



Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Finansal Piyasalarda Hisse Fiyatlarının Derin Öğrenme ve Yapay Sinir Ağı Yöntemleri ile Tahmin Edilmesi; S&P 500 Endeksi Örneği

 Ali İsa TAŞ^a,  Pelin GÜLÜM^b,  Gökalg TULUM^{c,*}

^a Elektrik-Elektronik Mühendisliği Bölümü, Fen Bilimleri Enstitüsü, İstanbul AREL Üniversitesi, İstanbul, TÜRKİYE

^b Endüstri Mühendisliği Bölümü, Mühendislik ve Mimarlık Fakültesi, İstanbul AREL Üniversitesi, İstanbul, TÜRKİYE

^c Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik ve Mimarlık Fakültesi, İstanbul AREL Üniversitesi, İstanbul, TÜRKİYE

* Sorumlu yazarın e-posta adresi: gokalptulum@gmail.com

DOI: 10.29130/dubited.820620

ÖZET

Gelişen teknolojiler sayesinde günümüzde bilgisayarların gücü artmış ve bununla birlikte farklı amaçlara hizmet eden birçok algoritma geliştirilmiştir. Bu algoritmalar birçok alanda olduğu gibi finans alanında da sıkça kullanılmakta ve karar vermenin farklı boyutlarında destekleyici bir rol üstlenmektedir. Özellikle ortaya çıkması muhtemel durumların önceden öngörülmesinin hayati önem taşıdığı borsa işlemlerinde tahmin yöntemlerine sıkça başvurulmaktadır. Bu çalışmada Yahoo Finans üzerinden elde edilen S&P 500 endeksine ait veriler kullanılarak derin öğrenme ve sığ öğrenme yöntemleri yardımıyla geleceğe yönelik fiyat tahminleme çalışması gerçekleştirilmiştir. Bu bağlamda 12.08.2000 ile 13.8.2020 tarihleri arasındaki günlük fiyat verileri ilk 19 sene (veri setinin %95'i) eğitim, son 1 sene (veri setinin %5'i) test olacak şekilde ayrılarak uzun kısa süreli bellek (LSTM) ve çok katmanlı algılayıcılar (MLP) yöntemleri tahmin gerçekleştirilmiştir. Eğitim, test ve tüm veri kök ortalama karesel hatalarının LSTM ağı için sırasıyla 17.3, 65.3 ve 22 dolar, MLP ağı için sırasıyla 16.1, 61.2 ve 20.6 dolar bulunmuştur. Bu da kullanılan her iki yöntemde elde edilen eğitim ve test hatalarının birbirine yakın sonuçlar verdiğini ve bu yöntemlerin tahmin çalışmaları için uygun seçenekler olduğunu göstermektedir.

Anahtar Kelimeler: Derin Öğrenme, LSTM, Sığ Yapay Sinir Ağı, Zaman Serileri

Deep Learning and Artificial Neural Network Estimation of Stock Prices in Financial Markets; S&P 500 Index Application

ABSTRACT

Thanks to the developing technologies, the power of computers has increased and many algorithms have been developed that serve different purposes. These algorithms are frequently used in finance as well as in many other fields and play a supportive role in different dimensions of decision-making. Forecasting methods are frequently used especially in stock exchange transactions where predicting possible situations in advance is vital. In this study, the future price estimation study was carried out with the help of deep learning and shallow learning methods using the data of the S&P 500 index obtained through Yahoo Finance. In this context, the daily price data between 12.08.2000 and 13.8.2020 is divided into the first 19 years (95% of the data set) training, the last 1 year (5% of the data set) as a test, and long short term memory (LSTM) and multilayer perceptron (MLP) methods have been estimated. Training, test, and all data root mean square errors were 17.3, 65.3 and 22 dollars for the LSTM network, 16.1, 61.2 and 20.6 dollars for the MLP network, respectively. This shows that the training and test errors obtained in both methods used give similar results and these methods are suitable options for prediction studies.

Keywords: Deep Learning, LSTM, Shallow Artificial Intelligence

I. GİRİŞ

Tahmin etmek, hangi değişkenlerin diğer değişkenleri öngördüğünü anlamak anlamına gelir. Bu, birçok değişken arasındaki olası gecikme ilişkilerinin zamanlamasının net bir şekilde anlaşılması, olası gecikme ilişkilerinin istatistiksel önemini ortaya konulması ve piyasa hareketlerinin tahmin edilmesi için sinyal olarak izlenecek değişkenlerden hangilerinin daha önemli olduğunu öğrenmek demektir. Tahmin etme, her zaman pazar profesyonellerinin ve araştırmacıların dikkatini çeken bir konu olduğundan bu bağlamda çeşitli modeller geliştirilmiştir. Söz konusu modellerden özellikle fiyat analizi ve teknik analizler aktif olarak kullanılmasına rağmen finansal varlıkların doğası gereği tutarlı sonuçlar elde etmek oldukça zordur. Teknolojik gelişmelere paralel olarak bilgisayarların hesaplama gücündeki artış özellikle makine öğrenmesi algoritmalarının gelişimini desteklemiş ve daha güçlü algoritmalar ortaya çıkmıştır. Finansal piyasalarda fiyat tahmini yapmak üzere kullanılan birçok tahmin algoritması da bu gelişmeler ışığında gün geçtikçe gelişim göstermiştir.

Literatürde birçok farklı yapay sinir ağı modelinin piyasa değerini tahmin etmek için istatistiksel modellere göre daha fazla tercih edildiği dikkat çekmektedir. Bunun esas nedeni yapay sinir ağı modellerinin çoğu durumda diğer yöntemlere kıyasla daha iyi sonuçlar vermesidir. Bununla birlikte farklı yapay sinir ağı modellerinin kendi içlerinde karşılaştırmasının yapıldığı az sayıda çalışma vardır. Bu çalışmanın, çok uzun bir süreyi kapsayan geniş bir veri seti yardımıyla iki farklı algoritmayı farklı açılardan mukayese ederek gerek araştırmacılara gerek bu konudaki profesyonellere karşılaştırmalı bir bakış açısı sunması hedeflenmektedir.

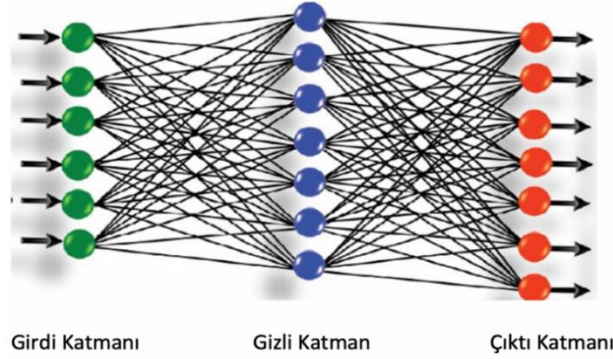
Bu amaçla finansal verilerin çok katmanlı algılayıcı ağları ve uzun-kısa süreli bellek yöntemi ile tahmin edilmesi çalışılmıştır. Bu tahminlerin gerçek yatırımlar ve diğer finansal piyasa endekslerinin genel yönünün analiz edilmesi açısından faydalı olduğu düşünülmektedir. Çalışmanın birincil amacı Uzun Kısa Vadeli Bellek (LSTM) ile Çok Katmanlı Algılayıcı (MLP) yöntemlerini tahmin performansı bakımından karşılaştırmaktır. Araştırmanın ikincil amacı ise daha iyi performans gösteren model hakkında ayrıntılı bilgi vererek modelin özelliklerini ve eksikliklerini analiz etmektir. Bu amaç doğrultusunda Yahoo Finance S&P günlük verilerini kullanılarak söz konusu iki model uygulanmış ve LSTM mimarisinin tahmin gücünün MLP yöntemine göre daha iyi olduğu görülmüştür.

Çalışmanın devamında literatürden örnek çalışmalara değinildikten sonra sığ öğrenme ve MLP, uzun kısa süreli bellek (LSTM) yöntemleri kısaca açıklanmış ve ardından materyal ve metod bölümünde söz konusu yöntemlerin kullanımına ilişkin açıklamalara yer verilmiştir. Yahoo Finance'den elde edilen S&P 500 verileri kullanılarak borsa fiyat tahmini uygulaması gerçekleştirilerek her iki yöntem için sonuçlar değerlendirilmiş ve tartışma bölümünün ardından gelecek çalışmalar için öneriler sunulmuş ve araştırma sonlandırılmıştır.

A. SIĞ ÖĞRENME VE MLP

Yakın zamana kadar, çoğu makine öğrenimi ve sinyal işleme tekniği sığ yapılmış mimarilerden yararlanmaktaydı. Sığ mimariler arasında, geleneksel gizli Markov modeller, doğrusal veya doğrusal olmayan dinamik sistemler, koşullu rastgele alanlar, maksimum entropi modelleri, Destek Vektör Makineleri (SVM), lojistik regresyon, çekirdek regresyon ve çoklu tek bir gizli katmana sahip katman algılayıcı sinir ağları vardır. Sığ öğrenme modellerinde giriş sinyallerini veya özelliklerini probleme özgü bir özellik alanına dönüştürmekten sorumlu yalnızca bir katmandan oluşan basit mimariler söz konusudur [1]. Şekil 1'deki diyagram, üç katmanlı bir MLP'yi göstermektedir. Soldaki birinci katmandaki (giriş katmanı) her algılayıcı, ikinci katmandaki (gizli katman) tüm algılayıcılara çıktılar gönderir ve ikinci katmandaki tüm algılayıcılar çıktıları sağdaki son katmana gönderirler. Her algılayıcı kendisinden sonraki katmana bir sinyal yollar ve her sinyal için farklı ağırlıklar kullanılır. Her katman çok sayıda algılayıcıya sahip olabilirken birden çok katman da olabilir ki böylece çok katmanlı algılayıcı hızla çok karmaşık bir sistem haline gelebilir. MLP'lerden gelişen sinir ağlarında, genellikle 0 ile 1 arasında veya -1 ile 1 arasında gerçek değerlerin çıktıları ile sonuçlanan diğer etkinleştirme işlevleri kullanılabilir. Bu, olasılığa dayalı tahminlere veya öğelerin çoklu etiketlere sınıflandırılmasına imkân verir [2].

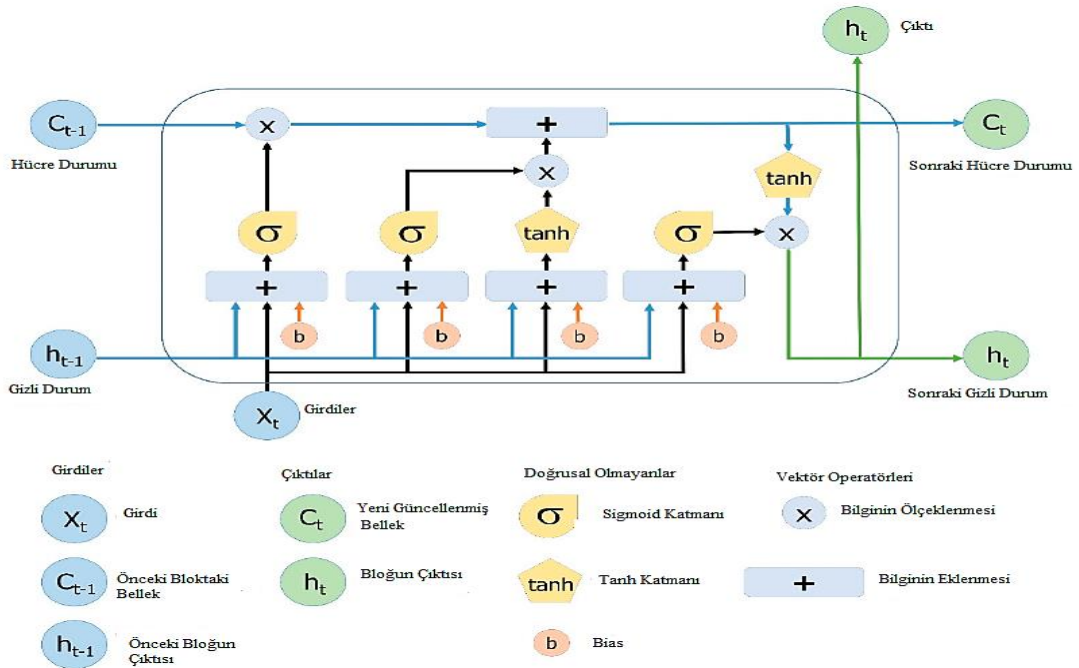
MLP algoritması denetimli öğrenme yöntemlerinden hatanın geri yayılımı algoritmasını kullanmaktadır. Bu yöntemde her bir nöronun ilk ağırlık değeri rastgele belirlenmektedir. Her iterasyonda ağırlık güncellemesi, gerçek çıkış değeri ile hesaplanan sonuç arasındaki hatanın ağına sonundan girişlere doğru yayılması ile gerçekleştirilmektedir. İterasyon işlemi hesaplanan sonuçlar ile denetimli eğitimde kullanılan gerçek değerler arasındaki hatanın belirlenen eşik değerinin altına inene kadar devam etmektedir.



Şekil 1. Sığ öğrenmeli MLP yapısı [3].

B. UZUN KISA SÜRELİ BELLEK (LSTM)

Yıllar içinde öz yinelemeli sinir ağlarının eğitime zorluklarının üstesinden gelmek için bir takım girişimlerde bulunulmuştur. Bu zorluklar, diziyeye dayalı görevlerin uzun vadeli bağımlılıklarını öğrenmek için geçitli yapıya sahip bir öz yinelemeli sinir ağı türü olan LSTM tarafından başarıyla aşılmıştır [4]. LSTM, Olah'a göre zincir yapısı şeklinde tasarlanmıştır. Ancak, yinelenen modül farklı bir yapıya sahiptir. Standart bir öz yinelemeli sinir ağı gibi tek bir sinir ağı yerine, benzersiz bir iletişim yöntemiyle etkileşen dört katmanı vardır [5].



Şekil 2. LSTM mimarisi [6].

LSTM ağı, Şekil 'den de görülebileceği üzere hücreler adı verilen bellek bloklarından oluşur. Hücre durumu ve gizli durum bir sonraki hücreye aktarılmaktadır. Hücre durumu, verilerin temelde değişmeden ileriye doğru akmasına izin veren ana akış veri zinciridir. Ancak, bazı doğrusal dönüşümler meydana gelebilir. Verilerin hücre durumuna eklenmesi veya hücre durumundan çıkarılması gibi olaylar sigmoid kapılar yolu ile gerçekleştirilebilir. Bir geçit, farklı ağırlıklar içeren bir katmana veya bir dizi matris işlemine benzer. LSTM yapıları uzun süreli bağımlılık probleminden kaçınmak amacıyla ezberleme sürecini kontrol eden kapılardan meydana gelmişlerdir.

Bir LSTM ağı oluşturmanın iki adımı vardır bunlardan ilki gerekli olmayan ve o adımda hücreden çıkarılacak bilgileri tanımlamaktır. Verileri tanımlama ve hariç tutma işlemine, (t-1) zamanında son LSTM biriminin (h_{t-1}) ve t zamanında akım girişinin (X_t) çıkışını alan sigmoid işlevi karar verir. Ek olarak, sigmoid işlevi eski çıktının hangi kısmının ortadan kaldırılması gerektiğini belirler. Bu kapıya unutmaya kapısı (veya f_t) denir; burada f_t , hücre durumundaki her sayıya (C_{t-1}) karşılık gelen, 0 ile 1 arasında değişen değerlere sahip bir vektördür.

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

Burada σ , sigmoid fonksiyonudur ve W_f ile b_f unutmaya geçidinin sırasıyla ağırlık matrisleri ve sapma değerleridir. Bir sonraki adım ise, hücre durumunda yeni girişin (X_t) bilgisine karar vermek, depolamak ve hücre durumunu güncellemektir. Bu adım, sigmoid katman ve tanh katmanı olarak iki bölümden oluşur. Birincisi, sigmoid katman yeni bilgilerin güncellenip güncellenmeyeceğine karar verir (0 veya 1) ve ikincisi tanh fonksiyonu, önem düzeylerine (-1 ile 1) karar vererek geçen değerlere ağırlık verir.

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i), \quad (2)$$

$$N_t = \tanh(W_n[h_{t-1}, X_t] + b_n), \quad (3)$$

$$C_t = C_{t-1}f_t + N_t i_t \quad (4)$$

Burada, C_{t-1} ve C_t , sırası ile (t-1) ve t zamanlarındaki hücre durumları iken, W ve b , sırasıyla hücre durumunun ağırlık matrisleri ve sapmasıdır. Son adımda, çıkış değerleri (h_t) çıkış hücresi durumuna (O_t) dayanır, ancak filtrelenmiş bir türüdür. İlk olarak, sigmoid bir katman, hücre durumunun hangi bölümlerinin çıktıya ulaşacağına karar verir. Daha sonra sigmoid geçidinin (O_t) çıkışı, tanh tabakası tarafından hücre durumundan (C_t) oluşturulan yeni değerlerle çarpılır ve -1 ile 1 arasında bir değer elde edilir.

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o), \quad (5)$$

$$h_t = O_t \tanh(C_t). \quad (6)$$

Burada W_o ve b_o , çıkış kapısının sırasıyla ağırlık matrisleri ve sapmasıdır [7].

II. GEÇMİŞ ÇALIŞMALAR VE BULGULAR

Geleneksel tahmin modellerine kıyasla daha az hata payı ile daha doğru sonuçlar sunan LSTM mimarisi son yıllarda oldukça fazla kullanılmaktadır. Standart ileri beslemeli ağlardan farklı olarak LSTM mimarisinin geri bildirim ağları olması özellikle zaman serisi verileri ile sınıflandırma ve tahmin konularında oldukça etkili sonuçlar vermektedir.

Bu bağlamda yapılan çalışmalar incelendiğinde, Liu ve arkadaşları [8], Otoregresif Hareketli Ortalama Modeline (ARMA) dayalı olarak, hem gizli durumları hem de sermaye varlığı fiyatlandırma modelini dikkate almışlardır. Orta vadeli hisse senedi tahmini için orta-LSTM kullanmış olup LSTM, gizli Markov modeli ve doğrusal regresyon ağları olmak üzere üç bileşenden oluşan bir ara LSTM tabanlı

derin sinir ağı tasarlamaşlardır. Ortaya konulan orta-LSTM ağı tasarımının, büyük tahmin hatalarını azaltmak için anormallikleri önleyebilme ve hisse senedi fiyatlarını etkileyen faktörler üzerinde açıklayıcı etkilere sahip olma gibi özellikleri ön plana çıkmaktadır.S&P 500 hisse senetleri üzerinde yaptıkları deneyde önerilen orta-LSTM'nin tahmin doğruluğunda %2-4 iyileşme sağladığını göstermişlerdir.

Kim ve arkadaşları [9], KOSPI 200 endeksinin eğilimlerini tahmin etmek için çeşitli derin öğrenme modellerini kullanmışlardır. Sıralı veri uygulamalarında yaygın olarak kullanılan modellerden MLP, Bir Boyutlu Konvolüsyonel Sinir Ağları (1D-CNN), LSTM ve dikkat ağlarını test etmişlerdir.

Çalışmalarında geçmişe dönük 60 işlem gününü incelemiş ve önerdikleri yaklaşımın çeşitli modellerde daha yüksek tahmin performansı gösterdiğini belirlemişlerdir. Girdi verilerini genişleterek daha fazla günü tahmin çalışmasına dâhil etmenin tahmin doğruluğunu arttıracığı sonucuna varmışlar ve özellikle LSTM içeren modellerin tahmin gücünün diğerlerinden daha iyi olduğunu keşfetmişlerdir.

Mevcut koşulları doğru analiz ederek gelecekteki risk ve fırsatları ön görmenin kritik öneme sahip olduğu alanlardan birisi finanstır. Siami ve arkadaşları [10] yaptıkları çalışma ile LSTM yöntemi gibi zaman serisi verilerini tahmin etmek amacıyla geliştirilmiş derin öğrenme tabanlı algoritmaların, finansal tahminlerde Otoregresif Süreçler (AR), Hareketli Ortalama (MA), Basit Exponansiyel Yumuşatma (SES) ve Bütünleşik Otoregresif Hareketli Ortalama (ARIMA) gibi geleneksel tahmin yöntemlerinden daha üstün performans gösterdiğini ortaya koymuşlardır. Bu bağlamda yaptıkları araştırma ile LSTM yöntemiyle yapılan tahminlerin ARIMA 'ya kıyasla kök ortalama karesel hata (RMSE) hata oranlarında %84-87 arasında azalmaya neden olduğu sonucuna ulaşmışlardır. Ayrıca çalışmada derin öğrenmedeki eğitim sürelerinin sayısının eğitilmiş ağı tahmin performansına bir etkisi olmadığı ve rastgele bir davranış sergilediği de ileri sürülmüştür. Bu çalışmanın aksine Roondiwala ve arkadaşları [11] hisse senetleri tahmini için Tekrarlı Sinir Ağı (RNN) ve LSTM yaklaşımlarını kullandıkları çalışmalarının sonunda RMSE 0.00983 eğitim, RMSE 0.00859 test sonucunun 500 devir ve açılış, kapanış, yüksek, düşük olmak üzere 4 parametre ile elde edildiğini gözlemleyerek devir sayısının önemine değinmişlerdir.

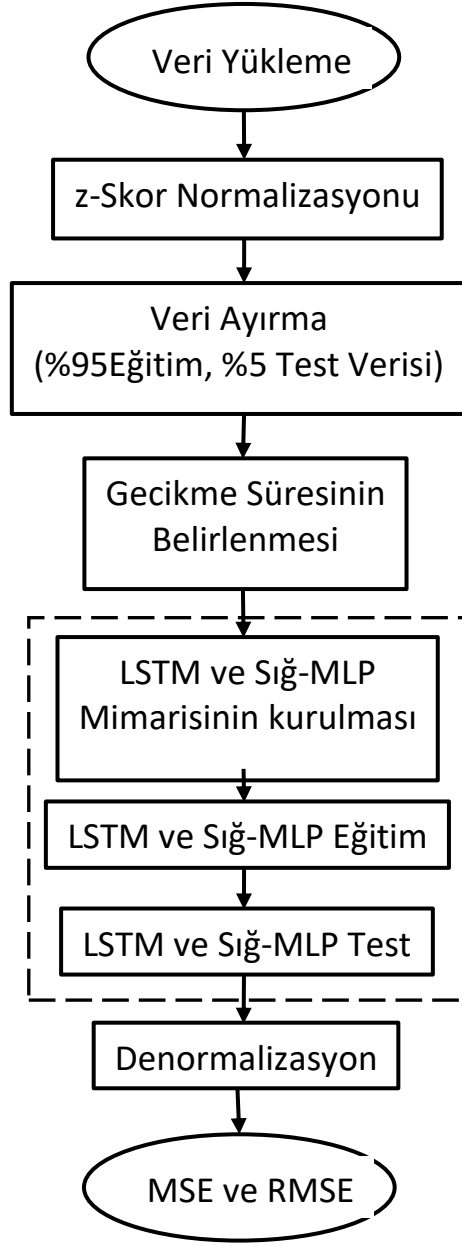
Manurung ve arkadaşları [12] hisse senedi fiyatları tahmini için Yahoo Finance 2013-2018 yılı Bank Central Asia (BCA) verilerini kullanarak verideki önemli parametreler için (açılış, yüksek, düşük, kapanış) LSTM ile tahmin çalışması gerçekleştirmişlerdir. Analizin sonucunda LSTM'deki en doğru tahminin 3 yıllık veya 5 yıllık eğitim verisi kullanmak yerine 1 yıldan az kısa vadeli veri kullanarak elde edildiği ve LSTM'nin %56 doğruluk oranına sahip geleneksel tahmin yöntemi ARIMA'ya kıyasla %94'e kadar kısa vadeli verilerde üstün olduğunu ortaya koymuştur. Böylece LSTM mimarisinin performansının dikkate alınacak zaman aralığından etkilendiği sonucu ortaya koyulurken bunun yanında geleneksel modellerden farkı da tespit edilmiştir [12]. Benzer konuda yaptıkları çalışma ile Zhuge ve arkadaşları [13] hisse senedi açılış fiyatlarını tahmin etmek amacıyla duygusal analiz (emotional analysis) ve LSTM modelinden oluşan iki aşamalı bir yaklaşım önermişlerdir. Verileri sınıflandırmak için ilk adımda Bayesian temelli sınıflandırıcı kullanılmış ve ardından önceki deneylerden elde edilen duygusal veriler, uzun davranış verileri LSTM öğrenme modeli için eğitim verisi olarak gerçek davranış verileri ile birleştirilerek kullanılmıştır. Sonuç olarak tahmin doğruluğu performansının büyük ölçüde iyileştiği ve önerilen iki aşamalı tekniğin geleneksel sinir ağı yaklaşımlarından daha üstün olduğu ortaya koyulmuştur. Fischer ve arkadaşları [14] uzun kısa süreli bellek (LSTM) ağı kullanarak 1992'den 2015'e kadar S&P 500'ün kurucu hisse senetleri için örneklem dışı yön hareketlerini tahmin etmişlerdir. İşlem maliyetlerinden önce yüzde 0,46'lık günlük getiri ve 5,8 keskinlik oranıyla, LSTM ağlarının daha iyi tahminleme yaptığını göstermişlerdir.

Kai ve arkadaşları, LSTM kullanarak Çin hisse senedi getirilerini modelleyip ve tahmin etmiştir. Çin hisse senedi piyasasının tarihsel verileri, 10 öğrenme özelliği ve 3 günlük kazanç oranı etiketlemesi ile 30 günlük dizilere dönüştürülmüştür. Rastgele tahmin yöntemiyle karşılaştırıldığında, LSTM modeli hisse senedi getirisi tahmininin doğruluğunu % 14,3'ten % 27,2'ye yükseltmiştir [15]. Ferdiansyah ve arkadaşları [16], Yahoo Finans borsasından bitcoin işlem verilerini kullanarak tahmin yaptıkları çalışmalarında zaman serisi teknikleri ile kurdukları model yardımıyla sonraki günler için fiyatı tahmini gerçekleştirmişlerdir. Piyasa dalgalanmaları, siyasi olaylar ,yerel veya küresel ekonomik sorunlar gibi birçok belirsizlik faktöründen etkilenen borsa için bitcoin yatırım kararlarını LSTM ile analiz

etmişlerdir. Jithin ve arkadaşları[17], zamansal verileri analiz etmede verilerden özelliklerin çıkarımı ve çift yönlü LSTM için birden fazla CNN hattını birleştiren yeni bir derin öğrenme modeli önermişlerdir. Geleneksel Destek Vektör Makineleri modeli ile karşılaştırıldığında iki yönlü LSTM ile birleştirilmiş CNN katmanlarının daha gelişmiş tahmin performansı gösterdiğini gözlemlemişlerdir. Ayrıca çalışmada kullanılan üç ayrı işlem hattındaki en iyi sonuçları birleştiren tek işlem hattının daha iyi bir sonuç sunduğunu ortaya koymuşlardır. Bu yaklaşım derin öğrenme modellerinin zaman serisi eğilimlerini tahmin etmede uygulanabilirliğini göstermiştir. Ioannis ve arkadaşları [18] altın fiyatı hareketlerinin tahmini için CNN-LSTM entegrasyonlu yeni bir tahmin modeli önermişlerdir. Bu çalışmadaki CNN-LSTM modelinin iki türünden birisi farklı sayıda filtrele ve iki evrişimli katmanlara sahip derin öğrenme ve diğeri makine öğrenimi temelli tahmin modelleridir. Birincisi en düşük MAE ve RMSE performansını veren regresyon problemleri için en iyi tahmin performansını gösterirken, ikincisi fiyat hareketleri için geleneksel zaman serisi modellerinden daha iyi performans göstermiştir. Sonuç olarak LSTM modellerinin altın fiyatı zaman serileri için yaygın olarak kabul gören bir yaklaşım olmasına rağmen ek evrişimli katmanlarla birlikte kullanımının tahmin performansında dikkate değer bir artışa neden olduğu belirlenmiştir.

III. MATERYAL VE YÖNTEM

Çalışmada kullanılan yöntemlere ilişkin akış Şekil 3'te verilmiştir. LSTM ve MLP yöntemlerinin tahmin gücünü karşılaştırmaya yönelik gerçekleştirilen bu çalışmanın ilk adımı tahmin için kullanılacak verilerin elde edilmesi ve uygun forma getirilmesidir. Ardından sırası ile LSTM ve MLP teknikleri ile tahmin yapılarının kurulması ve ağların eğitilerek sonuçlarının elde edilmesi adımları gelmektedir. Çalışmanın devamında tahmin hatalarını belirlemek için RMSE ve MSE değerlerinin elde edilmesi ve sonuçların analizi yer almaktadır.



Şekil 3. Geliştirilen tahmin algoritması.

A. VERİ SETİ

Bu çalışmada Finance Yahoo (<https://finance.yahoo.com/>) sitesinden alınan S&P veri seti kullanılmıştır. Veri setinin high, open, close, low gibi değerleri vardır. Burada veri setinin high, dönem içinde ulaşılan en yüksek değer. Open, işlem süresinin başlangıç zamanında aldığı ilk değerdir. Close dönemin sonunda alınan son değeri ifade eder. Low dönem içinde alınan en düşük değeri ifade eder [19]. Çalışma kapsamında ise close değerleri kullanılarak işlemler gerçekleştirilmiştir. Veriler 12.08.2000 ile 13.8.2020 tarihleri arasında 5051 günlük değerleri içermektedir.

B. VERİLERİN NORMALİZASYONU VE AYRILMASI

Veri setini tek bir düzeye indirgemek amacıyla normalizasyon uygulanmıştır. Literatürde z skor, min-max ve ondalık ölçekleme gibi farklı normalizasyon teknikleri bulunmaktadır. Bu çalışmada z-skor normalizasyonu tercih edilmiştir. İlk adım olarak finans verileri sisteme yüklenmiş daha sonra z-skor

normalizasyonu uygulanarak veriler -1 ile +1 değerleri arasına getirilmiştir. z-skor normalizasyon verinin ortalamasından çıkartılıp standart sapmasına bölünmesiyle elde edilmiştir [20].

Kurulan ağ yapılarını eğitmek amacıyla veri seti eğitim ve test olmak üzere ayrılmıştır. Bu ayırım için literatürde yer alan (80-20),(70-30), (90-10) gibi oranlar incelenerek ele alınan problem ve veri seti büyüklüğü özelinde en uygun oranın deneme yanılma yolu ile %95 eğitim ve %5 test olduğu görülmüştür[21]. Bu doğrultuda LSTM ve Sığ MLP ağlarının eğitimi ve eğitilmiş ağların test aşamaları için veri setinin başlangıçtan itibaren %95'lik kısmı eğitim (ilk 19 sene eğitim) ve %5'lik kısmı (son 1 sene test) test için ayrılmıştır.

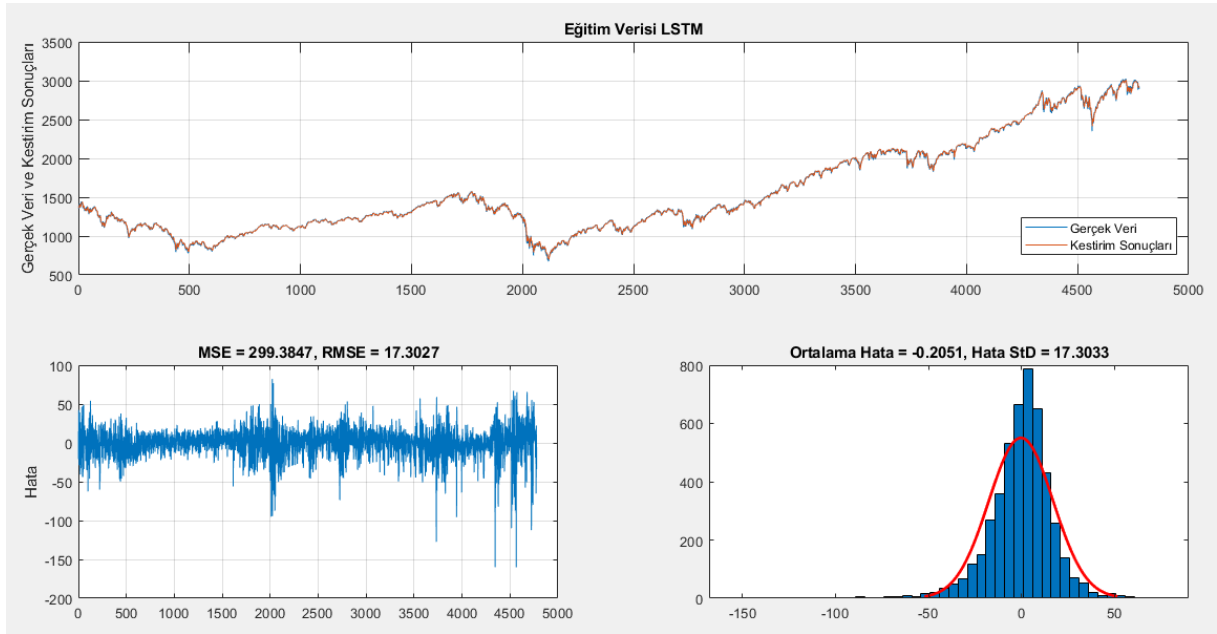
C. GECİKME SÜRELERİ VE AĞ YAPILARI

LSTM ve Sığ MLP ağlarının girişleri için Yahoo finans veri serisinin 11 farklı gün gecikme değerleri kullanılmıştır Bu girişler 1-5, 6, 10, 11, 15, 16 ve 20. gün olacak şekilde deneysel olarak belirlenmiştir. Gecikme seçimi tasarlanırken tahmini yapılacak günün önceki 5 günü, iki, üç ve dördüncü haftalarından da aralarında beşer gün olacak şekilde iki seçim yapılmıştır.

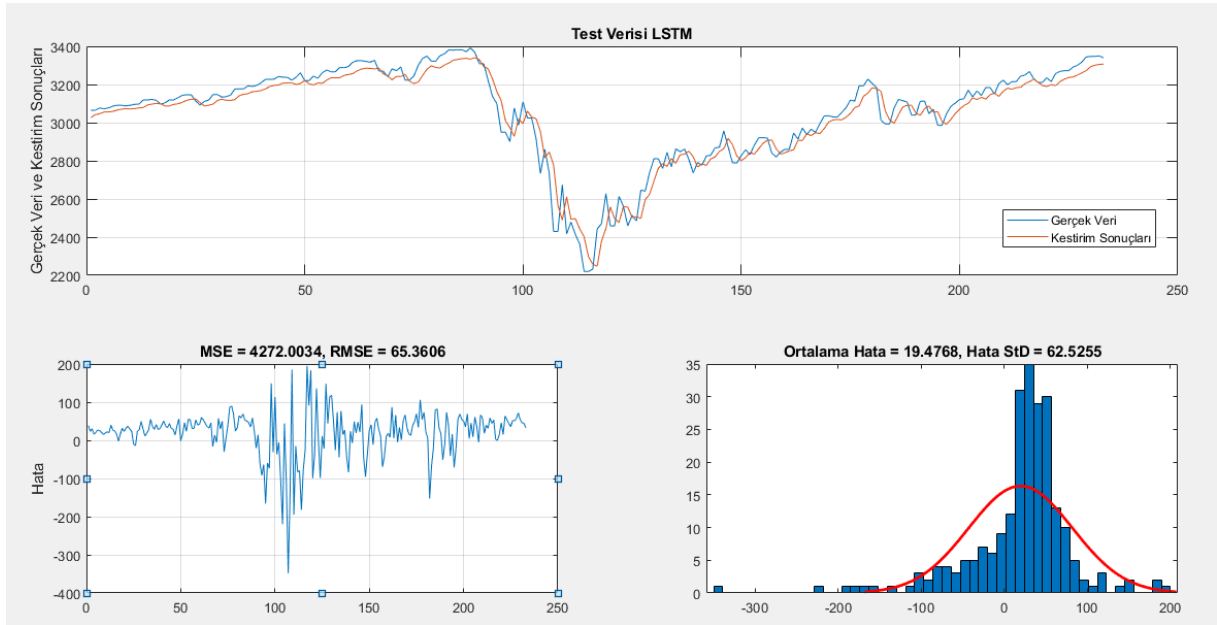
LSTM mimarisi iki gizli katman ve her gizli katmanlarda 300 nöron olacak şekilde tasarlanmıştır. Seyreltme değeri (dropout) 0.2 olarak seçilirken grup (batch size) boyutu 120 ve iterasyon sayısı 1000 olarak belirlenmiştir. Sığ MLP mimarisinde ise gizli katman sayısı iki olarak belirlenmiş (birinci gizli katman 15 nöron, ikinci gizli katman 30 nöron) maksimum iterasyon sayısı 100 olarak seçilerek eğitim için Bayes Düzenlemeli geri yayılım (Bayesian regularization packpropagation) algoritması kullanılmıştır. Eğitim ve test aşamalarından elde edilen çıktıları z-skor denormalizasyonu uygulanarak elde edilen sonuçlar ortalama MSE ve RMSE, cinsinden karşılaştırılmıştır. Ayrıca literatürde verilen sonuçların önerilen çalışmadaki elde edilen sonuçlarla tartışma bölümünde karşılaştırılabilmesi için ortalama mutlak yüzde yapması (MAPE) değerleri de hesaplanmıştır.

IV. UYGULAMA

Yöntem bölümünde detaylı olarak açıklandığı gibi LSTM ve MLP ağları kurulmuş ve belirlenene %95 eğitim %5 test verileri ile eğitilmiştir. Bu adımlar sonucunda elde edilen sonuçlar eğitim, test ve tüm veri olarak ayrı ayrı incelenmiştir. LSTM mimarisi ile elde edilen sonuçların MSE ve RMSE değerleri, ayrıca hatanın histogram ortalama ve standart sapma sonuçları eğitim için Şekil 4'te, test için Şekil 5'te ve tüm veri için Şekil 6'da verilmiştir. Şekil 4 incelendiğinde, ortalama karesel hatanın 300 dolar civarında olduğu ve hatanın çoğunlukla -50 dolar ile 50 dolar arasında dağılım gösterdiği görülmektedir. Şekil 5'te verilen test başarısında ise ortalama karesel hatanın 4727 dolar civarında olduğu ve en büyük hatanın Kovid sürecinin başladığı tarih olan 2020 Mart ayı civarında ortaya çıktığı görülmektedir. Eğitim ve test sonuçları bir arada incelendiğinde histogram eğrilerinden eğitim aşamasında ortalama hatanın 0'a yakın olduğu ve standart sapma değerinin 17 dolar civarı olduğu test aşamasında ise ortalama hata 19 dolara yakın iken standart sapma hatasının 62 dolar mertebesinde değiştiği söylenebilmektedir.

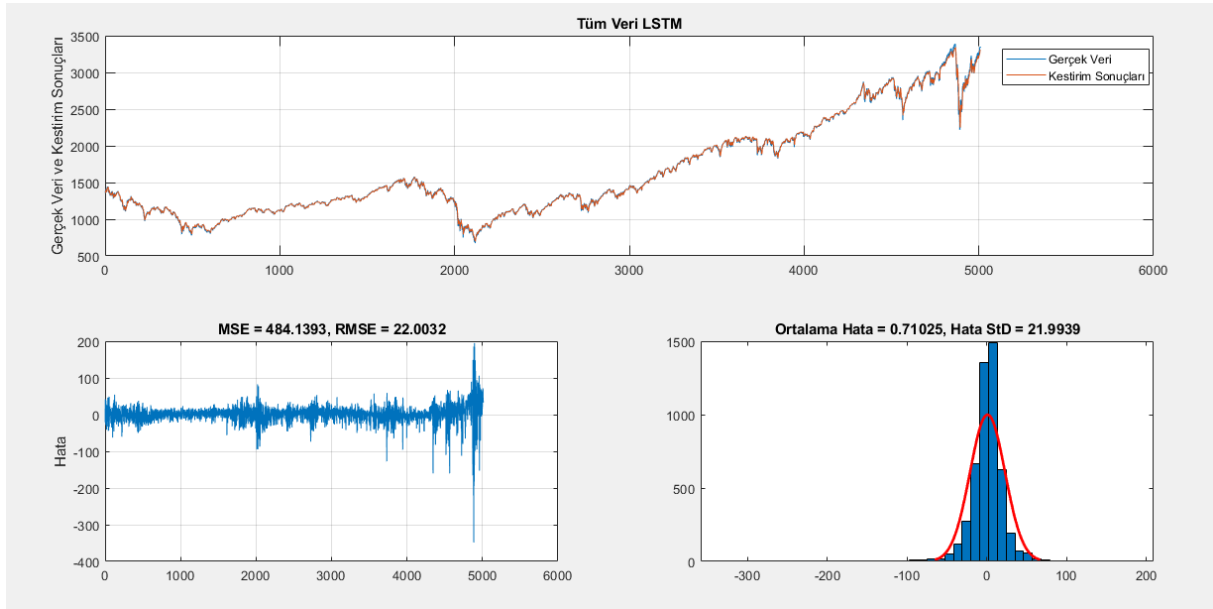


Şekil 4. LSTM eğitim sonuçları.



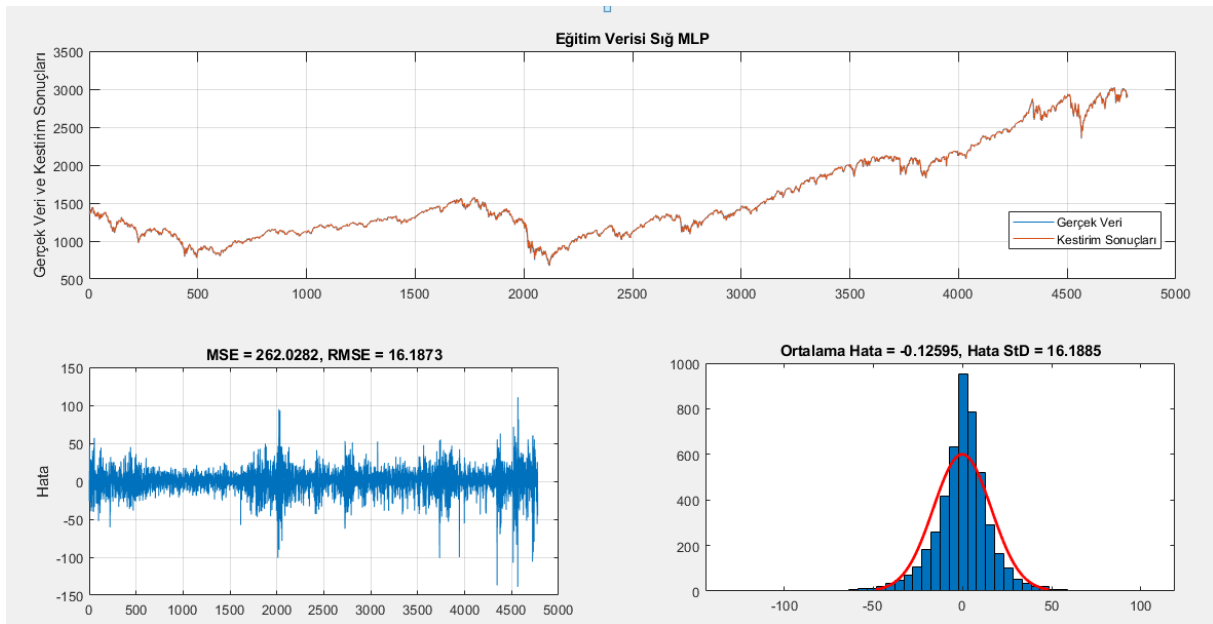
Şekil 5. LSTM test sonuçları.

LSTM mimarisi sonucunda elde edilen eğitim ve test aşamaları birleştirildiğinde ise kök ortalama hatanın 22 dolar mertebesine yakın olduğu ve hata histogram tablosundan ortalama hatanın 0, 7 dolar civarında ve standart sapmanın 22 doları geçmediği görülmüştür.

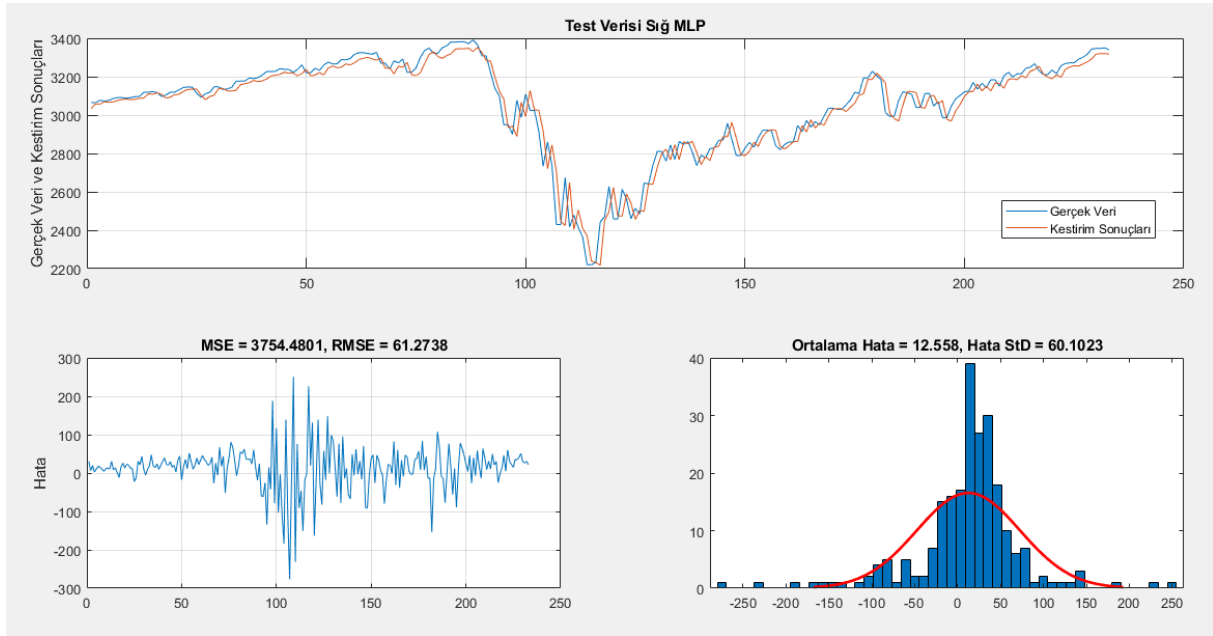


Şekil 6. LSTM tüm veri sonuçları.

Sığ MLP mimarisi ile elde edilen sonuçların MSE ve RMSE değerleri, ayrıca hatanın histogram ortalama ve standart sapma sonuçları eğitim için Şekil 7’de, test için Şekil 8’de ve tüm veri için Şekil 9’da verilmiştir. Şekil 7 incelendiğinde, ortalama karesel hatanın 260 dolar civarında olduğu ve hatanın çoğunlukla -40 dolar ile 40 dolar arasında dağılım gösterdiği görülmüştür. Şekil 8’de verilen test başarısında ise ortalama karesel hatanın 3755 dolar civarında olduğu ve yine en büyük hatanın Kovid sürecinin başladığı tarih olan 2020 Mart ayı civarında ortaya çıktığı görülmektedir. Eğitim ve test sonuçları bir arada incelendiğinde histogram eğrilerinden eğitim aşamasında ortalama hatanın 0 dolara yakın olduğu ve standart sapma değerinin 16 dolar civarı olduğu test aşamasında ise ortalama hatanın 12,5 dolara yakın olduğu fakat standart sapma hatasının 60 dolar mertebesine yükseldiği görülebilmektedir.

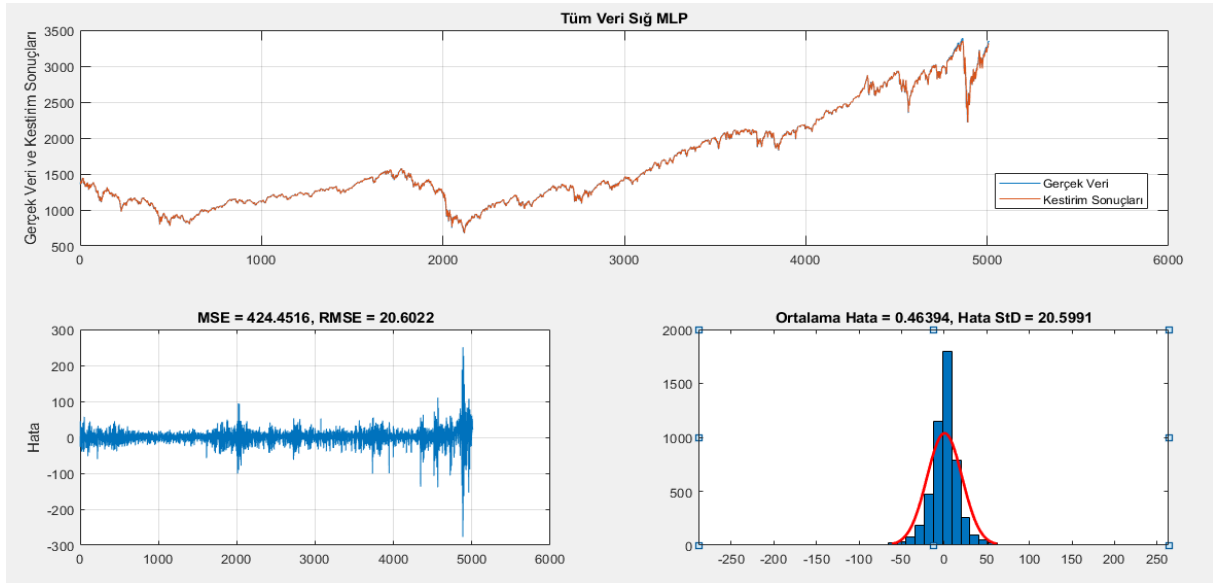


Şekil 7. Sığ MLP eğitim sonuçları.



Şekil 8. Sığ MLP test sonuçları.

Sığ MLP mimarisi sonucunda elde edilen eğitim ve test aşamaları bir arada incelendiğinde kök ortalama hatanın 20 dolar mertebesine yakın olduğu ve hata histogram tablosundan ortalama hatanın 0,5 dolar civarında ve standart sapmanın ise 20,5 dolar hesaplandığı görülmektedir.



Şekil 9. Sığ MLP tüm veri sonuçları.

Tasarlanan mimariler sonucunda elde edilen hata değerlerinin karşılaştırıldığı Tablo 1 incelendiğinde, eğitim başarısı açısından Sığ MLP mimarisinin LSTM mimarisine göre daha başarılı bir kestirim sonucu elde edildiği görülmektedir. Eğitilmiş ağların doğrulanması için gerçekleştirilen test adımında MSE ve RMSE değerleri açısından LSTM mimarisinin Sığ MLP'ye yakın değerler sağladığı görülmektedir. Sığ MLP ve LSTM ağlarının eğitim başarıları RMSE değerleri açısından karşılaştırıldığında Sığ MLP değerinin yaklaşık olarak 1 dolar kadar daha az hata yaptığı görülürken test aşamasında ise Sığ MLP kestirim başarısı yaklaşık olarak 5 dolar kadar daha iyi sonuç vermektedir. Eğitim ve Test adımları bir arada göz önüne alındığında ise Sığ MLP mimarisinin 1,5 dolar kadar daha iyi sonuç verdiği görülmektedir.

Tablo 1. LSTM ve sığ MLP yapılarının hata karşılaştırması.

HATA (\$)	Tüm Veri		Eğitim Verisi		Test Verisi	
	LSTM	Sığ MLP	LSTM	Sığ MLP	LSTM	Sığ MLP
MSE	484,14	424,45	299,38	262,08	4272.0	3754,48
RMSE	22,0	20,60	17.30	16.19	65.36	61,27

Tablo 2’ de her iki mimarinin hata değerlerinden elde edilen hata histogram sonuçları karşılaştırıldığında; Eğitim için beklenen değer (ortalama hata) sonuçlarının her iki mimari içinde yaklaşık olarak 0’a yakın olduğu gözlemlenmiştir. Test verisi için ise ortalama hata olarak MLP mimarisinin yaklaşık olarak 7 dolar kadar daha düşük olduğu görülmektedir. Her iki mimari standart sapma açısından incelendiğinde LSTM ve Sığ MLP mimarilerinin benzer bir dağılıma sahip olduğu görülmektedir. Doğrulama adımında her iki mimari karşılaştırılırsa, Sığ MLP ve LSTM mimarisinin yaklaşık 20 dolarlık bir sapma değerine sahip olduğu hesaplanmıştır.

Tablo 2. LSTM ve sığ MLP hata histogram sonuçları.

HATA HİSTOGRAM	Tüm Veri		Eğitim Verisi		Test Verisi	
	LSTM	Sığ MLP	LSTM	Sığ MLP	LSTM	Sığ MLP
Ortalama Hata	0.710	0.463	-0.205	-0.1259	19.476	12.558
Hata std	21.99	20.599	17.303	16.188	62.525	60.1023

V. SONUÇLAR VE TARTIŞMA

A. TARTIŞMA

Literatürde S&P veri seti ile gerçekleştirilen çalışmalar incelendiğinde farklı algoritmalar ve farklı zamanlardan elde edilen zaman serisi için gerçekleştirilen çalışmalar mevcuttur. Khaled ve arkadaşları 1 Ocak 2010 ile 30 Kasım 2017 arasındaki tarihlerden oluşturdukları S&P veri setini %80 eğitim ve %20 test verisi olarak kullanmış ve Bi-LSTM ve LSTM mimarileri ile gerçek değer kestirim işlemini gerçekleştirmişlerdir. Elde ettikleri sonuçlar incelendiğinde eğitim başarısı açısından 24.452 RMSE ve 0,0119 MAPE değerleri ile en iyi sonucu Bi-LSTM mimarisi ile hesapladıkları görülmektedir. Test verisi açısından yine Bi-LSTM mimarisinin %30 oranında daha düşük bir RMSE değeri sağladığı görülmektedir [22]. Önerilen çalışmada elde edilen sonuçlar ile Khaled ve arkadaşlarının elde ettiği sonuçlar incelendiğinde eğitim işleminde elde edilen RMSE değerinin (MLP için 16.187 ve LSTM için 17.302) daha az hatalı olduğu Tablo 3’ten görülmektedir. Test verisi açısından her iki çalışma karşılaştırıldığında ise LSTM mimarilerinin benzer sonuçlar verdiği gözlemlenmektedir. Cao ve arkadaşları ise 13 Aralık 2007 ve 12 Aralık 2017 tarihleri arasında oluşturdukları 10 senelik veri setini %90 eğitim ve %10 olacak şekilde bölüp farklı algoritmalar ile kestirim sonuçlarını hesaplamış; tüm veri için 18,2 ile 4.8 arasında değişen RMSE değerleri bulmuşlardır [23]. Aynı tarih aralığında Li ve arkadaşlarının gerçekleştirdiği çalışmada ise (9 Kasım 2007 ile 25 Nisan 2017 tarihleri arası) sadece tüm veri seti için MAPE değerleri hesaplanmıştır [8]. Çalışmaların gerçekleştirildiği zaman aralığı incelendiğinde bu süreçte dünya genelinde herhangi bir kriz süreci yaşanmadığı ve S&P borsa verisinin düzenli artış gösterdiği gözlemlenmiştir. Önerilen çalışmada RMSE ve MAPE değerini oldukça düşüren iki kriz faktörü bulunmaktadır. Bunlardan birincisi 2018’de yaşanan ekonomik kriz (Şekil 6 ve Şekil 9 tüm veri grafiği 4500. Adım civarı) ve Mart 2020 tarihli etkisini gösteren covid sürecidir (4.800. adım civarında 3400 \$ mertebesinden 2200 \$ mertebesine düşüş). Yaşanan bu süreçlere rağmen MLP ve LSTM mimarileri ile elde edilen MAPE değerleri (Sığ MLP 0.00815 ve LSTM 0,00897) Bu çalışmalara oldukça yakındır.

Tablo 3. Literatür karşılaştırması.

	Kullanılan Algoritma	Eğitim Verisi		Test Verisi		Tüm Veri	
		RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
Khaled ve diğerleri [22]	BLSTM	24,452	0,01196	46,409	0,02302		
	LSTM	26,418	6	61,393	3		
			0,01302		0,03100		
			0		1		
Cao ve diğerleri [23]	LSTM					18,232	0,00611
	SVM					1	5
	CEEMDAN-SVM					10,244	0,00399
	CEEMDAN-LSTM					8	3
	CEEMDAN-MLP					6,8118	0,00221
						6,1412	3
					4,8291	0,00205	
						1	0,00161
							7
Li ve diğerleri [8]	Linear Regression						0,0747
	Ridge Regression						0,0747
	Random Forest						0,0765
	LSTM						0,0742
Kılıç ve diğerleri [24]	ARIMA	120.152	0,07668	33,293	0,28405		
				5	9		
Önerilen Çalışma	MLP	16,1873	0.00784	61,273	0.01447	20,602	0.00815
	LSTM	17,3027	4	8	1	2	2
			0,00859	65,360	0,01679	22,003	0,00897
			4	6	3	2	5

Kılıç ve Arkadaşlarının gerçekleştirdiği çalışmada [24] ise 1 Ocak 1990 ile 1 Ocak 2011 arasındaki S&P borsa verisinden dalgacık katsayıları çıkartılarak ARIMA regresyon yöntemi kullanılmıştır. Önerilen çalışma ile Kılıç ve arkadaşlarının elde ettikleri RMSE sonuçları karşılaştırıldığında eğitim hatasının önerilen çalışmaya göre yaklaşık 5 dolarlık bir hata yaptığı gözlemlenmektedir. Fakat burada Kılıç ve arkadaşları tarafından eğitim için kullanılan veride borsa değerleri 1000 dolar civarında iken önerilen çalışmada borsa değeri 2000 dolar ile 3000 dolar mertebesinde değişmektedir. Bu neden ile dolar bazlı olarak gerçekleştirilen hata hesaplamasında değerler göz önünde bulundurulmalıdır.

Önerilen çalışmada tüm veri üzerinde başarıyı en çok etkileyen faktör test verisinin kestirilmesinde ortaya çıkmaktadır. Bunun nedeni küresel ölçekte yaşanan ve tüm dünya borsalarının %30 civarında değer kaybetmesine neden olan kovid sürecidir (Şekil 5 ve Şekil 8). Çalışmada önerilen algoritmalar bu süreçte kestirim adımını atlamayarak sadece artış trendi gösteren zaman dilimleri için değil, anlık düşüş ve yükseliş gösteren trendler için de başarılı bir kestirim tahmininde bulunduğu söylenebilir.

V. KAYNAKLAR

- [1] L. Deng, "An overview of deep-structured learning for information processing," in *Proc. Asian-Pacific Signal & Information Annual Summit & Conference (APSIPA-ASC)*, Xi'an, China, 2011
- [2] The deep learning platform. (2018, Sept. 18). *Perceptrons and multi-layer perceptrons: the artificial neuron at the core of deep learning*. [Online]. Available: <https://missinglink.ai/guides/neural->

[network-concepts/perceptrons-and-multi-layer-perceptrons-the-artificial-neuron-at-the-core-of-deep-learning/](#)

- [3] A. Serwa, "Studying the effect of activation function on classification accuracy using deep artificial neural networks," *Journal of Remote Sensing & GIS*, vol. 06, 2017.
- [4] S. Hochreiter, J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, pp. 1735-1780, 1997.
- [5] C. Olah. (2015, Aug. 27). Understanding LSTM Networks. [Online]. Available: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- [6] S. Yan. (2016, Mar. 13). *Understanding LSTM and its diagrams*. [Online]. Available: <https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714>.
- [7] X. H. Le, H. V. Ho, G. Lee, S.Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, pp. 1387-1406, 2019
- [8] X. Li, Y. Li, X.Y. Liu, C. D. Wang, "Risk management via anomaly circumvent: mnemonic deep learning for midterm stock prediction," in *Proc. 2nd KDD Workshop on Anomaly Detection in Finance*, Anchorage, AK, USA, 2019
- [9] S. Kim, M. Kang, "Financial series prediction using attention LSTM," 2019, *arXiv:1902.10877*.
- [10] S. Siami-Namini ve A. Siami Namin, "Forecasting economics and financial time series: ARIMA vs. LSTM," 2018, *arXiv:1803.06386*.
- [11] M. Roondiwala, H. Patel, S. Varma, "Predicting stock prices using LSTM," *International Journal of Science and Research (IJSR)*, vol. 6, pp. 1754-1756, 2017.
- [12] A. H. Manurung, W. Budiharto, H. Prabowo, "Algorithm and modeling of stock prices forecasting based on long short-term memory (LSTM)," *International Journal of Innovative Computing Information and Control*, vol. 12, pp. 1277-1283, 2018.
- [13] Q. Zhuge, L. Xu, G.Zhang, "LSTM neural network with emotional analysis for prediction of stock price," *Engineering Letters*, vol. 25, pp. 167-175, 2017
- [14] T. Fischer, C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, pp. 654-669, 2018
- [15] K. Chen, Y. Zhoui, F.Dai," A LSTM-based method for stock returns prediction : A case study of China stock market," in *2015 IEEE International conference on big data (big data)*, 2015, pp. 2823-2824.
- [16] F. Ferdiansyah, S.H. Othman, R.Z.R.M.Radzi," A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market," *International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2019, pp. 206-210.
- [17] J. Eapen, D. Bein, A. Verma," Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction," In *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*, 2019, pp. 0264-0270.

- [18] I. E. Livieris, E. Pintelas, P. Pintelas, "A CNN-LSTM model for gold price time-series forecasting," *Neural computing and applications*, vol.32, pp.17351-17360, 2020.
- [19] K. Bayındır, "Financial time series prediction with LSTM recurrent neural networks," Yüksek Lisans tezi, Bilgisayar Mühendisliği Bölümü, Bahçeşehir Üniversitesi, İstanbul, Türkiye, 2017.
- [20] Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462
- [21] D. Hammerstrom, 'Working with neural networks', *IEEE spectrum*, vol. 30, no. 7, pp. 46-53, 1993.
- [22] K. A. Althelaya, E.El-Alfy, S. Mohammed, "Evaluation of bidirectional lstm for short-and long-term stock market prediction," *9th international conference on information and communication systems (ICICS)*, 2018, pp. 151-156.
- [23] J. Cao, Z. Li, J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and its Applications*, vol. 519, pp. 127-139, 2018.
- [24] D. K. Kılıç, Ö. Uğur, "Multiresolution analysis of S&P500 time series," *Annals of Operations Research*, vol.260, pp. 197-216, 2018.