T.C.
İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

PERAKENDE SEKTÖRÜNDE LOKASYON ANALİZİ VE YER SEÇİMİ

Ahmet KAYIRAN

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Programı

DANIŞMAN
Dr. Öğr. Üyesi Funda Hatice SEZGİN

II. DANIŞMAN
Dr. Öğr. Üyesi Ersin NAMLI

Mayıs, 2018

İSTANBUL

Bu çalışma,25.06.2018 tarihinde aşağıdaki jüri tarafındanEndüstri Mühendisliği Anabilim Dalı,Endüstri Mühendisliği Programında Yüksek Lisans tezi olarak kabul edilmiştir.

Tez Jürisi

Dr. Öğr. Üyesi Funda Hatice Sezgin(Danışman)
İstanbul Üniversitesi
Mühendislik Fakültesi

Prof. Dr. Ş. Alp BARAY
İstanbul Üniversitesi
Mühendislik Fakültesi

Doç. Dr. Tarık Küçükdeniz
İstanbul Üniversitesi
Mühendislik Fakültesi

Doç. Dr. Ozan KOCADAGLI
Mimar Sinan Üniversitesi
Fen Edebiyat Fakültesi

Dr. Öğr. Üyesi Tuncay ÖZCAN
İstanbul Üniversitesi
Mühendislik Fakültesi

20.04.2016 tarihli resmî gazetede yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, İstanbul Üniversitesi'nin abonesi olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

ÖNSÖZ

"Perakende Sektöründe Lokasyon Analizi ve Yer Seçimi" adlı tez çalışmasının yürütülmesinde ve sonuçlandırılmasında yardım ve desteklerini esirgemeyen danışman hocam Sayın Dr. Öğr. Üyesi Funda Hatice Sezgin'e teşekkürlerimi sunarım.

Ayrıca çalışmanın her aşamasında sürekli görüş ve fikirlerine başvurduğum, gelişen yöntem ve analizler konusunda bana yön veren, önerileri ile araştırmanının gelişmesini sağlayan değerli Hocam ve ağabeyim Sayın Dr. Öğr. Üyesi Ersin Namlı 'ya teşekkür ederim.

İyi ve kötü günlerimde beni yalnız bırakmayan, değerli vakitlerini çalışmam için ayıran başta Elif Açıkgöz Okuducu ve Tüylü ailesi olmak üzere tüm değerli dostlarıma teşekkür ederim.

Eğitim hayatım boyunca beni destekleyen, hayat görüşümün ve ideallerimin şekillenmesinde yardımcı olan KAYIRAN ailesine sevgi saygı ve minneti bir borç bilirim.

Mayıs 2018 Ahmet KAYIRAN

İÇİNDEKİLER

Sayfa No

ÖNSÖZ ..... **iv**
İÇİNDEKİLER ..... **v**
ŞEKİL LİSTESİ ..... **vii**
TABLOLAR LİSTESİ ..... **ix**
SİMGE VE KISALTMA LİSTESİ ..... **xi**
ÖZET ..... **xiii**
SUMMARY ..... **xiv**
1. GİRİŞ **1**
2. GENEL KISIMLAR **3**
2.1.PERAKENDE SEKTÖRÜ 3
2.1.1. Perakende Nedir? 3
2.1.2.Perakende Sektörünün Tarihsel Gelişimi 4
2.1.3.Perakende Dönüşümü 6
2.1.4. Perakende Kanal Tipleri 9
2.1.5. Dünyada Perakende Sektörü 11
2.1.6. Türkiye'de Perakende Sektörü 12
2.2. LOKASYON ANALİZİ VE YER SEÇİMİ 13
2.2.1.LokasyonTeorisi 15
2.2.2. Perakende Sektöründe Lokasyon Analizi 23
2.3. MAKİNE ÖĞRENME 29
2.3.1. Tembel (Lazy) Algoritmalar 30
2.3.2. Bayes Ağları Algoritmaları 31
2.3.3. Kural Tabanlı Algoritmalar 31
2.3.4. Karar Ağacı Algoritmaları 32
2.3.5. Sınıflandırmada Kullanılan Fonksiyonlar 35
2.3.6. Bileşik Makine Öğrenme 36
3. MALZEME VE YÖNTEM **42**
3.1. WEKA 42

3.2. M5P ALGORİTMASI 43
3.3. M5RULES ALGORİTMASI 45
3.4. LWL (YEREL AĞIRLIKLI ÖĞRENME) ALGORİTMASI 46
3.5. KARAR TABLOSU ALGORİTMASI 46
3.6. DESTEK VEKTÖR REGRESYONU ALGORİTMASI 47
3.7 DESTEK VEKTÖR MAKİNELERİ ALGORİTMASI 47
3.8. DOĞRUSAL REGRESYON ALGORİTMASI 50
3.9. BAGGİNG ALGORİTMASI 51
3.10. VOTİNG (OYLAMA) ALGORİTMASI 52
3.11. RANDOMSUBSPACE (RASTGELE ALTUZAYLAR) ALGORİTMASI 53
3.12. VERİ ANALİZİ 54
3.12.1. Sonuç Değişkeni (Müşteri Gelişleri) 57
3.12.2. Kent Nüfus 60
3.12.3. Kent Gelişmişlik Değeri 62
3.12.4. Kent Ekonomi 64
3.12.4. Ortalama m2 değeri 64
3.12.5. Foot Traffic (Yaya trafik yoğunluğu) 65
3.12.6. Retail Index (Perakende harcaması yoğunluğu) 65
3.12.7. 1000 m, 500 m nokta verileri 65
4. BULGULAR 67
4.1. VAKA ÇALIŞMASI 67
4.2. SONUÇLARIN DEĞERLENDİRİLMESİ 68
4.2.1. Basit Makine Öğrenme Algoritmaları 68
4.2.2. Bileşik Makine Öğrenme Algoritmaları 79
5. TARTIŞMA VE SONUÇ 86
KAYNAKLAR 88
EKLER 94
ÖZGEÇMİŞ 96

ŞEKİL LİSTESİ

Sayfa No

Şekil 2.1: Perakende kanal fonksiyonlarının değişimi (Alpay, 2006). 4

Şekil 2.2: Müşteri davranış modeli 9

Şekil 2.3: Gelişen teknolojilere adapte olma süresi (comScore Dünya ve Türkiye Elektornik Pazar yeri trendleri raporu) 10

Şekil 2.4: Giyim sektöründe Kanal Tipleri dağılımı (PWC, 2013). 11

Şekil 2.5: Dünya geneli perakende sektörü (KPMG, 2018) 12

Şekil 2.6: Perakende satış ciroları (TÜİK, 2017) 13

Şekil 2.7: Ekonomik ve iktisadi coğrafyanın dağılımı (Taaffe ve Gauthier, 1994) 14

Şekil 2.8: Weberyen lojistik taşıma maliyetleri ve hammaddenin niteliğine göre optimum lokasyon (Weber, 1929; Yavan, 2006). 17

Şekil 2.9: Weber'in lokasyon üçgeni (Yavan, 2006). 18

Şekil 2.10: Weber'in eş maliyetli ulaşım eğrileri analizi çerçevesinde lojistik taşıma maliyetleri ve emek faktörlerine göre optimum lokasyon (Yavan, 2006). 18

Şekil 2.11: Weber'e göre yığılma ekonomilerinin oluşumu (Yavan, 2006). 19

Şekil 2.12: Lösch'ün mekânsal talep konisi ve Pazar alanları (Yavan, 2006). 20

Şekil 2.13: Hotelling modeli rakip firma stratejileri (Yavan, 2006; Hotelling, 1990). 21

Şekil 2.14: Christaller'in merkezi yer hiyerarşisi ve altıgen piyasa alanları (Peter Dicken ve Lloyd, 1990). 22

Şekil 2.15: Perakende Firması Mağaza Açılış Süreci İş Akışı 25

Şekil 2.16: Karar Ağacı Algoritması Çalışma Şekli (Ho, 2018) 33

Şekil 2.17: Makine Öğrenme Algoritma Seçme Kuralı (Pedregosa vd. 2011). 35

Şekil 2.18: Destek Vektör Fonksiyonları Sınıflandırma Desenleri (Pedregosa vd. 2011) 36

Şekil 2.19: Bileşik sınıflandırma makine öğrenme modelinin yapısı (Hsieh vd., 2012). 37

Şekil 3.2: M5P Yöntemi (Wang ve Witten, 1997; Quinlan, 1992) 38

Şekil 3.3: İki farklı veri sınıfının doğrusal olarak ayrılması (Steinwart ve Christmann, 2008; Shmilovici, 2009) 48

Şekil 3.4: Doğrusal olarak ayrılamayan verilerin sınıflandırılması (Steinwart ve Christmann, 2008)49

Şekil 3.5: Bagging Yöntemi (Aydogmus vd., 2015) 52

Şekil 3.6: Voting (Oylama) Yöntemi 53

Şekil 3.7: Saalik müşteri geliş dağılımları 59

Şekil 4.1: Uygulama Metodolojisi 67

Şekil 4.2: RandomTree ağaç yapısı 72

Şekil 4.3: J48 ağaç yapısı 74

TABLOLAR LİSTESİ

Sayfa No

Tablo 3.1: Lokasyonlara ait demografi verileri 57

Tablo 3.2: Lokasyonlara ait seçilen parametreler 57

Tablo 3.3: Müşteri gelişlerine ait değer segmentasyonu veri seti 58

Tablo 3.4: Değer segmentsayonu temel istatistiki göstergeleri 58

Tablo 3.5: Mağaza davranış segmentasyonu 60

Tablo 3.6: İl Türlerine göre ortalam nüfus verileri 60

Tablo 3.7: Kent nüfüsü ile küme değerleri eleman sayısı 62

Tablo 3.8: Kent gelişmişlik değeri temel istatistiki verileri 62

Tablo 3.9: Kent gelişmişlik değeri temel istatistiki verileri 62

Tablo 3.10: Pearson Korelasyon sonuçları 63

Tablo 3.11: Kent ekonomi katagorik verisi ile küme değerleri eleman sayısı 63

Tablo 3.12: Küme değerlerine göre ortalama yaya yoğunluğu 64

Tablo 3.13: Küme değerlerine göre toplam nokta sayıları 65

Tablo 4.1: RandomTree algoritma sonuçları (%90-%10) 70

Tablo 4.2: RandomTree algoritması kural ağacı 70

Tablo 4.3: RandomForest algoritma sonuçları (%90-%10) 72

Tablo 4.4: J48 algoritma sonuçları (%80-%20) 73

Tablo 4.5: J48 algoritması kural ağacı 73

Tablo 4.6: Hoeffding algoritma sonuçları (%90-%10) 75

Tablo 4.7: DecisionTable algoritma sonuçları (10 Fold) 77

Tablo 4.8: JRip algoritma sonuçları (20 Fold) 77

Tablo 4.9: LWL algoritma sonuçları (20 Fold) 77

Tablo 4.10: Multilayerperceptron algoritma sonuçları (10 Fold) 79

Tablo 4.11: Basit makine öğrenme algoritma sonuçları karşılaştırması 80

Tablo 4.12: RandoumSubSpace algoritma sonuçları 80

Tablo 4.13: Bagging algoritma sonuçları 82

Tablo 4.14: Vote algoritma sonuçları 82

Tablo 4.15: NaiveBayesUpdateable nominal değişken sonuçları 83

Tablo 4.16: NaiveBayesUpdateable numerik değişken sonuçları 84

SİMGE VE KISALTMA LİSTESİ

Simgeler	**Açıklama**
$\Longrightarrow$	: İse
$\wedge$	: Ve
$\vee$	: Veya
$\forall$	: Her
$\in$	: Elemanı
$\notin$	: Elemanı değil
%	: Yüzde
$=$	: Eşit
$\neq$	: Eşit değil
$<$	: Küçük
$>$	: Büyük
$\leq$	: Küçük eşit
$\geq$	: Büyük eşit
$\lvert \boldsymbol{x} \rvert$	: x'in mutlak değeri
$\sum$	: Toplam sembolü
$\left[\boldsymbol{c_{ij}}\right]_{\boldsymbol{N \times N}}$	: c_{ij} elemanlarından oluşan $N \times N$ boyutlu matris gösterimi
$[\boldsymbol{t_i}]_{\boldsymbol{N}}$	: t_i elemanlarından oluşan N boyutlu vektör gösterimi

Kısaltmalar	Açıklama
ABD	**:** Amerika Birleşik Devletleri
ANN	**:** Artificial Neural Network (Yapay Sinir Ağları)
BMÖ	**:** Bileşik Makine Öğrenme
DT	**:** Decision Table (Karar Tablosu)
DVM	**:** Destek Vektör Makinesi
EKKDVM	**:** En Küçük Kareler Destek Vektör Makinesi
EKKDVR	**:** En Küçük Kareler Destek Vektör Regresyonu
EKK-DVS	**:** En Küçük Kareler Destek Vektör Sınıflandırması
ENN	**:** Düzenlenmiş En Yakın Komşu
EPSO	**:** Gelişmiş Parçacık Sürü Optimizasyonu
GA	**:** Genetik Algoritmalar (Genetic Algorithm)
GSYİH	**:** Gayri Safi Yurt İçi Hasıla
ID3	**:** Iterative Dichotomiser 3 (İteratif İkileştirici 3)
IDTM	**:** Inducer of Decision Table Majority (Kara tablosu Çoğunluğunun Oluşturucusu)
kNN	**:** k En Yakın Komşu
LR	**:** Linear Regression (Doğrusal Regresyon)
LS-SVM	**:** En Küçük Kareler Destek Vektör Makineleri
MAE	**:** Ortalama Mutlak Hata
MKS	**:** Model Kurma Süresi
ML	**:** Machine Learning (Makine Öğrenme)
MÖ	**:** Makine Öğrenme
R	**:** Korelasyon Katsayısı
RAE	**:** Ortalama Bağıl Hata
RMSE	**:** Ortalama Hata Kareleri Kökü
RRSE	**:** Ortalama Bağıl Hata Kareleri kökü
RSS	**:** Random Sub Space (Rasgele Alt Uzay)
SMO	**:** Support Vector Machines (Destek Vektör Makineleri)
SMOreg	**:** Support Vector Regression (Destek Vektör Regresyonu)
SVM	**:** Support Vector Machines (Destek Vektör Makineleri)
SVR	**:** Support Vector Regression (Destek Vektör Regresyonu
WEKA	**:** Bilgi Analizi için Waikato Ortamı
YSA	**:** Yapay Sinir Ağları (Artificial Neural Networks)
0-R	**:** Zero Rule (0 Kural)
1-R	**:** One Rule (1 Kural)

ÖZET

YÜKSEK LİSANS TEZİ

PERAKENDE SEKTÖRÜNDE LOKASYON ANALİZİ VE YER SEÇİMİ

Ahmet KAYIRAN

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Endüstri Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Funda Hatice SEZGİN

Perakende Sektörü ülke ekonomileri için en önemli faaliyet alanıdır. Rekabetin hızlandığı ve şehirlerde yeni ticari alanların açıldığı bir zamanda firmalar için kuruluş yeri tercihi önemli hale gelmiştir. Lokasyonun çekim merkezleri içerisinde olması müşteriler için en önemli kriterler arasındadır. Firmaların rakiplerine, ulaşım noktalarına yakınlığı tercih nedenlerini arttıran parametrelerdendir. En iyi yatırım kararı için ilçe demografi ve ekonomi verileri ayrıca incelenmesi gereken göstergelerdendir. Bu nedenle Yatırım için en değerli nokta neresidir? Sorusunun cevabı firmalar için rekabetin sert olduğu ortamda en verimli karar destek göstergesidir.

Bu çalışmada müşteri gelişlerinin tahminlemesi yapılmıştır. Benzer lokasyonların bulunması için makine öğrenme algoritmaları kullanılarak yer seçimi için en uygun parametreler belirlenmiştir. Ayrıca elde edilen sonuçlar karşılaştırmalı olarak verilmiş algoritma çalışma parametreleri için en uygun değerler belirlenmiştir.

Mayıs 2018, 97 sayfa.

Anahtar kelimeler: Perakende Sektörü, Lokasyon Analizi, Yer Seçimi, Makine Öğrenme, Bileşik Makine Öğrenme

SUMMARY

M.Sc. THESIS

LOCATION ANALYSIS AND LOCATION SELECTION IN RETAIL MARKETING

Ahmet KAYIRAN

İstanbul University

Institute of Graduate Studies in Science and Engineering

Department of Industrial Engineering

Supervisor: Assist. Prof. Dr.Funda Hatice SEZGİN

The retail marketing is the most important activity area for the country's economies. At a time when competition has accelerated and new commercial areas in cities have opened up, the choice of establishment place for companies has become important. The fact that the location is within the shooting centers is one of the most important criteria for the customers. It is one of the parameters that increases the preference reasons for the closeness of the companies to their competitors and their transportation points. For the best investment decision, district demographics and economic data are also indicators that need to be analyzed. So what is the most valuable point for Investment? The answer to the question is the most efficient decision support indicator in the environment where competition is hard for companies.

In this study, estimations of customer arrivals were made. In order to find similar locations, machine learning algorithms were used to determine the most suitable parameters for location selection. In addition, the most suitable values are determined for the obtained algorithm parameters.

May 2018,97.pages.

Keywords: Retail Industry, Location Analysis, Location Selection, Machine Learning, Combined Machine Learning

1. GİRİŞ

İlk çağlardan itibaren insanların hayatlarını idame ettirebilmeleri için çeşitli sosyal faaliyetlerde bir araya gelme ve yardımlaşma ihtiyaçları oluşmuştur. Bu unsurların başında ise ortak üretim ihtiyaçları bulunmaktadır. Maslow ihtiyaçlar hiyerarşisindeki gereksinimlerin karşılanması amacı ile çeşitli ortaklıklar, yardımlaşmalar ve ikameler oluşmuştur. Paranın bir mübadele aracı olarak kullanılacağı dönemlere kadar takas yöntemleri ile karşılıklı alışveriş türleri oluşmuş ve geleneksel Pazar anlayışının temelleri atılmıştır.

Perakende sektörü olarak ekonomide yer alan bu faaliyet alanı, zaman içerisinde çeşitli evrimler geliştirmiştir. Böylece ödeme sistemlerinin ve müşteri iletişim kanallarının artması ile rekabetin en yoğun olduğu sektör halini almıştır.

Müşterilere her kanaldan satış fırsatı sunmak ve bunları en verimli şekilde yapmak özellikle küresel rekabetin yoğun olduğu bir pazarda en önemli verim arttırıcı metrikler haline gelmiştir. Müşterilerin en yoğun olduğu konumları seçmek veya müşterileri için mağazasını bir çekim merkezi haline getirmek perakendenin en büyük yatırım kalemleri arasındadır.

Faaliyette bulunulan lokasyonlar da insanların değer yargıları, tutum ve davranışları, kültürel alışkanlıkları, tatil günleri vb. gibi veriler müşterilerin sizi tercih edip etmeyeceğinin en önemli göstergeleridir. Sektörün büyüme hızı, rekabet durumu, karlılık derecesi, belirsizlikler ve riskler veriye dayalı karar verme stratejileri uygulayan firmalar için birinci derece araştırma konuları arasına girmiştir.

Perakende sektörüne baktığımızda, özellikle Türkiye'de oldukça değişen dinamik bir Pazar yapısı bulunmaktadır. Bu dinamik yapıya çevik bir şekilde uyum sağlamak firmaların sürdürülebilirlik ve rekabet ile ilgili stratejik planları arasındadır. Bu nedenle firmalar veriye dayalı kara vermek için değişen pazar dinamiklerini ve nüfus demografisini göz önünde bulundurmakta, bunun sonucu olarak da mağaza açma, kapama, birleştirme gibi kararları belli sıklıkta vermek zorunda kalmaktadırlar. Bu tür kararlar detaylı analizler gerektirmektedir. Müşteri profili ve pazar demografisi incelenmeli, ayrıca coğrafi unsurlar da dikkate alınmalıdır. Bu temel parametrelere dikkat edilmeden verilmiş kararlar geri dönüşü olmayan kayıplara yol

açabilmektedir. Ayrıca problemi zorlaştıran bir diğer faktör rakip mağazaların yerleri ve rakip firmaların büyüme stratejilerinde verilen kararların doğruluğunu sorgulanır hale getirmektedir.

Bu çalışmanın temel amacı ana hatlarıyla tarif edilen yer seçimi problemini sistematik ve nicel bir yaklaşımla modellemek. Bu amaçlar doğrultusunda Türkiye'de perakende sektöründe faaliyette bulunan bir firmanın mağazalarının bulunduğu lokasyonların il, ilçe ve çevre verileri incelenmiştir. Veri analizleri ile yer seçimi içim en uygun parametreler belirlenmiştir. Belirlenen parametreler referans alınarak Makine Öğrenme ve Birleşik Makine Öğrenme yöntemleri kullanılarak yer seçimi için en uygun noktanın tahminlenmesi çalışması yapılmıştır.

2. GENEL KISIMLAR

2.1.PERAKENDE SEKTÖRÜ

2.1.1. Perakende Nedir?

Perakende, katma değer çıktısı olan üretim veya hizmet çıktılarının, çeşitli satış kanalları ile parsiyel formlarda, müşterilerine sunulduğu bir finansal sektör türüdür. Türk Dil Kurumu, perakendeyi "düzenli olmayan, ayrı ayrı, dağınık, perişan, malların teker teker veya birkaç parça durumunda azar azar satılmasına dayanan (satış biçimi), toptan karşıtı" şeklinde tanımlamaktadır (Türk Dil Kurumu, 2018).

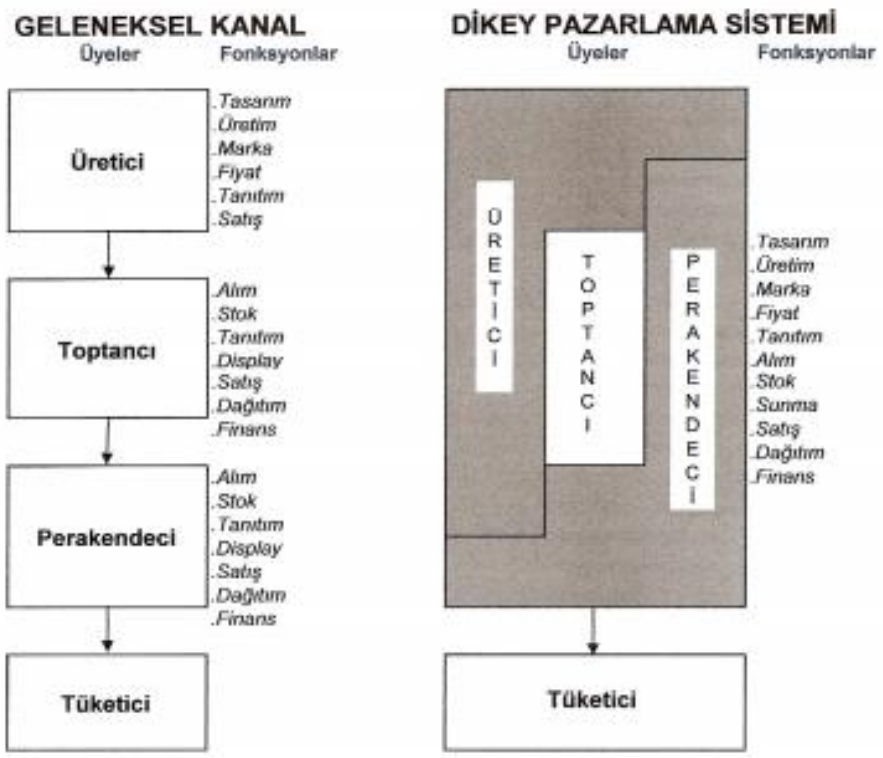
Perakende satış, ürünlerin bir işletme aracılığıyla tek tek ya da küçük miktarlarda tüketiciye satılmasına dayanan satış biçimidir.

Amerikan Pazarlama Derneğine (American Marketing Association) göre ise perakendecilik kavramı "Mal ve hizmetlerin kişisel, ailesel ya da ev kullanımı amacıyla satışı için; herhangi bir mağazada yapılan ya da mağazasız bazı formlarda yapılan bir takım işletme faaliyetleridir" şeklinde tanımlanmıştır (American Marketing Association, 2018).

Tüketim mallarının dağıtımı, dağıtım kanalının ilk üyesi olan üretici ile başlar ve nihai tüketici ile sona erer, ancak çoğu zaman arada en az bir aracı vardır ki o da perakendeci olmaktadır. Perakendecilik, mal ve hizmetlerin kişisel kullanımı için doğrudan doğruya nihai tüketicilere satışı ile ilgili tüm faaliyetleri kapsamaktadır (Mucuk, 1994).

Perakendecilik; tüketicilerin, kişisel veya ailevi ihtiyaçları için mal ve hizmet ya da her ikisini satarak değer yaratan faaliyetler kümesidir. Birçok insan perakendeciliğin, sadece mağazalarda ürün satışı yapmak olduğunu düşünmektedir. Fakat otelde konaklama hizmeti almak, kuaföre tıraş olmak, doktorda tedavi olmak veya pizza siparişi vermek gibi faaliyetler de perakendecilerin yerine getirdiği faaliyetlerden bazılarıdır. Ek olarak, perakendecilik sadece mağazalarda yapılmamaktadır. Kataloglar veya internet üzerinden yapılan satışlar da perakendeciliğe girmektedir (Levy ve Weitz, 2004).

Bir başka çalışmada geleneksel dağıtım kanalındaki üretici, toptancı ve perakendeci arasındaki bariyerlerin yıkıldığını belirtmektedir (McGoldrick, 1990). Perakende kanal fonksiyonlarının değişimi Şekil 2.1'de gösterilmektedir.



Şekil 2.1: Perakende kanal fonksiyonlarının değişimi (Alpay, 2006).

2.1.2.Perakende Sektörünün Tarihsel Gelişimi

İlk çağlardan itibaren insanların hayatlarını idame ettirebilmeleri için çeşitli sosyal faaliyetlerde bir araya gelme ve yardımlaşma ihtiyaçları oluşmuştur. Bu unsurların başında ise ortak üretim ihtiyaçları bulunmaktadır. Maslow ihtiyaçlar hiyerarşisindeki gereksinimlerin karşılanması amacı ile çeşitli ortaklıklar yardımlaşmalar ikameler oluşmuştur. Paranın bir mübadele aracı olarak kullanılacağı dönemlere kadar takas yöntemleri ile karşılıklı alışveriş türleri oluşmuş ve geleneksel Pazar anlayışının temelleri atılmıştır. Başlangıcı ilkel çağlara dayanan takasları gerçekleştirebilmek için takas olacak gereksinimlerin değerini ifade eden bir ölçü ortaya konması gerekliydi. Bu duruma örnek olarak ilk çağlarda metaller, işlenmiş parçalar (kazan, saç ayak, şiş bel) ya da külçeler verilebilir.

Zaman geçtikçe üretilen malların çeşitliliğinin artması ve insanların taleplerinin çoğalması ile takas yöntemi güncelliğini yitirmeye başladı. Ticaretin kolay yapılabilmesi için zaman içerisinde metal külçeler standart hale getirilerek ilk para icat edildi. Tarihte aynı ağırlıkta parçalar biçiminde metal birimlerin üretilmesi düşüncesi, İ.Ö VII. –VI. yüzyıllarda Çin'de, Hindistan'da ve Küçük Asya'da ortaya çıktı.

Coğrafi gelişmeler sayesinde farklı coğrafyada yer alan değişik ürünleri tanımaya başlayan insanlar bu ürünleri satın alma eğiliminde bulunmaya başladılar. Bu eğilim sayesinde ihtiyaçlar karşılanarak ticaret ve üretimin gelişmesi sağlanmıştır. İhtiyaç duyulan malların alınabilmesi için belirli satın alınabilecek yerler (dükkân, çarşı, pazar, agora vb.) oluşturulmuştur. Avrupa'da derebeylik sisteminin değişmesi ile şehirleşmenin önemi artmış ve İngiltere'de bu gelişme Kent Rönesans'ına sebep olmuştur. Şehirleri, yönetim ve savunma amaçlı geliştirmişlerdi. Fakat, sonradan ticaret aktiviteleri ağırlık kazanmıştır. Derebeylik sisteminin değişmesi ile şehirler, vergiden kurtulmanın yollarını aramışlardır. Serbest kalan şehirlerde ticaretin düzenli yapılabilmesi için çeşitlik teşkilatlar (borough) kurulmuştur. Bu teşkilatlar, lonca sisteminin kurulmasına öncülük etmişlerdir (Davies ve Ward, 2002).

Bugünkü gelişmiş araştırma metotları ve bilgisayar destekli karar verme tekniklerinin kullanıldığı Amerikan perakendeciliğinin kökleri XV. yüzyılda posta ticareti ile atılmıştır. O günlerde, satın almalarda para kullanılmamaktadır (Merçan, 2007).

Perakende ticaretin gelişimini incelerken, uluslararası ve ulusal ticaret kanunlarının anlaşılmasını kolaylaştıracağını düşündüğümüz, sınırlı sorumluluk (Limited liability) prensibinden de söz edilmesi gerekmektedir. Davies & Ward (2002), sınırlı sorumluluğun İngiliz hukukunda XV. yüzyıldan itibaren var olduğunu belirtmişlerdir. Daha sonra, İngiltere ve diğer ülkelerde sıklıkla, denizaşırı şirketlere (chartered company), tekel lisansı (monopoly licence) olarak verilen bir hak için kullanılmıştır. Bu haklar ile denizaşırı şirketler geniş haklara sahip olmuşlardır. The Dutch East India Firması 1602 yılında, Afrika ve Macellan Boğazı arasında ticari tekel garantisi ile kurulmuştur. Davies ve Ward, bu uygulamanın XIX. yüzyılda da örnekleri ile karşılaşıldığını, Heal's ve Harrods şirketlerinin de bu şekilde kurulan ve çok hızlı gelişme gösteren şirketler olduğunu belirtmişlerdir.

Davies ve Ward (2002), XIX. yüzyılın son yarısında, çok katlı yapısı olan Bon Marche'nin Fransa'da Paris'te açılmasından sonra Londra ve Şikago şehirlerinde de açıldığını belirtmiştir.

Bu mağazaların çok yükümlü ve dışarıdan gelen (exotic) bir alışveriş ortamı sunabilmek üzere düzenlediklerini belirtmiştir. Ancak 1900'lere gelindiğinde Dünya ticaretini derinden değiştirecek gelişmeler yaşanmıştır. XX. yüzyıl, 1914'de I. Dünya Savaşı ile başlayan ve neredeyse her on yılda bir dünyayı derinden etkileyen toplumsal olayların, savaşların, ekonomik krizlerin yaşandığı bir yüzyıl olmuştur. Bu yüzyılda ekonomik, sosyal değişimler ve Endüstri Devriminin de etkisi ile büyük şehirlere göç hızlanmıştır.

1908 yılında Amerika'da Ford-T ile seri üretime başlanması ile eski ekonominin yapısı değişmiş ve otomobil endüstrisinde düşen imalat maliyetleri her eve bir araba alınabilmesini sağlamıştır. Fordizm ve Taylorizm ile gelişen seri üretim neticesinde imalatçı firmaların elinde stoklar oluşmaya başlamıştır. Böylece stokların eritilmesi için klasik satış elemanlarının yanında farklı pazarlama teknikleri geliştirilmeye başlanmıştır. Bu pazarlama stratejileri ve teknikleri sonucunda farklı müşteri profilleri ve ihtiyaçları doğrultusunda mağazacılık sistemleri oluşturulmuştur.

XIX. yüzyılın ortalarında Birleşik Devletlerde çok çeşitli ürünler üretilmiş ve genel mağazalar, fabrikaların limitsiz mal arzına hizmet veremez noktaya geldiğinde perakendecilikte uzmanlaşma gerekmiştir. Böylece sınırlı ürün yelpazesine sahip mağazacılık (limitedlinestore) başlamıştır. Bu mağazalarda ürün kategorisinin geniş çeşidinin satışa sunulması esas alınmıştır. Bunlara örnek olarak, ayakkabı mağazaları, bayanlara yönelik mağazalar, mücevher mağazaları ve gıda mağazaları gösterilmiştir (Merçan, 2007).

1950'deki Avrupa perakendesi, savaş malları stoklarının ve iş gücü sıkıntısının yükseldiği, savaşın harap edip yıkmış olduğu fabrikaların olduğu bir perakendecilik dönemi olarak adlandırılmıştır. Bugün geçerli olan perakendeciliğin tohumları savaş sonrası kalkınma girişimleri sırasında ekilmiştir (Dawson, 2000).

2.1.3.Perakende Dönüşümü

Geleneksel perakende birçok faktörün tehdidi altındadır. Mevcut zincir mağaza yapıları, elektronik ticaret, çevresel etkenler, finansal kriz ve dinamik nüfus yapısı bu faktörlerin başlarında gelmektedir. Tüm olumsuz etkenlere rağmen çok sayıda insan geleneksel perakendenin gerçek dünyası diye tanımlayabileceğimiz mağazadan satın alma yöntemi şeklindeki alışverişten hala hoşlanmaktadır. Perakende sektörünün öncüleri çeşitli satış kanallarının tehdidini fırsata çevirme telaşı içerisindeler. Müşterilerin insanlarla ilişki kurma

ihtiyacı, anında tatmin olma arayışı, yaşanan deneyimleri hikâye olarak anlatma gereksinimleri gibi davranış modelleri geleneksel perakendeciliğin mağazacılık kültürünü güçlendirmiş ve aynı zamanda evrimleştirmiştir.

Trendwatching perakende sektörünün dönüşümünü dört ana etkende sıralamıştır.

1. OFF=ON: Müşterilerin sadece online alışverişleri sırasında sahip oldukları birtakım avantajlara, artık geleneksel alışverişte de sahip olmaları.

2. RETAIL SAFARI: Müşteriler tarafından gerçek deneyimlerin saltanatının devam ediyor oluşu.

3. INSTANT STATUS FIX: Geleneksel alışverişin anında sağladığı statü tatminini online olarak yapılan alışverişin hâlâ sağlayamıyor oluşu.

4. CITYSUMERS: Tüketiciliğin geleceğinin şehirlere bağlı olması ve şehir kültürünün, dünyanın her yerinde, perakende kültürü ile neredeyse eş anlamlı kabul edilmesi.

2.1.3.1. OFF=ON

Gerçek dünyanın giderek egemen hale gelmekte olan sanal dünyaya adapte olma çabasını anlatan OFF=ON isimli trend geleneksel perakendeciliğinin alışveriş deneyimlerini değiştirmiştir (“Retail Renaissance”, 2011).

Trendwatching bu değişimi şu şekilde açıklamaktadır. “Online dünya, bir dizi yepyeni iş yapış biçimi ve iş sürecinin ortaya çıkmasına neden olurken müşterinin sürece yeni dahil olma şekli ve yeni pazarlama teknikleri, tüketicilerin sanal dünyanın sağladığı olanakların tadını çıkartmasının önünü açmaktadır. Bundan böyle fiziksel dünya online dünyada yaşanan gelişmelere göre adapte olmak zorunda” (“Retail Renaissance”, 2011).

Her 10 tüketiciden 8’i satın alma faaliyetinden önce internette araştırma yapmaktadır. Bu tüketicilerin %42’si internette araştırma yaptıktan sonra online olarak, %51’i ise internette araştırma yaptıktan sonra mağazadan satın alarak alış veriş faaliyetini gerçekleştirmektedir (Google/IPSOS OTX MediaCT, 2011).

Alışveriş yapmadan önce birden fazla noktadan (mağaza, internet, mobil cihaz veya katalog) bilgi alan çoklu kanal deneyimlerine açık olan tüketiciler, sadece mağazadan alışveriş yapan

tüketicilere oranla satın alma faaliyeti başına %82 oranında daha fazla para bırakmaktadırlar (Deloitte, 2018).

Hazır giyim sektöründe, elektronik ticarette ziyaretin satın alma faaliyetine dönüşme oranı %2-3,5 civarında iken, aynı dönemde geleneksel mağazaların müşteri dönüşüm oranları %20-25'lerdedir (Vedat ve ÖZBEK, 2014).

Tüketicilerin %40'ının akıllı telefon sahibi olduğu Amerika'da, bu tüketicilerin %70'i akıllı telefonunu mağazada alışveriş yaparken kullanmaktadır. Akıllı telefonu olan tüketicilerden %74'ü akıllı telefonunu kullanarak alışveriş yapmıştır. Bu kişilerin %76'sı geleneksel mağazalardan, %59'u internetten ve sadece %35'i ise akıllı telefonu üzerinden bir alışveriş faaliyeti gerçekleştirmiştir (Google/IPSOS OTX MediaCT, 2011).

2010 yılında tüm dünyada mobil barkod tarama işlemlerinde (geleneksel UPC barkodları ve QR kodları dahil) %1600 oranında artış gözlemlenmiştir ("Retail Renaissance" 2011).

2.1.3.2. Perakende Safarisi

Müşterilerin internette alışveriş yaparken yaşamayacakları veya elde edemeyecekleri şeylere karşılaşma beklentisi artmış durumdadır. Mağazalarda ki ilgi çekici gösteriler, o mağazada satılan özel ürünler, ürünleri deneme şansı veya ürünleri ne şekilde kullanacağını öğrenme fırsatı gibi deneyimler geleneksel perakende de farklı formlarda yerini almış durumdadır ("Retail Renaissance", 2011).

2.1.3.3. Anında Statü

Müşteriler internet alışverişindeki teslim sürelerinden etkilenmek istemiyorlar. Aldıkları ürünlerdeki etkileşimi ve deneyimi anında yaşamak istiyorlar. Bir mağaza personelinin kendileri ile ilgileniyor olması bir statü beklentisi hissini tatmin etmektedir ("Retail Renaissance", 2011).

2.1.3.4. Citysumers

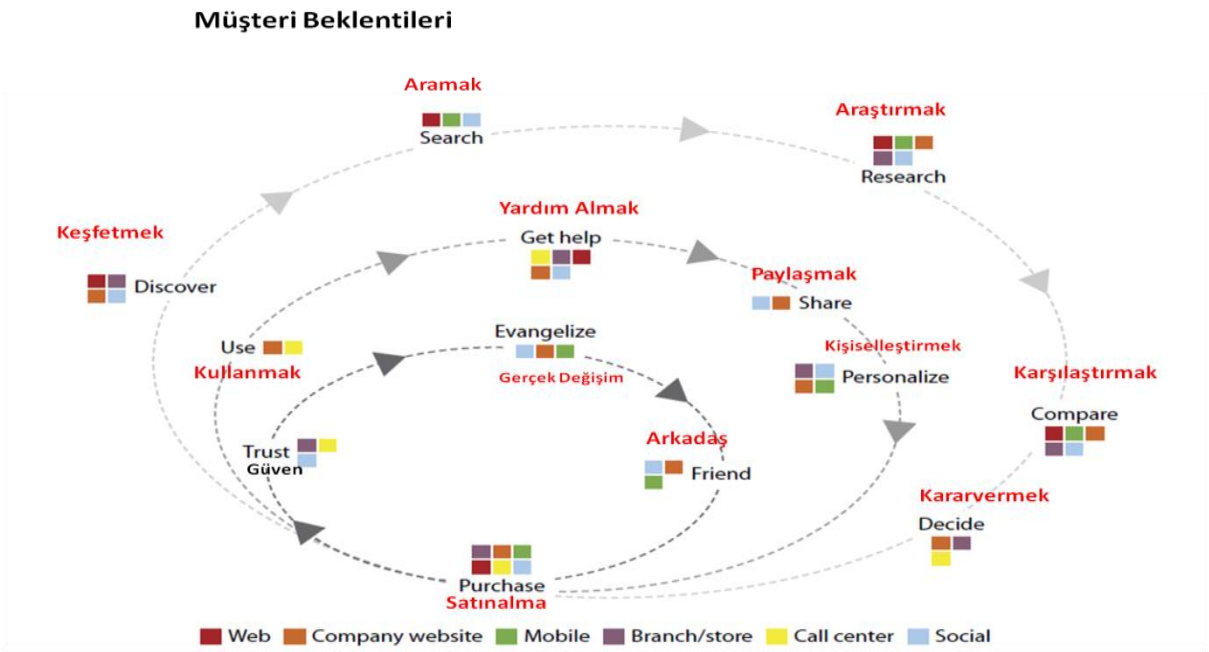
Bu trend başlığı şehirlerde yaşayan nüfusun tüketim alışkanlıklarının fazla olması ve perakende firmalarına ulaşımının kolay olmasının sağladığı avantajları kullanan bir müşteri profilinden

ortaya çıkmıştır. Bu yeni alışveriş demografisi mağazaların nerelere açılması ve online deneyimlerle desteklenmesi gerekliliğini ortaya koymuştur.

Asya'nın perakende satış hacminin 2010'da 5,4 trilyon USD iken 2014'de 8,5 trilyon seviyelerine çıkmıştır. Kuzey Amerika ve Batı Avrupa'nın perakende satış hacimlerinin ise aynı dönemde sırasıyla 4,5 trilyon USD ve 2.7 trilyon USD olarak gerçekleştiği gözlemlenmiştir (Merçan 2007).

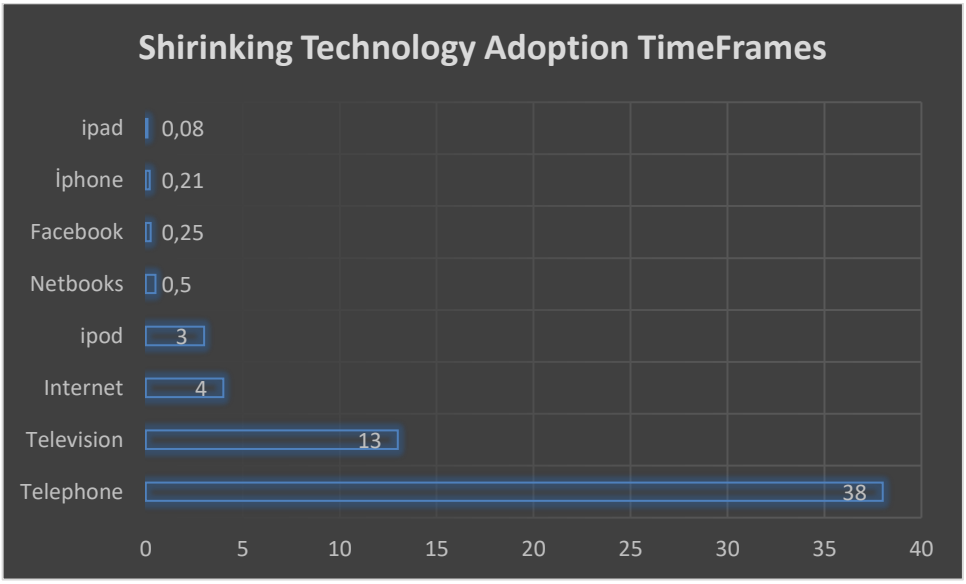
2.1.4. Perakende Kanal Tipleri

Müşteriler ürünlere her kanaldan ulaşma isteği eğilimdedirler. Şekil 2.2'de olduğu gibi müşterilerin bir satın alma deneyi mi için tatmin edilmesi gereken temel ihtiyaçları bulunmaktadır. Tüketiciler yeni ürün ve hizmetleri keşfetmek, tartışmak ve satın almak için artan oranda ve daha karmaşık yöntemler kullanarak, mevcut arkadaş grubu, fan ve takipçi ağlarına başvurmaktalar (Sheth, Mittal, ve Newman, 1999).



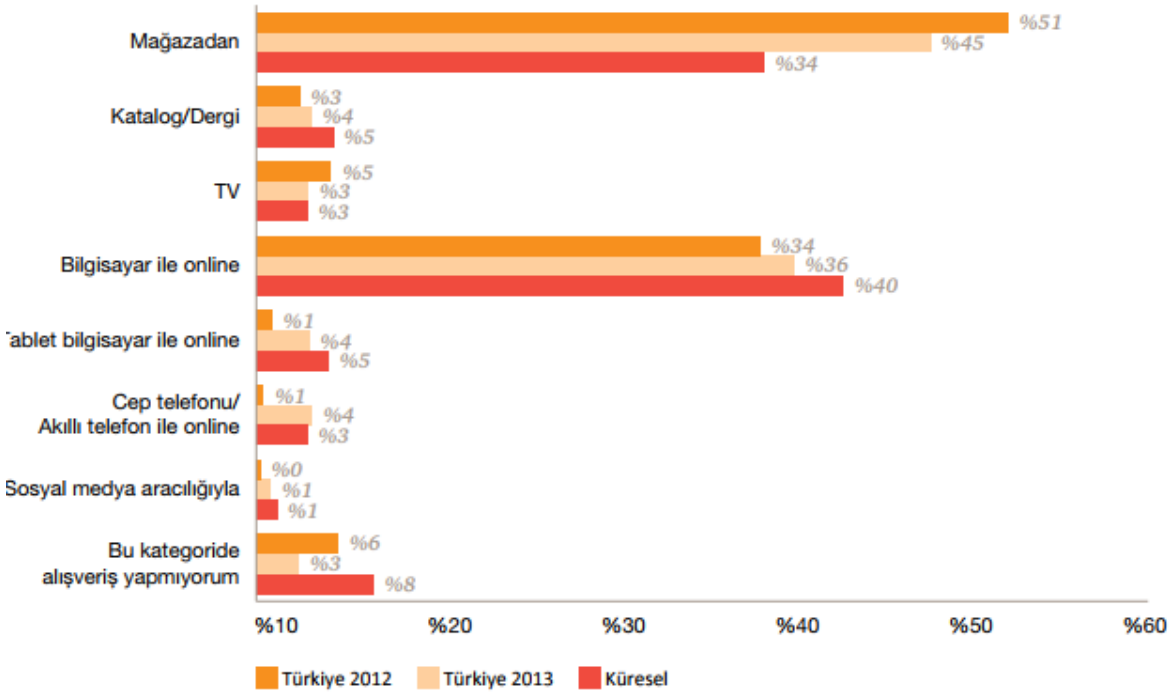
Şekil 2.2: Müşteri davranış modeli

Şekil 2.3'te gelişen teknolojilere adapte olma süresi gösterilmiştir. Buna göre gelişen teknolojilerle birlikte müşterilerin adaptasyon süreleri de kısalmış bulunmaktadır. Aynı zamanda telefona olan adaptasyon süresi 38 yılı bulurken artık yeni gelişen teknolojilere adepte olma süresi günlerle ifade edilmektedir.



Şekil 2.3: Gelişen teknolojilere adapte olma süresi (comScore Dünya ve Türkiye Elektornik Pazar yeri trendleri raporu)

PWC firmasının yaptığı araştırmaya göre Türkiye ve Dünya'da giyim sektöründe alışveriş kanallarının tercih edilme dağılımı Şekil 2.4'te gösterilmiştir.



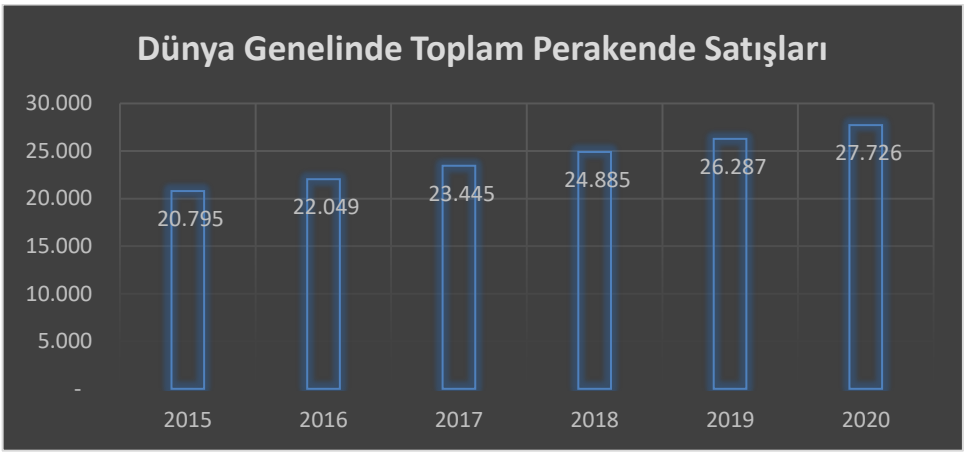
Şekil 2.4: Giyim sektöründe Kanal Tipleri dağılımı (PWC, 2013).

2.1.5. Dünyada Perakende Sektörü

Teknolojik gelişmeler, küresel ekonomideki değişim, hızlanan şehirleşme, iklim değişimi ve demografik değişiklikler gibi mega trendlerin diğer sektörlerde olduğu gibi perakende sektöründe de yarattığı yoğun dönüşüme ışık tutmaktadır. Perakende sektörü, değişen tüketici taleplerine karşı zamanlı ve doğru bir şekilde cevap verebilmek için hızlı olmak ve farklı satış kanallarını koordineli bir şekilde yürütmek zorundadır. Bu nedenle bilgi teknolojileri ve gerekli alt yapı yatırımları hayati önem taşımaktadır. Teknolojik gelişmeler, iletişimin çeşitlenmesi, tüketicilerin zaman darlığı ile birlikte elektronik ticaret gün geçtikçe önem kazanmaktadır. Entegre teknoloji platformları, müşteri analitiklerinin oluşturulmasına olanak sağlayan veri depolama sistemleri, teknolojik yatırımlar ile tedarik sistemlerinde yapılan verimlilik uygulamaları gelecekteki perakende iş modelinin temel yapı taşlarını oluşturmaktadır (Besfin, 2018).

Yıllara göre perakende gelişimi Şekil 2.5' te gösterilmiştir. Dünya çapında perakende satışlarının, e-ticaret de dahil olmak üzere 2017 yıl sonunda 23.445 milyar dolar olarak gerçekleştiği tahmin edilmektedir. E-ticaret sektörü her yıl, geleneksel perakende sektöründen

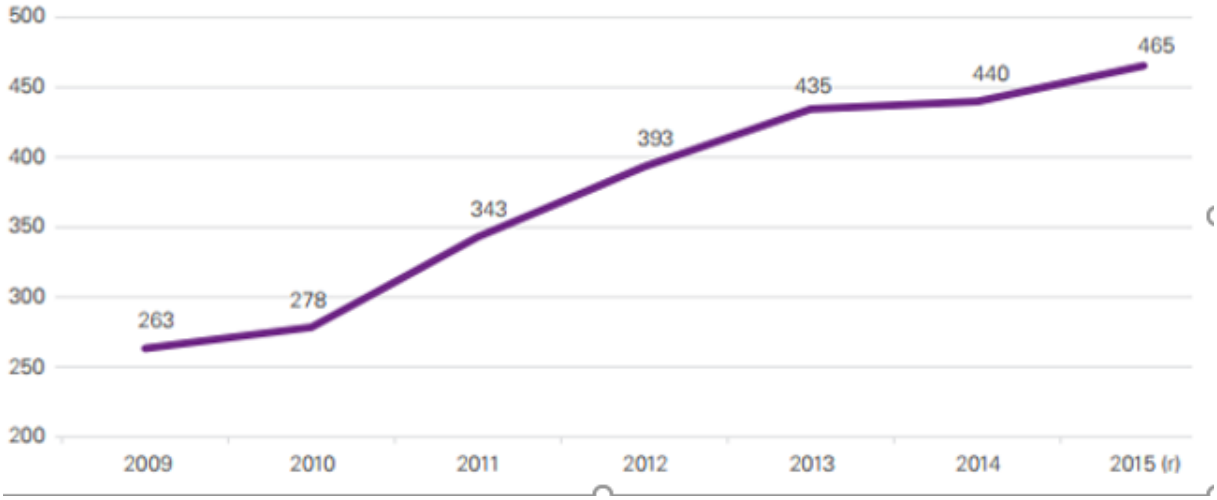
yaklaşık yüzde 1 oranında pay almaktadır. Perakende e-ticaret satışları 2017 yılsonunda dünya çapında 2.350 milyar dolara ulaşacaktır. Bu oran toplam perakende satışlarının yüzde 10,1'ini oluşturmaktadır. Bu noktada, küresel e-ticaretin toplam perakende satışları içerisindeki payı 2011 yılında yüzde 3,6 iken 2016'da bu oranın yüzde 8,7'ye yükseldiğini belirtmek gerekmektedir (KPMG, 2017).



Şekil 2.5: Dünya geneli perakende sektörü (KPMG, 2018)

2.1.6. Türkiye'de Perakende Sektörü

Türkiye perakende sektörünün yaklaşık %67'sini geleneksel perakende, %33'ünü ise organize perakende oluşturmaktadır. Organize perakende pazarı özellikle son 15 yılda büyük bir büyüme göstermiş olmasına rağmen, klasik esnaflık biçimi geleneksel perakende sektöründe hala önemini korumaktadır. Buna rağmen organize perakende sektörü büyüme trendini korurken, geleneksel sektörde düşüş yaşanmaktadır. Bu değişimin birkaç nedeni arasında; AVM yatırımlarında yaşanan büyük artış, geleneksel perakendenin içinde bulunduğu rekabet sorunu, müşteri davranışlarındaki değişim ve alışveriş alışkanlıklarını kökünden sarsan teknolojik yenilikler olarak sayılabilir. Bir başka önemli faktör de oldukça agresif büyüyen indirim marketlerinin, önemli buldukları lokasyonlarda satın alma veya devralma yolu ile geleneksel sektör aleyhinde büyüyor olmalarıdır (KPMG, 2018).



Şekil 2.6: Perakende satış ciroları (TÜİK, 2017)

Şekil 2.6'da 2009 ve 2015 yılları arasında Türkiye'deki perakende satış cirolarının değerleri gösterilmiştir. 6 yıl içerisinde ortalama %77 büyüme göstermiştir. Mağaza, metrekare ve sepet verimlilikleri doğru analiz edilip, yatırımlar bu yönde yapılmadığı sürece sektör, içinde bulunduğu bu sıkıntılı durumu mevcut yöntemlerle aşamayacaktır. Lokasyon neresi olursa olsun rakiplerin olduğu yerlerde mağaza açma eğilimi, yüksek kira giderleri ve fiyat dengesizlikleri sonucunda nakit akışı problemleri artırmaktadır. Fiyat indirimi ve kampanya stratejilerinin tutarsızlıkları, uzun vadeli nakit akışı problemine ve müşteri sadakati sorunlarına neden olmaktadır (KPMG, 2018).

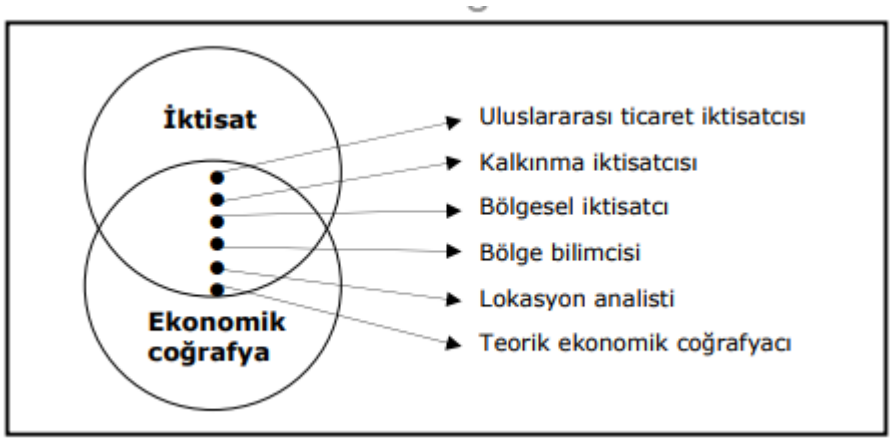
Mevsim etkisinden arındırılmış cari fiyatlarla 2017'nin hemen hemen tüm aylarında perakende ciro endeksi, bir önceki yıla göre artış göstermiştir. 2017 yılına, bir önceki yılla karşılaştırıldığında daha düşük bir satış hacmi ama daha yüksek bir ciro ile giren perakende sektörü, 2016'nın aynı dönemiyle karşılaştırıldığında %11 ciro büyümesi yakalamıştır. Dövizle ödenen kiralar ve asgari ücretin artışıyla yükselen iş gücü maliyetleri, sektörde ciro artışına rağmen kârlılığın baskı altında kalmasına neden olmuştur.

2.2. LOKASYON ANALİZİ VE YER SEÇİMİ

İnsanlığın faaliyet alanlarından en önemlisi olan ekonomik coğrafya, ticari faaliyetle ilişkili mal ve hizmetlerin üretim, tüketim ve dağıtımının mekânsal açıdan anlamlandırmaya ve açıklamaya çalışır. Bu nedenle ekonomik coğrafyanın ilgi alanı, ticari ürün ve hizmet faaliyetlerinin üretim, lojistik ve tüketim yerleridir. Bir farklı tanım ile ekonomik coğrafya, lokal, bölgesel ve bölgeler

arası boyutlarda ki ticari faaliyetlerin pazar alanlarını inceleyen bir araştırma alanıdır. Bu yönüyle ekonomik coğrafya, ürün ve hizmet faaliyetlerinin üretim, tüketim ve değişim davranışları ile ilişkili ticari aktivitelerin bölgelere göre farklılaşmasını anlamlandırmaya çalışır (Yavan, 2006).

Taafle ekonomik ve iktisadi coğrafyanın dağılımını Şekil 2.7'deki şema grafiği ile göstermiştir (Taaffe ve Gauthier, 1994).



Şekil 2.7: Ekonomik ve iktisadi coğrafyanın dağılımı (Taaffe ve Gauthier, 1994)

Bazı araştırmalarda lokasyon, bir şirketi bir şehire veya bir noktaya çeken parametrelerin anlamlı bir bileşiği (kombini) olarak tanımlanmaktadır. Bazılarında ise lokasyon lojistik faaliyeti maliyetlerinin önemli bir etkeni olan bir noktanın diğer bir noktaya olan mesafesi ve bu mesafenin etki alanları olarak ele alınmaktadır. Bunlara ek olarak lokasyonu, geleneksel fiziki mekân tanımının ötesinde, pazar alanları arasındaki göreceli kültürel yakınlık şeklinde tanımlayan araştırmalarda bulunmaktadır (Yavan, 2006).

Lokasyon Analizi ve Yer Seçimi birçok sektör de karşılaşılan temel problemlerden bir tanesidir. Mağaza yeri, depo yeri, üretim yeri, banka şubesi gibi fiziksel bir lokasyona tabi tüm faaliyet kollarında yer seçimi için bir takım temel kriterleri kontrol etmeleri gerekmektedir.

McCann ve Mudambi lokasyon analizi ile ilgili doğrudan yabancı yatırımcılar için iki farklı çalışma yürütmüşlerdir.

- Bölgeler arasındaki lokasyon seçimi (inter-countrylevel),
- Bölge içindeki lokasyon seçimi (subnationalgeographicalscale).

Görüldüğü gibi firmalar lokasyon seçimini birkaç farklı coğrafi boyutu analiz ederek yapmaktadır. Bu kararlar dikkate alındığında lokasyon yeri seçimi, bina seçimine karar verene kadar hiyerarşik olarak ülke, bölge, kent ve kentin içindeki herhangi bir noktanın seçilmesini ve bu seçilme nedenlerini içeren bir karar alma faaliyeti sürecidir (Yavan, 2006).

2.2.1.LokasyonTeorisi

Lokasyon teorisi çok uzun bir araştırma geçmişini içerir. Bu nedenle lokasyon teorisi ekonomik coğrafya ve bölgesel iktisadın temellerini oluşturan bir dizi çalışmalara öncülük yapmıştır. Lokasyon teorisi araştırmacılar üzerinde büyük bir yankı oluşturmuştur. Böylece çevre bilimi denilen yeni bir bilim dalının oluşmasına yol açmıştır. Lokasyon Teorisi ekonomik faaliyetlerin nerede yer aldığı ile ilgilenir. Ekonomik kararların coğrafi etkisini ve insan davranışlarını analiz etmektedir. Böylece lokasyon analizinde nerede sorusunun cevabını oluşturan, konum (space) ve yer (place) büyük bir öneme sahiptir. (Hoover, 1937; North, 1955)

Lokasyon teorilerinin açıklamaya cevap aradığı temel sorular, bir firma veya hane halkı/bireyin bir lokasyonu seçmesini sağlayan temel faktörler/motivasyonlar nelerdir? Lokasyon kararının/seçimin toplumsal, ekonomik, kültürel, siyasal ve mekânsal sonuçları ve uygulamaları hangisidir? sorularıdır. Böylece teori sadece firmaların değil aynı zamanda bireylerin/hane halklarının hatta devletin/kamunun yer seçim davranışları ve bunların sonuçları ile uğraşmaktadır (Yavan, 2006).

Lokasyon teorisi, üç büyük klasik kurucusunun çalışmalarına dayanmaktadır. Bunlardan ilki, tarımsal lokasyon teorisinin öncü araştırmalarını yapan VonThünen teorileridir. Diğeri, ekonomik lokasyon teorisini tanımlayan Alfred Weber'in teorisdir. Sonuncusu ise bir ekonomik sektör merkezi olarak kurulmuş şehir ve kentlerin konumlarını açıklamaya çalışan Christaller-Lösch teorisidir.

2.2.1.1.Weber'in En Az Maliyet Teorisi

Weber'in temel teorisi bir firmanın tesis yeri için seçilebilecek en uygun yer, üretim maliyetlerinin en az olduğu yerdir. Weber yer seçimindeki karar parametrelerinin karmaşıklığını azaltmak için aşağıdaki temel varsayımları kabul ederek analizlerine başlamıştır (Weber, 1929).

- Ekonomik faaliyet bölgelerinde tam rekabet koşulları hakimdir.
- Hiçbir firma lokasyon yer seçiminden dolayı monopolistik avantaja sahip olamaz.
- Tüm kararlar ekonomik rasyonel göstergelere göre verilmektedir.
- Düz ve geniş bir alan söz konusudur.
- Tüketim pazar yerlerinin büyüklüğü ve lokasyonu bellidir.
- Piyasa sadece özel yerlerde mevcuttur.
- Hammaddeler sadece belli yerlerde bulunmaktadır.
- İşgücü sadece birkaç lokasyonda bulunmaktadır. Buralarda da işgücü göç etkisinden uzak ve hareketsizdir fakat üretim yerlerinde sınırsız emek arzı vardır.

Dicken (Dicken, 1990), Weber'e göre hammaddeleri aşağıdaki şekilde açıklamıştır;

- Lokasyonu etkilemeyen girdiler (Ubiquitousmaterials/inputs).
- Belirli yerlerde bulunan ve lokasyonu etkileyen lokal girdiler (localizedmaterials/inputs).
    - Hurda çıktısı az olan ağırlığı değişmeyen girdiler (purelocalizedmetarials).
    - Hurda çıktısı çok olan ve ağırlığı değişen girdiler (impure/grosslocalizedmaterials).

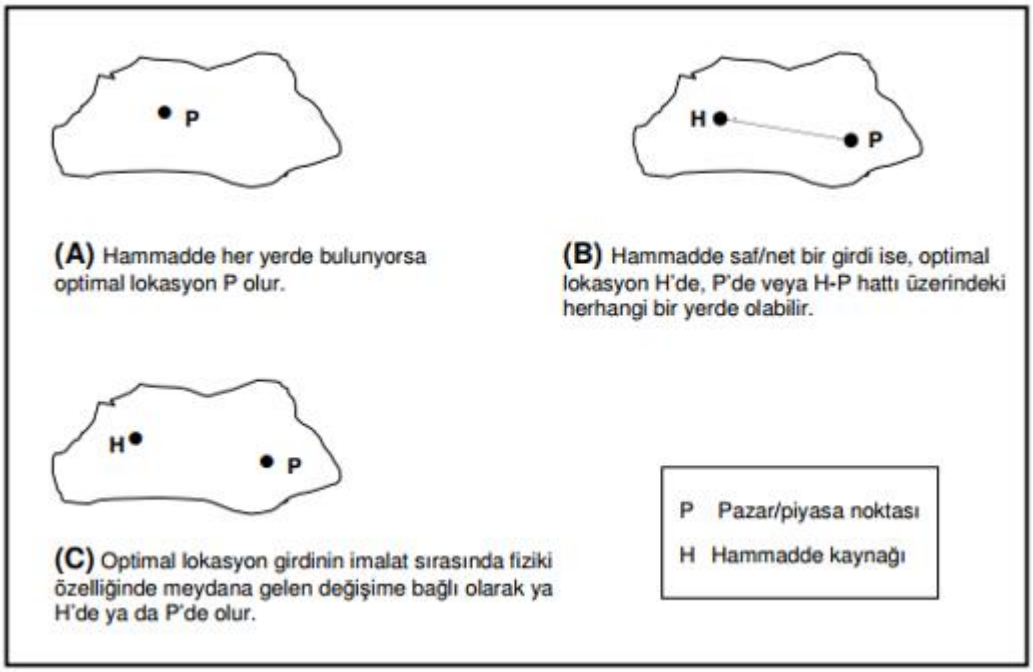
Weber bir hammadde indeks formülü geliştirmiş ve tesisin nereye kurulması gerektiğini açıklamaya çalışmıştır

$$Hammadde\ Indeks - HI = \frac{Hammaddelerin\ Ağırlıkları\ topamı}{Ürünlerin\ ağırlıkları\ toplamı}$$

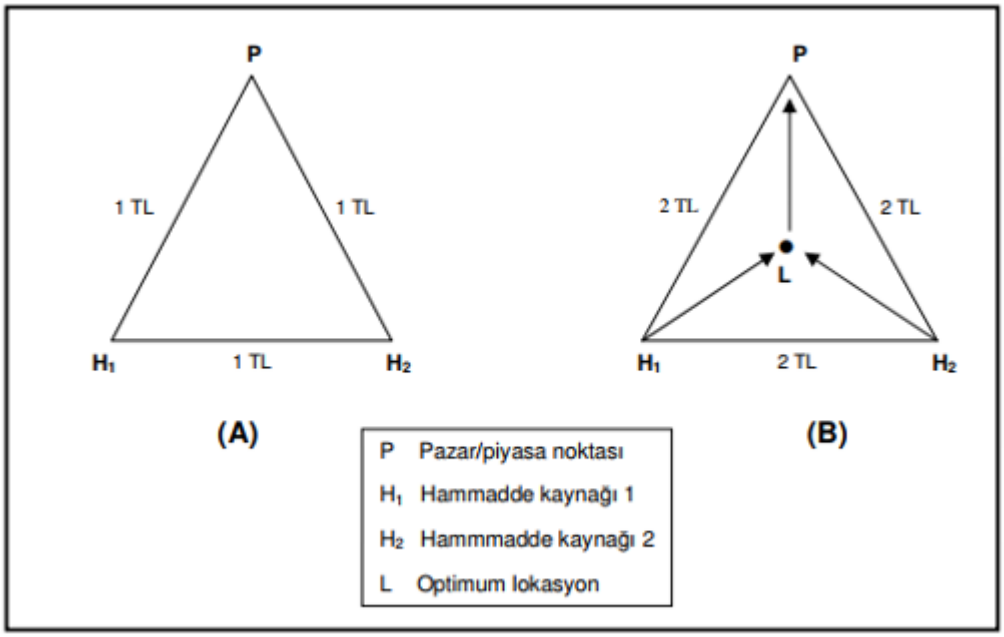
$$\mathbf{HI} > 1\ kuruluş\ yeri\ hammadde\ kaynaklarına\ yakın$$

$$\boldsymbol{HI} < 1\ kuruluş\ yeri\ ekonomik\ faaliyet\ alanlarına\ (pazar\ yerlerine)\ yakın$$

Weber'e göre "lojistik maliyeti" bir firmanın lokasyon seçimini etkileyen en önemli parametredir. Weber'in temel teorisi, "Bir firma lojistik maliyetlerinden taşıma maliyetini minimum seviyeye indirdiği yerde tesisini açma kararı verir" yaklaşımıdır.

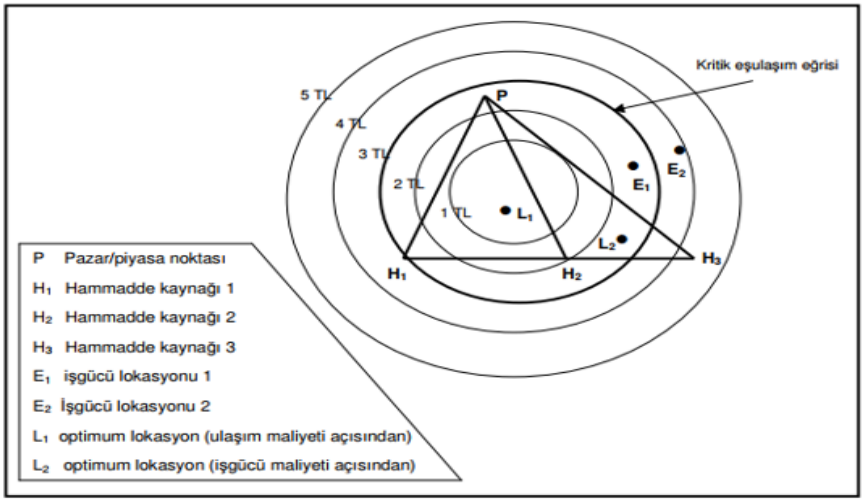


Şekil 2.8: Weberyen lojistik taşıma maliyetleri ve hammaddenin niteliğine göre optimum lokasyon (Weber, 1929; Yavan, 2006).



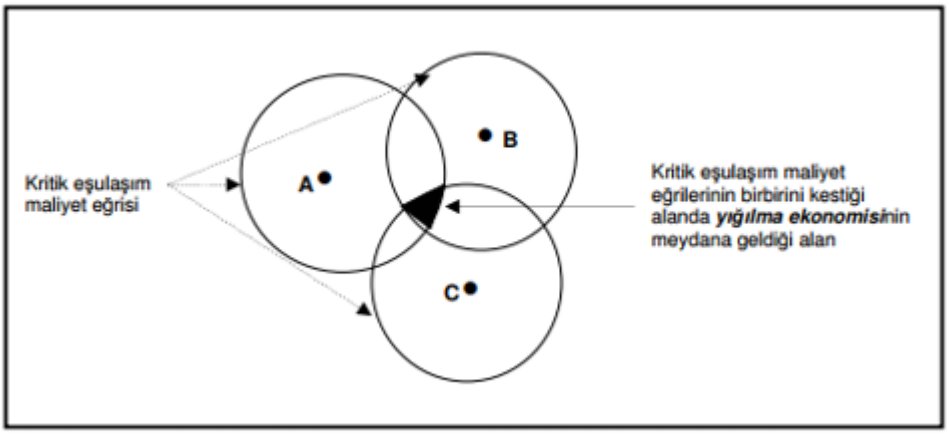
Şekil 2.9: Weber'in lokasyon üçgeni (Yavan, 2006).

Weber'in çalışmalarında ki diğer önemli parametre ise "iş gücü" maliyetlerinin bölgelere göre farklılık göstermesidir. Weber emeği, kuruluş yerini en düşük lojistik maliyeti noktasından en düşük iş güzü maliyetlerini oluşturan noktalara kaydıracak bir etken olarak görmektedir. Weber'in teorisine göre iş gücü maliyetlerinden elde edilen tasarruf, ulaşım maliyetlerini aşarsa, tesis emek noktasına yakın bir lokasyona açılacaktır.



Şekil 2.10: Weber'in eş maliyetli ulaşım eğrileri analizi çerçevesinde lojistik taşıma maliyetleri ve emek faktörlerine göre optimum lokasyon (Yavan, 2006).

Weber'in araştırmalarının son adımını "yığılma ekonomileri" (agglomeration) oluşturmaktadır. Weber yığılma ekonomisinin iş gücüne benzer bir etki gösterdiğini belirtmektedir. Yani tesisi en düşük işçi maliyetlerinin bulunduğu noktalardan yığılma ekonomilerin oluştuğu noktalara kayabileceğini düşünmektedir.



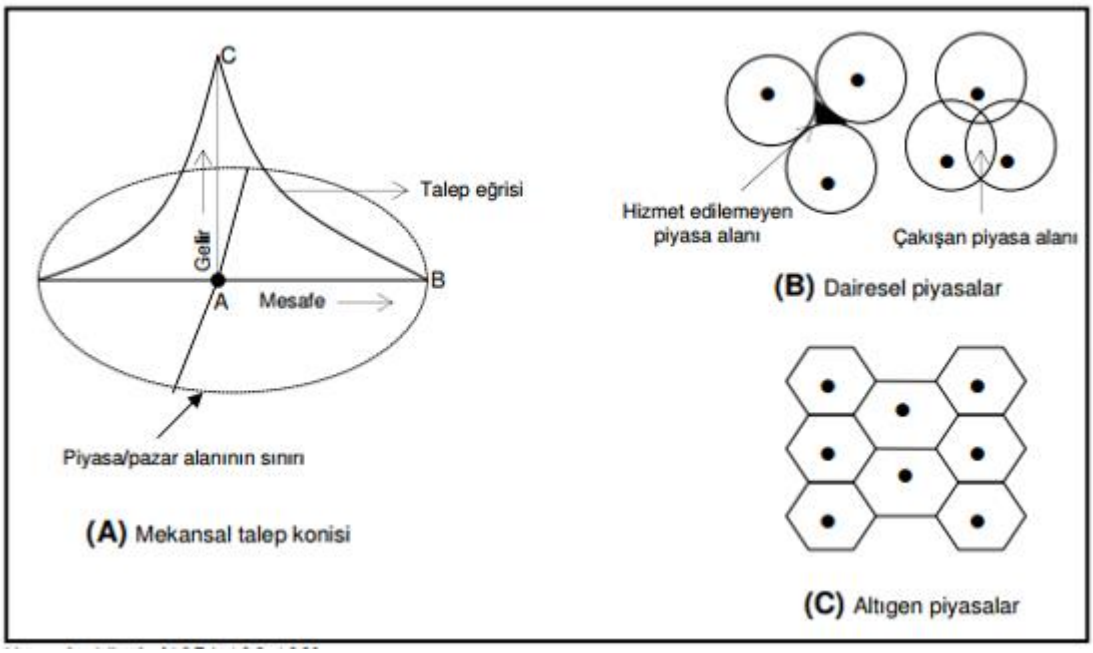
Şekil 2.11: Weber'e göre yığılma ekonomilerinin oluşumu (Yavan, 2006).

2.2.1.2.Lösch'ün En Yüksek Kazanç Teorisi

Lösch, firmalar için ekonomik pazar yerlerinin oluşumunu ve önemini göstermek için çalışmalarını üç aşamalı yolla ortaya koymuştur. İlk adım olarak bir mekânsal talep konisi geliştirerek bu koninin merkez noktasında (A noktası) bir firma üretim yapmakta olup, firma Şekil 2.12'deki BC talep eğrisine sahiptir. Fiyat mesafe ölçüsünün bir fonksiyonu olarak alındığında, AB lojistik taşıma maliyetini göstermekte ve mesafe ile birlikte taşıma maliyeti artmaktadır. Bu talep konisinde, piyasa merkezinden uzaklık arttıkça lojistik taşıma maliyetleri artacağı için talep edilen miktar azalacaktır ve en sonunda B noktasına ulaşıldığında talep sıfıra inecektir. Çünkü B noktasının ötesinde A'nın mallarının fiyatı çok pahalılaşır (her firmanın dairesel bir pazar alanı varsayımı altında) ve tüketiciler daha ucuz olduğu için diğer firmanın malını talep eder. ABC konisi firmanın toplam satış büyüklüğünü vermektedir. Lösch, talep eğrisini A noktası etrafından 360 derece döndürerek pazar alanının başlangıçta dairesel olduğunu ve her firmanın dairesel bir pazar alanına sahip olduğunu ortaya koymuştur (Yavan 2006).

Çok sayıda firmanın dairesel bir ekonomik pazar alanı içinde faaliyet gösterdiği ikinci aşamada, potansiyel pazarlardaki talebi karşılamak için rekabet artar ve bu artış bazı durumlarda dairelerin üst üste çakışmasına, bazı durumlarda ise birçok alana satış yapılamamasına neden olur. Rekabetin daha fazla firmayı aynı faaliyet alanına çekmesi zamanla her bir firmanın pazar alanının daralmasına yol açar. En sonunda ise başlangıçtaki dairesel piyasa altıgen şekilli bir

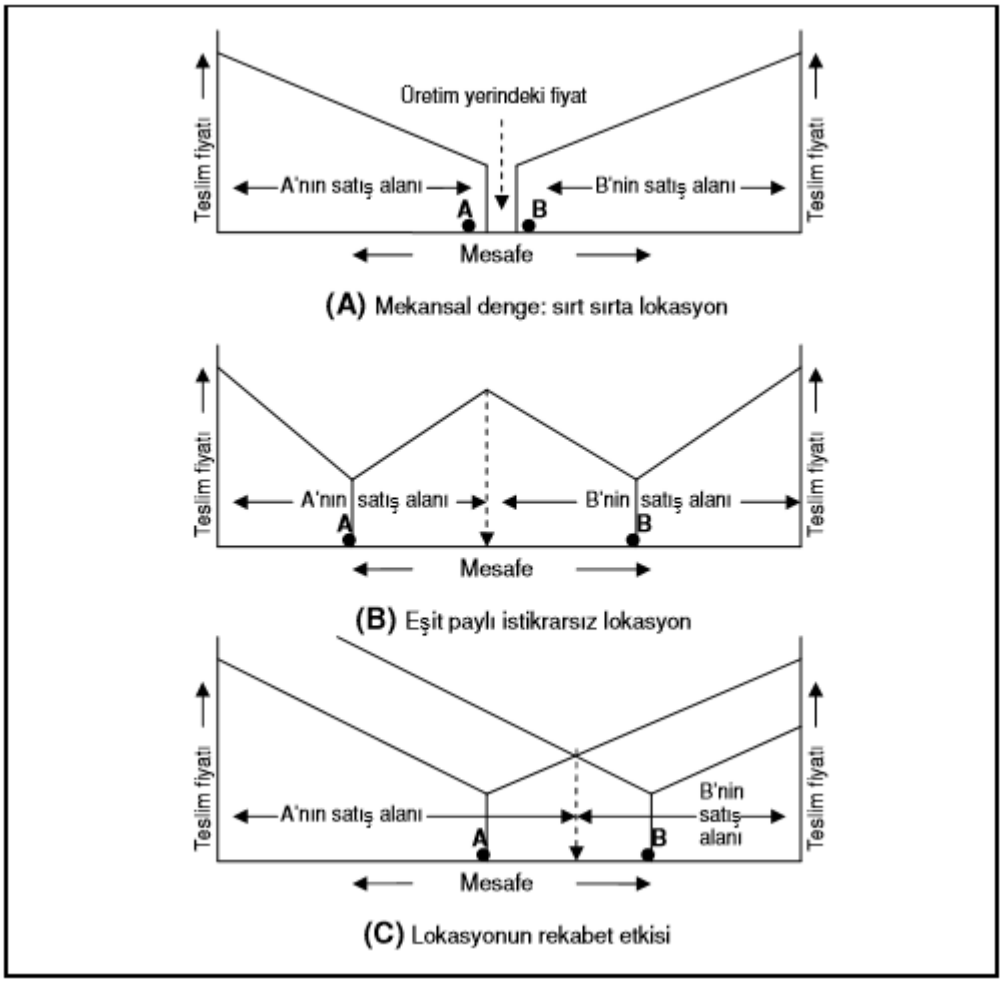
piyasaya dönüşür. Bu altıgen, piyasaların merkez noktalarında ise firmalar yer alır. (Yavan, 2006)



Şekil 2.12: Lösch'ün mekânsal talep konisi ve Pazar alanları (Yavan, 2006).

2.2.1.3. Hotelling'in Mekânsal Rekabet ve Lokasyonel Bağımlılık Modeli

Hotelling (1929) iki firmanın lokasyon yeri seçerken izlediği stratejiyi gösteren basit bir model geliştirdi. Hotelling, teorisinde iki firmanın (duapol) lokasyon er seçimi stratejisi üzerinde durarak dengeli ve istikrarlı bir lokasyon yerinin nasıl seçileceğini araştırmıştır. Bu modele göre, firmalar hareket serbestliğine sahip belirli büyüklükte talebin olduğu bir piyasada aynı mal ve hizmeti satmaktadır. Firmaların mal veya hizmetlerini sattığı piyasada tam rekabet koşulları bulunmakta ve firmalar arasında gizli bir anlaşma (collusion) veya kartel söz konusu değildir.

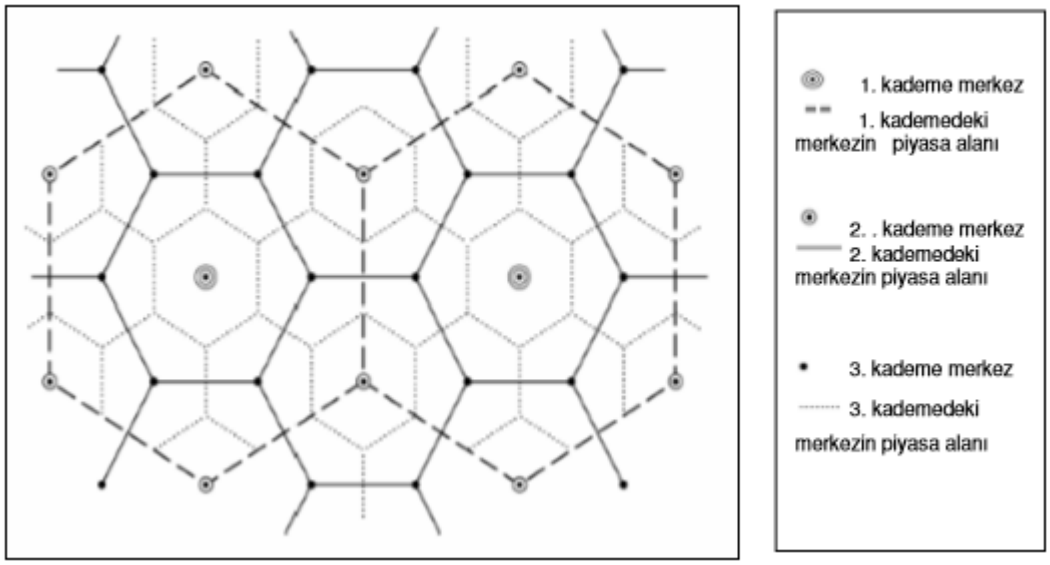


Şekil 2.13: Hotelling modeli rakip firma stratejileri (Yavan, 2006; Hotelling, 1990).

2.2.1.4. Christaller'in Merkezi Yer Teorisi

Christaller'in modelinde talep pazarda düzgün bir dağılıma sahiptir. Birçok firma aynı çıktıyı üretmekte, hammadde lokasyon bağımsız sağlanmakta, lojistik maliyetleri aynı ve tüketiciler en yakın lokasyonu tercih etmektedirler. Bu kabuller altında firmalar bir mal veya hizmetin faaliyet alanı içinde piyasa doygunluk eşiğinin minimum olduğu her lokasyonda bulunabilirler. Christaller'in modeline göre firmaların altıgen şekilli bir pazar içinde faaliyette bulunması, bu pazardaki tüm talebin karşılanması anlamına gelmektedir (Hayter, 1997).

Christaller modeli hiyerarşik bir üretim ve lokasyon yapısı öngörüsüne bulunmaktadır. Dicken (P Dicken, 1990) Şekil 2.14'te modeli şematize etmiştir.



Şekil 2.14: Christaller'in merkezi yer hiyerarşisi ve altıgen piyasa alanları (Peter Dicken ve Lloyd, 1990).

2.2.1.5. Isard ve Greenhut'un Mekân Ekonomisi

Lokasyon yer seçimi çalışmalarına en fazla katkıyı yapanlardan biri de Isard'dır. Isard, VonThünen, Weber, Christaller ve Lösch gibi lokasyon teorisi modellerini birleştirerek genel bir lokasyon yer seçimi teorisi oluşturmuştur. Diğerlerine benzer şekilde en önemli parametreyi sanayi faaliyetleri olarak belirlemiştir (Isard, 1975).

Weber ve onu takip edenlerin (Palander ve Hoover) en az maliyetli lokasyon yer seçimi modeli ile Hotelling ve Lösch'ün talebi ve pazar şartlarını dikkate alan en yüksek kazanç ve/veya pazar alanı modelini birleştiren ilk çalışma iktisatçı Greenhut tarafından yapılmıştır. Greenhut hem maliyet ve talep parametrelerini hem de fiyat ve piyasa parametrelerini göz önünde bulundurarak, firmaların lokasyon yer seçim modellerini oluşturmaya çalışmıştır.

Isard ve Greenhut araştırmaları ile mekân ekonomisinin (spaceeconomy) temellerini oluşturmuşlardır. Her iki bilim insanı da lokasyon teorisinin kapsamını genişleterek, genel ekonomik teorinin ilkeleri ile lokasyon teorisinin ilkelerini birleştirmiş ve bu nedenle genel dengeye oturan bir lokasyon teorisi inşa etmişlerdir. Fakat Isard ve Greenhut, teorilerini neoklasik iktisadın kabulleri ve kısıtları altında gerçekleştirmeye çalıştıkları için, özellikle 1960'lardan itibaren davranışsal iktisatçılar tarafından çok fazlar eleştiriye uğramışlardır (Yavan, 2006).

2.2.1.5. Gravity (Çekim) Modeli

Gravity (çekim) modeli, adını Newton'un "yerçekimi kanunu" teorisinden almaktadır. Buna göre iki kütle arasındaki çekim kuvveti, kütlelerin ağırlıkları ile doğru orantılı ve kütlelerin aralarındaki uzaklığın karesiyle de ters orantılıdır. Model uluslararası ticareti açıklamaya çalıştığında ise iki ülke arasındaki ticaret hacmi, ülkelerin milli gelirleriyle doğru orantılı, aralarındaki mesafeyle ters orantılı olarak değişir (Tatlıcı ve Kızıltan, 2011).

Bu modelde perakende firmaları arasındaki etkileşim firma büyüklükleri ile doğru orantılı ve uzaklık ile ters orantılıdır. Bir mağazayı etkileyen diğer mağaza ne kadar büyük ve yakın ise rekabet etkisi o denli fazla olmaktadır.

2.2.2. Perakende Sektöründe Lokasyon Analizi

Perakende sektörünü incelediğimizde, özellikle Türkiye'de oldukça dinamik bir yapı bulunmaktadır. Gıda, giyim, beyaz eşya, elektrik ve elektronik eşya gibi ürünleri satan firmalar, zincir mağaza modelinde, Türkiye ortamında sürekli olarak değişen pazar yapısını ve nüfus dinamiklerini göz önünde bulundurmakta, bunun sonucu olarak da mağaza açma, kapatma, büyütme ve birleştirme gibi kararları belirli sıklıklarda vermek zorunda kalmaktadırlar. Bu tür kararlar çoğu zaman detaylı araştırmalar yapılmadan, müşteri profili ve pazar yapısı incelenmeden, ayrıca mekânsal unsurlar da dikkate alınmadan verilmekte, bu da firmaları ciddi zararlara uğratmaktadır. Ayrıca yer seçimi problemini zorlaştıran bir diğer unsur da rakip mağazaların konumları ve rakip mağazaların büyüme stratejilerinin de incelenmesi gerekliliğidir (Bozkaya ve Yanık, 2008).

Bu amaçla firmalar lokasyon analizi konusu ile ilgili olarak çevre analizleri yaptıklarında aşağıdaki temel sorulara cevap aramaktadırlar.

- ✓ Bölge ticari olarak ne kadarlık değere sahip?
- ✓ Rakipler kimler?
    - o Olası rakiplerin tehdidi
    - o İkame ürünlerin tehdidi

- Tedarikçilerin gücü
- Müşterilerin gücü
- Rakipler arası rekabetin şiddeti

✓ Sosyal (sosyo kültürel yapı) yapı nasıl? (İnsanların değer yargıları, tutum ve davranışları, kültürel alışkanlıkları, Yaşam biçimlerini oluşturan ve etkileyen unsurların bulunduğu çevre, tatil günleri vs.)

✓ Ekonomik yapı nasıl? (Mal ve hizmet durumu, kaynakların üretim ve tüketim durumu, kişi başına gelir, kriz dönemi, canlanma dönemi, sezon, enflasyon vs.)

✓ Demografik yapı nasıl (ülkede bulunan nüfusun yapısını, özellikleri, durumunu, dinamik özelliklerini)

✓ Politik yapı nasıl?

✓ Teknolojik yapı nasıl? (Teknolojik çevre; yeni ürünler, süreçler, yeni yaşam biçimleri, bankacılık, iletişim, iş hayatı, vs.)

✓ Uluslararası yapı nasıl (Ülke dışındaki fırsat ve tehditleri barındıran politik, yasal, ekonomik ve sektörel olaylar ve oyuncuları kapsar. Savaş olasılıkları, Ülkeler arası birleşme / ayrılmalar, Serbest ticaret bölgeleri, Gümrük birliği antlaşmaları, Ortak Pazar antlaşmaları, Ambargolar ve kotalar)

Yöneylem araştırması teorilerinde yer alan, lojistik ve karar teorisini ilgilendiren çalışmalarda mağaza veya tesis yer seçimi problemlerinin uzun süredir araştırmalara konu olduğunu gözlemlemekteyiz. Lokasyon teorisi alanındaki araştırmalara 20. yüzyılın başlarında Reilly ve Converse'in yaptığı çalışmalar öncü olmuştur. Geliştirilen ilk "çekim modeli"nde, tüketicilerin şehirlerarası alışveriş amaçlı hareketleri, nüfus ve mesafe dikkate alınarak modellenmektedir (Reilly, 1929). Fakat sadece bu iki girdi parametresini dikkate alan modellerin içerdiği eksiklikleri aşmak için Huff tarafından müşteri tercihlerinin etkisinin de dikkate alındığı yeni ve daha temel bir model ortaya konmuştur (Huff, 1964). Huff, müşterilerin mağaza

tercihlerinin, mağazanın çekiciliğine ve müşteri ile mağaza arasındaki uzaklığa bağlı olduğu varsayımı ile bu modeli geliştirmiştir.

Lokasyon yer seçimi problemlerinde, müşterilerin her zaman en yakın mağazalara gitmeyip mağazaların çekiciliği ve kendi ihtiyaçları doğrultusunda tercihte bulundukları gerçeğini işleyen modellere son yıllarda daha fazla rastlanmaktadır. Örnek olarak Daşçı ve Laporte (2005), müşterilerin mesafe, hizmet türü ve tüketim özelliklerini dikkate alarak belirli bir olasılıkla mağaza tercih ettikleri bir model ortaya koymuşlardır (Dasci ve Laporte, 2005). Bu modelde yatırım bütçesi kısıtı da parametre olarak yer almaktadır. Fernandez (2006), tek bir tesis yer seçimi için hem ana firma hem de bayi için kar maksimizasyonu yapan bir model formüle etmiştir. Bu çalışmada da yatırım bütçesi kısıtı dikkate alınmakta ve problemin çözümü için iki algoritma önerilmektedir (Bozkaya ve Yanık, 2008). Silva ve Serra (2007), elde edilecek pazar payını hem uzaklığa bağlı erişim süresine hem de lokasyonda bekleme süresine bağlı bir fonksiyon şeklinde kurgulamış ve bundan yola çıkarak modelledikleri problemi karınca kolonisi sezgisel optimizasyon yöntemi ile çözmüşlerdir (Silva ve Serra, 2007). Drezner ve Drezner (2007), benzer problemi ulaşım ağı üzerinde p-medyan bazlı çekim modeli şeklinde modellemiş ve iki ayrı sezgisel algoritma yardımıyla çözmeye çalışmıştır (Drezner ve Drezner, 2007).

Bir perakende firmasının Mağaza Açılış Kararları aşağıdaki iş akış süreçlerini takip etmektedir.

Sürecin ikinci adımı olan lokasyon belirleme alt ardıl başlıkları ise Şekil 2.15'te verilmiştir.

Şekil 2.15: Perakende Firması Mağaza Açılış Süreci İş Akışı

Şekil 2.15'te gösterilen iş akış şemasının prosedürü ise aşağıda verilmiştir.

LOKASYONLARIN BELİRLENMESİ

- *Yeni mağaza açılması süreci iki şekilde başlatılır ve tüm süreç "Mağaza Açma Takip Çizelgesi" inde takip edilir.*
    - *Mağazacılık GMY tarafından mağaza açılması yönünde karara bağlanan yerlerde bölge yapılanması marifetiyle mağaza açmaya uygun yerler aranır.*
    - *Bölgelerde yeni açılan AVM'lerden ve talep edilmese dahi bölgeler tarafından uygun görülen yerler, durum bölge müdürü ve GMY aracılığı ile bildirilir.*
- *Mağazacılık GMY tarafından mağaza açılması yönünde karara bağlanan yerlerde bölge yapılanmasına mağaza açma talepleri, mağazanın yerleri ne zaman açılmayı planlandığı ve ne zamana kadar dönüş beklendiği yazılı talimatla iletilir.*
- *Bölge yapıları belirtilen tarihe kadar araştırmalarını yaparlar, belirlenen alternatifler inşaat ve emlak birimine "Lokasyon Değerlendirme Formu" ile iletilir.*
- *Bölgelerde yeni açılan AVM'lerden gelen talepler, talebin tarihi, geldiği bölge ve AVM projesinin adı ve diğer gerekli bilgileri içerecek şekilde inşaat ve emlak tarafından kayıt altına alınır.*
    - *İnşaat ve emlak birimi, AVM projelerinin yönetiminden gerekli olan bilgi ve belgeleri talep eder. (Mağaza bilgileri, resmî belgeler, vb.)*
- *İnşaat ve emlak birimi, bir hafta içerisinde AVM'den gerekli bilgileri toplayarak Lokasyon Değerlendirme Formu"na aktarır ve değerlendirme sürecine dahil eder.*

ADAY LOKASYONLARIN DEĞERLENDİRİLMESİ

- *Aday mağazanın tespitinin ardından mağaza yerinin uygunluğu "Lokasyon Değerlendirme Formu" ile alınan görsel ve çevre bilgilerine göre Mağazacılık GMY tarafından değerlendirilir. Değerlendirmelerde "Mağaza Açma Standartları"'nda belirlenen kurallar esas alınır. Gerekli görülmesi halinde görevlendirilen bölge müdürü tarafından aday mağaza ziyaret edilir.*

- *Aday lokasyonun binasının güvenliği özellikle eski binalarda (2000 yılı öncesi) inşaat ve emlak birimi tarafından tetkik ettirilir ve lokasyon değerlendirme formunda görüşlere yer verilir. Güvenli olmayan binalarda mağaza kesinlikle açılmaz.*

- *Aynı ilçe içerisinde birden fazla mağaza açılması, mevcut mağazanın performansını olumsuz etkileyebileceğinden ve başka bir yerde mağaza açmanın fırsatını elden kaçıracağından dolayı, aynı ilçe için önerilen mağazaların özel olarak bildirilir ve ilçedeki mevcut mağaza yöneticisinin görüşleri lokasyon değerlendirme formuna dahil edilir.*

- *Lokasyon olarak olumlu değerlendirilen aday mağazalar için bölge müdürlükleri bünyesinde detaylı bütçe ve fizibilite çalışmaları yapılması talep edilir.*

- *"Bütçe Hesaplama Tablosu" gelir projeksiyonlarını içerecek şekilde AVM mağazaları için 5 yıllık, diğer mağazalar için 3 yıllık olarak bölgeler tarafından hazırlanır ve Finans Birimi'ne iletilir.*

- *Finans birimi tarafından, giderler ve maliyetler bütçe hesaplama tablosuna işlenir. Gelir tahminlerinin tutarlılığı sorgulanarak gerekli düzeltmeler yapılır ve onaylanarak iskonto edilmiş nakit akışları ve bugünün değeri ile EBITDA (faiz vergi ve amortisman öncesi kar) ve mağaza yatırımının verimliliği Mağazacılık GMY'ye iletilir. Mağazacılık GMY, bütçe hesaplamalarını genel yatırım bütçesine göre kontrol eder ve yapılan projeksiyonların doğruluğunu ve yatırım karar standartlarına göre yatırımın uygunluğunu sorgular.*

- *Yüksek yatırım gerektiren veya kritik görülen mağazalarda fizibilite çalışmasına ek olarak lokasyon analiz anketi hizmeti alınır. Hangi yatırımlarda lokasyon analizi yapılması gerektiği ve kuralları Mağaza Açma Standartlarında belirlenmiştir.*

- *Genel Müdür Yardımcısının bütçe hesaplama tablosunun inceleyip onaylamasının ardından Genel Müdür projeyi yazılı olarak değerlendirir. Onaylanan mağazalar icra kuruluna sunulur.*

- *İcra kurulu tarafından onaylanan mağazaların bütçeleri, bütçe takibinde kullanılan ERP Bütçe Modülü 'ne Muhasebe Birimi tarafından işlenir. Sisteme tanımlanan bütçe,*

Mağazacılık GMY'nin onayına düşer ve bütçelerin onaylanmasının ardından bütçeler sistemde aktif hale gelir.

- *Mağaza açılış kararında onaylanan bütçeler yıllık olarak Genel Müdür Yardımcısı tarafından gerçekleşen veriler ile karşılaştırılarak analiz edilir.*

- *Olumlu sonuçlanan mağaza açma kararları için mağaza açılışı ile ilgili her türlü bilgi ve belgenin takip edildiği mağaza dosyaları oluşturulur ve "Mağaza Değerlendirme Takip Çizelgesi"inde durumu olumlu olarak işlenir.*

    - *Mağaza dosyasında, açılacak mağazanın değerlendirmesinde karşılaşılan mağaza açma standardına uygun olmayan her türlü durum, bu durumun değerlendirmeye etkisi ve nasıl karar verildiği, yazılı olarak dokümante edilmiş, GMY ve GM tarafından imzalanmıştır.*

- *Sürecin herhangi bir adımında olumsuz sonuçlanan mağaza değerlendirmeleri, olumsuz sonuçlanmasının nedenleri de dokümante edilerek dosyalanır ve arşivlenir. Olumsuz sonuçlanan mağaza değerlendirmeleri, "Mağaza Değerlendirme Takip Çizelgesi"inde olumsuz olarak işlenir.*

Son yıllarda Öklid uzaklıkları yöntemi ile lokasyon belirleme çalışmaları yapılmıştır. Öklid uzaklıkları yöntemi ile açılacak mağazanın önceden açılmış mağazalardan en çok hangisine benzediği keşfedilir. Bu yöntemde parasal herhangi bir değer kullanılmaz. Tamamen mağaza çevresinden elde edilen (Google Places vb. uygulamaların) verileri kullanılır. Bu yöntem literal anlamda birbirine kuş bakışı benzeyen mağaza çevrelerini teşhis etmektedir. Çevresi kamu ve özel kuruluşlar, okullar, üniversiteler, kafe ve restoranlar, mağazalar, ibadethaneler bakımından birbirine benzeyen mağazaların parasal değişkenler bakımından da birbirine benzeyeceği beklentisi üzerine geliştirilmiştir.

Bir diğer yöntemde Yapay Sinir Ağlarıdır. Bu yöntemlede yine elde edilmiş mağaza çevre verilerinin mağazaların cirosunu tahmin etme başarısına göre yeni açılacak mağazaların ne kadarlık bir ciro getireceği varsayımı üzerinden geliştirilmiştir.

2.3. MAKİNE ÖĞRENME

Makine öğrenmesi bilgisayarların geçmiş bilgilerinden elde edilen tecrübelerden yararlanarak, gelecekteki olayların tahmin edilmesine ve modelleme yapmasına imkân veren bir yapay zekâ alanıdır (Adeli ve Hung, 1994).

Makine öğrenimi yeni yöntem ve algoritmaların gerçek yaşam problemlerine uygulanabilirliği açısından en verimli araştırma alanlarından biridir. Teknolojinin dünyada beklenmedik bir hızda değişmesinden dolayı günümüzde nispeten ucuz bir fiyata yüksek kaliteli ve hızlı donanım altyapıları kullanılabilmektedir. Akıllı bir şekilde çevreye adapte edilebilen sistemlerin geliştirilmesi büyük bir pratik uygulamaya alanına sahiptir. Genellikle, bu sistemler maksimizasyon / minimizasyon kriterine göre belirli bir algoritma performansını optimize ederek, deneysel verileri kullanırlar (Olivas, 2009).

Makine öğrenme yöntemlerinin büyük veri tabanlarına uygulanmasına ise veri madenciliği denir. Veri madenciliği imalat, optimizasyon, kontrol ve sorun giderme vb. için öğrenme modelleri kullanmaktadır. Telekomünikasyonda, şebeke optimizasyonu ve hizmet kalitesini en üst düzeye çıkarmak için arama şablonları analiz edilmektedir. Fizik, astronomi ve biyoloji gibi bilim dallarında çok büyük miktarda veri olduğundan dolayı bu veriler ancak bilgisayarlar tarafından hızlı bir şekilde analiz edilebilmektedir (Alpaydin 2014).

Makine öğrenmenin en avantajlı özelliklerinden biri, desen tanıma (sinir ağları, destek vektör makineleri, karar ağaçları, takviye öğrenmesi, ...), veri madenciliği (zaman serisi tahmini, modelleme, ...), istatistik (Bayesci yöntemler, Montecarlo yöntemleri, önyükleme, ...) veya sinyal işleme (Markov modelleri) gibi farklı alanlardan gelen bilgileri bir araya getirmesidir. Bu nedenle, Makine öğrenme farklı bilgi alanları kullanan sağlam çözümler üretmektedir (Olivas, 2009).

Makine Öğrenmesi geçmiş öğrenme deneyimlerini ve örnek verileri kullanarak bilgisayarların öğrenme işlemini gerçekleştirmesini amaçlamaktadır. Bu öğrenme işlemi Denetimli ve Denetimsiz öğrenme olarak iki türdür (Subbulakshmi vd., 2010).

Denetimli öğrenme sınıflandırma ile eşanlamlıdır. Öğrenmedeki denetim, eğitim veri setindeki sınıflandırılmış örneklerden gelmektedir. Örneğin, posta kodu tanıma probleminde, sınıflandırma modelinin öğrenimini denetleyen eğitim örnekleri olarak bir dizi el yazısı posta

kodu görüntüsü ve bunlara karşılık gelen makine tarafından okunabilir çeviriler kullanılmaktadır (Han vd., 2011).

Denetimsiz öğrenmede uygun bir çıktı değişkenini tanımlamak için mevcut hiçbir bilgi veya sınıflandırılmış değişken bulunmamaktadır, sadece girdi verileri vardır. Sistem, girdi değişkenleri arasındaki ilişkileri bulmaya ve öğrenmeye çalıştığı için bu yönteme "bilimsel keşif" adı verilmiştir (Izenman, 2008).

Sınıflandırma kavramı, veriyi sahip olduğu niteliklere göre veri kümesinde bulunan sınıflara dağıtmaktır. Bu sınıfların özellikleri ve sayısı önceden belirlidir. Veri kümesindeki bu sınıfları belirten değerler etiket olarak adlandırılmaktadır. Veri setindeki her örneğin sınıf bilgisi de dahil olmak üzere çeşitli özellikleri mevcuttur. Sınıflandırmada eğitici bir öğrenme söz konusudur. Eğitim kümesindeki öğelerin sınıfları bellidir ve bu küme bir model oluşturmak için kullanılmaktadır. Sınıflandırma algoritmaları, verilen eğitim kümesindeki sınıf etiketleri ile diğer özellikler arasındaki ilişkileri analiz edip öğrenme işlemini gerçekleştirmektedir. Kurulan model öğrenme kümesinde olmayan öğelerle denenip modelin başarısı ölçülmektedir. Sınıflandırma algoritmalarının çalışması, eğitim verisi ile öğrenme ve öğrenme sonucu test verisinin sınıflandırılması aşamalarından oluşmaktadır. Sonuç olarak yeni gelen öğenin hangi sınıfa ait olduğuna kurulan ve test edilen bu model yardımı ile karar verilmektedir (Sheppard ve Young, 2006).

2.3.1. Tembel (Lazy) Algoritmalar

Tembel (lazy) algoritmaların diğer yöntemlerden en büyük farkı öğrenme kümesini saklamasıdır. Diğer yöntemlerde öğrenme aşamasında yapılan işlemler, bu yöntemde tahminleme aşamasında yapılmaktadır. Sınıflandırılacak yeni bir örnek geldiğinde, en benzer özelliklerin olduğu sınıfa atanmaktadır. Tembel algoritmaların yaygın olarak aşağıdaki yöntemleri kullanılmaktadır (Zhang ve Zhou, 2005).

- K-en yakın komşu
- K * (K Star)

K-en yakın komşu (KNN) algoritmaları, en temel örnek tabanlı sınıflandırma algoritmalarından biridir. Bu yöntem, sınıf türleri belli olan bir eğitim kümesinden faydalanarak, yeni bir gözlemin hangi sınıfa ait olduğunu belirlememizi sağlamaktadır. KNN algoritmasında her bir örnek n

boyutlu uzayda bir noktayı temsil eder. Eğitim setindeki tüm örnekler bu şekilde n boyutlu bir uzayda tutulur. Yeni bir gözlem bu uzaya dahil edileceğinde bu örnek uzaya dahil edilir. Örneğe en yakın k adet örnek seçilir ve çoğunluk oylamasına göre hangi sınıfa dahil edileceği belirlenir (Karagülle, 2008). KNN yönteminin performansını en yakın komşu sayısı, eşik değer ve benzerlik ölçütü etkilemektedir (Çalışkan ve Şoğukpınar, 2008).

K * (K Star), örnek tabanlı bir sınıflandırıcıdır, başka bir deyişle, bir test örneğinin sınıfı, bazı benzerlik fonksiyonlarıyla belirlendiği şekilde, benzer eğitim örneklerinin sınıfına dayanmaktadır. Entropi tabanlı bir mesafe işlevi kullanması ile diğer örnek tabanlı öğrenicilerden farklıdır. K* kesinlikle mesafe fonksiyonu değildir. Örneğin, K* genel olarak sıfırdan farklıdır ve işlev simetrik değildir. Uzaklık ölçüsü olarak entropiyi kullanmasının birçok yararı vardır. Sembolik nitelikleri, değerli öznitelikleri ve kayıp değerleri tutarlı bir şekilde ele almaktadır (Çölkesen ve Kavzoğlu, 2011).

2.3.2. Bayes Ağları Algoritmaları

NaiveBayes (NB) yöntemi, istatistiksel bir sınıflandırma algoritmasıdır. Bayes teoremine dayanan bu yöntem, bilinmeyen örnekleri sınıflandırmak için olasılıkçı bir yaklaşım sergiler (Karagülle, 2008).

T eğitim kümesi için her bir örnek n boyutlu uzayda, X = (x1, x2, …, xn) olarak tanımlansın. Sınıflamada veri kümesindeki m adet sınıf için (C1, C2, …, Cm) en büyük P(Ci|X) ifadesi aranır. P(Ci|X) olasılığının en büyütme işlemi için bayes teoremi kullanılır;

- P(X) : Herhangi bir örneğin X olma olasılığı,
- P(Ci) : Sınıf j‟ nin ilk olasılığı,
- P(X|Ci) : Sınıf i‟ den bir örneğin X olma olasığı,
- P(Ci|X) : X örneğinin sınıf i‟ den olma olasılığıdır.

$$P(Ci|X) = \frac{P(X|Ci)\ P(Ci)}{P(X)}$$

2.3.3. Kural Tabanlı Algoritmalar

Karar tabloları, eğitim setindeki verilerin özellikleri ile oluşturulan ve sınıflandırma yapan bir karar tablosudur. Mümkün olan en basit hipotez alanlarından biridir ve genellikle anlaşılması

kolaydır. Deneysel sonuçlar, yalnızca ayrık özellikler içeren suni ve gerçek dünya alanlarında, karar tablolarını kullanan bir algoritma olan IDTM'nin, bazen C4.5 gibi son teknoloji algoritmalardan daha iyi performansa sahip olabildiğini göstermektedir. Artımlı çapraz doğrulamayı kullanarak, belirli bir veri kümesini ve zaman içinde IDTM'yi, örneklerin sayısıyla, özellik sayısıyla ve sınıf değerleri sayısıyla lineer olarak çapraz doğrulama yapmak mümkündür. Sınıflandırılmış girdi değerlerinin örnek bir veri kümesi göz önüne alındığında, denetlenen makine öğrenme algoritmalarının tahmin değerleri ile gerçekleşen değerli arasındaki tutarlılığı aramaktadır (Kohavi, 1995).

0-R (Zero-R), sınıflandırıcı oluşturan ve kullanan için bir sınıflandırma türüdür. Ortalama değeri (sayısal bir sınıf için) veya modu (nominal bir sınıf için) öngörmektedir (Bouckaert, 2008).

1-R (One-R), adı verilen belirli kurallar, bir nesneyi tek bir nitelik temelinde sınıflandıran kurallardır (diğer bir deyişle, bunlar 1 düzeyli karar ağaçlarıdır). 1-R (1-düzeyli kurallar) sınıflandırıcı oluşturmak ve kullanmak için kullanılan bir sınıflandırma türüdür. Başka bir deyişle, sayısal öznitelikleri ayırma tahmini için minimum hata özniteliğini kullanmaktadır. 1-R her açıdan sıradandır, eğitim setindeki hata oranına göre nitelikleri sıralamaktadır (Holte, 1993).

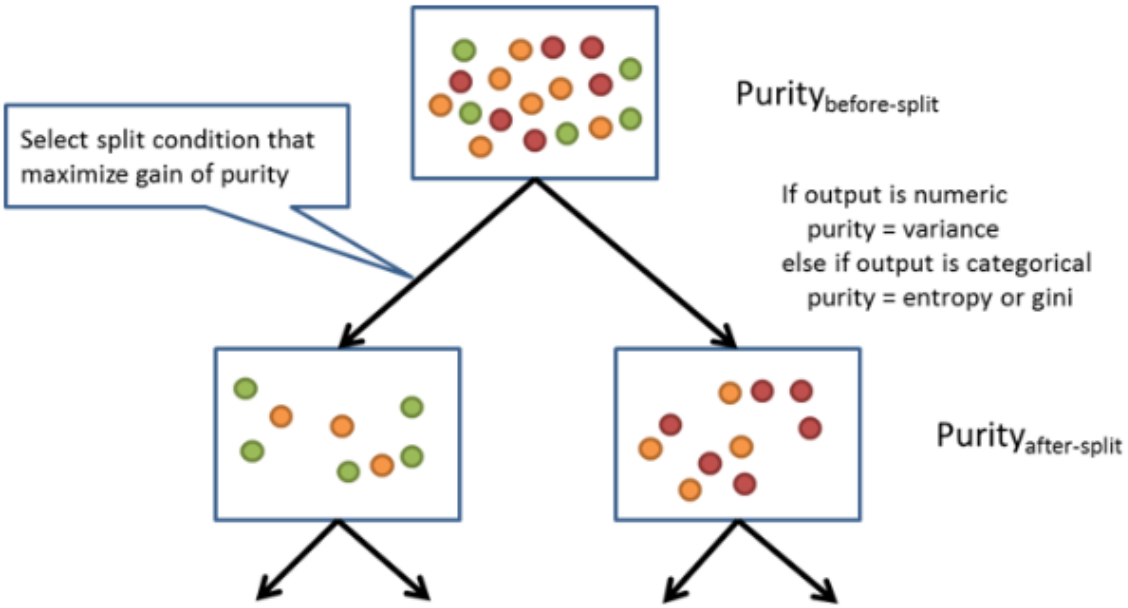
2.3.4. Karar Ağacı Algoritmaları

Karar ağacı algoritması, tüm veri kümesindeki girdileri belirli bir sınıfa atayana kadar veri kümesini yinelemeli olarak bölen bir yapay zekâ yöntemidir. Bir karar ağacı yapısı, kök, iç ve yaprak düğümlerinden oluşmaktadır. Ağaç yapısı, girdi değişkenlerini sınıflandırmada kullanılmaktadır. Ağacın her iç düğümünde, en iyi bölünme kararı, homojenlik ölçümleri kullanılarak yapılmaktadır. Ağaç yaprakları, veri kümesinin gruplandırıldığı sınıf değerlerinden oluşmaktadır. Karar ağacı sınıflandırma tekniği iki ilişkisel süreçten oluşmaktadır

1. Büyütme,
2. Budama

Ağaç büyütme yukarıdan aşağıya doğru hiyerarşik olarak yapılmaktadır. Bu aşamada, ağacın tüm veri kümeleri aynı sınıf değerlerine gelene kadar tekrar tekrar ayrılmaktadır. Eğitim verisi yinelemeli olarak değerlendirildiğinden hesaplama yoğunluğu vardır. Ağaç budama aşağıdan

yukarıya doğru (child-parent) yapılmaktadır. Sınıf içi aşırı uyum veya sınıflar arası uyumsuzluk algoritmanın tahmin ve sınıflandırma doğruluğunu iyileştirmek için kullanılmaktadır. Karar ağacı algoritmasında gereğinden fazla homojenlik yanlış sınıflandırma hatasına neden olmaktadır. Ağaç budamada, eğitim veri seti yalnızca bir kez değerlendirildiğinden, ağaç büyütme aşamasına kıyasla daha az hesaplama yoğunluğu içermektedir (Anyanwu ve Shiva, 2009).



Şekil 2.16: Karar Ağacı Algoritması Çalışma Şekli (Ho, 2018)

ID3 (İteratifDichotomiser 3) karar ağacı algoritması, veriler, ağaç büyütme aşamasında sırayla her düğümde sıralanır ve en iyi bölünen tek nitelik seçilmektedir. ID3, bölme özniteliğini seçerken bilgi kazancı ölçütü kullanılmaktadır. Ağaç modelinin oluşturulmasında yalnızca kategorik nitelikleri kabul etmektedir. Eğitim veri setinde çok fazla gürültü ya da detay varsa, ID3 doğru sonuç vermemekte, dolayısıyla ID3'le karar ağacı modeli oluşturulmadan önce yoğun bir veri ön işlemi uygulanmaktadır (Anyanwu ve Shiva, 2009).

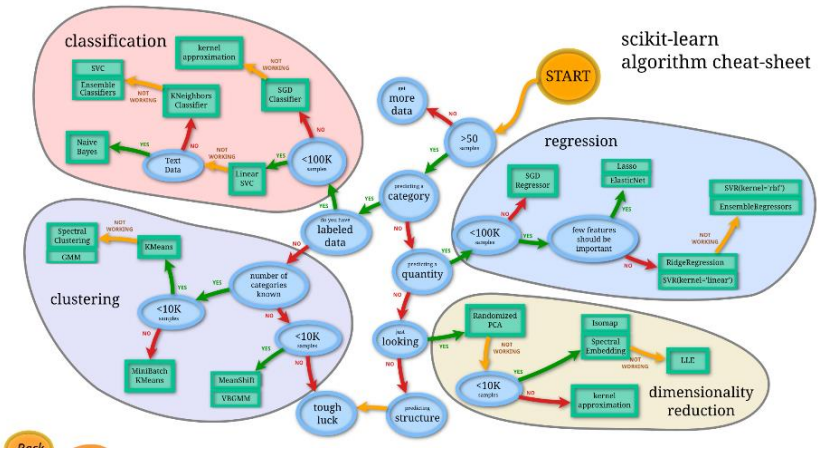
C4.5 ve C.5.0 algoritması, en fazla kullanılan karar ağacı algoritması Quinlan'ın ID3 algoritmasının geliştirilmiş hali olan C4.5 algoritmasıdır. C5.0 algoritması ise C4.5'in geliştirilmiş hali olup, özellikle büyük veri setleri için kullanılmaktadır. C5.0 algoritması, doğruluğu arttırmak için boosting algoritmasını kullandığından, boosting ağaçları olarak da bilinir. C5.0 algoritması C4.5'e göre çok daha hızlı olup, hafızayı daha verimli kullanmaktadır.

Her iki algoritmanın çıktıları aynı olsa da C5.0 şekil olarak daha düzgün karar ağaçları elde etmemizi sağlamaktadır (Quinlan, 1993).

CART Algoritması, Morgan ve Sonquist'in AID (AutomaticInteractionDetection) adlı karar ağacı algoritmasının devamı niteliğine Breiman ve diğerleri tarafından 1984 yılında önerilmiştir. Hem sayısal hem de nominal veri türlerini, girdi ve kestirimsel değişken olarak kabul edebilen CART algoritması, sınıflandırma ve regresyon problemlerinde bir çözüm olarak kullanılabilir. CART karar ağacı, ikili olarak özyinelemeli biçimde bölünen bir yapıya sahiptir. Dallanma kriteri olarak Gini indeksinden yararlanan CART ağacı, kuruluş aşamasında herhangi bir durma kuralı olmaksızın sürekli olarak bölünerek büyümektedir. Son olarak yeni bir bölünmenin gerçekleşmeyeceği durumda, bu sefer, uçtan köke doğru budama işlemi başlatılır. Olası en iyi sonucu veren karar ağacı, her budama işlemi sonrası bağımsızca seçilmiş bir test verisi ile değerlendiremeye tabi tutularak tespit edilmeye çalışılır (Çalış vd., 2014).

J48 algoritması, kesilmiş veya hazırlanmamış C4.5 karar ağacı oluşturmak için kullanılan, yukarıdan aşağıya doğru bir karar ağacı yapısı oluşturan ve yaprakları sınıf etiketini ifade eden bir sınıflandırma türüdür (Loh, 2008).

RandomForest (rasgele ormanlar) ağaç tahminleyicilerinin bir kombinasyonu olup, her ağaç, bağımsız olarak örneklenmiş rastgele bir vektörün değerlerine ve ormandaki tüm ağaçlar için aynı dağılıma bağlıdır. Sınıflandırma doğruluğundaki belirgin gelişmeler, bir grup ağaç yetiştirmesinden ve en popüler sınıf için oy kullanmalarından meydana gelmiştir. Bu toplulukları büyütmek için, çoğunlukla topluluktaki her ağacın büyümesini yöneten rastgele vektörler üretilmektedir. Çok sayıda ağaç üretildikten sonra en popüler sınıfa oy verilmektedir. Bu prosedürler "rasgele ormanlar" olarak adlandırılmaktadır. Rastgele orman, {k} bağımsız olarak eşit dağılmış rastgele vektörken ve her ağaç, x girişindeki en popüler sınıf için bir birim oy kullanırken, ağaç yapılı sınıflandırıcılardan {h (x, k), k = 1, ...} oluşan bir koleksiyonu içeren bir sınıflandırıcıdır (Breiman, 2001).

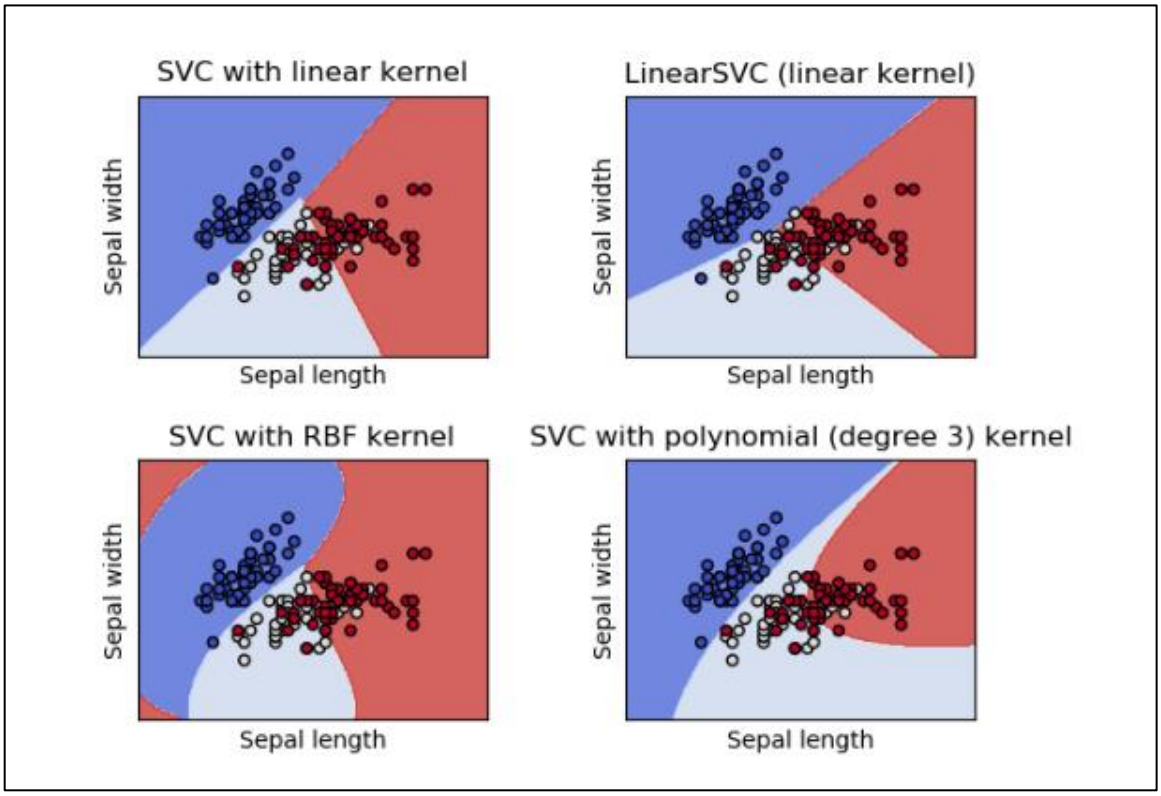


Şekil 2.17: Makine Öğrenme Algoritma Seçme Kuralı (Pedregosa vd. 2011).

2.3.5. Sınıflandırmada Kullanılan Fonksiyonlar

Yapay sinir ağları (YSA), insan beyninden esinlenerek geliştirilmiş, biyolojik sinir ağlarını taklit ederek öğrenen, ağırlıklı bağlantılar aracılığıyla birbirine bağlanan ve her biri kendi belleğine sahip işlem elemanlarından oluşan paralel ve dağıtılmış bilgi işleme yapılarında oluşan bilgisayar sistemleridir. Yapay sinir ağları, öğrenme yolu ile herhangi bir yardım almadan otomatik olarak yeni bilgiler elde edebilme ile geliştirilmiştir (Namlı, 2012).

Destek vektör makineleri (supportvectormachines), veriyi birbirinden ayırmak için en uygun fonksiyonun tahmin edilmesi esasına dayanan, tahmin edilen bu doğrusal veya doğrusal olmayan fonksiyon yardımıyla sınıflandırma ve regresyon problemlerini analiz eden istatistiksel bir yöntemdir. Destek Vektör Regresyonu (supportvectorregression), regresyon yapan destek vektör makineleridir. Öğrenme sonucu bir fonksiyonun tahmin edilmesi regresyon problemidir. Destek vektör regresyonlarında regresyon riskini minimize eden bir fonksiyon bulunmaya çalışılmaktadır (Namlı, 2012).



Şekil 2.18: Destek Vektör Fonksiyonları Sınıflandırma Desenleri (Pedregosa vd. 2011).

Lineer (Doğrusal) Regresyon, bir değişken ve bu değişkeni etkileyen bir veya daha fazla değişken arasındaki ilişkiyi doğrusal bir model ile ifade eden yöntemdir (Namlı, 2012).

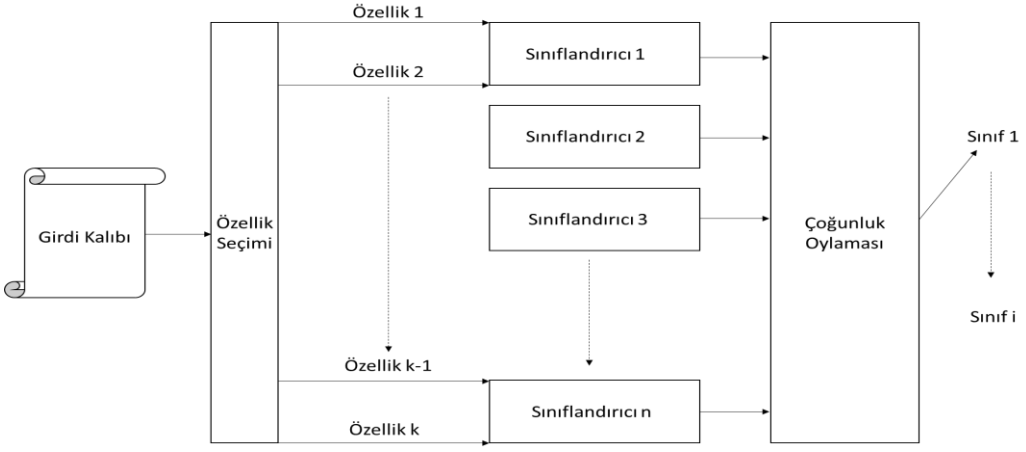
2.3.6. Bileşik Makine Öğrenme

Bileşik Makine Öğrenme sınıflandırmaya dayalı öğrenen sistemlerde, tahminlerdeki doğruluğu artırmak için kullanılan güçlü tekniklerdir. Temel sınıflandırıcı olarak makine öğrenmesi yöntemlerinden faydalanırlar (Özkan, 2008).

Bileşik Makine Öğrenme algoritmaları, klasik olarak uygulanan tek bir sınıflayıcı yerine birden çok sınıflayıcının içinde bulunduğu bir eğitim kümesi oluştururlar. Bu küme içerisindeki her bir sınıflayıcı birleştirilerek tek bir toplu sınıflayıcı elde edilir (Chaovalitwongse vd., 2007).

Tahmin edici performansındaki önemli iyileştirmeler, birden fazla modelin tek bir sınıflandırıcıya birleştirilmesi durumunda, bireysel modellerin sonuçlarının ortalaması alınarak veya oylama yapılarak gerçekleştirilebilir. Bireysel modellerin tahmin edilecek sınıflara göre tahmin performanslarında çeşitliliği önemlidir, böylece bir modelin zayıflığı başka bir modelin

gücüyle telafi edilmektedir. Bir veri kümesinden bir dizi tekrarlayıcı bölümleme sınıflandırıcısı geliştirerek ve her sınıflandırıcıdan gelen tahminlerin ortalamasını alarak son sınıflamayı oluşturmak suretiyle belirgin şekilde daha kuvvetli tahmini modellerin elde edilebileceği yapılan çalışmalarla gösterilmiştir (Sarmanova, 2013).



Şekil 2.19: Bileşik sınıflandırma makine öğrenme modelinin yapısı (Hsieh vd., 2012).

Bileşik öğrenmenin sezgisel kavramı, tek bir yaklaşımın veya sistemin diğerlerinden daha üstün olabileceğini iddia edemeyeceği ve birkaç tek yaklaşımın entegrasyonunun final sınıflandırıcısının performansını artıracağıdır (örn. doğruluk, güvenilirlik, anlaşılabilirlik). Dolayısıyla, bir bileşik sınıflandırıcı, bireysel taban sınıflandırıcılarından genel olarak daha iyi bir performans gösterebilir. Bileşik yöntemlerin etkililiği, temel öğrenen tarafından gerçekleştirilen hatanın bağımsızlığına oldukça bağımlıdır. Bileşik yöntemlerin performansı temel öğrenicilerin doğruluğuna ve çeşitliliğine güçlü bir şekilde bağlıdır (Tan ve Gilbert, 2003). Bileşik öğrenme, çeşitli temel sınıflandırmalar seti üretmek için tek veya çoklu öğrenme algoritmaları kullanmakta ve daha sonra bunları, yeni veri örneklerini sınıflandırmak için daha uygun kararlar vermek üzere birleştirmektedir. Pek çok araştırma, belirsizlikler altında verileri sınıflandırma doğruluğunu iyileştirmek için bileşik öğrenmenin umut verici yönünü göstermiştir. Bir topluluğun bireysel üyelerinden daha iyi performans göstermesi için gerekli ve yeterli şart, temel sınıflandırıcıların doğru ve çeşitli olmasıdır. Doğru sınıflandırıcı, yeni örnekler için sınıfları rastgele tahmin etmekten daha iyi bir hata oranına sahip olan bir sınıftır ve eğer iki sınıflandırıcı, ortak veri örnekleri üzerinde farklı (korelasyonsuz) hatalar yaparsa bunlar birbirinden farklıdır. Yapılan çalışmalarda çeşitli bileşik teknikler önerilmiştir. Buna ek olarak, bileşik yöntemlerin sınıflandırmaların doğruluğunu önemli ölçüde artırabilecekleri ve

daha iyi bir sınıflandırma performansı öngörmeleri konusunda güçlü kanıtlar elde edilmiştir (Hsieh vd., 2012).

Sınıflandırma problemleri için, bileşik öğrenme algoritmaları sınıflandırıcılar dizisi oluşturur ve daha sonra tahminlerini (ağırlıklı) oy alarak sınıflandırır. Genel olarak, bileşik öğrenme yöntemleri algoritmadan bağımsızdır ve temel öğrenicinin seçimine herhangi bir sınırlama getirmemektedir. Bir sınıflandırıcının tahmin hatasının iki bileşen halinde ayrışabileceği gösterilmiştir: önyargı ve varyans. Bireysel karar ağaçları, genelleme doğruluğu açısından yüksek değişkenlik göstermektedir. Böylece, karar ağacında bileşik öğrenmenin uygulanması varyansı azaltarak performansı artırabilir (Yılmaz, 2014).

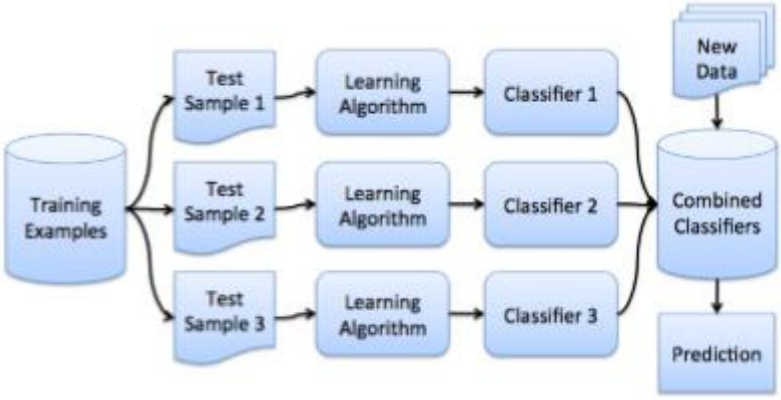
Bileşik öğrenme algoritmalarının tasarımında iki temel yaklaşım vardır: Ilk yaklaşım, her bir hipotezin bağımsız olarak, hipotez setinin sonucunun doğruluğunu ve çeşitliliğini gösterecek şekilde inşa edilmesidir- yani, her bir hipotez, yeni tahminler yapmak için makul düşük hata oranına sahiptir ve hipotezlerin uyumsuzlukları iptal edilmektedir. Bu tür bileşik yöntemler hem istatistiksel hem de hesaplama problemlerini aşabilir. Hipotezlerin ağırlıklı oyunun verilere iyi uyum sağlaması için bileşik yöntemleri tasarlamaya yönelik ikinci yaklaşım, hipotezleri birbirine bağlı bir şekilde yapılandırmaktadır (Dietterich, 2002).

Bileşik algoritmalar için öğrenme işlemi iki kısma ayrılabilir (Witten et al., 2016):

1. <u>Temel sınıflandırıcıları (modelleri) oluşturmak:</u> Bu bölümün temel görevleri şunlardır:
   a. Veri işleme, orijinal eğitim verilerini bozarak sınıflandırıcılar için girdi eğitim verilerini hazırlamak ve
   b. Temel sınıflandırıcı yapıları, temel öğrenci olarak bir öğrenme algoritması ile pertürbasyon verileri üzerine temel sınıflandırıcılar oluşturmaktır.
2. <u>Oylama:</u> Bileşik yöntemlerin ikinci aşaması, önceki aşamada oluşturulan temel modelleri, son bileşik modele birleştirmektir. Çeşitli oylama sistemleri vardır. Ağırlıklı oy kullanma ve ağırlıksız oy kullanma olmak üzere iki ana oylama sistemi genel olarak kullanılmaktadır. Ağırlıklı oylama sisteminde her temel sınıflandırıcı farklı oy verme gücüne sahip olmaktadır. Öte yandan, ağırlıksız sistemde bireysel taban sınıflandırıcı eşit ağırlığa sahip olmaktadır ve kazanan en çok oy alan gruptur (Adıgüzel Tüylü, 2017).

Bagging algoritması, özgün eğitim setlerinin alt kümeleri ile çalışmaktadır. Her bir alt küme, bir eğitim seti olarak kullanılmakta ve bir sınıflandırıcı oluşturmaktadır. Bu sınıflandırıcılar bir

birleşik sınıflandırıcı tarafından bileştirildiği için bu algoritmaya Bagging (torbalama) ismi verilmiştir (Bauer ve Kohavi, 1998).



Şekil 2.20: Bagging sınıflandırması süreci (Özkan, 2008).

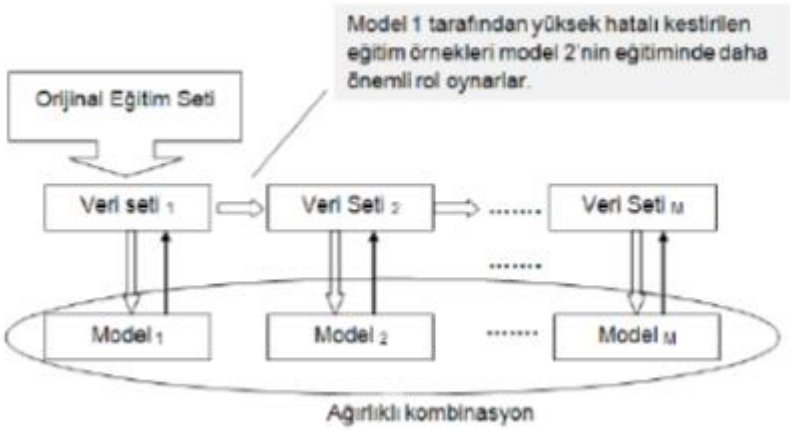
ViaRegression algoritması, karar ağaçlarının bir türü olan lineer regresyon fonksiyonları ile karar ağacının yapısını oluşturmaktadır (Özhan, 2014).

CostSensitiveClassifier algoritması, temel sınıflandırıcısını maliyete duyarlı hale getiren bir meta sınıflandırıcıdır. Maliyet hassasiyetini ortaya koymak için iki yöntem kullanılmaktadır; her sınıf için belirlenen toplam masrafa göre eğitim örneklerini yeniden ağırlıklandırmak veya minimum beklenen yanlış sınıflandırma maliyetiyle sınıfı tahmin etmektir. Temel sınıflandırıcının olasılık tahminlerini iyileştirmek için genellikle bir Bagging sınıflandırıcı kullanarak performansı iyileştirilebilmektedir (Turney, 1995).

AdditiveRegression algoritması, regresyon tabanlı bir sınıflandırıcının performansını arttıran bir meta sınıflandırıcıdır. Tahmin, her sınıflandırıcının tahminlerini ekleyerek gerçekleştirilmektedir. Öğrenme oranı parametresinin azaltılması, gereğinden fazla uyumu önlemeye yardımcı olmakta ve yumuşatma etkisi göstermektedir, ancak öğrenme süresini artırmaktadır (Stone, 1985).

FilteredClassifier algoritması, rastgele bir filtreden geçirilen veriler üzerinde rastgele sınıflandırıcı çalıştırmak için kullanılan bir yöntemdir. Sınıflandırıcı gibi, filtrenin yapısı da yalnızca eğitim verilerini temel almakta ve test örnekleri, yapılarını değiştirmeden filtre tarafından işlenmektedir (Witten vd., 2016).

Boosting algoritması, herhangi bir öğrenme algoritmasının performansını artırmak için genel bir yöntemdir. Teorik olarak, Boosting sürekli rastgele tahminden biraz daha iyi olması gereken sınıflandırıcıları sürekli üreten "zayıf" öğrenme algoritmasının hatasını önemli ölçüde azaltmak için kullanılabilmektedir. Boosting, eğitim verisi üzerindeki çeşitli dağılımlar üzerinde belirli zayıf bir öğrenme algoritmasını tekrar tekrar çalıştırarak ve zayıf öğrencinin ürettiği sınıflayıcıları tek bir bileşik sınıflandırıcıya birleştirerek çalışmaktadır. Teorik sonuçlarla vaat edilen ilerlemenin potansiyel yararlarına rağmen, Boosting'in gerçek pratik değeri ancak "gerçek" öğrenme problemleri üzerindeki metodu test ederek değerlendirilebilmektedir. WEKA'da bir Boosting algoritması olan AdaBoostM1 algoritması yer almaktadır (Witten vd., 2016).



Şekil 2.21: Boosting sınıflandırması süreci (Özkan, 2008).

RandomSub Space algoritması, eğitim verisi üzerinde en yüksek doğruluğu korumakta ve karmaşıklık arttıkça genelleme doğruluğunu geliştiren bir karar ağacı tabanlı sınıflandırıcıyı oluşturmaktadır. Sınıflandırıcı, sistematik olarak, özellik vektörünün bileşenlerinin alt gruplarını rastgele seçerek oluşturulan birden fazla ağaçtan yani, rasgele seçilen alt uzaylarda inşa edilmiş ağaçlardan oluşmaktadır (T. K. Ho, 1998).

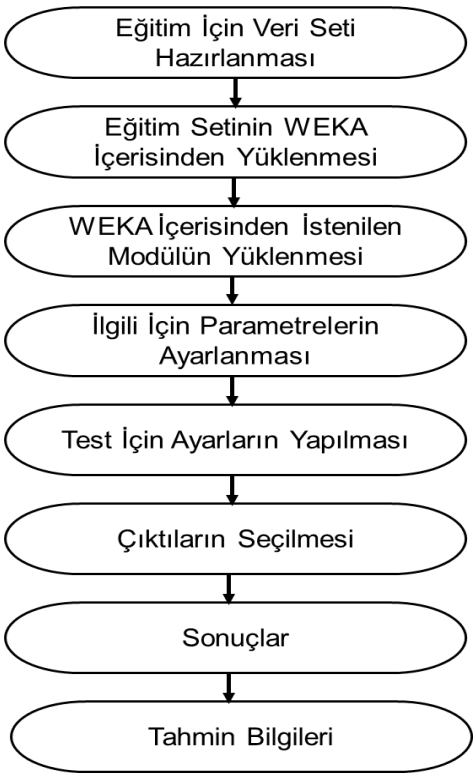
İstatistikteki en önemli görevlerden biri, bir tahminleyici veya sınıflandırıcının doğruluğunu değerlendirmektir. Genellikle kullanılan tahminin doğruluk ölçütü, genel olarak "tahmin hatası" olarak tanımlanan hata ölçüsüdür: Bir regresyon probleminde, hata, gerçek çıkış değeri ile buna karşılık gelen tahmini çıktı değeri arasındaki fark olan tahmin hatalarının karesinin ortalamasıdır; yani bir sınıflandırma probleminde, bir elemanın yanlış sınıflandırma olasılığıdır (Izenman, 2008). Gözlenen değerlerle gerçek değerler arasındaki farkın küçük olması sınıflandırıcın doğru sınıflandırdığını göstermektedir (Özhan, 2014).

3. MALZEME VE YÖNTEM

3.1. WEKA

WEKA, Waikato Üniversitesi'nde geliştirilen ve bilgisayar biliminin önemli alanlarından birisi olan makine öğrenmesinde kullanılan yazılımdır. Bu yazılım açık kaynak kodlu olarak Java dili üzerinde geliştirilmiştir ve GPL (General PublicLicense - Genel Kamu Lisansı) lisansı ile dağıtılmaktadır. İsmi "Waikato Environment for Knowledge Analysis" kelimelerinin baş harflerinden oluşur. WEKA (Bilgi Analizi için Waikato Ortamı), gerçek dünya veri setlerine standart makine öğrenme teknikleri uygulanmasını sağlayan bir programdır. Gerçek dünya veri setleriyle çalışmak için interaktif bir ara yüz vasıtasıyla çeşitli makine öğrenme tekniklerine kolay erişim sağlayan bütünleşik bir ortam sağlamak için geliştirilmiştir (Hall vd., 2009).

WEKA üzerinde makine öğrenmesi ve istatistik ile ilgili pek çok kütüphane hazır olarak gelmektedir. Örneğin veri ön işlemesi (data preprocessing), ilkelleme (regression), sınıflandırma (classification), gruplandırma (clustering), özellik seçimi veya özellik çıkarımı (featureextraction) bunlardan bazılarıdır. Ayrıca bu işlemler sonucunda çıkan neticelerinde görsel olarak gösterilmesini sağlayan görüntüleme (visualization) araçları bulunmaktadır. WEKA, makine öğrenme algoritmaları ve veri ön işleme araçlarının organize bir koleksiyonudur. Her makine öğrenme algoritması, verilerin kendi formatında bulunmasını gerektirir ve parametrelerin ve çıktıların belirlenmesinde kendine has bir yol vardır. WEKA sistemi, tutarlı bir yöntemle bu verilere kolayca uygulanabilmesi için ortak bir ara yüz altında bir dizi makine öğrenme tekniği veya şeması getirmek üzere tasarlanmıştır. Bu ara yüz, yeni şemaların eklenmesini teşvik edecek kadar esnektir (Yıldız and Şeker, 2016; Garner, 1995).



Şekil 3.1: WEKA'da temel işlem adımları (Singhal ve Jena, 2013)

3.2. M5P ALGORİTMASI

Bir karar ağacı, dal düğümlerinin çeşitli tercihler arasında bir seçimi temsil ettiği ve her yaprak düğümlerinin bir sınıflandırma sonucunu temsil ettiği bir ağaçtır. M5P modeli, karar ağacı ve çoklu doğrusal regresyonu algoritmalarını beraber kullanmaktadır. Karar ağacı, gözlemlenen girdileri kategorik veya kategorik olmayan veriler için geçerli olan öğrenme değeri ile tahmini çıktıları ilişkilendirmektedir (Graczyk vd., 2010). Karar ağacı girdi ve çıktılar arasındaki mantıksal ve regresyon tipi kuralları çıkarmak için girdi ve çıktı verilerini maksimum (Bayesçi) entropi ile sınıflandırmaktadır (Holmes vd., 1999). M5 olarak adlandırılan "model ağacı", öğrenme problemleriyle başa çıkmak için ortaya konulmuştur. Model ağaçları karar ağacı ve doğrusal regresyon fonksiyonlarının bir kombinasyonu olarak tasarlanmıştır. Young ve Witten, ağaç boyutunu azaltmak için orijinal M5 algoritması üzerinde veri kümesinde daha iyi performans gösteren M5P'yi geliştirmiştir. M5 ağacını kurmak için öncelikle bir karar ağacı algoritması kullanılmaktadır (Wang ve Witten, 1997; Quinlan, 1992).

Bir M5 ağacı oluşturmak için bölme ölçütleri tanımlanmalıdır. Dallanma ölçütlerinin bölünmesi, özellik değerlerinin standart sapmasına dayanmaktadır. Beklenen hatayı azaltan öznitelik, ağacın kökü olarak seçilmektedir. Standart sapma azaltma (SDR) formülü aşağıdaki gibi hesaplanmaktadır *(Adıgüzel Tüylü, 2017):*

Burada,

T: nitelik değerleri seti

T_i: seçilen özniteliğe göre bölünmüş düğümden alınan öznitelik değeri

$\bar{T}$: T özniteliğinin kümelerinin ortalama değeri

sd(T): T'nin standart sapması

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \text{x sd}(T_i) \tag{3.1}$$

$$sd(T) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(T_i - \bar{T})^2} \tag{3.2}$$

$$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i \tag{3.3}$$

Ağacı kurduktan sonra, sınıflandırma performansını artırmak için ağaç budama yapılmalıdır. Budama işlemi, öğrenme verilerinde hata oluşturan dalları silmektedir. Ağaç oluşturma işlemi başlar başlamaz, ağaç sürekli bir durma kuralı olmadan bölünerek büyümektedir. Yapı sona erdiğinde budama yapraklardan köke doğru başlamaktadır. Her budama işleminden sonra en başarılı ağaç belirlenmektedir (Adıgüzel Tüylü, 2017)

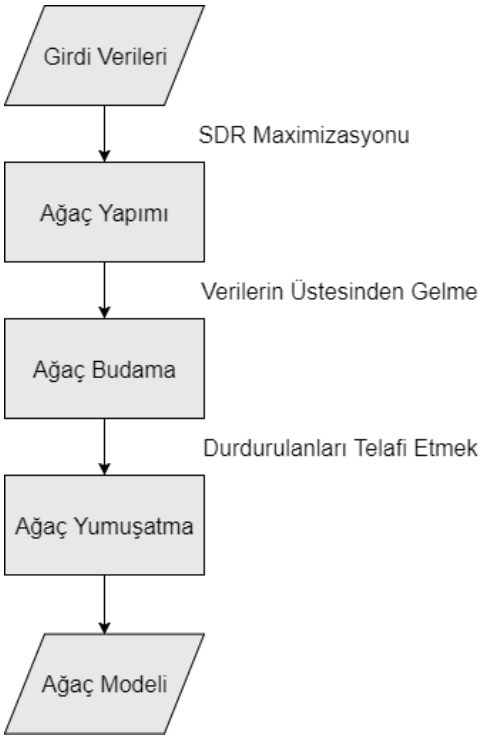
M5 algoritmasında, budama işlemi için her şeyden önce, gerçek sınıf değeri ile tahmin edilen değer arasındaki fark, her eğitim örneğinde düğüme ulaşıldığında ortalama alınmaktadır. Bu ortalama değer, bu katsayı ile çarpılmaktadır.

$$\frac{n+v}{n-v} \tag{3.4}$$

Burada,

n: düğümdeki eğitim örneklerinin sayısı

v: düğümdeki sınıf değerini temsil eden parametre sayısı



Şekil 3.2: M5P Yöntemi (Wang ve Witten, 1997; Quinlan, 1992)

3.3. M5RULES ALGORİTMASI

M5 kuralı, kural tabanlı bir öğrenme tekniğidir ve nominal ve sayısal değerleri tahmin edebilmektedir.M5 kural setleri model ağaçlarından oluşturulmaktadır. Model ağaçları sayısal ve nominal değerleri tahmin etmek için yeterlidir. Model ağaçları M5 kuralında kullanılmaktadır. Kural algoritması, model ağacı oluşturma işlemini tekrar ederek ve her çevrimde en iyi kuralı seçmeye çalışarak çalışmaktadır. Bir M5 kuralı oluşturmak için aşağıdaki adımlar uygulanmaktadır (Quinlan, 1992):

- Tüm eğitim verileri için bir M5 ağacı öğrencisi uygulanır,
- M5P gibi, ağaç budanır,
- En iyi yaprak kural haline getirilir,

- Bütün örnekler kurallara dahil edilinceye kadar devam eder. Bir örnek aynı anda farklı kurallara dahil edilebilir.
- Kategorik tahminde M5 Kuralları kısmen araştırılan ağaçlar yerine tam ağaçlar oluşturur.

3.4. LWL (YEREL AĞIRLIKLI ÖĞRENME) ALGORİTMASI

LWL, parametrik olmayan yöntemlerde kullanılan bir algoritma türüdür. Verilerin sadece bir alt kümesini ve belirli bir yerelliğe ait fonksiyonunu kullanarak girdi değişkenlerinin sınıflandırılmasını sağlar. Tüm girdi değerleri için bir fonksiyon oluşturmak yerine her bir ilgi alanı için yerel model oluşturur. Bu model oluşturma aşamasında referans noktanın komşu verilerine başvurur. LWL aynı zamanda bir tembel öğrenme algoritmasıdır (Englert, 2012).

3.5. KARAR TABLOSU ALGORİTMASI

Etiketli bir eğitim seti verildiğinde, tümevarım algoritması kullanılarak bir hipotez oluşturmaktadır. Burada gözlemlenen değer, çoğunluk sınıfına varsayılan kural eşlemesi olan bir Karar Tablosu (DecisionTable)'dur. Bu karar tablosunun iki bileşeni vardır (Kohavi, 1995):

1. Bir dizi özellik olan bir şema.
2. Etiketli örneklerin çoklu seti olan bir gövde.

Her örnek, şemadaki her bir özellik için bir değer ve etiket için bir değerden oluşmaktadır. Etiketsiz bir örnek (test örneği), $\tilde{X}$verildiğinde, bir karar tablosu sınıflandırıcısı tarafından örneğe atanan etiket aşağıdaki gibi hesaplanmaktadır (Adıgüzel Tüylü, 2017):

Karar tablosunda, sadece şablondaki özelliklerin eşleştirilmesi ve diğer tüm özelliklerin yok sayılması gereken, verilen örneğe tam olarak uyan etiketli örnekler kümesi olsun, $I = \emptyset$ ise; karar tablosundaki çoğunluk sınıfına geri dön, aksi halde, bağları keyfi olarak kırarak, çoğunluk sınıfını I'e geri döndür şeklindedir. Bilinmeyen/eksik değerler, eşleme işleminde farklı değerler olarak değerlendirilmektedir.

Tümevarım algoritmasının amacı, karar tablosu için en uygun özellik alt kümesini bulmaktır. $\tilde{X} = (X_1, \dots \dots, X_n)$bir dizi özellik ve D eğitim seti, özelliklerinin bir alt kümesi olsun. Tümevarım algoritmasının amacı, bir şema seçmektir. $\tilde{X}_{opt}$şu şekilde formülize edilmektedir:

$$\tilde{X}_{opt} = \underset{\tilde{X}' \subseteq \tilde{X}}{\arg\max}\, acc(DTM(\tilde{X}', D)) \quad (3.5)$$

Burada acc, karar tablosunun XxYuzayından rasgele seçilen bir örneği doğru olarak sınıflandırma ihtimalini göstermektedir.$\tilde{X}_{opt}$şeması, eğitim setindeki tüm örneklerin karar tablosunun gövdesinde saklandığı varsayımı altında bir karar tablosu için en uygun özellik alt kümesinden oluşmaktadır.

3.6. DESTEK VEKTÖR REGRESYONU ALGORİTMASI

Destek Vektör Makineleri (SupportVector Machine), denetimli bir öğrenme yöntemidir. Sınıflandırma problemlerinde kullanılmaktadır. Destek vektör makineleri, geleneksel öğrenme yöntemlerine kıyasla doğrusal olmayan problemleri çözmede daha iyi performansa sahiptir. Yaygın olarak sınıflandırma problemleri için kullanılan destek vektör makinelerinin regresyon için uyarlaması Destek Vektör Regresyonu olarak adlandırılmıştır (Basak vd., 2007).

3.7 DESTEK VEKTÖR MAKİNELERİ ALGORİTMASI

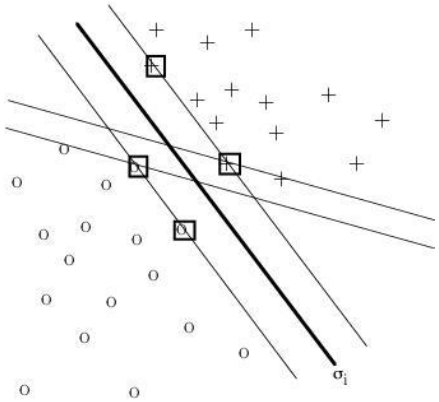
Destek Vektör Makineleri (DVM), doğrusal veya doğrusal olmayan bir fonksiyon yardımıyla sınıflandırma yapmaktadır. DVM, veriyi birbirinden ayıran en uygun fonksiyonu tahmin etmektedir (Namlı, 2012; Durmuşoğlu ve Kahraman, 2016).

DVM formülleri aşağıdaki şekilde açıklanmıştır (Shmilovici, 2009):

Bir H_0 hiper düzlemi üzerindeki noktaları bulmak için 4.1'de görülen denklemin, buhiper düzlemin iki yanındaki düzlemler arasındaki boşluğun $y_i(w^T x_i + b > 1)(i = 1,2, \dots \dots, n)$ kısıtı altında maksimum olması için 4.2'de görülen denklemin çözülmesi gerekmektedir.

$$H_0 = W^T X + b \quad (3.6)$$

$$w^T w/2 \quad (3.7)$$



Şekil 3.3: İki farklı veri sınıfının doğrusal olarak ayrılması (Steinwart ve Christmann, 2008; Shmilovici, 2009).

l tane örnekten oluşan bir eğitim kümesi olsun,

$$(x_1, y_1), \dots \dots, (x_n, y_n), \quad x_i \in R^d, \quad y_i \in \{-1, +1\} \tag{3.8}$$

(4.2)'deki eğitim kümesi uygun *w* ağırlık vektörü ve *b* eşik değeri için (4.4)'teki eşitsizlikleri sağlıyorsa o zaman denklem doğrusal olarak ayrılabilmektedir.

$$wx_i + b \geq 1 \qquad y_i = 1$$

$$wx_i + b \leq -1 \qquad y_i = -1 \tag{3.9}$$

Denklem (4.3)'teki eşitsizlik genelleştirilirse:

$$y_i(wx_i + b) \geq 1, \quad i = 1, \dots \dots, l \tag{3.10}$$

Karar sınırı denklemi aşağıda gösterilen en iyi ayırıcı doğrudur:

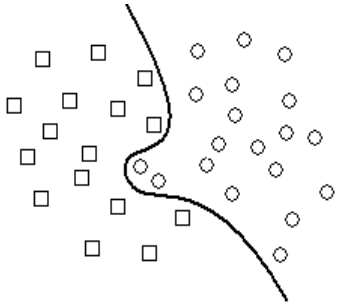
$$wx_i + b = 0 \tag{3.11}$$

Ayırt edici düzlem üzerindeki iki karar sınırı arasındaki mesafe aşağıdaki gibi yazılabilir.

$$d = \frac{2}{\|w\|} \tag{3.12}$$

w'nin normu $\|w\|$ ile gösterilir. Maksimum d mesafesi, minimum $\|w\|$'ye eşit olduğu için en uygun ayırt edici düzlem denklem (3.10)'daki kısıtları sağlayan ve minimum $\|w\|$elde eden optimizasyon problemi ile bulunabilir.

Verilerin doğrusal olarak ayrılamadığı durumlarda doğrusal olmayan sınıflandırıcılar kullanılır. Negatif olmayan ve hataları ifade eden ζ gevşek değişkenlerin optimizasyon modeline eklenmesi sağlanarak bu durum çözülmektedir.



Şekil 3.4: Doğrusal olarak ayrılamayan verilerin sınıflandırılması (Steinwart ve Christmann, 2008).

$\xi_i$gevşek değişkeni bir uygun sınıfın sınırdan sapma miktarıdır. Bir hatanın oluşması için $\xi_i >$ 1 olmalıdır. Ayırt edici düzlem kısıtları olan denklem (3.10)'daki ifadeye $\xi_i$gevşek değerleri eklenmektedir.

$$\begin{aligned} &y_i(wx_i + b) \geq 1 - \xi_i, \quad && i = 1, \dots \dots, l \\ &\xi_i \geq 0 && i = 1, \dots \dots, l \end{aligned} \tag{3.13}$$

$\sum_i \xi_i$Eğitim hataları sayısının üst sınırıdır. Buna göre en iyi ayırt edici düzlem,

$$P = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \tag{3.14}$$

Denklem (3.14)'deki parametresi sabit bir düzenleyicidir. Denklem (3.15) ile problem Lagrange çarpanları ile ikili probleme dönüştürülmelidir.

$$L(\alpha) = \sum_{i=1}^{l}\alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\alpha_i\alpha_j y_i y_j (x_i x_j) \tag{3.15}$$

$$\sum_{i=1}^{l} \alpha_i y_i \quad ve \quad 0 \leq \alpha_i \leq C \quad i = 1, \dots \dots, l \tag{3.16}$$

Sınıflandırma parametrelerini hesaplamak için (3.16) denklemi maksimum yapılmalıdır.

3.8. DOĞRUSAL REGRESYON ALGORİTMASI

Doğrusal regresyon analizi, bir sonuç değişkeni *y* ve tek bir tahmin değişkeni *x* içermektedir. Regresyonun en basit şeklidir ve *x*'in doğrusal bir fonksiyonu olarak *y*'leri modellemektedir (Kutner vd., 2004).

Doğrusal regresyon formülleri aşağıdaki şekilde açıklanmıştır:

$$y = b + wx \tag{3.17}$$

Burada *y*'nin varyansının sabit olduğu, *b* ve *w* değerlerinin Y-kesitini ve eğimini belirten regresyon katsayıları olduğu varsayılmaktadır. Regresyon katsayıların *w* ve *b*, aynı zamanda ağırlık olarak düşünülebileceği için eşdeğer olarak yazılabilir,

$$y = w_0 + w_1 x \tag{3.18}$$

Bu katsayılar, gerçek veri ile tahmin doğrusu arasındaki hatayı asgari düzeye indirgeyen doğruyu tahmin eden en küçük kareler yöntemiyle çözülebilir. *D*, tahmin değişkeni *x* ve bununla ilişkili sonuç değişkeni *y*'den oluşan bir eğitim setidir. Eğitim seti $|D|, (x_1, y_1), (x_2, y_2), \dots \dots, (x|D|, y|D|)$formunun veri noktalarını içermektedir. Regresyon katsayıları aşağıdaki denklemlerle bu yöntemi kullanarak hesaplanabilir:

$$w_1 = \frac{\sum_{i=1}^{|D|} (x_i - \bar{x})\,(y_i - \bar{y})}{\sum_{i=1}^{|D|} (x_i - \bar{x})^2} \tag{3.19}$$

$$w_0 = \bar{y} - w_1 \bar{x} \tag{3.20}$$

Burada x $x_1, x_2, \dots \dots, x|D|$'nin ortalama değeri,$y$ ise$y_1, y_2, \dots \dots, y|D|$'nin ortalama değeridir. Katsayılar $w_0$ve w_1, genellikle karmaşık regresyon denklemlerine iyi yaklaşımlar sağlamaktadır.

3.9. BAGGİNG ALGORİTMASI

Breiman'ın Bagging algoritması, karar ağaçlarının birleşmesinin ilk örneklerinden biridir. Bagging ve Boosting gibi topluluk öğrenme algoritmalarının çoğu denetlemeli öğrenme için tasarlanmıştır. Bagging yöntemi, aynı algoritmayı verinin iadeli örnekleme ile seçilen farklı altkümeleri üzerinde çalıştırmaktadır. Boyutu n olan bir yineleme veri kümesi, n kalıplarının orijinal veri kümesiyle değiştirilerek rastgele çizilmektedir. Bir önyükleme örneği D_1, D içine birden çok kez dahil olabilir, ancak diğerleri dahil edilmez. Önyükleme fikri, gerçek veri kümesi D'den örneklemenin bilinmeyen dağılım P'den örnekleme için mümkün olan en iyi yaklaşımıdır. Ardından, bu önyükleme veri kümesi kullanılarak bir model oluşturulmaktadır. Bu işlem T kez tekrarlanmakta ve böylece T modelleri elde edilmektedir. Daha sonra T modelleri, regresyon problemleri için ortalama kullanılarak toplanmaktadır. Bagging yöntemi, bir regresyon problemi için şu şekilde çalışmaktadır (Aydogmus vd., 2015).

Bagging yaklaşımı aşağıda verilen ifadelerle izah edilmiştir:

A,$i = 1,2, \dots \dots, n$ olmak üzere (X_i, Y_i)içeren D'nin eğitim setidir. buradaX_i çok boyutlu öngörü değişkeninin bir gerçekleştirmesidir ve $Y_i$gerçek değerli bir değişkenin gerçekleşmesidir.Bir tahminleyici$E(Y|X = x) = f(x)$şöyle gösterilir;

$$C_n(x) = h_n(D_1, \dots \dots, D_n)(x) \tag{3.21}$$

Teorik olarak Bagging aşağıdaki gibi tanımlanır: İlk olarak, önyüklemeli bir örnek oluşturulur. $D_i^* = (Y_i^*, X_i^*)$çiftlerin ampirik dağılımına göre$(i = 1,2, \dots \dots, n)$ olmak üzere$D_i^* = (Y_i^*, X_i^*)$ olur. İkinci olarak, ön yüklenmiş öngörme eklenti prensibi ile tahmin edilmektedir,

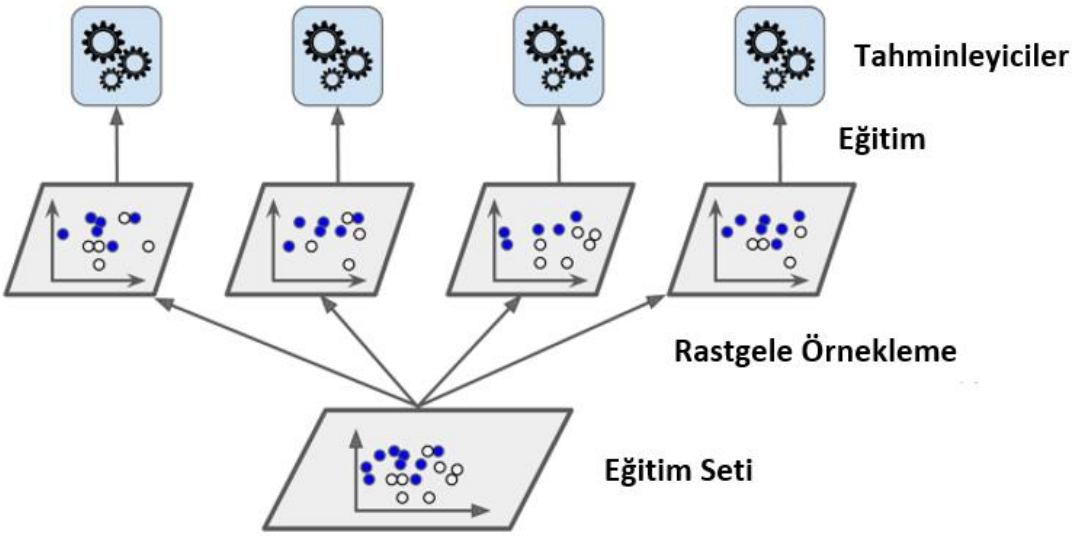
$$C_n^*(x) = h_n(D_1^*, \dots \dots, D_n^*)(x) \tag{3.22}$$

$$C_n\ (x) = h_n\big(D_1\ , \dots \dots, D_n\ \big)(x) \tag{2.23}$$

iken sonuç olarak Bagging öngörücü;

$$C_{n;B}(x) = E|D_n^*(x)| \tag{3.24}$$

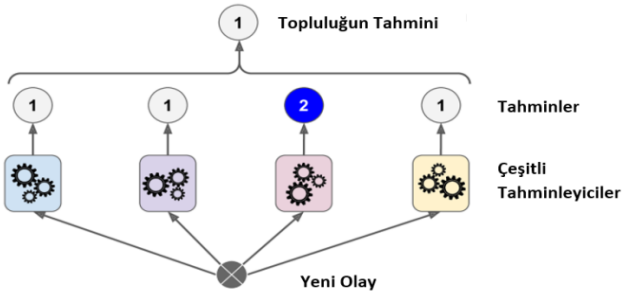
Bagging, tatmin edici bir performans ile temel öğreniciler ile kombine edildiğinde varyansı azaltabilen en basit uygulamalardan biridir (Aydogmus vd., 2015).



Şekil 3.5: Bagging Yöntemi (Aydogmus vd., 2015).

3.10. VOTİNG (OYLAMA) ALGORİTMASI

Birden çok sınıflandırıcıyı birleştirerek daha iyi bir sınıflandırıcı elde eden bileşik öğrenme yöntemlerinden en basit ve sık kullanılanlarından biri voting (oylama) yöntemidir (Koptur,2017). Birden çok (N adet) farklı sınıflandırıcının$C^1 \dots C^N$,aynı veri seti $\mathcal{L}$ üzerinde eğitilerek $(L^1 \dots L^N)$ elde ettikleri sınıflandırma sonuçları değerlendirilmektedir. En sık tekrarlanan yani en çok oyu alan sınıf tahmin sonucunu oluşturan sınıf olarak seçilmekte veya elde edilen sonuçlar olasılık değerleri şeklindeyse, bu sonuçların ortalaması alınarak tahmin sonucu elde edilmektedir. Sınıflandırıcıların oyları değerlendirilirken her birine eşit ağırlık verilebileceği gibi tüm veri seti kullanılarak eğitim yoluyla elde ettiği doğruluk değerlerine paralel olarak farklı ağırlık değerleri de verilebilmektedir.



Şekil 3.6: Voting (Oylama) Yöntemi

3.11. RANDOMSUBSPACE (RASTGELE ALTUZAYLAR) ALGORİTMASI

RandomSubspace algoritması aşağıdaki şekilde formülüze edilmiştir (T. K. Ho, 1998).

X, Y: Gözlemlenen veri

B: Alt küme ile çekilen veri sayısı

$|m| < M$: Alt uzayın büyüklüğü

TS_i^* : $\forall i$ için nihai t-tabanlı skor

$C_{i,k}$: Sayaç

- $m^* = \{i_1^*, \dots, i_{|m|}^*\}$modeli rassal olarak orijinal özellik alanından alınır.
- $\forall i \in m^*$içinT_{i,m^*}^2 hesaplanıp $y \sim x_{m^*}$modeli uyarlanır. $i \notin m^*$ise$T_{i,m^*=0}^2$ ayarlanır.
- $C_{i,k} = C_{i,k} + I\{i \in m^*\}$olarak güncellenir.

Bu işlemler $k = 1, \dots, B = B_n$'e kadar tekrarlanır; $n \to \infty$'a giderken $B_n \to \infty$ gider ve i için C_i, $0 = 0$'dan başlar.

TS_i^*şu şekilde hesaplanmaktadır:

$$TS_i^* = \frac{1}{C_i, B} \sum_{m^*: i \in m^*} \frac{T_{i,m^*}^2}{n - |m|} \tag{3.25}$$

$TS_i^*: TS_{i1}^* \geq TS_{i2}^* \dots \geq TS_{i,M}^*$ Skorlarına göre değişken listesi düzenlenir ve $\{i_1, \dots \dots, i_M\}$şeklinde sıralanır.

3.12. VERİ ANALİZİ

Bu çalışmada, lokasyon analizi ve kuruluş yeri seçimi kapsamında giyim perakendeciliğinde dünya çapında faaliyet gösteren, marka konumlandırması olarak orta sınıf segmente hitap eden ve düşük fiyat politikası izleyen öncü bir firmanın mağazalarının bulunduğu lokasyonlara ait il, ilçe ve çevre verileri incelenmiştir.

Verilerin elde edilmesi aşamasında lokasyon yeri seçimi modelinin girdi parametreleri, öncelikle işinde uzman olan iş analistleri ve mağaza çalışanları ile görüşülerek müşteri geliş sayılarını etkileyen, lokasyona ait özelliklerin neler olabileceğine yönelik yapılan kurum içi anket çalışması ile belirlenmiştir. Belirlenen bu parametreler veri analizi, deneme yanılma yöntemleri, parametre seçim algoritmaları ile de ayrıntılı olarak incelenmiş, böylece detaylı bir çalışma sonucu modelde kullanılacak girdi parametreleri ortaya konmuştur. Bu parametrelere göre TÜİK sisteminden ve googleplaces üzerinden veriler çekilmiş, elde edilen verilerde eksik ve aşırı uç verileri düzeltilmiştir. Yatırım için en değerli lokasyon nerelerdir? Sorusunun cevabının arandığı bu çalışmayla literatüre katkı sağlanması ve giyim perakende sektöründe zincir mağaza olarak faaliyette bulunan firmaların kullanabilecekler bir model ortaya konması amaçlanmıştır.

Bu amaçla perakende sektöründe faaliyette bulunan bir firmanın 61 il'in 96 ilçesinde bulunan 98 farklı mağazasının il, ilçe, 1000 m, 500 m, 250 m, 100m çevre verilerine ait aşağıdaki parametreler incelenmiştir.

- Toplam Nüfus
- Erkek Nüfusu
- Kadın Nüfusu
- Geçen Yıla Göre Nüf. Değişimi
- Geçen Yıla Göre Nüfus Artış Oranı
- Erkek Nüfusu Oranı
- Kadın Nüfusu Oranı
- Toplam Nüfus 2020
- Toplam Nüfus 2020 Değişimi
- Yaş Ortalaması
- AB Ses Grubu Nüfusu

- C Ses Grubu Nüfusu
- DE Ses Grubu Nüfusu
- C Oran
- DE Oran
- BC1
- Ortalama Hane Geliri
- Toplam Konut Sayısı
- Gelişmiş Konut Sayısı
- Orta Gelişmiş Konut Sayısı
- Gelişmemiş Konut Sayısı
- Toplam İşyeri Sayısı
- Yazlık Sayısı
- Toplam Yerleşim Sayısı
- Bölge Konutlaşma Oran
- Nüfus Yoğunluğu (kişi/ha)
- Bölge Yüzölçümü Km2
- Konut Yoğunluğu (Konut/Km2)
- İşyeri Yoğunluğu (işyeri/Km2)
- Toplam Hane Sayısı
- Ortalama Hane Büyüklüğü
- Toplam İstihdam
- Toplam İşsiz
- İş Gücünde Olmayanlar
- İstihdam Bilinmeyen
- Aktivite Oranı (Çalışan/İşgücü)
- Toplam Öğrenci
- Toplam Ev Kadını
- Toplam Emekli
- Toplam Irad (gelir) Sahibi
- Toplam Çalışan Nüfus
- İlmi ve Teknik Elemanlar, Serbest Meslek

- Direktörler ve Üst Kademe Yöneticiler
- İdari Personel ve Benzeri
- Ticaret ve Satış Personeli
- Hizmet İşlerinde Çalışanlar
- Otomobil Sayısı
- Seçmen Sayısı
- Geçerli Oy Sayısı
- AKP Oy Sayısı
- CHP Oy Sayısı
- MHP Oy Sayısı
- Diğer/Bağımsız Oy Sayısı
- AKP Oran
- CHP Oran
- MHP Oran
- Sol Oran
- Diğer/Bağımsız Oran
- Muhafazakâr Nüfus
- Ortalama Hane Harcaması
- Toplam Tasarruf
- Giyim Ve Ayakkabı ‘ya yapılan toplam harcama
- Sağlık Hizmetlerine yapılan harcama
- Ulaştırmaya yapılan harcama
- Haberleşmeye yapılan harcama
- Eğlence Ve Kültür faaliyetlerine yapılan harcama
- Yıl Geneli Ortalama Sıcaklık
- Zincir rakip sayısı
- Yerel rakip sayısı
- Perakende mağaza sayısı

Merkez	Kademe	Toplam Nüfus	Erkek Nüfusu	Kadın Nüfusu	Geçen Yıla Göre Nüf. Değişimi	Geçen Yıla G	Erkek Nufusı	Kadın Nüfusı	Toplam Nüfı	Toplam Nüfus 2020 Değişimi	Yaş Ortalaması
Lokasyon 1	1000 Metre	17752	8716	9036	-225	-1	49	51	20660	2908	35
Lokasyon 2	1000 Metre	25270	12583	12686	-634	-2	50	50	26620	1350	29
Lokasyon 3	1000 Metre	28840	14041	14798	840	3	49	51	21461	-7378	34
Lokasyon 22	1000 Metre	32339	15509	16830	-370	-1	48	52	35422	3083	38
Lokasyon 23	1000 Metre	9882	5044	4838	380	4	51	49	10617	735	35
Lokasyon 24	1000 Metre	55802	28630	27172	2047	4	51	49	60585	4783	32
Lokasyon 43	1000 Metre	12380	6435	5946	-387	-3	52	48	9453	-2927	26
Lokasyon 44	1000 Metre	34805	16859	17945	-294	-1	48	52	19024	-15780	32
Lokasyon 45	1000 Metre	10561	5395	5166	472	5	51	49	9174	-1387	28
Lokasyon 93	1000 Metre	24241	12030	12211	-533	-2	50	50	28028	3787	28
Lokasyon 94	1000 Metre	62251	31349	30902	-499	-1	50	50	68613	6362	33
Lokasyon 95	1000 Metre	18350	8839	9511	-194	-1	48	52	17055	-1294	36
Lokasyon 96	1000 Metre	19754	9764	9990	-122	-1	49	51	15457	-4297	32
Lokasyon 97	1000 Metre	15443	7664	7779	558	4	50	50	20306	4863	31
Lokasyon 98	1000 Metre	25497	12638	12859	-525	-2	50	50	22311	-3186	34

Tablo 3.1: Lokasyonlara ait demografi verileri

İncelenen bu parametrelerden lokasyon yer seçimine en uygun girdi değişkenleri aşağıdaki şekilde gruplandırılmıştır.

Tablo 3.2: Lokasyonlara ait seçilen parametreler

İlçe	1000 m	500 m
•Kent Nüfus •Kent Ekonomi •Kent Gelişmişlik	•Zincir Rakip Sayısı •Yerel Rakip Sayısı •Perakende Mağaza Sayısı •Ortalama m2 değeri •Retail Index •Foot Trafic •Ulaşım Noktası Sayısı	•Zincir Rakip Sayısı •Yerel Rakip Sayısı •Perakende Mağaza Sayısı

3.12.1. Sonuç Değişkeni (Müşteri Gelişleri)

Müşterilerin mağazaya geliş sayılarının fazla olması mağazanı iyi bir yerde konumlandırıldığını göstermektedir. Bu amaçla incelenen firmaya ait 97 mağazanın müşteri gelişlerinin büyüklüklerine göre değer segementasyonu çalışması yapılmıştır.

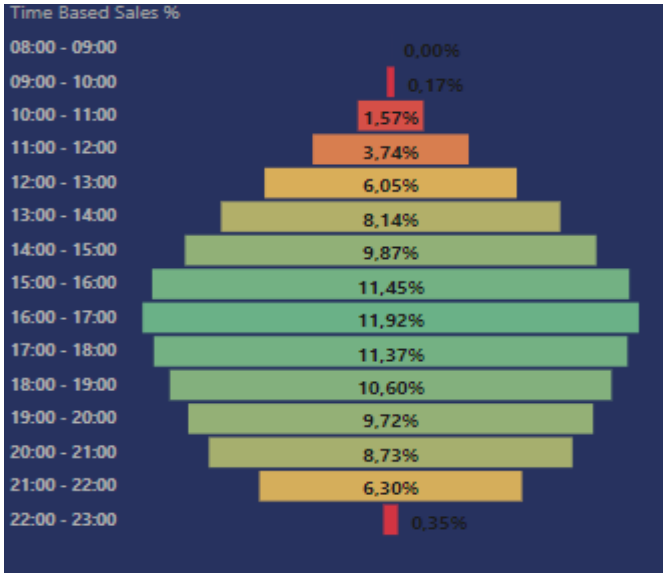
Tablo 3.3: Müşteri gelişlerine ait değer segmentasyonu veri seti

Lokasyon	Küme Değeri	Müşteri Gelişleri	Dönüşüm Oranı
Lokasyon 1	5	1.930.748	11,8%
Lokasyon 2	5	1.719.169	11,0%
Lokasyon 21	4	808.310	11,7%
Lokasyon 22	4	791.817	12,0%
Lokasyon 71	3	533.830	15,6%
Lokasyon 72	3	528.560	14,1%
Lokasyon 73	3	520.004	14,9%
Lokasyon 90	2	417.578	14,8%
Lokasyon 91	2	409.921	13,3%
Lokasyon 97	1	299.745	17,0%
Lokasyon 98	1	288.616	28,0%

Tablo 3.4 ‘de k-means yöntemi ile yapılan değer segmentasyonuna ait sonuç değerli bulunmaktadır. En fazla küme eleman sayısı %52 ile 3 no’lu küme de bulunmaktadır. 1 no’lu kümelerde ise gelen müşteri sayısı az olmasına rağmen müşteri dönüşüm oranı en yüksek mağazalar bulunmaktadır. Burada gezen müşteri profilinden daha çok marka algısına sahip müşterilerin alışveriş yaptığı söylenebilir.

Tablo 3.4: Değer segmentsayonu temel istatistiki göstergeleri

Küme Değeri	Müşteri Sayısı	Müşteri Geliş Dağılımı	Müşteri Dönüşüm Oranı	Mağaza Sayısı	Mağaza Sayısı %
1	2.316.520	3,47%	20,4%	7	7,14%
2	8.723.351	13,05%	19,6%	18	18,37%
3	32.665.517	48,87%	16,0%	51	52,04%
4	11.025.549	16,50%	12,6%	13	13,27%
5	12.105.796	18,11%	14,2%	9	9,18%
Toplam	**66.836.733**	**100,00%**	**16,3%**	**98**	**100,00%**



Şekil 3.7: Saatlik müşteri geliş dağılımları

Şekil 3.7'de müşterilerin mağazalara saatlik gelişlerine ait dağılım grafiği verilmiştir. Saat 16 – 17 arası mağazaların en yoğun olduğu saat aralıklarıdır.

Mağazanın çekim modeli için müşteri geliş sayıları referans alınarak bir inceleme yapılmıştır. Bunların satışa dönüş oranlarının yüksek olması mağaza içi davranış modeli ile açıklanmaktadır. Seçilen lokasyonlar üzerinde aşağıdaki kabuller dikkate alınarak mağaza davranış segmentasyonu oluşturulmuştur.

- Hafta sonu gün başı yoğunluk ortalamasının hafta içine oranı
- Akşam saatleri saat başı yoğunluk ortalamasının gün içine oranı
- Mağaza ürün kategori tercih oranları
- Ayların İlk Haftası ve Son Haftası durumları
- Özel Günlerdeki Satışların Normal Günlere Oranı

Tablo 3.5: Mağaza davranış segmentasyonu

Segment Adı	Mağaza Sayısı	Hafta İçi Ciro Oran	Sabah Ciro Oran	Öğlen Ciro Oran	Akşam Ciro Oran	Kapanış Ciro Oran	Erkek Ciro Oran	Kadın Ciro Oran	Çocuk Ciro Oran
AİLELERİN TATİL TERCİHİ	41	%51	%3	%20	%43	%24	%44	%43	%10
GENEL TERCİH	30	%54	%4	%21	%43	%21	%44	%45	%10
HAFTAİÇİ AİLE GİYİM TERCİHİ	17	%57	%5	%23	%42	%21	%42	%49	%4
ÇALIŞANLARIN TERCİHİ	9	%63	%6	%23	%37	%26	%39	%49	%9

3.12.2. Kent Nüfus

Nüfus perakende doygunluk verisi için en temel göstergelerden biridir. Pazar'ın büyüme hızını, açılacak olan mağazanın metrekare hacmini belirleyecek etkenler arasındadır. Nüfusla ilişkili parametreler arasında müşteri gelişleri ile arasında en fazla korelasyona sahip değerin ilçe kent nüfus parametresi olduğu belirlenmiştir.

Tablo 3.6: İl Türlerine göre ortalama nüfus verileri

İl Türü	Ortalama TOPLAM NUFUS	Ortalama Nufus 35 üstü	Ortalama SES BC1	Ortalama Kent Nufusu	Ortalama okuma yazma bilen
büyükşehir	370.274	153.886	134.736	345.238	322.843
büyükşehir değil	182.128	73.179	40.311	143.171	156.034
Grand Total	**312.742**	**129.207**	**105.862**	**283.449**	**271.835**

İl türü	StdSapmas TOPLAM NUFUS	StdSapmas Nufus 35 üstü	StdSapmas SES BC1	StdSapmas Kent Nufusu	StdSapmas okuma yazma bilen
büyükşehir	198.991	83.149	89.154	202.722	172.168
büyükşehir değil	80.554	30.965	23.161	74.560	67.413
Grand Total	**192.322**	**80.471**	**87.033**	**197.223**	**166.956**

	Toplam Nüfüs	Ortalama Nufu 35 üstü	Ortalma SES BC1	Ortalama Kent Nufus	Okuma Yazma Bilen
Büyükşehir	250.000	100.000	90.000	200.000	200.000
Büyükşehir Değil	140.000	50.000	25.000	100.000	100.000

Ciro, dönüşüm oranı, müşteri geliş sayıları incelenerek kent nüfusu aşağıdaki şekilde kategorize edilmiştir.

1- **Büyükşehirde mağaza açma kriteri**

a. Bir büyükşehirdeki toplam mağaza sayısı (özel projeler hariç) için üst sınır "300,000 kişiye 1 mağaza" dır.
b. Büyükşehir ilçeleri ile alakalı kriterler 2 no'lu başlıkta tanımlanmıştır.
c. Bir büyükşehirde sadece 1 adet cadde mağazası varsa, 2. mağaza AVM'de açılır.
d. Büyük şehirlerdeki "özel" projeler (sadece bir bölgeden değil, bütün şehirden müşteri çeken projeler) ayrıca değerlendirilir.

2- Büyükşehir ilçelerinde mağaza açma kriterleri

a. Bir büyükşehir ilçesinde nüfus> 200,000 ise mağaza açılabilir. *(1. derece öncelikli)*
b. Bir büyükşehir ilçesinde 2. mağaza nüfus> 450,000 ise açılabilir. *(2. derece öncelikli)*
c. Nüfusu> 100,000 ve <200,000 olan ilçeler için, ilçenin hane başına düşen geliri gelişmişlik oranı bulunduğu il ortalamasının üzerinde ise mağaza açılabilir. *(5. derece öncelikli)*
d. Nüfusu <100,000 olan büyükşehir ilçelerinde mağaza açmak uygun değildir.

3- İl Merkezlerinde mağaza açma kriterleri

a. Nüfusu 75,000'in üzerindeki il merkezlerinde 1 mağaza açılabilir. *(3. derece öncelikli)*
b. Nüfusu> 300,000 olan il merkezlerinde 2. mağaza açılabilir. 2. mağaza AVM'de açılır. *(4. derece öncelikli)*
c. Nüfusu> 50,000 ve <75,000 olan yerler için çok avantajlı kira ve lokasyon şartları var ise mağaza tutulabilir. *(7. derece öncelikli)*
d. Nüfusu <50,000 olan il merkezlerinde mağaza açılmaz.

4- İlçe merkezlerinde mağaza açma kriterleri

a. Nüfusu 75,000'in üzerindeki ilçe merkezlerinde 1 mağaza açılabilir. *(6. derece öncelikli)*
b. İlçe merkezlerinde en fazla 1 mağaza açılabilir.

Yukarıdaki kriterlere göre ilçeler kategorize edilmiş ve 5 ana gruba ayrılmıştır.

Tablo 3.7: Kent nüfusu ile küme değerleri eleman sayısı

Magaza Sayısı	Küme Değeri					
Kent Nüfus	1	2	3	4	5	Toplam
1	1	7	4			12
2	2	5	13			20
3		1	6			7
4			8	2	1	11
5	4	5	20	11	8	48
Toplam	**7**	**18**	**51**	**13**	**9**	**98**

Tablo 3.6'da kategorize edilmiş kent nüfusu verisi ile müşteri gelişlerine göre segmente edilmiş modeldeki eleman sayısı verileri bulunmaktadır. En fazla 5 no'lu kent nüfusu kategorisinde mağaza bulunmaktadır. Beklenen de mağaza açılış kararında bu davranışın sergilenmesidir. 1 ve 2 no'lu ilçelerde bulunulması gerekiyor fakat ciro ve karlılık az olacağından yatırım kararını riske etmeyecek diğer verilerin daha detaylı analizi gerektiğinden bu şekilde bir kural seti oluşturulmuştur.

Kent Nüfusu ile küme verileri arasındaki ilişki ki-kare testi ile analiz edildiğinde istatistiksel açıdan anlamlı bir data oluşmaktadır. Tablo 3.8 incelendiğinde kent nüfusu ile müşteri geliş sayılarına göre oluşturulmuş değer segmenti arasında anlamlı bir ilişki vardır.

3.12.3. Kent Gelişmişlik Değeri

İlgili veri TÜİK datası olarak alınmıştır. Gelişmişlik düzeyinin çok boyutlu bir kavram olduğu dikkate alındığında, bu tür endekslerin oluşturulmasında "temel bileşenler analizi"nin (principal components analysis-PCA) geniş ölçüde kullanıldığı görülmektedir. Bu sayede, gelişmişlik düzeyinin yorumlanmasında tek bir değişkene (GSYH, nüfus vb.) bağlı kalınması yerine daha fazla faktörün değerlendirmeye katılması sağlanmaktadır.

Gelişmişlik endeksi pazarın gelişme hızı ve hayat evresi hakkında önemli bilgiler vermektedir. Bölge ticari olarak ne kadar değere sahip bilgisi rekabetin düzeyini ve çekim merkezi yapısını kurgulatacak olan referans veri setini oluşturacaktır.

İncelenen mağazalara ait gelişmişlik değerleri aşağıdaki tabloda özet istatistiklerle verilmiştir. Mağazaların bulunduğu illerin ortalama gelişmişlik değerleri 46 civarındadır. Rakipler incelendiğinde rakiplerin bulunduğu lokasyonların gelişme endeksi ortalama 49 iken minimum 35 değerindeki yerlere mağaza açma kararı verilmiştir. İncelenen firmada ise bu değer 29'lara kadar düşmekle birlikte mağazalar bu yoğunluktaki alanlarda birikmiş durumdadır

Tablo 3.9: Kent gelişmişlik değeri temel istatistiki verileri

Gösterge Değerleri	Gelişmişlik Değeri
Ortalama	46
Standart Sapma	7
Mak.	67
Min.	29
En büyük 5 mağaza	51
En küçük 5 mağaza	45

Standart sapma değerleri göz önünde bulundurularak kent gelişmişlik değeri kendi içinde 5 kategoriye ayrılmıştır. Tablo 3.8 de bu ilişki modeli gösterilmektedir.

Tablo 3.10: Kent gelişmişlik kategorik verisi ile küme değerleri eleman sayısı

Magaza Sayısı	Küme Değeri					
Kent Gelismislik	1	2	3	4	5	Toplam
1		5	1			6
2	3	7	21	5		36
3	2	2	12	3	1	20
4	2	3	16	3	3	27
5		1	1	2	5	9
Toplam	**7**	**18**	**51**	**13**	**9**	**98**

3.12.4. Kent Ekonomi

İl ve İlçe merkezinde bulunan toplam iş yeri sayısı bölgeye ait aktivite oranını vermektedir. Analize dahil edilen ilçelerdeki iş yeri sayıları ile müşteri gelişleri arasında ilişki kurulmuştur.

İş yerindeki kent sayısı ve 50'den fazla personel çalıştıran iş yeri sayısı göz önünde bulundurularak kent iş yeri sayısı 5'li likert değerlerine ayrılmıştır. K-means yönteminin kullanıldığı yöntemde küme dışı heterojenlik verisini en iyi açıklayan küme sayısı 5 olarak değerlendirilmiştir. Hesaplanan bu değere kent ekonomi değişkeni adı verilmiştir.

Tablo 3.11: Kent ekonomi kategorik verisi ile küme değerleri eleman sayısı

Magaza Sayısı	Küme Değeri					
Kent Ekonomi	1	2	3	4	5	Toplam
1	3	13	15			31
2	3	2	30	7	3	45
3	1	2	6	4		13
4		1		1	6	8
5				1		1
Toplam	**7**	**18**	**51**	**13**	**9**	**98**

3.12.4. Ortalama m2 değeri

Ortalama metrekare değeri mağaza kira giderlerinin optimum yönetilmesi için incelenmesi gereken en önemli değişkenlerden bir tanesidir. Mağazanın çok merkezi bir yere açılması yapılacak kira yatırımının yüksek olmasına ve karlılığın düşük olmasına neden olacaktır. Cironun %7 'si ile %12'sini kira gideri olarak almayı göze alan firmanın 1000 m2 çevre verilerinden elde edilen değerler incelendiğinde

- Gelişmiş kenetlerde merkeze en yakın merkezi olmayan noktada
- Gelişmemiş kentlerde ise merkezi noktalarda

Mağazasını açtığı gözlemlenmiştir. Bu amaçla merkeze en yakın merkezi olmayan noktaların tespiti en önemli gösterge olarak seçilmiştir.

3.12.5. Foot Traffic (Yaya trafik yoğunluğu)

Yaya trafik yoğunluğu mağazanıza çekmeniz gereken müşteriler için daha az sermaye harcamanızı sağlayacaktır. Tek mağazada temel ihtiyaçları karşılamayı planlayan bir müşteri profili yaya olarak gezdiği lokasyonlarda bulunan mağazaları tercih etme davranışı sergilemektedir. 100'lük bir puantaja sahip olan veri setinde ortalama 40'lı yoğunluklarda mağaza açıldığı tespit edilmiştir.

Tablo 3.12: Küme değerlerine göre ortalama yaya yoğunluğu

Küme Değeri	Ortalama Yaya Yoğunluğu
1	17,286
2	45,500
3	40,314
4	39,154
5	35,778
Toplam	**39,051**

3.12.6. Retail Index (Perakende harcaması yoğunluğu)

TÜİK tarafından perakende mağazalarının bulunduğu yerlerdeki harcama yoğunluklarına göre hesaplanan bir değerdir. Bölgesel ağırlıklandırılması ile çevre verisi olarak hesaplanmıştır.

Tablo 3.13: Küme değerlerine göre ortalama perakende indeksi

Küme Değeri	Ortalama perakende indeksi
1	14,857
2	41,722
3	37,255
4	35,385
5	30,556
Toplam	**35,612**

3.12.7. 1000 m, 500 m nokta verileri

Mağazaların bulunduğu lokasyonların çevresinde ne kadar zincir rakip olduğu, perakende mağazalarının sayısı, ulaşım noktası sayısı mağazanın çekim etkisini arttıran değişkenlerdendir.

Mağazalara erişilebilirlik bakımından; yaya(toplu taşıma) ve araç ile olmak üzere iki farklı şekilde incelenmiştir. Belirlenen noktalara (ana arter, mevcut ve proje metro, durak, vb.) uzaklıklarına göre ulaşım noktası sayısı verilmiştir.

Tablo 3.14: Küme değerlerine göre toplam nokta sayıları

Küme Değeri	1000 m yerel rakip	500 m zincir rakip	1000 m zincir rakip	500 yerel rakip	1000 m ulaşım noktassı	1000 m perakande	500 m perakende
1	90	-	-	30	152	423	187
2	356	-	-	176	440	1.291	687
3	1.466	-	1	629	1.534	6.355	2.917
4	598	-	2	256	507	2.164	1.170
5	715	5	11	304	411	2.844	1.310
Toplam	**3.225**	**5**	**14**	**1.395**	**3.044**	**13.077**	**6.271**

İncelenen parametrelerden sonra Tablo 3.13'deki gibi bir veri seti elde edilmiştir. Bu veri seti ile makine öğrenme algoritmaları uygulanarak mağazaların benzerlik haritası çıkartılması hedeflenmiştir.

Tablo 3.15: Lokasyon analizi veri seti

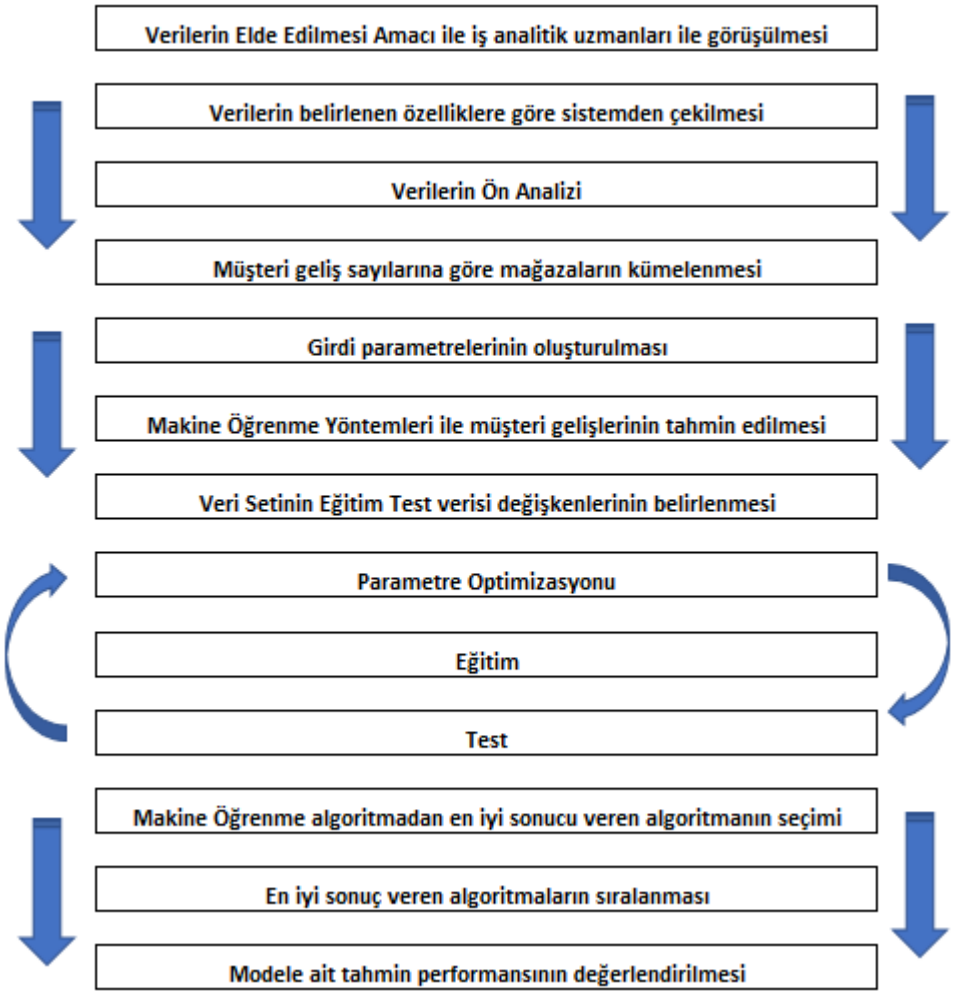
Lokasyon	Kent_Nuf	Kent_Geli	Kent_Eko	Foot_Traf	Retail_Ind	Ortalama_	YerelRaki	YerelRaki	Zincir_100	Zincir_500	Ulasimno	Perakend	500Perakende	Küme
Lokasyon 2	5	5	4	35	32	3991	27	8	5	1	51	539	161	5
Lokasyon 1	5	5	4	42	34	1223	221	79	0	0	84	340	189	5
Lokasyon 4	4	5	2	33	30	1405	29	19	2	2	22	421	304	5
Lokasyon 5	5	5	4	65	55	7475	31	10	1	1	14	240	108	5
Lokasyon 90	5	5	4	20	13	569	40	19	0	0	49	28	7	2
Lokasyon 57	5	5	3	17	15	139	47	29	0	0	37	100	57	3
Lokasyon 3	5	5	4	39	30	1174	184	90	0	0	60	291	87	5
Lokasyon 22	5	5	2	43	32	791	53	30	1	0	25	133	91	4
Lokasyon 11	5	5	3	34	31	565	68	22	0	0	58	181	40	4
Lokasyon 9	5	4	4	17	14	690	86	62	1	0	54	174	68	5
Lokasyon 21	5	4	2	25	22	3064	45	21	0	0	31	138	103	4
Lokasyon 49	5	4	3	31	27	2089	19	4	0	0	12	147	54	3
Lokasyon 59	5	4	2	20	20	2096	17	7	0	0	25	126	56	3
Lokasyon 17	4	4	2	41	33	529	24	15	1	0	55	164	112	4

4. BULGULAR

4.1. VAKA ÇALIŞMASI

Çalışmamızın ilk adımında perakende sektöründe faaliyet gösteren bir firmanın müşterilerine hizmet verdiği lokasyonların 1000 m, 500 m, 250 m çevre verileri ve ilçe verileri farklı özellikler açısından kıyaslanarak, ilgili lokasyonların müşteri çekme potansiyellerinin hangi kriterlere bağlı olduğu tespit edilmeye çalışılmıştır. Yaklaşık 100 lokasyona ait 250 farklı veri seti üzerinde inceleme yapılmıştır. Yatırım için en değerli lokasyonlar hangileridir? sorusunun cevabının arandığı çalışmada Weka programını ve makine öğrenme algoritmalarından faydalanılmıştır.

Düzenlenmiş veri seti WEKA programına girilerek veri tiplerine ait tanımlamalar yapılmıştır. Tanımlanan veriler üzerinde Makine Öğrenme ve Bileşik Makine Öğrenme teknikleri kullanılarak müşteri geliş sayılarına göre hesaplanmış küme değeri tahmin edilmiş ve bu tahminlere göre tahmin performansları elde edilmiştir. Her bir algoritma için en iyi sonucu veren test ve eğitim verileri deneme yanılma yöntemleri ile belirlenmiştir. Makine Öğrenme yöntemlerinden kategorik ve sayısal değerlerin ikisini de içeren veri setlerine uygun olan M5P, REPTree, Decision Stump, Random Tree, M5Rules, Decision Table, KStar, IBk, LWL, Linear Regression, SMOreg, Multilayer Perceptron yöntemlerinin müşteri gelişlerini tahmin etme performansları değerlendirilmiştir. Bileşik Makine Öğrenme yöntemleri için sınıflandırıcı olarak Makine Öğrenme yöntemlerinden en iyi performans elde ettiğimiz algoritmaların çeşitli kombinasyonları çalıştırılmıştır.



Şekil 4.1: Uygulama Metodolojisi

Uygulanan Makine Öğrenme ve Bileşik Makine Öğrenme tekniklerinin veri setinden elde edilen tahmin performansları tablolar halinde gösterilmiş ve yöntemler birbirleri ile korelasyon katsayısı ve hata değerleri bakımından kıyaslanmıştır. Uygulama metodolojisi Şekil 4.1'de görülmektedir.

4.2. SONUÇLARIN DEĞERLENDİRİLMESİ

4.2.1. Basit Makine Öğrenme Algoritmaları

Makine Öğrenme tekniklerinin tahmin performansı kullanılacak algoritma ve algoritma için seçilecek parametrelere göre değişmektedir. Bunların çoğu deneme yanılma yöntemleri ile elde

edilen ve herhangi bir yöntem için bu parametre daha iyidir şeklinde bir kural seti oluşturulacak sabit değişkenler değillerdir. Örnek olarak bir test için veri setinin %80 öğrenme %20'si test verisi olduğunda gösterdiği performans ile %90-%20 şeklinde ayarlanmış aynı veri setinin göstereceği performans aynı değildir. Bu amaçla uygulanan algoritmaların parametreleri deneme yanılma yöntemleri ile çalıştırılmış en iyi sonuçları veren parametreler seçilerek yöntem ve kısıtları değerlendirilmiştir. Öğrenme algoritmaları için en önemli parametre örneklem verinin seçilmesi yöntemidir. K Katlamalı çapraz doğrulama veya herhangi bir % aralığına göre belirlenen çalıştırma şekilleri ile algoritmalar çalıştırılmıştır. Her yöntem için farklı parametre seçiminin en iyi sonucu verdiği gözlemlenmiştir.

4.2.1.1 RandomTree

RandomTree her düğümde rastgele seçilen öz nitelikleri dikkate alan bir sınıflandırma algoritmasıdır. En çok tercih edilen karar ağaçları arasında bulunmaktadır. Literatür kısmında belirttiğimiz tanımlarda ki gibi bir dal yaprak modeli oluşturmaktadır. Bu tanımları burada yinelemekte aşağıdaki şeklin anlaşılması adına fayda bulunmaktadır. %90 eğitim- %10 test verisi parametresi seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %80 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %69,7 olarak bulunmuş ve iyi düzeyde bir uyum gözlendiği tespit edilmiştir. %80 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.1: RandomTree algoritma sonuçları (%90-%10)

Correctly Classified Instances	**80%**
Incorrectly Classified Instances	**20%**
Kappa statistic	**0.697**
Mean absolute error	**0.08**
Root mean squared error	**0.2828**
Relative absolute error	**28%**
Root relative squared error	**74%**

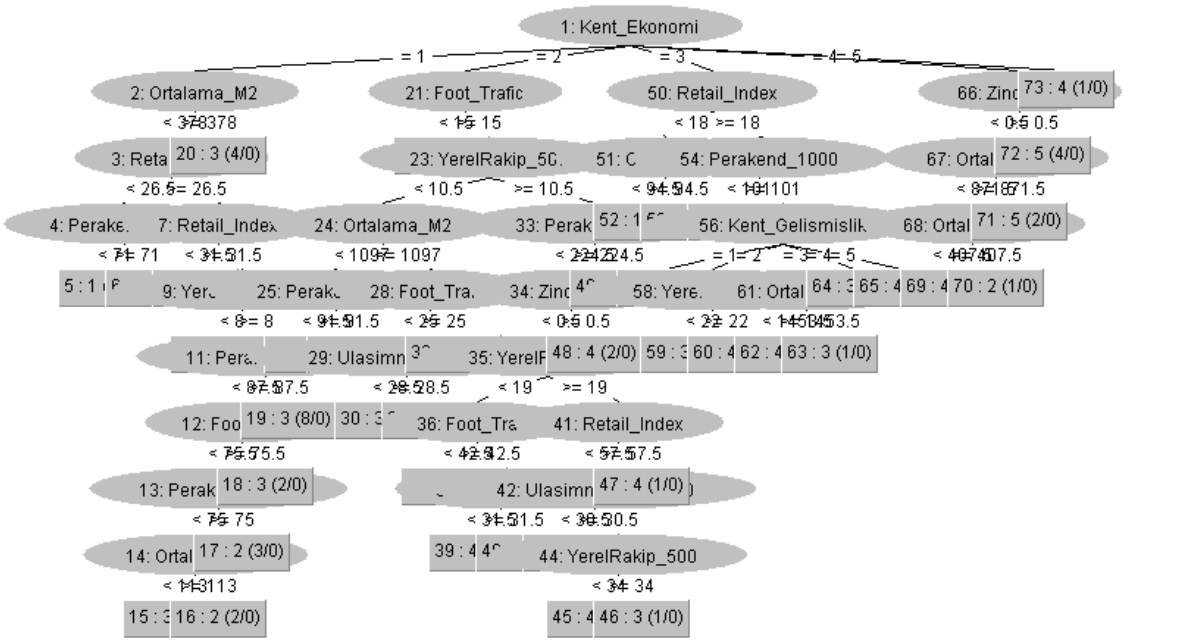
Tablo 4.1'de görüldüğü üzere her değişken sınıfı ile bir kural seti oluşturulmuştur. Örneğin Kent ekonomi değeri 5 (yüksek düzeyde) olan bir bölgeye açılan mağaza 4 numaraları sınıfa girecektir. Kent ekonomi değeri 1 (en düşük değer), ortalama metrekare değeri 378 TL'nin

altında, perakende indeksi 26,5'in altında ve 1000 metre çevresinde 71 adetin altında perakende mağaza bulunan bir bölgeye açılan mağaza ise 3 no'lu sınıfta bulunan mağazalara benzer özellik sergileyecektir. Kent ekonomi değeri 1 ve ortalama metrekare değeri 378 TL'nin üstünde bulunan bir bölgeye açılacak mağazalar ise 3 no'lu mağazalara benzer özellik sergileyecektir.

Tablo 4.2: Random Tree algoritması kural ağacı

Kent_Ekonomi = 1	Kent_Ekonomi = 2	Kent_Ekonomi = 3	Kent_Ekonomi = 4	Kent_Ekonomi = 5 : 4 (1/0)
| Ortalama_M2 < 378	| Foot_Trafic < 15 : 1 (3/0)	| Retail_Index < 18	| Zincir_1000 < 0.5	
| | Retail_Index < 26.5	| Foot_Trafic >= 15	| | Ortalama_M2 < 94.5 : 1 (1/0)	| | Ortalama_M2 < 871.5	
| | | Perakend_1000 < 71 : 1 (3/0)	| | YerelRakip_500 < 10.5	| | Ortalama_M2 >= 94.5 : 3 (3/0)	| | | Ortalama_M2 < 407.5 : 4 (1/0)	
| | | Perakend_1000 >= 71 : 2 (1/0)	| | | Ortalama_M2 < 1097	| Retail_Index >= 18	| | | Ortalama_M2 >= 407.5 : 2 (1/0)	
| | Retail_Index >= 26.5	| | | | Perakend_1000 < 91.5 : 2 (1/0)	| | Perakend_1000 < 101 : 2 (2/0)	| | Ortalama_M2 >= 871.5 : 5 (2/0)	
| | | Retail_Index < 31.5 : 2 (4/0)	| | | | Perakend_1000 >= 91.5 : 3 (12/0)	| | Perakend_1000 >= 101	| Zincir_1000 >= 0.5 : 5 (4/0)	
| | | Retail_Index >= 31.5	| | | Ortalama_M2 >= 1097	| | | Kent_Gelismislik = 1 : 1 (0/0)		
| | | | YerelRakip_1000 < 8 : 2 (3/0)	| | | | Foot_Trafic < 25	| | | Kent_Gelismislik = 2		
| | | | YerelRakip_1000 >= 8	| | | | | Ulasimnoktasi_1000 < 28.5 : 3 (1/0)	| | | | YerelRakip_500 < 22 : 3 (1/0)		
| | | | | Perakend_1000 < 87.5	| | | | | Ulasimnoktasi_1000 >= 28.5 : 2 (1/0)	| | | | YerelRakip_500 >= 22 : 4 (1/0)		
| | | | | | Foot_Trafic < 75.5	| | | | Foot_Trafic >= 25 : 4 (1/0)	| | | Kent_Gelismislik = 3		
| | | | | | | Perakend_1000 < 75	| | YerelRakip_500 >= 10.5	| | | | Ortalama_M2 < 1453.5 : 4 (2/0)		
| | | | | | | | Ortalama_M2 < 113 : 3 (1/0)	| | | Perakend_1000 < 224.5	| | | | Ortalama_M2 >= 1453.5 : 3 (1/0)		
| | | | | | | | Ortalama_M2 >= 113 : 2 (2/0)	| | | | Zincir_1000 < 0.5	| | | Kent_Gelismislik = 4 : 3 (1/0)		
| | | | | | | Perakend_1000 >= 75 : 2 (3/0)	| | | | | YerelRakip_500 < 19	| | | Kent_Gelismislik = 5 : 4 (1/0)		
| | | | | | Foot_Trafic >= 75.5 : 3 (2/0)	| | | | | | Foot_Trafic < 42.5 : 3 (10/0)			
| | | | | Perakend_1000 >= 87.5 : 3 (8/0)	| | | | | | Foot_Trafic >= 42.5			
| Ortalama_M2 >= 378 : 3 (4/0)	| | | | | | | Retail_Index < 31.5 : 4 (1/0)			
	| | | | | | | Retail_Index >= 31.5 : 3 (3/0)			
	| | | | | YerelRakip_500 >= 19			
	| | | | | | Retail_Index < 57.5			
	| | | | | | | Ulasimnoktasi_1000 < 30.5 : 3 (3/0)			
	| | | | | | | Ulasimnoktasi_1000 >= 30.5			
	| | | | | | | | YerelRakip_500 < 34 : 4 (2/0)			
	| | | | | | | | YerelRakip_500 >= 34 : 3 (1/0)			
	| | | | | | Retail_Index >= 57.5 : 4 (1/0)			
	| | | | Zincir_1000 >= 0.5 : 4 (2/0)			
	| | | Perakend_1000 >= 224.5 : 5 (3/0)			

Kent ekonomi değeri 2 fakat 1000 metre çevresinde bulunan perakende mağaza sayısı 224 yani ortalamanın üzerinde bir perakende yoğunluğuna sahip bir bölgede açılan mağaza 5 numaralı bir mağazaya benzeyecektir. Buda kötü bir ilçenin iyi bir noktasına mağaza açmanın daha tutarlı bir karar olduğunu göstermektedir.



Şekil 4.2: RandomTree ağaç yapısı

4.2.1.2 RandomForest

RandomForest ağaç tahminleyicilerinin bir kombinasyonu olup, her ağaç, bağımsız olarak örneklenmiş rastgele bir vektörün değerlerine ve ormandaki tüm ağaçlar için aynı dağılıma bağlıdır. Bu tanımdan da anlaşılacağı gibi bir dal yaprak modelinden ziyade çeşitli kombinasyonların sonuçlarını listeleyecektir. %90 eğitim- %10 test verisi seçildiğinde en iyi sonucu vermektedir. Algoritma 100 iterasyondan sonra öğrenmeye başlamış olup diğer algoritmalara göre daha hızlı bir sonuç vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %70 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %51,7 olarak bulunmuş ve orta düzeyde bir uyum gözlendiği tespit edilmiştir. %70 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.3: RandomForest algoritma sonuçları (%90-%10)

Correctly Classified Instances	**70%**
Incorrectly Classified Instances	**30%**
Kappa statistic	**0.5161**
Mean absolute error	**0.1967**
Root mean squared error	**0.3086**
Relative absolute error	**70%**
Root relative squared error	**81%**

4.2.1.3 J48

J48 karar ağacı algoritması C4.5 karar ağacı algoritmasının özel bir türüdür. %80 eğitim- %20 test verisi parametresi seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %60 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %45 olarak bulunmuş ve orta düzeyde bir uyum gözlendiği tespit edilmiştir. %80 oranındaki bu tahmin performansı orta düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.4: J48 algoritma sonuçları (%80-%20)

Correctly Classified Instances	**60%**
Incorrectly Classified Instances	**40%**
Kappa statistic	**0.45**
Mean absolute error	**0.17**
Root mean squared error	**0.40**
Relative absolute error	**61%**
Root relative squared error	**104%**

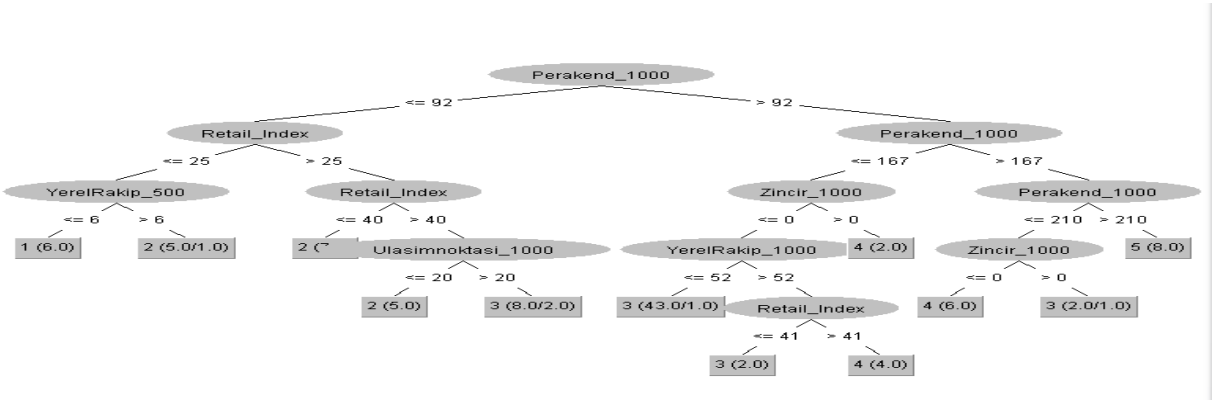
Sınıflandırma modelinde genel olarak nokta parametrelerini dikkate almıştır. 1000 metre çevredeki perakende mağaza sayısı 92'nin altında, perakende indeks değeri 25'in altında ve 500 metre çevredeki yerel rakip sayısı 6'nın üstünde ise 2 no'lu kategoriye girmektedir. Çevredeki yerel ve perakende nokta sayılarının düşüklüğü bölgenin bir çekim merkezi olmadığını

göstermektedir. Buraya açılan bir mağazadan yüksek performans beklemek yanlış olacaktır. Çevredeki perakende mağaza sayısı bu karar ağacından anlaşılacağı üzere mağazaya gelecek olan müşteri tahminleme de büyük bir gösterge sunmaktadır. Perakende mağaza sayısının 210'dan fazla olduğu yerlerde açılan mağazalar direk 5 no'lu sınıfa benzer bir değer segmenti sergileyeceği kuralı bu parametrenin iyi bir lokasyon için ne kadar önemli olduğunu göstermektedir. Bu kural seti ışığında 1000 metre çevresinde 90 adetten az perakende mağaza bulunduran herhangi bir bölgeye diğer çevre özellikleri ne olursa olsun mağaza açmak uygun değildir. Bölge bir çekim merkezi etkisi sergilemeyecektir.

Tablo 4.5: J48 algoritması kural ağacı

Perakend_1000 <= 92	Perakend_1000 > 92	| Perakend_1000 > 167
| Retail_Index <= 25	| Perakend_1000 <= 167	| | Perakend_1000 <= 210
| | YerelRakip_500 <= 6: 1 (6.0)	| | Zincir_1000 <= 0	| | | Zincir_1000 <= 0: 4 (6.0)
| | YerelRakip_500 > 6: 2 (5.0/1.0)	| | | YerelRakip_1000 <= 52: 3 (43.0/1.0)	| | | Zincir_1000 > 0: 3 (2.0/1.0)
| Retail_Index > 25	| | | YerelRakip_1000 > 52	| | Perakend_1000 > 210: 5 (8.0)
| | Retail_Index <= 40: 2 (7.0)	| | | | Retail_Index <= 41: 3 (2.0)	
| | Retail_Index > 40	| | | | Retail_Index > 41: 4 (4.0)	
| | | Ulasimnoktasi_1000 <= 20: 2 (5.0)	| | Zincir_1000 > 0: 4 (2.0)	
| | | Ulasimnoktasi_1000 > 20: 3 (8.0/2.0)		

J48 algoritmasının çıkarttığı kural setlerinden biride 1000 metre çevrede zincir rakip mağazasının bulunmasının etkisinin önemli olması sonucudur. Perakende mağaza sayısı 167'nin üstünde ve çevrede zincir rakip mağaza var ise mağaza 4 no'lu sınıfdaki mağazalara benzer özellik sergileyecektir. Burada ortaya çıkan sonuçlardan bir diğeri ise perakende mağaza sayısına göre merkeziyetlik derecesi belirlenecek olursa merkeziyetliği yüksek olan yerlerde zincir rakipler var ise tercih oralara kaymaktadır. Merkeziyeti yüksek yerlere gelen müşterilerin tercihi ile merkeziyeti düşük yerlere gelen müşterilerin tercihleri farklılaşmaktadır.



Şekil 4.3: J48 ağaç yapısı

4.2.1.4 HoeffdingTree

Küçük bir numunenin büyük veri setini optimum şekilde yorumlayacağı varsayımından oluşan bir karar ağacı algoritmasıdır. Gözlem sayısının nicel olarak değerlendirildiği bir yapıda model oluşturmaktadır. %90 eğitim- %10 test verisi parametresi seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %70 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %54 olarak bulunmuş ve orta düzeyde bir uyum gözlendiği tespit edilmiştir. %70 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.6: Hoeffding algoritma sonuçları (%90-%10)

Correctly Classified Instances	**70%**
Incorrectly Classified Instances	**30%**
Kappa statistic	**0,54**
Mean absolute error	**0,11**
Root mean squared error	**0,29**
Relative absolute error	**39%**
Root relative squared error	**77%**

4.2.1.5 PART

Bir karar listesi oluşturmak için kullanılan kural algoritmalarındandır. Her kural için basit bir C4.5 karar ağacı algoritması oluşturmaktadır. %80 eğitim - %20 test verisi parametresi seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %80 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %71 olarak bulunmuş ve iyi düzeyde bir uyum gözlendiği tespit edilmiştir. %80 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.7: PART algoritma sonuçları (%80-%20)

Correctly Classified Instances	**80%**
Incorrectly Classified Instances	**20%**
Kappa statistic	**0,71**
Mean absolute error	**0,09**
Root mean squared error	**0,28**
Relative absolute error	**34%**
Root relative squared error	**75%**

Part algoritması ile oluşturulan kurallar aşağıdaki gibidir.

- 1000 metre çevredeki perakende mağaza sayısı 92 ve 167 arasında, 1000 m çevresinde zincir mağaza bulunmayan ve yerel rakip sayısı 52 adedin altında olan bölgelere açılan mağaza 3 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 110'un altında, perakende indeksi 25 ile 40 arasında ve ulaşım noktası 20'nin üzerinde olan bölgelere açılan mağaza 3 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 110'un altında, yaya trafik yoğunluğu 19'un üstünde ve 1000 metredeki yerel rakip sayısı 6'dan büyük bölgelere açılan mağaza 2 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 101 adet ile 210 adet arasında, 500 metre çevresindeki perakende mağaza sayısı 38'in üzerinde olan bölgelere açılan mağaza 4 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 210 adetten fazla olan yerlere açılan mağazalar 5 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- Perakende indeksi 29'un altında olan bölgelere açılan mağazalar 1 no'lu kümedeki mağazalara benzer özellik sergileyecektir.

4.2.1.6 DecisionTable

K katlamalı çapraz doğrulama sayısı 20 olarak seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %72 olarak bulunmuştur. Kategorik

verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %58 olarak bulunmuş ve orta düzeyde bir uyum gözlendiği tespit edilmiştir. %72 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.7: DecisionTable algoritma sonuçları (10 - Fold)

Correctly Classified Instances	**72%**
Incorrectly Classified Instances	**27%**
Kappa statistic	**0,58**
Mean absolute error	**0,18**
Root mean squared error	**0,29**
Relative absolute error	**69%**
Root relative squared error	**80%**

4.2.1.7 JRip

IWP'nin optimize edilmiş bir versiyonudur. Artırımlı budama yöntemi ile çalışan bir algoritmadır. K katlamalı çapraz doğrulama sayısı 20 olarak seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %74 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %61 olarak bulunmuş ve iyi düzeyde bir uyum gözlendiği tespit edilmiştir. %74 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.8: JRip algoritma sonuçları (20 - Fold)

Correctly Classified Instances	**74%**
Incorrectly Classified Instances	**25%**

Kappa statistic	**0,61**
Mean absolute error	**0,12**
Root mean squared error	**0,29**
Relative absolute error	**46%**
Root relative squared error	**80%**

JRip algoritması ile oluşturulan kurallar aşağıdaki gibidir.

- 1000 metre çevredeki perakende mağaza sayısı 73'ün altında, yaya trafik yoğunluğu 30'un altında olan bölgelere açılan mağaza 1 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 240 adetin üstünde olan bölgelere açılan mağaza 5 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 168 ve 240 adet arasında olan bölgelere açılan mağaza 4 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki yerel rakip sayısı 53 adetin üstünde, kent nüfus değeri 5 olan bölgelere açılan mağaza 4 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 500 metre çevredeki perakende mağaza sayısı 103 adetten fazla olan yerlere açılan mağazalar 4 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- 1000 metre çevredeki perakende mağaza sayısı 87 adetten az olan bölgelere açılan mağazalar 2 no'lu kümedeki mağazalara benzer özellik sergileyecektir.

4.2.1.8 LWL

Tembel öğrenme algoritmalarında olan LWL için en iyi sonucu K katlamalı çapraz doğrulama sayısı 20 parametresi vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %65 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %41 olarak bulunmuş ve düşük düzeyde bir uyum gözlendiği tespit edilmiştir. %65 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.9: LWL algoritma sonuçları (20 - Fold)

Correctly Classified Instances	**65%**
Incorrectly Classified Instances	**34%**
Kappa statistic	**0,41**
Mean absolute error	**0,19**
Root mean squared error	**0,31**
Relative absolute error	**71%**
Root relative squared error	**85%**

4.2.1.9 MultilayerPerceptron

Yapay sinir ağı algoritması olan multilayerperceptron K katlamalı çapraz doğrulama sayısı 10 olarak seçildiğinde en iyi sonucu vermektedir. Doğru sınıflandırma yüzdesi (Correctly Classified Instances) %66 olarak bulunmuştur. Kategorik verilerin değerlendirilmesindeki uyumu ölçen istatistik değeri kappa istatistik katsayısı ise %47 olarak bulunmuş ve orta düzeyde bir uyum gözlendiği tespit edilmiştir. %66 oranındaki bu tahmin performansı iyi düzeyde bir tahmin başarısı elde edildiğini ve lokasyon yeri seçiminde kullanılabilecek bir model olduğunu göstermektedir.

Tablo 4.10: Multilayerperceptron algoritma sonuçları (10 - Fold)

Correctly Classified Instances	**66%**
Incorrectly Classified Instances	**33%**
Kappa statistic	**0,47**
Mean absolute error	**0,14**
Root mean squared error	**0,32**
Relative absolute error	**51%**
Root relative squared error	**89%**

4.2.1.10 Basit Makine Öğrenme Algoritma Sonuçlarının Karşılaştırılması

Basit makine öğrenme algoritmaları doğru sınıflandırma yüzde değerleri yani tahmin başarı performansı olarak karşılaştırıldığında en iyi sonucu RandomTree ve Part algoritmaları vermektedir. Kategorik değişkenlerin uyum istatistiği açısından kappa istatistik değerleri karşılaştırıldığında en iyi sonucu yine PART ve RandomTree algoritmalarının verdiği gözlemlenmiştir. Ortalama mutlak hata değeri en düşük olana algoritma ise 0,08 ile RandomTree algoritması olmuştur. Tüm sonuç değerlerinin bulunduğu veriler Tablo 4.11'de gösterilmiştir.

Tablo 4.11: Basit Makine Öğrenme Algoritma Sonuçları Karşılaştırması

Makine Öğrenme Algoritma Adı	Test Verisi	Corelation Coefficient	Mean Absolute Error	Kappa statistic
RandomTree	%90 - %10	80%	0,08	0,69
RAndomForest	%90 - %10	70%	0,19	0,51
J48	%80 - %20	60%	0,17	0,45
HoeffdingTree	%90 - %10	70%	0,11	0,54
Part	%80 - %20	80%	0,09	0,71
DecisionTable	20 Fold	72%	0,18	0,58
Jrip	20 Fold	74%	0,12	0,61
LWL	20 Fold	65%	0,19	0,41
MultilayerPerceptron	10 Fold	66%	0,12	0,47

4.2.2. Bileşik Makine Öğrenme Algoritmaları

4.2.1.1 RandomSubSpace

Rastgele uzaylar yöntemi topluluk yöntemleri arasında en fazla kullanılan algoritmalardandır. Eğitim esnasında sınıf tanımlayıcısına ait bazı öznitelikler seçilir bazıları seçilmez ve birden fazla alt uzay için seçilen algoritma türüne göre karar kuralları oluşturulur. En iyi sonucu veren RandomSubSpace algoritmaları Tablo 4.12'de verilmiştir.

Tablo 4.12: RandomSubSpace Algoritma Sonuçları

Makine Öğrenme Algoritma Adı	Test Verisi	Corelation Coefficient	Mean Absolute Error	Kappa statistic
RandomSubSpace (REPTree)	10 Fold	71%	0,19	0,52
RandomSubSpace (RandomForest)	10 Fold	72%	0,17	0,54
RandomSubSpace (LMT)	10 Fold	80%	0,13	0,69
RandomSubSpace (J48)	20-Fold	70%	0,16	0,51
RandomSubSpace (HoeffdingTree)	10-Fold	77%	0,16	0,66
RandomSubSpace (PART)	10-Fold	76%	0,15	0,62
RandomSubSpace (OneR)	10-Fold	65%	0,15	0,44
RandomSubSpace (JRıp)	10-Fold	70%	0,16	0,54
RandomSubSpace (DecisionTable)	10-Fold	70%	0,21	0,52
RandomSubSpace (KStar)	10-Fold	65%	0,17	0,44
RandomSubSpace (MultilayerPerceptron)	10-Fold	72%	0,15	0,57

Tablo 4.12'de görüldüğü gibi en iyi RandomSubSpace algoritması %80 doğru sınıflandırma yüzdesi ile LMT algoritmasıdır. Bu algoritma en iyi sonucu K katlamalı çapraz doğrulama katsayısı 10 olarak seçildiğinde vermiştir. Kappa istatistik katsayısı %69 olarak bulunmuş ve iyi düzeyde bir uyum gösterdiği tespit edilmiştir. RandomSubSpace (HoeffdingTree) ve RandomSubSpace (PART) algoritmalarıda basit makine öğrenme algoritmalarına göre daha iyi sonuç vermişlerdir.

RandomSubSpace (PART) algoritmasının oluşturduğu bazı kurallar aşağıda verilmiştir.

- Kent Ekonomisi 4 no'lu kategoride ve 1000 metre çevresinde zincir rakip bulunan herhangi bir bölgeye açılan mağaza 5 no'lu mağaza kümesindeki mağazalara benzer özellik sergileyecektir.
- Kent ekonomisi 1 no'lu kategoride, yaya trafik yoğunluğu 33'ün üzerinde, kent nüfusu 2 no'lu kategoride ve 1000 metre çevresinde 23 adetten fazla ulaşım noktası bulunan bölgelere açılan mağazalar 3 no'lu kümedeki mağazalara benzer özellik sergileyecektir.
- Kent nüfusu 5 no'lu kategoride ve 500 metre çevresinde 40 adetten az yerel rakip bulunan herhangi bir bölgeye açılan mağaza 4 no'lu kümedeki mağazalara benzer bir özellik sergileyecektir.

RandomSubSpace algoritmaları mağazaların açılmaması gereken noktaların tespitinde daha fazla bilgi veren kural setleri oluşturmuşlardır. Ulaşım noktası sayısı 13 'ün altında herhangi bir yere mağaza açmak veya yaya trafik yoğunluğu 22 'nin altında herhangi bir yere mağaza

açmak, Hem nüfus hem de ekonomi değeri aynı anda 1 olan kentlere mağaza açmak uygun değildir.

4.2.1.2 Bagging

Bileşik makine öğrenme algoritmalarından olan bagging, öğrenicilerinin her birini eğitim verisinden rastgele seçilen farklı alt kümelerle eğitir. Veri, ilk aşamada eğitim ve test olarak ayrılır. Daha sonra eğitim için ayrılan veri setinden rastgele seçim yapılır. Bu temel kuralla çalışan Bagging algoritmasında en iyi sonucu veren algoritma %81 doğru tahmin performansı ile Bagging (NaiveBayesUpdateable) algoritmasıdır. Algoritma sonuçlar Tablo 4.13'de verilmiştir.

Tablo 4.13: Bagging Algoritma Sonuçları

Makine Öğrenme Algoritma Adı	Test Verisi	Corelation Coefficient	Mean Absolute Error	Kappa statistic
Bagging (REPTree)	20 Fold	70%	0,17	0,51
Bagging (RandomTree)	90 Fold	69%	0,17	0,49
Bagging (LMT)	10 Fold	75%	0,11	0,62
Bagging (J48)	10 Fold	76%	0,12	0,64
Bagging (PART)	10 Fold	73%	0,14	0,59
Bagging (JRip)	10 Fold	74%	0,13	0,6
Bagging (DecisionTable)	10 Fold	74%	0,2	0,58
Bagging (LWL)	10 Fold	66%	0,18	0,43
Bagging (MultilayerPerceptron)	10 Fold	67%	0,15	0,49
Bagging (SimpleLogistic)	10 Fold	76%	0,12	0,64
Bagging (NaiveBayesUpdateable)	10 Fold	81%	0,11	0,71

Lojistik ve Bayes algoritmaları ile RandomSubSpace algoritmalarına göre dahi iyi sonuçlar üretmiştir. Tablo 4.15 'den de anlaşılacağı gibi en iyi sonucu veren ilk üç bagging algoritmaları Bagging (NaiveBayesUpdateable), Bagging (J48), Bagging (SimpleLogistic) 'dır.

4.2.1.3 Vote

Vote algoritması diğer algoritma yöntemlerine göre daha iyi sonuç üretmiştir. En iyi sonucu Vote (NaiveBayesUpdateable, SimpleLogistic, PART) algoritması vermiştir. Tahmin performansı k katlamalı çapraz doğrulama katsayısı 10 olarak seçildiğinde %83 olarak bulunmuştur. Ortalama mutlak hatası 0,11 ile en düşük algoritmalarından bir tanesidir. Kappa istatistik katsayısı 0,74 olarak bulunmuş ve iyi düzeyde bir uyum istatistiği elde edilmiştir.

Tablo 4.14 de vote algoritmasının sonuçları paylaşılmıştır. Bayes algoritmaları ile kombinasyon yapılan tüm algoritmalar iyi birer sonuç üretmiştir.

Tablo 4.14: Vote Algoritma Sonuçları

Makine Öğrenme Algoritma Adı	Test Verisi	Corelation Coefficient	Mean Absolute Error	Kappa statistic
Vote (NaiveBayesUpdateable, PART)	10 Fold	75%	0,11	0,62
Vote (NaiveBayesUpdateable, LMT)	10 Fold	82%	0,1	0,72
Vote (NaiveBayesUpdateable, LWL)	10 Fold	78%	0,14	0,66
Vote (NaiveBayesUpdateable, SimpleLogistic)	10 Fold	81%	0,1	0,7
Vote (NaiveBayesUpdateable, SimpleLogistic,PART)	10 Fold	83%	0,11	0,74
Vote (NaiveBayesUpdateable, SimpleLogistic,PART, LWL)	10 Fold	81%	0,11	0,7
Vote (NaiveBayesUpdateable, MultilayerPerceptron)	10 Fold	74%	0,12	0,61
Vote (LMT,HoeffdingTree,J48)	10 Fold	75%	0,11	0,62
Vote (NaiveBayesUpdateable, BayesNet)	10 Fold	81%	0,12	0,7

En iyi sonucu veren Vote (NaiveBayesUpdateable, SimpleLogistic, PART) algoritmasının sonuçları;

NaiveBayesUpdateable sonuçları

Tablo 4.15: NaiveBayesUpdateable nominal değişken sonuçları

		Class				
		1	**2**	**3**	**4**	**5**
		(0.08)	(0.18)	(0.5)	(0.14)	(0.1)
Kent_Nufus						
	1	2	8	5	1	1
	2	3	6	14	1	1
	3	1	2	7	1	1
	4	1	1	9	3	2
	5	5	6	21	12	9
Kent_Gelismislik						
	1	1	6	2	1	1
	2	4	8	22	6	1
	3	3	3	13	4	2
	4	3	4	17	4	4
	5	1	2	2	3	6
Kent_Ekonomi						
	1	4	14	16	1	1
	2	4	3	31	8	4
	3	2	3	7	5	1
	4	1	2	1	2	7
	5	1	1	1	2	1

Tablo 4.16: NaiveBayesUpdateable numerik değişken sonuçları

	Class				
	1	2	3	4	5
Foot_Trafic					
mean	17,26	45,46	40,27	39,06	35,86
std. dev.	6,53	20,13	16,72	13,08	13,02
weight sum	7,00	18,00	51,00	13,00	9,00
precision	1,53	1,53	1,53	1,53	1,53
Retail_Index					
mean	14,55	41,68	37,25	35,57	30,72
std. dev.	5,26	20,59	15,44	12,43	11,20
weight sum	7,00	18,00	51,00	13,00	9,00
precision	1,62	1,62	1,62	1,62	1,62
Ortalama_M2					
mean	527,00	264,15	425,96	870,23	1876,39
std. dev.	582,74	380,17	485,13	830,14	2253,84
weight sum	7,00	18,00	51,00	13,00	9,00
precision	81,98	81,98	81,98	81,98	81,98
YerelRakip_1000					
mean	12,76	20,09	28,78	45,82	78,95
std. dev.	8,51	12,27	14,26	18,29	68,52
weight sum	7,00	18,00	51,00	13,00	9,00
precision	4,25	4,25	4,25	4,25	4,25
YerelRakip_500					
mean	3,97	9,73	12,32	20,11	33,68
std. dev.	2,51	6,70	7,33	9,71	30,79
weight sum	7,00	18,00	51,00	13,00	9,00
precision	2,78	2,78	2,78	2,78	2,78
Zincir_1000					
mean	0,00	0,00	0,03	0,26	1,48
std. dev.	0,28	0,28	0,28	0,60	1,46
weight sum	7,00	18,00	51,00	13,00	9,00
precision	1,67	1,67	1,67	1,67	1,67
Zincir_500					
mean	0,00	0,00	0,00	0,00	0,56
std. dev.	0,17	0,17	0,17	0,17	0,68
weight sum	7,00	18,00	51,00	13,00	9,00
precision	1,00	1,00	1,00	1,00	1,00
Ulasimnoktasi_1000					
mean	21,85	24,28	30,06	38,92	45,38
std. dev.	8,19	10,50	13,49	18,78	20,28
weight sum	7,00	18,00	51,00	13,00	9,00
precision	1,68	1,68	1,68	1,68	1,68
Perakend_1000					
mean	61,32	71,92	124,38	166,14	316,44
std. dev.	8,14	16,25	28,23	23,23	100,81
weight sum	7,00	18,00	51,00	13,00	9,00
precision	6,81	6,81	6,81	6,81	6,81
500Perakende					
mean	27,64	37,75	57,00	89,31	145,50
std. dev.	16,57	14,63	23,55	27,22	68,64
weight sum	7,00	18,00	51,00	13,00	9,00
precision	4,50	4,50	4,50	4,50	4,50

SimpleLogistic algoritması sonuçları

Class 1 :38.09 + [Kent_Gelismislik=2] * 1.11 +[Foot_Trafic] * -0.31 +[Retail_Index] * -0.04 +[YerelRakip_1000] * -0.14 +[YerelRakip_500] * -0.2 +[Ulasimnoktasi_1000] * -0.27 +[Perakende_1000]* -0.22

Class 2 :17.11 + [Kent_Nufus=4] * -2.2 +[Kent_Gelismislik=3] * -0.82 +[Kent_Gelismislik=5] * -0.73 +[Kent_Ekonomi=1] * 1.45+[Foot_Trafic] * -0.03 +[Retail_Index] * -0.01 +[Ortalama_M2] * 0 +[YerelRakip_1000] * -0.05 +[Ulasimnoktasi_1000] * -0.07+[Perakende_1000] * -0.14

Class 3 :1.27 + [Kent_Nufus=3] * 0.65 +[Kent_Nufus=5] * -0.66 +[Kent_Gelismislik=1] * -2.29 +[Kent_Gelismislik=2] * -0.38 +[Kent_Gelismislik=3] * 0.26 +[Kent_Gelismislik=4] * 0.59 +[Kent_Ekonomi=3] * -0.59 +[Kent_Ekonomi=4] * -2.33 +[Retail_Index] * 0.07 +[Ortalama_M2] * -0 +[YerelRakip_1000] * -0.03 +[Zincir_1000] * -2.07 +[500Perakende] * -0.02

Class 4 :-20.54 + [Kent_Nufus=2] * -1.83 +[Kent_Nufus=4] * 1.25 +[Kent_Nufus=5] * 1.48 +[Kent_Ekonomi=3] * 0.67 +[Kent_Ekonomi=5] * 2.4 +[Foot_Trafic] * 0.13 +[Retail_Index] * 0.03 +[Ortalama_M2] * 0 +[YerelRakip_1000] * 0.06 +[Ulasimnoktasi_1000] * 0.05+[Perakende_1000] * 0.03 +[500Perakende] * 0.05

Class 5 :-34.41 + [Kent_Nufus=5] * 1.4 +[Kent_Ekonomi=3] * -2.94 +[Kent_Ekonomi=4] * 7.12 +[YerelRakip_500] * 0.05 +[Zincir_1000] * 2.56+[Perakende_1000] * 0.14 +[500Perakende] * 0.01

PART algoritması sonuçları

- Perakende_1000 > 92 && Perakende_1000 <= 167 && YerelRakip_1000 <= 52: 3
- Perakende_1000 <= 110 && Retail_Index > 40 && Ulasimnoktasi_1000 > 20: 3
- Perakende_1000 <= 110 && Foot_Trafic > 19 && YerelRakip_1000 > 6: 2
- Perakende_1000 > 101 && 500Perakende > 38 && Perakende_1000 <= 210: 4
- Perakende_1000 > 210: 5
- Retail_Index <= 29: 1

5. TARTIŞMA VE SONUÇ

Mağaza lokasyonu satışların hedefler çerçevesinde olması için en kritik faktördür. Mağazanın lokasyonu ciroları etkiler. Bu yönüyle perakende mağazacılıkta en kritik başarı faktörlerinin başında lokasyon seçimi gelmektedir. Mağaza lokasyonu için seçilen mevki, konum, pozisyonlama sonrasında mağazanın yapacağı tüm "ticari aktivitelerin kaderini" belirleyecek bir öneme ve önceliğe sahip olmaktadır. Mağazanın kuruluş yeri seçimi tam da bu bakış açısından dolayı son derece önemlidir. Çünkü perakende sektöründe ticari başarının büyük kısmı bu kritik seçim üzerinden devam etmektedir. Çalışmada yukarıda bahsedilen kısıtlar doğrultusunda mağaza açmak için en uygun lokasyon neresidir sorusunun cevabını veren bir karar destek sistemi kurulması hedeflenmiştir. Bu amaç doğtultusunda incelenen mevcut mağazaların üzerinde yapılan testlerde, müşteri geliş sayıları benzeyen mağazaların çevresel koşullarının da benzerlik gösterdiği ortaya çıkmıştır. Bu durum, bu tezin çıkış noktasını da oluşturmaktadır. Eğer açılması düşünülen mağazanın çevresel koşulları itibariyle önceden açılmış hangi mağazalara benzediği ortaya çıkartılırsa, onun cirosu hakkında da önsel bir fikir elde edilebilir.

Basit makine öğrenme algoritmalarından RandomTree, RandomForest, J48, HoeffdingTree, Part, DecisionTable, Jrip, LWL, MultilayerPerceptron algoritmaları kullanılmıştır. En iyi sonucu %80 tahmin performansı ile PART ve RandomTree algoritmaları vermiştir.

Bileşik makine öğrenme yöntemlerinden Vote, Bagging, RandomSubSpace algoritmaları kullanılmıştır. En iyi sonuç vote algoritmalarından Vote (NaiveBayesUpdateable, SimpleLogistic, PART) algoritması ile elde edilmiştir. %83'lük bir doğru sınıflandırma yüzdesi ile aynı zamanda en iyi tahmin performansını gösteren algoritma olmuştur.

Kullanılan algoritmalardan çeşitli kural setleri oluşturulmuştur. Mağazaların nerelere açılacağı bilgisi için belirlenen kuralların yanı sıra nerelere açılmaması gerekdiğine dair kural setleride oluşturulmuştur. Kent ekonomi değeri 1 yani en düşük aktivite değerine ve kent nüfus yoğunluğu 1 nolu kategoride olan ilçelere mağaza açmak ticari olarak uygun değildir kuralı oluşturulmuştur. Bu şekilde açılamayacak ilçeler belirlenmiş olmaktadır. Ortalama yaya trafik yoğunluğu 20'nin altında olan bölgeler ve perakende yoğunluğu 25'in altında olan bölgelerde ticari olarak uygun olmayan yerlerdir. Bunların yanı sıra en önemli parametrelerden bir tanesinin perakende mağaza sayısı olduğu tesbit edilmiştir. Perakende mağaza sayısı 200'ün

üstünde olan bölgeler ticari olarak uygun bölgelerdir. Merkeze en yakın merkezi olmayan noktaya açılması gerekliliği ise veri analizlerinden anlaşılmaktadır.

Değer segmentesyonların yanı sıra incelenen davranış segmentasyonları ile belirli bölgelerin harcama alışkanlıkları, cinsiyet bazlı harcama yoğunlukları gibi parametreler ile çevre özellikleri birlikte incelenerek çeşitli kampanya veya ürün çeşitliliği konularına referans oluşturabilecek analizler oluşturulmuştur.

Metodoloji uzman görüşlerden ziyade veriye dayalı karar verme sistemi şeklinde geliştirilmiştir. Lokasyon seçimi modeline girdi olacak parametrelerin uzman görüşleri, veri analizi, deneme yanılma ve parametre seçim algoritmaları kullanılarak seçilmesi şeklinde yapılan ön çalışma ile de literatüre ve perakende sektörüne katkı sağlanması beklenmektedir.

Çalışma sonuçlarına göre literatürde bu konuda yapılmış son dönem çalışmalarını da destekleyecek şekilde Bileşik Makine Öğrenme yöntemleri ile Makine Öğrenme yöntemlerine göre daha iyi sonuçlar elde edilmiştir. Elde edilen sonuçlar, kullanılan Makine Öğrenme ve Bileşik Makine Öğrenme yöntemlerinin perakende sektöründe lokasyon analizini tahmin edebilecek yeteneğe sahip olduğunu göstermektedir.

KAYNAKLAR

Adeli, Hojjat, ve Shih-Lin Hung. 1994. *Machine learning: neural networks, genetic algorithms, and fuzzy systems*. John Wiley & Sons, Inc.

Adigüzel Tüylü, 2017. "Tersine Lojistik Sürecinde Bir Optimizasyon Çalışması"

Alpay, H. 2006. "Perakendecilikte Mağaza Yer Seçimi ve Kocaeli'de Örnek Bir Uygulama". *Basılmamış Yüksek Lisans Tezi*.

Alpaydin, Ethem. 2014. *Introduction to machine learning*. MIT press.

American Marketing Association. 2018. "Retail". 2018. https://www.ama.org/resources/Pages/Dictionary.aspx?dLetter=R.

Anyanwu, Matthew N, ve Sajjan G Shiva. 2009. "Comparative analysis of serial decision tree classification algorithms". *International Journal of Computer Science and Security* 3 (3). Citeseer:230–40.

Aydogmus, Hacer Yumurtacı, H İ Erdal, Onur Karakurt, Ersin Namli, Yusuf S Turkan, ve Hamit Erdal. 2015. "A comparative assessment of bagging ensemble models for modeling concrete slump flow". *Computers and Concrete* 16 (5). Techno-Press:741–57.

Basak, Debasish, Srimanta Pal, ve Dipak Chandra Patranabis. 2007. "Support vector regression". *Neural Information Processing-Letters and Reviews* 11 (10):203–24.

Bauer, Eric, ve Ron Kohavi. 1998. "An empirical comparison of voting classification algorithms: Bagging, boosting, and variants". *Machine learning* 36 (1):2.

Besfin. 2018. "2020 Küresel Perakende Trendleri Raporu:Amazon'un Perakende Sektörü Üzerindeki Etkisi Dünya'da ve Türkiye'de Perakende Sektörü". http://www.besfin.com/var/uploads/files/2020 Küresel Perakende Trendleri Raporu.pdf.

Bouckaert, Remco R. 2008. "Bayesian network classifiers in weka for version 3-5-7". *Artificial Intelligence Tools* 11 (3):369–87.

Bozkaya, Burçin, ve Seda Yanık. 2008. "Mağaza Yer Seçimi Optimizasyonu İçin Cbs Tabanlı Karar Destek Sistemi". *Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*.

Breiman, L. 2001. "Random Forests Machine Learning. 45: 5–32". *View Article PubMed/NCBI Google Scholar*.

Chaovalitwongse, Wanpracha Art, Ya-Ju Fan, ve Rajesh C Sachdeo. 2007. "Support feature machine for classification of abnormal brain activity". Içinde *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 113–22. ACM.

Çalış, A, S Kayapınar, ve T Çetinyokuş. 2014. "Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama". *Endüstri Mühendisliği*

Dergisi 25 (3–4):2–19.

Çalışkan, Sibel Kırmızıgül, ve İ Şoğukpınar. 2008. "KxKNN: K-Means ve K En Yakin Komşu Yöntemleri İle Ağlarda Nüfuz Tespiti". *EMO Yayınları*, 120–24.

Çölkesen, İ, ve T Kavzoğlu. 2011. "örnek tabanlı k-star algoritması ile uzaktan algılanmış görüntülerin sınıflandırılması".

Dasci, Abdullah, ve Gilbert Laporte. 2005. "A continuous model for multistore competitive location". *Operations Research* 53 (2). INFORMS:263–80.

Davies, Barry, ve Philippa Ward. 2002. *Managing retail consumption*.

Dawson, John. 2000. "Retailing at century end: some challenges for management and research". *The International Review of Retail, Distribution and Consumer Research* 10 (2). Taylor & Francis:119–48.

Deloitte. 2018. "Global Powers of Retailing 2018". 2018. https://www2.deloitte.com/global/en/pages/consumer-business/articles/global-powers-of-retailing.html.

Dicken, P. 1990. "The geography of enterpriseelements of a research agenda-. Smidt, M. and Wever, B. eds.: The corporate firm in a changing world economy-case studies in the geography of enterprise". Routledge, London.

Dicken, Peter, ve Peter E Lloyd. 1990. *Location in space: Theoretical perspectives in economic geography*. Prentice Hall.

Dietterich, Thomas G. 2002. "Ensemble learning". *The handbook of brain theory and neural networks* 2. MIT Press: Cambridge, MA:110–25.

Drezner, Tammy, ve Zvi Drezner. 2007. "The gravity p-median model". *European Journal of Operational Research* 179 (3). Elsevier:1239–51.

Durmuşoğlu, A., ve Y. Kahraman. 2016. "Facial expression recognition using geometric features". Içinde *International Conference on Systems, Signals, and Image Processing*. C. 2016–June. https://doi.org/10.1109/IWSSIP.2016.7502700.

Englert, Peter. 2012. "Locally Weighted Learning". Içinde *Seminar Class on Autonomous Learning Systems*.

Garner, Stephen R. 1995. "Weka: The waikato environment for knowledge analysis". Içinde *Proceedings of the New Zealand computer science research students conference*, 57–64. Citeseer.

Google/IPSOS OTX MediaCT. 2011. "The Mobile Movement: Understanding Smartphone Users".2011.https://www.google.com/search?q=The+Mobile+Movement+Understanding+Smartphone+Users+Google/IPSOS+OTX+MediaCT+U.S.,+April+2011&spell=1&sa=X&ved=0ahUKEwiX94SomsbaAhWnBZoKHXMrBScQBQglKAA&biw=1280&bih=588.

Graczyk, Magdalena, Tadeusz Lasota, Bogdan Trawiński, ve Krzysztof Trawiński. 2010. "Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal". Içinde *Asian Conference on Intelligent Information and Database Systems*, 340–50. Springer.

Hall, Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, ve Ian H Witten. 2009. "The WEKA data mining software: an update". *ACM SIGKDD explorations newsletter* 11 (1). ACM:10–18.

Han, Jiawei, Jian Pei, ve Micheline Kamber. 2011. *Data mining: concepts and techniques*. Elsevier.

Hayter, Roger. 1997. *The dynamics of industrial location: the factory, the firm and the production system*. Wiley.

Ho, Ricky. 2018. "Machine Learning". 2018. https://dzone.com/refcardz/machine-learning-predictive?chapter=1.

Ho, Tin Kam. 1998. "The random subspace method for constructing decision forests". *IEEE transactions on pattern analysis and machine intelligence* 20 (8). IEEE:832–44.

Holmes, Geoffrey, Mark Hall, ve Eibe Prank. 1999. "Generating rule sets from model trees". Içinde *Australasian Joint Conference on Artificial Intelligence*, 1–12. Springer.

Holte, Robert C. 1993. "Very simple classification rules perform well on most commonly used datasets". *Machine learning* 11 (1). Springer:63–90.

Hoover, Edgar Malone. 1937. *Location theory and the shoe leather industries*. C. 55. Harvard University Press.

Hotelling, Harold. 1990. "Stability in competition". Içinde *The Collected Economics Articles of Harold Hotelling*, 50–63. Springer.

Hsieh, Sheau-Ling, Sung-Huai Hsieh, Po-Hsun Cheng, Chi-Huang Chen, Kai-Ping Hsu, I-Shun Lee, Zhenyu Wang, ve Feipei Lai. 2012. "Design ensemble machine learning model for breast cancer diagnosis". *Journal of medical systems* 36 (5). Springer:2841–47.

Huff, David L. 1964. "Defining and estimating a trading area". *The Journal of Marketing*. JSTOR, 34–38.

Isard, Walter. 1975. *Introduction to regional science*. Prentice-Hall.

Izenman, Alan Julian. 2008. "Modern multivariate statistical techniques". *Regression, classification and manifold learning*. Springer.

Karagülle, Fatih. 2008. "Destek vektör makinelerini kullanarak yüz bulma".

Kohavi, Ron. 1995. "A study of cross-validation and bootstrap for accuracy estimation and model selection". Içinde *Ijcai*, 14:1137–45. Montreal, Canada.

KPMG. 2018. "Sektörel Bakış: Perakende".

https://assets.kpmg.com/content/dam/kpmg/tr/pdf/2018/01/sektorel-bakis-2018-perakende.pdf.

Kutner, Michael H, Chris Nachtsheim, ve John Neter. 2004. *Applied linear regression models*. McGraw-Hill/Irwin.

Levy, Michael, ve Barton A Weitz. 2004. *Retailing management*. McGraw-Hill/Irwin.

Loh, Wei-Yin. 2008. "Classification and regression tree methods". *Encyclopedia of statistics in quality and reliability*. Wiley Online Library.

McGoldrick, P. J. 1990. *Retail Marketing*. McGraw-Hill. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=McGoldrick%2C+P.J.+%281990%29%2C+Retail+Marketing%2C+McGraw-Hill%2C+Maidenhead.+&btnG=.

Merçan, G Gülçin. 2007. "Perakende Ticaretin Gelişimi Ve Türkiye Deki Yapı Marketlerin Tedarikçi Firmaların Ürün Tasarımı Yeteneklerinin Gelişimine Etkisi". Fen Bilimleri Enstitüsü.

Mucuk, İ. 1994. "Pazarlama ilkeleri.(6 bs)". *İstanbul: DER Yayınları*.

Namlı, Ersin. 2012. "Performance Analysis of Seaport Projects and Its Economic Contribution For Countries of Balkan Peninsula". *Balkan Sutdies IV economıcs & law & health*.

North, Douglass C. 1955. "Location theory and regional economic growth". *Journal of political economy* 63 (3). The University of Chicago Press:243–58.

Olivas, Emilio Soria. 2009. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques: Algorithms, Methods, and Techniques*. IGI Global.

Özkan, Y. 2008. "Veri Madenciliği Yöntemleri, Papatya Yayıncılık ve Eğitim AŞ". *İstanbul, Türkiye*.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, ve Vincent Dubourg. 2011. "Scikit-learn: Machine learning in Python". *Journal of machine learning research* 12 (Oct):2825–30.

PWC. 2013. "Online alışveriş yapanların gizemi: Çok kanallı perakendecilikte 10 mit". https://www.pwc.com.tr/tr/publications/industrial/retail-consumer/pdf/online-alisveris-yapanlarin-gizemi.pdf.

Quinlan, J Ross. 1993. "C4. 5: Programming for machine learning". *Morgan Kauffmann* 38:48.

Quinlan, John R. 1992. "Learning with continuous classes". Içinde *5th Australian joint conference on artificial intelligence*, 92:343–48. Singapore.

Reilly, William J. 1929. "Methods for the study of retail relationships". University of Texas at Austin.

"Retail Renaissance". 2011. 2011. http://trendwatching.com/trends/retailrenaissance/.

Sarmanova, Akkenzhe. 2013. "Veri madenciliğindeki sınıf dengesizliği sorununun giderilmesi". YTÜ Fen Bilimleri Enstitüsü.

Sheppard, Jeremy M, ve Warren B Young. 2006. "Agility literature review: Classifications, training and testing". *Journal of sports sciences* 24 (9). Taylor & Francis:919–32.

Sheth, Jagdish N, Banwari Mittal, ve Bruce I Newman. 1999. "Consumer behavior and beyond". *NY: Harcourt Brace*.

Shmilovici, Armin. 2009. "Support vector machines". Içinde *Data mining and knowledge discovery handbook*, 231–47. Springer.

Silva, Francisco, ve Daniel Serra. 2007. "Incorporating waiting time in competitive location models". *Networks and Spatial Economics* 7 (1). Springer:63–76.

Singhal, Swasti, ve Monika Jena. 2013. "A study on WEKA tool for data preprocessing, classification and clustering". *International Journal of Innovative technology and exploring engineering (IJItee)* 2 (6):250–53.

Steinwart, Ingo, ve Andreas Christmann. 2008. *Support vector machines*. Springer Science & Business Media.

Stone, Charles J. 1985. "Additive regression and other nonparametric models". *The annals of Statistics*. JSTOR, 689–705.

Subbulakshmi, T, S Mercy Shalinie, ve A Ramamoorthi. 2010. "Detection and classification of DDoS attacks using machine learning algorithms". *European Journal of Scientific Research* 47 (3):334–46.

Taaffe, Edward J, ve Howard L Gauthier. 1994. "Transportation geography and geographic thought in the United States: An overview". *Journal of Transport Geography* 2 (3). Elsevier:155–68.

Tatlıcı, Özge, ve Alattin Kızıltan. 2011. "ÇEKİM MODELİ: TÜRKİYE'NİN İHRACATI ÜZERİNE BİR UYGULAMA". *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi* 25.

Turney, Peter D. 1995. "Cost-sensitive classification: Empirical evaluation of a hybrid genetic decision tree induction algorithm". *Journal of artificial intelligence research* 2:369–409.

TÜİK. 2017. "Perakende Satış Endeksi 2017". http://www.tuik.gov.tr/PreHaberBultenleri.do?id=27855.

Türk Dil Kurumu. 2018. "Perakende". 2018. http://www.tdk.gov.tr/index.php?option=com_gts&arama=gts&guid=TDK.GTS.5ad848852831f5.56385955.

Vedat, D A L, ve Ahmet ÖZBEK. 2014. "AVRUPA BİRLİĞİ VE TÜRK HAZIR GİYİM SANAYİİNDEKİ E-TİCARET UYGULAMALARININ KARŞILAŞTIRILMALI

OLARAK İNCELENMESİ". *Elektronik Sosyal Bilimler Dergisi* 5 (18):106–25.

Wang, Yong, ve Ian H Witten. 1997. "Inducing model trees for continuous classes". Içinde *Proceedings of the Ninth European Conference on Machine Learning*, 128–37.

Weber, Alfred. 1929. *Theory of the Location of Industries*. University of Chicago Press.

Witten, Ian H, Eibe Frank, Mark A Hall, ve Christopher J Pal. 2016. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Yavan, Nuri. 2006. "TÜRKİYE'DE DOĞRUDAN YABANCI YATIRIMLARIN LOKASYON SEÇİMİ ÜZERİNE UYGULAMALI BİR ARAŞTIRMA". Ankara University. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiu66jPxMbaAhUCYZoKHcIfBAoQFggoMAA&url=http%3A%2F%2Facikarsiv.ankara.edu.tr%2Fbrowse%2F1732%2F2387.pdf&usg=AOvVaw3Jq80eeKoJEYuN9Av-B3lo.

YAVAN, Nuri. y.y. "BiLiM FELSEFESi BAKIMINDAN COGRAFYADA PoziT'rviST YAKLASIM".

Yıldız, Merve, ve Şadi Evren Şeker. 2016. "Veri Madenciliği Araçları (Data Mining Tools)". *YBS Ansiklopedi* 3:10–19.

Yılmaz, Hülya. 2014. "Random Forests yönteminde kayıp veri probleminin incelenmesi ve sağlık alanında bir uygulama". Eskişehir Osmangazi Üniversitesi.

Zhang, Min-Ling, ve Zhi-Hua Zhou. 2005. "A k-nearest neighbor based algorithm for multi-label classification". Içinde *Granular Computing, 2005 IEEE International Conference on*, 2:718–21. IEEE.

EKLER

EK-1 1000 metre çevresinde 180 adetin üstünde perakende mağaza bulunan caddeler (Her ilçe için en büyük cadde alınmıştır)

CustomID	Cadde Adı	Cadde Adı	Tur	Cadde Türü	İl Adı	İlçe Adı	TOPLAM
CADDE_10147__1	Ebe Kızı Sk.	Ebe Kızı Sk.	Sokak	Lokal Segment	İstanbul	Şişli	779
CADDE_5446__1	Hakkı Yeten Cd.	Hakkı Yeten Cd.	Cadde	Lokal Segment	İstanbul	Beşiktaş	631
CADDE_102885__1	Mescit Sk.	Mescit Sk.	Sokak	Lokal Segment	İstanbul	Kağıthane	537
CADDE_3547__1	Zuhurat Baba Cd.	Zuhurat Baba Cd.	Cadde	Lokal Segment	İstanbul	Bakırköy	520
CADDE_6531__1	Halk Sk.	Halk Sk.	Sokak	Lokal Segment	İstanbul	Kadıköy	450
CADDE_5078__1	Doğanbey Toki Konutları	Doğanbey Toki Konutları	Sokak	Lokal Segment	Bursa	Osmangazi	440
CADDE_2597__1	Şair Fuzuli Cd.	Şair Fuzuli Cd.	Cadde	Lokal Segment	Eskişehir	Odunpazarı	398
CADDE_2602__1	Cengiz Topel-1 Cd.	Cengiz Topel-1 Cd.	Cadde	Lokal Segment	Eskişehir	Tepebaşı	392
CADDE_2497__1	Ahmet Paşa Cd.	Ahmet Paşa Cd.	Cadde	Lokal Segment	Kayseri	Kocasinan	376
CADDE_761__2	Mithatpaşa Cd.	Mithatpaşa Cd.	Cadde	Lokal Segment	Ankara	Çankaya	367
CADDE_481__1	Seyyid Burhanettin Blv.	Seyyid Burhanettin Blv.	Bulvar	Lokal Segment	Kayseri	Melikgazi	364
CADDE_5318__2	Sofular Cd.	Sofular Cd.	Cadde	Lokal Segment	İstanbul	Fatih	363
CADDE_135__3	3. Cadde	3. Cadde	Cadde	Genel Segment	Gaziantep	Şehitkamil	353
CADDE_5441__3	E-5 Yanyolu	E-5 Yanyolu	Sokak	Lokal Segment	İstanbul	Beylikdüzü	352
CADDE_80__2	Şarampol Caddesi	Şarampol Caddesi	Cadde	Genel Segment	Antalya	Muratpaşa	347
CADDE_5767__1	E-5 Yanyolu	E-5 Yanyolu	Sokak	Lokal Segment	İstanbul	Esenyurt	343
CADDE_102381__1	Prof. Dr. Turan Güneş Cd.	Prof. Dr. Turan Güneş Cd.	Cadde	Lokal Segment	İstanbul	Zeytinburnu	337
CADDE_299__1	Şht. Nevres Bulvarı	Şht. Nevres Bulvarı	Bulvar	Yeme İçme	İzmir	Konak	335
CADDE_100609__1	Ayşe Hatun Çeşme Sk.	Ayşe Hatun Çeşme Sk.	Sokak	Lokal Segment	İstanbul	Ataşehir	331
CADDE_378__5	İstiklal Caddesi	İstiklal Caddesi	Cadde	Genel Segment	Samsun	İlkadım	318
CADDE_103587__1	İsmet Paşa Cd.	İsmet Paşa Cd.	Cadde	Lokal Segment	İstanbul	Bayrampaşa	312
CADDE_590__5	Sütçü İmam Cd.	Sütçü İmam Cd.	Cadde	Genel Segment	İstanbul	Ümraniye	305
CADDE_6855__1	Mete Sk.	Mete Sk.	Sokak	Lokal Segment	İstanbul	Güngören	304
CADDE_324__6	Hürriyet Caddesi	Hürriyet Caddesi	Cadde	Genel Segment	Kocaeli	İzmit	301
CADDE_5743__1	Leyla Atakan Cd.	Leyla Atakan Cd.	Cadde	Lokal Segment	Kocaeli	İzmit	301
CADDE_258__15	Alemdağ Caddesi	Alemdağ Caddesi	Cadde	Genel Segment	İstanbul	Üsküdar	296
CADDE_5659__1	Gazi Muhtar Paşa Blv.	Gazi Muhtar Paşa Blv.	Bulvar	Lokal Segment	Gaziantep	Şahinbey	290
CADDE_6822__1	Hamalbaşı Cd.	Hamalbaşı Cd.	Cadde	Lokal Segment	İstanbul	Beyoğlu	290
CADDE_1936__2	Atatürk Cd.	Atatürk Cd.	Cadde	Lokal Segment	Kocaeli	Gebze	290
CADDE_101327__1	Sandalzade Blv.	Sandalzade Blv.	Bulvar	Lokal Segment	Kahramanmaraş	Onikişubat	286
CADDE_2751__1	Cumhuriyet Cd.	Cumhuriyet Cd.	Cadde	Lokal Segment	Adana	Seyhan	285
CADDE_159__2	Ahmet Yesevi Caddesi	Ahmet Yesevi Caddesi	Cadde	Lokal Segment	İstanbul	Bahçelievler	282
CADDE_7571__1	Zübeyde Hanım Blv.	Zübeyde Hanım Blv.	Bulvar	Lokal Segment	Kahramanmaraş	Dulkadiroğlu	273
CADDE_5387__1	Büyükdere Cd.	Büyükdere Cd.	Cadde	Lokal Segment	İstanbul	Sarıyer	272
CADDE_8900__1	Mithatpaşa Cd.	Mithatpaşa Cd.	Cadde	Lokal Segment	İzmir	Balçova	255
CADDE_103116__2	Atatürk Blv.	Atatürk Blv.	Bulvar	Lokal Segment	Ankara	Altındağ	252
CADDE_3671__1	Vali İzzetbey Cd.	Vali İzzetbey Cd.	Cadde	Lokal Segment	Konya	Karatay	248
CADDE_103039__1	Alaaddin Blv.	Alaaddin Blv.	Bulvar	Lokal Segment	Konya	Meram	247
CADDE_3599__3	İkitelli Cd.	İkitelli Cd.	Cadde	Lokal Segment	İstanbul	Küçükçekmece	245
CADDE_331__4	İnönü Caddesi	İnönü Caddesi	Cadde	Genel Segment	Malatya	Battalgazi	243
CADDE_5280__3	Çiğdem Cd.	Çiğdem Cd.	Cadde	Lokal Segment	İstanbul	Avcılar	237
CADDE_2240__3	Atatürk Blv.	Atatürk Blv.	Bulvar	Lokal Segment	Sakarya	Adapazarı	237
CADDE_5048__1	Büyükşehir Belediyesi Toptancı Hali	Büyükşehir Belediyesi Toptancı Hal	Sokak	Lokal Segment	Ankara	Yenimahalle	231
CADDE_103098__1	Hipodrum Cd.	Hipodrum Cd.	Cadde	Lokal Segment	Ankara	Yenimahalle	231
CADDE_103098__2	Hipodrum Cd.	Hipodrum Cd.	Cadde	Lokal Segment	Ankara	Yenimahalle	231
CADDE_103208__4	Buhara Blv.	Buhara Blv.	Bulvar	Lokal Segment	Malatya	Yeşilyurt	231
CADDE_328__1	Alaaddin Bulvarı	Alaaddin Bulvarı	Bulvar	Yeme İçme	Konya	Selçuklu	229
CADDE_145__3	Atatürk Caddesi	Atatürk Caddesi	Cadde	Genel Segment	Hatay	Antakya	225
CADDE_2102__2	Cumhuriyet Cd.	Cumhuriyet Cd.	Cadde	Lokal Segment	Hatay	Antakya	225
CADDE_103142__1	Gazi Mustafa Kemal Blv.	Gazi Mustafa Kemal Blv.	Bulvar	Lokal Segment	Mersin	Yenişehir	222
CADDE_6849__1	Çiçeksuyu Cd.	Çiçeksuyu Cd.	Cadde	Lokal Segment	İstanbul	Eyüp	218
CADDE_9405__1	Turan Güneş Cd.	Turan Güneş Cd.	Cadde	Lokal Segment	Denizli	Pamukkale	217
CADDE_101919__1	Bağlar Cd.	Bağlar Cd.	Cadde	Lokal Segment	İstanbul	Gaziosmanpaşa	215
AVM_100177__1	1835/1 Sk.	1835/1 Sk.	Sokak	Lokal Segment	İzmir	Karşıyaka	211
CADDE_5639__3	Atatürk Blv.	Atatürk Blv.	Bulvar	Lokal Segment	İzmir	Karşıyaka	211
CADDE_9015__2	Doç. Dr. Bahriye Üçok Blv.	Doç. Dr. Bahriye Üçok Blv.	Bulvar	Lokal Segment	İzmir	Karşıyaka	211
CADDE_101758__1	Namık Kemal Cd.	Namık Kemal Cd.	Cadde	Lokal Segment	İstanbul	Pendik	210
CADDE_6601__3	Saraylar Cd.	Saraylar Cd.	Cadde	Lokal Segment	İstanbul	Kartal	208
CADDE_1659__1	Fevzi Çakmak Cd.	Fevzi Çakmak Cd.	Cadde	Lokal Segment	Sivas	Merkez	207
CADDE_100520__1	Ankara Cd.	Ankara Cd.	Cadde	Lokal Segment	İzmir	Bornova	206
CADDE_743__1	2. Sk.	2. Sk.	Sokak	Lokal Segment	Şanlıurfa	Haliliye	205
CADDE_982__7	Akçay Cd.	Akçay Cd.	Cadde	Lokal Segment	İzmir	Gaziemir	200
CADDE_440__1	Kahramanmaraş Cd.	Kahramanmaraş Cd.	Cadde	Lokal Segment	Trabzon	Ortahisar	199
CADDE_396__4	Uzun Sokak	Uzun Sokak	Sokak	Genel Segment	Trabzon	Ortahisar	199
CADDE_101618__1	Muhittin Cami Sk.	Muhittin Cami Sk.	Sokak	Lokal Segment	Trabzon	Ortahisar	199
CADDE_100274__1	496. Sk.	496. Sk.	Sokak	Lokal Segment	Denizli	Merkezefendi	198
CADDE_5350__3	Turgut Özal Blv.	Turgut Özal Blv.	Bulvar	Lokal Segment	Diyarbakır	Bağlar	197
CADDE_7139__1	3. Sk.	3. Sk.	Sokak	Lokal Segment	İstanbul	Bağcılar	196
CADDE_113__3	Gazi Caddesi	Gazi Caddesi	Cadde	Genel Segment	Çorum	Merkez	195
CADDE_101016__1	Kulaksız Sk.	Kulaksız Sk.	Sokak	Lokal Segment	Çorum	Merkez	195
CADDE_10052__1	108. Cd.	108. Cd.	Cadde	Lokal Segment	Mersin	Akdeniz	195
CADDE_100882__1	Şehit Nusret Cd.	Şehit Nusret Cd.	Cadde	Lokal Segment	Şanlıurfa	Eyyübiye	192
CADDE_102359__1	Bülbül Deresi Cd.	Bülbül Deresi Cd.	Cadde	Lokal Segment	Ordu	Altınordu	187
CADDE_102358__1	Gündoğdu Sk.	Gündoğdu Sk.	Sokak	Lokal Segment	Ordu	Altınordu	187
CADDE_51__2	Etlik Caddesi	Etlik Caddesi	Cadde	Lokal Segment	Ankara	Keçiören	186
CADDE_95__1	Milli Kuvvetler Caddesi	Milli Kuvvetler Caddesi	Cadde	Genel Segment	Balıkesir	Karesi	185
CADDE_7563__1	100. Yıl Blv.	100. Yıl Blv.	Bulvar	Lokal Segment	Samsun	Canik	184
CADDE_103235__5	Şanlıurfa Blv.	Şanlıurfa Blv.	Bulvar	Lokal Segment	Diyarbakır	Kayapınar	182
CADDE_6988__1	Tugay Yolu Cd.	Tugay Yolu Cd.	Cadde	Lokal Segment	İstanbul	Maltepe	182

EK-2 500 metre çevresinde 20 adetin üstünde ulaşım noktası bulunan caddeler (Her ilçe için en büyük cadde alınmıştır)

CustomID	Cadde Adı	Cadde Adı	Tur	Cadde Türü	İl Adı	İlçe Adı	TOPLA
CADDE_9649__1	Menekşe-1 Sk.	Menekşe-1 Sk.	Sokak	Lokal Segment	Ankara	Çankaya	50
CADDE_2924__1	Gazi Osman Paşa Blv.	Gazi Osman Paşa Blv.	Bulvar	Lokal Segment	İzmir	Konak	48
CADDE_126__1	Ordu Caddesi	Ordu Caddesi	Cadde	Yeme İçme	Erzincan	Merkez	45
CADDE_7986__1	Abdül Ezel Paşa Cd.	Abdül Ezel Paşa Cd.	Cadde	Lokal Segment	Konya	Selçuklu	45
CADDE_2724__2	Hastane Cd.	Hastane Cd.	Cadde	Lokal Segment	Konya	Selçuklu	45
CADDE_2552__2	8 Eylül Cd.	8 Eylül Cd.	Cadde	Lokal Segment	Manisa	Şehzadeler	43
CADDE_8071__1	Yaşiltepe Cd.	Yaşiltepe Cd.	Cadde	Lokal Segment	Manisa	Şehzadeler	43
CADDE_3766__1	Şht. Binbaşı Cengiz Toytunç Cd.	Şht. Binbaşı Cengiz Toytunç Cd.	Cadde	Lokal Segment	Antalya	Muratpaşa	39
CADDE_326__2	Kazım Karabekir Caddesi	Kazım Karabekir Caddesi	Cadde	Genel Segment	Konya	Meram	37
CADDE_2842__1	Ankara Yolu Cd.	Ankara Yolu Cd.	Cadde	Lokal Segment	Bursa	Yıldırım	36
CADDE_111__2	Demircioğlu Caddesi	Demircioğlu Caddesi	Cadde	Genel Segment	Çanakkale	Merkez	35
CADDE_5152__1	Kural Sk.	Kural Sk.	Sokak	Lokal Segment	Antalya	Alanya	33
CADDE_6017__1	Sobucalı Sk.	Sobucalı Sk.	Sokak	Lokal Segment	Aydın	Kuşadası	33
CADDE_2331__1	Atatürk Cd.	Atatürk Cd.	Cadde	Lokal Segment	Balıkesir	Bandırma	33
CADDE_107__4	Fatih Sultan Mehmet Bulvarı	Fatih Sultan Mehmet Bulvarı	Bulvar	Yeme İçme	Bursa	Nilüfer	33
CADDE_145__1	Atatürk Caddesi	Atatürk Caddesi	Cadde	Genel Segment	Hatay	Antakya	33
CADDE_425__1	Forbes Sk.	Forbes Sk.	Sokak	Lokal Segment	İzmir	Buca	33
CADDE_3671__1	Vali İzzetbey Cd.	Vali İzzetbey Cd.	Cadde	Lokal Segment	Konya	Karatay	33
CADDE_2429__1	Kahramanmaraş Cd.	Kahramanmaraş Cd.	Cadde	Lokal Segment	Trabzon	Ortahisar	33
CADDE_396__1	Uzun Sokak	Uzun Sokak	Sokak	Genel Segment	Trabzon	Ortahisar	33
CADDE_7252__2	Yavuz Selim Blv.	Yavuz Selim Blv.	Bulvar	Lokal Segment	Trabzon	Ortahisar	33
CADDE_811__1	Sanayi Cd.	Sanayi Cd.	Cadde	Lokal Segment	Ankara	Altındağ	32
CADDE_632__2	Şht. Teğmen Kalmaz Cd.	Şht. Teğmen Kalmaz Cd.	Cadde	Lokal Segment	Ankara	Altındağ	32
CADDE_103120__1	Çavuş Sk.	Çavuş Sk.	Sokak	Lokal Segment	Ankara	Altındağ	32
CADDE_9134__1	1639. Sk.	1639. Sk.	Sokak	Lokal Segment	İzmir	Bayraklı	32
CADDE_5838__1	Zübeyde Hanım Cd.	Zübeyde Hanım Cd.	Cadde	Lokal Segment	Bursa	Osmangazi	31
CADDE_2551__1	8 Eylül Cd.	8 Eylül Cd.	Cadde	Lokal Segment	Manisa	Yunusemre	31
CADDE_10359__1	Battal Gazi Cd.	Battal Gazi Cd.	Cadde	Lokal Segment	Kocaeli	Darıca	30
CADDE_1940__3	İstasyon Cd.	İstasyon Cd.	Cadde	Lokal Segment	Kocaeli	Darıca	30
CADDE_5800__1	Baha Sk.	Baha Sk.	Sokak	Lokal Segment	Samsun	İlkadım	30
CADDE_9678__1	13-2. Sk.	13-2. Sk.	Sokak	Lokal Segment	Sivas	Merkez	30
CADDE_103493__1	13-3. Sk.	13-3. Sk.	Sokak	Lokal Segment	Sivas	Merkez	30
CADDE_10167__1	Kızılay Cd.	Kızılay Cd.	Cadde	Lokal Segment	Tekirdağ	Süleymanpaşa	30
CADDE_100570__1	Kolordu Cd.	Kolordu Cd.	Cadde	Lokal Segment	Tekirdağ	Süleymanpaşa	30
CADDE_101937__1	Orta Cami Sk.	Orta Cami Sk.	Sokak	Lokal Segment	Tekirdağ	Süleymanpaşa	30
CADDE_9953__1	Cumhuriyet 5. Sk. (Cumhuriyet Cc	Cumhuriyet 5. Sk.	Sokak	Lokal Segment	Van	İpekyolu	30
CADDE_95__1	Milli Kuvvetler Caddesi	Milli Kuvvetler Caddesi	Cadde	Genel Segment	Balıkesir	Karesi	29
CADDE_101239__1	Çiviciler Cd.	Çiviciler Cd.	Cadde	Lokal Segment	Balıkesir	Karesi	29
CADDE_7937__1	Tennuri Gç.	Tennuri Gç.	Sokak	Lokal Segment	Kayseri	Melikgazi	29
CADDE_100937__1	Yoğunburç Cd.	Yoğunburç Cd.	Cadde	Lokal Segment	Kayseri	Melikgazi	29
CADDE_3801__1	Bankalar Cd.	Bankalar Cd.	Cadde	Lokal Segment	Kırıkkale	Merkez	29
CADDE_5870__1	İzmir Cd.	İzmir Cd.	Cadde	Lokal Segment	Kırıkkale	Merkez	29
CADDE_7535__1	Zafer Cd.	Zafer Cd.	Cadde	Lokal Segment	Kırıkkale	Merkez	29
CADDE_874__2	Uzgörenler Sk.	Uzgörenler Sk.	Sokak	Lokal Segment	Ankara	Mamak	28
CADDE_8688__1	494. Sk.	494. Sk.	Sokak	Lokal Segment	İzmir	Karabağlar	28
CADDE_8704__1	502/2. Sk.	502/2. Sk.	Sokak	Lokal Segment	İzmir	Karabağlar	28
CADDE_838__2	Cumhuriyet Cd.	Cumhuriyet Cd.	Cadde	Lokal Segment	Ankara	Gölbaşı	27
CADDE_680__1	Çilingirler Cd.	Çilingirler Cd.	Cadde	Lokal Segment	Edirne	Merkez	27
CADDE_7641__1	Hoca Ahmet Yesevi Cd.	Hoca Ahmet Yesevi Cd.	Cadde	Lokal Segment	Edirne	Merkez	27
CADDE_102504__1	Zübeyde Hanım Cd.	Zübeyde Hanım Cd.	Cadde	Lokal Segment	Edirne	Merkez	27
CADDE_5063__2	İnönü Cd.	İnönü Cd.	Cadde	Lokal Segment	İstanbul	Beyoğlu	27
CADDE_5064__1	Miralay Şefikbey Sk.	Miralay Şefikbey Sk.	Sokak	Lokal Segment	İstanbul	Beyoğlu	27
CADDE_1924__3	Büyükdere Cd.	Büyükdere Cd.	Cadde	Lokal Segment	İstanbul	Şişli	27
CADDE_6019__1	Uzun Zeybek Sk.	Uzun Zeybek Sk.	Sokak	Lokal Segment	Kütahya	Merkez	27
CADDE_10270__1	Alemdar Sk.	Alemdar Sk.	Sokak	Lokal Segment	Zonguldak	Merkez	27
CADDE_67__2	Münir Özkul Liman Caddesi	Münir Özkul Liman Caddesi	Cadde	Lokal Segment	Antalya	Kemer	26
CADDE_6190__3	Bahabey Cd.	Bahabey Cd.	Cadde	Lokal Segment	Çorum	Merkez	26
CADDE_7352__2	Aksu Cd.	Aksu Cd.	Cadde	Lokal Segment	Isparta	Merkez	26
CADDE_462__3	Cumhuriyet Cd.	Cumhuriyet Cd.	Cadde	Lokal Segment	Isparta	Merkez	26
CADDE_5384__2	Zeytinoğlu Cd.	Zeytinoğlu Cd.	Cadde	Lokal Segment	İstanbul	Beşiktaş	26
CADDE_6692__1	Bahçelerarası Sk.	Bahçelerarası Sk.	Sokak	Lokal Segment	İstanbul	Kadıköy	26
CADDE_2264__2	Denizli Osb. / 2. Ticari Yol	Denizli Osb. / 2. Ticari Yol	Sokak	Sanayi	Denizli	Merkezefendi	25
CADDE_102196__1	Cumhuriyet Cd.	Cumhuriyet Cd.	Cadde	Lokal Segment	Denizli	Merkezefendi	25
CADDE_2262__2	Gazi Mustafa Kemal Blv.	Gazi Mustafa Kemal Blv.	Bulvar	Lokal Segment	Denizli	Pamukkale	25
CADDE_227__3	Eski Büyükdere Caddesi	Eski Büyükdere Caddesi	Cadde	Genel Segment	İstanbul	Kağıthane	25
CADDE_5482__2	Bosna Blv.	Bosna Blv.	Bulvar	Lokal Segment	İstanbul	Üsküdar	25
CADDE_478__3	Sivas Blv.	Sivas Blv.	Bulvar	Lokal Segment	Kayseri	Kocasinan	25
CADDE_372__2	Cumhuriyet Caddesi	Cumhuriyet Caddesi	Cadde	Genel Segment	Rize	Merkez	25
CADDE_8618__1	Ayvalı Cd.	Ayvalı Cd.	Cadde	Lokal Segment	Ankara	Keçiören	24
CADDE_2170__1	Mumcu Cd.	Mumcu Cd.	Cadde	Lokal Segment	Erzurum	Yakutiye	24
CADDE_2170__2	Mumcu Cd.	Mumcu Cd.	Cadde	Lokal Segment	Erzurum	Yakutiye	24
CADDE_9128__1	6800. Sk.	6800. Sk.	Sokak	Lokal Segment	İzmir	Çiğli	24
CADDE_925__2	6065. Sk.	6065. Sk.	Sokak	Lokal Segment	İzmir	Karşıyaka	24
CADDE_2561__2	Belediye Cd.	Belediye Cd.	Cadde	Lokal Segment	Manisa	Salihli	24
CADDE_2560__2	Mithatpaşa Cd.	Mithatpaşa Cd.	Cadde	Lokal Segment	Manisa	Salihli	24
CADDE_5826__1	İdman Sk.	İdman Sk.	Sokak	Lokal Segment	Manisa	Soma	24
CADDE_391__8	Gaziosmanpaşa Bulvarı	Gaziosmanpaşa Bulvarı	Bulvar	Genel Segment	Tokat	Merkez	24
CADDE_101426__1	Bosna 6. Sk.	Bosna 6. Sk.	Sokak	Lokal Segment	Tokat	Merkez	24
CADDE_101427__1	Fatih Cd. 3. Sk.	Fatih Cd. 3. Sk.	Sokak	Lokal Segment	Tokat	Merkez	24
CADDE_103497__1	803. Sk.	803. Sk.	Sokak	Lokal Segment	Van	Erciş	24
CADDE_8499__1	Ragıp Tüzün Cd.	Ragıp Tüzün Cd.	Cadde	Lokal Segment	Ankara	Yenimahalle	23
CADDE_103097__2	Mevlana Blv.	Mevlana Blv.	Bulvar	Lokal Segment	Ankara	Yenimahalle	23
CADDE_2789__1	Mehmet Akyolu Cd.	Mehmet Akyolu Cd.	Cadde	Lokal Segment	Bursa	İnegöl	23
CADDE_8332__1	Murat Cd.	Murat Cd.	Cadde	Lokal Segment	Bursa	İnegöl	23
CADDE_550__2	Gazi Mustafa Kemal Paşa Cd.	Gazi Mustafa Kemal Paşa Cd.	Cadde	Lokal Segment	İstanbul	Fatih	23
CADDE_232__1	Üsküdar Caddesi	Üsküdar Caddesi	Cadde	Lokal Segment	İstanbul	Kartal	23
CADDE_6658__3	Adem Yavuz Cd.	Adem Yavuz Cd.	Cadde	Lokal Segment	İstanbul	Ümraniye	23
CADDE_9325__1	853. Sk.	853. Sk.	Sokak	Lokal Segment	İzmir	Bornova	23
CADDE_103780__1	Özgürlük Cd.	Özgürlük Cd.	Cadde	Lokal Segment	Kocaeli	Çayırova	23
CADDE_618__1	Öncü Sk.	Öncü Sk.	Sokak	Lokal Segment	Ankara	Sincan	22
CADDE_96__2	Hükümet Caddesi	Hükümet Caddesi	Cadde	Genel Segment	Bartın	Merkez	22
CADDE_2951__17	Atatürk Blv.	Atatürk Blv.	Bulvar	Lokal Segment	Batman	Merkez	22
CADDE_2951__4	Atatürk Blv.	Atatürk Blv.	Bulvar	Lokal Segment	Batman	Merkez	22
CADDE_9155__1	Namık Kemal Cd.	Namık Kemal Cd.	Cadde	Lokal Segment	İzmir	Aliağa	22
CADDE_981__1	5. Sk.	5. Sk.	Sokak	Lokal Segment	İzmir	Gaziemir	22
CADDE_356__3	Kemal Seyfettin Elgin Bulvarı	Kemal Seyfettin Elgin Bulvarı	Bulvar	Yeme İçme	Muğla	Marmaris	22
CADDE_6072__1	İnönü Cd.	İnönü Cd.	Cadde	Lokal Segment	Zonguldak	Kilimli	22
CADDE_2045__3	Kadınana Cd.	Kadınana Cd.	Cadde	Lokal Segment	Afyonkarahisar	Merkez	21
CADDE_103__5	İzzet Baysal Caddesi	İzzet Baysal Caddesi	Cadde	Lokal Segment	Bolu	Merkez	21
CADDE_614__3	Şht. Cahar Dudayev Cd.	Şht. Cahar Dudayev Cd.	Cadde	Lokal Segment	İstanbul	Ataşehir	21
CADDE_1800__2	İncirli Cd.	İncirli Cd.	Cadde	Lokal Segment	İstanbul	Bakırköy	21
CADDE_10157__1	Hasan Ercan Cd.	Hasan Ercan Cd.	Cadde	Lokal Segment	Muğla	Menteşe	21

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Ahmet Kayıran
Doğum Yeri	Kahramanmaraş
Doğum Tarihi	25.02.1989
Uyruğu	×T.C. Diğer:
Telefon	05304127037
E-Posta Adresi	Kayiran.eng@gmail.com
Web Adresi	

Eğitim Bilgileri	
Lisans	
Üniversite	**İstanbul Üniversitesi**
Fakülte	Mühendislik Fakültesi
Bölümü	Endüstri Mühendisliği
Mezuniyet Yılı	21.07.2011

Yüksek Lisans	
Üniversite	İstanbul Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	Endüstri Mühendisliği Ana Bilim Dalı
Programı	Endüstri Mühendisliği
Mezuniyet Tarihi	04.05.2018