verilerle (bazı değişkenlerin veri örneklemin başında mevcut olmaması) başa çıkabilen bir yöntem olmasıdır. İlgili modelde PMI endüstriyel üretim, ithalat ve ihracat miktar endekslerinden yararlanılarak örnek bir çalışma yapılmıştır. İlgili çalışmanın iktisadi modeli temel alınarak, ülkemize ilişkin çeyreklik dönemlerde zincirleme hacim endeksi yöntemi ile hesaplanan GSYH (2009=100) büyüme oranının aylık olarak öngörülmesi için, tahmin edici değişken olarak aylık İhracat Değeri (IHR), İthalat Değeri(ITH) ve Sanayi Üretim Endeksi (SUE) (2010=100) değişkenleri seçilmiştir. Tüm değişkenler mevsim ve takvim etkisinden arındırılmış (sa) bir önceki aya göre büyüme oranlarıdır (TÜİK, 2025).

Tablo 1. Değişkenlerin başlangıç-bitiş tarihi, frekans türü ve gözlem sayıları

	Frekans	Başlangıç Tarihi	Bitiş Tarihi	Gözlem sayısı
GSYH(sa)	Çeyreklik	2002:Q1	2024:Q4	92
IHR (sa)	Aylık	2002:02	2024:12	263
ITH (sa)	Aylık	2002:02	2024:12	263
SUE (sa)	Aylık	2002:02	2025:01	264

Birinci aşamada GSYH büyümesinin anlık tahmini için ekonometrik model olarak TRAMO/SEATS yöntemi seçilmiştir. Temelleri Burman (1980) ve Hillmer ve Tiao (1982) tarafından atılan ARIMA modeline dayalı mevsimsel düzeltme yöntemini uygulayan TRAMO[3]/SEATS mevsimsel düzeltme prosedürü 1997 yılında İspanya Merkez Bankasından Gomez ve Maravall tarafından geliştirilmiştir.

[3] **TRAMO:** Time Series Regression with ARIMA Noise, Missing Observations, and Outliers/Signal Extraction in ARIMA Time Series

TRAMO kısmında Reg-ARIMA modelleme tekniği kullanılarak gözlenen zaman serisine;

- aykırı değer düzeltmesi,
- takvim etkisi düzeltmesi,
- kayıp değerlerin tahminleri uygulanarak seri "doğrusal" hale getirilmektedir.

SEATS aşamasında ise, Trend-Çevrim, Mevsimsel, Düzensiz bileşenler gibi gözlemlenemeyen bileşenleri tahmin etmek için ARIMA model tabanlı metod kullanılmaktadır.

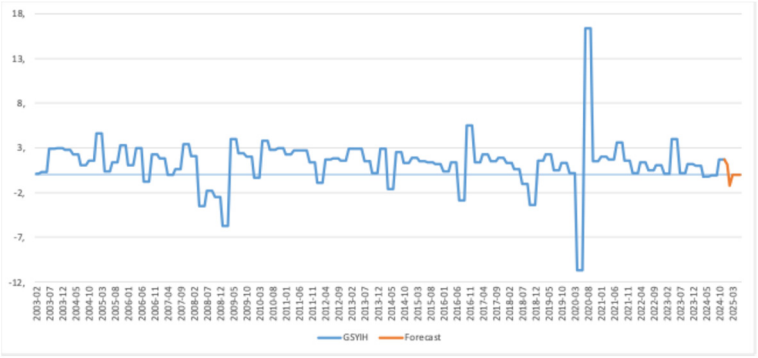
TRAMO sürecinde gözlenen bir zaman serisi deterministik ve stokastik kısım olarak iki parçaya ayrılmaktadır. Bu aşamada parametre tahmininde En Yüksek Olabilirlik Yöntemi kullanılmaktadır. Stokastik kısım ön düzeltmesi yapılmış zaman serisinin SARIMA (p,d,q) (P,D,Q)s olarak modellenmesidir. Reg-ARIMA kısmında ise parametre tahmininde Genelleştirilmiş En Küçük Kareler Yöntemi ile parametre spesifikasyonları için Kalman (1960) Filtresi kullanılmaktadır (Kaiser ve Maravall, 2001: 44; Marini ve Moauro, 2006: 4; Gomez ve Maravall, 1997: 2).

TRAMO modelinde öncelikle bağımlı değişken ile bağımsız değişkenlerin aynı frekans düzeyine getirilmesi gerekmektedir. Bu nedenle çeyreklik GSYH büyüme verisi üç aylık ortalamaları eşit olacak şekilde aylık düzeyde tekrar düzenlenmiştir. Bu aşamada birçok TRAMO modeli denemesi gerçekleştirilmiş ve ekonometrik & iktisadi olarak model varsayımlarını sağlayan en anlamlı model nihai model olarak seçilmiştir. Elde edilen TRAMO modeli ARIMA (0,1,1) (0,0,0) dağılımına sahip olup,

$$gsyh_t = \underset{\substack{0.4405E\text{-}01 \\ (t:2.56)}}{ith_{t-1}} + \underset{\substack{0.3764E\text{-}01. \\ (t:2.18).}}{ith_{t-2}} + \underset{\substack{0.5012E\text{-}01 \\ (t:2.99).}}{ihr_{t-1}} + \underset{\substack{0.7428E\text{-}01 \\ (t:3.97)}}{ihr_{t-2}} + \underset{\substack{0.3816E\text{-}01 \\ (t:2.52)}}{ihr_{t-3}} + \underset{\substack{0.1676 \\ (t:6.24).}}{sue_{t-1}} + \underset{\substack{0.3072 \\ (t:11.46)}}{sue_{t-2}}$$

şeklinde modellenmiştir (BIC: 0,782697). Tahmin edilen modele ilişkin zaman serisi grafiği Şekil 13'de belirtilmiştir. Bu modele ait TRAMO aşamasında belirlenen aykırı değerler[4] ise dönemleri ve katsayı anlamlılıkları ile birlikte AO01(042020, -7.72), AO02(092020, 6.69), TC01(042008, -4.76), LS01(072023, -7.18) şeklindedir.

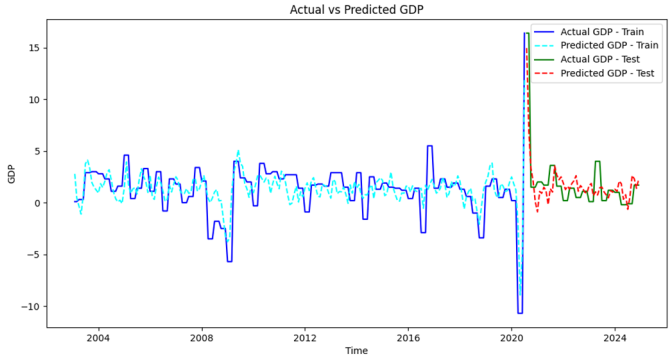
[4] AO: Toplamsal Aykırı Değer, LS: Seviye Kayması , TC: Geçici Değişim aykırı değerlerini ifade etmektedir.



Şekil 13. GSYH aylık tahminine ilişkin TRAMO modeli sonuçları

İkinci aşamada GSYH büyüme tahmini için LSTM modeli uygulanmıştır. LSTM modelinde ise veri seti %80 eğitim ve %20 test verisi olacak şekilde iki kısma ayrılmıştır. Bu kapsamda her bir değişkene ilişkin belirlenen eğitim ve test verisi kayıt sayıları ve tarih aralıkları Tablo 2'de belirtilmiştir. İlgili zaman serilerinin 12 aylık gecikmeli değerlerini hesaba alan LSTM modeli için lockback katsayısı 12 seçilmiştir. 100 gizli, 3 girdi , 1 çıktı katmanından oluşan 3 tabakalı LSTM modelinde aktivasyon fonksiyonu olarak ReLu seçilmiştir. Kayıp fonksiyonu için MSE değerleri hesaplanarak model en küçük MSE değerine erişen aşamada eğitimi durdurmaktadır. Başlangıç olarak epoch deneme sayısı 50 olarak belirtilmiştir.

Şekil 14'de 2020 Nisan / Covid dönemindeki büyüme oranındaki ani düşüş ve sonrasında ani yükşelişi model tahmin edebilmişken, 2021:1- 2023:9 döneminde tahmin edilen ve gerçekleşen büyüme oranlarının farklılaştığı görülmektedir. Bununla birlikte 2023:9 ve 2024:12 dönemi için büyüme oranlarında gerçekleşen değere yakın tahminler elde edilmiştir. Eğitim ve test modeline ilişkin MSE değerleri Tablo 2'de belirtilmiştir.

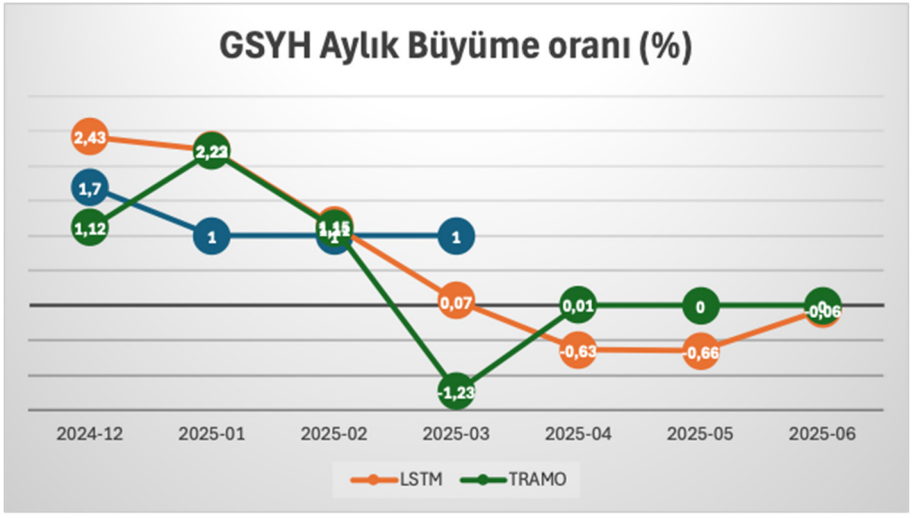


Şekil 14. GSYH aylık tahminine ilişkin LSTM eğitim ve test modeli sonuçları

Tablo 2. Modele ait test, eğitim ve şimdiki/ anlık tahmin dönemleri

	Tarih Aralığı	Gözlem sayısı	MSE
Eğitim	2003:02 - 2020:07	211	3,497682
Test	2020:08 - 2024:12	53	3,239892
Tahmin	2025:01 - 2028:12	48	

2025:01- 2025:04 dönemine ilişkin aylık GSYH büyüme oranı (%) tahmin sonuçları Şekil 15'de gösterilmektedir. 2024:4. çeyrek için gerçekleşen GSYH büyüme oranı 1.7 iken, 2024:12. ay için LSTM modeli 2.43, TRAMO modeli 1.12 olarak tahmin edilmiştir. 2025: 1. Çeyrek için gerçekleşen GSYH büyüme oranı 1.7 iken, 2025 yılı ilk üç ayın ortalama büyüme tahmini LSTM modeli için 1.15, TRAMO modeli için 0.7 olarak tahmin edilmiştir. 2025:01 - 2025:06 dönemine ilişkin aylık GSYH büyüme oranı (%) tahminlerine bakıldığında LSTM model sonuçlarının gerçekleşen çeyreklik büyüme oranına yakın tahminler ürettiği söylenebilir.



Şekil 15. GSYH büyümesine ilişkin TRAMO ve LSTM aylık öngörü değerleri

Sonuç

21. yüzyılın şimdiye kadarki belirleyici bir özelliği, üretilen ve depolanan verilerin hem hacimlerinde hem de çeşitlerinde yaşanan patlama olmuştur (Domo, 2017). Bu yeni veri miktarlarını analiz etmek için yeni yöntemlere duyulan ihtiyacın da etkisiyle Makine öğrenimi ve Yapay Zeka yöntemlerinde hızlı ilerlemeler kaydedilmiştir.

Makroekonomik serilerin zamanında ve doğru tahminleri, politika kararlarını bilgilendirmede oldukça önemlidir. Bu çalışma, şimdiki zaman tahmininde önerilen LSTM modeli ile klasik ekonometrik yöntemlerin kıyaslanması fikrine dayanmaktadır. Bu amaçla Türkiye Elektrik Tüketimi ve GSYH büyüme oranı anlık tahminleri için örnek ekonometrik yöntemler olarak TRAMO / SEATS yöntemi ile ARIMA modelleri seçilmiştir. Her iki uygulamada da LSTM model tahminleri görece daha yüksek tahmin performansı göstermiştir. Bunun sebebi LSTM modellerinin veriler arasındaki zamansal bağımlığı kolaylıkla öğrenebilmesidir. Son zamanlarda çalışmalarda tercih edilme sebepleri ise eksik gözlemler, düzensiz uç değerler ve karışık frekanslı büyük verilerle çalışabilmesidir.

Zaman serilerinin analizi ve tahminlerine yönelik Yapay Zeka algoritmaları ile gelecekte yapılacak araştırma ve geliştirmeler için çok fazla alan bulunmaktadır. LSTM'lerin daha geniş bir dizi ve frekans karışımındaki performansını doğrulamak için daha fazla veri setinden farklı testler yapılmalıdır. Model mimarisinin diğer yönlerinin daha iyi sonuçlarla sonuçlanıp sonuçlanmayacağını görmek için daha fazla model parametre ayarı yapılabilir. Pyhton yazılımında Keras ve PyTorch dahil olmak üzere birçok yaygın derin öğrenme algoritmaları gelişmeye devam etmektedir. Ayrıca Hoop tarafından PyPi kütüphanesi yayınlanmıştır (Hoop, 2021). Son olarak, LSTM'leri yorumlamak ve tahminlere olan katkısını belirlemek için klasik zaman serisi yöntemlerine ek olarak ele alınması modellerin yorumlanabilirliğini artıracaktır.

Kaynakça

Akkoyun, H. Ç., & Günay, M. (2012). Nowcasting Turkish GDP growth (Working Paper No. 12/33). Turkish Central Bank.

Antolin-Diaz, J., Drechsel, T., & Petrella, I. (2020). Advances in nowcasting economic activity: Secular trends, large shocks and new data. European Central Bank.

Amidi, A., & Amidi, S. (2019). Recurrent neural networks cheatsheet. Retrieved December 20, 2020, from https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real-time measurement of business conditions. Journal of Business and Economic Statistics, 27(4), 417–427.

Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the euro area GDP. International Journal of Forecasting, 20(3), 447–460.

Bańbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. Journal of Applied Econometrics, 25(1), 71–92.

Banbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). Now-casting and real-time data flow (Working Paper Series No. 1564). European Central Bank. https://doi.org/10.2139/ssrn.2201605 (eğer bu DOI geçerliyse – doğrulanması gerekir)

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157–166.

Benidis, K., Rangapuram, S. S., Flunkert, V., Wang, Y., Maddix, D. C., Turkmen, C., ... Januschowski, T. (2022). Deep learning for time series forecasting: Tutorial and literature survey. ACM Computing Surveys, 55(6), Article 121, 1–36. https://doi.org/10.1145/3533382

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford University Press.

Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M., & Tambalotti, A. (2018). Macroeconomic nowcasting and forecasting with big data. Annual Review of Economics, 10(1), 615–643.

Broomhead, D., & Lowe, D. (1988). Multivariate functional interpolation and adaptive networks. Complex Systems, 2, 321–355.

Burman, J. P. (1980). Seasonal adjustment by signal extraction. Journal of the Royal Statistical Society: Series A (General), 143, 321–337.

Cantú, F. (2018). Estimation of a coincident indicator for international trade and global economic activity (UNCTAD Research Paper No. 27). United Nations Conference on Trade and Development (UNCTAD).

Cimadomo, J., Giannone, D., Lenza, M., Sokol, A., & Monti, F. (2020). Nowcasting with large Bayesian vector autoregressions (Working Paper Series No. 2453). European Central Bank.

Domo. (2017). Data never sleeps 5.0. Retrieved September 1, 2021, from https://www.domo.com/learn/data-never-sleeps-5

Dematos, G., Boyd, M. S., Kermanshahi, B., Kohzadi, N., & Kaastra, I. (1996). Feedforward versus recurrent neural networks for forecasting monthly Japanese yen exchange rates. Financial Engineering and the Japanese Markets, 3(1), 59–75.

Dinh, L., Sohl-Dickstein, J., & Bengio, S. (2017). Density estimation using Real NVP. In International Conference on Learning Representations (ICLR 2017), Conference Track. Retrieved from https://openreview.net/forum?id=HkpbnH9lx

Dynan, K., & Sheiner, L. (2018). GDP as a measure of economic well-being (Hutchins Center Working Paper). Brookings Institution.

Elman, J. L. (1990). Finding structure in time. Cognitive Science, 14, 179–211.

Evans, M. D. D. (2005). Where are we now? Real-time estimates of the macroeconomy. International Journal of Central Banking, 1(2).

Farahani, M. A., Kalach, F. E., Harper, A., et al. (2025). Time-series forecasting in smart manufacturing systems: An experimental evaluation of the state-of-the-art algorithms. Robotics and Computer–Integrated Manufacturing, 95, 103010. Retrieved from https://www.elsevier.com/locate/rcim

Gomez, V., & Maravall, A. (1996). Programs TRAMO (Time series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series), Instructions for the user (Working Paper No. 9628). Servicio de Estudios, Banco de España.

Gökçin, G. M. (2019). Çimento ve taş kömürü örnek sektörlerinde stok miktarı tahmini için sınıflama tipi modeller ile çoklu regresyon modellerinin karşılaştırılması (Doctoral dissertation). Ankara.Graves, A, "Supervised Sequence Labelling with Recurrent Neural Networks", Springer 2012, 15-44.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics, 55(4), 665–676.

Grosse, R. (2017). Exploding and vanishing gradients. Retrieved December 19, 2020, from http://www.cs.toronto.edu/~rgrosse/courses/csc321_2017/readings/L15%20Exploding%20and%20Vanishing%20Gradients.pdf

Guichard, S., & Rusticelli, E. (2011). A dynamic factor model for world trade growth (OECD Economics Department Working Papers No. 874). OECD Publishing.

Han, J., & Kamber, M. (2001). Data mining: Concepts and techniques (3rd ed.). San Francisco, CA: Morgan Kaufmann Publishers.

Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen (Doctoral dissertation). Institut für Informatik, Technische Universität München.

Hochreiter, S., & Bengio, Y., Frasconi, P., & Schmidhuber, J. (2001). Gradient flow in recurrent nets: The difficulty of learning long-term dependencies. In S. C. Kremer & J. F. Kolen (Eds.), A field guide to dynamical recurrent neural networks. IEEE Press.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Hoop, D. (2021). Economic nowcasting with long short-term memory artificial neural networks (LSTM) (UNCTAD Research Paper No. 62).

Hoop, D. (2022). Benchmarking econometric and machine learning methodologies in nowcasting (UNCTAD Research Paper No. 83).

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences of the United States of America, 79(8), 2554–2558.

Houdt, G. V., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. Artificial Intelligence Review, 53, 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

International Monetary Fund (IMF). (2020). Gross domestic product: An economy's all. Retrieved February 13, 2022, from https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm

Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks (Technical Report GMD Report 148). German National Research Center for Information Technology.

Hillmer, S. C., & Tiao, G. C. (1982). An ARIMA-model-based approach to seasonal adjustment. Journal of the American Statistical Association, 77(377), 63–70. https://doi.org/10.2307/2287770

Hsu, D. (2017). Time series forecasting based on augmented long short-term memory. arXiv preprint arXiv:1707.00666.

Jordan, M. I. (1990). Attractor dynamics and parallelism in a connectionist sequential machine (pp. 112–127). IEEE Press.

Kaiser, R., & Maravall, A. (2000). Notes on time series analysis, ARIMA models and signal extraction (Banco de España Working Paper No. 0012).

Kapoor, A., & Debroy, B. (2019). GDP is not a measure of human well-being. Retrieved February 10, 2022, from https://hbr.org/2019/10/gdp-is-not-a-measure-of-human-well-being

Kohonen, T. (1989). Self-organization and associative memory (3rd ed.). Springer.

Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996). A comparison of artificial neural network and time series models for forecasting commodity prices. Neurocomputing, 10(2), 169–181.

Kuzin, V. N., Marcellino, M., & Schumacher, C. (2009). MIDAS versus mixed-frequency VAR: Nowcasting GDP in the euro area (Discussion Paper Series 1: Economic Studies No. 2009/07). Deutsche Bundesbank.

Lang, K. J., Waibel, A. H., & Hinton, G. E. (1990). A time-delay neural network architecture for isolated word recognition. Neural Networks, 3(1), 23–43.

Laptev, N., Yosinski, J., Li, E., & Smyl, S. (2017). Time-series extreme event forecasting with neural networks at Uber. In ICML Time Series Workshop.

Lara-Benítez, P., Carranza-García, M., & Riquelme, J. C. (2021). An experimental review on deep learning architectures for time series forecasting. International Journal of Neural Systems, 31(3), 2130001. https://doi.org/10.1142/S0129065721300011

McCulloch, W. S., & Pitts, W. (1988). A logical calculus of the ideas immanent in nervous activity (pp. 15–27). MIT Press.

Maravall, A. (2008). *Notes on programs TRAMO and SEATS, Part I: Some basic concepts in parametric time series analysis*. Bank of Spain.

Marcellino, M., & Schumacher, C. (2010). Factor MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP. *Oxford Bulletin of Economics and Statistics, 72*(4), 518–550.

Mariano, R., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics, 18*, 427–443.

Ozdemir, O., & Yozgatlıgil, C. (2022). Forecasting performance of machine learning, time series, and hybrid methods for low- and high-frequency time series. *Statistica Neerlandica*. https://doi.org/10.1111/stan.12326

Özdemir, H. (2013). Yapay sinir ağları ve dokuma teknolojisinde kullanımı. *Tekstil Teknolojileri Elektronik Dergisi, 7*(1), 51–68.

Öztemel, E. (2012). *Yapay sinir ağları*. İstanbul: Papatya Yayıncılık.

Papamakarios, G., Pavlakou, T., & Murray, I. (2017). Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems, 30*.

Patel, M., Patel, A., & Ghosh, R. (2018). *Precipitation nowcasting: Leveraging bidirectional LSTM and 1D CNN*. arXiv preprint. https://doi.org/10.48550/arXiv.1810.10485

Rahimi, R., Ravirathinam, P., Ebtehaj, A., Behrangi, A., Tan, J., & Kumar, V. (2023). *Global precipitation nowcasting of integrated multi-satellite retrievals for GPM: A U-Net convolutional LSTM architecture*. arXiv preprint. https://doi.org/10.48550/arXiv.2307.10843

Rasul, K., Sheikh, A. S., Schuster, I., Bergmann, U., & Vollgraf, R. (2020). *Multivariate probabilistic time series forecasting via conditioned normalizing flows*. arXiv preprint arXiv:2002.06103.

Rossiter, J. (2010). *Nowcasting the global economy* (No. 2010–12). Bank of Canada.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*, 386–408.

Rosenblatt, F. (1963). *Principles of neurodynamics*. Spartan.

Rothman, D. (2018). *Artificial intelligence by example*. Packt Publishing.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation (pp. 318–362). MIT Press.

Salinas, D., Bohlke-Schneider, M., Callot, L., & Gasthaus, J. (2019). High-dimensional multivariate forecasting with low-rank Gaussian copula processes. In *Advances in Neural Information Processing Systems*.

Stratos, K. (2020). *Neural networks research*. Retrieved December 20, 2020, from http://www1.cs.columbia.edu/~stratos/research/neural.pdf

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting, 23*(6), 405–430.

Werbos, P. J. (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks, 1*.

World Meteorological Organization (WMO). (2017). *Guidelines for nowcasting techniques* (WMO No. 1198).