



Research Article

ANALYSIS OF THE FACTORS AFFECTING THE CONSUMER IN ONLINE SHOPPING WITH DECISION TREES

KARAR AĞAÇLARI İLE İNTERNET ALIŞVERİŞLERİNDE TÜKETİCİYİ ETKİLEYEN FAKTÖRLERİN ANALİZİ

Burcu ÖZCAN ^{1,*} | Can TURNA ²

1 Doç. Dr., Kocaeli Üniversitesi, Türkiye, ORCID: 0000-0003-0820-4238

2 Endüstri Mühendisi, Türkiye, ORCID: 0000-0002-2294-7364

Article Info:

Received : Feb 18, 2021

Revised : July 7, 2021

Sep 15, 2021

Accepted : Sep 16, 2021

Keywords:

E-Commerce

Shopping

Consumer

Decision Trees

Sustainability

Anahtar Kelimeler:

E-Ticaret

Alışveriş

Tüketici

Karar Ağaçları

Sürdürülebilirlik

DOI: 10.46238/jobda.882832

ABSTRACT

The electronic commerce market, brought about by the increasingly widespread internet usage, has been a source of opportunity for both manufacturers and consumers. As a result of the shopping opportunities offered on the internet and its reliability day by day, it causes consumers to prefer shopping on the internet more. At this point, researchers examine consumer behavior using data mining methods. Data mining is the process of discovering meaningful, hidden and unpredictable patterns and rules from large databases. In the study, decision trees were created depending on the variables of gender, age and monthly income in order to examine the factors affecting the consumer in online shopping. Decision trees are a data mining method used to classify certain patterns and make predictions from large databases. As a result of the application, using the C5.0 Algorithm for the gender variable, it was concluded that the most important factor for consumers is consumer comments. Using the C&R Tree Algorithm for age variable and monthly income variable, it was concluded that the most important factor for consumers is free shipping.

ÖZ

Günümüzde giderek yaygınlaşan internet kullanımının beraberinde getirdiği elektronik ticaret pazarı hem üreticiler hem de tüketiciler açısından fırsat kaynağı olmuştur. İnternet üzerinden sunulan alışveriş imkânları ve günden güne güvenilirliğini kanıtlaması sonucunda tüketicilerin internet üzerinden alışverişi daha fazla tercih etmelerine sebep olmaktadır. Araştırmacılar bu noktada veri madenciliği yöntemlerini kullanarak tüketici davranışlarını incelemektedir. Veri madenciliği büyük veri tabanları içerisinden anlamlı, gizli kalmış ve önceden tahmin edilemeyen örüntüler ve kurallar keşfetme sürecidir. Yapılan çalışmada internet üzerinden gerçekleştirilen alışverişlerde tüketiciyi etkileyen faktörleri incelemek amacıyla cinsiyet, yaş ve aylık gelir durumu değişkenlerine bağlı olarak karar ağaçları oluşturulmuştur. Karar ağaçları, büyük veri tabanlarından belirli kalıplarda sınıflandırmalar yapmak ve tahminlerde bulunmak amacıyla kullanılan bir veri madenciliği yöntemidir. Uygulama sonucunda cinsiyet değişkeni için C5.0 Algoritması kullanılarak tüketiciler için en önemli faktörün tüketici yorumları olduğu sonucuna varılmıştır. Yaş değişkeni ve aylık gelir durumu değişkeni için ise C&R Tree Algoritması kullanılarak tüketiciler için en önemli faktörün ücretsiz kargo olduğu sonucuna varılmıştır.

© 2021 JOBDA All rights reserved

*** Corresponding Author.**

E-mail: burcu.ozcan@kocaeli.edu.tr (B. ÖZCAN)

1 | GİRİŞ

Veri madenciliğinin meydana gelişi, veri yığınlarının çok fazla yer tutması ile çok sayıdaki verilerin önemli ve yararlı bilgilere çevrilmesi ihtiyacına dayanmaktadır. Veri madenciliği; banka ve kredi kartı işlemleri, pazarlama, sigortacılık, risk yönetimi gibi birden çok alanda kullanılabilir olması sebebiyle, son dönemlerde, veri tabanı kullanıcılarının ve araştırmacılarının önemli derecede dikkatlerini çekebilmektedir (Özguven, 2011). Veri madenciliği, mevcut verileri analiz ederek yeni kalıplar ve belirli modellerin ortaya çıkarılmasını sağlamaktadır. Bir başka deyişle verilerdeki kalıpları keşfetme süreci olarak tanımlanabilmektedir (Witten, Frank, Hall ve Pal, 2016). Veri madenciliği, makina öğrenme, istatistiksel ve veri tabanlarıyla ilgili konulardaki yöntemleri birleştirip büyük veri tabanlarından yararlı ve önemli bilgileri çıkartabilmesine olanak tanımaktadır. Veri madenciliği, yıllardır geliştirilmekte olan çeşitli yöntemleri içermektedir (Emre ve Selçukcan Erol, 2017).

Veri madenciliği, pazarlama, eğitim, bankacılık, imalat, nakliye, perakende, sağlık, sigorta, internet ve kamu gibi birçok alanda uygulanabilmektedir. Veri madenciliği, istatistiksel ve matematiksel yöntemlerle örüntü teknolojilerinin kullanılıp depolanma yerlerinde sıkışan çok sayıdaki verinin elenmesi ile anlamlı örüntülerin bulunması sürecidir (Özden ve Üner, 2019). Veri madenciliği, çok sayıdaki veriden, anlamlı örüntüler ile yöntemler keşfetme süreçlerini kapsar. En basit tanımı ile veri madenciliği, veri içerisindeki yeni, saklı kalan ya da beklenilmeyen örüntülerin bulunabilmesi için kullanılan hareketlerin toplamıdır (Avunduk ve Kızgın, 2020).

Bu çalışmada internet üzerinden yapılan alışverişlerde tüketicileri etkileyen faktörler ele alınmış olup cinsiyet, yaş ve aylık gelir durumu değişkenlerine göre karar ağacı algoritmaları ile analizler yapılmış ve yorumlanmıştır.

2 | LİTERATÜR TARAMASI

Şeker (2020) tarafından gerçekleştirilen çalışmada, tüketicilerin internet üzerinden yapmış oldukları alışverişlerde davranışlarına yönelik bir çalışma ele alınmıştır. Bu çalışmada hedeflenen, tüketicilerin internet kullanım alışkanlıkları, internet üzerinde alışveriş yapma veya yapmama durumları, alışveriş için ayırdıkları bütçe, alışveriş yapma sıklıkları gibi detaylar ortaya konarak tüketicilerin internet üzerinden yaptıkları alışveriş davranışlarının demografik özelliklerine göre ne şekilde farklılıklar gösterdiği incelenmiştir. Çalışma için çevrimiçi anket formu hazırlanmıştır ve sosyal medya araçları ile

katılımcıları ulaşılmıştır. Katılımcılar 18 yaş üzeri ve internet kullanabilen tüketicilerden oluşmaktadır. Elde edilen veriler SPSS 21 programına aracılığıyla incelenmiştir. İncelemeler sonucunda tüketicilerin yaşları ilerledikçe internet üzerinden alışveriş yapma sıklıkları azalmaktadır. Bunun yanında eğitim durumu arttıkça da internet üzerinden alışveriş yapma sıklığı artmaktadır. İnternet kullanan katılımcıların büyük bölümünün internet üzerinden alışveriş yaptığı görülmüştür. İnternet üzerinden alışveriş yapmalarının en önemli nedeni fazla sayıda seçenek olması rahatlığı olduğu görülmüştür. Ayrıca yapılan çalışmada cinsiyet, medeni durum ve gelir düzeyinin davranışları etkilemediği görülmüştür. İnternet üzerinden alışveriş yapmayan müşteriler incelendiğinde ise en büyük neden ürünleri görmek ve denemek istemelerinden kaynaklanmaktadır. Çalışmada bu tür tüketicilerin iptal/iade konusunda sorun yaşayacaklarını ve güvenlik sebebiyle kimlik ve kart bilgilerini paylaşmak istemedikleri görülmüştür.

Özden ve Üner (2019) tarafından yapılan çalışmada, son yıllarda internet üzerinden yapılan alışverişlerin artmasıyla tüketicilerin etkileşim derecelerinin demografik özellikleri ile ilişkileri incelenmiştir. Araştırmada 18 yaş ve üzeri, internet alışverişini yapan ve Bolu ile Samsun illerinde yaşayan 194 tüketici ele alınmıştır. Veriler IBM SPSS Statistics 21 programı ile incelenmiştir. Araştırma sonucunda kadınların erkeklere göre, bekar tüketicilerin evlilere göre tüketici etkileşiminin daha fazla olduğu görülmüştür. 18-25 yaş grubundaki tüketicilerin diğer yaş grubundakilere göre internet üzerinden yaptıkları alışverişlerde tüketici etkileşiminin daha fazla olduğu ayrıca tüm yaş gruplarında da benzer durumların olduğu görülmüştür. Ayrıca eğitim durumu doktora düzeyinde olan tüketicilerin diğer tüketicilere göre internet üzerinden yapılan alışverişlerde tüketici etkileşiminin daha az olduğu, eğitim durumu lise olan tüketicilerde ise daha fazla olduğu görülmüştür. Eğitim durumunu daha yüksek olan kişilerin daha yüksek gelire sahip olduğu varsayılırsa bu tüketiciler müşteri etkileşiminden ziyade daha pahalı ve yüksek kaliteli ürünlere yöneldiği düşünülebilir. Geliri 2001-5000 ₺ arasında olan tüketicilerin, internet üzerinden yaptıkları alışverişlerde tüketici etkileşiminin daha fazla olduğu görülmüştür.

Bilgiler (2019) tarafından, Y kuşağının internet üzerinden gerçekleştirilen alışverişlere karşı tutumlarının demografik özelliklerine göre değişimi üzerine bir araştırma yapılmıştır. Araştırmada Y kuşağına uygun 547 tane İletişim Fakültesi üniversite öğrencisi kullanılmıştır. Veriler yüz yüze anket ile toplanmıştır. Araştırmada Silkü tarafından 2009 yılında geliştirilen "İnternette Alışverişe Yönelik

Tutum Ölçeği" (İAYTÖ) kullanılmıştır. Katılımcıların çoğunluğu kadın, 18-21 yaş arası, Gazetecilik bölümü öğrencisi, 501-1000 ₺ arası aylık gelire sahip ve 2.sınıf öğrencilerinden oluşmaktadır. Büyük çoğunluğu internette alışveriş yapan kişiler oluşturmuştur. İnternette alışveriş yapma nedenleri; Fiyat uygunluğu, çeşitlilik, kampanyalar şeklinde sıralanmıştır. İnternet üzerinden alışveriş yapmamanın nedenleri ise güvensizlik, ürünü görememe, ürün kalitesini bilemem gibi cevaplar sıralanmıştır. Katılımcıların büyük çoğunluğu giyecek ve ayakkabı almaktadır. Y kuşağı ödemelerini çoğunlukla kredi kartı ile gerçekleştirmektedir.

Eren ve Gödekmerdan Önder (2020) tarafından yapılan çalışmada, tüketicilerin internet üzerinden yaptıkları alışverişlerdeki deneyimlerinin belirlenmesi amaçlanmıştır. Çalışmada Z kuşağı olarak adlandırılan 18-20 yaş aralığında ve Ankara'da yaşayan 464 birey ele alınmıştır. Veri toplamak amacıyla katılımcılara anket uygulanmıştır. Katılımcıların çoğunluğu kadın, her gün alışveriş sitelerini ziyaret eden katılımcıların çoğunlukla 1-2 saat vakit geçirmekte, ayda bir internet üzerinde alışveriş yapmakta ve aylık 250 ₺ harcama yapmaktadır. Elde edilen veriler SPSS 20 ve Smart-PLS programlarında incelenmiştir. Araştırmada alışveriş sitelerinin tasarımının e-tatmin üzerinde bir etkisi olmadığı görülmüştür. Sipariş teslimat süresinin e-tatmin üzerinde etkisi olduğu görülmüştür. İnternet üzerindeki alışveriş mağazalarının tüketicilerle kurduğu iletişimin e-tatmin üzerinde bir etkisi olmadığı görülmüştür. İnternet üzerinden yapılan alışverişlerin satın alırken sağladığı kolaylıklar, müşterilerin gizliliklerine verilen önem ve yapılan promosyonların da e-tatmin üzerinde olumlu etkiye sebep olduğu görülmüştür. İnternet üzerinden yapılan alışveriş deneyimlerine ilişkin tatminin sadakat üzerinde etkisi olduğu da araştırma ile doğrulanmıştır.

Yaslıdağ ve Ferhadi (2016) tarafından yapılan çalışmada, Azerbaycan'daki müşterilerin internet üzerinden yapmış oldukları alışveriş davranışlarının incelenmesi amaçlanmıştır. Araştırma için Azerbaycan'da bulunan 400 adet katılımcıdan yüz yüze yapılan anketlerle veriler toplanmıştır. Toplanan veriler SPSS programına girilmiş ve ortaya çıkan sonuçlar ile değerlendirmeler yapılarak sonuca bağlanmıştır. Araştırma sonucunda müşterilerin internette yaptıkları alışverişlerde teslimata önem verdikleri görülmüştür. Diğer taraftan alışveriş sitelerinden gelen bilgilendirme mesajlarının müşteriler tarafından önemsenmediği görülmüştür. Ayrıca internette alışveriş için daha fazla vakit geçiren tüketicilerin ve alışveriş sitelerini takip eden müşterilerin fikirlerinin daha çok önemsenmesi gerektiği kabul edilmiştir. Müşterinin internet

üzerinden fazlaca alışveriş yapması sürekli aynı siteden alışveriş yapacağı anlamı çıkarılmaması gerektiği görülmüştür.

Sayılı ve Büyükköroğlu (2013) tarafından yapılan çalışmada, müşterilerin internet üzerinden gerçekleştirdikleri alışverişlerde gıda kategorisinde yer alan ürünleri satın almaya ait tutumlarını etkileyen unsurların incelenmesi amaçlanmıştır. Bu çalışma için Logit modeli kullanılarak Tokat'ta yaşayan 277 tüketici tarafından toplanan veriler kullanılmış ve model oluşturulmuştur. Tüketicilerin çoğunluğunun internet üzerinden gıda ürünleri alımını uygun bulmadıkları görülmüştür. Müşterilerin internet üzerinden yapılan alışverişlerde güvenlik faktöründen dolayı güvenmediği belirlenmiştir. Tüketicilerin çoğunun gıda ürünlerinin alınmasının uygun bulmadıkları görülmüştür. Tüketicilerin internet üzerinden gıda ürünleri almasını etkileyen faktörlerden olan aylık gıda harcaması, eğitim düzeyi ve güvenin gıda ürünleri alımını pozitif yönde etkilediği görülmüştür. Aylık gıda ürünlerine olan harcamalar arttıkça tüketiciler internet üzerinden gıda ürünleri satın almaya ait olumlu bir tutum sergilemektedirler. Sonuç olarak internet üzerinden gıda ürünleri satın alımının artırılması için tanıtım faaliyetleri artırılmalı ve müşterilere güven ortamı oluşturulmalıdır.

Sönmez (2019) tarafından yapılan yüksek lisans tezinde, internet üzerinden alışveriş yapan tüketicilerin satın alma tarzlarının kişilik tipleri ile ilişkisinin incelenmesi amaçlanmıştır. Veri toplamak amacıyla internet üzerinden anket çalışması yapılmıştır. Sakarya'da yaşayan ve 18 yaş üstü tüketiciler ele alınmıştır. Toplanan veriler SPSS 20.0 istatistikî analiz programı ile analiz edilmiştir. Çalışma sonucunda katılımcıların çoğunun yılda birkaç defa internet üzerinden alışveriş yaptıkları belirlenmiştir. Tüketicilerin interneti öncelikle sosyal medya için daha sonra ise alışveriş amacıyla kullandıkları belirlenmiştir. İnternet üzerinden gerçekleştirilen alışverişlerde en çok karşılaşılan sorun yanlış veya hatalı ürün olduğu görülürken, en az karşılaşılan sorunun ise ürün fiyatının belirtilenden daha çok çıkması durumudur. İnternet üzerinden gerçekleşen alışverişlerin en önemli avantajı zaman tasarrufu ve fiyat uygunluğudur. Çalışmada kişilik özellikleri ile satın alma alışkanlıkları arasında anlamlı bir ilişki olduğu görülmüştür.

Gençyürek Erdoğan (2020) tarafından gerçekleştirilen çalışmada, pandemi sürecinde hızla yaygınlaşan e-ticaret uygulamaları, değişen tüketim alışkanlıkları, e-ticaret harcamaları ve reklam yatırımları incelenmiştir. Çalışma sonucunda pandemi sürecinde market-gıda sektöründe bir artış görülmüştür. Turizm, eğlence merkezleri, restoran ve kafeler bu süreçten en çok

etkilenen sektörlerdir. Bu dönemde en sık görülen tüketici profili stoklayan tüketici profilidir. Tüketiciler ihtiyacından fazlasını alarak depolama yapmış ve bu şekilde kendini güvende hissetmiştir. E-ticaret sektöründe yaşanan artışın temel nedeni tüketicilerin pandemi sürecinde evde fazla zaman geçirmeleridir.

T. Bardak, Avcı, Kayahan ve S. Bardak (2018) tarafından yapılan çalışmada mobilya alımında mağazadan veya internet üzerindeki mağazaların seçimini etkileyen faktörlerin incelenmesi amaçlanmıştır. Bu çalışma için 217 kişiden anket yöntemiyle veriler toplanmıştır. Elde edilen veriler, Weka veri madenciliği platformu kullanılarak analiz edilmiştir. Birliktelik kurallarını elde etmek için Apriori Algoritması kullanılmıştır. Sonucunda 3064 adet kural oluşturulmuştur. Kurallara göre evli ve çocuğu olanlar daha çok internet üzerinden alışveriş yapmaktadır. 35-44 yaş aralığında ve çocuğu olan tüketiciler internet üzerindeki alışverişlerde en çok fiyata önem vermektedir. Sonuç olarak internet üzerinden gerçekleştirilen alışverişlerde evli ve çocuk sahibi olmak önemli bir unsurdur. Ayrıca internetten alışverişlerde en çok ürün fiyatının etkili olduğu belirtilmiştir.

Ağaç, Dengin Sevinir ve Yılmaz (2018) tarafından yapılan çalışmada, tüketicilerin internet üzerinden gerçekleştirdikleri giyim ürünleri alışverişlerinde yaşanan sorunlar tespit edilmek istenmektedir. Araştırma için Ankara'da bulunan, internet üzerinden alışveriş yapan ve bir problemle karşılaşmış ve çalışan 400 kişiden(200 kadın ve 200 erkek) anket ile veriler toplanmıştır. Toplanan veriler SPSS programıyla analiz edilmiştir. Tüketicilerin internet üzerinden alışveriş yapmama nedenlerinin başında alışveriş sitelerinde bulunan ürün bilgilerinin yeterli olmaması gelmektedir. Tüketicilerin, satın aldıkları ürünlerin lojistik ücretinin kendilerine yansıtılması konusunda şikayetçi oldukları görülmüştür. Tüketicilerin çoğunlukla yaşadıkları problemlerden birisi beden alternatiflerinin yetersiz olmasıdır. Çoğunlukla kadın tüketicilerde yaşanan sorun, alınan ürün ile site üzerinde yer alan ürüne ait görselin ve bilgilerin ürün ile uyuşmamasıdır. Sanal manken ve prova uygulamasının olmaması alışveriş sitelerinde eksik olarak görülen konular arasında yer almaktadır.

Demirdöğmez, Gültekin ve Taş (2018) tarafından yapılan çalışmada E-ticaret sektörünün ülkemizde yıllar içerisinde gelişmesini, özellikle son on yıl içerisindeki artışı incelenmiştir. E-ticaret sektörü etkisinin ortaya çıkabilmesi için perakendecilerin ve KOBİ'lerin e-ticaret sektöründe yer alması gerekmektedir. Aynı zamanda, algı ve teknik altyapının iyileştirilmesi, sektör ve idari yapı arasında etkin

iletişimin sağlanması önem arz etmektedir. Türkiye'nin e-ticaret potansiyelini tam anlamıyla etkinleştirebilmesi, ekonomisine ve ihracatına önemli katkılar sağlayacaktır sonucuna varılmıştır. E-ticaret sektörünün potansiyelini arttırmak için çeşitli önermeler yapılmıştır.

Kolata (2014) tarafından gerçekleştirilen yüksek lisans tezinde, internetten alışverişlerin incelenmesi, tüketicilerin bakış açıları, internetten ürün satın alma veya almama sebepleri internetten alışverişlerin tüketici eğilimlerinde oluşturmuş olduğu farklılık ve etkilerin tespit edilmesi amaçlanmıştır. Araştırma için anket ile elde edilen Konya'da yaşayan ve 18 yaş üzeri 273 kişiye ait veriler kullanılmıştır. Elde edilen veriler SPSS 15 programı aracılığıyla analiz edilmiştir. Tüketicilerin tüm siteleri gezerek planlı bir şekilde en uygun ürünü seçtikleri görülmüştür. Uygun fiyat, kalite, ürün bilgisi, ödeme yöntemi kolaylığı, kampanyalar ve iade süreci tüketicilerin internetten alışveriş yapmalarına neden olmaktadır. Alınan ürün ile teslim alınan ürünün uyuşmaması ve güvenlik problemleri müşterilerin internet üzerinden alışveriş yapmaları konusunda kaygı oluşturmaktadır.

3 | VERİ MADENCİLİĞİ KAVRAMI

Veri Madenciliği, büyük hacimli veri tabanları içerisinde gizli kalmış, faydalı ve anlamlı olabilecek bilgilerin ayrıştırılması ve dönüştürülmesi sürecidir (Doğan, 2017). Elde bulunan verilerin çok büyük ve karmaşık olması sebebiyle klasik yöntemler kullanılarak çözümlenmesi çok uzun vakit alabilmekte, iş yüküne sebep olabilmekte, yetersiz kalabilmekte hatta çoğu zaman istenen sonuca ulaşamamasına sebep olmaktadır (Demir ve Dinçer, 2020). Veri madenciliği büyük miktardaki verilerin analizlerini gerçekleştirebilmek için geliştirilmiş olan güçlü teknikler içerir.

Veri Madenciliği, satış tahminleri, kaynak planlama, pazarlama stratejileri, bankacılık işlemleri gibi birçok alanda uygulanabiliyor olması sebebiyle, önemli ölçüde dikkatleri üzerine çekmekte ve kullanıcılar tarafından ilgi görmektedir. Makine öğrenme, istatistik ve veri tabanı konularındaki metotları birlikte ele alarak büyük hacimli verilerden anlamlı ve gizli kalmış bilgilerin görülebilmesine olanak sağlamaktadır (Ching ve Michael, 2002). Tek bir adet veri madenciliği metodu olmamakla birlikte daha çok bu metotların bir arada kullanılması söz konusudur.

Veri tabanlarında Bilgi Keşfi olarak da tanımlanan veri madenciliği bir keşif alanıdır. Büyük miktardaki verilerden potansiyel olarak yararlı ve yeni bilgiler elde etmeye yaramaktadır. Son yıllarda araştırmacılar tarafından veri madenciliğine olan ilgi artmakla birlikte

büyük miktardaki verilerden benzersiz keşifler yapmak için yeni yöntemler geliştirilmektedir (Baker, 2010).

Özet olarak veri madenciliği, büyük hacimli ve ham verileri kullanarak; anlamlı, faydalı olabilecek, gizli kalmış ve önceden tahmin edilemeyen örüntüler ve kurallar keşfetme sürecidir.

4 | KARAR AĞAÇLARI VE KARAR AĞACI ALGORİTMALARI

Karar ağaçları, tahmin etmek ile sınıflandırmak amacıyla sıklıkla tercih edilen bir veri madenciliği yöntemidir (Quinlan, 1986). K-en yakın komşu algoritmaları, yapay sinir ağları ve regresyon gibi farklı yöntemler de tahmin ve sınıflandırma için kullanılabilirler. Fakat karar ağaçları, düşük maliyetli olması, anlaşılır olması, güvenilirliği ve kolaylıkla yorumlanabilmesi bakımından tercih edilmektedir (Çalış, Kayapınar ve Çetinyokuş, 2014).

Karar ağacındaki düğümler belirli bir özelliği test etmeyi içerir (Witten vd., 2016). Karar ağaçları yapısı itibarıyla kök, dallar ve yapraklardan oluşmakta ve incelendiğinde ağaca benzemektedir (Akman, 2010). Ağaç, tüm verileri içeren kök düğümlerle başlamaktadır ve aşağılara doğru bakıldığında büyük hacimli yığınları daha küçük hacimli kategorilere ayırarak dallandırmaktadır. Her düğüm en az iki dala ayrılmakta ve yeni bir soru yoksa dallanmalar sona ermektedir. Böylelikle bir sınıfı ifade eden yaprağa varılmaktadır (Oğuzlar, 2004; Pehlivan, 2006). Karar ağacındaki bir yol birden çok kez test edilerek en doğru sonuca varılmaya çalışılmaktadır (Brodley ve Utgoff, 1995).

C4.5 ve C5.0 Algoritmaları: C4.5 algoritması en yaygın kullanılan karar ağacı algoritmasıdır. C4.5 algoritmasının geliştirilerek oluşturulmuş olan C5.0 algoritması ise özellikle büyük ölçekli veri tabanları için kullanılmaktadır. C4.5 algoritmasına göre C5.0 algoritması daha hızlı olup, hafızayı çok daha verimli kullanmaktadır. C4.5 algoritmaları ve C5.0 algoritmalarının çıktıları benzer olsada C5.0 algoritması şeklen daha doğru karar ağaçları oluşturulmasını imkan tanımaktadır (Çalış vd., 2014).

CART Algoritması: CART ağacı dallanma ölçütü olarak Gini indeksinden yararlanır ve sürekli bölünerek büyür. Yeni bir bölünmenin gerçekleşmeyeceği durumda sondan başa doğru budama işlemi başlar. CART algoritması, en uygun dallanmanın belirlenmesinde C4.5 ve C5.0 algoritmaları gibi entropiden yararlanmaktadır (Demirel ve Giray, 2019).

CHAID Algoritması: CHAID, en iyi bölünmelerin belirlenebilmesi için ki-kare istatistiğinden yararlanan bir yöntemdir. Bölümlendirme işlemini gerçekleştirmek amacıyla kullanılan istatistiksel bir

yöntemdir (Çalış vd., 2014). CHAID algoritması kalabalık veri tabanlarının faydalı ve kolay anlaşılır bir biçimde yorumlanması için bağımsız değişkenler ile bağımlı değişkenleri alt dallara ayırmaktadır (Pehlivan, 2006).

QUEST Algoritması: Budama ve doğrudan durma kuralı gibi yöntemleri kullanması açısından İkili karar ağaçları yapısından yararlanan bir sınıflandırma algoritmasıdır. CHAID ve CART algoritması gibi değişken seçimi ve bölünmeyi eşzamanlı olarak değil eşzamanlı olarak yapar. QUEST algoritması, karar ağaçlarının dallara ayrılması durumunda önyargılı seçimlerin genelleştirilmesi ve maliyetlerinin daha aza indirilmesi amacı ile geliştirilmiştir (Çalış vd., 2014).

5 | UYGULAMA

Uygulamada, SPSS Clementine programı kullanılmaktadır. Program, veri madenciliği çözümlenmeleriyle istatistiksel algoritmaların ve yapay zeka temelli algoritmaların görsel bir program içinde sunulmasıdır.

Uygulanmış olan ankete ait veriler MS Excel üzerinde hazırlanmıştır. Bağımsız değişkenler, internet üzerinden alışverişi etkileyen faktörler ile ilgili 15 sorudan meydana gelirken, bağımlı değişkenler sırası ile yaş, cinsiyet ve aylık gelir durumudur.

Anket çalışmasına göre;

Ankete Kocaeli ilinde yaşayan 250 kişi katılmış olup, katılımcılardan %47,6'sı erkek, %52,4'ü ise kadındır.

Ankete göre katılımcıların %31,6'sı 18-22 yaş aralığı, %22'si 23-27 yaş aralığı, %18,4'ü 28-32 yaş aralığı, %15,2'si 33-39 yaş aralığı ve %12,8'i 40 ve üzeri yaş aralığındadır.

Ankete göre katılımcıların %23,2'si 0-1000, %16,4'ü 1000-2000, %16,8'i 2000-3000, %24'ü 3000-4000 ve %19,6'sı 4000 ve üzeri aylık gelir durumuna sahiptir.

Katılımcılar tarafından en çok tercih edilen üç önemli faktör; gizlilik/güvenlik, uygun fiyat ve tüketici yorumlarıdır. En az tercih edilen üç faktör ise beden seçeneklerinin yetersiz olması, taksit imkânı ve beden ölçülerinin uymaması düşüncesidir.

5.1. Karar Ağacı Algoritmalarının Uygulanması ve Karar Ağaçlarının İncelenmesi

Öncelikle cinsiyet, yaş ve aylık gelir durumu değişkenleri için ayrı ayrı SPSS Clementine programı ile C5.0 algoritması, C&R Tree algoritması, CHAID algoritması ve QUEST algoritması oluşturulmuş ve her değişken için dört algoritmaya ait doğruluk oranları hesaplanmıştır. Sonrasında en yüksek doğruluğa sahip

olan algoritmayla devam edilerek karar ağacı oluşturulmuş ve sonuçlar analiz edilmiştir.

5.1.1. Cinsiyet Değişkeni için Doğruluk Oranları

Cinsiyet değişkeninin SPSS Clementine Programı'nda C5.0, C&R Tree, CHAID ve QUEST algoritmalarına göre doğruluk oranları hesaplanmıştır. Tabloda görüleceği üzere, en yüksek doğruluk oranına C5.0 Algoritması ile ulaşılmıştır. Bundan dolayı karar ağaçlarıyla yorumlama yapılırken C5.0 Algoritmasının kullanılması sonucuna varılmıştır.

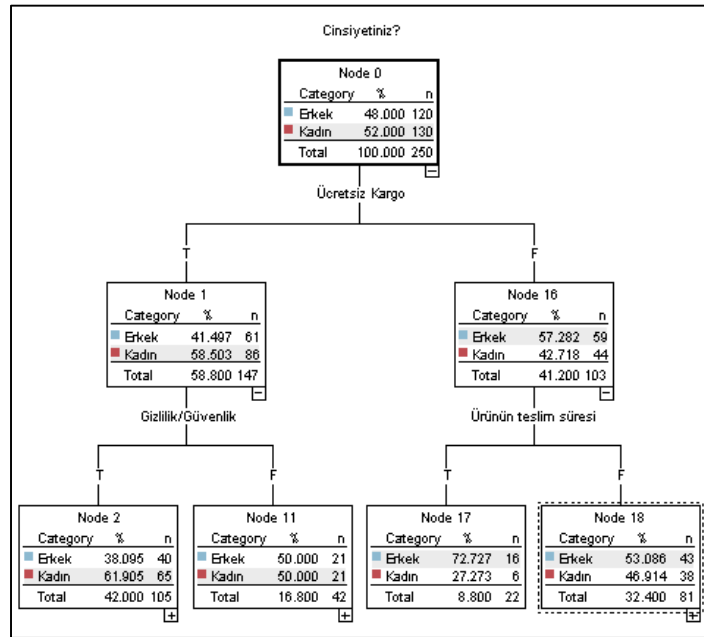
Tablo 1. Cinsiyet Değişkenine Ait Doğruluk Oranları

C&R Tree Algoritması			QUEST Algoritması		
Correct	154	61,6%	Correct	145	58%
Wrong	96	38,4%	Wrong	105	42%
CHAID Algoritması			C5.0 Algoritması		
Correct	145	58%	Correct	170	68%
Wrong	105	42%	Wrong	70	32%

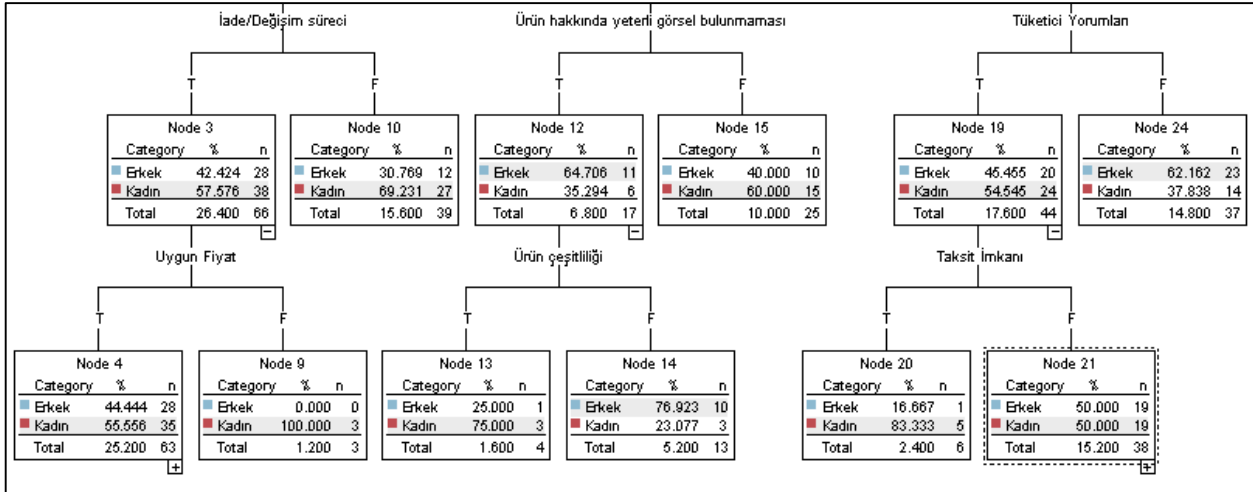
5.1.1.1. Cinsiyet Değişkeni için C5.0 Algoritması ile Hazırlanan Karar Ağacı

Cinsiyet değişkeni için C5.0 Algoritması ile oluşan karar ağacı incelendiğinde, karar ağacının öncelikle, ücretsiz kargo ile dallandığı görülmektedir. Yani, cinsiyet değişkeni için ağacın dallara ayrılmasındaki en önemli faktörün “ücretsiz kargo” olduğu görülmektedir. Ücretsiz kargonun önemli olduğu durumlarda “Gizlilik/Güvenlik” faktörü, önemsiz olduğu durumlarda ise “Ürünün Teslim Süresi” ön plana çıkmaktadır. Gizlilik/Güvenlik faktörünün önemli olduğu durumlarda “İade/Değişim Süreci” faktörü, önemsiz olduğu durumda “Ürün Hakkında Yeterli Görsel Bulunmaması” faktörünün ön plana çıktığı görülmüştür. Ürünün teslim süresinin önemsiz olduğu durumlarda “Tüketici Yorumları” faktörü ön plana çıkmaktadır. “İade/Değişim Süreci” faktörünün önemli olduğu durumlar “Uygun Fiyat” faktörü, “Ürün Hakkında Yeterli Görsel Bulunmaması” faktörünün önemli olduğu durumlarda “Ürün Çeşitliliği” faktörü, “Tüketici Yorumları” faktörünün önemli olduğu durumlarda da “Taksit İmkanı” faktörü önem arz etmektedir.

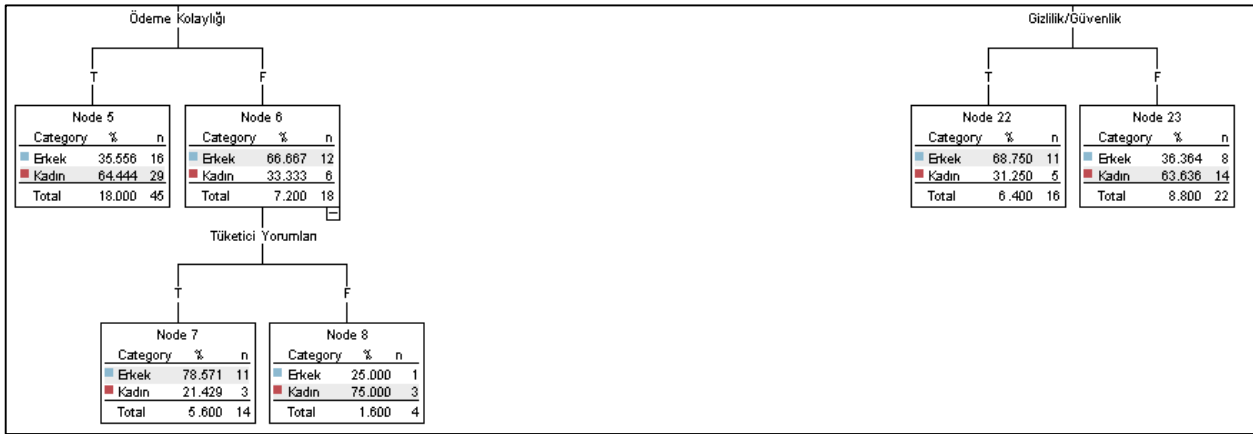
Şekil 1. Cinsiyet Değişkeni için Karar Ağacında Oluşan İlk Dal



Şekil 2. Cinsiyet Değişkeni İçin Karar Ağacında Oluşan İkinci Dal



Şekil 3. Cinsiyet Değişkeni İçin Karar Ağacında Oluşan Üçüncü Dal



5.1.2. Yaş Değişkeni için Doğruluk Oranları

Yaş değişkeninin SPSS Clementine Programı'nda C5.0, C&R Tree, CHAID ve QUEST algoritmalarına göre doğruluk oranları hesaplanmıştır. Tabloda görüleceği üzere, en yüksek doğruluk oranına C5.0 Algoritması ulaşılmıştır. Bu yüzden, karar ağacı ile yorumlama yapılırken C&R Tree Algoritmasının kullanılması sonucuna varılmıştır.

Tablo 2. Yaş Değişkenine Ait Doğruluk Oranları

C&R Tree Algoritması			QUEST Algoritması		
Correct	114	45,6%	Correct	79	31,6%
Wrong	136	54,4%	Wrong	171	68,4%
CHAID Algoritması			C5.0 Algoritması		
Correct	90	36%	Correct	120	48%
Wrong	160	64%	Wrong	130	52%

5.1.2.1. Yaş Değişkeni İçin C5.0 Algoritması ile Hazırlanan Karar Ağacı

Yaş değişkeni için C5.0 Algoritması ile oluşan karar ağacını incelediğinde, ağacın ilk başta, tüketici yorumlarından dallandığı görülmektedir. Yani, yaş değişkeni için ağacın dallara ayrılmasındaki en önemli faktörün "Tüketici Yorumları" olduğu görülmektedir. Tüketici yorumlarının önemli olduğu durumlarda "Taksit İmkkanı" faktörü, önemsiz olduğunda "Ürünün Teslim Süresi" faktörü ön plana çıkmaktadır. Taksit imkanı faktörünün önemli olduğu durumlarda "Uygun Fiyat" faktörü, önemsiz olduğu durumlarda "Beden Ölçülerinin Uymaması Düşüncesi" faktörü ön plana çıkmaktadır. Ürünün teslim süresi faktörü önemsiz olduğu durumlarda "Beden Seçeneklerinin Yetersiz Olması" faktörü ön plana çıkmaktadır.

5.1.3. Aylık Gelir Durumu Değişkeni için Doğruluk Oranları

Aylık gelir durumu değişkeninin SPSS Clementine Programı'nda C5.0, C&R Tree, CHAID ve QUEST algoritmalarına göre doğruluk oranları hesaplanmıştır. Tabloda görüleceği üzere, en yüksek doğruluk oranına C5.0 Algoritması ile ulaşılmıştır. Bu yüzden, karar ağacı ile yorumlama yapılırken C5.0 Algoritmasının kullanılması sonucuna ulaşılmıştır.

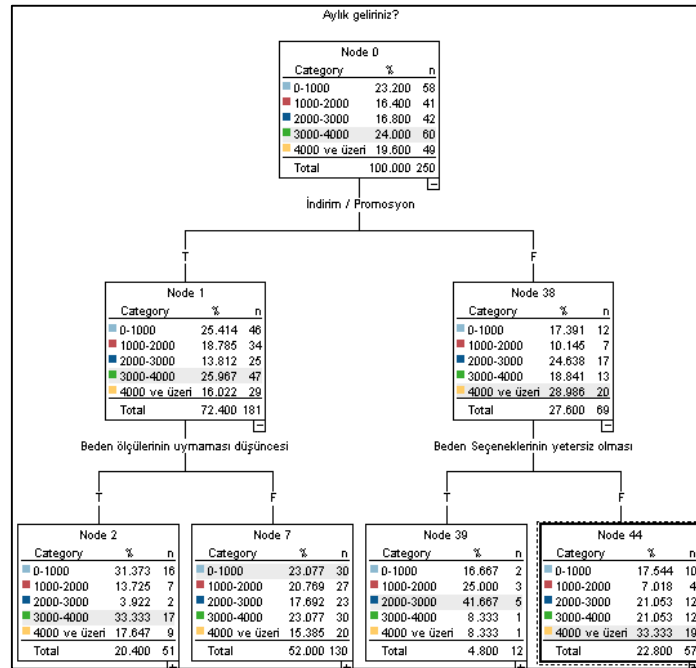
Tablo 3. Aylık Gelir Durumu Değişkenine Ait Doğruluk Oranları

C&R Tree Algoritması			QUEST Algoritması		
Correct	112	44,8%	Correct	58	23,2%
Wrong	138	55,2%	Wrong	192	76,8%
CHAID Algoritması			C5.0 Algoritması		
Correct	73	29,2%	Correct	125	50%
Wrong	177	70,8%	Wrong	125	50%

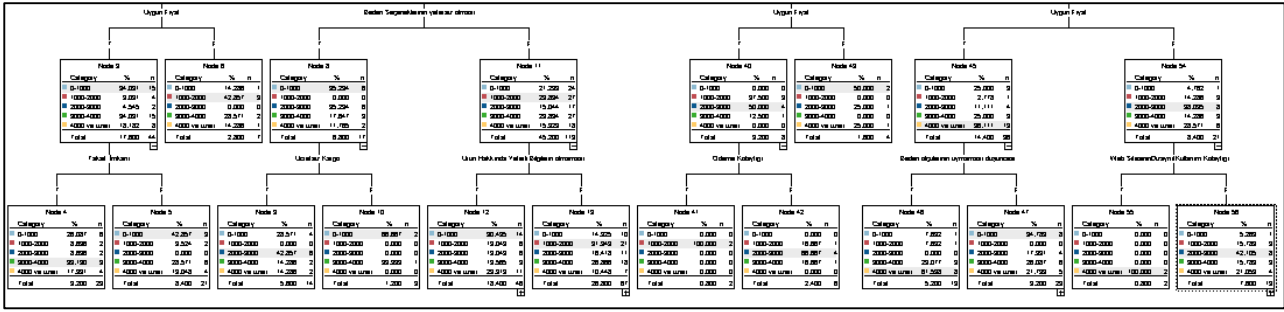
5.1.3.1. Aylık Gelir Durumu Değişkeni İçin C5.0 Algoritması ile Hazırlanan Karar Ağacı

Aylık gelir durumu değişkeni için C5.0 Algoritması ile hazırlanan karar ağacı incelendiğinde, ağacın öncelikle, indirim/promosyon faktöründen dallandığı görülmüştür. Yani, aylık gelir durumu değişkeni için ağacın dallanmasındaki en önemli değişkenin "İndirim/Promosyon" faktörü olduğu görülmektedir. İndirim/Promosyon faktörünün önemli olduğu durumlarda "Beden Ölçülerinin Uymaması Düşüncesi" faktörü ön plana çıkmaktadır. Beden Ölçülerinin Uymaması Düşüncesi faktörünün önemli olduğu durumlarda "Uygun Fiyat" faktörü, önemsiz olduğu durumlarda "Beden Seçeneklerinin Yetersiz Olması" faktörü ön plana çıkmaktadır. İndirim/Promosyon faktörünün önemsiz olduğu durumlarda "Beden Seçeneklerinin Yetersiz Olması" faktörü ön plana çıkmaktadır. Beden Seçeneklerinin Yetersiz Olması faktörünün önemli olduğu durumlarda "Uygun Fiyat" faktörü, önemsiz olduğu durumlarda yine "Uygun Fiyat" faktörü ön plana çıkmaktadır.

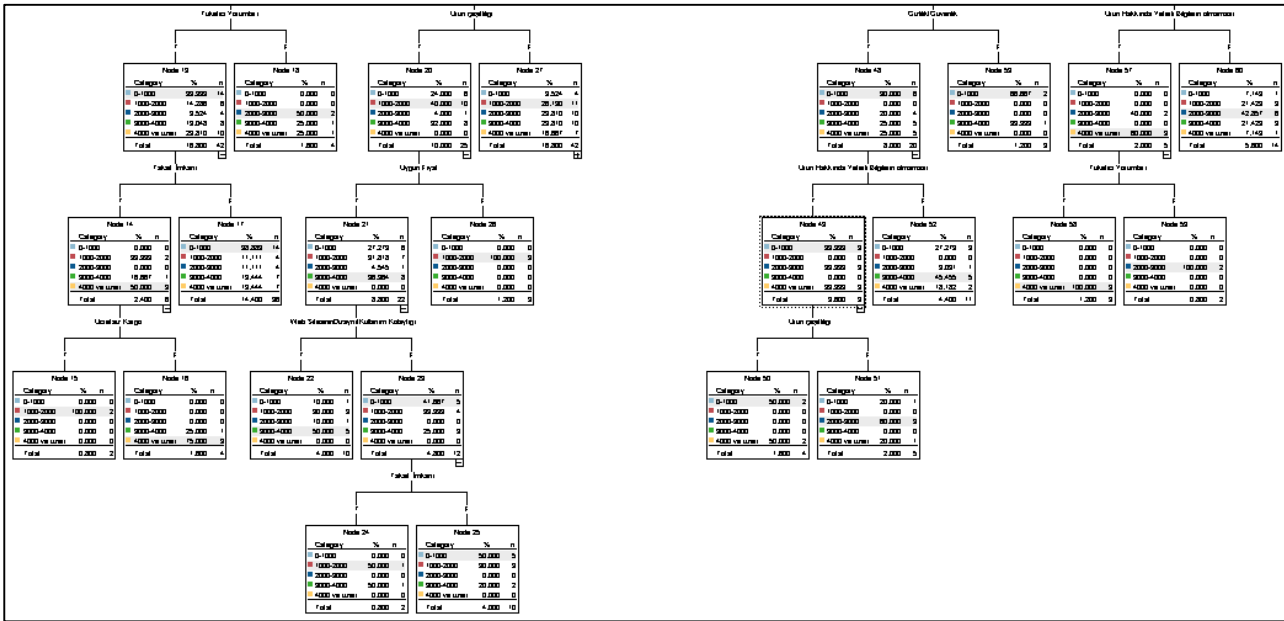
Şekil 7. Aylık Gelir Durumu Değişkeni İçin Karar Ağacında Oluşan İlk Dal



Şekil 8. Aylık Gelir Durumu Değişkeni İçin Karar Ağacında Oluşan İkinci Dal



Şekil 9. Aylık Gelir Durumu Değişkeni İçin Karar Ağacında Oluşan Üçüncü Dal



6 | SONUÇ

Veri madenciliği uygulaması içinde bulunduğğu yöntemlerle, büyük hacimli ve farklı tiplerdeki veriler içerisinde saklı kalmış olan anlamlı ve faydalı sonuçlara ve bilgilere ulaşmamızı hedefleyen bir prosestir. Finans, pazarlama, müşteri ilişkileri yöntemi, sağlık vb. alanlarda karar vermeme sürecinde duyulan ihtiyaçtan ötürü insanlar veri madenciliği uygulamasını yaygın olarak kullanmaktadırlar.

Çalışmamızı ele aldığımızda, ilk başta internet üzerinden alışveriş ile ilgili 18 sorudan oluşan anketimiz, farklı kişilik özelliklerine sahip ve Kocaeli ilinde yaşayan 250 kişiye uygulanmıştır. Akabinde SPSS Clementine programı vasıtası ile veri madenciliği uygulamalarında sınıflandırma metotlarından biri olan karar ağaçları metotları kullanılarak bireylerin yaş, aylık gelir ve cinsiyetine göre sorulara verdikleri yanıtların doğruluk oranlarını 4 farklı karar ağacı metoduna göre test edilmiş. Cinsiyet, yaş değişkeni ve aylık gelir durumu

değişkenleri için C5.0 algoritmasının doğruluk oranlarının diğer uygulanan algoritmalarından daha yüksek değerlere sahip olduğu görülmektedir. Bundan dolayı, uygulamaya tüm değişkenler için C5.0 algoritmasıyla devam edilmektedir.

Cinsiyet değişkeni için C5.0 Algoritmasıyla meydana gelen karar ağacı incelendiğinde, ağacın öncelikle ücretsiz kargo faktöründen dallandığı görülmektedir. Yani, cinsiyet değişkeni için karar ağacının dallanmasındaki en önemli faktör, **ücretsiz kargodur**.

Yaş değişkeni için C5.0 Algoritmasıyla oluşan karar ağacını incelendiğinde, ağacın öncelikle tüketici yorumları faktöründen dallandığı görülmektedir. Yani, yaş değişkeni için karar ağacının dallanmasındaki en önemli faktör, **tüketici yorumlarıdır**.

Aylık gelir durumu değişkeni için C5.0 Algoritmasıyla oluşan karar ağacı incelendiğinde, ağacın öncelikle, indirim/promosyon faktöründen dallandığı görülmektedir. Yani, aylık gelir durumu değişkeni için

ağacın dallanmasındaki en önemli faktör, **indirim/promosyondur.**

Cinsiyet, yaş, aylık gelir durumu olarak ifade edilen bağımlı değişkenler için karar ağaçları, ilk olarak sırası ile ücretsiz kargo, tüketici yorumları, ücretsiz kargo seçeneklerinden dallanmıştır. Yani yaş değişkeni bakımından karar ağacının oluşmasındaki en önemli faktör, “ücretsiz kargo” iken; cinsiyet değişkeni için, “tüketici yorumları”, aylık gelir durumu için ise “ücretsiz kargo” faktördür.

Çalışmanın Covid-19 pandemisi sürecinde yapıldığı göz önünde bulundurulmalıdır. Bu süreçte e-ticaret sektöründe artış gözlemlenmiştir. Pandeminin sona erdiği dönemde çalışma tekrar edilebilir ve elde edilen sonuçlar değerlendirilmelidir. Elde edilen sonuçların pandemi ile ilişki olup olmadığı değerlendirilebilir ve iki çalışma kıyaslanabilir.

KAYNAKLAR

Ağaç, S., Dengin Sevinir, S. ve Yılmaz, T. (2018). Online Giyim Alışverişinde Tüketicilerin Karşılaştıkları Sorunların Cinsiyet Değişkenine Göre İncelenmesi. Karadeniz Teknik Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Dergisi, 8 (15): 57-71.

Akman, M. (2010). Veri Madenciliğine Genel Bakış ve Random Forests Yönteminin İncelenmesi: Sağlık Alanında Bir Uygulama. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi, Ankara.

Avunduk, H. ve Kızgın, M. (2020). Büyük Veri Ve Sürekli Denetimde Veri Analizi. Journal of Business in The Digital Age, 3 (1): 76-83.

Baker, R.S.J.D. (2010). Data Mining For Education. International Encyclopedia Of Education, 7(3): 112-118.

Bardak, T., Avcı, Ö., Kayahan, K. ve Bardak, S. (2018). "Mobilya Alımında Geleneksel Mağaza ile Sanal Mağaza Tercihinin Veri Madenciliğine Dayalı Analizi". M. Türkmen (Ed.), 6. Uluslararası Bilim, Kültür ve Spor Kongresi, 25-27 Nisan, ss.645-652, Lviv/Ukrayna.

Bilgili, H. (2019). Y Kuşağının İnternette Alışverişe Yönelik Tutumları: Nicel Bir Araştırma. Erciyes İletişim Dergisi, 6 (1): 487-512.

Brodley, C.E. ve Utgoff, P.E. (1995). Multivariate Decision Trees. Machine Learning, 19: 45-77

Ching, W. K. ve Michael, K. P. 2002. Advances in Data Mining and Modeling, World Scientific, Hong Kong.

Çalış, A., Kayapınar, S. ve Çetinyokuş, T. (2014). Veri Madenciliğinde Karar Ağacı Algoritmaları İle Bilgisayar Ve İnternet Güvenliği Üzerine Bir Uygulama. Endüstri Mühendisliği Dergisi, 25 (3): 2-19.

Demir, E. ve Dinçer, S. (2020). Üretim Sektöründe Veri Madenciliği Uygulamaları: Literatür Taraması. Anadolu Üniversitesi İşletme Fakültesi Dergisi, 2 (1): 1-2.

Demirdöğmez, M., Gültekin, N. ve Taş, H. Y. (2018). Türkiye’de E-Ticaret Sektörünün Yıllara Göre Gelişimi. Uluslararası Toplum Araştırmaları Dergisi, 8 (15): 2216-2237.

Demirel, Ş. ve Giray Yakut, S. (2019). Karar Ağacı Algoritmaları ve Çocuk İşçiliği Üzerine Bir Uygulama. Sosyal Bilimler Araştırma Dergisi, 8 (4): 52-65.

Doğan, O. (2017). Türkiye’de Veri Madenciliği Konusunda Yapılan Lisansüstü Tezler Üzerine Bir Araştırma. Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 19 (3) , 929-951.

Emre, İ. ve Selçukcan Erol, Ç. (2017). Veri Analizinde İstatistik mi Veri Madenciliği mi?. Bilişim Teknolojileri Dergisi, 10 (2): 161-167.

Eren, B. ve Gödekmerdan Önder, L. (2020). Z Kuşağı Tüketicilerin On-Line Alışverişe Yönelik Deneyimlerinin Belirlenmesi: Ankara Uygulaması. Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (20): 203-228.

Gençyürk Erdoğan, M. (2020). Covid-19 Döneminde E-Ticaret Ve Dijital Reklam Yatırımları. Selçuk İletişim Dergisi, 13 (3): 1296-1318.

Kolata, S. (2014). İnternet Üzerinden Alışveriş Ve İnternet Üzerinden Alışverişin Tüketici Eğilimleri Üzerindeki Etkisi. Yüksek Lisans Tezi, KTO Karatay Üniversitesi.

Oğuzlar, A. (2010). CART Analizi İle Hanehalkı İşgücü Anketi Sonuçlarının Özetlenmesi. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 18 (3-4): 79-90.

Özden, A. ve Üner, T. (2019). İnternet Alışverişlerinde Tüketicilerin Etkileşim Düzeylerinin Belirlenmesi Üzerine Bir Araştırma. Abant Kültürel Araştırmalar Dergisi, 4 (8): 31-49.

Özguven, N. (2011). Tüketicilerin Online Alışverişe Karşı Tutumları İle Demografik Özellikleri Arasındaki İlişkinin Analizi. Karamanoğlu Mehmetbey

Üniversitesi Sosyal Ve Ekonomik Araştırmalar Dergisi, 2011 (2): 47-54.

Pehlivan, G. (2006). CHAID Analizi ve Bir Uygulama. Yayınlanmamış Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi, İstanbul.

Quinlan, J.R. (1986). Induction of Decision Trees. *Machine Learning*, 1: 81-106.

Sayılı, M. ve Büyükkoroğlu, A. (2013). E-Ticaret Yoluyla Gıda Maddeleri Satın Almaya Yönelik Tüketicilerin Tutumunu Etkileyen Faktörlerin Analizi. *Journal of Agricultural Sciences*, 18 (3): 246-255.

Sönmez, U. (2019) Online Alışverişe Yönelik Satın Alma Tarzlarının Kişilik Tipleri Açısından İncelenmesi: Sakarya İli Örneği. Yüksek Lisans Tezi, Sakarya Üniversitesi.

Şeker, A. (2020). Çevrimiçi Alışveriş Yapan Tüketici Davranışlarına Yönelik Çevrimiçi Bir Araştırma. *İşletme ve İktisat Çalışmaları Dergisi*, 8 (2): 11-27.

Witten, I.H., Frank, H., Hall, M.A. ve Pal, C.J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Cambridge: Elsevier

Yaslıdağ, B. ve Ferhadi, G. (2018). Tüketicilerin İnternet Üzerinden Satın Alma Davranışının İncelenmesi: Azerbaycan Örnekleri. *Anadolu Bil Meslek Yüksekokulu Dergisi*, 13 (50): 15-28.