

İKİNCİ EL OTOMOBİLLERİN GELECEKTEKİ FİYATLARININ YAPAY SİNİR AĞLARI İLE TAHMİN EDİLMESİ

FORECASTING THE FUTURE PRICES OF THE SECOND-HAND AUTOMOBILES USING ARTIFICIAL NEURAL NETWORKS

Dr.Özcan ASILKAN¹
Arş.Gör.Sezgin IRMAK²

ÖZET

Günümüzde ikinci el otomobiller dünyanın hemen her ülkesinde tüketicilerin en çok ilgi gösterdiği dayanıklı ürünler arasında yer almaktadır. Bu ürünlerin pazardaki güncel fiyatları ve gelecekteki fiyatları her zaman merak konusu olmuştur. Bu çalışmada son yıllarda sıkça kullanılan bir tahmin yöntemi haline gelen yapay sinir ağları kullanılarak ikinci el otomobillerin gelecekteki fiyatları tahmin edilmeye çalışılmıştır. Girdi olarak Avrupa kökenli çok sayıda ikinci el otomobil web sitesinden elde edilmiş olan ilan verileri kullanılmıştır. Yapay sinir ağları uygulaması ile bulunan sonuçlar, zaman serisi analizleri ile bulunmuş olan sonuçlarla karşılaştırılmıştır. Elde edilen sonuçlar, yapay sinir ağlarının ikinci el otomobillerin gelecekteki fiyatlarını tahmin etmede başarıyla kullanılabileceğini ortaya koymuştur.

ABSTRACT

Today, second-hand cars take place among the most attractive durable goods in almost all countries of the world. Current and future market prices of these goods have always been wondered by the consumers. This study has tried to forecast the future prices of the second-hand automobiles using Artificial Neural Networks (ANN) which became a very popular forecast technique in the recent years. The advertisement data gathered from Europe-based second-hand automobiles web sites have been used as the input of the study. The results obtained by Artificial Neural Networks have been compared to those obtained by some Time Series forecasting methods. As a conclusion, it has been shown that Artificial Neural Networks can be

¹ Akdeniz Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Antalya, Türkiye, oasilkan@akdeniz.edu.tr.

² Akdeniz Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Antalya, Türkiye, sezgin@akdeniz.edu.tr.

used successfully to forecast the future prices of the second-hand automobiles.

Yapay Sinir Ağları, İkinci El Otomobil, Fiyat Tahmini
Artificial Neural Networks, Second Hand Automobile, Forecasting Price

1. GİRİŞ

Otomotiv endüstrisi, sanayileşmiş ülkelerin hemen hepsinde ekonominin lokomotifi konumundadır. Bu endüstrinin dünya genelindeki toplam motorlu araç üretiminin yaklaşık % 70'ini de otomobillerin oluşturduğu hesaba katıldığında, otomobil sektörünün dünya ekonomisindeki baskın önemi açıkça anlaşılacaktır (Onat, 2007: 3).

Öte yandan, pek çok ülkede ikinci el otomobil satış hacmi, yeni otomobil satış hacminden çok daha büyük hale gelmiştir. İkinci el otomobil pazarındaki en önemli unsur ise hiç şüphesiz ikinci el otomobillerin fiyatlarında ortaya çıkan farklılıklar ya da değişimlerdir. İkinci el otomobillerin fiyatlarının yeni otomobil fiyatlarına göre çok daha değişken olması, tahmin edilmelerini de zorlaştırmaktadır. Bu nedenle, ikinci el otomobillerin fiyatlarını tahmin edebilmek için her şeyden önce uzun süreli ve nitelikli verilerin toplanması gerekmektedir.

Yapay Sinir Ağları (YSA) son yıllarda tahmin amacıyla hemen her alanda yaygın olarak kullanılmaktadır. Özellikle doğrusal olmayan zaman serilerinde YSA'nın geleneksel yöntemlerden daha üstün başarı göstermesi bu yöntemin tercih edilmesinde önemli bir faktör olmuştur (Zhang vd., 1998: 35-62). Literatürde bu konuda değişik alanlarda yapılmış çok sayıda çalışmaya rastlamak mümkündür. Ancak ne yazık ki otomobil sektöründeki durum hiç de iç açıcı değildir. Bu sektörde rastlanmıştır olan en kayda değer çalışmada İşeri ve Karlık (2008), YSA'nı kullanarak yeni otomobillerin özellikleri ile fiyatları arasındaki ilişkiyi ortaya çıkarıp fiyatlarını tahmin etmeye çalışmışlardır. Diğer yandan ikinci el otomobil fiyatlarının tahminine yönelik kayda değer bir çalışmaya rastlanamamıştır. Otomobil sektöründeki bu eksikliği azaltmaya yönelik katkı sağlayabilmek de bu çalışmanın hedefleri arasında yer almaktadır.

Çalışmada, YSA yardımıyla ikinci el otomobillerin 2005-2007 yılları arasındaki verileri kullanılarak 2008-2009 yılları arasındaki fiyatları tahmin edilmeye çalışılmıştır. Aynı yıllara ait veriler kullanılarak, zaman serileri analiz yöntemleri yardımıyla da tahminler yapılmış ve elde edilen sonuçlar karşılaştırılmıştır.

2. İKİNCİ EL OTOMOBİL PAZARI

Daha önce de belirtildiği gibi, dünya otomotiv endüstrisindeki toplam motorlu araç üretiminin yaklaşık % 70'ini otomobiller oluşturmaktadır. 2006 yılında Avrupa' da üretilen tüm motorlu araçların % 87' sini de yine otomobiller oluşturmuştur (Web-ACEA, 2007). Bu verilerden hareketle, otomobil sektörü için "otomotiv endüstrisinin temel yapı taşı" nitelemesi yapılabilir. Otomobil üretimi, güçlü yan sanayisi ile birlikte diğer taşıtların üretimine de destek olmaktadır (Onat, 2007; Web-Kobifinans, 2008).

İkinci el otomobil satış hacminin pek çok ülkede yeni otomobil satış hacminden çok daha büyük hale gelmiş olması, ikinci el araç pazarlarının artan önemini ortaya koymaktadır. Örneğin ABD ekonomisinde en büyük perakende sektörünü oluşturan ikinci el otomobil pazarının hacmi, yeni otomobil pazar hacminin iki katından daha fazladır ve sürekli olarak büyümeye devam etmektedir (Lee, 2006: 120). 2005 yılında ABD' de satılan ikinci el otomobil sayısı 44 milyon, satış değeri ise 370 milyar dolardır (Web-Wikipedia, 2008).

Avrupa ikinci el otomobil pazarı 2003-2006 yılları arasında yavaş fakat istikrarlı bir büyüme göstermiştir. Bu gidişatın sonraki beş yılda da devam edeceği beklenmektedir. Pazar değeri, 2003-2006 yılları arasındaki yaklaşık % 2 civarındaki bileşik yıllık büyüme oranıyla 2006 sonunda 263 milyar dolar değerine ulaşmıştır. 2007 – 2012 arasında % 11,3 büyüyerek 298,5 milyar dolar değerine ulaşacağı tahmin edilmektedir. Pazar hacmi, 2003-2006 yılları arasındaki yaklaşık % 1,3 civarındaki bileşik yıllık artış oranıyla 2006 sonunda yaklaşık 28,3 milyon birim hacme ulaşmıştır. 2007 – 2012 arasında % 7,9 büyüyerek 31 milyon birim hacme ulaşacağı tahmin edilmektedir. Ülkeler bazında Almanya, bölgesel pazar değerinin % 26,3 oranındaki değeri ile en geniş pazar konumundadır. Almanya' yı sırasıyla % 23,7 ile İngiltere, % 19,6 ile Fransa ve % 16,5 ile İtalya takip etmektedir (Datamonitor, 2007).

İnternet, Avrupa' daki ikinci el otomobil satışlarında çok önemli pazar haline gelmiştir. Neredeyse herkesin internet erişimi olduğu bu kıtada, ikinci el otomobil satıcıları ve alıcıları internet ortamında buluşmaktadır. Avrupa' daki en popüler ikinci el otomobil e-ticaret siteleri şunlardır: Autoscout24.de (Almanya), Mobiliti.de (Almanya), Paruvenu.fr (Fransa), Viaturasonline.pt (Portekiz), Segundamano.es (İspanya), DerStandard.at (Avusturya), Autoscout24.ch (İsviçre).

Günümüzde ikinci el otomobil pazarının potansiyelini fark eden neredeyse tüm büyük otomobil üreticileri, şimdiden bu pazara girmişler ya da girme hazırlığı yapmaktadırlar. Bu da ticari açıdan ikinci el otomobil pazarının cazibesini ortaya koymaktadır.

3. ZAMAN SERİLERİ ANALİZİ

Bir ya da daha fazla değişkenin belirli zaman aralıklarında yapılan ölçümlerinden elde edilen sıralı veri kümeleri “zaman serileri” olarak adlandırılmaktadır. Zaman serileri yardımıyla geçmiş gözlem dönemlerine ait veriler kullanılarak, gelecek dönemler hakkında tahminde bulunulabilmektedir (Yılmaz, 2002; Kutlar, 2000).

Zaman serileri, kendisine etki eden pek çok faktörün etkisinde dalgalanma göstermektedir. Bu dalgalanmalar; ekonomik, sosyal, demografik, psikolojik vb. çeşitli faktörlerin zaman serilerini farklı yön ve şiddette etkilemesinden kaynaklanmaktadır. Zaman serilerinin en önemli özelliği, gözlem değerlerindeki değişmelerin genelde bu unsurların ortak etkisi neticesinde oluşmasıdır. Bu unsurlar, aşağıda açıklanan dört ana grupta toplanabilir (Gürtan, 1977: 421; Serper, 1981: 203-206; Hamburg, 1970: 542-544; Chatfield, 1996: 12-13):

- Trend (Uzun dönemli eğilim) (T)
- Mevsimsel dalgalanma (S)
- Konjonktürel (Devirsel) dalgalanmalar (C)
- Tesadüfî dalgalanmalar (I)

Üstte anılan unsurlara sahip bir zaman serisi genel olarak alttaki denklemlerle gösterilebilir:

$$Y = T + S + C + I$$

Zaman serileri analizlerinde kullanılan yöntemler şunlardır:

- Trend Analizi
- Mekanik Tahmin
- Hareketli Ortalamalar
- Üstel Düzleştirme
- Uyarlayıcı Arındırma
- Box-Jenkins (ARIMA)

Zaman serilerinin gelecek değerlerinin tahmin edilmesinde kullanılan yöntemin seçimi; tahminin amacına, zaman serisinin tipine ve unsurlarına, geçmişe ait veri miktarına ve tahmin döneminin uzunluğuna bağlı olarak değişmektedir (Akgül, 2003). Günümüzde, gelişmiş bilgisayar yazılımları aracılığıyla en uygun modelin otomatik olarak belirlenmesi mümkün olmaktadır.

4. YAPAY SİNİR AĞLARI

YSA, karmaşık yapıda içsel ilişkilere sahip problemlere çözüm bulabilmek için geliştirilmiş bir yapay zekâ tekniğidir (Çakar vd., 1996: 77). Yapay zekâ teknikleri, karmaşık ilişkilerin modellenmesindeki üstünlükleri ve veriler hakkında ön kabullenmelere ihtiyaç duymamaları sebebiyle son yıllarda çok büyük popülariteye sahip olmuştur.

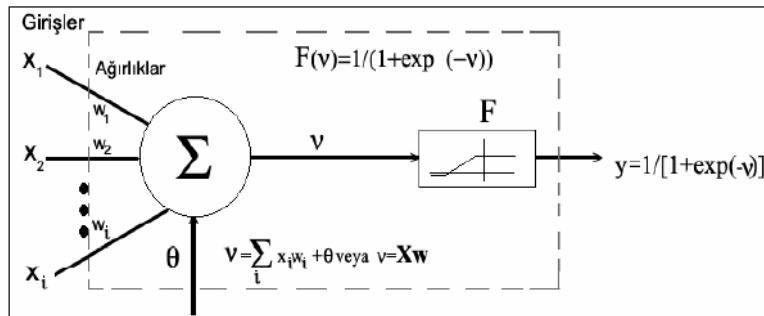
Haykin (1999), YSA için şu tanımları kullanmaktadır: “Bir sinir ağı, basit işlem birimlerinden oluşan, deneysel bilgileri biriktirmeye yönelik doğal bir eğilimi olan ve bunların kullanılmasını sağlayan yoğun bir şekilde paralel dağıtılmış bir işlemcidir. Bu işlemci iki şekilde beyin ile benzerlik göstermektedir: 1. Bilgi, ağ tarafından bir öğrenme süreciyle çevreden elde edilir. 2. Elde edilen bilgileri biriktirmek için sinaptik ağırlıklar olarak da bilinen nöronlar arası bağlantı güçleri kullanılır.”

İnsan beyninden esinlenerek ortaya çıkarılan YSA, deneme yolu ile öğrenme ve genelleştirme yapabilmektedir.

4.1. Yapısı

Bir yapay sinir ağı, kendisini oluşturan yapay sinir hücrelerinin birbirleriyle çeşitli katmanlar içerisinde paralel bağlantılar kurarak bir araya gelmelerinden oluşmaktadır. Yapay sinir hücreleri, biyolojik sinir sisteminin doğal yapısının matematiksel olarak modellenmesiyle oluşturulmaya çalışılmaktadır. Bu şekilde modellenen bir yapay sinir hücresi Şekil 1’de görülmektedir (Aydın, 2005). Yapay bir sinir hücresinde girdiler (X_1, X_2, \dots, X_n) kendilerine ait ağırlıklarla çarpılarak toplandıktan sonra aktivasyon fonksiyonuna gönderilirler. Aktivasyon fonksiyonuna gelen sinyal, burada işlenerek çıktıya dönüştürülür (Öztemel, 2003).

Şekil 1: Yapay sinir hücresi

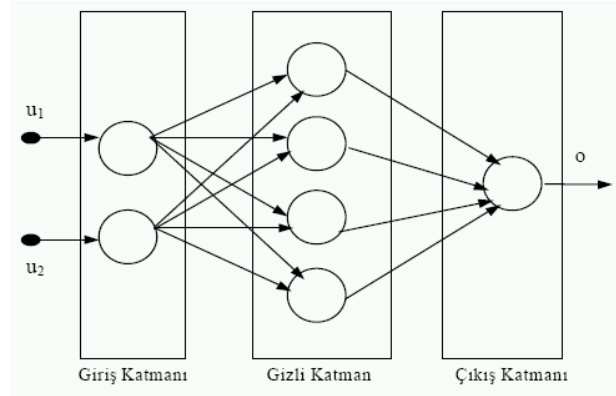


(Kaynak: Aydın, 2005)

Yapay sinir hücreleri, katmanlar içerisinde paralel şekilde bir araya gelerek YSA’ nı meydana getirmektedir. YSA’ nda üç ana katman

bulunmaktadır: girdi katmanı, gizli katman ve çıktı katmanı. Girdilerden oluşan girdi katmanı ile çıktılarından oluşan çıktı katmanı taktır. Gizli katman sayısı ise YSA' nın yapısına göre bir veya birden çok olabilmektedir. Girdi katmanından alınan girişler, girdi katmanı ve gizli katman arasında bulunan bağlantı ağırlıkları ile çarpılıp gizli katmana iletilmektedir. Gizli katmandaki nöronlara gelen girişler toplanarak aynı şekilde gizli katman ile çıktı katmanı arasındaki bağlantı ağırlıkları ile çarpılarak çıktı katmanına iletilir. Çıktı katmanındaki nöronlar da, kendisine gelen bu girişleri toplayarak uygun bir çıkış üretirler (Efe ve Kaynak, 2000: 8). Bağlantıların ağırlık değerleri öğrenme sürecinde belirlenmektedir. Şekil 2.4' de bir yapay sinir ağı oluşturulan katmanlar görülmektedir.

Şekil 2: İleri beslemeli sinir ağı yapısı



(Kaynak: Fırat ve Güngör, 2004)

YSA, hücrelerin bağlanma biçimlerine göre “ileri beslemeli” ve “geri beslemeli” olmak üzere iki mimari yapı altında sınıflandırılabilir (Slaughter, 2003):

İleri beslemeli ağlar: Verilerin sadece girdi birimlerinden çıktı birimlerine ileri doğru aktığı ağ yapısıdır. Bu yapıda nöronlar katmanlar şeklinde düzenlenir. Bir katmandaki nöronların çıkışları bir sonraki katmana ağırlıklar üzerinden giriş olarak verilir. Aynı katmandaki nöronlar arasında veya bir önceki katmana bağlantı yani geri besleme çevrimi yoktur. Uygulamalarda genellikle ileri beslemeli ağların tercih edildiği görülmektedir.

Geri beslemeli ağlar: Veri akışının sadece ileriye doğru değil geriye doğru da olabileceği ağ yapısıdır. Bu yapıda en az bir tane geri besleme çevrimi bulunur. Geri besleme, aynı katmandaki hücreler arasında olabileceği gibi farklı katmanlardaki nöronlar arasında da olabilir.

4.2. Süreçleri

YSA iki temel süreci içermektedir; İşlem ile eğitim-öğrenim süreci (Hosein, 2002):

4.2.1. İşlem Süreci

İşlem sürecinde toplama ve aktivasyon adlarında iki fonksiyon yer almaktadır.

Toplama (Birleştirme) Fonksiyonu, yapay sinir hücresine gelen girdilerin kendilerine ait ağırlıklarla çarpıldıktan sonra birleştirilmesi işlemini gerçekleştiren fonksiyondur. Bu fonksiyon, adından da anlaşılacağı gibi, genelde toplama işlemini kullanmakla birlikte farklı işlemleri de kullanabilir. Hatta araştırmacının kendi kurduğu işlemi de kullanması mümkündür. Toplama fonksiyonunda kullanılan işlem, genellikle seçilen ağ mimarisine de bağlıdır. En sık kullanılan fonksiyonlar şunlardır; Toplama, çarpım, maksimum, minimum, çoğunluk, kümülatif toplam (Öztemel, 2003).

Aktivasyon Fonksiyonu, yapay sinir hücresinin çıktısının büyüklüğünü sınırlandıran fonksiyondur. Bazı kaynaklarda transfer, eşik veya sıkıştırma fonksiyonu olarak da isimlendirilmektedir (Mandic ve Chambers, 2001). Bir ağdaki tüm hücrelerin aktivasyon fonksiyonu birbirinden farklı olabilir. Aktivasyon fonksiyonunda doğrusal fonksiyonlar genelde tercih edilmez. Zaman serileri için “Sigmoid”, ikili (binary) değişkenler için “Adım” fonksiyonu önerilmektedir (Tebelkis, 1995). En çok kullanılan aktivasyon fonksiyonları şunlardır: Doğrusal, adım, eşik değer, hiperbolik tanjant, sigmoid (Demuth ve Beale, 2000; Şen, 2004):

4.2.2. Eğitim ve Öğrenim Süreci

YSA’ nın en önemli özelliklerinden birisi, ilgili probleme ait örnekler verildiğinde girdiler ile çıktılar arasındaki ilişkiyi öğrenebilmesidir. YSA’ nın diğer bir süreci olan eğitim ve öğrenim sürecinin temel amacı girdi ve çıktı arasında bir eşleştirme yapmaktır. Danışmanlı öğrenme olarak adlandırılan bu yöntem ile ağ yeterince eğitildiğinde daha önceden görmediği girdilere karşı uygun çıktılar üretebilmektedir. Bu özellik YSA’ nın genelleme kapasitesi olarak adlandırılır ve çoğunlukla ağı yapısına ve eğitim için uygulanan yineleme sayısına bağlıdır (Alkan, 2001).

YSA’ nın, veri yapısındaki ilişkiyi öğrenmesi, probleme ait örnekler yardımı ile ağ ağırlıklarının en uygun değerlerinin belirlenmesine dayanır. Herhangi bir ağırlığın yeni değeri, eski değeri ile ağırlık değerlerinin değişim miktarının (ΔW) toplamı olarak belirlenir ($W_{\text{yeni}} = W_{\text{eski}} + \Delta W$). Belli bir kurala göre hesaplanan bu değişim miktarını (ΔW) belirlemek için tanımlanmış kurallara “öğrenme algoritmaları” denir.

YSA’ nın eğitimi için birçok değişik optimizasyon yöntemi

bulunmakla birlikte optimum sonucu garanti eden genel bir algoritma bulunmamaktadır. YSA' nın eğitiminde hatayı en aza indirmek için genellikle, hata fonksiyonunun yönünü bulmaya ve çıktı değerini azaltmaya çalışan dereceli azaltma (gradient descent) tabanlı algoritmalar kullanılır. Dereceli azaltma algoritmaları içerisinde en bilinen ve uygulamalarda en çok tercih edilen algoritma “Geri Yayılım Algoritması (GYA)” dır (Auclair, 2004).

4.3. Geri Yayılım Algoritması

Danışmanlı öğrenen, ileri beslemeli Geri Yayılım Ağları (Back Propagation Networks), adını hatayı yayma biçiminden alır. Hata azaltma işlemini hatayı tüm ağa yayarak gerçekleştirdiği için bu ağa “Hatayı Geriye Yayma – Error Back Propagation” da denmektedir. Bu algorithmada elde edilen çıktı ile olması gereken çıktı arasındaki fark yani hata, tüm ağırlıklara yansıtılarak dereceli olarak minimum düzeye indirilmeye çalışılır (Aydoğmuş ve Çöteli, 2005: 239-246). Geri yayılım algoritmasında eğitime rastgele bir ağırlık kümesi ile başlanır, birçok uygulamada ağırlık başarısı ağa atanan ilk ağırlık değerlerinin uygun seçilmesine bağlıdır (Nabiyev, 2003).

GYA, öğrenmesi ve uygulaması kolay bir modeldir. Günümüzdeki uygulamaların yaklaşık % 80'inde kullanıldığı tahmin edilen bu ağ, özellikle mühendislik alanında neredeyse her türlü problemin çözümü için kullanılabilir kadar güçlüdür (Kaastra ve Boyd, 1996). GYA, finansal zaman serilerinin gelecekteki davranışlarını tahmin etmede de sıklıkla tercih edilmektedir.

5. UYGULAMA

İkinci el otomobil pazarları ile yeni otomobil pazarları arasında çok yakın bir ilişki olduğu bilinmektedir. Her ne kadar ikinci el otomobil pazarları yeni otomobil pazarlarına belirli bir oranda bağımlı olsa da, yine de kendilerine özgü birtakım özellikleri mevcuttur. Genel olarak, tedarikleri esnek olmayan ve pazar dışı faktörlerden etkilenen ikinci el otomobillerin talep edilmeleri en az yeni araçlardaki kadar esnektir. Bu nedenle de, ikinci el otomobillerin fiyatlarının, yeni otomobil fiyatları ile karşılaştırıldığında çok daha değişken olduğu görülmekte, buna bağlı olarak da tahmin edilmesi zorlaşmaktadır (Lee, 2006: 120-121). Bu tahminleri yapabilmek için her şeyden önce uzun süreli ve nitelikli bir veri toplama işlemi gerçekleştirilmelidir. İnternet, bu konuda veri toplamaı kolaylaştırmış olsa da her gün milyonlarca ikinci el otomobilin sergilendiği yüzlerce web sitesinin sayfalarındaki ham verileri elde etmenin ve yapısal hale getirmenin zorluğu, bu konuda yapılabilecek çalışmalardaki en önemli sıkıntıyı oluşturmaktadır. Bu çalışmanın uygulamasında kullanılan verilerin internetten toplanarak yapısal hale getirilmiş olması, çalışmanın önemini bir kat daha artırmaktadır..

Aşağıda bu çalışmadaki uygulamanın aşamaları ve sonuçları anlatılmıştır. Uygulamada oluşturulan modellerin doğruluğunu ölçmek için literatürde yaygın olarak kullanılan üç istatistiğe başvurulmuştur: Ortalama Mutlak Hata (OMH) (Mean Absolute Error, MAE), Ortalama Mutlak Yüzde Hata (OMYH) (Mean Absolute Percentage Error, MAPE), Ortalama Hata Kareleri Kökü (OHKK) (Root Mean Squared Error, RMSE). Bilindiği gibi, tahmin hatası ile tahmin doğruluğu ters orantılıdır, yani tahmin hatası küçüldükçe modelin doğruluk derecesi artmaktadır. Bu nedenle, farklı tahmin modellerini karşılaştırırken en küçük hata değerine sahip model en iyi model olarak seçilmektedir. Bu uygulamada elde edilen sonuçların gösterildiği tablolarda, en iyi modeli vurgulayabilmek için en düşük hata değerleri ‘*’ işareti ile belirtilmiştir.

5.1. Verilerin Toplanması

Çalışmada Avrupa’ da ikinci el otomobil ilanı yayınlayan yaklaşık 40 adet web sitesinden 3 yıl boyunca sistematik olarak toplanmış olan veriler kullanılmıştır. Bu sitelerde yer alan veriler şu 8 ülkeyi içermektedir: Almanya, Fransa, İtalya, Portekiz, İspanya, Avusturya, Hollanda, İsviçre. Elde edilen ham veriler çeşitli bilgisayar teknikleri ve analiz yöntemleri ile temizlenip yapısal hale dönüştürülerek analize uygun hale getirilmiştir.

Örneklemdaki markalar, verilerin elde edildiği tarihteki ilan sayısına göre sıralanarak ilk 10 sırada yer alanları Tablo 1’ de listelenmiştir.

Tablo 1: İlan sayısının çokluğuna göre ilk 10 otomobil markası

No	Marka	İlan sayısı
1	VW	182689
2	BMW	126810
3	MERCEDES	126371
4	AUDI	102758
5	OPEL	94277
6	RENAULT	81744
7	PEUGEOT	67304
8	FORD	64474
9	CITROEN	34082
10	TOYOTA	28973

5.2. Verinin Hazırlanması

Çalışmadaki en fazla zaman ve enerjinin harcandığı bu aşamada; veri toplama, birleştirme, temizleme ve dönüştürme işlemleri neticesinde verilerin analiz ve modelleme çalışmasına uygun hale getirilmesi hedeflenmiştir. Bu amaçla öncelikle web sayfalarından düz metin olarak toplanmış olan veriler yapısal sorgulama dili (SQL, Structured Query Language) kullanılarak veritabanlarındaki uygun tablolara aktarılmıştır.

Örneğin herhangi bir ilanda rastlanan bir sözcük, bir araç markası ile eşleşiyor ise bu sözcük ilgili tablonun o ilana yönelik satırında BRAND (marka) alanına aktarılmış, şayet sayısal bir değer olup yanında para birimi bulunuyorsa (örneğin Euro sözcüğü ya da € simgesi), PRICE (fiyat) alanına aktarılmıştır.

Verilerin tablolara aktarılması sonrasında elde edilen kayıtlarda gözlenen sorunlar yine SQL kodları yardımıyla giderilmeye çalışılmıştır (Örneğin bir aracın üretim yılının 1900 olarak girilmesi gibi mantıksal hatalar). Sorunların giderilmesi sırasında, Avrupa’ da üretilen neredeyse tüm araçların detaylı teknik bilgilerini içeren bir model bankası referans olarak kullanılmıştır. Sorunların çözümleri fazla miktarda teknik bilgi gerektirdiği için burada anlatımına gerek duyulmamış, sadece sık rastlanan sorunlardan bir kısmı aşağıda ismen sıralanmıştır:

- Birden fazla girilmiş kayıtlar
- Yazım hataları
- Mantıksal hatalar
- Eksik bilgi
- Birim farkları

Verilerin analizlere uygun hale getirilmesinin ardından, araç sayısının çokluğundan dolayı uygulama girdisi olarak piyasada en çok işlem gören, diğer bir ifadeyle en popüler görünen 3 otomobil modelinin 1 yaşındaki (12 aylık) ve 20.000 km’ deki ilanlarının, Ocak 2005 – Aralık 2007 tarihleri arasındaki 3 yıllık günlük fiyat verilerinin aylık ortalamalarından oluşan zaman serileri kullanılmıştır. Bu otomobil modelleri Tablo 2’ de görülmektedir.

Tablo 2: Gelecekteki fiyatları tahmin etmede kullanılan otomobil modelleri

Marka	Temel Model	Detay Model	Kasa Tipi	Kapı Sayısı	Motor Gücü
VW	Golf	V 1.9 TDI Trendline	0	2	77
BMW	525	D DPF	0	4	130
MERCEDES	A150	Classic	0	3	70

Uygulamada günlük verilerin değil de aylık ortalamalarının kullanılmasının iki temel sebebi vardır:

1. Otomobil fiyatlarının dikkate değer ölçüde günlük değişimler göstermemesi,
2. Zaman zaman yaşanan teknik sorunlar veya bazı büyük hacimli web sitelerinin tamamının taranmasının birkaç gün sürebilmesi nedenleriyle sürekli olarak günlük veri elde edilememesi.

Analizlerde girdi olarak kullanılan veri kümeleri Tablo 3’ de verilmiştir.

Tablo 3: Veri kümeleri

	Veri Kümesi 1			Veri Kümesi 2			Veri Kümesi 3		
	2005	2006	2007	2005	2006	2007	2005	2006	2007
Ocak	17081	15794	16655	34875	37334	38833	15535	15481	16438
Şubat	16779	15802	16709	37543	37108	39001	16255	15613	16073
Mart	16771	15845	16739	37582	38261	39059	16741	15122	16262
Nisan	16819	15884	16853	37113	37637	39545	16733	15579	16341
Mayıs	16635	15890	16729	38352	37416	40432	16798	16311	15985
Haziran	16381	15854	16778	38258	38512	39862	16843	16372	16306
Temmuz	16066	15868	16997	38139	38593	39646	16769	16970	15944
Ağustos	15754	15996	16967	37550	38227	39739	18204	16449	15794
Eylül	15955	16229	16808	37417	38354	38830	17888	16311	16565
Ekim	16271	16357	16900	38205	38655	39364	16548	16361	16271
Kasım	16004	16338	17061	37907	38931	39032	16482	16607	16167
Aralık	15876	15710	16780	37672	39060	39135	16105	17795	15971

5.3. Zaman Serisi Analiz Sonuçları

Zaman serisi analizleri için R ve SPSS yazılımları kullanılarak veriye en uygun zaman serisi tahmin yöntemleri ve bu yöntemlerin optimum parametreleri tespit edilmeye çalışılmıştır. Elde edilen sonuçlar Tablo 4' de görülmektedir. Buna göre, optimum zaman serisi yöntemleri; 1., 2. ve 3. veri kümeleri için sırasıyla ARIMA(0,1,0)(0,0,0), Winter'ın Toplamsal ve ARIMA(1,0,0)(0,0,0) olarak bulunmuştur.

Tablo 4: Zaman serisi analiz sonuçları

Veri Kümesi	Hata Fonksiyonu	Üstel Düzleştirme	ARIMA
Veri Kümesi 1		Basit Mevsimsel	ARIMA(0,1,0)(0,0,0)
	OMH	142,873	134,314 *
	OMYH	0,875	0,821 *
	OHKK	192,694	170,455 *
Veri Kümesi 2		Winter'ın Toplamsal	ARIMA(0,1,0)(1,1,0)
	OMH	381.887 *	562,943
	OMYH	1,001 *	1,457
	OHKK	501,491 *	746,873
Veri Kümesi 3		Basit Mevsimsel	ARIMA(1,0,0)(0,0,0)
	OMH	376,699	259,547 *
	OMYH	2,278	1,605 *
	OHKK	485,931	356,927 *

5.4. Yapay Sinir Ağları Sonuçları

YSA uygulamasında SPSS ve Clementine yazılımları kullanılarak optimum parametre tespitleri yapılmaya çalışılmıştır. Kurulan Geri Yayılımlı Yapay Sinir Ağının tasarımı şu şekilde yapılmıştır:

- *Eğitim ve test* kümeleri için sırasıyla %70-%30, %80-%20 ve %90-%10 oranları denenmiş ve en başarılı sonuçların %90-%10 oranlarında elde edildiği gözlenmiştir.
- *Girdi nöronu sayısı, gizli katman sayısı ve gizli katmandaki nöron sayısı*, deneme-yanılma yöntemiyle tespit edilmiştir.
- *Çıktı nöronu sayısı*, zaman serisi tahmin problemlerinde tahmin döneminin uzunluğuna eşittir. Bu çalışmada tek dönemlik tahmin yapıldığı için çıktı nöron sayısı da bir olarak alınmıştır.
- *Aktivasyon Fonksiyonu* için çeşitli fonksiyonların denenmesinin ardından en iyi sonucu veren Hiperbolik Tanjant fonksiyonu kullanılmıştır.
- *Öğrenme Algoritması* olarak geri yayılım (back-propagation) algoritması tercih edilmiş, parametreleri yazılım aracılığıyla otomatik olarak belirlenmiştir.

Bu değerlendirilmelerin sonucunda optimum YSA modeli tüm veri kümelerinde aynı çıkmış ve Tablo 5’ de özetlenmiştir.

Tablo 5: YSA modeli

Girdi nöronu sayısı	Gizli katman sayısı	Gizli nöron sayısı	Çıktı nöronu sayısı	Aktivasyon fonksiyonu
1	1	2	1	Hiperbolik Tanjant

5.5. Modellerin Karşılaştırması

Zaman serileri yöntemlerinin uygulaması sonucunda elde edilen ve Tablo 4’ de gösterilmiş olan sonuçların içerisinde en yüksek doğruluğu veren yöntemler ile YSA’ nın uygulaması sonucu elde edilen değerler birleştirilerek Tablo 6’ da gösterilmiştir. Tablonun incelenmesiyle de anlaşılacağı üzere, tüm veri kümelerinde en küçük hata değerlerinin YSA’ nda elde edildiği görülmektedir.

Tablo 6: Modellerin karşılaştırmasına yönelik doğruluk ölçüleri

Veri Kümesi	Hata Fonksiyonu	Zaman Serisi Yöntemi	YSA
Veri Kümesi 1		ARIMA(0,1,0)(0,0,0)	
	OMH	134,314	106,659 *
	OMYH	0,821	0,653 *
	OHKK	170,455	151,127 *
Veri Kümesi 2		Winter'ın Toplamsal	
	OMH	381.887	297,795 *
	OMYH	1,001	0,778 *
	OHKK	501,491	369,775 *
Veri Kümesi 3		ARIMA(1,0,0)(0,0,0)	
	OMH	259,547	249,879 *
	OMYH	1,605	1,502 *
	OHKK	356,927	315,837 *

Bu sonuçlardan hareketle, ikinci el otomobil fiyatlarından oluşan zaman serilerinin ileriye yönelik tahminlerinde YSA' nın zaman serisi tahmin yöntemlerine göre daha güvenilir sonuçlar verdiği öne sürülebilir.

YSA ile elde edilen ileriye dönük iki yıllık tahmin değerleri Tablo 7' de verilmiştir.

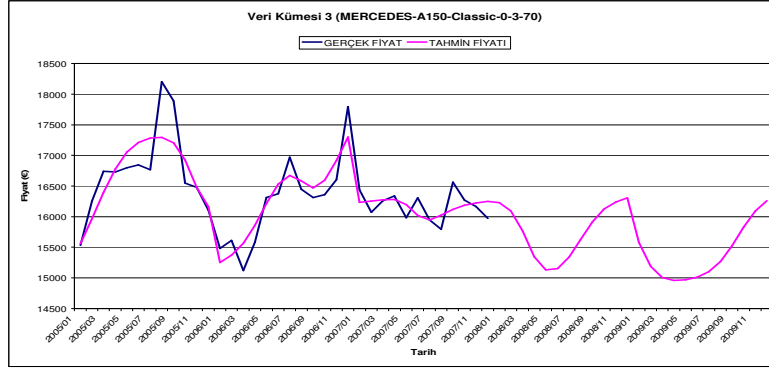
Tablo 7: Veri kümelerinin 2008-2009 yıllarında sahip olması öngörülen tahmin değerleri

	Veri Kümesi 1		Veri Kümesi 2		Veri Kümesi 3	
	2008	2009	2008	2009	2008	2009
Ocak	16847	16904	39707	39712	16227	15582
Şubat	16870	16921	39721	39705	16094	15187
Mart	16891	16936	39712	39693	15764	15004
Nisan	16910	16949	39690	39675	15345	14958
Mayıs	16927	16960	39659	39653	15129	14968
Haziran	16941	16970	39618	39624	15153	15012
Temmuz	16953	16978	39567	39588	15341	15101
Ağustos	16964	16984	39507	39543	15629	15265
Eylül	16972	16989	39437	39488	15916	15521
Ekim	16980	16994	39359	39424	16123	15828
Kasım	16986	16997	39276	39350	16242	16091
Aralık	16991	17000	39192	39269	16305	16262

Sonuçları görselleştirebilmek amacıyla, girdi olarak belirlenmiş olan otomobil modellerinden bir tanesinin (Mercedes A150 Classic, 3 no.lu veri kümesi) 2005-2007 yılları arasındaki gerçek fiyat verileri ile YSA modelinin ürettiği 2005-2009 arasındaki tahmini fiyat verileri ortak bir grafik üzerinde Şekil 3' de sunulmuştur. Bu grafikten de anlaşılacağı üzere, modelin ürettiği değerler geçmişteki gerçek değerler ile çok yakın paralellik göstermektedir. Bu sonuçtan hareketle, 2008-2009 yılları arasındaki tahmin değerlerinin de

tutarlılığının yüksek olacağı ifade edilebilir.

Şekil 3: Veri kümesi 3 için gerçek ve tahmini fiyatlar



6. SONUÇ

Bu çalışmada YSA ile ikinci el otomobillerin gelecekteki fiyatları tahmin edilmeye çalışılmıştır. YSA' nın tahmin doğruluğunu görebilmek için zaman serisi analiz yöntemleri de kullanılarak sonuçlar karşılaştırılmıştır.

Uygulamalarda kullanılan tüm yöntemlerde OMYH değerinin % 10' un altında çıkması, oluşturulan tüm modellerin “çok iyi” olduğunu göstermektedir. Bu nedenle, çalışmada oluşturulan tüm modeller ikinci el otomobil fiyatlarının tahmin edilmesinde başarıyla kullanılabilirler. Diğer yandan, kullanılan yöntemler arasında gözlenen hata değerleri arasındaki farklar çok yüksek olmasa da, YSA' nın tüm veri kümelerinde daha başarılı olması, bu yöntemin klasik istatistiksel yöntemlere alternatif oluşturabileceğini göstermiştir. Sonuç olarak, uygun parametrelerin seçildiği YSA' nın ikinci el otomobillerin güncel ve gelecek fiyatlarını tahmin etmede başarıyla kullanılacağı ifade edilebilir.

İleriki çalışmalara ışık tutması beklentisiyle uygulamada yaşanan bazı kısıtlamalar ve öneriler şu şekilde özetlenebilir: Öncelikle, girdi olarak kullanılan fiyat verileri gerçek satış fiyatlarından değil, web ilanlarından elde edilen arz fiyatlarından oluşmaktadır. Kapsamlı gerçek satış fiyatlarını edinmek pratikte mümkün görünmediği için bu verilere başvurulmak zorunda kalmıştır. Ayrıca, ikinci el otomobillerin fiyatlarını etkileyebilecek bazı ekonomik göstergeler (enflasyon, döviz kurları, vb) zaman kısıtından dolayı çalışmaya dahil edilmemiştir. İleride bu göstergelerin de dahil edilmesiyle yapılacak benzer çalışmalarda tahmin doğruluğu şüphesiz yükselecektir. Diğer yandan, Box-Jenkins yöntemlerinde etkin bir tahmin yapabilmek için en azından 50 veri bulunması önerilmesine rağmen (Box ve Jenkins, 1976) bu çalışmada farklı araç modellerini ifade eden her veri kümesi için sadece 36 adet (3 yıllık) veri kullanılabilmiştir. Çalışmada kullanılan Box-Jenkins

yönteminin YSA' na kıyasla daha az başarılı olmasının sebeplerinden birisi bu olabilir. YSA' nın az sayıda veriye rağmen güvenilir sonuçlar vermesi ise bu yöntemin güçlülüğünü ortaya koymaktadır.

Sonuç olarak, YSA' nın kara kutu (black box) olarak algılanmalarına rağmen, tasarımlarının kolay olması, probleme hızlıca adapte olabilmeleri ve az sayıda veriye rağmen başarılı sonuçlar üretebilmeleri nedeniyle pek çok alanda olduğu gibi otomobil sektöründe de kullanılmaları tavsiye edilebilir.

KAYNAKÇA

1. AKGÜL I., Zaman Serilerinin Analizi ve ARIMA Modelleri, Der Yayınları, İstanbul, 2003.
2. ALKAN A., Predictive Data Mining with Neural Networks and Genetic Algorithms, Ph.D. Thesis, İTÜ, İstanbul, 2001, 51.
3. AUCLAIR A., Feed-Forward Neural Networks Applied To The Estimation of Magnetic Distributions, M.S. Thesis, McGill University, Montreal, Canada, 2004, 128.
4. AYDIN Ö.,Yapay Sinir Ağlarını Kullanarak Bir Ses Tanıma Sistemi Geliştirilmesi, Yüksek Lisans Tezi, Trakya Üniversitesi, Edirne, 2005, 74.
5. AYDOĞMUŞ Z., ÇÖTELİ R., “Yapay Sinir Ağları Yardımıyla İzolatör Yüzeyinde Potansiyel Tahmini”, F.Ü. Fen ve Müh.Bil.Dergisi, 17/2, 2005, 239-246.
6. BOX G., JENKINS M., Time Series Analysis Forecasting and Control, Holden Day Inc., California, 1976.
7. CHATFIELD C., The Analysis of Time Series, Chapman & Hall, 5th Ed., New York, 1996.
8. ÇAKAR T., TÜRKER A. K., TORAMAN A., “İmalat Sistemlerinin Tasarlanmasında Yapay Sinirsel Ağların Kullanılması”, Birinci Ulusal Zeki İmalat Sistemleri Sempozyumu Zis'96 Bildirisi, Sakarya, Mayıs 1996, 10.
9. DATAMONITOR, “Used Cars in Europe”, Industry Profile, Reference Code: 0201-0750, December 2007.
10. DEMUTH H., BEALE M., Neural Network Toolbox for Use with Matlab, MathWorks Inc., 2000.

11. EFE Ö., KAYNAK O., Yapay Sinir Ağları ve Uygulamaları, Boğaziçi Üniversitesi Yayınları, Basım No: 696, İstanbul, 2000, 148.
12. FIRAT M., GÜNGÖR M., “Askı Madde Konsantrasyonu ve Miktarının Yapay Sinir Ağları ile Belirlenmesi”, İMO Teknik Dergi, 2004, 3267-3282.
13. GÜRTAN K., İstatistik ve Araştırma Metodları, Fatih Yayınevi Matbaası, İstanbul, 1977, 421.
14. HAMBURG M., Statistical Analysis for Decision Making, Harcourt, Brace and World Inc., New York, 1970.
15. HAYKIN S., Neural Networks: A Comprehensive Foundation, Prentice Hall International Inc., 1999.
16. HOSEIN P., “Stock Price Prediction By Artificial Neural Networks: A Study of Tehran’s Stock Exchange”, Management Studies Quarterly Journal of Management & Accounting School, No.31&32, 2002.
17. İŞERİ A., KARLIK B., “An Artificial Neural Networks Approach on Automobile Pricing”, Expert Systems with Applications, January 2008.
18. KAASTRA I., BOYD M., “Designing A Neural Network for Forecasting Financial and Econometric Time Series”, Neurocomputing, Vol. 10, 1996, 215-236.
19. KUTLAR A., Ekonometrik Zaman Serileri, Gazi Kitabevi, Ankara, 2000.
20. LEE J., Empirical Analysis of Wholesale Used Car Auctions, A Dissertation in University of California, 2006, 120-121.
21. MANDIC D. P., CHAMBERS J. A., Recurrent Neural Networks for Prediction - Learning Algorithms Architectures and Stability, John Wiley & Sons Ltd, 2001.
22. NABIYEV V.V.,Yapay Zeka, Seçkin Yayıncılık, Ankara, 2003.
23. ONAT M. G., Otomotiv Sektöründe Oranlar Yöntemi Aracılığı ile Finansal Analiz, Yüksek Lisans Tezi, Marmara Üniversitesi, İstanbul, 2007
24. ÖZTEMEL E., Yapay Sinir Ağları, Papatya Yayıncılık, Ağustos 2003.
25. SERPER Ö., İstatistik, Fikir Kitabevi, İstanbul, 1981, 23-206.

26. SLAUGHTER G. E. F., Artificial Neural Network for Temporal Impedance Recognition of Neurotoxins, M.S. Thesis, Virginia Commonwealth University, Virginia, 2003, 115.
27. ŞEN Z., Yapay Sinir Ağları İlkeleri, Su Vakfı, İstanbul, 2004.
28. TEBELKIS J., Speech Recognition Using Neural Networks, Ph.D. Thesis, Carnegie Mellon University, Pennsylvania, Mayıs 1995.
29. WEB-ACEA, Association Des Constructeurs Europeens D'automobiles, Economic Report, Acea's Position On Motor Vehicle Distribution in The European Union, 2001, www.acea.be, (15/09/2007).
30. WEB-KOBİFİNANS, www.kobifinans.com.tr/tr/sector/011001/12455, (01.06.2008).
31. WEB-WIKIPEDIA, en.wikipedia.org/wiki/Used_car, (11.06.2008).
32. YILMAZ L., A Decision Support Systems Using Data Mining, Yüksek Lisans Tezi, Yeditepe Üniversitesi, İstanbul, 2002, 16-22.
33. ZHANG G., PATUWO B.E., HU M.Y., "Forecasting with Artificial Neural Networks: The State of The Art", International Journal of Forecasting, Vol. 14, No.1, 1998, 35-62.