



SIVAS CUMHURİYET ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü
Yönetim Bilişim Sistemleri Ana Bilim Dalı

MÜŞTERİ TALEPLERİNİN ORGANİZASYONEL İŞ AKIŞLARINA
UYGUN SINIFLANDIRILMASI

Yüksek Lisans Tezi

Sümeyye ŞAHİN

Sivas

Ağustos 2019

SİVAS CUMHURİYET ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü
Yönetim Bilişim Sistemleri Ana Bilim Dalı

MÜŞTERİ TALEPLERİNİN ORGANİZASYONEL İŞ AKIŞLARINA
UYGUN SINIFLANDIRILMASI

Yüksek Lisans Tezi

Sümeyye ŞAHİN

Tez Danışman:


Dr. Öğr. Üyesi Halil ARSLAN

Sivas

Ağustos 2019

KABUL VE ONAY

Üniversite : Sivas Cumhuriyet Üniversitesi
Enstitü : Sosyal Bilimler Enstitüsü
Ana Bilim Dalı : Yönetim Bilişim Sistemleri Ana Bilim Dalı
Bilim Dalı :
Tezin Başlığı : Müşteri Taleplerinin Organizasyonel İş Akışlarına Uygun Sınıflandırılması
Savunma Tarihi : 02.08.2019
Danışmanı : Dr. Öğr. Üyesi Halil ARSLAN

	Unvanı - Adı Soyadı	İmza
Jüri Başkanı	Prof. Dr. Oğuz KAYNAR	
Üye	Doç. Dr. Galip AYDIN	
Üye	Dr. Öğr. Üyesi Halil ARSLAN	

Oy Birliği

Oy Çokluğu

Sümeyye ŞAHİN tarafından hazırlanan "Müşteri Taleplerinin Organizasyonel İş Akışlarına Uygun Sınıflandırılması" başlıklı yüksek lisans tezi, kabul edilmiştir.

.../.../ 2019

Prof. Dr. Ahmet ŞENGÖNÜL

Enstitü Müdürü

ETİK İLKELERE UYGUNLUK BEYANI

Sivas Cumhuriyet Üniversitesi Sosyal Bilimler Enstitüsü bünyesinde hazırladığım bu Yüksek Lisans tezinin bizzat tarafımdan ve kendi sözcüklerimle yazılmış orijinal bir çalışma olduğunu ve bu tezde;

- 1- Çeşitli yazarların çalışmalarından faydalandığımda bu çalışmaların ilgili bölümlerini doğru ve net biçimde göstererek yazarlara açık biçimde atıfta bulunduğumu;
- 2- Yazdığım metinlerin tamamı ya da sadece bir kısmı, daha önce herhangi bir yerde yayımlanmışsa bunu da açıkça ifade ederek gösterdiğimi;
- 3- Başkalarına ait alıntılanan tüm verileri (tablo, grafik, şekil vb. de dahil olmak üzere) atıflarla belirttiğimi;
- 4- Başka yazarların kendi kelimeleriyle alıntıladığım metinlerini, tırnak içerisinde veya farklı dizerek verdiğim yine başka yazarlara ait olup fakat kendi sözcüklerimle ifade ettiğim hususları da istisnasız olarak kaynak göstererek belirttiğimi,

beyan ve bu etik ilkeleri ihlal etmiş olmam halinde bütün sonuçlarına katlanacağımı kabul ederim.

26.08/2019

Sümeyye ŞAHİN



TEŐEKKÜR

Yüksek Lisans tez çalışmam süresince yardımlarını esirgemeyen, her konuda değerli fikirlerini sunarak bana yön gösteren, imkan ve destek sağlayan tez danışmanlarım Sayın Dr. Öğr. Üyesi Halil ARSLAN ve Prof. Dr. Oğuz KAYNAR'a teşekkürü borç bilirim.

Tez çalışmamın başlamasıyla birlikte değerli kaynaklarını benimle paylaşan Detaysoft firmasına ve yüksek lisans eğitim süresi boyunca maddi ve manevi desteklerini esirgemeyen her zaman yanımda olan aileme ve dostlarıma en içten şükranlarımı sunarım.

Sümeyye ŐAHIN

Ağustos, 2019

İÇİNDEKİLER

İÇİNDEKİLER	i
KISALTMALAR	v
TABLO LİSTESİ	vii
ŞEKİL LİSTESİ	ix
ÖZET	xi
ABSTRACT	xiii
BÖLÜM 1	1
GİRİŞ	1
Araştırmanın Amacı	3
Araştırmanın Önemi	4
BÖLÜM 2	5
BENZER ÇALIŞMALAR	5
BÖLÜM 3	9
YARDIM MASASI	9
3.1. Müşteri Kavramının Tanımı	9
3.1.1. İç Müşteri	10
3.1.2. Dış Müşteri	10
3.2. İş / Görev Yönetimi	11
3.3. Organizasyon Yapısı	12
3.4. Yardım Masası ve Yönetimi	14
BÖLÜM 4	15
YÖNTEM	15
4.1. Veri Setinin Toplanması	15
4.2. Veri Ön İşleme	15

4.3. Öz Nitelik Çıkarımı	16
4.3.1. Bag of Words (Bow) Modeli	16
4.3.1.1. Bitsel Tanımlama	17
4.3.1.2. Frekansa Göre Tanımlama	17
4.3.1.3. TF_IDF Ağırlıklandırma Yöntemine göre Tanımlama	18
4.3.2. Word2Vec Modeli	18
4.3.2.1. CBoW Öğrenme Modeli	20
4.3.2.2. Skip-Gram Öğrenme Modeli.....	21
4.3.3. Doc2Vec Modeli	21
4.3.3.1. DBoW Öğrenme Modeli	22
4.3.3.2. DM Öğrenme Modeli	22
4.4. Sınıflandırma Yöntemleri	23
4.4.1. Destek Vektör Makinaları.....	23
4.4.2. Naïve Bayes Algoritması	25
4.4.3. Çok Katmanlı Yapay Sinir Ağları.....	25
4.5. Değerlendirme Kriterleri	27
4.5.1. Doğruluk – Hata Oranı.....	27
4.5.2. Keskinlik	28
4.5.3. Duyarlılık	28
4.5.4. F – Ölçüsü	28
4.6. Cross Validation (Çapraz Doğrulama)	29
BÖLÜM 5.....	31
UYGULAMA.....	31
5.1. Taleplerin Toplanması.....	31
5.2. Ön İşlem Aşamaları.....	32
5.3. Talep Verilerinin Modele Göre Sınıflandırılması	33

5.3.1. Veri Setinin Eğitim-Test Olarak Ayrılması Yöntemi	33
5.3.1.1. BoW Öznitelik Çıkarım Yöntemiyle Sınıflandırma	33
5.3.1.2. Doc2Vec-DBoW Öznitelik Çıkarım Yöntemiyle Sınıflandırma	36
5.3.1.3. Doc2Vec-DM Öznitelik Çıkarım Yöntemiyle Sınıflandırma	38
5.3.2. 10-Kat Çapraz Doğrulama Yöntemi	40
5.3.2.1. BoW Öznitelik Çıkarım Yöntemiyle Sınıflandırma	41
5.3.2.2. Doc2Vec-DBoW Öznitelik Çıkarım Yöntemiyle Sınıflandırma	46
5.3.2.3. Doc2Vec-DM Öznitelik Çıkarım Yöntemiyle Sınıflandırma	51
BÖLÜM 6.....	57
SONUÇ VE DEĞERLENDİRME	57
KAYNAKÇA	61
ÖZGEÇMİŞ.....	69

KISALTMALAR

BoW : Bag of Words – Kelime Çantası

TF : Terim Frekansı

IDF : Ters Döküman Frekansı

CBoW : Sürekli Kelime Torbası

DBow : Dağıtık Kelime Torbası

DM : Dağıtık Hafıza

SVM : Destek Vektör Makinaları

NB : Naïve Bayes

MLP : Çok Katmanlı Yapay Sinir Ağı

A : Doğruluk

P : Keskinlik

R : Duyarlılık

TABLO LİSTESİ

Tablo 5.1. SVM Karmaşıklık Matrisi.....	34
Tablo 5.2. NB Karmaşıklık Matrisi.....	34
Tablo 5.3. MLP Karmaşıklık Matrisi	35
Tablo 5.4. BoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	35
Tablo 5.5. SVM Karmaşıklık Matrisi.....	36
Tablo 5.6. NB Karmaşıklık Matrisi.....	36
Tablo 5.7. MLP Karmaşıklık Matrisi	37
Tablo 5.8. DBoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	37
Tablo 5.9. SVM Karmaşıklık Matrisi.....	38
Tablo 5.10. NB Karmaşıklık Matrisi.....	39
Tablo 5.11. MLP Karmaşıklık Matrisi	39
Tablo 5.12. DM Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	40
Tablo 5.13. 10 Kat Çapraz Doğrulamada SVM Sınıflandırma Yönteminin Karmaşıklık Matrisleri.....	42
Tablo 5.14. 10 Kat Çapraz Doğrulamada NB Sınıflandırma Yönteminin Karmaşıklık Matrisleri	43
Tablo 5.15. 10 Kat Çapraz Doğrulamada MLP Sınıflandırma Yönteminin Karmaşıklık Matrisleri.....	44
Tablo 5.16. 10 Kat Çapraz Doğrulamanın Sınıflandırma Yöntemlerine Göre Her Parçası İçin Doğruluk Değeri	45
Tablo 5.17. BoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	45
Tablo 5.18. 10 kat çapraz doğrulamada SVM sınıflandırma yönteminin Karmaşıklık Matrisleri	47
Tablo 5.19. 10 kat çapraz doğrulamada NB sınıflandırma yönteminin Karmaşıklık Matrisleri	48

Tablo 5.20. 10 kat çapraz doğrulamada MLP sınıflandırma yönteminin Karmaşıklık Matrisleri.....	49
Tablo 5.21. 10 kat çapraz doğrulamada sınıflandırma yöntemlerine göre her parçası için doğruluk değeri.....	50
Tablo 5.22. DBoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	50
Tablo 5.23. 10 kat çapraz doğrulamada SVM sınıflandırma yönteminin Karmaşıklık Matrisleri.....	52
Tablo 5.24. 10 kat çapraz doğrulamada NB sınıflandırma yönteminin Karmaşıklık Matrisleri.....	53
Tablo 5.25. 10 kat çapraz doğrulamada MLP sınıflandırma yönteminin Karmaşıklık Matrisleri.....	54
Tablo 5.26. 10 kat çapraz doğrulamada sınıflandırma yöntemlerine göre her parçası için doğruluk değeri.....	55
Tablo 5.27. DM Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri.....	55
Tablo 5.28. SVM sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri	57
Tablo 5.29. NB sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri ...	58
Tablo 5.30. MLP sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri.	58
Tablo 5.31. SVM sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri	58
Tablo 5.32. NB sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri ...	59
Tablo 5.33. MLP sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri.	59
Tablo 5.34. Modelleme ve Sınıflandırma Yöntemlerinin Karşılaştırılması	60
Tablo 5.35. Modelleme ve Sınıflandırma Yöntemlerinin Karşılaştırılması	60

ŞEKİL LİSTESİ

Şekil 3.1. Organizasyon yapısını belirleyen unsurlar	12
Şekil 4.1. word2Vec örnek çalışma akışı.....	19
Şekil 4.2. Word2Vec vektör uzayı.....	20
Şekil 4.3. Word2Vec Model Yapısı (Mikolov vd. 2013)	20
Şekil 4.4. Skip-Gram çalışma yapısı (Büyükkınacı, 2018).....	21
Şekil 4.5. DBoW mimarisi.....	22
Şekil 4.7. Hiper Düzlemlerin belirlenmesi	24
Şekil 4.8. Optimum hiper düzlemde destek vektörleri	24
Şekil 4.9. Çok katmanlı yapay sinir ağ modeli	26
Şekil 4.10. Cross Validation (Temel vd. 2012)	29
Şekil 5.1. Detaysoft müşteri talep oluşturma ekranı	32
Şekil 5.2. Detaysoft müşteri taleplerinin listesi	32
Şekil 5.3. 10-kat çapraz doğrulama	41

ÖZET

Hızla gelişmekte olan teknolojik gelişmeler ve küreselleşme hem tüketicileri hem de üretim ve hizmet sektörlerini büyük oranda etkilemektedir. Küreselleşme sayesinde firmalar için yeni pazar alanları ortaya çıkarken, müşterilerin daha fazla seçeneğe ulaşmalarına olanak sağlamıştır. Teknolojik gelişmeler sayesinde müşterilerin benzer ürün veya hizmete kolay erişim sağlayabilmesi ve küreselleşme ile pazara daha fazla firmanın dahil olması firmalar arasında zorlu bir rekabete sebep olmuştur. Bu zorlu rekabetten ötürü firmaların varlığını sürdürebilmeleri hizmet sundukları müşteri sayısına bağlı hale gelmiştir. Mevcut müşterilerini korumak ve bünyelerine yeni müşteriler katabilmek için firmalar müşteri odaklı yenilikler ve düzenlemeler yapmaya yönelmiştir. Müşteri odaklılığını ön planda tutan firmalar için hizmet sonrası müşteri etkileşimi, müşteri memnuniyeti ve müşteri sadakati büyük önem taşımaktadır. Mevcut müşterilerini, sadık müşterilere dönüştürmek ve bu sadakati devam ettirebilmek için firmalar müşterilerini daha yakından tanıyarak onların talep ve ihtiyaçlarına tatmin edici dönüşler yapmalıdır. Gerek hizmet esnasında gerekse hizmet sonrası ortaya çıkan talep ve ihtiyaçların yönetimi için firmalar yardım masası uygulamaları kullanmaktadır.

Yardım masası uygulamaları, firmaların sunmuş oldukları hizmetler ile ilgili bilgi ve destek sağlamayı amaçlayan sistemlerdir. Müşteri ve firma personeli arasında merkezi koordinasyon sağlar. Yardım masası uygulamaları kullanılarak müşteri şikayet ve talepleri toplanmaktadır. Elde edilen bu şikayet ve talepler firmaların kendi eksikliklerini tespit edebilmesine olanak sağlamaktadır. Koordinasyonu doğru oluşturulmamış yardım masası uygulamaları gelen talep ve şikayetleri ele alırken zaman kayıpları ile karşılaşmaktadır. Bu çalışmada, müşteri taleplerinin ilgili birime aktarılması sürecinde ortaya çıkan verimsizliği minimum seviye indirgeyerek otomatik sınıflandırma yapan bir model önerilmiştir.

Anahtar Kelimeler: Müşteri Talepleri, Todo, Yardım Masası

ABSTRACT

Rapidly developing technological developments and globalization affect both consumers and production and service sectors to a great extent. Thanks to globalization, while new marketplaces have emerged for companies, enabling them to reach more options for consumers. Thanks to technological advances, customers have easy access to similar products or services, and globalization and the inclusion of more companies in the market have led to tough competition among firms. Because of this challenging competition, the sustainability of the companies depends on the number of customers they serve. In order to protect their existing customers and to add new customers to their structure, companies have focused on making customer-oriented innovations and arrangements. After-service customer interaction, customer satisfaction and customer loyalty are of great importance for the companies that prioritize customer focus. In order to turn existing customers into loyal customers and to maintain this loyalty, companies should make a better acquaintance with their customers and make satisfactory returns to their demands and needs. Help desk applications are used for the management of the demands and needs arising during and after the service.

Help desk applications are systems that aim to provide information and support about the services that companies provide. Provides central coordination between customer and company personnel. Using help desk applications, customer complaints and requests are collected. These complaints and demands enable companies to identify their own deficiencies. Loss of time occurs when handling demands and complaints from help desk applications that are not properly coordinated. In this study, it proposes a more automatic structure and alternatives by reducing time losses to a minimum level.

Keywords: Customer demands, Todo, Helpdesk.

BÖLÜM 1

GİRİŞ

Hızla gelişmekte olan teknoloji ve küresel rekabet firmaların mevcut konumlarını korumalarını ve büyümelerini zorlaştırmaktadır. Durumlar karşısında hızlı ve etkili çözümler sunamayan firmaların yaşama şansları azalmaktadır. Küreselleşme, firmalar için sadece tehdit değil, aynı zamanda onlara yeni pazar fırsatları da sunmuştur. Fakat bu durum aynı zamanda kendi pazarını da başkalarının yeni pazar listesine katmıştır (Ünver 2010). Küreselleşme sadece firmaları değil aynı zamanda müşteri profillerinde de değişikliklere neden olmuştur. Bilgi teknolojileri sayesinde müşterilerin bilgi düzeyi artmış ve bilinçlenmiştir, böylece müşteriler, daha seçici olmuş, talep edilen ürün ya da hizmetin ucuz maliyete ama yüksek kaliteye sahip olmasını bekler hale gelmişlerdir (Ünver 2010; Odabaşı 2000). Bir firmanın varlığını sürdürebilmesi hizmet sunduğu müşteri sayısına bağlı olmaktadır. Bu durum firmalar için müşterinin önemini arttırmakta ve müşterilerin sürekli değişen talep ve ihtiyaçlarını karşılayıp, yeni stratejiler geliştirmelerini gerektirmektedir (Cengiz 2013).

Gelişen teknoloji müşteriler için ürün çeşitliliği sağlayarak alternatif seçenekler sunmaktadır. Bir müşteri talep ettiği ürün veya hizmeti birden fazla firmada bulabilmekte ve bunlar arasında karşılaştırmalar yapabilmektedir. Bu durum firmalar için zorlu bir rekabet ortamının oluşmasına neden olmuştur. Zorlaşan rekabet şartları içerisinde varoluşunu sürdürmek isteyen firmalar, müşterileriyle olan ilişkilerini olumlu yönde geliştirdiklerinde rakiplerine karşı avantaj elde edip üstün konuma geçtiklerini görmüşlerdir (Çetiner 2012).

Rakiplerine karşı rekabet avantajı elde ederek maksimum karlılığı yakalamayı hedefleyen firmalar, müşteri beklentilerini doğru algılayarak bunu hizmetlerine yansıtmalıdırlar. Müşteri beklentilerini algılayabilmek için müşteri bilgilerinin ve davranışlarının analiz edilmesi gerekmektedir. Bu sayede müşteri beklentilerinin karşılanabilmesi ve dolayısıyla müşteri memnuniyetinin sağlanabilmesi mümkün olacaktır (Ersöz vd. 2008).

Mevcut müşteri sayısını arttırmak müşteri memnuniyetine bağlıdır. Çünkü memnun müşteri her zaman beraberinde yeni müşteriler de getirmektedir. Müşteri etkileşimi ile müşteri hakkında elde edilen bilgilerin analiz edilmesi ve müşteri ile tekrar görüşmelerinde bunu göz önünde bulundurarak kararlar verilmesi müşteri memnuniyetini kazanmanın temelini oluşturmaktadır (Sönmez 2007). Sonuç olarak müşteri memnuniyetini üst seviyelere çıkarabilmek için, müşteri beklentilerinin karşılanması büyük önem taşır (Alabay 2012). Fakat bazen müşteri beklentilerini bilmek onlar talep etmedikçe zordur. Firmalar tarafından, yardım masası gibi müşterilerin talep ve sorunlarını aktarabilecekleri iletişim araçları hizmete sunulmalıdır. Bu araçlar sayesinde sunulan ürün veya hizmetin sürekli iyileştirilmesi firmalar için bir fırsattır. Firmaların bu iyileştirmelerde göstermiş oldukları çabalar, müşterilerin hizmetten aldıkları tatmini ve firmaya olan bağlılıklarını olumlu yönde etkilemektedir (Alabay 2012).

Müşteri memnuniyetinin sağlanması ve artırılması müşteri odaklılıkla mümkün olmaktadır. Müşteri odaklılık, müşteri memnuniyetinin ve sadakatının sağlanmasını, sunulan ürün veya hizmette iyileştirmeler gerçekleştirilmesini ve etkin bir yardım masası yönetimini gerektirmektedir. Müşteri sadakati, müşteri taleplerinin karşılandığı hizmetin verilmesi, müşteri memnuniyetinin sağlanması, hizmet öncesi ve sonrası tüm ihtiyaçların karşılanması ile mümkün olabilir (Alabay 2012).

Müşteri memnuniyeti, müşteri sadakatini kazanmada büyük önem taşımaktadır. Firmalar müşterilerinin sadakatini kaliteli hizmet sunarak kazanmaya çalışırlar (Kotler, Armstrong 2010). Firmaya yeni müşteri kazandırmak, eski müşteriyi elde tutmaktan 5,6 kat daha masraflıdır (Blodgett vd. 1995; Wilson vd. 2012). Bundan dolayı varlığını devam ettirmek ve kazancını arttırmak isteyen firmalar mevcut müşterilerini ellerinde tutmak için stratejiler geliştirmeli ve müşterilerini sadık müşterilere dönüştürmelidirler. Firmaların müşterilerinin %5'ini ellerinde tutarak karlılıklarını %25'ten %125'e kadar çıkarabildikleri görülmüştür (Aksu 2004).

Firmaların kazanç elde edebilmesi müşterilerinin beklentilerini karşılamalarına ve onları memnun etmelerine bağlıdır (Alabay 2008). Beklentisi istediği gibi karşılanmış olan müşteri, ortalama memnun bir müşteriden daha fazla

sadıktır. Müşteri memnuniyeti beraberinde yeni müşteriler getirirken, memnuniyetsiz bir müşteri beraberinde daha fazla kişiyi olumsuz etkileyebilmektedir. Memnun olmayan müşterilerin sadece %4'ü bunu firmalarına bildirmektedir (Alabay 2012).

Müşteri talepleri, firmaların sunmuş olduğu ürün ve hizmet üzerinde iyileştirmeler yapmasına imkân sunar. Bu taleplerin doğru yönetilmesi (Barış 2008);

- Sunulan ürün veya hizmeti daha kaliteli bir hale getirir,
- Sadık müşteri sayısını artırır,
- Müşterilerin başkalarına olumlu mesajlar iletmesini sağlar,
- Firmaya yeni müşteriler kazandırır,
- Müşterinin zihnindeki kalite algısını artırır,
- Firmanın toplam gelirini artırır.

Araştırmanın Amacı

Gelişen bilişim teknolojileriyle birlikte müşterilerin aynı ürüne veya hizmete ulaşımının sağlanması firmalar arası rekabet artışına sebep olmuştur. Artan rekabetten dolayı firmaların varlıklarını sürdürmeleri ellerinde tutabildikleri müşterilere bağlıdır. Müşterilerinin memnuniyetini sağlayamayan firmalar müşterilerini rakiplerine kaptırarak yok olma düzeyine gelecektir. Bu yüzden firmalar için müşteri memnuniyetinin sağlanması ve müşterilerini birer sadık müşteriye dönüştürmeleri büyük önem taşımaktadır. Özellikle hizmet sektöründe yer alan kurumsal firmalar için hizmet sonrası müşteri etkileşimi, müşteri memnuniyeti ve sadakati açısından çok önemlidir. Bunu öncelik haline getirerek stratejilerinde düzenlemeye giden firmaların müşteri etkileşimi sağlamak için yardım masası adı altında birimleri veya sistemleri mevcuttur. Yardım masaları, müşterilerin karşılaştıkları aksaklık ve sorunları bildirmek, yeni bir ihtiyaç dâhilinde taleplerini dile getirmek ve bilgi almak gibi müşteri-firma etkileşimlerinde tercih edilmektedir.

Bu çalışma ile özel bir firmanın yardım masasına gönderilen istek, talep ve destek kayıtlarından oluşan veri seti üzerinde makina öğrenmesi/sınıflandırma yöntemleri ve metin madenciliği teknikleri kullanılarak alınan taleplerin organizasyonel iş akışlarına uygun olarak ilgili birime yönlendirilmesi amaçlanmıştır.

Araştırmanın Önemi

Hızla artan küresel rekabet ortamında varlığını sürdürmek ve kar elde etmek isteyen firmalar, mevcut stratejilerinde sürekli iyileştirmeler yapmalı ve hizmet sunduğu müşterilerinin memnuniyetlerini ön planda tutmalıdır. Memnun müşteri beraberinde 4 ya da 5 müşteriyi getirirken, memnun olmayan müşteri yanında 9 ya da 10 kişi götürebilmektedir.(Alabay 2012) Bundan dolayı mevcut müşterilerini korumanın yanında yeni müşteriler kazanmak için müşteri ile etkileşim halinde olunmalı ve müşteri taleplerine ikna edici dönüşlerde bulunulmalıdır. Müşteri taleplerine etkili ve ikna edici bir dönüşte bulunmanın ilk adımı talebinin doğru birime yönlendirilmesidir. Müşteri taleplerinin, uygun birime aktarılması durumunda, müşterinin doğru kişiyle etkileşime geçmesi sağlanacaktır. Bu durum ilgili talebin hızlı bir şekilde, daha net anlaşılmasına ve istenilen ölçüde karşılanmasına olanak sunacaktır. Bu sayede taleplere hızlı dönüşler yapılacak, doğru ekiplere sorun daha kısa sürede yönlendirildiği için çözümleme süresi kısalmaya ve müşteri memnuniyeti artırılarak müşteri kayıpları azaltılacaktır. Tüm bunlar firmaların karlılık oranlarına da olumlu katkılar sunacaktır.

BÖLÜM 2

BENZER ÇALIŞMALAR

Müşteri ilişkileri ve müşteri talep-şikâyetlerinin değerlendirilmesi üzerine yapılan çalışmalar incelendiğinde, daha çok talep ve şikâyetlerin nasıl toplanacağı, nasıl değerlendirileceği, müşteri ile iletişime geçerken nelere dikkat edilmesi gerektiği ve müşteri memnuniyetinin sağlanması için nelerin yapılması gerektiğine ilişkin konular üzerine yoğunlaştığı gözlemlenmiştir.

Chang vd. yaptıkları çalışmalarında internet ve ticaret ortamında etkili bir pazar stratejisi ve müşteri profili oluşturmak için, veri ambarı ve veri madenciliği tekniklerini kullanan bir büyüme modeli oluşturmayı amaçlamışlardır. Çalışmalarında modelin uygulanabilirliği ve etkinliği üzerinde durmuşlardır (Chang vd. 2009).

Ittoo vd. online alışverişlerde müşteriye doğru tercihleri önermek için müşteri karar verme sürecine yardımcı olacak metin tabanlı bir sistem önermektedirler. Önerilen bu sistem, ürün özelliklerini öğrenmek için metin madenciliği tekniklerini kullanarak, müşterilerin tercihlerine uygun ürünler sunmaktadır (Ittoo vd. 2006).

Thorleuchter vd. çalışmalarında e-ticaret şirketlerinin web siteleri üzerindeki metinsel bilgilerin ticari başarıları üzerine etkilerini analiz etmişlerdir. Dünya çapında en başarılı 100 e-ticaret şirketi üzerinde çalışarak, ticari başarısını tahmin etmede başarılı olan bir regresyon modeli oluşturmuşlardır (Van Den Poel, Thorleuchter 2012).

He vd. en büyük 3 pizza zincirinin sosyal medya üzerindeki yapılandırılmamış metinlerini analiz etmek için metin madenciliğini derinlemesine uygulayan bir vaka çalışması yapmışlardır. Şirketlerin sosyal medya analiz stratejilerini geliştirmelerine yardımcı olmak için önerilerde bulunmuşlardır (He vd. 2013).

Özyirmidokuz, popüler internet alışveriş firmalarını analiz etmek için metin madenciliği tekniğini kullanmıştır. Çalışmada 200 Türk firmasının web siteleri, doğal dil işleme ve web metin madenciliğiyle incelenmiş, benzerlik ve kümeleme

çalışmaları yapılmıştır. Çalışma sonucunda elde edilen kümeleme ile firmaların ilişkileri ve benzerlikleri tespit edilmiştir (Özyirmidokuz 2016).

Gamon, müşteri geri bildirimlerini çok gürültülü alanlarda bile otomatik olarak sınıflandırmanın mümkün olduğunu göstermiştir. Bir insanın bile sınıflandırmada zorlanacağı veriler üzerinde yüksek başarı elde etmede doğrusal vektör makinelerini ve doğal dil işleme tekniklerini kullanmıştır (Gamon 2004).

Özyirmidokuz vd. Türkiye'deki en büyük yedi ısıtma sistemi firmasının müşteri şikâyetlerini analiz etmek için web metin madenciliği tekniğini kullanmıştır. Çalışma sonucunda firmaların müşteri şikâyetleri açısından benzerliklerini belirlemiştir (Özyirmidokuz vd. 2014).

Stoica vd. Türkiye'deki bir telekomünikasyon firmasına bildirilen yapılandırılmamış müşteri geri bildirim belgelerinden alınan verileri özetlemek ve çıkarımda bulunmak için metin işleme tekniklerini ve kosinüs benzerliğini kullanmışlardır (Stoica, Özyirmidokuz 2015).

Caputo vd. Portekizce dilinde olan çağrı merkezi veri tabanlarında, müşterilerin taleplerine dayanarak yeni ve yararlı bilgileri tespit edebilmeyi hedeflemişlerdir. Bu sayede müşteriye yardımcı olmak ve müşteri memnuniyetini arttırmada firmanın eksikliklerini tamamlaması için bilgi çıkarımı sağlayan, metin madenciliği ve kümeleme yöntemlerinin kullanıldığı bir çalışma sunmuşlardır (Caputo vd. 2006).

Ordenes vd. müşteri geri bildirim verilerini analiz etmede metin madenciliği tekniğinin kullanıldığı dil bilimsel tabanlı bir yaklaşım sunmuşlardır. Çalışmalarında önerilen metin madenciliği modelinin, yüksek doğruluk oranına sahip olduğunu ve eğitim yoluyla değişen kapsamları hesaba katarak esneklik sağladığını ifade etmişlerdir (Ordenes 2014).

Filip, müşteri şikâyetlerinin firmalar için ne kadar önemli olduğu konusu üzerinde durmuştur. Müşteri memnuniyetsizliği hem müşteri kayıplarına neden olmakta hem de kurumun imajını olumsuz etkileyebilmektedir. Müşterilerin istenmeyen durumlarla karşılaşması halinde şikâyette bulunmalarını ve kurumun müşteriyi tatmin edecek şekilde sorunlarını gidermeleri üzerine imkân ve süreçlere sahip olması gerektiğini ifade etmiştir (Filip 2013).

Alabay, hazırladığı tez çalışmasında, kurumlar ve müşteriler için müşteri şikâyetlerinin önemi üzerinde durmuştur. Çalışmada müşteri şikâyetlerinin hangi kanallarla alabileceği ve müşterilerle iletişim kurarken nelere dikkat edilmesi gerektiği konuları derinlemesine incelenmiştir. Bu çalışmaya göre terk eden bir müşteriyi geri getirmek, yeni müşteri elde etmek için harcanan maliyetin 6 katına denk gelmektedir (Alabay 2012).

Coussement vd. gelen maillerin şikâyet olup olmadığını dil yapısı özelliklerini dikkate alarak sınıflandıran bir yöntem tanımlamışlardır. Çalışmalarında TF-IDF ve doğal dil işleme tekniklerinden faydalanmışlardır (Coussement, Van Den Poel 2008).

Weng vd. müşteri hizmetleri personelinin e-postalarla gelen sorulara yanıt vermedeki yükünü azaltmak ve e-posta yanıtlarına uygun olan şablonları sunan bir sistem önermişlerdir. Çalışmalarında metin sınıflandırma ve benzer doküman vektörü tekniğini kullanarak uygun yanıt şablonları sunmuşlardır (Weng, Liu 2004).

Gonzalez vd. bilgi yönetim merkezli yardım masası olarak adlandırılan bir yaklaşım önermektedirler. Temsilci merkezli yardım masası ile bilgi yönetim merkezli yardım masasının karşılaştırılması için bir simülasyon oluşturmuşlardır. Çalışmalarında bir takım iyileştirmelerde bulunmuş ve sonucunda problemin çözüm sürelerinde önemli ölçüde azalmalar meydana geldiğini ve yardım masasının çıktılarının geliştiğini göstermişlerdir (Gonzalez vd. 2005).

Tez kapsamına giren literatür araştırmasında ele alınan çalışmalarda da görüldüğü üzere müşteri geri bildirimlerinin, talep ve şikâyetlerinin müşteri odaklı firmalar için önemi büyüktür. Toplanan müşteri talep ve şikâyetlerinin zamanında değerlendirilmesi ve geri dönüşler sağlanması başarılı firmalar için bir ön koşul olmaktadır. Bu durum müşteri memnuniyetini ve bağlılığını beraberinde getirmektedir.

BÖLÜM 3

YARDIM MASASI

3.1. Müşteri Kavramının Tanımı

Teknoloji alanındaki gelişmelerden dolayı zorlaşan rekabet ortamında güçlü ve daimi olmayı hedefleyen firmalar müşteri ilişkilerini geliştirdiklerinde rakiplerine karşı avantaj elde ettiklerini fark etmişlerdir. Bu değişikliklerin hızla devam etmesi müşterilerin firmalara daha kısa sürede erişim sağlamalarının yanında, ihtiyaçlarının giderilmesinin yetersiz olduğunu hissettiklerinde alternatif firmalara erişimini getirmektedir (Türker, Özaltın 2010). Bu durum, firmalar açısından müşterilerini elde tutmanın oldukça zor bir hal aldığı göstermektedir. Firmaların varlıklarını sürdürebilmelerinin temelinde “müşteri” yatmaktadır. Firmalar için bu kadar önemli bir faktör haline gelen “müşteri, kimdir ve nedir?” sorusuna cevap aranmalıdır.

Müşteri, ihtiyaç ve gereksinimlerini karşılamak için ürün ve hizmet talep eden kişidir yani son kullanıcıdır. Varlıklarını devam ettirmeyi düşünen firmalar için müşteri önem verilmesi gereken bir kavramdır. Bir firmadan ürün ya da hizmet satın alan, bu ürün ve hizmetin pazardaki yerini belirleyen ve firmaların hedeflerine ulaşmasına yardımcı olan kişi ya da kuruluştur (Taşkın 2000; Çetiner 2012). Firmaların devamlılığını sürdürmeleri ellerinde tutabildikleri müşteri sayısına bağlıdır. Müşterilerini kaybetmemek için firmalar, müşterilerinin talep ve ihtiyaçlarını tatmin edici ölçüde karşılayarak müşteri memnuniyetini oluşturmak zorundadırlar. Memnun müşteri, beraberinde yeni müşteriler de getirebilmektedir. Bu da firmaların müşteri sayısını artmasını ve tanıtım maliyetlerini azalmasını sağlamaktadır (Alabay 2007).

Acuner (2003)' e göre “Bir firmanın müşterisi kimdir?” sorusuna verilen cevaplar aşağıdaki gibidir (Acuner 2003);

- Müşteri, firmanın ürün ve hizmetlerinin son kullanıcıdır.
- Müşteri, firmadaki en önemli kişidir.
- Müşteri, karşılanmayı bekleyen istek ve ihtiyaçları olan kişidir. İşletmenin görevi de hem kendisine hem de müşterilerine fayda sağlayacak şekilde hizmet etmektir.

- Müşteri bir istatistiksel veri değildir. Duyguları olan ve kendisine saygıyla davranılmasını hak eden kişidir.
- Müşteri işimiz için bir araç değil işimizin amacıdır. Hizmet vererek biz ona değil, bize böyle bir fırsat verdiği için, o bize iyilikte bulunur.

Müşteri kavramı sadece firma dışında yer alan nihai tüketicilerden oluşmamaktadır. Aynı zamanda firma içerisinde de müşteri konumunda yer alan firma çalışanlarını da kapsamaktadır. Bu da bize müşteri kavramını iki aşamada incelememizi gerektirmektedir.

3.1.1. İç Müşteri

İç müşteri, firma çalışanlarından oluşmaktadır. Kendi içerisinde bölümlere ayrılmış firmalarda, bir bölümün çıkartmış olduğu ürün ve hizmetin başka bir bölümün çalışanlarına girdi olarak temin edilmesi ya da onların kullanımına sunulması iç müşteri kavramını ortaya çıkarmıştır. Nihai tüketiciye yani dış müşteriye ürün ya da hizmeti sunmadan önce, ürün ya da hizmetin müşteriler için tatmin edici ölçüde olup olmadığının değerlendirmesi iç müşterilerin yapmış olduğu incelemelere bağlıdır (Akdın 2006). Sadece dış müşteri odaklı bir firma asla başarılı olamaz. İç müşteriyi memnun edemeyen bir firma dış müşteriyi de memnun edemez. Bu nedenle üretilen ürün ya da hizmetten öncelikle iç müşterinin memnuniyeti büyük önem taşımaktadır. Çünkü iç müşteriler bir ürün ya da hizmetin üretim süreci etkinliğinin iyileştirilmesine destek olabilmektedir. İç müşterilerden olumlu ya da olumsuz dönüşlerin alımı ve bunlar üzerinde öneri ve taleplerinin alınması dış müşteriye oranla daha kolaydır. Rakiplerine karşı rekabet üstünlüğü sağlamak isteyen firmalar iç müşterilerinin kullanım etkinliğini arttırdıklarında, dış müşteriye sunulacak olan ürün ya da hizmetin kalitesini de arttırmış olacaktır (Taşpınar 2005).

3.1.2. Dış Müşteri

Dış müşteri, firmaların üretmiş olduğu ürün ya da hizmeti kullanacak olan son kullanıcıdır. Dış müşteri, ürün ya da hizmetin nerede ve nasıl üretildiğiyle ilgilenmekten daha çok ürün ve hizmetin nasıl ve ne kadar sürede eline geçtiğiyle, ürün ve hizmetin belirtilen özelliklere sahip olup olmadığıyla ve ihtiyaçlarını karşılayıp karşılamadığıyla ilgilenir (Karakaya 2007)

3.2. İş / Görev Yönetimi

Küresel rekabet ortamında firmalar rakiplerine oranla en iyi olma çabası içerisinde. Bunu sağlamak çoğunlukla yüksek kalite, düşük maliyet ve müşteri ihtiyaçlarına tatmin edici karşılıklar sunmaya bağlıdır. Bunun için başarılı olmayı hedefleyen firmalar çalışanlarından, yüksek performans sergilemelerini, müşteri beklentilerini karşılamalarını ve rakiplerine karşı avantaj elde etmede katkıda bulunmalarını beklemektedirler (Biçer, Düztepe 2003). Verimli çalışma prensibini benimsemiş olan firmalar, çalışanlarının performanslarını ölçmek, değerlendirmek ve arttırmak için etkin iş planlaması ve görev takip sistemlerine ihtiyaç duymaktadırlar.

Büyük projelerin gerçekleştirilmesinde projenin daha verimli ve yönetilebilir olması için etkin iş planlaması kullanılmaktadır. Etkin iş planlaması sonrasında görev takibi gelmektedir. Proje gerçekleştirilmesinde çalışanlara atanan görevlerin bitip bitmediklerini veya hangi aşamalarda olduklarının takibi görev yönetiminden incelenebilmektedir. Görev yönetimi her bir çalışanın gerçekleştirdiği işle ilgili aktivitelerin izlenmesine ve her bir işin hangi çalışana atandığının tutulmasına izin verir (Whitley vd 2014). Görev yönetimi, bir çalışanın belirli zaman aralığında gerçekleştirmiş olduğu ya da gerçekleştireceği aktivitelerin yönetilmesine ve raporlanmasına olanak sağlar. İş yönetimi sürecinde önemli çıktılar sunan görev yönetim sistemleri, firma çalışanlarının ne tür görevler yaptıklarının takibini sağlayarak çalışanların verimlilik artışına önemli katkılar sağlayabilir.

Bu çalışmada kullanılmakta olan veri seti, Detaysoft firmasının kendi bünyesinde geliştirmiş olduğu görev yönetim sistemi (BNet To-Do Sistemi) üzerinden toplanmıştır. Geliştirilmiş olan bu görev yönetim sistemi, her projenin başından sonuna kadar gerçekleştirilen görevlerin hangi aşamalarla tamamlandığını takip imkânı sunmaktadır. Ayrıca bu uygulama üzerinde, proje tamamlandıktan sonra bakım/destek sürecinde son kullanıcıdan sorunların ve taleplerin toplanması ile yeni geliştirme isteklerinin alınması gibi süreçler yönetilmektedir. Müşteri taleplerinin toplandığı uygulama, firma çalışanlarının kullandığı To-Do uygulaması ile entegre bir yardım masası uygulamasıdır.

3.3. Organizasyon Yapısı

Çok sayıda çalışana sahip firmalar, kendi çalışanlarının uyumlu bir şekilde çalışmalarını yürütebilmesi ve bunların takip edilebilmesi için iç kontrol sistemlerinin kullanımını zorunlu kılmıştır. İç kontrol sisteminde görev ayrılık ilkesi büyük önem taşır. Görev ayrılığı ilkesine göre bir işlemi yapan, atayan, uygulayan kişiler farklı olmalıdır. Bir firmanın gerçekleştireceği tüm faaliyetlerin görev ayrılığı ilkesi gereği birbirinden ayrılması gerekmektedir (Dabbağoğlu 2009). Bu görev ayrılığı firma çalışanlarının organizasyonel bir yapı oluşturmaya neden olmuştur. Organizasyon, belirli bir amaca ulaşmada yapılması gereken görevlerin tanımlanması ve bu görevleri gerçekleştirecek kişilerin atanması yoluyla ortaya çıkan yapısal süreçtir (Şimşek, Çelik 2002). Organizasyon yapısı ise firmaların amaçları doğrultusunda organizasyonun içinde oluşan yetki ve sorumlulukların tanımlandığı, karşılıklı ilişkilerin oluşturulduğu ve bilgi, emir, karar ve rapor akışının gerçekleştirildiği haberleşme kanallarını kapsayan bir yapıdır (Özgener 2003). Firmaların çalışanlarından daha yüksek verim alabilmeleri için, iyi bir organizasyonel yapıya sahip olmaları gerekmektedir. Yanlış yapılanmalar firma performansına büyük zararlar verebilir (Demir, Tarhan 2011). Her firmanın kendine özgü bir organizasyon yapısı mevcuttur. Organizasyon yapıları kişilere göre değil, hedef, sistem ve süreç doğrultusunda oluşturulmalıdır. Organizasyon yapısı Şekil 3.1’de gösterildiği gibi iç ve dış faktörlerden etkilenebilmektedir.



Şekil 3.1. Organizasyon yapısını belirleyen unsurlar

Organizasyon yapıları çok sayıda personele sahip firmaların verimli bir şekilde yönetim ve iletişimlerinin sağlanmasında büyük rol oynamaktadır. Her firmanın kendine özgü bir organizasyon yapısı mevcuttur. Bir organizasyon yapısı oluşturulurken aşağıdaki unsurlar göz önünde bulundurulmalıdır.

Amaç Birliği: Her organizasyon ulaşılmak istenen amaçlara ve bu amaçlara ulaştıracak faaliyetlere göre tasarlanır. Firma içerisindeki bütün birimlerin kendi amaçları arasında bütünlük olmalı ve firmanın temel hedefine katkıda bulunacak nitelikte olmalıdır.

Yönetim Alanı: Kontrol alanı olarak da bilinen yönetim alanı, organizasyonda bir üste kaç ast bağlanacağı ile ilgili bir unsurdur. Bir yöneticiye bağlanacak olan ast sayısı ve üstleneceği sorumluluklar, yöneticinin bilgi, yetenek ve kavrama sınırlarını aşmamalıdır.

Yönetim birliği: Tek yöneticiye sahip bir grubun aynı amaca yöneltilmesidir. Yönetim birliği yönetim alanından farklıdır. Yönetim alanı tek kişiyi kapsarken yönetim birliği bir grubu kapsamaktadır.

Komuta birliği: Her astın bir üste bağlı olması ve sadece ondan emir alması olarak ifade edilir. Birden fazla üste bağlı olma durumunda yetiştirilemeyen faaliyetler, birbiriyle çelişebilecek görevlerle karşılaşma gibi sorunların ortaya çıkmasına ve faaliyet kargaşası oluşturmasına neden olabilir.

Yetki ve sorumluluk: Her çalışanın sorumlu olduğu iş ile sahip olduğu yetki ve sorumluluk açık bir şekilde belirtilmelidir. Aksi takdirde hem işlerin yürütülmesinde kargaşa yaşanacak, hem de yöneticiler arasında çatışmalar meydana gelecektir.

İletişim kanalları ve şekli: Organizasyonu oluşturan birimler arasındaki iletişim kanalları, işlerin zamanında ve doğru olarak yapılabilmesinde büyük rol oynar. Firmanın üst orta ve alt basamaklarında çalışanların kendi aralarında bir araya gelip görüşebilmesi ve ortaya çıkan sorunların çözümlenebilmesidir. İletişim noktasında, organizasyon yapısında yatay ve dikey iletişim kanalları olabilir. Yatay iletişim kanalı, aynı yönetim kademesindeki görev yapanlar arasındadır. Dikey iletişim kanalı ise üst ve astlar arasında gerçekleşir.

Personel ve kaynak dağılımında denge: Firma birimlerine gerekli miktarda çalışan sağlanmalı ve bu çalışanlar arasında nitelik yönünde denge olmalıdır. Çalışanlara sağlanan kaynaklarda, ihtiyaç önceliğine göre dağılımının dengelenmesi gerekmektedir.

Bölümlendirme: Organizasyondaki iş bölümü ve uzmanlaşma derecesine göre belirli iş ve görevler bir araya getirilerek organizasyonel birimler oluşturulur. Bir iş üzerinde uzmanlaşma dengesi sağlanırsa verimlilik elde edilir. Bir işte uzmanlaşmak için uzun süre aynı işe yoğunlaşmak gerekir. Eğer uzmanlık dengesi sağlanamazsa, bu durum monotonluk, stres ve verimsizlikle sonuçlanır. Bu nedenle firma içerisinde dengeli bir uzmanlaşma derecesi sağlanmalıdır.

3.4. Yardım Masası ve Yönetimi

Yardım masaları, son kullanıcı (müşteri) ile hizmet ekipleri arasındaki koordinasyon noktasıdır. Koordinasyon noktasının bulunmadığı organizasyonlarda yardım istemek ve sorunları çözümlenmek büyük zaman kayıplarına neden olmaktadır. Yardım Masası, gelen çağrılarını iş önceliklerine ve belirlenen servis seviyelerine göre çözümlenmesini sağlamakta, kullanıcı ve destek ekiplerinin verimliliğini artırmaktadır. Yardım masası kurumun ihtiyaçlarına ve kullandığı uygulamalara göre farklı şekillerde organize edilebilmektedir (Esener 2006).

BÖLÜM 4

YÖNTEM

4.1. Veri Setinin Toplanması

Detay Danışmanlık firması, müşteri etkileşimlerini optimum düzeyde gerçekleştirmelerini sağlamak ve onların talep ve isteklerini toplamak için bir yardım masası uygulaması geliştirmiştir. Bu yardım masası uygulaması danışmanlık hizmeti sundukları müşterilerin karşılaştıkları sorun ve talepleri bildirmelerini ve bu konuda etkileşim halinde olmalarını sağlamaktadır. Bu sistem aracılığıyla toplanan veri seti, ilgili talebi ve bu talebin hangi bölüm tarafından çözüme ulaştırıldığını içermektedir. Bu sistem üzerinde canlı/gerçek müşteriler tarafından oluşturulan görevler arasından daha önce başarıyla kapatılan görevler filtrelenmiştir. Bu görevlerden 886 tanesi veri seti olarak kullanılmak üzere ayrıştırılmıştır. Bu veri setinin;

- 305 Lojistik,
- 273 insan kaynakları,
- 308 finans bölümü tarafından çözümlenmiş görevleri içermektedir.

Oluşturulan bu veri seti tez çalışmasının amacı olan müşteri taleplerinin organizasyonel iş akışlarına uygun olarak sınıflandırılması sürecinde modelin eğitimi ve testi için kullanılmıştır.

4.2. Veri Ön İşleme

Veri seti olarak toplanan müşteri talepleri, doğal bir dil kullanılarak yazıldığı için belirli bir yapıya sahip değildir. Yapısal olmayan veri setleri ön işlem aşamalarından geçirilerek ham verilerin temizlenerek uygun bir formata dönüştürülmesi gerekmektedir (Feldman, Sanger 2007). Firmalara müşteri tarafından gönderilmekte olan talepler konu ve içerik şeklinde iki bölüm doldurularak oluşturulmaktadır. Daha verimli sonuçlar elde edilebilmesi için konu ve içerik kısmı bir bütün olarak ele alınmalıdır. Müşteri taleplerinin sınıflandırma işlemlerine tabi tutulmadan önce bazı ön işlem aşamalarından geçirilerek gereksiz veriler, ilgili birime aktarımını etkilememesi için temizlenmektedir. Bu ön işlem aşamaları (Karadağ, Takçı 2010);

- Simge ve Noktalama işaretlerinin temizlenmesi,
- Temizlenmiş metnin, token adı verilen kelime gruplarına ayrılması,
- Elde edilen bu kelime gruplarının Doğal Dil Araç Seti (Natural Language ToolKit - NLTK) ile köklerinin belirlenmesi,
- Belirlenen kelime köklerinin uzunluklarının hesaplanması, 3 ve daha az harf içeren kelimelerin veri setinden çıkartılması,
- Geriye kalan kelime listesinden “Durak Kelimeler” (stopwords) dediğimiz, cümle içinde kullanılan, tek başına bir anlam ifade etmeyen edat, bağlaç vs.’nin çıkartılması,
- Kelime listesindeki kelimelerin terim frekanslarının hesaplanması ve terim frekansı belirli aralıklarda olan kelimelerden “Kelimeler Çantasının” (BagOfWords - BoW) oluşturulması,
- Oluşturulan BagOfWords kullanılarak vektör uzayının oluşturulması, şeklinde sıralanmaktadır.

4.3. Öz Nitelik Çıkarımı

Öznitelik çıkarımı, müşteri taleplerini en iyi şekilde temsil eden vektör uzayının belirlenmesi ve her bir müşteri talebini temsil edecek vektörün oluşturulması için müşteri talep verileri içerisinde çeşitli teknikler kullanılarak anahtar kelimelerin belirlenmesi sürecidir. Bu tez çalışmasında literatürde yaygın olarak kullanılan 3 öznitelik çıkarım modelinden 2 tanesi üzerine yoğunlaşarak sonuçlar karşılaştırılmıştır.

4.3.1. Bag of Words (Bow) Modeli

Kelimeler Çantası (BoW), metin madenciliği uygulamalarında yaygın olarak kullanılan öznitelik çıkarım yöntemlerinden biridir. BoW modelindeki kelimeler vektör uzay modelini oluşturmakta ve her bir müşteri talebini bir vektör olarak temsil etmektedir.

Vektör uzay modeli, veri madenciliği ve metin madenciliği gibi bilgi çıkarımlarında kullanılan matematiksel bir modeldir. Ön işlem aşamalarından geçirilerek temizlenen terimlerden oluşan kelimeler çantasındaki her bir terim vektör uzayında bir eksene karşılık gelmektedir. Vektör uzay modeli sayesinde yapısal olmayan bu veri seti bir vektör haline dönüşerek yapısal hale getirilmiştir. Her bir

müşteri talebi vektör uzay modelinde bir vektöre karşılık gelmektedir. 3 farklı şekilde vektör oluşturulmaktadır (İlhan 2008).

4.3.1.1. Bitsel Tanımlama

Ön işlemde geçirilen yapısal olmayan veri seti kullanılarak bir kelime çantası oluşturulur. Bu yöntemde göre vektörün oluşturulması, kelimeler çantası içerisindeki terimlerin cümle içerisinde olup olmadığına göre şekillenmektedir (İlhan 2008).

Örnek: kelime çantası="müşteri, çek, iade, çıktı, sorun, fatura, al, form, sipariş, teslimat, blokaj" şeklinde tanımlandığında;

Talep 1: "Müşteri çek iadesinde sorun çıktı. Çek alımı sorunlu"

Vektör 1: 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0

Talep 2: "Fatura aldığımız müşteriye çek iade formu"

Vektör 2: 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0

Talep 3: "Sipariş teslimat blokaj sorunu"

Vektör 3: 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1

gibi talep içerisinde kaç defa geçtiğine bakılmaksızın sadece var yada yok şekilde 1 ve 0'lerden oluşan vektörler olarak tanımlanmaktadır.

4.3.1.2. Frekansa Göre Tanımlama

Oluşturulan kelime çantasındaki her bir terimin talep içerisindeki kullanım sıklığının hesaplanmasına terim frekansı denir (Karadağ, Takçı 2010, İlhan 2008). NLTK ile temizlenmiş örnek uygulama:

Örnek: kelime çantası = "müşteri, çek, iade, çıktı, sorun, fatura, al, form, sipariş, teslimat, blokaj" şeklinde olsun;

Talep 1: "Müşteri çek iadesinde sorun çıktı. Çek alımı sorunlu"

Vektör 1: 1, 2, 1, 1, 2, 0, 1, 0, 0, 0, 0

Talep 2: "Fatura aldığımız müşteriye çek iade formu"

Vektör 2: 0, 1,1, 0, 0, 1, 1, 1, 0, 0, 0

Talep 3: “Sipariş teslimat blokaj sorunu”

Vektör 3: 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1

4.3.1.3. TF_IDF Ağırlıklandırma Yöntemine göre Tanımlama

Terim Frekansı (Term Frequency - TF), bir terimin talep içerisindeki kullanım sıklığını belirler. Ters Doküman Frekansı (Inverse Document Frequency - IDF) ise bir terimin kaç farklı talep içerisinde geçtiğini belirler. Bu belirleme yönteminde orta değerler göz önüne alınmalıdır. Talepler içerisinde en çok kullanılan terim veya çok seyrek kullanılan terim o talepler için ayırt edici olmayabilir.

Bu nedenle ortalama bir terim ağırlığı belirlenerek vektör yapısı oluşturulmalıdır. Terim ağırlığı bu iki frekansın çarpımı sonucunda hesaplanmaktadır.

$$W_i = TF_i * IDF_i$$

$$IDF = \log(D/df_i)$$

df_i = terimin kaç farklı talepte geçtiğinin sayısı,

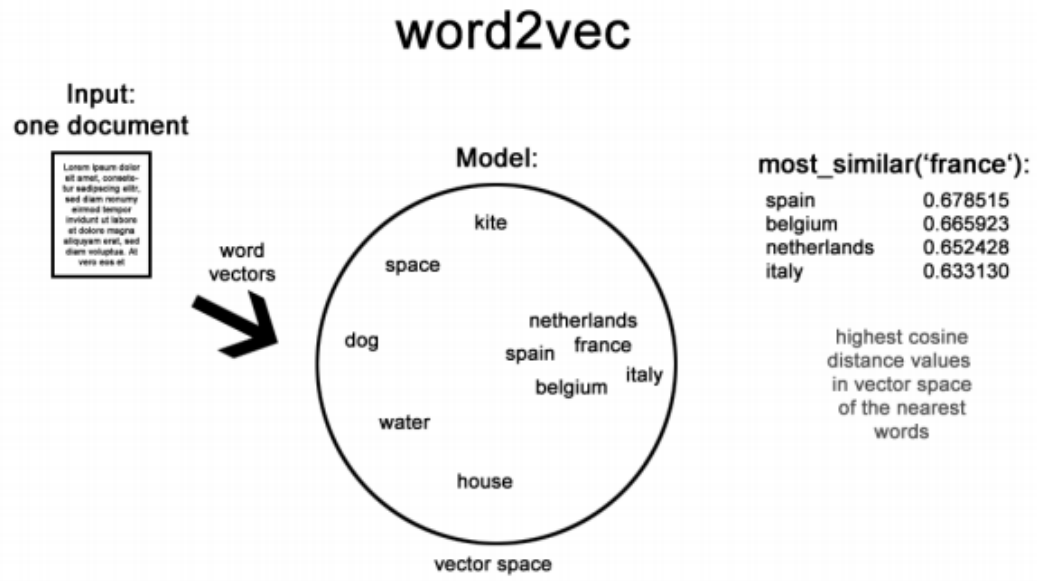
D = toplam talep sayısı

df_i değeri IDF ile ters orantılıdır, df_i değeri arttıkça IDF değeri azalacaktır. Yani IDF değeri “0” a yaklaşacaktır. “0” a eşit olması o terimin tüm dokümanlarda geçmiş olduğunu ifade etmektedir.

4.3.2. Word2Vec Modeli

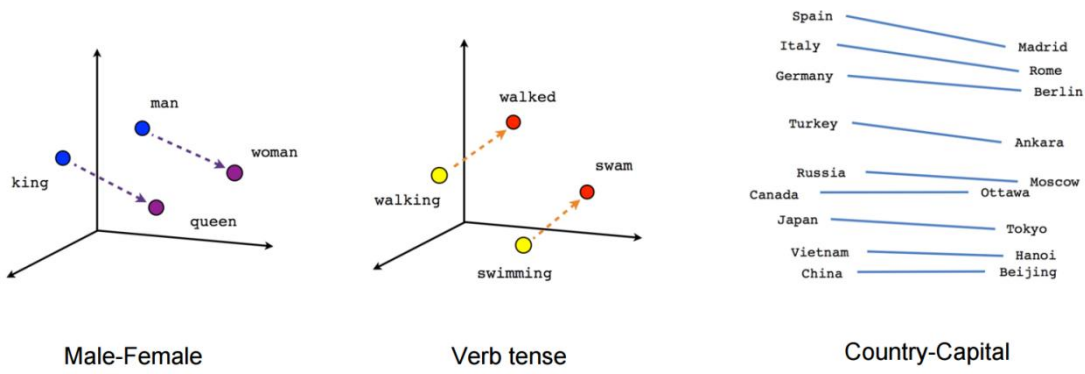
Metin madenciliği uygulamalarında kelimeleri temsil etmek için BoW modeli sıklıkla kullanılmaktadır. Bu modele göre vektör uzayı oluşturulurken kelimelere bağımsız değerler verilerek, kelimelerin birbirleriyle olan anlamsal bağlantıları göz ardı edilmektedir. BoW modelinde her kelimenin vektör uzayındaki karşılığı sayısal bir değer olmaktadır. Kelime sayısının artmasıyla vektör uzay boyutu oldukça büyük ve çoğu elemanı sıfırdan oluşan bir seyrek (sparse) vektör oluşturmaktadır. Bu gösterim şekli sınıflama performansında kötü sonuçlar verebilmektedir.

Bu soruna çözüm olarak Mikolov ve arkadaşları kelimelerin anlamsal ilişkilerini dikkate alan gömülü bir vektör modeli olan Word2Vec modelini geliştirmişlerdir (Mikolov vd. 2013, Tezgider vd. 2018). Word2Vec girdi olarak metin içerisindeki kelimeleri alarak n boyutlu bir vektör uzayı oluşturmaktadır (Ayata vd. 2017). Word2Vec tek gizli katmana sahip yapay sinir ağı yapısını kullanan denetimsiz bir vektör gösterim şekline sahip olup, geri yayılım modeliyle sürekli kendini yenileyebilmektedir (Mikolov vd. 2013, Yıldırım, Yıldız 2018).



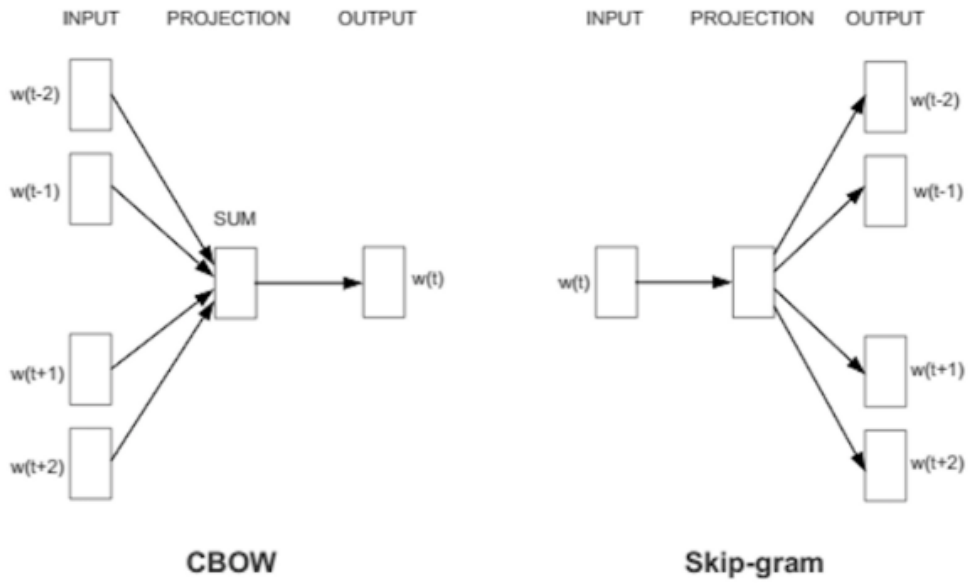
Şekil 4.1. word2Vec örnek çalışma akışı

Word2Vec modelinde, n boyutlu bir vektör uzayında her kelimeyi bir vektör temsil etmektedir. Model, vektörler arasındaki uzaklığı hesaplayarak kelimelerin birbirine olan benzerliklerini tespit etmeye çalışmaktadır (Çoban, Karabey 2017; Yuan vd. 2014). Şekil 4.2.'yi detaylandırarak olursak, $vec(\text{erkek}) - vec(\text{kral}) + vec(\text{kadın}) = vec(\text{kraliçe})$ gibi kelimelerin anlamsal ilişkilerini tespit ederek kelimeler üretebilmektedir.



Şekil 4.2. Word2Vec vektör uzayı

Word2Vec, Sürekli Kelime Torbası (Continuous Bag of Words - CBoW) ve Skip-Gram olmak üzere iki ana öğrenme modelini kullanmaktadır. Şekil 4.3.'te bu iki öğrenme modelinin yapısı verilmiştir.



Şekil 4.3. Word2Vec Model Yapısı (Mikolov vd. 2013)

4.3.2.1. CBoW Öğrenme Modeli

CBoW öğrenme modelinde, metin içerisindeki kelimelerin sırası modeli etkilemediği için kelimeler torbası olarak adlandırılmaktadır. Vektör uzay modelinde her bir kelime bir vektör tarafından temsil edilmektedir. CBoW modelinde hedef kelimenin komşu kelimelerine (sağındaki ve solundaki kelimeler) bakılmakta ve

hedef kelime, komşu kelimelerden tahmin edilmektedir (Karasoy, Ballı 2017; Şahin 2017). Şekil 4.3'te CBoW mimarisi gösterilmektedir.

4.3.2.2. Skip-Gram Öğrenme Modeli

Skip-Gram modeli, CBoW modelinin aksine hedef kelimeye bakılarak (sağında ve solundaki kelimeler) komşu kelimeler tahmin edilmektedir (Karasoy, Ballı 2017; Şahin 2017). Skip-Gram mimarisi Şekil 4.3'te gösterilmektedir. CBoW modelinin aksine skip-gram modelinde metin içerisindeki kelimelerin sıralamasından etkilenmektedir ve eş anlamlı kelimelerin öğrenilmesi CBoW modeline göre daha iyidir. Şekil 4.4'te görüldüğü üzere hedef kelime input sütunundaki kelime olarak belirtilmiş ve tahmin edilen komşu kelimeler ise output sütununda gösterilmektedir.

SKIP GRAMS WORD2VEC WITH WINDOW_SIZE=1							Training Samples	
							Input	Output
1)	bir	yandan	da	hiç	konuşmak	istemiyor	bir	yandan
2)	bir	yandan	da	hiç	konuşmak	istemiyor	yandan	bir
3)	bir	yandan	da	hiç	konuşmak	istemiyor	da	yandan
4)	bir	yandan	da	hiç	konuşmak	istemiyor	hiç	da
5)	bir	yandan	da	hiç	konuşmak	istemiyor	konuşmak	hiç
6)	bir	yandan	da	hiç	konuşmak	istemiyor	istemiyor	konuşmak

Şekil 4.4. Skip-Gram çalışma yapısı (Büyükkinacı, 2018)

Sonuç olarak küçük veri setlerinden CBoW öğrenme modeli daha iyi çalışırken, büyük veri setlerinde Skip-Gram öğrenme modeli daha iyi çalışmaktadır.(Büyükkinacı,2018)

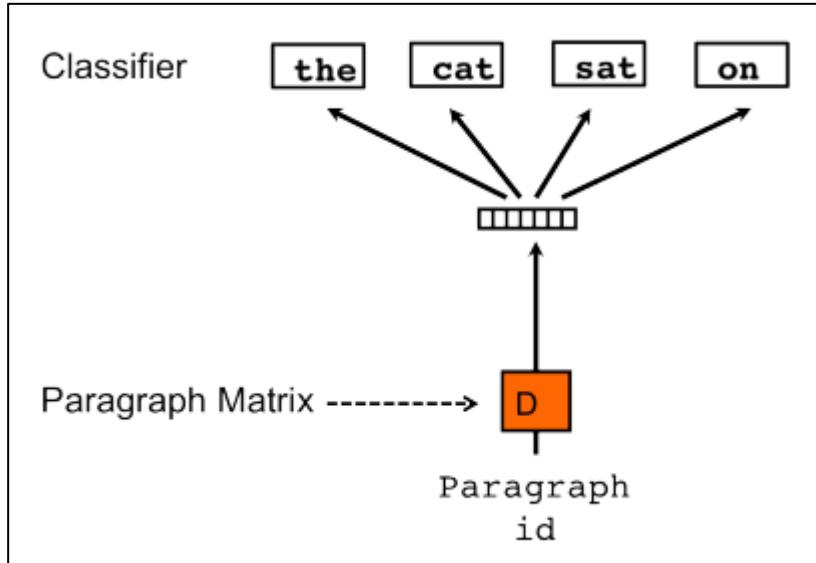
4.3.3. Doc2Vec Modeli

Word2Vec modelinde her kelime için bir vektör oluşturulurken, belgenin tamamını temsil etmek için vektörlerin ortalaması, birleştirilmesi yada kelime kümelerinin histogramları kullanılmaktadır. Bu durumda kelimelerin belge

içerisindeki konumları göz ardı edilmektedir. Bunun sonucunda her bir belgeye bir numara verilerek Doc2Vec modeliyle vektörler oluşturulabilir. Doc2Vec, Word2Vec'ten esinlenerek ortaya çıkmıştır. Doc2Vec modelinde amaç, n boyutlu bir vektör uzayında her belgeyi bir vektörün temsil etmesi ve vektörler arasındaki uzaklık hesaplanarak belgelerin birbiriyle olan anlamsal benzerliklerini tespit edebilmektir (Çoban, Karabey 2017; Yuan vd. 2014). Doc2Vec, Dağıtık Kelimeler Torbası (Distributed Bag Of Words - DBoW) ve Dağıtık Hafıza (Distributed Memory - DM) olmak üzere iki ana öğrenme modelini kullanmaktadır (Çoban, Karabey 2017).

4.3.3.1. DBoW Öğrenme Modeli

DBoW öğrenme modeli, Word2Vec yönteminin Skip-Gram modeli ile aynı mantıkta çalışmaktadır. Fakat ona göre daha basit ve daha hızlı işlev görmektedir. DBoW modeli, paragraftaki kelimelerden yola çıkarak kelimelerin olasılık dağılımını tahmin edebilmektedir (Kimothi vd. 2016; Stiebellehner vd. 2017). Şekil 4.5.'te DBoW mimarisi gösterilmektedir.

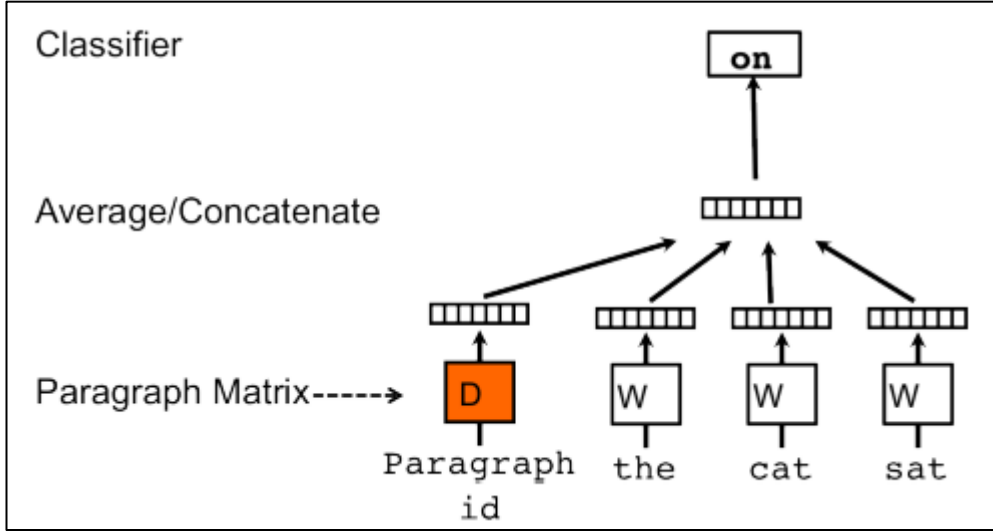


Şekil 4.5. DBoW mimarisi

4.3.3.2. DM Öğrenme Modeli

DM öğrenme modeli, Word2Vec yönteminin CBoW modeli ile aynı mantıkta çalışmaktadır. Şekil 4.6'da görüldüğü üzere hedef kelimenin tahmini, önceki kelimelerin vektörü ve paragraf vektörü ile sağlanmaktadır. Şekil 4.6'daki gibi her

paragraf için bir “id” tanımlanmaktadır. DM modelinde kelimelerin sırası önemlidir (Yıldırım, Yıldız 2018).



Şekil 4.6. DM mimarisi

4.4. Sınıflandırma Yöntemleri

4.4.1. Destek Vektör Makinaları

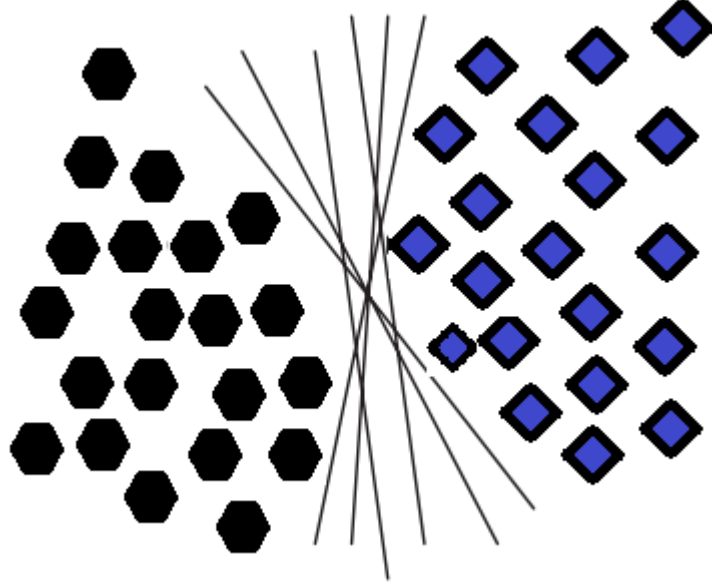
Destek Vektör Makinaları (DVM), sınıflandırma, kümeleme, tahminde bulunma gibi problemlerin çözümünde kullanılan, istatistiksel öğrenme teorisine dayalı bir sınıflandırma algoritmasıdır (Kavzoğlu, Çölkesen 2010; Çomak 2004). İki ve daha fazla sınıflandırma isteyen problemlerin çözümünde kullanılır. DVM, kullandığı kernel fonksiyonlar sayesinde doğrusal olmayan verileri sınıflandırmak için vektör uzayını daha yüksek boyutlu bir uzaya taşıyarak, doğrusal olmayan verilerin sınıflandırılmasında başarıyla kullanılmıştır (Arslan vd. 2015). DVM’de amaç, sınıfları birbirinden ayıran en iyi hiper-düzlemin tanımlanmasını sağlamaktır (Vapnik 2013). Sınıfları birbirinden ayırmak için Şekil 4.7.’deki gibi birden fazla hiper düzlem belirlenir ve bu hiper düzlemler arasında en uygun ayrımı yapan hiper düzlem, optimum hiper düzlem olarak kabul edilmektedir (Kavzoğlu, Çölkesen 2010). Optimum hiper düzlemde amaç iki sınır arasındaki uzaklığı maksimum yapmaktır. Şekil 4.8.’deki gibi sınır noktalarındaki noktalar, destek vektörleri olarak adlandırılmaktadır. N elemandan oluşan verinin hiper düzlemi belirlenirken;

$$i=1,2,3,\dots,N \text{ olsun,}$$

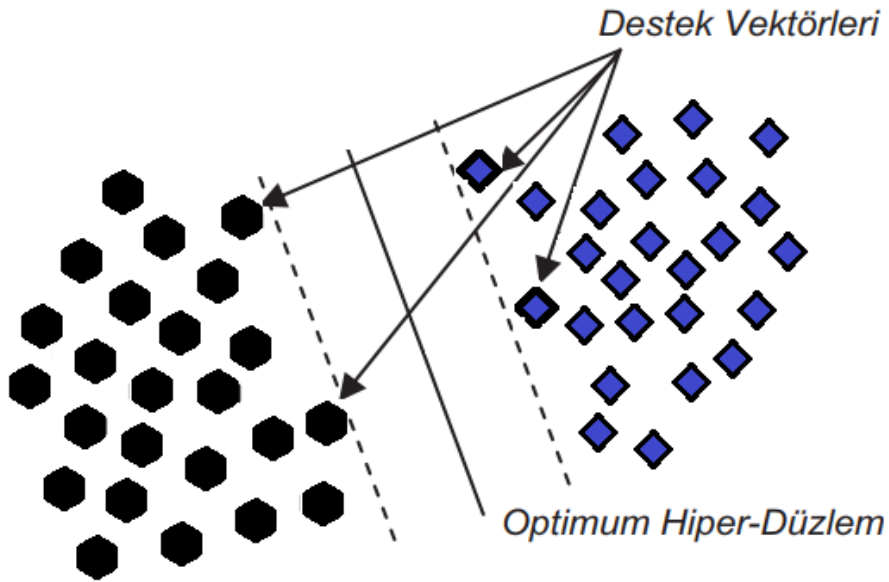
X_i = özellik vektörü

$Y_i \in \{-1, 1\}$ sınıf değerleri

$W \cdot X_i + b = \pm 1$ olarak ifade edilir (Kavzoğlu, Çölkesen 2010).



Şekil 4.7. Hiper Düzlemlerin belirlenmesi



Şekil 4.8. Optimum hiper düzlemde destek vektörleri

Yapısal olmayan veri yığınlarını sınıflandırmada DVM'in optimum hiper düzlemi oluşturması bu kadar kolay olamamaktadır. Bu tip veri türlerini ayırmada

Şekil 4.8'deki gibi girdi uzayından özellik uzayına taşınır ve kernel fonksiyonu ile veri yüksek bir boyuta dönüştürülür (Kavzoğlu, Çölkesen 2010). Kernel fonksiyonu:

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$$

şeklinde ifade edilir. En çok kullanılan kernel fonksiyonları Gauss, Polynomial ve Sigmoid fonksiyonlarıdır.

4.4.2. Naïve Bayes Algoritması

Naïve Bayes algoritması, olasılık tabanlı olup istatistiğe dayalı bir sınıflandırma algoritmasıdır. Mevcut veri seti kullanılarak önce model eğitilir daha sonra gelen yeni verinin hangi sınıfa ait olduğu konusunda olasılıkları hesaplar (Ur-Rahman, Harding 2012; Özyurt, Köse 2010). Girdiyi, olasılıksal olarak en yüksek çıkan sınıfa dâhil eder. Naïve Bayes algoritması, iyi performansı, basit yapısı ve eksik veriye olan duyarsızlığından dolayı yaygın olarak kullanılan bir sınıflandırma yöntemidir. Bu algorithmada;

$$P(A/B) = (P(B/A) * P(A)) / P(B)$$

$$P(A) = A \text{ olayının bağımsız olasılığı,}$$

$$P(B) = B \text{ olayının bağımsız olasılığı,}$$

$$P(A/B) = B \text{ olayı bilindiğinde A olayının olasılığı,}$$

$$P(B/A) = A \text{ olayı olduğu bilindiğinde B olayının olasılığı,}$$

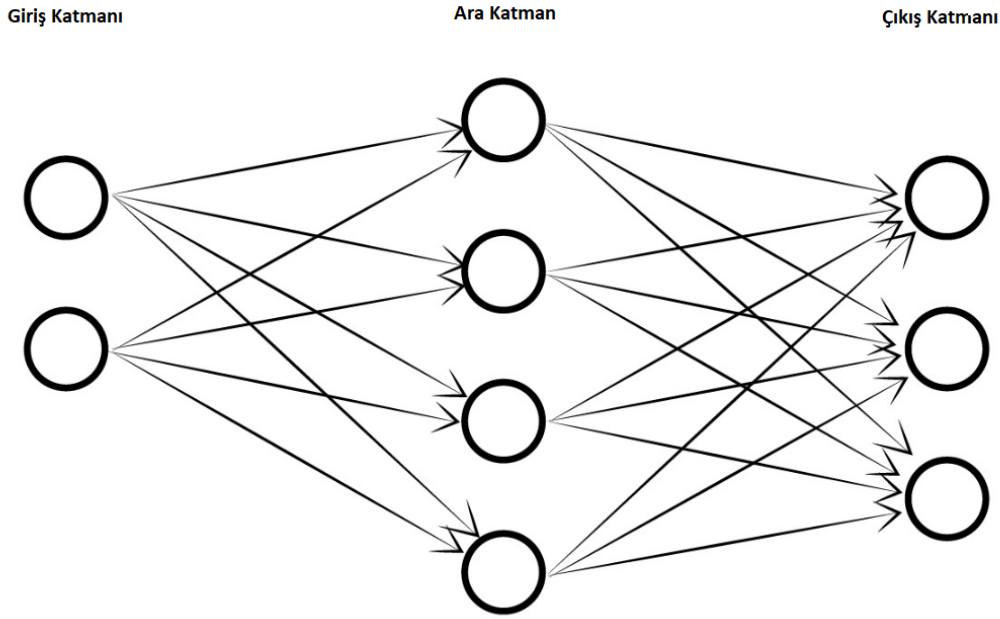
$P(A/B)$ olasılığını en yüksek yapan durumlar hesaplanarak yeni gelen verinin sınıfı tahmin edilir (Çalış vd. 2013).

4.4.3. Çok Katmanlı Yapay Sinir Ağları

Yapay sinir ağları insan beyninden esinlenerek ortaya çıkmış olup, farklı işlemleri gerçekleştiren birkaç nöronun paralel olarak dağıtıldığı yapılardır. Yapay sinir ağlarındaki nöronlar problemin çözümüne yardımcı olacak işlemleri içermektedir ve bu nöronların sayısı çözümlenmesi istenen problemlere göre değişiklikler göstermektedir (Çuhadar, Kayacan 2005). Yapay sinir ağları geçmiş bilgileri inceleyip öğrenme yoluyla gelecek bilgiler üzerine tahminde bulunma, birbiriyle ilişkilendirme, keşfetme gibi insan beyninin yapabilecekleri işlemleri

yardım almadan yapabilmesi için geliştirilmiştir (Uğur, Kınacı 2006). Mühendislik Problemlerinin %95'inin çözümüne destek olan çok katmanlı yapay sinir ağları günümüzde en yaygın kullanılan yapay sinir ağı modelidir. Çok katmanlı yapay sinir ağları, XOR problemlerinin çözümlenmeleri sonucunda ortaya çıkmıştır (Öztemel 2003). Yapay sinir ağları 3 temel katmandan oluşmaktadır:

- **Girdi Katmanı:** Dış ortamdan veri girişini sağlar.
- **Ara Katman:** Girdi katmanından gelen verileri işleyerek çıktı katmanına gönderir. Dış ortamla bir bağlantısı yoktur ve bu katman gizli katman olarak da adlandırılır. Bir problemin çözümünde birden fazla ara katman kullanılabilir. Ara katman sayısı, ağın performansını etkilediği için önemlidir (Benli 2002).
- **Çıktı Katmanı:** Ara katmandan gelen bilgileri işleyerek sonuçları dış ortama aktarır.



Şekil 4.9. Çok katmanlı yapay sinir ağı modeli

Bu katmanlardaki sinir ağlarının sayısı probleme bakılarak belirlenir, fakat ara katman sayısı deneme yanılma yoluyla da belirlenebilir. Ağın performansı, ara katman sayısına bağlıdır. Performansı ölçmek için önce belirli bir miktar veri seti ile

eđitilir ve daha sonra hi sisteme sunulmamıř bir veride alınan sonu incelenir. Eđer verilen sonu dođru ise performans iyi olarak nitelendirilir (Öztemel 2003).

ok katmanlı yapay sinir ađları

- Sınıflandırma,
- Tahminde bulunma,
- Yorumlama,
- Teřhis etme,
- Veri filtreleme,
- Veri ilişkilendirme,

gibi mühendislik problemlerinde daha fazla kullanılmakta ve başarılı sonular vermektedir.

4.5. Deđerlendirme Kriterleri

Sınıflandırma yöntemlerine sunulan eđitim verisi kullanılarak problemin özümüne yönelik model oluřturulmaktadır. Modele uygulanan test verisi sonucunda elde edilen ıktılar, deđerlendirme kriterleri kullanılarak incelenir ve modelin dođruluk düzeyi belirlenir.

4.5.1. Dođruluk – Hata Oranı

Seilen sınıflandırma yönteminin ve modelin dođruluk derecesinin belirlenmesinde kullanılan en yaygın yöntemdir. Modelin dođruluk oranı, dođru sınıflandırılmıř örnek sayısının toplam örnek sayısına oranı sonucunda belirlenir. Hata oranı ise “1 - dođruluk oranı” řeklinindedir. Yani yanlış sınıflandırılmıř örnek sayısının toplam örnek sayısına oranıdır.

$$\text{Dođruluk (A)} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Hata Oranı} = \frac{TP + FN}{TP + FP + FN + TN}$$

TP : Dođru Pozitif

TN: Dođru Negatif

FP: Yanlıř Pozitif

FN: Yanlıř Negatif

4.5.2. Keskinlik

Keskinlik, sınıflandırma tahmininin dođru pozitif olan örnek sayısının, pozitif olan tüm örnek sayısına oranıyla belirlenmektedir.

$$\text{Keskinlik (P)} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

TP: Dođru Pozitif

FP: Yanlıř Pozitif

4.5.3. Duyarlılık

Duyarlılık, sınıflandırma tahmininin dođru pozitif örnek sayısının toplam pozitif örnek sayısına oranıyla belirlenmektedir.

$$\text{Duyarlılık (R)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

TP: Dođru Pozitif

FN: Yanlıř Negatif

4.5.4. F – Ölçüsü

Keskinlik ve dođruluk ayrı ayrı deđerlendirildiđinde modelin dođruluđu üzerine yeterli bir sonuç çıkarılamamaktadır. Bundan dolayı keskinlik ve duyarlılık ölçütlerinin birbiriyle iřlenmesi sonucu elde edilen F-ölçütü ile modelin deđerlendirilmesi daha dođru sonuç verebilmektedir.

$$\text{F – Metric} = \frac{2 * \text{P} * \text{R}}{\text{P} + \text{R}}$$

P: Keskinlik

R: Duyarlılık

4.6. Cross Validation (Çapraz Doğrulama)

Cross Validation yöntemi ilk kez 1948 yılında Kurtz tarafından önerilmiş (Kurtz 1948) ve 1982 yılında Krus ve Fuller tarafından geliştirilerek multicross validation haline getirilmiştir (Krus 1982). Cross Validation, modelin eğitilmesi için kullanılan veri setinin rastgele parçalara ayrılması ile uygulanan bir yöntemdir. Öncelikle eğitim için ayrılan veri seti kullanılarak model eğitilir, test verisi kullanılarak da modelin doğruluğu test edilir. Genel olarak veri setinin %20-%33'lük kısmı test verisi olarak kullanılırken kalanı modelin eğitimi için kullanılmaktadır. Fakat kullanılan bu klasik yöntem verilerin dağılımına bağlı olarak bazı sapma ve hatalar oluşturabilmektedir.

Cross Validation veri setini k parçaya böler, her bir parçayı bir kez test verisi olarak kullanır ve geriye kalan $k-1$ 'lik veri parçasıyla da model eğitilir. Şekil 4.10.'da gösterildiği gibi bu işlem k kez tekrarlanır (Temel vd. 2012). Bu sayede verinin tamamı hem test hem de eğitim verisi olarak kullanılır. Bu yöntemin doğruluk değeri, modelin test sonucu vermiş olduğu doğruluk değerlerinin ortalaması olarak kabul edilmektedir (Refaeilzadeh vd. 2009; Hastie vd. 2009).

1	<u>$n-1$</u>	<u>n</u>
Eğitim Seti			Test Seti

1	<u>$n-1$</u>	<u>n</u>
Eğitim Seti			Test Seti

⋮

1	<u>$n-1$</u>	<u>n</u>
Test Seti	Eğitim Seti		

Şekil 4.10. Cross Validation (Temel vd. 2012)

BÖLÜM 5

UYGULAMA

Günümüz firmaları birbirleriyle iletişim halinde olarak, mevcut sorun ve düzenlemeler için çoğu kez danışmanlık desteği almaktadırlar. Bu durum danışmanlık hizmeti sunan firma sayısını arttırmakta ve bu danışmanlık hizmeti kullanım oranını artırmaktadır. Danışmanlık hizmeti sunan firmaların mevcut müşterileriyle iletişimde olmaları, onların talep ve ihtiyaçlarının tespiti ve toplanması, yardım masası gibi uygulamalar üzerinden sağlanmaktadır. Müşteri devamlılığını ve memnuniyetini sağlama açısından, toplanan müşteri talep ve ihtiyaçların ilgili birime yönlendirimi ve çözümlene süreleri önemli rol oynamaktadır.

Bu tez çalışmasında, Türkiye'nin en büyük SAP çözüm ortaklarından biri olan Detay Danışmanlık firmasının mevcut müşterilerinin taleplerine daha hızlı ve doğru sonuçlar aldırılmayı hedefleyen problem temelinde alınan taleplerin organizasyonel iş akışlarına uygun olarak sınıflandırılması hedeflenmiştir.

5.1. Taleplerin Toplanması

Müşteri talepleri Türkiye merkezli yerel bir firmanın mevcut müşterileri için hazırlanmış olduğu bir yardım masası uygulaması aracılığıyla toplanmaktadır. İlgili kuruluş, müşteri odaklı bir firma olmasından dolayı müşteri ilişkilerini sıkı tutmayı amaçlamakta ve gelen talepleri anlık olarak değerlendirerek geri dönüşler yapmayı hedeflemektedir. Fakat kullanılan yardım masası uygulamasında müşteri oluşturduğu talebin hangi birime aktarılması gerektiğini kendisi belirlemek zorundadır.

Bu çalışmada önerilen yöntem, müşterilerin buldukları talebin aktarılacağı birimi kendileri seçmek yerine sistemin içerikten yola çıkarak ilgili birimi tahmin etmesini ve yönlendirmesini hedeflemektedir. Firma bünyesinde müşterilerin kullanımına sunulan yardım masası uygulamasına ait ekran görselleri Şekil 5.1. ve 5.2.'de sunulmuştur.

Şekil 5.1. Detaysoft müşteri talep oluşturma ekranı

Mesaj Numarası	Durum	Öncelik	Kategori	Mesaj Tipi	Tanım	Yaratma Tarihi	Bildiren	İşle
TK0000168324	Yeni	Orta		Servis Talebi	deneme bir ki	17.12.2018	ESRA AKAR	den
TK0000168322	Test	Orta		Efor Bilgi Talebi	nasılın	17.12.2018	ESRA AKAR	den
TK0000168317	Müşteri Aksiyonu	Orta		Problem Talebi	merhaba	17.12.2018	deneme deneme	ES

Şekil 5.2. Detaysoft müşteri taleplerinin listesi

5.2. Ön İşlem Aşamaları

Müşteri talepleri doğal dil ile yazıldığı için belirli bir yapıya sahip değildir ve bu nedenle bu talepler ön işlem teknikleri uygulanarak yapısal hale getirilmesi gerekmektedir. Sırasıyla uygulanan ön işlem aşamaları:

- Simge ve noktalama işaretlerinin temizlenmesi,
- Metnin tamamının küçük harfe çevrilmesi,
- Metnin kelimelerine (token) ayrılması,
- Her bir kelimenin NLTK kütüphanesinden geçirilerek köklerinin bulunması,
- Eldeki kök kelimelerden edat, bağlaç gibi durak kelimelerin çıkarılması,

- Terim frekanslarının hesaplanması ve terim frekansı 2'den uzun olan kelimeler ve frekans oranı 15 ile 400 arasında olan kelimelerin vektör uzayı ve kelimeler çantasının oluşturulması.
- Her bir talebin vektör uzayındaki vektörel yapısının oluşturulması, şeklinde gerçekleştirilmiştir.

5.3. Talep Verilerinin Modele Göre Sınıflandırılması

Toplanan talep verileri daha önceden ilgili bölümler tarafından başarıyla çözümlenmiş talepleri içermektedir. 886 adet talep verisinin bir kısmıyla sınıflandırma modeli eğitilmekte, kalan kısmıyla da modelin doğruluk oranını tespit edilmektedir. Mevcut veri setinin eğitim ve test verisi şeklinde ayrıştırılmasında, “Veri setinin eğitim-test olarak ayrılması” ve “k-kat çapraz doğrulama” olarak 2 yöntem kullanılmıştır.

5.3.1. Veri Setinin Eğitim-Test Olarak Ayrılması Yöntemi

Veri setini, eğitim ve test seti olarak 2 parçaya ayırma işlemi sınıflandırma problemlerinde yaygın olarak kullanılan bir yöntemdir. Veri setinin %67'si eğitim verisi, %33'ü de test verisi olarak ayrıştırılmıştır. Eğitim seti kullanılarak sınıflandırma algoritması eğitilmiş ve bir model oluşturulmuştur. Daha önce sistemin görmediği test verileri modele sunularak modelin sınıflandırma performansı gözlemlenmiştir. Ancak gerçekleştirilen test sonuçlarının doğruluğu, modele sunulan verilerin problemi homojen olarak temsil etme oranına bağlıdır.

5.3.1.1. BoW Öznitelik Çıkarım Yöntemiyle Sınıflandırma

BoW öznitelik çıkarım yöntemiyle elde edilen vektörler MLP, SVM ve NB sınıflandırma yöntemlerine uygulanmıştır. Talep verilerinin %67'lik dilimi ile model eğitilmiş, veri setinin %33'lük dilimiyle oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur.

Tablo 5.1. SVM Karmaşıklık Matrisi

SVM	SAPFI	SAPHR	SAPLO
SAPFI	103	0	0
SAPHR	89	0	0
SAPLO	0	0	101

Tablo 5.1’de, SVM sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPHR sınıfı ile etiketlenmiş 89 adet müşteri talebi, SAPFI sınıfında gösterilmiştir.

Toplamda 89 adet müşteri talebi yanlış sınıflandırılmıştır. Bu durumun modelin SAPHR etiketine sahip veriler için önemli bir performans sorunu olduğunu göstermektedir.

Tablo 5.2. NB Karmaşıklık Matrisi

NB	SAPFI	SAPHR	SAPLO
SAPFI	93	7	3
SAPHR	8	78	3
SAPLO	1	13	87

Tablo 5.2’de, NB sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 10 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 11 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 14 müşteri talebi birbirlerinden etkilenerek yanlış sınıflarda gösterilmiştir.

Toplamda 35 adet müşteri talebi NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.3. MLP Karmaşıklık Matrisi

MLP	SAPFI	SAPHR	SAPLO
SAPFI	98	3	2
SAPHR	7	74	8
SAPLO	0	14	87

Tablo 5.3'te, MLP sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 5 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 15 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 14 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 34 adet müşteri talebi MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.4. BoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.696	0.533	0.696	0.590	0.303
NB	0.880	0.884	0.880	0.881	0.119
MLP	0.883	0.884	0.883	0.883	0.116

Tablo 5.4'te veri setinin eğitim ve test olarak ayrıştırılarak BoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmalarına ait metrik değerleri sunulmuştur.

BoW öznitelik çıkarma yöntemi kullanılarak oluşturulan modellerin sınıflandırma çalışmalarında en iyi sonucu, %88.3 doğruluk, %88.4 keskinlik, %88.3 duyarlılık, %88.3 F-ölçütü ve %11.6 hata oranıyla MLP sınıflandırma algoritmasının verdiği gözlemlenmiştir.

5.3.1.2. Doc2Vec-DBoW Öznitelik Çıkarma Yöntemiyle Sınıflandırma

Doc2Vec öznitelik çıkarma yönteminin DBoW öğrenme modeli kullanılarak elde edilen vektörler MLP, SVM ve NB sınıflandırma yöntemlerine uygulanmıştır. Talep verilerinin %67'lik dilimi ile model eğitilmiş, veri setinin %33'lük dilimiyle oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur.

Tablo 5.5. SVM Karmaşıklık Matrisi

SVM	SAPFI	SAPHR	SAPLO
SAPFI	66	0	1
SAPHR	72	6	0
SAPLO	42	0	35

Tablo 5.5'te, SVM sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 1 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 72 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 42 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 115 adet müşteri talebi SVM algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.6. NB Karmaşıklık Matrisi

NB	SAPFI	SAPHR	SAPLO
SAPFI	41	11	15
SAPHR	11	60	7
SAPLO	8	2	67

Tablo 5.6'da, NB sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 26 müşteri talebi,

gerçekte SAPHR sınıfı ile etiketlenmiş 18 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 10 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 54 adet müşteri talebi NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.7. MLP Karmaşıklık Matrisi

MLP	SAPFI	SAPHR	SAPLO
SAPFI	51	7	9
SAPHR	12	60	6
SAPLO	6	4	67

Tablo 5.7’de, MLP sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 16 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 18 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 10 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 44 adet müşteri talebi MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.8. DBoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.481	0.799	0.481	0.426	0.518
NB	0.756	0.756	0.756	0.754	0.243
MLP	0.801	0.803	0.801	0.801	0.198

Tablo 5.8’de veri setinin eğitim ve test olarak ayrıştırılarak Doc2Vec öznitelik çıkarma yönteminin DBoW öğrenme tekniği ile yapılan tüm sınıflandırma çalışmalarının metrik değerleri sunulmuştur.

Doc2Vec öznitelik çıkarma yönteminin DBoW öğrenme tekniği kullanılarak oluşturulan modelin sınıflandırma çalışmalarında en iyi sonucu, %80.1 doğruluk, %80.3 keskinlik, %80.1 duyarlılık, %80.1 F-ölçütü ve %19.8 hata oranıyla MLP sınıflandırma algoritmasının verdiği gözlemlenmiştir.

5.3.1.3. Doc2Vec-DM Öznitelik Çıkarma Yöntemiyle Sınıflandırma

Doc2Vec öznitelik çıkarma yönteminin 2. öğrenme modeli olan DM öğrenme modeli kullanılarak elde edilen vektörler MLP, SVM ve NB sınıflandırma yöntemlerine uygulanmıştır. Talep verilerinin %67'lik dilimi ile model eğitilmiş, veri setinin %33'lük dilimiyle oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur.

Tablo 5.9. SVM Karmaşıklık Matrisi

SVM	SAPFI	SAPHR	SAPLO
SAPFI	67	0	0
SAPHR	76	2	0
SAPLO	77	0	0

Tablo 5.9'da, SVM sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPHR sınıfı ile etiketlenmiş 76 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 77 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 153 adet müşteri talebi SVM algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.10. NB Karmaşıklık Matrisi

NB	SAPFI	SAPHR	SAPLO
SAPFI	45	5	17
SAPHR	32	36	10
SAPLO	33	3	41

Tablo 5.10'da, NB sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 22 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 42 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 36 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 100 adet müşteri talebi, NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.11. MLP Karmaşıklık Matrisi

MLP	SAPFI	SAPHR	SAPLO
SAPFI	51	6	10
SAPHR	24	46	8
SAPLO	24	8	45

Tablo 5.11'de, MLP sınıflandırma algoritması kullanılarak eğitilen modele, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisi sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 16 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 32 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 32 müşteri talebi yanlış sınıfa dâhil edilmiştir.

Toplamda 80 adet müşteri talebi, MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.12. DM Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.310	0.443	0.310	0.158	0.689
NB	0.549	0.620	0.549	0.556	0.450
MLP	0.639	0.672	0.639	0.642	0.360

Tablo 5.12’de veri setinin eğitim ve test olarak ayrıştırılarak Doc2Vec öznitelik çıkarma yönteminin DM öğrenme tekniği ile yapılan tüm sınıflandırma çalışmalarının metrik değerleri sunulmuştur.

Doc2Vec öznitelik çıkarma yönteminin DM öğrenme tekniği kullanılarak oluşturulan modelin sınıflandırma çalışmalarında en iyi sonucu, %63,9 doğruluk, %67,2 keskinlik, %63,9 duyarlılık, %64,2 F-ölçütü ve %36 hata oranıyla MLP sınıflandırma algoritmasının verdiği gözlemlenmiştir.

5.3.2. 10-Kat Çapraz Doğrulama Yöntemi

Sınıflandırıcıların performansı, genellikle tahmin hatası ile belirlenir. Tahmin hatasının düşük olması, sınıflandırıcının performansının yüksek olduğunu göstermektedir (Kırlioğlu, Ceyhan 2014). Müşteri taleplerinin, modelin eğitimi için hangi dilimini kullanılacağı ve problemi temsil edecek test verileri için hangi diliminin alınacağı homojen olarak belirlenememektedir. Seçilen ilk dilim eğitim ve test verileri için mükemmel sonuç veren bir model, eğitim ve test verileri için veri setinin başka bir dilimi aldığı anda aynı sonucu vermediği gözlemlenmiştir. Bundan dolayı ortalama bir değerlendirme elde edebilmek için veri setinin bölünmesinde k-kat çapraz doğrulama yöntemi kullanılmıştır. K değeri 10 alınarak mevcut veri seti 10 eşit parçaya bölünmüş olup 9 parçası eğitim, 1 parçası da test verisi olarak kullanılmıştır. Bu işlem sırasıyla her dilim test olmak koşuluyla 10 defa tekrarlanmıştır. Sınıflandırma sonucunda elde edilen sonuçların ortalaması sınıflandırma modelinin doğru tespit oranı olarak kabul edilmiştir. Her çapraz doğrulama işleminde ortaya çıkan hatanın oranı da hatanın ortalama değeri olarak kabul edilmiştir. Bu durumda hata oranı eğitim setine ve çapraz doğrulama sayısına bağlıdır.

1.kat	Test Seti	Eđitim Seti		
2.kat	Eđitim Seti	Test Seti	Eđitim Seti	
3.kat	Eđitim Seti		Test Seti	Eđitim Seti
4.kat	Eđitim Seti		Test Seti	Eđitim Seti
5.kat	Eđitim Seti		Test Seti	Eđitim Seti
6.kat	Eđitim Seti		Test Seti	Eđitim Seti
7.kat	Eđitim Seti		Test Seti	Eđitim Seti
8.kat	Eđitim Seti		Test Seti	Eđitim Seti
9.kat	Eđitim Seti		Test Seti	Eđitim Seti
10.kat	Eđitim Seti			Test Seti

Şekil 5.3. 10-kat apraz dođrulama

5.3.2.1. BoW znitelik ıkarım Yöntemiyle Sınıflandırma

BoW znitelik ıkarım yöntemiyle elde edilen vektörler MLP, SVM ve NB sınıflandırma yöntemlerine uygulanmıştır. Talep verileri 10 kat apraz dođrulama yöntemine göre 10 eşit paraya bölünmüş olup 9 para ile model eğitilmiş, 1 parasıyla da oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur. Her sınıflandırma yöntemi için 10 adet karmaşıklık matrisi oluşturulmuştur.

Tablo 5.13. 10 Kat Çapraz Doğrulamada SVM Sınıflandırma Yönteminin Karmaşıklık Matrisleri

SVM								
1 – 2		1	2	3		1	2	3
	1	29	0	0	1	38	0	0
	2	24	0	0	2	0	0	26
	3	0	0	36	3	0	0	25
3 – 4		1	2	3		1	2	3
	1	23	0	0	1	32	0	0
	2	31	0	0	2	14	0	0
	3	0	0	35	3	0	0	43
5 – 6		1	2	3		1	2	3
	1	29	0	0	1	29	0	0
	2	27	0	0	2	33	0	0
	3	0	0	33	3	0	0	27
7 – 8		1	2	3		1	2	3
	1	38	0	0	1	21	0	0
	2	0	0	20	2	39	0	0
	3	0	0	30	3	0	0	28
9 – 10		1	2	3		1	2	3
	1	36	0	0	1	33	0	0
	2	0	0	30	2	0	0	29
	3	0	0	22	3	0	0	26
1:SAPFI 2:SAPHR 3:SAPLO								

Tablo 5.13'te, SVM sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPHR sınıfı ile etiketlenmiş müşteri talepleri yanlış sınıfa dahil edilmiştir.

Toplamda 10 karmaşıklık matrisi incelendiğinde SAPHR sınıfı ile etiketlenmiş müşteri talepleri SVM algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.14. 10 Kat Çapraz Doğrulamada NB Sınıflandırma Yönteminin Karmaşıklık Matrisleri

NB									
1 – 2		1	2	3		1	2	3	
	1	26	2	1	1	33	2	3	
	2	1	21	2	2	1	23	2	
	3	0	4	32	3	0	2	23	
3 – 4		1	2	3		1	2	3	
	1	20	1	2	1	29	2	1	
	2	1	30	0	2	0	13	1	
	3	0	3	32	3	1	1	41	
5 – 6		1	2	3		1	2	3	
	1	28	1	0	1	27	2	0	
	2	4	23	0	2	4	27	2	
	3	0	4	29	3	0	8	19	
7 – 8		1	2	3		1	2	3	
	1	33	4	1	1	21	0	0	
	2	4	16	0	2	3	31	5	
	3	0	5	25	3	1	2	25	
9 – 10		1	2	3		1	2	3	
	1	36	0	0	1	32	0	1	
	2	1	28	1	2	6	23	0	
	3	0	1	21	3	1	1	24	
		1:SAPFI		2:SAPHR		3:SAPLO			

Tablo 5.14'te, NB sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler

sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 23 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 38 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 34 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 95 adet müşteri talebi NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.15. 10 Kat Çapraz Doğrulama MLP Sınıflandırma Yönteminin Karmaşıklık Matrisleri

MLP											
1 – 2		1	2	3		1	2	3			
	1	29	0	0	1	36	2	0			
	2	1	22	1	2	0	20	6			
	3	1	9	26	3	0	2	23			
3 – 4		1	2	3		1	2	3			
	1	20	3	0	1	28	3	1			
	2	0	26	5	2	0	12	2			
	3	0	5	30	3	1	2	40			
5 – 6		1	2	3		1	2	3			
	1	28	0	1	1	27	2	0			
	2	2	23	2	2	1	28	4			
	3	0	6	27	3	0	2	25			
7 – 8		1	2	3		1	2	3			
	1	35	1	2	1	20	0	1			
	2	1	17	2	2	4	29	6			
	3	0	6	24	3	0	1	27			
9 – 10		1	2	3		1	2	3			
	1	33	2	1	1	31	1	1			
	2	1	25	4	2	3	21	5			
	3	0	1	21	3	0	5	21			
1:SAPFI				2:SAPHR				3:SAPLO			

Tablo 5.15'te, MLP sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 21 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 50 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 41 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 112 adet müşteri talebi MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.16. 10 Kat Çapraz Doğrulamanın Sınıflandırma Yöntemlerine Göre Her Parçası İçin Doğruluk Değeri

A	1	2	3	4	5	6	7	8	9	10
SVM	0.73	0.70	0.65	0.84	0.69	0.62	0.77	0.55	0.65	0.67
NB	0.88	0.88	0.92	0.93	0.89	0.82	0.84	0.87	0.96	0.89
MLP	0.86	0.88	0.85	0.89	0.87	0.89	0.86	0.86	0.89	0.82

Tablo 5.16'da veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak BoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmaları sonucunda elde edilen doğruluk değerleri sunulmuştur.

Tablo 5.17. BoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.691	0.544	0.691	0.592	0.308
NB	0.892	0.899	0.892	0.893	0.107
MLP	0.873	0.881	0.873	0.874	0.126

Tablo 5.17'de veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak BoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmalarına ait metrik değerleri sunulmuştur.

BoW öznitelik çıkarma yöntemi kullanılarak oluşturulan modellerin sınıflandırma çalışmalarında en iyi sonucu, %89.2 doğruluk, %89.9 keskinlik, %89.2 duyarlılık, %89.3 F-ölçütü ve %10.7 hata oranıyla NB sınıflandırma algoritmasının verdiği gözlemlenmiştir.

5.3.2.2. Doc2Vec-DBoW Öznitelik Çıkarma Yöntemiyle Sınıflandırma

Doc2Vec öznitelik çıkarma yönteminin DBoW öğrenme modeli kullanılarak elde edilen vektörler MLP, NB ve SVM sınıflandırma yöntemlerine uygulanmıştır. Talep verileri 10 kat çapraz doğrulama yöntemine göre 10 eşit parçaya bölünmüş olup 9 parça ile model eğitilmiş, 1 parçasıyla da oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur. Her sınıflandırma yöntemi için 10 adet karmaşıklık matrisi oluşturulmuştur.

Tablo 5.18. 10 kat çapraz doğrulamada SVM sınıflandırma yönteminin Karmaşıklık Matrisleri

SVM								
1 – 2		1	2	3		1	2	3
	1	29	0	0	1	27	2	9
	2	9	15	0	2	1	23	2
	3	8	1	27	3	0	0	25
3 – 4		1	2	3		1	2	3
	1	20	0	3	1	29	2	1
	2	6	23	2	2	4	10	0
	3	6	3	26	3	18	2	23
5 – 6		1	2	3		1	2	3
	1	24	2	3	1	22	1	6
	2	8	16	3	2	3	28	2
	3	8	1	24	3	5	1	21
7 – 8		1	2	3		1	2	3
	1	32	2	4	1	21	0	0
	2	1	18	1	2	26	9	4
	3	4	1	25	3	4	1	23
9 – 10		1	2	3		1	2	3
	1	27	1	8	1	23	4	6
	2	6	24	0	2	9	18	2
	3	2	1	19	3	3	1	22
1:SAPFI 2:SAPHR 3:SAPLO								

Tablo 5.18’de SVM sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 54 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 88 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 70 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 212 adet müşteri talebi SVM algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.19. 10 kat çapraz doğrulamada NB sınıflandırma yönteminin Karmaşıklık Matrisleri

NB										
1 – 2		1	2	3		1	2	3		
	1	20	0	9	1	28	3	7		
	2	3	21	0	2	1	24	1		
	3	3	1	32	3	0	0	25		
3 – 4		1	2	3		1	2	3		
	1	16	1	6	1	24	3	5		
	2	3	25	3	2	3	10	1		
	3	5	3	27	3	3	0	40		
5 – 6		1	2	3		1	2	3		
	1	23	3	3	1	20	1	8		
	2	4	19	4	2	0	29	4		
	3	7	1	25	3	4	1	22		
7 – 8		1	2	3		1	2	3		
	1	25	3	10	1	13	4	4		
	2	1	17	2	2	3	27	9		
	3	1	3	26	3	2	0	26		
9 – 10		1	2	3		1	2	3		
	1	25	3	8	1	22	4	7		
	2	3	27	0	2	1	27	1		
	3	4	0	18	3	3	1	22		
		1:SAPFI			2:SAPHR			3:SAPLO		

Tablo 5.19’da NB sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 92

müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 47 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 42 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 181 adet müşteri talebi NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.20. 10 kat çapraz doğrulamada MLP sınıflandırma yönteminin Karmaşıklık Matrisleri

MLP									
1 – 2		1	2	3		1	2	3	
	1	24	2	3	1	27	5	6	
	2	1	21	2	2	2	23	1	
	3	4	0	32	3	2	3	20	
3 – 4		1	2	3		1	2	3	
	1	19	1	3	1	24	6	2	
	2	5	23	3	2	2	10	2	
	3	7	1	27	3	6	2	37	
5 – 6		1	2	3		1	2	3	
	1	22	2	5	1	22	1	6	
	2	5	17	5	2	2	29	2	
	3	4	0	29	3	1	3	23	
7 – 8		1	2	3		1	2	3	
	1	27	5	6	1	15	2	4	
	2	2	18	0	2	3	30	6	
	3	3	2	25	3	3	0	25	
9 – 10		1	2	3		1	2	3	
	1	25	3	8	1	27	5	1	
	2	5	25	0	2	5	19	5	
	3	2	0	20	3	6	2	18	
		1:SAPFI		2:SAPHR		3:SAPLO			

Tablo 5.20’de sunulan modelin MLP sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 76 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 58 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 51 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 185 adet müşteri talebi MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.21. 10 kat çapraz doğrulamanın sınıflandırma yöntemlerine göre her parçası için doğruluk değeri

A	1	2	3	4	5	6	7	8	9	10
SVM	0.79	0.84	0.77	0.69	0.71	0.79	0.85	0.60	0.79	0.71
NB	0.82	0.86	0.76	0.83	0.75	0.79	0.77	0.75	0.79	0.80
MLP	0.86	0.78	0.77	0.79	0.76	0.83	0.79	0.79	0.79	0.72

Tablo5.21’de veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak DBoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmaları sonucunda elde edilen doğruluk değerleri sunulmuştur.

Tablo 5.22. DBoW Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.759	0.801	0.759	0.757	0.240
NB	0.795	0.806	0.795	0.795	0.204
MLP	0.793	0.803	0.793	0.793	0.206

Tablo 5.22’de veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak BoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmalarına ait metrik değerleri sunulmuştur.

DBoW öznitelik çıkarma yöntemi kullanılarak oluşturulan modellerin sınıflandırma çalışmalarında en iyi sonucu, %79.5 doğruluk, %80.6 keskinlik, %79.5

duyarlılık, %79.5 F-ölçütü ve %20.4 hata oranıyla NB sınıflandırma algoritmasının verdiği gözlemlenmiştir.

5.3.2.3. Doc2Vec-DM Öznitelik Çıkarma Yöntemiyle Sınıflandırma

Doc2Vec öznitelik çıkarma yönteminin 2. öğrenme modeli olan DM öğrenme modeli kullanılarak elde edilen vektörler MLP, SVM ve NB sınıflandırma yöntemlerine uygulanmıştır. Talep verileri 10 kat çapraz doğrulama yöntemine göre 10 eşit parçaya bölünmüş olup 9 parça ile model eğitilmiş, 1 parçasıyla da oluşturulan model test edilmiştir. Sonuçlar aşağıdaki tablolarda sunulmuştur. Her sınıflandırma yöntemi için 10 tane karmaşıklık matrisi oluşturulmuştur.

Tablo 5.23. 10 kat çapraz doğrulamada SVM sınıflandırma yönteminin Karmaşıklık Matrisleri

SVM													
1 – 2			1	2	3			1	2	3			
	1	29	0	0		1	8	1	29				
	2	23	1	0		2	7	1	18				
	3	33	0	3		3	1	0	24				
3 – 4			1	2	3			1	2	3			
	1	23	0	0		1	32	0	0				
	2	31	0	0		2	13	1	0				
	3	34	0	1		3	40	0	3				
5 – 6			1	2	3			1	2	3			
	1	28	0	1		1	26	0	3				
	2	23	3	1		2	24	3	6				
	3	26	0	7		3	16	0	11				
7 – 8			1	2	3			1	2	3			
	1	31	0	7		1	21	0	0				
	2	13	2	5		2	36	2	1				
	3	10	0	20		3	24	1	3				
9 – 10			1	2	3			1	2	3			
	1	17	0	19		1	19	0	14				
	2	11	0	19		2	18	0	11				
	3	2	0	20		3	4	0	22				
		1:SAPFI				2:SAPHR				3:SAPLO			

Tablo 5.23’de SVM sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 74 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 260 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 191 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 525 adet müşteri talebi SVM algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.24. 10 kat çapraz doğrulamada NB sınıflandırma yönteminin Karmaşıklık Matrisleri

NB									
1 – 2		1	2	3		1	2	3	
	1	19	5	5	1	29	5	4	
	2	19	5	0	2	8	17	1	
	3	11	3	22	3	6	2	17	
3 – 4		1	2	3		1	2	3	
	1	17	1	5	1	30	0	2	
	2	20	10	1	2	10	3	1	
	3	17	3	15	3	18	3	22	
5 – 6		1	2	3		1	2	3	
	1	23	5	1	1	20	2	7	
	2	11	14	2	2	6	21	6	
	3	12	1	20	3	8	3	16	
7 – 8		1	2	3		1	2	3	
	1	20	6	12	1	14	1	6	
	2	4	14	2	2	15	16	8	
	3	3	5	22	3	8	0	20	
9 – 10		1	2	3		1	2	3	
	1	25	2	9	1	27	4	2	
	2	20	8	2	2	12	17	0	
	3	10	0	12	3	9	1	16	
		1:SAPFI		2:SAPHR		3:SAPLO			

Tablo 5.24'te NB sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 84

müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 148 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 123 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 355 adet müşteri talebi NB algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.25. 10 kat çapraz doğrulamada MLP sınıflandırma yönteminin Karmaşıklık Matrisleri

MLP									
1 – 2		1	2	3		1	2	3	
	1	17	3	9	1	16	9	13	
	2	5	14	5	2	2	16	8	
	3	9	2	25	3	4	1	20	
3 – 4		1	2	3		1	2	3	
	1	15	3	5	1	22	4	6	
	2	4	24	3	2	2	10	2	
	3	14	4	17	3	9	3	31	
5 – 6		1	2	3		1	2	3	
	1	23	2	4	1	17	3	9	
	2	6	13	8	2	6	24	3	
	3	18	2	13	3	8	1	18	
7 – 8		1	2	3		1	2	3	
	1	22	7	9	1	12	2	7	
	2	2	18	0	2	2	22	15	
	3	7	3	20	3	5	4	19	
9 – 10		1	2	3		1	2	3	
	1	25	5	6	1	22	5	6	
	2	5	17	8	2	10	17	2	
	3	8	4	10	3	5	3	18	
		1:SAPFI	2:SAPHR			3:SAPLO			

Tablo 5.25’de MLP sınıflandırma algoritması kullanılarak eğitilen modellere, test verilerinin uygulanması sonucunda elde edilen karmaşıklık matrisler sunulmuştur. Sonuçlar analiz edildiğinde, gerçekte SAPFI sınıfı ile etiketlenmiş 117 müşteri talebi, gerçekte SAPHR sınıfı ile etiketlenmiş 98 müşteri talebi ve gerçekte SAPLO sınıfı ile etiketlenmiş 114 müşteri talebi yanlış sınıfa dahil edilmiştir.

Toplamda 329 adet müşteri talebi MLP algoritması kullanılarak oluşturulan modelde doğru sınıflandırılmamıştır.

Tablo 5.26. 10 kat çapraz doğrulamanın sınıflandırma yöntemlerine göre her parçası için doğruluk değeri

A	1	2	3	4	5	6	7	8	9	10
SVM	0.37	0.37	0.26	0.40	0.42	0.44	0.60	0.29	0.42	0.46
NB	0.51	0.70	0.47	0.61	0.64	0.64	0.63	0.56	0.51	0.68
MLP	0.62	0.58	0.62	0.70	0.55	0.66	0.68	0.60	0.59	0.64

Tablo5.26’da veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak DM öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmaları sonucunda elde edilen doğruluk değerleri sunulmuştur.

Tablo 5.27. DM Modelinde Sınıflandırma Yöntemlerinin Metrik Değerleri

Sınıflandırma Yöntemi	A	P	R	F-Metric	Hata
SVM	0.405	0.576	0.407	0.314	0.592
NB	0.599	0.659	0.599	0.597	0.400
MLP	0.628	0.647	0.628	0.628	0.371

Tablo 5.27’de veri setinin 10 kat çapraz doğrulama ile ayrıştırılarak BoW öznitelik çıkarma yöntemi ile yapılan tüm sınıflandırma çalışmalarına ait metrik değerleri sunulmuştur.

DM öznitelik çıkarma yöntemi kullanılarak oluşturulan modellerin sınıflandırma çalışmalarında en iyi sonucu, %62.8 doğruluk, %64.7 keskinlik, %62.8 duyarlılık, %62.8 F-ölçütü ve %37.1 hata oranıyla MLP sınıflandırma algoritmasının verdiği gözlemlenmiştir.

BÖLÜM 6

SONUÇ VE DEĞERLENDİRME

Bu çalışmada, yazılım danışmanlığı hizmeti veren Detaysoft firmasına ait son 3 yıl içerisindeki müşterileri taleplerini içeren veri seti kullanılmıştır. Firmaya ait yardım masası uygulaması aracılığıyla alınan müşteri taleplerinin ilgili birimlere aktarılması işlemi, talebi tanımlayan kişinin tercihinine bırakılmıştır. Önerilen yöntem ile müşteri taleplerinin yönlendirileceği birimin otomatik seçilebilmesi probleminin çözülmesi hedeflenmiştir. Müşteri taleplerinin ilgili birime doğru yönlendirilmesi, müşterilere etkili ve hızlı geri dönüşler sağlanması açısından büyük önem taşımaktadır.

Çalışma kapsamında, 2 veri ayrıştırma yöntemi, 3 öznitelik çıkarım modeli ve 3 sınıflandırma algoritması üzerine odaklanılmıştır. Eğitim ve test verisi olarak ayrıştırma işleminde %67'lik veri seti ile 3 model eğitilmiş ve 3 sınıf kullanılarak elde edilen değerler karşılaştırılmıştır. Tablo 5.28'de, SVM sınıflandırma yönteminin eğitim ve test verisi şeklinde ayrıştırılarak, BoW, DBoW ve DM öznitelik çıkarım modelleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. Sonuçlar incelendiğinde bu problem için SVM sınıflandırma yöntemi beklenen başarımdan uzak durmaktadır. Ortaya çıkan sonuçlar müşteri taleplerinin doğru sınıflandırılması açısından istenilen oranda elde edilememiştir.

Tablo 5.28. SVM sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

SVM	A	P	R	F
BoW	0.696	0.533	0.696	0.590
DBoW	0.481	0.799	0.481	0.426
DM	0.310	0.443	0.310	0.158

Tablo 5.29'da, NB sınıflandırma yönteminin eğitim ve test verisi şeklinde ayrıştırılarak, BoW, DBoW ve DM öznitelik çıkarım modelleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. Bu modelde BoW öznitelik çıkarım yöntemiyle elde edilen sonuçların başarımdan daha verimli olduğu görülmektedir.

Tablo 5.29. NB sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

NB	A	P	R	F
BoW	0.880	0.884	0.880	0.881
DBoW	0.756	0.756	0.756	0.754
DM	0.549	0.620	0.407549	0.556

Tablo5.30’da, MLP sınıflandırma yönteminin eğitim ve test verisi şeklinde ayrıştırılarak, BoW, DBoW ve DM öznitelik çıkarım modelleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. Bu modelde, MLP sınıflandırma algoritmasının BoW öznitelik çıkarım yöntemiyle %88,3 doğruluk oranı ile en başarılı sonucu elde ettiği görülmektedir.

Tablo 5.30. MLP sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

MLP	A	P	R	F
BoW	0.883	0.884	0.883	0.883
DBoW	0.801	0.803	0.801	0.801
DM	0.639	0.672	0.639	0.642

Kullanılan ikinci veri ayrıştırma yönteminde, 10 kat çapraz doğrulama işleminde veri seti 10 eşit parçaya bölünmüştür. Elde edilen bu 10 parçanın 1 parçası test, 9 parçası ise modelin eğitiminde kullanılmıştır. Tablo 5.31’de, SVM sınıflandırma algoritmasının 10 kat çapraz doğrulama veri ayrıştırma yönteminde BoW, DBoW ve DM öznitelik çıkarım teknikleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. SVM modelinde 10 kat çapraz doğrulama yapılması durumunda, Doc2Vec tekniklerinden biri olan DBoW yöntemiyle % 75’lik doğruluk oranı elde edilmiştir.

Tablo 5.31. SVM sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

SVM	A	P	R	F
BoW	0.691	0.544	0.691	0.592
DBoW	0.759	0.801	0.759	0.757
DM	0.405	0.576	0.407	0.314

Tablo5.32’de NB sınıflandırma algoritmasının 10 kat çapraz doğrulama veri ayrıştırma yönteminde BoW, DBoW ve DM öznitelik çıkarım teknikleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. NB modelinde 10 kat çapraz doğrulama yapılması durumunda, BoW yöntemiyle % 89,2’lik doğruluk oranı elde edilmiştir.

Tablo 5.32. NB sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

NB	A	P	R	F
BoW	0.892	0.899	0.892	0.893
DBoW	0.795	0.806	0.795	0.795
DM	0.599	0.659	0.599	0.597

Tablo5.33’te MLP sınıflandırma algoritmasının 10 kat çapraz doğrulama veri ayrıştırma yönteminde BoW, DBoW ve DM öznitelik çıkarım teknikleriyle elde edilen doğruluk (A), keskinlik (P), duyarlılık(R) ve F-metrik değerleri gösterilmiştir. MLP modelinde 10 kat çapraz doğrulama yapılması durumunda, BoW yöntemiyle % 87,3’lük doğruluk oranı elde edilmiştir.

Tablo 5.33. MLP sınıflandırmasının Öznitelik çıkarımlarına göre metrik değerleri

MLP	A	P	R	F
BoW	0.873	0.881	0.873	0.874
DBoW	0.793	0.803	0.793	0.793
DM	0.628	0.647	0.628	0.628

İlgili problem için kullanılan sınıflandırma algoritmalarının 3 farklı öznitelik çıkarma yöntemi ile veri setinin eğitim ve test olarak ayrıştırılması durumunda elde edilen sonuçlar Tablo 5.34’te karşılaştırılmıştır. Bu yöntemde en iyi sonucu, MLP sınıflandırma algoritmasının kullanıldığı BoW modeli vermiştir.

Tablo 5.34. Modelleme ve Sınıflandırma Yöntemlerinin Karşılaştırılması

A	BoW	DBoW	DM
SVM	0.696	0.481	0.310
NB	0.880	0.756	0.549
MLP	0.883	0.801	0.639

İlgili problem için kullanılan sınıflandırma algoritmalarının 3 farklı öznitelik çıkarma yöntemi ile 10 kat çapraz doğrulama yapılması durumunda elde edilen sonuçlar Tablo 5.35'te karşılaştırılmıştır. Bu yöntemde en iyi sonucu, %89,2 doğruluk oranı ile NB sınıflandırma algoritmasının kullanıldığı BoW modeli vermiştir.

Tablo 5.35. Modelleme ve Sınıflandırma Yöntemlerinin Karşılaştırılması

A	BoW	DBoW	DM
SVM	0.691	0.759	0.405
NB	0.892	0.795	0.599
MLP	0.873	0.793	0.628

Her iki veri ayrıştırma yönteminde de en iyi sonuçlar, BoW öznitelik çıkarım tekniğinde elde edilmiştir. İki veri ayrıştırma yönteminde de DBoW öznitelik çıkarım tekniği, BoW tekniğine yakın sonuçlar vermiştir. DBoW tekniği, müşteri taleplerini anlamsal olarak değerlendirmekte ve buna göre modellemektedir. DBoW öznitelik çıkarım tekniğinde düşük sonuçlar alınmasının temel sebebinin veri setinin problemi modelleyebilmesi için yeterli büyüklükte olmamasından kaynaklandığı ön görülmektedir.

Sonraki çalışmalarda müşteri taleplerinin yer aldığı veri seti artırılarak DBoW öznitelik çıkarım tekniği ile problem yeniden ele alınabilir. Ayrıca bu çalışmada hedeflenen 3 farklı sınıf firmaların organizasyon yapısına göre artırılarak sonuçlar yeniden değerlendirilebilir.

Sonuç olarak müşteri taleplerinin organizasyonel iş akışlarına uygun olarak sınıflandırılabilmesi problemi için önerilen modellerin kullanılabilir/uygulanabilir olduğu görülmektedir.

KAYNAKÇA

- Acuner, Şebnem Akın (2003). Müşteri İlişkilerinde Hareket Noktası, "Müşteri Memnuniyeti ve Ölçümü.", *Milli Produktivite Merkezi Yayınları* , 655.
- Akdın, Funda (2006). "Lojistik Rekabette Müşteri İlişkileri Yönetiminin Yeri ve Önemi." *Diss. DEÜ Sosyal Bilimleri Enstitüsü.*
- Aksu, Akın (2004). "Otel İşletmelerinde Müşteri Sadakatine Genel Bir Bakış.", *Standart Dergisi* 43.507, 70-75
- Alabay, Mehmet Nurettin (2012). "Müşteri Şikâyetleri Yönetimi.", *Uluslararası Yönetim İktisat ve İşletme dergisi* 8.16, 137-157
- Alabay, Mehmet Nurettin (2008). "*CRM Rekabet Stratejisi Olarak Müşteri İlişkileri Yönetimi.*", İlke Yayınevi, Ankara.
- Alabay, Mehmet Nurettin (2007). "Müşteri İlişkileri Yönetimi (MİY): Turizm Sektörü Üzerinde Bir Uygulama.", *Gazi Üniversitesi Sosyal Bilimler Enstitüsü*, Yayınlanmamış Doktora Tezi, Ankara.
- Arslan Halil, Kaynar Oğuz ve Yüksek Ahmet Gürkan (2015). "Kurumsal Kolektif Süreçler için e-posta İletilerinden Görev Keşfi ve Gerçek Zamanlı Görev Yönetim Sisteminin Geliştirilmesi.", *Bilişim Teknolojileri Dergisi* 10.4, 381-388.
- Ayata Değer, Saraçlar Murat ve Özgür Arzucan (2017). "Turkish Tweet Sentiment Analysis with Word Embedding and Machine Learning.", *25th signal processing and Communications Applications Conference (SIU). IEEE.*
- Barış, Gülfidan (2008). "*Kusursuz Müşteri Memnuniyeti için Şikâyet Yönetimi.*", İstanbul: Mediacat.
- Benli, Yasemin Keskin (2002). "Finansal Başarısızlığın Tahmininde Yapay Sinir Ağı Kullanımı ve İMKB’de Bir Uygulama.", *Muhasebe Bilim Dünyası Dergisi* 4.4, 17-30.

- Biçer Gülkibar ve Düztepe Şerafettin (2003). "Yetkinlikler ve Yetkinliklerin İşletmeler Açısından Önemi.", *Journal of aeronautics and Space Technologies* 1.2, 13-20.
- Blodgett Jeffrey G., Wakefield Kirk L., and Barnes James H., (1995). The Effects of Customer Service on Consumer Complaining Behavior.", *Journal of Services Marketing* 9.4, 31-42.
- Büyükkınacı, Muhammed (2019). Word2vec Nedir? (türkçe), <https://medium.com/@mubuyuk51/word2vec-nedir-t%c3%bcrk%c3%a7e-f0cfab20d3ae> , (07-01-2018/20-03-2019)
- Caputo Graziella M., Bastos Valeria M. And Ebecken Nelson F.F., (2006). "Using Text Mining to Understand The Call Center Customers' Claims.", *WIT Transactions on Information and- Communication Technologies*37.
- Chang Che-Wei, Lin Chin-Tsai and Wang Lian-Qing, (2009). "Mining The Text Information to Optimizing The Customer Relationship Management", *Expert Systems WITH Applications* 36.2, 1433-1443
- Cengiz, Ilgın (2013). "Müşteri Memnuniyetini Oluşturan Faktörlerin Müşteri Sadakatine Etkisinin Lojistik Regresyon Analizi ile İncelenmesi.", *Afyon Kocatepe Üniversitesi Sosyal Bilimler Enstitüsü*, Afyon.
- Coussement Kristof and Van Den Poel Dirk, (2008). "Improving Customer Complaint Management by Automatic Email Classification Using Linguistic Style Features As Predictors.", *Decision Support Systems* 44.4, 870-882.
- Çalış Kübra, Gazdağı Oya ve Yıldız Oktay (2013). "Reklam İçerikli Epostaların Metin Madenciliği Yöntemleri İle Otomatik Tespiti.", *Bilişim Teknolojileri Dergisi* 6.1, 1-7.
- Çetiner, Volkan (2012). "Müşteri Sadakati Oluşturmada Müşteri İlişkileri Yönetiminin Etkisi ve Önemi.", *Niğde Üniversitesi, Sosyal Bilimler Enstitüsü*, Yayınlanmamış Yüksek Lisans Tezi, Niğde.

- Çoban Önder ve Karabey Işıl (2017). "Music Genre Classification With Word And Document Vectors." *25th Signal Processing and Communications Applications Conference (SIU)*, IEEE
- Çomak, Emre (2004). "Destek Vektör Makineleri Çoklu Sınıf Problemleri İçin Çözüm Önerileri.", *Diss. Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya.
- Çuhadar Murat ve Kayacan Cengiz (2005). "Yapay Sinir Ağları Kullanılarak Konaklama İşletmelerinde Doluluk Oranı Tahmini: Türkiye'deki Konaklama İşletmeleri Üzerine Bir Deneme.", *Anatolia: Turizm Arastirmalari Dergisi 16.1*, 24-30
- Dabbağoğlu, Kadir (2009). "İç Kontrol Sistemi.", *Journal of Qafqaz University 26*, 109-115.
- Demir Halis ve Tarhan Okan (2011). "Teknoloji, Örgüt Yapısı Ve Performans Arasındaki İlişkiler Üzerine Bir Araştırma." *Doğuş Üniversitesi Dergisi 10.1*, 57-72.
- Ersöz Süleyman, Yaman Nevra, ve Birgören Burak (2018). "Müşteri İlişkileri Yönetiminde Verilerin Yapay Sinir Ağları İle Modellenmesi Ve Analizi.", *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi 23.4*.
- Esener, Harun (2006). "Hizmet Yönetimi Sistemi (HYS)."
- Feldman Ronen and Sanger James (2007). *"The Text Mining Handbook Advanced Approaches In Analyzing Unstructured Data."*, Cambridge university press.
- Fılıp, Alina (2013). "Complaint Management: A Customer Satisfaction Learning Process." *Procedia-Social And Behavioral Sciences 93*, 271-275.
- Gamon, Michael (2004). "Sentiment Classification On Customer Feedback Data: Noisy Data, Large Feature Vectors, and The Role Of Linguistic Analysis.", *Proceedings of the 20th International Conference On Computational Linguistics*. Association For Computational Linguistics, 841

- González Luz Minerva vd., (2005). "Knowledge Management-Centric Help Desk: Specification And Performance Evaluation.", *Decision Support Systems* 40.2, 389-405.
- He Wu, ZHA Shenghua and Li Ling, (2013). "Social Media Competitive Analysis And Text Mining: A Case Study İn The Pizza İndustry." *International Journal of Information Management* 33.3, 464-472.
- Hastie Trevor vd., (2005). "The Elements of Statistical Learning: Data Mining, Inference and Prediction." *The Mathematical Intelligencer* 27.2, 83-85.
- Ittoo Ashwin Ravi, Zhang Yiyang, ve Jiao Jianxin, (2006). "A Text Mining-Based Recommendation System For Customer Decision Making İn Online Product Customization." *2006 IEEE International Conference On Management Of Innovation And technology*. Vol. 1. IEEE, 473-477
- İlhan, Sevinç vd., (2008). "Metin Madenciliği İle Soru Cevaplama Sistemi." *Elektronik ve Bilgisayar Mühendisliği Sempozyumu (ELECO)*, Bursa , 26-30.
- Karadağ Anıl ve Takçı Hidayet (2010). "Metin Madenciliği İle Benzer Haber Tespiti." *Akademik Bilişim*.
- Karakaya, Ertuğrul (2007). "Rekabet Avantajı Sağlaması Açısından Müşteri İlişkileri Yönetimi Ve Bir Süpermarket Uygulaması.", *Kırıkkale Üniversitesi Sosyal Bilimler Enstitüsü*, Kırıkkale.
- Karasoy Onur ve Ballı Serkan (2017). "Classification Turkish Sms With Deep Learning Tool Word2vec." *International Conference On Computer Science and Engineering (UBMK)*, IEEE, 294-297.
- Kavzoğlu Taşkın ve Çölkesen İsmail (2010). "Destek Vektör Makineleri İle Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi." *Harita Dergisi* 144.7, 73-82.
- Kimothi Dhananjay vd., (2016). "Distributed Representations For Biological Sequence Analysis." *Arxiv Preprint Arxiv:1608.05949*, 2016

- Kotler Philip ve Armstrong Gary (2010). "*Principles of Marketing.*" Pearson Education.
- Kırliođlu Hilmi ve Ceyhan İsmail Fatih (2014). "Mali Tablo Denetiminde Ön Analitik İnceleme Tekniđi Olarak Veri Madenciliđinin Kullanımı: Borsa İstanbul Uygulaması." *Akademik Yaklaşımlar Dergisi* 5.1, 13-36.
- Kurtz Albert K. (1948) "A Research Test Of The Rorschach Test." *Personnel Psychology*, .
- Krus David J., and Fuller Ellen A. (1982). "Computer Assisted Multicrossvalidation In Regression Analysis." *Educational and Psychological Measurement* 42.1, 1982, 187-193.
- Mikolov Tomas vd., (2013). "Efficient Estimation Of Word Representations In Vector Space.", *Arxiv Preprint Arxiv:1301.3781*,
- Mikolov Tomas vd. (2013). "Distributed Representations Of Words And Phrases And Their Compositionality." *Advances In Neural Information Processing Systems*.
- Odabaşı, Yavuz (2000). "Satıřta Ve Pazarlamada Müřteri İliřkileri Yönetimi." *Sistem Yayıncılık*, İstanbul.
- Ordenes Francisco Villarroel vd., (2014). "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach." *Journal of Service Research* 17.3, 278-295.
- Özgener, řevki (2003). "Büyüme Sürecindeki Kobi'lerin Yönetim Ve Organizasyon Sorunları: Nevşehir Un Sanayii Örneđi." *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi* 20, 137-161.
- Öztemel, Ercan (2012). "*Yapay Sinir Ağları.*" Papatya yayıncılık, İstanbul.
- Özyirmidokuz Esra Kahya ve Özyirmidokuz Mustafa Hakan (2014). "Analyzing Customer Complaints: A Web Text Mining Application." *International Conference On Education And Social Sciences*, 507-515

- Özyirmidokuz, Esra Kahya (2016). "Analyzing Unstructured Facebook Social Network Data Through Web Text Mining: A Study Of Online Shopping Firms İn Turkey." *Information Development* 32.1, 70-80.
- Özyurt Özcan ve Köse Cemal (2010). "Chat Mining: Automatically Determination Of Chat Conversations' Topic İn Turkish Text Based Chat Mediums." *Expert Systems With Applications* 37.12, 8705-8710.
- Refaeilzadeh Payam, Tang Lei and Liu Huan. (2009). "Cross Validation, Encyclopedia Of Database Systems (edbs).", *Encyclopedia of database systems*, 532-538..
- Sönmez, Derya (2007). Müşteri İlişkileri Yönetiminde Bilişim Teknolojilerinin Müşteri Odaklılığını Sağlamadaki Rolü. Diss. Yüksek Lisans Tezi, Gazi Üniversitesi, Ankara.
- Stiebellehner Simon, Wang Jun and Yuan Shuai (2017). "Learning Continuous User Representations Through Hybrid Filtering With Doc2vec." *Arxiv Preprint Arxiv:1801.00215*
- Stoica Eduard Alexandru ve Özyirmidokuz Esra Kahya (2015). "Mining Customer Feedback Documents." *International Journal Of Knowledge Engineering* 1.1, 68-71.
- Şahin, Gürkan (2017). "Turkish Document Classification Based On Word2vec And Svm Classifier." *25th Signal Processing And Communications Applications Conference (SIU)*. IEEE, 2017, 1-4
- Şimşek Mehmet Şerif ve Çelik Adnan (2002). "Yönetim Ve Organizasyon (7. Baskı)." Günay Ofset, Konya.
- Taşkın, Erdoğan (2000). "Müşteri İlişkileri Eğitimi", Papatya Yayıncılık, İstanbul.
- Taşpınar, Hakan (2005). "Bilişim Altyapısıyla CRM:(Customer Relationship Management) Müşteri İlişkileri Yönetimi", Seçkin Yayıncılık.
- Tezgider Murat, Yıldız Beytullah ve Aydın Galip (2018). "Improving Word Representation By Tuning Word2vec Parameters With Deep Learning

Model." *International Conference On Artificial Intelligence And Data Processing (IDAP)*, IEEE, 2018,1-7

Temel Gülhan, Erdoğan Semra ve Ankaralı Handan (2012). "Sınıflama Modelinin Performansını Değerlendirmede Yeniden Örnekleme Yöntemlerinin Kullanımı." *Bilişim Teknolojileri Dergisi* 5.3, 1-8.

Thorleucher Dirk and Van Den Poel Dirk (2012). "Predicting E-Commerce Company Success By Mining The Text Of Its Publicly-Accessible Website." *Expert Systems With Applications* 39.17, 13026-13034.

Türker Ali ve Özaltın Gülay (2010). "Konaklama İşletmelerinde Müşteri İlişkileri Yönetimi: İzmir İli Örneği." *Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 25, 81-104

Uğur Aybars ve Kınacı Ahmet Cumhur (2006). "Yapay Zeka Teknikleri Ve Yapay Sinir Ağları Kullanılarak Web Sayfalarının Sınıflandırılması." *XI. Türkiye'de İnternet Konferansı Bildirileri*, 362.

Ur-Rahman Nadeem and Harding Jennifer A., (2012). "Textual Data Mining For Industrial Knowledge Management And Text Classification: A Business Oriented Approach." *Expert Systems With Applications* 39.5, 4729-4739.

Ünver, Ceyda (2010). "Müşteri İlişkileri Yönetimi: Müşteri Şikâyetleri Ve İnternet." *Sosyal Bilimler Enstitüsü Halkla İlişkiler Ve Tanıtım Anabilim Dalı*, Yayınlanmamış Yüksek Lisans Tezi, Ankara.

Vapnik, Vladimir (2013). "*The Nature Of Statistical Learning Theory*", Springer Science & Business Media.

Wilson Alan vd., (2016). "*Services Marketing: Integrating Customer Focus Across The Firm*"

Weng Sung-Shun ve Liu Chih-Kai (2004). "Using Text Classification And Multiple Concepts To Answer E-Mails.", *Expert Systems with Applications* 26.4, 529-543.

Whitley Stephen vd., "Task Management System And Method.", *U.S. Patent Application no. 14/146,552*, 2014

Yuan Yanhong vd (2014). "A New Study Based On Word2vec And Cluster For Document Categorization.", *Journal Of Computational Information Systems* 10.21, 9301-9308.

Yıldırım Savaş ve Yıldız Tuğba (2018). "Türkçe İçin Karşılaştırmalı Metin Sınıflandırma Analizi.", *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi* 24.5, 879-886.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Sümeyye ŞAHİN
Uyruğu : T.C
Doğum Tarihi ve Yeri : 01.01.1993 BASKİL
e-posta : sumeyyekayaokay@gmail.com

EĞİTİM

Derece	Kurum	Mezuniyet Yılı
Lisans	Fırat Üniversitesi	2014
Yüksek Lisans	Sivas Cumhuriyet Üniversitesi	2019

İŞ TECRÜBESİ

Tarih	Kurum	Görev
2014 - 2017	Detay Danışmanlık Ltd. (Detaysoft)	Web Uzmanı
2018 - halen	Hakan Bilgi Teknolojileri	Proje Yöneticisi

YABANCI DİL BİLGİSİ

İngilizce YÖKDİL (67.5)