

## RFM Metodu ile Elektrik Piyasası İçin Müşteri Skorlama

**Abstract:** Bu çalışma da Takasbank platformu kullanılarak günlük elektrik piyasası ile ilgili sözleşmeler yapan müşteri verileri işlenmiştir ve veri madenciliği tekniklerinden RFM modeli uygulanmıştır. Her bir müşteri için RFM (güncelik, sıklık ve parasal) puanları hesaplanılmıştır. RFM puanlamasına göre müşteri kategorize edilmesi kümeleme tekniklerinden k-means algoritması ile yapılmış ve 4 ayrı model çıktısı elde edildikten sonra elde edilen sonuçlara göre 4 model karşılaştırması yapılarak müşteri doğru sınıflandırılması değerlendirilmiştir.

**Anahtar kelimeler:** Veri Madenciliği, Elektrik, RFM , K-means

### 1. Giriş

RFM, Güncelik, Frekans (Sıklık) ve Parasal değer anlamına gelir. RFM modeli, müşterilerin sadakatini ve katkısını ölçmek için tüketim aralığı (günceliği), sıklık ve para miktarı gibi üç değişkenin kullanıldığı, müşterilerin tüketim davranış analizi için etkili bir yöntemdir. RFM değerine göre, müşteriler farklı gruplara ayrılabilir ve her bir grup bilgisi müşteri ile ilgili alınacak karar sürecinde çok faydalıdır. RFM analizi, müşteri davranışını analiz etmek için kullanılan iyi bir pazarlama tekniğidir. RFM parametrelerinin anlamları şu şekildedir;

**Yenilik ya da güncellik:** Bir müşteri en son ne zaman işlem yaptı?

**Frekans (Sıklık):** Müşteri ne sıklıkla işlem yapıyor?

**Para:** Müşteri ne kadar harcıyor?

Sorular bağlamında müşteri segmentasyonu belirlenir ve puanlama sistemini kullanarak her bir müşteri için bir puanlama değeri üretilir. Müşteri için üretilen puanlama sistemine göre ilgili kampanya ve stratejiler belirlenir. (Elmas, 2018)

RFM modeli ile ilgili bir çok çalışma yapılmıştır. (bkz; (Aggelis ve Christodoulakis, 2005) & Chang ve Tsai, 2011) & (Birant, 2011) & (Naik ve diğerleri, 2013) & (Dursun, 2015)) RFM'nin davranış temelli bir model olması, kullanımının kolay ve hızlı bir şekilde uygulanabilmesi birçok işletme için müşterilerini sınıflandırıp, ilgili aksiyon alınmasını kolaylaştırmıştır. RFM model kapsamında yapılan çalışmalara kısaca bakarsak bu konuda RFM modeline dair yapılan ilk çalışmalardan olan (Aggelis ve Christodoulakis, 2005) çalışmasında, aktif internet bankacılığı kullancılarını gelirlerine göre sıralamak için RFM puanlama modeli kullanmış ve müşteri gruplamak için ise k-means metodu kullanarak piramit modeline göre müşteri gruplandırması yapmıştır.

RFM modeline ek olarak yeni bir model önerisinde bulunan (Chang ve Tsai, 2011) çalışmalarında, yüksek sadakat ve katkı sağlayan müşterileri belirlemek için yeni bir RFM model olan Grup RFM (kısaca GRFM) önerdiklerini ve ürünlerin tanıtımından hareketle potansiyel müşterileri keşfettiklerini belirtmektedirler. GRFM, müşterilerin satın aldıkları tüm ürünlerdeki RFM değerlerini hesaplamak yerine, müşterilerin satın aldıkları ürünleri müşterilerin analizindeki özellikleri gibi GRFM değerini de hesaplar. Özel olarak, GRFM modeli her bir işlem veri setinde sıkça satın alınan, bir ürün seti sunan yaygın modelleri keşfeder. Daha sonra, keşfedilen sık rastlanan kalıplara dayanarak müşteriler gruplara ayrılır. Her sık kullanılan model için, müşterileri sürekli model satın almışlarsa, bir grup olarak kabul edilirler.

Sınıflandırma algoritmalarını RFM model ile çalıştıran (Birant, 2011) ise, Türkiye'de bir spor mağazası tarafından toplanılan veri kümelerini kullanarak, RFM değerleri ve gelecekteki potansiyel müşteri davranışlarını, sıklıkla birlikte satın alınan ürünleri, kümeleme metodu ile gruplandırmıştır. Ayrıca, müşteri demografik verilerini, müşteri işlem verilerini ve ürün özelliklerini kullanarak sınıflandırma çalışmalarından farklı olarak, kümeleme algoritması tarafından keşfedilen müşteri segmentlerini kullanarak da bir sınıflandırma algoritması uygulanmaya çalışılmıştır.

Daha önceki çalışmaların bağlamında bir başka çalışma olan (Naik ve diğerleri, 2013), RFM parametreleri olan güncellik, sıklık ve miktar kavramlarını dikkate alarak RFM değerlerini hesaplamışlardır. Daha sonra RFM değerine dayanarak, kümeleme, sınıflandırma ve birliktelik kuralları ile müşteri davranışlarının tahmininde bulunulmuş ve ürün önerisi yapılmıştır. Bir başka çalışmada (Nimbalkar ve Shah, 2013), RFM analizini kullanarak müşteri segmentasyonu yapılmıştır. Ardından bu analizden elde edilen değerler ile kümelemeyi gerçekleştirerek ve kümelemeden sonra elde edilen verilerle her müşteri segmenti için geçerli kuralları elde etmek adına sınıflandırma yapılarak müşteri ilişkisi daha hassas olacak şekilde yönetilmeye çalışılmıştır.

RFM model ile kümeleme algoritmalarını kullanan (Dursun, 2015), Otel işletmelerinde otel müşterilerinin profilini oluşturmada veri madenciliği tekniklerinden RFM (güncelik(R), sıklık(F) ve miktar/parasal(m)) RFM modeli uygulamasıyla ilgili çalışmalar yapmıştır. Veri setine dair ise Antalya'da faaliyet göstermekte olan bir beş yıldızlı zincir otel işletmesinin mevcut müşteri verileri kullanılmış ve çalışmada küme sayılarını belirlemek için Kendini Düzenleyen Haritalar'dan (self-organizing maps) faydalanılmış ve kümeleme analizi için K-means metodu kullanılmıştır.

Müşteri ilişkileri uzunluğu hakkında önerileri ile (Daoud ve diğerleri, 2015) farklı bir model önerisinde bulunulmuş ve yaptıkları çalışmada, E-ticarette Müşteri Sınıflandırma Modeli Kullanımını Kümeleme Teknikleri ve LRFM Modeli ile, Fas'ta aktif olarak hizmet veren bir online e-ticaret sitesi, günlük müşteri davranışlarını tahmin etmek için kullanılmıştır. RFM modele ek olarak değişken müşteri ilişkileri uzunluğu (L), RFM modelini benimseyerek çalışmalar yapılmıştır. Çünkü LRFM modeli ile şirketlerin kısa ömürlü ve uzun ömürlü müşterileri etkin olarak ayırt edebileceklerini belirtilmişlerdir. Metodolojik olarak veri madenciliği, Kendi Kendini Düzenleyen Haritalar ve K-means kümeleme gibi teknikler ile tüm müşterileri gruplandırmak için kullanılmıştır.

Bir başka çalışma olan, (Baydarman, 2017) ise yaptığı çalışmada, kritik finansal sistemlerde iş süreçleri yönetimine entegre, iş önceliklendirme üzerine kurgulanan bir otomatik iş sahiplendirme modeli geliştirmiştir. Belirlenen bir skorum modeli ile işlere göre belli skorlar üretilmekte ve bu skora göre iş ataması gerçekleştirilmektedir. Skorum modeli için incelenen çalışmalara benzer şekilde RFM analizi ve AHP tekniği kullanılmıştır. RFM analiz yardımı ile müşterilerin yaptıkları işlemler güncellik, frekans ve tutar bakımından değerlendirilerek skor üretilmektedir. Bu iki yöntem ile en önemli işlem ve müşteri tespit edilmekte ve önceliklendirilmektedir.

Yapıla gelen çalışmalardan farklı bir bakışla (Huang ve diğerleri, 2017) çalışmalarında Standart Sapma değerini de RFM model analizine ekleyerek yapmışlardır. Çin'de hızla büyümekte olan Mikrokredi endüstrisine de RFM model uyarlanması yapılmıştır. Mikrokredi endüstrisinde geleneksel kredilerin aksine, tipik olarak teminat eksikliği önem kazanmakta ve bu da kredi puanlamasını çok önemli kılmaktadır. Bu bağlamdan hareketle RFMS yöntemi uygulanmıştır. Burada “R” günceliği, “F” frekansı, “M” parasal değeri belirtmektedir. Bununla birlikte, “S” (Standart Sapma) dikkate alınarak süreci daha da uzatma ve hassasiyet ölçümü yapılabilmektedir. Veri seti olarak Çin mikro kredi şirketinden alınan gerçek bir veri örneği ile çalışma yapılmıştır.

Son olarak daha güncel bir çalışma olan Müşteri Analitiği ve Öneri Sistemleri Uygulaması (Elmas, 2018) çalışmasında klasik RFM model analizi ile ilgili bir araştırma yapılmıştır. Elmas, yaptığı çalışmada, müşteri değer yönetimi bağlamında mevcut müşterileri değerlerine göre kategorize edilip, bu veriler doğrultusunda benzer müşteriler bularak ilgili müşteriye pazarlama kampanyaları oluşturmak amacıyla öneri sistemi çalışması yapmıştır. Araştırmada UCI Machine Learning Repository'den alınan, İngiltere merkezli bir e-ticaret şirketinin, 01/12/2010 ve 09/12/2011 tarihleri arasında gerçekleşen 541.910 adet işlemi içeren veri seti kullanılmıştır. Bununla gelecek müşteri

davranışlarını daha iyi tahmin etmek ve hedef müşterileri daha iyi belirlemek için RFM analizi yapılmıştır.

Bu çalışmada ise, açık kaynak olan R ile elektrik piyasası sözleşme verileri kullanılarak her müşteri için RFM puanlamaları hesaplanmıştır. RFM ile hesaplanan müşteri puanlanması veri madenciliği tekniklerinden k-means algoritması kullanılarak birbirine yakın ödeme alışkanlığı gösteren müşterileri aynı grup içerisinde olacak şekilde hesaplamalar yapılmıştır. Bu metodoloji ile veri setinden 4 ayrı model çıkarımı yapıp, modellerin RFM puanlaması yapılmıştır. Daha sonra 4 model kendi içerisinde karşılaştırılarak müşteri gruplandırılmasının hassasiyeti hakkında değerlendirmelerde bulunulmuştur. Makalenin izlenci hakkında ise ikinci bölümde, RFM analizinin temellerini tanıtmakta ve güncellik (recency), sıklık (frequency) ve parasal (monetary) kavramlarını detaylı olarak açıklamak ile birlikte RFM modelinin ne olduğu ve hangi problem kategorisinde uygulandığı ile ilgili bilgi verilmiştir. Üçüncü bölümde, RFM model sonuçları K-means kümeleme algoritması kullanılarak benzer müşteri grupları belirlenmeye çalışılmaktadır. Dördüncü bölümde, problem tanımı, veri seti, ve ön işlem adımlarından bahsedilmiştir. Beşinci bölümde ise önerilen modelin veri ön işleme, RFM analizi, müşteri segmentasyonu, müşteri davranışı tahmini ve ürün önerisini içeren bir vaka çalışması olarak gerçek dünyadaki verileri analiz etmek için nasıl kullanılabileceği gösterilmiştir. Son bölümde ise sonuçların değerlendirmesini ve gelecekteki çalışmalar belirtilmiştir.

## **2. RFM Modeli Nedir?**

Müşteri segmentasyonu kavramı, 1950' lerin ortalarında bir Amerikan pazarlama uzmanı Wendell R. Smith tarafından geliştirilmiştir. (Cheng ve Chen, 2009) Müşterileri benzer özellikleri paylaşan ve benzer modelleri sergileyen gruplara ayırabilmek adına kullanılan bir modeldir. RFM modeli ilk olarak Hughes (1994) tarafından önerilmektedir ve önemli müşterileri büyük işlem verilerinden ayıran bir modeldir. RFM yöntemi, müşteri bölümlendirmesi için çok etkili niteliklerdir (Newell, 1997). RFM'nin açılımı; Yenilik, Frekans ve Parasal değer anlamına gelmektedir. RFM analizinin, müşterilerin satın alma davranışlarını modellemek ve bağlılıklarını, katkılarını ve satın alma potansiyellerini ölçmek için tüketim eğilimi (R), sıklık (F) ve parasal (M) olmak üzere üç önemli özelliği bulunmaktadır. RFM modelinde, tüketim eğilimi (R) genel olarak, en son tüketimin mevcut olduğu zamandan günümüze kadar olan aralık olarak tanımlanır, sıklık (F) belirli bir süre içindeki tüketim sayısıdır ve parasal (M) ise belli bir süre içinde harcanan para miktarıdır.(Baydarman,2017) (Wu ve Lin, 2005) tarafından yapılan bir çalışma, daha büyük R, F ve M değerlerine sahip

müşterilerin yeni bir işlem yapma ihtimalinin daha yüksek olduğunu göstermiştir. (Cheng ve Chen, 2009) 'de belirtildiği gibi, R, F ve M değerlerinin önemi hakkında iki görüş vardır. Üç parametre Miglautsch'ta (2000) eşit derecede önemli olarak kabul edilirken, Tsai ve Chiu'daki (2004) endüstrinin özellikleri nedeniyle eşit derecede ağırlıklandırılmıştır. Miglautsch'ta (2000) R, F, M boyutlarının her biri beş eşit parçaya bölünerek, müşteriler R, F, M değerlerine göre 125 gruba ayrılmaktadır.

Müşteri davranışını analiz etmede RFM modelini uygulamanın yaygın bir yolu, RFM'nin her bir boyutuyla müşteri verisini analiz etmek ve bu verileri beş eşit bölüme ayırarak kodlamaktır. Kodlama yönteminde ilgili parametreler bazında tüm müşterilerle her bir müşteri kıyaslanır ve her müşteri için RFM skoru oluşturularak, güncellik için müşteri veri tabanı satın alma tarihleri azalan şekilde sıralanmaktadır. Firmadan en güncel olarak alışveriş yapan müşterilerin %20'si 5, diğer %20'si 4 olacak şekilde numaralandırılarak devam edilmektedir. Sıklık ve tutar için müşterinin ziyaret sıklığı verisi ve satın alma verisi de yine azalan şekilde sıralanmaktadır. Tüm müşteriler 555,554,553,...111 şeklinde 125 mümkün şekilde kodlanmakta (5 x 5 x 5) böylece veri tabanı 125 eşit kümeye bölünmektedir. Müşteriler içerisinde en yüksek RFM skoruna sahip olanlar genellikle en kârlı olanlardır (Hosseini, Maleki, ve Gholamian, 2010, s. 5260; Wei, Lee, Chen, ve Wu, 2013, s. 7515).

RFM analizi ve veri madenciliği tekniklerinin entegrasyonu, mevcut ve yeni müşteriler için faydalı bilgiler sağlamaktadır. Elektrik piyasası müşterilerini skorlama çalışmamızda ise RFM özelliklerine dayalı kümeleme metodu ile benzer davranışları olan müşterileri aynı grup içinde sınıflandırmaya çalışılmaktadır. RFM skor belirleme değeri için yaygın olarak müşteri boyutlarını 5 eşit bölüme ayırarak RFM değeri 555....111 arasında skorlanacaktır. Bu bölümde kısaca RFM model ve analizine dair kavramsal bir çerçeve sunulmakta ve RFM model parametrelerinin hesaplamasına dair bilgi verilmektedir.

### **3. K-means Algoritması**

## 4. Problem Tanımı, Veri Seti ve Ön İşlem Adımı

### 4.1 Problem:

Yapılan günlük avans dosyası takas işlemleri ve fatura dosyası takas işlemlerinden müşterilerin ödeme alışkanlıklarına göre risk grubu belirlenip EPIAŞ (Enerji Piyasaları İşletme A.Ş.)'a raporlanacaktır. EPIAŞ ise bu rapora göre müşterileri elektrik piyasasına dahil etmeden önce her bir müşterinin yapacağı sözleşmeler için teminat yatırma miktarını belirleyecektir. Bu puanlama sisteminden hareketle EPIAŞ her gün müşterilerin elektrik piyasasına girmeden önce yatıracağı teminat miktarına göre, yapacağı sözleşmelerin takasında karşılıklı mağduriyet riskini azaltmış olacaktır.

### 4.2 Veri Seti

Günlük veya aylık olacak şekilde iki ayrı sözleşme takasına göre oluşan Avans ve Fatura dosyalarından oluşmaktadır.

**Avans Dosyası:** Günlük yapılan takas işlemleridir. Her müşterinin günlük yaptığı sözleşmelerin bilgisini içerir.

**Fatura Dosyası:** Aylık takas işlemlerinin EPIAŞ tarafından verilen dosya bilgileridir. Verilen bilgiler, EPIAŞ'ın bir önceki ayını referans alarak her müşteri bazında borç ve alacak bilgilerine dayanarak hesapladığı işlemleri içerir. Örneğin: Her ayın 15 inde fatura bilgilerini verir, 5 günlük bir çağrı yapılır. 5 gün içinde borçların kapatılması istenir ve 5 günden sonra ilgili müşteri temerrüte (ceza) girer.

### 4.3 Veri ön işlem adımı

Dört model üzerinde çalışma yapabilmek için ilgili modellerde veri temizleme, veri düzeltme işlemleri yapılmıştır. Her bir dört model için veri ön işlem süreci izlenmiştir.

**Model 1:** Avans dosyası temerrütsüz ödenen tutar bilgisi.

Veri Seti Kolonları	Açıklamaları
<b>Musterino</b>	Takas İstanbul tarafında müşterilere verilen benzersiz numaradır.
<b>Tarih</b>	Her gün sonunda, o güne ait işlemlerin gerçekleştiği tarih bilgisidir.
<b>Avans Temerrütsüz Ödenen Tutar</b>	Takas tarihindeki avans net borç tutarının saat 15:00 'a kadar ödenen kısmıdır. Aynı gün Takas tarihli borçların 15:00'a kadar ödeme süresi vardır.

**Model 2:** Avans dosyası ceza tutarına göre hesaplanmıştır. Ceza bilgisi avans gecikmeli ödeme ve avans temerrütlü ödeme tutarların toplamıdır.

Veri Seti Kolonları	Açıklamaları
<b>Musterino</b>	Takas İstanbul tarafında müşterilere verilen benzersiz numaradır.
<b>Tarih</b>	Her gün sonunda, o güne ait işlemlerin gerçekleştiği tarih bilgisidir.
<b>Avans Gecikmeli Ödenen Tutar</b>	Takas tarihindeki avans net borç tutarının saat 15:00-17:00 arası ödenen tutarıdır. Aynı gün Takas tarihli borçların 15:00 - 17:00 arası ödemesi halinde temerrüt uygulanır. Fakat yarım ceza (0.7) hesaplanır
<b>Avans Temerrütlü Ödenen Tutar</b>	Takas tarihi geçmiş olan avans net borç tutarının ödenen tutarıdır. İlgili tarihte ödenen temerrüt tutarlarının takas tarihi, en az 1 gün öncesine aittir. Takas tarihini geçirmiş tüm borçlara tam ceza (1.4) hesaplanır.

**Model 3:** Fatura dosyası temerrütsüz ödenen tutar bilgisi.

Veri Seti Kolonları	Açıklamaları
<b>Musterino</b>	Takas İstanbul tarafında müşterilere verilen benzersiz numaradır.
<b>Tarih</b>	Her gün sonunda, o güne ait işlemlerin gerçekleştiği tarih bilgisidir.
<b>Fatura Temerrütsüz Ödenen Tutar</b>	Aylık takas işlemlerinin EPIAŞ tarafından verilen süre içerisinde ödenen kısımdır.

**Model 4:** Fatura dosyası ceza tutar bilgisi.

Veri Seti Kolonları	Açıklamaları
Musterino	Takas İstanbul tarafında müşterilere verilen benzersiz numaradır.
Tarih	Her gün sonunda, o güne ait işlemlerin gerçekleştiği tarih bilgisidir.
Fatura Ceza Tutarı	Aylık takas işlemlerinin EPIAŞ tarafından verilen süre içerisinde ödenmeyen işlemlerin ceza uygulanmış miktarıdır.

şeklinde dört ayrı model oluşturulmuştur. Ve her bir model için aynı veri ön işlem adımı uygulanmıştır.

**Veri ön işlem:** Veri setinde 'null' olan kayıtlar veritabanı tarafında hatalı kayıt statüsü 'H' olarak kaydedildiğinden, ilgili kayıtları veri setinden silinmiştir. Daha sonra R studio ile kayıtların ortalaması alınmış ve gereksiz olan kayıtlar ile sıfırlı kayıtlar çıkarılmıştır. İlgili uzman görüşüne danışarak veri temizleme yapılmıştır.

## 5. Uygulama / Model Sonuçları



## 6. Değerlendirme ve Sonuç

### Kaynakça

Aghdaie H. M. & Tafreshi F. P., *A New Perspective on RFM Analysis*. Handbook of Research on Intelligent Techniques and Modeling Applications in Marketing Analytics Chapter: 1, 2016.

Aggelis V. & Christodoulakis N D., *Customer clustering using RFM analysis*. Computer Engineering and Informatics Department University of Patras, Patras, GREECE. ISBN: 960-8457-29-7 (2005).

Akarsu E., *Customer Retention Via Hybrid Modeling For Banking Industry*. Yayınlanmış Yüksek Lisans Tezi, Bahçeşehir Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul. (2010).

Baydarman V. M., *Kritik Finansal Sistemlerde İş Süreçleri Yönetimindeki İşlerin Otomatik Sahiplendirilmesi*. Yayınlanmış Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya. (2017).

Birant D., *Data Mining Using RFM Analysis*. DOI: 10.5772/13683, Dokuz Eylul University, Izmir, Turkey, 2011.

Chang H. & Tsai H., *Group RFM analysis as a novel framework to discover better customer consumption behavior*. Expert Systems with Applications 38 (2011) 14499–14513.

Chen D., Sain L. S. & Guo K., *Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining*. Journal of Database Marketing & Customer Strategy Management. Volume 19, Issue 3, pp 197–208, September 2012.

Cheng C. & Chen Y., *Classifying the segmentation of customer value via RFM model and RS theory*. Expert Systems with Applications 36 (2009) 4176–4184.

Daoud A. R., Amine A., Bouikhalene B. & Lbibb R., *Customer Segmentation Model in E-commerce Using Clustering Techniques and LRFM Model: The Case of Online Stores in Morocco* . World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering Vol:9, No:8, 2015.

Dursun A., *Otel İşletmelerinde Müşteri Profili Oluşturmada Veri Madenciliği Tekniklerinin Kullanılması: RFM Modeli Örneği*. Yayınlanmış Yüksek Lisans Tezi, Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Antalya, (2015).

Elmas Ö., *Müşteri Analitiği ve Öneri Sistemleri Uygulaması*. Yayınlanmış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi , Fen Bilimleri Enstitüsü, İstanbul. (2018)

Naik C., Prof. Kharwar A. & Desai N., *A Review: RFM Approach on Different Data Mining Techniques*. International Journal of Emerging Technology and Advanced Engineering. Journal, Volume 3, Issue 10, October 2013.

Newell F., *The new rules of marketing: how to use one-to-one relationship marketing to be the leader in your industry*. New York: McGraw-Hills Companies Inc. (1997).

Roshan H. & Afsharinezhad M., *The New Approach in Market Segmentation by Using RFM Model*. Journal of Applied Research on Industrial Engineering. Ayandegan Institute of Higher Education, Tonekabon, Iran. Vol. 4, No. 4 (2017) 259–267.

Wu J. & Lin Z., *Research on customer segmentation model by clustering*. ACM International Conference Proceeding Series, p. 113, (2005).

Zhou J., Huang D. & Wang H., *RFMS Method for Credit Scoring Based on Bank Card Transaction Data*. DOI: 10.5705/ss.202017.0043 , SSRN Electronic Journal, January 2018.