T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
SAYISAL YÖNTEMLER BİLİM DALI

K-MEANS ALGORİTMASI İLE MÜŞTERİ PROFİLİNİN BELİRLENMESİ

Yüksek Lisans Tezi

Merve GÜNAYDIN

İstanbul, 2021

T.C.
MARMARA ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
SAYISAL YÖNTEMLER BİLİM DALI

K-MEANS ALGORİTMASI İLE MÜŞTERİ PROFİLİNİN BELİRLENMESİ

Yüksek Lisans Tezi

Merve GÜNAYDIN

Danışman: Doç. Dr. Özgür ÇAKIR

İstanbul, 2021

GENEL BİLGİLER

İsim ve Soyadı	: Merve GÜNAYDIN
Anabilim Dalı	: İşletme
Programı	: Sayısal Yöntemler
Tez Danışmanı	: Doç. Dr. Özgür ÇAKIR
Tez Türü ve Tarihi	: Yüksek Lisans - Ocak 2021
Anahtar Kelimeler	: Veri Madenciliği, Kümeleme Analizi, K- Ortalama ve Ward Algoritması

ÖZET

K-MEANS ALGORİTMASI İLE MÜŞTERİ PROFİLİNİN BELİRLENMESİ

Bilgi teknolojisinde ve veri boyutunda meydana gelen hızlı gelişim ve değişim ile büyük veri kümelerinden değerli bilgileri çıkarmak için güçlü ve çok yönlü araçlara ihtiyaç duyulmaktadır. Veri Madenciliği, büyük ve karmaşık veri yığınları içerisindeki faydalı bilgileri keşfetme sürecidir.

Bu tez çalışmasında veri madenciliğinin bankacılık ve pazarlama alanındaki fonksiyonu incelenmiştir. Temel amaç, heterojen veri kümesi içerisinden homojen müşteri gruplarının ve oluşan müşteri kümelerine özel pazarlama stratejisinin belirlenmesidir. Çalışmada, 18 davranış değişkeni ve 8950 adet müşterinin kredi kartı kullanım davranışını içeren bir banka veri kümesi üzerinde veri madenciliği yöntemlerinden K-Ortalama algoritması kullanılarak kümeleme çalışması yapılmıştır. Analiz sonucunda sekiz müşteri kümesi oluşmuştur. Oluşan sekiz kümenin grup içinde ve gruplar arasındaki anlamlı farklılığın tespiti için SPSS programı kullanılarak Kruskal-Wallis sınaması yapılmıştır. Kümeler arasında bulunan anlamlı farklılıkların hangi küme gruplarından kaynaklandığını bulmak için Tamhane's T^2 sınaması uygulanmıştır ve oluşan sekiz kümenin karakteristik özellikleri belirlenmiştir.

Veri hazırlama aşamasında MS Excel, veri ön işleme ve modelleme aşamasında Python ve sonuçların analizi aşamasında IBM SPSS yazılımları kullanılmıştır.

GENERAL KNOWLEDGE

Name and Surname	: Merve GÜNAYDIN
Field	: Business Administration
Programme	: Quantitative Methods
Supervisor	: Doç. Dr. Özgür ÇAKIR
Degree Awarded and Date	: Master – January 2021
Keywords	: Data Mining, Clustering Analysis, K- Means and Ward Algorithms

ABSTRACT

CUSTOMER PROFILE EXTRACTION WITH K-MEANS ALGORİTHM

Powerful and versatile tools are needed to extract valuable information from large data sets. With the rapid development and change in information technology and data size. Data Mining is the process of discovering useful information within large and complex data heaps.

In this thesis, the function of data mining in banking and marketing has been examined. The main purpose is to determine the homogeneous customer groups from the heterogeneous data set and the marketing strategy specific to the formed customer clusters. In the study, a clustering study was conducted using the K-Average algorithm, one of the data mining methods, on a bank data set containing 18 behavioral variables and the credit card usage behaviors of 8950 customers. As a result of the analysis, eight customer clusters were formed. Kruskal-Wallis test was performed using the SPSS program to determine the significant difference between and within the group of the eight clusters formed. Tamhane's T^2 test was applied to find out which cluster groups caused the significant differences between clusters and the characteristics of the eight clusters formed were determined.

MS Excel in the data preparation phase, Python in the data pre-processing and modeling phase, and IBM SPSS software in the analysis of the results.

ÖNSÖZ

Beni yetiştirip bugünlere getiren, hayatımın her döneminde benden desteklerini esirgemeyen sevgili anne ve babama ve her zaman yanımda olan kardeşlerime,

Tez çalışmam süresince bilgi ve tecrübelerini paylaşan, desteğini eksik etmeyen değerli danışman hocam Doç. Dr. Özgür Çakır'a,

Tezimin düzenlenme aşamasında yardımlarını ve desteklerini esirgemeyen arkadaşlarım Hicran Tiryakioğlu ve Neslihan Ekici'ye sonsuz teşekkürlerimi sunarım.

İstanbul, 2021 Merve GÜNAYDIN

İÇİNDEKİLER

Sayfa No.

ÖZET ..... **i**
ABSTRACT ..... **ii**
ÖNSÖZ ..... **iii**
İÇİNDEKİLER ..... **iv**
TABLO LİSTESİ ..... **vi**
ŞEKİL LİSTESİ ..... **vii**
KISALTMALAR ..... **viii**
GİRİŞ **1**
1. VERİ MADENCİLİĞİ VE KÜMELEME **3**
1.1. VERİ MADENCİLİĞİ 3
1.1.1. Veri Madenciliği Süreci ve Veri Tabanlarında Bilgi Keşfi 4
1.1.2. Veri Madenciliği Uygulama Alanları 8
1.1.3. Veri Madenciliği Fonksiyonları 9
1.2. KÜMELEME ANALİZİ 11
1.2.1. Kümeleme Analizine Genel Bakış 11
1.2.2. Kümeleme Analizinde Benzerlik ve Uzaklık Ölçüleri 15
1.2.2.1. Minkowski Ölçüsü 15
1.2.2.2. Karesel Öklid (Squared Euclidean) Ölçüsü 18
1.2.2.3. Mahalanobis Ölçüsü 18
1.2.2.4. Chebyshev (Chebychev) Ölçüsü 18
1.2.2.5. Hotelling T^2 Ölçüsü 19
1.2.2.6. Canberra Ölçüsü 19
1.2.2.7. Korelasyon Ölçüsü 20
1.2.2.8. Karl Pearson Ölçüsü 20
1.2.2.9. Karesel Pearson Ölçüsü 20
1.2.3. Küme Sayısının Belirlenmesi 21
1.2.3.1. Elbow (Dirsek) Metodu 22
1.2.3.2. Krzanowski and Lai (KL) Endeksi 23
1.2.3.3. Calinski and Harabatsz Endeksi 23
1.2.3.4. Gap Endeksi 24
1.2.3.5. Silhouette Endeksi 25
1.2.4. Veri Madenciliğinde Kümeleme Analizi Yöntemleri 25
1.2.4.1. Hiyerarşik Yöntemler (Hierarchical Methods) 26
1.2.4.1.1. Tek Bağlantı (Single-Linkage) 27

1.2.4.1.2. Tam Bağlantı (Complete- Linkage) 27
1.2.4.1.3. Ortalama Bağlantı (Average- Linkage) 28
1.2.4.1.4. Merkezi Bağlantı (Centroid- Linkage) 28
1.2.4.1.5. WARD Algoritması 28
1.2.4.2. Bölümleme Yöntemleri (Partitioning Methods) 31
1.2.4.2.1. K-Means (K-Ortalama) Algoritması 31
1.2.4.2.2. Diğer Bölümleme Algoritmaları 35
1.2.4.3. Yoğunluk Tabanlı Yöntemler (Density-Based Methods) 36
1.2.4.4. Izgara Tabanlı Yöntemler (Grid-Based Methods) 38
1.3. K-MEANS ALGORİTMASI İLE İLGİLİ ÇALIŞMALAR 38
2. MÜŞTERİ PROFİLİ BELİRLEME UYGULAMASI 42
2.1. AMAÇ VE KAPSAM 42
2.2. VERİ KÜMESİ 43
2.3. VERİ ÖN İŞLEME 45
2.4. DEĞİŞKENLER ARASINDAKİ İLİŞKİLERİN İNCELENMESİ 45
2.5. KÜMELEME YÖNTEMİNİN UYGULANMASI 47
2.5.1. Kümeler Arası Farklılıkların İncelenmesi 48
2.5.2. Kümelerin Tanımlanması 52
2.5.3. Kümeler Arası Farklılıkların Değerlendirilmesi 53
SONUÇ 54
KAYNAKLAR 56
EKLER 66

TABLO LİSTESİ

Sayfa No.

Tablo 2.1 : Veri Kümesi Tanımı ve Nitelik Tipleri..... 43

Tablo 2.2 : Tanımlayıcı İstatistiksel Analiz..... 44

Tablo 2.3 : Python'da Kullanılan Kütüphane ve Fonksiyonlar..... 47

Tablo 2.4: Kruskal-Wallis Sınaması Sonuçları..... 50

Tablo 2.5 : Kümeler Arası Farklılıklara İlişkin Bulgular..... 51

Tablo 2.6 : Kümeler Arası Anlamlı Farklılık Düzeyleri..... 51

ŞEKİL LİSTESİ

Sayfa No.

Şekil 1.1 : Veri Madenciliğinin Tarihsel Gelişimi…….. 4

Şekil 1.2 : Bilgi Keşif Süreci……..……..6

Şekil 1.3 : CRISP-DM Süreç Aşamaları……..7

Şekil 1.4 : Kümeleme Analizi Karar Süreci……...14

Şekil 1.5 : Farklı p Değerlerine Ait Minkowski Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği ……...16

Şekil 1.6 : Öklid Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği……....17

Şekil 1.7 : Manhattan Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği……....17

Şekil 1.8 : Küme Sayısının Seçimi…….21

Şekil 1.9 : Elbow (Dirsek) Metodu ile Küme Sayısının Belirlenmesi ……..22

Şekil 1.10 : Farklı Hiyerarşik Yöntemlerin Karşılaştırılması…..……..30

Şekil 1.11 : K-Means Kümeleme Algoritma Süreci……..33

Şekil 1.12 : K-Means Algoritması Kümeleme Örneği……...34

Şekil 1.13 : Yoğunluk Tabanlı Kümeleme ……... 37

Şekil 2.14 : Veri Kümesi İçerisindeki Değişkenlerin Korelasyon Matrisi……..... 46

Şekil 2.15 : Elbow Metodu (Dirsek Yöntemi)……...48

Şekil 2.16 : Kredi Kartı Kullanım Davranışına Göre Müşteri Segmentasyonu…….....49

KISALTMALAR

CRISP - DM	Cross Industry Standard Processing – Data Mining
SOM	Sef Organizing Maps
BIRCH	Balanced Iterative Reducing and Clustering Using Hierarchies
CURE	Clustering Using REpresentatives
ROCK	Robust Clustering Using Links
CAMELON	Hierarchical Clustering Using Dynamic Modeling
PAM	Partitioning Around Medoids
CLARA	Clustering Large Applications
CLARANS	Clustering Large Applications Based Upon Randomized Search
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
OPTICS	Ordering Points To Identify The Clustering Structure
DENCLUE	Density-Based Clustering
STING	Statistical Information Grid
CLIQUE	Clustering High-Dimensional Space

GİRİŞ

İçinde bulunduğumuz ve yaşadığımız bilgi çağında veri miktarı her gün hızla artmaktadır. Mevcut veri hacmindeki bu büyüme teknolojinin hızla gelişmesi ile birlikte bilgisayar ortamına geçme süreci, veri toplama araçlarının hızlı değişimi ve gelişiminin bir sonucu olarak karşımıza çıkmaktadır. Dünya çapında işletmeler, tıp ve sağlık endüstrisi, web arama işlemleri, sosyal medya, videolar, bloglar, dijital fotoğraflar vb. büyük miktarda veri üreten kaynaklardır. Veri boyutunda meydana gelen hızlı büyüme büyük veri kümelerinden değerli bilgileri otomatik çıkarmak ve organize bilgiye dönüştürmek için güçlü ve çok yönlü araçlara ihtiyaç duymuştur. Bu durum, veri madenciliği kavramının ortaya çıkmasına neden olmuştur.

Veri madenciliğinin disiplinler arası bir konu olarak birden farklı şekilde tanımlanması mümkündür. En genel tanımıyla veri madenciliği, büyük veri yığınları içerisindeki gizli bilgileri (örüntüleri) farklı yöntemler ile analiz ederek keşfetme süreci olarak tanımlanmaktadır. Bu süreç sonucunda ortaya çıkan büyük miktarlardaki verilerden anlamlı bilgiler üretmek için veri madenciliği alanında birçok yöntem geliştirilmiştir.

Veri madenciliği süreci; iş süreçlerini anlama, veriyi anlama, ön işleme (veriyi kullanıma hazır hale getirme), modelleme, değerlendirme ve geliştirme adımlarından oluşmaktadır.

Sınıflandırma, regresyon, kümeleme ve birliktelik kuralları veri madenciliğinin en önemli fonksiyonlarını oluşturmaktadır.

Veri madenciliği, astronomi, biyoloji, finans, bankacılık, pazarlama-satış, tıp, bilim ve mühendislik gibi pek çok alanda kullanılmaktadır. Bankacılık ve pazarlama alanında müşteri kredi risklerinin araştırılması, kredi taleplerinin değerlendirilmesi, kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi, hedeflenen pazarlama kampanyaları gibi vb. konularda veri madenciliği fonksiyonları kullanılmaktadır.

Bu tez çalışmasında veri madenciliğinin bankacılık ve pazarlama alanındaki fonksiyonu incelenmiştir. Makine öğrenimi topluluğu ve halka açık bir veri platformu

olan kaggle.com veri tabanında bulunan, kredi kartı kullanım verileri üzerinde python yazılım dili kullanılarak potansiyel müşteri kümesinin belirlenmesi ve pazarlama stratejisinin geliştirilmesi amaçlanmıştır.

Tez çalışması iki bölümden oluşmaktadır. İlk bölümde veri madenciliğinin tanımı, süreci, uygulama alanları ve fonksiyonları üzerinde durulmuştur. Tezin ana odağı olan kümeleme analizi ve teknikleri hakkında bu kısımda ayrıntılı bilgi verilmiştir. İkinci bölümde kredi kartı kullanım verileri üzerinde geleneksel ve popüler kümeleme yöntemlerinden biri olan K-Means algoritması uygulanmıştır. Uygulama sonucunda elde edilen bulgular sunulmuş, küme yapıları tanımlanmış ve sonuç kısmında değerlendirilerek yorumlanmıştır.

1. VERİ MADENCİLİĞİ VE KÜMELEME

Bu bölümde, veri madenciliği ve kümele analizi kavramları tanımlanacaktır. Kümeleme analizine genel bakış, kümeleme analizinde benzerlik ve uzaklık ölçüleri, küme sayısının belirlenmesinde kullanılan metotlar ve kümeleme analiz yöntemleri hakkında bilgi verilecektir.

1.1. VERİ MADENCİLİĞİ

Veri toplama ve depolama teknolojisindeki hızlı gelişmeler kuruluşların çok miktarda veri toplamasına imkan sağlamıştır. Ancak yararlı ve faydalı bilgilerin dağınık ve toplanan veri kümesinden çıkarılması son derece zordur. Genellikle bir veri kümesinin büyük boyutu nedeniyle geleneksel veri analiz araçları ve teknikleri kullanılamaz ve bu durum mevcuttaki soruları açıklamada yetersiz kaldığından yeni analiz tekniklerinin geliştirilmesini gerektirmektedir (Kumar, 2014, s.1).

Veri madenciliği, büyük hacimli verileri işlemek için geleneksel veri analiz yöntemlerini gelişmiş algoritmalar ile birleştiren bir teknolojidir (Kumar, 2014, s.1). Aynı zamanda veri tabanlarında bilgi keşfi olarak da bilinen veri madenciliği, büyük bilgi havuzlarını analiz etme ve kesinliği bilinmeyen potansiyel faydalı bilgiler keşfetme süreci olarak da tanımlanmaktadır (Han ve diğerleri, 2012). Veri madenciliği, büyük miktarlarda veriyi inceleyerek gizli ilişkileri, bilinmeyen kalıpları ve eğilimleri ortaya çıkarma yeteneğine sahiptir (Sumathi ve Sivanandam, 2006).

Literatürde veri madenciliği tanımları aşağıdaki gibi sıralanabilir;

Veri madenciliği, veri depolarında bulunan büyük miktarda ki verilerden, örüntü tanıma teknolojilerinin yanı sıra, istatistiksel ve matematiksel yöntemler kullanılarak, anlamlı yeni korelasyonlar, kalıplar ve eğilimler keşfetme sürecidir (Larose, 2005, s.1-26).

Veri madenciliği, veri tabanlarındaki gizli kalıp ve bilgileri bulma, bu bilgileri tahmin modelleri oluşturmak için kullanma sürecidir (Diwani ve diğerleri, 2013, s.73-77).

Veri madenciliği, mevcut verilerden gizli modeller keşfetme sürecidir (PhridviRaj ve GuruRao, 2014, s.255-263).

Veri madeciliğindeki ilk çalışmalar 1960 yılında, mantıklı veri ilişkileri bulmak için veri tarama işlemleri ile başlamıştır. 1989 yılında bilgi keşfi ve veri madenciliği tanıtılmış, makine öğreniminin gelişimi için ilerleyerek günümüze kadar gelişimini sürdürmüştür. Şekil 1.1'de veri madenciliğinin tarihsel gelişimi gösterilmiştir. (Özyazar, 2019, s.8)



Şekil 1.1: Veri Madenciliğinin Tarihsel Gelişimi

Bu bölümde, veri madenciliğine genel bakış, veri madenciliği süreci, uygulama alanları ve fonksiyonları özetlenmiştir.

1.1.1. Veri Madenciliği Süreci ve Veri Tabanlarında Bilgi Keşfi

Bilgi teknolojilerindeki gelişim çeşitli alanlarda çok sayıda veri tabanı ve büyük verinin ortaya çıkmasına neden olmuştur. Veri tabanları ve bilgi teknolojisi alanındaki

araştırmalar bu değerli verilerin daha ileri kararlar için saklanması ve manipüle edilmesi yaklaşımını doğurmuştur (Ramageri, 2010, s. 301).

Veri tabanı yönetim sistemlerinin kurulmasından sonra veri tabanı teknolojisi, web tabanlı veri tabanları, gelişmiş veri analizi için veri ambarı ve veri tabanı sistemlerinin geliştirilmesine doğru ilerlemiştir (Han ve diğerleri, 2012 s.4).

Bilgisayar teknolojisinin son otuz yılda istikrarlı ve göz kamaştırıcı ilerlemesi veri toplama ekipmanlarının ve depolama ortamlarının büyük miktarda tedarik edilmesine yol açmıştır. Bu teknoloji, işlem yönetimi, bilgi çıkarımı ve veri analizi için çok sayıda veri tabanı ve bilgi havuzunun bulunmasını sağlar. Bu durumda, veriler farklı türde veri tabanında ve bilgi deposunda saklanabilir (Han ve diğerleri, 2012 s.4).

Veri madenciliği ve veri tabanlarındaki bilgi keşifleri genellikle eş anlamlı olarak değerlendirilir ancak veri madenciliği bilgi keşif sürecinin bir parçasıdır. Şekil 1.2'de bilgi keşif sürecinde veri madenciliğinin adımları gösterilmiştir. Bu süreç, ham veri toplanması, veri ön işlemeden, yeni bilginin elde edilmesi için veri madenciliği sonuçlarının işlenmesine kadar birkaç yinelemeli adımdan oluşur. Bu adımlar aşağıdaki gibi sıralanabilir: (Zaiane, 1999, s.4).

Veri Temizleme: Gürültülü ve gereksiz verilerin veri kümesinden çıkarılması

Veri Entegrasyonu: Heterojen (farklı türde) olan çoklu veri kaynaklarının ortak bir kaynakta toplanması

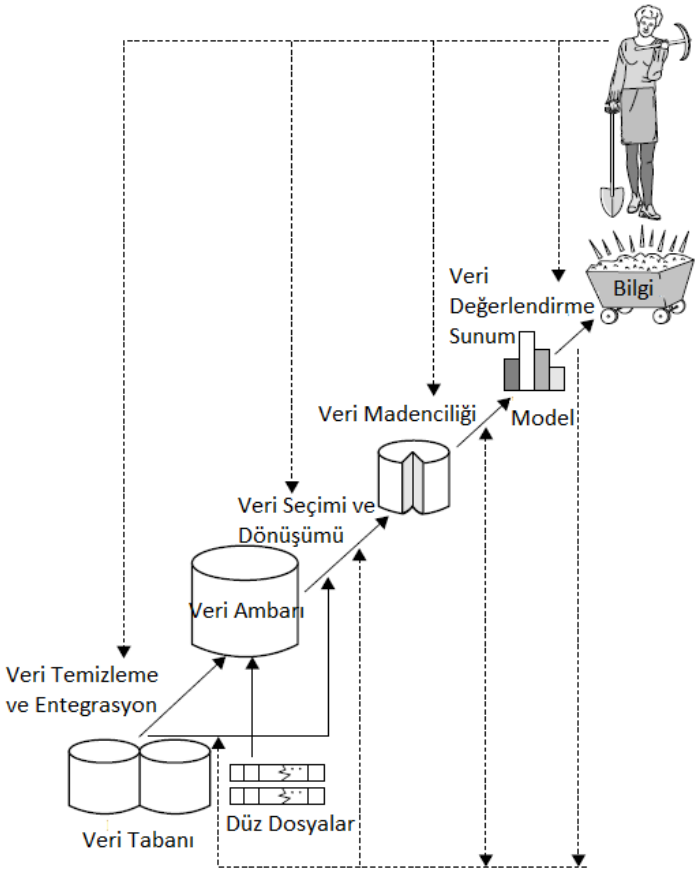
Veri Seçimi: Analizde kullanılacak verilerin seçilmesi ve toplanması

Veri Dönüşümü: Verinin konsolidasyonu olarak da adlandırılır. Seçilen verilerin veri madenciliği süreci için uygun formlara dönüştürülmesi

Veri Madenciliği: Veri kümesindeki faydalı ve gizli modelleri çıkarmak için uygun analiz yöntemlerinin kullanılması

Model Değerlendirme: Yapılan ölçümlere göre elde edilmiş bilgiyi ifade eden ilginç örüntülerin bulunması

Bilgi gösterimi: Keşfedilen bilginin görsel olarak kullanıcıya sunulması. Bu önemli son aşama, kullanıcıların veri madenciliği sonuçlarını anlamasına ve yorumlamasına yardımcı olmak için görselleştirme tekniklerini kullanır.



Şekil 1.2: Bilgi Keşif Süreci

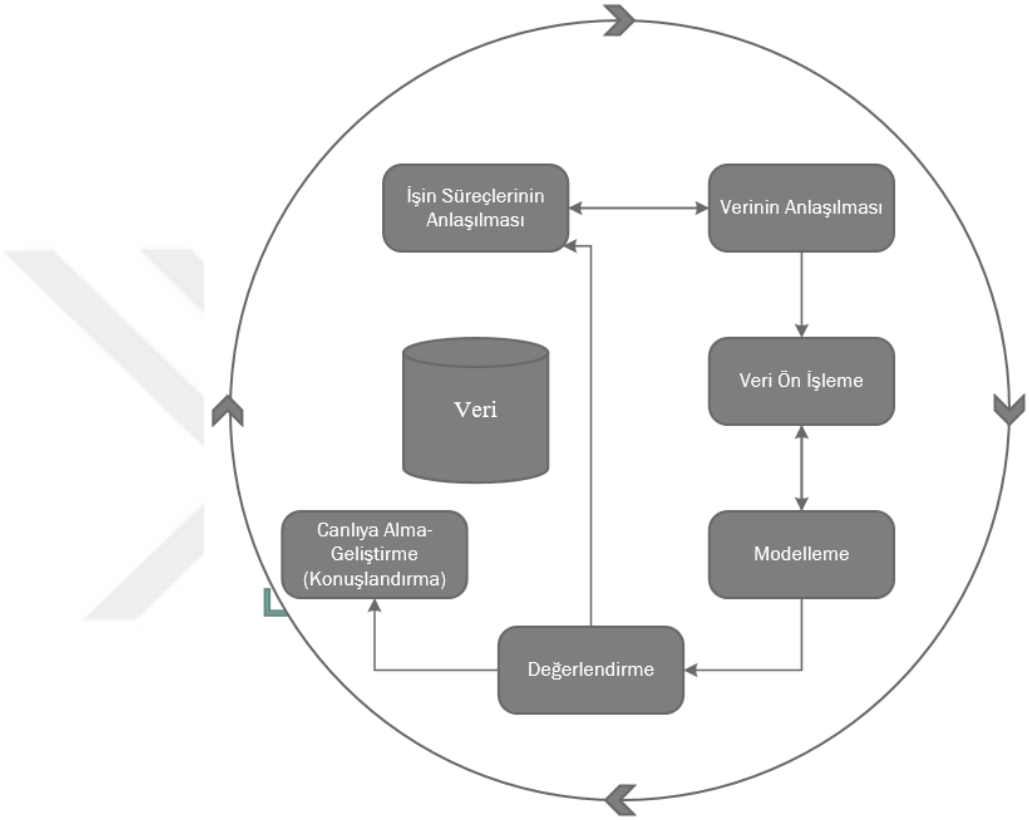
Kaynak: Han, J., M. Kamber ve J. Pei. (2012). *Data Mining Concepts and Techniques.* Waltham: Elsevier. s.7

Veri madenciliği metodolojileri ve süreç modelleri açısından 2000 yılı en önemli kilometre taşlarından birini işaret etmektedir: Endüstriler Arası Standart İşleme Süreci (CRISP-DM). CRISP-DM, veri madenciliğini geliştirmek için kullanılan en yaygın yöntemdir (Chapman ve diğerleri, 2000)

CRISP-DM süreci, 1999 yılında Daimler Chrysler (sonra Daimler-Benz), SPSS ve NCR tarafından geliştirilmiştir ve 2000 yılında CRISP–DM 1.0 dökümanı yayımlanmıştır. Veri madenciliği için tek tip bir çerçeve ve yönerge sağlar (Shearer,

2000, s.13-22). CRISP-DM süreç modeli, bir veri madenciliği projesi geliştirmek için takip edilmesi gereken adımları belirtmektedir (Chapman ve diğerleri, 2000).

CRISP-DM modeli, iyi yapılandırılmış ve tanımlanmış altı adımlı bir süreçten oluşur. Bu süreçler Şekil 1.3'te gösterilmiştir.



Şekil 1.3: CRISP-DM Süreç Aşamaları

1. **İş Süreçlerinin Anlaşılması:** Bu ilk aşama, proje hedeflerini ve gereksinimlerini bir iş perspektifinden anlamaya ve daha sonra bu bilgiyi bir veri madenciliği problemi tanımına ve hedeflere ulaşmak için tasarlanmış bir ön proje planına dönüştürmeye odaklanmaktadır (Chapman ve diğerleri, 2000).
2. **Verinin Anlaşılması:** Veri toplamaya, veri kalite sorunlarını belirlemeye ve veri kümesi içerisindeki gizli bilgileri keşfetmek için hipotezler oluşturmak üzere veri hakkında fikir edinmek amacıyla verilerin incelenmesine odaklanan CRISP-DM sürecinin ikinci aşamasıdır (Chapman ve diğerleri, 2000).

3. **Veri Ön İşleme:** Veri kümesinin seçimi ve hazırlanmasına odaklanan CRISP-DM sürecinin üçüncü aşamasıdır. Bu aşama, tablo ve öznitelik seçiminin yanı sıra modelleme araçları için verilerin temizlenmesini ve dönüştürülmesini içerir (Chapman ve diğerleri, 2000).
4. **Modelleme Aşaması**: Çeşitli modelleme tekniklerinin seçildiği ve uygulandığı dördüncü aşamadır. Farklı parametreler belirlenir ve aynı veri madenciliği problemi için farklı modeller oluşturulur (Chapman ve diğerleri, 2000).
5. **Değerlendirme Aşaması:** Elde edilen modellerin değerlendirildiği ve sonuçların nasıl kullanılması gerektiğine karar verildiği aşamadır. Modelin yorumlanması algoritmaya bağlıdır ve modeller hedeflere doğru bir şekilde ulaşılıp ulaşılmadığını gözden geçirmek için değerlendirilebilir (Chapman ve diğerleri, 2000).
6. **Konuşlandırma:** CRISP-DM sürecinin altıncı ve son aşamasıdır. Bu aşama, veri madenciliği sonuçlarının işe uygulanması için değerlendirme sonuçlarını alarak bir strateji geliştirilmesini ifade eder. Böylece, elde edilen bilgi ve sonuçların kullanımının belirlenmesine odaklanır (Chapman ve diğerleri, 2000).

1.1.2. Veri Madenciliği Uygulama Alanları

İstatistik teknikleri, büyük veri tabanlarındaki sayısal veriler ve veri ilişkileri ile ilgilenir. Örneğin, regresyon ve kümeleme analizi vb. (David J. Hand, 1998, s.112-118). Yapay zeka tekniği, istatistiksel problemlere insan gibi düşünme sürecini uygular ve genetik algoritma, bulanık mantık gibi çeşitli teknikler kullanır. (Girija ve Srivatsa, 2006, s.560-600). Makine öğrenimi ise veri analizi ve bilgi keşfi için kullanılan gelişmiş veri madenciliği tekniğidir (Kononenko ve Kukar, 2007).

Veri madenciliği, astronomi, biyoloji, finans, pazarlama-satış, tıp, fen bilimleri ve mühendislik gibi pek çok alanda kullanılmaktadır. Tıp ve mühendislik araştırmalarında ortaya çıkan yeni keşifler sonucu bilgiler hızla birikmekte ve veri madenciliği teknikleri kullanılarak analiz edilmektedir. Örneğin, moleküler biyolojideki araştırmalar genlerin yapısını ve işlevini daha iyi anlamak için şu anda toplanan büyük miktardaki veriyi kullanmaktadır. Veri madenciliği teknikleri kullanılarak yapılan araştırmalarda, her genin fonksiyonunun belirlenmesine ve belirli hastalıklardan sorumlu

genlerin izole edilmesi amaçlanmaktadır veya küresel ısınmaya neden olan kasırga ve ekosistem bozulmalarının sıklığı ve yoğunluğu gibi araştırma konularında kullanılır (Kumar, 2014, s.2). Veri madenciliği farklı alanlar arasında köprü görevi gören bir disiplindir ve uygulama alanları aşağıdaki gibi sıralanabilir (Baykal, 2006, s.97):

- Müşteri kredi risklerinin araştırılması,
- Pazar araştırması,
- Risk analizi rekabet – sahtekarlık analizi,
- Kaynak optimizasyonu,
- Müşteri profili oluşturma- müşteri ilişkileri yönetimi,
- Pazar sepet analizi,
- Hedeflenen pazarlama kampanyaları,
- İş akış yönetimi,
- Mağaza düzeni,
- Müşteri kayıp analizi,
- Müşteri satın alma davranışlarının belirlenmesi,
- Satış tahmini,
- Kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi,
- Kredi taleplerinin değerlendirilmesi,

1.1.3. Veri Madenciliği Fonksiyonları

Veri madenciliği alanında birçok yaklaşım, yöntem ve algoritma geliştirilmiştir. Veri tabanlarında bilgi keşfi için sınıflandırma, regresyon, kümeleme, birliktelik kuralları, yapay öğrenme, yapay sinir ağları vb gibi çeşitli algoritma ve yöntemler kullanılmaktadır (Colak ve diğerleri, 2015, s.242). Kullanılan bu tekniklerden bazıları aşağıda kısaca açıklanmaktadır.

- **Sınıflandırma (Classification):** Veri tabanı kayıtlarını genel olarak sınıflandırabilen bir model geliştirmek için önceden tanımlanmış bir dizi kategorik örnek kullanan en popüler veri madenciliği tekniğidir. Dolandırıcılık tespiti ve kredi risk uygulamalarının analizinde kullanılır (Ramageri, 2010, s. 302). Sınıflandırma tekniğinde yaygın olarak kullanılan algoritmaların başında,

Karar Ağaçları (Decision Tree), K-En Yakın Komşu Algoritması (K-Nearest Neighbor), Naive Bayesion Sınıflandırması, Yapay Sinir Ağları (Neural Networks), Genetik Algoritmalar (Genetic Algorithms) ve Destek Vektör Makineleri (Support Vector Machine) gelmektedir (Sahu ve diğerleri, 2011, s. 115).

- **Regresyon (Regression):** Veri madenciliği topluluğunun kullandığı en eski ve bilinen tekniktir. Verileri en az hata ile modelleyerek tahmin etmeye çalışır. Temel olarak sayısal bir veri kümesini ele alır ve matematiksel bir formül geliştirerek süreklilik gösteren değerlerin öngörülmesinde kullanılır (Sahu ve diğerleri, 2011, s. 115). En yaygın kullanılan regresyon tekniklerine, Doğrusal Regresyon (Linear Regression), Doğrusal Olmayan Regresyon (Non-Linear Regression), ve Lojistik Regresyon (Logistic Regression) örnek gösterilebilir (Guzman ve diğerleri, 2015, s. 11).
- **Kümeleme (Clustering):** Kümeleme analizi, herhangi bir veri kümesinde benzer grupları kümelemek için kullanılır (Kumar, 2014, s.1-37). Bu yaklaşım, önceden sınıflandırılmış kayıtlardan oluşan bir eğitim setinin mevcut olmadığı durumlarda kullanılır (Chen ve Liu, 2004, s.550-558).

  Popüler kümeleme teknikleri arasında K-Means Algoritması ve Beklenti Maksimizasyonu (Expectation Maximization) Kümelemesi yer almaktadır (Sahu ve diğerleri, 2011, s. 115). Literatürde birçok kümeleme algoritması mevcuttur. Bu algoritmaların seçimi ise kullanılacak olan verinin niteliğine ve problemin amacına bağlı olarak değişmektedir. Bu kısım, tezin ilerleyen bölümlerinde "Kümeleme Analizi" başlığı altında detaylıca anlatılacaktır.
- **Birliktelik Kuralları (Association Rule):** Veri tabanındaki öznitelikler ve değişkenler gibi veri öğeleri arasındaki mevcut veya potansiyel ilişkileri bulmayı amaçlamaktadır (Lunfeng, Huan ve Li, 2012, s.761-764). Birliktelik kuralları müşterilerin satın alma davranışlarının tanımlanmasında ve pazar sepet analizinde yaygın olarak kullanılmaktadır (Akpınar, 2000, s.7). Örneğin; X ürününü satın alan kişi aynı zamanda Y ürününü de alıyorsa bu durum birliktelik kuralına örnek verilebilir. Birliktelik kurallarının türetilmesi için kullanılan en yaygın tekniğe Apriori Algoritması örnek olarak verilebilir (Guzman ve diğerleri, 2015, s. 10).

1.2. KÜMELEME ANALİZİ

Bu başlık altında, kümeleme ve kümeleme analiz kavramı tanımlanacaktır. Kümeleme analizinde kullanılan benzerlik ve uzaklık ölçüleri ile küme sayısının belirlenmesinde kullanılan metotlar anlatılacaktır. Daha sonra kümeleme analizinde kullanılan yöntemlerin tanımı ve uygulanabilirliği hakkında bilgi verilecektir.

1.2.1. Kümeleme Analizine Genel Bakış

Kümeleme (clustering) analizi üzerine ilk kitap 1939'da Tryon tarafından yayımlanmıştır. Tryon'un kümeleme analizi üzerinde yaptığı incelemeler 1950 ve 1960'lı yıllardan sonra bilgisayarların üniversitelerde kullanıma sunulması ile birlikte birçok disiplinde hızla büyümüştür. 1973 yılında Peter Senath ve Robert Sokal'ın yazdığı "Sayısal Sınıflandırma" adlı kitap kümeleme analizinin gelişimine önemli bir katkı sağlamıştır (Anderberg, 1973 s.553-555).

Kümeleme (clustering), benzer grupların veya kümelerin ortaya çıkarılmasını amaçlayan bir yöntemdir. Kümelenme faaliyetlerinin varlığı yüzlerce yıl önce farklı ülkelerde ki farklı disiplinlerde görülebilir. Kümeleme analizi, pazar araştırması, iş zekası, web taraması, biyoloji, tıp, kimya, örüntü tanıma, veri analizi ve görüntü işleme vb. gibi pek çok alanda yaygın olarak kullanılmaktadır (Bansal, ve diğerleri, 2017 s.35).

Kümeleme analizi (clustering analysis) veya basitçe kümeleme (clustering), bir dizi veri nesnesini alt kümeler veya gözlemler halinde bölümleme işlemidir. Her alt küme bir kümedir. Öyle ki, kümedeki nesneler birbirine benzer ancak diğer kümelerdeki nesnelere benzemez. Bir küme analizinden elde edilen sonuç kümeleme olarak adlandırılır (Han ve diğerleri, 2012 s.444).

Kümeleme analizi çalışmalarının kullanım alanlarının ve amaçlarının farklılaşmasıyla birlikte yapılacak tanımlamalar da değişkenlik göstermektedir. Kümeleme analizine ait literatürde yapılan çalışmalar incelendiğinde farklı tanımlamalar olduğu görülmektedir. Bu tanımlamalardan bazıları aşağıdaki gibidir:

Mac Quenn (1967), nesnelere ait birden fazla boyutlu gözlemlerin sınıflandırılması olarak belirtir.

Kaufman ve Rousseeuw (1990) kümeleme analizini, gözlemlenen nesnelerin birbirine benzeyeceği şekilde gruplar oluşturmak isterken, farklı gruplardaki nesnelerin mümkün olduğunca birbirinden farklı olduğu küme grupları oluşturması olarak tanımlar.

Hair ve Black (2000) kümeleme analizini, nesneleri sahip oldukları özelliklerine göre gruplayan çok değişkenli teknikler grubu olarak tanımlamaktadır.

Webb (2002) kümeleme analizini, veri içerisindeki yapıyı keşfetmek ve doğallığını tespit etmek amacıyla bir popülasyonda ki nesnelerin belirli yöntemler aracılığıyla gruplandırılması olarak adlandırmaktadır.

Kantardzic (2003) kümeleme analizini, seçilen yöntemler ile analiz edilen verinin önceden etiketlenmemiş gruplara atanması olarak belirtmektedir.

Mirkin (2005) kümeleme analizini, veri kümesi içerisinde ki homojen varlık gruplarını bulmaya ve tanımlamaya adanmış bir yöntem olarak tanımlamaktadır.

Gan ve ark (2007) kümeleme analizini, segmentasyon analizi, taksonomi analizi veya denetimsiz sınıflandırma olarak da adlandırırlar. Ayrıca, nesnelerin küme içi homojenliğinin ve kümeler arası heterojenliğinin maksimum olacak şekilde küme grupları oluşturulması yöntemi olarak da tanımlamaktadırlar.

Yapılan bu tanımlamalar neticesinde bir istatistik yöntemi ve veri madenciliği fonksiyonu olarak kümeleme analizi, verilerin dağılımını kavramak, her bir kümenin özelliklerini gözlemlemek ve daha fazla analiz için belirli bir veri kümesine odaklanmayı amaçlayan bağımsız bir araç olarak kullanılır ve mesafe ölçüsüne dayanır (Han ve diğerleri 2012 s.445).

Kümeleme analizindeki genel amaç birbirine benzer olan nesnelerin aynı kümelerde toplanmasını sağlamaktır (Çakmak, 1999 s.187-205). Böylece her grup içerisindeki gözlemler, ilgili değişkenler veya nitelikler küme içi maksimum homojenlik ve kümeler arası maksimum heterojenlik göstermelidir (Hair ve diğerleri, 2014 s.418).

Kümeleme analizi, her bir gözlemin birkaç gruptan birine ait olduğu bilinen sınıflandırma probleminin aksine yeni bir gözlemin ait olduğu grubu tahminlemeye, küme gruplarının sayısını ve niteliğini keşfetmeye çalışır (Tryfos, 1999).

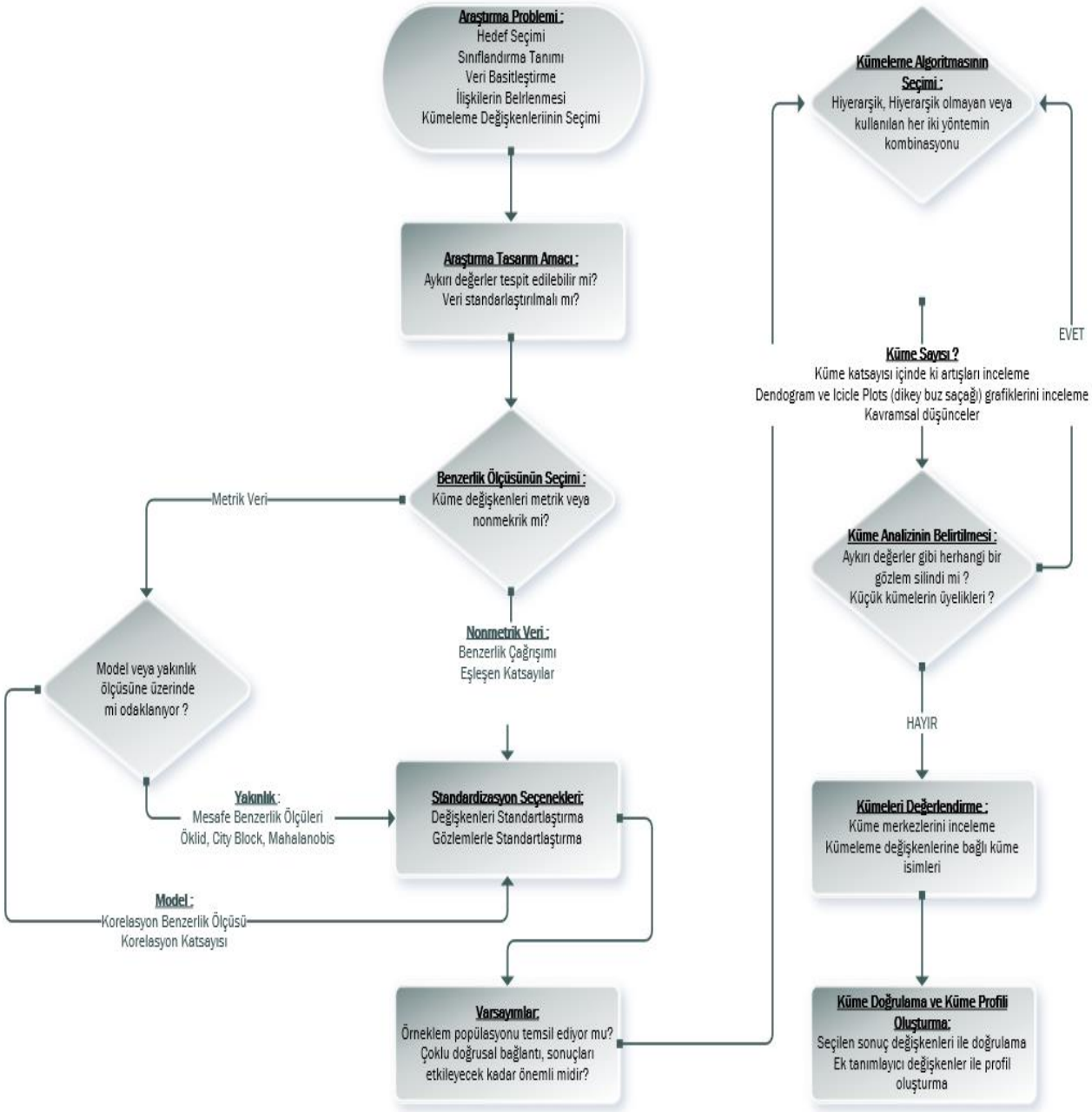
Kümeleme, denetimsiz öğrenme (unsupervised learning) yöntemidir. Denetimsiz öğrenme yönteminde ise ilgili örneklerin gözlenmesi ve gözlenen bu örneklerin özellikleri arasındaki benzerliklerden yararlanarak grupların tanımlanması hedeflenmektedir (Akpınar, 2000 s.7). Denetimsiz öğrenme yönteminde bağımsız değişkenlerden yararlanılarak model kurulur. Bu analiz yönteminde, örneklerin gruplara atanması için daha önce belirlenmiş kurallar, fonksiyonlar ve sınıflar yoktur. Yani; model bağımsız değişkenlerden yararlanılarak verilerden üretilmektedir (Koldere, 2008, s.67).

Kümeleme analizi, nesne (gözlem) yapılarının değerlendirilmesinde faktör analizi ile karşılaştırılır. Kümeleme analizinin faktör analizinden farkı değişkenleri değil nesneleri gruplandırmasıdır. Aynı şekilde, faktör analizi gruplandırma işlemini veri içerisindeki varyasyon (korelasyon) modellerine göre yaparken, kümeleme analizi gruplandırma işlemini veri içerisindeki nesnelerin birbirlerine olan mesafe (yakınlık) ölçüsüne göre yapar.

Bir kümeleme analizinin tasarımı aşağıdaki gibi sıralanır (Rendon ve diğerleri, 2011 s.27-34):

- Bir küme çözümü seçmek ve kümeler oluşturmak için veri kümesini bölümlemek,
- Oluşan her kümenin karakteristik özelliklerini anlamak için kümeleri değerlendirmek ve doğal yapılarına uygun bir etiket veya isim geliştirmek,
- Nihai küme çözüm sonuçlarının doğrulanması ve demografik veriler gibi ilgili boyutlarda nasıl farklılıklar oluşacağının gösterilebilmesi için her bir kümenin özelliklerinin tanımlanması,

Şekil 1.4'te kümeleme analizinin, karar sürecine ilişkin bir akış diyagramı gösterilmiştir.



Şekil 1.4: Kümeleme Analizi Karar Süreci

Kaynak: Hair, J. F., W. C. Black, B. J. Babin ve R. E. Anderson. (2014). *Multivare Data Analysis*. Edinburg: Pearson.

1.2.2. Kümeleme Analizinde Benzerlik ve Uzaklık Ölçüleri

Kümeleme analizinde benzer nesnelerin/özelliklerin gruplandırılması söz konusu olduğundan iki nesnenin benzer ya da farklı olup olmadığını belirlemek için kullanılan iki ölçü türü vardır. Bunlardan biri uzaklık ölçüleri diğeri ise benzerlik ölçüleridir (Rokach ve Maimon, 2005 s.322).

Nicel veri özellikleri arasındaki ilişkiyi tanımlamak için uzaklık ölçüleri, niteliksel veri özellikleri arasındaki ilişkiyi tanımlamak için ise benzerlik ölçüleri kullanılır (Xu ve Tian, 2015, s.165-193).

Uzaklık ölçüleri, küme değişkenlerindeki gözlemlerin birbirine yakınlığı olarak benzerliği temsil etmektedir (Hair ve diğerleri, 2014 s.418.).

Her kümeleme algoritması, veri noktaları arasındaki benzerlik ve farklılık endeksine dayanır. (Jain ve Dubes, 1988) Veri nokta çiftleri arasında benzerlik ve farklılık ölçüsü yoksa anlamlı bir kümeleme analizi mümkün değildir. Sokal and Sneath (1973), Legendre and Legendre (1983), Anderberg (1973), Gordon (1999) and Everitt (2001) tarafından çeşitlik, benzerlik ve farklılıklar tartışılmıştır. (Gan ve diğerleri, 2007, s.466).

Tez çalışmasının bu bölümünde, kümeleme analizinde yaygın olarak kullanılan uzaklık ve benzerlik ölçüleri anlatılacaktır.

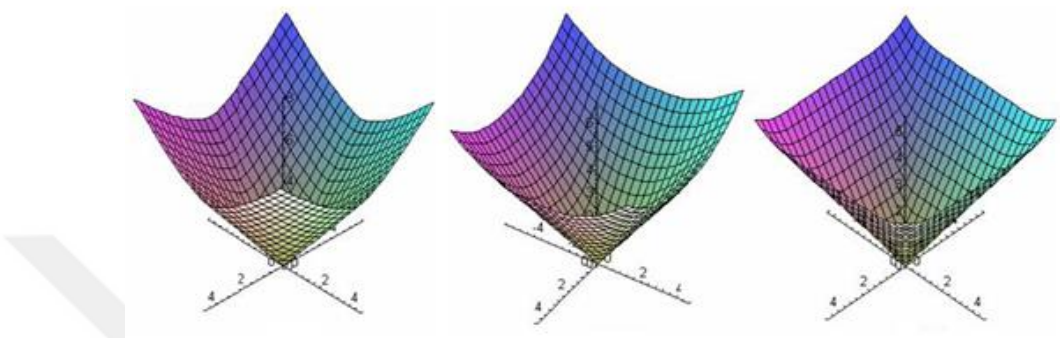
1.2.2.1. Minkowski Ölçüsü

Minkowski uzaklık öcüsü, öklid ve manhattan uzaklık ölçülerinin genelleştirilmiş hali olarak tanımlanır (Han ve diğerleri, 2012 s.72).Veri kümesinin, sıkıştırılmış veya ayrılmış küme gruplarından oluştuğu zamanlarda Minkowski uzaklık ölçüsü kullanılır. Aksi durumda, büyük özeliklere sahip nesneler diğer nesneleri/değişkenleri etkileyeceğinden doğru sonuçlar elde edilmez (Tekin, 2015, s. 396).

Minkowski uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = \left(\sum_{k=1}^{d} \left|X_{ik} - X_{jk}\right|^{1/p}\right)^{p} \quad (1)$$

Yukarıdaki formülde belirtilen p değişkeni p = 1 olduğu durumda uzaklık ölçüsü öklid, p = 2 olduğu durumda ise uzaklık ölçüsü Manhattan (city – block) olmaktadır. chebyshev uzaklık ölçüsü, $p \rightarrow \infty$ olduğu durumda Minkowski uzaklık ölçüsünün bir çeşididir. Sıralı ve nicel değişkenlerin her ikisi için de bu uzaklık ölçüsü kullanılır (Singh ve diğerleri, 2013, s.14).



p = 1,5 p = 2,5 p = 7,0

Şekil 1.5: Farklı p Değerlerine Ait Minkowski Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği

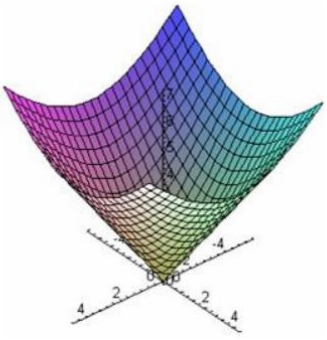
Kaynak: Gündüz, S., (2011). Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar. *Yayınlanmış Yüksek Lisans Tezi.* Adana: Çukurova Üniversitesi. FBE.

Öklid (Euclidean) uzaklık ölçüsü, iki nokta arasında ki doğrusal mesafenin hesaplanmasında en yaygın kullanılan uzaklık ölçüsüdür. Öklid uzaklık ölçüsünde işlenmemiş sayısal veriler kullanılır. Öklid uzaklık ölçüleri, kümeleme analizine eklenecek sıra dışı nesnelerden etkilenmezler. Fakat boyutlar arasında olan ölçek farklılıkları öklid uzaklık ölçüsünü önemli bir derecede etkiler (Demiralay ve Çamurcu, 2005, s.1-18).

m boyutlu bir uzayda iki nokta (X_i, X_j) arasındaki öklid uzaklığı aşağıdaki gibi formüle edilir (Singh ve diğerleri, 2013, s.13):

$$Dij = \sqrt{\sum_{k=1}^{m} \left(X_{ik} - X_{jk}\right)^2} \qquad (2)$$

$X_i = X_1, X_2, \ldots\ldots X_i, X_j = X_1, X_2, \ldots\ldots X_j, k = 1,2,3,\ldots\ldots\ldots m$



Şekil 1.6: Öklid Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği

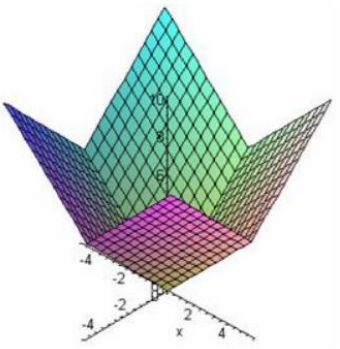
Kaynak: Gündüz, S., (2011). Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar. *Yayınlanmış Yüksek Lisans Tezi.* Adana: Çukurova Üniversitesi. FBE.

Manhattan uzaklık ölçüsü "City – Block" olarak da adlandırılır. Manhattan uzaklık ölçüsü, boyutlar arasındaki nesnelerin ortalama mutlak farkına eşittir (Demiralay ve Çamurcu, 2005, s.1-18).

Bu uzaklık ölçütü, hesaplama için kullanılan en basit yöntemlerden biridir. Ancak kümeleme analizinde yüksek korelasyona sahip nesneler var ise geçersiz kümeler oluşabilir (Hair ve diğerleri, 2014 s.431).

Manhattan uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$D_{ij} = |Xik - Xjk| \qquad (3)$$



Şekil 1.7: Manhattan Uzaklık Fonksiyonunun Üç Boyutlu Uzayda Grafiği

Kaynak: Gündüz, S., (2011). Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar. *Yayınlanmış Yüksek Lisans Tezi.* Adana: Çukurova Üniversitesi. FBE.

1.2.2.2. Karesel Öklid (Squared Euclidean) Ölçüsü

Karesel öklid uzaklık ölçüsü, daha önce anlatılan öklid uzaklık ölçüsü ile aynı şekilde formüle edilir. Tek fark iki nesne arasındaki mesafenin hesaplanmasında bu ölçü kullanılırken karekök alınmaz ve nesnelerin farklarının karelerinin toplamını ifade eder. Bu durum yapılan hesaplamaları hızlandırması açısından bir avantaj olarak görülür. Bu ölçü, kümeleme algoritmalarından Ward ve Centroid metotları için önerilir (Hair ve diğerleri, 2014 s.431).

Karesel öklid uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = X \sum_{k=1}^{m} (X_{ik} - X_{jk})^2 \quad (4)$$

1.2.2.3. Mahalanobis Ölçüsü

Mahalanobis uzaklık ölçüsü, nesneler arasında her nesnenin niteliğini varyanslarına ve ikili korelasyonlarına göre eşit olarak ağırlıklandıran genelleştirilmiş bir mesafe ölçüsüdür. Bu ölçüt, küme içinde bulunan üyelerin çok değişkenli Gauss dağılımından geldiğini varsaymaktadır (Gündüz, 2011, s.33).

Mahalanobis uzaklık ölçüsü aşağıdaki gibi formüle edilir (Uslu, 2019, s.5):

$$Dij = (X_i - X_j)^T . S^{-1} . (X_i - X_j) \quad (5)$$

Bu denklemde, Xi, i. kümenin ortalama vektörünü, *S*, pxp kovaryans matrisini, *T*, matrisin transpozunu, S^{-1}, kovaryans matrisinin tersini göstermektedir.

1.2.2.4. Chebyshev (Chebychev) Ölçüsü

Chebyshev uzaklık ölçüsü, maksimum metrik olarak da adlandırılmaktadır. Bir nesne çiftinin koordinatları arasındaki farkların mutlak büyüklüğü alınarak bu ölçüt hesaplanır (Singh ve diğerleri, 2013, s.13).

Veri kümesinde bulunan nesnelerin ölçü birimleri farklı ise bu ölçüt de terim büyüklüğü yüksek olan diğer nesnelere baskın gelecektir. Bu durumda benzerlik

hesaplanırken terim büyüklüğünün yüksek olduğu boyutun dışındaki nesneler dikkate alınmaz ve bu veri kümeleri için Chebyshev uzaklık ölçütünün kullanılması önerilmez (Gündüz, 2011, s.28).

Chebyshev uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = \max \left| X_{ik} - X_{jk} \right| \quad (6)$$

1.2.2.5. Hotelling T^2 Ölçüsü

Hotelling T^2 uzaklık ölçüsü, iki nesne kümesinin ortalama vektörlerinin birbirlerinden uzaklıklarının karşılaştırılmasında kullanılır (Dinler, 2014, s.16).

Hotelling T^2 uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$T^2 = (n_1 n_2 / n).\ (\overline{X}_\iota - \overline{X}_J).S^{-1}.(\overline{X}_\iota - \overline{X}_J) \quad (7)$$

n_1 = 1. Nesne kümesinin gözlem sayısı

n_2 = 2. Nesne kümesinin gözlem sayısı

S^{-1} = pxp kovaryans matrisinin tersi

$\overline{X\iota}$ = i. gözlemin ortalama vektör değeri

$\overline{X_J}$ = j. Gözlemin ortalama vektör değeri

1.2.2.6. Canberra Ölçüsü

Canberra uzaklık ölçüsü, veri kümesinde bulunan değişkenlerin tamamının pozitif olduğu durumlarda kullanılır (Johnson ve Wichern, 1988, s.727-799).

Canberra uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = \sum_{i=1}^{n} \frac{|Xi - Yi|}{Xi + Yi} \quad (8)$$

1.2.2.7. Korelasyon Ölçüsü

Nesneler arasındaki uzaklık korelasyon katsayısı ile de hesaplanabilir. Korelasyon katsayısı [1,-1] aralığında değer alır. Korelasyon katsayısı +1 değerine yakın ise nesneler (değişkenler) arasında pozitif yönlü bir ilişki, -1 değerine yakın ise değişkenler arasında negatif yönlü bir ilişki, 0 ise değişkenler arasında bir ilişki olmadığını gösterir (Servi, 2009, S.37-38).

Korelasyon katsayısı uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = (1 - r_{ij}) / 2 \quad (9)$$

1.2.2.8. Karl Pearson Ölçüsü

Karl Pearson uzaklık ölçüsü, öklid uzaklık ölçüsünün standartlaştırılması ile elde edilen bir mesafe ölçüsüdür. (Alpar, 2011, s.167).

Karl Pearson uzaklık ölçüsü aşağıdaki gibi formüle edilir:

$$Dij = \sqrt{\sum_{k=1}^{n} . W_k{}^2 \left(X_{ik} - X_{jk}\right)^2} \quad (10)$$

W_k = *k. Değişkenin standart sapma değeri*

1.2.2.9. Karesel Pearson Ölçüsü

Karesel Pearson uzaklık ölçüsü, Karl Pearson uzaklık ölçüsü formülünün karesinin alınması ile elde edilir.

Karesel Pearson uzaklık ölçüsü aşağıdaki gibi formüle edilir:

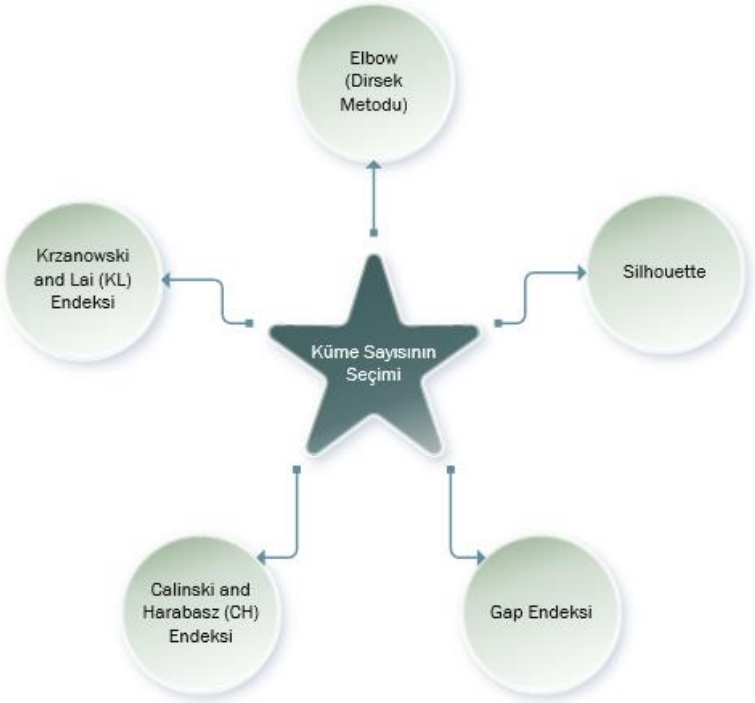
$$Dij = \sum_{k=1}^{n} . W_k{}^2 \left(X_{ik} - X_{jk}\right)^2 \quad (11)$$

1.2.3. Küme Sayısının Belirlenmesi

Kümeleme analizindeki en önemli problemlerden biri küme sayısının belirlenmesi ve doğru değerin seçilmesidir. Küme sayısının belirlenmesi, sadece K-Means gibi bazı kümeleme algoritmalarında bir parametre gerektiği için değil aynı zamanda küme sayısının uygunluğu kümeleme analizinin ayrıntı düzeneğini de kontrol ettiği için önemlidir (Han ve diğerleri, 2012 s.486).

Doğru küme sayısının belirlenmesi, genellikle veri kümesindeki dağılımın şekline, ölçeğine ve kullanıcının gerektirdiği küme çözünürlülüğüne bağlıdır (Han ve diğerleri, 2012 s.486).

Literatürde küme sayısının belirlenmesinde kullanılan birçok farklı yöntem mevcuttur. Bu bölümde, Şekil 1.8'de belirtilen kümeleme yöntemleri incelenecektir.



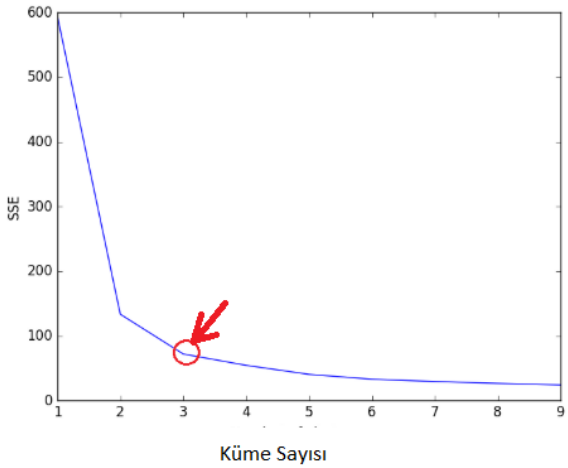
Şekil 1.8 : Küme Sayısının Seçimi

1.2.3.1. Elbow (Dirsek) Metodu

Elbow (Dirsek) Metodu, küme sayısının arttırılmasının, her kümenin küme içi varyans toplamını azaltmaya yardımcı olabileceği gözlemine dayanmaktadır. Bunun nedeni daha fazla kümeye sahip olmanın birbirine daha benzeyen veri nesnesi gruplarının belirlenmesine yardımcı olmasıdır. Çok fazla küme oluşması durumunda ise küme içi varyans toplamını azaltmanın marjinal etkisi düşer. Sonuç olarak, doğru küme sayısının bir sezgisel tarama ile seçilmesi küme sayısına göre küme içi varyans toplamını gösteren eğrinin dönüm noktasını kullanmaktır (Han ve diğerleri, 2012 s.486).

Teknik olarak, verilen sayı k>0 olduğunda K-Means gibi bir kümeleme algoritması kullanılarak söz konusu veri kümesi içerisinde k kümeleri oluşur ve küme içi varyans toplamı hesaplanır. k'ya göre varyans eğrisi çizilir. Eğrinin ilk dönüm noktası doğru küme sayısını verir (Han ve diğerleri, 2012 s.486).

Şekil 1.9'da küme sayısı ve küme içi varyans toplamı arasındaki ilişkiyi gösteren bir grafik verilmiştir. Küme içi varyans toplamı genel bir azalış gösterirken k=3 noktasında bir kırılma mevcuttur. Bu noktadan sonra eğri yatay olarak devam etmektedir. Bu durumda, k= 3 değeri optimum küme sayısı göstermektedir.



Şekil 1.9 : Elbow (Dirsek) Metodu ile Küme Sayısının Belirlenmesi

1.2.3.2. Krzanowski and Lai (KL) Endeksi

Krzanowski ve Lai kümelemenin kalitesini değerlendirmek için 1985 yılında aşağıda tanımlanan formülü önermiştir. (Üstünel, 2018, s. 29)

Her küme sayısı k ≥ 2 için, Krzanowski and Lai endeksinin tanımı:

$$DIFF\ (k) = (k-1)^{\frac{2}{p}}.WSS(k-1) - k^{\frac{2}{p}}\,\mathrm{WSS}(k)$$

$$KL\ (k) = \left|\frac{DIFF\ (k)}{DIFF\ (k)+1}\right| \qquad (12)$$

Krzanowski and Lai (KL) endeksine göre KL(k) değerini maksimum yapan k değeri küme sayısı olarak seçilir. Ancak k=1 için KL(k) değeri tanımlanmamıştır (Xu ve diğerleri, 2016, s.1495). Küme içi kareler toplamı (WSS) değerinin azalışı tabanındaki KL endeksine uygun küme sayısının belirlenmesindeki ana düşünce, WSS(k) değerinin uygun küme sayısına kadar hızlı bir şekilde azaldığı daha sonra yavaş azaldığı şeklindedir.

1.2.3.3. Calinski and Harabatsz Endeksi

Milligan ve Cooper, yaptıkları 30 farklı yaklaşımın deneysel karşılaştırılmasından sonra aşağıdaki endeksi önermişlerdir (Patil ve Baidari, 2019, s.133).

Calinski ve Harabatsz (CH) önerilen endeks aşağıdaki gibi tanımlanır:

$$CH\ (k) = \frac{(B_k)/(k-1)}{(W_k)/(n-k)} \qquad (13)$$

B_k ve W_k sırasıyla küme içi ve kümeler arası kareler toplamıdır. Karesel uzaklık ölçüleri kullanılarak B_k ve W_k değerleri hesaplanmaktadır. Calinski and Harabatsz endeksine göre CH(k) değerini maksimum yapan k değeri küme sayısı olarak seçilir (Patil ve Baidari, 2019, s.133-134).

1.2.3.4. Gap Endeksi

Gap istatistiği, veri kümesi içerisindeki optimum küme sayısını belirlemek için kullanılan en popüler yöntemlerden biridir. Gap İstatistiğindeki temel mantık, küme içi dağılımı, verinin yapısına uygun bir referans dağılım altındaki beklenti ile kıyaslamaktır. (Tibshirani ve diğerleri, 2001, s.11-423).

Gap istatistiğinin küme sayısını fazla tahmin etme eğilimi olduğu Dudoit and Fridlyand tarafından belirtilmiştir (Dudoit ve Fridlyand, 2002). Ayrıca verilerin üstel dağılımlardan elde edildiği durumlarda Gap istatistiği doğru sonuçlar vermez Sugar ve James, 2003, s.750-763). Yin ve diğ. (2008) göre bir veri kümesi farklı yoğunluklarda kümeler içerdiği durumlarda gap istatistiği başarısız olabilir. Bunun için herhangi bir dağılım yerine normal dağılımdan örneklenmiş referans veri kümelerinin kullanılması önerilir.

Gap istatistiğinin genel formülü aşağıdaki gibi tanımlanır (Patil ve Baidari, 2019, s.134):

$$Gap(q) = \frac{1}{B} \sum_{b=1}^{B} \log W_{qb} - \log W_q \qquad (14)$$

Burada B değeri, düzgün dağılım kullanılarak türetilen referans veri kümesinin sayısıdır. W_q, ağırlık değeri, W_{qb} ise dağılım içi matrisi gösterir. Optimum küme sayısı en küçük q değerinin bulunması ile seçilir (Patil ve Baidari, 2019, s.134):

$$Gap\ (q) \geq Gap\ (q+1) - s_{q+1}, (q = 1, \dots \dots n-2)$$

$$s_d = s_{dq}\sqrt{1 + 1/B}$$

$$s_{dq} = \sqrt{\frac{1}{B}\sum_{b-1}^{B} (\log W_{qb} - \bar{I})^2}$$

$S_{dq} =$ log (W_{qb})'nin standart sapması, b = 1,……,B

$$\bar{I} = \frac{1}{B}\sum_{b-1}^{B} \log W_{qb}$$

1.2.3.5. Silhouette Endeksi

Kaufman and Rousseeuw tarafından tanıtılan silhouette endeksi, verilen kümeleme çıktısı içerisinde her bir nesnenin nasıl sınıflandırıldığını gösterir (Patil ve Baidari, 2019, s.134).

n nesneden oluşan d veri kümesi için d'nin k kümelerine bölündüğünü varsayalım. c_1,……..c_k, her nesne o € d ise, a(o); o ve o'nun ait olduğu kümedeki tüm nesneler ile arasındaki ortalama uzaklık ve b(o); o ve o'nun ait olmadığı tüm kümeleri arasındaki minimum ortalama uzaklık hesaplanarak silhoutte endeksi aşağıdaki şekilde formüle edilir (Han ve diğerleri, 2012 s.449-490):

$$s\,(o) = \frac{b(o) - a\,(o)}{mak\,\{a(o), b(o)\}} \qquad (15)$$

Silhouette endeksinin değeri +1 ve -1 arasındadır. a(o) değeri, ait olduğunu kümenin yoğunluğunu gösterir. Bu değer ne kadar küçük olursa küme o kadar yoğun (kompakt) demektir. b(o) değeri ise o'nun diğer kümelerden ayrılma derecesini gösterir. b(o) değeri büyüdükçe "o" diğer kümelerden daha fazla ayrışır. "o"nun silhouette endeks değeri 1'e yaklaştığında, "o" içeren küme yoğun ve diğer kümelerden ayrışmış anlamına gelmektedir. Silhouette endeksi negatif olursa, "o" başka kümedeki nesnelere, bulunduğu kümedeki nesnelerden daha yakın ve kümeye uygun olmadığı anlamına gelmektedir. (Han ve diğerleri, 2012 s.449-490).

1.2.4. Veri Madenciliğinde Kümeleme Analizi Yöntemleri

Literatürde birçok kümeleme algoritması vardır. Kümeleme yöntemleri, nesnel fonksiyon seçiminde, kullanılan mesafe matrisi ve benzerlik matrisini oluşturma yaklaşımlarında farklılık gösterir (Jatain ve diğerleri, 2013, s.171).

Farley ve Raftery (1998) kümeleme yöntemlerini genel olarak iki ana gruba ayırmayı önerir: hiyerarşik ve hiyerarşik olmayan yöntemler.

Han ve Kamber (2012) kümeleme yöntemlerini dört ana kategoriye ayırmayı önermektedir. Bunlar;

- Hiyerarşik Kümeleme Yöntemleri
- Bölümleme Kümeleme Yöntemleri
- Yoğunluk Tabanlı Kümeleme Yöntemleri
- Izgara Tabanlı Kümeleme Yöntemleri

Estivill Castro (2000) tarafından ise çeşitli kümeleme yöntemlerinin tümevarım ilkesine dayanan alternatif bir sınıflandırma önerilmiştir.

1.2.4.1. Hiyerarşik Yöntemler (Hierarchical Methods)

Hiyerarşik yöntem, verilen veri kümesi içerisindeki nesneleri hiyerarşik ayrıştırır ve veri nesnelerini bir küme hiyerarşisine veya küme ağacına gruplamaya çalışır (Han ve diğerleri, 2012 s.449).

Hiyerarşik kümele yöntemlerinin birleştirme ve bölme noktalarının seçimlerinde zorlanması dezavantajlarından biridir. Çünkü bir grup nesne birleştirildiğinde veya ayrıldığında daha önce yapılan işlemler geri alınamadığından veya kümeler arasında nesne değişimi yapılamadığından bir sonraki süreç yeni oluşan kümeler üzerinden devam edecektir (Han ve diğerleri, 2012 s.449).

Hiyerarşik yöntem, hiyerarşik ayrışmanın nasıl oluştuğuna bağlı olarak birleştirici (agglomerative), ve bölümleme (divisive) olmak üzere iki kategoride sınıflandırılır. Kullanıcı istediği sayıda kümeyi bir sonlandırma koşulu olarak belirleyebilir (Han ve diğerleri, 2012 s.449).

Bölümleme (Divise) Kümeleme Yöntemleri, yukarıdan – aşağıya yaklaşım olarak da adlandırılır. Tüm nesneler başlangıçta bir kümeyi temsil eder. Daha sonra küme, istenen küme yapısı elde edilene kadar arka arkaya bölünerek kendi alt kümelerine ayrılır (Rokach ve Maimon, 2005 s.331).

Birleştirici (Agglomerative) Kümeleme Yöntemleri, aşağıdan – yukarıya yaklaşım olarak da adlandırılır. Her nesnenin ayrı bir grup oluşturması ile başlar ve her nesne başlangıçta kendi kümesini temsil eder. Daha sonra kümeler istenen küme yapısı elde edilene kadar ard arda birleştirilir (Rokach ve Maimon, 2005 s.331).

Hiyerarşik kümeleme yöntemlerinin sonucu nesnelerin iç içe gruplanmasını ve gruplamaların değiştiği benzerlik düzeylerini temsil eden bir dendogramdır. Dendogramın istenen benzerlik düzeyinde kesilmesiyle veri nesnelerinin bir kümelenmesi elde edilir. Kümelerin birleştirilmesi veya bölünmesi, bazı ölçütleri (karelerin toplamı gibi) optimize edecek şekilde seçilen bazı benzerlik ölçütlerine göre gerçekleştirilir. Hiyerarşik kümeleme yöntemleri benzerlik ölçüsünün hesaplanma şekline göre ayrıca gruplandırılabilir (Rokach ve Maimon, 2005 s.331).

Kümeler arasındaki benzerliği ölçmek için kullanılan en popüler birleştirici (agglomerative) beş yöntem aşağıdaki gibi tanımlanmaktadır.

1.2.4.1.1. Tek Bağlantı (Single-Linkage)

En yakın komşu yöntemi olarak da bilinir. Kümeler arasındaki benzerliği, bir kümedeki herhangi bir nesne ile diğer kümede bulunan nesne arasındaki en kısa mesafe olarak tanımlar (Han ve diğerleri, 2012 s.460).

Veriler benzerliklerden oluşuyorsa bir çift küme arasındaki benzerliğin, bir kümenin herhangi bir üyesinden diğer kümenin herhangi bir üyesine olan en büyük benzerliğe eşit olduğu kabul edilir. (Sneath ve Sokal, 1973).

Tek bağlantı kümelemenin "Zincirleme Etkisi" olarak bilinen bir dezavantajı vardır. Bu etki, iki küme arasında köprü oluşturan noktaların bu iki kümeyi bir araya getirmesine engel olur (Rokach ve Maimon, 2005 s.331).

1.2.4.1.2. Tam Bağlantı (Complete- Linkage)

En uzak komşu veya çap yöntemi olarak da bilinir. Küme benzerliği, her kümedeki nesneler arasındaki maksimum mesafe olarak tanımlanır. Bir kümedeki tüm nesneler maksimum mesafeden birbirine bağlanır ve böylece küme içerisindeki benzerlik küme çapına eşitlenir. Bu yöntem, tek bağlantılı kümeleme yöntemlerinden daha kompakt kümeleme çözümlerini üretebilir (Hair ve diğerleri, 2014 s.441).

1.2.4.1.3. Ortalama Bağlantı (Average- Linkage)

Minimum varyans yöntemi olarak da bilinir. Ortalama bağlantı yönteminde, iki küme arasındaki mesafenin bir kümenin herhangi bir üyesinden diğer kümenin herhangi bir üyesine olan ortalama mesafeye eşit olduğunu tanımlanır (Rokach ve Maimon, 2005 s.331).

Ortalama bağlantı kümelemenin de tek bağlantı kümeleme gibi bir dezavantajı vardır. Uzun kümelerin bölünmesine ve oluşan komşu uzun kümelerin birleştirilmesine neden olur (Rokach ve Maimon, 2005 s.331).

Bu yöntem, tek bağlantı veya tam bağlantı gibi uç değerlere bağlı değildir. Ortalama bağlantı yöntemleri küme içi minimum varyasyona ve grup içi eşit varyansa sahip kümeler oluşturma eğilimindedir (Hair ve diğerleri, 2014 s.441).

1.2.4.1.4. Merkezi Bağlantı (Centroid- Linkage)

İki küme arasındaki benzerlik küme merkezleri arasındaki mesafe olarak tanımlanır. Küme merkezleri, küme değişkeni içinde bulunan değişkenler üzerindeki nesnelerin ortalama değerleridir. Bu yöntemde, nesneler her gruplandığından yeni bir küme merkezi oluşur (Hair ve diğerleri, 2014 s.442).

Bu yöntemin avantajı, ortalama bağlantı yöntemi gibi diğer hiyerarşik yöntemlere göre aykırı değerlerden daha az etkilenmesidir (Hair ve diğerleri, 2014 s.442).

1.2.4.1.5. WARD Algoritması

Ward'ın yönteminde, iki küme arasındaki benzerlik ölçütü tüm değişkenler üzerinde toplanmış kümeler içindeki kareler toplamıdır. Hangi iki kümenin birleştirileceğinin seçimi, hangi küme kombinasyonunun küme içi kareler toplamını minimize ettiğine bağlıdır (Hair ve diğerleri, 2014 s.442).

Küme içindeki her nesne bağımsızdır ve çoğu adım adım bir kümeye atanacaktır. Her bir analiz adımı, mevcut kümeye en yakın olan nesneleri içerir. Adım sayısı 1'den ….n'ye kadar olabilir. (analiz edilen nesnelerin sayısı) Adım sayısının bir olması

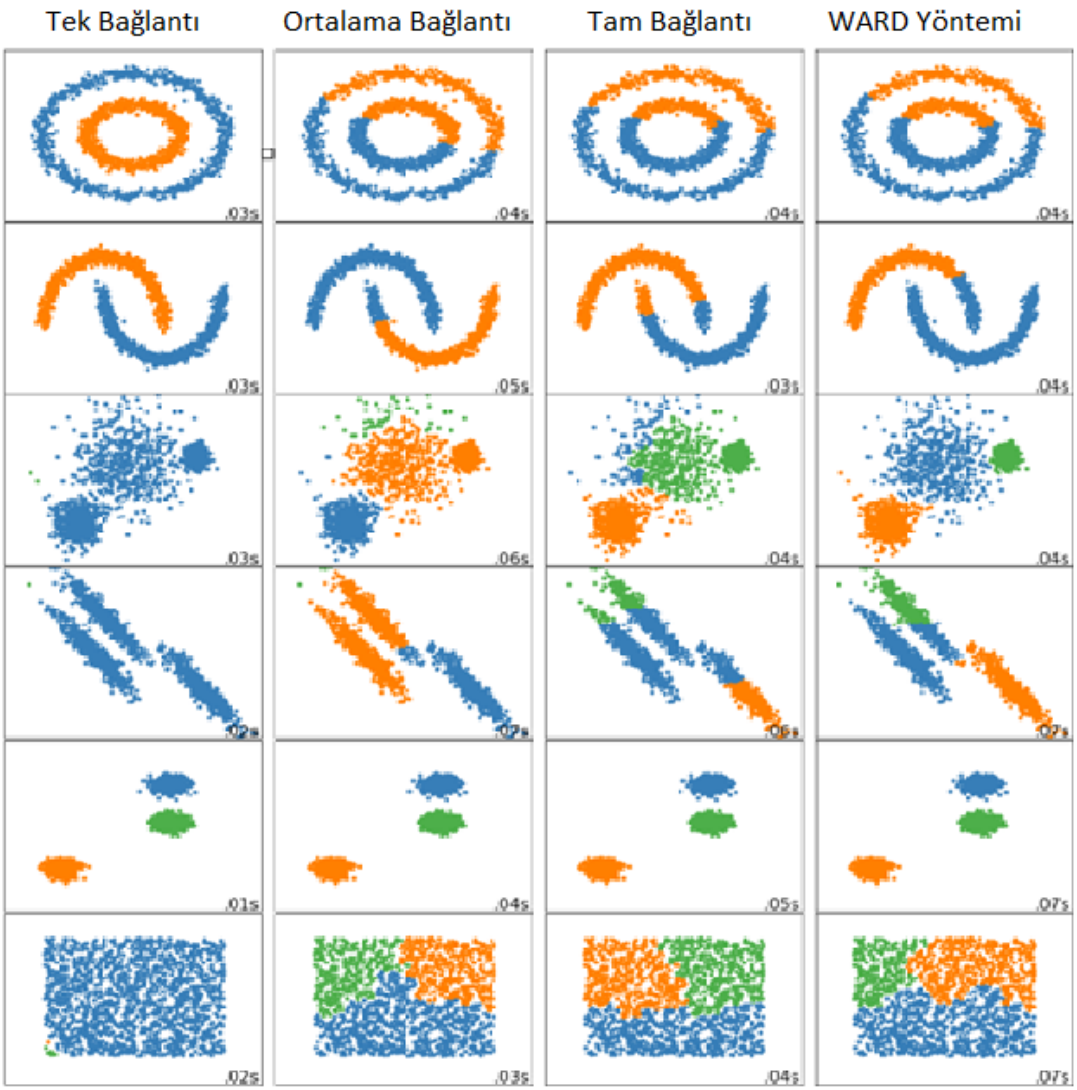
durumunda yalnızca tek bir küme tüm nesneleri içerirken n olması durumunda ise tüm öğeler kendi kümesini oluşturur. Çalıştırılan bir adım sonucunda yeni kümeler oluşturulduktan sonra oluşan yeni kümenin nesneleri ayrıştırılamaz (Eszergár-Kiss ve Caesar, 2016, s.25-34).

Ward yöntemi diğer yöntemler ile kıyaslandığında sonuçlarda yüksek doğruluk payı sağlar ve nesneler arasındaki varyansı minimum yapar. Hands ve Everitt ve ark. (1987) Ward'ın yönteminin diğer hiyerarşik kümeleme yöntemlerinden daha iyi sonuç verdiğini belirtmiştir (Eszergár-Kiss ve Caesar, 2016, s.25-34).

Ward yönteminin amacı, veri kümesinin aykırı değerler içermediği, eşit boyutta kümeler oluşturmak için en uygun kümeleme adımlarını bulmaya çalışmaktır. Monotondur ancak aykırı değerlere duyarlıdır. (Almeida ve diğerleri, 2007, s.208-217).BIRCH Algoritması, (Balanced Iterative Reducing and Clustering Using Hierarchies) yöntemi, toplayıcı hiyerarşik kümeleme yöntemlerindendir. Sayısal verileri kümelemek için kullanılır. Bu yöntemde, küme sayısı önceden bilinmelidir (Jatain, ve Diğerleri, 2013, s.172).

Şekil 1.10'da kümeler arasındaki benzerliği ölçmek için kullanılan farklı hiyerarşik bağlantı yöntemlerini karşılaştıran bir görsel verilmiştir.

BIRCH yöntemi, küme özelliği (clustering features) kavramına dayanır. Küme özeliği kavramı da üç bileşenden oluşur. Bunlar, n, LS ve SS'dir. n küme içerisindeki veri nesnelerinin sayısını, LS kümedeki nesnelerin öznitelik değerlerini, SS ise küme içindeki nesnelerin özellik değerlerinin karelerinin toplamını gösterir. Bunlar CF-Tree adı verilen bir ağaçta depolanır. Elde edilen bilgi bir ağaç formunda saklandığından veri tabanı içerisinde oluşan tüm veri kayıtları veya kümeler ana bellekte tutulmaz (Jatain ve Diğerleri, 2013, s.172).



Şekil 1.10 : Farklı Hiyerarşik Yöntemlerin Karşılaştırılması

Kaynak: Overview of Clustering Methods. (t.y). https://scikit-learn.org/stable/modules/clustering#k-means (21 June 2020)

CURE Algoritması, temsilcileri kullanarak kümeleme olarak da bilinir. Cure (Clustering Using Representatives) yöntemi, büyük veri kümelerini, aykırı değerleri, küresel olmayan şekilleri ve düzgün olmayan boyutlara sahip kümeleri işleyebilen bir yöntemdir. Bu yöntemde önce veri bölümlenir daha sonra küme oluşturmak için hiyerarşik kümeleme kullanılır. Küme sayısı ve veri özelliği önceden bilinmelidir BIRCH yöntemine göre küme kalite performansı daha iyidir (Jiman ve Khanuja, 2016 s. 672).

ROCK Algoritması, kategorik veri kümeleme tekniğidir. Küme sayısı önceden bilinmelidir. ROCK (Robust Clustering Using Links) yöntemi, herhangi bir mesafe

fonksiyonuna bağlı değildir Çok sayıda özniteliği sahip veri nesnelerinin olduğu durumda bu yöntem iyi sonuç vermez (Jiman ve Khanuja, 2016 s. 672).

CAMELON Algoritması, (Hierarchical Clustering Using Dynamic Modeling) yönteminin çıkış amacı, CURE ve ROCK yöntemlerindeki zayıf noktaları iyileştirmektir. Verileri çok sayıda, nispeten küçük alt kümelere ayırmak için bir grafik bölümleme algoritması kullanılır. İkinci aşamada ise bu alt kümeleri tekrar bir araya getirerek gerçek kümeleri bulmak ve sayısal verileri kümelemek için toplayıcı hiyerarşik kümeleme algoritması kullanılır. Küme sayısı önceden bilinmelidir (Jiman ve Khanuja, 2016 s. 672).

1.2.4.2. Bölümleme Yöntemleri (Partitioning Methods)

Bölümleme yöntemleri veri kümeleri üzerinde tek düzeyde bölümleme gerçekleştirir. Temel bölümleme yöntemleri genellikle tek bir küme oluşturmayı benimser. Bu yöntemde, her nesne bir gruba ait olmalıdır. Çoğu bölümleme yöntemi mesafeye dayalıdır. Oluşturulacak bölüm sayısı k olarak verildiğinde bölümleme yöntemi bir başlangıç kümesi oluşturur. İyi bir bölümlemenin genel kriteri aynı kümedeki nesnelerin yakın veya birbiri ile ilişkili olduğu, farklı kümelerdeki nesnelerin uzak veya birbirinden farklı olduğudur (Han ve diğerleri, 2012 s.448-449).

Bölümleme yöntemlerinde küresel optimallik elde etme hesaplama açısından zordur ve bu durum oluşan tüm olası kümelerin kapsamlı numaralandırılmasını gerektirir. Bunun yerine çoğu zaman popüler olan K-Means ve K-Medoids gibi sezgisel yöntemler benimsenir (Han ve diğerleri, 2012 s.449-449).

En popüler bölümleme kümeleme algoritmaları K-Means, K-Medoids, Pam, Clara, Clarans olarak sıralanabilir.

1.2.4.2.1. K-Means (K-Ortalama) Algoritması

K-Means algoritması en yaygın kullanılan denetimsiz öğrenme yöntemlerinden biridir (Lloyd ve diğerleri, 2005, s. 1161-1168). K-Means algoritmasının uygulamadaki kolaylığı, basitliği, verimliliği ve deneysel başarısı popülerliğinin ana nedenlerindendir (Jain, 2010, s. 653). Kümeleme algoritmalarının bir çeşidi olan bölümleme temeline dayanan bir algoritmadır ve (MacQueen, 1967, s. 281-297). 1957 yılında Cox tarafından

ortaya çıkmış (Koldere, 2008, s. 78), 1967 yılında J.B.MacQueen tarafından önerilmiş (Jain ve Dubes, 1988), K- J.A. Hargiton ve M.A Wong tarafından ise geliştirilmiştir. K-Means algoritması sayısal özeliklere sahip büyük veri kümelerini bölümlemek için kullanılan çok popüler bir teknik olmuştur (San ve diğerleri, 2004, s.241). Denetimsiz öğrenme kategorisinde olan bu algoritma, genellikle veri madenciliğinde örüntü tanıma ve veri ön işleme alanlarında kullanılır (Kazaz, 2019, s.30-31). K-Means algoritması, büyük veri kümeleri ile uğraşırken hızlı yaklaşım, ölçeklenebilirlik ve yüksek verimliliğe sahiptir. Bu algoritmanın amacı, k değişkeni tarafından temsil edilen grup sayıları ile verideki grupları bulmaktır. K-Means algoritması, hızlı ve etkili bir yöntemdir çünkü bir yinelemenin karmaşıklığı k*n*d'dir. k (küme sayısı), n (örnek sayısı), d (iki nokta arasındaki Öklid mesafesini hesaplayan zaman) (Kaur, 2014).

Aşağıda ki değer K-Means kümeleme algoritmasını temsil eder:

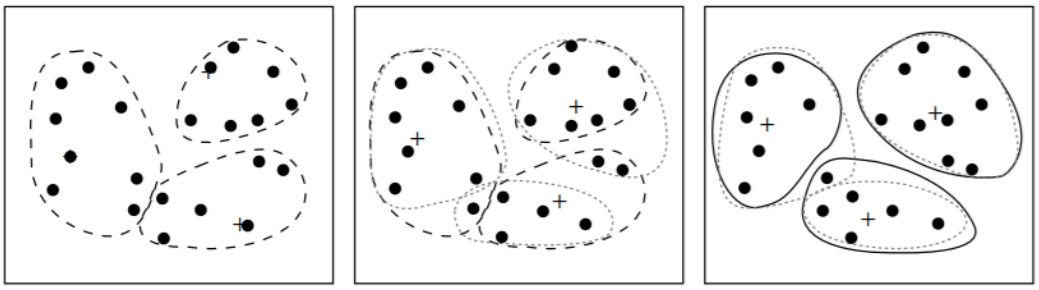
$$J(V) = \sum_{k=1}^{c} . \sum_{j=1}^{n} (IIXi - VjII)^2 \qquad (16)$$

(IIXi – VjII) Xi, Vj, bir nokta ve bir merkez arasındaki Öklid uzaklığıdır.

K-Means kümeleme algoritmasının genel çalışma mantığında parametre olarak "k" kullanılır ve amaç n adet nesneden oluşan bir veri setini "k" adet kümeye bölmektir. Böylece aynı küme içerisindeki nesnelerin birbirine benzerlikleri maksimum, farklı kümelerin birbirine (kümeler arası) benzerlikleri minimum olacaktır (https://scikit-learn.org/stable/modules/clustering#k-means, 2007). Çalıştırılan kümeleme algoritmasının sonucunda, K-Means algoritması her bir nesnenin sadece bir kümeye ait olmasını sağlamaktadır (Jain ve Dubes, 1988).

K-Means algoritması, bir kümenin merkezini küme içindeki noktaların ortalama değeri olarak tanımlar. İlk olarak, her biri başlangıçta bir küme ortalamasını veya küme merkezini temsil eden n veri kümesi içerisindeki nesnelerin k değerlerini rastgele veya belirli yöntemler kullanarak seçer. Diğer nesnelerin her biri için nesne ile küme ortalaması arasındaki öklid uzaklığına bağlı olarak kendisine en çok benzeyen kümeye bir nesne atar. Bu durumda, K-Means algoritması tekrarlanarak küme içi varyasyonu geliştirir. Her küme için bir önceki yinelemede kümeye atanan nesneler kullanılarak hesaplanan yeni bir ortalama değer ile küme merkezleri yeniden belirlenir. Atamalar sabit

olup küme merkezleri değişmeyene kadar iterasyonlar devam eder. Bir önceki iterasyonda oluşturulan kümeler ile yeni yapılacak iterasyonda oluşacak kümeler aynı olur. K-Means algoritma süreci Şekil 1.11'de özetlenmiştir (Han ve diğerleri, 2012 s.486).



(a) 1. İterasyon (b) 2. İterasyon (c) Kümeleme Sonucu

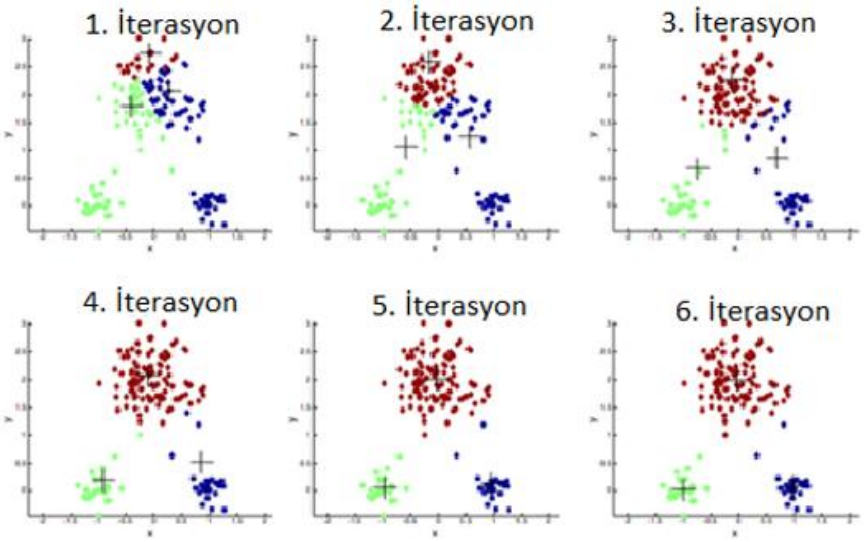
Şekil 1.11:K-Means Kümeleme Algoritma Süreci

Kaynak: Han, J., M. Kamber ve J. Pei. (2012). *Data Mining Concepts and Techniques.* Waltham: Elsevier. s.486.

K-Means algoritması genellikle lloyd algoritması olarak adlandırılır ve temel olarak üç adımı vardır (Bansal ve diğerleri, 2017 s.37).

1. Adım: Başlangıç merkezlerini seçmek. En temel yöntem n veri kümesinden k adet örneklem seçmektir. Veri kümesindeki her nesne kendisine en yakın olan kümenin ağırlık merkezine atanır.
2. Daha önce her bir ağırlık merkezine atanan tüm nesnelerin ortalama değerleri alınarak yeni bir ağırlık merkezi oluşturulur. K-Means algoritması iki adım arasında döngü oluşturmayı içerir.
3. Eski ve yeni ağırlık merkezleri arasındaki fark (uzaklık) hesaplanır ve algoritma bu değer bir eşik değerinden daha az olana kadar son iki adımı tekrarlar. Yani ağırlık merkezlerinin değerleri önemli ölçüde değişmeyene kadar bu adım tekrarlanır.

Yeterli bir zaman verildiğinde K-Means algoritması yerel bir minimum noktasına doğru yaklaşacaktır. Bu durum, ilk yapılan rastgele başlangıç merkezlerinin başlangıç seçimine bağlıdır ve ağırlık merkezlerinin farklı başlangıçları ile birkaç kez algoritma tekrarlanıp hesaplama yapılır (Bansal ve diğerleri, 2017 s.37).



Şekil 1.12: K-Means Algoritması Kümeleme Örneği

Kaynak: Tan, P. N., M. Steinbach, ve V. Kumar. (2006). *Introduction to Data Mining.* Boston MA: Pearson Addison Wesley.

K-Means Algoritması aşağıdaki 5 adımı takip eder (Ullman ve diğerleri, 2014):

1. Önceden tanımlanmış olan k küme sayısının değeri seçilir (rastgele veya farklı yöntemler ile),
2. Başlangıç merkezi olarak seçilen "k" veri noktaları küme merkezi olarak belirlenir,
3. Seçilen veri noktalarının küme merkezlerine olan öklid uzaklığı hesaplanır ve kendisine en yakın küme merkezine her bir veri noktası atanır,
4. Merkeze atanan her veri noktası için mevcut küme üyelikleri kullanılıp nesnelerin ortalama değerleri alınarak her küme için ağırlık merkezleri yeniden hesaplanır,
5. Veri noktalarının farklı kümelere yeniden atanmaması, kümelerdeki ağırlık merkezlerinin değişmemesi veya hata kareler toplamında minimum azalmanın (SSE) olması durumunda kümeleme işlemi tamamlanır. Aksi durumda 2. ve 3. adımlar tekrar edilir.

K-Means algoritmasının diğer kümeleme algoritmalarına göre sahip olduğu avantaj ve dezavantalara dair literatürde yer alan bazı tespitler şu şekilde özetlenebilir:

Avantajları (Ullman ve diğerleri, 2014):

- ✓ Basit, anlaşılması ve uygulanması kolay bir algoritma olması,

- ✓ Etkinliğinin ve zaman karmaşıklığının = 0 (tnk) minimum olması,
  n - veri noktası sayısı,
  k – küme sayısı,
  t - iterasyon sayısı,
- ✓ k ve t parametrelerinin küçük olmasından dolayı K-Means'in doğrusal bir algoritma olarak kabul edilmesi,

Dezavantajları (Han ve diğerleri, 2012 s.486):

- ✓ Küme sayısı olan k'nın önceden belirlenmesi,
- ✓ Dış bükey şekilli kümeler ve çok farklı büyüklükteki kümeler için uygun olmaması,
- ✓ Gürültülü ve aykırı değerlere duyarlı olması çünkü bu tür verilerin bir kısmı ortalama değeri önemli ölçüde etkileyebilir.

1.2.4.2.2. Diğer Bölümleme Algoritmaları

K-Medoids (PAM-Partitioning Around Medoids) Kaufman ve Housseeuw tarafından geliştirilmiştir (Nagpal ve diğerleri, 2013, s.172).

K-Medoids (PAM) algoritmasının temel stratejisi her kümenin küme merkezinin yakın lokasyonlardan biri ile temsil edilmesidir. PAM ilk K – Medoids algoritması olarak bilinir (Madhulatha, 2012, s.719-725).

K-Medoids (PAM) algoritması hata kare toplamını (SSE) minimum yapmaya çalışır. (Kaufmann and Rousseeuw, 1990). K-Means algoritmasına çok benzer (Rokach ve Maimon, 2005 s.334).

K-Medoids (PAM) algoritması, önce her küme için temsili bir nokta belirleyerek k kümelerini bulur. Bu temsili nesne noktası, küme içinde medoid adı verilen en merkezi konumda bulunur. Medoidler seçildikten sonra seçili olmayan tüm nesneler için ortalama benzerlik hesaplanır ve ait olma olasılığı en yüksek olan medoidde gruplandırılır. Algoritma her bir iterasyonda daha iyi bir medoid hesaplayarak sürekli iyileştirme sağlar (Nagpal ve diğerleri, 2013, s.172).

K-Medods (PAM) algoritması, gürültülü ve aykırı değerlerin varlığında K-Means algoritmasından daha iyi sonuç verir. Çünkü medoidler ortalama ve aykırı değerlerden daha az etkilenir (Madhulatha, 2012, s.719-725).

CLARA Algoritması, (Clustering Large Applications) algoritması, Kaufman ve Rousseeuw tarafından büyük verileri işleyecek şekilde tasarlanmıştır. CLARA algoritması, tüm veri kümeleri için temsili nesneler bulmak yerine veri kümesinin bir örneğini çizer, bu örneğe PAM uygular ve medoidleri bulur. CLARA algoritması, en iyi kümeleme sonucu bulmak için tekrar tekrar çalıştırılır. Burada bir kümenin kalitesi tüm veri kümesindeki nesnelerin ortalama farklılığı dikkate alınarak ölçülür (Kaufman ve Rousseeuw, 1990).

CLARANS Algoritması, CLARA ve PAM algoritmalarından daha etkili ve verimlidir (Kaufman ve Rousseeuw, 1990).

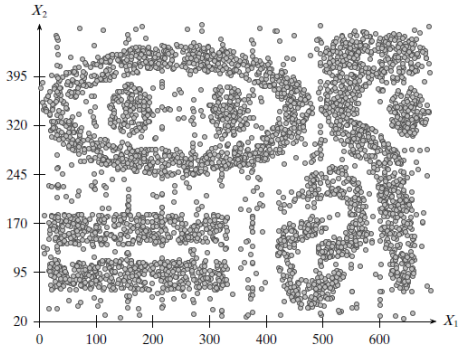
CLARAS (Clustering Large Applications Based Upon Randomized Search) algoritması binlerce nesnenin oluşturduğu veri tabanlarında verimli bir şekilde çalışır. CLARANS algoritması, CLARA ve PAM algoritmalarının çalışma mantığını birlikte uygulamaya çalışır. Bu algoritmanın dezavantajı, kümelenecek tüm nesnelerin aynı anda ana bellekte bulunabileceğini varsaymasıdır (Nagpal ve diğerleri, 2013, s.172).

1.2.4.3. Yoğunluk Tabanlı Yöntemler (Density-Based Methods)

Kümeleme yöntemlerinin çoğu nesneleri nesneler arasındaki mesafeye göre bölümler ancak yoğunluğa dayalı yöntemler her kümeye ait noktaların belirli bir olasılık dağılımından alındığını varsayar (Banfield ve Raftery, 1993). Bu yöntemin amacı, kümeleri ve dağılım parametrelerini tanımlamaktır. Yoğunluk tabanlı kümeleme yöntemleri, dışbükey olmayan, seyrek bölgelere ayrılmış rastgele şekilli yoğun kümeleri keşfetmek için tasarlanmıştır (Rokach ve Maimon, 2005 s.335).

Yoğunluk tabanlı kümeleme, kümeleri belirlemek için noktaların yerel yoğunluğunu kullanır (Zaki ve Meira JR., 2014, s.385). Yoğunluk tabanlı kümeleme yöntemleri, bir nesne kümesini birden çok özel kümeye veya küme hiyerarşisine bölebilir ve genellikle özel kümeleri dikkate alır. Bulanık kümeleri dikkate almaz (Han ve

diğerleri, 2012 s.471). Şekil 1.13'te bir veri kümesi üzerinde yoğunluk tabanlı kümeleme yöntemi gösterilmiştir. DBSCAN, OPTICS ve DENCLUE yoğunluk tabanlı kümeleme algoritmalarından bazılarıdır.



Şekil 1.13: Yoğunluk Tabanlı Kümeleme

Kaynak: Zaki M.J. ve M.JR. Meira (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. New York: Cambridge University Press.

DBSCAN Algoritması, (Density-Based Spatial Clustering of Applications with Noise) algoritması, gürültülü uygulamaların yoğunluğa dayalı uzaysal kümelenmesi olarak adlandırılır. DBSCAN algoritması, büyük uzaysal veri tabanlarında etkilidir ve gelişi güzel şekle sahip kümeleri keşfeder. Veri tabanındaki her bir nesnenin komşuluğunu arayarak kümeleri araştırır ve minimum nesne sayısından daha fazlasını içerip- içermediğini kontrol eder (Ester ve diğerleri, 1996).

OPTICS Algoritması, DBSCAN algoritmasının bir uzantısıdır. OPTICS (Ordering Points To Identify The Clustering Structure) algoritması, kümelenme yapısını tanımlamak için sipariş noktaları olarak adlandırılır. Çekirdek mesafesini ve her nesne için uygun bir erişilebilirlik mesafesini saklayan bir veri tabanı sırası ve çok çeşitli olası değerleri tanımlayan bir kümeleme yapısı oluşturur. Bu sayede verileri etkileşimli ve otomatik olarak kümelendirir (Ankerst ve Markus, 1996).

DENCLUE Algoritması, (Density-Based Clustering) algoritması, bölümleme ve hiyerarşik yöntemlerin toplamıdır (Hinneburg ve Keim, 1998). DENCLUE algoritması, bir dizi yoğunluk dağılım fonksiyonunu temel alan bir kümeleme yöntemidir (Han ve diğerleri, 2012). Aynı girdi kümesi için DECLUE algoritması diğer yaklaşımlardan daha etkilidir. Bu algoritma, rastgele gürültülü seviyelerde ve yüksek boyutlu multimedya veri

kümeleri üzerinde çalışır. DBSCAN algoritmasından daha iyi performans gösterir (Nagpal ve diğerleri, 2013, s.173).

1.2.4.4. Izgara Tabanlı Yöntemler (Grid-Based Methods)

Bu yöntemler, bir veri tabanı alanını kümeleme işlemi için tüm işlemlerin gerçekleştirildiği bir ızgara yapısı oluşturan sınırlı sayıda hücreye ayırır. Izgara tabanlı kümeleme yöntemlerinin avantajı hızlı işlem süresidir (Han ve diğerleri, 2012). STING ve CLIQUE, ızgara tabanlı kümeleme yöntemlerinden bazılarıdır.

STING Algoritması, Wang, Yang ve Murtz tarafından 1997 yılında önerilmiştir. STING algoritması, yüksek derecede ölçeklenebilir algoritmalardan biridir ve veri kümesini çeşitli ayrıntılı düzeylere böler. STING (Statistical Information Grid) algoritması, uzaysal verileri alır ve dikdörtgen hiyerarşik yapıyı kullanarak hücre sayısına ayrıştırır ve her bir hücrede ortalama, varyans, minimum ve maksimum değer hesaplar. Hiyerarşinin her düzeyinde uygun hücreleri ziyaret ederek uzaysal veriler hakkında bilgi verir (Nagpal ve diğerleri, 2013, s.173).

CLIQUE Algoritması, R. Agrawal, J. Gehrke, D. Gunopulos ve P. Raghavan tarafından önerilen CLIQUE algoritması, yoğunluk tabanlı ve ızgara tabanlı kümele tekniklerini birleştirir. (Agrawal ve diğerleri, 1998, s.94-95).

CLIQUE (Clustering High-Dimensional Space) algoritması, çakışmayan aralıklar içerisinde her boyutu bölümler ve bu sayede veri nesnelerinin içine yerleştirilecek alanı hücrelere böler. Yoğun hücreleri ve seyrek olanları tanımlamak için bir yoğunluk eşiği kullanır. Eşlenen nesnelerin sayısı yoğunluk eşiğini aştığında hücre yoğun olarak tanımlanır (Han ve diğerleri, 2012 s. 481-483).

1.3. K-MEANS ALGORİTMASI İLE İLGİLİ ÇALIŞMALAR

Literatürde bankacılık sektörüne ilişkin K-Means algoritması kullanılarak yapılan çalışmalarda genellikle müşteri segmentasyonu, müşteri satın alma davranışlarının belirlenmesi, müşteri kayıp analizi ve kredi risklerinin araştırılması vb. konular incelenmiştir.

Aşan (2007), kredi kartı kullanan müşterilerin sosyo - ekonomik özelliklerini gruplandırmayı amaçlamıştır. Öncelikle, bireysel bankacılık ve kredi kartları tanımlanmıştır; bu kavramın ülkedeki yeri ve önemi anlatılmış ve kredi kartı kullanan banka müşteri verileri üzerinde kümeleme analizi gerçekleştirilmiştir. Yapılan analiz soncunda banka müşterileri sosyo - ekonomik özelliklerine göre üç gruba ayrılmıştır. Her bir grupta yer alan müşterilerin sosyo - ekonomik göstergeye göre farklılık gösterdiği görülmektedir.

Doğan (2008), (1998–2006) döneminde Türk Bankacılık Sektöründe faaliyet gösteren ticari bankaların finansal rasyolarını ele alarak kümeleme analiz çalışması gerçekleştirmiştir. Finansal oranlara dayalı ve aktif ticari bankalara ait uygulama sonuçlarının, konu tarihi itibarıyla uyumluluğu tartışılarak analiz yapılmış ve çalışmada açıklanmıştır. Bankalar için yapılan analiz sonuçları ile uygulama sonuçlarının uyarlana bilirliği tartışılarak, kümeleme analiz tekniğinin kullanım amacı doğrultusunda bankaların finansal performansları belirlenmiş ve finansal açıdan benzer özellik gösteren bankalar tanımlanmıştır.

Keçek ve Cinser (2008), benzer performans gösteren bankaları, kümeleme ve diskriminant analizini kullanarak finansal oranlar açısından sınıflandırmıştır. Çalışmada, kullanılan finansal oranlar işletmenin içinde bulunduğu yıla ait performansı gösterdiğinden bu çalışma performans değerlendirmesi olarak da ele alınabilir.

Bilen (2009), banka çalışanlarının performanslarını değerlendirerek ve K-Means kümeleme algoritmasını kullanarak çalışanları performans başarı seviyelerine göre sınıflandırmıştır. Analiz sürecinde çalışanların demografik, eğitim ve öğretim durumu hakkındaki bilgiler dikkate alınarak WEKA programı kullanılmıştır.

Farajian ve Mohammadi (2010), K-Means ve Apriori Algoritmasını birlikte kullanarak iki aşamalı müşteri davranış analiz çalışması gerçekleştirmiştir. K-Means algoritması, müşteri gruplarını yenilik, sıklık ve parasal davranış puanlama belirleyicilerine göre tanımlamak için kullanılmıştır. Müşteriler üç ana gruba ayrılmıştır. Analiz sonucunda, yapılan çalışma banka veri tabanlarını analiz etmek için yeni bir yöntem önerisi sunmuştur.

Khajvand ve Tarokh (2011), veri madenciliği yaklaşımı olarak K-Means ve Karar Ağacı algoritmalarını kullanmıştır. K-Means algoritmasını kullanma amacı, banka müşterilerini kredi ödeme geçmişindeki benzerliklere göre gruplandırmaktır.

Tudor (2012), müşteri (borçlular ve iyi ödeme yapanlar) profilini belirlemek için borçlu grupları K-Means algoritması ile kümelere ayırarak veri madenciliği tekniklerine dayalı bankacılık alanı için yararlı bir model oluşturmuştur. Kümeleme sonucunda, veri madenciliği teknikleri kötü borçluların bulunduğu kümede uygulanmakta ve uygulama sonrasında çok yüksek bir doğruluğa ulaşmaktadır.

Bansal ve Bohra (2013), internet bankacılığı endüstrisinde müşteri veri tabanı üzerinde, K-Means kümeleme algoritmasını kullanarak müşteri özelliklerini ve davranışlarını, uygun çalışma süresini, çalışma tarihini, işlem durumunu ve işlem kriterlerini belirlemeye çalışmıştır.

Kandemir ve Arıcı (2013), mülkiyet yapılarına göre Türkiye'de faaliyet gösteren bankaları gruplara ayırmıştır. 2001-2010 yıllarına ait banka performanslarını incelemiştir. Yapılan analiz sonucunda, mevduat bankalarının 2001 krizinden sonra yüksek likidite oranı ve yüksek sermayeye sahip olduğu görülmektedir. En iyi performans gösteren banka grubunun, yabancı sermayeli mevduat bankaları olduğu tespit edilmiştir.

Boyaci ve Baynal (2015), mevcut bireysel kredi müşterilerinin analizi ile karar verme sürecinde risk oranının minimuma indirilmesi ve potansiyel müşterilerin ödeme performanslarını öngörmek için veri madenciliğinin sınıflandırma modellerinden biri olan karar ağaçları yöntemi ile kümeleme yöntemlerinden biri olan K-Means algoritmasını kullanmıştır. Araştırmada SPSS Clementine yazılımı kullanılmıştır. Analiz sonucunda, en uygun küme sayısı üç olarak belirlenmiştir. İlk kümede, vadesi gelmiş ödeme yapmayı geciktiren müşteri grubunun yer aldığı görülmektedir. Bu kümedeki müşterilerin kredi başvuruları olumsuz değerlendirilebilir. İkinci kümedeki müşterilerin %92,73'ünün normal ödeme durumu ve kendine ait evi olmadığı görülmektedir. Bu grupta yer alan müşterilere konut kredisi kullanmaları önerilebilir. Üçüncü kümede yer alan müşterilerin %98'inin düzenli ödeme yaptığı görülmektedir.

Sharahi ve Aligholi (2015), Tahran Sepah Banka Şubelerinin veri kümesi üzerinde iki adım ve K-Means kümeleme algoritmalarını kullanarak bir sınıflandırma modeli sunmuştur. Sepah Bankasının müşterisi olan 60 şirketin segmentasyonu bir tür demografik ve davranışsal segmentasyondur ve sadık müşterilerin belirlenmesine yardımcı olmuştur.

Çalış ve Baynal (2016), bir bankanın bireysel kredi müşterilerine ait satış stratejilerinin belirlenmesine yönelik bir çalışma yapmıştır. Bu çalışmada, K-Means kümeleme algoritması ve analiz sürecinde SPSS programı kullanılmıştır. Yapılan analiz sonucunda, 1. Kümenin çoğunlukla 45-51 yaş aralığında, emekli, ilköğretim mezunu, evi-arabası olmayan, aylık 750-1400 TL aralığında geliri erkek müşterilerden oluştuğu, 2. Kümenin genellikle bekar, 24-30 yaş aralığında, özel ve kamu sektöründe çalışan müşterilerden oluştuğu, 3. Kümenin 38-44 yaş aralığında, emekli, evi-arabası bulunan, 1401-2050 TL aralığında geliri olan erkek müşterilerden oluştuğu görülmektedir.

Bansal ve ark. (2017), K-Means kümeleme algoritmasında küçük bir değişiklik yapmıştır. Bu değişiklik normalleşmeye dayanmaktadır. Araştırmacı, sonuçları bulmak için kanser veri kümesi kullanmıştır. Orijinal veriler oldukça boyutlu olduğundan gereksinimlere dayalı olarak yalnızca beş özellik değerlendirilmiştir. Bu makale, mevcut algoritma için doğruluk oranının %57.14'e eşit olduğunu, geliştirilmiş algoritmanın ise %92.86 değerini kaydettiğini görülmektedir.

Çınar ve Silahtaroğlu (2018), müşteri memnuniyetini etkileyen ve gizli sebepleri bulmak için 301 kişinin katıldığı anket sonuçlarını değerlendirerek, müşteri memnuniyeti veya memnuniyetsizliğini ortaya çıkarmak için K-Means kümeleme ve C5.0 sınıflandırma algoritmalarını kullanmışlardır.

Tekin ve Temelli (2020), Türkiye Bankalar birliğinden elde edilen yedi farklı sermaye rasyosunu kullanarak, Türkiye'de faaliyet gösteren ticari bankaları sermaye rasyolarına göre gruplamak için kümeleme çalışması yapmıştır. Yapılan analiz sonucunda, en uygun küme sayısının dört olduğu belirlenmiştir. Bankaların çoğu tek bir kümede toplanmıştır ve sermaye rasyolarına göre anlamlı farklar olduğu tespit edilmiştir.

2. MÜŞTERİ PROFİLİ BELİRLEME UYGULAMASI

Bu bölümde, veri madenciliğinin kümeleme fonksiyonuna ilişkin tekniklerinden K-Means algoritması bir bankanın pazarlama stratejisini tanımlamak için müşteri segmentasyonu geliştirmeyi amaçlayan bir veri kümesi üzerinde uygulanacaktır. Bu şekilde, veri madenciliğinin kümeleme süreci gerçekleştirilmiş olacak ve uygulamada ortaya çıkan sonuçlar yorumlanarak potansiyel müşteri profil kümesi belirlenecektir. Yapılan kümeleme analiz çalışmasında python programlama dili kullanılacaktır.

2.1. AMAÇ VE KAPSAM

Uygulamanın amacı, veri madenciliğinin kümeleme fonksiyonuna ilişkin yöntemlerinden K-Means algoritmasının kullanılarak heterojen veri kümesinden homojen müşteri gruplarını belirlemek, oluşan bu müşteri gruplarını özellik ve niteliklerine göre tanımlamak, potansiyel olarak birbirinden farklı bu müşteri gruplarına farklı satış ve pazarlama yöntemlerinin uygulanmasının uygunluğunu ortaya koymaktır.

8950 aktif kredi kartı sahibinin kullanım davranışlarını tanımlayan bir banka veri kümesi üzerinde veri madenciliğinin kümeleme yöntemlerinden biri olan K-Means Algoritması uygulanacaktır. Analiz sonucunda kümeleme tanımları ve olası müşteri profili belirlenmeye çalışılacaktır. Çalışmanın veri hazırlama aşamasında MS Excel, veri ön işleme ve modelleme aşamasında Python ve sonuçların analizi aşamasında IBM SPSS yazılımlarından faydalanılmıştır.

Python, 1991 yılında Hollandalı bilgisayar bilimcisi Guido Van Rossum tarafından geliştirilmiştir. Açık kaynak kod lisansına sahiptir. Unix, Linux, Mac, Windows, Amiga, Symbian gibi birçok işletim sistemi ile uyumludur ve işletim sistemleri üzerinde farklı yöntemler ile çalıştırılması mümkündür. Bu platformlar içinde ağ ve ara birim programlaması, veri tabanı yazılımı ve uygulama gibi farklı alanlarda kullanılabilir. Yüksek seviyeli, nesne yönelimli, modüler, etkileşimli ve yorumsal bir dil olan python çok sayıda işi az sayıda kod satırı kullanarak yapmaktadır. Çok büyük yazılımların seri bir şekilde programlanması gerektiğinde C veya C++ gibi dillere göre daha kullanışlıdır.

Son birkaç yılda python, modern yazılım geliştirme, altyapı yönetimi ve veri analizinde birinci sınıf bir programlama dili olarak ön plana çıkmıştır. Veri analizleri ve makine öğrenimi konularında sıkça kullanılan yazılım dilidir. Ara yüzünde birçok kütüphane bulundurur. Python 1990 yılından itibaren geliştirilmektedir. İlk versiyonu olan 1.0 Ocak 1994'te yayınlanmıştır. En güncel versiyonu 3.9.0'dır.

Bu tez çalışmasında kullanılan veri kümesi, çevrimiçi veri bilimcileri makine öğrenimi topluluğu ve halka açık bir veri platformu olan Kaggle.com web sitesi üzerinden alınmıştır. Kaggle, 2010 yılında Anthony Goldbloom, Jeremy Howard, Nicholas Gruen ve Max Levchin tarafından kurulmuştur. Çıkış noktası makine öğrenimi yarışmalarıdır. Günümüzde ise makine öğrenimi yarışmalarının yanı sıra veri bilimi için bulut tabanlı bir çalışma tezgahı ve yapay zeka eğitimi sunan web tabanlı bir platformdur.

Tablo 2.1: Veri Kümesi Tanımı ve Nitelik Tipleri

Değişken	Tanım
X1	Müşteri No
X2	Kredi Kartı Aylık Ortalama Bakiye
X3	Bakiye Güncelleme Sıklığı
X4	Yıllık Toplam Satın Alma Tutarı
X5	Tek Çekim Satın Alım Tutarı
X6	Taksitli Satın Alma Tutarı
X7	Nakit Avans Tutarı
X8	Satın Alma Sıklığı
X9	Tek Çekim Satın Alma Sıklığı
X10	Taksitli Satın Alma Sıklığı
X11	Nakit Avans Sıklığı
X12	Nakit Avans İşlemi Başına Ortalama Tutar
X13	Satın Alma İşlemi Başına Ortalama Tutar
X14	Kredi Limiti
X15	Ödeme Miktarları
X16	Minimum Ödeme Miktarı
X17	Müşteriler Tarafından Yapılan Toplam Ödeme Yüzdesi
X18	Kredi Kart Hizmet Süresi

2.2. VERİ KÜMESİ

Uygulama kısmında kullanılan veri kümesi, 6 ay içerisinde aktif kredi kartı kullanan yaklaşık 8950 adet müşterinin kullanım davranışını göstermekte ve 18 davranış

değişkeni içermektedir. Veri kümesi sayısal niteliklerden oluşmaktadır. Veri kümesi içerisindeki nitelikler ve özellikler Tablo 2.1'de gösterilmiştir.

Analiz sürecine başlamadan önce verilerin dağılımı ve veri kümesinin yapısını anlamak için tanımlayıcı istatistiksel analiz yapılmıştır. Yapılan analiz Tablo 2.2'de gösterilmiştir.

Tablo 2.2: Tanımlayıcı İstatistiksel Analiz

	N	Minimum	Maksimum	Ortalama	Standart Sapma	Çarpıklık	Basıklık	Değişkenlik Katsayısı
X2	8950	0	19043,14	1564,475	2081,5319	2,393	7,675	133%
X3	8950	0	1	0,8773	0,2369	-2,023	3,092	27%
X4	8950	0	49039,57	1003,205	2136,6348	8,144	111,39	213%
X5	8950	0	40761,25	592,4374	1659,8879	10,045	164,19	280%
X6	8950	0	22500	411,0676	904,33812	7,299	96,575	220%
X7	8950	0	47137,21	978,8711	2097,1639	5,167	52,899	214%
X8	8950	0	1	0,4904	0,40137	0,06	-1,639	82%
X9	8950	0	1	0,2025	0,29834	1,536	1,162	147%
X10	8950	0	1	0,3644	0,39745	0,509	-1,399	109%
X11	8950	0	1,5	0,1351	0,20012	1,829	3,335	148%
X12	8950	0	123	3,2488	6,82465	5,721	61,647	210%
X13	8950	0	358	14,7098	24,85765	4,631	34,793	169%
X14	8949	50	30000	4494,45	3638,8157	1,522	2,837	81%
X15	8950	0	50721,48	1733,144	2895,0638	5,908	54,771	167%
X16	8637	0,02	76406,21	864,2065	2372,4466	13,623	283,99	275%
X17	8950	0	1	0,1537	0,2925	1,943	2,432	190%
X18	8950	6	12	11,5173	1,33833	-2,943	7,695	12%
Geçerli N	8636							

Tanımlayıcı istatistik değerlerinden çarpıklık düzeyine ve sapmanın yüksekliğine bakıldığında verilerin uç değerler barındıracağı ve değişkenlerin çoğunun normallik varsayımından uzak olduğunu söyleyebiliriz. Tablo 2.2 incelendiğinde; X14 (kredi limiti) değişkeninin en yüksek minimum değere sahip olduğunu görülmektedir. Değişkenlerin çoğunun çarpıklık değerlerinin sıfırdan büyük olduğunu söyleyebiliriz.

Basıklık değerlerini incelediğimizde ise X8 (satın alma sıklığı), X9 (tek seferlik satın alma sıklığı), X10 (taksitli satın alma sıklığı), X14 (kredi limiti) ve X17 (müşteriler tarafından yapılan toplam ödeme yüzdesi) değişkenlerinin basık, diğer değişkenlerin ve veri kümesinin sivri bir yapıda olduğu söylenebilir. En yüksek değişkenliği X5 (tek

seferlik satın alım tutarı) ve X16 (minimum ödeme miktarı) değişkenlerinin gösterdiğini söyleyebiliriz.

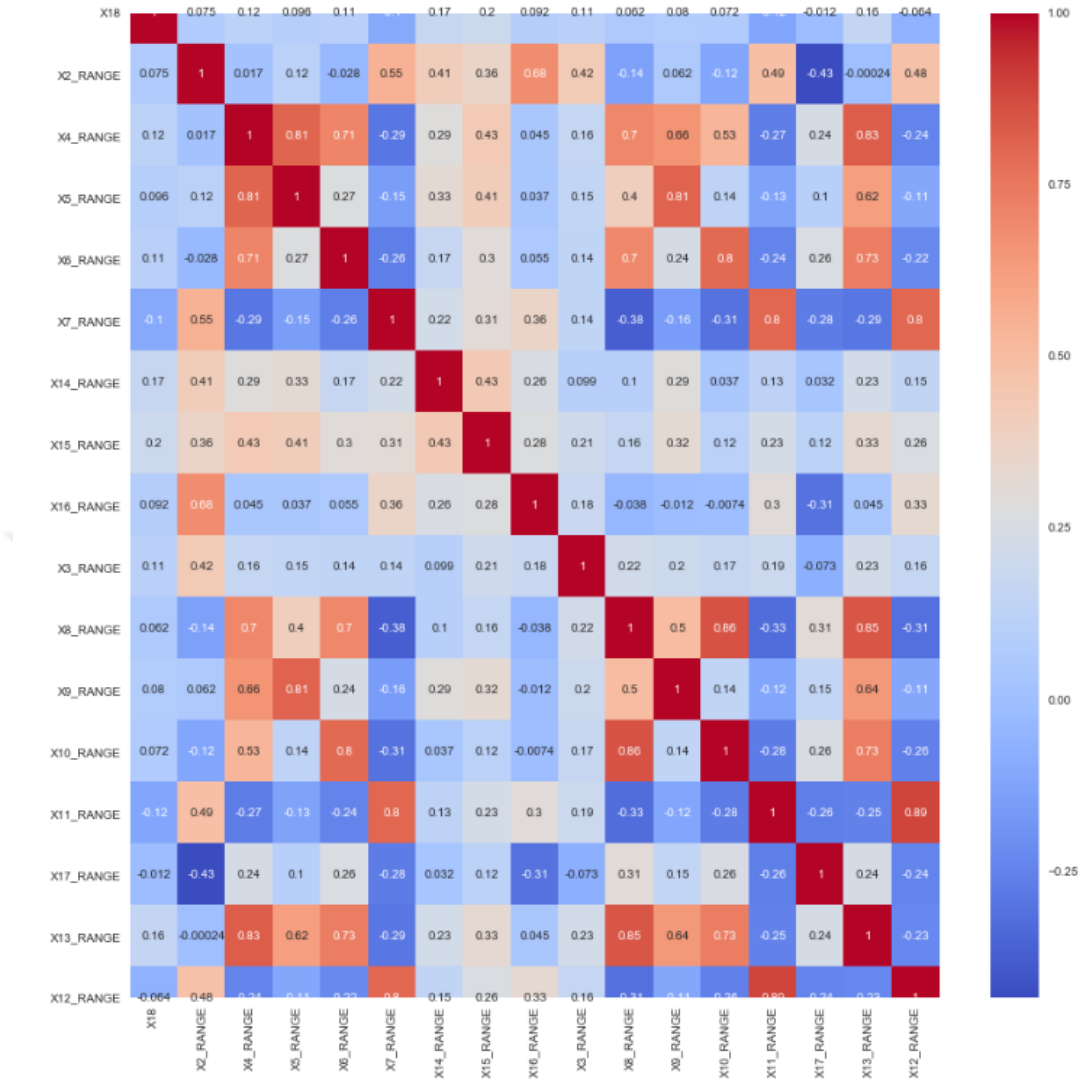
2.3. VERİ ÖN İŞLEME

Veri kümesinin orijinal hali analiz öncesinde kontrol edilerek bazı ön işleme sürecinden geçirilmiş ve uygulamaya hazır hale getirilmiştir. Yapılan veri ön işleme süreci aşağıdaki adımları takip etmektedir.

- Öncelikle veri kümesindeki eksik değerlerin varlığı kontrol edilmiştir. Kontrol sonucunda 'X16' (Minimum Ödeme Miktarı) ve 'X14' (Kredi Limiti) değişkenlerinde eksik değerler saptanmıştır. 'X16' değişkeninde 313 satır, 'X14' değişkeninde 1 satır boş gelmektedir. Bu satırlar veri kümesinden çıkarılmadan ilgili değişkenlerin sahip olduğu ortalama değeri ile doldurularak düzeltilmiştir.
- Analiz sürecinin ilk adımı olan veriyi anlama aşaması için tanımlayıcı istatistiksel analiz yapılmıştır. İstatistiksel veri analizi yapılarak aykırı (uç) değerlerin varlığı kontrol edilmiştir. Yapılan analiz Tablo 2.2'de gösterilmiştir. Veri kümesinde bulunan aykırı değerler ile başa çıkmak için eşit aralıklar oluşturulmuştur.
- Veri kümesi ölçek olarak farklı değişkenler içerdiğinden giriş değişkenleri standartlaştırılarak normalize edilmiştir. Tüm nitelikler 0-1 aralığına ölçeklendirilmiştir.

2.4. DEĞİŞKENLER ARASINDAKİ İLİŞKİLERİN İNCELENMESİ

Kümeleme algoritmalarını veri kümesine uygulama aşamasında birbirlerini etkileme düzeyi ve aralarındaki ilişki oranı yüksek olan değişkenlerin belirlenmesi için korelasyon matrisi kullanılmıştır. Kümeleme analizinde bazı değişkenlerin veri kümesi içerisindeki ağırlık oranı diğer değişkenlere göre daha fazladır. Şekil 2.14'te değişkenler arasındaki korelasyon değerleri görülmektedir. Grafik üzerinde koyu kırmızı noktalar "+1" değerine daha yakın olduğundan değişkenler arasındaki güçlü pozitif ilişkileri, koyu mavi noktalar ise "-1" değerine daha yakın olduğundan değişkenler arasındaki güçlü negatif ilişkileri göstermektedir.



Şekil 2.14: Veri Kümesi İçerisindeki Değişkenlerin Korelasyon Matrisi

Korelasyon matrisine göre, X11-X12 (nakit avans sıklığı – nakit avans işlemi başına ortalama tutar), X8-X13 (satın alma sıklığı – satın alma işlemi başına ortalama tutar) ve X8-X10 (satın alma sıklığı – taksitli satın alma sıklığı) değişkenleri arasında pozitif yönlü güçlü bir ilişki vardır.

X2-X17 (ödeme yüzdesi – aylık ortalama bakiye), X7-X8 (satın alma sıklığı – nakit avans tutarı) ve X4-X7 (yıllık toplam satın alma tutarı – nakit avans tutarı) değişkenleri arasında negatif yönlü güçlü bir ilişki vardır.

2.5. KÜMELEME YÖNTEMİNİN UYGULANMASI

Bu tez çalışmasında, bir bankanın pazarlama stratejisini tanımlamak ve olası müşteri profilini belirlemek için kümeleme analiz çalışması yapılacaktır.

Tablo 2.3: Python'da Kullanılan Kütüphane ve Fonksiyonlar

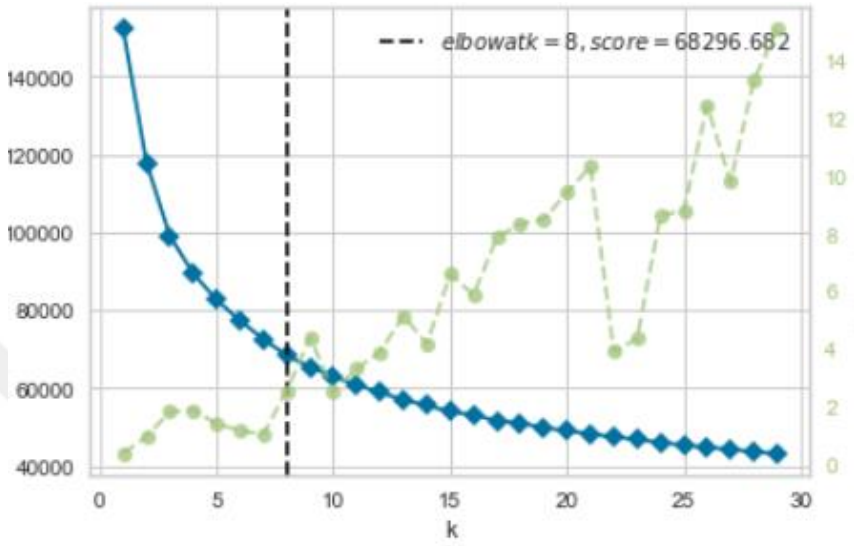
Kullanılan Algoritmalar (Fonksiyonlar)	Kullanılan Kütüphane (Paket)	Kullanım Amacı
StandardScaler	sklearn.preprocessing	Veri Ön İşleme
K-Means	sklearn.cluster	Kümeleme Algoritması
Sns	seaborn	İstatistiksel Veri Görselleştirme
Plt	matplotlib.pyplot	Veri Görselleştirme
Np	numpy	Nümerik Hesaplamalar
Pd	pandas	Veri Ön İşleme ve Analiz
KelbowVisualizer	yellowbrick.cluster	Veri Analizi Görselleştirme

Bu çalışmada K-Means kümeleme algoritması kullanılacaktır. Algoritma uygulanırken kodlar çeşitli kaynaklardan bir araya getirilerek ve tarafımca düzenlenerek amaca uygun hale getirilmiştir[1]. Kullanılan algoritmalar, python kütüphaneleri ve kullanım amaçları Tablo 2.3'te gösterilmiştir. Bu başlık altında K-Means Algoritması ile bankacılık veri kümesi üzerinde veri madenciliği tekniklerinden kümeleme çalışması yapılmıştır. Veri ön işleme adımları ile veri kümesi analize hazır hale geldikten sonra K-Means algoritmasının uygulanabilmesi için öncelikle küme sayısına karar verilmesi gerekmektedir. Küme sayısının belirlenmesi doğru bir kümeleme analizi yapılması ve etkin analiz sonuçları alınması açısından oldukça önemlidir. Küme sayısının (k) seçimi için en çok kullanılan yöntemlerden Elbow Metodu (Dirsek Yöntemi) kullanılmıştır. Elbow Metodu (Dirsek Yöntemi) sonucunda çıkan "k" değeri belirlenerek analiz yapılmıştır. Ulaşılan model üzerinde en iyi kümeleme algoritmasının doğruluğu hedeflenmektedir.

Küme sayısının belirlenmesinde kullanılan Elbow Metoduna (Dirsek Yöntemi) göre Şekil 2.15'te gösterilen eğrinin ilk kırılma noktası k=8 noktasına denk gelir. Grafiğin dikey sol ekseni her bir noktadan merkeze atanan mesafelerin karesinin toplamını, dikey

[1] Kaggle, Clustering & Visualization of Clustering using PCA, 2020, Clustering & Visualization of Clusters using PCA | Kaggle (18 Nisan 2020)

sağ eksen ise kümeleme modelini eğitmek için gerekli süreyi ifade eder. Grafiğin en üstünde hesaplanan k değeri ve puan değerleri gösterilmektedir. Puan değeri, kümeler arasındaki ve kümeler içindeki dağılım oranını ifade eder. Bu durumda ilk etapta küme sayısı k=8 olarak seçilir ve K-Means algoritması uygulanır.

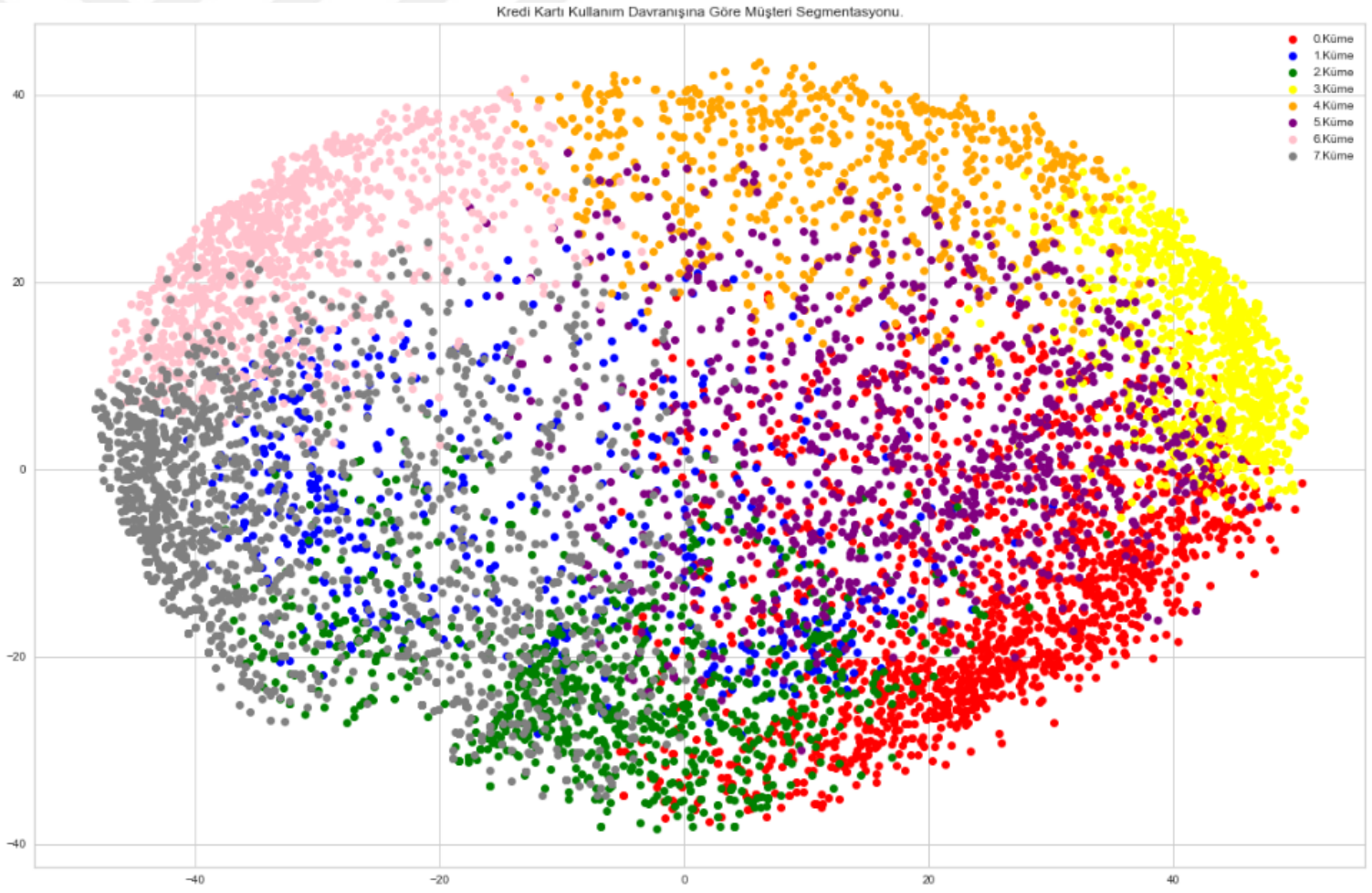


Şekil 2.15: Elbow Metodu (Dirsek Yöntemi)

Elbow Metoduna (Dirsek Yöntemi) göre karar verilen küme sayısına uygun olarak (k=8) K-Means algoritması çalıştırılmıştır. Yapılan analiz sonucunda ortaya çıkan grafik sonuçları tezin *"Ekler"* bölümünde gösterilmiştir. Oluşturulan sekiz küme iki boyuta indirgenerek Şekil 2.16'da belirtilmiştir.

2.5.1. Kümeler Arası Farklılıkların İncelenmesi

Her bir küme için değişkenler bazında anlamlı farklılıkların olması beklenir. Bu beklentinin gerçekleşip gerçekleşmediğini tespit etmek için Varyans Analizi (ANOVA) tekniği kullanılabilir. Bilindiği üzere, Varyans Analizi bir değişkene ait birden fazla gözlem grubu arasında ortalamalar açısından anlamlı farkın belirlenmesinde kullanılır. Ancak analize konu olan değişkenlerin Normal dağılıma uygunluktan uzak olduğu tanımlayıcı istatistiklerden dahi kolayca anlaşılmaktadır. Bu nedenle, Varyans Analizine karşılık gelen ve parametrik olmayan bir yöntem olarak Kruskal-Wallis sınaması tercih edilmiştir (Armutlulu İ.H. 2000).



Şekil 2.16: Kredi Kartı Kullanım Davranışına Göre Müşteri Segmentasyonu

Kruskal-Wallis sınamasında hipotezler aşağıdaki gibidir;

H_0: Küme grupları açısından X_k değişkeninin ortancaları arasında fark yoktur.

H_1: En az bir küme grubunda X_k değişkenin ortancası diğerlerinden farklıdır.

Analize dahil edilen 17 değişken için sekiz küme grubu üzerinde uygulanan Kruskal-Wallis sınaması sonuçları aşağıdaki tabloda (Tablo 2.4) gösterilmiştir. Bu sonuçlara göre, tüm değişkenler için H_0 hipotezinin 0,05 anlamlılık düzeyinde reddedildiği görülmektedir. Bir başka deyişle, yapılan analiz sonucunda kümelerin tüm değişkenlerde anlamlı farklılıklara sahip olduğu tespit edilmiştir.

Tablo 2.4: Kruskal-Wallis Sınaması Sonuçları

Değişken	H istatistiği	Serbestlik Derecesi	Anlamlılık
X2	6337,443	7	0,000E+00
X3	5151,446	7	0,000E+00
X4	5338,697	7	0,000E+00
X5	4895,171	7	0,000E+00
X6	5468,035	7	0,000E+00
X7	4511,654	7	0,000E+00
X8	3189,038	7	0,000E+00
X9	3956,843	7	0,000E+00
X10	3460,451	7	0,000E+00
X11	4596,452	7	0,000E+00
X12	6226,945	7	0,000E+00
X13	4970,858	7	0,000E+00
X14	5979,554	7	0,000E+00
X15	4403,855	7	0,000E+00
X16	857,8006	7	0,000E+00
X17	6491,041	7	0,000E+00
X18	4205,722	7	0,000E+00

Ancak bu anlamlı farkların her bir değişken düzeyinde hangi gruplardan kaynaklandığını belirlemek için ikili karşılaştırmaların yapılması gerekmektedir. Bu amaçla, ikili karşılaştırmalara dayalı birçok teknikten eşit olmayan varyanslar halinde etkili sonuçlar veren Tamhane's T^2 istatistiği uygulanmıştır.

Tamhane's T^2 sınamasında hipotezler aşağıdaki gibidir;

H_0: X_k değişkeninin K_i ve K_j kümelerindeki ortalamaları arasında fark yoktur.

H_1: X_k değişkeninin K_i ve K_j kümelerindeki ortalamaları arasında fark vardır.

SPSS programında yapılan Tamhane's T^2 sınamasının sonuçları tezin ekler bölümünde (Ek 2) gösterilmiştir. Yapılan ikili karşılaştırma analizi 0,05 anlam düzeyinde yorumlanmıştır. Veri kümesine uygulanan K-Means algoritması sonucunda elde edilen 8 kümeye ait ikili karşılaştırma bulguları Tablo 2.5'te beş kategorili bir ölçekle kodlanarak tanımlanmıştır.

Tablo 2.5: Kümeler Arası Farklılıklara İlişkin Bulgular

Ort.	K0	K1	K2	K3	K4	K5	K6	K7
X2	0	-	0	+	0	0	0	+
X3	-	-	-	0	+	-	+	0
X4	0	--	-	++	++	+	--	--
X5	-	0	0	++	++	+	0	0
X6	+	0	0	++	++	0	--	-
X7	--	+	-	--	++	-	++	0
X8	-	--	0	++	++	+	++	-
X9	-	--	--	++	++	0	+	--
X10	-	-	-	0	++	0	+	0
X11	0	-	--	++	++	0	+	+
X12	+	0	-	++	+	+	--	--
X13	-	0	-	++	+	++	-	-
X14	++	+	0	++	++	0	-	-
X15	--	+	--	--	++	-	++	0
X16	++	+	++	++	0	+	0	-
X17	+	0	0	++	++	+	-	-
X18	-	+	0	-	++	0	++	+

Tablo 2.5'te her değişken sekiz kategorili olarak düşünülmüş ve ikili karşılaştırma analizi yapılmıştır. Yapılan analiz sonucunda anlamlı farklar belirlenmiştir. Değişkenlerin her bir kümedeki ortalamaları arasında tespit edilmiş olan anlamlı farklılıklar H_0 hipotezini ölçü alarak pozitif (+) ya da negatif (-) olmak üzere sınıflandırılmıştır. Tablo 2.6'da bu sınıflandırma için kullanılan semboller ve bu sembollere ait tanımlamalar belirtilmiştir.

Tablo 2.6: Kümeler Arası Anlamlı Farklılık Düzeyleri

İşaret Tanımı	Potansiyel
++	Anlamlı Şekilde Çok Yüksek
+	Anlamlı Şekilde Yüksek
0	Anlamlı Fark Yok
-	Anlamlı Şekilde Düşük
--	Anlamlı Şekilde Çok Düşük

2.5.2. Kümelerin Tanımlanması

Analizler sonucunda oluşan küme gruplarının renkleri ve karakteristik özellikleri aşağıdaki gibi tanımlanır:

0.Küme (Kırmızı): Bu grupta yer alan müşteriler yüksek kredi limiti ve minimum ödeme miktarına, ortalama kredi kartı bakiyesi ve nakit avans sıklığına sahiptir. Yüksek taksitli satın alma işlemi gerçekleştirirler. Düşük nakit avans tutarı ve ödeme miktarına sahip müşteriler bu kümede yer alır. (Taksit Sevenler)

1.Küme (Mavi): Bu grupta yer alan müşteriler, düşük tek çekim satın alma sıklığı, düşük yıllık toplam satın alma tutarı ve sıklığına sahiptir. Ortalama seviyede tek çekim ve taksitli satın alma işlemi gerçekleştirirler. Yüksek kredi limitine sahip olan bu müşterilerin yüksek nakit avans tutarı ve ödeme miktarı bulunmaktadır. (Kredi Limiti Yüksek Nakit Çekenler)

2. Küme (Yeşil): Bu grupta yer alan müşteriler, düşük tek çekim satın alma sıklığı, düşük nakit avans sıklığı ve ödeme miktarına sahiptirler. Çok sık bakiye güncelleme işlemi gerçekleştirmezler. Düşük seviyede satın alma işlemi gerçekleştirirler. Nakit avans almayı ve taksitli işlem yapmayı sevmezler. Ancak yüksek ödeme miktarına sahiptirler. (Tüketim Alışkanlığı Düşük Olanlar)

3. Küme (Sarı): Bu grupta yer alan müşteriler, yüksek kredi limiti ve yıllık toplam satın alma tutarına sahiptirler. Aynı zamanda, yüksek tek çekim ve taksitli satın alma işlemi gerçekleştirirler. Müşteriler Tarafından Yapılan Toplam Ödeme Yüzdesi çok yüksektir. Aylık kredi kartı bakiyeleri yüksek seviyededir ve alışveriş yapmayı severler. (Alışveriş Yapmayı Sevenler)

4. Küme (Turuncu): Bu grupta yer alan müşteriler, 3. Kümeden sonra 2. yüksek satın alma tutarı, kredi limiti ve toplam ödeme yüzdesine sahiptirler. Her türlü satın alma işlemi gerçekleştirirler. (Taksitli, Nakit Avans, Tek Çekim vs.)
Büyük ödemeler yaparlar. (Büyük Harcama Yapanlar)

5.Küme (Mor): Bu grupta yer alan müşteriler, 3. yüksek satın alma ve tek çekim satın alma tutarına sahiptirler. Sık sık satın alma işlemi ve yüksek ödeme miktarı gerçekleştirirler. Bakiye güncelleme sıklığı yavaştır ve nakit avans almayı sevmezler. (Tek Seferlik Satın Alma Gerçekleştirenler)

6. Küme (Pembe): Bu grupta yer alan müşteriler, yüksek nakit avans tutarı, ödeme miktarı ve kredi kartı hizmet süresine sahip olmalarına rağmen en düşük seviyede satın alma işlemi gerçekleştirirler. (Nakit Avans Sevenler)

7. Küme (Gri): Bu grupta yer alan müşteriler, en düşük satın alma tutarına sahiptirler. Düşük seviyede ödeme gerçekleştirirler. (Az Para Harcayanlar)

2.5.3. Kümeler Arası Farklılıkların Değerlendirilmesi

K-Means algoritması ile kredi kartı verileri kullanılarak müşterileri kümelemek mümkündür. Kümeler arasında çeşitli değişkenler düzeyinde anlamlı farklılıkların olduğu tespit edilmiştir. Bu farklılıklar, her bir kümede yer alan müşteri gruplarının birbirlerinden farklı kredi kartı kullanım davranışları gösterdiğini kanıtlamaktadır. Bu farklılıklar ilgili işletmeler tarafından satış ve pazarlama süreçlerinde kullanılabilir. Bu anlamda aşağıdaki önerilerde bulunulmuştur.

- Eldeki değişkenler ile bankacılık müşterileri homojen alt gruplara ayrılarak müşteri sadakat sistemi oluşturulabilir.
- Müşteri segmentasyon çalışması ölçülebilir, ulaşılabilir, gerçekçi, savunulabilir, dayanıklı ve rekabetçi yapıldığı durumda her bir gruptaki homojen müşteri yapısının yönetimininde kolaylık sağlanabilir.
- Müşterilerin talepleri, şikayetleri, ihtiyaçları vb. benzerlik göstereceğinden, gruplandırma işlemi ile müşteri memnuniyetini sağlamak kolaylaşabilir.
- Bankalar, müşterilerini homojen alt gruplara bölerek, müşteri beklentilerine uygun ürün öneri sistemi geliştirerek satış ve karlılığı artırabilir.
- Bankalar, başarılı müşteri segmentasyonu ile ileriye yönelik projeksiyon ve hedef odaklı kampanya stratejisi gerçekleştirebilir.
- Müşteri segmentasyon çalışması, daha kolay pazarlama, niş pazar bulabilme imkanı sağlayarak verimliliği artırma yönünde yardımcı olabilir.

SONUÇ

Günümüzde gelişen teknoloji ile birlikte veri sayısı ve türleri hızla artmaktadır. Karmaşık ve birikmiş büyük veri yığınları içerisinden anlamlı bilgiler elde etmek için veri madenciliği teknikleri kullanılmaktadır. Geleceğe yönelik çıkarımlar ve analizler yapmak için bu teknikler gelişmiş bilgisayar programları aracılığıyla uygulanmaktadır. Veri madenciliğinin pazarlama, satış, sigorta, tıp, finans, bankacılık vb. olmak üzere pek çok uygulama alanı bulunmaktadır.

Bu tez çalışmasında kaggle.com web sitesi üzerinden elde edilen bankacılık sektörüne ait bir veri kümesi kullanılmıştır. Veri madenciliğinin kümeleme yöntemlerinden biri olan K-Means algoritması kullanılarak heterojen veri kümesi içerisinden homojen müşteri gruplarının belirlenmesi, oluşan bu müşteri gruplarının özelliklerine göre potansiyel müşteri profil kümesinin oluşturulması ve oluşan müşteri kümelerine özel pazarlama stratejisinin belirlenmesi amaçlanmıştır.

Elde edilen veri kümesi sayısal niteliklerden oluşmaktadır. Veri madenciliği sürecinin ilk aşaması olan verinin anlaşılması için veri kümesine tanımlayıcı istatistiksel analiz testi uygulanmıştır. Bunun sonucunda veri içerisindeki değişkenlerin özellikleri ve verinin dağılım yapısı incelenmiştir. Veri kümesi aykırı, gürültülü ve uç değerlerden temizlenerek standardize edilmiş ve analize hazır hale getirilmiştir. Daha sonra python programında K-Means algoritması uygulanmıştır. Analiz sonucunda sekiz müşteri kümesi oluşmuştur. Oluşan sekiz kümenin grup içinde ve gruplar arasındaki anlamlı farklılığın tespiti için SPSS programı kullanılarak Kruskal-Wallis sınaması yapılmıştır. Bunun sonucunda her bir değişkende anlamlı farklılıklar tespit edilmiştir. Kümeler arasında bulunan anlamlı farklılıkların hangi küme gruplarından kaynaklandığını bulmak için Tamhane's T^2 sınaması uygulanmıştır ve oluşan sekiz kümenin karakteristik özellikleri belirlenmiştir. Böylece, farklı nitelikteki müşteri profilleri gruplandırılarak, işletmeler tarafından satış ve pazarlama süreçlerinde kullanılabilir. Bu sayede, müşterilere yönelik özel pazarlama stratejileri geliştirilebilir.

Bu çalışma temelinde, farklı zaman periyodları için de kümele analizi yapılarak sonuçların tutarlılığı kıyaslanabilir. Aynı zamanda, bu analize farklı parametreler de

eklenebilir. Örneğin; kriz dönemleri, doğal afetler, salgın hastalıklar vb. süreçlerde müşterilerin kredi kartı kullanım davranışları için ayrı analizler yapılarak farklı müşteri profilleri ve pazarlama stratejileri belirlenebilir.

KAYNAKLAR

Kitaplar:

Anderberg, Michael R. (1973). *Cluster analysis for applications*. New York: Academic Press.

Alpar, R. (2011). *Çok Değişkenli İstatistiksel Yöntemler.* Ankara: Detay Yayıncılık.

Armutlulu İ.H. (2000). *İşletmelerde Uygulamalı İstatistik*. 2. Baskı. Alfa.

Chapman, P., J. Clinton., R. Kerber., T. Khabaza ve T. Reinartz. (2000). *CRISP-DM 1.0 Step-by-Step Data Mining Guide*. Technical report

Gan. G., C. Ma ve J. Wu. (2007). *Data Clustering: Theory, Algorithms, and Application*. Philadelphia: Society For Industrial & Applied Mathematics.

Han, J., M. Kamber ve J. Pei. (2012). *Data Mining Concepts and Techniques: Third Edition.* Waltham: Elsevier.

Hair, J. F., W. C. Black, B. J. Babin ve R. E. Anderson. (2014). *Multivare Data Analysis.* Edinburg: Pearson.

Hair, J.F. ve W.C. Black. (2000). *Cluster Analysis. In Laurence G.Grimm & Paul R. Yarnold (Eds.). Reading and Understanding More Multivariate Statistics.* Washington, D.C.: American Psychological Association.

Jain, A.K ve R.C.Dubes. (1988). *Algorithms for Clustering Data.* New Jersey: Prentice-Hall.

Johnson, A.R ve D.W. Wichern. (1992). *Applied Multivare Statistical Analysis*. New Jersey: Prentice-Hall.

Kaufmann, L. ve P.J. Rousseeuw. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis.* New York: John Wiley & Sons, Inc.

Kantardzic, M. (2003). *Data Mining: Concepts, Models, Methods, and Algorithms.* New Jersey: John Wiley & Sons Inc.

Kononenko, I ve M. Kukar. (2007). *Machine Learning and Data Mining.* Chichester: Woodhead Publishing.

Kumar, T.S. (2014). *Introduction to Data Mining: First Edition*. Harlow: Pearson Education Limited.

Larose, D. T. (2005). *In Discovering Knowledge in Data: An Introduction to Data Mining.* New Jersey: John Wiley & Sons, Inc.

Mirkin, B. (2005). *Clustering for Data Mining: A Data Recovery Approach.* London: Chapman &Hall/CRC.

Romesburg, H.C. (1984). *Cluster Analysis for Researcher*. California: Lifetime Learnibg Publications.

Senath, P.H.A. ve R.R. Sokal. (1973). *Numerical Taxonomy*. San Francisco: W.H Freeman and Company.

Sumathi, S., S.N. Sivanandam. (2006). *Introduction to Data Mining and its Applications.* Berlin: Springer.

Tan, P. N., M. Steinbach, ve V. Kumar. (2006). *Introduction to Data Mining.* Boston MA: Pearson Addison Wesley.

Tryfos, P. (1997). *Methods for Business Analysis and Forecasting: Text & Cases.* New Jersey: Wiley.

Zaki, M.J. ve Meira. M.JR. (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. New York: Cambridge University Press.

Webb, A.R. (2002). *Statistical Pattern Recognition, Second Edition.* Chichester: John Wiley & Sons, Ltd.

***Makaleler ve Süreli Yayınlar*:**

Akpınar, H. (2000). Veri Tabanlarında Bilgi Kesfi ve Veri Madenciligi. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*. 29.1, 7.

Almeida J., L.M.S. Barbosa., A. Pais ve S. Formosinho. (2007). Improving Hierarchical Cluster Analysis: A new Method With Outlier Detection and Automatic Clustering. *Chemometrics and Intelligent Laboratory Systems.* 87.2, 208-217.

Aşan, Z. (2007). Examining the Socioeconomic Characteristics of Customers Using Credit Cards, with Clustering Analysis. *Dumlupınar University The Journal of Social Sciences. 17.*, 256-267.

Bansal, A., M. Sharma ve S. Goel. (2017). Improved K-mean Clustering Algorithm for Prediction Analysis Using Classification Technique in Data Mining. *International Journal of Computer Applications.* 157.6, 35.

Bansal, K. ve A. Bohra (2013). K-Means Clustering Algorithm Implemented To E-Banking. *International Journal of Engineering Research & Technology 2.4*, 1644-1650.

Bansal, A., M. ve S. Goel (2017). Improved K-Mean Clustering Algorithm for Predistion Analysis Using Classification Technique in Data Mining. *Int. J. Comput. Appl. 157.6*, 1-7.

Banfield, J.D ve A.E. Raftery. (1993). Model-Based Gaussian and Non-Gaussian Clustering. *International Biometric Society.* 49, 803-821.

Baykal, A. (2006). Veri Madenciliği Uygulama Alanları. *D.Ü.Ziya Gökalp Eğitim Fakültesi Dergisi.* 7, 95-107.

Çalış, A. ve K. Baynal. (2016). Kümeleme Analizi İle Bankacılık Sektöründe Satış Stratejilerinin Belirlenmesi. *Beykent Üniversitesi Fen ve Mühendislik Bilimleri Dergisi.* 9., 13-41.

Çınar, A. ve G. Silahtaroğlu. (2018). Veri Madenciliği Teknikleri İle Müşteri Memnuniyetine Etki Eden Gizli Nedenlerin Keşfi. *ResearchGate.* 3.

Chen, S.Y ve X. Liu. (2004). The Contribution of Data Mining to Information Science. *Journal of Information Science.* 30.6, 550-558.

Colak, I., S. Sagiroglu ve M. Yesilbudak. (2012). Data Mining and Wind Power Prediction: A Literature Review. *Renewable Energy.* 46, 241-247.

Çakmak, Z. (1999). Kümeleme Analizinde Geçerlilik Problemi ve Kümeleme Sonuçlarının Değerlendirilmesi. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi.* 3, 187-205.

Demiralay, M ve A.Y. Çamurcu. (2005). Cure, Agnes ve K-MeansAlgoritmalarındaki Kümeleme Yeteneklerinin Karşılaştırılması. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi.* 4.8, 1-18.

Diwani, S., S. Msihol., D.S. Kayange ve D. Machuve. (2013). Overview Applications of Data Mining In Health Care: The Case Study of Arusha Region. *International Journal of Computational Engineering Research.* 3.8, 73-77.

Dudoit, S. ve J. Fridlyand. (2002). A Prediction-Based Resampling Method For Estimating The Number of Clusters In A Dataset. *Genome Biology.* 3.7, 1-20.

Estivill-Castro, V ve Yang, J. A. (2000). Fast and Robust General Purpose Clustering Algorithms. *Pacific Rim International Conference on Artificial Intelligence.* 208-218.

Faajian, M.A ve S.M. Mohammadi. (2010). Mining The Banking Customer Behavior Using Clustering and Association Rules Methods. *International Journal of Industrial Engineering & Production Research*. 21.4, 239-245.

Fraley, C. ve A.E. Raftery. (1998). How Many Cluster? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal.* 41.8, 578-588.

Girija, N ve S. Srivatsa. (2006). A research study: Using Data Mining in Knowledge Base Business Strategies. *Information Technology Journal.* 5.3, 590-600.

Guzman, L.S., V. Saquicela, E.A. Ordonez ve J. Vandewalle. (2015). Literature Review of Data Mining Applications in Academic Libraries. *The Journal of Academic Librarianship.* 41.4, 11.

Hand, D. J. (1998). Data Mining: Statistics and more?. *The American Statistician.* 52.2, 112-118.

Jain, A.K., M.N. Murty ve P.J. Flynn. (1999). Data Clustering: A Review. *ACM Computing Surveys.* 31.3, 164-323.

Jain, A.K. (2010). Data clustering: 50 Years Beyond K-means. *Pattern Recognition Letters*. 31, 653.

Jatain, A., A. Nagpal ve D. Gaur. (2013). Review Based on Data Clustering Algorithms. *IEEE International Conference on Information and Communication Technologies.* 6.13, 171.

Jiman, A.V. ve H.K. Khanuja. (2016). Literature Survey: Clustering Technique. *International Journal of Computer Applications Technology and Research.* 5.10, 671-674.

Kaur, M. ve N. Kaur. (2014). "Adaptive K-MeansClustering Techniques For Data Clustering. International Journal of Innovative Research in Science Engineering and Technology. 3.9.

Kandemir, T. ve N.D. Arıcı (2013). Mevduat Bankalarında CAMELS Performans Değerleme Modeli Üzerine Karşılaştırmalı Bir Çalışma. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. 18.1., 61-87.*

Keçek, T. ve V. Cinsek (2013). Türkiye'de Faaliyette Bulunan Ticaret Bankalarının Performanslarına Göre Sınıflandırılmasında Etkili Olan Değişkenlerin Belirlenmesi ve Bir Uygulama Denemesi. *Dumlupınar Üniversitesi Sosyal Bilimler Fakültesi Dergisi. 22., 189-206.*

Khajvand, M., ve M.J. Tarokh (2011). Analyzing Customer Segmentation Based on Customer Value Components (Case Study: A Private Bank) (Technical note. *Journal of Industrial Engineering, University of Tehran.* 79-93.

Lloyd, J.F., R.S. Evans, G.J. Stoddard ve J.R. Nebeker. (2005). Risk Factors For Adverse Drug Events. *The Annals of Pharmacotherapy*. 39, 1161-1168.

MacQuenn, J. (1967). Some Methods For Classification and Analysis of Multivariate Observation. *In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability.* 1, 281-297.

Madhulatha, T.S. (2012). An Overview on Clustering Methods. *IOSR Journal of Engineering*. 2.4, 719-725.

Patil, C. ve I. Baidari. (2019). Estimating the Optimal Number of Clusters k in a Dataset Using Data Depth. *Data Science and Engineering.* 4, 133-134.

PhridviRaj, M ve C. GuruRao. (2014). Data Mining – Past, Present and Future A Typical Survey on Data Streams. *Procedia Technology.* 12, 255 – 263.

Ramageri, B.M. (2010). Data Mining Techniques and Applications. *Indian Journal of Computer Science and Engineering.* 1.4, 301-305.

Rendon, E., I. Abundez, A. Arizmendi ve E.M. Quiroz. (2011). Internal Versus External Cluster Validation Indexes. *International Journal Of Computers and Communcation.* 5.1, 27-34.

Rokach, L. ve O. Maimon. (2005). *The Data Mining and Knowledge Discovery Handbook.* Boston: Sipringer.

San, O.M., V.N. Huynh ve Y. Nakamori. (2004). An Alternative Extension of The K-MeansAlgorithm For Clustering Categorical Data. *Int. J. Appl. Math. Comput. Sci.* 14.2, 241-247.

Sahu, H., S. Shrma ve S. Gondhalakar. (2011). A Brief Overview on Data Mining Survey. *International Journal of Computer Technology and Electronics Engineering.* 1.3, 115.

Shearer, C. The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*. 5.4, 12-22

Sharahi, M. ve M. Aligholi (2015). Classify The Data of Bank Customers Using Data Mining and Clustering Techniques a Retail Bank Portfolio. *J. Appl. Environ. Biol. Sci.* 5., 458-464.

Singh, A., A. Yadav ve A. Rana. (2013). K-Meanswith Three Different Distance Metrics. *International Journal of Computer Applications.* 67.10, 13.

Sugar, C.A. ve G.M. James. (2003). Finding the Number of Clusters In A Data Set - An Information Theoretic Approach. *Journal of the American Statistical Association.* 98.463, 750-763.

Tibshirani, R., G. Walther ve T. Hastie. (2001). Estimating The Number of Clusters In A Data Set Via The Gap Statistic. *Royal Statistical Society.* 63.2, 411-423.

Tekin, B (2015). Temel Sağlık Göstergeleri Açısından Türkiye'de ki İllerin Gruplandırılması: Bir Kümeleme Analizi Uygulaması. *Çankırı Karatekin Üniversitesi, İktisadi ve İdari Bilimler Fakültesi.* 5.2, 389-416.

Tekin, B. ve F. Temelli (2020). K-Ortalamalar Yöntemi İle Bankaların Sermaye Yeterlilik Rasyolarına Göre Kümelenmesi. *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi. 10.1*, 11-36.

Tudor, A.I., A. ve E. Andrei (2012). Clustering Analysis for Credit Default Probabilities in a Retail Bank Portfolio. *Bucharest Academy of Economic Studies, Database Systems Journal.* 3.2, 23-30.

Yin, Z., X. Zhou, C. Bakal ve F. Li. (2008). Using Iterative Cluster Merging With Improved Gap Statistics To Perform Online Phenotype Discovery In The Context of High-Throughput RNAi Screens. *BMC Bioinformatics.* 9.264, 1-20.

Xu, D. ve Y. Tian. (2015). A Comprehensive Survey of Clustering Algorithms. *Ann. Data. Sci.* 2.2, 165-193.

Xu, S., X. Qiao, L. Zhu ve Y. Zhang. (2016). Reviews on Determining the Number of Clusters. *Applied Mathematics & Information Science An International Journal.* 10.4, 1495.

Diğer Kaynaklar:

Ankerst, M., M.M. Breuning,., H.P. Kriegel ve J. Sander. (June 1996). OPTICS: Ordering Points To Identify the Clustering Structure. *In Proceedings of the International Conference on Management of Data (SIGMOD).* Munich: Institute for Computer Science, University of Munich. 49-60.

Aggrawal, R., J. Gehke, D. Gunopulos ve P. Raghavan. (June 1998). Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications. *Proc. ACM SIGMOD International Conference on Management of Data Seattle.* San Jose: IBM Almaden Research Center. 94-105.

Boyaci, A., A. ve K. Baynal (2015). Data Mining Application in Banking Sector with Clustering and Classification Methods. *Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management.* Dubai: IEEE.

Bilen, H (2009). Bankacılık Sektöründe Personel Seçimi ve Performans Değerlendirilmesine İlişkin Veri Madenciliği Uygulaması. *Yayınlanmış Yüksek Lisans Tezi*. Ankara: Gazi Üniversitesi. FBE.

Dinler, M. (2014). Kümeleme analizi Yöntemlerinin Hayvancılık Verilerinde Karşılaştırmalı Olarak İncelenmesi. *Yayımlanmamış Yüksek Lisans Tezi*, Bingöl: Bingöl Üniversitesi. FBE.

Doğan, B (2008). Clustering Analysis As a Tool Under the Supervision of Banks: An Application for Turkish Banking Sector. *PhD in Thesis.* İstanbul: Kadir Has University. Graduate School of Social Sciences.

Ester, M., H.P. Kriegel, S. Sander ve X. Xu. (August 1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96 Proceeding of the 2nd International Conference on Knowledge Discovery and Data Mining.* Menlo Park: University of Munich, 226-231.

Eszergár-Kiss, D ve B. Caesar. (5-7 September 2016). Definition of User Group Applying Ward's Method. *19th Euro Working Group on Transportation Meeting.* İstanbul, 25-34.

Gündüz, S (2011). Uzaklık Fonksiyonlarının Çok Boyutlu Ölçekleme Algoritmalarındaki Etkinliğinin İncelenmesi ve Uygulamalar. *Yayınlanmış Yüksek Lisans Tezi*. Adana: Çukurova Üniversitesi. FBE.

Hinneburg, A ve D.A. Keim. (August 1998). An Efficient Approach to Clustering in Large Multimedia Databases with Noise. *KDD-98 Proceeding of the Fourth*

International Conference on Knowledge Discovery and Data Mining. Halle: Institute of Computer Science, University of Hall. 58-65.

Kazaz, N (2019). Veri Madenciliğinde Kümeleme Analizi Yöntemlerinin İncelenmesi ve Sağlık Bilimleri Alanındaki Uygulamaları. *Yüksek Lisans Tezi.* İstanbul: İstanbul Üniversitesi. SBE.

Kaggle. Clustering & Visualization of Clustering using PCA. 2020. Clustering & Visualization of Clusters using PCA | Kaggle (18 Nisan 2020).

Koldere, Y (2008). Veri Madenciliğinde Kümeleme Algoritmaları ve Kümeleme Analizi. *Doktora Tezi.* İstanbul: Marmara Üniversitesi. SBE.

Lunfeng, G., L. Huan ve Z. Li. (May 2012). The Application of Association Rules of Data Mining in Book-Lending Service. *International Conference on Fuzzy Systems and Knowledge Discovery.* Sichuan: IEEE. 761-764.

Li Y ve H. Wu. (December 2012). A Clustering Method Based On K-MeansAlgorithm. *International Conference on Solid State Devices and Materials Science.* China: Physics Procedia

Miljković, D. (22-26 Mayıs 2017). Brief Review of Self-Organizing Maps. *International Convention on Information and Communication Technology, Electronics and Microelectronics.* Hırvatistan: IEEE, 1252-12.

Overview of Clustering Methods. (t.y). https://scikit-learn.org/stable/modules/clustering#k-means (21 June 2020)

Özyazar, Ö (2019). Data Mining Applications For Sustainable Medical Systems: A Study On Diabetes. *Master Thesis.* İstanbul: Marmara University. Institute for Graduate Studies in Pure and Applied Sciences.

Servi, T (2009). Çok Değişkenli Karma Dağılım Modeline Dayalı Kümeleme Analizi. *Doktora Tezi.* Adana: Çukurova Üniversitesi. FBE.

Ullman, S., T. Poggio, D. Harari ve D. Zysman. (2014). Unsupervised Learning Clustering. *Center for Brains, Minds & Machines.* https://www.ccs.neu.edu/home/vip/teach/DMcourse/2_cluster_EM_mixt/notes_slides/Class13.pdf (20 June 2020)

Uslu, T (2019). Kümeleme Analizi ve Sağlık Alanında Bir Uygulama. *Yüksek Lisans Tezi.* Malatya: İnönü Üniversitesi. SBE.

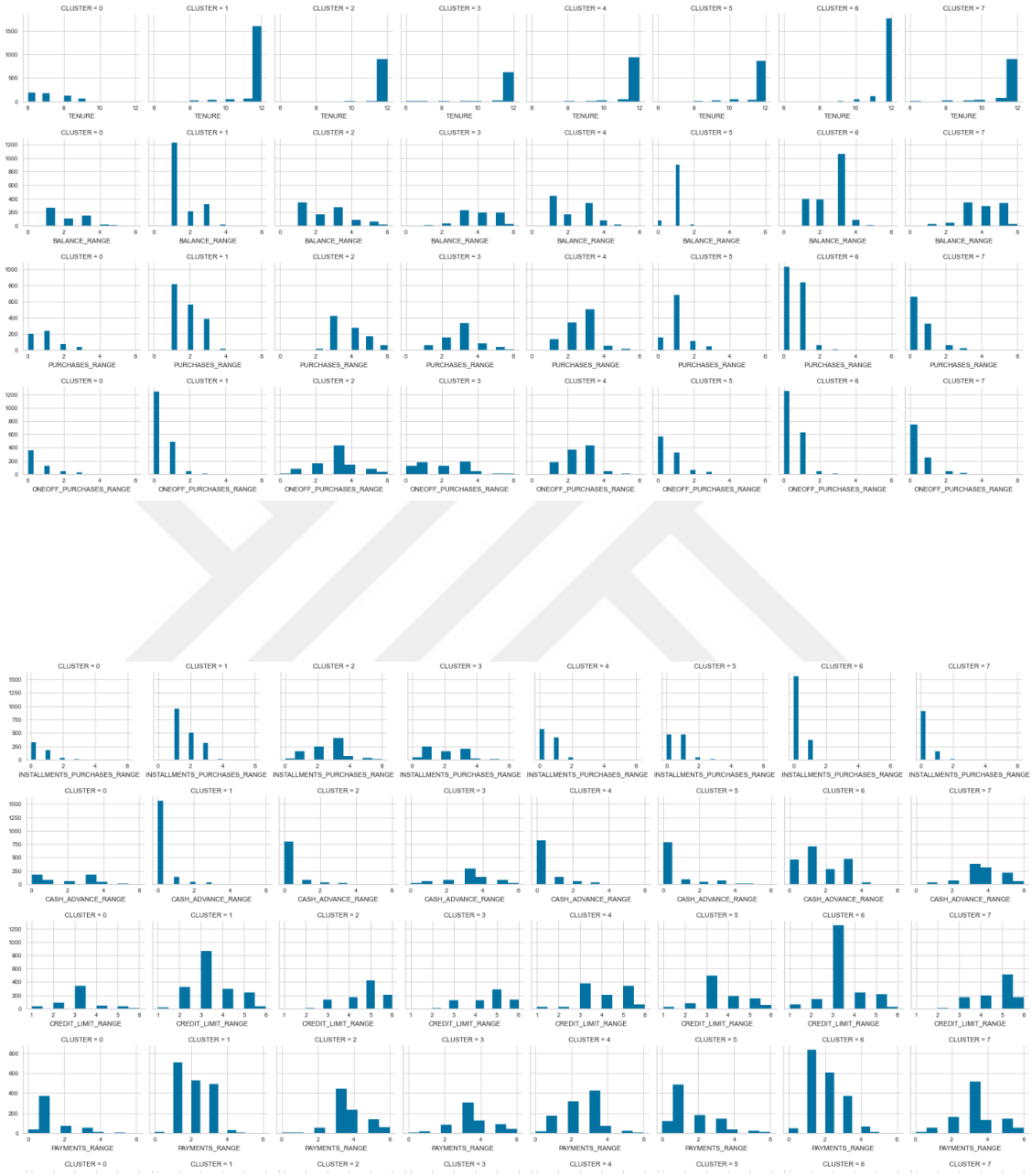
Üstünel, M (2018). K-Meanslar Algoritmasına Dayalı Kümeleme Analizi Sistemi ve Perakendecilik Sektöründe Uygulaması. *Yüksek Lisans Tezi.* İstanbul: Yıldız Teknik Üniversitesi. FBE.

Zaiane, O. (1999). Introduction To Data Mining. *University of Alberta. Department of Computing Science. https://webdocs.cs.ualberta.ca/~zaiane/courses/cmput690/notes/Chapter1/index.html (9 March 2020).*

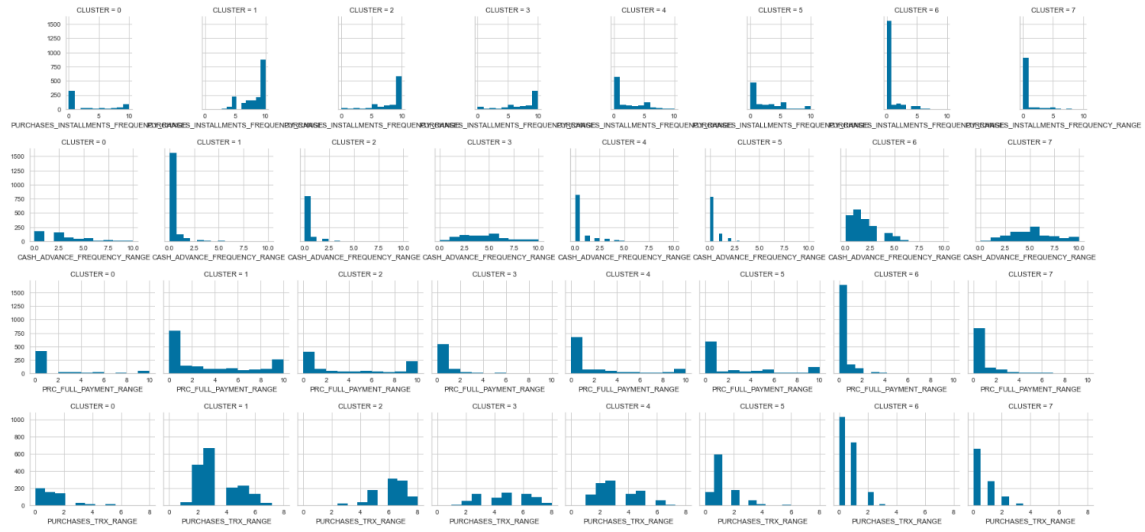
Wang, J ve X. Su. (Juen 2011). An Improved K-MeansClustering Algorithm. ICCSN International Conference *Communication Software and Networks.* Xi'an : IEEE

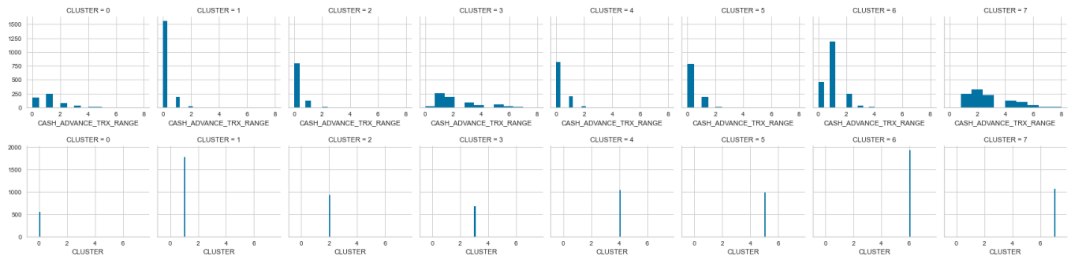
EKLER

EK 1 : K-MeansAlgoritması Sonucu Elde Edilen Bulgular




CLUSTER = 0
CLUSTER = 1
CLUSTER = 2
CLUSTER = 3
CLUSTER = 4
CLUSTER = 5
CLUSTER = 6
CLUSTER = 7
MINIMUM_PAYMENTS_RANGE
BALANCE_FREQUENCY_RANGE
PURCHASES_FREQUENCY_RANGE
ONEOFF_PURCHASES_FREQUENCY_RANGE


CLUSTER = 0
CLUSTER = 1
CLUSTER = 2
CLUSTER = 3
CLUSTER = 4
CLUSTER = 5
CLUSTER = 6
CLUSTER = 7
PURCHASES_INSTALLMENTS_FREQUENCY_RANGE
CASH_ADVANCE_FREQUENCY_RANGE
PRC_FULL_PAYMENT_RANGE
PURCHASES_TRX_RANGE


CLUSTER = 0
CLUSTER = 1
CLUSTER = 2
CLUSTER = 3
CLUSTER = 4
CLUSTER = 5
CLUSTER = 6
CLUSTER = 7
CASH_ADVANCE_TRX_RANGE
CLUSTER

EK 3 : Tamhane's T^2 (İkili Karşılaştırma) Sınaması Sonuçları

Dependent Variable	(I) Clust	(J) Clust	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
XX2	,00	1,00	3,51143*	,03424	,000	3,4043	3,6185
		2,00	,04296	,02198	,767	-,0256	,1115
		3,00	-,09873*	,01739	,000	-,1530	-,0445
		4,00	,01387	,02776	1,000	-,0729	,1006
		5,00	-,01053	,02020	1,000	-,0736	,0525
		6,00	,10280*	,02470	,001	,0257	,1799
		7,00	-,06273*	,01433	,000	-,1074	-,0180
	1,00	,00	-3,51143*	,03424	,000	-3,6185	-3,4043
		2,00	-3,46847*	,03665	,000	-3,5830	-3,3539
		3,00	-3,61016*	,03409	,000	-3,7168	-3,5035
		4,00	-3,49756*	,04038	,000	-3,6237	-3,3714
		5,00	-3,52196*	,03561	,000	-3,6333	-3,4106
		6,00	-3,40863*	,03834	,000	-3,5284	-3,2888
		7,00	-3,57416*	,03264	,000	-3,6763	-3,4720
	2,00	,00	-,04296	,02198	,767	-,1115	,0256
		1,00	3,46847*	,03665	,000	3,3539	3,5830
		3,00	-,14169*	,02175	,000	-,2096	-,0738
		4,00	-,02909	,03067	1,000	-,1249	,0667
		5,00	-,05349	,02405	,525	-,1286	,0216
		6,00	,05983	,02794	,602	-,0273	,1470
		7,00	-,10569*	,01938	,000	-,1662	-,0452
	3,00	,00	,09873*	,01739	,000	,0445	,1530
		1,00	3,61016*	,03409	,000	3,5035	3,7168

		2,00	,14169*	,02175	,000	,0738	,2096
		4,00	,11259*	,02758	,001	,0264	,1988
		5,00	,08820*	,01995	,000	,0259	,1505
		6,00	,20152*	,02450	,000	,1251	,2780
		7,00	,03600	,01398	,247	-,0076	,0796
	4,00	,00	-,01387	,02776	1,000	-,1006	,0729
		1,00	3,49756*	,04038	,000	3,3714	3,6237
		2,00	,02909	,03067	1,000	-,0667	,1249
		3,00	-,11259*	,02758	,001	-,1988	-,0264
		5,00	-,02440	,02943	1,000	-,1163	,0675
		6,00	,08893	,03268	,169	-,0131	,1910
		7,00	-,07660	,02575	,081	-,1571	,0039
	5,00	,00	,01053	,02020	1,000	-,0525	,0736
		1,00	3,52196*	,03561	,000	3,4106	3,6333
		2,00	,05349	,02405	,525	-,0216	,1286
		3,00	-,08820*	,01995	,000	-,1505	-,0259
		4,00	,02440	,02943	1,000	-,0675	,1163
		6,00	,11333*	,02656	,001	,0304	,1962
		7,00	-,05220	,01735	,072	-,1064	,0020
	6,00	,00	-,10280*	,02470	,001	-,1799	-,0257
		1,00	3,40863*	,03834	,000	3,2888	3,5284
		2,00	-,05983	,02794	,602	-,1470	,0273
		3,00	-,20152*	,02450	,000	-,2780	-,1251
		4,00	-,08893	,03268	,169	-,1910	,0131
		5,00	-,11333*	,02656	,001	-,1962	-,0304
		7,00	-,16553*	,02243	,000	-,2356	-,0955

	7,00	,00	,06273*	,01433	,000	,0180	,1074
		1,00	3,57416*	,03264	,000	3,4720	3,6763
		2,00	,10569*	,01938	,000	,0452	,1662
		3,00	-,03600	,01398	,247	-,0796	,0076
		4,00	,07660	,02575	,081	-,0039	,1571
		5,00	,05220	,01735	,072	-,0020	,1064
		6,00	,16553*	,02243	,000	,0955	,2356
XX3	,00	1,00	-,25246*	,03457	,000	-,3605	-,1444
		2,00	,43319*	,01643	,000	,3819	,4844
		3,00	-,65923*	,03587	,000	-,7713	-,5472
		4,00	-1,83450*	,03339	,000	-1,9388	-1,7302
		5,00	-,44133*	,02963	,000	-,5338	-,3489
		6,00	-1,79832*	,02827	,000	-1,8865	-1,7101
		7,00	-,70652*	,02118	,000	-,7726	-,6405
	1,00	,00	,25246*	,03457	,000	,1444	,3605
		2,00	,68565*	,03208	,000	,5852	,7861
		3,00	-,40678*	,04523	,000	-,5480	-,2655
		4,00	-1,58204*	,04329	,000	-1,7172	-1,4468
		5,00	-,18887*	,04046	,000	-,3152	-,0625
		6,00	-1,54586*	,03947	,000	-1,6692	-1,4226
		7,00	-,45406*	,03475	,000	-,5627	-,3454
	2,00	,00	-,43319*	,01643	,000	-,4844	-,3819
		1,00	-,68565*	,03208	,000	-,7861	-,5852
		3,00	-1,09242*	,03348	,000	-1,1970	-,9878
		4,00	-2,26769*	,03080	,000	-2,3640	-2,1714
		5,00	-,87452*	,02668	,000	-,9579	-,7912

		6,00	-2,23151*	,02516	,000	-2,3101	-2,1529
		7,00	-1,13971*	,01681	,000	-1,1922	-1,0873
	3,00	,00	,65923*	,03587	,000	,5472	,7713
		1,00	,40678*	,04523	,000	,2655	,5480
		2,00	1,09242*	,03348	,000	,9878	1,1970
		4,00	-1,17527*	,04434	,000	-1,3137	-1,0369
		5,00	,21790*	,04158	,000	,0881	,3477
		6,00	-1,13908*	,04062	,000	-1,2659	-1,0123
		7,00	-,04729	,03605	,997	-,1599	,0653
	4,00	,00	1,83450*	,03339	,000	1,7302	1,9388
		1,00	1,58204*	,04329	,000	1,4468	1,7172
		2,00	2,26769*	,03080	,000	2,1714	2,3640
		3,00	1,17527*	,04434	,000	1,0369	1,3137
		5,00	1,39317*	,03945	,000	1,2700	1,5163
		6,00	,03618	,03844	1,000	-,0838	,1562
		7,00	1,12798*	,03358	,000	1,0231	1,2329
	5,00	,00	,44133*	,02963	,000	,3489	,5338
		1,00	,18887*	,04046	,000	,0625	,3152
		2,00	,87452*	,02668	,000	,7912	,9579
		3,00	-,21790*	,04158	,000	-,3477	-,0881
		4,00	-1,39317*	,03945	,000	-1,5163	-1,2700
		6,00	-1,35699*	,03523	,000	-1,4669	-1,2471
		7,00	-,26519*	,02984	,000	-,3583	-,1720
	6,00	,00	1,79832*	,02827	,000	1,7101	1,8865
		1,00	1,54586*	,03947	,000	1,4226	1,6692
		2,00	2,23151*	,02516	,000	2,1529	2,3101

			3,00	1,13908*	,04062	,000	1,0123	1,2659
			4,00	-,03618	,03844	1,000	-,1562	,0838
			5,00	1,35699*	,03523	,000	1,2471	1,4669
			7,00	1,09180*	,02849	,000	1,0029	1,1807
		7,00	,00	,70652*	,02118	,000	,6405	,7726
			1,00	,45406*	,03475	,000	,3454	,5627
			2,00	1,13971*	,01681	,000	1,0873	1,1922
			3,00	,04729	,03605	,997	-,0653	,1599
			4,00	-1,12798*	,03358	,000	-1,2329	-1,0231
			5,00	,26519*	,02984	,000	,1720	,3583
			6,00	-1,09180*	,02849	,000	-1,1807	-1,0029
	XX4	,00	1,00	,65217*	,03069	,000	,5562	,7481
			2,00	,53793*	,02135	,000	,4713	,6046
			3,00	-1,53943*	,02752	,000	-1,6254	-1,4535
			4,00	-,79226*	,03220	,000	-,8929	-,6916
			5,00	-,55464*	,02380	,000	-,6289	-,4804
			6,00	,96990*	,02131	,000	,9034	1,0364
			7,00	,94002*	,01729	,000	,8861	,9939
		1,00	,00	-,65217*	,03069	,000	-,7481	-,5562
			2,00	-,11424*	,03145	,008	-,2125	-,0159
			3,00	-2,19160*	,03593	,000	-2,3038	-2,0794
			4,00	-1,44443*	,03962	,000	-1,5682	-1,3207
			5,00	-1,20681*	,03316	,000	-1,3104	-1,1032
			6,00	,31773*	,03143	,000	,2195	,4159
			7,00	,28785*	,02885	,000	,1976	,3781
		2,00	,00	-,53793*	,02135	,000	-,6046	-,4713

		1,00	,11424*	,03145	,008	,0159	,2125
		3,00	-2,07736*	,02837	,000	-2,1659	-1,9888
		4,00	-1,33019*	,03292	,000	-1,4330	-1,2273
		5,00	-1,09257*	,02477	,000	-1,1699	-1,0153
		6,00	,43197*	,02239	,000	,3621	,5018
		7,00	,40209*	,01860	,000	,3440	,4601
	3,00	,00	1,53943*	,02752	,000	1,4535	1,6254
		1,00	2,19160*	,03593	,000	2,0794	2,3038
		2,00	2,07736*	,02837	,000	1,9888	2,1659
		4,00	,74717*	,03722	,000	,6309	,8634
		5,00	,98479*	,03025	,000	,8904	1,0792
		6,00	2,50933*	,02834	,000	2,4209	2,5978
		7,00	2,47945*	,02545	,000	2,4000	2,5589
	4,00	,00	,79226*	,03220	,000	,6916	,8929
		1,00	1,44443*	,03962	,000	1,3207	1,5682
		2,00	1,33019*	,03292	,000	1,2273	1,4330
		3,00	-,74717*	,03722	,000	-,8634	-,6309
		5,00	,23762*	,03456	,000	,1297	,3456
		6,00	1,76216*	,03290	,000	1,6594	1,8649
		7,00	1,73228*	,03044	,000	1,6371	1,8275
	5,00	,00	,55464*	,02380	,000	,4804	,6289
		1,00	1,20681*	,03316	,000	1,1032	1,3104
		2,00	1,09257*	,02477	,000	1,0153	1,1699
		3,00	-,98479*	,03025	,000	-1,0792	-,8904
		4,00	-,23762*	,03456	,000	-,3456	-,1297
		6,00	1,52454*	,02474	,000	1,4473	1,6017

		7,00	1,49466*	,02137	,000	1,4280	1,5614
	6,00	,00	-,96990*	,02131	,000	-1,0364	-,9034
		1,00	-,31773*	,03143	,000	-,4159	-,2195
		2,00	-,43197*	,02239	,000	-,5018	-,3621
		3,00	-2,50933*	,02834	,000	-2,5978	-2,4209
		4,00	-1,76216*	,03290	,000	-1,8649	-1,6594
		5,00	-1,52454*	,02474	,000	-1,6017	-1,4473
		7,00	-,02988	,01855	,959	-,0878	,0280
	7,00	,00	-,94002*	,01729	,000	-,9939	-,8861
		1,00	-,28785*	,02885	,000	-,3781	-,1976
		2,00	-,40209*	,01860	,000	-,4601	-,3440
		3,00	-2,47945*	,02545	,000	-2,5589	-2,4000
		4,00	-1,73228*	,03044	,000	-1,8275	-1,6371
		5,00	-1,49466*	,02137	,000	-1,5614	-1,4280
		6,00	,02988	,01855	,959	-,0280	,0878
	XX5	,00	1,00	-,12743*	,02877	,000	-,2174	-,0374
			2,00	-,18030*	,02180	,000	-,2483	-,1122
			3,00	-2,15227*	,03263	,000	-2,2542	-2,0503
			4,00	-1,17248*	,04141	,000	-1,3020	-1,0430
			5,00	-1,63397*	,02352	,000	-1,7074	-1,5605
			6,00	-,03355	,01899	,895	-,0928	,0257
			7,00	-,03974	,01416	,132	-,0839	,0044
		1,00	,00	,12743*	,02877	,000	,0374	,2174
			2,00	-,05287	,03309	,962	-,1562	,0505
			3,00	-2,02484*	,04104	,000	-2,1530	-1,8967
			4,00	-1,04506*	,04832	,000	-1,1960	-,8941

		5,00	-1,50654*	,03425	,000	-1,6135	-1,3996
		6,00	,09388	,03131	,075	-,0040	,1917
		7,00	,08769	,02864	,062	-,0019	,1773
	2,00	,00	,18030*	,02180	,000	,1122	,2483
		1,00	,05287	,03309	,962	-,0505	,1562
		3,00	-1,97198*	,03650	,000	-2,0859	-1,8580
		4,00	-,99219*	,04452	,000	-1,1313	-,8531
		5,00	-1,45367*	,02865	,000	-1,5431	-1,3643
		6,00	,14674*	,02506	,000	,0685	,2249
		7,00	,14056*	,02163	,000	,0730	,2081
	3,00	,00	2,15227*	,03263	,000	2,0503	2,2542
		1,00	2,02484*	,04104	,000	1,8967	2,1530
		2,00	1,97198*	,03650	,000	1,8580	2,0859
		4,00	,97979*	,05071	,000	,8214	1,1381
		5,00	,51831*	,03755	,000	,4011	,6355
		6,00	2,11872*	,03489	,000	2,0098	2,2277
		7,00	2,11253*	,03252	,000	2,0109	2,2141
	4,00	,00	1,17248*	,04141	,000	1,0430	1,3020
		1,00	1,04506*	,04832	,000	,8941	1,1960
		2,00	,99219*	,04452	,000	,8531	1,1313
		3,00	-,97979*	,05071	,000	-1,1381	-,8214
		5,00	-,46148*	,04539	,000	-,6033	-,3197
		6,00	1,13893*	,04321	,000	1,0038	1,2740
		7,00	1,13274*	,04132	,000	1,0035	1,2620
	5,00	,00	1,63397*	,02352	,000	1,5605	1,7074
		1,00	1,50654*	,03425	,000	1,3996	1,6135

		2,00	1,45367*	,02865	,000	1,3643	1,5431
		3,00	-,51831*	,03755	,000	-,6355	-,4011
		4,00	,46148*	,04539	,000	,3197	,6033
		6,00	1,60041*	,02657	,000	1,5175	1,6833
		7,00	1,59423*	,02337	,000	1,5213	1,6672
	6,00	,00	,03355	,01899	,895	-,0257	,0928
		1,00	-,09388	,03131	,075	-,1917	,0040
		2,00	-,14674*	,02506	,000	-,2249	-,0685
		3,00	-2,11872*	,03489	,000	-2,2277	-2,0098
		4,00	-1,13893*	,04321	,000	-1,2740	-1,0038
		5,00	-1,60041*	,02657	,000	-1,6833	-1,5175
		7,00	-,00619	,01880	1,000	-,0648	,0525
	7,00	,00	,03974	,01416	,132	-,0044	,0839
		1,00	-,08769	,02864	,062	-,1773	,0019
		2,00	-,14056*	,02163	,000	-,2081	-,0730
		3,00	-2,11253*	,03252	,000	-2,2141	-2,0109
		4,00	-1,13274*	,04132	,000	-1,2620	-1,0035
		5,00	-1,59423*	,02337	,000	-1,6672	-1,5213
		6,00	,00619	,01880	1,000	-,0525	,0648
XX6	,00	1,00	1,08488*	,03107	,000	,9878	1,1820
		2,00	1,00464*	,02482	,000	,9272	1,0821
		3,00	-,81955*	,03729	,000	-,9360	-,7031
		4,00	-,22937*	,04261	,000	-,3625	-,0962
		5,00	1,07235*	,02423	,000	,9967	1,1480
		6,00	1,39829*	,02018	,000	1,3353	1,4612
		7,00	1,35201*	,01929	,000	1,2918	1,4122

	1,00	,00	-1,08488*	,03107	,000	-1,1820	-,9878
		2,00	-,08023	,03133	,257	-,1781	,0176
		3,00	-1,90443*	,04190	,000	-2,0353	-1,7736
		4,00	-1,31425*	,04670	,000	-1,4601	-1,1684
		5,00	-,01253	,03087	1,000	-,1090	,0839
		6,00	,31341*	,02780	,000	,2265	,4004
		7,00	,26713*	,02715	,000	,1822	,3521
	2,00	,00	-1,00464*	,02482	,000	-1,0821	-,9272
		1,00	,08023	,03133	,257	-,0176	,1781
		3,00	-1,82420*	,03750	,000	-1,9413	-1,7071
		4,00	-1,23401*	,04280	,000	-1,3678	-1,1003
		5,00	,06771	,02456	,152	-,0089	,1444
		6,00	,39364*	,02057	,000	,3294	,4579
		7,00	,34736*	,01969	,000	,2859	,4089
	3,00	,00	,81955*	,03729	,000	,7031	,9360
		1,00	1,90443*	,04190	,000	1,7736	2,0353
		2,00	1,82420*	,03750	,000	1,7071	1,9413
		4,00	,59019*	,05105	,000	,4308	,7496
		5,00	1,89191*	,03712	,000	1,7760	2,0078
		6,00	2,21784*	,03461	,000	2,1097	2,3260
		7,00	2,17156*	,03409	,000	2,0650	2,2781
	4,00	,00	,22937*	,04261	,000	,0962	,3625
		1,00	1,31425*	,04670	,000	1,1684	1,4601
		2,00	1,23401*	,04280	,000	1,1003	1,3678
		3,00	-,59019*	,05105	,000	-,7496	-,4308
		5,00	1,30172*	,04246	,000	1,1690	1,4344

		6,00	1,62765*	,04028	,000	1,5017	1,7536
		7,00	1,58137*	,03984	,000	1,4568	1,7060
	5,00	,00	-1,07235*	,02423	,000	-1,1480	-,9967
		1,00	,01253	,03087	1,000	-,0839	,1090
		2,00	-,06771	,02456	,152	-,1444	,0089
		3,00	-1,89191*	,03712	,000	-2,0078	-1,7760
		4,00	-1,30172*	,04246	,000	-1,4344	-1,1690
		6,00	,32593*	,01986	,000	,2639	,3879
		7,00	,27966*	,01895	,000	,2205	,3388
	6,00	,00	-1,39829*	,02018	,000	-1,4612	-1,3353
		1,00	-,31341*	,02780	,000	-,4004	-,2265
		2,00	-,39364*	,02057	,000	-,4579	-,3294
		3,00	-2,21784*	,03461	,000	-2,3260	-2,1097
		4,00	-1,62765*	,04028	,000	-1,7536	-1,5017
		5,00	-,32593*	,01986	,000	-,3879	-,2639
		7,00	-,04628*	,01338	,015	-,0880	-,0045
	7,00	,00	-1,35201*	,01929	,000	-1,4122	-1,2918
		1,00	-,26713*	,02715	,000	-,3521	-,1822
		2,00	-,34736*	,01969	,000	-,4089	-,2859
		3,00	-2,17156*	,03409	,000	-2,2781	-2,0650
		4,00	-1,58137*	,03984	,000	-1,7060	-1,4568
		5,00	-,27966*	,01895	,000	-,3388	-,2205
		6,00	,04628*	,01338	,015	,0045	,0880
XX7	,00	1,00	-,96791*	,04107	,000	-1,0965	-,8393
		2,00	-,14872*	,02080	,000	-,2137	-,0838
		3,00	-,03571	,01611	,532	-,0860	,0146

		4,00	-1,90511*	,03295	,000	-2,0082	-1,8020
		5,00	-,10437*	,01719	,000	-,1580	-,0507
		6,00	-2,22940*	,02321	,000	-2,3019	-2,1569
		7,00	-,78282*	,01838	,000	-,8402	-,7255
	1,00	,00	,96791*	,04107	,000	,8393	1,0965
		2,00	,81919*	,04449	,000	,6801	,9583
		3,00	,93220*	,04250	,000	,7992	1,0652
		4,00	-,93720*	,05131	,000	-1,0975	-,7769
		5,00	,86354*	,04292	,000	,7293	,9978
		6,00	-1,26149*	,04567	,000	-1,4042	-1,1187
		7,00	,18508*	,04341	,001	,0493	,3209
	2,00	,00	,14872*	,02080	,000	,0838	,2137
		1,00	-,81919*	,04449	,000	-,9583	-,6801
		3,00	,11301*	,02350	,000	,0397	,1863
		4,00	-1,75639*	,03713	,000	-1,8724	-1,6404
		5,00	,04434	,02425	,859	-,0313	,1200
		6,00	-2,08068*	,02883	,000	-2,1707	-1,9907
		7,00	-,63411*	,02511	,000	-,7125	-,5558
	3,00	,00	,03571	,01611	,532	-,0146	,0860
		1,00	-,93220*	,04250	,000	-1,0652	-,7992
		2,00	-,11301*	,02350	,000	-,1863	-,0397
		4,00	-1,86940*	,03472	,000	-1,9779	-1,7609
		5,00	-,06866*	,02037	,021	-,1322	-,0051
		6,00	-2,19369*	,02566	,000	-2,2738	-2,1136
		7,00	-,74711*	,02139	,000	-,8138	-,6804
	4,00	,00	1,90511*	,03295	,000	1,8020	2,0082

		1,00	,93720*	,05131	,000	,7769	1,0975
		2,00	1,75639*	,03713	,000	1,6404	1,8724
		3,00	1,86940*	,03472	,000	1,7609	1,9779
		5,00	1,80074*	,03523	,000	1,6906	1,9108
		6,00	-,32429*	,03853	,000	-,4446	-,2040
		7,00	1,12228*	,03583	,000	1,0103	1,2342
	5,00	,00	,10437*	,01719	,000	,0507	,1580
		1,00	-,86354*	,04292	,000	-,9978	-,7293
		2,00	-,04434	,02425	,859	-,1200	,0313
		3,00	,06866*	,02037	,021	,0051	,1322
		4,00	-1,80074*	,03523	,000	-1,9108	-1,6906
		6,00	-2,12503*	,02635	,000	-2,2073	-2,0428
		7,00	-,67845*	,02221	,000	-,7477	-,6092
	6,00	,00	2,22940*	,02321	,000	2,1569	2,3019
		1,00	1,26149*	,04567	,000	1,1187	1,4042
		2,00	2,08068*	,02883	,000	1,9907	2,1707
		3,00	2,19369*	,02566	,000	2,1136	2,2738
		4,00	,32429*	,03853	,000	,2040	,4446
		5,00	2,12503*	,02635	,000	2,0428	2,2073
		7,00	1,44657*	,02714	,000	1,3619	1,5313
	7,00	,00	,78282*	,01838	,000	,7255	,8402
		1,00	-,18508*	,04341	,001	-,3209	-,0493
		2,00	,63411*	,02511	,000	,5558	,7125
		3,00	,74711*	,02139	,000	,6804	,8138
		4,00	-1,12228*	,03583	,000	-1,2342	-1,0103
		5,00	,67845*	,02221	,000	,6092	,7477

			6,00	-1,44657*	,02714	,000	-1,5313	-1,3619
	XX8	,00	1,00	,28562*	,03869	,000	,1647	,4065
			2,00	-,19255*	,03600	,000	-,3049	-,0802
			3,00	-1,23378*	,03407	,000	-1,3401	-1,1275
			4,00	-1,13595*	,03916	,000	-1,2583	-1,0136
			5,00	-,57553*	,03551	,000	-,6863	-,4647
			6,00	-1,13482*	,03202	,000	-1,2347	-1,0349
			7,00	,04452	,02673	,941	-,0389	,1279
		1,00	,00	-,28562*	,03869	,000	-,4065	-,1647
			2,00	-,47817*	,04437	,000	-,6167	-,3396
			3,00	-1,51940*	,04281	,000	-1,6531	-1,3857
			4,00	-1,42156*	,04696	,000	-1,5682	-1,2749
			5,00	-,86115*	,04397	,000	-,9984	-,7239
			6,00	-1,42043*	,04120	,000	-1,5491	-1,2917
			7,00	-,24109*	,03724	,000	-,3575	-,1247
		2,00	,00	,19255*	,03600	,000	,0802	,3049
			1,00	,47817*	,04437	,000	,3396	,6167
			3,00	-1,04123*	,04040	,000	-1,1673	-,9152
			4,00	-,94339*	,04477	,000	-1,0832	-,8036
			5,00	-,38298*	,04162	,000	-,5129	-,2531
			6,00	-,94227*	,03869	,000	-1,0630	-,8215
			7,00	,23708*	,03444	,000	,1296	,3446
		3,00	,00	1,23378*	,03407	,000	1,1275	1,3401
			1,00	1,51940*	,04281	,000	1,3857	1,6531
			2,00	1,04123*	,04040	,000	,9152	1,1673
			4,00	,09784	,04323	,490	-,0372	,2328

			2,00	-,23708*	,03444	,000	-,3446	-,1296
			3,00	-1,27831*	,03241	,000	-1,3795	-1,1771
			4,00	-1,18047*	,03772	,000	-1,2983	-1,0626
			5,00	-,62006*	,03392	,000	-,7259	-,5142
			6,00	-1,17934*	,03025	,000	-1,2737	-1,0849
	XX9	,00	1,00	,44354*	,03322	,000	,3397	,5474
			2,00	,23655*	,03419	,000	,1298	,3433
			3,00	-1,36135*	,03154	,000	-1,4598	-1,2629
			4,00	-1,21813*	,03855	,000	-1,3386	-1,0977
			5,00	-,42950*	,02902	,000	-,5201	-,3389
			6,00	-1,05523*	,03297	,000	-1,1581	-,9523
			7,00	,09017*	,02303	,003	,0183	,1620
		1,00	,00	-,44354*	,03322	,000	-,5474	-,3397
			2,00	-,20699*	,04173	,000	-,3373	-,0767
			3,00	-1,80489*	,03959	,000	-1,9285	-1,6812
			4,00	-1,66167*	,04538	,000	-1,8034	-1,5199
			5,00	-,87304*	,03761	,000	-,9905	-,7556
			6,00	-1,49877*	,04074	,000	-1,6260	-1,3716
			7,00	-,35337*	,03321	,000	-,4572	-,2496
		2,00	,00	-,23655*	,03419	,000	-,3433	-,1298
			1,00	,20699*	,04173	,000	,0767	,3373
			3,00	-1,59790*	,04041	,000	-1,7240	-1,4718
			4,00	-1,45468*	,04609	,000	-1,5986	-1,3108
			5,00	-,66604*	,03847	,000	-,7861	-,5460
			6,00	-1,29178*	,04153	,000	-1,4214	-1,1622
			7,00	-,14638*	,03418	,001	-,2531	-,0397

	3,00	,00	1,36135*	,03154	,000	1,2629	1,4598
		1,00	1,80489*	,03959	,000	1,6812	1,9285
		2,00	1,59790*	,04041	,000	1,4718	1,7240
		4,00	,14322*	,04416	,033	,0053	,2811
		5,00	,93186*	,03614	,000	,8191	1,0447
		6,00	,30612*	,03938	,000	,1832	,4290
		7,00	1,45152*	,03154	,000	1,3531	1,5500
	4,00	,00	1,21813*	,03855	,000	1,0977	1,3386
		1,00	1,66167*	,04538	,000	1,5199	1,8034
		2,00	1,45468*	,04609	,000	1,3108	1,5986
		3,00	-,14322*	,04416	,033	-,2811	-,0053
		5,00	,78864*	,04240	,000	,6562	,9210
		6,00	,16290*	,04519	,009	,0218	,3040
		7,00	1,30830*	,03855	,000	1,1878	1,4288
	5,00	,00	,42950*	,02902	,000	,3389	,5201
		1,00	,87304*	,03761	,000	,7556	,9905
		2,00	,66604*	,03847	,000	,5460	,7861
		3,00	-,93186*	,03614	,000	-1,0447	-,8191
		4,00	-,78864*	,04240	,000	-,9210	-,6562
		6,00	-,62573*	,03739	,000	-,7424	-,5090
		7,00	,51967*	,02901	,000	,4291	,6102
	6,00	,00	1,05523*	,03297	,000	,9523	1,1581
		1,00	1,49877*	,04074	,000	1,3716	1,6260
		2,00	1,29178*	,04153	,000	1,1622	1,4214
		3,00	-,30612*	,03938	,000	-,4290	-,1832
		4,00	-,16290*	,04519	,009	-,3040	-,0218

		5,00	,62573*	,03739	,000	,5090	,7424
		7,00	1,14540*	,03296	,000	1,0425	1,2483
	7,00	,00	-,09017*	,02303	,003	-,1620	-,0183
		1,00	,35337*	,03321	,000	,2496	,4572
		2,00	,14638*	,03418	,001	,0397	,2531
		3,00	-1,45152*	,03154	,000	-1,5500	-1,3531
		4,00	-1,30830*	,03855	,000	-1,4288	-1,1878
		5,00	-,51967*	,02901	,000	-,6102	-,4291
		6,00	-1,14540*	,03296	,000	-1,2483	-1,0425
	XX10 ,00	1,00	,07597	,03216	,404	-,0245	,1764
		2,00	,14037*	,02441	,000	,0642	,2165
		3,00	-,21371*	,03623	,000	-,3268	-,1006
		4,00	-1,57372*	,04781	,000	-1,7231	-1,4243
		5,00	-,08635	,03056	,125	-,1817	,0090
		6,00	-1,27028*	,03602	,000	-1,3827	-1,1579
		7,00	-,17940*	,02804	,000	-,2669	-,0920
	1,00	,00	-,07597	,03216	,404	-,1764	,0245
		2,00	,06440	,02899	,529	-,0262	,1550
		3,00	-,28968*	,03946	,000	-,4129	-,1665
		4,00	-1,64969*	,05030	,000	-1,8069	-1,4925
		5,00	-,16232*	,03432	,000	-,2695	-,0551
		6,00	-1,34625*	,03926	,000	-1,4688	-1,2237
		7,00	-,25538*	,03210	,000	-,3556	-,1551
	2,00	,00	-,14037*	,02441	,000	-,2165	-,0642
		1,00	-,06440	,02899	,529	-,1550	,0262
		3,00	-,35407*	,03345	,000	-,4585	-,2496

		4,00	-1,71409*	,04574	,000	-1,8571	-1,5711
		5,00	-,22671*	,02720	,000	-,3116	-,1418
		6,00	-1,41065*	,03322	,000	-1,5144	-1,3069
		7,00	-,31977*	,02434	,000	-,3957	-,2439
	3,00	,00	,21371*	,03623	,000	,1006	,3268
		1,00	,28968*	,03946	,000	,1665	,4129
		2,00	,35407*	,03345	,000	,2496	,4585
		4,00	-1,36001*	,05300	,000	-1,5255	-1,1945
		5,00	,12736*	,03816	,024	,0082	,2465
		6,00	-1,05657*	,04266	,000	-1,1897	-,9234
		7,00	,03430	,03618	1,000	-,0786	,1472
	4,00	,00	1,57372*	,04781	,000	1,4243	1,7231
		1,00	1,64969*	,05030	,000	1,4925	1,8069
		2,00	1,71409*	,04574	,000	1,5711	1,8571
		3,00	1,36001*	,05300	,000	1,1945	1,5255
		5,00	1,48738*	,04930	,000	1,3334	1,6414
		6,00	,30344*	,05285	,000	,1384	,4685
		7,00	1,39432*	,04777	,000	1,2450	1,5436
	5,00	,00	,08635	,03056	,125	-,0090	,1817
		1,00	,16232*	,03432	,000	,0551	,2695
		2,00	,22671*	,02720	,000	,1418	,3116
		3,00	-,12736*	,03816	,024	-,2465	-,0082
		4,00	-1,48738*	,04930	,000	-1,6414	-1,3334
		6,00	-1,18393*	,03796	,000	-1,3024	-1,0655
		7,00	-,09306	,03050	,063	-,1882	,0021
	6,00	,00	1,27028*	,03602	,000	1,1579	1,3827

		1,00	1,34625*	,03926	,000	1,2237	1,4688
		2,00	1,41065*	,03322	,000	1,3069	1,5144
		3,00	1,05657*	,04266	,000	,9234	1,1897
		4,00	-,30344*	,05285	,000	-,4685	-,1384
		5,00	1,18393*	,03796	,000	1,0655	1,3024
		7,00	1,09087*	,03597	,000	,9786	1,2031
	7,00	,00	,17940*	,02804	,000	,0920	,2669
		1,00	,25538*	,03210	,000	,1551	,3556
		2,00	,31977*	,02434	,000	,2439	,3957
		3,00	-,03430	,03618	1,000	-,1472	,0786
		4,00	-1,39432*	,04777	,000	-1,5436	-1,2450
		5,00	,09306	,03050	,063	-,0021	,1882
		6,00	-1,09087*	,03597	,000	-1,2031	-,9786
XX11	,00	1,00	,62476*	,04502	,000	,4839	,7657
		2,00	2,62227*	,03056	,000	2,5268	2,7177
		3,00	-,14144*	,01436	,000	-,1862	-,0966
		4,00	-,14758*	,01414	,000	-,1917	-,1035
		5,00	-,00338	,01894	1,000	-,0625	,0557
		6,00	-,06157*	,01735	,011	-,1157	-,0075
		7,00	-,08201*	,01408	,000	-,1259	-,0381
	1,00	,00	-,62476*	,04502	,000	-,7657	-,4839
		2,00	1,99752*	,05199	,000	1,8351	2,1600
		3,00	-,76620*	,04444	,000	-,9053	-,6271
		4,00	-,77234*	,04437	,000	-,9112	-,6334
		5,00	-,62814*	,04612	,000	-,7724	-,4838
		6,00	-,68633*	,04549	,000	-,8287	-,5440

		7,00	-,70677*	,04435	,000	-,8456	-,5679
	2,00	,00	-2,62227*	,03056	,000	-2,7177	-2,5268
		1,00	-1,99752*	,05199	,000	-2,1600	-1,8351
		3,00	-2,76371*	,02970	,000	-2,8565	-2,6709
		4,00	-2,76985*	,02959	,000	-2,8623	-2,6774
		5,00	-2,62566*	,03217	,000	-2,7261	-2,5252
		6,00	-2,68385*	,03125	,000	-2,7814	-2,5863
		7,00	-2,70428*	,02956	,000	-2,7966	-2,6119
	3,00	,00	,14144*	,01436	,000	,0966	,1862
		1,00	,76620*	,04444	,000	,6271	,9053
		2,00	2,76371*	,02970	,000	2,6709	2,8565
		4,00	-,00614	,01216	1,000	-,0441	,0318
		5,00	,13806*	,01752	,000	,0834	,1927
		6,00	,07987*	,01577	,000	,0306	,1291
		7,00	,05943*	,01209	,000	,0217	,0971
	4,00	,00	,14758*	,01414	,000	,1035	,1917
		1,00	,77234*	,04437	,000	,6334	,9112
		2,00	2,76985*	,02959	,000	2,6774	2,8623
		3,00	,00614	,01216	1,000	-,0318	,0441
		5,00	,14419*	,01734	,000	,0901	,1983
		6,00	,08601*	,01558	,000	,0374	,1346
		7,00	,06557*	,01183	,000	,0287	,1025
	5,00	,00	,00338	,01894	1,000	-,0557	,0625
		1,00	,62814*	,04612	,000	,4838	,7724
		2,00	2,62566*	,03217	,000	2,5252	2,7261
		3,00	-,13806*	,01752	,000	-,1927	-,0834

		4,00	-,14419*	,01734	,000	-,1983	-,0901
		6,00	-,05819	,02004	,099	-,1207	,0043
		7,00	-,07863*	,01729	,000	-,1326	-,0247
	6,00	,00	,06157*	,01735	,011	,0075	,1157
		1,00	,68633*	,04549	,000	,5440	,8287
		2,00	2,68385*	,03125	,000	2,5863	2,7814
		3,00	-,07987*	,01577	,000	-,1291	-,0306
		4,00	-,08601*	,01558	,000	-,1346	-,0374
		5,00	,05819	,02004	,099	-,0043	,1207
		7,00	-,02044	,01552	,997	-,0689	,0280
	7,00	,00	,08201*	,01408	,000	,0381	,1259
		1,00	,70677*	,04435	,000	,5679	,8456
		2,00	2,70428*	,02956	,000	2,6119	2,7966
		3,00	-,05943*	,01209	,000	-,0971	-,0217
		4,00	-,06557*	,01183	,000	-,1025	-,0287
		5,00	,07863*	,01729	,000	,0247	,1326
		6,00	,02044	,01552	,997	-,0280	,0689
XX12	,00	1,00	1,14601*	,04299	,000	1,0115	1,2806
		2,00	1,46233*	,02227	,000	1,3928	1,5318
		3,00	-,24573*	,01231	,000	-,2841	-,2073
		4,00	-,00199	,02020	1,000	-,0651	,0611
		5,00	,47806*	,02245	,000	,4080	,5481
		6,00	1,88189*	,01749	,000	1,8273	1,9365
		7,00	1,86343*	,01394	,000	1,8199	1,9069
	1,00	,00	-1,14601*	,04299	,000	-1,2806	-1,0115
		2,00	,31632*	,04614	,000	,1720	,4606

		3,00	-1,39174*	,04225	,000	-1,5240	-1,2595
		4,00	-1,14800*	,04518	,000	-1,2893	-1,0067
		5,00	-,66795*	,04623	,000	-,8125	-,5234
		6,00	,73589*	,04404	,000	,5981	,8737
		7,00	,71742*	,04275	,000	,5836	,8512
	2,00	,00	-1,46233*	,02227	,000	-1,5318	-1,3928
		1,00	-,31632*	,04614	,000	-,4606	-,1720
		3,00	-1,70806*	,02080	,000	-1,7730	-1,6431
		4,00	-1,46432*	,02625	,000	-1,5463	-1,3824
		5,00	-,98427*	,02802	,000	-1,0717	-,8968
		6,00	,41956*	,02423	,000	,3439	,4952
		7,00	,40110*	,02181	,000	,3330	,4692
	3,00	,00	,24573*	,01231	,000	,2073	,2841
		1,00	1,39174*	,04225	,000	1,2595	1,5240
		2,00	1,70806*	,02080	,000	1,6431	1,7730
		4,00	,24374*	,01857	,000	,1857	,3018
		5,00	,72379*	,02100	,000	,6582	,7894
		6,00	2,12762*	,01559	,000	2,0790	2,1763
		7,00	2,10916*	,01146	,000	2,0734	2,1449
	4,00	,00	,00199	,02020	1,000	-,0611	,0651
		1,00	1,14800*	,04518	,000	1,0067	1,2893
		2,00	1,46432*	,02625	,000	1,3824	1,5463
		3,00	-,24374*	,01857	,000	-,3018	-,1857
		5,00	,48005*	,02641	,000	,3976	,5625
		6,00	1,88389*	,02235	,000	1,8141	1,9537
		7,00	1,86542*	,01969	,000	1,8039	1,9269

	5,00	,00	-,47806*	,02245	,000	-,5481	-,4080
		1,00	,66795*	,04623	,000	,5234	,8125
		2,00	,98427*	,02802	,000	,8968	1,0717
		3,00	-,72379*	,02100	,000	-,7894	-,6582
		4,00	-,48005*	,02641	,000	-,5625	-,3976
		6,00	1,40384*	,02441	,000	1,3277	1,4800
		7,00	1,38537*	,02200	,000	1,3167	1,4541
	6,00	,00	-1,88189*	,01749	,000	-1,9365	-1,8273
		1,00	-,73589*	,04404	,000	-,8737	-,5981
		2,00	-,41956*	,02423	,000	-,4952	-,3439
		3,00	-2,12762*	,01559	,000	-2,1763	-2,0790
		4,00	-1,88389*	,02235	,000	-1,9537	-1,8141
		5,00	-1,40384*	,02441	,000	-1,4800	-1,3277
		7,00	-,01847	,01691	1,000	-,0712	,0343
	7,00	,00	-1,86343*	,01394	,000	-1,9069	-1,8199
		1,00	-,71742*	,04275	,000	-,8512	-,5836
		2,00	-,40110*	,02181	,000	-,4692	-,3330
		3,00	-2,10916*	,01146	,000	-2,1449	-2,0734
		4,00	-1,86542*	,01969	,000	-1,9269	-1,8039
		5,00	-1,38537*	,02200	,000	-1,4541	-1,3167
		6,00	,01847	,01691	1,000	-,0343	,0712
XX13	,00	1,00	-,18414*	,02925	,000	-,2757	-,0926
		2,00	-,03742	,01459	,254	-,0829	,0081
		3,00	-2,11709*	,03163	,000	-2,2159	-2,0183
		4,00	-1,04680*	,04118	,000	-1,1756	-,9180
		5,00	-1,72709*	,02960	,000	-1,8195	-1,6347

		6,00	-,00730	,01526	1,000	-,0549	,0403
		7,00	,00137	,01177	1,000	-,0353	,0381
1,00	,00		,18414*	,02925	,000	,0926	,2757
	2,00		,14672*	,03009	,000	,0526	,2408
	3,00		-1,93296*	,04115	,000	-2,0614	-1,8045
	4,00		-,86266*	,04887	,000	-1,0153	-,7100
	5,00		-1,54296*	,03960	,000	-1,6666	-1,4193
	6,00		,17684*	,03042	,000	,0817	,2720
	7,00		,18551*	,02882	,000	,0953	,2757
2,00	,00		,03742	,01459	,254	-,0081	,0829
	1,00		-,14672*	,03009	,000	-,2408	-,0526
	3,00		-2,07968*	,03241	,000	-2,1809	-1,9785
	4,00		-1,00939*	,04178	,000	-1,1400	-,8787
	5,00		-1,68968*	,03042	,000	-1,7847	-1,5947
	6,00		,03012	,01681	,881	-,0223	,0826
	7,00		,03879	,01371	,124	-,0040	,0816
3,00	,00		2,11709*	,03163	,000	2,0183	2,2159
	1,00		1,93296*	,04115	,000	1,8045	2,0614
	2,00		2,07968*	,03241	,000	1,9785	2,1809
	4,00		1,07029*	,05033	,000	,9131	1,2275
	5,00		,39000*	,04139	,000	,2608	,5192
	6,00		2,10980*	,03271	,000	2,0076	2,2120
	7,00		2,11847*	,03124	,000	2,0209	2,2161
4,00	,00		1,04680*	,04118	,000	,9180	1,1756
	1,00		,86266*	,04887	,000	,7100	1,0153
	2,00		1,00939*	,04178	,000	,8787	1,1400

			3,00	-1,07029*	,05033	,000	-1,2275	-,9131
			5,00	-,68029*	,04907	,000	-,8335	-,5270
			6,00	1,03950*	,04201	,000	,9081	1,1709
			7,00	1,04818*	,04087	,000	,9203	1,1760
		5,00	,00	1,72709*	,02960	,000	1,6347	1,8195
			1,00	1,54296*	,03960	,000	1,4193	1,6666
			2,00	1,68968*	,03042	,000	1,5947	1,7847
			3,00	-,39000*	,04139	,000	-,5192	-,2608
			4,00	,68029*	,04907	,000	,5270	,8335
			6,00	1,71980*	,03075	,000	1,6238	1,8158
			7,00	1,72847*	,02917	,000	1,6373	1,8196
		6,00	,00	,00730	,01526	1,000	-,0403	,0549
			1,00	-,17684*	,03042	,000	-,2720	-,0817
			2,00	-,03012	,01681	,881	-,0826	,0223
			3,00	-2,10980*	,03271	,000	-2,2120	-2,0076
			4,00	-1,03950*	,04201	,000	-1,1709	-,9081
			5,00	-1,71980*	,03075	,000	-1,8158	-1,6238
			7,00	,00867	,01442	1,000	-,0363	,0537
		7,00	,00	-,00137	,01177	1,000	-,0381	,0353
			1,00	-,18551*	,02882	,000	-,2757	-,0953
			2,00	-,03879	,01371	,124	-,0816	,0040
			3,00	-2,11847*	,03124	,000	-2,2161	-2,0209
			4,00	-1,04818*	,04087	,000	-1,1760	-,9203
			5,00	-1,72847*	,02917	,000	-1,8196	-1,6373
			6,00	-,00867	,01442	1,000	-,0537	,0363
	XX14	,00	1,00	1,36638*	,04036	,000	1,2401	1,4927

		2,00	1,55238*	,02395	,000	1,4776	1,6271
		3,00	,01621	,02310	1,000	-,0559	,0883
		4,00	,27281*	,03173	,000	,1736	,3720
		5,00	1,60898*	,02219	,000	1,5397	1,6782
		6,00	1,93234*	,01557	,000	1,8838	1,9809
		7,00	1,91003*	,01383	,000	1,8669	1,9532
	1,00	,00	-1,36638*	,04036	,000	-1,4927	-1,2401
		2,00	,18600*	,04394	,001	,0486	,3234
		3,00	-1,35017*	,04349	,000	-1,4861	-1,2142
		4,00	-1,09356*	,04862	,000	-1,2455	-,9417
		5,00	,24261*	,04301	,000	,1081	,3771
		6,00	,56596*	,04000	,000	,4408	,6911
		7,00	,54365*	,03935	,000	,4205	,6669
	2,00	,00	-1,55238*	,02395	,000	-1,6271	-1,4776
		1,00	-,18600*	,04394	,001	-,3234	-,0486
		3,00	-1,53617*	,02891	,000	-1,6264	-1,4459
		4,00	-1,27956*	,03618	,000	-1,3926	-1,1666
		5,00	,05661	,02818	,722	-,0313	,1446
		6,00	,37996*	,02333	,000	,3071	,4528
		7,00	,35765*	,02221	,000	,2883	,4270
	3,00	,00	-,01621	,02310	1,000	-,0883	,0559
		1,00	1,35017*	,04349	,000	1,2142	1,4861
		2,00	1,53617*	,02891	,000	1,4459	1,6264
		4,00	,25660*	,03563	,000	,1453	,3679
		5,00	1,59277*	,02747	,000	1,5070	1,6785
		6,00	1,91613*	,02246	,000	1,8460	1,9863

		7,00	1,89382*	,02130	,000	1,8273	1,9603
	4,00	,00	-,27281*	,03173	,000	-,3720	-,1736
		1,00	1,09356*	,04862	,000	,9417	1,2455
		2,00	1,27956*	,03618	,000	1,1666	1,3926
		3,00	-,25660*	,03563	,000	-,3679	-,1453
		5,00	1,33617*	,03504	,000	1,2267	1,4456
		6,00	1,65953*	,03127	,000	1,5618	1,7573
		7,00	1,63722*	,03044	,000	1,5420	1,7324
	5,00	,00	-1,60898*	,02219	,000	-1,6782	-1,5397
		1,00	-,24261*	,04301	,000	-,3771	-,1081
		2,00	-,05661	,02818	,722	-,1446	,0313
		3,00	-1,59277*	,02747	,000	-1,6785	-1,5070
		4,00	-1,33617*	,03504	,000	-1,4456	-1,2267
		6,00	,32336*	,02152	,000	,2562	,3905
		7,00	,30105*	,02030	,000	,2377	,3644
	6,00	,00	-1,93234*	,01557	,000	-1,9809	-1,8838
		1,00	-,56596*	,04000	,000	-,6911	-,4408
		2,00	-,37996*	,02333	,000	-,4528	-,3071
		3,00	-1,91613*	,02246	,000	-1,9863	-1,8460
		4,00	-1,65953*	,03127	,000	-1,7573	-1,5618
		5,00	-,32336*	,02152	,000	-,3905	-,2562
		7,00	-,02231	,01274	,903	-,0621	,0174
	7,00	,00	-1,91003*	,01383	,000	-1,9532	-1,8669
		1,00	-,54365*	,03935	,000	-,6669	-,4205
		2,00	-,35765*	,02221	,000	-,4270	-,2883
		3,00	-1,89382*	,02130	,000	-1,9603	-1,8273

			4,00	-1,63722*	,03044	,000	-1,7324	-1,5420
			5,00	-,30105*	,02030	,000	-,3644	-,2377
			6,00	,02231	,01274	,903	-,0174	,0621
	XX15	,00	1,00	-,99212*	,04534	,000	-1,1341	-,8502
			2,00	-,03979*	,01183	,022	-,0767	-,0029
			3,00	-,01910	,01288	,985	-,0593	,0211
			4,00	-1,72944*	,04170	,000	-1,8599	-1,5990
			5,00	-,10033*	,01537	,000	-,1483	-,0523
			6,00	-2,11258*	,03241	,000	-2,2138	-2,0114
			7,00	-,69924*	,01748	,000	-,7538	-,6447
		1,00	,00	,99212*	,04534	,000	,8502	1,1341
			2,00	,95233*	,04572	,000	,8092	1,0955
			3,00	,97302*	,04600	,000	,8290	1,1170
			4,00	-,73732*	,06074	,000	-,9270	-,5476
			5,00	,89179*	,04676	,000	,7455	1,0381
			6,00	-1,12046*	,05478	,000	-1,2916	-,9493
			7,00	,29288*	,04750	,000	,1443	,4415
		2,00	,00	,03979*	,01183	,022	,0029	,0767
			1,00	-,95233*	,04572	,000	-1,0955	-,8092
			3,00	,02069	,01415	,987	-,0235	,0649
			4,00	-1,68965*	,04211	,000	-1,8214	-1,5580
			5,00	-,06054*	,01645	,007	-,1119	-,0092
			6,00	-2,07279*	,03293	,000	-2,1757	-1,9699
			7,00	-,65945*	,01844	,000	-,7170	-,6019
		3,00	,00	,01910	,01288	,985	-,0211	,0593
			1,00	-,97302*	,04600	,000	-1,1170	-,8290

		2,00	-,02069	,01415	,987	-,0649	,0235
		4,00	-1,71034*	,04241	,000	-1,8430	-1,5777
		5,00	-,08123*	,01723	,000	-,1350	-,0275
		6,00	-2,09348*	,03333	,000	-2,1976	-1,9894
		7,00	-,68014*	,01913	,000	-,7398	-,6205
	4,00	,00	1,72944*	,04170	,000	1,5990	1,8599
		1,00	,73732*	,06074	,000	,5476	,9270
		2,00	1,68965*	,04211	,000	1,5580	1,8214
		3,00	1,71034*	,04241	,000	1,5777	1,8430
		5,00	1,62911*	,04324	,000	1,4939	1,7643
		6,00	-,38314*	,05180	,000	-,5449	-,2214
		7,00	1,03020*	,04403	,000	,8926	1,1678
	5,00	,00	,10033*	,01537	,000	,0523	,1483
		1,00	-,89179*	,04676	,000	-1,0381	-,7455
		2,00	,06054*	,01645	,007	,0092	,1119
		3,00	,08123*	,01723	,000	,0275	,1350
		4,00	-1,62911*	,04324	,000	-1,7643	-1,4939
		6,00	-2,01225*	,03437	,000	-2,1196	-1,9049
		7,00	-,59891*	,02089	,000	-,6641	-,5338
	6,00	,00	2,11258*	,03241	,000	2,0114	2,2138
		1,00	1,12046*	,05478	,000	,9493	1,2916
		2,00	2,07279*	,03293	,000	1,9699	2,1757
		3,00	2,09348*	,03333	,000	1,9894	2,1976
		4,00	,38314*	,05180	,000	,2214	,5449
		5,00	2,01225*	,03437	,000	1,9049	2,1196
		7,00	1,41334*	,03536	,000	1,3030	1,5237

		7,00	,00	,69924*	,01748	,000	,6447	,7538
			1,00	-,29288*	,04750	,000	-,4415	-,1443
			2,00	,65945*	,01844	,000	,6019	,7170
			3,00	,68014*	,01913	,000	,6205	,7398
			4,00	-1,03020*	,04403	,000	-1,1678	-,8926
			5,00	,59891*	,02089	,000	,5338	,6641
			6,00	-1,41334*	,03536	,000	-1,5237	-1,3030
	XX16	,00	1,00	,53467*	,05033	,000	,3775	,6919
			2,00	,26498*	,04643	,000	,1201	,4099
			3,00	-,16012	,05312	,071	-,3259	,0057
			4,00	,89372*	,03165	,000	,7950	,9925
			5,00	,44222*	,04257	,000	,3094	,5750
			6,00	,87403*	,03085	,000	,7778	,9703
			7,00	,93268*	,02923	,000	,8415	1,0239
		1,00	,00	-,53467*	,05033	,000	-,6919	-,3775
			2,00	-,26969*	,05506	,000	-,4416	-,0978
			3,00	-,69479*	,06081	,000	-,8846	-,5049
			4,00	,35908*	,04333	,000	,2235	,4946
			5,00	-,09245	,05184	,887	-,2544	,0695
			6,00	,33937*	,04275	,000	,2056	,4731
			7,00	,39802*	,04159	,000	,2678	,5282
		2,00	,00	-,26498*	,04643	,000	-,4099	-,1201
			1,00	,26969*	,05506	,000	,0978	,4416
			3,00	-,42510*	,05762	,000	-,6049	-,2452
			4,00	,62874*	,03873	,000	,5078	,7497
			5,00	,17724*	,04807	,006	,0272	,3272

		6,00	,60906*	,03808	,000	,4901	,7280
		7,00	,66771*	,03678	,000	,5528	,7826
	3,00	,00	,16012	,05312	,071	-,0057	,3259
		1,00	,69479*	,06081	,000	,5049	,8846
		2,00	,42510*	,05762	,000	,2452	,6049
		4,00	1,05384*	,04655	,000	,9084	1,1993
		5,00	,60234*	,05456	,000	,4320	,7726
		6,00	1,03415*	,04601	,000	,8904	1,1779
		7,00	1,09280*	,04494	,000	,9524	1,2332
	4,00	,00	-,89372*	,03165	,000	-,9925	-,7950
		1,00	-,35906*	,04333	,000	-,4946	-,2235
		2,00	-,62874*	,03873	,000	-,7497	-,5078
		3,00	-1,05384*	,04655	,000	-1,1993	-,9084
		5,00	-,45151*	,03401	,000	-,5577	-,3453
		6,00	-,01969	,01721	1,000	-,0734	,0340
		7,00	,03896	,01410	,152	-,0051	,0830
	5,00	,00	-,44222*	,04257	,000	-,5750	-,3094
		1,00	,09245	,05184	,887	-,0695	,2544
		2,00	-,17724*	,04807	,006	-,3272	-,0272
		3,00	-,60234*	,05456	,000	-,7726	-,4320
		4,00	,45151*	,03401	,000	,3453	,5577
		6,00	,43182*	,03326	,000	,3279	,5357
		7,00	,49047*	,03177	,000	,3912	,5897
	6,00	,00	-,87403*	,03085	,000	-,9703	-,7778
		1,00	-,33937*	,04275	,000	-,4731	-,2056
		2,00	-,60906*	,03808	,000	-,7280	-,4901

			3,00	-1,03415*	,04601	,000	-1,1779	-,8904
			4,00	,01969	,01721	1,000	-,0340	,0734
			5,00	-,43182*	,03326	,000	-,5357	-,3279
			7,00	,05865*	,01220	,000	,0206	,0967
		7,00	,00	-,93268*	,02923	,000	-1,0239	-,8415
			1,00	-,39802*	,04159	,000	-,5282	-,2678
			2,00	-,66771*	,03678	,000	-,7826	-,5528
			3,00	-1,09280*	,04494	,000	-1,2332	-,9524
			4,00	-,03896	,01410	,152	-,0830	,0051
			5,00	-,49047*	,03177	,000	-,5897	-,3912
			6,00	-,05865*	,01220	,000	-,0967	-,0206
	XX17	,00	1,00	,99844*	,02669	,000	,9150	1,0818
			2,00	,98478*	,01887	,000	,9259	1,0436
			3,00	-1,28823*	,02226	,000	-1,3577	-1,2188
			4,00	-,66154*	,03271	,000	-,7638	-,5593
			5,00	,04880	,02526	,785	-,0300	,1276
			6,00	1,29230*	,01775	,000	1,2369	1,3477
			7,00	1,26280*	,01597	,000	1,2130	1,3126
		1,00	,00	-,99844*	,02669	,000	-1,0818	-,9150
			2,00	-,01366	,02564	1,000	-,0938	,0665
			3,00	-2,28667*	,02823	,000	-2,3749	-2,1985
			4,00	-1,65997*	,03704	,000	-1,7757	-1,5443
			5,00	-,94963*	,03065	,000	-1,0453	-,8539
			6,00	,29386*	,02484	,000	,2162	,3715
			7,00	,26436*	,02359	,000	,1905	,3382
		2,00	,00	-,98478*	,01887	,000	-1,0436	-,9259

		1,00	,01366	,02564	1,000	-,0665	,0938
		3,00	-2,27301*	,02098	,000	-2,3385	-2,2075
		4,00	-1,64631*	,03186	,000	-1,7459	-1,5467
		5,00	-,93597*	,02414	,000	-1,0113	-,8606
		6,00	,30752*	,01613	,000	,2572	,3579
		7,00	,27802*	,01414	,000	,2339	,3222
	3,00	,00	1,28823*	,02226	,000	1,2188	1,3577
		1,00	2,28667*	,02823	,000	2,1985	2,3749
		2,00	2,27301*	,02098	,000	2,2075	2,3385
		4,00	,62670*	,03398	,000	,5205	,7328
		5,00	1,33704*	,02687	,000	1,2532	1,4209
		6,00	2,58053*	,01999	,000	2,5181	2,6429
		7,00	2,55103*	,01842	,000	2,4935	2,6086
	4,00	,00	,66154*	,03271	,000	,5593	,7638
		1,00	1,65997*	,03704	,000	1,5443	1,7757
		2,00	1,64631*	,03186	,000	1,5467	1,7459
		3,00	-,62670*	,03398	,000	-,7328	-,5205
		5,00	,71034*	,03602	,000	,5979	,8228
		6,00	1,95383*	,03122	,000	1,8562	2,0514
		7,00	1,92433*	,03024	,000	1,8298	2,0189
	5,00	,00	-,04880	,02526	,785	-,1276	,0300
		1,00	,94963*	,03065	,000	,8539	1,0453
		2,00	,93597*	,02414	,000	,8606	1,0113
		3,00	-1,33704*	,02687	,000	-1,4209	-1,2532
		4,00	-,71034*	,03602	,000	-,8228	-,5979
		6,00	1,24349*	,02328	,000	1,1708	1,3162

		7,00	1,21399*	,02195	,000	1,1454	1,2825
	6,00	,00	-1,29230*	,01775	,000	-1,3477	-1,2369
		1,00	-,29386*	,02484	,000	-,3715	-,2162
		2,00	-,30752*	,01613	,000	-,3579	-,2572
		3,00	-2,58053*	,01999	,000	-2,6429	-2,5181
		4,00	-1,95383*	,03122	,000	-2,0514	-1,8562
		5,00	-1,24349*	,02328	,000	-1,3162	-1,1708
		7,00	-,02950	,01262	,424	-,0689	,0099
	7,00	,00	-1,26280*	,01597	,000	-1,3126	-1,2130
		1,00	-,26436*	,02359	,000	-,3382	-,1905
		2,00	-,27802*	,01414	,000	-,3222	-,2339
		3,00	-2,55103*	,01842	,000	-2,6086	-2,4935
		4,00	-1,92433*	,03024	,000	-2,0189	-1,8298
		5,00	-1,21399*	,02195	,000	-1,2825	-1,1454
		6,00	,02950	,01262	,424	-,0099	,0689
XX18	,00	1,00	-,70571*	,03502	,000	-,8153	-,5961
		2,00	-,07987*	,01472	,000	-,1258	-,0339
		3,00	-,02477	,01335	,842	-,0664	,0169
		4,00	-1,71689*	,04828	,000	-1,8679	-1,5658
		5,00	-,08807*	,01417	,000	-,1323	-,0439
		6,00	-2,11284*	,03792	,000	-2,2313	-1,9944
		7,00	-,64912*	,01449	,000	-,6943	-,6039
	1,00	,00	,70571*	,03502	,000	,5961	,8153
		2,00	,62584*	,03656	,000	,5115	,7402
		3,00	,68094*	,03604	,000	,5682	,7937
		4,00	-1,01118*	,05875	,000	-1,1947	-,8277

		5,00	,61764*	,03635	,000	,5039	,7313
		6,00	-1,40713*	,05058	,000	-1,5650	-1,2492
		7,00	,05659	,03647	,973	-,0575	,1707
	2,00	,00	,07987*	,01472	,000	,0339	,1258
		1,00	-,62584*	,03656	,000	-,7402	-,5115
		3,00	,05510*	,01699	,033	,0021	,1081
		4,00	-1,63702*	,04941	,000	-1,7916	-1,4825
		5,00	-,00820	,01764	1,000	-,0632	,0468
		6,00	-2,03297*	,03935	,000	-2,1558	-1,9101
		7,00	-,56925*	,01790	,000	-,6251	-,5134
	3,00	,00	,02477	,01335	,842	-,0169	,0664
		1,00	-,68094*	,03604	,000	-,7937	-,5682
		2,00	-,05510*	,01699	,033	-,1081	-,0021
		4,00	-1,69212*	,04902	,000	-1,8455	-1,5388
		5,00	-,06331*	,01652	,004	-,1149	-,0118
		6,00	-2,08807*	,03886	,000	-2,2094	-1,9667
		7,00	-,62435*	,01680	,000	-,6768	-,5719
	4,00	,00	1,71689*	,04828	,000	1,5658	1,8679
		1,00	1,01118*	,05875	,000	,8277	1,1947
		2,00	1,63702*	,04941	,000	1,4825	1,7916
		3,00	1,69212*	,04902	,000	1,5388	1,8455
		5,00	1,62882*	,04925	,000	1,4748	1,7829
		6,00	-,39595*	,06052	,000	-,5849	-,2070
		7,00	1,06777*	,04935	,000	,9134	1,2221
	5,00	,00	,08807*	,01417	,000	,0439	,1323
		1,00	-,61764*	,03635	,000	-,7313	-,5039

		2,00	,00820	,01764	1,000	-,0468	,0632
		3,00	,06331*	,01652	,004	,0118	,1149
		4,00	-1,62882*	,04925	,000	-1,7829	-1,4748
		6,00	-2,02477*	,03915	,000	-2,1470	-1,9025
		7,00	-,56104*	,01745	,000	-,6155	-,5066
	6,00	,00	2,11284*	,03792	,000	1,9944	2,2313
		1,00	1,40713*	,05058	,000	1,2492	1,5650
		2,00	2,03297*	,03935	,000	1,9101	2,1558
		3,00	2,08807*	,03886	,000	1,9667	2,2094
		4,00	,39595*	,06052	,000	,2070	,5849
		5,00	2,02477*	,03915	,000	1,9025	2,1470
		7,00	1,46373*	,03926	,000	1,3411	1,5863
	7,00	,00	,64912*	,01449	,000	,6039	,6943
		1,00	-,05659	,03647	,973	-,1707	,0575
		2,00	,56925*	,01790	,000	,5134	,6251
		3,00	,62435*	,01680	,000	,5719	,6768
		4,00	-1,06777*	,04935	,000	-1,2221	-,9134
		5,00	,56104*	,01745	,000	,5066	,6155
		6,00	-1,46373*	,03926	,000	-1,5863	-1,3411