Avrupa Bilim ve Teknoloji Dergisi
Sayı 16, S. 382-394, Ağustos 2019
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 16, pp. 382-394, March 2019
Copyright © 2019 EJOSAT
Research Article

Bankacılık Sektöründe Yeni Özniteliklerle ve Makine Öğrenmesi Yöntemleriyle Müşteri İlişkileri Yönetiminin Zenginleştirilmesi

Faruk Bulut [*1]

[1] Haliç Üniversitesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0003-2960-8725)

(İlk Geliş Tarihi 31 Ocak 2019 ve Kabul Tarihi 15 Mayıs 2019)

(**DOI:** 10.31590/ejosat.520295)

ATIF/REFERENCE: Bulut, F. (2019). Bankacılık Sektöründe Yeni Özniteliklerle ve Makine Öğrenmesi Yöntemleriyle Müşteri İlişkileri Yönetiminin Zenginleştirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 382-394.

Öz

Bir banka müşterilerinden kendilerini tanımlayıcı detaylı kişisel verileri kolaylıkla alamayabilir. Fakat müşteriler zaman içerisinde, davranışsal özellikleri ve yaşamsal alışkanlıkları ile ilgili birçok bilgiyi farkında olmadan banka ile paylaşırlar. Bu bilgiler sadece bir birey için yüzlerce, hatta binlerce olabilir. Milyonlarca bireysel ve kurumsal müşterisi olan bir bankanın tüm müşterilerine ait belirli kategorilerdeki bilgileri milyarlara ulaşabilir. Milyonlarca müşterisi olan büyük kuruluşlar genel olarak Bilgi Teknolojileri alanında "veri zengini ve bilgi yoksunu" olarak nitelendirilirler. Bu işlenmemiş veriler içerisine müşteriler ve piyasa hakkında gizlenmiş birbirinden farklı oldukça fazla değerli bilgiler bulunabilmektedir. Veri Bilimi esasen bu gizli bilgileri ham veriler içerisinden çıkarmayı amaçlar. Bu çalışmada, kurumsal ve bireysel müşterilerin sayısal verilere dönüştürülebilen farklı yaşamsal ve davranışsal alışkanlıklarının neler olabileceği ile ilgili önerilerde bulunulmuştur. Bu veriler ile piyasanın daha iyi analiz edilmesi, müşteri segmentasyonunun daha iyi yapılması, hizmet ve ürünlerin doğru müşteri kitlelerine daha az emek ile satılmasını amaçlayan yapay zekâ tabanlı bir sunulan Müşteri İlişkileri Yönetimi yani CRM (Customer Relationship Management) uygulamasının yapısı önerilmektedir. Projede bilişim alanındaki Veri Madenciliği disiplinlerinde bulunan algoritmalar kullanılarak özgün bir yapay destek modelinin önerilmesi amaçlanmaktadır. Farklı veri madenciliği uygulamaları ve süreçlerin tanımlanmasıyla bankacılık ve müşteri ilişkileri yönetimi alanına yeni bir perspektif kazandırılmaya çalışılmıştır.

Anahtar Kelimeler: Müşteri Tanıma, Bankacılık Sektörü, CRM, Makine Öğrenmesi.

Enriching Customer Relationship Management with New Attributes and Machine Learning Methods in Banking Sector

Abstract

A bank may have difficulty in acquiring detailed personal information from their clients. Over time, however, customers share a variety of information about their behavioral characteristics and life habits with the bank. This information for people can be hundreds or even thousands. It can reach billions of information in certain categories of all customers of a bank with millions of individual and corporate clients. Large organizations with millions of customers are generally referred to as "rich in data rich and lack in knowledge" in the field of Information Technology. This unprocessed data might contain many valuable information that is hidden from customers and the market. There might be a lot of valuable hidden information about both customers and market. Data Science essentially aims to extract this confidential information from raw data. In this study, suggestions have been made about the different vital and behavioral habits of institutional and individual customers which can be converted into numerical data. With this data, the structure of CRM (Customer Relationship Management), an artificial intelligence-based system which aims to better analyze the market, better

[*] Sorumlu Yazar: Haliç Üniversitesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye (ORCID: 0000-0003-2960-8725), farukbulut@halic.edu.tr

customer segmentation, and offer services and products with less labor to the right customers, is recommended. In this project, it is aimed to propose an original artificial support model by using algorithms in Data Mining disciplines. With the identification of different data mining applications and processes, we have tried to give a new perspective to the field of banking and customer relationship management.

Keywords: Customer recognition, Banking sector, CRM, Machine Learning.

1. Giriş

Bilişim dünyasındaki yazılımsal ve donanımsal gelişmeler finansal alandaki iş, ticaret ve kârlılık gibi birçok kavramın değişmesine neden olmuştur. Özellikle milyonlarca müşterisi olan ve bu müşterilere ait milyarlarca işlenmemiş ham verilere sahip bankalar bilişim odaklı çözümlere daha çok önem verir olmuştur (Zerbino vd., 2018).

Veri madenciliği, mevcut veri yığınlarından işe yarayacak önemli bilgileri çıkarmak ve bankacılık sektöründe daha iyi kararlar alabilmek için kullanılan bir araçtır. Verilerin toplanabilmesi ve veri ambarının oluşturulabilmesi için veri tabanlarındaki çeşitli verilerin uygun bir biçimde birleştirilmesi ve hazırlanması gerekir (Nguyen vd., 2018). Kuşkusuz, veri madenciliği birçok alanda kullanılmaktadır. Tıbbi sonuçları analiz etme, kredi kartı dolandırıcılığını tespit etme, müşteri satın alma davranışını tahmin etme, Web kullanıcılarının kişisel çıkarlarını öngörme, üretim süreçlerini optimize etme gibi alanlarda pratik uygulamaları çok başarılı olmuştur (Amani vd., 2011).

Bankacılık, milyonlarca kişi ile uğraşan ve milyarlarca TL işlem hacmi olan oldukça büyük ve karlı bir sektördür. Teknolojinin donanımsal ve yazılımsal olarak ilerlemesinin ardından 4. Jenerasyon Sanayi 4.0'ın bu alanda somut uygulamaları her alanda görülmektedir (Tam vd., 2017).

Bankalar veri açısından oldukça zengin kurumlar arasında kabul edilir. Burada asıl amaç işlenmemiş yüz milyonlarca ham etiketli veriden değerli, işe yarar bilgilerin elde edilmesidir. Veriler, herhangi bir şirketin en değerli varlıklarından biri olabilir, ancak yalnızca ham verilerde bulunan değerli bilgilerin nasıl açığa çıkarılacağı bilinmelidir. Veri madenciliği, tarihsel veriden bilgi elmaslarının çıkarılmasına ve gelecekteki durumların sonuçlarının tahmin edilmesine olanak tanır. Veri madenciliği bu durumda, iş kararlarını optimize etmeye, her bir müşterinin ve iletişimin değerini artırmaya ve müşteri memnuniyetini artırmaya yardımcı olur. Veri madenciliği, daha önce bilinmeyen bilgilerin tipik olarak desenler ve dernekler şeklinde geniş veri tabanlarından çıkartılması işlemidir. Bankalar veri madenciliği ile sağlanan birçok avantajı fark eder. Toplanan büyük miktarda veriden potansiyel olarak yararlı bilgileri belirleyerek değerli bir araçtır. Bir banka, iyi bir veri madenciliği destekli CRM uygulaması ile rakiplerine karşı açık bir avantaj kazanabilir (Kasemsap vd., 2018).

Müşteri ilişkileri yönetimi (CRM), müşteriler ile uzun süreli ilişkiler kurulmasına, gelirlerin ve kazançların artırılmasına yardımcı olan bir stratejidir. CRM, bankacılık sektöründe daha da fazla bir öneme sahiptir. CRM'in odağında, müşteri kazanımından müşteri sadakatine geçmekte ve zaman, para ve yönetim kaynaklarının uygun miktarlarının bu kilit görevlerin her ikisine de yönlendirilmesini sağlamaktadır (Hasan, 2018).

Müşterinin firmaya olan bağlılığın devam ettirilebilmesi için çeşitli önlemlerin alınması, gerekli hizmetlerin sunulması ve olası memnuniyetsizliklerin giderilmesi gibi faktörlere dikkat edilmesi gereklidir. Bu nedenle çalışmamızda CRM konusunda müşterilere ait kullanılabilecek etkin özniteliklerin belirlenmesi ve etkin modellerin inşa edilmesine olanak sağlamaya çalışma vardır.

Yaşanılan rekabetçi ortamda müşterilerin ihtiyaçlarına uygun bankayı birçok seçenek içerisinden bulabilmesi hiç de zor değildir. Bankaların mevcut müşterilerinin ihtiyaçlarını karşılayamaması durumunda müşteri kaybetmeleri kaçınılmazdır. Bu nedenle bankacılıkta CRM uygulamalarının veri bilimi ile desteklenmesi gerekmektedir. CRM uygulamaları temel olarak dört aşamadan oluşmaktadır (Anshari vd., 2018).

1. Müşteri tanımlama (Customer Identification)
2. Müşteri Çekme (Customer Attraction)

Müşteri tanımlama aşamasından sonra yapılabilecek bir işlemdir. Hedef müşteri kitlesine bire bir münasebet ile Müşteri Tanıma aşamasında yapılan işlemlerle belirlenen en uygun ürünlerin satılması (Direct Marketing) sağlanacaktır. Bu bankanın yapacağı bir çalışmadır.

3. Müşteri Koruma (Customer Retention)

CRM uygulamalarının esas odağı mevcut müşterilerin korunması amacını gütmektedir. Müşteri memnuniyetinin sağlanması, müşteri memnuniyet derecesinin belirlenmesi bu sayede yapılabilir. Bu alan daha çok bankayı ilgilendirdiği için detaylara giremiyoruz.

4. Müşteri Geliştirme (Customer Development)

Müşteri sayısının artırılması olarak da tanımlanabilir.

CRM uygulamalarında karşılaşılabilecek problemler şu şekilde sıralanabilir:

- Milyonlarca müşteri

- Her bir müşteriye ait binlerce işlenmemiş ham veri
- Çok veri az bilgi
- Zayıf ve yetersiz tahminler
- Karlılık oranının olması gerekenden düşük olması
- Kişiye özgü çözüm eksikliği
- Yetersiz ve eksik sektör analizi
- Zamana ve mekâna bağlı olmayan genel değerlendirme ve çözümler
- Müşteriye uygun olmayan ürünün sunulması
- Müşteri gereksinim analizinin eksik veya hatalı yapılışı

Tanımlanan bu probleme karşı önerilebilecek çözümlerin neler olacağı ise şu şekilde sıralanabilir:

- Bireysel bazlı çözümler
- Sektörel analizlerin doğruluğu
- Mekânsal ve zamansal çözümlerin doğruluğu
- Veri ambarlarında gizli bulunan değerli bilgilerin açığa çıkarılması (Veri Madenciliği)
- Karlılık oranın artırılması
- Dikey büyümenin temin edilmesi
- Spesifik bir ürünü alacak doğru müşteri ile temasa geçilmesi. Kısaca doğru bir şekilde ürün-müşteri ikilisinin oluşturulması
- Daha az gayret ile daha fazla kazanç.

Banka yönetiminin müşteriyi tanımlayan genel öznitelikleri (yaş, cinsiyet, gelir, iş bilgisi vs.) dikkate alarak olası potansiyel müşteri gruplarını belirlemesi ve sıra ile bu gruptaki müşterilere ulaşarak ürün veya hizmet satmak istemesi doğru bir strateji olmayabilir. Bazen de banka çalışanının temel grupta olamayan müşteriye ulaşıp ürün veya hizmet satmak istemesinin birçok olumsuz yönü olabilir. Hem müşteride oluşan imajda hem de bankada oluşan olumsuz sonuçlar şu şekilde sıralanabilir:

- Müşteride bankaya karşı iticilik oluşturma
- Müşterinin gereksiz yere rahatsız edilmesi
- Müşterinin saldırgan ve kızgın davranış sergilemesine neden olma olasılığı
- Banka personelinin gereğinden fazla mesai harcaması
- Banka personelinin yanlış müşteri kitlesi ile muhatap edilmesi
- Mesai ve kaynak israfı vs.

CRM (Customer Relationship Management) uygulamalarında segmentasyonun (müşteri kümeleme, bölümlendirme) sunduğu en büyük fayda temel müşteri grubunu doğru belirlemektedir. Segmentasyon ile aynı zamanda pazarın durumu, hedef müşteri kitlesi, mevcut ve olası tüketici grupları, satış hedefi, satış tutarı gibi bazı analizlerin yanı sıra pazarlama stratejilerinin de yapılması sağlanabilir. Ayrıca segmentasyon yapılırken amaç sistemin ölçülebilirliğinin sağlanması, yüksek kâr elde edilmesi, en az maliyet en fazla kazanç oranının sağlanması, sürdürülebilir olması ve belirlenen hedef müşteri gruplarının kendi içinde homojen bir yapıda olmaları gibi öğeleri içermelidir (Das vd., 2019).

Pazar ve müşteri segmentasyonu ile banka, benzer müşterileri hedef almış olur. Banka, Coğrafi Segmentasyon (yaşanılan semt, çalışılan iş yeri konumu) ve Demografik Segmentasyon (yaş, cinsiyet, meslek) ile hedef müşteri grubunu belirlemeye çalışabilir. Nitekim bu guruplar içerisinde de potansiyel ve hedef müşteri grubuna girmeyen birçok kişi olabileceği gibi bu grupların dışında da çokça temel müşteriler olabilir.

Çalışma müşteriye ait davranışsal ve yaşamsal özelliklerin segmente edilmesi temeline dayanmaktadır. Kısaca sunulan bu projede Müşteri Tanı Sisteminin geliştirilmesi, müşteri segmentasyonunun en iyi şekilde yapılabilmesi, belirli bir ürünü veya hizmeti doğru belirlenmiş hedef müşteri grubuna daha az emek ile satmaya dayanan Karar Destek Sisteminin (Decision Support System) için uygun ortamın oluşturulması amaçlanmıştır. Projede genel amaç müşteri sayısını artırmak değil, var olan müşteriler ile banka kârlılık oranının artırılması için gerekli sayısal verilerin nasıl hazırlanacağına ilişkin ortamın oluşturulmasıdır. Maksat bir veri madenciliği algoritmasının belirli bir veri seti üzerinde denenmesi ve elde edilen sonuçların irdelenmesi ve diğer uzaman sistemlerle kıyaslanması değildir. Amacımız test model önermek, bu konuda sektörü bilgilendirmek ve belirli bir prototip çalışma ile ileri aşama uygulamalara yardımcı olmaktır.

Makalenin bundan sonraki kalan kısmında dört bölüm daha vardır. Takip eden ikinci bölüde yapılan literatür taraması, üçüncü bölümde banka müşterilerinin daha farklı tanınması ve tanımlanması için önerilen farklı tanımlayıcılar bulunmaktadır. Sonraki bölümde önerilen projenin somut uygulama alanları ve elde edilebilecek getirilerinin neler olabileceğine yer verilmiştir. Makine öğrenmesi ve veri madenciliği alanındaki yöntemlerin bankacılık sektörüne nasıl uyarlanabileceği ile alakalı teorik ve açıklayıcı bilgilere yer verilmiştir. Son bölümde de özete ve değerlendirmelere yer verilmiştir.

2. Benzer Çalışmalar

Bilgi teknolojileri kullanılarak bankacılıktaki CRM alanında hem akademik çalışmalar vardır hem de ticari uygulamalar vardır. Bu alanda yapılan çalışmalar iki farklı başlık altında incelenmiştir.

2.1. Akademik Çalışmalar

Literatürde 2010 yılından bu yana yayınlanmış bulunan 23 adet araştırma makalesi, 31 konferans ve sempozyum bildirileri, 4 doktora ve yüksek lisans tezleri ile 2 kitap çalışmaları bir araya getirilmiş ve içlerinden değerli bulunan çalışmalar incelemeye tabi tutulmuştur. Bu çalışmaların çoğu birbirleri ile benzerlik göstermektedir. Bu çalışmalardan bir kaçının açıklamaları kronolojik sıralamayla yapılmıştır.

Tsiptsis ve arkadaşı 2010 yılında "Data Mining Techniques in CRM: Inside Customer Segmentation" isimli bir kitap hazırlamışlar ve kitaplarında Veri Madenciliği ile CRM uygulamaları, Veri Madenciliği teknikleri, müşteri gruplandırma, telekomünikasyon sektöründe ve perakende satış piyasasında müşteri ve ürünleri kümeleme çalışmalarından geniş bir şekilde bahsedilmiştir (Tsiptsis vd., 2010). Kitabın 6. bölümü ise bankacılıkta CRM ve veri madenciliği uygulamalarına ayrılmıştır. Bu bölümde detaylı olarak bankacılıkta gruplara ayırma, kredi çekimleri, ATM cihazlarının kullanımı, kredi kartı kullanımı gibi veriler sayesinde veri setinin nasıl oluşturulacağı, madencilik algoritmalarının nasıl kullanılacağı gibi detaylardan bahsedilmektedir. Daha çok finansal alanda hangi veriler ile hangi bilgi elde edilir başlığı altında somut uygulamalara yer verilmiştir.

2011 yılında Vivek Bhambri isimli bir araştırmacı Veri Madenciliği yöntemlerinin bankacılıkta kullanılan yöntemlerini genel olarak özetlemiş ve somut uygulamaların neler olduğunu göstermiştir (Bhambri, 2011) . Sınıflandırma, kümeleme, ilişkilendirme gibi Veri Madenciliği yöntemlerinde neler yapıldığını anlatmıştır. Ayrıca Veri Madenciliği tekniklerinin bankalara ve finansal kurumlara, daha iyi hedef belirlemede, yeni müşteri edinmede, dolandırıcılık tespitini gerçek zamanlı olarak gerçekleştirmede, müşterilere için segment tabanlı doğru ürünler sunmada, müşterilerin satın alma modellerini analiz etme gibi farklı CRM konularında yardımcı olabileceğini önermiştir. İlgili sektördeki CRM, Pazar, risk yönetimi, sahtecilik, yazılımsal destekler konusunda toparlayıcı bilgiler vermiştir. Aynı araştırmacı Krishan ile birlikte aynı yıl stratejik CRM uygulamalarını veri madenciliği teknikleri ile farklı bir çalışmada anlatmıştır (Chopra vd., 2011). Bu çalışma bir öncekinin uygulaması niteliğindedir.

Oliveira, 2012 yılında veri madenciliği destekli analitik CRM üzerine kapsamlı bir doktora tezi hazırlamıştır (Oliveira, 2012). 2012 yılına kadar bu alandaki en kapsamlı çalışmanın Oliveira tarafından İspanya'da yapılan ve İngilizce dilinde hazırlanmış olan bir doktora tezi olduğu görülmektedir. Oliveira, bu çalışmasında perakende satış sektöründe veri madenciliği uygulamasının yapılabilmesi için öncelikle sektörün iyi analiz edilmesi gerektiğini belirtmiş ve müşterilerin nasıl karakterize edilmesi gerektiği konusunda öneriler sunmuştur. Müşterilerin yaşam tarzlarına ve davranışsal özelliklerine göre nasıl gruplara ayrılabileceğini belirtmiştir. Müşterilerin rakip firmalara kaptırılmaması için oluşturulabilecek tahmin modellerinin neler olabileceği vurgulanmıştır.

Ellaban, 2013 yılında hazırladığı bir yüksek lisans tezinde, bankacılık ve telekomünikasyon sektöründe CRM'in detaylarını endüstri mühendisliği disiplini içerisinde değerlendirerek veri madenciliği yöntemleri ile neler yapılabileceğini listelemiştir (Ellaban, 2013). CRM'in dört temel boyutu olan Müşteri tanımlama (*customer identification*), müşteri çekme (*customer attraction*), müşteri koruma (*customer retention*),müşteri geliştirme (*customer development*) üzerine veri madenciliği uygulamalarının neler olabileceği hakkında detaylı açıklamaları kaynakçalar eşliğinde anlatmıştır.

Jayasreeve arkadaşı 2013 yılında o ana dek "A review on data mining in banking sector" isimli bir çalışma ile literatür taraması yapmışlardır (Jayasree vd., 2013). Bankacılık, risk yönetimi, dolandırıcılık tespiti, kârlılık oranı tespiti gibi başlıklar altında yapılan çalışmaları incelemişlerdir.

Sharahi ve Aligholi 2015 yılında özel bir bankadan aldıkları sınırlı bir veri tabanı ile 60 kadar kurumsal banka müşterisi hakkında veri madenciliği uygulamaları yapmışlardır (Sharahi vd., 2015). Şirketlere ait öznitelik olarak da şirket türü, şirketin çalışma alanı, banka ile çalışma süresi, kredi geçmişi, çekilen kredi türleri, çekilen toplam kredi miktarları seçilmiştir. K-means metodu ile veri seti üzerinde kümeleme yaparak müşterileri çeşitli gruplara ayırmışlardır. Çeşitli sayılarda gruplara ayrılan firmalar kendi içlerinde gösterdikleri karakteristik özellikler manuel olarak incelenerek banka stratejisinin belirlenmesi amacıyla çeşitli tavsiyelerde bulunulmuştur.

2015 yılında Bahari ve Elayidom tarafından hazırlanan bir araştırma makalesinde, müşterilerin davranışlarını önceden kestirmeye yönelik bir CRM – Veri Madenciliği uygulaması yapılmıştır (Bahari vd., 2015). Çalışmalarında Irvine'deki California Üniversitesinin ortak kullanıma sunduğu UCI veri setlerinden bankacılık ile ilgili olanı tercih etmişler (Moro vd., 2011) ve Pratik bir kaç uygulama yapmışlardır. Veri setinde, Portekiz'de bulunan 17 bankanın 2008-2010 yılları arasında 79.354 müşteri hakkında elde ettiği veriler vardır. Her bir müşteriye ait 58 öznitelik bulunmaktadır. Fakat bu verilerden 45.211 müşteri ve 16 öznitelik seçilmiştir. Müşterilerin birikimlerini uzun vadeli olarak çalıştıkları bankada tutup tutmayacakları ile ilgili banka çalışanları tarafından telefon ile aranmış ve sözlü olarak kontrat yapılmak istenmiştir. Müşteriler teklifi ya kabul etmiş ya da reddetmiştir. Verilen bu kararlar veri setinin etiketi olarak belirlenmiştir. Ayrıca veri setinde her bir müşteriye ait 16 adet öznitelik bulunmaktadır. Bunlar müşterinin demografik

özellikleri, yıllık ortalama işlem hacmi, meslek türü, borç ödeme durumu, müşteri ile son görüşme tarihi, müşterinin bankaya olan borç durumu gibi öznitelikerdir. Sadece Naïve Bayes ve Neural Networks sınıflandırıcılarını kullanarak hazır veri seti üzerinde doğruluk (*accuracy*), hassasiyet (*sensitivity*) ve netlik (*precision*) testlerini yine hazır bir uygulama olan Weka yazılım ortamında yapmıştır. Çalışmasında Neural Networks'in daha iyi sonuçlar verdiğini göstermiştir. Bu çalışma hazır veri seti üzerinde yapılmış deneysel uygulamaların ve karşılaştırmaların yapıldığı bir çalışma olmuştur.

Mitik'in 2016 yılında sunduğu bir konferans bildirisinde bankaya ait ürünlerin nasıl pazarlanması gerektiği ile ilgili veri madenciliği tabanlı bir satış stratejisi önerilmiştir (Mitik vd., 2016). Çalışma iki aşamada yapılmıştır. İlki müşterilerin doğru kümelere ayrılması ve sonra ürün ve iletişim kanalları ile ilgili bir sınıflandırma modelinin ortaya konulmasıdır. Sunulan ürünler hakkında doğrudan bilgilendirme ve satış kanallarının müşteri karakteristiğine uygun seçilmesi gerektiği hipotezi ile sınıflandırma çalışması yapılmıştır. Müşteri ile 3 farklı kanaldan (telefon, SMS mesajları ve e-posta) iletişime geçilebileceği gerçeği ile çalışmasını yapmış ve umut vadeden sonuçlar alındığını göstermiştir.

John, 2016 yılında gerçek zamanlı olarak dolandırıcılık tespitinin nasıl yapılabileceği üzerine bir çalışma gerçekleştirmiştir (John vd., 2016). Sigorta, kredi kartı ve hesaplar üzerinde yapılan dolandırıcılıklar içerisinde veri madenciliği teknikleri ile hangi belirtilerin dolandırıcılık girişimini işaret ettiğini belirlemeye çalışmışlardır.

Amania ve Fadlallab tarafından 2017 yılında yayınlanmış olan çalışmada finansal alanda veri madenciliğinin teorik yaklaşımlarından ve pratik uygulamalarından bahsedilmiştir (Amani vd., 2017). Öncelikle veri madenciliği teknikleri kategorilere ayrılıp açıklanmış ve finansal sektörün içeriği tanımlanmıştır. Kullanılan madencilik teknikleri ile bankacılık alanında tespit, tahmin ve reçetelendirme aşamalarının nasıl olacağı anlatılmıştır. Ayrıca bankacılık alanında değerli bilgilerin nasıl elde edilebileceği konusunda teorik yöntemler bilimsel bir dille açıklanmıştır.

Birçok çalışma veri madenciliği tabanlı CRM alanına yeni bir yaklaşım getirmekten ziyade tekrar niteliğindedir. Bu alandaki birçok çalışmanın bir birine benzediği, birinin diğerlerinden anlatım farklılıkları olduğu görülmüştür.

2.1. Ticari Uygulamalar

Bu alanda faaliyet gösteren yazılım firmalarının dünya genelinde çok da fazla olmadığı görülmektedir. Yapılan araştırmalara göre yaklaşık olarak 21 firma bu alanda hizmet sunmaktadır. Bu firmaların 12 tanesi ABD'de 5 tanesi de Avrupa Birliği ülkelerindedir. Genel olarak bankalar kendi içlerinde bir birim kurarak konusunda uzman bilişim personelini istihdam etmektedir. Bankalar genel olarak müşterilerine ait gizli kalması gerektiğini düşündüğü özel bilgilerin, işlem hacimlerinin, mali yapılarının ve hitap ettiği sektöre ait detaylı verilerin başka yazılım firmaları tarafından elde edilmesini istememektedir. Gerçi bu tür yazılım firmaları her ne kadar verilerin kendi firmaları tarafından yedeklenmediğini, ulaşılmadığına garanti verse de bankalar bu konuda oldukça dikkatli davranmaktadır. Hatta bu konuda yeri gelmişken işe yarayabilecek bir tavsiye de şu olabilir: bankalar müşterisini tanımlayıcı kimlik bilgileri ile veri madenciliği işlemleri yapmak yerine her bir müşterisine sadece bu iş için kullanılacak bir tanımlayıcı numara verebilir. Yazılım firmaları da sadece bu numaralar üzerinden işlem yaparlar.

StatSoft firmasının dünyaca ünlü yazılımı STATISTICA Data Miner (Statistica, 2019) bu alandaki en bilinen ve yaygın olan yazılımdır. Eş zamanlı olarak çalışan ve görsel sonuçlar sunan bu yazılım farklı veri tabanları ile senkronize çalışabilmektedir. İlişkisel veri tabanları, Oracle, DB2, MySQL, Teradata, SAP, Hadoop, OSI PI, LIMSs (*Laboratory Information Systems*) ve MESs (*Manufacturing Execution Systems*) gibi sistemlere uyumludur. Ayrıca STATISTICA C#, C/C++, Java ve VB.NET gibi programlama dilleri ile etkileşim içerisinde olabilir. Ekseriyet finansal kuruluşlar için tasarlanan bu yazılım, sahtecilik ve dolandırıcılığın tespit edilebilmesi, risk durumlarının anlaşılması, sofistike ve otomatik risk modellerinin oluşturulması, homojen müşteri gruplarının tahmin ve tespit edilmesi, ikili indikatörler arasında bulunan ve tahmin edilmesi zor korelasyonların bulunması gibi temel özelliklere sahiptir. Yazılım paketi içerisinde farklı bölümler vardır.

Örneğin, *STATISTICA Data Miner's General Classifier and Regression Tool* isminden de anlaşılacağı müşterileri sınıflara veya gruplara ayırmak için kullanılan bir araçtır. Ayrıca genel sınıflandırıcısı ve regresyon araçlarıyla müşteri davranışını öngörme, en kârlı ve sadık müşterilerinin kimler olduğunu tespit etme, kimlerin borçlarını vaktinde ve tam ödeyebileceğini kestirmeye yardımcı olur. *STATISTICA Data Miner's Cluster Analysis Tool* aracı, veriler içerisindeki gizli bulunan şekilleri, kümeleri ve kalıpları bulmak için kullanılır. Bu araç içerisinde Generalized EM (*Expectation Maximization*) ve K-means gibi çeşitli kümeleme algoritmaları vardır. *STATISTICA General Neural Networks Explorer* modülü ise gelişmiş analizler ve Akıllı Sihirbaz (*Intelligent Wizard*) sayesinde gizli yapıyı tespit etme ve tahmin etme gibi araçları sunar.

Data Miner's Feature Selection and Variable Filtering modülü de PCA (*Principal Components Analysis*) ve Partial Least Squares gibi yöntemlerle en önemli değişkenleri ortaya çıkaran ve veri setini basit ve anlaşılabilir hale getiren özelliklere sahiptir. Ayrıca STATISTICA Data Miner, Bağlantı fonksiyonları ile Doğrusal ve Doğrusal Olmayan Çoklu Regresyon, Yapay Sinir Ağı, ARIMA, Fourier Analizi ve bunun gibi gelişmiş özelliklerle çeşitli tahminler yapar.

STATISTICA yazılım paketi banka, bölge, ülke ve işlem hacmine dayalı bir fiyatlandırma politikası olduğu için bu firmadan alınacak hizmetin ücreti konusunda bir netlik yoktur. Ayrıca bu firmadan yazılıma ait talep edilecek ekstra niteliklerle yazılımın fiyatı değişmektedir. Yapılan araştırmalara göre tahmini yıllık kurulum, uyarlama, personel eğitimi, bakım, destek gibi anahtar çözüm masraflarının 2 milyon USD civarında olduğu görülmektedir.

Yeni Zelanda'da kurulmuş 11Ants Analytics Ltd (www.11antsanalytics.com) firması bilişim konusunda uzman olmayan personelin de kullanabileceği bir finans sektörüne güçlü ve yeni özellikleri olan yazılım sunmaktadır.

Data Mining with SAS® Enterprise Miner (www.sas.com) ürünü ise benzer bir şekilde yolsuzluk tespiti yaptığı gibi müşteri ve piyasa taleplerinin önceden tahmin edilebilmesine de yardımcı olmaktadır. Ayrıca bu firma, yaptığı ürün tanıtımında mevcut müşterileri korumayı da taahhüt ettiğini söylemektedir.

3. Müşteri Tanıma/Tanımlama İle İlgili Yeni Yöntemler

Banka müşterileri zaman içerisinde kendileri ile alakalı birçok veriyi bilerek ya da bilmeyerek banka ile paylaşır. Bu veriler mudilerin yaşamsal alışkanlıklarını, kişilik özelliklerini, hayat tarzlarını, mali durumlarını, alım güçlerini, yönelimlerini içermektedir. Bu veriler önerilen projenin temelini oluşturmaktadır.

Periyodik olarak müşteri bilgilerinin güncellenmesi gerekmektedir. Bu nedenle verilerin dinamik bir yapıda olduğu ve doğal olarak da elde verilen sonuçlarla birlikte sistem tarafından önerilenlerin zamana bağlı olarak değişkenlik göstereceği bir gerçektir. Ayrıca oluşturulacak veri bütünlüğü (*data integrity*) sayesinde tüm işlemler kolaylıkla gerçekleştirilebilecektir.

Müşteriyi en iyi tanımlama ve tanıma yöntemi müşteriden alınabilecek bilgilerle mümkündür.

Yeteri kadar veri olmadan hiçbir yapay karar destek sisteminin başarılı olması mümkün değildir. Ne kadar çok veri var ise o kadar doğru müşteri sınıflandırması ve kararlar verilir. Mümkün olduğu kadar müşteriye ait bilgilerin toplanması gerekmektedir. Elbette ki bu bilgiler önceden kategorize edilmelidir.

Banka müşterilerini Gerçek ve Tüzel olmak üzere iki sınıfa ayırıp bu kişilerin olası özelliklerini (öznitelikleri) ayrı ayrı inceleyelim.

3.1. Gerçek Kişiliklerin Öznitelikleri

İsim, yaş, cinsiyet, gelir bilgisi, çocuk sayısı, medeni durum, eğitim durumu, iş bilgisi, yaşanılan semt gibi temel öznitelikler bir müşteriyi daha iyi tanımak ve tanımlamak için yeterli değildir. Mutlaka daha farklı ve detaylı verilerin de alınması gereklidir. Üzerinde çalışılması planlanan projede müşterilere ait olası öznitelikler şunlar olabilir:

- Sağlık durumu
- Davranış durumları:
    - Belirli bir zaman diliminde telefon bankacılığını kullanma sıklığı
    - Kullanım süresi
    - Gece veya gündüz kullanım durumu
    - Internet bankacılığına bağlanma süresi
- Banka şubesinde bekleme süresi. Bu süre şubeye giriş ve işlem bitimine kadar geçen süreyi kapsamaktadır
- Banka şubesine gitme sıklığı
- Gidilen banka şubesinden müşterinin yaşadığı veya çalıştığı bölgenin tespiti
- Aylık ve yıllık İşlem potansiyeli
- Bankamatik kullanım sıklığı
- Hangi şehirdeki ve semtteki bankamatikleri kullandığı
- Kredi kartı ile mekâna bağlı alış veriş hareketleri
- Kredi kartı ile zamana bağlı alış veriş hareketleri
- Kredi kartı ile alınan tüketim malzemeleri türü (Ek kredi kartı için de aynı analiziler yapılabilir)
- Sanal kredi kartı kullanım miktarı, oranı
- Müşteri temsilcisini arama sıklığı
- Müşteri temsilcisini arama nedeni
- Bankaya E-posta gönderme sıklığı
- Bankaya E-posta gönderme nedeni
- Yazılı iletişimde ifade gücü, imla kurallarına uyup uymadığı vs.
- Hangi numaradan (şehir ya da cep numarası) aradığı
- Ses tonu ile kişisel özellik analizi (nazik, agresif, kaba, medeni, eğitimsiz vs.)
- Sosyal medya aracılığı ile ulaşıp ulaşmadığı

- Fatura ödeme türleri
- Ödemelerini ne düzende yapabildiği
- Hangi şubeleri ziyaret ettiği
- Şubeleri ziyaret etme nedeni
- Bankaya yaptığı şikâyetlerin türü
- Bankaya yaptığı şikâyetlerin nedeni
- Problemin hukuki süreçte olup olmadığı
- Şikayetvar.com gibi sitelere banka hakkında şikâyette bulunup bulunmadığı
- Bankadan alınan hizmet türleri
- Alınan hizmetlerden duyulan memnuniyet düzeyi
- Bankanın önerdiği ürün ve hizmetlerden hangilerini reddettiği
- Bankadan alınan ürün ve hizmet bilgileri ve türleri
- Müşteri temsilcisinin müşteriyi tel ile arayıp ürün veya hizmet satmak istemesi karşısında müşterinin tutumu (Bu bilgi özellikle Makine Öğrenmesi ile yapılan sınıflandırma algoritmalarında kullanılabilir.)
- Kredi kartı ve diğer faturalarını düzenli ödeme niteliği
- Mobil cihazlar ile Internet Bankacılığı kullanıyor mu?
- Kullandığı Internet Bankacılığı servisleri
- Aylık Internet Bankacılığı işlem kapasitesi
- Turist olarak yurt içinde nerelere gittiği
- Turist olarak yurt dışında nerelere gittiği
- Turistik gezilerinde kaç gün kaldığı ve ne kadar harcama yaptığı
- Ne sıklıkta tatile çıktığı
- Yaz turizme gidip gitmediği
- Kış turizme gidip gitmediği
- Yurt dışı turizme gidip gitmediği
- Akaryakıt alım sıklığı
- Hangi akaryakıt firmasından alış veriş yaptığı
- Birim zaman diliminde banka ile geçirdiği süre (Telefonda, şubede, webde)
- Vadesiz hesap miktarı
- Vadeli hesap miktarı
- Altın hesabı
- Son birkaç aylık veya yıllık mevduat ortalaması
- Son birkaç aylık veya yıllık işlem miktarı
- Son birkaç aylık veya yıllık işlem hacmi

Yukarıda bahsi geçen birçok özniteliği alabileceği cevaplar kategorize edilmelidir. Ayrıca bazılarının cevapları ise reel sayılara dönüştürülmelidir. Elbette ki Makine Öğrenmesi gibi yapay zeka alanında elde edilen verilerin tamamı kategorize edilmiş nominal veya reel olmalıdır ki gerekli yapay karar destek sistemleri inşa edilebilsin.

3.2. Tüzel Kişiliklerin Öznitelikleri

Gerçek kişilikteki özniteliklerden uygun olanların seçilerek Tüzel kişiliklere ilave edilmesinin yanında olası diğer öznitelikleri şunlar olabilir:

- Şirket kredi kartı harcama bilgileri
- Şirket kredi kartı ile harcama türleri

- Şirket kredi kartı ile harcama yapılan mekân bilgisi
- Şirket kredi kartı ile yapılan harcama miktarı
- Şirketin bankada bulunan nakit durumu
- Şirket çalışanlarına ödenen maaş bilgileri
- Şirket çalışan sayısı
- Şirket çalışanlarının yukarıda bahsedilen (Gerçek Kişilik Öznitelik) bilgileri
- Şirket araçlarının akaryakıt alım bilgileri

(alım miktarı, alım yapılan mekân bilgileri, alım sıklıkları)

- Şirket araç sayısı

Elbette ki gerçek ve tüzel kişilikleri tanımlayıcı çok daha fazla veri müşteriler tarafından farklı zaman, mekân ve durumlarda banka ile paylaşılmaktadır. Burada sadece bir prototip çalışma amaçlandığı için yukarıda listelenenlerin yeterli olduğu düşünülmektedir. Ayrıca proje aşamasında kişilere ait her bir öznitelik sayısal değerlere dönüştürülecek, kullanılması uygun olmayan öznitelikler silinecek, temin edilemeyen veriler en uygun verilerle doldurulup sistemin sürekliliği ve çalışması sağlanmalıdır. Bu arada projenin *Big Data* alanına girdiğini belirtmekte fayda vardır.

Burada gerçek ve tüzel kişilere ait tanımlayıcı bilgilerin (özniteliklerin) fazla oluşu yapay öğrenme mekanizmaları için genel bir problem teşkil etse de, proje büyük veri kapsamında ele alındığında ve gerekli yazılımsal ve donanımsal araçlar da kullanıldığında başarılı sonuçların elde edileceği kesindir.

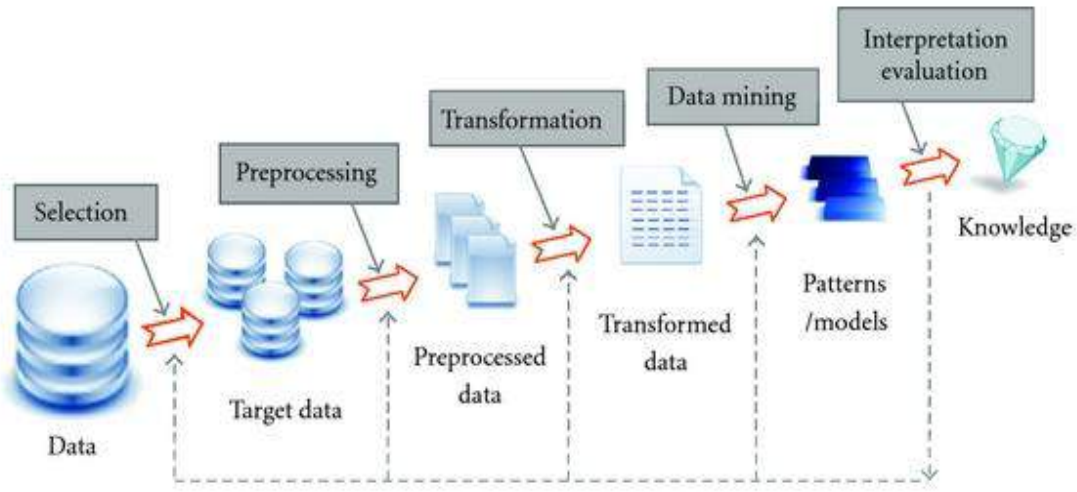
Ayrıca çok boyutlu düzlemlerde yani müşterilere ait çok fazla özniteliklerin bulunduğu ver setlerinde öğrenme ve sınıflandırma araçlarının başarılı sonuçlar veremeyeceği ya da bazı sapmaların meydana geleceği literatürde tanımlanmıştır. Bu gibi durumlarda boyut indirgeme algoritmaları kullanılmaktadır (Ho vd., 2009). LDA (*Linear Discriminant Analysis*), PCA (*Principal Component Analysis*) ve Faktör Analizi gibi yöntemler ile sınıflandırma başarısına etki etmeyen ya da çok az etkisi olan öznitelikler uzaydan çıkarılmaktadır.

Yine aynı şekilde k-NN gibi örnek tabanlı ve uzaydaki uzaklığı temel alarak sınıflandırma yapan öğreniciler çok boyutlu ortamlarda yeterince başarı sergileyememektedir. Bu gibi veri setleri için tembel öğrenicilerden (*lazy learners*) daha çok daha çok akıllı öğrenicilerin (*eager learners*) tercih edilmesi gerektiği bir gerçektir (Bulut vd., 2017). Karar ağaçları, Çok Katmanlı Algılayıcılar, Naive Bayes sınıflandırıcısı, kolektif öğreniciler (*Ensemble Learners*) bunlardan bir kaçıdır. Büyük veri için literatürde önerilmiş ve gerekli testleri yapılmış bazı sınıflandırıcılar vardır (Fernández vd., 2018).

4. Bankacılıkta Bilgi Teknolojileri Ve Uygulamaları

4.1. Bilgi Teknolojileri Aşamaları

Veri madenciliği adından da anlaşılacağı üzere, büyük verilerin bulunduğu ortamdan anlamlı bilgilerin çıkarılmasıdır. Diğer bir deyişle veri madenciliği, veri tabanlarından bilgi keşfinin (*Knowledge Discovery in Database, KDD*) yapılmasıdır. Çıkarılan bilgilerin yeni, bariz bir şekilde bilinebilir olmaması, ilgili kişiler tarafından kullanılabilir ve uygulanabilir olması gerekmektedir (Narang vd., 2017). Bilgi keşfi süreci elbette ki veri tabanı yönetimi (*Structure Query Language, SQL*) ve OLAP (*Online Analytical Processing*) ile başlamaktadır. Şekil 1'de görüldüğü üzere bu aşamadan sonra adım adım yapılması gereken bazı işlemler vardır.



Şekil 1. Bilgi keşfi adımları (Terzi vd., 2012)

KDD süreci içerisinde ilgili verilerin seçimi (*Data Selection*) oldukça önemlidir. Araştırma konusu ile ilgili olmayan, aralarında korelasyonun bulunmadığı sebep-sonuç ikililerinin seçimi hatalı veya eksik bilgilerin elde edilmesine neden olacaktır. Her bir bilgi keşfi uygulamasında ilgili ve alakalı verilerin seçilmesi şarttır. Seçim aşamasında zaman, mekân, sektör gibi faktörlerin de göz önünde bulundurulması gerekir. Örneğin 10 yıl önceki statik verilerle günümüz hakkında varsayımlarda bulunmak ya da Türkiye'nin doğu bölgesindeki bireysel müşteri verileri ile batı bölgesini yorumlamak yanlış olacaktır.

Veri Hazırlama (*Data Preprocessing*) aşamasında ise verilerin kullanıma hazırlanabilmeleri için verilerin temizlenmesi ve veri bütünlüğünün sağlanması gerekmektedir. Gürültülü ya da sisteme uymayan veriler silinmelidir. Nominal, tam sayı, reel, mantıksal değişken veriler veri setine nasıl yerleştirilecek bu önemlidir. Esasen bu bölüm veri tabanı bilgisi de gerektirmektedir.

Veri Temizleme (*Data Cleansing*) işlemi, gereksiz verilerin (*noise, outlier*) veri setinden çıkarılmasıdır. Bütün bu işlemler veri bütünlüğünün (*Data Integration*) oluşturulması için önemlidir.

Veri Dönüşümü (*Data Transformation*) ile elde edilen veriler madencilik için uygun formata dönüştürülür. En yaygın dönüşüm şekli 3 tanedir: veri normalizasyonu, verileri birleştirme (*data aggregation*) ve veri ayrıştırma (*data discretization*)

- Bazen nominal değerlere sahip veriler sayısallaştırılabilir, isimsel değerler ile matematiksel işlemlerin yapılması imkansızdır, mutlaka rasyonel veya tam sayı değerlerinin olması gerekir. Bazı durumlarda ise veriler kategorilere ayrılabilir, mantıksal 1-0 (doğru-yanlış) değerlerine de sahip olabilir.
- Kayıp değerler (*missing value*) en uygun değerlerle tolore edilebilir,
- Bazen de normalizasyon yapılabilir. Normalize edilmemiş bir veri seti üzerinde yapılan madencilik işlemleri yanlış sonuçlar verebilmektedir. Normalizasyon için literatürde önerilmiş Min-Max ve Z-Score gibi yöntemler mevcuttur (Bulut, 2017).

Veri Madenciliği: En önemli aşamadır. En uygun metotlar ile veri ambarında bulunan gizli ve değerli bilgi kalıpları açığa çıkarılmalıdır.

Yorum ve değerlendirme ile elde edilen bilgilerin, yargıların, yorumlamaların, sınıflandırmaların gerektiğinde görsel olarak kullanıcıya yorumlama ve değerlendirme amaçlı sunulması gerekmektedir.

Ayrıca yapay öğrenme alanında sınıflandırma uygulamalarının başarımı öğrenicilere sunulan veri setleri ile doğrudan alakalıdır. En başarılı sınıflandırıcının bile veri setinin yapısına bağlı olduğu bilinmektedir.

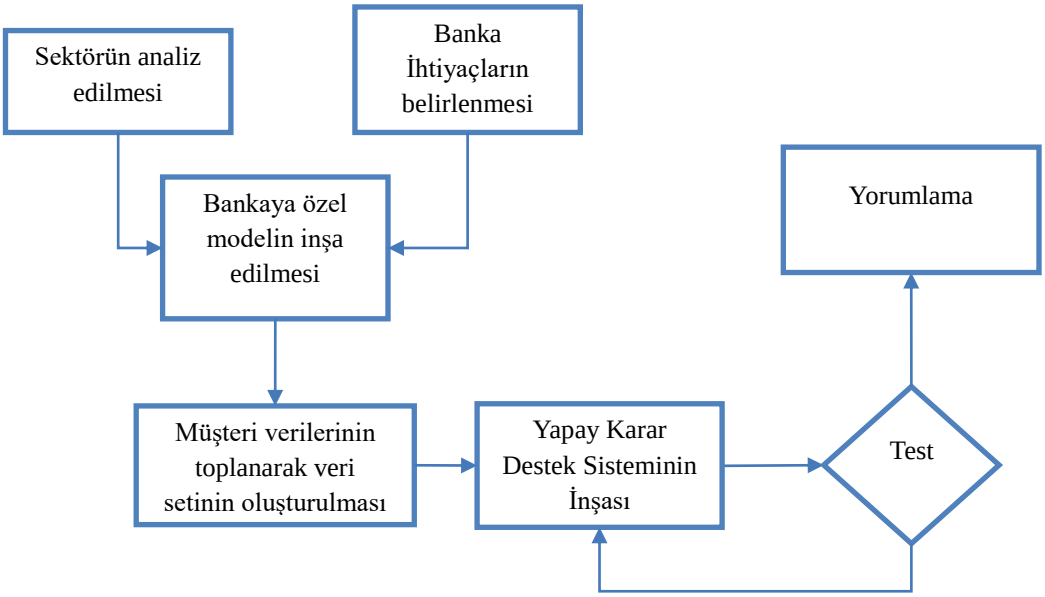
Bu nedenle şu iki hususa dikkat etmek gerekmektedir:

1. En başarılı sınıflandırma algoritmasının belirlenmesi için bir takım deneysel işlemlerin yapılması şarttır. Bu konuda elbette ki Meta Öğrenme ile veri setine uygun sınıflandırıcı yöntemlerinin bulunması bir tercih olabilir. Fakat daha çok literatürde önerilen sınıflandırıcıların başarılar deneysel olarak tercih edilmesi daha doğru olabilir.
2. Esasen sınıflandırma işlemin veri uzayının seyrek (*sparse*) bir yapıda olması sınıflandırma işlemini başarısız kılar. Nitekim banka sahip olduğu teknolojik yapı ile milyonlarca müşterisine ait gün içerisinde yine milyonlarca reel, etiketli, gerçek veri toplayabilmektedir. Bu veriler –belirtmek isteriz ki- doğru ve gerçek verilerdir. İçlerinde sisteme uymayan, yapıyı bozmayan, veriler bulunmamaktadır. Ayrıca kayıp değer de yoktur. Bu nedenle oluşturulacak veri seti oldukça zengin, yoğun (*dense*) ve filtre gerektirmeyen bir yapıdadır.

Bilinmelidir ki sınıflandırma başarısı veri setinin yapısı ile doğrudan alakalıdır. Bu gibi durumlar bir banka için oldukça büyük bir avantajdır.

4.2. Sürecin yaşam döngüsü

Bu çalışmada önerilen sistem dinamik bir yapıda olacak ve belirli periyotlarla müşterilerden elde edilen güncel veriler ışığında kendini yenilemelidir. Doğal olarak değişen ülke ekonomisine, kriz dönemlerine ve siyasal durumlara göre banka en uygun stratejisini zamana bağlı olarak geliştirebilecektir. Ayrıca coğrafi bölgelerin diğer bölgelerden farklı olması nedeni ile o bölgelere uygun farklı ve benzer çalışmalar yapılabilir. Şekil 2'de veri bilimi ışığında yapılabilecek müşteri ilişkileri yönetiminin yaşam döngüsü grafik şeklinde gösterilmektedir.



Şekil 2. Proje Yaşam Döngüsü Aşamaları

Elbette ki önerilen bu model sektör ve banka ihtiyaçları göz önüne alınarak özgün bir şekle bürünecektir. Projenin başarılı bir şekilde hayata geçirilmesi amacıyla üç ana aşamadan oluşmaktadır:

1. Banka müşterilerine ait birçok veri bir veri setinde (havuzunda) toplanarak sayısallaştırılır, normalize edilir ve kullanılabilir hale getirilebilir.
2. Bu veriler, Veri Madenciliği ve Makine Öğrenmesi alanındaki Sınıflandırma (*Classification*), Tahmin (*Prediction*), Kümeleme (*Clustering*), İlişkilendirme (*Association*), Yarı Denetimli Öğrenme (*Semi Supervised Learning*) gibi alt başlıklardaki algoritmalar ile en uygun model oluşturulabilir.
3. Daha sonra kullanılan model ile müşteriler hakkında gerekli tahminler, yönlendirmeler ve kararlar elde edilebilir.

4.3. Sektörde Kullanılabilecek Veri Madenciliği ve Makine Öğrenmesi Teknikleri

Projede Yapay Zekâ alanındaki Makine Öğrenmesi (*Machine Learning*), Veri Madenciliği ve Bilgi Keşfi (*Data Mining and Knowledge Discovery*) ve Şekil Tanıma (*Pattern Recognition*) disiplinlerinde tanımlanan metotlar seçilerek özgün ve ihtiyaca uygun bir karar destek modelinin geliştirilmesi hedeflenmelidir. Teknik anlamda yapılabilecek projede kullanılacak algoritmaların bulunduğu ana başlıklar şu şekilde sıralanabilir:

1. Sınıflandırma (*Classification*)
2. Kümeleme (*Clustering*)
3. İlişkilendirme (*Association*)
4. Tahmin (*Forecasting*) ve Reel değerlerle Sınıflandırma (*Regression*)
5. Görselleştirme (*Visualization*)
6. İstatistiksel Analiz (*Statistical Analysis*)
7. Zamana Bağlı Dinamik Yöntemler

Yukarıda bahsedilen ana başlıklar için teorik olarak geliştirilmiş ve pratiğe dökülmüş birçok algoritmik yardımcı araç (*tool*) literatürde vardır ve kullanımdadır. Bu araçlardan en uygunları proje için modifiye edilip kullanılabilir. Görüldüğü üzere proje, bilinen istatistiksel ve matematiksel teoremlerden çok bilgi teknolojileri disiplini ile geliştirilebilir.

Veri madenciliği alanında en çok kullanılan teknik olan sınıflandırmada önceden tanımlanmış ve etiketli verilerin bulunduğu bir veri seti gereklidir. Veri setinde bulunan her bir örneğin etiket bilgisi isim veya sayı olabilir. Sınıflandırma sonucu yine nominal bir sınıf etiketidir. Bu metot ile müşteri gruplandırma ve dolandırıcılık tespiti gibi birçok uygulama yapılabilir.

Kümeleme, benzer özellikler gösteren örneklerin bir araya getirilmesi amacıyla kullanılır. Sınıflandırmadan farklı olarak kümelemede örneklerin sınıf etiketlerinin bilinmesine gerek yoktur. Kümeleme işlemi örneklerin öznitelik değerlerine bakılarak

yapılır. Müşterilerin bölgesel, demografik, gelir düzeyleri, işlem hacimleri, harcama miktarları ve yaşam tarzları bakımında gruplara ayrılması belirli aşamalarda yeterli olabilir ama ileri düzey için daha karmaşık işlemlere gerek vardır.

Geliştirilmiş bir CRM uygulaması ile müşteri segmentasyonu yapılabilir. Mevcut müşteri odaklı doğrudan Pazar ve piyasa araştırması amacıyla Müşteri Segmentasyonu yapılabilir. Birbiri ile benzeşen homojen yapıların tespit edilmesi sağlanabilir. Müşteri karakteristiği ve davranışları analiz edilebilir, en iyi müşteri segmentasyonu bu sayede yapılmış olur. Ayrıca sunulması planlanan her bir ürün için hedef gruplarının belirlenmesi ve bu sayede karlılığı en üst seviyeye çıkarmak için segmente edilmiş homojen gruplara belirlenen en uygun ürün ve hizmetlerin satılması sağlanmış olur. Örneğin belirli bir coğrafi bölgede yaşayanlar ve belirli bir meslek grubundaki çalışanların kendilerine özgü talepleri vardır. İşte birbiri ile benzerlik gösteren bu müşteriler için gerekli özel çözümlerin geliştirilip sunulması gerekmektedir.

Satılacak belirli bir ürün veya hizmet için homojen yapıdaki temel müşteri grubunun belirlenmesi oldukça önemlidir. Kümeleme ile müşteri profil ve segmentlerine en uygun gruplama yapılarak bir ürün benzer öznitelikler sergileyen diğer müşterilere de satılabilir. Banka personeli doğru hedef müşteri grubunu bu sayede bulabilir, az mesai ile ürün satışı gerçekleştirilebilir.

İlişkilendirme (Association), iki veri, bilgi, davranış veya olayın aralarındaki korelasyonun diğer bir deyişle ilişkinin bulunması için kullanılan bir yöntemdir. Sepet Analiz Yöntemi (*Basket Analysis*) ile belirli bir ürün için yeni potansiyel müşteri gruplarını tespit etme işlemidir. Diğer bir deyişle bunu alan müşteri şunu da alır mı sorusunun cevabını bulma yöntemidir. Bir ürünü veya hizmeti alan müşteri (birey ya da KOBİ) başka hangi ürünü almış mıdır? İlişkilendirme bize korelasyon analizi yapmamızı sağlar. Bankadan bir ürünü veya hizmeti alan bir müşteri, büyük ihtimalle başka bir ürünü ya da hizmeti de alabilecektir. Ya da bir müşterinin belirlenen bir ürünü alma olasılığı nedir. Bu müşteri kredi kullanacak olsa ne kadar kredi kullanabilir? İşte bunun gibi büyük veri setlerinde gizli bir takım ikili işlem ilişkileri vardır. Büyük verilerin içerisinden bunu çıkarılması gerekmektedir.

Elde edilen bu gibi değerli bilgiler daha sonra fiyat belirlemede, ürün ve hizmet sunmada, kampanya belirleme gibi işlemlerde kullanılabilir. Ayrıca ücretlendirme, satış ve market stratejisinin belirlenmesi gibi alanlarda kesin kararların verilmesini sağlar. Dikey ve çapraz satışın artırılmasında bu yöntemler oldukça faydalıdır. Bilindiği üzere dikey satış (*Up Selling*), bir ürünü alan müşterinin o ürünün diğer eklenti ürünlerini de alma ihtimalini belirleme ve bu sayede doğru satış stratejisi belirleme yöntemidir. Çapraz satış (*Cross Selling*) ise bir ürünü alan müşterinin, alma ihtimali yüksek diğer ürünlerin belirlenmesini sağlama ve bu sayede hem doğru satış stratejisi belirleme hem de kârlılığı artırma tekniğidir.

Banka bir ürünü veya hizmeti, müşteriye ulaşarak satmak isteyen banka çalışanı belki önceden 100 kişiden 2-3 kişiye satabilirken bu yöntemler ile doğru bir şekilde belirlenmiş temel kitledeki müşterilere ulaşmış olacak ve satış oranı yükselebilecektir.

Tahmin (*Forecasting*) için Regresyon tekniği sistemine uyarlanabilir. Regresyon analizleri bağımlı veya bağımsız değişkenlerin aralarındaki ikili ilişkilerin incelenmesi için kullanılabilir. Veri madenciliğinde değişkenler öznitelik (*attribute, feature*) olarak karşımıza çıkmaktadır. Örneğin satış hacmi, depolama maliyetleri, başarısızlık riski oranı gibi durumların önceden tahmin edilmesi oldukça zordur. Çünkü bunlar karmaşık ve belirlenmesi gün nedenlerden ötürü meydana gelmiş olabilir. Geleceğe yönelik analiz çalışmaları için daha karmaşık (örneğin *Logistic Regression, Decision Trees,* ya da *Neural Networks*) yöntemler tercih edilebilir. Bu projede bir ürününüzü veya hizmetinizi satın alan müşterinin sunulan diğer bir ürünü alma olasılığını bu sayede bulunabilir. Eğer olasılık belirli bir oranın üzerinde ise bu müşteri sunulan ürünü büyük ihtimalle alacaktır. Alınan bir ürününün diğer bir ürün üzerindeki alım etkisinin analizi yapılabilir. Satılan ürünlerin birbirlerini etkileme durumlarının incelenebilir. Bir müşterinin ne kadar süre banka ile çalışacağını tahmin etme (*Customer Lifetime Value*) ve mevcut her bir müşteriden elde edilebilecek kâr oranı tahmin etme bu kategoride incelenebilir.

Reel değerlerle Sınıflandırma işleminde ise sınıflandırmada olduğu gibi önceden tanımlanmış bir veri seti gereklidir. Etiket değerleri sınıflandırmadan farklı olarak sayısal değerlere sahiptir. Elde edilen hesaplamaların sonucu da doğal olarak reel değerler olacaktır. Çekilen bireysel veya ticari kredilerin bankaya ödenebilmesi ile ilgili risk analizinde bu yöntem kullanılabilir.

Ayrıca işletme tarafında kendi müşterilerine ait bir takım analizler de şu şekilde yapılabilir:

- Müşteri alım davranışları, gereksinimleri nelerdir.
- Müşteri banka değiştirmeden önce hangi tür davranışlar sergiler.
- Hangi ürün grupları hangi müşteri kitleleri tarafından satın alınır.
- Hangi tür davranışlar sahtecilik veya dolandırıcılık göstergesidir. Bununla ilgili daha önce sahtecilik veya dolandırıcılık yapan müşterilerin ön alım davranışları ve işlemleri sistemin eğitilmesinde kullanılmalıdır.
- Borç ödemede yüksek risk taşıyan müşteri tanımının yapılması.
- Mevcut müşterilerin ihtiyacı olabilecek servis, ürün ve hizmetler nelerdir?
- Sadık müşteri tanımı için tüm ihtiyaçlarını aynı bankadan karşılayan müşterilerin tespit edilmesi.

Veri madenciliği, makine öğrenmesi ve şekil tanımada gibi veri bilimi dallarında her geçen gün içerisinde mevcut öğrenme metotlarının iyileştirilmiş bir versiyonları ve bunun yanında yeni yeni teknikler literatürde yerini almaktadır. Elbette ki bu çalışmalar makaleye ilave edilebilir. Burada önemli olan yeni metotlarla var olan bir sistemi hızlandırmaya ya da daha doğru sonuçlar almaya çalışmak değildir. Amaç var olan yöntemlerin yanı sıra farklı bir modelin inşasına zemin hazırlamaya çalışmaktır.

5. Sonuç

Sonuç olarak banka personeli daha az bir mesai ile alım yapacak hedef müşteriye önerilen bu yazılım sayesinde ulaşabilecektir. Önerilen sistem dinamik bir yapıya sahip olacaktır. Zamana bağlı olarak tüzel ve gerçek kişilerden elde edilen veriler veri havuzunu zenginleştirecek, satılması planlanan ürün ve hizmetlerin en az mesai ile doğru zamanda doğru kişilere satışının yapılması sağlanmış olacaktır.

Esasen bu çalışma hem bir araştırma makalesi hem de bir derleme makalesi özelliği taşımaktadır. Çalışmada önerilen model uygulanabilir bir yapıdadır fakat bilindiği üzere uygulama yapabilmek için gerçek verilere ihtiyaç vardır. Türkiye'de bir bankanın kendi müşterilerine ait verileri üçüncü şahıs ve şirketlerle –ki bu akademik çalışmalar olsa dahi– paylaşması mümkün gözükmemektedir. Bu nedenle çalışma ağırlık olarak teorik bir modelin önerilmesi şeklindedir. Bunun yanında farklı olarak müşterilerin yaşam alışkanlıkları ve davranışlarının modellenmesi üzerine kurulu bir veri toplama modeli de şekillendirilmeye çalışılmıştır.

Bu çalışmada önemli olan veri madenciliği, makine öğrenmesi, şekil tanıma ve istatistik alanında yeni algoritmaların geliştirilmesi, kullanılan tekniklerin iyileştirilmesi değildir. Ya da hızlı ve kapasitesi yüksek bilgisayarların bu alanda kullanılması da değildir. Önemli olan bankacılık sektöründeki ihtiyaçların en iyi şekilde belirlenmesi, kullanılması gereken modelin en uygun şekilde oluşturulması ve elde edilen sayısal değerlerin iyi yorumlanması gerekmektedir.

Üzerinde çalışılan proje, bankalara ait finansal işlemlerden ve milyonlarca müşterilere ait ham verilerden makine öğrenmesi ve veri madenciliği metotları ile sektöre ait değerli bilgilerin çıkarılması amacıyla gerekli alt yapının hazırlanmasını içermektedir. Bu çalışmada önerilen veri toplama yöntemleri, elde edilen verilerin sayısallaştırılması, her bir yöntemin sektöre nasıl uyarlanacağı ile ilgili yapay destek modelinin ufuk açıcı özellikte olduğu düşünülmektedir. Burada belirtmek isteriz ki projede asıl amaç müşterileri fişlemek, belirli kategorilere ayırmak, kişisel verileri elde etmek, özel yaşam alanları ile ilgili verileri istiflemek olmamalıdır. Proje kapsamında yapılan işlemler sadece banka tarafında gerçekleşecek ve hiçbir şekilde müşteriden elde edilen veriler başka kurum ve kuruluşlarca paylaşılmamalıdır. Çalışmada ele alınan konunun büyük projeler için bir prototip olması nedeniyle daha farklı bir öneme sahip olduğu düşünülmektedir.

Kaynakça

Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. International Journal of Accounting Information Systems, 24, 32-58.

Bahari, T. F., & Elayidom, M. S. (2015). An efficient CRM-data mining framework for the prediction of customer behaviour. Procedia Computer Science, 46, 725-731.

Bhambri, V. (2011). Application of data mining in banking sector. IJCST, 2(2), 199-202.

Bulut, F., & Amasyali, M. F. (2017). Locally adaptive k parameter selection for nearest neighbor classifier: one nearest cluster. Pattern Analysis and Applications, 20(2), 415-425.

Bulut, F. (2017). Örnek tabanlı sınıflandırıcı topluluklarıyla yeni bir klinik karar destek sistemi. Journal of the Faculty of Engineering and Architecture of Gazi University, 32(1), 65-76.

Chopra, B., Bhambri, V., & Krishan, B. (2011). Implementation of data mining techniques for strategic CRM issues. Int. J. Comput. Technol. Appli, 2, 879-883.

Das, S., & Mishra, M. (2019). The Impact of Customer Relationship Management (CRM) Practices on Customer Satisfaction. In Business Governance and Society (pp. 43-54). Palgrave Macmillan, Cham.

Ellaban, Mahmoud Ayesh Abu, (2013), The Role of Data Mining Technology in Building Marketing and Customer Relationship Management (CRM) for Telecommunication Industry, Islamic University, BusinessAdministration.

Anshari, M., Almunawar, M. N., Lim, S. A., & Al-Mudimigh, A. (2018). Customer relationship management and big data enabled: personalization & customization of services. *Applied Computing and Informatics*.

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). Imbalanced Classification for Big Data. In Learning from Imbalanced Data Sets (pp. 327-349). Springer, Cham.

Hasan, A. A. T. (2018). Customer Relationship Management (CRM) Practices of City Bank in Customer Retention Perspective in Bangladesh. Global Journal of Management And Business Research.

Ho, C. T. B., & Wu, D. D. (2009). Online banking performance evaluation using data envelopment analysis and principal component analysis. Computers & Operations Research, 36(6), 1835-1842.

Kasemsap, K. (2018). Facilitating customer relationship management in modern business. In Encyclopedia of Information Science and Technology, Fourth Edition (pp. 1594-1604). IGI Global.

John, S. N., Anele, C., Kennedy, O. O., Olajide, F., & Kennedy, C. G. (2016, December). Realtime Fraud Detection in the Banking Sector Using Data Mining Techniques/Algorithm. In Computational Science and Computational Intelligence (CSCI), 2016 International Conference on (pp. 1186-1191). IEEE.

Jayasree, V., & Balan, R. V. S. (2013). A review on data mining in banking sector. American Journal of Applied Sciences, 10(10), 1160.

Mitik, M., Korkmaz, O., Karagoz, P., Toroslu, I. H., & Yucel, F. (2016, December). Data Mining Based Product Marketing Technique for Banking Products. In Data Mining Workshops (ICDMW), 2016 IEEE 16th International Conference on (pp. 552-559). IEEE.

Moro, S., Laureano, R., & Cortez, P. (2011). Using data mining for bank direct marketing: An application of the crisp-dm methodology. In Proceedings of European Simulation and Modelling Conference-ESM'2011 (pp. 117-121). EUROSIS-ETI.

Narang, S. K., Kumar, S., & Verma, V. (2017). Knowledge discovery from massive data streams. In Web semantics for textual and visual information retrieval (pp. 109-143). IGI Global.

Nguyen, N. T., Hoang, D. H., Hong, T. P., Pham, H., & Trawiński, B. (Eds.). (2018). Intelligent Information and Database Systems: 10th Asian Conference, ACIIDS 2018, Dong Hoi City, Vietnam, March 19-21, 2018, Proceedings (Vol. 10751). Springer.

Oliveira, V. L. M. (2012). Analytical customer relationship management in retailing supported by data mining techniques.

Sharahi, M., & Aligholi, M. (2015). Classify the data of bank customers using data mining and clustering techniques (Case study: Sepah bank branches Tehran). Journal of Applied Environmental and Biological Sciences, 5(5), 458-464.

Tam, P. T., & Van Thuy, M. B. (2017). THE INDUSTRY 4.0 FACTOR AFFECTING THE SERVICE QUALITY OF COMMERCIAL BANKS IN DONG NAI PROVINCE. European Journal of Accounting Auditing and Finance Research, 5(9), 81-91.

STATISTICA Data Miner, (2019). A Trademark for industries, URL: www.statsoft.com, USA.

Tsiptsis, K. K., & Chorianopoulos, A. (2011). Data mining techniques in CRM: inside customer segmentation. John Wiley & Sons.

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2018). Big Data-enabled customer relationship management: A holistic approach. Information Processing & Management, 54(5), 818-846.