

ENDÜSTRİ 4.0 İŞİĞİNDA VERİ MADENCİLİĞİ VE PAZARLAMA: LİTERATÜR TARAMASIYLA SON GELİŞMELER, YENİ TRENDLER

Emrah BİLGİÇ¹
M.Fevzi ESEN²

ÖZ

Endüstri 4.0 ile başlayan gelişmeler firmaların sadece üretim bölümlerinde değişimlere sebep olmamış, Endüstri 4.0'ın üretimi etkileyen bütün süreçlerin daha verimli, müşterinin isteklerine daha hızlı uyarlanabilir, müşteriye tam istediği özelliklerde ve tam zamanında kusursuz hizmet gibi hedefleri, diğer departmanların da değişime ayak uydurmasını zorunlu kılmıştır. Endüstri 4.0'dan önce de kullanılan, fakat günümüzde hepsinin ortak bir amaç için kullanıldığı Wifi, RFID, sensörler, POS cihazları vb. akıllı sistemler, şirketlerle alakalı her türlü veriyi üretim, satın alma, lojistik, pazarlama, satış vb. süreçlerde toplanmaktadır. Böylece Bulut Sistemlerde depolanan çok büyük boyutlu verilerin bilgiye nasıl dönüştürülebileceği hususu da önemli hale gelmiştir. Bu durum, Veri Bilimi alanının doğmasını ve kullanımının hızla yaygınlaşmasını beraberinde getirmiş, firmaların farklı departmanlarında Veri Bilimcisi istihdamını zorunlu kılmıştır. İşletmeyle alakalı her türlü karar verme süreçlerini destekleyen İşletme Zekâsının içerisinde hali hazırda yer alan ve Büyük Veri Biliminin (Analitiği) alt kolları olan Veri Madenciliği, Metin Madenciliği, Makine Öğrenmesi ve Yapay Zeka gibi kavramlar sıkça başvurulan teknikler olmakla beraber, Endüstri 4.0'la birlikte önemleri daha da artmıştır. Bu çalışmada, özellikle Veri Madenciliği olmak üzere, Metin Madenciliği ve Makine Öğrenmesi tekniklerinin pazarlama alanındaki güncel uygulamaları, 2015 yılından günümüze kadar yapılmış çalışmaların ayrıntılı incelenmesi ile ortaya konulmaktadır. Yapılan literatür çalışmasında görülmüştür ki, bilgisayar bilimcileri tarafından geliştirilmekte olan farklı algoritmalar hızlıca farklı pazarlama uygulamalarında kullanılmaya başlamıştır.

Anahtar Kelimeler: Endüstri 4.0, Pazarlama, Veri Madenciliği, Metin Madenciliği, Makine Öğrenmesi

Jel Kodları: M3, M30, C38

DATA MINING AND MARKETING IN THE LIGHT OF INDUSTRY 4.0: LAST ISSUES, NEW TRENDS

ABSTRACT

The improvements launched with the Industry 4.0 did not result in changes in the production departments of companies only, at the same time other departments had to keep up with changes since one of the aim of Industry 4.0 is to make all departments productive. Furthermore, Industry 4.0 claims to have smarter, more adaptable processes to customer needs and wants, and just in time, customized and perfect service to customers make all departments take part in change. Smart tools such as Wifi, RFID, sensors which were used before Industry 4.0 but now being used for a common objective, to collect data from different departments of the company such as marketing, production, logistics. The data which stored in cloud systems is big, hard to manage and analyze that's why Data Science is popular nowadays and companies have to hire Data Scientist in different departments. Business Intelligence which supports many decision making processes already have tools such as Data Mining, Text Mining, Machine Learning and Artificial Intelligence gain more importance with Industry 4.0. In this paper recent applications of marketing especially with Data Mining will be discussed by searching papers published since 2015. One can conclude that different algorithms which developed by computer scientists are being used to solve different marketing problems.

Keywords: Industry 4.0, Marketing, Data Mining, Machine Learning, Text Mining

JEL Codes: M3, M30, C38

GİRİŞ

Veri Madenciliği ve diğer tekniklerin pazarlama alanındaki son uygulamalarına geçmeden önce teknolojik gelişmelerle birlikte ortaya çıkan yeni veri kaynaklarına dikkat çekmek önemlidir. Pazarlamanın en yaygın uygulama alanı olan perakende sektöründeki mahalli küçük işletmelerin kapanması, çok uluslu veya ulusal kurumsal perakende zincirlerinin mağaza sayısını artırma politikaları ve tüketicilerin artan ihtiyaçlarını söz konusu perakende zincirlerinden karşılaması gibi sebeplerle, perakendecilerin topladıkları veri günümüzde büyük boyutlarla ifade edilmektedir. Üreticilerin akıllı sistemlerle stok uygulamaları ve lojistik çözümleri gibi faaliyetlerle perakendeci ile birebir ilişki kurması da verinin boyutlarını daha çok arttırmaktadır. Rekabete ayak uydurmak ve müşterilerine daha iyi hizmet sunmak amacı ile hareket eden perakende sektörü, POS sistemleri dışında RFID, akıllı sensörler, kablosuz ağlar, vb. sistemler ile veri toplayarak kendi bulut sistemlerini oluşturmaya başlamıştır. Ayrıca firmalar, sosyal medya ve tüketici bloglarında bulunan tüketici yorumları, firma web sitelerini ziyaret verileri, müşterilerin konum tabanlı kullanıcı bilgilerine ilişkin dış kaynaklı verileri de elde edebilmekte, müşterilerini sürekli gözetim altında tutmakta ve bu verilerden faydalı çıkar-samalarda bulunmak amacıyla çeşitli analizler yapmaktadır.

Endüstri 4.0'ın müşterilerden toplanan anlık verileri en hızlı biçimde üretim süreçlerine yansıtma hedefi müşterilerin tutum ve davranışlarının sürekli takibini ve sürekli analizi gerektirmektedir. Bu amaçla Veri Madenciliği ve benzeri tekniklerin pazarlama alanındaki kullanılışı hakkında genel bilgiler de vermekte olan çalışmamız özellikle "Veri Madenciliği ve Pazarlama" literatüründeki güncel yaklaşımları incelemekte ve gelecekte yapılabilecek potansiyel çalışmaları da tespit etmeyi amaçlamaktadır.

YÖNTEM

Çalışma için Elsevier Science Direct ve Web of Science veri tabanlarında yer alan pazarlama çalışmalarının yoğun olduğu dergilerde, Cho vd. (2017) tarafından oluşturulan pazarlama konu dizinleri sözlüğü ile uyumlu olarak, "data mining & marketing", "big data & marketing", "data mining & customer", "data mining & retail", "data mining & CRM" ve "marketing & industry 4.0" gibi arama terimleri kullanılarak literatür taraması gerçekleştirilmiştir.

LİTERATÜR TARAMASI SONUCU BULGULAR

Endüstri 4.0 ile birlikte, ürünün tasarım sürecinden başlayarak, üretim sürecinde, depolama sürecinde, lojistiği sürecinde, mağazalarda veya web sitelerinde yerini alması sürecinde, pazarlanması, satılması süreçlerinde hatta müşterinin kullanımı sürecinde yapısal olmayan çok büyük boyutlu veriler,

firmalar tarafından eş anlı olarak elde edilmektedir. Bu tip veriler firmalar tarafından hızlıca analiz edilerek, üretim departmanları başta olmak üzere işletmedeki departmanlara iletilmekte ve firma müşteri istek ve taleplerine hızlıca karşılık verebilmektedir (Galletta vd., 2017). Gartner (2017) tarafından yayınlanan teknoloji araştırma raporuna göre, önümüzdeki beş yılda pazarlama başta olmak üzere birçok alanda paralel veri işleme ve analiz sistemleri kullanımının itici güç olacağı, ayrıca Veri Madenciliği ile Yapay Zekâ teknolojilerinin, ilk beş teknolojik faaliyet alanlarının içerisinde olacağı belirtilmektedir.

Bu bağlamda çalışmamız kapsamında yapılan literatür taraması sonucu elde edilen bulgular kısaca şu şekilde özetlenebilir:

Alborzi ve Khanbabaie (2016), İran'daki bir bankanın 4459 müşterisine ait banka kartı kullanım bilgileri ile yaptığı hibrit analizlerle müşterilerin kredi skorlarını oluşturmuş, müşterilerin davranışlarını tespit etmiş, ayrıca yapay sinir ağları yaklaşımı ile müşterileri sınıflara ayırmışlardır. Bu bağlamda 10. segmentte yer alan 28 müşteri bankanın en değerli müşterileri olarak belirlenmiştir.

Culotta ve Cutler (2016), müşterilerin marka algılarını, sosyal medya verileri kullanarak, zahmetsiz ve maliyetsiz bir yöntemle analiz etmişlerdir. Veri madenciliği tekniklerinin kullanıldığı çalışmada önceki çalışmalardan farklı olarak sosyal medya kullanıcılarının oluşturduğu içerikleri analiz etmek yerine, markanın müşterilerinin oluşturmuş oldukları ağlar analiz edilmiştir.

Ghazavi ve Lotfi (2016), müşterilerin mağaza içindeki gezintileri için geliştirmiş oldukları matematiksel formülle, raf tasarımı konusunda karar vericilere bir fikir sunmaktadır. Çalışmada Yapay Zekâ teknikleri kullanılmıştır. Qi vd. (2016), firmaların yeni ürün geliştirme sürecinde kolay karar vermeleri hususunda, müşteriler tarafından yapılan çevrimiçi yorumların sentiment analizi ile faydalı bilgilere dönüştürülebileceğini göstermişlerdir. Kaneko ve Yada (2016), 1000 süpermarket müşterisine ait mağazada gezinme ve alışveriş işlemleri verilerini kullanarak, mağaza içerisindeki müşterilerin gezinmelerinin artması ile müşterilerin daha fazla alışveriş yaptıklarını tespit etmişlerdir. Veriler, müşterilerin alışveriş sepetine takılan bir RFID etiketi ile mağazadaki dolaşma verileri ve müşterilerin POS verilerini eşleştirerek elde edilmiştir.

Jin vd. (2016), ürün tasarımcılarının karar verme süreçlerini desteklemek için, müşteri yorumlarını analiz etmiştir. Bu amaçla Amazon.com sitesindeki benzer ürünlerin yorumları duygu analizi ile incelenmiştir. Çalışmanın amacı ürün dizaynı süreçlerini pazarın ihtiyaçları doğrultusunda şekillendirmektir. Birçok cep telefonu markasının binlerce yorumlarını inceleyen çalışmada müşteriler en çok telefonların ekranını, bataryasını, internet ve kablosuz ağını,

uygulamalarını ve hafızasını yorumlamışlardır. Yorumlar ayrıca Bayesçi yöntem kullanılarak olumlu ve olumsuz olarak sınıflandırılmıştır.

Rodrigues ve Ferreiea (2016), bir parfümeri zincirinin müşterilerini RFM değişkenleri ile üç segmente ayırmış ve birliktelik kuralları analizi ile her bir segmente en uygun ürünleri öneren bir algoritma yazmışlardır. Weng (2016), pazara daha sonradan giren bir ürünün, kendisinden daha önce pazara girmiş başka bir ürünle olan birlikteliğinin destek değerinin düşük olacağını belirtmiş, bu bağlamda pazara farklı zamanlarda girmiş olan ürünlerin birliktelik kuralları analizi için yeni bir algoritma önermiştir. Vidden vd. (2016), simülasyonla oluşturdukları 10 değişkenli bir pazar veri setine farklı farklı kümeleme teknikleri ile segmentasyon yaparak kümelenecekleri karşılaştırmış, Gizli Sınıf (Latent Class) ile yapılan kümeleme analizi diğer kümeleme tekniklerine üstün gelmiştir. Backiel vd. (2016), bir milyon dört yüz bin mobil telefon müşterisinin 30 milyonu aşkın işlem verisini kullanarak, firmaların müşteri segmentlerinin belirlenmesi ve segmentlere ait stratejilerin oluşturulmasında kümeleme modellerinden nasıl etkin yararlanabileceklerini gösterilmiştir.

Azeem vd. (2017), yapay sinir ağları, doğrusal regresyon, C4.5 ve destek vektör makinaları gibi veri madenciliği tekniklerinin kayıp müşterilerin sınıflandırılması analizindeki performanslarını, bulanık bir sınıflandırma algoritmasının performansı ile karşılaştırmışlardır. Cho vd. (2017), pazarlama literatüründe, müşteri değeri ve müşteri sadakatini ölçen önemli çalışmalarda kullanılan değişkenlerle network analizini kullanarak bir pazarlama kavram sözlüğü oluşturmuştur. Fu vd. (2017), müşteri ilişkileri yönetimi kapsamında, online oyun sitelerindeki aktif ve kayıp müşterilerin sınıflandırılması için kümeleme analizini kullanarak, yüksek boyutlu müşteri segmentlerini tespit etmişlerdir.

Tanaka vd. (2017), klasik RFM analizini geliştirerek, Japonya'daki bir süpermarket zinciri müşterilerini, önem derecelerine göre lojistik regresyon yöntemi ile sınıflandırılmıştır. Tsai vd. (2017) çalışması, müşterilerin satın alma davranışlarını tahmin eden akıllı bir sistem önermektedir. Sistem dört ana içerikten oluşmaktadır: Mobil bir işlem (transaction) veri tabanı, davranışları araştıran bir modül, benzerlikleri ortaya çıkaran bir modül ve davranış tahminini yapan bir modül. İlk içerik müşterilerin satın alma işlemlerini diğer çalışmalardan farklı olarak satın alınan ürün adedini de dikkate alarak kaydettiği gibi müşterilerin mağaza içi hareketlerini de kaydetmektedir. Bilgisayar ortamında üretilmiş verilerle test edilmiş olan sistem son olarak hedef kullanıcıya bir alışveriş listesi sunmaktadır.

Singh ve Tucker (2017) internet ortamındaki ürünle alakalı yorumların müşterilerin satın alma kararlarında çok önemli olduğunu vurgulamış, bu amaçla makine öğrenmesi (machine learning) yaklaşımı ile müşteri yorumlarını

otomatik olarak iki ana grupta sınıflandırmışlardır. Analizde kullanılacak olan veri Amazon.com sitesinden, 3 farklı model cep telefonu için yapılmış yorumlardan elde edilmiş olup, sınıflandırma işi Naif Bayes, Destekçi Vektör Makineleri ve Karar Ağaçları yöntemleri ile yapılmıştır. Başarıyla uygulanan sınıflandırma ile yorumlar, iki ana gruptan birincisi için üç (ürünün formu, fonksiyonu ve davranışı) ve diğeri için iki (hizmet ve diğeri) alt gruba sınıflandırılmıştır.

Aguwa vd. (2017) işletmelerin karar verme süreçlerine destek olması amacıyla, bulanık mantık kullanarak, bir otomobil firmasının müşterilerin sesi (VoC) verilerine Metin Madenciliği ve Birliktelik Kuralları Analizi uygulamış, ortaya çıkan sonuçlarla müşterilerin memnuniyeti endeksi oluşturulmuştur.

Griva vd. (2018), temel veri madenciliği tekniklerinden kümeleme ve birliktelik kuralları analizi ile satış verilerini kullanarak müşterileri mağazayı ziyaretlerine göre segmente etmektedirler. Çalışma daha önce yapılan benzer çalışmaların aksine bir müşterinin bir ziyaretinde almış olduğu ürünleri değil, çoklu ziyaretlerde aldığı ürünleri dikkate almaktadır. Böylece, örneğin müşterinin son ziyaretinin amacının kahvaltı olduğu (süt, mısır gevreği, ekmek, peynir alışverişi) tespit edilebilmektedir. Boix ve Moreno (2018), mobil telefon şebekelerinin pazarlama kampanyalarında kendilerini doğru konumlandırabilmeleri için gerekli olan müşteri kullanım detaylarına ilişkin önemli değişkenlerin seçilmesi ve müşteri kayıp analizinin gerçekleştirilmesi için bir veri madenciliği modeli önermişlerdir.

Nakano ve Kondo (2018), yaklaşık 2500 Japon müşteriye ait mağaza içi alışveriş, internet üzerinden alışveriş, internet tıklama verileri, sosyal medya verileri ve demografik verileri kullanarak müşterileri Gizli Sınıf kümeleme analizi ile segmente etmiştir.

Vincent vd. (2018), 1000 müşteriye ait verileri moda tasarımı üzerinde başarıyı yakalayabilmek ve hedef pazarlama amacıyla veri madenciliği tekniklerinden olan k-ortalamlar tekniği ile sınıflandırma yapmışlardır. Ortaya çıkan sonuçlar müşterilerin yeni siparişlerinde ne tür bir tasarım istedikleri hususunda firmalara yardımcı olacaktır.

SONUÇ VE GELECEK ÇALIŞMALAR İÇİN ÖNGÖRÜLER

Söz konusu çalışmalar yürütülürken, Endüstri 4.0 ile ilgili yaşanan gelişmelere paralel olarak, işletmelerin her türlü karar verme süreçlerini doğrudan etkileyen veri ve bilgi yönetimi konusundaki teknik ve teknolojileri inceleyen çalışmalara önem verilmeye başlandığı tespit edilmiştir. Özellikle 2015 yılından sonra yapılan çalışmalarda dijital pazarlama, sosyal medya

temelli müşteri ilişkileri yönetimi, nesnelerin interneti, akıllı sistemlerin kullanımı gibi konularda da veri madenciliği tekniklerinin halen kullanıldığı gözlemlenmektedir.

E-ticaret alışveriş hacminin son yıllarda gösterdiği gelişmeye paralel olarak e-ticaretle alakalı çalışmalarda artış gözlenmesinde karşın, müşterilerin büyük çoğunluğunun halen gerçek mağazalardan (bricks and mortar) alışverişe yapmakta olduğu hususu göz önünde bulundurulduğunda, mağaza içi uygulamalarla alakalı olan çalışmalarda da son yıllarda artış olduğu belirlenmiştir. Ayrıca yürütülen Veri Madenciliği uygulamalarında, müşterilerin alışveriş ve sosyo-demografik verilerinin yanı sıra, internette gezinme verileri ile ürün ve hizmetler hakkındaki tutum ve davranışlarına yönelik sosyal medya paylaşımları ve akıllı aletlerle mağaza içinde elde edilmiş veriler de kullanılmaktadır.

Ayrıca gerçekleştirilen literatür çalışmasından hareketle, temel Veri Madenciliği ve Makine Öğrenmesi Teknikleri algoritmaları farklı pazarlama problemleri için halen geliştirilmektedir. Örneğin, satın alınan ürün adedini de dikkate alan Birliktelik Kuralları algoritmaları, firmaların kârlılığına en çok katkıyı yapacak ürün ilişkilerini veren geliştirilmiş Birliktelik Kuralları algoritmaları ve müşterilerin satın alma örüntülerini inceleyip ürün öneren otomatik algoritmalar (automated recommender systems) bu kapsamda önerilen tekniklerdendir.

Daha önce vurgulandığı gibi, çalışmalarda sadece müşterilerin satın alma verileri ile ilgilenilmediği, ayrıca, dış veri kaynaklarından da verinin sağlandığı görülmektedir. Örneğin, müşterilerin internet üzerinden satın alma davranışları, internet kullanım verileri ve sosyal medyadaki yorumları gibi veriler hibrit tekniklerle analiz edilmektedir. Bu bakımdan gelecekte yapılacak çalışmalarda sadece mağaza içi alışveriş verileri değil, çoklu veri kaynaklarından da (sensörler, RFID, kablosuz ağlar, web...) yararlanarak, yapılandırılmış ve yapılandırılmamış verilerin birlikte analizini sağlayacak modeller ve teknikler geliştirilebilir. Nitekim günümüzde vurgulandığı gibi kablolu ve kablosuz akıllı sistemler vasıtasıyla müşterilerin mağaza içi hareketleri takip edilmekte, hangi reyonlarda ne kadar duraksadıkları, hangi ürünlerle etkileşim sağladıkları kaydedilip bulut sistemlerinde tutulmaktadır. Söz konusu verilerle, hizmetin/ürünün müşteriye sunumuna ilişkin tesis yerleşimi ile reyon ve raf düzeni hakkındaki araştırmalar halen devam etmektedir. Bu hususta yapılacak olan çalışmalar, hem akademik hem de sektörel katkı açısından önem arz etmektedir.

KAYNAKÇA

- Aguwa, C., Olya, M. H., & Monplaisir, L. (2017). Modeling of fuzzy-based voice of customer for business decision analytics. *Knowledge-Based Systems*, 125, 136-145.
- Alborzi, M., & Khanbabaei, M. (2016). Using data mining and neural networks techniques to propose a new hybrid customer behaviour analysis and credit scoring model in banking services based on a developed RFM analysis method. *International Journal of Business Information Systems*, 23(1), 1-22.
- Azeem, M., Usman, M., Fong, A.C.M. (2017). A churn prediction model for prepaid customers in telecom using fuzzy classifiers. *Telecommunication Systems*, 66 (4), 603-614.
- Backiel, A., Baesens, B., Claeskens, G. (2016). Predicting time-to-churn of prepaid mobile telephone customers using social network analysis. *Journal of the Operational Research Society* , 67(9), 1135-1145.
- Boix, M.P., Moreno, J.L.M. (2018). A social model based on customers' profiles for analyzing the churning process in the mobile market of data plans. *Physica A: Statistical Mechanics and its Applications*, 496, 571-592.
- Cho, Y.C., Fu, P.W., Wu, C.C. (2017). Popular Research Topics in Marketing Journals, 1995–2014. *Journal of Interactive Marketing*, 40, 52-72.
- Culotta, A. & Cutler, J. (2016). Mining Brand Perceptions from Twitter Social Networks. *Marketing Science*, 35(3) ,343-362.
- Fu, X., Chen, X., Shi, Y.T., Bose, I., Cai,S. (2017). User segmentation for retention management in online social games. *Decision Support Systems*, 101, 51-68.
- Gartner Group (2017). "Advanced Technology Research Note 2017, <https://www.gartner.com/technology/research/>, 20.02.2018
- Ghazavi, E., & Lotfi, M. M. (2016). Formulation of customers' shopping path in shelf space planning: A simulation-optimization approach. *Expert Systems with Applications*, 55, 243-254.
- Griva, A., Bardaki, C., Pramadari, K., & Papakyriakopoulos, D. (2018). Retail Business Analytics: Customer Visit Segmentation Using Market Basket Data. *Expert Systems with Applications*.
- Jin, J., Liu, Y., Ji, P., & Liu, H. (2016). Understanding big consumer opinion data for market-driven product design. *International Journal of Production Research*, 54(10), 3019-3041.

- Kaneko, Y., & Yada, K. (2016). Fractal Dimension of Shopping Path: Influence on Purchase Behavior in a Supermarket. *Procedia Computer Science*, 96, 1764-1771.
- Nakano, S., & Kondo, F. N. (2018). Customer segmentation with purchase channels and media touchpoints using single source panel data. *Journal of Retailing and Consumer Services*, 41, 142-152.
- Qi, J., Zhang, Z., Jeon, S., Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. *Information & Management*, 53 (8), 951-963.
- Rodrigues, F., & Ferreira, B. (2016). Product recommendation based on shared customer's behaviour. *Procedia Computer Science*, 100, 136-146.
- Singh, A., & Tucker, C. S. (2017). A machine learning approach to product review disambiguation based on function, form and behavior classification. *Decision Support Systems*, 97, 81-91.
- Tanaka, T., Hamaguchi, T., Saigo, T., & Tsuda, K. (2017). Classifying and Understanding Prospective Customers via Heterogeneity of Supermarket Stores. *Procedia Computer Science*, 112, 956-964.
- Tsai, C. Y., Li, M. H., & Kuo, R. J. (2017). A shopping behavior prediction system: Considering moving patterns and product characteristics. *Computers & Industrial Engineering*, 106, 192-204.
- Vidden, C., Vriens, M., & Chen, S. (2016). Comparing clustering methods for market segmentation: A simulation study. *Applied Marketing Analytics*, 2(3), 225-238.
- Vincent, O. R., Makinde, A. S., Salako, O. S., & Oluwafemi, O. D. (2018). A self-adaptive k-means classifier for business incentive in a fashion design environment. *Applied Computing and Informatics*, 14(1), 88-97.
- Weng, C. H. (2016). Identifying association rules of specific later-marketed products. *Applied Soft Computing*, 38, 518-529.