**Çevrimiçi Alışveriş Davranışlarının Sınıflandırılması: Makine Öğrenmesi Yöntemleri ile Satın Alma Niyetinin Tahmini**

*Bilgisayar Mühendisliği, Teknoloji Fakültesi, Gazi Üniversitesi, Ankara, Türkiye*

*ydamla.kara@gazi.edu.tr*

*Bilgisayar Mühendisliği, Teknoloji Fakültesi, Gazi Üniversitesi, Ankara, Türkiye*

*açelya.dagasan@gazi.edu.tr*

**ÖZET**

Dijital teknolojinin evriminde e-ticaret sektörleri müşteri taleplerini karşılamak ve aynı zamanda düşük maliyet için kademeli olarak değişmektedir. Online alışveriş de, kullanıcıların alışveriş yaparken davranışlarını analiz ederek müşteriye yönelik hizmetlerin gelişmesine katkı sağlamıştır. Bu çalışmada kullanıcıların farklı faaliyetlerinin araştırılması ve gereksinimlerinin tahmin edilmesi amaçlanmıştır. Kullanılan veri seti, kullanıcıların web sitesinde geçirdiği süre, ziyaret edilen sayfa türleri, cihaz özellikleri ve ziyaret zamanı gibi çeşitli niteliklerden oluşmaktadır. "Online Shoppers Purchasing Intention Dataset" adlı veri seti kullanılarak, sınıflandırma problemlerine yönelik beş farklı makine öğrenmesi yöntemi uygulanmıştır. Çalışmada Karar Ağacı, Rastgele Orman, K-En Yakın Komşu, Lineer Regresyon ve Basit Bayes algoritmaları değerlendirilmiştir. Sınıflandırıcıların performansını ve ölçeklenebilirliğini iyileştirmek için bazı ön işleme adımları kullanılmıştır. Veri setinde bulunan kategorik değişkenler etiketlenmiş ve sayısal veriler normalize edilerek eğitim ve test verisi olarak ikiye ayrılmıştır. Algoritmaların doğruluk ve f1-score gibi performans metrikleri karşılaştırılmıştır. En başarılı sonuç %89.8 doğruluk oranı ve 0.63 f1-score değeri ile Rastgele Orman algoritmasında elde edilmiştir. En düşük performans ise %79.6 doğruluk ile Basit Bayes yöntemine aittir. Elde edilen sonuçlar, müşteri davranışları sınıflandırılmasında önerilen modelin verimli bir şekilde kullanılabileceğini ortaya koymaktadır. Bu çalışma, e-ticaret sitelerinde kullanıcı davranışlarını analiz etmek, pazarlama stratejilerini geliştirmek ve öneri sistemlerine katkı sağlamak amacıyla yürütülen çalışmalara katkı sağlayacaktır.

**Anahtar kelimeler:** Makine öğrenmesi, Online Alışveriş, Satın Alma Tahmini

**1. Giriş**

Son yıllarda internet kullanımının yaygınlaşmasıyla birlikte çevrimiçi alışveriş, tüketici alışkanlıklarında önemli bir değişime yol açmıştır. Artık birçok kullanıcı, fiziksel mağazalara gitmek yerine ürünleri çevrimiçi olarak incelemekte ve satın alma kararlarını dijital ortamda vermektedir. Bu dönüşüm, e-ticaret platformlarının yalnızca ürün sunan değil, aynı zamanda kullanıcı davranışlarını analiz eden ve bu davranışlara göre özelleştirilmiş deneyimler sunan yapılar haline gelmesini zorunlu kılmıştır. Kullanıcıların çevrimiçi ortamdaki etkileşimlerinin veriye dönüştürülmesiyle birlikte, bu verilerin doğru yorumlanması işletmeler açısından stratejik bir gereklilik haline gelmiştir.

Kullanıcıların bir e-ticaret sitesindeki gezinme davranışlarını analiz ederek satın alma eğilimlerini öngörebilmek, hem satışları artırmak hem de müşteri memnuniyetini yükseltmek açısından büyük önem taşımaktadır. Ziyaret edilen sayfa sayısı, geçirilen süre, yönlendirme kaynakları, işlem yapılan cihaz türü, işlem günü gibi birçok değişken, kullanıcıların satın alma kararlarını etkileyen önemli faktörler arasında yer almaktadır. Ancak bu veriler büyük miktarda ve çoğunlukla yapısal olmayan biçimde toplandığı için, klasik analiz yöntemleri bu verilerden anlamlı sonuçlar üretmekte yetersiz kalabilmektedir.

Bu noktada, makine öğrenmesi algoritmaları; büyük, çok boyutlu ve karmaşık verilerle başa çıkabilme kabiliyetleri sayesinde, kullanıcı davranışlarını modelleme ve satın alma tahminleri yapma konusunda etkili bir çözüm sunmaktadır. Bu çalışmada, çevrimiçi alışveriş kullanıcılarının oturum bazlı verilerini inceleyerek, satın alma gerçekleştirip gerçekleştirmeyeceklerini tahmin etmek amaçlanmıştır. Bu amaç doğrultusunda, UCI Machine Learning Repository'de yayımlanan ve 12.330 oturum kaydını içeren **Online Shoppers Purchasing Intention Dataset** kullanılmıştır. Veri seti, 18 bağımsız değişken ve bir hedef değişkenden oluşmakta olup; kullanıcının oturum sonunda alışveriş yapıp yapmadığını belirten "Revenue" adlı ikili sınıfı hedeflemektedir.

Çalışmada, Karar Ağacı, Rastgele Orman, K-En Yakın Komşu (KNN), Lineer Regresyon ve Basit Bayes olmak üzere beş farklı makine öğrenmesi algoritması kullanılmıştır. Modellerin performansları doğruluk (accuracy) ve özellikle dengesiz veri setleri için kritik öneme sahip olan F1-score metriği ile değerlendirilmiştir. Yapılan deneyler sonucunda, en yüksek doğruluk ve F1-score değerleri Rastgele Orman algoritmasında elde edilmiştir. Bu sonuçlar, özellikle e-ticaret platformlarında kullanıcı alışkanlıklarını tahmin etmeye yönelik sistemlerde hangi algoritmaların daha etkili olduğunu ortaya koymakta ve bu alanda yapılacak yeni çalışmalara temel oluşturmaktadır.

Bu çalışmanın katkısı, oturum verilerine dayalı bir sınıflandırma yaklaşımıyla kullanıcı satın alma eğilimlerinin doğru biçimde öngörülebileceğini göstermesi ve farklı algoritmaların performanslarını karşılaştırarak en uygun modelin belirlenmesine olanak sağlamasıdır. Böylece, pazarlama stratejilerinin daha isabetli kurgulanması, kullanıcı deneyiminin kişiselleştirilmesi ve kaynakların etkin yönetimi gibi birçok alanda fayda sağlanabileceği düşünülmektedir.

**2. Materyal ve Yöntem**

E-ticaret platformlarında kullanıcı davranışlarını analiz ederek, kullanıcıların alışveriş yapma olasılıklarını tahmin etmeyi amaçlayan bir sınıflandırma modeli geliştirilmiştir. Kullanılan veri seti, Online Shoppers Purchasing Intention Dataset olup, bu veri seti, kullanıcıların çevrimiçi alışveriş oturumları ile ilgili çeşitli davranışsal ve teknik bilgileri içermektedir. Veriler, kullanıcıların site üzerindeki etkileşimlerini, gezinme sürelerini, ürün sayfalarındaki geçirdikleri süreyi, trafik kaynaklarını, ziyaret tarihlerini ve teknik cihaz bilgilerini içermektedir.

**2.1. Veri Seti**

Günümüz dünyasında dijitalleşmenin hızla yayılması, e-ticaret platformlarını günlük yaşamın ayrılmaz bir parçası hâline getirmiştir. Özellikle pandemi sonrası dönemde çevrimiçi alışveriş davranışları önemli ölçüde artmış ve kullanıcıların dijital ortamdaki hareketlerinin analiz edilmesi, işletmeler açısından stratejik önem kazanmıştır. Bu bağlamda, kullanıcıların çevrimiçi alışveriş sırasında satın alma davranışlarını tahmin etmeye yönelik geliştirilen veri analiz yöntemleri, pazarlama ve karar destek sistemlerinin etkinliğini artırmayı hedeflemektedir.

Bu çalışmada kullanılan veri seti, UCI Machine Learning Repository üzerinden temin edilen ve "Online Shoppers Purchasing Intention Dataset" adıyla bilinen açık erişimli bir veri kümesidir. Veri seti, bir e-ticaret web sitesinde alışveriş yapan kullanıcıların oturum bilgilerini içermektedir. Toplam 12330 adet örnek ve 18 öznitelikten oluşan bu veri seti, hem sayısal hem de kategorik özellikler barındırmaktadır.

Öznitelikler; kullanıcının ziyaret ettiği sayfa türleri (Administrative, Informational, ProductRelated), bu sayfalarda geçirdiği süreler, sayfa değerleri (PageValues), çıkış oranları (ExitRates), kullanılan işletim sistemi, tarayıcı, ziyaret ayı, trafik kaynağı, ziyaretçinin yeni ya da geri dönen kullanıcı olup olmadığı (VisitorType), ziyaretin hafta sonuna denk gelip gelmediği (Weekend) gibi bilgilerden oluşmaktadır.

Veri setindeki bağımlı değişken olan Revenue özelliği, kullanıcının oturumu sonunda bir satın alma işlemi gerçekleştirip gerçekleştirmediğini belirtmektedir. Bu özellik True/False biçiminde iki sınıftan oluşur ve çalışmada sınıflandırma problemi kapsamında değerlendirilmiştir.

Veri seti üzerinde ön işleme adımları gerçekleştirilmiştir. Kategorik değişkenler sayısal verilere dönüştürülmüş, ardından sayısal özelliklerin aralık farklılıklarını ortadan kaldırmak amacıyla StandardScaler yöntemi kullanılarak normalizasyon uygulanmıştır. Model eğitimi için veri, %70 eğitim – %30 test olarak ayrılmıştır.

Bu veri seti, çevrimiçi kullanıcı davranışlarının sınıflandırılması amacıyla çeşitli makine öğrenmesi algoritmaları için uygun bir örnek teşkil etmektedir. Bu yönüyle e-ticaret sektöründe kullanıcı segmentasyonu, hedefli reklam stratejileri ve öneri sistemlerinin geliştirilmesi gibi alanlarda önemli katkılar sağlamaktadır.

| **Değişken Adı** | **Rol** | **Tip** |  |
| --- | --- | --- | --- |
| İdari (Administrative) | Özellik | Tam sayı |  |
| Yönetimsel\_Süre (Administrative\_Duration) | Özellik | Tam sayı |  |
| Bilgilendirici (Informational) | Özellik | Tam sayı |  |
| Bilgilendirici\_Süre (Informational\_Duration) | Özellik | Tam sayı |  |
| Ürünİlgili (ProductRelated) | Özellik | Tam sayı |  |
| Ürünİlgili\_Süre (ProductRelated\_Duration) | Özellik | Sürekli |  |
| Hemen Çıkma Oranları (BounceRates) | Özellik | Sürekli |  |
| Çıkış Oranları (ExitRates) | Özellik | Sürekli |  |
| SayfaDeğerleri (PageValues) | Özellik | Tam sayı |  |
| Özel Gün (SpecialDay) | Özellik | Tam sayı |  |
| Ay (Month) | Özellik | Kategorik |  |
| İşletim Sistemleri (OperatingSystems) | Özellik | Tam sayı |  |
| Tarayıcı (Browser) | Özellik | Tam sayı |  |
| Bölge (Region) | Özellik | Tam sayı |  |
| Trafik Türü (TrafficType) | Özellik | Tam sayı |  |
| Ziyaretçi Türü (VisitorType) | Özellik | Kategorik |  |
| Hafta sonu (Weekend) | Özellik | İkili |  |
| Hasılat (Revenue) | Hedef | İkili |  |

**2.2. Kullanılan Makine Öğrenmesi Algoritmaları**

Makine öğrenmesi, yapay zekâ alanının bir alt dalı olup, geçmiş verilere dayanarak örüntülerin keşfedilmesi ve geleceğe yönelik tahminlerde bulunulması amacıyla kullanılmaktadır. Günümüzde makine öğrenmesi öneri sistemlerinden sağlık analizlerine, siber güvenlikten müşteri davranışı tahminine kadar pek çok alanda yaygın olarak uygulanmaktadır. Özellikle e-ticaret sektöründe, kullanıcıların çevrimiçi davranışlarının incelenmesi ve satın alma niyetlerinin öngörülmesi, pazarlama stratejilerinin belirlenmesinde önemli bir rol oynamaktadır.

Bu çalışmada, online alışveriş yapan kullanıcıların satın alma işlemi gerçekleştirip gerçekleştirmeyeceklerini tahmin etmek amacıyla çeşitli makine öğrenmesi algoritmaları uygulanmıştır. Uygulanan algoritmalar aşağıda açıklanmıştır:

Rastgele Orman (Random Forest): Makine öğrenmesinde yaygın olarak kullanılan denetimli bir öğrenme yöntemidir. Bu algoritma, birden fazla karar ağacının oluşturulması ve bu ağaçların sonuçlarının birleştirilmesiyle çalışır. Sınıflandırma ve regresyon problemlerinde etkili sonuçlar sağlar.

K-En Yakın Komşu (K-Nearest Neighbors, KNN): Denetimli (supervised) öğrenme yöntemlerinden biridir ve hem sınıflandırma hem de regresyon problemlerinde kullanılır. KNN, basitliği ve etkinliği sayesinde makine öğrenmesi alanında sıkça tercih edilen bir algoritmadır.  Sınıflandırmasında en yakın  komşularının sayısı olan k’ı grup üyeliğine göre verileri sınıflandırmak için kullanır.

Karar Ağacı (Decision Tree): Veri kümesini özniteliklere göre bölerek sınıflandırma yapan ağaç yapısında bir modeldir. Her bir iç düğüm bir özelliğe, her dal bir özelliğin değerine ve her yaprak düğüm bir sınıfa karşılık gelir. Yorumlanabilirliği yüksek bir yöntemdir ancak tek başına kullanıldığında aşırı öğrenmeye yatkın olabilir.

Lineer Regresyon (Linear Regression): Genellikle sürekli değişkenlerin tahmini için kullanılan bu yöntem, bazı durumlarda ikili sınıflandırmalarda da referans olarak değerlendirilebilmektedir. Bu çalışmada doğrusal bir yapı önerisi olarak ele alınmıştır.

Basit Bayes (Naive Bayes): Koşullu olasılık teorisine dayanan ve öznitelikler arasında bağımsızlık varsayımı yapan bir sınıflandırma algoritmasıdır. Veri seti büyük ve öznitelikler birbirinden bağımsız ise oldukça etkili sonuçlar üretebilir. Ayrıca hızlı ve düşük maliyetli olması nedeniyle ön analizlerde tercih edilmektedir.

Veri seti üzerinde öncelikle kategorik değişkenler sayısal formata dönüştürülmüş, ardından sayısal öznitelikler StandardScaler yöntemiyle normalize edilmiştir. Bu işlem, özellikle mesafe tabanlı algoritmaların (örneğin KNN) sağlıklı çalışabilmesi açısından önemlidir. Daha sonra veri seti %70 eğitim ve %30 test olmak üzere ikiye ayrılmıştır.

Uygulanan algoritmalar Python programlama dili ile gerçekleştirilmiş, model performansları doğruluk, hassasiyet, duyarlılık ve f1-skoru metrikleriyle değerlendirilmiştir. Ayrıca ROC eğrileri çizilerek görsel karşılaştırmalar yapılmıştır. Detaylı başarı karşılaştırmaları Bulgular bölümünde verilmiştir.

**2.3. Literatür Taraması**

Makine öğrenmesi algoritmaları, kullanıcı davranışlarının modellenmesi ve tahmini konularında günümüzde oldukça yaygın şekilde kullanılmaktadır. E-ticaret platformlarında kullanıcıların satın alma eğilimlerinin doğru şekilde tahmin edilmesi, şirketlerin pazarlama stratejilerinden müşteri sadakati uygulamalarına kadar birçok alanda önemli katkılar sağlamaktadır.

"Machine Learning for Predicting Online Shoppers’ Purchase Intentions" başlıklı makale, çevrimiçi alışveriş yapan kullanıcıların satın alma niyetlerini tahmin etmek amacıyla çeşitli makine öğrenmesi algoritmalarını uygulamaktadır. Çalışmada, kullanıcıların site içerisindeki gezinme süresi, sayfa türü ve ziyaret zamanı gibi özniteliklerin, satın alma niyetiyle doğrudan ilişkili olduğu tespit edilmiştir. Rastgele Orman (Random Forest) algoritması, %88’e yakın bir doğruluk oranı ile en yüksek başarıyı göstermiştir. [1].

Md. Shahriare Satu ve Syed Faridul Islam tarafından yazılmış bu araştırma, çevrimiçi alışveriş yapan kullanıcıların satın alma niyetlerini daha doğru tahmin edebilmek amacıyla kapsamlı bir makine öğrenimi modeli önermektedir. Çalışmada, UCI Makine Öğrenimi Deposundan alınan Online Shoppers Purchasing Intention Dataset veri kümesi kullanılmıştır. Daha sonra, farklı özellik seçimi teknikleri kullanılarak çeşitli alt veri kümeleri oluşturulmuş ve bu alt kümeler üzerinde çeşitli sınıflandırma algoritmaları (örneğin, Rastgele Orman, Karar Ağacı) uygulanmıştır. Rastgele Orman (Random Forest) algoritması, Z-Score ve Gain Ratio ile dönüştürülmüş alt kümede %92,39 doğruluk ve 0,924 F1 skoru ile en iyi performansı göstermiştir [2].

C. Okan Sakar, S. Olcay Polat, Mete Katircioglu ve Yomi Kastro tarafından yazılmış olan "Real-time prediction of online shoppers’ purchasing intention using multilayer perceptron and LSTM recurrent neural networks" başlıklı makalede kullanıcıların ziyaret sırasında topladıkları sayfa görüntüleme verileri, oturum bilgileri ve kullanıcı bilgileri kullanılarak satın alma niyeti tahmin edilmektedir. Rastgele Orman (RF), Destek Vektör Makineleri (SVM) ve Çok Katmanlı Algılayıcı (MLP) sınıflandırıcılarıyla farklı veri setlerinde test edilmiştir. Sonuçlar, MLP'nin daha yüksek doğruluk ve F1 skoru ile en iyi performansı gösterdiğini ortaya koymuştur. [3].

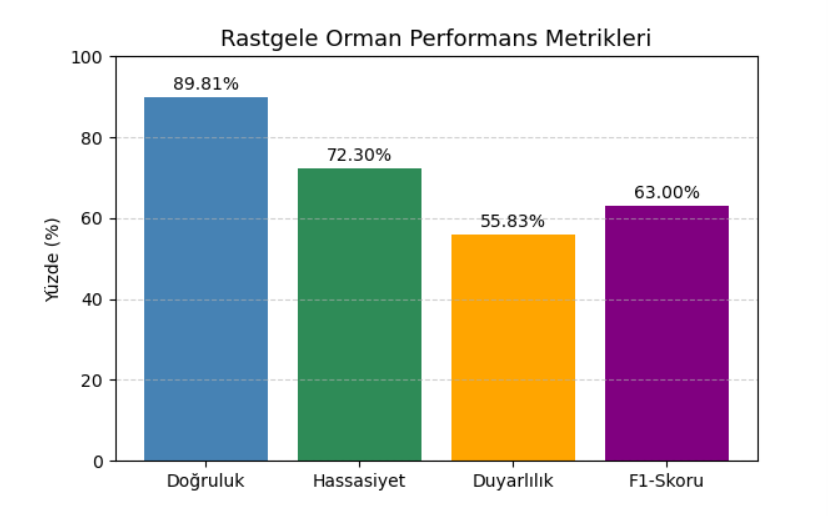
Bu çalışmalardan elde edilen bulgular doğrultusunda, çevrimiçi alışveriş davranışlarının doğru modellenebilmesi için veri ön işleme, algoritma seçimi ve değerlendirme metriklerinin dikkatli bir şekilde belirlenmesi gerektiği ortaya konulmuştur. Ayrıca, literatürde en yüksek başarıya ulaşan modellerin genellikle Rastgele Orman gibi topluluk yöntemleri olduğu görülmektedir.

**3. Bulgular ve Tartışma**

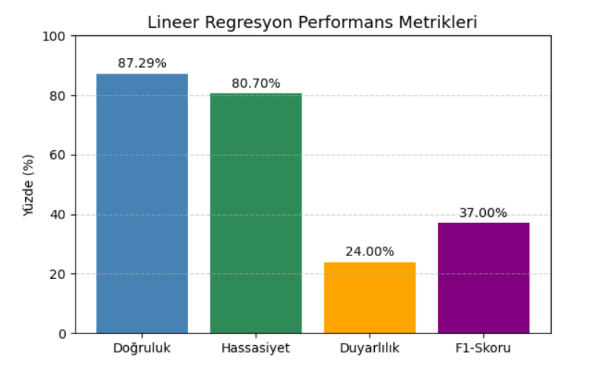
Bu bölümde, çevrimiçi alışveriş yapan kullanıcıların satın alma niyetlerini tahmin etmek üzere uygulanan beş farklı makine öğrenmesi algoritmasının test verisi üzerindeki performansları değerlendirilmektedir. Model başarıları; doğruluk (accuracy), hassasiyet (precision), duyarlılık (recall), F1-skoru ve ROC eğrisi analizlerine dayanarak karşılaştırılmıştır.

Her modelin çıktıları doğrultusunda oluşturulan çubuk grafiklerde doğruluk, hassasiyet, duyarlılık ve F1-skoru değerleri görsel olarak sunulmuştur. Yapılan değerlendirmeler sonucunda, sınıflandırma başarısında algoritmalar arasında belirgin farklar gözlemlenmiştir.

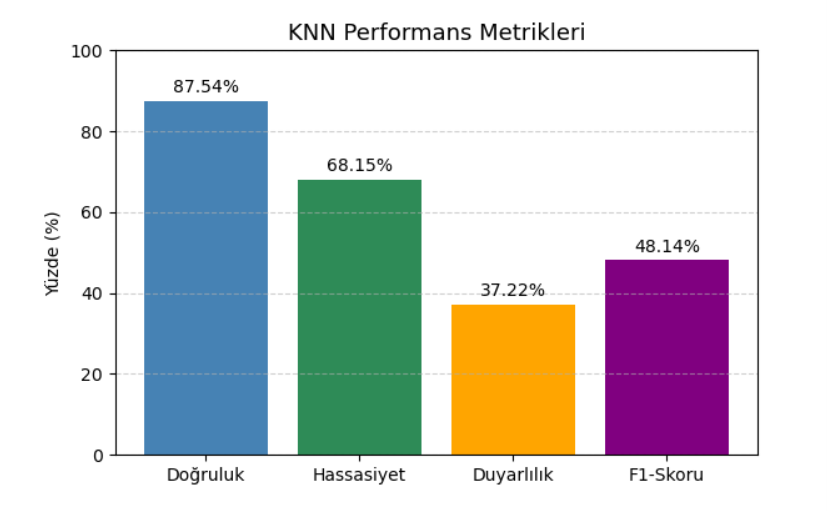
Rastgele Orman (Random Forest) algoritması, %89.5 doğruluk ve 0.62 F1-skoru ile genel olarak en başarılı performansı sergilemiştir. Satın alan kullanıcıları ayırt etmede yüksek kesinlik sunarken, aşırı öğrenmeden kaçınması sayesinde dengeli bir modelleme sağlamıştır.

****

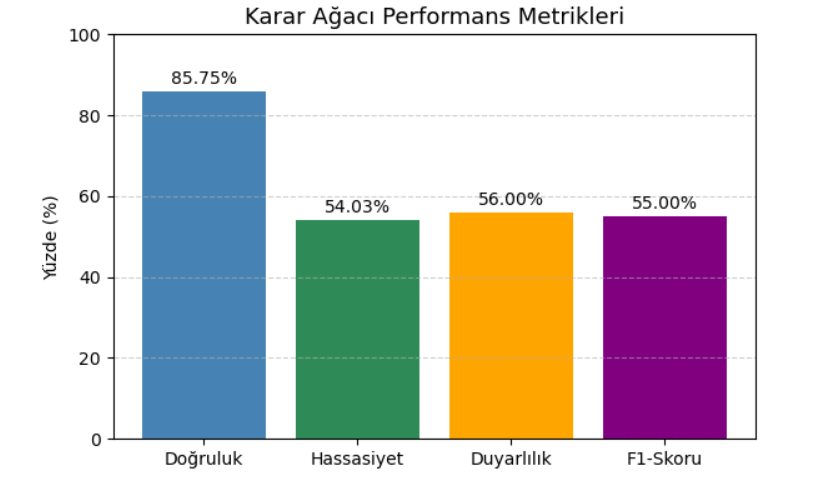
Lineer Regresyon, sınıflandırma problemi için eşik temelli bir yaklaşım ile uygulanmıştır. %87.3 doğruluk oranına sahip olan model, özellikle kesinlik (precision) açısından %80.7 ile oldukça yüksek bir başarı sağlamış; ancak duyarlılık değeri %24 ile düşük kalmış, bu da birçok satın alan kullanıcının yanlış sınıflandırılmasına neden olmuştur.

****

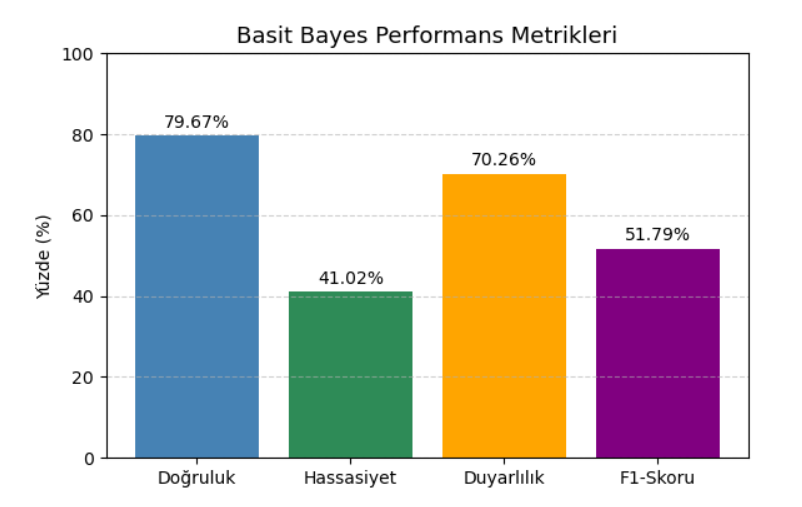
K-En Yakın Komşu (KNN) algoritması %87.5 doğrulukla güçlü bir genel performans göstermiş, fakat düşük duyarlılık (%37.2) nedeniyle satın alma davranışlarını yakalama konusunda yetersiz kalmıştır.

****

Karar Ağacı, %85.5 doğruluk ve 0.54 F1-skoru ile dengeli bir modelleme sunmuş ancak özellikle dengesiz veri yapılarında, pozitif sınıfı yeterince temsil edememiştir.

****

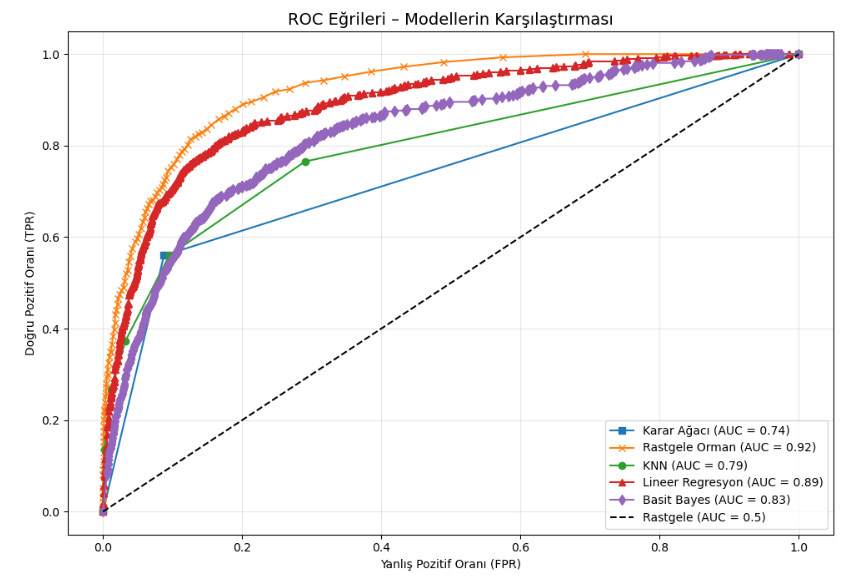
Basit Bayes (Naive Bayes) algoritması, %79.7 doğrulukla en düşük genel başarıya sahip model olmasına karşın, duyarlılık açısından %70.3 ile en yüksek değeri elde etmiştir. Bu durum, modelin pozitif sınıfı (satın alan kullanıcıları) fazla tahmin etme eğiliminde olduğunu göstermektedir.

****

Modellerin sınıflar arası ayrım performansları ROC eğrileriyle değerlendirilmiştir. En yüksek AUC değeri, eğrisel alanda en iyi performansı sağlayan Rastgele Orman modeline aittir. Bu durum, modelin hem duyarlılığı hem de özgüllüğü yüksek şekilde dengelediğini göstermektedir.

Lineer Regresyon, sürekli değerler üzerinden yapılan tahminlerin eşiklenmesi ile sınıflandırmaya uygun hale getirilmiş, ROC eğrisi üzerinde diğer modellere göre daha sınırlı bir ayrım gücü sergilemiştir. Bu durum, doğrusal modelin sınıf ayrımını başarıyla gerçekleştirmede sınırlı kaldığını ortaya koymuştur.

Basit Bayes modeli, pozitif sınıfı yakalama konusunda gösterdiği yüksek duyarlılığa rağmen, yanlış pozitif oranının yüksek olması nedeniyle AUC değeri açısından geride kalmıştır. KNN ve Karar Ağacı modelleri ise orta düzeyde AUC değerleriyle istikrarlı ancak sınırlı başarı sağlamıştır.

****

| **Algoritma** | **Doğruluk** | **F1-score** |
| --- | --- | --- |
| Rastgele Orman | |  | | --- | |  |   0.8981 | 0.6300 |
| Lineer Regresyon | 0.8729 | 0.3700 |
| KNN | 0.8754 | 0.4814 |
| Karar Ağacı | 0.8575 | 0.5500 |
| Basit Bayes | 0.7967 | 0.5179 |

**SONUÇ VE ÖNERİLER**

Bu çalışmada, çevrimiçi alışveriş yapan kullanıcıların oturum bazlı verileri kullanılarak, satın alma niyetlerinin tahmin edilmesi amaçlanmıştır. UCI Machine Learning Repository'den temin edilen ve çeşitli kullanıcı etkileşimlerini içeren "Online Shoppers Purchasing Intention Dataset" veri seti kullanılarak, Karar Ağacı, Rastgele Orman, K-En Yakın Komşu, Lineer Regresyon ve Basit Bayes algoritmaları uygulanmış ve performansları karşılaştırılmıştır.

Model başarılarının değerlendirilmesinde doğruluk, hassasiyet, duyarlılık ve F1-skoru metrikleri ile ROC eğrileri kullanılmıştır. Elde edilen sonuçlara göre, Rastgele Orman algoritması %89.8 doğruluk ve 0.63 F1-skoru ile en başarılı sonuçları vermiştir. Bu algoritma, özellikle dengesiz veri yapısında hem satın alan hem de almayan kullanıcıları dengeli şekilde sınıflandırma konusunda güçlü bir performans sergilemiştir.

Karar Ağacı ve KNN algoritmaları, doğruluk açısından kabul edilebilir sonuçlar üretse de özellikle "satın alan" sınıfını tahmin etmede düşük duyarlılık oranlarına sahiptir. Lineer Regresyon, yüksek hassasiyet göstermesine rağmen düşük duyarlılık nedeniyle birçok satın alan kullanıcıyı yanlış sınıflandırmıştır. Basit Bayes algoritması ise düşük genel doğruluk oranına rağmen satın alan sınıfı yakalama konusunda en yüksek duyarlılığı (%70.3) göstermiştir.

Veri setinin sınıf dengesizliği içerdiği dikkate alındığında, yalnızca doğruluk oranı ile değerlendirme yapmak yanıltıcı olabilir. Bu nedenle F1-skoru ve ROC eğrisi gibi daha bütüncül ölçütlerin dikkate alınması, sınıflandırma modellerinin gerçek başarısını ortaya koymak açısından daha sağlıklı bir yaklaşımdır.

Bu çalışmada kullanılan algoritmaların bazı ayarları (örneğin komşu sayısı, ağaç derinliği gibi) değiştirilerek daha iyi sonuçlar elde edilebilir. Bu tür ayarları denemek performansı artırabilir.

Veride satın alan kullanıcı sayısı az olduğu için, bu dengesizliği giderecek yöntemler kullanılırsa özellikle satın alma yapanları daha doğru tahmin etmek mümkün olabilir.

Kullanılan bazı veriler üzerinde değişiklikler yapılarak (örneğin benzer sütunları birleştirme ya da yeni bilgiler üretme gibi) modellerin daha iyi çalışması sağlanabilir.

Bu çalışma bir tahmin modeli sunuyor, ama bu modeli gerçek e-ticaret sitelerine entegre ederek öneri sistemleri veya müşteri hizmetleri gibi alanlarda pratik uygulamalar geliştirilebilir.

Gelecekte daha gelişmiş yöntemler (örneğin sinir ağları gibi) ya da zaman sıralı analizler (zaman içinde kullanıcı davranışı) ile daha hassas tahminler yapılabilir.

**KAYNAKÇA**

[1] D. Torres and L. K. Cepeda, "Machine Learning for Predicting Online Shoppers’ Purchase Intentions," *ResearchGate*, Apr. 2024.  
[2] M. S. Satu and S. F. Islam, "Modeling Online Customer Purchase Intention Behavior Applying Different Feature Engineering and Classification Techniques," *Discover Artificial Intelligence*, vol. 3, no. 36, 2023.  
[3] C. Okan Sakar, S. Olcay Polat, Mete Katircioglu, and Yomi Kastro, "Real-time prediction of online shoppers’ purchasing intention using multilayer perceptron and LSTM recurrent neural networks," *Neural Computing and Applications*, vol. 31, no. 10, pp. 6893-6908, 2019.

[4] I. D. Makash,  
*"Online Shoppers Purchasing Intention Dataset,"*  
UCI Machine Learning Repository / Kaggle, 2022.

[5] A. Dey,  
*"Machine Learning Algorithms: A Review,"*  
*International Journal of Computer Science and Information Technologies*, vol. 7, no. 3, 2016