

## IVR Sistemlerindeki Çağrı Sayılarının Tahmininde Doğrusal Regresyon ve Makine Öğrenmesi Modelleri \*

*Linear Regression and Machine Learning Models in Prediction the Number of Calls in IVR Systems*

### ÖZET

Çağrı hacmi tahminleri, çağrı merkezlerinin operasyonlarının optimizasyonunda daha verimli ve etkili bir şekilde yönetilmesine olanak tanımaktadır. Ayrıca, çağrıların etkin yönetimi, çağrı hizmeti verilen işletmelerin müşteri memnuniyeti ve marka itibarını artırmada önemli bir rol oynamaktadır. Bu çalışmada, çağrı merkezi operasyonlarının optimizasyonunda yapay zeka ve makine öğrenmesi modellerinin uygulanabilirliğini ve etkinliğini araştırılması amaçlanmıştır. Bu amaç için, REST API kullanarak C# diliyle çağrı merkezlerinin karşılaştığı mevcut sorunlara yapay zeka destekli çözümler sunma potansiyeline sahip AGI tabanlı bir sistem geliştirilmiş ve IVR çağrıları kullanılarak aylık bir zaman serisi veri seti oluşturulmuştur. Bu veri setine doğrusal regresyon, rassal orman ve yapay sinir ağları modelleri uygulanarak, model performansları MAE, MAPE, RMSPE hatalara dayalı kriterleri ile karşılaştırılmıştır. Sonuç olarak, doğrusal regresyon modelinin basit bir eğilim çizgisi çizmekle sınırlı kaldığı ve yüksek hata kriteri değerleri ile veri setindeki değişkenliği yakalayamadığı gözlemlenmiştir. Yapay sinir ağları modeli ile 500 epoch için öğrenme süreci artırılmış, daha karmaşık ilişkiler öğrenilmeye çalışılmış olsa da, hata kriterlerine göre değerlendirildiğinde çok yüksek bir tahmin performansı elde edilememiştir. Rassal ormanlar modeli, veri setindeki karmaşıklıkları modelleme konusunda diğer iki modele kıyasla daha düşük hata kriteri değeri ile değişkenliği yakalamada en başarılı model seçilmiştir. Bu bulgular ışığında, çağrı hacmi tahmininde rassal orman dışındaki iki modelin oluşturulan IVR çağrıları veri setindeki takvim etkilerini, tatil dönemleri ve diğer etkenleri yansıtmada yetersiz kaldığı sonucuna ulaşılmıştır.

**Anahtar Kelimeler:** Makine Öğrenmesi, Rassal Orman, Yapay Sinir Ağları, IVR, AGI.

### ABSTRACT

The call volume predictions provide call centers with an opportunity to be managed more efficiently and effectively in optimizing their operations. In addition, effective management of calls plays an important role in increasing customer satisfaction and brand reputation of businesses that provide call services. This study aimed to investigate the applicability and effectiveness of artificial intelligence and machine learning models in optimizing call center operations. To this aim, an AGI-based system with the potential to provide artificial intelligence-supported solutions to the current problems faced by call centers is developed in C# using REST API, and a monthly time series dataset is created using IVR calls. The linear regression, random forest and artificial neural network models are applied to this dataset and the performances of them are compared with MAE, MAPE, RMSPE error-based criteria. As a result, it is observed that the linear regression model is limited to drawing a simple trend line and could not capture the variability in the dataset with high error criterion values. Although the learning process is increased with 500 epochs with the artificial neural networks model and more complex relationships are tried to be learned, a very high prediction performance could not be achieved when evaluated according to the error criteria. The random forest model is chosen as the most successful model in capturing variability with lower error criteria value compared to the other two models in modeling the complexities in the dataset. The random forest model is selected as the most successful model in capturing variability with lower error criteria values compared to the other two models in modeling the complexities in the dataset. In the light of these findings, it has been concluded that two models other than random forest in call volume estimation are insufficient to reflect calendar effects, holiday periods and other factors in the created IVR calls dataset.

**Keywords:** Machine Learning, Random Forest, Artificial Neural Networks, IVR, AGI.

### GİRİŞ

Kurumların müşterilerinden gelen telefon aramalarını karşılamak, müşterilerin talepleri yönlendirmek ve onlara çeşitli hizmetleri sunmak amacıyla oluşturdukları çağrı merkezleri günümüzde, müşteri memnuniyetine odaklanmaları, potansiyel müşteri bulmaları, yeni müşteri yaratmaları, sipariş ve ödeme gibi işlemlerde kolaylıklar sağlamaları gibi geniş kapsamlı destek sağlamaları ile kurumsal şirketlerin müşteri memnuniyeti ve marka itibarını artırmada önemli bir hizmet sunmaktadır. 1960'lardan günümüze kadar gelen ve giderek daha fazla gelişen çağrı merkezleri çağrının yönüne ve türüne göre; gelen çağrı merkezi hedefi, giden çağrı merkezi hedefi, karışık çağrı merkezi, sanal çağrı merkezi, çok kanallı ve tüm kanallı iletişim merkezleri, içeride (in-house) ve dış kaynaklı (outsourced) çağrı merkezleri, off-shore çağrı merkezi, PBX (özel santraller) olarak çeşitlenmektedir. Günümüzde PBX'ler en yaygın kullanım alanı olan çağrı merkezlerinden biridir.

Merve Baştuğ<sup>1</sup>   
Kadriye Hilal Topal<sup>2</sup> 

**How to Cite This Article**  
Baştuğ, M & Topal, K. H. (2024).  
"IVR Sistemlerindeki Çağrı Sayılarının Tahmininde Doğrusal Regresyon ve Makine Öğrenmesi Modelleri", International Academic Social Resources Journal, (e-ISSN: 2636-7637), Vol:9, Issue:5; pp:460-465. DOI: <https://doi.org/10.5281/zenodo.13995435>

Arrival: 25 July 2024  
Published: 26 October 2024

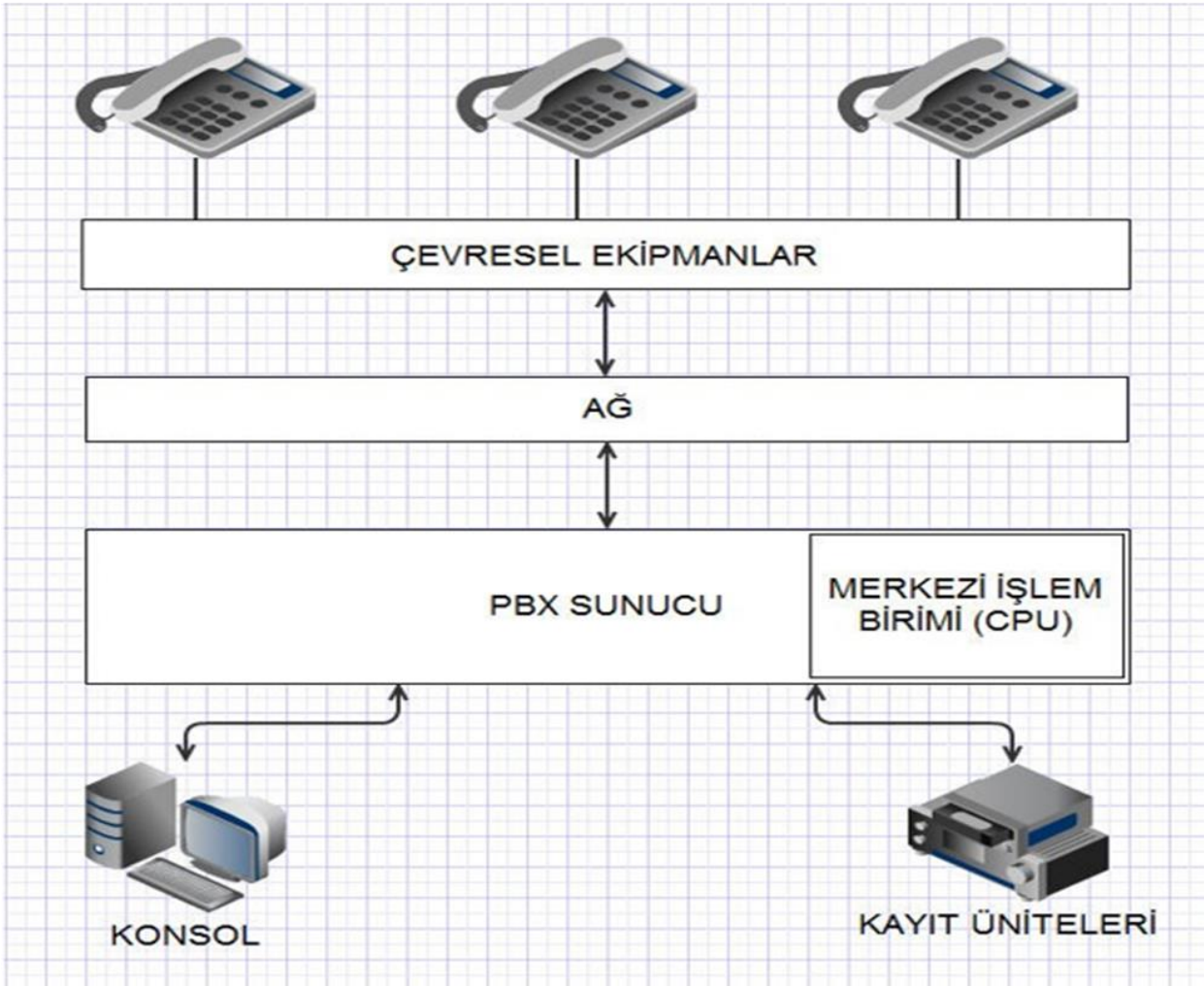
Academic Social Resources Journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

\* Yüksek Lisans tezinden üretilmiş bu çalışma, IERFM2024 Kongresinde sunulan bildirinin gözden geçirilmiş ve düzenlenmiş halidir.

<sup>1</sup> Yüksek Lisans Öğrencisi, İstanbul Nişantaşı Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yazılım Mühendisliği ABD, İstanbul, Türkiye.

<sup>2</sup> Dr. Öğr. Üyesi, İstanbul Nişantaşı Üniversitesi, Bilgisayar Programcılığı Pr., İstanbul, Türkiye.

PBX sistemleri, modern işletmelerin iletişim ihtiyaçlarını karşılamak için işletmelerin, hem iç hem de dış iletişimlerini etkili bir şekilde yönetmelerine olanak sağlamaktadırlar. PBX sistemlerinin özellikleri arasında çağrı yönetimi, çağrı tamamlama, çağrı transferi, gelişmiş çağrı özellikleri, yerel varlık hissi ve detaylı çağrı kontrolü bulunur. Çağrı merkezleri, gelen ve giden çağrı hacmi düşünüldüğünde, PBX gibi teknolojik altyapılar sayesinde işletmelere dijital çözümler sunarak etkin bir iletişim ağı kurulmasına olanak sağlamaktadır. Tipik bir santralin ana bileşenleri Şekil 1 ile gösterilmektedir.



Şekil 1: Tipik Bir Santralin Ana Bileşenleri

Kaynak: Karabayır, 2010:7

Tipik bir PBX santralde; CPU (Merkezi İşlem Birimi), ağ ve çevresel ekipmanlar olmak üzere üç ana bileşen bulunmaktadır. CPU, hafızasında sakladığı talimatları uygular, çevresel ekipmanların parçası olan hat gruplarına ve PBX istasyonlarına bağlanır ve anahtarlama fonksiyonlarını kontrol eder. Ağ, yapılacak olan görüşmelerde iletimin gerçekleştirileceği ortamı sağlar. Çevresel ekipmanlar ise, görüşme ortamının oluşturulmasında kullanılacak her çeşit teçhizatır.

Günümüzde çağrı merkezleri, yapay zeka'nın etkili kullanımı ile müşterileri anonsları sayesinde yönlendiren IVR (İnteraktif Yanıt Sistemi) kullanmaktadır. IVR sayesinde basit sorular sistem tarafından çözülürken, müşteri temsilcileri daha komplike sorunlara rahatlıkla çözüm sunabilecek zaman kazanırlar. Burada interaktif işlemler yapay zekanın özel bir dalı olan insan zekasına benzer şekilde problem çözebilen ve farklı alanlarda esnek bir şekilde kullanılabilen AGI (Yapay Genel Zeka) ile gerçekleştirilir. AGI sorunları anlayıp çözüm üretme işlemlerini NLP (Doğal Dil İşleme) teknolojisini kullanarak yapmaktadır. AGI büyük verilerin analizi ve raporlanması işlevlerini de yerine getirebildiğinden yöneticilere performans izleme ve hizmet kalitesini artırma konusunda daha fazla bilgi sağlamaktadır (Lee ve Moray, 1992: 1250). IVR ve AGI'nin kullanımı ile işletmelere daha hızlı ve etkin 7/24 hizmet, verilerini analiz ederek müşterilere özel hizmet sunma, NLP sayesinde müşterilerin daha karmaşık sorunların otomatik olarak çözülmesi, daha az insan müdahalesi, veri toplama ve analizi, sürekli iyileştirme, insan hatalarının azalması, maliyet tasarrufu gibi avantajlar sağlamaktadır.

SOAP (Basit Nesne Erişim Protokolü) ve REST API (Temsili Durum Aktarımı Uygulama Programlama Arayüzü), web tabanlı uygulamalar arasında veri alışverişi ve iletişim için kullanılan iki iletişim protokolüdür. Her ne kadar SOAP bir çok şirket tarafından veri iletimi sırasında hataların tespit edilmesini ve düzeltilmesini sağlayan güvenilir

bir iletişim protokolü olsa da, geniş bant aralığı gerektirmesi, yavaş olması ve insanlar tarafından okunabilir olmaması nedeniyle günümüzde REST API daha çok kullanılmaktadır (Box, Ehnebuske, Kakivaya, Layman, Mendelsohn, Nielsen, Thatte ve Winer, 2000: 50). AGI de REST API kullanımı; veri erişimi ve entegrasyon kolaylığı, AGI'nin yeteneklerini genişletme, özelleştirme, hızlandırma, ölçeklendirme ve bu sayede müşteri hizmetlerini iyileştirme, veri güncellemeleri ve senkronizasyon, mobil uygulamalar ve çevrimdışı kullanım, AGI'nin geliştirilmesini ve evrilmesi avantajlarını sağlamaktadır. Ayrıca REST API, üçüncü taraf geliştiricilerin kendi uygulamalarını sosyal medya platformlarına entegre etmelerini, hasta bilgilerinin güvenli bir şekilde farklı sağlık kurumları arasında entegrasyonunu, bankalarda ödeme ağ geçitleri ve finansal hizmet sağlayıcıları arasındaki veri alışverişininin kolaylaşmasına olanak sağlamaktadır.

IVR sistemlerinin üzerinde REST API ile AGI çözümlerinin etkin kullanımı işletmelerin operasyonel verimliliğini ve müşteri memnuniyetini artırmada oldukça önemlidir. Bu nedenle bu çalışmada, çağrı merkezlerinin operasyonel performansını doğrudan etkileyen çağrı sayıları üzerinde daha doğru tahminler yapılmasını sağlayacak makine öğrenmesi modelleri değerlendirilmiştir.

Literatürde çağrı merkezlerinde verimlilik ile ilgili yapılmış bazı çalışmalar bulunmaktadır. Bu çalışmalardan Ortakaya ve Tuntaş (2018), MEBİM (Millî Eğitim Bakanlığı İletişim Merkezi) müşteri temsilcilerinin performansını ölçmek amacıyla, toplam çalışma sürelerine kıyasla gerçekleştirdikleri görüşme sürelerini YSA (Yapay Sinir Ağları) kullanarak tahmin etmişlerdir. Analiz sonuçlarına göre YSA modelinin tahmin başarısı yüksek bulunmuştur. Baquirin ve Fernandez (2018), çağrı merkezi temsilcilerinin veya iş başvurularının telaffuzlarını değerlendirmek amacıyla, küçük bir konuşma kaydı veri seti üzerinde YSA kullanmış, YSA sınıflandırması sonucunda, belirlenen 10 kelimenin telaffuzunun 'Nötr' ya da 'Nötr Değil' olduğunu %89,60 tahmin doğruluğuyla tespit etmişlerdir. Ayrıca, bu çalışmada MLP (Çok Katmanlı Algılayıcı) ile yaptıkları tahmin başarısının da performansı yüksek bulunmuştur. Rausch ve Albrecht (2020) çalışmasında Almanya'da faaliyet gösteren bir çevrimiçi perakendecinin çağrı merkezinden alınan 174,5 haftalık verilerine ARIMA (Oto-regresif Entegre Hareketli Ortalama), ETS (Hata, Trend, Mevsimsel; İnovasyon Durumu Uzay Modeli), TBATS (Trigonometrik Mevsimsellik, Box-Cox Dönüşüm, ARMA hataları, Trend ve Mevsimsel Bileşenler), NBGLM (Negatif Binom Genelleştirilmiş Doğrusal Modeller), ZINB (Sıfır- Yükseltilmiş Negatif İki Terimli Modeller) gibi modelleri uygulayarak müşterilerin bekleme sürelerinin hizmet kalitesine etkisini incelemiş ve RO modelinin diğer modellerden daha iyi tahmin performansı gösterdiği bulgusuna ulaşmıştır. Yuce (2020), Vanad Laboratuvarları adlı bir çağrı merkezinden 2014 yılına ait çağrı günlükleri ve temsilci etkinlik günlükleri olarak tanımlanan iki farklı veri setini, Lineer Regresyon, Beta Regresyon, SVR, Random Forest ve XGBoost modelleri ile tahmin etmiş, MSE ve R2 değerleri açısından Random Forest modelinin daha iyi performans gösterdiğini ortaya koymuştur. Ballouch, Akay, Erdem, Tartuk, Nurdağ ve Yurdağül (2021) çalışmasında bir çağrı merkezindeki gelen çağrı sayısını tahmin etmek için zaman gecikmeleri ile entegreli MLP ve LTSM (Uzun Kısa Vadeli Bellek) tabanlı modeller kullanmış, hatalara dayalı kriterlere göre MLP tabanlı modellerin daha iyi sonuçlar verdiğini göstermişlerdir.

Literatürde çağrı merkezlerindeki çağrı tahminlerini konu alan sınırlı sayıda çalışma bulunmamaktadır. Çağrı merkezi operasyonlarının optimizasyonunda AI (yapay zeka) ve ML (makine öğrenmesi) modellerinin uygulanabilirliğini ve etkinliğinin araştırılması, bu alanda yeni bir yaklaşımların sunulması ve özellikle çağrı hacmi tahminleri yaparak, çağrı merkezlerinin daha verimli ve etkili bir şekilde yönetilmesine olanak tanınması çağrı merkezleri açısından büyük önem arz etmektedir. Çağrı merkezlerinin operasyonel performansını doğrudan etkileyen çağrı sayıları üzerinde daha doğru tahminler yapılmasını sağlayacak makine öğrenmesi modellerini geliştirmek ve değerlendirmek amacıyla bu çalışmada, çağrı merkezlerinin operasyonel performansını doğrudan etkileyen IVR verimliliğinin artırılması üzerinde yapay zeka sistemlerinin etkinliği incelenmiştir. Bu amaçla, çağrı merkezlerinin IVR sistemlerinin verimliliğini artırmak amacıyla REST API kullanarak C# diliyle çağrı merkezlerinin karşılaştığı mevcut sorunlara yapay zeka destekli çözümler sunma potansiyeline sahip AGI tabanlı bir sistem geliştirilmiştir. Sistem geliştirmede nesne tabanlı programlama dilinden yararlanılmıştır. IVR gelen çağrılar kullanılarak aylık bir veri seti oluşturulmuştur. Elde edilen veri seti doğrusal regresyon modeli tahmin yöntemlerinden LR (Doğrusal Regresyon), RO (Rassal ormanlar) ve YSA makine öğrenmesi modelleri ile tahmin edilmiş ve bu modellerin çağrı hacmi tahminlerindeki etkinliği hatalara dayalı kriterler ile karşılaştırılmış, RO regresyonu modelinin AGI ile IVR sisteminde müşteri verimliliği tahmininde en başarılı model olduğu sonucuna ulaşılmıştır.

Çalışmanın 2. Bölümünde çalışmada kullanılan makine öğrenmesi yöntemlerine ait metodoloji, 3. Bölümünde veri seti ve ampirik bulgular ve 4. Bölümünde ise sonuç kısmına yer verilmiştir.

## METODOLOJİ

Bu çalışmada, klasik modellerden LR, makine öğrenmesi modellerinden RO ve YSA tahmin modelleri kullanılmıştır.

## Doğrusal Regresyon Modeli

Doğrusal regresyon modeli, yaygın olarak kullanılan klasik bir tahmin yöntemidir. Yöntem, hata kereleri toplamının minimizasyonunu temel almaktadır. Genellikle en küçük kareler yöntemi kullanılarak hesaplanır. Maliyet fonksiyonunu minimize etmeyi hedefleyen yöntemin fonksiyonu;

$$J(\beta) = \sum_{i=1}^n (Y_i - (\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}))^2 \quad (1)$$

şeklinde ifade edilebilir. Burada  $y_i$  gerçek değerleri,  $X_{i1}, \dots, X_{ip}$  bağımsız değişkenleri,  $\beta_0$  regresyon sabitini,  $\beta_1, \dots, \beta_p$  regresyon katsayılarını temsil eder. Yöntemin genel formülü;

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon \quad (2)$$

dır. Burada  $Y$  bağımlı değişken,  $(X_1 + \dots + X_n)$  bağımsız değişkenler,  $\beta_0$  abit parametre,  $(\beta_1 + \dots + \beta_n)$  katsayı vektörü,  $\varepsilon$  hata terimidir. LR modelinde her bir katsayı, ilgili bağımsız değişkenin bağımlı değişken üzerindeki ortalama etkisini ölçer. Model, bağımlı ve bağımsız değişkenler arasındaki ilişkinin doğrusal olduğu varsayımına dayanır. Yöntemin geçerliliği için bağımsız değişkenler arasında çoklu doğrusallık (birbirleriyle yüksek korelasyon) olmaması, hata terimlerinin normal dağılıma uygunluk göstermesi ve birbirinden bağımsız olması (otokorelasyon olmamalıdır) ve hata terimlerinin varyansının sabit olması (homoskedastisite) temel varsayımlarının sağlanması gerekir (Güriş ve Çağlayan, 2010: 83).

## Rassal Orman Modeli

Ho (1998) ve Breiman (2001) tarafından geliştirilen RO regresyonu, birden fazla karar ağacının aynı bağımlı değişken için oluşturulması ve sonuçların birleştirilmesi sürecini içeren ve yüksek tahmin başarısı sağlayan bir ensemble öğrenme tekniğidir (Çağlayan-Akay, Topal, Kızıllarslan ve Bülbül, 2019: 8). Her bir karar ağacının olası tahmin değerlerinin ortalaması alınarak toplu bir karar üretmektedir. RO regresyon modeli, bireysel karar ağaçlarının aşırı uyum (overfitting) sorununu azaltarak daha doğru ve kararlı tahminler sağlamaktadır. Ayrıca yöntem, ensemble yaklaşımı sayesinde farklı ağaçların hatalarını dengeleyerek, tek bir karar ağacının gözden kaçırabileceği veri yapısını yakalamayı amaçlar. Hem sınıflandırma hem de regresyon için kullanılmaktadır. Sınıflandırma için;

$$Y_i = \text{mod}(\{Y_1, Y_2, \dots, Y_n\}) \quad (3)$$

şeklinde formüle edilebilir. Burada  $Y_i$  i'nci karar ağacının tahmini ve  $Y_i$  mod, yani en sık rastlanan sınıf etiketidir. Regresyon için;

$$Y_i = \frac{1}{n} \sum_{i=1}^n Y_i \quad (4)$$

Burada  $Y_i$  i'nci karar ağacının tahmini ve  $Y$  bu tahminlerin ortalamasıdır. Her bir karar ağacı temelde birer tahmincidir. Veri setinden rassal seçimle alınan alt örneklem eğitilir. Burada tekrar yerine koyma ile bootstrap örnekleme yapılmaktadır. Her ağaç, tüm özelliklerin rastgele seçilen bir özellik alt kümesi kullanılarak oluşturulur. Modeldeki karar ağaçlarının sayısı önemli bir hiperparametredir. Tahminlerin birleştirilmesi ortalamayla gerçekleştirilebildiği gibi, oylama ile de yapılabilir.

RO algoritma kodu ile yaprak düğümlerini döngü ile dönerek gerekli formülasyonlar yapılabilir (Cutler, Edwards, Beard, Cutler, Hess, Gibson ve Lawler, 2007: 2788). RO algoritması, veri kümesinin belirli bir kritere göre belirlenmiş bir durma koşuluna ulaşıncaya kadar yinelemeli şekilde devam eder.

## Yapay Sinir Ağları

Yapay sinir ağları, karmaşık ve doğrusal olmayan veri setlerinde özellikleri öğrenmek ve regresyon ve sınıflama tahminleri yapmak için en yaygın kullanılan derin öğrenme yöntemlerinden biridir. YSA'ların makine öğrenmesi aşamasında, verilen girişlere karşılık gelen çıktıları üretebilecek şekilde ağırlık ve sapma değerleri ayarlanmaktadır (Geurts, Ernst ve Wehenkel, 2006:40)

Veriler giriş katmanından alınır ve her girişte veri setindeki bir özelliğe karşılık gelen birer giriş nöronu yer almaktadır. Bu nöronlar ağırlıkları ve aktivasyon fonksiyonlarını kullanır. Giriş ve çıkış katmanları arasında modelin karmaşık ilişkileri öğrenme yeteneğini belirleyen gizli katmanlar bulunur. En sonda ise çıkış katmanı yer almaktadır. YSA'lar için geliştirilmiş çeşitli öğrenme algoritmaları bulunsun da en temel geri yayılım (backpropagation) ve gradyan inişi (gradient descent) algoritmalarıdır. Tek katmanlı bir YSA'nın denklemi;

$$Y = f(w \cdot x + b) \quad (5)$$

şeklinde. Burada  $Y$  çıkış vektörü,  $x$  giriş vektörü,  $w$  Ağırlık vektörü,  $b$  sapma terimi,  $f$  aktivasyon fonksiyonu ve “.” vektörlerin nokta çarpımıdır. Çok katmanlı YSA denklemi;

$$a^J = f^J(W^{[J]}.a^{[J-1]} + b^{[J]}) \quad (6)$$

şeklinde ifade edilir. Burada J katman,  $a^J$  J. katmanın aktivasyon çıktıları,  $a^{[J-1]}$  bir önceki katmanın aktivasyon girdileri,  $W^{[J]}$  J. katmanın ağırlık matrisi,  $b^{[J]}$ , J. katmanın sapma vektörü,  $f^J$  J. katmanın aktivasyon fonksiyonudur (Zhang, Patuwo ve Hu, 1998: 50). YSA tahmininde nöronlar ile giriş sinyallerini alınıp ağırlıklandırılır. Elde edilen ağırlıkların toplanmasına eklenerek aktivasyonun fonksiyonunda işlenir. Bu işlem her bir gizli katman için tekrarlanır.

## AMPİRİK BULGULAR

Çalışmada, çağrı merkezlerinin karşılaştığı mevcut sorunlara yapay zeka destekli çözümler sunma potansiyeline sahip, çağrı merkezlerinin IVR sistemlerinin verimliliğini artırmayı amaçlayan REST API kullanarak C# diliyle AGI tabanlı bir sistem geliştirilmiş, bir aylık süreçte IVR sistemine giren veriler kullanılarak kapsamlı bir veri seti oluşturulmuştur. Elde edilen veri setine, LR, RO ve YSA olmak üzere 3 farklı makine öğrenmesi yöntemi uygulanmış ve tahmin performansları karşılaştırılmıştır.

### Veri Seti

Çalışmadaki veri seti bir aylık süreçte REST API kullanarak C# diliyle AGI tabanlı geliştirilen IVR sisteminde web servislerden geçerek müşteri temsilcisine düşen 01-12-1023 ile 31-12-2023 tarihleri arasındaki günlük çağrı sayılarını içermektedir. Veri yaratma sürecinde nesne tabanlı programlama yaklaşımından yararlanılmış ve geniş class yapısı ile bir çok web servisi entegrasyonu gerçekleştirilmiştir. Elde edilen günlük çağrı sayılarına ait tanımlayıcı istatistikler Tablo 1'de sunulmuştur.

**Tablo 1:** Tanımlayıcı İstatistikler

N	Min.	Max.	Ortalama	Standart Sapma	Medyan
31	10783	38345	22299,645	6515,010	22060

Tablo 1'de görüldüğü üzere veri seti 31 gözlemden oluşmaktadır. Elde edilen en düşük çağrı miktarı 31-12-2023 günü için 10783 iken en yüksek çağrı miktarı 04-12-2023 günü için 38345'dir. Sisteme aylık ortalama 22299.645 çağrı düşmektedir. Model tahmini için veri seti %80'e %20 oranında bölünmüştür. Burada veri setinin %80'i model tahmini amacıyla öğrenme ve kalan %20'si modelin geçerliliğininin test edilmesi amacıyla test veri seti olarak kullanılmıştır.

Bu çalışma, Intel(R) Core(TM) i7-10510U CPU @ 1.80GHz işlemci ve 16 GB RAM özelliklerine sahip bir bilgisayarda gerçekleştirilmiştir. Kodlamada, Visual Studio Code ve Google Colab araçları kullanılmıştır. Analiz sürecinde açık kaynak kodlu Python programlama dili kullanılmıştır. Python'da LR modeli için sklearn.linear\_model kütüphanesi altındaki LinearRegression sınıfı, RO regresyonu için sklearn.ensemble kütüphanesi altındaki RandomForestRegressor sınıfı, YSA için TensorFlow ve Keras kütüphaneleri kullanılmıştır.

### Model Tahmin Sonuçları

Çağrı sayıları veri setine LR, RO ve YSA makine öğrenmesi yöntemleri uygulanmış, tahmin performansları karşılaştırılmıştır. Çalışmada eğitim veri seti, modelin öğrenme sürecinde parametre tahminleri için kullanılırken, test veri seti modelin örneklem dışı tahmin performansını değerlendirmek için kullanılmıştır. Test veri seti kullanılarak elde edilen örneklem dışı tahmin sonuçları Tablo 2 ile sunulmuştur.

**Tablo 2:** Model Karşılaştırma Tablosu

Model	LR	RO	YSA
MAE	5038,583	4906,251	19988,855
MAPE	31,39	27,68	92,31
RMSPE	39,40	38,71	92,35

Tablo 2'deki MAE (Ortalama Mutlak Hata), MAPE (Ortalama Yüzdese Mutlak Hata) ve RMSPE (Kök Ortalama Kare Yüzde Hata) hatalara dayalı kriterler incelendiğinde RO regresyonunun LR ve YSA modelinden daha iyi tahmin performansı gösterdiği görülmektedir. RO Regresyonu ile MAE (4906,251), MAPE (27,68) ve RMSPE (38,71) ile hatalara dayalı en düşük değerler elde edilmiştir. Hatalara dayalı kriterlerin düşük olması RO regresyonunun çağrı sayılarının tahmin için LR ve YSA'ye kıyasla gerçeğe daha yakın tahmin sonuçları elde edildiğini göstermektedir. LR ve YSA modelleri veri setinin değişkenlik ve dönemsel etkilerini yansıtmada yetersiz kalmıştır. RO modeli, doğrusal olmayan ilişkileri yakalama konusunda diğer iki modele kıyasla daha etkili bulunmuştur.

## SONUÇ

Bu çalışmada çağrı merkezlerindeki çağrı hacimlerinin tahmininde literatürde yaygın olarak kullanılan bazı makine öğrenmesi yöntemlerinin karşılaştırmalı olarak değerlendirilmesi gerçekleştirilmiştir. Çağrı merkezi veri seti üzerinde LR, RO ve YSA makine öğrenmesi modelleri kullanılarak tahminler gerçekleştirilmiş ve her bir modelin tahmin performansı hesaplanarak karşılaştırılmıştır.

Çalışmada veri setlerini oluştururken geniş class yapısı bir çok web servise entegre edilerek, web servisler sayesinde gelen müşterilerin bir çoğu azaltılarak IVR akışında sorunlarının giderilmesi ve müşteri temsilcisine düşen çağrı oranının azaltılması hedeflenmiştir. Elde edilen modellerin örneklem dışı performansını ölçmek için MAE, MAPE, RMSPE hata metriklerinden yararlanılmış ve en iyi model bu metriklere göre belirlenmiştir. Hatalara dayalı kriterler karşılaştırıldığında RO regresyonunun LR ve YSA modellerinden daha iyi tahmin performansı gösterdiği sonucu elde edilmiştir. Bulgulara dayanarak LR regresyonunun en iyi ikinci tahmin modeli olduğu sonucuna ulaşılmıştır. YSA ise aylık çağrı merkezi veri setinin tahmininde başarısız bir model olarak belirlenmiştir.

LR regresyonu modeli, genel eğilimi yakalamış ancak doğrusal bir model olması nedeniyle dalgalanmaları yansıtmada yetersiz kalmıştır. YSA ise, 500 epoch ile eğitildiğinde, artan bir trend yakalayabilmiş fakat yine de hatalara dayalı kriterlerin yüksek değerleri ile tahmin sonuçlarının gerçek değerlerden oldukça uzaklaştığı görülmektedir. RO regresyonunun tahmin performansı diğer iki yöntemden daha başarılı bulunmuştur.

Çalışmada elde edilen bulgular, önceki çalışmalarla karşılaştırıldığında, kullanılan geleneksel yöntemlerin kimi zorlukları aşmada yetersiz kaldığı söylenebilir. Makine öğrenmesi modelleri klasik yöntemlere kıyasla doğrusal olmama ve değişken veri seti yapısında çok farklı avantajlar sunmuş olsa da çağrı merkezinin dinamik ve dönemsel yapısını yansıtma konusunda sınırlı kalmıştır.

Sonuç olarak, çağrı merkezleri için daha etkili tahmin modellerinin geliştirilmesi gerekmektedir. Ayrıca, veri ön işleme aşamalarının iyileştirilmesi ve tahmin performansını etkileyebilecek farklı parametrelerin modellere dahil edilmesi tahmin başarısını artırabilir. Zaman serisi verilerinin dönemsel ve trend özelliklerini daha iyi yakalayabilecek ve hiperparametre ayarlamalarına optimizasyon sağlanabilen modellerin kullanılması daha iyi tahmin performanslarına ulaşmada etkili olabilir.

## KAYNAKÇA

- Çağlayan-Akay, E., Topal, H.K., Kızıllarlan, S. & Bülbül, H. (2019). “Türkiye Konut Fiyat Endeksi Öngörüsü: ARIMA, Rassal Orman ve ARIMA-Rassal Orman”, *Journal of Economic Perspectives*. DOI: 10.17261/Pressacademia.1134.
- Ballouch, M., Akay, F.M., Erdem, S., Tartuk, M., Nurdağ, F.T. & Yurdagül, H.H. (2021). “Makine Öğrenmesi Kullanarak Çağrı Merkezine Gelen Çağrıların Tahmin Edilmesi”, *Journal of Call Center Studies*, 4(1):96-101.
- Baquirin, R. B. M. & Fernandez, P.L. (2018). “Artificial Neural Network in a Small Dataset to Determine Neutrality in the Pronunciation of English as a Foreign Language in Filipino Call Center Agents”, *Journal of Language and Linguistic Studies*, 21(62):134-144.
- Box, D., Ehnebuske, D., Kakıvaya, G., Layman, A., Mendelsohn, N., Nielsen, H. F., Thatte, S. & Winer, D. (2000). Simple Object Access Protocol (Soap) 1.1. W3c Note”. Erişim Adresi: <https://Www.W3.Org/Tr/2000/Note-Soap-20000508/>. Erişim Tarihi: 23.12.2023.
- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984). *Classification and Regression Trees*. CRC Press.
- Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007). “Random Forests for Classification in Ecology”, *Ecology*, 88(11):2783-2792.
- Geurts, P., Ernst, D. & Wehenkel, L. (2006). “Extremely Randomized Trees”, *Machine Learning*, 63:3–42.
- Güriş, S. & Çağlayan, E. (2000). *Ekonometri*. Der yayınları, İstanbul.
- Ho, T. K. (1998). “The Random Subspace Method for Constructing Decision Forests”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):832–844. doi:10.1109/34.709601
- Karabayır, M. (2010). SIP protokolü ile internet üzerinden IVR etkileşimli çağrı merkezi hizmeti. Yüksek Lisans Tezi, Haliç Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Lee, J. H. & Moray, N. (1992). “Trust, Control Strategies And Allocation of Function in Human-Machine Systems”, *Ergonomics*, 35(10):1243-1270.
- Ortakaya, S. & Tuntaş, R. (2018). “Çağrı Merkezi Çalışanlarının Aylık Performanslarının Yapay Sinir Ağları Yardımı ile Tahmin Edilmesi”, *Journal of Call Center Performance*, 4(3):161-172.
- Rausch, M.T. & Albrecht, T. (2020). The Impact of Lead Time And Model Selection on the Accuracy of Call Center Arrivals' Forecasts. Yüksek Lisans Tezi, Bayreuth Üniversitesi, Almanya.
- Yüce, I.E. (2020). Validation Of Call Center Assumptions and Prediction of a Single-Skill Call Center Performance Measures via Machine Learning Approach. Yüksek Lisans Tezi, Vrije Üniversitesi, Amsterdam.
- Zhang, G., Patuwo, B. E. & Hu, Y. M. (1998). “Forecasting with Artificial Neural Networks: The State of the Art”, *International Journal of Forecasting*, 14(1):35–62.