

**SURİYELİ GÖÇMENLERİN TÜRK HALKI İLE
KAYNAŞTIRILMASI ÇALIŞMALARI ÜZERİNE VERİ
MADENCİLİĞİ ARAŞTIRMASI**

**DATA MINING RESEARCH ON STUDIES OF THE
INTEGRATION OF SYRIAN IMMIGRANTS WITH THE
TURKISH PEOPLE**

DİDEM YILMAZ

Dr.Öğr.Üyesi İbrahim ZOR

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı İçin Öngördüğü

Yüksek Lisans Tezi Olarak Hazırlanmıştır.

2020

Didem YILMAZ'ın hazırladığı “**Suriyeli Göçmenlerin Türk Halkı İle Kaynaştırılması Çalışmaları Üzerine Veri Madenciliği Araştırması**” adlı bu çalışma aşağıdaki jüri tarafından **İSTATİSTİK ANABİLİM DALI**'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof. Dr. XXXXXX YYYYYYY

Başkan

.....

Doç. Dr. XXXXXX YYYYYYY

Danışman

.....

Prof. Dr. XXXXX YYYYYYY

Üye

.....

Doç. Dr. XXXXX YYYYYYY

Üye

.....

Dr. Öğr. Üyesi XXXXX YYYYYYY

Üye

.....

Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından **YÜKSEK LİSANS TEZİ** olarak / /..... tarihinde onaylanmıştır.

Prof. Dr. Salih Bülent ALTEN

Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

___ / ___ / 2020

DİDEM YILMAZ

YAYINLANMA FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan “*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*” kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H. Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.
- Enstitü / Fakülte yönetim kurulu gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ay ertelenmiştir.
- Tezim ile ilgili gizlilik kararı verilmiştir.

..... / /

(İmza)

DİDEM YILMAZ

ÖZET

SURİYELİ GÖÇMENLERİN TÜRK HALKI İLE KAYNAŞTIRILMASI ÇALIŞMALARI ÜZERİNE VERİ MADENCİLİĞİ ARAŞTIRMASI

DİDEM YILMAZ

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Dr.Öğr.Üyesi İbrahim Zor

Ağustos 2020

Türkiye oldukça fazla göç alan bir ülkedir. Özellikle Suriye krizinden sonra yaklaşık 3.5 milyon Suriyeli ülkemize göç etmek zorunda kalmıştır. Devlet, insani yardım dernekleri ve sivil toplum kuruluşları yardımları ile ülkemizdeki Suriyeli göçmenler Türk halkı ile bir arada yaşamaya başlamıştır. Suriyeli göçmenlerin ülkemizde ihtiyaç duyabilecekleri hizmetlerin karşılanabilmesi ve iki farklı kültürün kaynaştırılması amacıyla Türk Kızılay tarafından Toplum Merkezleri kurulmuştur. Toplum Merkezleri sayesinde birçok Suriyeli göçmen eğitim, psikolojik destek vb. hizmetlere ulaşabilmektedir. Bu çalışma, Toplum Merkezleri tarafından verilmekte olan hizmetlerin yararlanıcılar üzerindeki etkisini araştırmak amacıyla yapılmaktadır. Çalışmada yöntem olarak veri madenciliği uygulaması seçilmiştir. Veri madenciliği, çalışmada kullanılacak veri seti üzerinden tahminlerde bulunulacak olması ve alınan sonuçların kolay yorumlanabilmesi amacıyla seçilmiştir. Araştırmacılar tarafından kolay ulaşılabilir olması, açık kaynak kodlu birçok yazılım olması, internet aracılığıyla yazılımların kolayca öğrenilebilmesi gibi sebepler de bu seçimde etkili olmuştur. Veri madenciliği uygulamasında KNIME yazılımı kullanılmıştır. Bu yazılımın seçiminde de veri madenciliği yönteminin seçilme sebepleri ile aynı sebepler etkili olmuştur.

Çalışmada karar ağaçları ve destek vektör makineleri modelleri uygulanmıştır. Karar ağaçları modeline ilişkin sonuçlar farklı K-katlı çapraz doğrulama durumları ve ayırma kriterlerine aittir. Bu sonuçlara göre, karar ağacı modeline ait yapılan tüm denemelerde elde edilen doğruluk yüzdeleri %99'un üzerindedir. Aynı şekilde destek vektör makineleri modelinde de farklı K-katlı çapraz doğrulama ve çekirdek fonksiyonlara ait sonuçlar elde edilmiştir. Bu sonuçlar doğrultusunda, yapılan tüm denemelerde doğru sınıflandırma yüzdeleri %92 ve üzerindedir.

Bu çalışma kapsamında Toplum Merkezleri'nde verilen eğitim hizmetlerinin yararlanıcılar üzerindeki etkisi araştırılmıştır. Yararlanıcıların çalışma durumları, hangi kurslara katılım sağladıkları, hangi Toplum Merkezleri'nden hizmet aldıkları vb. değişkenlerin bir araya gelerek eğitim aldıkları kurslar sonucu istihdam edildikleri alan üzerinde etkili olup olmadığı sorusu araştırılmaktadır.

Anahtar Kelimeler: Suriye Krizi, Veri Madenciliği, Karar Ağaçları, Destek Vektör Makineleri

ABSTRACT

DATA MINING RESEARCH ON STUDIES OF THE INTEGRATION OF SYRIAN IMMIGRANTS WITH THE TURKISH PEOPLE

DİDEM YILMAZ

M.Sc Thesis, Department of Statistics

Supervisor: Asst. Prof. Dr. İbrahim ZOR

August 2020

Turkey is a country rather than immigration. Especially after the Syrian crisis, approximately 3.5 million Syrian had to migrate to our country. Syrian immigrants in our country have started to live together with the Turkish people with the aid of state, humanitarian organizations and non-governmental organizations. Community Centers were established by the Turkish Red Crescent in order to meet the services that Syrian migrants may need in our country and to integrate two different cultures. Thanks to Community Centers, many Syrian immigrants receive education, psychological support etc. can access services. This study is carried out to investigate the effects of the services provided by Community Centers on the beneficiaries. Data mining application was chosen as the method in study. Data mining was chosen in order to make predictions on the data set to be used in the study and to interpret the results easily. The reasons such as being easily accessible by researchers, having many open source software, and being able to learn software easily through the internet have also been effective in this choice. KNIME software was used in data mining application. The same reasons for choosing the data mining method were effective in the selection of this software.

Decision trees and support vector machines models were applied in the study. The results of the decision trees model belong to different K-fold cross validation cases and separation criteria. According to these results, the accuracy percentages obtained in all trials of the decision tree model are over 99%. Similarly, results of different K-fold cross validation and functions were obtained in the support vector machine model. These articles have correct classification percentages of 92% and above in the experiments.

Within the scope of this study, the effect of educational services provided in Community Centers on the beneficiaries was investigated. Employment status of the beneficiaries, which courses they attended, from which Community Centers they received service, etc. The question of whether the variables are effective on the field in which they are employed as a result of the courses they have taken is investigated.

Keywords: Syrian Crisis, Data Mining, Decision Trees, Support Vektor Machine

TEŐEKKÜR

Lisans ve Yüksek lisans eğitimim boyunca desteğini hiç eksik etmeyen değerli öğretmenim Dr.Öğr.Üyesi İbrahim ZOR'a,

Türk Kızılay çatısı altında bilgi paylaşımından, veri seti paylaşımına her türlü desteğini sağlayan değerli yöneticilerime, Araş.Gör.Selman Salim KESGİN'e ve tüm Toplum Temelli Göç Programları Raporlama ve İzleme Birimi ekibine,

Lisans ve Yüksek lisans hayatımda arkamda olduklarını hep hissettiğim aileme,

Teşekkür ederim.

Didem YILMAZ

Ağustos 2020, Ankara

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR.....	v
İÇİNDEKİLER	vi
ŞEKİLLER.....	viii
TABLolar	ix
SİMGELER VE KISALTMALAR.....	x
1.GİRİŞ.....	1
2.VERİ MADENCİLİĞİ	4
2.1. Veri Madenciliği Tanımı	4
2.2. Veri Madenciliği Uygulama Alanları.....	4
2.3. Veri Madenciliğinde Kullanılan Yazılımlar	5
2.4. Çalışmada Kullanılan Modeller	5
2.4.1. Karar Ağaçları.....	5
2.4.2. Karar Ağacı Algoritmaları.....	6
2.4.2.1. CART (Classification and Regression Tree) Algoritması	7
2.4.2.2. CHAID (Chi- Squared Automatic Interaction Detector) Algoritması	8
2.4.2.3. QUEST (Quick, Unbiased, Efficient, Statistical Tree) Algoritması	9
2.4.2.4. C4.5 ve C5.0 Algoritmaları.....	11
2.4.3. Karar Ağaçlarında Durdurma ve Budama Kuralları	11
2.4.4. Karar Ağaçlarının Avantaj ve Dezavantajları.....	12
2.4.5. Karar Ağaçlarının Uygulama Alanları	12
2.4.6. Sınıflandırma İşleminin Doğruluğu	13
2.4.7. Sınıflandırma Performanslarının Değerlendirilmesi	14
2.4.7.1. Karışıklık Tabloları	14
2.4.7.2. ROC Eğrileri	15
2.5. Destek Vektör Makineleri	17
2.5.1. Doğrusal Destek Vektör Makineleri	17
2.5.2. Doğrusal Olarak Ayrılabilen Verilerde DVM	18

2.5.3.Doğrusal Olarak Ayrılamayan Verilerde DVM.....	20
3. TOPLUM MERKEZLERİ.....	22
3.1.Toplum Merkezleri Nedir, Nasıl Çalışır.....	25
3.2. Toplum Merkezleri Tarafından Verilen Hizmetler	25
3.3.Uygulanan Programlar	26
3.3.1.Geçim Kaynağını Geliştirme Programı	26
3.3.2.Koruma Programı.....	28
3.3.3.Psikososyal Sağlık ve Destek Programı	29
3.3.4.Aile Bağlarının Yeniden Tesisi Programı	30
3.3.5.Sosyal Uyum Programı.....	32
4.UYGULAMA.....	34
4.1.Çalışmada Yer Alan Değişkenler.....	34
4.2. Veri Madenciliği Modellemeleri.....	43
4.2.1.Karar Ağaçları.....	43
4.2.2.Destek Vektör Makineleri.....	49
5.SONUÇ.....	54
ÖZGEÇMİŞ	61

ŞEKİLLER

Şekil 1: Karar Ağacı Yapısı	6
Şekil 2: ROC Eğrisi	16
Şekil 3: Doğrusal Olarak Ayrılabilen Durum- Optimum Hiper-Düzlem	18
Şekil 4: Belirli Bir Hata İle Doğrusal Ayrılma Durumu	19
Şekil 5: Doğrusal Olmayan Haritalama Yaklaşımı Örneği	21
Şekil 6: Ülkemizde Şehirlere Göre Suriyeli Göçmen Yoğunlukları	23
Şekil 7: Türk Kızılay Toplum Merkezleri'ne Ait İstatistikler	24
Şekil 8: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 1	26
Şekil 9: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 2	28
Şekil 10: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 3	29
Şekil 11: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 4	31
Şekil 12: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 5	32
Şekil 13: Karar Ağaçları Model Akış Şeması	44
Şekil 14: Karar Ağacı Diyagramı 1	45
Şekil 15: K=5 Katlı Çapraz Doğrulama ve Gini İndeksi Kullanıldığı Durum İçin ROC Eğrisi	46
Şekil 16: Karar Ağacı Diyagramı 2	47
Şekil 17: K=5 Katlı Çapraz Doğrulama ve Kazanç Oranı Kullanıldığı Durum İçin ROC Eğrisi	48
Şekil 18: Destek Vektör Makineleri Model Akış Şeması	49

TABLolar

Tablo 1:Veri Madenciliđi alıřmalarında Kullanılan Yazılımlardan Bazıları	5
Tablo 2:Model Deđerlendirmesi	13
Tablo 3:Karışıklık Matrisi	15
Tablo 4:Destek Vektör Makineleri'nde Kullanılan Temel ekirdek (Kernel) Fonksiyonlar ...	21
Tablo 5:Cinsiyet Deđerkenine İliřkin Sıklık Tablosu.....	34
Tablo 6:Toplum Merkezi Deđerkenine İliřkin Sıklık Tablosu	35
Tablo 7:Kurs Adı Deđerkenine İliřkin Sıklık Tablosu	35
Tablo 8:Eđitim Seviyesi Deđerkenine İliřkin Sıklık Tablosu.....	36
Tablo 9:Bildiđi Dil Deđerkenine İliřkin Sıklık Tablosu.....	37
Tablo 10:Sađlık ve Engel Durumu Deđerkenine İliřkin Sıklık Tablosu	37
Tablo 11:Kurs Ücret Durumu Deđerkenine İliřkin Sıklık Tablosu.....	38
Tablo 12:Kurs Bırakma Durumu Deđerkenine İliřkin Sıklık Tablosu	38
Tablo 13:alıřma Durumu Deđerkenine İliřkin Sıklık Tablosu	39
Tablo 14:Gelir Türü Deđerkenine İliřkin Sıklık Tablosu	39
Tablo 15:Medeni Hali Deđerkenine İliřkin Sıklık Tablosu	40
Tablo 16:Sertifika Alma Durumu Deđerkenine İliřkin Sıklık Tablosu	40
Tablo 17:İstihdam Alanı ile Kurs İlgisi Durumu Deđerkenine İliřkin Sıklık Tablosu	41
Tablo 18:Yař Deđerkenine İliřkin Sıklık Tablosu.....	41
Tablo 19:Cinsiyet*İstihdam Alanı ile Kurs İlgisi Durumu Deđerkenlerine Ait apraz Tablo ...	42
Tablo 20:Karar Ađaçları Model Sonuları	44
Tablo 21:Destek Vektör Makineleri Model Sonuları (ekirdek Fonksiyon: Polinom)	50
Tablo 22:Destek Vektör Makineleri Model Sonuları (ekirdek Fonksiyon: RBF)	52

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar açıklamaları ile birlikte aşağıdaki sunulmaktadır.

Simgeler	Açıklama
%	Yüzde
Kısaltmalar	Açıklama
CART	Classification and Regression Tree
QUEST	Quick, Unbiased, Efficient Statistical Tree
CHAID	Chi-Square Automatic Interaction Detector
TP	True Positive
FP	False Positive
TN	True Negative
FN	False Negative
ROC	Receiver Operating Characteristic Curve

1.GİRİŞ

Türkiye, tarih boyunca konumu itibariyle dikkatleri üzerine çeken bir ülke olmuştur. Gerek kıtaları birbirine bağlayan bir köprü oluşu, gerekse bünyesinde barındırdığı sosyo-ekonomik, coğrafik vb. özellikleri ile hem vatandaşlarına hem de farklı ülkelerin vatandaşlarına pek çok fırsat sunan bir ülke konumundadır. 1960'lı yıllarda başlayan ve 1970'lerde de devam eden göç olgusu bağlamında ağırlıklı olarak Türkiye'den, başta Almanya olmak üzere, Batı Avrupa ülkelerine ciddi bir göç dalgası yaşanmıştır. Avrupa Birliği'ne geçiş yapmak için Asya ülkelerinden ülkemize yönelen göç dalgası içerisinde yer alan başlıca ülkeler: Afganistan, Bangladeş, Irak, İran ve Pakistan'dır [1]. Türkiye'ye 2018 yılında gelen yabancı uyruklu nüfusun içerisinde ilk sırayı %23,6 ile Irak vatandaşları almaktadır. Bu sırasıyla %9,6 ile Afganistan, %8,4 ile Suriye, %7,5 ile Türkmenistan ve %6,8 ile İran vatandaşları izlemektedir [2].

Bu çalışmada özellikle Suriyeli halkın, Türkiye'ye yaptığı göç hareketi üzerinde durulacaktır. 2011 yılında Suriye'de meydana gelen siyasi olaylar ve iç karışıklıklar, yerel halkın kitleler halinde ülkelerini terk edip, yeni yaşam alanları aramalarına sebep olmuştur. Suriyeli yerel halk, öncelikli olarak komşu ülkelere göç etmeyi tercih etmiştir. Başta Türkiye olmak üzere, Lübnan, Ürdün gibi ülkeler 2011 yılından itibaren sürekli olarak Suriyeli göçmenlere kapılarını açmaktadır. Suriyeli göçmenler, genel olarak ülkemizde Hatay, Şanlıurfa, Gaziantep gibi Suriye ile sınırı olan şehirlerimize göç etmişlerdir. Türkiye'de kayıtlı Suriyeli göçmen sayısı 23 Eylül 2020 tarihi itibariyle bir önceki aya göre 12 bin 84 kişi artarak 3 milyon 621 bin 968 kişi olmuştur. Suriyeli erkek göçmenler, toplam Suriyeli göçmenlerin %53,8'ini oluşturmaktadır. Göç idaresinin 23 Eylül 2020 tarihli verilerine göre en çok Suriyeli göçmenin yaşadığı şehir İstanbul'dur. Sonrasında ise sırasıyla Gaziantep, Hatay, Şanlıurfa illeri gelmektedir. Suriyeli göçmenlerin en az ikamet ettikleri şehir ise 25 kişilik Suriyeli nüfus ile Bayburt ilimizdir [3].

Göçmenler açısından göç ettikleri ülke, kültürleşmeye yönelik gösterdikleri eğilim, göçmenlerin kişisel özellikleri gibi çeşitli etkenler uyum sürecinin belirlenmesinde etkili olabilmektedir. Nitekim gidilen ülkede azınlık olarak yaşayan göçmenlerin gittikleri kültürü içselleştirerek kendi hayatlarına adapte edebilmeleri sosyo-kültürel uyumlarını kolaylaştırmaktadır [4].

Uyum ve sosyalleşme boyutuyla ele alındığında günümüzde çoğu ülke göçmen ailelerine karşı olumsuz bir tutum izlemekte, diğer taraftan göçmenlerle ilgili uyguladıkları prosedürleri zorlaştırmaya çalışmaktadır. Göçmenlerin ülkelerine girişleri konusunda uluslararası hukuki denetimleri arttıran yabancı ülkelerin aksine Türkiye açık kapı politikası izleyerek başta Suriyeliler olmak üzere pek çok göçmene ev sahipliği yapmaktadır [5].

Tam da bu noktada devreye Türk Kızılay'ı girmektedir. Kurulduğu 1868 yılından bu yana toplumsal dayanışmayı sağlamak, sosyal refahın gelişmesine katkıda bulunmak, yoksul ve muhtaç insanlara barınma, beslenme ve sağlık yardımı ulaştırmak için önemli görevler üstlenen Türk Kızılay, kan, afet, uluslararası yardım, göç ve mülteci hizmetleri, sosyal hizmetler, sağlık, ilk yardım, eğitim, gençlik ve mineralli su işletmeleri alanlarında faaliyet sunmaktadır. Özellikle birçok şehirde hizmet vermekte olan Toplum Merkezleri, göçmenlere çeşitli alanlarda hizmet/yardım sağlamaktadır.

20 Ocak 2015 tarihinde Türk Kızılay'ı ilk Toplum Merkezini yereldeki işbirlikleriyle Şanlıurfa'da kurmuştur. Hali hazırda Şanlıurfa, İstanbul (Anadolu ve Avrupa yakası), Konya, Ankara, Kilis, Bursa, İzmir, Adana, Mersin, Gaziantep, Hatay, Kayseri, Kahramanmaraş ve Mardin'de olmak üzere 15 toplum merkezi hem yerel nüfusa hem de sığınmacı nüfusa hizmet sağlamaktadır.

Toplum Temelli Göç Programları kapsamında yürütülen Toplum Merkezlerinin temel amacı; Suriye krizinden etkilenmiş olan ve Türkiye'de kamp dışında geçici koruma altında yaşayan Suriyelilerin, diğer uluslararası ve geçici koruma altında bulunan göçmenlerin ve yerel halkın, psikolojik, sosyal ve ekonomik iyilik hallerinin sağlanması, yeniden toparlanma ve yerel toplum ile uyum içerisinde birlikte yaşama kültürünün geliştirilmesini sağlayarak kriz sonrası iyileştirme yardım faaliyetlerini yürütmektir.

Toplum Merkezleri çatısı altında; Türkiye'ye uyum eğitimi (yasal hakları ve insani hizmetlerle ilgili bilgilendirme), Türkçe dil eğitimi, mesleki eğitimler, bireysel ve grup psikososyal destek aktiviteleri, grup terapileri, bireysel terapiler, çocuklar için drama, müzik ve atölye çalışmaları, çocuk ve genç dostu alan faaliyetleri, sağlık bilgilendirme ve yönlendirme çalışmaları, kadınlar için bilgilendirme toplantıları, el sanatları, giyim üretim, sanat ve tasarım eğitimleri, işe yönlendirme, saha çalışmaları (ihtiyaç tespiti vb.), vaka takibi, koruma gibi faaliyetler yürütülmektedir.

Çok sayıda analiz yöntemiyle, verideki ilişkileri bularak, bunları geçerli tahmin ve/veya sınıflama yapmak için kullanan bir analiz sürecidir. Temel amacı ise, önceki bilgilerin analizinden yararlanarak, geleceğe yönelik tahmin ve karar verme modelleri geliştirmektir [6].

Bu çalışma kapsamında Türk Kızılay Toplum Merkezleri bünyesinde tutulan ve merkez ofis raporlama ve izleme birimi tarafından çalışılmakta olan verilerden yararlanılacaktır. Çalışmada yöntem olarak veri madenciliği seçilmiştir. İleriye dönük tahmin algoritmaları ve veri setinin yapısı göz önüne alındığında veri madenciliği yöntemlerinin uygulanmasına karar verilmiştir. Gerekli izinler sonucu elde edilen veri seti üzerinde veri madenciliğinin tahmin edici modellerinden olan karar ağaçları ve destek vektör makineleri modelleri uygulanmıştır. Çalışmanın amacı, Toplum Merkezleri'nde, yararlanıcılara sağlanan eğitim hizmetlerinin, yararlanıcılar üzerindeki etkisinin araştırılması ve veri madenciliği yöntemleri uygulanarak yorumlanmasıdır. Yararlanıcıların Toplum Merkezleri'nde aldıkları eğitim ile bu eğitimler sonrası istihdam edildikleri iş kolları arasında ilişki olup olmadığı ve eğer ilişki var ise yararlanıcılara ait hangi değişkenlerin bu ilişkiye katkısı olduğu soruları karar ağaçları ve destek vektör makineleri modelleri kullanılarak araştırılacaktır.

2.VERİ MADENCİLİĞİ

Bu başlık altında veri madenciliği hakkında bilgi verilecektir. Veri madenciliğinin tanımı, uygulama alanları ve çalışmada kullanılan modeller olan karar ağacı ve destek vektör makineleri üzerinde durulacaktır.

2.1.Verit Madenciliği Tanımı

Günlük hayatlarımızda herhangi bir internet sitesinde okuduğumuz haberden, yemek tariflerine, sosyal medya hesaplarımızdaki bilgilerden, hava durumu tahminlerine kadar sınırsız veri her gün, her saat hatta her dakika kendini yenilemekte ve depolanmaktadır. Depolanan tüm bu veriler aslında kendi aralarında bir örüntülere/desenlere sahiptir. Bu gizli örüntüleri ortaya çıkarmak amacıyla veri madenciliği yöntemlerine başvurulmaktadır. Veri madenciliği, büyük miktarlardaki verinin içinden geleceği tahmin edilmesinde yardımcı olacak anlamlı ve yararlı bağlantı ve kuralların bilgisayar programlarının aracılığıyla aranması ve analizidir [7].

2.2.Verit Madenciliği Uygulama Alanları

Veri madenciliği veri tabanı teknolojisi, istatistik, yapay zeka (artificial intelligence), makine öğrenimi (machine learning), örüntü tanımlama (pattern recognition) ve veri görselleştirme (data visualization) gibi pek çok teknik alan arasında köprü görevi gören çok disiplinli bir alandır.

Veri madenciliği, günümüzde pek çok alanda kullanılmaktadır. Son yıllarda Amerika Birleşik Devletleri'nde çeşitli veri madenciliği algoritmalarının gizli dinlemeden, vergi kaçakçılıklarının ortaya çıkarılmasına kadar çeşitli uygulamalarda kullanıldığı bilinmektedir [7]. Veri madenciliği bankacılık, genetik, istihbarat, sağlık çalışmaları, sigortacılık, mühendislik çalışmaları vb. gibi alanlarda sıkça başvurulmuş bir analizdir. Örneğin sağlık alanında, ilaç/ürün geliştirme, tahlil sonuçlarının tahmin edilmesi, hastalık tanısı ya da ön teşhis koyma gibi amaçlarla veri madenciliğine başvurulmaktadır. Aynı şekilde yazılım sektöründe de, zararlı maillerin ayıklanması, internet sitelerindeki yasal olmayan içeriklerin tespiti gibi durumlarda veri madenciliği kullanılmaktadır [8].

2.3. Veri Madenciliğinde Kullanılan Yazılımlar

Veri madenciliğinde kullanılmakta olan yazılımlardan bazıları Tablo 1'deki gibidir.

Tablo 1: Veri Madenciliği Çalışmalarında Kullanılan Yazılımlardan Bazıları

MATLAB	IBM SPSS STATISTICS
ORANGE	WEKA
RAPIDMINER	STATICA
SPSS CLEMENTİNE	KNIME
SAS ENTERPRISE GUIDE	TANAGRA
SAS ENTERPRISE MINER	RATTLE

Bu çalışma kapsamında açık kaynak kodlu olması ve araştırmacılar tarafından kolay ulaşılabilir bir yazılım olması sebebiyle KNIME yazılımından yararlanılmıştır. Çalışmada KNIME yazılımının 4.2.2 sürümü kullanılmıştır.

2.4. Çalışmada Kullanılan Modeller

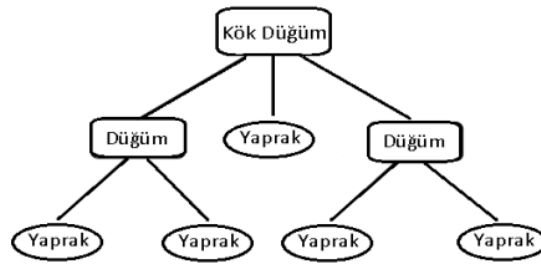
Bu bölümde karar ağaçları ve destek vektör makineleri modellerine ilişkin ayrıntılı bilgiler verilmeye çalışılacaktır.

2.4.1. Karar Ağaçları

Karar ağaçları sınıflandırma ve regresyon için kullanılmakta olan parametrik olmayan bir yöntemdir. Kolay yorumlanabilmesi, veri tabanına kolaylıkla dahil olabilmesi ve güvenilir olması sebepleriyle sınıflandırma problemlerinde sıkça başvurulan bir yöntemdir. Kategorik ve sürekli değişkenler üzerinde çalışabilmektedir. Eğer cevap değişkeni sürekli yapıda ise regresyon ağacı (regression tree), kategorik yapıda ise sınıflandırma ağacı (classification tree) olarak isimlendirilmektedir.

Karar ağaçlarında amaç üzerinde çalışılan verileri kullanarak basit karar kurallarını öğrenmek ve bu kurallardan yararlanarak hedef değişkenin değerini tahmin eden bir model oluşturmaktır [9].

Karar ağaçlarında veriyi sınıflandırma işlemi iki aşamada gerçekleşmektedir. Birinci aşama öğrenme aşamasıdır, bu aşamada bilinen bir öğrenme veri seti, model oluşturmak için sınıflandırma algoritması ile belirlenir. Bu model, sınıflandırma kurallarını oluşturur. İkinci aşama ise test aşamasıdır, sınıflandırma kurallarının doğruluğu değerlendirilir. Sınıflandırma kurallarının doğruluğu kabul edilebilir bir düzeyde ise bu kurallar yeni verilerin sınıflandırılması amacıyla kullanılabilir [10].



Şekil 1: Karar Ağacı Yapısı [11]

Karar ağaçları, Şekil 1'de görülebileceği üzere karar düğümleri, dallar ve yapraklardan meydana gelmektedir. Her bir değişken bir düğümü ifade etmektedir. Kök ve yapraklar arasında kalan kısımlar dal olarak isimlendirilmektedir. Karar düğümleri gerçekleştirilmesi planlanan testtir, soru sorar ve alınan cevaplara göre dallara ayrılır. Dallara ayrılma bölümünde herhangi bir veri kaybı meydana gelmemektedir. Karar ağaçları kök düğümünden başlar ve yukarıdan aşağıya doğru yaprak elde edilinceye kadar dallanır. Eğer dal artık sorusuna cevap alamıyorsa, o dalın sonunda karar düğümü oluşur, dal sorusuna belirgin bir cevap alıyorsa o dalın sonunda yaprak oluşur. Bu yaprak veri setinden alınması istenen cevaplardan birisidir. Dallanma sonucu oluşan karar düğümlerinin sayısı derinlik olarak isimlendirilmektedir [12].

Karar ağaçlarının oluşturulması sürecinde en önemli nokta, veri setinde yer alan değişkenleri sınıflandırması beklenen dallanmanın hangi kritere ya da değişkene göre yapılacağına belirlenmesidir [12]. Literatürde bu belirleme işlemi için geliştirilmiş yaklaşımlar bulunmaktadır. Bunlardan bazıları Twoing kuralı, bilgi kazancı, bilgi kazanç oranı ve Gini indeksi yaklaşımlarıdır [13].

2.4.2. Karar Ağacı Algoritmaları

Literatürde geliştirilmiş birçok karar ağacı algoritması yer almaktadır. Bu bölümde sıklıkla kullanılmakta olan karar ağacı algoritmalarından, CART, CHAID, QUEST, C4.5 ve C5 algoritmaları hakkında bilgi verilecektir.

2.4.2.1.CART (Classification and Regression Tree) Algoritması

Eğer bağımlı değişken kategorik yapıda ise yöntem sınıflama ağaçları (Classification Trees, CT), sürekli yapıda ise regresyon ağaçları (Regression Trees, RT) olarak isimlendirilmektedir. Bağımsız değişkenler de kategorik ya da sürekli bir yapıya sahip olabilmektedir.

Kök düğüm ile başlanarak, iki alt düğüme ve daha sonra her bir düğüm tekrar iki alt düğüme bölünmektedir. Kök düğüm tüm eğitim verilerini içermektedir. Bağımlı değişkene ait tüm veriler burada toplanmaktadır [14].

Karar ağaçlarının oluşum sürecinde dikkat edilmesi gereken en önemli nokta, veri setindeki değişkenlerin sınıflandırılmasını sağlayacak dallanmanın hangi kritere veya hangi değişkene göre yapılacağını belirlemesidir. Bu aşamada, belirsizlik oranı en düşük olan değişken işleme alınır ve kök düğümde test için kullanılır. Hangi değişkene göre dallanmanın yapılacağını belirlemede, yaygın olarak “entropi” sürecini içeren “bilgi kazancı” ve “bilgi kazanç oranı” kullanılmaktadır [15].

$$\text{Bilgi Kazancı } (D, X) = Ent(D) - \sum_{i=1}^n p(D_i) \cdot Ent(D_i) \quad (8)$$

$Ent(D)$ değeri, veri setinin X adet sınıfa bölünmeden önceki entropi değeri,

$Ent(D_i)$ değeri, i alt bölümünün X 'e bölümünden sonra hesaplanan entropi değeri,

$p(D_i)$ değer, i alt bölümünün X 'e bölümünden sonra elde edilen olasılık değeridir $\left(\frac{|D_i|}{|D|}\right)$.

Dallanma bilgi kazancı en yüksek olan değişken üzerinden başlar. Eşitlik 8'de yer alan D değeri veri setini, X değeri ise sınıf değerine göre alt kümeleri simgeliyor olsun. D veri seti kullanılarak, X sınıf değerine bölünmesiyle oluşan kazanç Eşitlik 8'deki gibi hesaplanmaktadır [15] [16] [17].

CART algoritmasında, karar ağacının büyümesinin durdurulup durdurulmayacağı sorusu “durdurma kuralları” ile cevaplanmaya çalışılmaktadır. Eğer bir karar ağacında aşağıdaki durumlardan bir ya da birkaçı gözlemleniyorsa durdurma kuralları uygulanır [15].

- Bir düğümdeki bütün gözlemler, hedef değişken için aynı cevabı veriyorsa,
- Düğüm ve dal sayısı, kullanıcının belirlediği minimum sayıdan küçükse,
- Ağacın derinliği, kullanıcının belirlediği maksimum derinlik sınırına ulaşıyorsa.

CART algoritması parametrik olmaması yani istatistiksel varsayımlar gerektirmemesi sebebiyle araştırmacıya emek, zaman ve maliyet açılarından avantaj sağlayan bir algoritmadır. Hem kategorik hem de sürekli değişkenler üzerinde çalışılabilmesi sık tercih edilmesini sağlamaktadır.

2.4.2.2.CHAID (Chi- Squared Automatic Interaction Detector) Algoritması

CHAID algoritması ismini, karar ağacındaki dallanmaların belirlenmesinde ki-kare test istatistiği kullanılmasından almaktadır. Algoritmada çok sayıda çapraz tablo yer almaktadır. Bu algoritma, kategorik yapıdaki bir hedef değişkeni ve diğer kategorik değişkenler arasındaki ilişkiyi keşfetmek amacıyla kullanılmaktadır [18]. Açıklayıcı değişken sürekli bir yapıda ise bu değişken kategorik değişkene dönüştürülmelidir [19].

CHAID algoritması bağımlı değişkendeki varyasyonu bölümler içi minimum, bölümler arası maksimum olacak şekilde alt gruplara tekrarlı olarak bölen bir tekniktir [13]. Buradaki amaç veriyi daha homojen olan alt gruplara bölmektir. Eğer hedef değişken sürekli bir yapıda ise F testi, kategorik bir yapıda ise ki-kare testi kullanılmaktadır [13]. Bonferroni düzeltmeleri ya da diğer düzenlemeler, yapılan çoklu testleri hesaba katmak amacıyla kullanılmaktadır [20]. Bonferroni p değerleri ile χ^2 istatistikleri hesaplanır [19]. Açıklayıcı değişkenler birbiri ile karşılaştırılıp, en küçük Bonferroni p değerine sahip olan açıklayıcı değişkenin kategorilerine göre, veriler alt gruplara ayrılır [19].

CART, C4.5 gibi karar ağacı algoritmaları ikili dallanma üretmekte iken, CHAID algoritması çoklu dallanma üretmektedir. CHAID algoritması aşağıdaki şekilde çalışmaktadır [13][20]:

- Tüm bağımsız değişkenler için, bağımlı değişkenin kategorileri ile bağımsız değişkenin kategorileri arasında çapraz tablo oluşturulur. Bunların arasından en az öneme sahip yani p değeri en büyük olan kategori çiftleri bulunur. Eğer bağımlı değişken sürekli ise F testi, kategorik ise ki-kare testi kullanılarak p değeri hesaplanır.
- En büyük p değerini alan bağımsız kategori çifti için elde edilen p değeri ile alfa değeri karşılaştırılır. Elde edilen p değeri alfadan büyük ise bu iki kategori birleştirilir ve önceki adıma dönülür, küçük ise sıradaki adıma geçilir.
- Bağımlı ve bağımsız kategori kümeleri için uygun Bonferroni düzeltmesi kullanılarak, düzeltilmiş p değeri hesaplanır.
- En önemli bağımsız değişken, en küçük düzeltilmiş p değerine sahip olan arasından seçilir ve bu değer alfa ile karşılaştırılır. Eğer p değeri alfadan küçük ya da eşit ise düğüm belirlenmiş olan bağımsız kategori kümesi temel alınarak bölünür, büyük ise bölünme olmaz, bu düğüm uç düğüm olur.
- Bu işlem durma kuralları gerçekleşinceye kadar devam eder.

2.4.2.3. QUEST (Quick, Unbiased, Efficient, Statistical Tree) Algoritması

QUEST (hızlı, yansız, etkili istatistik ağacı) algoritması, bir değişkenin iki öğrenme basamağını içeren kök düğümden başlayarak art arda iki düğüme bölünmesiyle oluşturulmuş, ikili karar ağacı sağlayan bir sınıflandırma algoritmasıdır [12]. Hedef değişken kategorik yapıda iken, bağımsız değişkenler kategorik ya da sürekli yapıda olabilir. Karar ağacı oluşturulurken değişken seçimi ve bölünme eş zamanlı olarak yapılmaktadır [21].

QUEST algoritmasında, her bir bölme işlemi için, ANOVA F-testi ya da Levene's test (sıralı ve sürekli değişkenler için) ya da Pearson's Chi-Square (nominal değişkenler için) kullanılarak her bir girdi değişkeni ile hedef değişken arasındaki bağlantılar hesaplanır. Bu algorithmada, karar ağaçlarını budama işlemi için on katlı çapraz doğrulama işlemi yapılmaktadır [22].

QUEST algoritmasının adımları aşağıda yer almaktadır;

Bağımsız değişken seçimi;

1. Her bir belirleyici bağımsız değişken (X) için, eğer bağımsız değişken kategorik alan ise, nominal ya da sıralı, bağımsız değişken ve bağımlı değişken arasındaki bağımsızlığın bir pearson ki-kare testinin p değeri hesaplanır. Eğer bağımsız değişken sürekli ise bu durumda F testi ile p değeri bulunur.

2. En küçük p değeri önceden belirlenen bonferroni düzeltilmiş alfa değeri (α_B) ile kıyaslanır;
 - Hesaplanan en küçük p değeri α_B 'den küçük ise, karşılık gelen ilgili değişken düğümü bölmek için seçilir.
 - Hesaplanan en küçük p değeri α_B 'den büyük ise, sürekli olan her bir X için bir p değeri hesaplanır, eşit olmayan varyanslar için Levene's testi kullanılır.
 - Levene's testinden elde edilen küçük p değeri yeni bir Bonferroni düzeltilmiş alfa düzeyi α_L ile kıyaslanır.
 - Eğer p değeri α_L 'den küçük ise düğümü bölmek için Levene's testinden en küçük p değerli olan ilgili alan seçilir.
 - Eğer p değeri α_L 'den büyük ise düğüm bölünmez.

Bölünme Noktası Seçim Adımları;

1. Eğer bağımlı değişken (Y) sadece iki kategoriye sahip ise, bir sonraki adıma geçer, değilse Y'nin kategorilerini iki üst sınıfa aşağıdaki gibi ayırır.
 - Y'nin her bir kategorisi için X'in ortalaması hesaplanır.
 - Eğer tüm ortalamalar aynı ise, en büyük ağırlıklı frekans kategori bir üst sınıf olarak seçilir ve tüm diğer kategoriler öteki üst sınıf oluşturmak için birleştirilir.
 - Eğer ortalamalar tamamen aynı değil ise, başlangıç küme merkezleri iki en uç sınıf ortalaması olarak ayarlanması ile, k- ortalamalar kümesinde k=2 olarak alınarak, iki-ortalama kümesi algoritması Y'nin iki üst sınıfını elde etmek için uygulanır.

2. Bölme noktasının belirlenmesinde Kuadratik Diskriminant Analizi (QDA) uygulanır. QDA genel olarak iki kesme noktası üretir, birinci üst sınıfın örnek ortalamasına en yakın olanı seçilir [20].

Bu algoritma; ağacın ikili bölünmeyle sınırlandırılması, yansız ağaç tahmininin önemli olması, hesaplama maliyetinin düşürülmek istenmesi ve büyük veri setiyle çalışılması gibi durumlarda tercih edilir [12].

2.4.2.4.C4.5 ve C5.0 Algoritmaları

Ross Quinlan tarafından 1993 yılında geliştirilmiştir. C4.5 algoritması ID3 algoritmasının, C5.0 algoritması ise C4.5 algoritmasının geliştirilmiş halidir. C5.0 algoritması, özellikle büyük veri setlerinde başarılı ağaçlar oluşturmaktadır [14]. Bölme kriteri olarak kazanç oranı (gain ratio) kullanılmaktadır. Bölme işlemi bölünmesi gereken örneklerin sayısı belli bir eşik değerinden düşük olduğu zaman sonlanır [14]. Budama işlemi ise her bir yaprağın hata oranına dayanmaktadır.

Yapraklar sınıflandırılırken; değişken sürekli yapıda ise ikili bölme yöntemini, kategorik yapıda ise çoklu bölme yönetimi kullanılır. Yaprakların en iyisini belirlemede entropi değeri kullanılmaktadır. Dallanmanın durdurulmasında ise Geriye Doğru Budama Kuralı esas alınmaktadır [23].

2.4.3.Karar Ağaçlarında Durdurma ve Budama Kuralları

Üzerinde çalışılan veri setinde çok sayıda değişken olduğu durumlarda, çok sayıda dallanma meydana gelmiş büyük karar ağaçları oluşmaktadır. Bu tür ağaçlarda aşırı öğrenme (overfitting) sorunu görülebilmektedir. Aşırı öğrenme probleminin önüne geçmek amacıyla, araştırmacı modelin maksimum derinliğini belirleyebilir ya da her yaprakta kullanılacak olan minimum eğitim girdisi belirleyebilir [9].

Budama işlemi, önem düzeyi düşük değişkenlerden beslenen dalların kaldırılması işlemidir. Bu sayede ağacın karmaşıklığı ve aşırı öğrenme sorununun yarattığı etki azaltılmış olmaktadır [9].

2.4.4.Karar Ağaçlarının Avantaj ve Dezavantajları

Karar ağaçları kullanıcılar tarafından kolay yorumlanabilen bir yöntemdir. Kolay yorumlanabilmesinin yanı sıra sahip olduğu avantajların bazıları aşağıdaki gibidir:

- Büyük veriler üzerinde uygulanabilmektedir,
- Parametrik olmayan bir yöntemdir, istatistiksel varsayımları gerektirmemektedir,
- Sürekli ve kategorik veriler ile çalışabilmektedir,
- Literatürde birçok kaynak yer almaktadır,
- Araştırmacılar tarafından kolay yorumlanabilmektedir,
- İstatistiksel testler kullanılarak modelin doğrulanması sağlanabilmektedir.

Karar ağaçlarının sahip olduğu bazı dezavantajlar ise aşağıdaki gibidir:

- Karmaşık sistemler oluşturabilmektedir, bu durum literatürde aşırı öğrenme (overfitting) olarak geçmektedir,
- Büyük veriler üzerinde uygulanırken, ağaç oluşturma ve ağaç budama karmaşıklığı fazladır.

2.4.5.Karar Ağaçlarının Uygulama Alanları

Karar ağaçları ekonomi, tıp, felsefe, istatistik vb. birçok disiplin tarafından kullanılmaktadır. Karar ağaçları, gelecekte meydana gelebilecek olayların tahmin edilmesi amacıyla kurallar oluşturma, verilerin kategorilere ayrılması, sürekli değişkenlerin kategorik değişkenlere dönüştürülmesi vb. alanlarda kullanılmaktadır. Uygulama alanlarından bazıları ise aşağıdaki gibidir:

- İşletmeye en faydalı olan bireylerin özellikleri kullanılarak işe alım sürecinin belirlenmesi,
- Satışları hangi değişkenlerin etkilediğinin belirlenmesi,
- Kredi geçmişleri kullanılarak bireylere ilişkin kredi kararlarının verilmesi [23].

2.4.6.Sınıflandırma İşleminin Doğruluğu

Sınıflandırma sürecinde karar ağacının eğitilmesi için gerekli verilerin seçilmesi gerekmektedir. Genellikle veriler eğitim ve test verisi olacak şekilde ikiye bölünmektedir. Bu işlem oluşturulan ağacın doğruluğunu ölçmek için yapılmaktadır. Ancak bazı durumlarda veri seti, bu bölme işleminin gerçekleştirilebileceği kadar büyük olmayabilir. Bu durumda doğruluğu test etmek amacıyla geliştirilmiş yöntemler bulunmaktadır. Bu yöntemler Dışarıda Tutma (Hold-Out), K-Katlı Çapraz Doğrulama (K-Fold Cross Validation), Leave-One-Out ve Bootstrap olmak üzere dört grup altında incelenmektedir.

Tablo 2:Model Değerlendirmesi

Model Değerlendirmesi	İşleyişi
Dışarıda Tutma (Hold-Out)	Veri seti eğitim ve test verisi olarak ikiye bölünmektedir. Genellikle veri setinin 2/3'ü eğitim, 1/3'ü ise test verisi olarak ayrılmaktadır. Sınıflandırmanın doğruluğu test verisi üzerinden hesaplanmaktadır [24].
K-Katlı Çapraz Doğrulama (K-Fold Cross Validation)	Veri seti k adet alt kümeye ayrılmaktadır. Uygulamada genellikle k=10 alınmaktadır. Alt kümeler belirlendikten sonra test kümesinin yeri değiştirilerek hata oranları sırayla elde edilmektedir. k=10 olarak alındığında ilk grup test, kalan dokuz grup ise model oluşturmada kullanılmaktadır. Bu yöntem her seferinde test verisi değiştirilerek 10 kez tekrarlanmaktadır. Elde edilen 10 adet bağımsız hata değerinin ortalaması, kurulan modelin hata oranı olmaktadır [25].

<p>Tek Çıkışlı Çapraz Doğrulama (Leave-One-Out)</p>	<p>K-Katlı Çapraz Doğrulama yönteminin özel bir halidir. Bu yöntemde k sayısı, veri sayısına yani N değerine eşittir. N-1 adet örnek eğitim verisi olarak kullanılır, kalan 1 örnek ise test için kullanılır. Bu işlem, her örnek bir kere sınanacak şekilde gerçekleştirilir [26]. Model başarısı ise denemeler sonucunda elde edilen hataların ortalaması alınarak bulunur [27].</p>
<p>Yeniden Örnekleme (Bootstrap)</p>	<p>Bu yöntem yerine koyarak istatistiksel örnekleme yöntemini temel almaktadır [24]. Genellikle küçük veri setlerinde kullanılmaktadır. Model tüm veri setinden oluşturulmaktadır, sonrasında bootstrap ismi verilen çok sayıda örnekleme çekilmektedir. Her gözlem örnekleme seçildikten sonra yerine konur ve tüm örnekleme seçilene kadar işlem tekrar edilir. Bir kayıt birden fazla veri setinde yer alabilir. Model bu veri setinden oluşturulur ve hata oranı hesaplanır [28]. Model için tüm veri setinden oluşan son hata tahmin değeri, bütün bootstrap örneklemlerinin hata tahmin değerinin ortalaması alınarak hesaplanır [25].</p>

Tablo 2’de doğruluğu test etmek amacıyla kullanılan yöntemlerin özet halinde açıklamaları yer almaktadır.

2.4.7.Sınıflandırma Performanslarının Değerlendirilmesi

Sınıflandırma algoritmalarından hangisi ile çalışılması gerektiği sorusuna, aşağıdaki kriterler göz önüne alınarak karar verilebilmektedir.

2.4.7.1.Karışıklık Tabloları

Karışıklık tablolarından yararlanarak doğruluk, kesinlik, hata oranı, anma ve F-ölçütü gibi değerlere ulaşılabilir. Tablo 3’te karışıklık tablosu gösterilmektedir.

Tablo 3:Karışıklık Matrisi [24]

	TAHMİN		
	POZİTİF	NEGATİF	
GERÇEK POZİTİF	TP (DOĞRU POZİTİF)	FN (YANLIŞ NEGATİF)	TP+FN
GERÇEK NEGATİF	FP (YANLIŞ POZİTİF)	TN (DOĞRU NEGATİF)	FP+TN
	TP+FP	FN+TN	TP+FN+FP+TN=N (KİTLE)

$$Doğruluk = \frac{TP + TN}{N (Kitle)} \quad (1)$$

$$Hata Oranı = \frac{FN + FP}{N (Kitle)} \quad (2)$$

$$Kesinlik = \frac{TP}{TP + FP} \quad (3)$$

$$Duyarlılık = \frac{TP}{TP + FN} \quad (4)$$

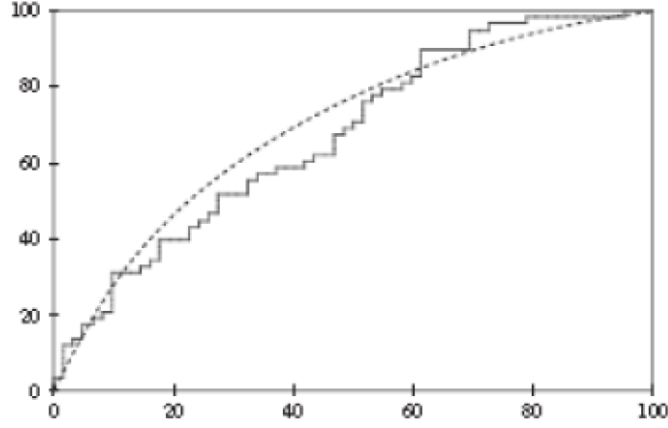
Burada TP (True Positive) doğru pozitif, TN (True Negative) doğru negatif, FN (False Negative) yanlış negatif ve FP (False Positive) ise yanlış pozitif değerlerini temsil etmektedir. N değeri ise çalışmada yer alan veri sayısıdır. N değeri TP, TF, FP ve FN değerlerinin toplamına eşittir. Modelin başarı durumu incelenmek istenildiğinde doğruluk ve hata oranı değerleri kullanılmaktadır [29].

$$F - \text{Ölçütü} = \frac{2 * Kesinlik * Anma}{Kesinlik + Anma} \quad (5)$$

F ölçütü anma ve kesinlik değerlerinin harmonik ortalaması ile hesaplanmaktadır [30].

2.4.7.2.ROC Eğrileri

Sınıflandırma algoritmalarının karşılaştırılmasında kullanılan diğer bir kriter ise ROC (Receiver Operating Characteric) eğrileridir.



Şekil 2:ROC Eğrisi [24]

Şekil 2’de yer alan ROC Eğrisi çiziminde yatay ekseninde yanlış pozitif değerler, dikey ekseninde ise doğru pozitif değerler yer almaktadır. Bu çizimde “doğru pozitif” ve “yanlış negatif” değerlerinden yararlanılmaktadır.

$$\text{Doğru Pozitif Oranı (TPR)} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{Yanlış Pozitif Oranı (FPR)} = \frac{FP}{TN + FP} \quad (7)$$

Doğru pozitif oranı ve yanlış pozitif oranı (6) ve (7) numaralı işlemler aracılığıyla elde edilmektedir.

TP ve FP değerleri ile çizilen ROC eğrisi grafikte yer alan değerler için;

(0,0): Bütün değerler negatif sınıflandırıldığını,

(1,1): Bütün örnekler pozitif sınıflandırıldığını,

(0,1): Bütün negatif ve pozitif değerlerin doğru sınıflandırıldığını (en ideal durum),

(1,0): bütün negatif ve pozitif değerlerin hatalı sınıflandırıldığını ifade eder.

Bu durumlardan da anlaşılacağı üzere ROC eğrisi altındaki alanı 1’e daha yakın olan modelin sınıflandırma işleminde kullanılması tercih edilir [31].

2.5. Destek Vektör Makineleri

Destek vektör makineleri (DVM), istatistik öğrenme teorisine dayanan, sınıflandırma ve regresyon problemlerinin çözümü amacıyla Vapnik tarafından ortaya atılmış, parametrik olmayan bir sınıflandırma yöntemidir [32] [33].

DVM'nin amacı iki sınıflı bir veriyi en uygun karar fonksiyonunu tahmin ederek birbirinden ayırmaktır [33]. DVM optimizasyon temelli olmaları sebebiyle sınıflandırma performansı ve kullanılabilirlik gibi özellikleri ile diğer tekniklere göre daha başarılı görülmektedir [35]. Bu algoritmalar, yüz-ses-el yazısı tanıma, görüntülerin sınıflandırılmasında vb. kullanılmaktadır. Özellikle biyoloji biliminde yaygın olarak kullanılmaktadır.

DVM algoritmaları veriye ilişkin birleşik dağılım fonksiyonu gerektirmemesi sebebiyle dağılımdan bağımsız algoritmalar [36]. DVM algoritmaları ilk olarak doğrusal olan iki sınıflı verilerin sınıflandırılmasında kullanılırken, sonrasında doğrusal olmayan ve çok sınıflı verilerin sınıflandırılmasında da kullanılabilir şekilde geliştirilmiştir [37]. Gerçek veriler üzerinde çalışılırken, doğrusal olarak ayrılması mümkün olmayan veriler ile çalışılması gerekebilmektedir. Bu gibi durumlarda, verileri doğrusal olarak ayrılabilirleri yüksek boyutlu bir uzaya taşımak için kernel (çekirdek) fonksiyonları kullanılmaktadır, ancak bu taşıma işlemi beraberinde çoklu öğrenme ve hesaplama hataları sorunlarını da beraberinde getirmektedir. Bu hataları en aza indirmek amacıyla iç çarpım formülleri kullanılmaktadır.

Destek vektör makineleri iki alt başlık halinde incelenmektedir.

1. Doğrusal Destek Vektör Makineleri
2. Doğrusal Olmayan Destek Vektör Makineleri

2.5.1. Doğrusal Destek Vektör Makineleri

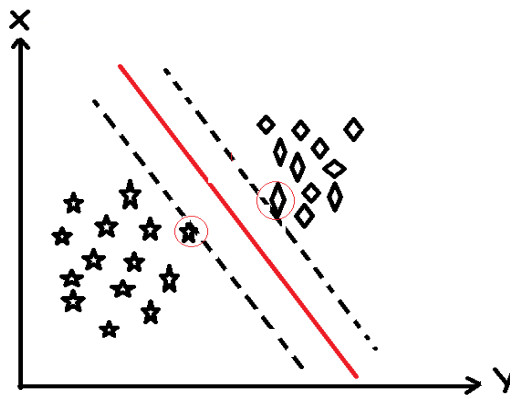
Doğrusal olan DVM'leri doğrusal olarak ayrılabilen ve doğrusal olarak ayrılamayan veriler olmak üzere iki gruba ayrılmaktadır.

2.5.2. Doğrusal Olarak Ayrılabilen Verilerde DVM

Doğrusal olarak ayrılabilen DVM'leri doğrusal ayrılma ve belirli bir hata ile doğrusal ayrılma durumlarına göre iki ayrı başlık altında incelenmektedir.

- Doğrusal Ayrılma Durumu

Doğrusal olarak ayrılabilen verilerde DVM anlatımında aşağıdaki şekilden faydalanılacaktır.



Şekil 3: Doğrusal Olarak Ayrılabilen Durum- Optimum Hiper-Düzlem

Şekil 3'de X düzlemi üzerinde yer alan veriler internet abonelik ücretini (aylık), Y düzlemi üzerinde yer alan veriler ise ne kadar süredir abone olduğunu belirtmekte olsun. Kırmızı ile renklendirilmiş doğru literatürde hiper düzlem olarak, bazı kaynaklarda ise karar doğrusu olarak geçmektedir. Hiper düzlem doğrusunun sağında yani yukarısında kalan verilerin aboneliğini sonlandırmaya eğilimi olan müşterilere ait, doğrunun solunda yani aşağısında kalan verilerin ise aboneliğini devam ettirmeye eğilimi olan müşterilere ait veriler olduğunu varsayalım.

Buradaki amaç abonelik süresi ve abonelik ücreti olarak bahsedilen iki sınıfı, eğitim verileri kullanarak elde edilen hiper düzlem ile birbirinden ayırmaktır. İlk olarak sınıflar arasında birbirine en yakın iki nokta seçilir, Şekil 1'de kırmızı daireler içine alınan veriler destek vektörleri olarak isimlendirilmektedir. Kesik çizgiler ile belirtilen doğrular ise sınır çizgileridir. Sınır çizgileri, destek vektörlerinin bulunduğu noktadan geçecek şekilde çizilmektedir. Sınır çizgileri arasında herhangi bir eğitim verisi yer almamaktadır [38]. Hiper düzlem, her sınıfın sınır çizgisine eşit uzaklıkta olacak şekilde çizilmektedir.

Hiper düzlemin üst kısmında kalan sınır çizgisinin formülü: $w \cdot x + b = 1$

Hiper düzlemin alt kısmında kalan sınır çizgisinin formülü: $w \cdot x + b = -1$

Hiper düzleme ait formül: $w \cdot x + b = 0$ olacak şekildedir.

Hiper düzlemin bulunabilmesi için veri setindeki tüm verilerin (1) ve (2) numaralı eşitsizlikleri sağlaması gerekmektedir [36].

$$f(X_i) = \langle w, x_i \rangle + b \geq +1, \quad y_i = +1 \quad (1)$$

$$f(X_i) = \langle w, x_i \rangle + b \leq -1, \quad y_i = -1 \quad (2)$$

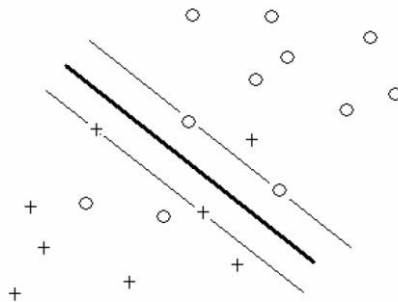
Bu eşitsizliklerde x bir vektör noktası, w ağırlık vektörü, b ise yan (bias) olmak üzere bir sabit sayıdır [31]. En uygun hiper düzlemin seçiminde aşağıdaki iki durum önemli rol oynamaktadır.

1. Sınır çizgileri arasındaki mesafe (margin) en büyük olmalıdır,
2. Sınıflara ait örnekler birbirlerinden en uygun şekilde ayrılabilirler.

Eğer bu iki durum sağlanıyorsa seçilen hiper düzlem en uygun düzlem (optimal hyperplane) olarak ifade edilir.

- Belirli Bir Hata İle Doğrusal Ayrılma Durumu

Gerçek veriler üzerinde yapılan çalışmaların çoğunda ayırıcı hiper düzlem bulunamamaktadır. Hatalı girilmiş veriler, verilerin çok boyutlu olması ya da gürültülü veriler hiper düzlemin belirlenmesinde problemler yaratmaktadırlar. Bu durumda verilerin minimum düzeyde hata ile ayrılmasına izin verilir [39].



Şekil 4: Belirli Bir Hata İle Doğrusal Ayrılma Durumu [39]

Şekil 4'te iki sınıfa ait verilerin doğrusal olarak ayrılamadığı, iki sınır çizgisinin de gerisinde diğer sınıfa ait verinin/verilerin yer aldığı görülmektedir.

Belirli bir hata ile doğrusal ayrılma durumu söz konusu olduğunda yumuşak sınır (soft margin) yaklaşımı kullanılmaktadır. Bu yaklaşımda modele ε_i aylak değişkeni eklenmektedir.

Bu değişken, bir örneğin yanlış sınıflandırılması durumunda ait olduğu karar sınırına olan uzaklığının ölçüsüdür [40] [38].

$$f(x_i) = \langle w, x_i \rangle + b \geq +1 - \varepsilon_i, \quad y_i = +1 \quad (3)$$

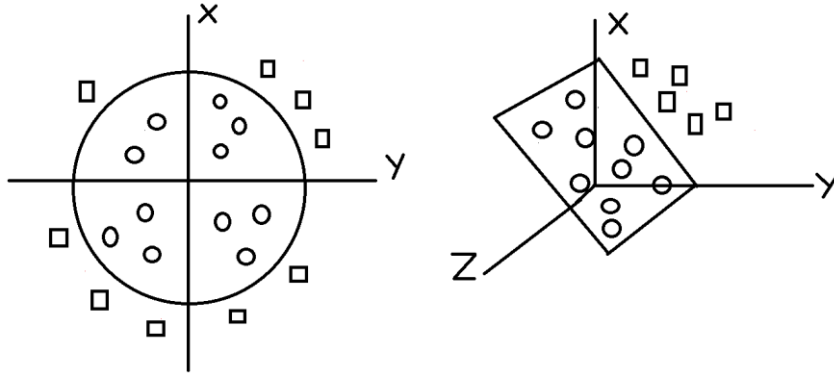
$$f(x_i) = \langle w, x_i \rangle + b \leq -1 + \varepsilon_i, \quad y_i = -1 \quad (4)$$

Belirli bir hata ile doğrusal ayrılma durumunda, hiper düzlemin bulunabilmesi için veri setindeki tüm verilerin (3) ve (4) numaralı eşitsizlikleri sağlaması gerekmektedir [40].

2.5.3. Doğrusal Olarak Ayrılamayan Verilerde DVM

Doğrusal olarak ayrılamayan DVM'nde veriler öncelikle çekirdek fonksiyonlar aracılığı ile daha yüksek boyutlu ve doğrusal olarak ayrılabilirler bir uzaya alınmaktadır. Doğrusal DVM'nden doğrusal olmayan DVM'ne geçişi sağlayan bu işlem Kernel Hilesi olarak adlandırılmaktadır [14].

p boyutlu girdi vektörü x 'in p boyutlu özellik vektörü ϕ 'ye dönüştürülmektedir [40]. Bu dönüşüm işlemi için hiper düzlemin özellik uzayında tanımlanması gerekmektedir. Hiper düzlemin özellik uzayında tanımlanabilmesi amacıyla doğrusal olmayan haritalama yaklaşımından yararlanılmaktadır [41].



Şekil 5:Doğrusal Olmayan Haritalama Yaklaşımı Örneği

Doğrusal olmayan haritalama yaklaşımı ile iki boyutlu bir veri seti, üç boyutlu özellik uzayına taşınarak doğrusal olarak ayrımı sağlanmaktadır [42]. Şekil 5’de doğrusal olmayan haritalama yaklaşımına ilişkin bir örnek yer almaktadır.

Tablo 4:Destek Vektör Makineleri’nde Kullanılan Temel Çekirdek (Kernel) Fonksiyonlar

Kernel Tipi	Fonksiyon
Doğrusal (Lineer)	$K(x_i, x_j) = x_i^T \cdot x_j$
Polinom	$K(x_i, x_j) = (\alpha \cdot x_i^T \cdot x_j + b)^d, \alpha > 0$
Radyal Tabanlı Fonksiyon	$K(x_i, x_j) = \exp(-g \cdot \ x_i - x_j\ ^2), g > 0$
Sigmoid	$K(x_i, x_j) = \tan\alpha(g \cdot x_i^T \cdot x_j + b)$

Tablo 4’te destek vektör makinelerinde kullanılmakta olan temel kernel fonksiyonları yer almaktadır.

Polinom fonksiyonuna ait denklemde yer alan d indisi polinom derecesini ifade etmekte ve 1 ile 6 değerleri arasında değer almakta iken b değeri sabit terimi ve α değeri ise eğilimi ifade etmektedir [43]. Sigmoid fonksiyonunda da yer α değeri eğim ifade etmektedir, α için ortak bir değer $1/N$ ’dir. Burada N veri setinin büyüklüğüdür [44].

Çekirdek fonksiyonlarının seçiminde üzerinde çalışılan veri setinin yapısı göz önüne alınmaktadır. Örneğin sınıflandırma problemi doğrusal olarak ayrılabilirse Doğrusal (Lineer) kernel kullanılabilir, ya da sınıflandırma problemi kuadratik bir yapıdaysa Polinom kernel kullanılabilir.

3.TOPLUM MERKEZLERİ

Türkiye’de kayıtlı Suriyeli göçmen sayısı 23 Eylül 2020 tarihi itibarıyla bir önceki aya göre 12 bin 84 kişi artarak 3 milyon 621 bin 968 kişi oldu [3].

Suriye’de dokuzuncu yılına giren iç karışıklık sebebiyle milyonlarca Suriyeli, ülkesinden göç etmek zorunda kalmıştır. Türkiye devlet politikası gereği birçok Suriyeli göçmene kapılarını açmış ve 6458 sayılı kanunun 91.maddesindeki “Geçici Koruma” maddesine istinaden 13/10/2014 tarihli Geçici Koruma Yönetmeliği kapsamında göçmenlere Geçici Koruma Kimliği sağlamıştır. Bu sayede Suriyeli göçmenlerin temel hak ve hizmetlerden yararlanabilmesi garanti altına alınmıştır [45].

Türkiye’deki kayıtlı Suriyeli sayısı 25 Aralık 2019 tarihi itibarıyla 3 milyon 571 bin 30 kişi oldu. Bu kişilerin %53,9’u erkeklerden, %46,1’i ise kadınlardan oluşuyor [3].

Ülkemizdeki Suriyelilerin %48,12’si 0-18 yaş grubu aralığındadır. Ülkemizde bulunan kayıtlı Suriyelilerin yaş ortalaması ise yaklaşık olarak 21,66’dır.

Elde edilen son istatistikler sonucunda da görüldüğü gibi ülkemizde çok sayıda genç, çalışma çağında Suriyeli mülteci bulunmaktadır. Bu açıdan bakıldığında, Suriyeli mültecilerin Türk halkı ile kaynaştırılarak gerek sosyal hayata gerekse iş gücüne katılımlarının sağlanmasında yardıma ihtiyaçları bulunmaktadır.

Türk Kızılay, gerek sınır içi gerekse sınır dışında insani yardım müdahaleleri gerçekleştirmekte ve birçok hizmet sunmaktadır. Bu amaçla ülkemize göç eden Suriyeli mültecilerin Türk halkı ile birlikte psikolojik, sosyal ve ekonomik açılardan kaynaştırılması ve uyum içerisinde birlikte yaşamın sağlanması için 15 ilde 16 farklı Türk Kızılay Toplum Merkezleri kurulmuştur. Bu Toplum merkezleri aracılığı ile Koruma, Psikososyal Destek, Geçim Kaynağını Geliştirme ve Sosyal Uyum programları yürütülmektedir [45].



Şekil 6:Ülkemizde Şehirlere Göre Suriyeli Göçmen Yoğunlukları [46]

Şekil 6’da ülkemizin hangi şehirlerinde daha fazla Suriyeli mülteci olduğu görülmektedir. Haritadaki kırmızı rengin açık tonlu olduğu şehirlerde, koyu tonlu olan şehirlere göre daha az sayıda Suriyeli mülteci bulunmaktadır. Türkiye’de Suriyeli mültecilerin en yoğun olduğu şehirler Hatay, Gaziantep, Şanlıurfa ve İstanbul şehirleridir. Suriye ile Türkiye arasında sınır kapısı bulunan illerin daha yoğun göç aldığı söylemek yanlış olmayacaktır. Toplum merkezlerinin bulunduğu konumlara dikkat edildiğinde Suriyeli nüfusun yoğun olduğu iller olduğu görülecektir.

BAGCILAR		102.768
Koruma	19.653	
Geçim Kaynağını Geliştirme	7.712	
Çocuk ve Gençlik	19.046	
Sosyal, Kültürel ve Uyum	30.161	
Psikososyal Destek ve Sağlık	26.196	

SULTANBEYLİ		253.672
Koruma	34.498	
Geçim Kaynağını Geliştirme	5.958	
Çocuk ve Gençlik	22.213	
Sosyal, Kültürel ve Uyum	50.437	
Psikososyal Destek ve Sağlık	140.566	

BURSA		89.010
Koruma	16.503	
Geçim Kaynağını Geliştirme	3.782	
Çocuk ve Gençlik	17.319	
Sosyal, Kültürel ve Uyum	27.319	
Psikososyal Destek ve Sağlık	24.087	

İZMİR		98.023
Koruma	18.352	
Geçim Kaynağını Geliştirme	3.793	
Çocuk ve Gençlik	11.233	
Sosyal, Kültürel ve Uyum	39.192	
Psikososyal Destek ve Sağlık	25.453	

MERSİN		90.507
Koruma	16.189	
Geçim Kaynağını Geliştirme	3.325	
Çocuk ve Gençlik	9.722	
Sosyal, Kültürel ve Uyum	31.074	
Psikososyal Destek ve Sağlık	30.197	

KOCAELİ		33.443
Koruma	4.668	
Geçim Kaynağını Geliştirme	1.966	
Çocuk ve Gençlik	11.636	
Sosyal, Kültürel ve Uyum	6.062	
Psikososyal Destek ve Sağlık	9.111	

ANKARA		85.805
Koruma	15.244	
Geçim Kaynağını Geliştirme	3.642	
Çocuk ve Gençlik	20.805	
Sosyal, Kültürel ve Uyum	25.076	
Psikososyal Destek ve Sağlık	21.038	

KONYA		90.625
Koruma	16.710	
Geçim Kaynağını Geliştirme	4.583	
Çocuk ve Gençlik	23.068	
Sosyal, Kültürel ve Uyum	19.517	
Psikososyal Destek ve Sağlık	26.747	

ŞANLIURFA		113.231
Koruma	24.630	
Geçim Kaynağını Geliştirme	3.840	
Çocuk ve Gençlik	14.940	
Sosyal, Kültürel ve Uyum	12.621	
Psikososyal Destek ve Sağlık	57.200	

GAZİANTEP		88.595
Koruma	16.625	
Geçim Kaynağını Geliştirme	3.148	
Çocuk ve Gençlik	15.289	
Sosyal, Kültürel ve Uyum	32.293	
Psikososyal Destek ve Sağlık	21.240	

KAYSERİ		67.238
Koruma	12.334	
Geçim Kaynağını Geliştirme	2.321	
Çocuk ve Gençlik	10.519	
Sosyal, Kültürel ve Uyum	21.816	
Psikososyal Destek ve Sağlık	20.248	

K.MARAS		82.515
Koruma	13.092	
Geçim Kaynağını Geliştirme	2.093	
Çocuk ve Gençlik	8.528	
Sosyal, Kültürel ve Uyum	35.482	
Psikososyal Destek ve Sağlık	23.320	

MARDİN		58.709
Koruma	11.252	
Geçim Kaynağını Geliştirme	1.804	
Çocuk ve Gençlik	17.889	
Sosyal, Kültürel ve Uyum	11.383	
Psikososyal Destek ve Sağlık	16.381	



ADANA		96.178
Koruma	13.745	
Geçim Kaynağını Geliştirme	3.530	
Çocuk ve Gençlik	9.100	
Sosyal, Kültürel ve Uyum	32.567	
Psikososyal Destek ve Sağlık	37.236	

HATAY		82.923
Koruma	15.074	
Geçim Kaynağını Geliştirme	1.942	
Çocuk ve Gençlik	9.723	
Sosyal, Kültürel ve Uyum	34.421	
Psikososyal Destek ve Sağlık	21.763	

KİLİS		91.557
Koruma	21.249	
Geçim Kaynağını Geliştirme	3.622	
Çocuk ve Gençlik	21.721	
Sosyal, Kültürel ve Uyum	20.968	
Psikososyal Destek ve Sağlık	23.997	

Şekil 7: Türk Kızılay Toplum Merkezleri'ne Ait İstatistikler [48]

Şekil 7’de aktif olarak hizmet vermekte olan Toplum Merkezleri’nin programlar bazında ulaştıkları yararlanıcıların sayıları görülmektedir. Şekilde yer alan rakamlar kümülatif olmakla birlikte en çok yararlanıcıya ulaşılan Toplum Merkezi’nin İstanbul Sultanbeyli Toplum Merkezi olduğu görülmektedir.

3.1. Toplum Merkezleri Nedir, Nasıl Çalışır

Toplum merkezleri Suriyeliler ve birlikte yaşadıkları diğer savunmasız gruplar için psikososyal destek sağlamak, yönlendirme, koruma ve eğitim hizmetleri sunmaktadır [45]. Toplum Merkezleri’nin ilki Ocak 2015’te Şanlıurfa ilinde hizmete girmiştir. Yıllar içerisinde sayıları artarak bugün 15 ilde 16 farklı Toplum merkezi haline gelmiştir. Bu Toplum merkezleri İstanbul (Sultanbeyli ve Bağcılar), Şanlıurfa, Ankara, Konya, Kayseri, Kahramanmaraş, Gaziantep, Mardin, Mersin, Kilis, Adana, Bursa, Hatay, İzmir ve Kocaeli illerinde hizmet vermektedirler. Toplum merkezleri Suriye’deki iç karışıklıktan etkilenerek ülkemize göç etmek zorunda kalmış Suriyeli nüfusun, diğer milletlerden ülkemize göç etmiş nüfusun ve yerel halkın hep birlikte hizmet alabileceği, birleştirici bir yapıdır. Bahsedilen tüm bu insanların psikolojik, sosyal ve ekonomik açıdan iyileştirilmesi ve yerel halk ile uyum içerisinde yaşayabilmesi adına hizmet vermeyi amaçlamaktadırlar.

3.2. Toplum Merkezleri Tarafından Verilen Hizmetler

Toplum merkezleri bünyesinde, Suriyeli halkın kendilerini geliştirebilmeleri ve olabildiğince hızlı bir şekilde yerel halk ile kaynaşabilmeleri amacıyla belirlenmiş pek çok hizmet verilmektedir. Bu hizmetlerden bazıları aşağıdakiler gibidir:

- Türkçe dil eğitimi,
- Meslek edindirme kursları,
- Bireysel ve grup psikososyal destek etkinlikleri,
- Çocuklar için müzik ve atölye çalışmaları,
- Çocuk ve genç dostu alanlar ve etkinlikler,
- Hukuki danışmanlık ve seminer hizmetleri,
- Saha çalışmalar (ihtiyaç tespiti vb.),

- Vaka takibi ve koruma hizmetleri.

3.3.Uygulanan Programlar

Toplum merkezleri tarafından Suriyeli göçmenler ve yerel halka verilen hizmetler genel olarak 5 ana program adı altında birleşmektedir.

3.3.1.Geçim Kaynağını Geliştirme Programı

Bu program kapsamında ihtiyaç sahibi göçmenlere gerekli eğitimler verilerek kişilerin ekonomik güçlerini kendilerinin kazanması ve iş gücüne katılmalarının sağlanması hedeflenmektedir.

İşe yönlendirme ve İstihdam Danışmanlık Hizmetleri



Bireysel iş danışmanlığı, firmalar ve stajyerler arasında iş birliğinin kurulması, işveren ve işçinin eşleşmesi, işveren için çalışma izni ücretinin karşılanması ve çalışma izni konusunda bilginin yaygınlaşması amacıyla bilgilendirme seminerlerini kapsar.

19.699

Mesleki Eğitimler



Hayat Boyu Öğrenme Genel Müdürlüğü ve özel sektör ile yapılan iş birlikleri aracılığıyla kendi yaşamsal becerilerini geliştirme, piyasa ihtiyaçları ve istihdam alanları doğrultusunda saça şekil kuaförlük, biçki – dikiş, bilgisayar, takı tasarımı, kaynak teknikerliği, açılış, hasta ve yaşlı bakımı gibi kurslar yürütülmüştür.

19.383

Tarım ve Hayvancılık Eğitimleri



Tarım ve Orman Bakanlığı ile iş birliği yapılarak proje kapsamındaki illerde paydaşlar aracılığıyla yararlanıcılara teorik ve uygulamalı mesleki eğitimlerin verilmesi, malzeme ve ekipman ihtiyaçlarının karşılanması ve yararlanıcıların istihdam edilmesine yönelik işbirlikleri gerçekleştirilmektedir.

2.103

Türkçe Dil Kursları



Bireysel ve piyasa ihtiyaçları çerçevesinde mesleki eğitim faydalanıcıları belirli sektörlerde istihdamlarını pekiştirmek için Türkçe kurslarına da yönlendirilirler. (A1'den C2'ye kadar)

15.876

Şekil 8:Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 1 [48]

Şekil 8'de Geçim Kaynağını Geliştirme Programı kapsamında, Toplum merkezlerinin kurulumundan günümüze kadar geçen sürede kaç kişiye hizmet götürüldüğü belirtilmektedir.

Geçim kaynağını geliştirme programı üç alt programa ayrılmaktadır.

- İstihdama Yönlendirme

İş gücüne katılımı sağlamak amacıyla çeşitli analizler yapılmaktadır. Bu analizler sonucunda elde edilen bilgiler aşağıdaki gibidir:

1. Boş pozisyonların tespiti,
2. İşverenlerin talepleri,
3. Yetkinlik ihtiyaçları.

Elde edilen bu ve benzeri bilgiler ışığında yararlanıcılara mesleki eğitimler, Türkçe dil kursları ve kişisel gelişim adına eğitimler verilmektedir. Eğitimler sonucunda istihdam edilmeye hak kazanan yararlanıcılar için çalışma şartları hakkında bilgilendirmeler verilmekte ve çalışma izinleri sürecinde danışmanlık hizmetleri verilmektedir. Ek olarak çalışma izni harç bedeli ücretleri de Toplum merkezleri tarafından karşılanmaktadır.

- Girişimcilik Programı

Çeşitli ihtiyaç analizleri sonucunda üniversiteler, diğer sivil toplum kuruluşları ve benzeri kurumlar ile iş birliği gerçekleştirilmektedir. Bu kapsamda temel ve ileri düzeyde girişimcilik eğitimleri verilmektedir.

Eğitimler sonucunda yararlanıcılar için girişimcilik kampları düzenlenmektedir. Bu kamplara katılan yararlanıcılar arasında sürdürülebilirlik açısından uygun görülen iş planları belirlenerek bu iş planlarına aynı destekler verilmektedir.

Uygun bulunan iş planlarına hukuki danışmanlık, pazarlama desteği gibi konularda destek verilmeye devam edilmektedir.

- Tarım ve Hayvancılık Programı

Tarım ve Orman Bakanlığı'na bağlı yerel kurumlar, Ziraat odaları, birlikler ve Üniversiteler ile yapılan işbirlikleri doğrultusunda yerel ihtiyaçları karşılamak üzere yararlanıcılara teorik ve uygulamalı tarım ve hayvancılık eğitimleri verilmektedir. Eğitimleri başarıyla tamamlayan yararlanıcılar istihdama yönlendirilmekte veya üretim gerçekleştirmek isteyen yararlanıcılara Hane İçi Üretim, Küçük Tarımsal ve Hayvancılık Üretimi destekleri verilmektedir.

3.3.2.Koruma Programı

Koruma Programı kapsamında hem Suriyeli göçmenlere hem de diğer milletlerden Türkiye'ye sığınmış olan göçmenlerin karşılaşılabilecekleri tehditler, sömürü ve istismar gibi durumlardan doğabilecek olumsuz sonuçları gerek önlemek gerekse azaltmak amaçlanmaktadır.



Şekil 9:Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 2 [48]

Şekil 9'da Koruma Programı tarafından, Toplum merkezlerinin kurulumundan günümüze kadar geçen sürede kaç kişiye hizmet götürüldüğü belirtilmektedir.

Koruma Programı tarafından verilen hizmetlerin bazıları aşağıdaki gibidir:

- Temel Hak ve Hizmetlere Erişim

Yararlanıcıların erişebilecekleri eğitim, sağlık, evlenme/boşanma, hukuki seminer ve danışmanlık gibi hizmetlere dair bilgilendirmeler yapılmaktadır.

- Çocuk Koruma ve Eğitime Erişim

Bu kapsamda çocuk işçiliğın önlenmesi, istismar, çocuk yaşta evlilik gibi tehditlerden korumak amaçlanmaktadır. Söz konusu risk altındaki çocukların tespiti ve değerlendirilmesi hususunda Aile, Çalışma ve Sosyal Hizmetler Bakanlığı gibi kurum ve kuruluşlarla birlikte çalışılmaktadır.

- Şiddetin Önlenmesi ve Şiddetle Mücadele

Şiddetin tüm türlerini önlemek ve şiddete maruz kalmış yararlanıcılar için ilgili kurum ve kuruluşlara gerekli başvuruların yapılmasını sağlamak gibi hizmetler verilmektedir.

- İnsan Ticaretinin Önlenmesi ve İnsan Ticareti Mağdurlarına Yönelik Müdahale

İnsan Ticaretinin tespiti ve ilgili birimlere yönlendirilmesi gibi hizmetler verilmektedir. Ek olarak insan ticareti mağdurlarına yönelik sığınma evleri hizmetleri de verilmektedir.

3.3.3. Psikososyal Sağlık ve Destek Programı

Özellikle Suriye'deki iç karışıklıktan etkilenerek ülkemize göç etmek zorunda kalan Suriyeli göçmenler ve diğer milletlerden ülkemize sığınan göçmenlere yönelik sosyal, psikolojik ve fiziksel durumlarının değerlendirilmesi ve gerekli desteklerin verilmesi amaçlanmaktadır.



Şekil 10: Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 3 [48]

Şekil 10'da Psikososyal Sağlık ve Destek Programı tarafından, Toplum merkezlerinin kurulumundan günümüze kadar geçen sürede kaç kişiye hizmet götürüldüğü belirtilmektedir. Psikososyal Sağlık ve Destek Programı tarafından verilen hizmetlerden bazıları aşağıdaki gibidir:

- Psikolojik Destek

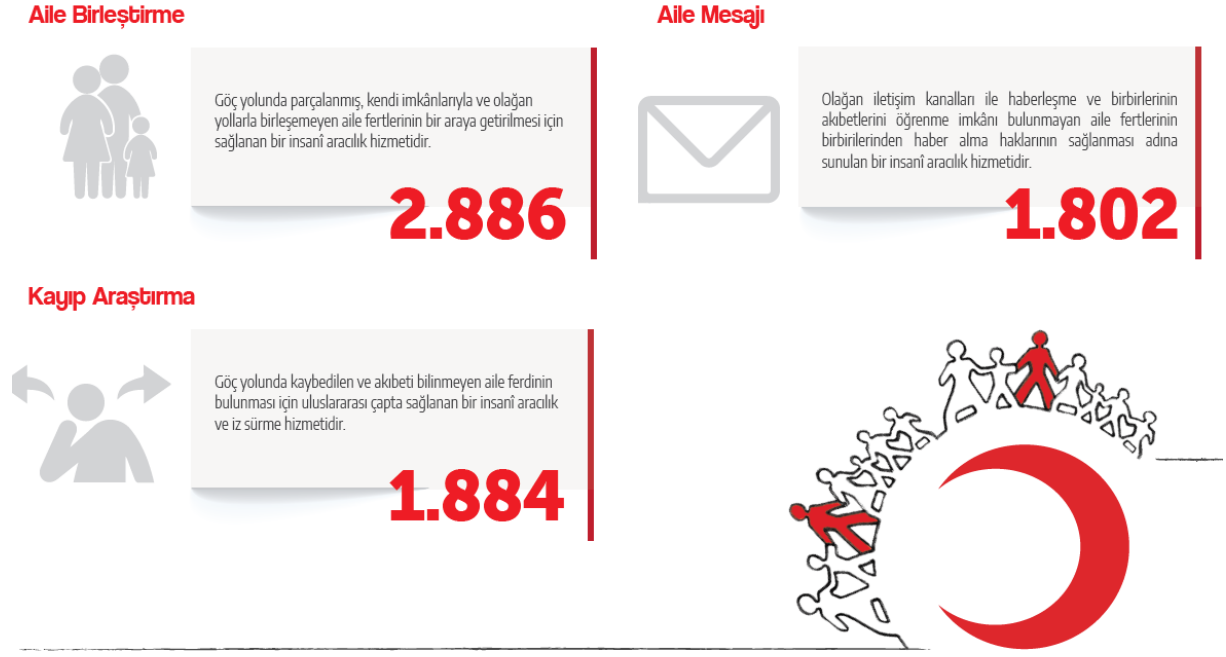
Travma, depresyon vb. gibi rahatsızların gözlemlendiği yararlanıcılar için bireysel danışmanlık, grup terapisi/grup çalışması gibi hizmetler verilmektedir. İleri düzey rahatsızlar için gerekli kurum ve kuruluşlara yönlendirmeler yapılmaktadır.

- Sağlık Eğitim Merkezleri

Yararlanıcılar için sağlık eğitim ve seminerleri verilmektedir.

3.3.4.Aile Bağlarının Yeniden Tesisi Programı

Aile Bağlarının Yeniden Tesisi, Uluslararası İnsancıl Hukuk, İnsan Hakları Beyannamesi gibi uluslararası yasal çerçeveler ve Türkiye Cumhuriyeti kanunlarının sağladığı haklar ile kişilerin aile birleşimi, kayıp araştırma, aile mesajı iletimi gibi haklardan yararlanması için aracı olmaktadır [45].



Şekil 11:Toplum Merkezleri Hizmetlerine İliřkin İstatistikler 4 [48]

Şekil 11’de Aile Bađlarının Yeniden Tesisi Programı tarafından, Toplum merkezlerinin kurulumundan günümüze kadar geçen sürede kaç kişiye hizmet götürüldüğü belirtilmektedir.

- Aile Birleřimi

Göç sırasında ya da sonrasında aile bireylerinden ayrılmak zorunda kalmış ya da aile bireyine ulařamayan yararlanıcılar için verilen bir hizmettir. Bu hizmet için aile bireyinin geçici ya da uluslararası koruma altında olması şartı aranmaktadır.

- Aile Mesajı İletimi

Aile bireyleri ile olağın kanallardan haberleřme sađlayamayan yararlanıcılar için verilen insani bir hizmettir. Bu kapsamdaki haberleřmeler genellikle salamat (selamların iletilmesi), iyilik durumu, telefon ya da video-konferans şeklindedir.

3.3.5.Sosyal Uyum Programı

Bu program kapsamında özellikle Suriye'deki savaş ortamında kaçarak ülkemize göç eden Suriyeli göçmeler ile yerel halkı kaynaştırmak ve kültürel bir ortak payda sağlamak amacıyla çeşitli hizmetler verilmektedir.

Sosyal ve Kültürel Etkinlikler



Sosyal Uyum Programı kapsamında sosyal uyum atölyeleri, spor, kültürel aktiviteleri, farkındalığı artırma ve gönüllülük çalışmaları düzenlenmektedir.

430.389

Danışma Kurulu



Danışma kurulu, yerel sosyal yapıyı bilen ve toplumdaki uyumun güçlendirilmesine katkı sağlayabilecek göçmen ve yerel halk üyelerinden oluşmaktadır. Katılımcılar aylık olarak düzenlenen toplantılarda Türk Kızılay Toplum Merkezleri tarafından sunulan hizmetlerle ilgili görüş, şikâyet ve geri bildirimlerini paylaşabilmektedir.

Danışma Kurulu Toplantı Sayısı	156	Genç Danışma Kurulu Toplantı Sayısı	28
Güncel Üye Sayısı	171	Güncel Üye Sayısı	77
Yerel Halk	67	Yerel Halk	5
Göçmen	104	Göçmen	72

Doğru Bilinen Yanlışlar



Toplumda yayılan doğru bilinen yanlışlar ya da söylentiler, yerel halk ile göçmenler arasında nefret söylemi ve çatışmalara neden olmaktadır. Doğru bilinen yanlışların giderilmesi ve söylentilerin ortadan kaldırılması amacıyla Türk Kızılay Toplum Merkezleri broşür, yüz yüze bilgilendirme, video ve seminerlerle toplumun doğru bilgiye ulaşmasını sağlamaktadır.

Seminer Sayısı	205
Ulaşılan Kişi Sayısı	7.876

Şekil 12:Toplum Merkezleri Hizmetlerine İlişkin İstatistikler 5 [48]

Şekil 12'de Sosyal Uyum Programı tarafından, Toplum merkezlerinin kurulumundan günümüze kadar geçen sürede kaç kişiye hizmet götürüldüğü belirtilmektedir.

Sosyal Uyum Programı kapsamında dört alt program bulunmaktadır.

- Bir Arada Yaşama Kültürünün Geliştirilmesi

Bir arada yaşama kültürünün geliştirilmesi, yerel halk ve göçmenlerin birlikte yaşamaları ve birbirlerini tanımalarını hedeflemektedir.

- Gönüllülük

Sosyal Uyum Programı kapsamında farklı alanlarda gönüllülük iş birlikleri ve kapasite geliştirme çalışmalar gerçekleştirilmektedir.

- Kamu ve Sivil Toplumun Güçlendirilmesi

Türkiye’de göçmenlerin yoğun yaşadığı bölgelerde yasal izinli olan ve bu minvalde faaliyet gösteren devletin güç konusundaki politikalarına katkı sunma amacıyla programlar yürüten yerelde bazı sivil toplum kuruluşları mevcuttur.

- Topluluk Katılımı ve Hesap Verilebilirlik

Toplum Merkezleri bünyesinde yürütülen programlarda iletişimi ve hesap verilebilirliği arttırmak çalışmaların özünü teşkil etmektedir.

4.UYGULAMA

Türk Kızılay Toplum Temelli Göç Programları Koordinatörlüğü'ne bağlı Raporlama ve İzleme Birimi'nden temin edilen ilk veri seti toplam 17.300 adet veri ve 24 adet değişkenden meydana gelmektedir. Çalışma kapsamında uyuşu Suriye olan göçmenler üzerinde çalışılacağı için uyruk değişkeni Suriye seçilerek, diğer uyuşuğa sahip yararlanıcıların verileri çalışma dışı bırakılmıştır. 24 değişken arasında yararlanıcılara ait ad, soyad, anne ismi, baba ismi, vatandaşlık numaraları ve adres bilgileri gibi kişisel veriler, kişisel verilerin korunması amacıyla çalışma dışı bırakılmıştır. Uyruk ve kişisel bilgilerin yer aldığı gözlemlerin çıkarıldığı veri setinde 10.372 veri ve 14 değişken yer almaktadır.

Veri setinde yer alan tüm değişkenler üzerinde etiketleme yapılmıştır. Bu etiketler değişkenler açıklanırken belirtilecektir.

4.1.Çalışmada Yer Alan Değişkenler

Bu bölümde çalışma kapsamında yer alan 14 değişken hakkında bilgi verilecektir.

- Cinsiyet

Cinsiyet değişkeni Kadın ve Erkek olacak biçimde etiketlenmiştir.

Tablo 5:Cinsiyet Değişkenine İlişkin Sıklık Tablosu

Cinsiyet	Sıklık	Yüzde (%)
Erkek	3.830	36,9
Kadın	6.542	63,1
Toplam	10.372	100

Tablo 5'de Cinsiyet değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada yer alan 10.372 yararlanıcının 6.542'si kadın, 3.830'u ise erkek yararlanıcılardır.

- Toplum Merkezleri

Etiketleme işlemi sonrasında her Toplum Merkezi'ne hizmet verdiği şehre göre hangi coğrafi bölgede yer alıyor ise o bölgenin ismi verilmiştir. Örneğin Mersin Toplum Merkezi etiketlenirken, Akdeniz etiketi almıştır. Aynı şekilde Ankara Toplum Merkezi'nin etiketi İç Anadolu olarak verilmiştir.

Tablo 6:Toplum Merkezi Değişkenine İlişkin Sıklık Tablosu

Toplum Merkezi	Sıklık	Yüzde(%)
Akdeniz	3.445	33,2
Güney Doğu Anadolu	2.720	26,2
İç Anadolu	2.040	19,7
Marmara	2.167	20,9
Toplam	10.372	100

Tablo 6'da Toplum Merkezleri değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada kullanılan veri setinde yer alan Toplum Merkezleri arasında, en çok yararlanıcıya Akdeniz bölgesinde hizmet vermekte olan Toplum Merkezleri'nde ulaşıldığı görülmektedir. Akdeniz bölgesinde hizmet almakta olan yararlanıcılar veri setinin %33,2'sini, İç Anadolu bölgesinde hizmet almakta olan yararlanıcılar %19,7'sini oluşturmaktadır. Akdeniz ve Marmara bölgesindeki yararlanıcı sayısının fazla olmasının sebebi deniz yolu ile farklı ülkelere göç imkanı vermesiyle, Güney Doğu Anadolu bölgesinde yer alan Toplum Merkezleri'ndeki yararlanıcı sayısının fazla oluş sebebi ise Suriye ile olan sınır komşuluğu durumuna bağlanabilmektedir.

- Hizmet Vermekte Olan Kurslar

Veri setinde yer alan tüm kurslar türlerine göre etiketlenmiştir. Örneğin Şiş Örgü, Takı Tasarımı gibi kurslar El Sanatları vb. Kurslar olarak etiketlenirken, Mantar Yetiştiriciliği, Sert Çekirdekli Meyve Yetiştiriciliği vb. kurslar Yetiştiricilik/Tarımsal Kurslar olarak etiketlenmiştir.

Tablo 7:Kurs Adı Değişkenine İlişkin Sıklık Tablosu

Kurs Adı	Sıklık	Yüzde (%)
El Sanatları vb. Kurslar	1.554	15
Girişimciliği Destekleyen Kurslar	7.732	74,5

Yemek/Pastacılık Kursları	127	1,2
Yetiştiricilik/Tarımsal Kurslar	959	9,2
Toplam	10.372	100

Tablo 7’de Kurs Adı değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. En çok tercih edilen kursun %74,5’lik bir oranla Girişimciliği Destekleyen Kurslar olduğu görülmektedir. Girişimciliği Destekleyen Kurslar istihdama yönelik olmaları sebebiyle yararlanıcılar tarafından sıkça tercih edilmektedir. %1,2’lik bir oranla en az tercih edilen kursun Yemek/Pastacılık Kursları olduğu görülmektedir. Bu kurslara katılımın düşük olması, genellikle ev hanımı olan yararlanıcılar tarafından tercih edilmesi ile ilişkilendirilebilmektedir.

- Eğitim Seviyesi

Eğitim seviyesi değişkeni İlk/Ortaokul ve Dengi, Lise ve Dengi, Üniversite ve Dengi ve Diğer olacak biçimde 4 farklı etikete sahiptir.

Tablo 8:Eğitim Seviyesi Değişkenine İlişkin Sıklık Tablosu

Eğitim Seviyesi	Sıklık	Yüzde (%)
Diğer	435	4,2
İlk/Ortaokul ve Dengi	9.382	90,5
Lise ve Dengi	547	5,3
Üniversite ve Dengi	8	0,1
Toplam	10.372	100

Tablo 8’de Eğitim Seviyesi değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Burada yer alan “Diğer” etiketi altında yararlanıcının ülkesinde ya da Türkiye’de MEB onaylı olmayan, belgelenemeyen bir kurum aracılığı ile eğitim aldığı durum belirtilmektedir. Yararlanıcıların %90,5’i ilk ve ortaokul dengi eğitim seviyesine sahip iken, sadece %0,1’lik bir kısmı üniversite ve dengi eğitim seviyesine sahip olduğu görülmektedir.

- Bildiği Dil

Çalışmada yer alan yararlanıcılar Arapça, Türkçe, İngilizce ve Kürtçe dillerini bilmektedir.

Tablo 9:Bildiği Dil Değişkenine İlişkin Sıklık Tablosu

Bildiği Dil	Sıklık	Yüzde (%)
Arapça	7.635	73.6
İngilizce	308	3
Türkçe	2.207	21.3
Kürtçe	222	2.1
Toplam	10.372	100

Tablo 9’da Bildiği Dil değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Yararlanıcıların %73,6’sı anadili olan Arapça dilini kullanmaktadır. Türkçe dilini kullanan yararlanıcılar ise veri setinin %2,1’ini oluşturmaktadır.

- Sağlık ve Engel Durumu

Bu değişken ise Akut Sağlık Sorunu, Engelli, Kronik Sağlık Sorunu ve Sağlık Sorunu Yok olacak biçimde etiketlenmiştir.

Tablo 10:Sağlık ve Engel Durumu Değişkenine İlişkin Sıklık Tablosu

Sağlık ve Engel Durumu	Sıklık	Yüzde (%)
Akut Sağlık Sorunu	103	1
Engelli	50	0,3
Kronik Sağlık Sorunu	95	0,9
Sağlık Sorunu Yok	10.144	97,8
Toplam	10.372	100

Tablo 10’da Sağlık ve Engel Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada yer alan yararlanıcıların %97,8’inde herhangi bir sağlık sorunu görülmemektedir, %1’lik bir kısımda ise Akut Sağlık Sorunu görülmektedir.

- Kurs Ücret Durumu

Toplum Merkezleri'nde hizmet vermekte olan kursların ücretli olup olmadıklarına göre etiketleme işlemi yapılmıştır. Kurs ücretli ise Evet, ücretli değil ise Hayır etiketi verilmiştir.

Tablo 11:Kurs Ücret Durumu Değişkenine İlişkin Sıklık Tablosu

Kurs Ücret Durumu	Sıklık	Yüzde (%)
Evet	2.061	19,9
Hayır	8.311	80,1
Toplam	10.372	100

Tablo 11'de Kurs Ücret Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Toplum Merkezleri'nde hizmet vermekte olan kursların %80,1'i ücretsiz kurslardır. Bu şekilde yararlanıcıların daha çok katılım sağlamaları ve zorlanmadan eğitim alabilmeleri sağlanmaya çalışılmaktadır.

- Kurs Bırakma Durumu

Kurs Bırakma Durumu değişkeni Evet ve Hayır olacak biçimde etiketlenmiştir. Eğer yararlanıcı kursu bıraktıysa Evet, kursa devam ediyorsa Hayır cevabı alınmıştır.

Tablo 12:Kurs Bırakma Durumu Değişkenine İlişkin Sıklık Tablosu

Kurs Bırakma Durumu	Sıklık	Yüzde (%)
Evet	425	4,1
Hayır	9.947	95,9
Toplam	10.372	100

Tablo 12'de Kurs Bırakma Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Burada elde edilen "Evet" ve "Hayır" cevapları, veriler toplanırken hizmet vermekte olan kurslar tamamlandıktan sonra elde edilmiştir. Yararlanıcıların %95,9'luk kısmı hizmet almakta oldukları kurslara devam etmektedir.

- Çalışma Durumu

Yararlanıcıların çalışma durumları Çalışıyor, Çalışmıyor, İş Arıyor ve İşveren olarak etiketlenmiştir.

Tablo 13:Çalışma Durumu Değişkenine İlişkin Sıklık Tablosu

Çalışma Durumu	Sıklık	Yüzde (%)
Çalışıyor	2.197	21,2
Çalışmıyor	5.901	56,9
İş Arıyor	2.225	21,5
İşveren	49	0,5
Toplam	10.372	100

Tablo 13’de Çalışma Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada yer alan yararlanıcıların %56,9’u çalışmıyor, %21,5’i iş arıyor iken %0,5’lik bir kısım işveren konumundadır.

- Gelir Türü

Gelir Türü değişkeni Düzenli, Düzensiz ve Geliri Yok şeklinde etiketlenmiştir.

Tablo 14:Gelir Türü Değişkenine İlişkin Sıklık Tablosu

Gelir Türü	Sıklık	Yüzde (%)
Düzenli Gelir	181	1,7
Düzensiz Gelir	10.062	97
Geliri Yok	129	1,2
Toplam	10.372	100

Tablo 14’te Gelir Türü değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. “Düzensiz Gelir” cevabına ilişkin olarak; yararlanıcıların tarlalarda ya da benzeri geçici/mevsizlik iş kollarında elde ettikleri gelir türünü temsil etmektedir. 10.372 yararlanıcının %97’sinin düzenli bir gelir bulunmamaktadır. Bu durum yararlanıcıların %56,9’unun çalışmıyor olması ile ilişkilendirilebilmektedir.

- Medeni Hali

Bu deęişken ise Bekar/Nişanlı, Bilmiyor, Boşanmış/Eşinden Ayrı ve İmam/Resmi Nikahlı olacak biçimde etiketlenmiştir.

Tablo 15:Medeni Hali Deęişkenine İlişkin Sıklık Tablosu

Medeni Durumu	Sıklık	Yüzde (%)
Bekar/Nişanlı	2.484	23,9
Bilmiyor	293	2,8
Boşanmış/Eşinden Ayrı	294	2,8
İmam/Resmi Nikahlı	7.301	70,4
Toplam	10.372	100

Tablo 15’de Medeni Durum deęişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Yararlanıcıların %70,4’ü imam nikahı ya da resmi nikah ile evli durumda iken, %23,9’unun bekar ya da nişanlı olduęu görülmektedir. Yararlanıcıların %2,8’inin medeni durumunu bilmedięi görülmektedir, bu durum resmi ya da imam nikahı kıyılmadığını ancak hayatlarında birinin olduęunu temsil etmektedir.

- Sertifika Alma Durumu

Eđer yararlanıcı kursu tamamlayıp sertifika almaya hak kazanmış ise Evet, kursu tamamlayamamış ise Hayır cevapları etiketlenmiştir.

Tablo 16:Sertifika Alma Durumu Deęişkenine İlişkin Sıklık Tablosu

Sertifika Alma Durumu	Sıklık	Yüzde (%)
Evet	10.178	98,1
Hayır	194	1,9
Toplam	10.372	100

Tablo 16’da Sertifika Alma Durumu deęişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada yer alan yararlanıcıların %98,1’i kursu tamamlayarak sertifika almaya hak kazanmıştır.

- İstihdam Alanı ile Kurs İlgisi Durumu

Eğitimi alınan kurs ile yararlanıcının istihdam edildiği iş kolu arasında bir ilişki var ise Evet, yok ise Hayır olacak biçimde etiketleme işlemi yapılmıştır.

Tablo 17:İstihdam Alanı ile Kurs İlgisi Durumu Değişkenine İlişkin Sıklık Tablosu

İstihdam Alanı ile Kurs İlgisi Durumu	Sıklık	Yüzde (%)
Evet	10.348	99,8
Hayır	24	0,2
Toplam	10.372	100

Tablo 17’de İstihdam Alanı ile Kurs İlgisi Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Eğitim aldıkları kurs ile istihdam edildikleri iş kolu arasında ilişki olan yararlanıcılar veri setinin %99,8’ini oluşturmaktadır.

- Yaş

Yaş değişkeni, en küçük ve en büyük yaş değerlerini kapsayacak biçimde kategorize edilmiştir. Etiketlemeler Tablo 18’deki gibidir.

Tablo 18:Yaş Değişkenine İlişkin Sıklık Tablosu

Yaş	Sıklık	Yüzde (%)
18-27	2.385	23
28-37	5.031	48,5
38-47	1.767	17
48-57	927	8,9
58-67	230	2,2
68-77	28	0,3
78-87	4	0,0
Toplam	10.372	100

Tablo 18’de Yaş Durumu değişkenine ait sıklık ve yüzdeler bilgileri yer almaktadır. Çalışmada yer alan yararlanıcıların %48,5’i 28-37 yaş aralığında, %23’ü ise 18-27 yaş aralığındadır.

Tablo 19:Cinsiyet*İstihdam Alanı ile Kurs İlgili Durumu Değişkenlerine Ait Çapraz Tablo

Cinsiyet*İstihdam Alanı ile Kurs İlgili Durumu					
			İstihdam Alanı ile Kurs İlgili Durumu		Toplam
			Evet	Hayır	
Cinsiyet	Erkek	Değer	3.816	14	3.830
		% Cinsiyet	99,6	0,4	100
		% İaikid	36,9	58,3	36,9
		%Toplam	36,8	0,1	36,9
	Kadın	Değer	6.532	10	6.542
		% Cinsiyet	99,8	0,2	100
		% İaikid	63,1	41,7	63,1
		%Toplam	63	0,1	63,1
Toplam	Değer	10.348	24	10.372	
	% Cinsiyet	99,8	0,2	100	
	% İaikid	100	100	100	
	%Toplam	99,8	0,2	100	

Tablo 19'da Cinsiyet değişkeni ile İstihdam Alanı ile Kurs İlgili Durumu değişkenlerinin oluşturduğu çapraz tablo yer almaktadır. Tabloda yer alan "İaikid" ifadesi İstihdam Alanı ile Kurs ilgili Durumu değişkeninin baş harflerini temsil etmektedir. Çapraz tablodan yola çıkarak; kadın yararlanıcıların %99,8'inin istihdam edildikleri iş ile hizmet aldıkları kurs arasında ilgi bulunmaktadır ve erkek yararlanıcıların %99,6'sının istihdam edildikleri iş ile hizmet aldıkları kurs arasında ilgi bulunmaktadır yorumları yapılabilmektedir. Çapraz tabloya ait pearson ki kare değeri 4,733 ve p değeri ise 0,03 olarak elde edilmiştir ($\alpha=0,05$). Cinsiyet değişkeni ile İstihdam Alanı ile Kurs İlgili Durumu değişkenlerine ait gruplar istatistiksel olarak anlamlı bir şekilde farklılaşmaktadır.

4.2. Veri Madenciliği Modellemeleri

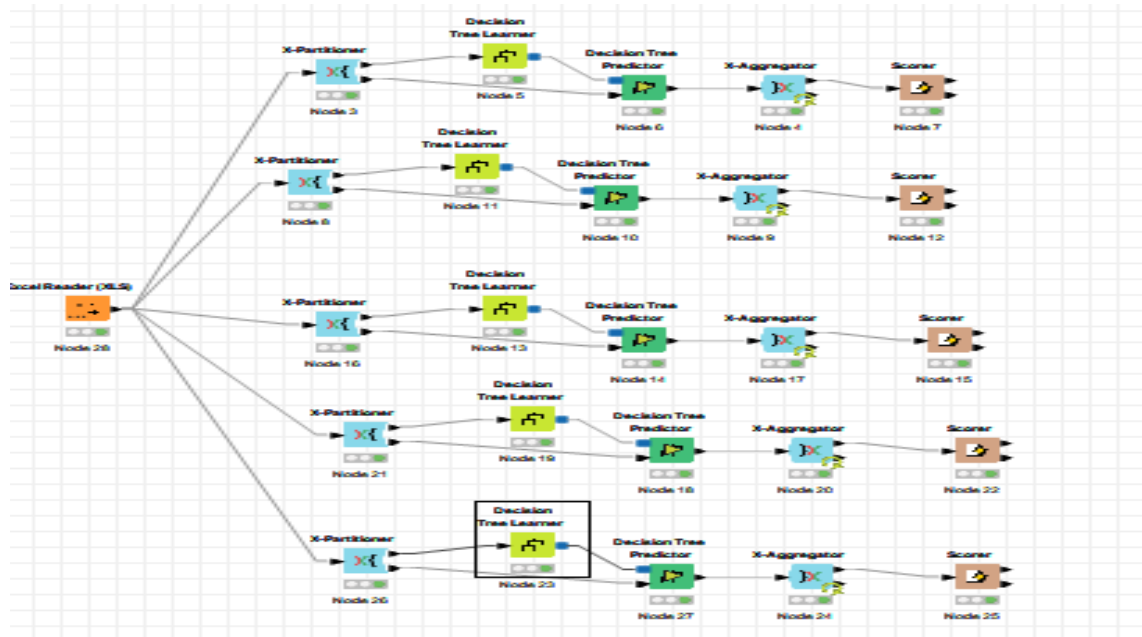
Veri madenciliği modellemeleri KNIME yazılımının 4.2.2 sürümü ile gerçekleştirilmiştir. Bu bölümde karar ağaçları ve destek vektör makineleri modellerine ait sonuçlar üzerinde durulacaktır.

Çalışma kapsamında hedef değişken olarak yararlanıcının eğitim aldığı kurs sonrası istihdam edildiği iş kolu ile arasındaki ilişki durumu seçilmiştir. Çalışma kapsamında Toplum Merkezleri'nde verilen eğitimin yararlanıcılar üzerindeki etkisi araştırıldığı için bu değişken hedef değişken olarak seçilmiştir. Çalışmada yararlanıcıların eğitim aldıkları kurs ile istihdam edildikleri iş kolu arasında ilişki olması durumuna odaklanılmaktadır, bu sebeple hedef değişkene ait "Evet" cevabına ait sonuçlar üzerinde durulacaktır. "Evet" cevabı yararlanıcının aldığı kurs ile ilişkili bir iş kolunda çalıştığı anlamına gelmektedir.

KNIME yazılımı üzerinde veri setini eğitim ve test kümeleri olarak ikiye ayırma işlemi için "X-Partitioner" düğümü kullanılmıştır. Bu düğüm sayesinde veri setinde yer alan veriler algoritmaya alınırken eğitim ve test kümesi olarak ikiye ayrılmaktadır, sonrasında ise "X-Aggregator" düğümü kullanılarak elde edilen sonuçları değerlendirmeden önce veri setinin tekrar birleşmesi sağlanmıştır. Örneğin K=5 katlı çapraz doğrulama işlemi sırasında "X-Partitioner" düğümü kullanılarak veri seti eğitim kümesi= 8.298, test kümesi= 2.074 adet olacak şekilde ikiye bölünmüştür. Algoritma sonuçları incelenmeye başlamadan önce "X-Aggregator" düğümü kullanılarak tekrar veri seti kadar yani 10.372 adet veri ile sonuçlar incelenebilmiştir.

4.2.1. Karar Ağaçları

Bu bölümde karar ağaçları modeline ilişkin KNIME yazılımı aracılığı ile elde edilen çıktılar ve yorumları üzerinde durulacaktır.



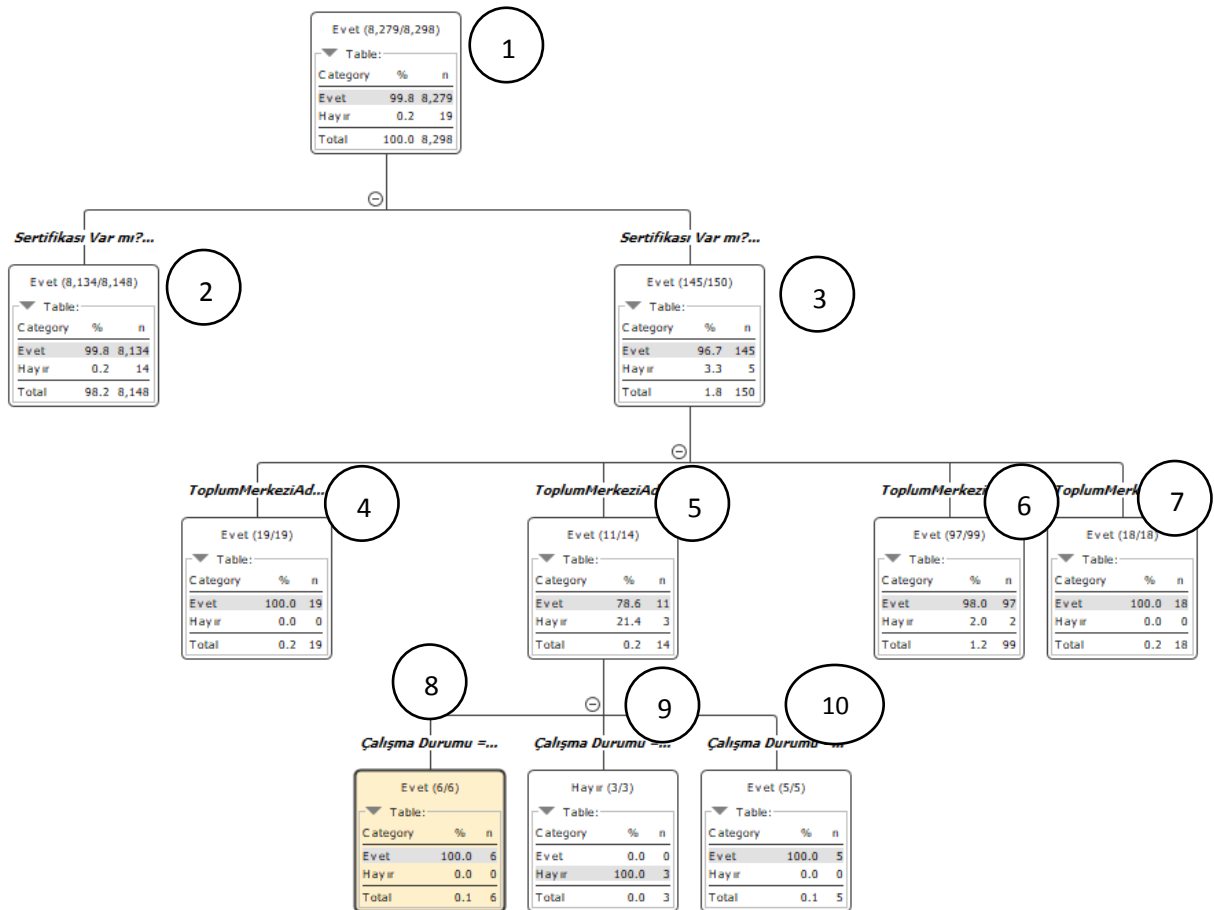
Şekil 13: Karar Ağaçları Model Akış Şeması

Şekil 13’de KNIME yazılımı üzerinde çalıştırılan algoritmaların bir kısmı yer almaktadır. Her algoritma için ayırma kriteri ve K-katlı çapraz doğrulama sayısı değişiklik göstermektedir.

Tablo 20: Karar Ağaçları Model Sonuçları

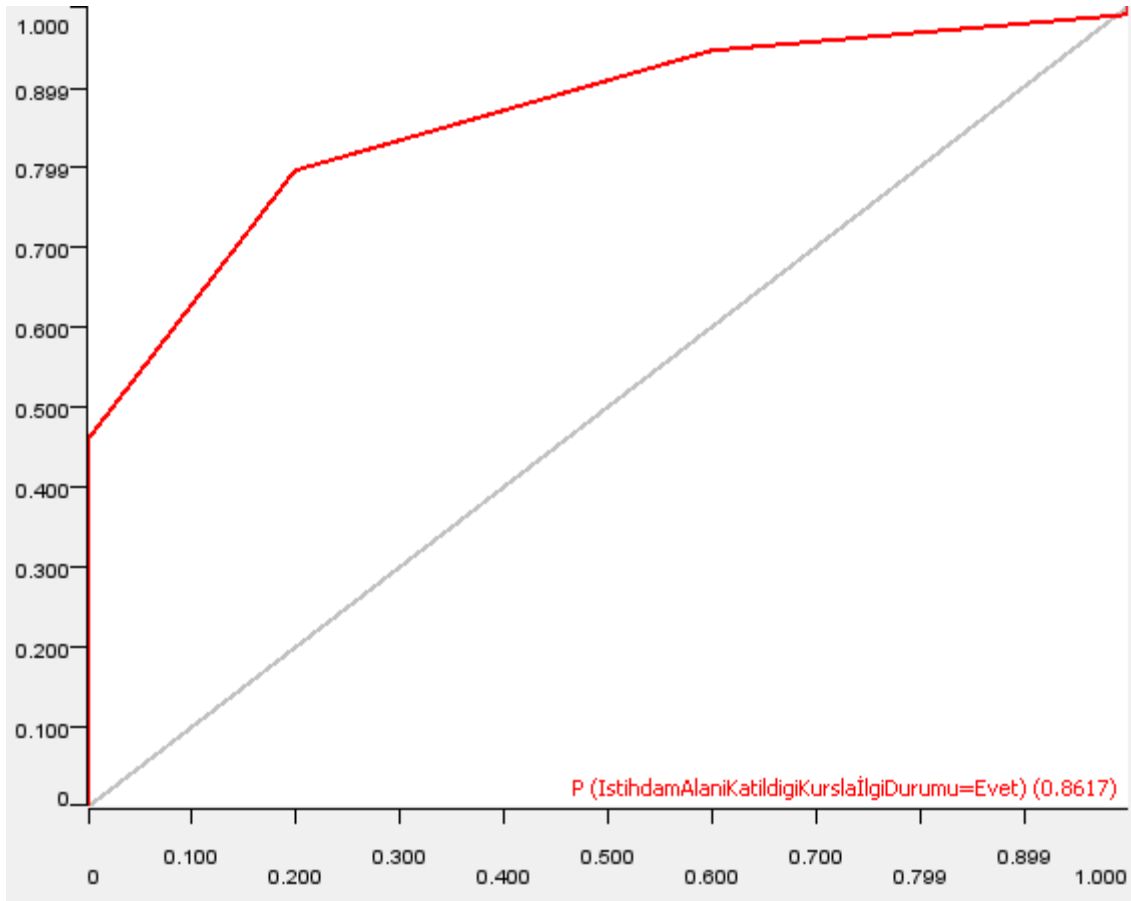
K-Katlı Çapraz Doğrulama	Ayırma Kriteri	Eğitim Kümesi	Test Kümesi	Doğruluk	Hata Oranı	Kesinlik
K=5	Gini İndeksi	8.298	2.074	99,77%	0,0022	99,93%
	Kazanç Oranı (Gain Ratio)	8.298	2.074	99,79%	0,0020	99,95%
K=10	Gini İndeksi	9.335	1.037	99,77%	0,0022	99,93%
	Kazanç Oranı (Gain Ratio)	9.335	1.037	99,79%	0,0020	99,95%
K=20	Gini İndeksi	9.854	518	99,79%	0,0020	99,95%
	Kazanç Oranı (Gain Ratio)	9.854	518	99,79%	0,0020	99,95%
K=30	Gini İndeksi	10.027	345	99,79%	0,0020	99,95%
	Kazanç Oranı (Gain Ratio)	10.027	345	99,79%	0,0020	99,95%

Tablo 20’de farklı ayırma kriterleri ve K-katlı çapraz doğrulama sayısına ait algoritmalar için doğruluk, hata oranı ve kesinlik değerleri yer almaktadır. Ayırma kriteri olarak gini indeksi ve K=5 katlı çapraz doğrulama durumunda doğruluk değeri %99,77 iken, ayırma kriteri olarak kazanç oranı (gain ratio) seçildiğinde doğruluk değeri %99,79 olarak hesaplanmaktadır. K değeri arttıkça eğitim kümesindeki veri sayısının arttığı, test kümesindeki veri sayısının ise azaldığı görülmektedir. K değerindeki artışların doğruluk, hata oranı ve kesinlik değerlerinde etkisinin %0,02 düzeyinde olduğu görülmektedir. K=5 ve K=10 katlı çapraz doğrulama durumları göz önüne alındığında ayırma kriteri olarak kazanç oranı (gain ratio) tercih edildiğinde doğruluk değerinin gini indeksi’ne göre daha yüksek olduğu görülmektedir. Elde edilen doğruluk ve hata oranı değerleri göz önüne alındığında karar ağacı modellerinin çalışma kapsamında kullanılan veri seti için uygun bir modeller olduğu görülmektedir.



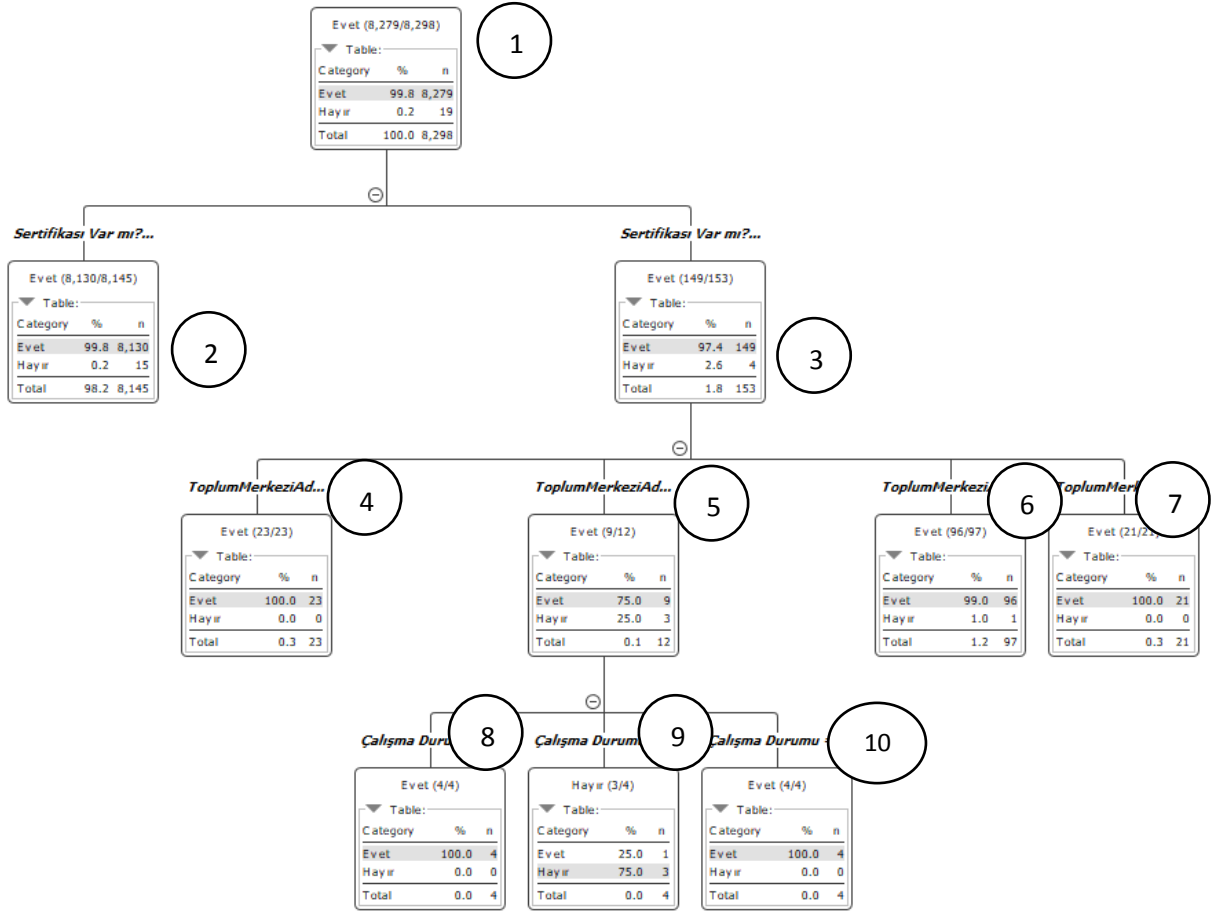
Şekil 14: Karar Ağacı Diyagramı 1

Şekil 14'te K=5 katlı çapraz doğrulama ve ayırma kriteri olarak gini indeksi seçilmiş algoritmaya ait karar ağacı diyagramı yer almaktadır. Karar ağacında kök düğümün, hedef değişkeni olarak belirlenen “İstihdam Alanı ile Kurs İlgisi Durumu” değişkeni olduğu görülmektedir. Bu değişkene ait “Evet” cevabına ilişkin olarak (eğitim aldığı kurs ile ilişkili bir iş kolunda istihdam edilmiş yararlanıcılara ilişkin olarak); eğitim kümesinde yer alan 8298 yararlanıcının % 99,8'inin sertifika almış olduğu görülmektedir (2 numaralı düğüm). Sertifika alamamış olan 155 yararlanıcının (3 numaralı düğüm) 21'i İç Anadolu bölgesindeki toplum merkezlerinde (4 numaralı düğüm), 15'i Marmara bölgesindeki toplum merkezlerinde (5 numaralı düğüm), 97'si Güney Doğu Anadolu bölgesindeki toplum merkezlerinde (6 numaralı düğüm) ve 22'si Akdeniz bölgesindeki toplum merkezlerindeki (7 numaralı düğüm) kurslarda eğitim almaktadırlar. Marmara bölgesindeki toplum merkezlerindeki kurslarda eğitim alan 15 yararlanıcının 5'i çalışmıyor (8 numaralı düğüm), 5'i çalışıyor (9 numaralı düğüm) ve 5'i iş arıyor (10 numaralı düğüm) oldukları görülmektedir.



Şekil 15:K=5 Katlı Çapraz Doğrulama ve Gini İndeksi Kullanıldığı Durum İçin ROC Eğrisi

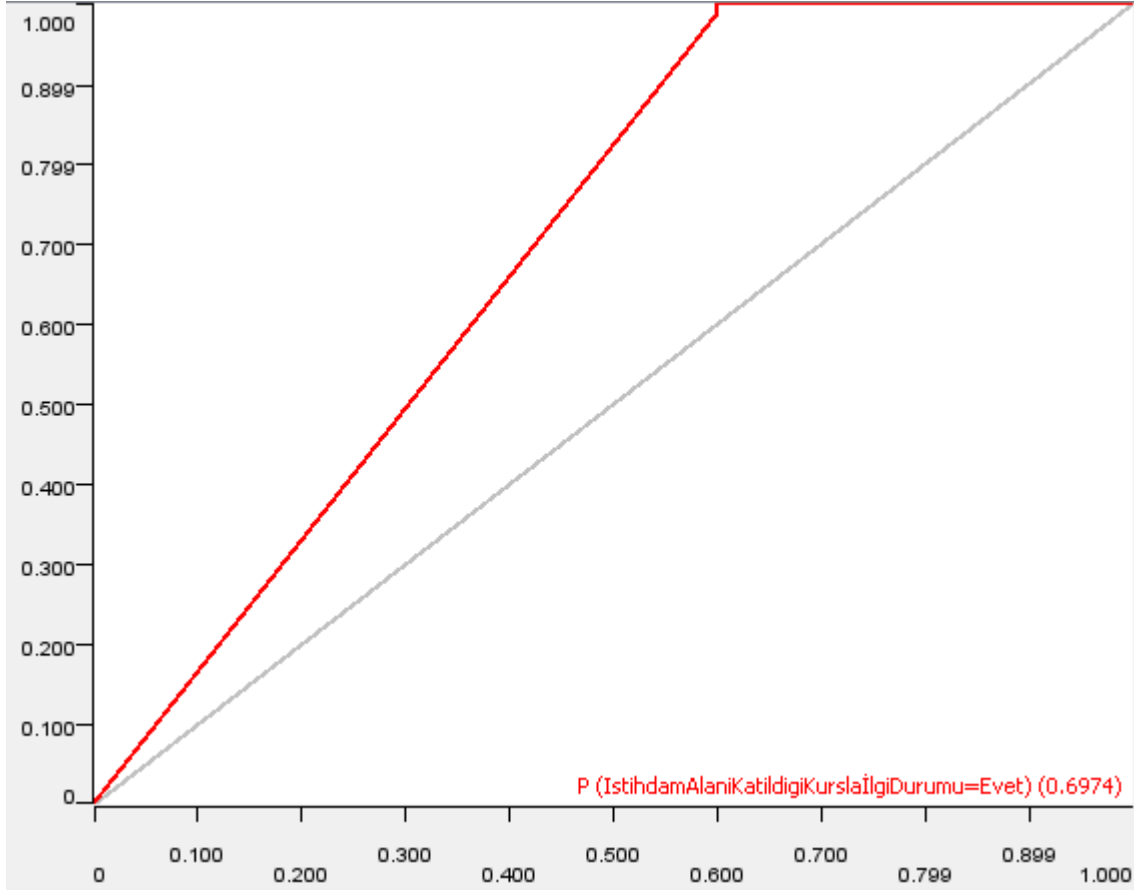
Şekil 15’de hedef değişken ve hedef değişkenin tahmin değerlerine ait ROC eğrisi yer almaktadır. Eğrinin altında kalan alan 0,8617 olarak elde edilmiştir.



Şekil 16: Karar Ağacı Diyagramı 2

Şekil 15’de K=5 katlı çapraz doğrulama ve ayırma kriteri olarak kazanç oranı (gain ratio) seçilmiş algoritmaya ait karar ağacı diyagramı yer almaktadır. Hedef değişkene ait “Evet” cevabına ilişkin olarak; eğitim kümesinde yer alan 8298 yararlanıcının % 99,8’inin sertifika almış olduğu görülmektedir (2 numaralı düğüm). Sertifika alamamış olan 153 yararlanıcının (3 numaralı düğümde görülmektedir) 23’ü İç Anadolu bölgesindeki toplum merkezlerinde (4 numaralı düğüm), 12’si Marmara bölgesindeki toplum merkezlerinde (5 numaralı düğüm), 97’si Güney Doğu Anadolu bölgesindeki toplum merkezlerinde (6 numaralı düğüm) ve 21’i Akdeniz bölgesindeki toplum merkezlerindeki (7 numaralı düğüm) kurslarda eğitim almaktadırlar. Marmara bölgesindeki toplum merkezlerindeki kurslarda eğitim alan 12 yararlanıcının 4’ü çalışmıyor (8 numaralı düğüm), 4’ü çalışıyor (9 numaralı düğüm) ve 4’ü iş arıyor (10 numaralı düğüm) oldukları görülmektedir.

K=5 katlı çapraz doğrulama ve farklı ayırma kriterlerine göre elde edilen iki karar ağacında alınan sonuçların aynı doğrultuda oldukları, sadece yüzdesel olarak farklılık göstermekle birlikte dallanma şekillerinin aynı olduğu görülmektedir. Ayırma kriteri olarak gini indeksi seçildiğinde sertifika alamamış yararlanıcı sayısı 155 iken, ayırma kriteri kazanç oranı (gain ratio) seçildiğinde bu sayının 153 olduğu görülmektedir.



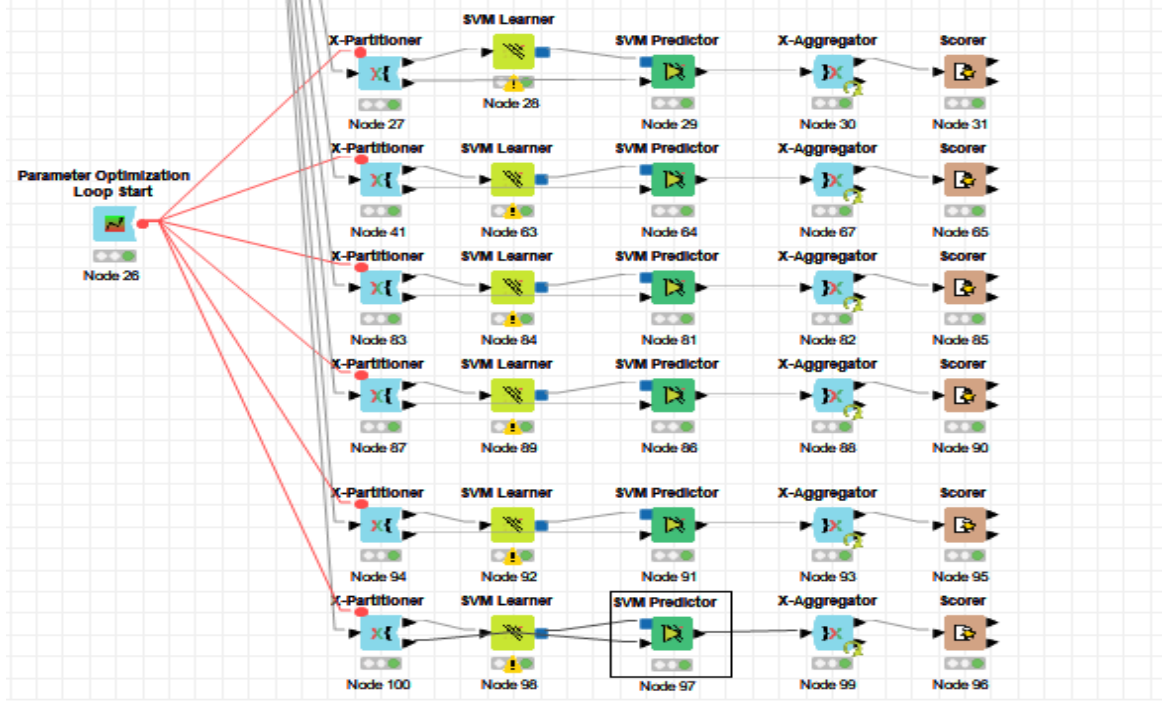
Şekil 17:K=5 Katlı Çapraz Doğrulama ve Kazanç Oranı Kullanıldığı Durum İçin ROC Eğrisi

Şekil 17’de hedef değişken ve hedef değişkenin tahmin değerlerine ait ROC eğrisi yer almaktadır. Eğrinin altında kalan alan 0,6974 olarak elde edilmiştir.

Elde edilen iki ROC eğrisi altında kalan alanlar değerlendirildiğinde, 1’e daha yakın bir değer olması sebebiyle K=5 ve gini indeksinin kullanıldığı durum tercih edilmelidir şeklinde yorumlanabilir.

4.2.2. Destek Vektör Makineleri

Bu bölümde destek vektör makineleri modeline ilişkin KNIME yazılımı aracılığı ile elde edilen çıktılar ve yorumları üzerinde durulacaktır.



Şekil 18: Destek Vektör Makineleri Model Akış Şeması

Şekil 16'da KNIME yazılımı üzerinde çalıştırılan algoritmaların bir kısmı yer almaktadır. Her algoritma için çekirdek fonksiyon ve K-katlı çapraz doğrulama sayısı değişiklik göstermektedir.

Tablo 21: Destek Vektör Makineleri Model Sonuçları (Çekirdek Fonksiyon: Polinom)

K-Kath Çapraz Doğrulama	Çekirdek Fonksiyon	Sapma	Gama Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=5	Polinom (Polynomial)	1	1	8.298	2.074	99,77%	0,23%
		0,9	1	8.298	2.074	99,77%	0,23%
		0,5	1	8.298	2.074	99,77%	0,23%
		1	0,5	8.298	2.074	99,77%	0,23%
		1	1,5	8.298	2.074	99,77%	0,23%
K=10	Polinom (Polynomial)	1	1	9.335	1.037	99,77%	0,23%
		0,9	1	9.335	1.037	99,77%	0,23%
		0,5	1	9.335	1.037	99,77%	0,23%
		1	0,5	9.335	1.037	99,77%	0,23%
		1	1,5	9.335	1.037	99,77%	0,23%
K=20	Polinom (Polynomial)	1	1	9.854	518	99,77%	0,23%
		0,9	1	9.854	518	99,77%	0,23%
		0,5	1	9.854	518	99,77%	0,23%
		1	0,5	9.854	518	99,77%	0,23%
		1	1,5	9.854	518	99,77%	0,23%

		1	1,5	9.854	518	99,77%	0,23%
K-Katlı Çapraz Doğrulama	Çekirdek Fonksiyon	Sapma	Gama Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=30	Polinom (Polynomial)	1	1	10.027	345	99,17%	0,83%
		0,9	1	10.027	345	99,77%	0,23%
		0,5	1	10.027	345	99,77%	0,23%
		1	0,5	10.027	345	99,77%	0,23%
		1	1,5	10.027	345	99,77%	0,23%

Tablo 21’de farklı K-katlı çapraz doğrulama sayılarına ilişkin polinom çekirdek türüne ait sonuçlar yer almaktadır. Her K-katlı çapraz doğrulama denemesinde sapma değeri ve gama değeri farklılık göstermektedir. K=5 katlı çapraz doğrulama için; sapma değeri için sırasıyla 0,5, 0,9 ve 1 değerleri, gama değeri için ise sırasıyla 1, 0,5, 1,5 değerleri denemiştir. Bu denemeler sonucunda elde edilen doğru sınıflandırma yüzdesi açısından herhangi bir farklılık gözlemlenmemiştir.

Her bir denemede elde edilen doğru sınıflandırma yüzdesi K=5 katlı çapraz doğrulama için %99,77 olarak elde edilmiştir. K=10 ve K=20 değerleri için yapılan denemeler sonucu elde edilen doğru sınıflandırma yüzdesinin, K=5 değeri için elde edilen sonuçlar ile aynı olduğu görülmektedir. K=30 katlı çapraz doğrulamaya ait sonuçlarda, sapma değeri=1 ve gama değeri=1 iken doğru sınıflandırma yüzdesi %99,17 olarak elde edilmiştir. Eğitim kümesinde yer alan veri sayısı arttıkça doğru sınıflandırma yüzdesinde herhangi bir artış gözlemlenmemiştir, sadece K=30, sapma değeri=1 ve gama değeri=1 iken alınan sonuçta doğru sınıflandırma yüzdesi %99,17’e gerilemiştir.

Tablo 22: Destek Vektör Makineleri Model Sonuçları (Çekirdek Fonksiyon: RBF)

K-Katlı Çapraz Doğrulama	Çekirdek Fonksiyon	Sigma Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=5	Radyal	0,1	8.298	2.074	99,77%	0,23%
	Tabanlı Fonksiyon (RBF)	0,3	8.298	2.074	99,77%	0,23%
		0,5	8.298	2.074	99,16%	0,84%
		0,7	8.298	2.074	99,34%	0,66%
K-Katlı Çapraz Doğrulama	Çekirdek Fonksiyon	Sigma Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=10	Radyal	0,1	9.335	1.037	99,77%	0,23%
	Tabanlı Fonksiyon (RBF)	0,3	9.335	1.037	99,77%	0,23%
		0,5	9.335	1.037	99,57%	0,43%
		0,7	9.335	1.037	99,77%	0,23%
K-Katlı Çapraz Doğrulama	Çekirdek Fonksiyon	Sigma Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=20	Radyal	0,1	9.854	518	92,26%	7,74%
	Tabanlı Fonksiyon (RBF)	0,3	9.854	518	98,29%	1,71%
		0,5	9.854	518	99,45%	0,55%
		0,7	9.854	518	99,77%	0,23%
K-Katlı Çapraz Doğrulama	Çekirdek Fonksiyon	Sigma Değeri	Eğitim Kümesi	Test Kümesi	Doğruluk (%)	Hata Oranı (%)
K=30	Radyal	0,1	10.027	345	99,62%	0,38%
	Tabanlı Fonksiyon (RBF)	0,3	10.027	345	99,72%	0,28%
		0,5	10.027	345	99,43%	0,57%
		0,7	10.027	345	99,77%	0,23%

Tablo 22’de farklı K-katlı çapraz doğrulama sayılarına ve sigma değerlerine ilişkin radyal tabanlı fonksiyon çekirdek türüne ait sonuçlar yer almaktadır. Her K-katlı çapraz doğrulama denemesinde sigma değeri farklılık göstermektedir. Her bir K-katlı çapraz doğrulama

denemesinde, sigma deęerleri sırasıyla 0,1, 0,3, 0,5 ve 0,7 olarak alınmıştır. K=5, 10 ve 30 deęerleri için yapılan denemelerde elde edilen doęruluk yüzdeleri %99'un üzerinde iken K=20 deęeri için yapılan denemelerde, sigma=0,1 ve 0,3 deęerleri için elde edilen doęru sınıflandırma yüzdelerinin sırasıyla %92,26 ve %98,29'a geriledięi görölmektedir. Elde edilen doęru sınıflandırma yüzdeleri göz önüne alındığında destek vektör makineleri modellerinin çalışma kapsamında kullanılan veri seti için uygun bir modeller olduęu görölmektedir.

5.SONUÇ

2011 Yılında Suriye’de meydana gelen iç karışıklıklar sebebi ile milyonlarca Suriye vatandaşı memleketlerini terk etmek zorunda kaldı. Birçok ülke göçmenlere kapılarını açtı ancak en çok göçmene ev sahipliği yapan ülkenin Türkiye olduğunu söylemek yanlış olmayacaktır. 2011 yılından itibaren, ülkemizde ikamet etmekte olan göçmenlerin Türkiye’de yaşadıkları hayatın iyileştirilmesi ve yerel halk ile kaynaştırılmaları amacıyla pek çok hizmet farklı kanallar aracılığı ile yerine getirilmektedir. Verilmekte olan bu hizmetlerin gerek iyileştirilmesi gerekse farklı hizmet ihtiyaçlarının belirlenmesi hususunda veri madenciliğini kullanmak yararlı olacaktır.

Çalışmada kullanılan veri setinde, Türk Kızılay tarafından hizmet vermekte olan Toplum Merkezleri’nde, yararlanıcı olarak hizmet alan bireylerin bilgilerinden yararlanılmıştır. Bireysel bilgilerin gizliliğini koruma amacı ile hiçbir kişisel veri çalışmada kullanılmamıştır. Bireylerin hangi Toplum Merkezleri’nde hizmet aldıklarından, hangi kurs hizmetinden yararlandıklarına kadar pek çok bilgi çalışmada yer almaktadır. Bu bilgiler ışığında veri madenciliği yöntemleri uygulanarak, çalışmada kullanılan veri seti gibi bir veri setinde hangi modelin kullanılması daha uygun olacaktır sorusunun cevabı aranmıştır. Karar ağaçları ve destek vektör makineleri modelleri çalışılmış ve elde edilen sonuçlar tablolar haline getirilmiştir.

Karar ağacı modeli uygulamasında, farklı ayırma kriterleri ve K-katlı çapraz doğrulama durumları altında elde edilen doğruluk, hata oranı ve kesinlik değerleri tablo haline getirilmiştir. Uygulama bölümünde yer alan Tablo 20’de bu değerler yer almaktadır. Tablo 20’de yer alan bilgilere göre, ayırma kriteri olarak gini indeksi ve K=5 katlı çapraz doğrulama durumunda doğruluk değeri %99,77 iken, ayırma kriteri olarak kazanç oranı (gain ratio) seçildiğinde doğruluk değeri %99,79 olarak hesaplanmaktadır. K=5 ve K=10 katlı çapraz doğrulama durumları göz önüne alındığında ayırma kriteri olarak kazanç oranı (gain ratio) tercih edildiğinde doğruluk değerinin gini indeksi’ne göre daha yüksek olduğu görülmektedir. Doğruluk ve hata oranı değerleri göz önüne alındığında karar ağacı modelinin çalışma kapsamında kullanılan veri seti için uygun bir model olduğu söylenebilmektedir.

Destek vektör makineleri modeli uygulamasında, farklı çekirdek türleri ve K-katlı çapraz doğrulama durumları altında elde edilen doğru sınıflandırma ve hatalı sınıflandırma yüzdeleri, uygulama bölümünde yer alan Tablo 21 ve Tablo 22’de yer almaktadır.

Polinom (Polynomial) kernel türü için Tablo 21’de yer alan bilgilere göre, K=5 katlı çapraz doğrulama için; sapma değeri adına sırasıyla 0,5, 0,9 ve 1 değerleri, gama değeri adına ise sırasıyla 1, 0,5, 1,5 değerleri denemiştir. Bu denemeler sonucunda elde edilen doğru sınıflandırma yüzdesi açısından herhangi bir farklılık gözlemlenmemiştir. Her bir denemede elde edilen doğru sınıflandırma yüzdesi K=5 katlı çapraz doğrulama için %99,77 olarak elde edilmiştir. K=10 ve K=20 değerleri için yapılan denemeler sonucu elde edilen doğru sınıflandırma yüzdesinin, K=5 değeri için elde edilen sonuçlar ile aynı olduğu görülmektedir.

K=30 katlı çapraz doğrulamaya ait sonuçlarda, sapma değeri=1 ve gama değeri=1 iken doğru sınıflandırma yüzdesi %99,17 olarak elde edilmiştir. Eğitim kümesinde yer alan veri sayısı arttıkça doğru sınıflandırma yüzdesinde herhangi bir artış gözlemlenmemiştir, sadece K=30, sapma değeri=1 ve gama değeri=1 iken alınan sonuçta doğru sınıflandırma yüzdesi %99,17’e gerilemiştir.

Radyal tabanlı fonksiyon (RBF) kernel türü için Tablo 22’de yer alan bilgilere göre, K=5, 10 ve 30 değerleri için yapılan denemelerde elde edilen doğru sınıflandırma yüzdeleri %99’un üzerinde iken K=20 değeri için yapılan denemelerde, sigma=0,1 ve 0,3 değerleri için elde edilen doğru sınıflandırma yüzdeslerinin sırasıyla %92,26 ve %98,29’a gerilediği görülmektedir. Elde edilen doğru sınıflandırma yüzdeleri göz önüne alındığında destek vektör makineleri modelinin çalışma kapsamında kullanılan veri seti için uygun bir model olduğu görülmektedir.

K=5, 10, 20 ve 30 katlı çapraz doğrulama ile uygulamaları yapılan iki modele (Karar ağaçları ve destek vektör makineleri) ait sonuçlar incelendiğinde, doğruluk ve doğru sınıflandırma yüzdeleri tüm denemelerde %92 ve üzerinde elde edilmiştir.

Çalışmada kullanılan veri seti gibi bir seti için veri madenciliği modellerinden, karar ağaçları ve destek vektör makineleri modellerinin kullanışlı olabileceği görülmektedir.

Toplum Merkezleri’nde kurs hizmeti alan yararlanıcıların, kurs sonrasında eğitimini aldıkları kurs doğrultusunda bir iş kolunda istihdam edilmeleri Türk Kızılay açısından oldukça

sevindirici bir durumdur. Bu durum Toplum Merkezleri'nde verilen eğitim hizmetlerinin yararlanıcıların ülkemizde iş sahibi olmalarını ve dolaylı olarak yerel halk ile kaynaşmalarını kolaylaştırmaktadır. Bu çalışma Toplum Merkezi çalışanları tarafından incelenme fırsatı bulunduğu takdirde, sonraki süreçlerde yararlanıcılara verilmekte olan hizmetlerin daha da iyileştirileceğine ve yerel halk ile Suriyeli göçmenlerin daha fazla kaynaştırılacağına inanılmaktadır.

Ek olarak yapılan çalışmaya dair belirtilmesi gereken birkaç husus bulunmaktadır. Bu çalışma, kullanılan veri seti sebebiyle dengesiz sınıf problemi (imbalanced class problem) içermektedir. İkili sınıflandırma durumlarında kullanılan verilerde tahmin edilecek olan sınıfa ait örnek sayısı diğer sınıfa ait örnek sayısından daha fazla ise bu veri dengesiz veri olarak isimlendirilmektedir [47]. Literatürde bu problemin çözülmesine ilişkin pek çok yöntem yer almaktadır. Bu yöntemlere örnek olarak; doğru performans ölçütünün seçilmesi ya da yeniden örnekleme yöntemleri ile veriyi dengeli bir hale getirme işlemi gösterilebilir [48].

Çalışmada seçilen hedef değişkenin dengesiz veri olması sebebiyle dengesiz sınıf problemi ile karşılaşmıştır, bu sebeple elde edilen sonuçlar sınırlıdır. Dengesiz sınıf problemi, elde edilen doğruluk oranlarının da oldukça yüksek elde edilmesine sebep olmuştur. Bu handikaplara rağmen yapılan çalışmanın farklı bir uygulama alanında veri madenciliği tekniklerinin gösterilmesi açısından önemli olabileceğine inanılmaktadır.

KAYNAKLAR

- [1] KİRİSCİ. K, The Question of Asylum and Illegal Migration in European Union-Turkish Relations, **2003**.
- [2] Anonim, TÜİK Haber Bülteni, Sayı: 30711, **2019**.
- [3] Anonim, Türkiye'deki Suriyeli Sayısı Ekim 2020, <https://multeciler.org.tr/turkiyedeki-suriyeli-sayisi/> (Erişim tarihi: **28 Ekim 2020**).
- [4] S. Saygın, D. Hasta, Göç, Kültürleşme ve Uyum, Psikiyatride Güncel Yaklaşımlar, **2018**.
- [5] B. Akıncı, A. Nergiz, E. Gedik, Uyum Süreci Üzerine Bir Değerlendirme: Göç ve Toplumsal Kabul, Göç Araştırmaları Dergisi, **2015**.
- [6] Bhatt. C., Mining the Medical Literature 2004,. <https://slideplayer.com/slide/5334684/>, (Erişim Tarihi: **08.01.2018**).
- [7] N. Şahin, Yapay Sinir Ağları ve Karar Ağaçları Modelleri İle İşletmelerin Finansal Başarısızlıklarının Tahminlenmesi, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, **2019**.
- [8] S. Savaş, N. Topaloğlu, M. Yılmaz, Veri Madenciliği ve Türkiye'deki Uygulama Örnekleri, İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, **2012**.
- [9] J. Han, M. Kamber, Data Mining: Concepts and Techniques, **2000**.
- [10] G. Silaharoğlu, Kavram ve Algoritmalarıyla Temel Yeri Madenciliği, İstanbul, Papatya Yayıncılık, **2008**.
- [11] S. Elasan, Veri Madenciliğinde Farklı Karar Ağaçları ve K-En Yakın Komşuluk Yöntemlerinin İncelenmesi: Kadın Hastalıkları ve Doğum Verisinde Bir Uygulama, Doktora Tezi, Van Yüzüncü Yıl Üniversitesi Sağlık Bilimleri Enstitüsü, Van, **2019**.
- [12] F. Köktürk, K En Yakın Komşuluk, Yapay Sinir Ağları ve Karar Ağaçları Yöntemlerinin Sınıflandırma Başarılarının Karşılaştırılması, Doktora Tezi, Bülent Ecevit Üniversitesi Sağlık Bilimleri Enstitüsü, Zonguldak, **2012**.
- [13] A. Oğuzlar, CART analizi ile hane halkı işgücü anketi sonuçlarının özetlenmesi, Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, **2004**.
- [14] Breiman, L., Friedman, J. ve Olshen, R., Classification and Regression Trees, Wadsworth Int, **1984**.
- [15] Quinlan JR., Simplifying decision trees, Int J Man-Mach Stud., **1987**;27(3):221-34.
- [16] Quinlan JR. Decision trees and decision-making. IEEE Trans Syst Man Cybern. 1990;20(2):339-46.

- [17] M. Şahin, Karar Ağaçları ve Yapay Sinir Ağları Kullanarak Kasko Sigortalarında Risk Değerlendirme, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 2018.
- [18] Kass GV, An Exploratory Technique For Investigating Large Quantities of Categorical Data, Appl Statist, 1980.
- [19] Littler, S., CHAID (Chi-square Automatic Interaction Detector), Select Statistical Services: www.select-statistics.co.uk/blog/chaid-chi-square-automatic-interaction-detect, (Erişim Tarihi: 23 Nisan 2019).
- [20] S. Diler, Veri Madenciliği Süreçleri ve Karar Ağaçları Algoritmaları ile Bir Uygulama, Yüksek Lisans Tezi, Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü, Van, 2016.
- [21] Rokach, L., Maimon, O. Decision Trees: Data Mining and Knowledge Discovery Handbook, Boston, 2005.
- [22] Farid D, Sadeghi. H, Hajigol. E, ve Parirooy. N, Classification of Bank Customers by Data Mining: A Case Study of Mellat Bank Branches in Shiraz, International Journal of Management Accounting and Economics, 2016.
- [23] Witten. I.H, Frank, E. ve Hall, A. M, Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, 2011.
- [24] Babadağ. K. K, Veri Madenciliği Yaklaşımı ve Veri Kalitesinin Artması İçin Kullanılması, Uzmanlık Tezi, Türkiye İstatistik Kurumu, Ankara. 2003.
- [25] Özbudak. Ö, Yüz Resimlerinden Cinsiyet Tayini, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen bilimleri Enstitüsü, İstanbul. 2009.
- [26] Vercellis. C, Business intelligence: Data Mining and Optimization for Decision Making, Wiley Publication, 2009.
- [27] Kavurkacı. Ş, Aydın. G. Z, Şamlı. R, Büyük Ölçekli Veri Tabanlarında Bilgi Keşfi, XIII.Akademik Bilişim Konferans Yayımları, Malatya, 2011.
- [28] Şeker, E., S., WEKA ile Veri Madenciliği, Cinius Yayınları, 2013.

- [29] Dalkılıç, H. ve Dalkılıç, F., Karar Ağaçları Destekli Vadeli Mevduat Hesabı, <http://ab.org.tr/ab15/kabul.html> 17, Akademik Bilişim Konferansı (Erişim Tarihi: 21 Aralık 2015).
- [30] Dondurmacı, G., Veri Madenciliğinde Regresyon ağaçları ile Sınıflandırma Ve Bir Uygulama, Doktora Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul. 2011.
- [31] S. Ramazanov, Türkiyede Fertlerin Elektronik Ticaret Kullanım Sürelerinin Destek Vektör Makineleri ve Karar Ağaçları Yöntemleri ile Karşılaştırmalı Analizi, Yüksek Lisans Tezi, Kafkas Üniversitesi Sosyal Bilimler Enstitüsü, Kars, 2020.
- [32] N. V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, 1995.
- [33] N. V. Vapnik. Statistical Learning Theory, John Wiley & Sons, 1998.
- [34] Nitze. I, Schulthess. U, and Asche. H, Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification, Proceedings of the 4th GEOBIA, 2012.
- [35] Soman. K.P, Loganathan. R, and Ajay. V, Machine Learning with SVM and Other Kernel Methods, PHI Learning Pvt. Ltd., 2011.
- [36] T. Kavzoğlu, İ. Çölkesen, Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi, Harita Dergisi, 2010.
- [37] Burges, Christopher JC., A Tutorial on Support Vector Machines For Pattern Recognition, Data Mining and Knowledge Siscovery 2.2, 1998.
- [38] E. Çomak, Destek Vektör Makinelerinin Etkin Eğitimi İçin Yeni Yaklaşımlar, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya, 2008.
- [39] Cortes. C, and Vapnik. V, Support Vector Networks, Machine Learning, 1995.

- [40] Busuttil, S, Support Vector Machines In Proceedings of The Computer Science Annual Research Workshop, Villa Bighi, Kalkara, University of Malta, 2003.
- [41] S. Ayhan, Ş. Erdoğan, Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi, Eskişehir Osmangazi Üniversitesi İibf Dergisi, 2014.
- [42] O. Doğan, Kentsel Dokunun Destek Vektör Makineleri Yöntemi İle Sınıflandırılması: İstanbul Örneği, Yüksek Lisans Tezi, Sivas Cumhuriyet Üniversitesi Fen Bilimleri Üniversitesi, Sivas, 2019.
- [43] E. Güldoğan, Çeşitli Çekirdek Fonksiyonları İle Oluşturulan Destek Vektör Makinesi Modellerinin Performanslarının İncelenmesi: Bir Klinik Uygulama, Ortak Doktora Tezi, İnönü Üniversitesi ve Mersin Üniversitesi Sağlık Bilimleri Enstitüsü, Malatya, 2017.
- [44] R. Varçın, İ. Savcı, B. Gülçubuk, M. Pişkin, Türkiye’de Geçici Koruma Altındaki Suriyelilere Yönelik İşgücü Piyasası Araştırması, Türk Kızılay Toplum Merkezi, Ankara, 2018.
- [45] Anonim, UNCHR Turkey: Provincial Breakdown Syrian Refugees in Turkey, https://www.unhcr.org/tr/wp-content/uploads/sites/14/2019/08/9.2-UNHCR_Provincial-Breakdown-Syrian-Refugees-in-Turkey-August-2019.png, (Erişim Tarihi: 15 Nisan 2020).
- [46] Anonim, Türk Kızılay Toplum Merkezleri Ekim Ayı Raporu, Türk Kızılay Toplum Temelli Göç Programları Koordinatörlüğü, Ankara, 2020.
- [47] Kartal, E., Özen, Z., “Dengesiz Veri Setlerinde Sınıflandırma”, Mühendislikte Yapay Zekâ Uygulamaları, Sakarya, 2017.
- [48] M.A.Aydın, Müşteri Kaybı Tahmininde Sınıf Dengesizliği Problemi, Politeknik Dergisi.

ÖZGEÇMİŞ

1995 yılında Ankara'da doğdum. Lise öğrenimimi 50.Yıl Lisesi'nde gördüm. 2013 yılında Hacettepe Üniversitesi İstatistik bölümüne girdim, 2017 yılında mezun oldum. 2017 yılında Hacettepe Üniversitesi İstatistik bölümünde yüksek lisans öğrenimime başladım. 2018 yılının Ekim ayından itibaren Türk Kızılay Toplum Temelli Göç Programları Koordinatörlüğünde finans asistanı olarak görev yapmaktayım.

