

SATIŞ TAHMİNİ İÇİN DERİN ÖĞRENME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Begüm EROL *
Tülin İNKAYA **

Alınma: 15.11.2023; kabul: 14.07.2024

Öz: Dijital dönüşüm ile tedarik zinciri yönetiminde büyük veri analitiğinin önemi gün geçtikçe artmaktadır. Özellikle müşteri taleplerinin hızlı ve doğru tahmin edilmesinde büyük verinin kullanımı firmalara rekabet avantajı sağlamaktadır. Bu doğrultuda, yapay zekâ tekniklerinden biri olan derin öğrenme modelleri büyük verideki karmaşık örüntülerin keşfedilmesinde öne çıkmaktadır. Son yıllarda literatürde çok sayıda derin öğrenme yöntemi önerilmiştir. Bu çalışmada, satış tahmini problemi için derin öğrenme yöntemlerinin performansları karşılaştırılmıştır. Bu kapsamda derin sinir ağı (DNN), derin otokodlayıcı (Deep AE), evrişimli sinir ağı (CNN), tekrarlayan sinir ağı (RNN), uzun kısa-süreli bellek (LSTM) ağı, çift yönlü LSTM (Bi-LSTM) ağı, kapalı tekrarlayan birim (GRU), CNN-LSTM ve evrişimli LSTM (ConvLSTM) yöntemleri uygulanmıştır. Çeşitli sektörlerde ait satış verileri kullanılarak deneysel çalışmalar gerçekleştirilmiştir. Hiperparametre optimizasyonu ardından ele alınan yöntemlerin performansları tahmin doğruluğu ve eğitim süreleri açısından karşılaştırılarak sonuçların istatistiksel anlamlılığı değerlendirilmiştir. Sonuç olarak, LSTM ve GRU modellerinin tahmin doğruluğunda başarılı sonuçlar verdiği, CNN modelinin ise eğitim süresini kısalttığı görülmüştür.

Anahtar Kelimeler: Derin öğrenme, Satış tahmini, Zaman serisi, Evrişimli sinir ağı, Tekrarlayan sinir ağı, Uzun kısa-süreli bellek ağı

Comparison of Deep Learning Methods for Sales Forecasting

Abstract: With the digital transformation, the importance of big data analytics in the supply chain management has been increasing day by day. Especially, the use of big data in fast and accurate estimation of customer demand provides companies competitive advantage. In this direction deep learning models, which is one of the artificial intelligence techniques, stand out in the discovery of complex patterns in big data. In the recent years, several deep learning methods have been proposed in the literature. In this study, the performances of deep learning methods for the sales forecasting problem are compared. In this context, deep neural network (DNN), deep autoencoder (DAE), convolutional neural network (CNN), recurrent neural network (RNN), long short-term memory (LSTM) network, bi-directional LSTM (Bi-LSTM) network, gated recurrent unit (GRU), CNN-LSTM and convolutional LSTM (ConvLSTM) methods have been applied. Experimental studies were carried out using sales data from various sectors. After hyperparameter optimization, the performances of the methods discussed were compared in terms of forecasting accuracy and training time, and the statistical significance of the results was evaluated. As a result, it has been seen that LSTM and GRU models gave successful results in the prediction accuracy, and CNN model shortens the training time.

Keywords: Deep learning, Sales forecasting, Time series, Convolutional neural network, Recurrent neural network, Long short-term memory network

* İstanbul Beykent Üniversitesi, Mühendislik-Mimarlık Fakültesi, Endüstri Mühendisliği Bölümü, Ayazağa Kampüsü, 34398, Sarıyer/İstanbul

** Bursa Uludağ Üniversitesi, Mühendislik Fakültesi, Endüstri Mühendisliği Bölümü, 16059, Nilüfer/Bursa

İletişim Yazarı: Begüm EROL (begumerol@beykent.edu.tr)

1. GİRİŞ

Satış tahmini, tedarik zinciri yönetiminin temel görevlerinden biridir ve tedarik zincirindeki planlama ile karar verme süreçlerinde kritik bir rol oynamaktadır (Dharshini ve Vijila, 2021). Örneğin; seri üretim yapan bir firma üretim miktarlarını ve stok seviyelerini belirlemek için satış tahminlerine ihtiyaç duymaktadır. Siparişe göre üretim yapan bir firma ise deposunda bulundurması gereken malzeme miktarını belirlemek için talebi öngörmek istemektedir.

Doğru satış tahmini firmanın kârlılığını ve verimliliğini artırırken eş zamanlı olarak stok ve işletme maliyetlerini azaltmaktadır (Acı ve Doğansoy, 2022). Öte yandan, satış tahminindeki hatalar firmalarda yok satmaya bağlı gelir kaybına ya da fazla üretim sonucu artan stok maliyetlerine neden olmaktadır. Ayrıca, tahminin hatalı yapıldığı durumlarda müşteri kaybı ve müşteri memnuniyetsizliği de yaşanabilmektedir.

Satış tahmini, tüketici davranışlarının belirsizliği ve özelleştirilmiş ürünlere artan ilgi nedeniyle çözümü zor bir problemdir. Etkin tahmin modellerinin geliştirilmesi amacıyla istatistiksel yöntemler ve yapay zekâ teknikleri kullanılmaktadır (Dharshini ve Vijila, 2021; Acı ve Doğansoy, 2022). Özellikle son yıllarda yapay zekâ tekniklerinden derin öğrenme yöntemleri öne çıkmaktadır (Dharshini ve Vijila, 2021; Acı ve Doğansoy, 2022). Derin öğrenme yöntemleri, büyük hacimli verilerde karmaşık ve doğrusal olmayan ilişkileri keşfetme özellikleri nedeniyle tercih edilmektedir.

Donanımsal gelişmeler ve farklı alanlara uygulama esnekliği nedeniyle literatürde çok sayıda derin öğrenme yöntemi önerilmiştir (Sarker ve diğ., 2021). Özellikle çeşitli sektörlerde ait satış tahmini problemlerinde başarılı sonuçlar veren derin öğrenme yöntemleri arasında çok katmanlı derin sinir ağı, zamansal bağımlılıkları öğrenebilen tekrarlayan sinir ağları ve türevleri (tekrarlayan sinir ağı, uzun kısa-sürelilik bellek ağı, kapılı tekrarlayan birim, çift yönlü uzun kısa-sürelilik bellek ağı), öznelik çıkarma özelliğine sahip yöntemler (derin oto kodlayıcılar, evrişimli sinir ağı) ve hibrit yaklaşımlar bulunmaktadır. Derin öğrenme yöntemlerinin her birinin kendilerine has güçlü ve zayıf yönleri bulunmaktadır. Ancak bu yöntemlerin performansını kapsamlı bir şekilde karşılaştıran bir çalışmaya literatürde rastlanmamıştır.

Bu çalışmada, literatürdeki boşluğu doldurmak amacıyla sık kullanılan ve güncel derin öğrenme yöntemlerinin satış tahminindeki performanslarının karşılaştırılması ele alınmıştır. Bu kapsamda, çeşitli sektörlerde ait talep verileri kullanılmıştır. Performans karşılaştırılması tahmin doğruluğu ve eğitim süresine göre yapılmıştır. Performanslar arasındaki farkın istatistiksel olarak anlamlılığını incelemek için parametrik olmayan testler uygulanmıştır. Sonuç olarak, bu çalışma satış tahmini problemlerinde güncel derin öğrenme yöntemlerine ilişkin performansları kapsamlı bir şekilde karşılaştırması itibarıyla literatüre katkı sağlamaktadır.

Makale beş bölüm içermektedir. Giriş bölümü sonrasında ikinci bölümde literatür taraması sunulmaktadır. Üçüncü bölümde çalışmada kullanılan veri kümeleri ile derin öğrenme yöntemleri açıklanmaktadır. Dördüncü bölümde çalışmanın akış şeması, deneysel koşullar, karşılaştırmalı sonuçlar ve istatistiksel analizler detaylarıyla verilmektedir. Son olarak, beşinci bölümde sonuçlar ile gelecek çalışmalar belirtilmektedir.

2. LİTERATÜR TARAMASI

Satış tahmini, tedarik zincirinde temel bir araştırma konusudur (Ingle ve diğ., 2021). Bu alanda kullanılan yöntemler istatistiksel yöntemler ve yapay zekâ yöntemleri olmak üzere iki başlıkta incelenebilir (Ingle ve diğ., 2021). İstatistiksel yöntemler arasında doğrusal regresyon, hareketli ortalama, ağırlıklı ortalama, üstel düzeltme, ARIMA gibi yöntemler yer almaktadır (Ingle ve diğ., 2021; Acı ve Doğansoy, 2022; Erol ve İnkaya, 2023). Yapay zekâ yöntemleri ise makine öğrenmesi ve derin öğrenme yöntemleri olmak üzere iki alt başlıkta ele alınabilir. Makine öğrenmesi yöntemleri arasında k-en yakın komşuluk, destek vektör makinesi, karar ağacı, yapay sinir ağı gibi yöntemler; derin öğrenme yöntemleri arasında ise çok katmanlı derin

sinir ağı (DNN), evrişimli sinir ağı (CNN), tekrarlayan sinir ağı (RNN), uzun kısa-süreli bellek (LSTM) ağı gibi yöntemler bulunmaktadır (Ingle ve diğ., 2021; Dharshini ve Vijila, 2021; Acı ve Doğansoy, 2022; Erol ve İnkaya, 2023).

Özellikle son senelerde satış tahmini alanında derin öğrenme yöntemlerini kullanan çalışmalarda artış olmuştur (Dharshini ve Vijila, 2021; Eglite ve Birzniece, 2022). Dharshini ve Vijila (2021) satış ve talep tahmininde kullanılan makine öğrenmesi ve derin öğrenme yaklaşımlarının artı ve eksi yönlerini içeren bir derleme çalışması sunmuştur ve derin öğrenme yaklaşımlarının tahmin doğruluğu açısından makine öğrenmesi yaklaşımlarına göre daha üstün performans gösterdiğini belirtmiştir. Eglite ve Birzniece (2022) derin öğrenme yöntemleri ile perakende satışlarının tahmin edilmesi konusunda sistematik bir literatür tarama çalışması gerçekleştirmiştir. Ingle ve diğ. (2021) ait bir derleme çalışmasında ise istatistiksel, makine öğrenmesi, derin öğrenme ve hibrit modeller incelenmiştir. Karmaşık veri kümelerinde makine öğrenmesi ile derin öğrenme modellerinin kullanılmasının uygun olduğu belirtilmiştir.

Derin öğrenme yöntemlerini uygulayan çalışmalarda gıda ile temel ihtiyaçlar (Qi ve diğ., 2019, Jiang ve diğ., 2020); perakende (Bousqaoui ve diğ., 2021; Pacella ve Papadia, 2021); moda ile tekstil (Wang ve diğ., 2020); otomotiv (Chandriah ve diğ., 2021); sağlık (Buyar ve diğ., 2019, Qi ve diğ., 2019); e-ticaret (Qi ve diğ., 2019; Wang ve diğ., 2020; Acı ve Doğansoy, 2022); lojistik (Jiang ve diğ., 2020; Ren, 2020; Peköz ve İnkaya, 2020) gibi çeşitli sektörlerdeki talebin tahmin edilmesi ele alınmıştır. Bu çalışmalarda kullanılan derin öğrenme yöntemleri ve veriler Tablo 1’de verilmiştir. Buna göre, satış tahmininde CNN, RNN ve LSTM yöntemlerinin daha sık kullanıldığı görülmektedir.

Satış verisindeki zamansal bağımlılıkların keşfedilmesinde sıralı verileri işleyen RNN, LSTM, çift yönlü LSTM (Bi-LSTM) ve kapılı tekrarlayan birim (GRU) gibi tekrarlayan sinir ağlarını kullanan çok sayıda çalışma bulunmaktadır. Örnek olarak Muhaimin ve diğ. (2021) aralıklı talep verilerinin tahmininde RNN yönteminin performansını çok katmanlı derin sinir ağı ve diğer geleneksel tahmin yöntemleri ile kıyaslamıştır. Wang ve diğ. (2019) bir perakende zincirinde bozulabilir ve dayanıklı ürünlere olan talebi RNN ve LSTM yöntemleri ile tahmin etmiştir. Chandriah ve diğ. (2021) yedek parça talebinin tahmini için LSTM modeli geliştirmiştir. Pacella ve Papadia (2021) tedarik zincirinde son müşterinin talep tahmini için LSTM ve Bi-LSTM modellerini uygulamıştır. Qi ve diğ. (2019) ise promosyon kampanyalarında ikame edilebilir ürünler arasındaki rekabetçi ilişkilerin etkisini incelemiştir ve e-ticaret satışlarının tahmini için GRU tabanlı derin sinir ağı çerçevesi sunmuştur.

Çalışmaların bir diğer bölümünde özneliklerin çıkarılması amacıyla CNN uygulanmıştır. Buyar ve diğ. (2019) bir ilaç firmasına ait büyük veriyi kullanarak CNN ile gelecek satışları tahmin etmiştir. Belas ve Bidyuk (2021) perakendecilik sektöründeki satış verisini zaman serisi olarak ele almıştır ve önerdikleri CNN modelinin performansını LSTM modelinin performansı ile karşılaştırmıştır. Rizvi ve diğ. (2021) ise IoT cihazı kullanan bir akaryakıt istasyonu için genişletilmiş tek boyutlu CNN (1D-CNN) modeli önermiştir ve önerdikleri modelin diğer tahmin yöntemlerine göre daha etkili sonuç verdiğini göstermiştir.

Bir diğer grup çalışmada ise derin öğrenme yöntemlerinin hibrit versiyonları geliştirilerek tahmin performansının iyileştirilmesi hedeflenmiştir. Ren ve diğ. (2020) envanter optimizasyonu ile talep tahmini sürecini entegre ederek CNN-LSTM hibrit modelini önermiştir. Peköz ve İnkaya (2023) COVID-19 sürecinde lojistik sektöründeki talebi tahmin etmek için CNN-LSTM hibrit modeli geliştirmiştir. Bir diğer hibrit yöntem çalışmasında, Liu ve diğ. (2021) LSTM’ye evrişim işlemini entegre ederek evrişimli uzun kısa-süreli bellek (ConvLSTM) modelini önermiştir ve perakendecilik sektöründeki toplam satışları tahmin etmiştir. Önerilen modelin performansı geleneksel RNN ile karşılaştırılmıştır. Wang ve diğ. (2020) ise ürünlerin geçmiş satış bilgileri, görsel ve temel özelliklerini çok modlu ele almıştır ve DNN, CNN ile LSTM yöntemlerini bütünlük kullanarak moda ürünleri için satış tahmin modeli önermiştir.

Literatürde farklı derin öğrenme yöntemlerinin performanslarını karşılaştıran çalışmalar da bulunmaktadır. Örneğin; Bousqaoui ve diğ. (2021) süpermarket satış tahmini için CNN, LSTM,

yapay sinir ağı ve ARIMA yöntemlerinin performanslarını kıyaslamalı olarak değerlendirmiştir. Acı ve Doğanşoy (2022) e-perakende satış tahmininde makine öğrenmesi ve derin öğrenme yöntemlerini karşılaştırmıştır. Jiang ve diğ. (2020) ise bir alkollü içecek dağıtım şirketinin satış tahmini problemini ele almıştır ve DNN, LSTM, Bi-LSTM, CNN-LSTM, derin otokodlayıcılar (Deep AE) gibi derin öğrenme yaklaşımlarının performansını naif yöntemler ve ARIMA modelleri ile kıyaslamıştır.

Bu çalışma literatürde yer alan çalışmalardan birkaç yönden farklılaşmaktadır. İlk olarak, literatürde sıklıkla kullanılan ve güncel derin öğrenme yöntemlerine odaklanılmıştır ve bu yöntemlerin satış tahmini alanındaki performansı irdelenmiştir. İkinci olarak, performans karşılaştırmasında perakende, eczane, üretim, hizmet gibi çeşitli sektörlerle ait satış verileri kullanılmıştır. Böylece derin öğrenme modellerinin farklı özelliklere sahip veri kümelerindeki performansı ele alınmıştır. Çeşitli sektörlerle ait satış verilerinde güncel derin öğrenme yöntemlerinin kullanılması ile kapsamlı bir karşılaştırma gerçekleştirilmiştir. Bunlara ek olarak çalışmanın literatüre bir diğer katkısı da derin öğrenme yöntemlerinin performansının istatistiksel testler yardımıyla değerlendirilmesidir.

Tablo 1. Satış ve talep tahmininde derin öğrenme yöntemlerini kullanan çalışmalar

| Çalışma | Yöntem | Veri |
|---------------------------|--------------------------------|--|
| Muhaimin ve diğ. (2021) | DNN, RNN | M5 yarışma satış verileri |
| Buyar ve diğ. (2019) | CNN | Eczane satış verileri |
| Belas ve Bidyuk (2021) | CNN, RNN | Farklı bölgelerde bulunan 45 perakende mağazanın geçmiş satış verileri |
| Wang ve diğ. [14] | RNN, LSTM | Süpermarketlere ait satış verileri |
| Chandriah ve diğ. (2021) | LSTM | Yaklaşık 11 yıla ait Norveç'teki aylık yeni binek otomobil satış verileri |
| Bousqaoui ve diğ. (2021) | CNN, LSTM | Fas'taki bir süpermarketin satış verileri |
| Wang ve diğ. (2020) | DNN, CNN, LSTM | Bir giyim e-ticaret şirketindeki 4.354 ürüne ait satış kayıt verileri |
| Qi ve diğ. (2019) | GRU | Atıştırmalık ve kişisel bakım ürünleri satan iki farklı e-ticaret sitesine ait geçmiş satış verileri |
| Rizvi ve diğ. (2021) | CNN, LSTM, GRU | Akaryakıt dağıtım ünitelerinden toplanan 15 günlük akaryakıt satış verileri |
| Pacella ve Papadia (2021) | LSTM, BiLSTM | Bir pazardaki 10 spesifik ürüne ait talep verileri |
| Acı ve Doğanşoy (2022) | DeepAE, CNN, RNN, LSTM, BiLSTM | Türkiye'de yer alan yerel bir süpermarketin mobil ve internet satış verileri |
| Jiang ve diğ. (2020) | LSTM, BiLSTM, CNN-LSTM | Alkollü içecek dağıtım şirketinin yedi yıllık satış verileri |
| Ren ve diğ. (2020) | CNN-LSTM | Üçüncü parti lojistik firmasına ait lojistik hizmet talebi verileri |
| Peköz ve İnkaya (2023) | DNN, LSTM, CNN-LSTM | Bir lojistik firmasına ait sevkiyat verileri |
| Liu ve diğ. (2021) | RNN, ConvLSTM | Bir bölgedeki perakende mağazaların 26 haftalık satış verileri |

3. MATERYAL VE METOT

Bu bölümde, çalışma kapsamında kullanılan veri kümeleri ve derin öğrenme yöntemleri sunulmuştur.

3.1. Materyal

Bu çalışmada satış tahmini problemi tek değişkenli zaman serisi olarak modellenmiştir. Zaman serisi $Y = \{y_t: t \in T\}$ olarak tanımlanmış olup y_t t zaman adımıdaki gözlem değerini (satış miktarını), T ise zaman serisinin uzunluğunu (dönem sayısını) ifade etmektedir.

Bu çalışma kapsamında kullanılan ve çeşitli sektörlerle ait satış verileri Tablo 2'de sunulmuştur. Bu tabloda verilen Rossmann, Tv_Sales ve Favorita veri kümelerinde tek bir zaman serisi bulunmaktadır. Diğer veri kümelerinde ise birden çok bölgeye, mağazaya veya ürüne ait satış verileri bulunmaktadır. Bu veri kümelerinde bir zaman serisi seçilerek çalışma kapsamına alınmıştır. Her satış verisinin özelliklerinin belirlenmesi için otokorelasyon ve kısmi otokorelasyon fonksiyonları kullanılmıştır. Örneğin, Walmart veri kümesinde 60 dönemlik mevsimsellik gözlemlenmiştir.

Tablo 2. Veri kümeleri

| Veri Kümesi | Zaman Serisinin Uzunluğu (T) | Verinin Özellikleri | Açıklama |
|-------------|------------------------------|--|--|
| Walmart | 143 | Mevsimsellik vardır. | Walmart mağazasına ait 2010-2012 yılları arasındaki haftalık satış verileri (Kaggle, 2020) |
| Forklift | 276 | Çok sayıda sıfır içermektedir ve azalma eğilimine sahiptir. | Bir forklift üreticisine ait dört yıllık günlük satış verileri (Puspita ve diğ., 2019) |
| LearnX | 882 | Mevsimsellik vardır. | Çevrimiçi bir eğitim platformuna ait 2014-2016 yılları arasındaki yıllık satış verileri (Kaggle, 2020) |
| Rossmann | 942 | Mevsimsellik vardır. | Rossmann mağazasına ait 2010-2012 yılları arasındaki günlük satış verileri (Kaggle, 2020) |
| Tv Sales | 974 | Artma eğilimi vardır. | Bangladeş'teki iki yıllık televizyon satış verileri (Kaggle, 2020) |
| Favorita | 1688 | Mevsimsellik vardır. | Favorita marketlerine ait 2013-2017 yılları arasındaki haftalık satış verileri (Kaggle, 2020) |
| Pharma | 2106 | Mevsimsellik vardır ve sırasıyla azalma ve artma eğilimine sahiptir. | 2014-2019 yılları arasındaki eczane günlük ilaç satış verileri (Kaggle, 2020) |

3.2. Derin Öğrenme Yöntemleri

Bu bölümde, çalışma kapsamında ele alınan dokuz adet derin öğrenme yöntemi açıklanmıştır.

3.2.1. Derin Sinir Ağı

DNN, girdi ve çıktı katmanları arasında çok sayıda katman bulunan yapay bir sinir ağı yapısıdır. Ağın derinleşmesi ile genelleme yeteneğinin sığ ağlara göre daha iyi olduğu öne sürülmektedir (Goodfellow ve diğ., 2016). DNN, öznelik mühendisliği gereksinimini azaltabilme ve karmaşık problemleri birden çok katmanla çözebilme özelliklerine sahiptir. Ayrıca DNN mimarisi sayesinde geçmiş ve mevcut verileri kullanarak geleceği tahmin etme konusunda iyi performans göstermektedir (Gashler ve diğ., 2016).

3.2.2. Derin Otokodlayıcı

Otokodlayıcı (Autoencoders, AE), karmaşık veri temsillerini öğrenmek için eğitilmiş denetimsiz bir sinir ağıdır. Otokodlayıcı, girdi veri kümesinin boyutunu azaltarak en az kayıp ile girdiye en yakın çıktıyı üreten, öznelikleri ve yapısal bilgileri öğrenebilen bir yapay sinir ağı mimarisidir (Goodfellow ve diğ., 2016). Otokodlayıcı genellikle veri sıkıştırma ve gürültü azaltma amaçlarıyla kullanılmaktadır. Bununla birlikte son yıllarda derin üretken modellerin yapı taşları olarak da sıklıkla kullanılmaya başlanmıştır (Goodfellow ve diğ., 2016).

Her bir katmandaki çıktıların ardışık katmanın girdilerine bağlandığı ve birden fazla otokodlayıcı katmanından oluşan sinir ağı, derin otokodlayıcı olarak adlandırılmaktadır (Goodfellow ve diğ., 2016). Altı tam bağlantılı katmandan oluşan standart bir derin otokodlayıcı ağı için ilk üç katman kodlayıcıdır ve girdi uzayını bir öznelik uzayına eşlemekle sorumludur. Son üç katman ise kod çözücüdür ve ters eşlemeye karşılık gelir. Merkezi katmana ise kod denir. Eğitim süreci, girdi ile çıktı arasındaki farkı ölçen bir kayıp fonksiyonunun minimize edilmesini içerir. Bu fark değeri ile otokodlayıcının başarısı değerlendirilir.

Eğitim sonrasında girdinin (x) en önemli özellikleri kod (z) ile temsil edilebilir. Kod çözücü doğrusal olduğunda ve kayıp fonksiyonu ortalama karesel hata olduğunda, otomatik kodlayıcı temel bileşen analizi (PCA) ile aynı alt uzayı kapsayacak niteliktedir. Bu nedenle, otokodlayıcılar PCA'nın doğrusal olmayan genellemeleri olarak yorumlanabilir (Goodfellow ve diğ., 2016).

3.2.3. Evrişimli Sinir Ağı

CNN görüntü sınıflandırma ve nesne tanıma gibi genellikle görsel görevlerde kullanılmak üzere geliştirilmiş bir sinir ağı türüdür (LeCun ve diğ., 2015). Özellikle girdi verisindeki özneliklerin çıkarılmasında ve karmaşık temsillerin öğrenilmesinde etkilidir. Bir CNN mimarisi girdi katmanı, evrişim katmanı, ortaklama katmanı, tam bağlantılı katman ve çıktı katmanı içermektedir.

İşlenen girdi verisinin boyutuna göre CNN üç gruba ayrılır: Tek boyutlu, iki boyutlu ve üç boyutlu CNN. Tek boyutlu CNN (1D-CNN) zaman serisi ya da ses sinyali gibi tek boyutlu ilişkilerin olduğu veride kullanılır. İki boyutlu CNN görüntülerin, üç boyutlu CNN ise video tanıma ya da üç boyutlu medikal görüntülerin işlenmesinde kullanılır (LeCun ve diğ., 2015).

Bu çalışmada tek boyutlu zaman serileri olarak ele alındığı için 1D-CNN kullanılmıştır. 1D-CNN girdi dizisindeki zamansal ilişkilerin öğrenilmesini sağlar. 1D-CNN'nin ağ yapısında girdi katmanından sonra ardışık gelen evrişim ve ortaklama (havuzlama) katmanları girdi dizisinden öznelik çıkarma için kullanılır. Bu katmanlar sonucunda oluşan çıktı dizisi tam bağlantılı katmana gönderilerek nihai çıktı elde edilir (Kiranyaz ve diğ., 2021).

1D-CNN'de evrişim işlemi çekirdeğin girdi verisi üzerinde kaydırılmasını ve girdinin her bir alt dizisi ile çekirdeğin iç çarpımının hesaplanmasını içerir. Elde edilen çıktı dizisi, ileri seviyeli öznelikleri çıkarmak ve modelin nihai çıktısını oluşturmak için ek katmanlardan geçirilen bir öznelik haritasıdır. Evrişim katmanındaki ağırlık paylaşımı eğitimdeki parametre sayısını azaltarak 1D-CNN'nin daha erken ve daha hızlı yakınsamasını sağlar. 1D-CNN'de evrişim işlemi Denklem (1) ile ifade edilir (Kiranyaz ve diğ., 2021):

$$x_k^l = \sum_{i=1}^{N_{l-1}} conv1D(w_{ik}^{l-1}, s_i^{l-1}) + b_k^l \quad (1)$$

Burada sırasıyla x_k^l ve b_k^l l . evrişim katmanında k . nörondaki girdi ve sapma, w_{ik}^{l-1} ($l-1$). evrişim katmanındaki i . nöronun l . evrişim katmanındaki k . nörona kadar evrişim çekirdeği, s_i^{l-1} ($l-1$). evrişim katmanında i . nörondaki çıktı, N_{l-1} ($l-1$). evrişim katmanındaki nöron sayısı ve conv1D ise evrişim operatörüdür. x_k^l girdisine aktivasyon fonksiyonu uygulandığında ara çıktı elde edilir (Kiranyaz ve diğ., 2021).

Evrişim katmanından sonra gelen ortaklama katmanı, ana öznelikleri korurken boyut azaltma amacıyla öznelik haritasından alt örneklem oluşturur. Bu amaçla, ortaklamada ortalama ya da maksimum yaklaşımları kullanılabilir (Wang ve diğ., 2017). Ortalama ortaklama yaklaşımında belirli bir komşulukta ortalama değer alınır. Maksimum ortaklama yaklaşımında ise komşulukta en büyük değer alınır. Ortaklama yaklaşımları aynı zamanda aşırı uyum sorununu çözmeyi amaçlamaktadır (Wang ve diğ., 2017). Sonrasında elde edilen öznelikler tam bağlantılı katmana girdi olarak verilir. Tam bağlantılı katman düzleştirilmiş girdi ile çalışır ve girdiler bu katmandaki tüm nöronlar ile bağlıdır (Wang ve diğ., 2017). Nihai çıktının elde edilmesi için uygun bir aktivasyon fonksiyonu seçilir. Buna göre sınıflandırma problemlerinde sınıf etiketlerinin olasılık dağılımı, tahmin problemlerinde ise sürekli bir çıktı değeri oluşturulur.

3.2.4. Tekrarlayan Sinir Ağı

RNN sıralı verinin işlenmesinde kullanılır. Geçmiş girdi verisini hafızada tutarak mevcut verinin işlenmesini sağlayan bir yapay sinir ağıdır (Acı ve Doğansoy, 2022). Bu amaçla, tekrarlayan bağlantı adı verilen bir yapı kullanılır ve bir zaman adımındaki çıktının bir sonraki zaman adımına girdi olarak kullanılmasına izin verir.

Klasik sinir ağı mimarisinde girdi-çıkıtı bağımsızlığı olduğu varsayılır. Fakat RNN sıralı şekilde gelen verileri kullanır ve dizideki her verinin çıktısı önceki hesaplamalara bağlıdır (Acı

ve Doğanşoy, 2022). Böylece ileri beslemeli sinir ağından farklı olarak RNN hafızaya sahiptir. Bir dizinin her elamanı için aynı görevi yerine getirmeleri ve çıktının önceki hesaplamalara bağlı olması nedeniyle tekrarlayan olarak adlandırılır.

RNN'nin tekrarlayan birimlerine ait matematiksel işlemler Denklem (2)-(3)'te verilmiştir. h_t , t zaman adımındaki gizli durumu gösterir ve bu yapı RNN hücresinin hafızasıdır. RNN hücresinin t zaman adımındaki girdi ve çıktısını sırasıyla x_t ve y_t göstermektedir. W_{xi} , W_{hi} ve W_o sırasıyla girdi, gizli durum ve çıktı için ağırlık matrislerini; b_i ve b_o ise sapma vektörlerini ifade etmektedir. Mevcut gizli durum, önceki zaman adımının gizli durumuna ve mevcut girdiye bağlıdır. Ayrıca, mevcut durumu bir sonraki duruma bağlayan geri bildirim döngüleri ile desteklenir.

Uygulama alanlarına ve kullanılan veri türlerine göre aktivasyon fonksiyonları değişebilmektedir. Genellikle aktivasyon fonksiyonu olarak gizli katmanda sigmoid fonksiyonu (σ), çıktı katmanında ise hiperbolik tanjant fonksiyonu (\tanh) kullanılmaktadır.

$$h_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$y_t = \tanh(W_o h_t + b_o) \quad (3)$$

RNN'nin eğitimi için geriye yayılım algoritması kullanılır. Ancak uzun dizilerde geriye yayılımdaki türev işlemleri esnasında türevlerin yok olması ya da aşırı büyümesi ile karşılaşılabilir. Bu problemlere kaybolan gradyan (vanishing gradient) ve patlayan gradyan (exploding gradient) adı verilir. Bu durumda ağı doğru bir şekilde eğitimi zorlaşmaktadır. RNN'de yaşanan bu problemlerin üstesinden gelmek için LSTM ve GRU gibi derin öğrenme yöntemleri önerilmiştir (Hochreiter ve Schmidhuber, 1997).

3.2.5. Uzun Kısa-Sürelili Bellek Ağı

RNN'de uzun zaman dizilerinin eğitiminde karşılaşılabilen kaybolan/patlayan gradyan problemlerine çözüm olarak LSTM ağı önerilmiştir (Hochreiter ve Schmidhuber, 1997). LSTM uzun dönemli zaman bağımlılığı bilgilerini depolayabilen ve girdi ile çıktı dizilerini uygun şekilde eşleştirebilen tekrarlayan sinir ağı mimarisidir (Hochreiter ve Schmidhuber, 1997).

LSTM birimleri bir blok içinde yer alır ve bir veya birden fazla blok bir araya gelerek LSTM mimarisini oluşturur. LSTM bloğu veri girişi, çıkışı ile akışını kontrol eden üç kapı ve bilgiyi tutan bir hafıza hücresi içermektedir (Greff, 2016).

LSTM ağındaki işlemler Denklem (4)-(10) ile ifade edilebilir. Burada zaman adımı t , girdi kapısı i , unutma kapısı f , hafıza hücresi c , çıktı kapısı o , aktivasyon fonksiyonları \tanh ve σ , noktasal çarpımı \times , eleman bazında çarpımı ise \otimes ile gösterilmektedir. Ayrıca, t zaman adımında girdi vektörü x_t , çıktı vektörü y_t , hafıza hücresi durumu c_t , girdi kapısının aday değeri \tilde{c}_t , unutma kapısının aday değeri f_t , çıktı kapısının aday değeri o_t ve gizli katmanın aday değeri h_t ile gösterilmektedir. Sırasıyla W_f, W_i, W_c, W_o ağırlık matrislerini; b_f, b_i, b_c, b_o ise sapma vektörlerini temsil etmektedir (Greff, 2016; Erol ve İnkaya, 2024).

$$f_t = \sigma(W_f \otimes [h_{t-1}, x_t] + b_f) \quad (4)$$

$$i_t = \sigma(W_i \otimes [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{c}_t = \tanh(W_c \otimes [h_{t-1}, x_t] + b_c) \quad (6)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \quad (7)$$

$$o_t = \sigma(W_o \otimes [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t \times \tanh(c_t) \quad (9)$$

$$y_t = \sigma(h_t) \quad (10)$$

LSTM ağında Denklem (4)'te yer alan unutmama kapısı ile veri silinmesi, Denklem (5)'te yer alan girdi kapısı ile veri güncellemesi, Denklem (6) ile ise hücre durumuna eklenecek yeni aday değerlerin oluşturulması sağlanır. Denklem (7)'de unutmama kapısı ile mevcut durum, girdi kapısı ile aday girdiler güncellenerek mevcut hafıza hücresinin durumu oluşturulur. Denklem (8)-(10) ile ağı'nın nihai çıktısı; çıktı kapısı, önceki ve mevcut hafıza hücresi durumlarına bağlı olarak hesaplanır (Greff, 2016). Özetle, LSTM ağının ileri yayılım işlemleri Denklem (4)-(10)'da verilmiştir.

LSTM ağının eğitimi için geriye yayılım algoritması kullanılır. Çıktı katmanından girdi katmanına doğru geriye yayılırken ağıdaki hatayı en küçükleyecek şekilde ağırlıklar üzerinde güncellemeler yapılır. Tahmin problemleri için kayıp fonksiyonu olarak genellikle ortalama karesel hata (Mean Squared Error, MSE) kullanılmaktadır. Her iterasyonda kayıp fonksiyona ait gradyanlar (kısmi türevler) hesaplanır ve ardından her bir katmanın ağırlıkları ile sapma değerleri güncellenir. İleri yöndeki ve geri yöndeki bu döngü belirlenen bir durdurma kriteri sağlanana kadar tekrarlanır (Greff, 2016).

LSTM'de RNN'nin aksine sabit bir W matrisi yerine Denklem (7)'deki gibi her bir adımda değişen farklı bir unutmama kapısı değeri kullanılmaktadır. Böylece kaybolan ve patlayan gradyan problemlerinin önüne geçilebilmektedir. Hata terimine bağlı olarak her kapının ağırlığı bir optimizasyon algoritması kullanılarak güncellenir.

3.2.6. Kapılı Tekrarlayan Birim

RNN, kısa süreli hafızaya sahip olduğu için uzun dönemli bağımlılıkları taşıyamaz. Bu duruma çözüm olarak Cho ve diğ. (2014) tarafından GRU yöntemi önerilmiştir. Her tekrarlayan birimin farklı zaman adımlarındaki bağımlılıkları yakalaması amacıyla GRU birimindeki bilgi akışını ayarlayan, ancak ayrı bir hafıza hücresine sahip olmayan kapı adı verilen dahili mekanizmalara (LSTM birimine benzer) sahiptir (Chung ve diğ., 2014).

GRU bilgi aktarmak için gizli durumu kullanır ancak hücre durumu içermez. Bu nedenle GRU, LSTM'ye kıyasla daha az tensör işlemine sahiptir ve eğitim süresi daha hızlıdır. GRU yapısında bulunan kapılar, bir dizideki hangi verilerin saklanması veya unutulmasının önemli olduğunu öğrenir ve tahmin için yalnızca ilgili verileri tutmayı, ilgisiz verileri ise unutmayı sağlayabilir. Böylece uzun diziler boyunca ilgili bilgileri iletebilir.

GRU, güncelleme ve sıfırlama adında iki kapı içerir. Güncelleme kapısı, önceki gizli durumun ne kadarının saklanması gerektiğini ve onu güncellemek için yeni girdinin ne kadarının kullanılması gerektiğini belirler. Sıfırlama kapısı ise yeni gizli durum hesaplanırken önceki gizli durumun ne kadarının unutulması gerektiğini ve yeni girdinin ne kadarının dikkate alınması gerektiğini belirler (Chung ve diğ., 2014).

3.2.7. Çift Yönlü Uzun Kısa-Süreli Bellek Ağı

Çift yönlü LSTM (Bi-LSTM) iki ayrı LSTM modelinden oluşur. LSTM modellerinden biri girdiyi ileri yönde alır, diğeri ise geri yönde alır. Böylece, ağıdaki bilgi miktarını artırarak zaman serisindeki iki yönlü uzun vadeli bağımlılıkları öğrenebilir. Bi-LSTM sıralı verinin

işlenmesinde etkilidir ancak girdi dizisi ileri ve geri yönlü iki defa işlendiği için hesap karmaşıklığı yüksektir (Graves ve Schmidhuber, 2005).

3.2.8. CNN-LSTM Hibrit Modeli

CNN-LSTM sıralı girdi verisinin işlenmesi için CNN ve LSTM ağını birleştiren bir sinir ağı mimarisidir. CNN evrişim işlemleri ile sıralı verideki özniteliklerin çıkarılmasını sağlar. CNN ile elde edilen öznitelikler LSTM ağına girdi olarak verilir ve LSTM sıralı verideki uzun dönemli bağımlılıkların öğrenilmesini sağlar.

Hibrit modelde, her bir girdi dizisi CNN modeli tarafından işlenen çoklu alt dizilere bölünür. Örneğin, 36 zaman adımı içeren bir dizi 12 zaman adımı içeren üç alt diziyeye bölünerek aynı CNN katmanında işlenir. Bu amaçla, zamana göre yayılan (TimeDistributed) katman sarmalayıcı olarak kullanılabilir. CNN alt modelinin çıktısı LSTM katmanı tarafından işlenir. Bunu, tam bağlantılı bir katman ve tek adımlı tahminler yapmak için bir çıktı katmanı izleyebilir. CNN-LSTM modeli için alt dizi sayısı ve uzunluğu iki önemli parametredir (Wang ve diğ., 2016).

3.2.9. Evrişimli Uzun Kısa-Süreli Bellek Ağı

Evrişimli (convolutional) LSTM veya kısaca ConvLSTM, LSTM hücresi içinde evrişim işlemi içeren bir LSTM katmanı türüdür. ConvLSTM'de LSTM hücresinde yer alan her kapıdaki matris çarpımı evrişim işlemiyle değiştirilir. Böylece, çok boyutlu verilerde uzamsal özellikler yakalanır. ConvLSTM, girdiden duruma ve durumdan duruma geçişlerde evrişimli yapılar içeren, zaman-mekansal (spatiotemporal) tahmin için kullanılabilen tekrarlayan bir sinir ağı türüdür (Xingjian ve diğ., 2015).

ConvLSTM ve LSTM arasındaki temel fark, girdi boyutlarının sayısıdır. LSTM girdi verileri tek boyutlu olduğundan video, uydu, radar görüntü veri seti gibi mekansal dizi verileri için uygun değildir. ConvLSTM ise üç boyutlu veriler için tasarlanmıştır. Durumların girdilerle aynı sayıda satıra ve aynı sayıda sütuna sahip olmasını sağlamak için evrişim işlemi uygulamadan önce dolgu yapılması gerekebilir. CNN-LSTM'de olduğu gibi, girdi verileri her bir dizinin sabit sayıda zaman adımına sahip olduğu alt dizilere bölünür, ancak her bir alt dizideki satır sayısı da belirtilmelidir. CNN'dekine benzer şekilde ConvLSTM katmanı, girdi dizilerini okurken filtre sayısını ve kullanılan çekirdeğin boyutunu belirlemeye olanak verir. Bu katmanın çıktısı, bir öznitelik haritası olup önce düzleştirilir ve sonrasında bir çıktı katmanına girdi olarak verilir (Xingjian ve diğ., 2015).

4. DENEYSEL ÇALIŞMA

Bu bölüm altında deneysel çalışmalar ve sonuçları verilmiştir.

4.1. Deneysel Koşulları ve Parametre Optimizasyonu

Deneysel çalışmada izlenen akış şeması Şekil 1'de verilmiştir. Buna göre, öncelikle Tablo 2'de yer alan her satış verisine $[0, 1]$ normalizasyonu uygulanır. Her satış verisinin %64'ü eğitim, %16'sı doğrulama ve %20'si test verisi olarak ayrılmıştır. Sonraki aşamada, her satış verisine ait eğitim ve doğrulama verileri kullanılarak dokuz adet derin öğrenme modeli geliştirilmiştir. En iyi hiperparametre değerlerini belirlemek için modellerin ve veri kümelerinin özellikleri dikkate alınmıştır ve Tablo 3'te verilen arama uzayı kullanılmıştır. Izgara arama yöntemiyle gerçekleştirilen hiperparametre optimizasyonu sonucunda her satış verisi için en uygun hiperparametre değerleri belirlenmiştir. Sonrasında her satış verisi için geliştirilen derin öğrenme modelleri ile test kümesinde tahmin yapılmıştır. Son aşamada, derin öğrenme yöntemlerinin performansları arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını

tespit etmek için parametrik olmayan istatistiksel testler uygulanmıştır ve elde edilen sonuçlar yorumlanmıştır.

Derin öğrenme yöntemlerine ait performansların karşılaştırılmasında Denklem (11) ve (12)'de yer alan ortalama karesel hatanın karekökü (Root Mean Squared Error, RMSE) ile ortalama mutlak hata (Mean Absolute Error, MAE) ölçütleri seçilmiştir. Burada, \hat{y}_t tahmin edilen satış değerini, y_t gerçek satış değerini ve n dönem sayısını belirtmektedir. Bunlara ek olarak, derin öğrenme yöntemlerinin eğitim süreleri de karşılaştırma ölçütü olarak kullanılmıştır.

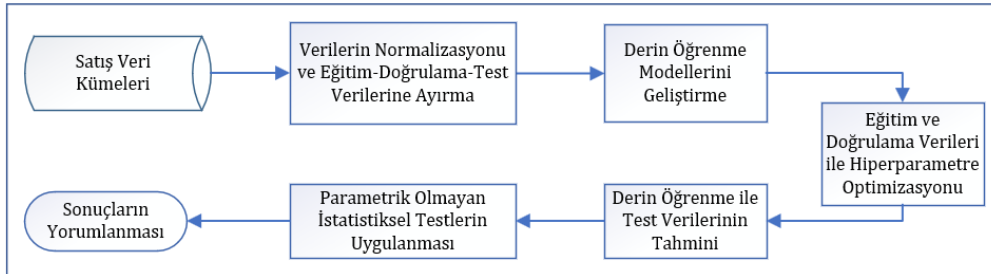
$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |\hat{y}_t - y_t| \quad (12)$$

Tüm deneylerde MSE kayıp fonksiyonu olarak ve Adam optimizasyon algoritması (Kingma ve Ba, 2021) olarak kullanılmıştır. Her veri kümesinde, derin öğrenme yöntemleri ile 30'ar tekrar yapılmıştır.

Çalışma kapsamında yapılan deneylerde Intel(R) Core(TM) i7-11800HQ CPU 4.60 GHz 64.0 GB RAM özellikli bilgisayar kullanılmıştır. Deneyler, Google Colab ve Anaconda Spyder ortamında Python kodlama dilinde Keras kütüphanesi kullanılarak gerçekleştirilmiştir.

Örnek olarak seçilen ve Şekil 2 (a)'da gösterilen Pharma veri kümesi için hiperparametre optimizasyonu sonucunda elde edilen LSTM ağı mimarisi ve kayıp fonksiyonunun dönem sayısına göre değişimi Şekil 2 (b) ve (c)'de sırasıyla verilmiştir. Buna göre, Şekil 2 (b)'de iki gizli ve bir çıktı katmanı görülen bir LSTM ağı oluşturulmuştur. Şekil 2 (c)'de ise eğitim ve doğrulama verileri için kayıp fonksiyonlarının sıfıra yakınsaması model eğitiminin başarılı bir şekilde gerçekleştiğini göstermektedir.



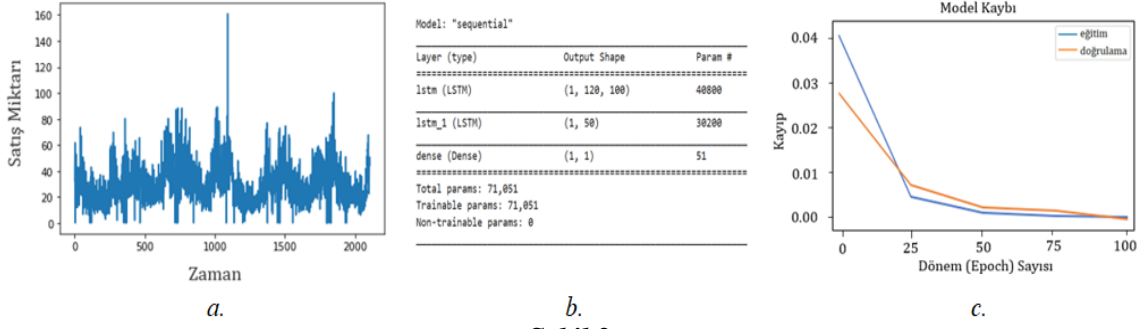
Şekil 1:

Derin öğrenme yöntemlerinin performanslarının karşılaştırılması için akış şeması

4.2. Deney sonuçları

Her satış verisinde derin öğrenme yöntemleri ile elde edilen test verisi performansları (ortalama RMSE ve MAE değerleri) ve eğitim süreleri Tablo 4'te verilmiştir. Tablo 4'te her veri kümesindeki en iyi performans ölçütleri koyu olarak işaretlenmiştir. Her bir yöntemde test süresi yaklaşık 0,4 saniyedir.

Tablo 4'teki sonuçlara göre incelenen yedi adet veri kümesinden dördünde (Walmart, Rossmann, Tv_Sales ve Pharma) hem RMSE hem MAE performans ölçütlerine göre en başarılı sonuçları LSTM yöntemi vermiştir. Ayrıca iki adet veri kümesinde (Forklift ve LearnX) en başarılı sonuçlar GRU, bir adet veri kümesinde (Favorita) ise ConvLSTM ile elde edilmiştir. Eğitim süresi açısından tüm veri kümelerinde en iyi sonuçlar sırasıyla CNN ve Deep AE yöntemleri ile elde edilmiştir. Bi-LSTM ve LSTM yöntemleri ise en uzun eğitim sürelerine sahip yöntemlerdir.



Şekil 2:

Pharma veri kümesi için **a.** veri setinin zamana bağlı grafiği, **b.** LSTM ağı mimarisi ve **c.** dönem sayısına göre kayıp fonksiyonundaki değişim.

Tablo 2’de yer alan veri özelliklerine göre derin öğrenme yöntemlerinin RMSE ve MAE açısından tahmin performansları değerlendirildiğinde şu gözlemler yapılmıştır:

- Mevsimsellik içeren veri kümelerinden Walmart veri kümesinde LSTM, DNN ve DeepAE; LearnX veri kümesinde GRU, Bi-LSTM ve Deep AE; Rossmann veri kümesinde LSTM, CNN-LSTM ve CNN; Favorita veri kümesinde ise ConvLSTM, CNN-LSTM ve CNN en yüksek performans sıralamasında ilk üçte yer alan yöntemlerdir.
- Sadece trend içeren Tv_Sales veri kümesinde LSTM, DNN ve GRU en iyi performansa sahiptir.
- Hem mevsimsellik hem de trend içeren Pharma veri kümesinde en iyi üç performans LSTM, CNN-LSTM ve GRU yöntemleri ile elde edilmiştir.
- Hem trend içeren hem de seyrek bir veri kümesi olan Forklift verisinde ise en iyi üç performansa sahip yöntem GRU, ConvLSTM ve RNN yöntemleridir.

Bu gözlemlere göre, mevsimselliğe sahip veri kümelerinde LSTM ve varyantları ile LSTM içeren hibrit yöntemlerin başarılı tahminler ürettiği görülmektedir. Bir başka deyişle verideki mevsimselliğin keşfedilebilmesi için uzun dönemli zamansal bağımlılıkları öğrenebilen LSTM ağı ön plana çıkmıştır. Benzer şekilde trend içeren veri kümelerinde tekrarlayan sinir ağı yapılarının, hem mevsimsellik hem trend içeren veri kümelerinde ise LSTM ve varyantlarının daha iyi sonuçlar verdiği görülmektedir. Rossmann, Favorita ve Pharma gibi mevsimsellik içeren büyük veri kümelerinde evrişim işlemleri ile LSTM ağının bütünleşik kullanıldığı hibrit yaklaşımlar tahmin doğruluğunu iyileştirmektedir.

Şekil 3’te örnek bir satış veri kümesi için test verisindeki tahmin performansı görselleştirilmiştir. Bu kapsamda, Pharma verisi seçilmiştir ve en iyi üç performansa sahip derin öğrenme yöntemlerinin (sırasıyla LSTM, CNN-LSTM ve GRU) tahmin değerleri ile gerçek satış değerleri Şekil 3’te verilmiştir. Buna göre, LSTM yöntemi ile gerçek satış miktarlarına daha yakın tahmin değerleri elde edilmiştir. Ayrıca LSTM, CNN-LSTM ve GRU’nun verideki mevsimselliği yakalamada başarılı olduğu gözlenmiştir.

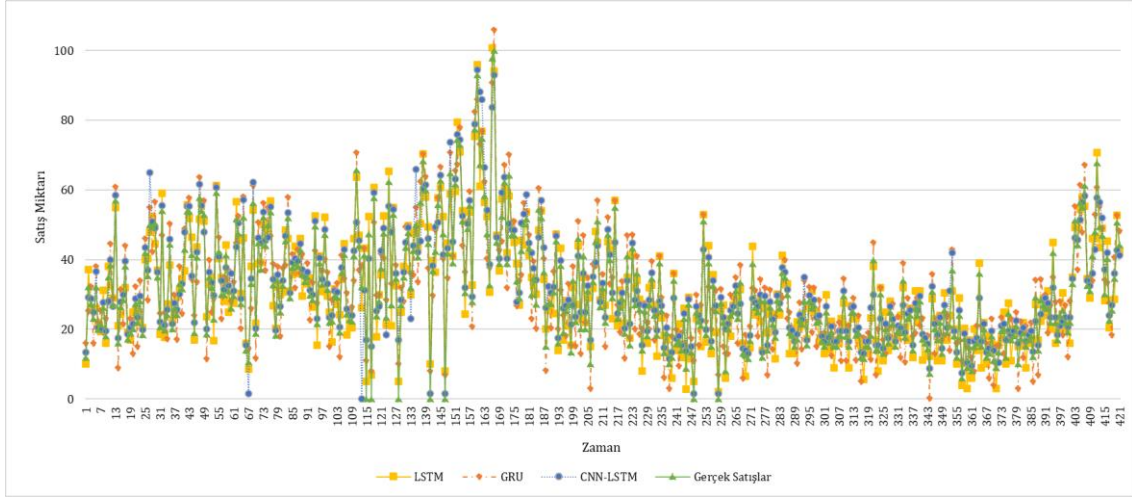
Tablo 3. Derin öğrenme modellerindeki hiperparametreler

| Yöntem(ler) | Hiperparametreler | Değer Aralığı |
|------------------------------------|--|---------------------------------------|
| DNN, RNN, LSTM, GRU, Bi-LSTM | Girdi boyutu | [1-370] |
| | Gizli katman sayısı | {2, 3, 4, 5} |
| | Gizli katman nöron sayısı | {4, 8, 16, 32, 64, 128, 256} |
| | Öğrenme oranı | {0,001; 0,01; 0,1} |
| | Parti (Batch) büyüklüğü | {1, 2, 4, 8, 16, 32, 64} |
| | Dönem (Epoch) sayısı | {25, 30, 40, 50, 75, 100, 150, 250} |
| Deep AE | Aktivasyon Fonksiyonu | {tanh, ReLu} |
| | Gizli (Kodlayıcı-Kod çözücü) katman sayıları | {1, 2, 3, 4, 5} |
| | Gizli katman nöron (kod) sayısı | {4, 8, 16, 32, 64, 128, 256} |
| | Öğrenme oranı | {0,001; 0,01; 0,1} |
| | Parti büyüklüğü | {1, 2, 4, 8, 16, 32, 64} |
| | Dönem sayısı | {25, 30, 40, 50, 75, 100, 150, 250} |
| CNN | Aktivasyon Fonksiyonu | {tanh, ReLu} |
| | Evrişim katmanı sayısı | {1, 2, 3} |
| | Kernel boyutu | {3, 5, 7} |
| | Filtre boyutu | {3, 4, 5, 7, 8, 16, 32, 64, 128, 256} |
| | Havuzlama | {Maksimum, Ortalama} |
| | Havuzlama boyutu | {2, 3, 4, 5, 6} |
| | Öğrenme oranı | {0,001; 0,01; 0,1} |
| | Parti büyüklüğü | {1, 2, 4, 8, 16, 32, 64} |
| | Dönem sayısı | {25, 30, 40, 50, 75, 100, 150, 250} |
| CNN-LSTM, ConvLSTM | Aktivasyon Fonksiyonu | {tanh, ReLu} |
| | Girdi boyutu | [1-370] |
| | Gizli katman sayısı | {2, 3, 4, 5} |
| | Gizli katman nöron sayısı | {4, 8, 16, 32, 64, 128, 256} |
| | Evrişim katmanı sayısı | {1, 2, 3} |
| | Kernel boyutu | {3, 5, 7} |
| | Filtre boyutu | {3, 4, 5, 7, 8, 16, 32, 64, 128, 256} |
| | Havuzlama | {Maksimum, Ortalama} |
| | Havuzlama boyutu | {2, 3, 4, 5, 6} |
| | Alt dizilerin sayısı | {2, 3, 4, 5, 6} |
| | Alt dizideki zaman adımı sayısı | {2, 3, 4, 5, 7, 12, 52} |
| | Öğrenme oranı | {0,001; 0,01; 0,1} |
| | Parti büyüklüğü | {1, 2, 4, 8, 16, 32, 64} |
| | Dönem sayısı | {25, 30, 40, 50, 75, 100, 150, 250} |
| Aktivasyon Fonksiyonu | {tanh, ReLu} | |

Tablo 4. Derin öğrenme yöntemlerinin test verilerindeki performansları

| Veri Kümesi | Ölçüt | Yöntemler | | | | | | | | |
|-------------|-------|--------------|----------------|----------------|-------------|---------------------|---------------|-------------|-------------|------------------|
| | | DNN | Deep AE | CNN | RNN | LSTM | GRU | BiLSTM | CNN-LSTM | ConvLSTM |
| Walmart | RMSE | 90.493,18 | 96.416,28 | 96.436,58 | 97.770,50 | 90.000,71* | 96.745,35 | 111.716,18 | 109.141,18 | 111.478,22 |
| | MAE | 74.722,67 | 76.422,03 | 76.644,88 | 77.141,59 | 71.829,10* | 76.689,38 | 91.796,85 | 90.151,42 | 90.988,57 |
| | Süre | 79,95 | 71,23 | 62,12* | 82,51 | 91,24 | 84,77 | 96,22 | 88,26 | 87,97 |
| Forklift | RMSE | 0,88 | 0,75 | 0,72 | 0,52 | 0,69 | 0,44* | 0,54 | 0,75 | 0,50 |
| | MAE | 0,82 | 0,73 | 0,71 | 0,50 | 0,65 | 0,42* | 0,53 | 0,72 | 0,48 |
| | Süre | 64,02 | 55,03 | 50,11* | 71,25 | 75,14 | 67,98 | 74,18 | 72,77 | 73,15 |
| LearnX | RMSE | 25,12 | 23,88 | 28,89 | 27,16 | 28,18 | 23,05* | 23,48 | 24,03 | 23,74 |
| | MAE | 20,44 | 19,15 | 23,61 | 22,49 | 23,07 | 18,88* | 18,95 | 19,02 | 18,95 |
| | Süre | 255,84 | 210,45 | 187,52* | 298,12 | 312,55 | 287,14 | 315,48 | 259,74 | 241,77 |
| Rossmann | RMSE | 1.758.434,92 | 1.124.172,4 | 1.706.163,1 | 1.721.218,1 | 1.478.745,3* | 1.871.191,3 | 2.110.472,4 | 1.646.852,7 | 1.742.350,9 |
| | MAE | 1.282.280,6 | 1.674.952,2 | 1.199.796,4 | 1.282.312,5 | 857.379,9* | 1.321.701,8 | 1.648.692,2 | 1.197.538,5 | 1.271.196,9 |
| | Süre | 579,10 | 476,23 | 399,28* | 522,04 | 715,23 | 602,48 | 752,89 | 571,77 | 501,29 |
| Tv_Sales | RMSE | 1.379,42 | 1.497,46 | 2.403,75 | 1.530,08 | 1.320,58* | 1.383,65 | 1.789,25 | 1.940,89 | 1.525,55 |
| | MAE | 715,58 | 801,77 | 1.978,47 | 1.101,21 | 711,96* | 777,58 | 1.245,41 | 1.402,23 | 1.097,52 |
| | Süre | 776,41 | 654,23* | 683,79 | 975,65 | 986,79 | 874,50 | 1.003,25 | 746,87 | 757,57 |
| Favorita | RMSE | 12.093,45 | 11.487,66 | 11.065,27 | 12.138,09 | 13.771,48 | 12.969,86 | 13.648,80 | 10.547,63 | 9.832,65* |
| | MAE | 9.012,31 | 7.896,36 | 7.785,25 | 9.102,45 | 9.124,39 | 9.114,23 | 9.174,13 | 6.578,93 | 5.140,25* |
| | Süre | 745,22 | 685,33 | 601,45* | 786,55 | 855,49 | 802,25 | 897,63 | 798,87 | 1.150,26 |
| Pharma | RMSE | 14,01 | 13,54 | 16,41 | 13,42 | 11,47* | 13,15 | 13,76 | 12,82 | 14,79 |
| | MAE | 12,44 | 11,25 | 13,87 | 11,21 | 9,61* | 11,20 | 11,28 | 10,11 | 12,74 |
| | Süre | 633,08 | 558,25 | 491,44* | 697,55 | 715,25 | 657,63 | 714,87 | 685,22 | 675,43 |

* Her ölçütte en iyi performansa sahip yöntem koyu olarak işaretlenmiştir.



Şekil 3:

Pharma test verisinde en yüksek performansa sahip ilk üç derin öğrenme yönteminin tahmin performanslarının karşılaştırılması

4.3. Parametrik olmayan istatistiksel testler

Bu çalışmada, derin öğrenme yöntemlerinin performansları arasındaki farkların istatistiksel olarak anlamlılığını değerlendirmek için parametrik olmayan istatistiksel testler kullanılmıştır. Parametrik olmayan istatistiksel testlerde varyans homojenliği ve normallik gibi varsayımlar bulunmamaktadır. Ayrıca yöntemlerin birden fazla veri kümesindeki performansları karşılaştırılabilmektedir (Demšar, 2006).

Bu çalışmada ele alınan derin öğrenme yöntemlerinin hem ikili hem de çoklu karşılaştırmaları gerçekleştirilmiştir (Demšar, 2006). Çoklu karşılaştırmalarda Friedman post-hoc testi ve Nemenyi post-hoc testi uygulanmıştır (Demšar, 2006). İkili karşılaştırmalarda Wilcoxon eşleştirilmiş işaretli sıra testi ile kazanç, kayıp ve eşitlik sayıları (counts of wins, losses and ties) işaret testi uygulanmıştır (Demšar, 2006).

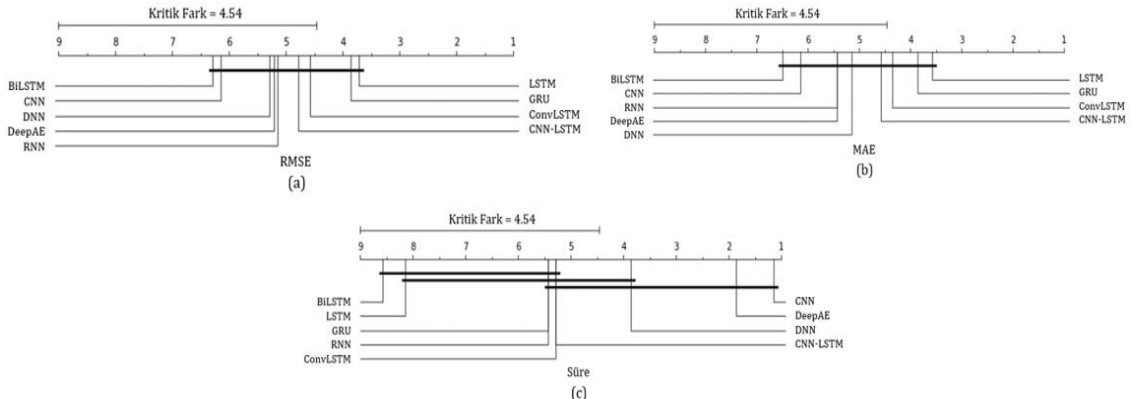
Friedman testi için kurulan hipotezler şu şekildedir: (i) H_0 : Derin öğrenme yöntemlerinin performansları arasında farklılık bulunmamaktadır. (ii) H_1 : En az bir derin öğrenme yönteminin performansında farklılık bulunmaktadır. Friedman testi için Demšar'ın (2006) çalışmasında önerilen Iman ve Davenport istatistiği (1980) uygulanmıştır. Bu doğrultuda, RMSE, MAE ve eğitim süresi ölçütlerine göre yapılan Friedman testleri için Ki-kare değerleri (χ^2_F) %5 anlamlılık düzeyinde ($\alpha = 0,05$) sırasıyla 54,0; 54,0 ve 52,76'dır. Tüm ölçütler için ki-kare tablo değeri serbestlik derecesi 6 ve $\alpha = 0,05$ anlamlılık düzeyinde $\chi^2 = 12,59$ 'dur ve bu değer hesaplanan ki-kare değerinden (χ^2_F) küçük olduğu için H_0 hipotezleri reddedilmektedir. Böylece hem RMSE ve MAE, hem de eğitim süresine göre derin öğrenme yöntemleri arasında istatistiksel olarak anlamlı bir farklılık olduğu görülmektedir.

Friedman testleri sonrasında istatistiksel olarak anlamlı farka sahip olan derin öğrenme yöntemlerinin tespit edilmesi amacıyla Nemenyi testi uygulanmıştır. Bu testte her yöntemin ortalama sıralama değeri (OSD) belirlenir. İki yöntemin karşılaştırılması için OSD'leri arasındaki fark hesaplanır ve bu fark kritik fark değeriyle kıyaslanır. Eğer iki yöntem arasındaki OSD farkı, kritik farktan büyük olursa bu yöntemler arasında istatistiksel olarak anlamlı bir fark bulunduğu sonucu çıkarılır.

Derin öğrenme yöntemlerinin RMSE, MAE ve eğitim süresine göre OSD'leri sırasıyla Şekil 4 (a), (b) ve (c)'de yer almaktadır. Kritik fark %5 anlamlılık düzeyinde 4,54 olarak hesaplanmıştır. Derin öğrenme yöntemleri Şekil 4'te aynı doğru parçası üzerinde OSD'ye göre sıralanarak gösterilmiştir. Buna göre, küçük OSD'ye sahip yöntemler diyagramın sağında, büyük OSD'ye sahip yöntemler ise diyagramın solunda yer almaktadır.

Şekil 4 (a) ve (b)'de incelenen derin öğrenme yöntemleri arasında RMSE ve MAE'ye göre anlamlı bir fark olmadığı görülmektedir. OSD büyüklüklerine göre değerlendirme yapıldığında ise Şekil 4 (a) ve (b)'de diyagramların sağında yer alan LSTM, GRU, ConvLSTM ve CNN-LSTM yöntemlerine ait OSD'ler, diyagramın solunda yer alan Bi-LSTM, CNN, DNN, DeepAE ve RNN yöntemlerine ait OSD'lerden daha küçüktür. Bu nedenle, diyagramın sağındaki yöntemler diyagramın solundaki yöntemlerden daha iyi performansa sahiptir. Örnek olarak, LSTM yöntemi RMSE ve MAE ölçütlerine göre en küçük OSD'lere (sırasıyla 3,72 ve 3,58) sahiptir, ikinci sırada ise GRU yöntemi (sırasıyla 3,86 ve 3,77) gelmektedir. Bu sıralama, LSTM ve GRU gibi uzun dönemli bağımlılıkları öğrenebilen ve tekrarlayan sinir ağlarının tahmin doğruluğunda öne çıktığını göstermektedir. RMSE ve MAE ölçütlerine göre en büyük OSD'lere sahip iki yöntem olan Bi-LSTM (sırasıyla 6,36 ve 6,50) ve CNN (sırasıyla 6,20 ve 6,22) diğer yöntemlere kıyasla daha düşük performansa sahip yöntemlerdir.

Şekil 4 (c)'de eğitim süresine göre değerlendirme yer almaktadır. Buna göre, en küçük üç OSD'nin sırasıyla 1,17; 1,85 ve 3,87 olduğu ve bu değerlerin CNN, Deep AE ve DNN yöntemlerine ait olduğu görülmektedir. CNN (OSD = 1,17) yönteminin eğitim süresi, Bi-LSTM (OSD = 8,58) ile LSTM (OSD = 8,21) yöntemlerinin eğitim sürelerinden; Deep AE (OSD = 1,85) yönteminin eğitim süresi ise Bi-LSTM ile LSTM yöntemlerinin eğitim sürelerinden anlamlı olarak daha kısadır. Ayrıca DNN (OSD = 3,87) yöntemi, Bi-LSTM yöntemine göre istatistiksel olarak daha kısa sürede eğitilmiştir. Hesaplamalara örnek vermek gerekirse CNN (OSD = 1,17) ile Bi-LSTM (OSD = 8,58) yöntemlerinin OSD'leri arasındaki fark $8,58 - 1,17 = 7,41$ 'dir ve bu fark kritik fark değerinden ($7,41 > 4,54$) büyüktür. Bu sebeple, süre açısından CNN yönteminin performansı Bi-LSTM yönteminin performansından istatistiksel olarak üstündür. Diğer tüm yorumlar benzer hesaplamalara göre yapılmıştır. Diğer yöntemlerin arasında süre açısından istatistiksel olarak anlamlı fark bulunmamaktadır.



Şekil 4:

a. RMSE, b. MAE ve c. eğitim süresi için Nemenyi testi sonuçları

Derin öğrenme yöntemlerinin ikili karşılaştırmalarında Wilcoxon eşleştirilmiş işaretli sıra testi kullanılmıştır ve bu amaçla kurulan tek taraflı hipotezler şu şekildedir: (i) H_0 : İki derin öğrenme yönteminin performans ölçütü eşit veya birinci derin öğrenme yönteminin performans ölçütü, ikinci derin öğrenme yönteminin performans ölçütünden daha büyüktür. (ii) H_1 : Birinci derin öğrenme yönteminin performans ölçütü, ikinci derin öğrenme yönteminin performans ölçütünden daha küçüktür. Burada, performans ölçütü olarak RMSE, MAE ve eğitim süresi kullanılmıştır.

Ele alınan derin öğrenme yöntemlerinin ikili karşılaştırma sonuçları Tablo 5'te verilmektedir. Tablo 5'teki satırlar hipotezdeki birinci derin öğrenme yöntemini, sütunlar ise hipotezdeki ikinci derin öğrenme yöntemini ifade etmektedir. Buna göre, H_0 hipotezi reddedilen

ikili karşılaştırmalar %5 anlamlılık düzeyinde ($\alpha = 0,05$) ** işareti ile %10 anlamlılık düzeyinde ($\alpha = 0,10$) ise * işareti ile gösterilmiştir. Bu tabloya göre aşağıdaki gözlemler yapılmıştır:

- Tekrarlayan sinir ağlarından RNN ve GRU yöntemleri tüm ölçütlere göre Bi-LSTM'den istatistiksel olarak üstün performansa sahiptir. LSTM ağı MAE ve eğitim süresine göre, ConvLSTM yöntemi ise RMSE ve MAE ölçütlerine göre Bi-LSTM'den daha iyi sonuç vermiştir. DNN yöntemi ise MAE ölçütüne göre RNN'den daha iyi performansa sahiptir.
- CNN, diğer tüm derin öğrenme yöntemlerinden; Deep AE ise CNN dışındaki tüm derin öğrenme yöntemlerinden istatistiksel olarak daha kısa eğitim süresine sahiptir.
- Eğitim süresi açısından DNN yöntemi tekrarlayan sinir ağlarından (RNN, LSTM, GRU ve Bi-LSTM); RNN, GRU ve CNN-LSTM yöntemleri ise LSTM ve Bi-LSTM yöntemlerinden istatistiksel olarak daha iyi performans göstermiştir.
- Diğer ikili karşılaştırmalarda istatistiksel olarak anlamlı bir farklılık bulunmamaktadır.

Yukarıdaki istatistiksel testlere ilave olarak derin öğrenme yöntemlerinin karşılaştırılması amacıyla kazanç, kayıp ve eşitlik sayıları (counts of wins, losses and ties) için işaret testi uygulanmıştır (Demšar, 2006). Bu test, bir yöntemin diğer bir yöntemle göre üstün performans gösterdiği veri kümelerinin sayısını dikkate alır (Demšar, 2006). Kazanç, kayıp ve eşitlik sayıları için kurulan hipotezler şu şekildedir: (i) H_0 : İki derin öğrenme yöntemi eşit sayıda veri kümesinde iyi performansa sahiptir (performansları eşdeğerdir). (ii) H_1 : Bir derin öğrenme yöntemi diğer derin öğrenme yönteminden en az w_α veri kümesinde daha iyi performansa sahiptir. H_0 hipotezinde karşılaştırılan iki yöntem eşdeğer ise her biri N veri kümesinden N/2 kadarını kazanmalıdır. Bir veri kümesinde berabere kalan eşleşmeler (eşitlik durumu) iki yöntem arasında eşit olarak bölünür (Demšar, 2006). Kazanç sayısının dağılımı binom dağılımı olduğunda yedi adet veri kümesi için %95 ($\alpha = 0,05$) ve %90 ($\alpha = 0,10$) güven aralığında kritik kazanç değerleri (w_α) sırasıyla yedi ve altıdır (Demšar, 2006).

Derin öğrenme yöntemleri için kazanç, kayıp ve eşitlik sayıları Tablo 6'da verilmektedir. Tablo 6'da satırdaki ve sütundaki yöntemler karşılaştırılan yöntemleri göstermektedir. Buna göre, %95 güven aralığında aralarında istatistiksel olarak anlamlı fark olan derin öğrenme yöntemleri (kazanç sayısı 7) ** işareti ile, %90 güven aralığında aralarında istatistiksel olarak anlamlı fark olan derin öğrenme yöntemleri ise (kazanç sayısı 6) * işareti ile belirtilmiştir. Buna göre, aşağıdaki gözlemler yapılmıştır:

- %95 güven aralığında GRU yöntemi tüm ölçütlere göre Bi-LSTM yönteminden, eğitim süresine göre ise LSTM yönteminden istatistiksel olarak üstündür.
- %95 güven aralığında DNN yöntemi LSTM, GRU ve Bi-LSTM yöntemlerinden; Deep AE yöntemi ise CNN dışındaki tüm yöntemlerden istatistiksel olarak daha kısa eğitim süresine sahiptir. Ayrıca CNN yöntemi Deep AE dışındaki tüm yöntemlerden; RNN ve CNN-LSTM yöntemleri ise LSTM ve Bi-LSTM'den istatistiksel olarak daha kısa sürede eğitilmiştir.
- %95 güven aralığında diğer ikili karşılaştırmalarda istatistiksel olarak anlamlı bir farklılık bulunmamaktadır.

4.4. Tartışma

Bu bölümde, dokuz farklı derin öğrenme yönteminin yedi adet satış veri kümesindeki performansları tartışılmıştır.

Tahmin performansına göre ele alınan yedi veri kümesinden dördünde LSTM yönteminin diğer derin öğrenme yöntemlerine kıyasla en yüksek performansa sahip olduğu gözlemlenmektedir. Öte yandan tüm veri kümelerinde en yüksek performansa sahip ilk üç yöntemle bakıldığında LSTM'ye ek olarak GRU ve CNN-LSTM yöntemlerinin öne çıktığı görülmektedir. Bu durum, tahmin performansı dikkate alındığında tekrarlayan sinir ağı varyantlarının (LSTM, GRU, hibrit CNN-LSTM) diğer yöntemlere üstünlük sağladığını

göstermektedir. Bu yöntemler sıralı verileri işleyerek belirli aralıklar ile geçmiş verilerin hatırlanmasını sağlamaktadır. Bu şekilde, verideki zamansal bağımlılıkların keşfedilmesi ile başarılı tahmin sonuçları elde edilebilmektedir. Diğer yandan, tahmin performansına göre her veri kümesinde en düşük performansa sahip üç yöntem arasında Bi-LSTM yöntemi yer almaktadır. Şöyle ki, Bi-LSTM Walmart veri kümesinde en kötü performansa, Rossmann ile Favorita verilerinde en kötü ikinci performansa, Tv_Sales ile Pharma verilerinde ise en kötü üçüncü performansa sahiptir. Bi-LSTM'deki karmaşık ağ yapısı modelin eğitimini zorlaştırmaktadır. Bir diğer kritik bulgu ise derin öğrenme yöntemleri arasında tahmin performansına göre istatistiksel olarak öne çıkan bir yöntemin bulunmamasıdır.

Eğitim süresine göre değerlendirme yapıldığında CNN yönteminin diğer yöntemlerden istatistiksel olarak daha kısa eğitim süresine sahip olduğu görülmektedir. Bunun yanı sıra, Deep AE ve DNN yöntemlerinin de eğitim sürelerinin diğer derin öğrenme yöntemlerinden daha kısa olduğu gözlenmektedir. Bu durumda, CNN yöntemindeki ortaklama katmanı, Deep AE yöntemindeki otokodlayıcı özelliği boyut azaltmaya katkı sağlamaktadır. DNN yönteminde ise ağ yapısı diğer derin öğrenme yöntemlerinden daha basit niteliktedir. Tahmin performansı açısından üstünlük gösteren LSTM'nin diğer yöntemlere kıyasla daha uzun eğitim sürelerine sahip olduğu gözlemlenmiştir. Tahmin performansı iyi olan bir diğer yöntem olan GRU eğitim süreleri açısından LSTM'ye avantaj sağlamaktadır ancak CNN, Deep AE ve DNN yöntemlerine göre bu yöntemin eğitim süresi daha uzundur. Ayrıca CNN, Deep AE ve DNN dışında kalan yöntemler arasında eğitim süresine göre istatistiksel olarak anlamlı bir fark bulunamamıştır.

Tablo 5. Wilcoxon testlerine ait p-değerleri

| Yöntemler | Ölçüt | DNN | Deep AE | CNN | RNN | LSTM | GRU | BiLSTM | CNN-LSTM | ConvLSTM |
|-----------|-------|---------------|---------------|--------------|---------------|---------------|---------------|---------------|---------------|---------------|
| DNN | RMSE | | 0,31 | 0,43 | 0,25 | 0,80 | 0,09* | 0,09* | 0,75 | 0,57 |
| | MAE | | 0,31 | 0,37 | 0,03** | 0,88 | 0,09* | 0,09* | 0,75 | 0,63 |
| | Süre | | 0,99 | 0,99 | 0,06* | 0,01** | 0,01** | 0,01** | 0,12 | 0,43 |
| Deep AE | RMSE | 0,69 | | 0,43 | 0,25 | 0,84 | 0,69 | 0,25 | 0,60 | 0,57 |
| | MAE | 0,69 | | 0,37 | 0,25 | 0,84 | 0,69 | 0,37 | 0,75 | 0,57 |
| | Süre | 0,01** | | 0,95 | 0,01** | 0,01** | 0,01** | 0,01** | 0,01** | 0,01** |
| CNN | RMSE | 0,57 | 0,57 | | 0,25 | 0,94 | 0,31 | 0,25 | 0,88 | 0,57 |
| | MAE | 0,63 | 0,63 | | 0,31 | 0,94 | 0,31 | 0,25 | 0,84 | 0,57 |
| | Süre | 0,01** | 0,05** | | 0,01** | 0,01** | 0,01** | 0,01** | 0,01** | 0,01** |
| RNN | RMSE | 0,75 | 0,75 | 0,75 | | 0,84 | 0,63 | 0,03** | 0,69 | 0,43 |
| | MAE | 0,97 | 0,75 | 0,69 | | 0,88 | 0,69 | 0,03** | 0,69 | 0,80 |
| | Süre | 0,94 | 0,99 | 0,99 | | 0,01** | 0,69 | 0,01** | 0,63 | 0,75 |
| LSTM | RMSE | 0,20 | 0,16 | 0,06* | 0,16 | | 0,20 | 0,16 | 0,16 | 0,20 |
| | MAE | 0,12 | 0,16 | 0,06* | 0,12 | | 0,16 | 0,05** | 0,16 | 0,20 |
| | Süre | 0,99 | 0,99 | 0,99 | 0,99 | | 0,99 | 0,03** | 0,99 | 0,88 |
| GRU | RMSE | 0,91 | 0,31 | 0,69 | 0,37 | 0,80 | | 0,01** | 0,50 | 0,37 |
| | MAE | 0,91 | 0,31 | 0,69 | 0,31 | 0,84 | | 0,01** | 0,57 | 0,37 |
| | Süre | 0,99 | 0,99 | 0,99 | 0,31 | 0,01** | | 0,01** | 0,75 | 0,57 |
| BiLSTM | RMSE | 0,91 | 0,75 | 0,75 | 0,97 | 0,85 | 0,99 | | 0,88 | 0,94 |
| | MAE | 0,91 | 0,63 | 0,75 | 0,97 | 0,95 | 0,99 | | 0,88 | 0,96 |
| | Süre | 0,99 | 0,99 | 0,99 | 0,99 | 0,97 | 0,99 | | 0,99 | 0,88 |
| CNN-LSTM | RMSE | 0,25 | 0,40 | 0,12 | 0,31 | 0,84 | 0,50 | 0,12 | | 0,37 |
| | MAE | 0,25 | 0,25 | 0,16 | 0,31 | 0,84 | 0,43 | 0,12 | | 0,43 |
| | Süre | 0,88 | 0,99 | 0,99 | 0,37 | 0,01** | 0,25 | 0,01** | | 0,57 |
| ConvLSTM | RMSE | 0,43 | 0,43 | 0,43 | 0,57 | 0,80 | 0,63 | 0,06* | 0,63 | |
| | MAE | 0,37 | 0,43 | 0,43 | 0,20 | 0,80 | 0,63 | 0,04** | 0,57 | |
| | Süre | 0,57 | 0,99 | 0,99 | 0,25 | 0,12 | 0,43 | 0,12 | 0,43 | |

* %90 güven aralığında H_0 hipotezi reddedilen testler koyu olarak * ile işaretlenmiştir ($\alpha = 0,10$).

** %95 güven aralığında H_0 hipotezi reddedilen testler koyu olarak ** ile işaretlenmiştir ($\alpha = 0,05$).

Tablo 6. Kazanç, kayıp ve eşitlik sayıları (kazanç/kayıp/eşitlik)

| Yöntem | Ölçüt | DNN | Deep AE | CNN | RNN | LSTM | GRU | BiLSTM | CNN-LSTM | ConvLSTM |
|----------|-------|----------------|---------------|---------------|----------------|----------------|----------------|----------------|----------------|----------------|
| DNN | RMSE | 0/0/7 | 3/4/0 | 4/3/0 | 4/3/0 | 2/5/0 | 4/3/0 | 4/3/0 | 2/5/0 | 3/4/0 |
| | MAE | 0/0/7 | 3/4/0 | 4/3/0 | 5/2/0 | 2/5/0 | 4/3/0 | 4/3/0 | 2/5/0 | 3/4/0 |
| | Süre | 0/0/7 | 0/7/0 | 0/7/0 | 6/1/0* | 7/0/0** | 7/0/0** | 7/0/0** | 5/2/0 | 4/3/0 |
| Deep AE | RMSE | 4/3/0 | 0/0/7 | 4/3/0 | 4/3/0 | 2/5/0 | 2/5/0 | 4/3/0 | 3/4/0 | 3/4/0 |
| | MAE | 4/3/0 | 0/0/7 | 4/3/0 | 4/3/0 | 2/5/0 | 2/5/0 | 4/3/0 | 2/5/0 | 3/4/0 |
| | Süre | 7/0/0** | 0/0/7 | 1/6/0 | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** |
| CNN | RMSE | 3/4/0 | 3/4/0 | 0/0/7 | 3/4/0 | 1/6/0 | 3/4/0 | 3/4/0 | 2/5/0 | 2/5/0 |
| | MAE | 3/4/0 | 3/4/0 | 0/0/7 | 3/4/0 | 1/6/0 | 3/4/0 | 3/4/0 | 2/5/0 | 2/5/0 |
| | Süre | 7/0/0** | 6/1/0* | 0/0/7 | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** | 7/0/0** |
| RNN | RMSE | 3/4/0 | 3/4/0 | 4/3/0 | 0/0/7 | 3/4/0 | 2/5/0 | 6/1/0* | 3/4/0 | 3/4/0 |
| | MAE | 2/5/0 | 3/4/0 | 4/3/0 | 0/0/7 | 3/4/0 | 2/5/0 | 6/1/0* | 3/4/0 | 2/5/0 |
| | Süre | 1/6/0 | 0/7/0 | 0/7/0 | 0/0/7 | 7/0/0** | 3/4/0 | 7/0/0** | 3/4/0 | 3/4/0 |
| LSTM | RMSE | 5/2/0 | 5/2/0 | 6/1/0* | 4/3/0 | 0/0/7 | 4/3/0 | 4/3/0 | 5/2/0 | 4/3/0 |
| | MAE | 5/2/0 | 5/2/0 | 6/1/0* | 4/3/0 | 0/0/7 | 4/3/0 | 5/2/0 | 5/2/0 | 4/3/0 |
| | Süre | 0/7/0 | 0/7/0 | 0/7/0 | 0/7/0 | 0/0/7 | 0/7/0 | 5/2/0 | 0/7/0 | 1/6/0 |
| GRU | RMSE | 3/4/0 | 5/2/0 | 4/3/0 | 5/2/0 | 3/4/0 | 0/0/7 | 7/0/0** | 4/3/0 | 5/2/0 |
| | MAE | 3/4/0 | 5/2/0 | 4/3/0 | 5/2/0 | 3/4/0 | 0/0/7 | 7/0/0** | 4/3/0 | 5/2/0 |
| | Süre | 0/7/0 | 0/7/0 | 0/7/0 | 4/3/0 | 7/0/0** | 0/0/7 | 7/0/0** | 3/4/0 | 4/3/0 |
| BiLSTM | RMSE | 3/4/0 | 3/4/0 | 4/3/0 | 1/6/0 | 3/4/0 | 0/7/0 | 0/0/7 | 3/4/0 | 2/5/0 |
| | MAE | 3/4/0 | 3/4/0 | 4/3/0 | 1/6/0 | 2/5/0 | 0/7/0 | 0/0/7 | 3/4/0 | 1/6/0 |
| | Süre | 0/7/0 | 0/7/0 | 0/7/0 | 0/7/0 | 2/5/0 | 0/7/0 | 0/0/7 | 0/7/0 | 1/6/0 |
| CNN-LSTM | RMSE | 5/2/0 | 3/4/0 | 5/2/0 | 4/3/0 | 2/5/0 | 3/4/0 | 4/3/0 | 0/0/7 | 3/4/0 |
| | MAE | 5/2/0 | 5/2/0 | 5/2/0 | 4/3/0 | 2/5/0 | 3/4/0 | 4/3/0 | 0/0/7 | 3/4/0 |
| | Süre | 2/5/0 | 0/7/0 | 0/7/0 | 3/4/0 | 7/0/0** | 4/3/0 | 7/0/0** | 0/0/7 | 3/4/0 |
| ConvLSTM | RMSE | 4/3/0 | 4/3/0 | 5/2/0 | 4/3/0 | 3/4/0 | 2/5/0 | 5/2/0 | 4/3/0 | 0/0/7 |
| | MAE | 4/3/0 | 4/3/0 | 5/2/0 | 5/2/0 | 3/4/0 | 2/5/0 | 5/2/0 | 4/3/0 | 0/0/7 |
| | Süre | 3/4/0 | 0/7/0 | 0/7/0 | 4/3/0 | 6/1/0* | 3/4/0 | 6/1/0* | 4/3/0 | 0/0/7 |

* %90 güven aralığında istatistiksel olarak anlamlı farka sahip yöntemler koyu olarak * ile işaretlenmiştir ($\alpha = 0,10$).

** %95 güven aralığında istatistiksel olarak anlamlı farka sahip yöntemleri koyu olarak ** ile işaretlenmiştir ($\alpha = 0,05$).

5. SONUÇLAR

Etkili talep tahmini, tedarik zincirinde karar verme ve planlama süreçlerini iyileştirerek müşteri taleplerinin doğru zamanda, doğru yerde ve doğru miktarda karşılanmasını sağlamaktadır. Ayrıca müşteri memnuniyetinin artırılmasına katkı sağlayarak firmalara rekabet avantajı kazandırmaktadır. Ancak talep tahmini çözülmesi zor ve karmaşık bir problemdir. Bu doğrultuda, bu çalışma son yıllarda sıklıkla kullanılan ve karmaşık ilişkilerin çıkarılmasını sağlayan derin öğrenme yöntemlerinin satış tahmini probleminde karşılaştırılmasını ele almıştır.

Derin öğrenme yöntemlerinin karşılaştırılmasında çeşitli sektörlerde ait satış verileri kullanılmıştır. Bu veriler mevsimsellik, trend gibi farklı özelliklere sahiptir. Çalışma kapsamında çok katmanlı yapay sinir ağları (DNN), zamansal bağımlılıkları çıkarabilen tekrarlayan sinir ağları (RNN, LSTM, GRU ve BiLSTM), öznelik çıkarma özelliğine sahip yöntemler (Deep AE ve CNN) ve hibrit yaklaşımlar (CNN-LSTM ve ConvLSTM) ele alınmıştır. Performanslar arasındaki farkların istatistiksel olarak anlamlılığını değerlendirmek için parametrik olmayan testler kullanılmıştır.

DeneySEL sonuçlara göre tahmin doğruluğunda LSTM ve GRU yöntemleri; eğitim süresinde ise DNN, Deep AE ve CNN yöntemleri öne çıkmıştır. Mevsimsellik veya trend içeren değişkenliği yüksek veri kümelerinde LSTM ve GRU yöntemlerinin, değişkenliği düşük veri kümelerinde CNN yönteminin ve birden çok farklı özelliğe (bileşene) sahip zaman serilerinde ise CNN-LSTM veya ConvLSTM gibi hibrit yöntemlerin daha iyi performans gösterdiği gözlemlenmiştir. Diğer yandan, tahmin performansına göre derin öğrenme yöntemleri arasında istatistiksel olarak anlamlı bir fark bulunmadığı gözlemlenmiştir.

Bu çalışma, çeşitli sektörlerde ait satış tahmini problemi için güncel derin öğrenme yöntemlerini kapsamlı bir şekilde karşılaştırması ile literatüre katkı sağlamaktadır. Gelecek çalışmalarda tatil olup olmaması, hava durumu, ürün özellikleri gibi satışı etkileyen diğer

faktörler öznitelik olarak eklenebilir. Böylece, satış tahmini problemi çok değişkenli zaman serisi olarak ele alınabilir. Ek olarak, topluluk öğrenme ve transfer öğrenme yaklaşımlarının satış problemi için geliştirilen derin öğrenme yöntemlerindeki performansı analiz edilebilir.

TEŞEKKÜR

Bu araştırma Bursa Uludağ Üniversitesi (BUÜ) Bilimsel Araştırma Projeleri (BAP) Birimi tarafından desteklenmiştir (Proje Kodu: FDK2021-518).

ÇIKAR ÇATIŞMASI

Yazarlar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

YAZAR KATKISI

Begüm Erol'un literatür araştırması, veri toplanması ve analizi, yöntemlerin geliştirilmesi ve uygulanması, deneylerin yapılması, deneysel sonuçların analiz edilmesi ve makale yazımında; Tülin İnkaya'nın konunun oluşturulması, sonuçların irdelenmesi ve yorumlanması, yazım ve içerik denetimi, eleştirel gözden geçirme hususlarında katkıları bulunmaktadır.

KAYNAKLAR

1. Acı, M., and Doğansoy G. A. (2022) Demand forecasting for e-retail sector using machine learning and deep learning methods, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 37(3), 1325-1339. doi: 10.17341/gazimmfd.944081
2. Belas, A., and Bidyuk, P. (2021) Convolutional neural networks for modeling and forecasting nonlinear nonstationary processes, *ScienceRise*, (3), 12-20. doi:10.21303/2313-8416.2021.001924
3. Bousqaoui, H., Slimani, I., and Achchab, S. (2021) Comparative analysis of short-term demand predicting models using ARIMA and deep learning, *International Journal of Electrical & Computer Engineering*, 2088-8708, 11(4). doi:10.11591/ijece.v11i4.pp3319-3328
4. Buyar, V., and Abdel-Raouf, A. (2019) A convolutional neural networks-based model for sales prediction, In Proceedings of the 2019 *International Conference on Artificial Intelligence, Robotics and Control*, 61-67. doi:10.1145/3388218.3388228
5. Chandriah K. K., Naraganahalli, R. V. (2021) RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting, *Multimedia Tools and Applications*, 1-15. doi:10.1007/s11042-021-10913-0
6. Cho, K., Van M. B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv*. doi:10.48550/arXiv.1406.1078
7. Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling, *arXiv*. doi: 10.48550/arXiv.1412.3555
8. Demšar J. (2006) Statistical comparisons of classifiers over multiple data sets, *The Journal of Machine Learning Research*, 7, 1-30.

9. Dharshini MPA, and Vijila SA. (2021) Survey of machine learning and deep learning approaches on sales forecasting, *In 2021 4th International Conference on Computing and Communications Technologies (ICCCT)*, 59-64. doi:10.1109/ICCCT53315.2021.9711878
10. Eglite, L., and Birzniece, I. (2022) Retail Sales Forecasting Using Deep Learning: Systematic Literature Review, *Complex Systems Informatics and Modeling Quarterly*, (30), 53-62. doi:10.7250/csimq.2022-30.03
11. Erol, B., and İnkaya, T. (2024) Satış tahmini için uzun kısa-sürelili bellek ağı tabanlı derin transfer öğrenme yaklaşımı, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 39(1), 191-202. doi:10.17341/gazimmfd.1089173
12. Gashler, M. S., and Ashmore, S. C. (2016) Modeling time series data with deep Fourier neural networks, *Neurocomputing*, 188, 3-11. doi:10.1016/j.neucom.2015.01.108
13. Graves, A., and Schmidhuber, J. (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural networks*, 18(5-6), 602-610. doi:10.1016/j.neunet.2005.06.042
14. Goodfellow, I., Bengio, Y., and Courville, A. (2016) Deep learning, *USA: MIT Press*, Vol 1.
15. Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. (2016) LSTM: A search space odyssey, *IEEE transactions on neural networks and learning systems*, 28(10), 2222-2232. doi:10.1109/TNNLS.2016.2582924
16. Hochreiter, S., and Schmidhuber, J. (1997) LSTM can solve hard long time lag problems, *Advances in neural information processing systems*, 473-479.
17. Iman, R. L., Davenport, J. M. (1980) Approximations of the critical region of the fbietkan statistic, *Communications in Statistics-Theory and Methods*, 9(6), 571-595. doi:10.1080/03610928008827904
18. Ingle, C., Bakliwal, D., Jain, J., Singh, P., Kale, P., and Chhajed, V. (2021) Demand forecasting: Literature review on various methodologies, *In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-7. doi:10.1109/ICCCNT51525.2021.9580139
19. Jiang, L., Rollins, K. M., Ludlow, M., and Sadler, B. (2020) Demand forecasting for alcoholic beverage distribution, *SMU Data Science Review*, 3(1), 5.
20. Kaggle, (2020). <https://www.kaggle.com/datasets> (Erişim tarihi: 5.10.2020).
21. Kingma, D. P., and Ba, J. (2014) Adam: A method for stochastic optimization, *arXiv*. doi:10.48550/arXiv.1412.6980
22. Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2021) 1D convolutional neural networks and applications: A survey, *Mechanical systems and signal processing*, 151, 107398. doi:10.1016/j.ymssp.2020.107398
23. LeCun, Y., Bengio, Y., and Hinton, G. (2015) Deep learning, *nature*, 521(7553), 436-444.
24. Liu, Y., Lan, K., Huang, F., Cao, X., Feng, B., and Zhu, B. (2021) An aggregate store sales forecasting framework based on ConvLSTM, *In 2021 The 5th International Conference on Compute and Data Analysis*, 67-72. doi:10.1145/3456529.3456540

25. Muhaimin, A., Prastyo, D. D., and Lu, H. H. S. (2021) Forecasting with recurrent neural network in intermittent demand data, *In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 802-809. doi: 10.1109/Confluence51648.2021.9376880
26. Pacella, M., and Papadia, G. (2021) Evaluation of deep learning with long short-term memory networks for time series forecasting in supply chain management, *Procedia CIRP*, 99, 604-609. doi:10.1016/j.procir.2021.03.081
27. Peköz, A. Z., and İnkaya, T. (2023) Derin öğrenme ile talep tahmini: Bir üçüncü parti lojistik firması için COVID-19 döneminde vaka analizi, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 24(4), 705-712. doi:10.5505/pajes.2022.73537
28. Puspita, P. E., İnkaya, T., and Akansel, M. (2019) Clustering-based sales forecasting in a forklift distributor, *International Journal of Engineering Research and Development*, 11 (1), 25-40. doi:10.29137/umagd.473977
29. Qi, Y., Li, C., Deng, H., Cai, M., Qi, Y., and Deng, Y. (2019) A deep neural framework for sales forecasting in e-commerce, *In Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 299-308. doi:10.1145/3357384.3357883
30. Ren, S., Choi, T. M., Lee, K. M., and Lin, L. (2020) Intelligent service capacity allocation for cross-border-E-commerce related third-party-forwarding logistics operations: A deep learning approach, *Transportation Research Part E: Logistics and Transportation Review*, 134, 101834. doi:10.1016/j.tre.2019.101834
31. Rizvi, S. M., Syed, T., and Qureshi, J. (2021) Real-time forecasting of petrol retail using dilated causal CNNs, *Journal of Ambient Intelligence and Humanized Computing*, 1-12. doi:10.1007/s12652-021-02941-3
32. Sarker, I. H. (2021) Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions, *SN Computer Science*, 2(6), 420. doi:10.1007/s42979-021-00815-1
33. Wang, J., Yu, L. C., Lai, K. R., and Zhang, X. (2016) Dimensional sentiment analysis using a regional CNN-LSTM model, *In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Volume 2: Short Papers, 225-230.
34. Wang, J., Liu G. Q., and Liu, L. (2019) A selection of advanced technologies for demand forecasting in the retail industry, *In 2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)*, 317-320. doi:10.1109/ICBDA.2019.8713196
35. Wang, S., Jiang, Y., Hou, X., Cheng, H., and Du, S. (2017) Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling, *IEEE Access*, 5, 16576-16583. doi:10.1109/ACCESS.2017.2736558
36. Wang, T., Li, L., and Huang, W. (2020) Research on the construction of sales forecasting model of fashion products based on feature representation of multimodal and deep learning, *WHICEB 2020 Proceedings*, 33.
37. Xingjian, S. H. I., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C. (2015) Convolutional LSTM network: A machine learning approach for precipitation nowcasting, *In Advances in neural information processing systems*, 28, 802-810.