



Journal of Economics, Finance and Accounting

Year: 2017 Volume: 4 Issue: 3



ANALYSIS OF FINANCIAL TIME SERIES WITH MODEL HYBRIDIZATION

DOI: 10.17261/Pressacademia.2017.700

JEFA- V.4-ISS.3-2017(12)-p.331-341

Huseyin Ince¹, Fatma Sonmez Cakir²,

1 Gebze Technical University, Department of Economics, Kocaeli, Turkey. h.ince@gtu.edu.tr

2 Bartin University, Department of Management Information Systems, Bartin, Turkey. fonmez@bartin.edu.tr

To cite this document

Ince, H. and Sonmez Cakir F. (2017). Analysis of financial time series with model hybridization. Journal of Economics, Finance and Accounting (JEFA), V.4, Iss.3, p.331-341.

Permament link to this document: <http://doi.org/10.17261/Pressacademia.2017.700>

Copyright: Published by PressAcademia and limited licenced re-use rights only.

ABSTRACT

Purpose- The aim of this study is to obtain better estimation results by hybridizing models that reveal linear and nonlinear relationships used in the intended financial time series.

Methodology- ARIMA and Artificial Neural Networks (ANN) models were used in estimating NASDAQ stock market index values between 03.01.2012 and 30.06.2017 comparison of hybrid model results with different ways of error determination in literature.

Findings- ARIMA residues have been tested different models where only residues are used with basic indications, only residues and basic. The calculation of residues was done separately with the addition and multiplication function. These residues were modeled with ANN, and the obtained results are collected and established hybrid model with ARIMA forecasts. When the results obtained at the end of the operations are compared, it is seen that the product function of some of the addition functions gives better results in some models.

Conclusion- The hybridization of the ANN and NASDAQ index estimates with the ARIMA method resulted in processing for both addition and multiplication functions. Residues calculated with the addition model showed better results in ANN hybrid. What variables are used to calculate residuals is that the hybrid model gives better estimation results than single models.

Keywords: ARIMA, artificial neural network, hybrid models, time series analysis, forecasting

JEL Codes: C13, C15, E27

MODEL MELEZLEME İLE FİNANSAL ZAMAN SERİSİ ANALİZİ

ÖZET

Amaç-Bu çalışmanın amacı finansal zaman serilerinde kullanılan doğrusal ve doğrusal olmayan ilişkileri ortaya çıkartan modelleri melezleyerek daha iyi tahmin sonuçları elde etmektir.

Yöntem- Çalışmada 03.01.2012 ile 30.06.2017 tarihleri arasındaki NASDAQ borsa endeks değerlerinin tahmin edilmesinde ARIMA ve Yapay Sinir Ağları (YSA) modelleri kullanılmış aynı zamanda literatürdeki farklı hata belirleme yolları ile melez model sonuçlarının karşılaştırması yapılmıştır.

Bulgular- ARIMA artıkları sadece temel göstergeler, sadece artıklar ve temel göstergelerle birlikte artıkların kullanıldığı farklı modellerde denemiştir. Artıkların hesaplanması toplama ve çarpma fonksiyonu ile ayrı ayrı yapılmıştır. Bu artıklar YSA ile modellenmiştir ve elde edilen sonuçlar ARIMA tahminleri ile toplanarak melez model kurulmuştur. İşlemler sonunda elde edilen sonuçlar karşılaştırıldığında bazı modellerde toplam fonksiyonunun bazılarında çarpım fonksiyonunun daha iyi sonuç verdiği görülmüştür.

Sonuç- YSA ile NASDAQ endeks tahmininin ARIMA yöntemi ile melezlenmesi sonucunda hem toplama hem çarpma fonksiyonu için işlemler yapılmıştır. Toplama modeli ile hesaplanan artıkların YSA melezinde daha iyi sonuç verdiği görülmüştür. Hangi değişkenler kullanılarak artıklar hesaplanırsa hesaplanırsa melez modelin tekil modellerden daha iyi tahmin sonucu verdiği ortaya çıkmıştır.

Anahtar Kelimeler: ARIMA, yapay sinir ağları, melez modeller, zaman serisi analizi, öngörü

JEL Kodları: C13, C15, E27

1.GİRİŞ

İleri bir dönem veya zamanda finansal serilerin gelecek değerlerinin öngörülmesi ya da en azından küçük bir hata payıyla tahmin edilmesi istenilen bir durumdur. Her birey yapacağı bir eylemden önce bu eylemin kendisine olan getirisinin ya da

kaybının ne olacağını bilmek ister. Zaman serileri analizleri bu nedenle her alan için tercih edilen bir analiz tekniği olmuştur. Zaman serilerinin analiz edilmesi bu analiz sonuçlarının anlaşılır hale getirilip öngöründe kullanılması için modellenmesi önemli bir araştırma alanıdır. Ancak zaman serisi tahmininde bir uygulamada iyi sonuç veren bir model başka uygulamalarda da aynı başarıda sonuçlar vermeyebilir. Bu sebeple öngörü yapılmadan önce verinin durumu tespit edilmelidir. Aynı zamanda verinin değişken sayıları, doğrusal olup olmadığı vb. konularda da bilgi sahibi olunmalıdır. Aksi takdirde serinin analizinden istenilen başarılı sonuç elde edilemeyebilir.

Pek çok alanda olduğu gibi finans alanında da ileriye görebilmek önemli bir araştırma konusudur. Özellikle finans alanında tam ve kesin tahminler yapmak mümkün değildir. Çünkü finansal veriler çok farklı değişkenlerden ve çevre şartlarından etkilenir. Finansal piyasalarda dalgalanmalar mevcuttur ve bu sebeple ne zaman ne olacağını belirlenmesi pek mümkün görülmemektedir. Finansal piyasalara yatırım yapanlar ise geçmiş durumlara bakarak karar vermeye çalışmaktadır. Ancak açıktır ki oynaklığın bu denli çok olduğu bir piyasada kaybetme olasılığı da söz konusudur. Yatırımcılar muhtemel getirileri ile olası kayıplarını daha önceden bilmek ve buna göre yatırım kararları almak istemektedirler. Bu amaca ulaşabilmek için çok çeşitli yöntemler geliştirilmiştir. Her yöntem farklı şartlar altında değişik ve birbirinden daha iyi sonuçlar verebilmektedir. Öngörü ve sınıflandırmada kullanılan geleneksel modellerin yerine son zamanlarda daha modern yöntemler geçmiştir. Bilgisayar teknolojisinin hızla gelişmesi ile bu öngörü ve sınıflandırmayı kendi başlarına yapan yöntemler geliştirilmiştir. Bu yöntemler uygulama açısından diğer yöntemlere göre çok daha avantajlıdır.

Bu araştırmanın uygulanmasında öncelikle model melezlemeye başvurulmasının nedeni birden fazla modelin birbirlerini tamamlayarak tekil olarak elde edecekleri faydanın üzerine çıkılmasını sağlamaktır. Bu amaçla önce tekil olarak ele alınan modeller denenmiş ve bunların sonuçlarına göre daha iyi sonuç vermelerini sağlamak için modeller arasında melezlemeler yapılmıştır. Bu melezlemelerle modellerin her birinin avantajlı ve dezavantajlı oldukları noktalar belirlenerek; bunların ışığında birbirlerinin eksiklerini tamamlayarak daha iyi tahmin yapan melezlenmiş modeller ortaya konulmaya çalışılmıştır.

2. bölümde literatür taraması ile hem ARIMA hem YSA hem de Melez modeller için bilgi ve bazı çalışmalara yer verilmiştir. 3. bölümde analiz için kullanılacak veri ve uygulanacak yöntem hakkında bilgi verilmiştir. 4. bölümde uygulama sonuçlarında elde edilmiş olan bulgulara yer verilmiştir. Bu bölümde kurulan farklı modeller için ARIMA, YSA ve melez model çıktıları sunulmuştur. Sonraki bölümde ise uygulama sonuçları verilmiştir.

2. LİTERATÜR

Öngörü de finansal verilerin kullanıldığı araştırmalara literatürde sıklıkla rastlanmaktadır. Özellikle kambiyo borsaları, iflas tahminleri, hisse senedi fiyatı belirlenmesi gibi alanlarda yapılan birçok çalışma bulunmaktadır. Hisse senedi fiyatları önemli bir ekonomik aktivite olarak ülkelerin ekonomik gelişmelerinde çok önemli bir görev arz etmektedir. Hisse senetlerinin fiyatlarındaki artış şirketleri karlılıklarını artırmakta ve hisse senedi fiyatı artan şirketlere yapılan yatırım kararlarını etkilemektedir. Hisse senetlerindeki fiyat düşüşleri yatırımcıları bu şirketlerden uzaklaştırmakta bu ise şirketin piyasa fiyatının ve karlılığının düşmesine sebep olmaktadır.

Pai, P. F., & Lin, C. S. (2004) Zaman serisi tahmini olarak hisse senedi fiyatlarının seçildiği çalışmada ARIMA modeli ile DVM modelleri hibridlenmiştir. Sonuçta elde edilen Hibrid modelin sonuçları iki yaklaşımla elde edilen sonuçlar ile karşılaştırılmış ve Hibrid modelin tahmin öngörüsünün daha başarılı olduğu ortaya çıkarılmıştır. Fernández ve Gómez (2007) ise çalışmalarında Hong-Kong, Almanya, İngiltere, Amerika, Japonya borsalarını kullanmışlardır. Holpfield yapay sinir ağı modeli ile tavlama benzetimi algoritması, genetik algoritma ve tabu algoritmasını karşılaştırmış ve portföy seçiminde yapay sinir ağlarının daha etkin sonuç verdiğini göstermişlerdir. Yudong ve Lenan (2009) 1998 ile 2008 yılları arasındaki Standard's and Poor's 500 hisse senetleri için etkili bir tahmin modeli kurmak üzerine bir çalışma yapmışlardır. Verilen tarihler arasındaki günlük değerler dikkate alınmıştır. Sutheebanjard P. ve Premchaiswadi W. (2010) Tayland hisse senedi fiyatlarının hareketlerini tahmin etmek için bir çalışma yapmışlardır. Çalışmada Tayland'ın hisse senedi piyasasına odaklanılmış ve piyasa tahmini için geri yayımlı yapay sinir ağı kullanılmıştır. Aghababaeyan, Siddiqui ve diğerleri (2011) Tahran Borsa Endeksi (TSE) üzerinde ileri beslemeli ve geri yayımlı bir YSA modeli kurarak Tahran Borsa Endeksi'nin yönünü %81 olasılıkla tahmin etmişlerdir. Güresen, E., Kayakutlu, G., & Daim, T. U. (2011) hisse senedi fiyat tahmin için bir çalışma yapmıştır. Model çok katmanlı algılayıcı yapısı, dinamik yapay sinir ağları ve GARCH modeli ile hibritlenmiş yapay sinir ağları ile kurulmuştur. NASDAQ 100 endeksine uygulama yapılmıştır. Yakut E., Elmas B. & Yavuz S. (2014) yaptıkları çalışmada ileri beslemeli yapay sinir ağları ve destek vektör makineleri yöntemiyle BIST100 endeksi için etkin bir öngörü modeli kurmaya çalışmışlardır. Bu amaçla endeks değerlerinin 1-2 ve 3 gün öncesine ait değerleri, Dolar kuru, gecelik faiz oranı, Japonya-Brezilya-İngiltere-Fransa-Almanya borsaları için 2005-2012 tarihleri arasındaki borsa endeks değerleri kullanılmış ve BIST 100 değerleri tahmin edilmiştir.

Literatüre bakıldığında zaman serilerinin analizinde kullanılmak üzere oldukça fazla modelin geliştirilmeye devam ettiğini görebiliriz. Çünkü bilişim teknolojileri her geçen gün gelişmekte ve daha karmaşık algoritmalara çözüm üretilebilmektedir. Dolayısıyla yeni algoritma, yöntem ve modellerin ortaya çıkması yeni melez modellerinde araştırılması ihtiyacını beraberinde getirmektedir. Zaman serisinden kastedilen belirli zaman aralıklarında ölçülmüş değerlerdir. Borsalarda işlem

gören hisse senetlerinin fiyatlarının tahmin edilmesi, işsizlik oranının tahmin edilmesi, enerji tüketiminin ya da üretiminin tahmin edilmesi gibi belirli periyodlarla ölçülen değişkenlerin gözlemleri zaman serisi oluşturur. Zaman serilerinin tahmininde girdi olarak hangi değişkenlerin alınacağı modelin başarısı açısından çok önemlidir. Analiz konumuz olan borsa endeks değerlerinin tahmini için kullanılacak birçok değişken vardır. Bu değişkenler hacim, fiyat veya günlük fiyat değişimi gibi direk göstergeler olabileceği gibi bunlardan türetilen hareketli ortalamalar, eğim çizgisi göstergeleri vb. teknik göstergeler veya içsel piyasa değeri, ekonomik çevre vb. ekonomik göstergeler de olabilir (Lawrence, 1997). Lawrence 1997'deki çalışmasında Johannesburg Menkul Kıymetler Borsasının performansını modellemeye çalıştığı araştırmasında belirlenebilecek gösterge sayısının 63 adet olduğundan bahsetmiştir. Ancak uygulaması sonunda bu göstergelerin bazılarının kullanılmasının gerekli olmadığı kanaatine ulaşmıştır. Lawrence tarafından belirlenen göstergeler aşağıdaki gibidir:

1. Temel Göstergeler (3 adet): Hacim, Verim, Fiyat / Kazanç
2. Teknik Göstergeler (17 adet) - Hareketli ortalamalar, Hacim eğilimleri vb.
3. JSE endeksleri (20 adet) - Çeşitli sektörler için pazar endeksleri: Altın, Metal vb.
4. Uluslararası Endeksler (9 adet) - DJIA, NASDAQ, S&P vb.
5. Altın fiyatı / Döviz kurları (3 adet)
6. Faiz Oranları (4 adet)
7. Ekonomik İstatistikler (7 adet) - İhracat, ithalat vs.

2.1.ARIMA Süreci

Zaman serileri için veriler farklı formlara ve rassal süreçlere sahip olabilir. Modelleme yapılırken bir zaman serisi Otoregresif (AR), Hareketli Ortalama (MA) ve Bütünleşik model olarak 3 sınıfa ayrılabilir. Bu üç sınıflandırmada da veriler bir önceki noktalara doğrudan bağlıdır (Gershenfeld, 1999). Bu üç yöntemin farklı kombinasyonları ile Otoregresif ve Hareketli Ortalama (ARMA), Otoregresif Bütünleşik Hareketli Ortalama (ARIMA), Kesirli Birleştirilmiş Otoregresif Hareketli Ortalama (ARFIMA) modelleri oluşturulmuştur.

AR modelleri Yule (1927) tarafından geliştirilmiştir. AR modelleri; bir zaman serisinin herhangi bir dönemdeki gözlem değerini, çalışılmakta olan serinin önceki belli sayıda geçmiş dönem gözlem değerine ve hata terimine bağlı olarak açıklar. Bu modeller, içlerinde bulunan geçmiş dönem gözlem değeri sayısına göre isimlendirilirler. Genel olarak, “ p. dereceden” model AR(p) ile gösterilir ve aşağıdaki şekildedir:

$$w_t = m + \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + u_t \quad (1)$$

Denklemden w_t herhangi bir zaman, u_t hata, w_{t-1} herhangi bir zamandan bir önceki zaman, w_{t-p} herhangi bir zamandan p önceki zamanı gösterir. Denklemden ϕ katsayıları ise modelin parametrelerini gösterir.

MA (Moving Averages) Modelleri ise ilk defa Slutsky (1937) tarafından ortaya atılmıştır ve bu modeller bir zaman serisinin herhangi bir dönemdeki gözlem değerini aynı dönemdeki hata terimi ve ondan önceki belirli sayıda dönemin hata terimine bağlı olarak açıklayan modellerdir. Örneğin bir zaman serisinin şimdiki değeri sabit terim ile iki dönemlik hataların hareketli ortalamalarının toplamına eşittir. Genel olarak, “q.dereceden” model MA(q) ile gösterilir ve aşağıdaki şekilde yazılır:

$$w_t = \mu + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q} \quad (2)$$

Zaman serileri çoğu zaman AR veya MA süreçleri ile gösterilemeyebilir. Bu durumda AR+MA=ARMA olarak tanımlanan AR ve MA modeli ile süreç gösterilir. Bu model Wold (1954) tarafından geliştirilmiştir. Aynı zamanda bu tanımlanan 3 süreç serinin durağan olduğu varsayımına dayanır. Genel olarak ARMA(p,q) şeklinde ifade edilen süreç iki modelin toplamından oluşur ve aşağıdaki şekilde gösterilir:

$$w_t = m + \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad (3)$$

AR(p), MA(q) ve ARMA(p,q) süreçlerindeki temel varsayım serinin durağan olduğu şeklindedir ancak bazı durumlarda serinin durağan olmadığı zamanlar olabilir. Bir zaman serisinin durağan olması zaman içinde serideki değişimlerin ve ortalamasının aynı olmasıdır. Zaman aralıklarındaki değişim zamana bağlı değildir. Bu birlikte değişimler değişkenler arasındaki gecikmelere bağlıdır (Gujarati,1995). Eğer süreç zamana göre değişiyorsa dizi durağan değildir denir. Durağan serileri modellerle ifade etmek mümkündür. Ancak durağan olmayan serileri basit matematiksel modellerle ifade etmek mümkün değildir. Zaman serilerinin durağanlığını etkileyen nedenlere rağmen genelde serilerin homojen olduğu görülür (Box ve Jenkins, 1976). Bu durağan olmayan seriyi durağan hale getirebilmek için uygun derecede fark(d) alma işlemi yapılır. Bu durumda ARIMA(p,q,d) süreci kullanılmalıdır. Yöntem AR ve MA fonksiyonlarına fark alma işleminin eklenmesi ile aşağıdaki gibi gösterilir (Box ve Jenkins, 1976);

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (4)$$

Bu denklemde $Z_t, Z_{t-1}, \dots, Z_{t-p}$ farkı alınmış gözlem değerlerini, $\phi_1, \phi_2, \dots, \phi_p$ gözlem değerlerinin katsayılarını, δ sabit değerini, $a_t, a_{t-1}, \dots, a_{t-q}$ hatayı ve $\theta_1, \theta_2, \dots, \theta_q$ hata terimlerinin katsayılarını göstermektedir.

2.2.Yapay Sinir Ağları

YSA birbirlerinden ayrı olarak tasarlanmış birden fazla yapay sinir hücresinin (nöron) birbirleriyle paralel bağlanması sonucu oluşan bir yapıdır (Rojas R. 2013). Bir YSA hücresinde girdiler ($X_1, X_2, X_3 \dots$), ağırlıklar ($W_1, W_2, W_3 \dots$), birleştirme fonksiyonu, aktivasyon Fonksiyonu ve çıktı (Y) olmak üzere 5 bölüm vardır. Girdiler rasgele olan veya başta belirlenmiş ağırlık değerleri ile çarpılarak birleştirme fonksiyonu aşamasına gelirler. Burada hangi birleştirme metodunun kullanılacağı belirlenmiştir. En çok kullanılan birleştirme metodları toplama ve çarpım olarak belirlenen metodlardır. Birleştirme fonksiyonu olarak toplama işleminin kullanılması aşağıdaki denklemle gösterilebilir:

$$NET = \sum_{i=1}^n X_i \cdot W_i \quad (5)$$

Bu formülde 1. girdiden n. girdiye kadar olan girdilerin her birinin kendilerine verilmiş rasgele ağırlık değerleri ile çarpılması ve toplanması anlatılmaktadır. X değerleri her bir girdiyi W değerleri ise her bir girdinin kendi ağırlıklarıdır. Birleştirme fonksiyonu olarak toplama işleminin kullanılması aşağıdaki denklemle gösterilebilir:

$$NET = \prod_{i=1}^n X_i \cdot W_i \quad (6)$$

Bu formülde ise 1. girdiden n. girdiye kadar olan girdilerin her birinin kendilerine rasgele verilmiş ağırlık değerleri ile çarpılması ve sonra elde edilen değerlerin birbirleri ile çarpılması ile NET girdi elde edilmiş olur. Hangi yöntem kullanılırsa kullanılsın birleştirme fonksiyonundan sonra elde edilen (NET) değeri aktivasyon fonksiyonuna alınır. Burada da en çok kullanılan fonksiyonlar Sigmoid ve Hiperbolik Tanjant (tansig) fonksiyonlarıdır.

Sigmoid fonksiyonu aşağıdaki gibi gösterilebilir.

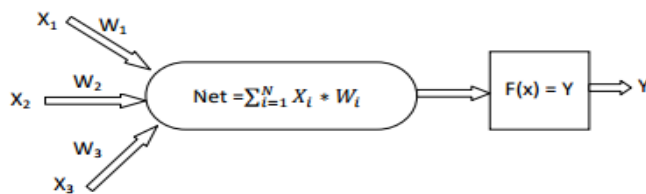
$$F(Net) = \frac{1}{1+e^{-Net}} \quad (7)$$

YSA'da kullanılacak bir aktivasyon fonksiyonu Hiperbolik Tanjant fonksiyonudur. Bu fonksiyonu kullanabilmek için girdi değerleri önce [-1,1] aralığında normalize edilir ve çıktı değerleri de [-1,1] aralığında elde edilir. Çıktısının [-1,1] aralığında değerler olmasının arzu edildiği uygulamalar için bu fonksiyon çok kullanışlıdır (Hamzaçebi,2011).

$$F(Net) = \frac{e^{Net} - e^{-Net}}{e^{Net} + e^{-Net}} \quad (8)$$

Genel olarak bir Yapay Sinir ağı hücresinin yapısı aşağıdaki gibidir.

Şekil 1. Yapay Sinir Hücresinin Yapısı



En son aşamada ağın ürettiği çıktı elde edilir. Elde edilen çıktının gerçek çıktıya uygun olup olmadığının belirlenmesinde bazı hata hesaplama yöntemlerine bakılabilir. Bunun için en çok tercih edilenler ise Ortalama Hata Kareler (Mean Square Error-MSE), Ortalama Hata Karelerinin Karekökü (Root Mean Square Error-RMSE) vb. değerlerdir.

Ortalama Hata Kareleri (Mean Square Error-MSE); ağın her bir çıktısı ile her bir gerçek çıktının farklarının karelerinin toplamının ortalaması alınarak bulunur

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2 \quad (9)$$

Ortalama Hata Kareleri Karakökü (Root Mean Square Error-RMSE); ağın her bir çıktısı ile her bir gerçek çıktının farklarının karelerinin toplamının ortalamasının karakökü alınarak bulunur

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \quad (10)$$

2.3. Melez Modeller

Melez yöntemler finansal zaman serisi tahmin işinde büyük başarılar sağlamış durumdadır. İstatistik ve makine öğrenmesinde bu metodlar; kurucu öğrenme algoritmalarından elde edilenden daha iyi tahmin performansı elde etmek için çoklu öğrenme algoritmaları kullanılır (Opitz d at all., 1999; Rokach L., 2010). Eğitilebilir ve tahmin amacıyla kullanılabilir olduklarından bu metodlar birer denetimli öğrenme algoritmalarıdır.

Deneyisel olarak melezleme, modeller arasında önemli bir çeşitlilik olduğunda daha iyi sonuçlar elde etme eğilimindedir (Kuncheva L. at all, 2003). Bu modellerin amacı, modellerin birleşimiyle ortaya çıkacak çeşitliliği teşvik etmek ve daha iyi sonuçlar elde edilmesini sağlamaktır (Adeva J. at all., 2005).

Melezleme genel sistem performansını artırmak için birlikte çalışan öğrenme algoritmalarının uygulandığı örneklerin kümesidir. Birleştirmeler genellikle çoklu iki sınıflı çok sayıda problemin çok sınıflı bir probleme dönüşümü için kullanılır (Adeva J. at all, 2005). Melezlenmiş tahmin modelleri tahmin etme görevinin en başarılı modelleri olarak tanınırlar. Melez modeller hakkındaki daha önceden yapılmış çalışmalar bu başarının nedenini gruplar arasındaki çeşitlilik olduğunu göstermiştir. Çeşitlilik oluşturmak için değişik yöntemler mevcuttur (Oliveira M., & Torgo L. 2014). Melez modellerle elde edilen sonuçlar ile modellerin tekil halde kullanılmaları sonucu elde edilen sonuçlar karşılaştırıldığında birbirinden ilişkisiz bile olsalar melezleme yapmanın genel varyans ve ya hatayı düşürebildiği görülmüştür (Khashei, M., ve diğerleri, 2008).

Ortaya atılan her bir modelde bir hata vardır. Gerçek değer ile model sonucu elde edilen tahmin değeri arasındaki fark hata olarak adlandırılır ve bu hatanın belirlenmesinde birkaç farklı yöntem kullanılabilir. Bunlardan biri toplam bir diğeri çarpım modelidir. Bu metodlar aşağıdaki gösterilebilir;

Toplama metoduna göre gerçek veriler (r_t) hata terimleri (ε_t) ile tahmin edilen doğrusal verilerin (\hat{r}_t) toplamına eşittir.

$$r_t = \hat{r}_t + \varepsilon_t \quad (11)$$

Buna göre hata terimlerini belirlemek için;

$$\varepsilon_t = r_t - \hat{r}_t \quad (12)$$

işlemini yapmak gerekmektedir. Çarpma metodunda ise gerçek veriler (r_t) hata terimleri (ε_t) ile tahmin edilen doğrusal verilerin (\hat{r}_t) çarpımına eşittir.

$$r_t = \hat{r}_t * \varepsilon_t \quad (13)$$

Buna göre hata terimlerini belirlemek için;

$$\varepsilon_t = \frac{r_t}{\hat{r}_t} \quad (14)$$

işlemleri yapılmalıdır. Melez modellerin doğrusal kısımlarındaki hata bu metodlardan biriyle veya her ikisiyle belirlendikten sonra bu hata değerleri doğrusal olmayan kısmı gösterecektir. Daha sonra yapılan melezleme işlemi ile bu doğrusal olmayan yapı açıklanmaya çalışılır.

$$y_t = L_t + N_t \quad (15)$$

Burada y_t gözlem değerlerini t zamanı L_t ve N_t ise doğrusal ve doğrusal olmayan bileşenleri göstermektedir. Gözlem değerleri doğrusal ve doğrusal olmayan bileşenlerin toplamından oluşur.

ARIMA-YSA melez çalışmalarında öncelikle ARIMA ile doğrusal olan bileşen t zamanında tahmin edilir ve bu \hat{L}_t olarak alınır. ARIMA işlemi bittikten sonra elde edilen artıklar sadece doğrusal olmayan bir bileşen içerir ve bu aşamadan sonra başka yöntemlerle işleme devam edilmesi gerekir. YSA aşamasına gelindiğinde P sayıda girdi kullanılarak artıklar için bir YSA tasarlanmıştır.

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-p}) + \varepsilon_t \quad (16)$$

Bu gösterimde f doğrusal olmayan bir fonksiyonu göstermektedir. Artıklar arasında doğrusal olmayan fonksiyon belirlendiğinde ε_t bu fonksiyon sonunda ortaya çıkan artığı göstermektedir. Eğer doğrusal olmayan bileşende YSA ile tahmin edilirse o zaman model;

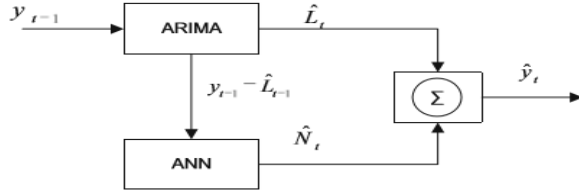
$$y_t = \hat{L}_t + \hat{N}_t \quad (17)$$

şekline dönüşür ve bu melez tahmin modelini verir.

Zhang (2003)'de geliştirdiği melez modelinde; ne ARIMA ne de YSA'nın tüm gerçek zaman serilerine uygun olmadığını çünkü bu serilerde gözlemler arasında doğrusal ve doğrusal olmayan korelasyon yapıları bulunduğunu ve bu nedenle bir zaman

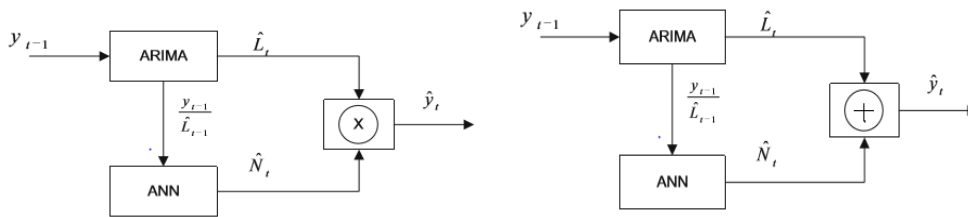
serisinin hem doğrusal hem doğrusal olmayan bileşenlerinin tahmini için bir melez model kullanılması gerektiğini göstermiştir. Model aşağıdaki gibi gösterilebilir.

Şekil 2: Zhang Melez Modeli



Zhang (2003) kurduğu melez modelinde sadece toplama metodu ile elde edilen hatalar için melez model tasarlamıştır. (Wang vd., 2013) çalışmalarında zaman serisinin analizi için iki farklı model düşünmüşlerdir. Hatanın tespit edilmesi için hem toplama hem de çarpma metodunu kullanmışlardır.

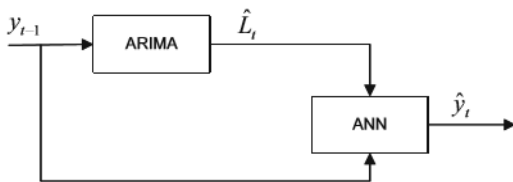
Şekil 3: Wang Melez Modelleri



Khashei ve Bijari (2010) yaptıkları çalışmada Zhang'ın melez modeline benzer bir model önermişlerdir. Ancak doğrusal olan kısım ARIMA ile işlendikten sonra elde edilen hatalar tek başına YSA ile işlenmemiş; hem gerçek değerler hem hata değerleri geliştirilen çok katmanlı YSA'ya girdi değerleri olarak verilmiştir. Çalışma sonucunda elde edilen değerlerin Zhang'ın 2003'teki modelinden daha iyi olduğunu iddia etmiştir.

Khashei ve Bijari (2011) bir ARIMA-YSA melez modeli daha önermiştir. Bir önceki çalışmalarından farklı olarak YSA girdileri gerçek gözlem değerleri, ARIMA'nın tahmini değerleri ve hata değerlerinden oluşmaktadır. Elde edilen tahmin değeri üzerinde herhangi bir işlem yapmadan tahmin değerlerini verir (Khashei, Bijari, 2011).

Şekil 4: Khashei ve Bijari (2011) Melez Modeli



3. VERİ VE YÖNTEM

Veri olarak Ocak 2012'den Haziran 2017 sonuna kadar olan günlük NASDAQ borsası endeks değerleri öngörü amaçlı kullanılmıştır. Veriler diğer değişkenlerde dâhil edilerek işleme alınmadan önce tarihsel olarak düzenlenmiş her bir değişkenin aynı tarihi gösteren değerleri bulunmuştur. Aynı tarihte ölçüm değeri olmayan günler işleminden çıkartılmıştır. Verilerin üzerinde başka herhangi bir hesaplama veya dönüştürme işlemi yapılmamıştır. Modellerde farklı inputlar ve 2 farklı output kullanılmıştır. Melez model için eldeki mevcut veriler ilk aşamada ARIMA modeli ile analiz edilmiş ve artık miktarları belirlenmiştir. Elde edilen artıklar Yapay Sinir Ağında farklı inputlar kullanılarak tahmin edilmiştir. Kullanılan YSA kodları Matlab Function kullanılarak yazılmış ve eğitim aşaması ile en iyi model bulunmuştur. Modeller kurulurken en iyi olduğu düşünülen sonuçlar üretilene kadar yinelemeler yapılmıştır. Farklı gizli nöron sayıları, farklı aktivasyon ve öğrenme fonksiyonları kullanılarak ağ sürekli eğitilmiştir. Hangi fonksiyonun kullanılacağı deneme yanılma yolu ile belirlenmiştir ve içlerinden tahmin için en iyi olan seçilmiştir. Eğitim sonunda MSE değerleri ile ağın performansına bakılmıştır. Levenberg-Marquardt en iyi öğrenme fonksiyonu olarak bulunmuş aktivasyon fonksiyonu olarak 'tansig' kullanılmıştır. Uygulama verilerinin %70'i eğitim, %30'u doğrulama olarak alınmıştır. Verilerin bu gruplara seçiminde rasgelelik kullanılmıştır. NASDAQ Borsası değeri için 6 farklı model kurulmuştur. Bu modeller aşağıdaki gibidir:

- 1) Saf ARIMA

- 2) Temel göstergeler ile NASDAQ borsa değerinin tahmini için YSA
- 3) Temel göstergeler ile kalıntı tahmini için YSA
- 4) Kalıntı değerleri ile kalıntı tahmini için YSA
- 5) Temel göstergeler ve kalıntı değerleri ile kalıntı tahmini için YSA
- 6) Temel göstergeler ve kalıntı değerleri ile NASDAQ borsa değerinin tahmini için YSA

1. ve 2. modeller için denklem (9) kullanılarak MSE değerleri hesaplandıktan sonra kalıntıların tahminlerinin yapıldığı 3., 4. ve 5. modeller kurulmuştur. ARIMA'dan elde edilen tahmin değerleri ile YSA kurulan artık veri tahminlerinin kullanılması ile melez modeller oluşturulmuştur. Kalıntılar ilk olarak toplama ve sonrasında çarpma metoduyla elde edilmiştir. Sonuçlar karşılaştırılmıştır.

4. BULGULAR

1. Model : NASDAQ borsası verileri kullanılarak 03.01.2012 ile 30.06.2017 tarihleri arası günlük kapanış verilerine göre uygun ARIMA modeli kurulmuştur. Durağanlık şartı modelin oluşturulmasında önemli bir kuraldır. Seri incelendiğinde zaman içinde otokorelasyonu ve ortalaması sabit olmadığı görülmüştür. Seri durağan olmadığı için öncelikle 1 günlük farkları alınarak trend bileşeni kaldırılarak seri durağan hale getirilmiştir. Çözümleme Eviews 9 programında yapılmıştır. Sonuçta ARMA(3,3) borsa değeri için uygun sonuç vermiştir.

ARMA(3,3) için hesaplanan F istatistiği (Prob(F-Statistic))=0,0273 olarak bulunmuştur. Bu değer 0,05'den küçük olması modelin bir bütün olarak anlamlı olduğunu göstermektedir. Elde edilen model;

$$\text{Nasdaq} = \text{Nasdaq}_t - \text{Nasdaq}_{t-1}$$

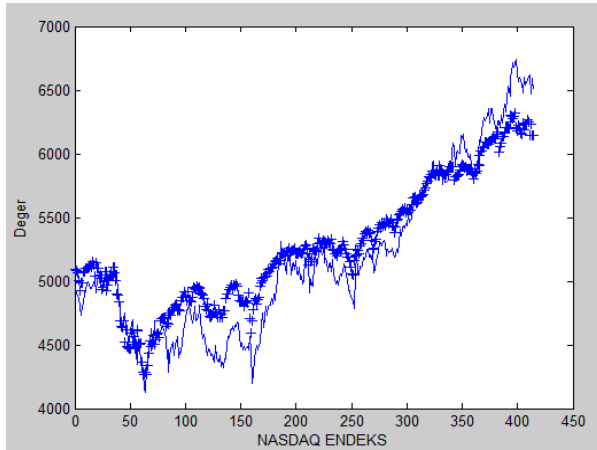
$$\text{Nasdaq} = 2,493 + 0,278\text{Nasdaq}_{t-1} - 0,308\text{Nasdaq}_{t-2} + 0,902\text{Nasdaq}_{t-3} - 0,272U_{t-1} + 0,280U_{t-2} - 0,933U_{t-3}$$

şeklinde gösterilebilir.

Elde edilen tahmin değerleri gerçek değerlerle karşılaştırılmış ve denklem (12) ve denklem (14) kullanılarak hem toplama yöntemi hem de çarpma yöntemi ile artıklar belirlenmiştir. Bu artık değerleri diğer aşamalarda YSA için farklı modellerde input ve output olarak kullanılacaktır. Tahmin değerleri ise melez modelimizin doğrusal kısmını oluşturacaktır. ARIMA tahminine göre MSE değeri toplam fonksiyonu ile hesaplanınca 62883,8; çarpım fonksiyonu ile hesaplanınca ise 59843,6 olarak bulunmuştur.

2. Model Uygulama verilerinin %70'i eğitim %30'u doğrulama olarak alınmıştır. Verilerin bu gruplara seçiminde rasgelelik kullanılmıştır. NASDAQ borsası için ARIMA'da kullanılan aynı tarih aralığı kullanılmıştır. Burada ilk olarak kurulacak ağ için NASDAQ endeks değerleri YSA'nın output değeri olacaktır. Input değerleri ise 1 gün gecikmeli temel göstergelerdir. Böylece 1 bağımlı 10 bağımsız değişkenli bir YSA tasarlanmıştır. En iyi YSA sonucu 10 gizli nöron %70 eğitim oranında aşağıdaki gibi bulunmuştur.

3.Şekil 5: NASDAQ Borsası İçin Endeks Tahmini

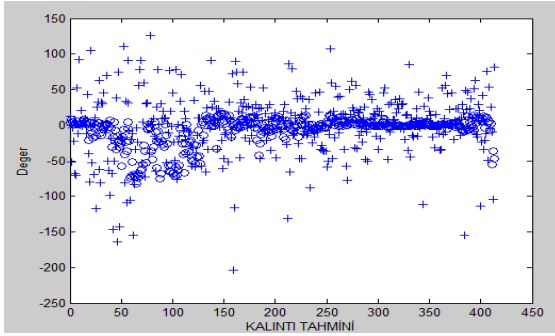


MSE : 47463.7

+++ : Gerçek NASDAQ Değerleri

----- : Tahmini NASDAQ Değerleri

4. Model: ARIMA sonucunda elde edilen tahmin değerleri ve gerçek değerler kullanılarak artıklar toplam ve çarpım metodları ile ayrı şekilde belirlenmiştir. Elde edilen artık değerleri temel göstergeler kullanılarak tahmin edilmiştir. Yine her iki durumda da artık değerleri YSA'nın output değeri olacaktır. Input değerleri ise 1 ve 2 gün gecikmeli temel değerlerden oluşmaktadır. Burada artıkların tahmin edilmesinde sadece temel göstergeler kullanılmıştır. Böylece 1 bağımlı 8 bağımsız değişkenli bir YSA tasarlanmıştır. Sonuçlar aşağıdaki gibi bulunmuştur.

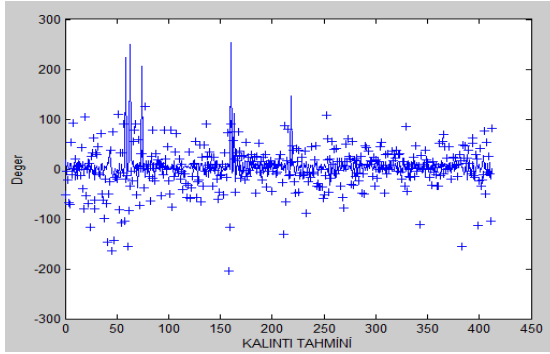
Şekil 6: Toplam Fonksiyonu için Kalıntı Tahmini

MSE : 2307,8

+++ : Gerçek Kalıntı Değerleri

ooo : Tahmini Kalıntı Değerleri

5. Model: Kurulan bu modelde hesaplanan artık değerleri YSA'nın output değeri olacaktır. Input değerleri bu artık değerlerinin 1, 2, 3, 4, 5 ve 6 günlük gecikmeli değerleridir. Burada artıkların tahmin edilmesinde sadece artıkların gecikmeli değerleri kullanılmıştır. Böylece 1 bağımlı 6 bağımsız değişkenli bir YSA tasarlanmıştır. Sonuçlar aşağıdaki gibi bulunmuştur.

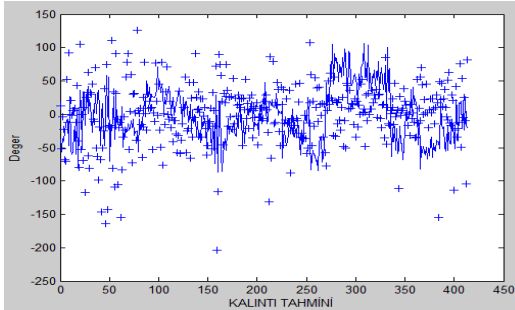
Şekil 8. Toplama Fonksiyonu için Kalıntı Tahmini

MSE : 2757,6

+++ : Gerçek Kalıntı Değerleri

---- : Tahmini Kalıntı Değerleri

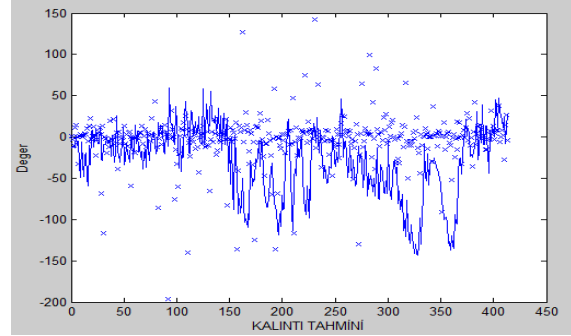
6. Model: Bu modelde ARIMA ile elde edilen artık değerleri YSA'nın output değeri olacaktır. Input değerleri ise temel göstergeler ve 1-2 gün gecikmeli artık değerleridir. Yani ARIMA'dan elde edilen artık değerleri YSA'ya hem input hem output olarak verilmiştir. Böylece 1 bağımlı 10 bağımsız değişkenli bir YSA tasarlanmıştır.

Şekil 10: Toplama Fonksiyonu için Kalıntı Tahmini

MSE : 2008,4

+++ : Gerçek Kalıntı Değerleri

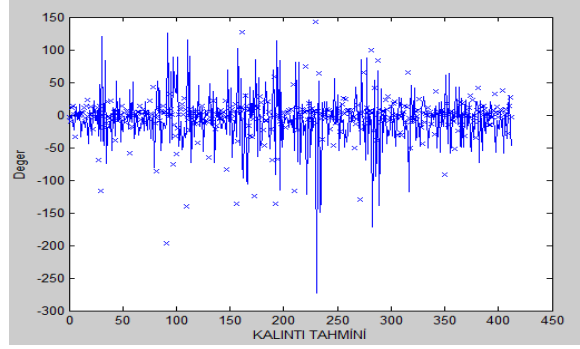
---- : Tahmini Kalıntı Değerleri

Şekil 7: Çarpma Fonksiyonu için Kalıntı Tahmini

MSE : 3389,4

+++ : Gerçek Kalıntı Değerleri

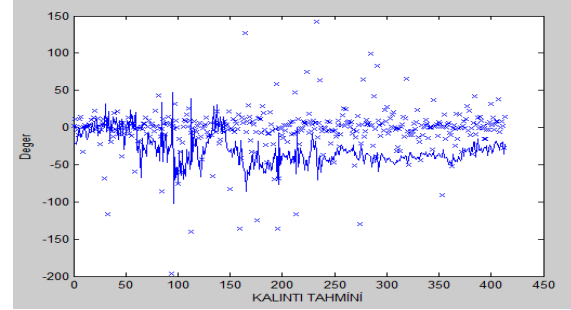
---- : Tahmini Kalıntı Değerleri

Şekil 9. Çarpma Fonksiyonu Sonucu için Kalıntı Tahmini

MSE : 2172,3

+++ : Gerçek Kalıntı Değerleri

---- : Tahmini Kalıntı Değerleri

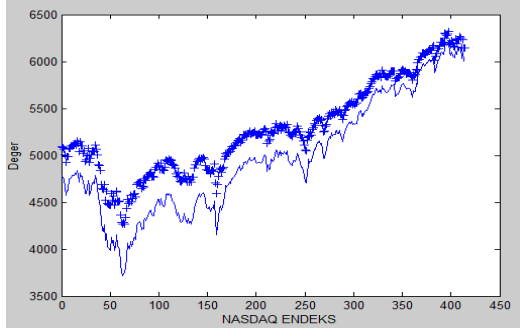
Şekil 11. Çarpma Fonksiyonu için Kalıntı TahminiMSE :
3382,
2

+++ : Gerçek Kalıntı Değerleri

---- : Tahmini Kalıntı Değerleri

7. Model: NASDAQ endeks değerleri YSA'nın output değeri olacaktır. Input değerleri ise temel göstergeler ve 1-2 gün gecikmeli kalıntı değerleridir. Böylece 1 bağımlı 12 bağımsız değişkenli bir YSA tasarlanmıştır. Bu ağ ile NASDAQ borsa endeks değeri hem temel göstergeler hem de ARIMA artıkları ile tahmin edilmeye çalışılmıştır.

Şekil 12: Toplama Fonksiyonu Artıkları için NASDAQ

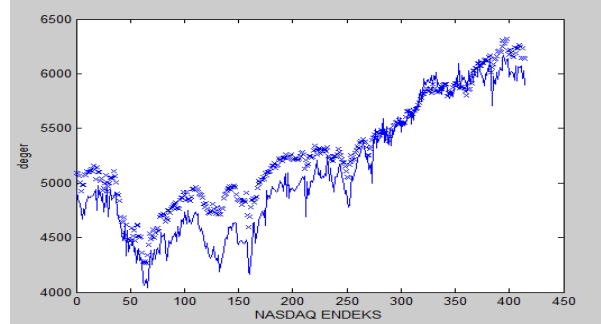


MSE : 57560,1

+++ : Gerçek Kalıntı Değerleri

----- : Tahmini Kalıntı Değerleri

Şekil 13. Çarpma Fonksiyonu Artıkları için NASDAQ



MSE : 52318,3

+++ : Gerçek Kalıntı Değerleri

----- : Tahmini Kalıntı Değerleri

6 farklı model kurulma işlemleri tamamlandıktan sonra elde edilen sonuçları aşağıdaki gibidir. Tüm hesaplanan değerler %30 doğrulama kısmında yer alan 414 günlük değerler içindir.

Tablo 1: Modellerin Toplama ve Çarpma Fonksiyonları için MSE Değerleri

MSE DEĞERLERİ	1. Model	2. Model	3. Model	4. Model	5. Model	6. Model
Toplama M.	82883,8	47643.7	2307.8	2757.6	3382.2	57560.1
Çarpma M.	79843,6	47643.7	3389.4	2172.3	2008.4	52313.8

1, 2 ve 6. modellerde NASDAQ borsa endeksi tahmin edilmeye 3, 4 ve 5. modellerde endeks için hesaplanan hata terimleri tahmin edilmeye çalışılmıştır. Melezleme modellerinde modelin doğrusal olmayan kısmını 3, 4 ve 5. modellerden elde edilen artık tahmin değerleri oluşturacakken; modelin doğrusal olan kısmını 1.model olan ARIMA'dan elde edilen tahmin sonuçları oluşturacaktır. 2. ve 6. modeller karşılaştırma yapmak için hesaplanmıştır. Denklem (17) kullanılarak yapılan melezleme sonuçları aşağıdaki gibi elde edilmiştir.

3. modelden elde edilen artık tahmin değerlerine ARIMA ile tahmin edilen değerler eklenince ortaya çıkan melez modelde MSE değerleri aşağıdaki gibi bulunmuştur:

Tablo 2: ARIMA tahminleri ile 3.Model YSA Melezleme Sonuçları

MSE DEĞERLERİ	2. Model
Toplama M.	43954.1
Çarpma M.	44212.3

4. modelden elde edilen artık tahmin değerlerine ARIMA ile tahmin edilen değerleri eklenince ortaya çıkan melez modelde MSE değerleri aşağıdaki gibi bulunmuştur:

Tablo 3: ARIMA tahminleri ile 4.Model YSA Melezleme Sonuçları

MSE DEĞERLERİ	6. Model
Toplama M.	45128,4
Çarpma M.	45237.2

5. modelden elde edilen artık tahmin değerlerine ARIMA ile tahmin edilen değerleri eklenince ortaya çıkan melez modelde MSE değerleri aşağıdaki gibi bulunmuştur:

Tablo 4: ARIMA tahminleri ile 5.Model YSA Melezleme Sonuçları

MSE DEĞERLERİ	6. Model
Toplama M.	51487,1
Çarpma M.	46337.2

5. SONUÇ

Bu arařtırmada modellerin birbirleri ile karřılařtırması yapıldığında temel göstergelerin kullanıldığı NASDAQ endeks tahmininin ARIMA yöntemi ile melezlenmesi sonucunda hem toplama hem çarpma fonksiyonu için işlemler yapılmıştır. Toplama modeli ile hesaplanan artıkların YSA melezinde daha iyi sonuç verdiği görülmüştür. Hangi deęişkenler kullanılarak artıklar hesaplanırsa hesaplanırsın melez modelin daha iyi tahmin sonucu verdiği açıkça görülecektir. Saf ARIMA deęerlerinde hesaplanan MSE, artıkların hesaplandığı her YSA ile melezde daha iyi sonuçlar vermiştir. Dahası kurulan melez model sadece YSA'nın kullanıldığı 2. ve 6. model sonuçlarından da daha iyi sonuçlar vermiştir. Melezlemede toplama ve çarpma fonksiyonları için net bir karar verilememiştir. Bazı modellerde toplama bazı modellerde ise çarpma fonksiyonu daha başarılı sonuç vermiştir.

2. modelde kurulacak aę için NASDAQ endeks deęerleri YSA'nın output deęeri olmuş input deęerleri ise 1 gün gecikmeli temel göstergeler olarak alınmıştır. Bu sonuç 6. modelde yer alan inputlardan daha iyi bir sonuç vermiştir.

Melez modellerle tekil modeller karřılařtırıldığında melez modellerin her zaman daha iyi sonuç verdiği daha iyi tahmin performansı olduęu görülmüştür. İlerleyen çalışmalarda YSA'nın tahmin performansını daha da iyileřtirmek için birden fazla modelle melezleme yapılmaya çalışılacaktır.

KAYNAKÇA

Aghababaeian, R. et al. (2011). Forecasting the Tehran Stock Market by artificial neural network. International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence.

Yakut, E., Elmas, B., & Yavuz, S. (2014). Yapay Sinir Aęları ve Destek Vektör Makineleri. Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 19(1).

Box, G. E. P. ve Jenkins, G. M., Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, 1976.

Fernández, A. and S. Gómez (2007). Portfolio selection using neural networks. Computers & Operations Research, 34(4): 1177-1191.

Gershenfeld, N. A. (1999). The nature of mathematical modeling. Cambridge university press.

Gujarati, D. N. (1995). Basic Econometrics 3rd edition, New York: Me Graw Hill. Gupta KL (1970), 'Personal savings in developing nations. Further evidence', the economic record, 46, 243-249.

Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. Expert Systems with Applications, 38(8), 10389-10397.

Hamzaçebi, c. (2011). Yapay Sinir Aęları:tahmin amaçlı kullanımı MATLAB ve Neurosolutions uygulamalı. Ekin Basım Yayın Daęıtım, 2011

Khashei, M., Hejazi, S.R., Bijari, M., A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, Fuzzy Sets and Systems 159, 769 – 786, 2008.

Khashei, M., Bijari, M., A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, Applied Soft Computing 11, 2664–2675, 2011

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, 37(1), 479-489.

Kuncheva, L. and Whitaker, C., Measures of diversity in classifier ensembles, Machine Learning, 51, pp. 181-207, 2003

J. J. García Adeva, Ulises Cerviño, and R. Calvo, CLEI Journal, Vol. 8, No. 2, pp. 1 - 12, December 2005.

Lawrence, R. (1997). Using neural networks to forecast stock market prices. University of Manitoba.

Oliveira, M., & Torgo, L. (2014). Ensembles for Time Series Forecasting. In Proceedings of the Sixth Asian Conference on Machine Learning (pp. 360-370).

Opitz, D.; Maclin, R. (1999). "Popular ensemble methods: An empirical study". Journal of Artificial Intelligence Research 11: 169–198.

Pai, P. F., & Lin, C. S. (2004). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33(6), 497-505.

Rokach, L. (2010). "Ensemble-based classifiers". Artificial Intelligence Review 33 (1-2)

Rojas, R. (2013). *Neural networks: a systematic introduction*. Springer Science & Business Media.

Slutsky, E.,(1937) "The Summation of Random Causes As The Source of Cyclic Processes".

Sutheebanjard, P., & Premchaiswadi, W. (2010). Forecasting The Thailand Stock Market Using Evolution Strategies. *Asian Academy of Management Journal of Accounting & Finance*, 6(2).

Yule, G.U., "On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wölfer's Sunspot Numbers", *Phil. Trans.*, A226, 267, 1927.

Wang, L., Zou, H., Su, J., Li, L., & Chaudhry, S. (2013). An ARIMA-ANN hybrid model for time series forecasting. *Systems Research and Behavioral Science*, 30(3), 244-259.

Wold, H.O., *A Study in The Analysis of Stationary Time Series*, Almquist and Wicksell, Uppsala, 1954.

Yudong, Z., & Lenan, W. (2009). Stock market prediction of S& P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications*, 36, 8849–8854

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.