



**TÜRKİYE SAĞLIK ARAŞTIRMASI VERİLERİNİN YAPAY SİNİR
AĞLARI VE ÇOK KATEGORİLİ LOJİSTİK REGRESYON YÖNTEMİYLE
İNCELENMESİ**

Özge YAMANTAŞ

**YÜKSEK LİSANS TEZİ
İSTATİSTİK ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

TEMMUZ 2019

Özge YAMANTAŞ tarafından hazırlanan “TÜRKİYE SAĞLIK ARAŞTIRMASI VERİLERİNİN YAPAY SİNİR AĞLARI VE ÇOK KATEGORİLİ LOJİSTİK REGRESYON YÖNTEMİYLE İNCELENMESİ” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi İstatistik Ana Bilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Doç. Dr. Meral EBEGİL

İstatistik Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.

.....

Başkan: Prof. Dr. Mehtap AKÇİL OK

İstatistik ve Bilgisayar Bilimleri Ana Bilim Dalı, Başkent Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.

.....

Üye: Prof. Dr. Fikri GÖKPINAR

İstatistik Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum.

.....

Tez Savunma Tarihi: 23/07/2019

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

.....

Prof. Dr. Sena YAŞYERLİ
Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Özge YAMANTAŞ

23/07/2019

TÜRKİYE SAĞLIK ARAŞTIRMASI VERİLERİNİN YAPAY SİNİR AĞLARI VE ÇOK KATEGORİLİ LOJİSTİK REGRESYON YÖNTEMİYLE İNCELENMESİ

(Yüksek Lisans Tezi)

Özge YAMANTAŞ

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Temmuz 2019

ÖZET

Sağlıklı olmak, kaliteli yaşam için gerekli olan temel faktördür. Kaliteli yaşam için ise birey, sağlığına zarar verecek tutum ve davranışlardan kaçınmalıdır. Bireylerin sağlık durumlarını etkileyen birçok etken vardır. Bu çalışmada bireylerin genel sağlık durumları ve genel sağlık durumlarına etki eden faktörler belirlenmeye çalışılmıştır. Bu amaçla TÜİK tarafından yürütülmekte olan Türkiye Sağlık Araştırması (TSA) kapsamında yer alan anket verileri kullanılmıştır. Çalışmada, yapay sinir ağları ve çok kategorili lojistik regresyon analizi ayrı ayrı uygulanmış olup sınıflandırma başarıları karşılaştırılmıştır. Yapılan analizler sonucunda, bireylerin genel sağlık durumlarının doğru sınıflandırma oranları, çok kategorili lojistik regresyon analizi için %76,29 ve yapay sinir ağları tekniği için de %75,37 olarak hesaplanmıştır. Çalışmaya alınan veri seti için çok kategorili lojistik regresyon analizinin yapay sinir ağları tekniğine göre daha iyi bir “doğru sınıflandırma oranına” sahip olduğu görülmüştür. Bu çalışma ile yapılan iki farklı analiz yöntemi ile elde edilen sınıflandırma başarılarının yorumlanması sonucunda literatüre katkı sağlamak hedeflenmektedir.

Bilim Kodu : 20513
Anahtar Kelimeler : Yapay Sinir Ağları, Çok Kategorili Lojistik Regresyon Analizi, Türkiye Sağlık Araştırması, Genel Sağlık Durumu
Sayfa Adedi : 86
Danışman : Doç. Dr. Meral EBEGİL

INVESTIGATION OF TURKEY HEALTH SURVEY DATA WITH ARTIFICIAL NEURAL NETWORKS AND MULTINOMIAL LOGISTIC REGRESSION

(M. Sc. Thesis)

Özge YAMANTAŞ

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

July 2019

ABSTRACT

Being healthy is the essential factor for quality life. For the quality of life, the individual should avoid attitudes and behaviors that will harm his / her health. There are many factors that affect the health status of individuals. In this study, the factors affecting the general health status of individuals and their general health status were tried to be determined. For this purpose, Turkey Health Survey conducted by Turkish Statistical Institute (TurkStat) within the scope of the data is used. In this study, artificial neural networks and multinomial logistic regression analysis were applied separately and classification successes were compared. As a result of the analyzes, the correct classification rates of the general health status of the individuals were calculated as 76,29% for the multinomial logistic regression analysis and 75,37% for the artificial neural network. Multinomial logistic regression has a better correct classification rate than artificial neural network for the data set. The aim of this study is to contribute to the literature as a result of the interpretation of the classification successes obtained by two different analysis methods.

Science Code : 20513
Key Words : Artificial Neural Network, Multinomial Logistic Regression, Turkey
Health Survey, General Health Status
Page Number : 86
Supervisor : Assoc. Prof. Dr. Meral EBEGİL

TEŞEKKÜR

Tez çalışmam süresince büyük emeği olan, üstün bilgi ve tecrübesiyle beni her konuda yönlendiren danışmanım Sayın Doç. Dr. Meral EBEGİL'e, tez yazımının uygulama aşamasında bana yardımcı olan, verdiği tavsiyeler ve yardımlarından dolayı değerli iş arkadaşlarım Levent AHİ ve Fatih YALABUK'a, çalışmalarım boyunca beni destekleyen değerli mesai arkadaşlarıma, yardımlarını esirgemeyen bölüm hocalarıma ve arkadaşlarıma, bana olan güvenleri ve inançları ile daima yanımda olan, emeklerini hiçbir zaman ödeyemeyeceğim sevgili annem Hacıkadın ERKİLİÇ'a, sevgili babam İsmet ERKİLİÇ'a ve kardeşlerime, birlikte çıktığımız hayat yolunda desteğini en derinden hissettiğim, beni sevgisiyle her daim yüreklendiren, her an sabırla yanımda olan kıymetli eşim, yol arkadaşım Ahmet Yaşar YAMANTAŞ'a teşekkürlerimi sunmayı borç bilirim.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	xi
SİMGELER VE KISALTMALAR.....	xii
1. GİRİŞ.....	1
2. YAPAY SİNİR AĞLARI.....	9
2.1. Yapay Sinir Ağlarının Tanımı ve Özellikleri.....	9
2.2. Biyolojik Sinir Hücresi ve Yapay Sinir Hücresi	11
2.3. Yapay Sinir Ağları Bileşenleri	12
2.4. Yapay Sinir Hücresinin Çalışma Prensibi.....	15
2.5. Yapay Sinir Ağlarının Yapısı.....	16
2.6. Yapay Sinir Ağlarında Öğrenme.....	17
2.6.1. Öğrenme yöntemleri	18
2.7. Yapay Sinir Ağlarında Öğrenme Kuralları	19
2.7.1. Hebb kuralı.....	20
2.7.2. Hopfield kuralı	20
2.7.3. Delta kuralı.....	20
2.7.4. Kohonen kuralı.....	21
2.8. Yapay Sinir Ağı Modelleri.....	21
2.8.1. Tek katmanlı algılayıcı modeli	21
2.8.2. Çok katmanlı algılayıcı modeli	23

	Sayfa
2.8.3. LVQ ağırları modeli	23
2.8.4. ART ağırları modeli	25
2.9. Yapay Sinir Ağlarının Avantajları ve Dezavantajları	26
3. LOJİSTİK REGRESYON ANALİZİ.....	29
3.1. Lojistik Regresyon Analizinin Tanımı ve Özellikleri	29
3.2. Çok kategorili (Multinomial) Lojistik Regresyon Analizi.....	31
4. UYGULAMA.....	37
4.1. Kullanılan Veri Setine İlişkin Genel Bilgiler	37
4.2. Veri Setinin Düzenlenmesi	39
4.3. Veri indirgeme	44
4.4. Tanımlayıcı İstatistikler	47
4.5. Yapay Sinir Ağları Yönteminin Uygulanması	51
4.6. Çok kategorili Lojistik Regresyon Yönteminin Uygulanması	54
4.7. Yöntemlerin Karşılaştırılması	57
5. SONUÇ	61
KAYNAKLAR	63
EKLER.....	69
EK-1. SPSS paket programında elde edilen ve VIF değerleri neticesinde veri setinden çıkarılan/çıkarılmayan değişkenlere ilişkin tablo	70
EK-2. Değişkenler frekans yüzde dağılımı (%).....	74
ÖZGEÇMİŞ	86

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. Sinir sistemi ile yapay sinir ağları arasındaki ilişki	12
Çizelge 2.2. Toplama fonksiyonlarının toplama işlevleri	13
Çizelge 2.3. Aktivasyon fonksiyonu örnekleri	14
Çizelge 4.1. Genel sağlık durumu değişkeni	39
Çizelge 4.2. Yaş grubu değişkeni	41
Çizelge 4.3. Eğitim durumu değişkeni.....	42
Çizelge 4.4. 2014 ve 2016 yıllarına ait hanelerin gelir bilgisi seçenekleri	42
Çizelge 4.5. Uluslararası standart meslek sınıflaması	43
Çizelge 4.6. Avrupa topluluğunda ekonomik faaliyetlerin istatistiki sınıflaması.....	43
Çizelge 4.7. Kullanılan bağımsız değişkenler.....	45
Çizelge 4.8. Yaş değişkeni kategorilerine göre genel sağlık durumu (%).....	47
Çizelge 4.9. Cinsiyet değişkeni kategorilerine göre genel sağlık durumu (%).....	48
Çizelge 4.10. Eğitim durumu değişkeni kategorilerine göre genel sağlık durumu (%)	48
Çizelge 4.11. Medeni durum değişkeni kategorilerine göre genel sağlık durumu (%)	49
Çizelge 4.12. Hanenin gelir durumu değişkeni kategorilerine göre genel sağlık durumu (%).....	49
Çizelge 4.13. Uluslararası standart meslek sınıflaması kodlarına göre genel sağlık durumu (%)	49
Çizelge 4.14. Hastalık sağlık durumu değişkeni kategorilerine göre genel sağlık durumu (%).....	50
Çizelge 4.15. Yaşamsal faaliyetin kısıtlanma durumu değişkeni kategorilerine göre genel sağlık durumu (%).....	50
Çizelge 4.16. Günlük faaliyet durum değişkeninin kategorilerine göre genel sağlık durumu (%).....	51
Çizelge 4.17. Yapay sinir ağı eğitim veri seti sonucu.....	53
Çizelge 4.18. Yapay sinir ağı test veri seti sonucu	53

Çizelge	Sayfa
Çizelge 4.19. Yapay sinir ağı doğrulama veri seti sonucu.....	54
Çizelge 4.20. Çok kategorili lojistik regresyon model uyum ölçütü	55
Çizelge 4.21. Çok kategorili lojistik regresyon Pseudo R^2 (sözde R^2) değerleri	56
Çizelge 4.22. Çok kategorili lojistik regresyon eğitim veri seti sonucu	56
Çizelge 4.23. Çok kategorili lojistik regresyon test veri seti sonucu.....	56
Çizelge 4.24. Çok kategorili lojistik regresyon doğrulama veri seti sonucu	57
Çizelge 4.25. Yapay sinir ağları modelinin alt veri setleri içindeki doğruluk oranlarının karşılaştırılması	58
Çizelge 4.26. Çok kategorili lojistik regresyon modelinin alt veri setleri içindeki doğruluk oranlarının karşılaştırılması.....	59

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Basit bir yapay sinir ağı modelinin yapısı	10
Şekil 2.2. Biyolojik sinir sisteminin blok gösterimi	12
Şekil 2.3. Sigmoid fonksiyonu.....	14
Şekil 2.4. Bir yapay sinir hücresinin çalışması örneği.....	15
Şekil 2.5. Yapay sinir ağı katmanlarının birbirleriyle ilişkisi.....	16
Şekil 2.6. Yapay sinir ağı örneği.....	17
Şekil 2.7. Tek katmanlı algılayıcı modeli genel yapısı	22
Şekil 2.8. Çok katmanlı algılayıcı modeli genel yapısı	23
Şekil 2.9. LVQ modeli genel yapısı.....	24
Şekil 2.10. ART modeli genel yapısı	25
Şekil 4.1. Yapay sinir ağları çıktı görünümü	53
Şekil 4.2. Yapay sinir ağları ve çok kategorili lojistik regresyon karşılaştırılması	58

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar	Açıklamalar
ART	Adaptif Rezonans Teori
GRNN	Genel Regresyon Ağları
GSS	Genel Sağlık Sigortası
ILO	Uluslararası Çalışma Örgütü
ISCO-08	Uluslararası Standart Meslek Sınıflaması
KDH	Kısa Dönemli Hafıza
LVQ	Doğrusal Vektör Parçalama Yöntemi
MLP	Çok Katmanlı Ağlar
NACE-Rev.2	Avrupa Topluluğunda Ekonomik Faaliyetlerin Sınıflaması
PNN	Olasılıklı Ağlar
RBF	Radyal Tabanlı Fonksiyon
SGK	Sosyal Güvenlik Kurumu
TSA	Türkiye Sağlık Araştırması
TÜİK	Türkiye İstatistik Kurumu
UDH	Uzun Dönemli Hafıza
VIF	Varyans Büyütme Faktörü

1. GİRİŞ

Sınıflandırma uygulamalarında pek çok model kullanılmaktadır. Her bir model için farklı algoritmalar yer almaktadır. Bu algoritmaların hangisinin daha iyi sonuçlar ürettiği ve hangi alanda daha başarılı olduğu sorularına verilen cevaplar uygulama başarısına etki ederek yapılan işin veriminin artmasını sağlamaktadır. Bu nedenle aynı veri üzerinde farklı algoritmaların uygulanması ve karşılaştırılması büyük önem taşımaktadır. Çok fazla algoritmanın varlığı, her algoritmanın çalışma prensibinin farklı olması, birden çok versiyona sahip olması, her bir algoritmanın kullanım amacının farklı olması, kullanılan veri kaynağının değişmesi, her bir algoritmanın kullanacağı veri tiplerinin farklı olması ve verinin önışlem sürecinden geçirilip geçirilmemesi sebepleri ile sonuçlar farklı şekilde elde edilebilmektedir.

Yapay sinir ağıları insan beyninden esinlenilerek oluşturulmuş olup bir takım bağlantılar sayesinde birbirine bağlanan ve her biri kendi işlem verisinden oluşmuş olan yapılardır. İnsanların biyolojik sinir yapılarının taklit edilmesiyle oluşan programlardır. Yapay sinir ağıları ilave bir programcı yeteneği olmaksızın kendiliğinden öğrenme özelliğine sahiptir. Biyolojik sinir ağını taklit eden bu ağılar, öğrenebilme yeteneğinin yanı sıra ezber yapma ve ilişki kurma özelliklerine de sahiptir [1].

İlk yapay sinir ağı modeli Warren McCulloch ve Walter Pitts tarafından 1943 yılında geliştirilmiştir [2]. İsmi geçen araştırmacılar, ilk yapay sinir ağı hücresinin yapısını oluşturmuş ve yapay sinir ağlarının temelini atmışlardır. McCulloch ve Pitts, elektrik devrelerini kullanarak insan beyninin hesaplama ve analiz yeteneğinden yola çıkarak bu hücreyi geliştirmişlerdir. Böylece yapay sinir hücreleri ile birlikte her çeşit mantıksal ifadenin formülize edilebilirliği gösterilmiştir. Daha sonra birden fazla hücrenin birlikte çalışması gerektiğini düşünüp paralel olarak çalıştırma yöntemi ile öğrenme kurallarını belirlemeye başlamışlardır [3].

1949 yılında Hebb kitabında öğrenebilen ve uyum sağlayabilen sinir ağıları modeli için temel oluşturacak “Hebb kuralını” ileri sürmüş olup öğrenme ile ilgili temel teoriyi incelemiştir. Bu teoriye göre sinir ağının bağlantı sayısı değiştirildiği zaman, öğrenebileceği ön görülmektedir [4].

1957 yılında “Perceptron” Frank Rosentblatt tarafından keşfedildikten sonra yapay sinir ağı alanındaki çalışmalar hız kazanmıştır [5]. Perceptron, tek katmanlı eğitilebilen ve tek bir çıkışa sahip yapay sinir ağı olarak tanımlanır. 1958 yılında Rosenblatt ve 1961 yılında Caianiello tarafından, rassal ağlar ile adaptif tepki üretme kavramı üzerine çalışmalar devam etmiştir [5,6].

1956 yılında yapay zekâ kavramının ortaya çıkmasıyla yapay sinir ağları geri planda kalmıştır. 1960 ile 1980 yılları arasında bilimsel alanda yapılan çalışmaların büyük bir kısmı yapay zekâyâ çevrilmiştir. O dönemde bazı bilim adamlarının yapay sinir ağları üzerine çalışmaları devam etmiş ve ortaya yeni fikirlerin atılması ile yapay sinir ağları tekrar hız kazanmaya başlamıştır [7].

1988 yılında, Broomhead ve Lowe radyal tabanlı fonksiyonlar modelini (Radial Basis Functions (RBF)) geliştirmişler ve özellikle filtreleme alanında başarılı sonuçlar elde etmişlerdir. Ardından Spect, bu ağların daha gelişmiş hali olan olasılıklı (probabilistik) ağlar (PNN) ve Genel Regresyon Ağlarını (GRNN) geliştirmiştir [8].

Veri setinde yer alan değişkenler ve bu değişkenlere ait özellikler kullanılacak olan yöntemin belirlenmesinde büyük paya sahiptir. Bir bağımlı değişken ve bir veya daha çok bağımsız değişken arasındaki ilişkinin incelenmek istenmesi durumunda regresyon analizi yöntemleri uygulanabilmektedir. Regresyon analizinin basit ve çoklu doğrusal regresyon olmak üzere iki alt başlığı vardır. Bu durum bağımlı değişkenin metrik (aralık ya da oran) ölçekli olması durumunda geçerlidir. Ancak bağımlı değişkenin kategorik veri tipinde olması durumu da söz konusudur. Bu tip durumlarda doğrusal regresyonda parametre tahmini yapmak üzere en küçük kareler yöntemi uygulamak bu yöntem ile ilgili varsayımların sağlanamamasından dolayı uygun bulunmamaktadır. Bu sebeple bağımlı değişken iki ya da ikiden fazla kategorili olduğunda lojistik regresyon yöntemi uygulanabilmektedir [9].

Lojistik regresyon yönteminin amacı bir veya daha fazla sayıda bağımsız değişken ile sonuç değişkeni arasında bir model elde etmektir. Diğer regresyon yöntemlerinden farkı ise sonuç değişkeninin sürekli değerler almayıp kesikli iki ya da daha çok değer almasıdır [10].

Bu nedenle lojistik regresyon analizinde sonuç değişkeni yani bağımlı değişken iki ya da ikiden fazla kategorili olduğu durumlar için genelleştirilebilir. Literatürde cevap değişkeni

ikiden fazla kategorili olduđuunda analiz çok kategorili ya da multinomial lojistik regresyon analizi olarak isimlendirilmektedir. Bu analiz ile cevap deęiřkeni üç veya daha fazla kategorili iken cevap deęiřkeni ile açıklayıcı yani bağımsız deęiřkenler arasındaki ilişki belirlenebilmektedir.

Çok kategorili lojistik regresyon analizinde, cevap deęiřkeninin herhangi bir kategorisi referans kategori olarak alınmakta ve dięer kategoriler bu referans kategoriye göre analiz edilmektedir [11].

Lojistik regresyonda bağımlı deęiřken kategorik iken bağımsız deęiřkenler karışık ölçekli olabilir. Herhangi bir dağılım varsayımı olmaksızın bağımlı deęiřken ile bağımsız deęiřkenler arasında neden-sonuç ilişkisinin belirlenmesi amacıyla kullanılmaktadır [12].

Lojistik modelin kullanılmaya başlanması 1845’li yıllara dayanmaktadır. Çoęunlukla sosyoekonomik konular için kullanılmakta olan lojistik analiz, ilk başlarda nüfus artışının matematiksel ifadelerle açıklanmasını amaçlayan çalışmalar için kullanılmıştır [13].

1944 yılında lojistik modeli biyolojik deneyleri analiz etmek için kullanmayı öneren ilk kiři Berkson olmuřtur [14]. Cox, 1970 yılında biyolojik deneylerin analiz edilmesi konusunda çeřitli modeller kurarak uygulamalar yapmıřtır. Yaygın bir biçimde kullanılır hale gelen lojistik regresyon modelleri, katsayı tahmin metodlarının gelişmesini ve lojistik regresyon modellerinin kapsamlı bir şekilde incelenmesini beraberinde getirmiřtir [15]. 1962 yılında Cornfield, lojistik regresyondaki katsayı tahmin işlemlerinde diskriminant fonksiyonu yaklaşımını ilk kez kullanarak popüler hale getirmiřtir [16]. 1984 yılında Lee, basit dönüşümlü (cross-over) deneme planları için doğrusal lojistik modeller üzerinde çalışmalar yapmıřtır [17]. 1987 yılında Bonney lojistik regresyon modelini kullanmaya ve geliřtirmeye yönelik çalışmalar yapmıřtır [18].

Lojistik regresyon analizi son yirmi yıldır askeri alanlarda, meteoroloji konularında, iç göç hareketlerinde, eğitim ve saęlık alanlarında artarak uygulanmaktadır. İstatistik paket programlarının yaygınlařması bu artışı açıklayan en önemli etkenlerden biri olmaktadır. Farklı alanlarda uygulanmakta olsa da yine de en yaygın kullanım alanının tıp olduęu görölmektedir [19].

Yapay sinir ağıları sınıflama ve bilgilerin yorumlanması amacıyla farklı problemlerin çözümünde kullanılmasının yanı sıra sağlık, finans, endüstri, iş hayatı ve eğitim alanlarında da uygulanmaktadır [20].

Lojistik regresyon modelleri de, son yıllarda tıp, ekonomi, biyoloji, veterinerlik ve ziraat gibi birçok farklı alanda yaygın olarak kullanılmaktadır [21].

1990'ların başlarında yayınlanmış olan kaynaklara göre yapay sinir ağıları ve bazı istatistiksel yöntemlerin benzer olduğu hatta bazılarının aynı olduğu tespit edilmiştir. Daha sonra ortaya atılan çalışmalar bu benzerliğin tesadüfi olmadığını ve birbirleriyle ileri düzeyde ilişki olduklarını göstermiştir. Yapay sinir ağı yöntemi ile istatistiksel yöntemlerin karşılaştırılması, birinin diğerinin geliştirilmesi sürecine katkı sağladığını göstermiştir. Bilim adamları tarafından bazı yapay sinir ağı modellerinin istatistiksel uygulamalar için yararlı olabileceği ve benzer şekilde tahmin ölçütü, diagnostik yöntemler, güven aralıkları gibi istatistiksel yöntemlerin de yapay sinir ağı uygulamalarına uygulanabilirliği ileri sürülmüştür. Yapay sinir ağıları ve istatistiksel yöntemler arasındaki ilişkinin ilerlemesi her iki alana büyük fayda sağlamaktadır [22].

Yapay sinir ağıları ve lojistik regresyon yöntemlerinin sınıflandırma oranlarının karşılaştırılması çalışmaları uzun yıllardan beri devam etmektedir.

Literatür incelendiğinde ikili lojistik regresyon ile ilgili birçok çalışma yer almasına rağmen, çok kategorili lojistik regresyon analizi ile ilgili ayrıntılı bilgi veren çalışmaların çok fazla olmadığı söylenebilir. Bu konuda Türkçe literatürün geliştirilmesi gerekmektedir.

Daha evvel yapılmış olan yapay sinir ağı ve lojistik regresyon analizi / çok kategorili lojistik regresyon analizi yöntemlerinin karşılaştırıldığı bazı çalışmalar incelenmiş olup aşağıda özetlenmiştir.

2009 yılında Karakuş tarafından yapılan çalışmada davranışsal Kredi Skorum sistemi ile şüpheli alacaklara neden olacak abonelerin önceden tahmin edilmesi ve bu sayede zamanında müdahale edilerek zararın minimize edilmesi amaçlanmıştır. Uygulama aşamasında farklı algoritmalar denenmiştir. Uygulama kısmında denenilen algoritmalar arasında yapay sinir ağıları ve lojistik regresyon modelleri de yer almaktadır. Birbirine yakın

ve tutarlı sonuçlar elde edilmiş olmasına rağmen lojistik regresyon modelinin uygulanan yöntemler arasında en başarılı sonuçlar verdiği görülmüştür [23].

2010 yılında Kayri ve Çokluk tarafından yapılan çalışmada çok kategorili lojistik regresyon ve yapay sinir ağlarının sınıflandırma oranları karşılaştırılmıştır. Bu kapsamda okul türü (kamu-özel), cinsiyet, sınıf, anne mesleği, baba mesleği, annenin eğitim durumu, babanın eğitim durumu, erkek kardeş / kız kardeş sayısı, aylık gelir gibi değişkenler bağımsız değişken olarak verilmiştir. İnternet kullanım süresi ise bağımlı değişken olarak ele alınmıştır. Yapılan çalışmada yapay sinir ağları bağımlı değişkeni yüksek doğruluk seviyesinde sınıflandırmış ve modelin çok kategorili lojistik regresyondan daha yüksek bir seviyede uygun olduğunu göstermiştir. Ayrıca, yapay sinir ağları, çok kategorili lojistik regresyondan farklı olarak parametre katsayıları elde etmiştir. Modelin yapay sinir ağları ile daha tutarlı ve doğru bir şekilde tahmin edildiği görülmüştür [24].

2011 yılında Pourshahriar tarafından yapılan çalışmada lojistik regresyon analizi ve yapay sinir ağları yönteminin psikolojik araştırmalardaki yeteneklerini incelemek için Tahran'da eğitim alan rasgele seçilmiş 456 öğrenci ile bir araştırma yapılmıştır. Bu araştırma sonucunda yapay sinir ağları yönteminin lojistik regresyondan daha başarılı olduğu sonucuna ulaşılmıştır. Tahmin edici değişken sayısını azalttıktan sonra ise yapay sinir ağlarının üstünlüğünün lojistik regresyon lehine değiştiğini gözlemlemiştir [25].

2012 yılında Budak ve Erpolat tarafından yapılan çalışmada kredi riskinin öngörülmesine yönelik kullanılan istatistiksel yöntemlerden biri olan lojistik regresyon analizi ve yapay sinir ağları kullanılarak, yöntemlerin kredi riskini tahmin etme güçleri karşılaştırılmıştır. Analiz sonuçlarına göre yöntemlerin sınıflandırma başarıları incelenmiş olup yapaysinir ağları modelinin doğru sınıflandırma oranının %70,3 ve lojistik regresyon modelinin doğru sınıflandırma oranının %65,1 olduğu sonucuna varılmıştır. Yapay sinir ağları yönteminin kredi riskinin öngörülmesinde lojistik regresyon analizine göre daha üstün olduğu tespit edilmiştir [26].

2015 yılında Asogwa ve Oladugba tarafından yapılan çalışmada ise her iki model deneysel olarak karşılaştırılmış ve sonuçlar özetlenmiştir. Akademik performanslarına göre öğrencilerin sınıflandırılması amacıyla yapay sinir ağları ve lojistik regresyon yöntemi uygulanmış olup her model için tahmin doğruluğu ve ortalama doğru sınıflandırma oranı

bulunmuştur. Elde edilen oranlara göre yapay sinir ağıları lojistik regresyondan daha başarılı bulunmuştur [27].

2015 yılında Arundina ve arkadaşları tarafından yapılan çalışmada çok kategorili lojistik regresyon ve yapay sinir ağıları gibi hem istatistiksel hem de istatistiki olmayan teknikler kullanılarak, sukuk derecelendirme tahminindeki sonuçlar tartışılmıştır. Doğru sınıflandırma oranları incelendiğinde, sukuk derecelendirme tahmini için kullanılan örnekte istatistiksel yöntem olmayan yapay sinir ağıları metodunun istatistiksel yöntem olan çok kategorili lojistik regresyondan daha güçlü olduğu sonucuna varılmıştır. Yapay sinir ağı eğitim modelini %96,18 doğruluk oranı ile sınıflandırırken çok kategorili lojistik regresyon modeli %91,72 doğruluk oranı ile sınıflandırmıştır [28].

2019 yılında Hassanipour ve arkadaşları tarafından yapılan çalışmada rasgele etki modellerindeki doğruluk oranları yapay sinir ağıları için %90,5 bulurken lojistik regresyon için %83,2 bulunmuştur. Çalışma sonucunda travmatik hastaların terminal sonuçlarının tahmininde kullanılan hem eğri altında kalan alanın hem de doğruluk oranının yapay sinir ağlarında lojistik regresyondan daha iyi performans gösterdiği görülmüştür. Travma hastalarının etkilerinin tahmin edilmesinde yapay sinir ağlarının kullanılmasının daha doğru klinik kararlar sağlayabileceği sonucuna varılmıştır [29].

Bu çalışmada, yapay sinir ağıları ve çok kategorili lojistik regresyon analizi tekniklerinin sınıflama özellikleri, Türkiye İstatistik Kurumu (TÜİK) tarafından iki yılda bir yapılmakta olan “Türkiye Sağlık Araştırması (TSA)” verileri üzerine uygulanarak karşılaştırılmış olup iki yöntemin sınıflama özellikleri tartışılmıştır.

TSA, hane halkı fertlerinin demografik özellikleri ve sağlık durumları ile ilgili bilgileri elde etmeyi amaçlayan bir araştırmadır. Hane halkı araştırmaları için geliştirilen bu araştırma Türkiye tahmini verecek şekilde tasarlanmıştır. 2008 yılından itibaren anket yılı içerisinde belirli bir uygulama periyodu çerçevesinde yapılmakta olup soru kâğıdı örneğe çıkan tüm hanelere uygulanmaktadır.

Bu kapsamda TÜİK’ ten alınan 2014 ve 2016 yıllarına ait TSA verileri kullanılarak kişilerin genel sağlık durumunun tahmin edilmesi amaçlanmıştır. Bu amaç doğrultusunda, IBM SPSS Modeler paket programından yararlanılmıştır. Kişilerin genel sağlık durumu amaç

değişkenimiz olarak belirlendikten sonra 2014-2016 yılı verileri birleştirilerek iki farklı yöntem uygulanmıştır. Modellerin kişilerin genel sağlık durumunu doğru sınıflandırma oranlarının karşılaştırılması amaçlanmaktadır.

2. YAPAY SINİR AĞLARI

Bu bölümde yapay sinir ağlarına ilişkin genel bilgiler verilecek olup ardından yapay sinir ağları modelleri özetlenecektir.

2.1. Yapay Sinir Ağlarının Tanımı ve Özellikleri

Yapay sinir ağları, insan beyninin çalışma biçiminden esinlenilerek geliştirilmiştir. Bilgisayar ortamında insan beyninin yaptığı işlemleri yapabilen, karar verip sonuç çıkarabilen, öğrenen ve hatırlayan bir algoritmadır. Yapay sinir ağları veri girişini sürekli olarak kabul eder, yetersiz veri olması durumunda var olan bilgi sayesinde sonuca ulaşır.

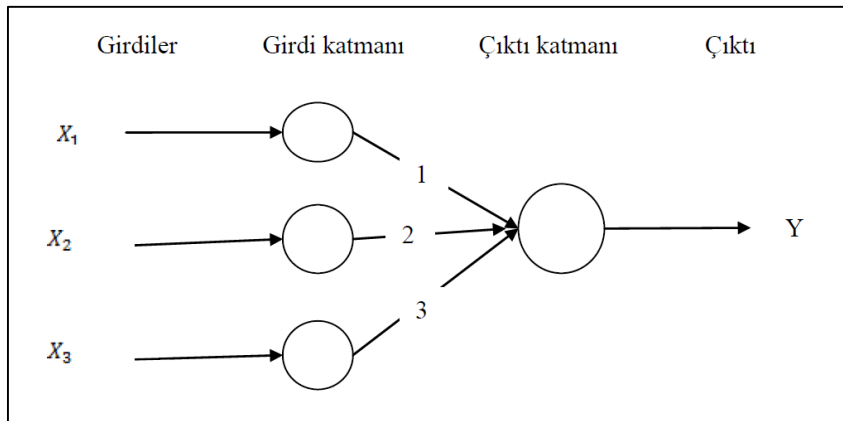
İnsan beyninin üstün özellikleri, daima bilim adamlarının merak konusu olmuştur. Beynin yapısından ilham alarak matematiksel modelleri geliştirmeye çalışmışlardır. Yapay sinir ağları insan beyninin çalışma şeklini yani biyolojik sinir sistemini temel almaktadır. Bu konudaki ilk çalışmalar beyni oluşturan biyolojik üniteler olan nöronların modellenerek bilgisayar sistemine aktarılması ile başlamıştır. Bilgisayar sistemlerinin gelişimi ile birlikte birçok alanda kullanılır hale gelmiştir [30].

Yapay sinir ağları gerçek hayattaki problemlerin çözümünde oldukça kapsamlı uygulama alanına sahiptir. Günümüzde farklı alanlarda başarılı bir şekilde uygulanmaktadır. Uygulama alanları içerisinde tahminleme de yer almaktadır. Bu sayede geleneksel tahmin yöntemlerinin dışında ileriye yönelik tahmin yapılabilmesine imkân vermektedir. Yapay sinir ağlarının tahmin yapmak amacıyla kullanılmasının birçok sebebi vardır. Geleneksel tahmin yöntemlerinde varsayımların sağlanması gerekmekte iken yapay sinir ağları için herhangi bir varsayıma gerek duyulmaz. Değişkenler arasında yer alan fonksiyonel ilişkilerin açıklanması mümkün olmadığında bile, yapay sinir ağları örneklerden hareket ederek öğrenme yolu ile bu ilişkiyi tespit edebilmektedir. Veride yer alan eğilim veya yapının (pattern) tanımlanmasında en iyi yöntem olmalarından ötürü, öngöründe bulunmak için oldukça uygundurlar. Aynı zamanda yapay sinir ağları genelleme yapma imkânı da verir. Diğer bir ifade ile yapay sinir ağları Kendi kendine öğrenme yetenekleri sayesinde mevcut örnekleri kullanarak daha önce hiç karşılaşılmayan durumlar için genelleme yapabilirler.

Yapay sinir ağlarının önemli bir diğer özelliği ise gerçek hayatta olası doğrusal olmayan yapıları da hesaba katmasıdır. Bu özellikleri sayesinde yapay sinir ağları doğrusal olmayan karmaşık problemlerin çözümü için önemli bir araç haline gelmiştir [31].

Yapay sinir ağları ile talep tahmini yapılırken geleneksel yöntemlerdeki gibi geçmiş veriler kullanılır. Talep tahmini için kullanılan tipik bir yapay sinir ağı; girdi katmanı, gizli katman ve çıktı katmanı olmak üzere üç katmandan oluşmaktadır. Girdi katmanı her bir bağımsız değişken için bir tane hücre taşır. Çıktı katmanı bağımlı değişken sayısı kadar hücre taşır. Gizli katman ise herhangi bir sayıda hücre taşıyabilir. Ancak çoğu zaman girdi sayısının iki katını geçmez. Girdi katmanında yer alan hücreler ile ara katmanda yer alan hücreler arasında ve ara katmanda yer alan hücreler ile çıktı katmanında yer alan hücreler arasında ağırlıklar bulunmaktadır. Sinir ağları, uygun ağırlık değerlerinin belirlenmesi amacıyla eğitilir. Uygun bir çözüm bulununcaya kadar bir dizi iterasyonla eğitime devam edilir [31].

Yapay sinir ağlarında bir takım işlemlere tabi tutulan sinyaller vardır. Bu sinyaller birleştirilip sayısal bir sonuç üretilir. Yapay sinir ağları; öğrenme, ilişkilendirme, sınıflandırma, genelleme, tahmin, özellik belirleme ve optimizasyon işlemlerini yapabilmektedir. Öğrenme işlemi süresince verilen bilgiler doğrultusunda kendini düzenleyip daha sonraki girdiler için doğru karar verebilme yeteneğine sahiptirler [32].



Şekil 2.1. Basit bir yapay sinir ağı modelinin yapısı

Şekil 2.1’de görüldüğü gibi her bir girdi değişkeni (X_1, X_2, X_3) girdi düğümünden geçmekte ve bir ağırlık katsayısı ile çarpıldıktan sonra çıktı düğümünde toplanıp çıktıyı (Y çıktı değişkeni) oluşturmaktadır. Şekilde gösterilen model basitleştirilmiş bir yapay sinir ağı

modelidir. Yapay sinir ağı modeli düğümler arası ağırlıkların düzeltilmesi süreci ile birlikte doğrusal olmayan karmaşık ilişkileri modellemekte ve öğrenmektedir [33].

Genel olarak yapay sinir ağları canlı beyninin yapısını gerçekleştirmeyi amaçlar. Buna bağlı olarak yapay sinir ağları:

- Öğrenme
- İlişkilendirme
- Sınıflandırma
- Genelleme
- Tahmin
- Özellik belirleme
- Optimizasyon

İşlemlerini gerçekleştirirler. Bu işlemleri gerçekleştiren sinir ağlarının ortak özelliği ise bir müdahale olmaksızın, elinde bulunan bilgilere göre sonuç üretebilmesidir. Yapay sinir ağları öğrenme esnasında verilen bilgiler ile kendini düzenleyerek daha sonraki girdiler için doğru kararlar verebilme yeteneğine sahiptir [32].

Öğrenme sürelerinin çok uzun olması ve ağ yapısı içerisinde yer alan parametrelerin yorumlanabilirliğinin düşük olması, yapay sinir ağlarının en zayıf yönleri arasında yer alır. Modelde bulunan gizli katmanın ve öğrenilen ağırlıkların arkasında yatan sembolik anlamın insanlar tarafından yorumlanması bir hayli zordur [34].

Buna karşılık gürültülü (hatalı) verilere karşı dirençli olması, araştırmacının veri hakkındaki bilgisinin kısıtlı olduğu durumlarda dahi kullanılabilmesi, pek çok karar ağacı algoritmasından farklı bir biçimde girdi veya çıktı olarak sürekli değişkenlerin kullanılmasına uyum sağlayabilmesi ise yapay sinir ağı modelinin güçlü yanları olarak karşımıza çıkmaktadır [34].

2.2. Biyolojik Sinir Hücresi ve Yapay Sinir Hücresi

Biyolojik sinir sistemi verilerin alınması, yorumlanması ve karar üretilmesi gibi işlevlerin yürütüldüğü beyin merkezli 3 katmanlı bir yapıdır. Uyarılar alıcı sinirler ile beraber elektriksel sinyallere dönüştürülerek beyne iletilmektedir. Beyin tarafından oluşturulan

çıktılar ise, tepki sinirleri aracılığı ile belirli tepkilere dönüştürülür. Biyolojik sinir sisteminin blok gösterimi Şekil 2.2'deki gibidir [12].



Şekil 2.2. Biyolojik sinir sisteminin blok gösterimi

Tahmin edici algoritmaların esas amacı girdi ve çıktı değişkenleri arasındaki ilişkileri modellemektir. Yapay sinir ağı yönteminin yaklaşımı, biyolojik bir nöronun yapısına benzeyen matematiksel bir ifade geliştirmektir. Bu sebeple yapay sinir ağı tekniğini geliştirenler modelleme sürecinin çalışmasını anlatırken çok sayıda biyolojik terim kullanmışlardır [33]. Sinir sistemi ile yapay sinir ağı arasındaki ilişki Çizelge 2.1'de verilmiştir [30].

Çizelge 2.1. Sinir sistemi ile yapay sinir ağları arasındaki ilişki

SİNİR SİSTEMİ	YAPAY SİNİR AĞI SİSTEMİ
Nöron	İşlem elemanı
Dendrit	Toplama fonksiyonu
Hücre gövdesi	Aktivasyon fonksiyonu
Aksonlar	Eleman çıktısı
Sinapslar	Ağırlıklar

2.3. Yapay Sinir Ağları Bileşenleri

Bu kısımda yapay sinir ağlarını oluşturan bileşenler kısaca açıklanacaktır. Bunlar girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktılarıdır.

Girdiler: Yapay sinir ağı hücresine dışarıdan veya diğer hücrelerden verilen bilgilerden oluşmaktadır.

Ağırlıklar: Hücreler arasındaki bağlantıların sayısal değerini gösterir. Hücreye gelen bilginin etkisini ifade eder.

Toplama fonksiyonu: Hücreye gelen net girdinin hesaplanması için kullanılan fonksiyondur. Toplam fonksiyonu için değişik fonksiyonlardan yararlanılır. En yaygın olarak kullanılan toplama fonksiyonu, ağırlıklı toplamı bulan fonksiyondur. Bu fonksiyonda gelen her bir girdi değeri ağırlığı ile çarpılarak toplanır. Diğer bir ifade ile Eş. 2.1'deki gibi hesaplanır.

$$NET = \sum_i^n G_i * A_i \quad (2.1)$$

Burada G girdileri, A ağırlıkları, n ise hücredeki toplam girdi sayısını ifade etmektedir. Her model ve her uygulama için bu toplama fonksiyonunun kullanılması şart değildir. Bazı modellerde kullanılacak toplama fonksiyonu farklılık gösterebilir. Daha karmaşık olmakla birlikte farklı toplama fonksiyonları da kullanılabilir. Bunlar Çizelge 2.2'de gösterilmiştir [35].

Çizelge 2.2. Toplama fonksiyonlarının toplama işlevleri

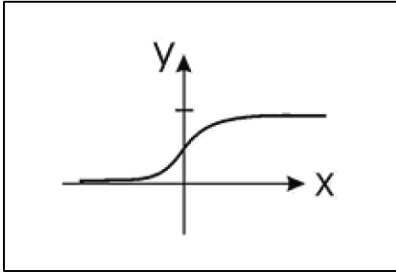
Net giriş	Açıklama
Çarpım Net Girdi= $\prod_i G_i * A_i$	Ağırlık değerleri ve girdiler çarpılır. Bulunan değerler de birbirleri ile çarpılarak net girdiye ulaşılır.
Maksimum Net Girdi= $\text{Max} (G_i * A_i) , i=1,2,\dots,n$	n adet girdi ve ağırlıklar birbiriyle çarpıldıktan sonra en büyüğü yapay sinir hücresinin net girdisi olarak belirlenir.
Minimum Net Girdi= $\text{Min} (G_i * A_i) , i=1,2,\dots,n$	n adet girdi ve ağırlıklar birbiriyle çarpıldıktan sonra en küçüğü yapay sinir hücresinin net girdisi olarak belirlenir.
Çoğunluk Net Girdi= $\sum_i \text{sgn}(G_i * A_i)$	n adet girdi ve ağırlıklar birbiriyle çarpıldıktan sonra pozitif ve negatif olanların sayısı bulunur. Bunlardan büyük olan sayı hücrenin net girdisi olarak belirlenir.
Kümülatif Toplam Net Girdi= $\text{Net}(\text{eski}) + \sum_i (G_i * A_i)$	Hücreye gelen bilgiler ağırlıklı olarak toplanır ve daha önce gelen bilgilere eklenerek hücrenin net girdisi bulunur.

Aktivasyon fonksiyonu: Transfer fonksiyonu olarak da bilinir. Hücrenin net girdi bilgisine göre çıktı bilgisini üreten fonksiyondur. Bu fonksiyon genellikle doğrusal olmaz. En yaygın olarak sigmoid fonksiyonu kullanılmakla birlikte çeşitli aktivasyon fonksiyonları vardır [30].

Sigmoid fonksiyonunun matematiksel ifadesi Eş. 2.2’de verildiği gibidir.

$$F(NET) = \frac{1}{1 + e^{-NET}} \quad (2.2)$$

Buradaki NET girdi değeri (Eş. 2.1 veya Çizelge 2.2’den biri alınabilir) toplama fonksiyonu kullanılarak belirlenen değerdir. Sigmoid fonksiyonu şekilsel olarak Şekil 2.3’te gösterilmiştir.



Şekil 2.3. Sigmoid fonksiyonu

Aktivasyon fonksiyonu olarak kullanılabilecek diğer fonksiyon örnekleri Çizelge 2.3’te gösterilmiştir [35].

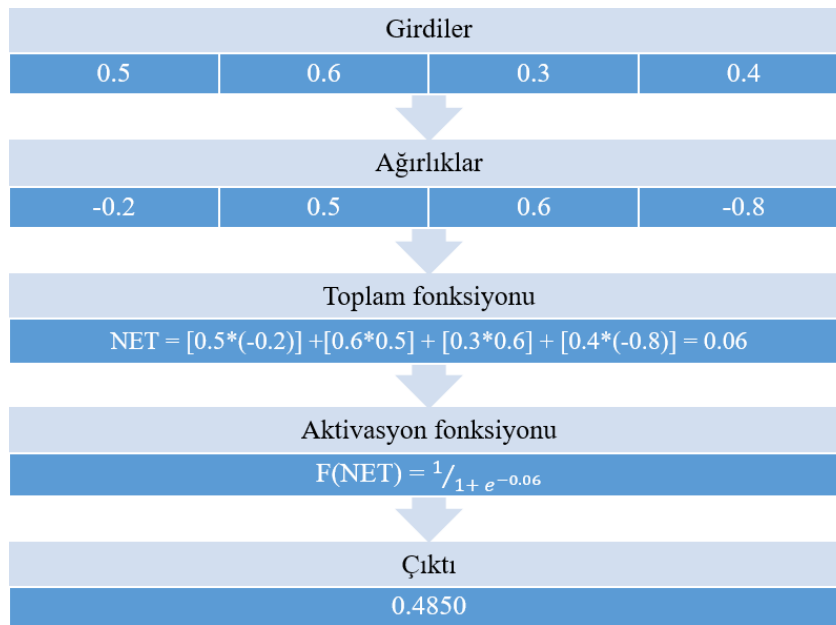
Çizelge 2.3. Aktivasyon fonksiyonu örnekleri

Aktivasyon fonksiyonu	Açıklama
Doğrusal fonksiyon $F(NET) = NET$	Hücreye gelen girdileri olduğu gibi çıkışa veren fonksiyondur.
Adım aktivasyon fonksiyonu $F(NET) = \begin{cases} 1 & \text{if } NET > \text{eşik değeri} \\ 0 & \text{if } NET \leq \text{eşik değeri} \end{cases}$	NET girdi değerinin belirlenen eşik değerin altında veya üstünde olmasına göre hücrelere ait çıktının 1 veya 0 değerini almaktadır.
Sinüs fonksiyonu $F(NET) = \sin(NE T)$	Öğrenilmesi düşünülen olaylar sinüs fonksiyonuna uygun dağılım gösteriyorsa aktivasyon fonksiyonu olarak sinüs fonksiyonu kullanılmaktadır.
Eşik değeri fonksiyonu $F(NET) = \begin{cases} 0 & \text{if } NET \leq 0 \\ NET & \text{if } 0 < NET < 1 \\ 1 & \text{if } NET \geq 1 \end{cases}$	Gelen bilgilerin 0 veya 1’den büyük ya da küçük olması durumuna göre değeri alır. 0 ve 1 arasında değeri alabilir. Bunların dışında bir değeri alamaz.

Çıktılar: Aktivasyon fonksiyonundan sonra çıkan değeri, çıktı değeri ifade etmektedir.

2.4. Yapay Sinir Hücresinin Çalışma Prensibi

Bu kısımda yapay sinir hücresinin çalışma prensibini daha iyi anlamak için bir örnek verilmiştir. Bu örnekte dört girdi değeri ve bu girdilere ait ağırlıklar aşağıdaki biçimde verilmiş olsun. Toplama fonksiyonu olarak Eş. 2.1’de verilen ağırlıklı toplamı bulan fonksiyon ve aktivasyon fonksiyonu olarak ise Eş. 2.2’de verilen sigmoid fonksiyonu kullanılsın. Bu verilen bilgiler ile oluşacak olan çıktı Şekil 2.4’de verilmiştir [35].



Şekil 2.4. Bir yapay sinir hücresinin çalışması örneği

Hücreye gelen NET girdi şu şekilde hesaplanmaktadır.

$$NET = 0.5*(-0.2) + 0.6*0.5 + 0.3*0.6 + 0.4*(-0.8) = 0.06$$

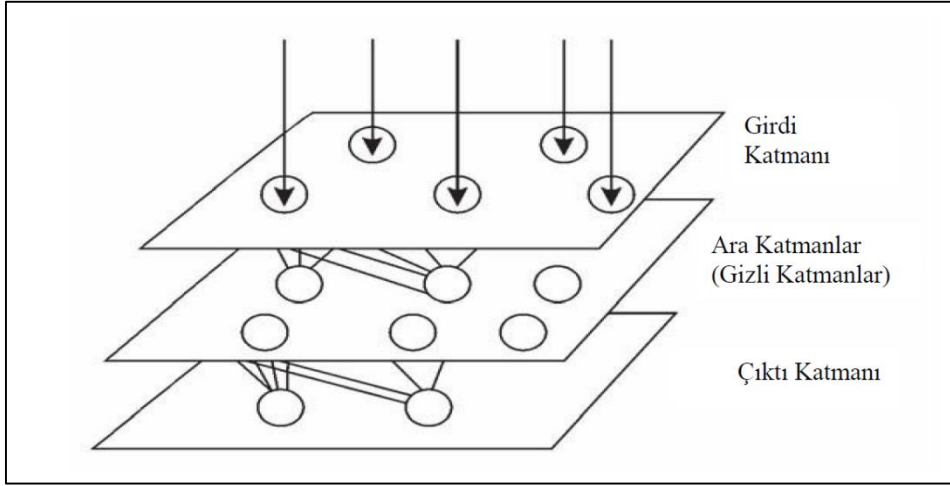
Hücrenin sigmoid fonksiyonu uygulandıktan sonra çıktısı ise;

$$F(NET) = \frac{1}{1 + e^{-0.06}}$$

$F(NET) = 0.4850$ bulunur. Şekil 2.4’de verilen örnek ile girdilere karşılık çıktıların nasıl üretildiği görülmektedir.

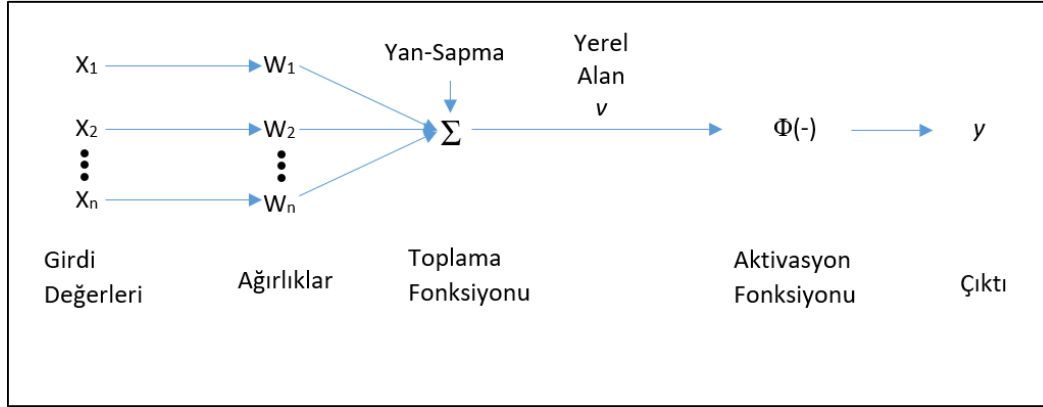
2.5. Yapay Sinir Ağlarının Yapısı

Yapay sinir ağı üç katmandan oluşmaktadır. Bu katmanlar girdi katmanı, ara katmanlar (gizli katmanlar) ve çıktı katmanı olarak isimlendirilir. Bu durum Şekil 2.5’de özetlenmiştir [36].



Şekil 2.5. Yapay sinir ağı katmanlarının birbirleriyle ilişkisi

Bir yapay sinir ağı modeli çoğunlukla doğrusal olmayan karmaşık ilişkilerin modellenmesi amacıyla kullanılır. Modellemede girdi ve çıktı katmanları dışında başka katmanların da yer almasına imkân verir. Bu katmanlar ara (gizli) katmanlar olarak isimlendirilir. Öğrenme safhasında ağırlıklar, girdi sınıfını en başarılı biçimde tahmin edene kadar değiştirilir. Daha açık bir ifade ile öğrenme setinde yer alan her bir gözlem için tahmin değeri hesaplanarak gerçek değerler ile karşılaştırılır. Daha sonra model, ilk durumda ortaya çıkan hataları kullanarak ağırlıkları düzeltir. Hata makul bir seviyeye inene kadar düzeltme işlemi devam eder. Ağırlıkların belirlenmesinden sonra girdiler eş zamanlı olarak girdi katmanına girer ve ağırlık katsayıları ile çarpılır. Bu işlemin ardından girdikleri ara (gizli) katman, başka ara (gizli) katmanlar için girdi olabilir. Son gizli katmandan elde edilen ağırlıklı çıktılar, çıkış katmanına girdi olarak dâhil edilirler. Burada yapay sinir ağının tahmini yer alır [34].



Şekil 2.6. Yapay sinir ağı örneği

Teknik açıdan yapay sinir ağının görevi, girdi seti olarak kendisine verilen bilgilerden Şekil 2.6’da verildiği gibi bir çıktı üretebilmektir. Bunun yapılabilmesi için ağ belirli örneklerle eğitilir. Sonra ağ genelleme yapabilecek ve karar verebilecek seviyeye ulaşır. Bu kazanılan yetenek ile çıktılar belirlenir [3].

2.6. Yapay Sinir Ağlarında Öğrenme

Yapay sinir ağlarında işlem elemanlarının bağlantılarının ağırlık değerlerinin belirlenmesine “ağın eğitilmesi” denir. Başlangıçta bu ağırlık değerleri rasgele olarak atanmaktadır. Bu ağırlık değerleri yapay sinir ağlarının kendilerine örnekler göstermesiyle değişmektedir. Amaç ağa gösterilen örnekler için doğru çıktıları üretecek şekilde ağırlık değerlerini belirlemektir. Örnekler ağa defalarca (iterasyonlarda) gösterilerek en doğru ağırlık değerleri bulunur. Ağın doğru ağırlık değerlerine ulaşması ile birlikte örneklerin temsil ettiği olay hakkında genellemeler yapabilme yeteneğine kavuşur. Bu genelleştirme özelliğine kavuşması işlemi “ağın öğrenmesi” olarak ifade edilir [37].

$X_i (i = 1, \dots, m)$ değişkenleri sinir girişlerinden ağa girer.

$w_i (i = 1, \dots, m)$ ağırlıklardır.

$$y_i = f(w, x) = f(w^T x) = f(\sum_{j=1}^m w_{ij} x_i)$$

Burada;

$w_i = (w_1, \dots, w_m)^T$ olarak karşımıza çıkar.

Eğitim aşaması yapay sinir ağları için en önemli aşama olmaktadır. Eğitim ve öğrenme birbirinden farklı kavramlardır. Eğitme ağı öğrenme aşamasında gerçekleşen adım olurken, öğrenme ise eğitim işlemi sonucunu ifade etmektedir. Eğitme aşaması, ele alınan problemin özelliği dikkate alınarak öğrenme kuralının yapay sinir ağlarına nasıl uyarlanabileceğini gösterir [38].

Yapay sinir ağının eğitimi tamamlandıktan sonra performansının (öğrenip öğrenmediğinin) test edilmesi için bir takım denemeler yapılmakta ve bu denemelere “ağın test edilmesi” denilmektedir. Bazı paket programlarda ağ eğitilirken “doğrulama seti” (validation set) ile ağın her iterasyonunda ne kadar öğrenmiş olduğu test edilir. Öğrenim kriteri olarak doğrulama setinin hata değerleri kullanılır. Eğitim setinde yer alan verilerle istenilen başarı sağlanmışsa ağ daha evvel görmediği veriler ile test edilir. Test işleminin yapıldığı veri seti “test seti” olarak adlandırılır. Test işlemi sırasında ağın ağırlık değerlerinin değiştirilmemesi gerekir. Örnekler ağa gösterildikten sonra eğitim sonucunda belirlenmiş olan ağırlık değerleri kullanılarak daha evvel görülmeyen bu örnekler için çıktılar üretilmektedir. Karşımıza çıkan çıktı değerlerinin doğruluk derecesi ağın öğrenmesi hakkında bilgi vermektedir. Sonucun iyi olması eğitimin performansının da iyi olduğunu göstermektedir. Test seti ile istenen sonuçlara ulaşılması durumunda ağın eğitimi tamamlanmış olur [39].

2.6.1. Öğrenme yöntemleri

Bu kısımda yapay sinir ağlarındaki öğrenme yöntemleri açıklanacaktır. Öğrenme yöntemleri danışmanlı öğrenme, destekleyici öğrenme ve danışmansız öğrenme olmak üzere üç başlık altında ele alınmaktadır.

Danışmanlı öğrenme

Buradaki öğrenmede ağ kullanılmadan önce mutlak manada eğitilmesi gerekmektedir. Eğitim elemanları giriş ve çıkış verilerinden oluşmaktadır. Bir giriş grubu için uygun çıkış grubu belirlenmelidir. Burada gerçek bilgilerin yapay sinir ağına sunulması zorunludur. Eğitim işlemi bu durumdan ötürü süre olarak çokça uzamaktadır. Eğitim işlemi, ancak istenen istatistiksel doğruluk sonucunda sonuçlanır. Eğitim bittiğinde ise elde edilen ağırlık değerleri sabit kalır. Yapay sinir ağları yalnızca matematiksel giriş verileri ile çalışmaktadır. Eldeki bilgilerin ölçeklendirilerek sayısal verilere dönüştürülmesi şartıyla giriş yapılabilir.

Yapay sinir ağında elde edilen çıktılar ile olası beklenen çıktı arasındaki hata hesaplanarak doğru ağırlıklar bu pay doğrultusunda yeniden hesaplanır. Daha sonra birer birer tüm ağırlıklar güncellenir [40].

Destekleyici öğrenme

Destekleyici öğrenme yönteminde amaç, gerçek bir çıkış değeri vermekten ziyade ağı çıkış değerlerinin ne kadar doğru verildiğine dair ağı performansına ait değerin iyi veya kötü olarak derecelendirilmesidir. Denetim hareketlerinin başarısı ile ilgilenilmektedir. Danışmanlı öğrenme yönteminde ağa giriş verileri ile birlikte çıkması istenilen sonuçlarda yüklenir. Ancak bazı durumlarda ağı nasıl bir sonuca ulaşması gerektiği bilgisi ağa tanıtılamamaktadır. Çünkü bu konuda elde edilmiş araştırma sonuçları mevcut olmayabilir. Ancak ağ sonucunun doğru veya yanlış olup olmadığı sorgulanabilir. Bu şekildeki öğrenme yöntemine destekleyici öğrenme denir. Kohonenin ağlarındaki LVQ (Linear Vektör Quantitation - Doğrusal Vektör Parçalama Yöntemi) bunun en yaygın kullanım şeklidir [20], [41], [42].

Danışmansız öğrenme

Danışmansız öğrenme yönteminde ise sistemin öğrenmesine yardımcı olan herhangi bir öğretmen yoktur. Sisteme yalnızca girdi değerleri gösterilir. Örneklerdeki parametreler arasındaki ilişkileri sistemin kendi kendisine öğrenmesi beklenir. Bu, daha çok sınıflandırma problemleri için kullanılan bir yöntemdir. Fakat kullanıcı tarafından, sistemin öğrenmesi sona erdikten sonra çıktıların ne anlama geldiğini belirten etiketlendirmenin yapılması gerekmektedir [41].

2.7. Yapay Sinir Ağlarında Öğrenme Kuralları

Ağırlık değerlerinin değişmesi belirli kurallar çerçevesinde yürütülmektedir. Bu kurallar öğrenme kuralları olarak isimlendirilir [37].

Literatürde birçok öğrenme algoritması kullanılmaktadır. Bu öğrenme algoritmalarının büyük bir kısmı matematik tabanlı olmakla birlikte ağırlıkların güncelleştirilmesinde kullanılır. Öğrenme işlemi çok parametrelili, karmaşık ve matematiksel olarak ifadesi zor bir işlemdir. Bugün kullanılan öğrenme kuralları bu işlemin basitleştirilmiş veya matematiksel

olarak ifade edilmiş biçimleridir. Mevcut olan öğrenme algoritmalarının birçoğu Hebb, Delta, Kohonen ve Hopfield olmak üzere dört farklı öğrenme kuralından ilham alarak geliştirilmiştir [39].

2.7.1. Hebb kuralı

1949 yılında Donald Hebb tarafından geliştirilen Hebb kuralı, ilk ve en bilinen öğrenme kuralıdır [4]. Yapay sinir ağlarında öğrenme kurallarının çoğu Hebb kuralına dayanır. Tanımı, “The Organization of Behavior” kitabında anlatılmıştır. Kurala göre, bir hücre diğer bir hücreden bilgi alırsa ve her iki hücre de aktif ise her iki hücre arasındaki bağlantı kuvvetlendirilmelidir. Eğer A hücresi B hücresini uyarmaya yetecek kadar yakınsa ve B hücresinin aktifleştirilmesinde devamlı olarak yer alıyorsa, A hücresinin etkinliği arttırılacak şekilde bir veya her iki hücrede de değişiklikler yapılır. Bu kural bir hücre kendisi aktif ise bağlı olduğu hücreyi aktif yapmaya, kendisi pasif ise bağlı olduğu hücreyi pasif yapmaya çalışır şeklinde özetlenebilir [43].

2.7.2. Hopfield kuralı

Bu kural Hebb kuralı ile benzerlik göstermektedir. Yapay sinir ağına ait elemanlara ilişkin bağlantıların kuvvetlendirilmesi veya zayıflatılması gerektiğine karar vermektedir. Beklenen çıktı ve girdi değerlerinin her ikisi de pasif veya aktif çıkmışsa öğrenme katsayısı kadar ağırlıkları kuvvetlendirilmektedir/ zayıflatılmaktadır. Genel olarak, ağırlıkların öğrenme katsayısı ile zayıflatılıp/kuvvetlendirildiği görülür. Öğrenme katsayısı ise çoğu zaman 0 ve 1 arasında olup kullanıcı tarafından atanmış olan pozitif ve sabit bir değerdir [44].

2.7.3. Delta kuralı

Delta kuralı Widrow ve Hoff tarafından 1960 yılında ortaya atılmış olan yinelemeli bir öğrenme süreci olup Widrow ve Hoff kuralı olarak da bilinmektedir. Delta kuralında amaç, toplam hatanın küçültülmesidir. Yani girdi numuneleri için çıktı ve hedefin farkların karelerinin toplam değerlerinin küçültülmesi hedeflenmektedir. Tüm eğitim numunelerinin hatalarının en aza indirilmesidir. Çok sayıda eğitim numunesi ile birlikte biriktirilebilen ağırlık düzeltmeleri, yığın güncelleştirilmesi olarak isimlendirilir [45].

2.7.4. Kohonen kuralı

Yarışmacı Öğrenme Kuralı olarak da bilinen Kohonen kuralı, Teuvo Kohonen (1982) tarafından biyolojik sistemlerde yer alan öğrenme yönteminden esinlenilerek geliştirilmiştir [46]. Bu kural kapsamında ağırlık elemanları ağırlıkları değiştirmek için kendi içlerinde yarışır [47]. Hebb kuralından farkı ise, tek seferde sadece bir işlemci elemanın, yani kazanan nöronun bağlantı ağırlıklarının değiştirilmesidir. En uygun çıktıya sahip olan işlemci elemanın kazandığı bu kural, kazanan işlemci elemanın, kendisine komşu olan başka bir işlemci elemanın veya elemanların ağırlıklarının değiştirilmesine olanak sağlamaktadır. Komşu sayısı eğitim boyunca farklılık göstermektedir. Eğitim süresince en kapsamlı komşu tanımından en dar komşu tanımına doğru inilmektedir. Girdi desenini en iyi bir şekilde ifade eden eleman kazanan eleman olarak adlandırılmaktadır [47].

2.8. Yapay Sinir Ağı Modelleri

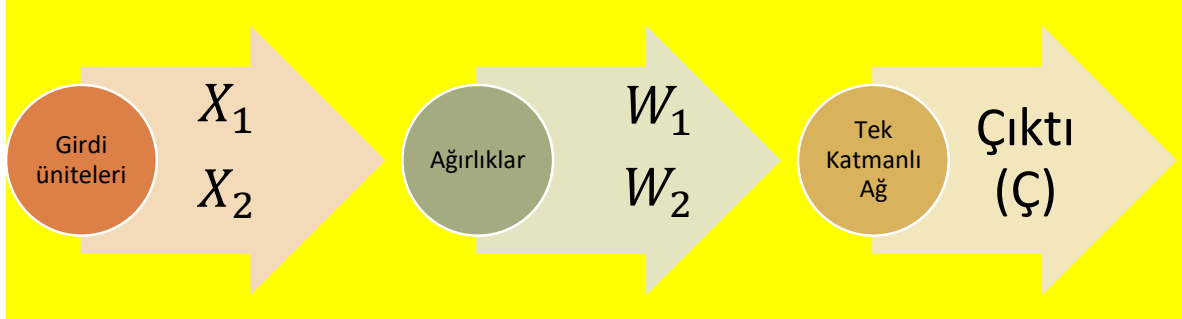
Yapay sinir ağlarına ait süreç elemanlarının birbirine bağlanması sonucu oluşmuş olan topoloji, süreç elemanlarına ait toplama ve aktivasyon fonksiyonları, öğrenme stratejileri ve kullanılmakta olan öğrenme kuralı kapsamında çeşitli ağ modelleri yer almaktadır [48]. Yapay sinir ağları temelde tek katmanlı ya da çok katmanlı olmak üzere sınıflandırılmaktadır. Katman sayısının belirlenme aşamasında, girdi katmanı bir katman olarak ele alınmamaktadır; çünkü bu girdi katmanı herhangi bir hesaplama işleminde yer almaz [12]. Yapay sinir ağının tek katmandan oluşması durumunda tek katmanlı yapay sinir ağı, birden çok katmandan oluşması durumunda ise çok katmanlı yapay sinir ağı olarak isimlendirilmektedir [49].

2.8.1. Tek katmanlı algılayıcı modeli

Yapay sinir ağlarının önemli bir sınıfını oluşturan tek katmanlı yapay sinir ağları, kendisinden daha karmaşık yapıya sahip olan çok katmanlı yapay sinir ağlarına ışık tutarlar [50].

Tek katmanlı yapay sinir ağları en basit ağ yapısına sahiptir. Bir giriş ve bir çıkış katmanından oluşmaktadır. Her ağda bir veya birden fazla girdi (X) ve çıktı (Ç) yer

almaktadır. Çıktı üniteleri tüm girdi ünitelerine bağlanır. Her bağlantıya ait bir ağırlık (W) vardır [43]. İki girdi ve bir çıktıya sahip tek katmanlı ağ yapısı Şekil 2.7’de verilmiştir.



Şekil 2.7. Tek katmanlı algılayıcı modeli genel yapısı

Bu ağlar, süreç elemanlarına ait değerlerin sıfırdan farklı olmasını sağlayan bir eşik değere (Φ) sahiptir. Eşik değere ait girdi her daim 1 değerini alır. Ağın çıktısı, ağırlıklandırılmış girdi değeri ile eşik değerinin toplanması sonucu bulunmaktadır. Çıktı fonksiyonu doğrusal bir fonksiyon olmaktadır. Ağa gösterilmiş olan örnekler iki sınıf arasında paylaştırılır. Bu sayede iki sınıfı birbirinden ayıran doğru belirlenmeye çalışılır. Bunun için eşik değer fonksiyonundan yararlanılır. Sınıf ayırıcı olarak isimlendirilen doğru Eş. 2.3’deki gibi tanımlanır.

$$W_1 * X_1 + W_2 * X_2 + \Phi = 0 \quad (2.3)$$

Bu ağlarda öğrenme, ağın sınıf ayırıcı doğrusunun sahip olduğu pozisyonu, her iki sınıfı en iyi biçimde ayıracak olan pozisyonu belirlemektir. Bu yüzden ağırlık değerlerinin değiştirilmesi gerekmektedir. Bu durumda t zaman biriminde ağırlık değerleri ΔW kadar değiştirilerek, yeni ağırlık değerleri Eş. 2.4’deki gibi elde edilir.

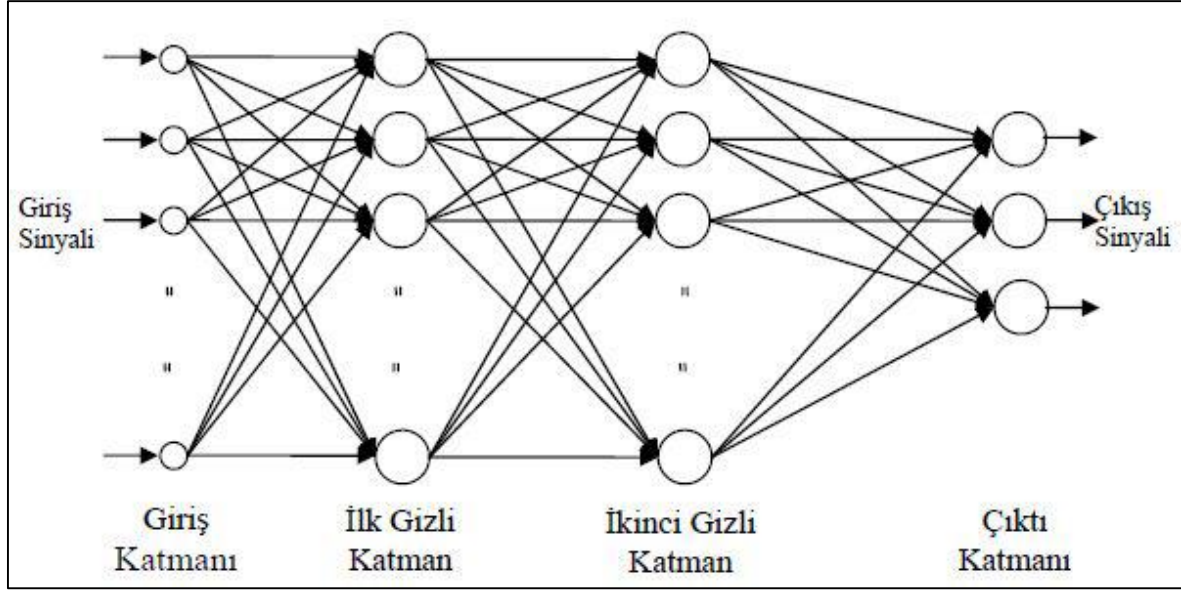
$$W_i(t + 1) = W_i(t) + \Delta W_i(t) \quad (2.4)$$

Ağırlıklar değiştirildiğinde doğrunun eğimi değişmektedir. Ağırlıkların değiştirilmesi yeterli olmayabilir. Bu durumda eşik değerinin de değiştirilmesi gerek. Buna göre t zaman biriminde eşik değeri $\Delta \Phi$ kadar değiştirildiğinde yeni eşik değeri Eş. 2.5’deki hale gelir [41].

$$\Phi_i(t + 1) = \Phi_i(t) + \Delta \Phi_i(t) \quad (2.5)$$

2.8.2. Çok katmanlı algılayıcı modeli

Çok katmanlı ağların (Multi Layer Perceptron - MLP) çalışma prensibi ile tek katmanlı ağların çalışma prensibi aynı biçimdedir. Çok katmanlı yapay sinir ağlarının farkı ise girdi ve çıktı katmanları arasında yer alan gizli katmanlardır [51]. Çok katmanlı ağ yapısı Şekil 2.8'de özetlenmiştir.



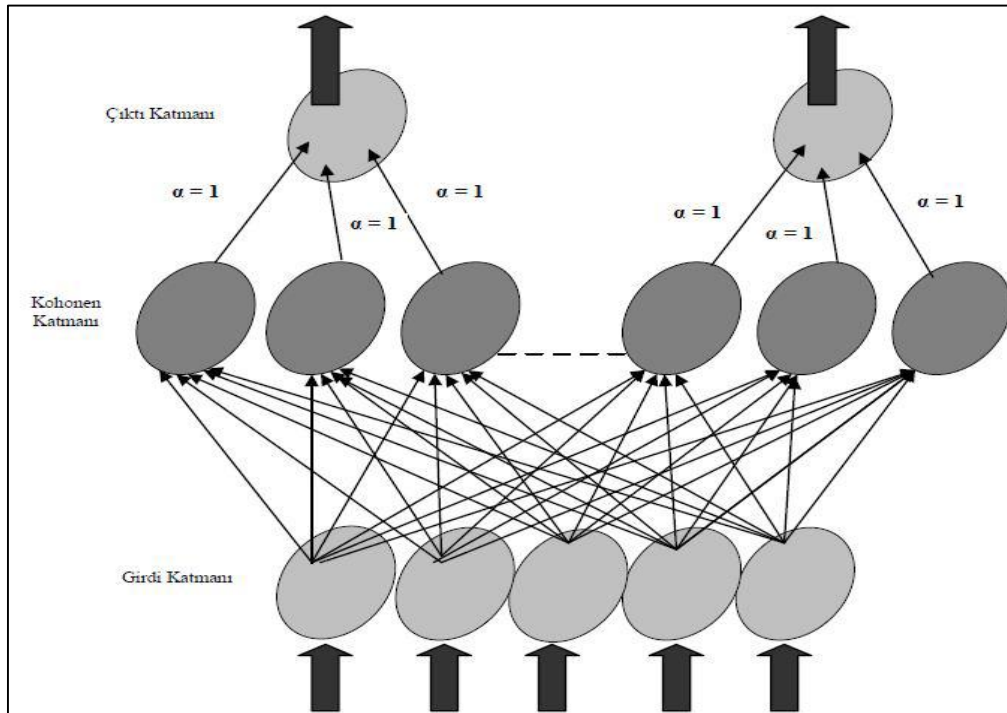
Şekil 2.8. Çok katmanlı algılayıcı modeli genel yapısı

Bir yapay sinir ağında ağın öğrenmesi istenen olaylara ait girdi ve çıktı arasındaki ilişkiler doğrusal olmadığında tek katmanlı algılayıcılar ile öğrenme gerçekleştirilemez. Günlük meydana gelen olayların çoğu doğrusal olmayan niteliktedir ve doğrusal olmayan problemlerin çözümünde yeni modellerin geliştirilmesine gereksinim duyulmuştur. Bu sorunun çözülebilmesi için çok katmanlı algılayıcılar geliştirilmiştir [40].

2.8.3. LVQ ağları modeli

LVQ (Learning Vector Quantization) ağı Kohonen tarafından 1984 yılında geliştirilmiştir [52]. Danışmansız öğrenme kuralına dayalı modellerde eğitim süresince ağı hem girdi değerleri hem de girdi değerleri için üretilecek olan çıktı değerlerinin ne olması gerektiği verilmektedir. Ağı, çıktı değerlerinin ne olduğunu vermek bazen mümkün değildir. Sadece ağın ürettiği çıktı değerlerinin doğru ya da yanlış olduğu belirtilmektedir. LVQ modeli, destekleyici öğrenme olarak belirlenmiş olan bu yöntemi kullanan modellerden biridir [47].

LVQ algoritmasına ait ağ mimarisi, her bir çıkış biriminin bir sınıfla ilişkilendirilmesi bakımından farklı olmakla beraber, rekabetçi öğrenme ağlarına oldukça benzer bir yapıdadır. LVQ yöntemi iki aşamadan oluşmaktadır. İlk aşamada, veri kümesi için danışmansız bir öğrenme yöntemi kullanılır ve kümelenme merkezlerinin sınıf bilgisi kullanmadan konumlandırılması hedeflenir. İkinci aşamada ise, sınıf bilgileri kümelerin merkezlerinin konumunu iyileştirmek ve bu sayede yanlış sınıflandırılmış vakaların sayısını en aza indirmek hedeflenir [53].



Şekil 2.9. LVQ modeli genel yapısı

Şekil 2.9’da bir LVQ ağı yapısı gösterilmiştir. Çok katmanlı ağlardaki gibi LVQ ağları da üç katmandan oluşmaktadır. Görüldüğü gibi bir girdi katmanı, bir çıktı katmanı ve bir de Kohonen katmanı bulunmaktadır. Girdi katmanındaki nöronlar ara katmandaki tüm nöronlarla bağlıdır. Fakat ara katmandaki süreç elemanları yalnızca çıktı katmanındaki bir tek süreç elemanına bağlıdır. Kohonen katmanı ile çıktı katmanı arasındaki ağırlıklar (α) sabittir ve 1’e eşittir. Bu ağırlıklar değiştirilememektedir. Sadece girdi katmanı ile Kohonen katmanı arasındaki ağırlıklar değiştirilmektedir. Bu ağırlıklar üzerinden öğrenme gerçekleşmektedir. Bu ağın temel prensibi n boyutlu bir vektörün, bir vektörler setine haritalanmasıdır. LVQ ağlarının öğrenme hızı çok katmanlı ağlara göre daha yüksektir. Bir

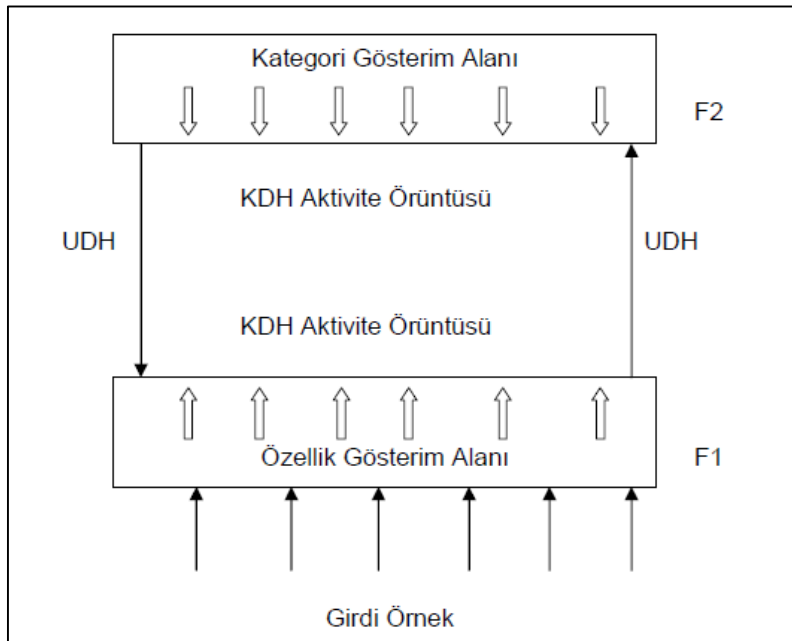
diğer ifade ile LVQ ağıları bir olayı çok katmanlı ağlardan daha kısa zamanda öğrenebilmektedir [41] .

2.8.4. ART ağları modeli

ART (Adaptive Resonance Theory) ağları danışmansız öğrenmeye dayalı bir yöntemdir. Bu ağlar beklenen çıktıları bir danışmandan almaktansa kendi kendine öğrenmeye çalışır. Öğrenme süresince kullanılmış olan örneklerden öğrendikleri bilgilerle daha sonra görülmemiş örnekler ile ilgili yorumlar yapabilmektedir.

ART ağlarında bilgilerin hem kısa dönemli hafızada (KDH) hem de uzun dönemli hafızada (UDH) saklanma durumu söz konusudur. Bazı bilgiler öğrenilirken bazı bilgiler de unutulmaktadır. Zaman içerisinde öğrenilen bilgiler ağıın olay ile ilgili karar vermesine neden olmaktadır.

ART ağlarının amacı sınıflandırma yapmaktır. Bu amaçla geliştirilmiş olan bu ağlar başarılı bir şekilde kullanılmaktadır. LVQ ağları da ART ağları gibi sınıflandırma için kullanılmakta olup aralarında bir takım farklılıklar söz konusudur. ART ağlarını LVQ ağlarından ayırt eden özelliği, ağa yapılacak sınıflandırma ile alakalı herhangi bir bilginin verilmemesidir. ART ağları sınıflandırmayı kendi başlarına yaparlar [41].



Şekil 2.10. ART modeli genel yapısı

ART ağırları yapısı Şekil 2.10'da gösterildiği gibi genellikle iki katmandan oluşmaktadır. F1 ve F2 olarak adlandırılan iki katman vardır. F1 katmanı girdinin özelliklerini göstermekte iken, F2 katmanı kategorileri ifade etmektedir. Bu iki katman birbirlerine UDH ile bağlanmaktadır. Girdilerin özellikleri F1 katmanının aktivasyonu tarafından belirlenmektedir. UDH'da bağlantı değerleri ile gelen bilgiler kategorilere ayrılır ve F2 katmanına gönderilir. F2 katmanındaki sınıflandırma ile F1 katmanından gelen sınıflandırma birbirleri ile eşleştirilir. Örneğin belirlenmiş bir sınıfa uyması durumunda o kategoride gösterilir. Aksi halde, ya yeni bir sınıf oluşturulur ya da girdinin sınıflandırılması yapılmaz [41].

2.9. Yapay Sinir Ağlarının Avantajları ve Dezavantajları

Yapay sinir ağlarının doğrusal ya da doğrusal olmayan problemler için tercih edilen aktivasyon fonksiyonu sayesinde, iyi çözümlere ulaşılabilir. Günlük hayatta karşılaşılan problemlerin birçoğu doğrusal olmadığından dolayı, yapay sinir ağlarının bu gibi konularda önemi anlaşılabilmektedir [54].

Yapay sinir ağlarının öğrenme sistemi, insan beyninin çalışma şekline benzerlik göstermektedir. Yapay sinir ağları yönteminde eldeki probleme ilişkin veriler kullanılmakta ve veri yapısındaki saklı ilişkiler ortaya çıkarılmaya çalışılmaktadır [49].

Günümüzde kullanılan bilgi işleme teknikleri genelde seri, sıralı işlemlerden oluşmaktadır. Seri işlemcilerde herhangi bir birimin yavaş olması tüm sistemi yavaşlamasına sebep olurken, yapay sinir ağlarının sahip olduğu paralellik sayesinde yavaş bir birim sistemin çalışması esnasında herhangi bir probleme yol açmamaktadır. Bu durum yapay sinir ağlarının hem daha hızlı hem de güvenilir olması sonuçlarını doğurmaktadır [36].

Yapay sinir ağları herhangi bir matematiksel modele gerek duymaz. Eksik veriler söz konusu olduğunda, bu eksik verilerin giderilmesi durumunda yüksek performans gösterirler. İlişkilendirme ve sınıflandırma yapabilirler. İlk kez karşılaşılmış durum hakkında bilgi verirler. Hücrelerden bir kısmının bozulmasına ya da kaybolmasına rağmen, anlamlı olan bilgiye zarar gelmez. Olayları öğrenerek benzer sorunlarla karşılaşıldığında gerçekçi kararlar verebilirler [1].

Yapay sinir ağlarının oluşturulmasında, model seçiminde, ağ topolojisinin belirlenmesinde bir kurallar bütünü yoktur. Kullanıcı tecrübesine dayanmaktadır. Basit olarak görülen modelleme yapıları olmasına rağmen ara sıra zor ve karmaşık olabilmektedir. Problemlere optimum sonuç garanti etmezler. Ancak üretilen sonuçların iyi sonuçlardan biri olduğu söylenebilir. Geleneksel yöntemler ise optimum sonuçlar üretirler.

Yapay sinir ağlarının eğitilmesine ve test edilebilmesine yetecek genişlikte veri setine ihtiyaç duyulur. Problemleri doğru temsil eden örnekler bulunamadığında sağlıklı sonuçlar üretebilmek mümkün olmaz.

Ağın davranışları açıklanamamaktadır. Bu sorun yapay sinir ağlarının en temel sorunudur. Yapay sinir ağları bir probleme çözüm ürettiğinde bunun neden ve nasıl olduğuna dair hiç bir ipucu vermez. Bu ağa olan güvenin azalmasına sebep olmaktadır [41].

Yapay sinir ağları ile geleneksel yöntemlerin tahmin başarılarının karşılaştırılması konusunda literatüre bakıldığı zaman çok farklı sonuçların elde edildiği gözükmemektedir. Bunun temel nedeni, yapay sinir ağlarının performansını etkileyen ağ topolojisi, eğitim algoritması, veri setleri vb. çeşitli faktörlerin olmasıdır. Bazı durumlarda yapay sinir ağlarının tahmin yapma başarısı, doğrusal istatistiksel yöntemlere göre düşük çıkmaktadır. Bunun nedeni, veriler arasındaki ilişkinin doğrusal olması ya da uygun ağ yapısının oluşturulamaması olabilir. Kıyaslamalarda kullanılacak tahmin yöntemleri seçiminde genelde iki durum söz konusudur. Bağımlı değişkeni etkileyen bağımsız değişkenler belirlenmişse yapay sinir ağları ile regresyona dayanan yöntemler kıyaslanırken, zaman serisi analizleri yapılması durumunda yapay sinir ağları ile hareketli ortalamalar, üstel durağanlaştırma ve Box-Jenkins yöntemleri kıyaslanmaktadır [55].

3. LOJİSTİK REGRESYON ANALİZİ

Bu kısımda çok kategorili lojistik regresyon analizi anlatılmadan önce genel olarak lojistik regresyon analizi hakkında bilgi verilecektir.

3.1. Lojistik Regresyon Analizinin Tanımı ve Özellikleri

Lojistik regresyon, açıklayıcı değişkenlerin değerlerine göre bağımlı değişkenin daha önce tanımlanmış gruplara atanma olasılıklarını hesaplayabilen çok değişkenli bir sınıflama yöntemidir [56].

Çok değişkenli istatistiksel yöntemlerden biri olan lojistik regresyon analizi, sınıflama ve atama işlemlerinde kullanılabilen bir regresyon yöntemidir. Lojistik regresyon analizi, normallik, ortak kovaryansa sahip olma gibi çeşitli varsayımların sağlanmadığı durumlarda diskriminant ve çapraz tablo analizlerine alternatif olarak uygulanmaktadır. Bağımsız değişken yapısı ve kombinasyonu yönünden diskriminant analizinden ayrılan lojistik regresyon analizinin klasik doğrusal regresyon analizinden üç önemli farkı vardır. Bunlar;

- Doğrusal regresyon analizinde tahmin edilecek olan bağımlı değişken sürekli iken, lojistik regresyon analizinde bağımlı değişken kesikli bir değer almaktadır.
- Doğrusal regresyon analizinde bağımlı değişkenin değeri, lojistik regresyon analizinde ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı tahmin edilir.
- Doğrusal regresyon analizinden bağımlı değişkenin normal dağılım göstermesi koşulu aranırken, lojistik regresyon analizinde böyle bir koşul yoktur [57].

Temel olarak lojistik regresyonda bağımsız değişkenler ile iki ya da çok sınıflı kategorik bağımlı değişken arasındaki ilişkinin tanımlanması için matematiksel modelleme yapmak amaçlanmaktadır [58].

Son yıllarda tıp, biyoloji, tarım ve ekonomi gibi birçok alanda kolay kullanımı ve yorumlanmasından dolayı lojistik regresyon yaygın olarak kullanılmakta ve tercih edilen bir yöntem haline gelmektedir. İstatistiksel uygulamalarda araştırmacılar tarafından genellikle bağımsız değişkenlerle bağımlı değişken arasında ilişki olup olmadığı analiz edilmek istenir. Yapılan istatistiksel analiz yöntemlerinde, verilerin yapısına göre en uygun yöntemin

belirlenmesi büyük önem arz etmektedir. Bağımlı değişken sürekli olduğunda, genellikle doğrusal regresyon modeli kullanılmaktadır. Doğrusal modellerin önemli bir varsayımı hata terimlerinin normal dağılıma sahip olmasıdır [59]. Tipik doğrusal regresyon modeli Eş. 3.1'deki gibi ifade edilmektedir.

$$Y_i = \sum_{k=0}^p \beta_k * X_{ik} + \varepsilon_i \quad (3.1)$$

Burada, Y; bağımlı değişken, β ; regresyon katsayıları, X; bağımsız değişkenler, ε ; hata terimi olmaktadır. Diğer bir şekilde; k bağımsız değişken ve n gözlem olduğu durumda doğrusal regresyon modeli,

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i, \quad i=1, \dots, n \quad (3.2)$$

şeklinde gösterilmektedir. Eş. 3.2'de verilen doğrusal regresyon modelinde bağımsız değişkenlerin kesikli ya da sürekli olmaları, modelin tahmininde kullanılacak yöntemi ve bu yöntemle elde edilen parametre tahminlerinin özelliklerini etkilememektedir. Bu yüzden modele girecek bağımsız değişkenler hem kesikli hem de sürekli değişkenler olabilirler. Buna karşın modeldeki bağımlı değişkenin kesikli bir yapıya sahip olmasının etkisi büyüktür. Bağımlı değişkenin kategorik olması durumunda normallik varsayımı bozulur. Bu nedenle klasik doğrusal model uygulanamaz [59].

Regresyonda, bağımsız değişken x verildiğinde, bağımlı değişken y nin koşullu ortalaması yani beklenen değeri Eş. 3.3'deki $E(Y|x)$ ile gösterilir. Doğrusal regresyonda bu ortalamanın x 'e göre doğrusal bir fonksiyon olduğu varsayılır.

$$E(Y|x) = \beta_0 + \beta_1 x \quad (3.3)$$

Doğrusal regresyon analizinde kullanılan teknikler, lojistik regresyona uyarlanabilir. Lojistik regresyona ait formül Eş. 3.4'deki gibidir [60].

$$\pi(x) = E(Y|x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad 0 \leq E(Y|x) \leq 1 \quad (3.4)$$

Burada, $g(x) = \ln \frac{\pi(x)}{1 - \pi(x)}$ logit dönüşümü uygulanırsa, Eş. 3.5 elde edilmektedir [61].

$$g(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \beta_0 + \beta_1 x \quad (3.5)$$

Eş. 3.5'teki, $\pi / (1-\pi)$ oranı odds adını almaktadır [62].

3.2. Çok kategorili (Multinomial) Lojistik Regresyon Analizi

Bağımlı Y değişkeninin ikiden fazla değer aldığı durumlarda çok kategorili lojistik regresyon analizi kullanılmaktadır. Çok kategorili lojistik regresyon modellerinde bağımlı değişken sınıflayıcı ölçme düzeyinde ölçülmüş olup en az üç kategoriye sahip olmalıdır. Çok kategorili lojistik model, esas olarak bağımlı değişkenin iki değer alabildiği durumun bir uzantısı olarak ifade edilebilir. Model ikili lojistik modele benzemekte ve dolayısıyla varsayımları da benzer olmaktadır. Genel olarak çok kategorili lojistik modelin varsayımlarını aşağıdaki şekilde sıralayabiliriz:

- Y bağımlı değişkenine ait tekrarlı değerler birbirinden bağımsızdır. Aynı değerler ile kodlanmış değerlerin birbirleri ile ilişkisi yoktur.
- Bağımsız X değişkenleri arasında çoklu doğrusal bağlantı yoktur.
- X' in verilen herhangi bir gözlem değerine ait Y sonucunun olasılığı için verilen fonksiyon kullanılarak bulunmaktadır.
- Çok kategorili lojistik regresyon analizinde temel bir grup seçilerek diğer gruplar arasında karşılaştırma yapılabilir. İkili lojistik regresyon modelinden ayrıştığı en temel nokta bu karşılaştırma yöntemidir [50],[54].

Y bağımlı değişkeninin ikiden fazla değer aldığı durumlarda çok kategorili lojistik regresyon analizi kullanılmaktadır. Örneğin bağımlı değişken iyi, orta, kötü gibi üç farklı düzeyde değer aldığı durumlarda ikili (binary) lojistik regresyon analizi yerine çok kategorili lojistik regresyon analizi kullanılır [68].

Kategori sayısının 2'den fazla olması durumunda karşılaştırma ve çözümleme yapma amacı ile bir referans kategorisi belirlenmelidir. Referans kategorisinin seçimi isteğe bağlı olmakla birlikte araştırmacı tarafından seçilir [60].

Çıktı değişkeni vektörü $Y = (0, 1, \dots, c-1)$ ve bağımsız değişken vektörü $x = (x_1, x_2, \dots, x_p)$ iken her bir çıktı değerinin çok terimli olasılıkları bağımsız değişken x verildiği koşulu altında:

$$P(Y = 0 \setminus x) = \frac{1}{1 + e^{g_1(x)} + \dots + e^{g_{c-1}(x)}} \quad (3.6)$$

$$P(Y = 1 \setminus x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + \dots + e^{g_{c-1}(x)}} \quad (3.7)$$

$\vdots$

$$P(Y = c-1 \setminus x) = \frac{e^{g_{c-1}(x)}}{1 + e^{g_1(x)} + \dots + e^{g_{c-1}(x)}} \quad (3.8)$$

Eş. 3.6, Eş. 3.7 ve Eş. 3.8 elde edilmektedir. Burada Eş. 3.9'daki gibi logit dönüşüm uygulanması durumunda:

$$g_j(x) = \ln \left[\frac{P(Y=j \setminus x)}{P(Y=0 \setminus x)} \right], \quad i = 1, \dots, p, \quad j = 1, \dots, c-1 \quad (3.9)$$

$$= \beta_{j0} + \beta_{j1}x_1 + \dots + \beta_{jp}x_p \quad (3.10)$$

Eş. 3.10 elde edilmektedir [69].

Bağımsız değişkenlerin cinsiyet, meslek gibi farklı nominal ölçüm düzeyinde olması durumunda, nominal ölçümlerin seviyelerini gösteren ve sayısal olarak anlamlılığı olmayan sayılar verilir (1, 2, ...) ve dummy değişkenleri kullanılır [60].

n bağımsız gözlem için (x_i, y_i) , $i=1, 2, \dots, n$ örneğini ele alalım. Tek değişkenli analiz durumunda olduğu gibi, model çıkarımı için $\beta' = (\beta_0, \beta_1, \dots, \beta_p)$ gibi tahmin vektörü gerekmektedir. Modelin tahmini için, tek değişkenlide kullanıldığı gibi en çok olabilirlik yöntemi kullanılmaktadır. Olabilirlik eşitlik sistemi Eş. 3.11 ve Eş. 3.12'deki gibi verilebilir:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (3.11)$$

ve

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0 \quad j=1,2, \dots, p \quad (3.12)$$

Tahmin modelinde varyans ve kovaryansın tahmini katsayılarını bulmak için en çok olabilirlik teorisinden yararlanılır. Bu teoride, log olabilirlik fonksiyonunun ikinci kısmı türevinden elde edilen kestirimler yer alır. Bu türevler Eş. 3.13 ve Eş. 3.14'deki gibi ifade edilir:

$$\frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n x_{ij}^2 \pi_i (1 - \pi_i) \quad (3.13)$$

ve

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \pi_i (1 - \pi_i), \quad j, l = 0, 1, \dots, p \quad (3.14)$$

Burada, $\pi(x_i) = \pi_i$ dir. $(p+1) \times (p+1)$ matrisi, Eş. 3.11 ve Eş. 3.12'deki negatif terimleri içeren matrisi ifade etmektedir. Bu matris gözlenen bilgi matrisi olarak adlandırılmaktadır. Varyans ve kovaryansın kestirim değerleri bu matrisin tersinden elde edilmiş olan $Var(\beta) = I^{-1}(\beta)$ matrisi ile bulunur. Burada $Var(\beta_j)$, $\hat{\beta}_j$ nin varyansı olan j. köşegen elemanı olup $Cov(\beta_j, \beta_l)$, $\hat{\beta}_j$ ile $\hat{\beta}_l$ arasındaki kovaryansı ifade etmektedir. Kestirim katsayılarının standart hata tahmini için Eş. 3.15'ten faydalanılır:

$$\widehat{SE}(\hat{\beta}_j) = [\widehat{Var}(\hat{\beta}_j)]^{1/2} \quad j=0,1, \dots, p \quad (3.15)$$

Model için parametre kestirimi $\hat{\beta} = (X'VX)^{-1}X'VY$ matrisinden elde edilir. Burada X, $p+1$ satır, n sütunu olan bağımsız değişken matrisi ve V, $\hat{\pi}_i(1 - \hat{\pi}_i)$ elemanlarından oluşan matrisin gösterimi Eş. 3.16 ve Eş. 3.17'deki gibidir:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (3.16)$$

ve

$$V = \begin{bmatrix} \hat{\pi}_1(1 - \hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2(1 - \hat{\pi}_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \hat{\pi}_n(1 - \hat{\pi}_n) \end{bmatrix} \quad (3.17)$$

şeklinde gösterilir. [70].

Lojistik regresyon analizinde tahmin edilen katsayıların modele uygun olup olmadığının sınanmasında olabilirlik oran testi (likelihood ratio test) ve Wald testi yaygın olarak kullanılan yöntemlerdir. Katsayıların anlamlılığı için Eş. 3.18'deki hipotez çifti kurulur [71].

$$H_0 : \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad H_1 : \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (3.18)$$

Lojistik regresyon modelinde yer alan her bir bağımsız değişkenin modeldeki varlığının anlamlı olup olmadığı Wald testi ile belirlenir. Wald testi, t değerlerinin karelerinden oluşan, "1" serbestlik dereceli, asimptotik X^2 dağılımına sahip bir istatistiktir [71]. Wald test istatistiği Eş. 3.19'daki gibidir:

$$W_i = \left(\frac{\hat{\beta}_j}{s(\hat{\beta}_j)} \right)^2, \quad i=1,2, \dots, p \quad (3.19)$$

Lojistik regresyon modelindeki regresyon katsayılarının anlamlılığını test etmek için kullanılan Olabilirlik Oran Testleri, (bağımlı değişkenin kategori sayısı-1) (tahmin edilen parametre sayısı-1) serbestlik dereceli, X^2 dağılımına sahip bir istatistiktir [63]. Lojistik regresyonda analiz yaparken, R^2 istatistiği elde edilmemektedir, çünkü minimum varyans için model tahmini yapılmamaktadır. Bunun yerine en çok olabilirlik yöntemi kullanılarak model tahmini yapılır. Lojistik regresyonda, modelin uyum iyiliğini değerlendirmek amacıyla R^2 den geliştirilen Pseudo R^2 (sözde R^2) değeri kullanılır. Pseudo R^2 değerleri arasında en yaygın kullanılanları Cox and Snell's R^2 ve Nagelkerke's R^2 'dir [72].

Cox and Snell's R^2 değeri Eş. 3.20'deki gibidir.

$$\frac{1 - L_{final}}{L_{null}} \quad (3.20)$$

Burada, L_{final} değeri, sabit değer ve tahmin değerlerinin yer aldığı son model için log olabilirlik değeridir. L_{null} değeri ise, geçersiz model için log olabilirlik değerini ifade etmektedir.

Nagelkerke's R^2 , en büyük pozitif değer görüldüğü Cox and Snell's R^2 değerine bölünmesiyle türetilmekte olup 0 ile 1 arasında değer alır. Pseudo R^2 , model tarafından açıklanan bağımlı değişkendeki varyans miktarının tahminini göstermektedir [73].

Çok kategorili lojistik regresyonda, modelin yeterliliğinin test edilmesi için Pseudo R^2 istatistikleri de incelenebilir. Mcfadden R^2 , Cox-Snell ve Nagelkerke R^2 istatistikleri en sık kullanılan Pseudo R^2 istatistikleri olarak karşımıza çıkmaktadır. Bu katsayılar, doğrusal regresyondan farklı olarak küçük çıkar. Daha çok model oluşturma aşamasında farklı modellerin performansının karşılaştırılmasında kullanılmaları önerilir. Olabilirlik oran indeksi olarak da adlandırılan Mcfadden R^2 , çok kategorili lojistik regresyonda elde edilen R^2 ye göre oldukça küçük değerler alma eğiliminde olduğu için, 0.20-0.40 arasındaki bir değer çok yüksek olduğu söylenebilir. Cox-Snell R^2 en küçük 0 değerini alır fakat en büyük değeri 1 değildir. Nagelkerke R^2 değeri Cox-Snell R^2 istatistiğinin en büyük değerinin 1 olabilmesi için yapılan düzeltmeyle oluştuğu için bu değeri yorumlamak daha kolaydır [74].

4. UYGULAMA

Hane halklarına uygulanmak üzere geliştirilen, hane halkı fertlerinin demografik özellikleri ve sağlık durumları ile ilgili bilgileri elde etmeyi amaçlayan Türkiye Sağlık Araştırması (TSA) Türkiye tahmini verecek şekilde tasarlanmıştır. Soru kâğıdı örneklerle çıkan tüm hanelere uygulanmaktadır. 2008, 2010, 2012, 2014 ve 2016 yıllarında uygulanmış olan araştırma, 2008 yılının Nisan ayında, 2010 ve 2012 yıllarının Mayıs-Haziran döneminde bir ay uygulanmış iken 2014 ve 2016 yıllarının ise metodoloji gereği Ağustos-Ekim dönemindeki üç ayda uygulanmıştır. Soru formu için, Eurostat'ın (Avrupa İstatistik Ofisi) geliştirmiş olduğu metodoloji el kitabı temel kabul edilmekle birlikte, zamanla ülke ihtiyaçlarına cevap verebilecek bir takım ilave sorular da eklenmiştir. Türkiye İstatistik Kurumu (TÜİK) tarafından uygulanan ve yürütülen TSA' ya ait detaylı bilgiler ışığında genel sağlık durumlarının belirlenmesi amacıyla 2014 ve 2016 yıllarına ait veri setleri TÜİK'ten talep edilmiştir. 2014 ve 2016 yıllarına ait soru formlarında değişikliklerin olması sebebiyle 2014 ve 2016 yıllarına ait veriler birleştirilmiş olup 15 yaş ve üzeri bireyler çalışma kapsamına alınmıştır. Birleştirilen veriler üzerine yapay sinir ağları ve çok kategorili lojistik regresyon yöntemi uygulanmıştır. Uygulanan iki farklı yöntemin sonuçları karşılaştırılmış ve yorumlanmıştır. IBM SPSS Modeler paket programı kullanılmıştır.

4.1. Kullanılan Veri Setine İlişkin Genel Bilgiler

İlk defa 2008 yılında gerçekleştirilmiş olan TSA iki yılda bir yapılmaktadır. Bu araştırma ile bebek, çocuk ve yetişkinlere sorulan sağlık durumları sorularının yanı sıra 15 yaş ve üzeri bireylerin sağlık hizmetinden faydalanma durumları ve memnuniyet düzeyleri, günlük aktivitelerini gerçekleştirirken yaşadıkları zorlukların derecesi ile sigara ve alkol kullanma alışkanlıkları başta olmak üzere sağlık alanındaki birçok gösterge elde edilmektedir.

Araştırma ile ülkelerin gelişmişlik düzeylerini gösteren kalkınma göstergeleri içerisinde önemli bir paya sahip olan sağlık göstergelerine yönelik bilgilerin elde edilmesi amaçlanmaktadır. Araştırma, ülke genelini yansıtan bir çalışma olmasının yanı sıra hem uluslararası karşılaştırmalara imkân veren hem de ulusal ihtiyaçlara ışık tutan bir çalışma olması açısından da önemlidir.

Türkiye Cumhuriyeti sınırları dâhilinde yer alan tüm yerleşim yerleri kapsama dahil edilmiştir. TSA'nın coğrafi çerçevesi Türkiye sınırları içerisindeki tüm yerleşim yerlerinde bulunan hane halklarıdır. Kurumsal olarak nitelendirilen nüfus (asker, yurtlarda, hapisanede, hastanede sürekli olarak kalan, huzurevleri vb.) kapsam dışı olup, ayrıca yeterli örnek hane sayısına ulaşamayacağı düşünülen adres sayısı 20'den küçük (küçük köyler, oba, mezra vb.) yerleşim yerleri kapsam dışı bırakılmıştır. Araştırmanın örneklem çalışmalarında kullanılan çerçeve 2007 yılında kurulan Adrese Dayalı Nüfus Kayıt Sistemine altlık oluşturan Ulusal Adres Veri Tabanıdır.

TSA soru formu; kapak sayfası, hane halkı soru formu, 0-6 yaş grubu soru formu, 7-14 yaş grubu soru formu, 15+ yaş grubu soru formu olmak üzere beş ana kısımdan oluşmaktadır.

Bölümler ise; hane halkı bireylerinin temel özellikleri, 0-6 yaş grubu, 7-14 yaş grubu, 15+ yaş grubu, sağlık durumu, kronik hastalıklar, kazalar ve yaralanmalar, işten uzak kalma, fiziksel ve duyuşsal kısıtlamalar, kişisel bakım faaliyetleri, hane halkı faaliyetleri, ağrı, zihinsel sağlık, yatan hasta tedavisi/ günöbirlik bakım, ayakta tedavi/ evde bakım, ilaç kullanımı, koruyucu hizmetler, sağlık bakımından karşılanamayan ihtiyaçlar, boy ve kilo, fiziksel aktivite/ egzersiz, meyve ve sebze tüketimi, tütün mamulü kullanımı, alkol kullanımı, sosyal destek, resmi olmayan bakım veya yardım sağlanması olarak ifade edilmektedir.

TSA ile bireylerin genel sağlık profilinin ortaya çıkarılması ve ölkelerin gelişmişlik düzeylerini gösteren kalkınma göstergeleri içerisinde önemli bir paya sahip olan sağlık göstergelerine yönelik bilgilerin elde edilmesi amaçlanmaktadır. Bu çalışma, hem uluslararası karşılaştırmalara imkân veren hem de ulusal ihtiyaçlara ışık tutan bir araştırma olma özelliğini taşımaktadır.

2014 yılı TSA'nın örnek hacmi 9 740 hane olarak belirlenmiştir. Örneğe seçilen 9 740 hanenin 8 634'ü ile anket sonuçlandırılmıştır. Toplam 26 075 bireye anket uygulanmıştır [75].

2016 yılı TSA'nın örnek hacmi yine 9 740 hane olarak belirlenmiştir. Örneğe seçilen 9 740 hanenin 8 325'i ile anket sonuçlandırılmıştır. Toplam 23 606 bireye anket uygulanmıştır [76].

4.2. Veri Setinin Düzenlenmesi

Hane halkı soru formunda fertlerin; cinsiyet, yaş, medeni durum, çalışma durumu, gelir ve öğrenim durumları ile ilgili sorular sorulmuştur. 0-6 ve 7-14 yaş gruplarında yer alan fertlere; başlarından geçen kazalar, hastalıklar ve sağlık hizmetlerinden faydalanma durumu ile ilgili sorular sorulmuştur. 15+ yaş grubundaki fertlerden ise genel sağlık durumu, sağlık problemlerine dayalı günlük yaşamsal faaliyetlerini gerçekleştirmedeki kısıtlamalar, günlük bedensel aktiviteler, tütün kullanımı ve son bir yıl içinde almış oldukları sağlık hizmetleri konularında bilgiler alınmıştır.

Bağımlı değişken olan “genel sağlık durumunuz nasıldır?” sorusu kategorik bir değişken olup beş seçenekten oluşmaktadır. Bu beş seçenek analiz aşamasında düzenlenerek Çizelge 4.1’deki üç kategorili hale getirilmiştir.

Çizelge 4.1. Genel sağlık durumu değişkeni

Genel Sağlık Durumu	GENELSAGLIKDURUM	GSD
Çok iyi	1	3
İyi	2	3
Orta	3	2
Kötü	4	1
Çok kötü	5	1

TSA verilerinin kullanılmasındaki amaç, genel sağlık durumunun detaylı olarak ele alınması, yapay sinir ağları ve çok kategorili lojistik regresyon yöntemleri ile modelin tanımadığı genel sağlık durumu bilinmeyen bir bireyin model sonucunda genel sağlık durumunun tahmin edilmesidir. Elde edilen modeller sonucunda genel sağlık durumu değişkeni için doğruluk oranlarına göre yapay sinir ağları ve çok kategorili lojistik regresyon modelleri karşılaştırılmıştır.

Veri setinin düzenlenmesi aşamasında yapılan işlemler sırasıyla aşağıda yer almaktadır:

- 26 075 fert ile yapılan 2014 yılı TSA verileri veri seti içerisinde yer alan değişken sayısı 471 iken bu değişkenlerin 245 tanesini 0-6 ve 7-14 yaş grubunda yer alan fertlere sorulan sorular oluşturmaktadır.

- Anketin uygulanması esnasında bağımlı değişkene ait “Genel olarak sağlık durumunuz nasıldır?” sorusu, 15 yaşından büyük fertlere sorulmuştur. 0-6 ve 7-14 yaş grubunda yer alan fertlerin bağımlı değişkene ait hücreleri boştur. Bu nedenle 0-6 ve 7-14 yaş grubunda yer alan 6 946 fert, veri setinden çıkartılmıştır.
- 2014 yılı TSA verileri ile 2016 yılı TSA verileri birleştirileceği için kullanılan sınıflama ve veri setinde yer alan değişken isimlerinin aynı olması beklenmektedir. Bu durumda 2014 ve 2016 yılı TSA verileri değişken isimleri ve seçenekler bakımından aynı olacak şekilde düzenlenmiştir.
- 2014 yılı veri setinde yer alan “Tedavi masraflarınız hangisi tarafından karşılanıyor?” sorusunun seçenekleri arasında yeşil kart mevcut iken 2016 yılına ait bu sorunun seçeneklerinde yeşil kart olmadığı için bu değişkene ait bilgiler genel sağlık sigortasına (GSS) aktarılmıştır.
- 2016 yılı veri setinde yer alıp 2014 yılı veri setinde yer almayan “Son 12 ay içerisinde aşağıdaki hastalıklardan ya da durumlardan birini yaşadınız mı?” sorusuna ait seçenekler arasında diğer seçeneği yer almadığı için bu seçeneğe ait bilgiler 2016 yılı veri setinden çıkarılmıştır. (1 376 anket)
- “İlk kez kaç yaşında tütün mamulü kullanmayı denediniz?” sorusu 2016 yılı veri setinde olmadığı için 2014 yılı veri setinden çıkarılmıştır.
- 2014 yılı TSA kapsamında “Tütün mamulü kullanıyor musunuz?” sorusuna yanıt olarak “sadece denedim” diyenler “ bıraktım” seçeneği altında yer alırken, 2016 yılında aynı şekilde cevap verenler “hiçbir zaman” seçeneği altında yer almaktadır.
- 2014 yılı veri setinde “Alkol kullanmanızın ya da kullanmayı denemenizin nedeni nedir?” sorusunun seçenekleri 2016 yılı veri setine göre daha detaylı olduğu için seçeneklere ilişkin bilgiler düzenlenerek seçenek sayısı ve içerik aynı hale getirilmiştir.
- 2014 yılı veri setinde hafta içi alkol kullanım detayları yer alırken 2016 yılı veri setinde bu soru yer almadığı için 2014 yılı veri setinden çıkarılmıştır.
- 2014 yılı veri setinde yer alıp 2016 yılı veri setinde yer almayan “Alkol kullanımının zarar verdiğini düşünüyor musunuz?” sorusu veri setinden çıkarılmıştır.
- “Ne kadar süredir tütün mamulü kullanmıyorsunuz?” sorusuna ait eksik gözlem sayısı her iki yılın verisinde de fazla olduğu için iki yılın veri setinden de çıkarılmıştır.
- 2014 yılında sorulan “Tütün mamulü kullanmaya başlama nedeniniz nedir?” sorusuna ait şıklar arasından birden fazla seçenek işaretlenebildiği ancak 2016 yılında sorulan aynı soruya yalnızca tek seçenek işaretlenebildiği için bu soruya yalnız bir cevap verenlerin bilgisi veri setinde tutulup geri kalan bilgiler veri setinden çıkarılmıştır.

- 2014 yılı TSA da “Normal bir hafta içerisinde spor, fitness veya boş zaman aktivitesi gerçekleştirirken toplamda ne kadar zaman harcarsınız?” sorusu dakika cinsinden alınmakta iken 2016 yılı TSA’da sorulan aynı sorunun cevabı saat cinsinden veri setinde yer almaktadır. 2016 yılı veri setinde yer alan saat cinsinden ifadeler 60 ile çarpılarak dakikaya çevrilmiş ve böylece 2014 yılında yer alan değişkenle aynı hale getirilmiştir.
- 2014 yılında sorulan “Alkol kullanmanızın ya da kullanmayı denemenizin nedeni nedir?” sorusuna ait şıklar arasından birden fazla seçenek işaretlenebildiği ancak 2016 yılında sorulan aynı soruya yalnızca tek seçenek işaretlenebildiği için bu soruya yalnız bir cevap verenlerin bilgisi veri setinde tutulup geri kalan bilgiler veri setinden çıkarılmıştır.

Böylelikle iki yıla ait veriler aynı hale getirilmiştir. Veriler aynı hale getirildikten sonra verilerin programda uygun olarak yorumlanabilmesi için bir takım düzenlemeler yapılmıştır.

2014 ve 2016 yıllarına ait TSA verilerinde yaş grubu değişkeni 0-6 yaş arası, 7-14 yaş arası, 15-24 yaş arası, 25-34 yaş arası, 35-44 yaş arası, 45-54 yaş arası, 55-64 yaş arası, 65-74 yaş arası ve 75 yaş ve üzeri olmak üzere yaş grubu değişkeni dokuz kategoriden oluşurken 0-6 yaş arası ile 7-14 yaş arasına ait bilgiler veri setinden çıkarılınca geriye kalan veriler yedi kategori şeklinde yer almıştır. Yaş grubu değişkeni revize edilerek Çizelge 4.2’de özetlenmiştir.

Çizelge 4.2. Yaş grubu değişkeni

YAS_GRUP	YAS_SIRA
15-24	1
25-34	2
35-44	3
45-54	4
55-64	5
65-74	6
75 ve üzeri	7

2014 ve 2016 yılı veri setinde yer alan “En son hangi okulu bitirdiniz?” sorusuna ait cevap seçeneklerinin kodları ordinal olacak şekilde yeniden düzenlenmiştir. Eğitim durumu arttıkça o seçeneğe karşılık gelen kod da artacak şekilde tekrar kodlanarak Çizelge 4.3’de gösterilmiştir.

Çizelge 4.3. Eğitim durumu değişkeni

EGITIM_ACIKLAMA	EGITIM	EGITIM_SIRALI
Okuma-Yazma bilmeyen	8	0
Bir okul bitirmedi	0	1
İlkokul	1	2
Genel ortaokul	21	3
Mesleki veya teknik ortaokul	22	3
İlköğretim	23	3
Lise veya mesleki lise	3	4
2 veya 3 yıllık yüksekokul	4	5
4 yıllık yüksekokul veya fakülte	5	6
Yüksek lisans (5 veya 6 yıllık fakülteler dahil)	6	7
Doktora	7	8

2014 ve 2016 yılına ait “Hanenizin ortalama aylık geliri aşağıdaki gelir gruplarından hangisinde yer almaktadır?” sorusuna ait seçenekler ifade edilen miktarlar, yıl değiştiği için farklı olmakla birlikte aynı grubu temsil etmektedir. Bu durum Çizelge 4.4’de özetlenmiştir.

Çizelge 4.4. 2014 ve 2016 yıllarına ait hanelerin gelir bilgisi seçenekleri

2014	2016	HANE GELİR
0 – 1 080 TL	0 – 1 264 TL	1
1 081 – 1 550 TL	1 265 – 1 814 TL	2
1 551 – 2 170 TL	1 815 – 2 540 TL	3
2 171 – 3 180 TL	2 541 – 3 721 TL	4
3 181 + TL	3 722 + TL	5

2014 ve 2016 yıllarına ait veri setinde yer alan Uluslararası Standart Meslek Sınıflaması-ISCO 08 ve Avrupa Topluluğunda Ekonomik Faaliyetlerin İstatistiksel Sınıflaması-NACE Rev.2 kodları 4 haneli olacak şekilde veri setinde yer almakta iken bu alt kodlara karşılık gelen üst kodlar veri setine yansıtılmıştır [77].

Uluslararası Çalışma Örgütü (ILO)’nün önemli bir yayını olan ISCO-88’in güncellenmiş hali olan ISCO-08, hem ülkelerin kendi ulusal sınıflamalarının geliştirilmesinde hem de meslek bilgilerinin uluslararası karşılaştırmasını mümkün kılma yönünde kullanışlı bir modeldir. ISCO-88, bir işi yapmak için gerekli olan beceri, uzmanlık ve seviyeyi dikkate almaktaydı, ISCO-08 de aynı temele dayanmaktadır. Yine aynı şekilde, ISCO-08, ISCO-88 gibi, yapılan iş türüne göre dört tabakalı hiyerarşik bir yapıya sahiptir. Ana grup sayısı

değişmemesine rağmen, alt ana grup, grup ve birim grup sayılarında önemli bir artış olmuştur.

Çizelge 4.5. Uluslararası standart meslek sınıflaması

Kod	Tanım
1	Silahlı kuvvetlerle ilgili meslekler
2	Yöneticiler
3	Profesyonel meslek mensupları
4	Teknisyenler, teknikerler ve yardımcı profesyonel meslek mensupları
5	Büro hizmetlerinde çalışan elemanlar
6	Hizmet ve satış elemanları
7	Nitelikli tarım, ormancılık ve su ürünleri çalışanları
8	Sanatkârlar ve ilgili işlerde çalışanlar
9	Tesis ve makine operatörleri ve montajcılar
10	Nitelik gerektirmeyen işlerde çalışanlar

2014 ve 2016 yıllarına ait veri seti içerisinde yer alan birimlere karşılık gelen kodlar üst kodlara dönüştürülmüştür.

Çizelge 4.6. Avrupa topluluğunda ekonomik faaliyetlerin istatistiki sınıflaması

Kod	Tanım
A	Tarım, ormancılık ve balıkçılık
B	Madencilik ve taş ocaklığı
C	İmalat
D	Elektrik, Gaz, Buhar ve İklimlendirme Üretimi ve Dağıtımı
E	Su Temini; Kanalizasyon, Atık Yönetimi ve İyileştirme Faaliyetleri
F	İnşaat
G	Toptan ve Perakende Ticaret; Motorlu Kara Taşıtlarının ve Motosikletlerin Onarımı
H	Ulaştırma ve Depolama
I	Konaklama ve Yiyecek Hizmeti Faaliyetleri
J	Bilgi ve İletişim
K	Finans ve Sigorta Faaliyetleri
L	Gayrimenkul Faaliyetleri
M	Mesleki, Bilimsel ve Teknik Faaliyetler
N	İdari ve Destek Hizmet Faaliyetleri
O	Kamu Yönetimi ve Savunma; Zorunlu Sosyal Güvenlik
P	Eğitim
Q	İnsan Sağlığı ve Sosyal Hizmet Faaliyetleri
R	Kültür, Sanat, Eğlence, Dinlence ve Spor
S	Diğer Hizmet Faaliyetleri
T	Hanehalklarının İşverenler Olarak Faaliyetleri; Hanehalkları Tarafından Kendi Kullanımlarına Yönelik Olarak Ayrım Yapılmamış Mal ve Hizmet Üretim Faaliyetleri
U	Uluslararası Örgütler ve Temsilciliklerinin Faaliyetleri

NACE Rev. 2, ISIC Rev. 4'e karşılık gelen Avrupa seviyesinde ekonomik faaliyetler sınıflamasıdır. NACE Rev. 2, ISIC Rev. 4'ten daha çok parçalara ayrılmış olmasına rağmen, ISIC Rev. 4 ile tamamen uyumludur ve bu yüzden ISIC Rev. 4'ün Avrupa versiyonu olarak görülebilir.

Ulusal ekonomik yapılar ciddi şekilde değiştikleri için, NACE Rev. 2'de olan bazı ekonomik faaliyetler bütün Üye Devletler için önemli olmayabilir veya görülmeyebilir. Örneğin, bir ülkede uzay aracı imalatı yapılırken başka bir ülkede yapılmayabilir. NACE Rev. 2'nin Yönetmeliği Üye Devletlerin NACE Rev. 2'den türetilmiş olan ulusal bir versiyonunu ulusal amaçlar için kullanmalarına izin verir. Fakat bu ulusal versiyonlar, NACE Rev. 2'de belirtilen yapısal ve hiyerarşik çerçeveye uymak zorundadırlar [78].

2014 ve 2016 yılı TSA veri seti yukarıdaki üst kodlarda yer alan harflere karşılık rakamlar gelecek şekilde düzenlenmiştir.

Bu işlemlerin sonucunda analiz öncesi düzenlenen veri setinde 2014 yılı için 19 129 satır ve 226 sütun, 2016 yılı için ise 15 866 satır ve 226 sütun yer almaktadır.

4.3. Veri indirgeme

Veri setinde yer alan tüm değişkenler modele dâhil edildiğinde anlamlı sonuçlar ortaya çıkmamaktadır. Veri setindeki bağımsız değişkenlerden bazıları benzer sorular olması sebebi ile veri setinden çıkarılmıştır.

IBM SPSS Modeler paket programı kullanılarak bağımsız değişkenler arasında çoklu bağlantı (multi-collinearity) olup olmaması durumu incelenmiştir.

Çoklu bağlantının tespit edilmesi amacıyla doğrusal regresyon analizi uygulanarak doğrusallık (çoklu ilişki) olup olmadığı tespit edilmeye çalışılmıştır. Bu analizin amacı, değişkenlerin birbirleri ile arasında yüksek korelasyona sahip olup olmadığının tespit edilmesidir. Değişkenler arasında yüksek korelasyon olması durumunda tahmin gücü azalmaktadır. Tahmin gücünün azalmaması için bu değişkenler belirlenip analiz dışında bırakılmalıdır. Analiz dışında bırakılmayıp modele dahil edildiklerinde katsayılar tanımsız olabilecek ve bu katsayıların standart hatası da sonsuz olarak karşımıza çıkabilecektir.

Katsayıların varyans ve kovaryansları da artma eğiliminde olacaktır. Bu durumun engellenmesi için analiz yapıp verilerin incelenmesi önem taşımaktadır [79].

Analiz sonucu aralarında çoklu bağlantı olan değişkenlerden daha az detayda bilgi barındıranlar veri setinden çıkarılmıştır. Bu sorunun en önemli göstergelerinden olan varyans büyütme faktörü (Variance Inflation Factor-VIF) kullanılmıştır. Uygulamalarda farklı eşikler kullanılmasına rağmen VIF değerinin 5'ten küçük olması durumunda, bağımsız değişkenler arasında çoklu bağlantının olmadığı sonucuna varılabilir. Bağımsız değişkenlerin VIF değerlerine bakılıp uygun görülen değişkenler veri setinden çıkarılmıştır (EK-1).

Analize girecek bağımsız değişkenlerin son hallerinin belirlenmesi için IBM SPSS Modeler paket programında yer alan feature selection tekniğinden faydalanılmıştır.

Feature selection: Veri setinde bulunan değişkenlerden hangilerinin bağımlı değişkenin açıklanmasında önemli olduğunu ortaya koyan işlemcidir. Bu sayede analiz işleminde gereksiz olan değişkenler elenir. Bu işlemcinin kullanılması sonucunda yalnızca önemli değişkenlerin modelleme aşamasında yer alması sağlanır. Ayrıca bu işlem sonucunda modelin tahmin doğruluğu artmaktadır. Feature selection tekniğinde bağımsız değişkenler 0 ile 1 arasında önem değerleri almaktadır.

- Önem değeri $> 0,95$ ise değişken, bağımlı değişken üzerinde önemli etkiye sahiptir,
- $0,95 \geq$ önem değeri $> 0,9$ ise değişken, bağımlı değişken üzerinde az önemli etkiye sahiptir,
- Önem değeri $< 0,9$ ise değişken, bağımlı değişken üzerinde önemsiz etkiye sahiptir şeklinde yorumlanmaktadır.

Çizelge 4.7. Kullanılan bağımsız değişkenler

	Değişken Adı	Değişken Türü	Durum	Önem Değeri
1	ALKOL_KULLANMANEDEN_1	Nominal	Önemli	1,0
2	ALKOLKUL_DUZENLI_YAS	Continuous	Önemli	1,0
3	BEDENSELAGRI	Nominal	Önemli	1,0
4	BUNALIM	Nominal	Önemli	1,0
5	CALISMAMA_NEDEN	Nominal	Önemli	1,0
6	CINSIYET	Flag	Önemli	1,0
7	DEGERSIZHISSETME	Nominal	Önemli	1,0

Çizelge 4.7. (devam) Kullanılan bağımsız değişkenler

8	DESTEKSİZ12BASAMAK	Nominal	Önemli	1,0
9	DIS_HEK_HIZALMAZAMAN	Nominal	Önemli	1,0
10	EGITIM_SIRALI	Ordinal	Önemli	1,0
11	GOZLUKKULLANMA	Nominal	Önemli	1,0
12	GRIPASISIOLMA_ZAMAN	Nominal	Önemli	1,0
13	GUNLUK_FAALİYETDURUM	Nominal	Önemli	1,0
14	HANE_GELİR	Ordinal	Önemli	1,0
15	HASTALIK_SAGLIKDURUM	Flag	Önemli	1,0
16	HATIRLAMA	Nominal	Önemli	1,0
17	HUZURSUZLUK	Nominal	Önemli	1,0
18	ISCO_GRUP	Nominal	Önemli	1,0
19	ISTAHSIZLIK	Nominal	Önemli	1,0
20	ISTEKİ_DURUM	Nominal	Önemli	1,0
21	KENDI_ILAC	Flag	Önemli	1,0
22	KILO	Continuous	Önemli	1,0
23	KONSANTRASYONPRB	Nominal	Önemli	1,0
24	KRONIKHAST_ALERJİ	Flag	Önemli	1,0
25	KRONIKHAST_BELBOLGE	Flag	Önemli	1,0
26	KRONIKHAST_BOYUNBOLGE	Flag	Önemli	1,0
27	KRONIKHAST_HIPERTANSİYON	Flag	Önemli	1,0
28	KRONIKHAST_SEKERHAST	Flag	Önemli	1,0
29	MEDENİ_DURUM	Nominal	Önemli	1,0
30	MEYVE_YEMESİKLİK	Nominal	Önemli	1,0
31	NACE_GRUP	Nominal	Önemli	1,0
32	ODEMEGUCLUGU_DISBAKIM	Nominal	Önemli	1,0
33	ODEMEGUCLUGU_ILAC	Nominal	Önemli	1,0
34	ODEMEGUCLUGU_RUHTEDV	Nominal	Önemli	1,0
35	ODEMEGUCLUGU_TIBBİBAKIM	Nominal	Önemli	1,0
36	OGRENME	Nominal	Önemli	1,0
37	RECETEEDİLEN_ILACKULLANIM	Flag	Önemli	1,0
38	S_BIRAKMAYONTEM	Nominal	Önemli	1,0
39	SAGKURUM_TERCİH	Nominal	Önemli	1,0
40	SEBZE_SALATA_YEMESİKLİK	Nominal	Önemli	1,0
41	SIGARA_ADET	Continuous	Önemli	1,0
42	SON12AY_6VEUZERİ	Nominal	Önemli	1,0
43	SPOR_ZAMAN	Continuous	Önemli	1,0
44	TEDAVI_MASRAFI_GSS	Flag	Önemli	1,0
45	TEDAVI_MASRAFI_KENDİ	Flag	Önemli	0,991
46	TEDAVI_MASRAFI_SGK	Flag	Önemli	1,0
47	TUTUN_BASLAMANEDEN_1	Nominal	Önemli	1,0
48	TUTUNDUMAN_MARUZ	Nominal	Önemli	1,0
49	TUTUNKUL_DUZENLI_DURUM	Flag	Önemli	1,0
50	TUTUNKUL_DUZENLI_YAS	Continuous	Önemli	1,0
51	TUTUNKULLANIM_DURUM	Nominal	Önemli	1,0
52	UYKUSORUN	Nominal	Önemli	1,0
53	UZMAN_HEK_HIZALMADURUM	Nominal	Önemli	1,0
54	YAS_SIRA	Ordinal	Önemli	1,0

Çizelge 4.7. (devam) Kullanılan bağımsız değişkenler

55	YASAMFAAL_KISITLANMADURUM	Nominal	Önemli	1,0
56	YORGUNLUK	Nominal	Önemli	1,0
57	YURUME_ZAMAN	Nominal	Önemli	1,0
58	ZORLUKCEKME_AGIREVIS	Nominal	Önemli	1,0
59	ZORLUKCEKME_HAFIFEVIS	Nominal	Önemli	1,0
60	ZORLUKCEKME_IDARIMALIKONU	Nominal	Önemli	1,0

Feature selection işlemcisi sonrası modele giren değişkenler ve değişkenlerin özelliklerine göre ölçüm değerleri Çizelge 4.7’de gösterilmiştir. Bu değişkenler analiz için önemli olan değişkenleri ifade eder. Ele alınan değişkenlere ait frekans tabloları Ek-2’de yer almaktadır.

4.4. Tanımlayıcı İstatistikler

Veri setinde yer alan bağımlı ve bağımsız değişkenlere ait tanımlayıcı istatistikler elde edilmiştir. Bu istatistiklere ait tablolar Çizelge 4.8 - 4.16’da özetlenmiştir. Seçilmiş bağımsız değişkenlere göre kişilerin genel sağlık durumları;

Çizelge 4.8. Yaş değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAGLIK DURUMU		
YAS_SIRA	Kötü	Orta	İyi
15-24	1,9	10,1	88,0
25-34	3,4	18,5	78,1
35-44	6,7	28,3	65,0
45-54	12,7	37,5	49,8
55-64	19,1	40,7	40,2
65-74	29,7	42,7	27,6
75 ve üzeri	43,6	39,4	17,0

Çizelge 4.8’de bireylerin yaş gruplarına göre genel sağlık durumları incelendiğinde; 15-24 yaş grubunda yer alan bireylerin %88’inin, 25-34 yaş grubunda yer alan bireylerin %78’inin, 35-44 yaş grubunda yer alan bireylerin %65’inin, 45-54 yaş grubunda yer alan bireylerin yaklaşık %50’sinin, 55-64 yaş grubunda yer alan bireylerin %40’ının, 65-74 yaş grubunda yer alan bireylerin yaklaşık %28’inin, 75 ve üzeri yaş grubunda yer alan bireylerin ise %17’sinin genel sağlık durumunu “iyi” olarak nitelendirdiği görülmektedir. Kişilerin yaş grubu arttıkça genel sağlık durumlarını iyi olarak nitelendirme oranlarının azaldığı, kötü olarak nitelendirme oranlarının ise arttığı görülmektedir.

Çizelge 4.9. Cinsiyet değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
CINSİYET	Kötü	Orta	İyi
Erkek	8,1	24,5	67,4
Kadın	14,8	31,5	53,7

Çizelge 4.9’da bireylerin cinsiyetlerine göre genel sağlık durumları incelendiğinde; erkeklerin %67’sinin genel sağlık durumunu iyi olarak nitelendirdiği, %25’inin genel sağlık durumunu orta olarak nitelendirdiği, %8’inin ise genel sağlık durumunu kötü olarak nitelendirdiği görülmektedir. Kadınların %54’ünün genel sağlık durumunu iyi olarak nitelendirdiği, %32’sinin genel sağlık durumunu orta olarak nitelendirdiği, %15’inin ise genel sağlık durumunu kötü olarak nitelendirdiği görülmektedir. Erkeklerin genel sağlık durumlarını iyi olarak nitelendirme oranlarının kadınlardan fazla olduğu, kadınların genel sağlık durumlarını kötü olarak nitelendirme oranlarının ise erkeklerden fazla olduğu görülmektedir.

Çizelge 4.10. Eğitim durumu değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
EGITIM_SIRALI	Kötü	Orta	İyi
Okuma-Yazma bilmeyen	40,8	35,9	23,3
Bir okul bitirmedi	24,4	37,7	37,9
İlkokul	13,4	37,7	48,9
Genel ortaokul, mesleki veya teknik ortaokul, ilköğretim	4,2	19,7	76,1
Lise veya mesleki lise	3,7	20,6	75,7
2 veya 3 yıllık yüksekokul	3,4	17,5	79,1
4 yıllık yüksekokul veya fakülte	1,5	15,4	83,1
Yüksek lisans (5 veya 6 yıllık fakülteler dahil)	2,6	12,0	85,4
Doktora	0,0	20,9	79,1

Çizelge 4.10’da bireylerin eğitim durumlarına göre genel sağlık durumları incelendiğinde; okuma-yazma bilmeyen bireylerin yaklaşık %41’inin genel sağlık durumunu kötü, %36’sının genel sağlık durumunu orta, %23’ünün ise genel sağlık durumunu iyi olarak nitelendirdiği görülmektedir. Genel ortaokul, mesleki veya teknik ortaokul, ilköğretim ve daha ileri düzeydeki eğitim durumunda yer alan bireylerin çoğunun genel sağlık durumunu iyi olarak nitelendirdiği görülmektedir.

Çizelge 4.11. Medeni durum değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
MEDENI_DURUM	Kötü	Orta	İyi
Hiç evlenmedi	3,0	11,7	85,3
Evli	12,0	32,1	55,9
Boşandı	12,8	31,3	55,9
Eşi öldü	37,1	40,9	22,0

Çizelge 4.11’de bireylerin medeni durumlarına göre genel sağlık durumları incelendiğinde; hiç evlenmeyen bireylerin %85’i genel sağlık durumunu iyi olarak nitelendirirken, evli ve boşanmış olan bireylerin %56’sı genel sağlık durumunu iyi olarak nitelendirmekte olduğu görülmektedir. Eşi ölen bireylerin ise %41’i genel sağlık durumunu orta, %37’si genel sağlık durumunu kötü, %22’si ise genel sağlık durumunu iyi olarak nitelendirmektedir.

Çizelge 4.12. Hanenin gelir durumu değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
HANE_GELİR	Kötü	Orta	İyi
1	21,2	32,0	46,8
2	11,7	30,5	57,8
3	9,0	28,5	62,5
4	6,8	25,4	67,8
5	4,3	22,0	73,7

Çizelge 4.12’de bireylerin hanelerine ait gelir durumlarına göre genel sağlık durumları incelendiğinde; hanelerin gelir durumları arttığında bireylerin genel sağlık durumlarını iyi olarak nitelendirme oranlarının arttığı, hanelerin gelir durumları azaldığında ise bireylerin genel sağlık durumlarını kötü olarak nitelendirme oranlarının arttığı görülmektedir.

Çizelge 4.13. Uluslararası standart meslek sınıflaması kodlarına göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
ISCO_GRUP	Kötü	Orta	İyi
-	15,0	28,4	56,6
1	7,8	13,6	78,6
2	5,4	24,9	69,7
3	3,3	18,6	78,1
4	3,8	23,2	73,0
5	5,0	25,3	69,7
6	7,1	26,6	66,3
7	21,9	36,4	41,7
8	8,7	27,7	63,6
9	7,3	29,5	63,2
10	16,6	34,3	49,1

Çizelge 4.13’de bireylerin Uluslararası Standart Meslek Sınıflaması (ISCO-08) kodlarına göre genel sağlık durumu incelendiğinde; yöneticiler grubunda yer alıp 1 kodunu alan silahlı kuvvetlerle ilgili meslekler grubunda yer alan bireylerin ve 3 kodunu alan profesyonel meslek mensupları grubunda yer alan bireylerin genel sağlık durumlarını iyi olarak nitelendirme oranının %78 civarında olduğu görülmektedir.

Nitelikli tarım, ormancılık ve su ürünleri çalışanlarının ve nitelik gerektirmeyen işlerde çalışanların genel sağlık durumlarını kötü olarak nitelendirme oranları sırasıyla %22 ve yaklaşık %17’dir.

Çizelge 4.14. Hastalık sağlık durumu değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
HASTALIK_SAGLIK DURUMU	Kötü	Orta	İyi
Evet	23,5	46,1	30,4
Hayır	0,8	11,6	87,6

Çizelge 4.14’de bireylerin ‘6 ay ya da uzun süren/sürmesi beklenen hastalığınız/sağlık probleminiz var mı?’ sorusuna verdikleri cevaplara göre genel sağlık durumları yer almaktadır. Bu soruya ‘Hayır’ diyenlerin yaklaşık %88’i genel sağlık durumunu iyi olarak nitelendirirken, ‘Evet’ diyenlerin ise yalnızca %30’unun genel sağlık durumunu iyi olarak nitelendirdiği görülmektedir.

Çizelge 4.15. Yaşamsal faaliyetin kısıtlanma durumu değişkeni kategorilerine göre genel sağlık durumu (%)

	GENEL SAĞLIK DURUMU		
YASAMFAAL_KISITLANMADURUM	Kötü	Orta	İyi
Ciddi ölçüde kısıtlandı	55,2	33,5	11,3
Kısıtlandı fakat ciddi ölçüde değil	15,2	51,9	32,9
Kısıtlanmadı	1,2	16,5	82,3

Çizelge 4.15’de bireylerin ‘son 6 ay içinde herhangi bir sağlık problemi nedeniyle günlük yaşam faaliyetlerinizin kısıtlanma durumu nedir?’ sorusuna verdikleri cevaplara göre genel sağlık durumları yer almaktadır. Bu soruya ‘kısıtlanmadı’ diyenlerin %82’si genel sağlık durumunu iyi olarak belirtirken, ‘kısıtlandı fakat ciddi ölçüde değil’ diyenlerin %33’ü genel sağlık durumunu iyi olarak belirtmiştir. ‘Ciddi ölçüde kısıtlandı’ diyenlerin ise yalnızca

%11'i genel sağlık durumunu iyi olarak nitelendirmiştir. Yaşamsal faaliyetin kısıtlanma durumuna verilen cevaplar ile genel sağlık durumuna verilen cevapların birbirini destekleyici olduğu görülmektedir.

Çizelge 4.16. Günlük faaliyet durum değişkeninin kategorilerine göre genel sağlık durumu (%)

GUNLUK_FAALİYETDURUM	GENEL SAĞLIK DURUMU		
	Kötü	Orta	İyi
Çoğunlukla oturan veya duran	19,7	29,7	50,6
Çoğunlukla yürüyüş veya orta derecede fiziksel güç gerektiren işler	6,5	27,5	66,0
Çoğunlukla ağır iş veya fiziksel güç gerektiren işler	4,7	26,2	69,1

Çizelge 4.16'da bireylerin 'çalışıyorken/günlük faaliyetlerinizi gerçekleştirirken, aşağıdakilerden hangisi durumunuzu en iyi tanımlar?' sorusuna verdikleri cevaplara göre genel sağlık durumları yer almaktadır. Bu soruya 'çoğunlukla oturan veya duran' cevabını verenlerin %50' si genel sağlık durumunu iyi olarak nitelendirirken, 'çoğunlukla yürüyüş veya orta derecede fiziksel güç gerektiren işler' cevabını verenlerin %66'sının genel sağlık durumunu iyi olarak nitelendirdiği ve 'çoğunlukla ağır iş veya fiziksel güç gerektiren işler' diyenlerin yaklaşık %70'inin genel sağlık durumunu iyi olarak nitelendirdiği görülmektedir. Bu durumda bireylerin günlük faaliyetleri arttıkça genel sağlık durumlarının da artış gösterdiği sonucuna varılabilmektedir.

4.5. Yapay Sinir Ağları Yönteminin Uygulanması

IBM SPSS Modeler paket programında yer alan yapay sinir ağı işlemcisi default olarak çok katmanlı algılayıcıları kullanmaktadır. Birçok katmanlı algılayıcının eğitilmesi ise geri yayılım metodu ile yapılmaktadır. Bunun anlamı, eğitim verilerinin girdi katmanından girerek çıktı katmanında tahminler üretmesi ve üretilen bu tahminlerin yine eğitim verisindeki gerçek değerler ile karşılaştırılmasıdır. Bu karşılaştırma sonucunda oluştuğu belirlenen hatalar geriye yayılarak bağlantı ağırlıklarının düzeltilmesinde kullanılır [80].

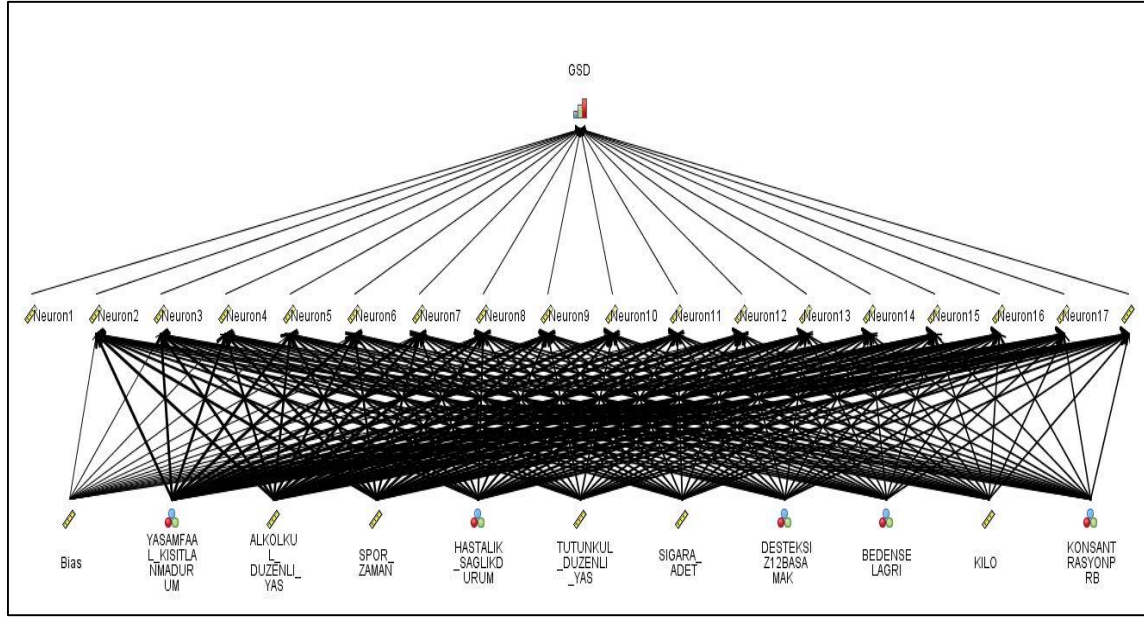
Model oluşum aşamasında RBF yöntemi de uygulanmıştır. Ancak default olarak belirlenen çok katmanlı algılayıcıların yapay sinir ağı için daha başarılı sonuçlar ürettiği gözlemlenmiştir. Elimizde büyük veri setinin olması durumunda eğitim, doğrulama ve test

olmak üzere üç farklı veri seti elde edilebilir. Çoğu çalışmada doğrulama veri setine gerek duyulmaz. Ancak modelin performansının değerlendirilmesi ve aşırı uyum (over fitting) durumundan kaçınmak için doğrulama veri setinin oluşturulması önem taşımaktadır. Test veri seti ise daha önce hiç görülmeyen durumlar için gelecekteki performansların değerlendirilmesi amacıyla kullanılmaktadır. Test veri seti sonuçları eğitim aşamasına göre kötü çıkar ise aşırı uyum durumunun olduğu söylenebilir.

IBM SPSS Modeler paket programında yer alan Partition sekmesi ile veri setinin %65'i eğitim, %25'i test ve %10'u ise doğrulama veri seti olmak üzere dallara ayrılmıştır.

Veri setinde düzenlemeler yapılarak yapay sinir ağları ve çok kategorili lojistik regresyon metodu veri setine uygulanmış ve sonuçlar yorumlanmıştır.

Yapay sinir ağları uygulanarak genel sağlık durumu bağımlı değişkeni üzerinde etkili olan bağımsız değişkenler önem seviyelerine göre belirlenmiştir. Buna göre genel sağlık durumu değişkeni üzerinde etkili değişkenler sırasıyla kişilerin günlük yaşamsal faaliyetlerinin kısıtlanma durumu, düzenli olarak alkollü içecek kullanmaya başlama yaşı, normal bir hafta içerisinde spor yapma zamanı, altı ay ya da uzun süren/sürmesi beklenen hastalık sağlık durumu, düzenli olarak tütün kullanmaya başlama yaşı, içilen sigara adedi, desteksiz on iki basamak merdiven inip çıkmakta zorluk durumu, son dört hafta içerisindeki bedensel ağrı durumu, kişilerin kilosu ve son iki hafta içerisindeki konsantrasyon durumu olduğu görülmektedir. Yapay sinir ağlarına ait çıktı görüntüsü önem seviyesine göre sıralanmış olarak Şekil 4.1'de gösterilmiştir.



Şekil 4.1. Yapay sinir ağı çıktı görünümü

Yukarıdaki Şekil 4.1’de görüldüğü üzere yapay sinir ağı sonucu 17 adet katman üretilmiş olup bu katmanlar hiperbolik tanjant fonksiyonu kullanılarak üretilmiştir.

Çizelge 4.17. Yapay sinir ağı eğitim veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	1,533	1,009	137	2,679
	Orta	682	3,707	2,138	6,527
	İyi	101	1,549	11,827	13,477
	Toplam	2,316	6,265	14,102	22,683

Elimizde bulunan veri setinin %65’lik kısmını eğitim veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.17’de gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

Çizelge 4.18. Yapay sinir ağı test veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	608	392	48	1,048
	Orta	257	1,314	837	2,408
	İyi	42	552	4,715	5,309
	Toplam	907	2,258	5,600	8,765

Elimizde bulunan veri setinin %25’lik kısmını test veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.18’de gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

Çizelge 4.19. Yapay sinir ağı doğrulama veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	218	159	18	395
	Orta	105	553	320	978
	İyi	21	249	1,904	2,174
	Toplam	344	961	2,242	3,547

Elimizde bulunan veri setinin %10’luk kısmını doğrulama veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.19’da gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

Bu durumda;

Eğitim veri seti için doğru sınıflandırma oranı	: %75,24
Eğitim veri seti için yanlış sınıflandırma oranı	: %24,76
Test veri seti için doğru sınıflandırma oranı	: %75,72
Test veri seti için yanlış sınıflandırma oranı	: %24,28
Doğrulama veri seti için doğru sınıflandırma oranı	: %75,42
Doğrulama veri seti için yanlış sınıflandırma oranı	: %24,58

olarak karşımıza çıkmaktadır.

4.6. Çok kategorili Lojistik Regresyon Yönteminin Uygulanması

Veri seti üzerinde çok kategorili lojistik regresyon yöntemi uygulama aşamasında bağımlı değişkene ait değerler arasından bir tane referans kategori belirlenmesi gerekmektedir. Sonuçlar seçilen referans kategoriye göre elde edilmektedir. Bağımlı değişken olarak ele aldığımız genel sağlık durumu değişkeni kategorilerinden en çok frekansa sahip olan genel sağlık durumunun “iyi” olarak nitelendirildiği kategori, referans kategorisi olarak alınmıştır. Model oluşturulurken IBM SPSS Modeler paket programında default olarak belirlenen ‘enter’ metodu denenmiş fakat ‘stepwise’ metodu ile daha başarılı sonuçlar elde edildiğinden

uygulama ‘stepwise’ metodu ile gerçekleştirilmiştir. Yapay sinir ağlarında olduğu gibi eğitim, test ve doğrulama veri seti aynı oranlar kullanılarak oluşturulmuştur.

Çok kategorili lojistik regresyon analizi uygulanarak genel sağlık durumu bağımlı değişkeni üzerinde etkili olan bağımsız değişkenler önem seviyelerine göre belirlenmiştir. Buna göre genel sağlık durumu değişkeni üzerinde etkili değişkenler sırasıyla kişilerin günlük yaşamsal faaliyetlerinin kısıtlanma durumu, altı ay ya da uzun süren/sürmesi beklenen hastalık sağlık durumu, son dört hafta içerisindeki bedensel ağrı durumu, kişilerin yaşı, desteksiz on iki basamak merdiven inip çıkmakta zorluk durumu, kişilerin eğitim düzeyi, reçete edilen ilaç kullanım durumu, kişilerin çalışmama nedenleri, ağır ev işleri yapmaktaki zorluk çekme durumu ve yorgunluk durumu olduğu görülmektedir.

Olabilirlik oran testi modele dahil edilen herhangi bir açıklayıcı değişken için seviyelerinin hepsini eş zamanlı olarak test eder. Açıklayıcı değişkenin herhangi bir tek seviyesi için test yapılmak istendiğinde Wald istatistiği kullanılır [21].

Wald istatistiği sonuçlarına göre genel sağlık durumunun kötü olarak nitelendirilmesinde en çok kişilerin günlük yaşamsal faaliyetlerinin kısıtlanma durumu, altı ay ya da uzun süren/sürmesi beklenen hastalık sağlık durumu, son dört hafta içerisindeki bedensel ağrı durumunun etkili olduğu görülmektedir. ($p < 0,05$)

Wald istatistiği sonuçlarına göre genel sağlık durumunun orta olarak nitelendirilmesinde en çok altı ay ya da uzun süren/sürmesi beklenen hastalık sağlık durumu, günlük yaşamsal faaliyetlerinin kısıtlanma durumu, reçete edilen ilaç kullanım durumunun etkili olduğu görülmektedir.

Çizelge 4.20. Çok kategorili lojistik regresyon model uyum ölçütü

Model	Model Uyum Ölçütü	Olabilirlik Oran Testi		
	-2 Log Likelihood	Ki-kare	Ser. Der.	p
Sabit	41739,659			
Son model	24981,857	16757,801	166	,000

Model uyum bilgileri çizelgesi incelendiğinde, Çizelge 4.20’de son modelin anlamlılık (p) değerinin 0.05’ten küçük olduğu görülmektedir. Bu durumda modelin anlamlı olduğu

söylenbilir. Yani, modelde yer alan bağımsız değişkenler ile bağımlı değişkenler arasında anlamlı bir ilişki olduğu söylenbilir.

Çizelge 4.21. Çok kategorili lojistik regresyon Pseudo R^2 (sözde R^2) değerleri

Cox and Snell	,522
Nagelkerke	,621
McFadden	,401

Çizelge 4.21’de çok kategorili lojistik regresyon modeline ilişkin log-olabilirlik değeri ve belirtme katsayıları gösterilmektedir. Cox Snell, Nagelkerke ve McFadden belirtme katsayılarına bakıldığında genel sağlık durumu değişkeninin bağımsız değişkenler tarafından orta düzeyde açıklanabildiği sonucuna varılabilir.

Çizelge 4.22. Çok kategorili lojistik regresyon eğitim veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	1,491	1,055	133	2,679
	Orta	558	3,775	2,194	6,527
	İyi	73	1,425	11,979	13,477
	Toplam	2,122	6,255	14,306	22,683

Elimizde bulunan veri setinin %65’lik kısmını eğitim veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.22’de gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

Çizelge 4.23. Çok kategorili lojistik regresyon test veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	594	399	55	1,048
	Orta	210	1,352	846	2,408
	İyi	29	480	4,800	5,309
	Toplam	833	2,231	5,701	8,765

Elimizde bulunan veri setinin %25’lik kısmını test veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.23’te gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

Çizelge 4.24. Çok kategorili lojistik regresyon doğrulama veri seti sonucu

		Model Sonucu			
		Kötü	Orta	İyi	Toplam
Gerçek Durum	Kötü	210	169	16	395
	Orta	75	579	324	978
	İyi	11	244	1,919	2,174
	Toplam	296	992	2,259	3,547

Elimizde bulunan veri setinin %10'luk kısmını doğrulama veri setine ayırdığımızda elde ettiğimiz sonuç Çizelge 4.24'de gösterilmektedir. Gerçek durumdaki genel sağlık durumları ve modelin sonucunda elde edilen genel sağlık durumlarının tahmin değerleri yer almaktadır.

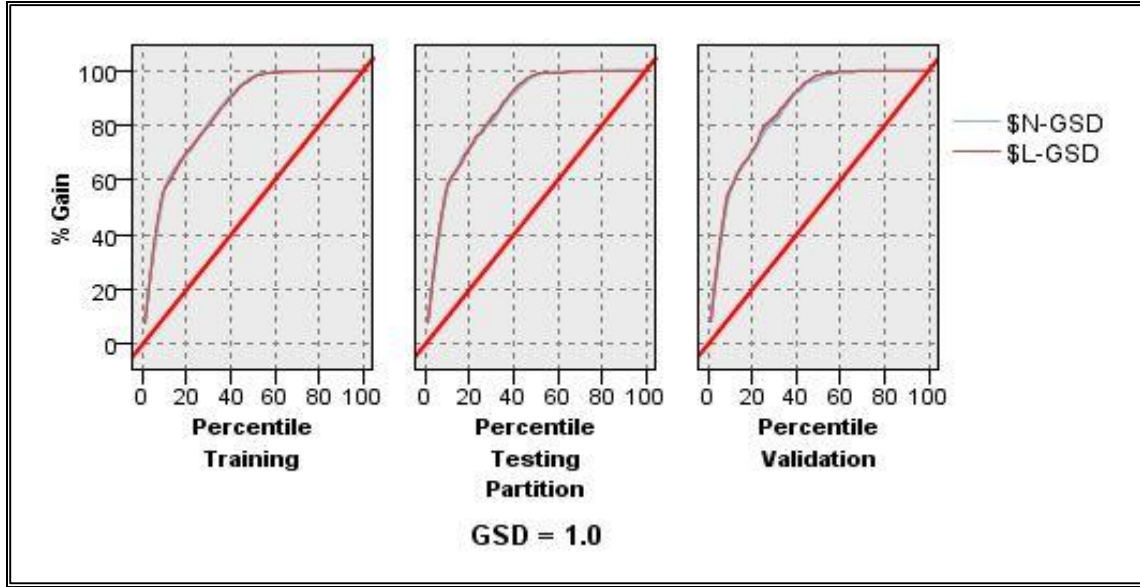
Bu durumda;

Eğitim veri seti için doğru sınıflandırma oranı	: %76,03
Eğitim veri seti için yanlış sınıflandırma oranı	: %23,97
Test veri seti için doğru sınıflandırma oranı	: %76,97
Test veri seti için yanlış sınıflandırma oranı	: %23,03
Doğrulama veri seti için doğru sınıflandırma oranı	: %76,35
Doğrulama veri seti için yanlış sınıflandırma oranı	: %23,65

olarak karşımıza çıkmaktadır.

4.7. Yöntemlerin Karşılaştırılması

Bu çalışmada; TSA verilerine çok kategorili lojistik regresyon ve yapay sinir ağı modeli uygulanmıştır. Her iki yöntem ile “genel sağlık durumu” bağımlı değişkeni ele alınan bağımsız değişkenler ile tahmin edilmeye çalışılmış olup ayrıca bireylerin genel sağlık durumuna etki eden bağımsız değişkenler önem seviyesine göre sıralanmıştır. Her iki modele ait sonuçlar, eğitim, test ve doğrulama veri setlerini içermektedir. Bu veri setleri ile elde edilen doğru ve yanlış sınıflandırma oranları hesaplanmıştır. Hesaplanan bu oranlar birbirleri ile karşılaştırılmıştır.



Şekil 4.2. Yapay sinir ağları ve çok kategorili lojistik regresyon karşılaştırılması

Şekil 4.2’de yapay sinir ağlarına ait oluşturulan model ile çok kategorili lojistik regresyona ait sonuçlar karşılaştırılarak gösterilmiş ve birbirine yakın sonuçların olduğu görülmüştür.

Lojistik regresyon analizinde veri setini ikiye ayrılma zorunluluğu yoktur. Yapay sinir ağları yönteminde ise genellikle veri seti eğitim ve test veri seti olmak üzere ikiye ayrılmaktadır. Aşırı uyum durumunun engellenmesi için ise bir de doğrulama veri setinin oluşturulmasında fayda vardır. Bu çalışmada yapay sinir ağları ve çok kategorili lojistik regresyon yöntemi uygulama aşamasında her iki yöntem için veri seti üç gruba ayrılmış ve bu yüzden veri setinin hacminde azalma meydana gelmiştir. Modeller, veri bölümlemesine göre her bir alt veri seti için doğruluk oranları belirlenerek değerlendirilmiştir. Modellerin doğruluk oranlarının farklı veri setlerine göre doğru ve yanlış tahmin oranları Çizelge 4.25 ve Çizelge 4.26’da yer almaktadır.

Çizelge 4.25. Yapay sinir ağları modelinin alt veri setleri içindeki doğruluk oranlarının karşılaştırılması

Model Sonucu	Eğitim Verisi		Test Verisi		Doğrulama Verisi	
	Adet	Oran (%)	Adet	Oran (%)	Adet	Oran (%)
Doğru	17 067	75,24%	6 637	75,72%	2 675	75,42%
Yanlış	5 616	24,76%	2 128	24,28%	872	24,58%
Toplam	22 683	100,00%	8 765	100,00%	3 547	100,00%

Çizelge 4.26. Çok kategorili lojistik regresyon modelinin alt veri setleri içindeki doğruluk oranlarının karşılaştırılması

Model Sonucu	Eğitim Verisi		Test Verisi		Doğrulama Verisi	
	Adet	Oran (%)	Adet	Oran (%)	Adet	Oran (%)
Doğru	17 245	76,03%	6 746	76,97%	2 708	76,35%
Yanlış	5 438	23,97%	2 019	23,03%	839	23,65%
Toplam	22 683	100,00%	8 765	100,00%	3 547	100,00%

Eğitim veri setinde yer alan doğruluk oranı ile doğrulama ve test veri setinde yer alan doğruluk oranlarının birbirine yakın olması modelin tutarlılığı bakımından önem taşımaktadır. Eğitim veri setinde doğruluk oranına baktığımızda %76,03 ile çok kategorili lojistik regresyon modelinin yapay sinir ağları modelinin doğruluk oranına (%75,24) göre daha yüksek olduğu görülmektedir. Test ve doğrulama veri setindeki doğruluk oranlarına baktığımızda da sırasıyla %76,97 ve %76,35 ile çok kategorili lojistik regresyon modelinin yapay sinir ağları modelinin doğruluk oranına (%75,72 ve %75,42) göre daha yüksek olduğu görülmektedir.

5. SONUÇ

Yapay sinir ağlarında veri seti için bilgisayarla en uygun modelin belirlenmesi süreci daha fazla zaman gerektirirken lojistik regresyon analizi ile uygun modelin oluşturulması bilgisayarla çok kısa bir sürede gerçekleştirilebilmektedir. Lojistik regresyon analizi ile elde edilen modelin genel yapısı yapay sinir ağına göre daha basittir. Yapay sinir ağları modelinde gizli tabaka ile gizli nöron sayıları fazla olduğu için parametre tahminlerinin yorumu zorlaşmaktadır ve detaylı bilgi elde edilememektedir. Ayrıca lojistik regresyon analizinde bağımlı değişken üzerinde etkisi olan değişkenler modele alınarak etkili olmayan değişkenlerin elenebilme durumu söz konusudur. Yapay sinir ağlarında ise etkisi olmayan değişkenler modelden çıkarılamamaktadır.

Yapay sinir ağlarının farklı istatistikî işlemleri gerçekleştirilmesindeki önemi kaçınılmazdır. Yapay sinir ağları karmaşık problemlerin modellenmesi ve çözümlenmesinde kullanılabilen matematiksel bir modeldir. Doğrusal olan veya doğrusal olmayan modeller için seçim imkânına sahip olması yapay sinir ağlarının lojistik regresyon analizine göre üstünlüğü arasındadır.

Çok kategorili lojistik regresyon analizi yönteminde cevap değişkeninin herhangi bir kategorisi, referans kategori olarak seçilir ve diğer kategoriler bu referans kategoriye göre analiz edilir. Bu referans kategorisi tamamen araştırmacının tercihine bağlı olarak belirlenir. İncelenen çalışmaların bazılarında en çok frekansa sahip olan değer referans kategorisi olarak belirlenmiştir. Zira referans kategorisinin değişmesi modelin yapısı hakkında herhangi bir değişikliğe neden olmamaktadır. Yalnızca parametre tahminlerinde ve yorumlarda değişikliğe neden olmaktadır. Cevap değişkeninin M kategoriden oluştuğu durumda, cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişkinin incelenmesinde referans kategorisi ile her bir kategorinin tek tek incelendiği $M-1$ tane denklem hesaplanır.

Her iki yöntemin avantajı, dezavantajı, üstünlükleri ve eksiklikleri söz konusudur. Araştırma kapsamında ele alınan veri setinin de uygulanan yöntemin performansına etki ettiği söylenebilir.

Bu çalışmada genel olarak yapay sinir ağları ile lojistik regresyon analizi kapsamında yer alan çok kategorili lojistik regresyon analizi özetlenmiştir. İki farklı yöntemin sınıflandırma

başarılarının karşılaştırılması amacıyla TÜİK tarafından iki yılda bir yapılan Türkiye Sağlık Araştırması (TSA) kapsamında yer alan 2014 ve 2016 yılına ait veri seti ele alınmıştır.

Uygulama kısmında veri seti eğitim, test ve doğrulama verisi olmak üzere üç kısma ayrılmıştır. Her iki model için her bir veri kümesine ait doğruluk oranlarına bakılmıştır. Çok kategorili lojistik regresyon analizi ile elde edilen doğru sınıflandırma oranları eğitim verisi, test verisi ve doğrulama verisi için sırasıyla %76,03, %76,97 ve %76,35'tir. Yapay sinir ağları yöntemi ile elde edilen doğru sınıflandırma oranları eğitim verisi, test verisi ve doğrulama verisi için ise sırasıyla %75,24, %75,72 ve %75,42 olarak bulunmuştur. Her iki model için alt veri setlerindeki doğruluk oranlarının tutarlı olduğu görülmüştür. Her iki model için alt veri setlerinden en yüksek doğruluk oranının test verisi için elde edildiği görülmektedir. Test veri seti için oluşturulan doğruluk oranı, eğitim veri seti için oluşturulan doğruluk oranından yüksek olduğu için aşırı uyum durumunun gözlemlenmediği görülmüştür.

Her bir alt veri seti için doğruluk oranına baktığımızda çok kategorili lojistik regresyon analizi sonuçları yapay sinir ağı ile elde edilen doğruluk oranlarına göre daha başarılıdır. Ancak beklendiği üzere performansları arasındaki farklılık oldukça düşük bir seviyede gerçekleşmiştir. Literatüre bakıldığında bu amaçla yapılmış olan çalışmaların çoğunda yapay sinir ağı çok kategorili lojistik regresyon analizine göre daha iyi bir sınıflandırma oranına sahip olmuştur. Ancak uygulama sırasında veri setinin bölümlenmesi işleminin çok kategorili lojistik regresyon analizinde de uygulanmış olması çok kategorili lojistik regresyon analizini yapay sinir ağlarına göre başarılı kılmıştır.

Analiz sonuçlarına göre; model karmaşıklığı, aşırı uyum ve eğitim algoritmasının seçimi gibi eğitim performansını etkileyen faktörlerin her iki yöntem için göz önünde bulundurulmasıyla çok kategorili lojistik regresyon analizinin yapay sinir ağlarına göre daha iyi bir sınıflandırma performansı sağladığı görülmüştür.

Uygulanan yöntemlerin doğru sınıflandırma başarılarının ilgilenilen veri seti ile ilişkili olduğu sonucuna varılabilir. Literatür incelendiğinde her zaman aynı yöntemin yüksek doğruluk oranına sahip olmadığı görülmüştür. Bu çalışmanın ileride yapılacak olan çalışmalara ışık tutması ümit edilmektedir.

KAYNAKLAR

1. Aslan, H. (2018). *Yapay sinir ağları kullanarak alışveriş merkezi lokasyon belirlemeye ilişkin model önerisi*. Ankara: Türk Hava Kurumu Üniversitesi.
2. McCulloch, W. S., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.
3. Keskenler, M. F., Keskenler, E. F. (2017). Geçmişten günümüze yapay sinir ağları ve tarihçesi. *Takvim-i Vekayi*, 5(2), 8-18.
4. Hebb, D. O. (1949). *The first stage of perception: growth of the assembly*. 'The Organization of Behavior'. Newyork: Psychology Press, 4-78, 102-120.
5. Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386.
6. Caianiello, E. R. (1961). Outline of a theory of thought-processes and thinking machines. *Journal of Theoretical Biology*, 1(2), 204-235.
7. Irwin, G. W., Warwick, K., Hunt, K. J. (1995). *Neural network applications in control*. London: Institution of Electrical Engineers.
8. Broomhead, D. S., Lowe, D. (1988). Radial basis functions, multi-variable functional interpolation and adaptive networks (No. RSRE-MEMO-4148). Royal Signals and Radar Establishment Malvern DoD S&T Reports, *United Kingdom*, 41-48.
9. Alpar, R. (2013). *Uygulamalı çok değişkenli istatistiksel yöntemler*. Ankara: Detay Yayıncılık.
10. Ürük, E. (2007). *İstatistiksel uyg. logistik regresyon analizi*. İstanbul: Marmara Üniversitesi.
11. Kleinbaum, D. G., Klein, M. (2002). *Logistic Regression*. (2nd Edition). USA: Department of Epidemiology Emory University.
12. Kuyucu, Y. E. (2012). *Lojistik regresyon analizi (LRA), yapay sinir ağları (YSA) ve sınıflandırma ve regresyon ağaçları (C&RT) yöntemlerinin karşılaştırılması ve tıp alanında bir uygulama*. Tokat: Gaziosmanpaşa Üniversitesi.
13. Gürcan, M. (1998). *Lojistik regresyon analizi ve bir uygulama*. Yüksek Lisans Tezi. Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Samsun.
14. Berkson, J. (1944). Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, 39(227), 357-365.
15. Cox, R. A. (1970). Conformation of nucleic acids and the analysis of the hypochromic effect. *Biochemical Journal*, 120(3), 539.
16. Cornfield, J. (1962, July). Joint dependence of risk of coronary heart disease on serum cholesterol and systolic blood pressure: A discriminant function analysis. *Federation Proceedings*, 21, 58-61.
17. Le, C. T. (1984). Logistic models for cross-over designs. *Biometrika*, 71(1), 216-217.
18. Bonney, G. E. (1987). Logistic regression for dependent binary observations. *Biometrics*, 951-973.

19. Seven, Z. (1997). *Değişken seçimi yöntemi olarak adımsal lojistik regresyon ile adımsal diskriminant analizinin karşılaştırılması*. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
20. Elmas, Ç. (2003). *Yapay sinir ağları*. Ankara: Seçkin Yayınevi.
21. Elasan, S. (2010). *Multinomiyal logistik regresyon analizi ve uygulaması*, Yüksek Lisans Tezi, Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Van.
22. Mehmedov, M., Eryılmaz, H. (2003). *Yapay sinir ağları ile bazı istatistiksel modeller arasındaki ilişki*. International XII. Turkish Symposium on Artificial Intelligence and Network (TAINN 2003), Çanakkale.
23. Karakuş, K. (2009). *Veri madenciliği teknikleri ile mobil telekom sektöründe müşterilerin kredi skorlamasına ilişkin istatistiksel bir analizi*. Yüksek Lisans Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
24. Kayri, M., Çokluk, Ö. (2010). Using multinomial logistic regression analysis in artificial neural network: An application. *Ozean Journal of Applied Sciences*, 3(2), 259-268.
25. Pourshahriar, H. (2012). Correct vs. accurate prediction: A comparison between prediction power of artificial neural networks and logistic regression in psychological researches. *Procedia-Social and Behavioral Sciences*, 32, 97-103.
26. Budak, H., Erpolat, S. (2012). Comparison of artificial neural networks and logistic regression analysis in the credit risk prediction. *Online International Journal of Information Technology*, 3(9), 23.
27. Asogwa, O. C., Oladugba, A. V. (2015). On the comparison of artificial neural network (ANN) and multinomial logistic regression (MLR). *West African Journal of Industrial and Academic Research*, 13(1), 3-9.
28. Arundina, T., Omar, M. A., Kartiwi, M. (2015). The predictive accuracy of Sukuk ratings; Multinomial Logistic and Neural Network inferences. *Pacific-Basin Finance Journal*, 34, 273-292.
29. Hassanipour, S., Ghaem, H., Arab-Zozani, M., Seif, M., Fararouei, M., Abdzadeh, E., Sabetian, G., Paydar, S. (2019). Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and meta-analysis. *Injury*, 50(2), 244--250.
30. Baş, C. (2010). *Cevap yüzeyi tasarımları ve sinir ağları yaklaşımı*. Doktora Tezi, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
31. Zhang, G., Patuwo, B. E., Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14(1), 35-62.
32. İnternet: Kakıcı, A. *Yapay sinir ağlarına giriş*. [Online]. URL: <https://yapayzeka.ai/yapay-sinir-aglarına-giris-1/>, Son Erişim Tarihi: 13.01.2019.
33. Kotu, V., Deshpande, B. (2015). *Predictive analytics and data mining: concepts and practice with Rapidminer*. Waltham: Morgan Kaufmann.
34. Han, J., Kamber, M., Pei, J. (2011). *Data mining: concepts and techniques*. USA: Morgan Kaufmann.
35. Öztemel, E. (2003). *Yapay sinir ağları*. İstanbul: Papatya Yayıncılık.

36. Yurtoğlu, H. (2005). *Yapay sinir ağları metodolojisi ile öngörü modellemesi*. Ankara: Ekonomik Modeller ve Stratejik Araştırmalar Genel Müdürlüğü Yayınları.
37. Sağiroğlu, S., Beşdok, E., Erler, M. (2003). *Muhendislikte yapay zeka uygulamaları-I: yapay sinir ağları*. Kayseri: Ufuk Kitabevi, 10-100.
38. Özcan, B. (2007). *Yapay sinir ağı yaklaşımıyla, peçete makinesi imalatı yapan işletmede makine işleme sürelerinin tahmin edilmesi*. Yüksek Lisans Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli.
39. Doğan, M. İ. (2012). *Ekonomik bir büyüme modelinin yapay sinir ağları ile tahmini ve türkiye uygulaması*. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
40. Elmas, Ç. (2016). *Yapay zeka uygulamaları, yapay sinir ağı, bulanık mantık, sinirsel bulanık mantık, genetik algoritma*. Ankara: Seçkin Yayıncılık.
41. Öztemel, E. (2012). *Yapay sinir ağları*. İstanbul: Papatya Yayıncılık.
42. Ocakoğlu, G. (2006). *Lojistik regresyon analizi ve yapay sinir ağları tekniklerinin sınıflama özelliklerinin karşılaştırılması ve bir uygulama*. Yüksek Lisans Tezi, Uludağ Üniversitesi Sağlık Bilimleri Enstitüsü, Bursa.
43. Haykin, S. (1999). *Neural networks: A comprehensive foundation*. New Jersey: Prentice Hall.
44. Fausett, L. (1994). *Fundamentals of neural networks: Architectures, algorithms and applications*. New Jersey: Prentice Hall.
45. Abdi, H., Valentin, D., Edelman, B. (1999). *Neural networks*. California: Sage Publications.
46. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1), 59-69.
47. Saraç, T. (2004). *Yapay sinir ağları*. Basılmamış Seminer Projesi, Gazi Üniversitesi, Ankara.
48. Karasu, F. (2012). *Petrokimya sektöründe talep tahmininde yapay sinir ağlarının kullanılması 'petkim a.ş. örneği'*. Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
49. Hamzaçebi, C. (2011). *Yapay sinir ağları: Tahmin amaçlı kullanımı matlab ve neurosolutions uygulamalı*. Trabzon: Ekin Yayınevi.
50. Bishop, M. (1995). *Neural networks for pattern recognition*. UK: Clarendon Press.
51. Güngör, E. (2007). *Yapay sinir ağları yardımı ile makine arızalarının önceden tahmin edilmesi*. Yüksek Lisans Tezi, Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü, Kocaeli.
52. Kohonen, T. (1984). *Self organization and associative memory*. New York: Springer-Verlag.
53. Jang, J. S. R., Sun, C. T., Mizutani, E. (1997). Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. *The Institute of Electrical and Electronics Engineers Transactions on Automatic Control*, 42(10), 1482-1484.

54. Haykin, S. (1994). *Neural network, a comprehensive foundation*. New York: Macmillan College Publishing Company.
55. Kuo, C., Reitsch, A. (1995). Neural networks vs. conventional methods of forecasting. *The Journal of Business Forecasting*, 14(4), 17.
56. Powers, D., Xie, Y. (2000). *Statistical methods for categorical data analysis*. USA: Emerald Group Publishing.
57. Elhan, A. H. (1997). *Lojistik regresyon analizinin incelenmesi ve tipta bir uygulaması*. Yüksek Lisans Tezi, Ankara Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara.
58. Kleinbaum, D. G., Klein, M. (2010). *A self-learning text logistic regression*. Newyork: Springer.
59. Güner, Z. B. (2014). Veri madenciliğinde cart ve lojistik regresyon analizinin yeri: ilaç provizyon sistemi verileri üzerinde örnek bir uygulama. *Sosyal Güvençe Dergisi*, 6, 53-99.
60. Hosmer, D. W., Lemeshow, S., Sturdivant, R. X. (2000). Introduction to the logistic regression model. *Applied Logistic Regression*, 2, 1-30.
61. Andersen, E. B. (1997). *Introduction to the statistical analysis of categorical data*. Berlin: Springer-Verlag.
62. Montgomery, D. C., Peck, E. A., Vining, G. G. (2012). *Introduction to linear regression analysis*. (Fifth Edition). New Jersey: John Wiley & Sons.
63. Agresti, A. (1996). *An introduction to categorical data analysis*. New York: John Wiley & Sons.
64. Özdamar, K. (2004). *Paket programları ile istatistiksel veri analizi 1*. Eskişehir: Kaan Kitabevi.
65. Özdamar, K. (2004). *Paket programları ile istatistiksel veri analizi 2*. Eskişehir: Kaan Kitabevi.
66. Korkmaz, T. K. (2004). Bankalarda kredi riski ölçümünde alternatif yöntemler. *Active Bankacılık ve Finans Dergisi*, 5(11), 49-62.
67. Turangil, N. (2006). *Kredi skorlamada kullanılan yöntemler ve uygulamaları*. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Üniversitesi, İstanbul.
68. Şahin, E. (2007). *Lojistik regresyon analizi ve bir uygulama*. İstanbul: Yıldız Teknik Üniversitesi.
69. Fagerland, M. W., Hosmer, D. W., Bofin, A. M. (2008). Multinomial goodness-of-fit tests for logistic regression models. *Statistics in Medicine*, 27(21), 4238-4253.
70. Özer, E. (2015). *Epilepsi Hastalığının dalgacık dönüşümleri ve yapay sinir ağları ile tanınması*. Yüksek Lisans Tezi, Mimar Sinan Güzel Sanatlar Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
71. Sharma, S. (1996). *Applied multivariate techniques*. New York; John Wiley&Sons. Inc.
72. Meyers, L. S., Gamst, G., Guarino, A. J. (2016). *Applied multivariate research: Design and interpretation*. United Kingdom: Sage publications.

73. Mertler, C. A., Reinhart, R. V. (2016). *Advanced and multivariate statistical methods: Practical application and interpretation*. Abingdon: Routledge.
74. Alpar, R. (2011). *Uygulamalı çok değişkenli istatistiksel yöntemler*. Ankara: Detay Yayıncılık.
75. İnternet: TÜİK. (2015). *Türkiye Sağlık Araştırması*, 2014. <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=18854>, Son Erişim Tarihi: 03.03.2019.
76. İnternet: TÜİK. (2017). *Türkiye Sağlık Araştırması*, 2016. URL: <http://www.tuik.gov.tr/PreHaberBultenleri.do?id=24573>, Son Erişim Tarihi: 03.03.2019.
77. İnternet: TÜİK. *Meslek Sınıflamaları*. URL: <https://biruni.tuik.gov.tr/DIESS/SiniflamaSurumListeAction.do?turId=41&turAdi=9>. Meslek Sınıflamaları&guncel=Y, Son Erişim Tarihi: 03.03.2019.
78. İnternet: TÜİK, *Faaliyet Sınıflamaları*. URL: <https://biruni.tuik.gov.tr/DIESS/SiniflamaSurumListeAction.do?turId=1&turAdi=1>. Faaliyet Sınıflamaları&guncel=Y, Son Erişim Tarihi: 03.03.2019.
79. Burmaoğlu, S. (2009). *Birleşmiş milletler kalkınma programı beşeri kalkınma endeksi verilerini kullanarak diskriminant analizi, lojistik regresyon analizi ve yapay sinir ağlarının sınıflandırma başarılarının değerlendirilmesi*. Doktora Tezi, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü, Erzurum.
80. International Business Machines Corporation. (2012). *IBM SPSS modeler 15 algorithms guide*. New York: IBM Corporation.

EKLER

EK-1. SPSS paket programında elde edilen ve VIF değerleri neticesinde veri setinden çıkarılan/çıkarılmayan değişkenlere ilişkin tablo

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5,613	,510	10,995	,000		
	YAS_SIRA	,059	,003	,119	,000	,307	3,255
	CINSİYET	-,012	,010	-,007	,260	,356	2,807
	DOGUM_YER	,000	,000	,001	,300	,764	1,765
	VATANDASLIK	-,001	,001	-,006	,204	,565	1,770
	TEDAVI_MASRAF_SGK	,068	,014	,030	4,706	,000	3,172
	TEDAVI_MASRAF_OZELSAGSIG	,055	,022	,009	2,538	,011	1,084
	TEDAVI_MASRAF_OZELSANDIK	,036	,043	,003	,837	,403	1,054
	TEDAVI_MASRAF_GSS	,026	,017	,009	1,546	,122	2,820
	TEDAVI_MASRAF_KENDI	,020	,008	,009	2,384	,017	1,224
	TEDAVI_MASRAF_DIGER	,093	,026	,013	3,517	,000	1,049
	EGITIM_SIRALI	-,025	,003	-,048	-8,682	,000	2,465
	MEDENI_DURUM	,002	,006	,001	,269	,788	1,944
	CALISMA_DURUM	,008	,020	,009	,428	,669	38,960
	CALISMAMA_NEDEN	,000	,001	-,003	-,766	,444	1,187
	ONCEDEN_CALISMA_DURUM	-,028	,027	-,028	-1,061	,289	55,711
	ISTEKI_DURUM	,003	,005	,003	,560	,575	2,923
	IS_SUREKLILIK	-,002	,012	-,001	-,128	,898	5,759
	CALISMA_SEKLI	-,013	,017	-,008	-,778	,437	9,042
	ISCO_GRUP	,001	,002	,005	,548	,584	5,517
	NACE_GRUP	-2,533E-005	,001	,000	-,031	,976	2,419
	GELIR_KAYNAK_UCRET	,016	,010	,015	1,528	,127	7,895
	GELIR_KAYNAK_MUTESEBBIS	-,009	,014	-,010	-,658	,510	18,480
	GELIR_KAYNAK_GAYRIMENK	,005	,017	,005	,291	,771	27,539
	GELIR_KAYNAK_MENKUL	,022	,032	,024	,669	,503	107,356
	GELIR_KAYNAK_EMEKLIMAAS	,019	,011	,020	1,728	,084	10,438
	GELIR_KAYNAK_DIGERSOSYAL_TRNS	-,026	,021	-,030	-1,236	,216	46,280
	GELIR_KAYNAK_HANE_TRNS	-,032	,028	-,037	-1,179	,238	78,093
	HANE_GELIR	-,021	,003	-,034	-7,785	,000	1,517
	HASTALIK_SAGLIKDURUM	-,263	,009	-,152	-29,132	,000	2,176
	YASAMFAAL_KISITLANMADURUM	-,268	,006	-,218	-44,216	,000	1,944
	KRONIKHAST_ASTIM	-,057	,013	-,018	-4,517	,000	1,344
	KRONIKHAST_KRONIKBRONSIT	-,043	,013	-,014	-3,331	,001	1,331
	KRONIKHAST_ENFARKTUS	-,068	,022	-,012	-3,068	,002	1,199

EK-1. (devam) SPSS paket programında elde edilen ve VIF değerleri neticesinde veri setinden çıkarılan/ çıkarılmayan değişkenlere ilişkin tablo

KRONIKHAST_KORONERKALPHAST	-,089	,013	-,028	-6,979	,000	,763	1,311
KRONIKHAST_HIPERTANSİYON	-,032	,010	-,015	-3,267	,001	,628	1,593
KRONIKHAST_INMEFELC	-,093	,033	-,010	-2,784	,005	,933	1,072
KRONIKHAST_ARTHROSIS	,012	,012	,004	,975	,330	,808	1,238
KRONIKHAST_BELBOLGE	-,044	,008	-,024	-5,473	,000	,667	1,500
KRONIKHAST_BOYUNBOLGE	-,039	,009	-,019	-4,443	,000	,706	1,416
KRONIKHAST_SEKERHAST	-,097	,011	-,035	-8,753	,000	,796	1,256
KRONIKHAST_ALERJİ	-,006	,010	-,002	-,595	,552	,895	1,117
KRONIKHAST_KARACİGERİYETMZ	-,107	,024	-,016	-4,404	,000	,939	1,065
KRONIKHAST_IDRARKACIRMA	-,012	,013	-,004	-,929	,353	,774	1,292
KRONIKHAST_BOBREK	-,052	,013	-,016	-4,187	,000	,868	1,152
KRONIKHAST_DEPRESYON	-,036	,012	-,012	-2,932	,003	,712	1,404
KRONIKHAST_ALZHEIMER	-,031	,030	-,004	-1,030	,303	,865	1,156
KRONIKHAST_COLYAK	,013	,053	,001	,240	,810	,987	1,013
KRONIKHAST_MADDEKUL	-,438	,126	-,012	-3,480	,001	,987	1,013
KRONIKHAST_DOWN	-,392	,140	-,010	-2,796	,005	,983	1,017
YARALANMA_TRAFIKKAZA	,054	,057	,008	,955	,340	,163	6,140
TEDAVI_TRAFIKKAZA	,014	,019	,006	,728	,466	,163	6,120
YARALANMA_EVKAZA	,003	,050	,000	,050	,960	,136	7,376
TEDAVI_EVKAZA	-,012	,016	-,007	-,721	,471	,135	7,410
YARALANMA_GUNDELIKKAZA	,076	,049	,014	1,534	,125	,145	6,888
TEDAVI_GUNDELIKKAZA	-,002	,016	-,001	-,100	,920	,144	6,927
YARALANMA_ISKAZA	,021	,069	,003	,302	,763	,146	6,856
TEDAVI_ISKAZA	-,011	,025	-,004	-,449	,653	,148	6,754
SON12AY_ISEARAVERME	-,004	,016	-,004	-,262	,794	,044	22,712
SON12AY_ISEARAVERME_TOPGUN	,000	,000	,004	1,094	,274	,798	1,253
GOZLUKKULLANMA	,004	,007	,002	,572	,568	,727	1,376
ISITMECIHAZIKULLANMA	,035	,016	,008	2,252	,024	,988	1,012
DESTEKSİZ12BASAMAK	,083	,007	,066	12,087	,000	,419	2,385
OGRENME	,004	,007	,003	,574	,566	,502	1,993
HATIRLAMA	,038	,007	,024	5,201	,000	,610	1,640
ZORLUKCEKME_BESLENME	,021	,018	,006	1,181	,238	,490	2,039
ZORLUKCEKME_YATMOTURMKALKM	,075	,019	,026	3,976	,000	,284	3,522
ZORLUKCEKME_GIYINMESOYUNMA	-,022	,024	-,007	-,933	,351	,198	5,059
ZORLUKCEKME_TUVALETKLLANIM	,047	,023	,015	2,052	,040	,224	4,466
ZORLUKCEKME_BANYOYAPMA	,010	,020	,004	,504	,614	,218	4,597
ZORLUKCEKME_YEMEKHAZIRLAMA	-,006	,006	-,005	-1,142	,254	,591	1,691
ZORLUKCEKME_TELEFONKULLANM	-,011	,007	-,008	-1,515	,130	,505	1,979
ZORLUKCEKME_ALISVERIS	,003	,007	,002	,422	,673	,399	2,507
ZORLUKCEKME_ILACKULLANIM	,024	,012	,010	2,018	,044	,511	1,959
ZORLUKCEKME_HAFİFEVİS	,009	,005	,009	1,764	,078	,461	2,170
ZORLUKCEKME_AGİREVİS	,011	,004	,015	2,844	,004	,471	2,121
ZORLUKCEKME_IDARİMALIKONU	-,007	,005	-,007	-1,398	,162	,564	1,774
BEDENSELAGRI	,058	,003	,106	21,991	,000	,538	1,858

EK-1. (devam) SPSS paket programında elde edilen ve VIF değerleri neticesinde veri setinden çıkarılan/çıkarılmayan değişkenlere ilişkin tablo

BUNALIM	,025	,006	,022	4,346	,000	,485	2,063
UYKUSORUN	,010	,005	,009	1,987	,047	,564	1,773
YORGUNLUK	,022	,005	,021	4,030	,000	,473	2,115
ISTAHSIZLIK	,026	,006	,020	4,464	,000	,625	1,599
DEGERSIZHISSETME	,006	,007	,004	,788	,431	,447	2,237
KONSANTRASYONPRB	-,009	,008	-,006	-1,220	,222	,491	2,038
HUZURSUZLUK	,026	,008	,015	3,090	,002	,520	1,923
YATAKLIHIZMET_TOPGECE	,001	,001	,009	2,284	,022	,894	1,118
GUNUBIRLIK_TOPGUN	,003	,000	,030	7,624	,000	,831	1,203
SAGKURUM_TERCIH	,000	,000	-,003	-,820	,412	,979	1,022
DIS_HEK_HIZALMAZAMAN	,008	,003	,009	2,313	,021	,901	1,110
AILE_PRTS_HEK_HIZALMA_KERE	-,001	,002	-,002	-,485	,628	,856	1,168
UZMAN_HEK_HIZALMADURUM	-,029	,006	-,020	-5,015	,000	,787	1,271
UZMAN_HEK_HIZALMA_KERE	,006	,002	,013	3,245	,001	,781	1,280
SON12AY_FIZYOTERAPIST	,073	,026	,011	2,850	,004	,804	1,243
SON12AY_FIZIKTEDVUZMAN	-,028	,015	-,007	-1,840	,066	,771	1,297
SON12AY_PSIKOLOG	-,021	,022	-,004	-,981	,327	,799	1,252
SON12AY_PSIKOTERAPIST	,002	,038	,000	,050	,960	,821	1,219
SON12AY_PSIKIYATRIST	-,047	,019	-,010	-2,498	,013	,748	1,337
SOSYALYARDIMHIZMETI	,013	,043	,001	,292	,770	,947	1,055
RECETEEDILEN_ILACKULLANIM	-,066	,008	-,037	-8,553	,000	,684	1,461
KENDI_ILAC	-,004	,007	-,002	-,575	,565	,930	1,075
KENDI_GIDATAKVIYE	,001	,014	,000	,081	,936	,938	1,066
GRIPASISOLMA_ZAMAN	,003	,007	,002	,420	,675	,936	1,068
ODEMEGUCLUGU_TIBBIBAKIM	-,030	,010	-,017	-3,114	,002	,438	2,285
ODEMEGUCLUGU_DISBAKIM	-,011	,008	-,007	-1,403	,161	,555	1,802
ODEMEGUCLUGU_ILAC	-,032	,011	-,016	-2,977	,003	,450	2,220
ODEMEGUCLUGU_RUHTEDV	,007	,007	,004	1,020	,308	,717	1,394
BOY	-,002	,001	-,021	-4,044	,000	,458	2,185
KILO	,001	,000	,017	3,997	,000	,687	1,456
GUNLUK_FAALİYETDURUM	-,025	,006	-,017	-4,224	,000	,772	1,295
YURUME_ZAMAN	-,012	,003	-,015	-3,923	,000	,857	1,166
BISIKLET_ZAMAN	-,025	,007	-,012	-3,294	,001	,945	1,058
SPOR_ZAMAN	,000	,000	-,014	-3,684	,000	,841	1,189
KASGUCLENDİRME	-,021	,005	-,017	-4,516	,000	,850	1,176
MEYVE_YEMESIKLIK	,018	,003	,023	5,470	,000	,721	1,387
SEBZE_SALATA_YEMESIKLIK	,014	,004	,014	3,503	,000	,744	1,345
TUTUNKULLANIM_DURUM	,004	,007	,005	,571	,568	,194	5,156
TUTUN_BASLAMANEDEN_1	,000	,000	-,006	-1,548	,122	,922	1,084
TUTUNKUL_DUZENLI_DURUM	-,001	,006	,000	-,105	,916	,826	1,210
TUTUNKUL_DUZENLI_YAS	,001	,000	,011	2,062	,039	,453	2,206
SIGARA_ADET	,001	,001	,013	2,153	,031	,319	3,130
S_BIRAKMAYONTEM	,002	,001	,009	2,591	,010	,968	1,033
S_BIRAKMADENEME	,003	,010	,002	,275	,783	,159	6,299

EK-1. (devam) SPSS paket programında elde edilen ve VIF değerleri neticesinde veri setinden çıkarılan/çıkarılmayan değişkenlere ilişkin tablo

TUTUNDUMAN_MARUZ	,009	,005	,007	1,881	,060	,878	1,138
ALKOLKULLANIM_DURUM	-,012	,055	-,006	-,215	,830	,015	66,408
ALKOL_KULLANMANEDEN_1	,000	,000	-,001	-,363	,716	,916	1,092
SON12AY_ALKOLKULLANIM_SIKLIK	-,003	,007	-,013	-,498	,619	,019	52,023
ALKOLKUL_DUZENLI_YAS	-,001	,001	-,003	-,466	,641	,371	2,694
SON12AY_6VEUZERI	-,005	,002	-,014	-2,169	,030	,306	3,268

a. Dependent Variable: GENELSAGLIKDURUM

EK-2. Değişkenler frekans yüzde dağılımı (%)

<i>Değişkenler Frekans Yüzde dağılımı (%)</i>			
	Değişkenler	Frekans	Yüzde dağılımı
YASAMFAAL_KISITLANMADURUM	Son 6 ay içinde herhangi bir sağlık problemi nedeniyle günlük yaşam faaliyetlerinizin kısıtlanma durumu nedir?		
	<i>Ciddi ölçüde kısıtlandı</i>	4335	12,4
	<i>Kısıtlandı fakat ciddi ölçüde değil</i>	9625	27,5
	<i>Kısıtlanmadı</i>	21035	60,1
HASTALIK_SAGLIKDURUM	6 ay ya da uzun süren /sürmesi beklenen hastalığınız/sağlık probleminiz var mı?		
	<i>Evet</i>	16960	48,5
	<i>Hayır</i>	18035	51,5
DESTEKSIZ12BASAMAK	12 basamak merdiven inip çıkmakta zorluk çekiyor musunuz?		
	<i>Hiç zorlanmıyorum</i>	25716	73,5
	<i>Biraz zorlanıyorum</i>	5817	16,6
	<i>Çok zorlanıyorum</i>	3081	8,8
	<i>Hiç merdiven inip çıkamıyorum</i>	381	1,1
BEDENSELAGRI	Son 4 hafta içerisinde ne kadar bedensel ağrınız oldu?		
	<i>Hiç</i>	16134	46,1
	<i>Çok az</i>	4909	14,0
	<i>Az</i>	4013	11,5
	<i>Orta</i>	5050	14,4
	<i>Fazla</i>	3548	10,1
	<i>Çok fazla</i>	1341	3,8
ZORLUKCEKME_AGIREVIS	Yardım almadan ağır ev işlerini yapmada zorluk çekiyor musunuz?		
	<i>Hiç zorlanmıyorum</i>	24202	69,2
	<i>Biraz zorlanıyorum</i>	4360	12,5
	<i>Çok zorlanıyorum</i>	2441	7,0
	<i>Kendi kendime yapamıyorum</i>	1931	5,5
	<i>Hiçbir zaman denemedim/Yapmaya ihtiyaç duymadım</i>	2061	5,9

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

YAS_SIRA	Görüşülen Kişinin Yaş Grubu		
	17-24	6152	17,6
	25-34	6450	18,4
	35-44	6891	19,7
	45-54	6064	17,3
	55-64	4713	13,5
	64-74	2915	8,3
	74+	1810	5,2
CALISMAMA_NEDEN	Çalışmama nedeniniz nedir?		
	Çalıştı veya çalışmadı fakat iş ile ilgisi devam ediyor	13685	39,1
	İş bulamama / İşsiz olup iş arıyor	1542	4,4
	Mevsimlik çalışıyor	70	0,2
	Eğitim ve öğretime devam ediyor	3074	8,8
	Ev işleri ile meşgul (çocuk, yaşlı, hasta, vb. kişilerin bakımı dâhil)	10587	30,3
	Emekli veya işi bırakmış	3967	11,3
	Engelli veya hasta (çalışamaz halde)	567	1,6
	Yaşlı (emekli değil, ancak çalışmak için yaşlı olduğunu düşünmüyor, 65+)	1207	3,4
	İrad sahibi	4	0,0
	Ailevi ve kişisel nedenler	163	0,5
	Diğer	129	0,4
EGITIM_SIRALI	En son hangi okulu bitirdiniz?		
	Okuma-Yazma bilmeyen	3433	9,8
	Bir okul bitirmedi	1841	5,3
	İlkokul	12548	35,9
	Genel ortaokul, Mesleki veya teknik ortaokul, İlköğretim	5972	17,1
	Lise veya mesleki lise	6223	17,8
	2 veya 3 yıllık yüksekokul	1635	4,7
	4 yıllık yüksekokul veya fakülte	2886	8,2
	Yüksek lisans (5 veya 6 yıllık fakülteler dâhil)	390	1,1
	Doktora	67	0,2

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

ZORLUKCEKME_HAFIFEVIS	Yardım almadan hafif ev işlerini yapmada zorluk çekiyor musunuz?		
	<i>Hiç zorlanmıyorum</i>	30503	87,2
	<i>Biraz zorlanıyorum</i>	1740	5,0
	<i>Çok zorlanıyorum</i>	598	1,7
	<i>Kendi kendime yapamıyorum</i>	733	2,1
	<i>Hiçbir zaman denemedim/Yapmaya ihtiyaç duymadım</i>	1421	4,1
OGRENME	Yaştlarınıza göre yeni bilgi beceri öğrenmede (okuma-yazma, hesaplama, basit dört işlem yapmada vb.) güçlük çekiyor musunuz?		
	<i>Zorluk çekmiyor</i>	29551	84,4
	<i>Biraz zorlanıyor</i>	3575	10,2
	<i>Çok zorlanıyor</i>	1144	3,3
	<i>Hiç yapamıyor</i>	725	2,1
KRONIKHAST_HIPERTANSİYON	Son 12 ay içerisinde hipertansiyon yaşadınız mı?		
	<i>Evet</i>	6548	18,7
	<i>Hayır</i>	28447	81,3
ZORLUKCEKME_IDARIMALIKONU	Yardım almadan günlük sorumluluklar ve mali konularla ilgilenmede zorluk çekiyor musunuz?		
	<i>Hiç zorlanmıyorum</i>	30992	88,6
	<i>Biraz zorlanıyorum</i>	1887	5,4
	<i>Çok zorlanıyorum</i>	569	1,6
	<i>Kendi kendime yapamıyorum</i>	767	2,2
	<i>Hiçbir zaman denemedim/Yapmaya ihtiyaç duymadım</i>	780	2,2
HATIRLAMA	Yakın zamanda yaşanan olayları (gün içindeki olaylar, yakınlarınızın adları gibi) hatırlamakta güçlük çekiyor musunuz?		
	<i>Zorluk çekmiyor</i>	27375	78,2
	<i>Biraz zorlanıyor</i>	6253	17,9
	<i>Çok zorlanıyor</i>	1268	3,6
	<i>Hiç yapamıyor</i>	99	0,3

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

RECETEEDİLEN_ILACKULLANIM	Son 2 hafta içerisinde bir doktor tarafından size reçete edilen bir ilaç kullandınız mı?		
	<i>Evet</i>	12861	36,8
	<i>Hayır</i>	22134	63,2
YORGUNLUK	Son 2 hafta içerisinde yorgun/enerjisiz hissetme ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	26289	75,1
	<i>Bazı günler</i>	7016	20,0
	<i>Bir haftadan fazla</i>	506	1,4
	<i>Neredeyse her gün</i>	1184	3,4
KRONIKHAST_BELBOLGE	Son 12 ay içerisinde bel bölgesi problemleri (bel ağrısı, bel fıtığı ve diğer bel defektleri) yaşadınız mı?		
	<i>Evet</i>	11414	32,6
	<i>Hayır</i>	23581	67,4
MEDENI_DURUM	Medeni durumunuz nedir?		
	<i>Hiç evlenmedi</i>	7511	21,5
	<i>Evli</i>	24079	68,8
	<i>Boşandı</i>	994	2,8
	<i>Eşi öldü</i>	2411	6,9
HUZURSUZLUK	Son 2 hafta içerisinde başkalarının fark edeceği derecede ağır hareket etme/konuşma ya da huzursuz/ fazla hareket etme ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	30083	86,0
	<i>Bazı günler</i>	4137	11,8
	<i>Bir haftadan fazla</i>	245	0,7
	<i>Neredeyse her gün</i>	530	1,5
UYKUSORUN	Son 2 hafta içerisinde uykuya dalmakta veya uyumakta zorluk çekme / çok fazla uyuma ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	22626	64,7
	<i>Bazı günler</i>	9299	26,6
	<i>Bir haftadan fazla</i>	892	2,5
	<i>Neredeyse her gün</i>	2178	6,2

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

BUNALIM	Son 2 hafta içerisinde morali bozuk, bunalımlı veya umutsuz hissetme ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	19583	56,0
	<i>Bazı günler</i>	12917	36,9
	<i>Bir haftadan fazla</i>	874	2,5
	<i>Neredeyse her gün</i>	1621	4,6
DEGERSIZHISSETME	Son 2 hafta içerisinde kendini değersiz ve kötü hissetme ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	26703	76,3
	<i>Bazı günler</i>	7035	20,1
	<i>Bir haftadan fazla</i>	381	1,1
	<i>Neredeyse her gün</i>	876	2,5
KRONIKHAST_SEKERHAST	Son 12 ay içerisinde şeker hastalığı (diyabet) yaşadınız mı?		
	<i>Evet</i>	3718	10,6
	<i>Hayır</i>	31277	89,4
KRONIKHAST_BOYUNBOLGE	Son 12 ay içerisinde boyun bölgesi problemleri (boyun ağrısı, boyun fıtığı, diğer boyun defektleri) yaşadınız mı?		
	<i>Evet</i>	7606	21,7
	<i>Hayır</i>	27389	78,3
KONSANTRASYONPRB	Son 2 hafta içerisinde konsantrasyon problemi (gazete okurken, TV izlerken vb.) ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	27885	79,7
	<i>Bazı günler</i>	6020	17,2
	<i>Bir haftadan fazla</i>	370	1,1
	<i>Neredeyse her gün</i>	720	2,1
ISTAHSIZLIK	Son 2 hafta içerisinde iştahsızlık/aşırı yemek yeme ne sıklıkla sizi rahatsız etti?		
	<i>Hiç</i>	26289	75,1
	<i>Bazı günler</i>	7016	20,0
	<i>Bir haftadan fazla</i>	506	1,4
	<i>Neredeyse her gün</i>	1184	3,4
YURUME_ZAMAN	Normal bir günde, bir yerden bir yere giderken yürüyüş için ne kadar zaman harcarsınız?		
	-	7713	22,0
	<i>Günde 10-29 dakika</i>	14640	41,8
	<i>Günde 30-59 dakika</i>	8095	23,1
	<i>Günde 1 saatten fazla, 2 saatten az</i>	2984	8,5
	<i>Günde 2 saatten fazla, 3 saatten az</i>	762	2,2
	<i>Günde 3 saat veya daha fazla</i>	801	2,3

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

HANE_GELIR	Hanenin ortalama aylık geliri aşağıdaki gelir gruplarından hangisinde yer almaktadır?		
2014 - 2016	2014 (0-1080 TL) - 2016 (0-1264 TL)	9313	26,6
	2014 (1081-1550 TL) - 2016 (1265-1814 TL))	8213	23,5
	2014 (1551-2170 TL) - 2016 (1815-2540 TL)	5990	17,1
	2014 (2171-3180 TL) - 2016 (2541-3721 TL)	5982	17,1
	2014 (3181 + TL) - 2016 (3722 + TL)	5497	15,7
ISCO_GRUP	İşyerinizdeki yaptığınız iş(görev ve sorumluluğunuz) nedir/neydi?		
	-	12114	34,6
	Silahlı kuvvetlerle ilgili meslekler	103	0,3
	Yöneticiler	1120	3,2
	Profesyonel meslek mensupları	2591	7,4
	Teknisyenler, teknikerler ve yardımcı profesyonel meslek mensupları	1454	4,2
	Büro hizmetlerinde çalışan elemanlar	1473	4,2
	Hizmet ve satış elemanları	4065	11,6
	Nitelikli tarım, ormancılık ve su ürünleri çalışanları	3187	9,1
	Sanatkârlar ve ilgili işlerde çalışanlar	3192	9,1
	Tesis ve makine operatörleri ve montajcılar	1966	5,6
	Nitelik gerektirmeyen işlerde çalışanlar	3730	10,7

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

NACE_GRUP	İşyerinin ana faaliyeti?		
	-	12114	34,62
	Tarım, ormancılık ve balıkçılık	4556	13,02
	Madencilik ve taş ocakçılığı	158	0,45
	İmalat	4767	13,62
	Elektrik, gaz, buhar ve iklimlendirme üretimi ve dağıtımı	139	0,40
	Su temini; kanalizasyon, atık yönetimi ve iyileştirme faaliyetleri	62	0,18
	İnşaat	1407	4,02
	Toptan ve perakende ticaret; motorlu kara taşıtlarının ve motosikletlerin onarımı	2877	8,22
	Ulaştırma ve depolama	974	2,78
	Konaklama ve yiyecek hizmeti faaliyetleri	1214	3,47
	Bilgi ve iletişim	226	0,65
	Finans ve sigorta faaliyetleri	353	1,01
	Gayrimenkul faaliyetleri	153	0,44
	Mesleki, bilimsel ve teknik faaliyetler	615	1,76
	İdari ve destek hizmeti faaliyetleri	812	2,32
	Kamu yönetimi ve savunma; zorunlu sosyal güvenlik	1677	4,79
	Eğitim	1271	3,63
	İnsan sağlığı ve sosyal hizmet faaliyetleri	848	2,42
	Kültür, sanat, eğlence, dinlence ve spor	95	0,27
	Diğer hizmet faaliyetleri	502	1,43
	Hane halklarının işverenler olarak faaliyetleri; hanehalklarını tarafından kendi kullarımlarına yönelik olarak ayırım yapılmamış mal ve hizmet üretim faaliyetleri	167	0,48
	Uluslararası örgütler ve temsilciliklerinin faaliyetleri	8	0,02
GUNLUK_FAALIYETDURUM	Çalışırken/günlük faaliyetlerinizi gerçekleştirirken, aşağıdakilerden hangisi durumunuzu en iyi tanımlar?		
	Çoğunlukla oturan veya duran	14381	41,1
	Çoğunlukla yürüyüş veya orta derecede fiziksel güç gerektiren işler	18331	52,4
	Çoğunlukla ağır iş veya fiziksel güç gerektiren işler	2283	6,5
UZMAN_HEK_HIZALMADURUM	Kendiniz için, en son ne zaman hekimden hizmet aldınız?		
	12 aydan az süre önce	23695	67,7
	12 ay ve daha fazla süre önce	9219	26,3
	Hiçbir zaman	2081	5,9

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

GOZLUKKULLANMA	Gözlük ya da lens kullanıyor musunuz?		
	<i>Evet</i>	13780	39,4
	<i>Hayır</i>	21178	60,5
	<i>Hiç göremiyorum</i>	37	0,1
ODEMEGUCLUGU_TIBBIBAKIM	Son 12 ay içerisinde, tıbbi bakıma ihtiyacınız olduğu halde ödeme gücünüz nedeniyle karşılayamadığınız oldu mu?		
	<i>Evet</i>	3665	10,5
	<i>Hayır</i>	27059	77,3
	<i>İhtiyacım olmadı</i>	4271	12,2
ODEMEGUCLUGU_ILAC	Son 12 ay içerisinde, reçete edilen ilaca ihtiyacınız olduğu halde ödeme gücünüz nedeniyle karşılayamadığınız oldu mu?		
	<i>Evet</i>	2763	7,9
	<i>Hayır</i>	28703	82,0
	<i>İhtiyacım olmadı</i>	3529	10,1
ISTEKI_DURUM	Şu anda veya en son çalıştığınız işyerinde işteki durumunuz nedir/neydi?		
	-	12114	34,6
	<i>Ücretli, maaşlı veya yevmiyeli</i>	16634	47,5
	<i>İşveren</i>	706	2,0
	<i>Kendi hesabına</i>	3234	9,2
	<i>Ücretsiz aile işçisi</i>	2307	6,6
CINSİYET	Cinsiyetiniz nedir?		
	<i>Erkek</i>	15904	45,4
	<i>Kadın</i>	19091	54,6
KRONIKHAST_ALERJİ	Son 12 ay içerisinde alerji(alerjik rinit, dermatit, yiyecek vb. alerjisi, alerjik astım hariç) yaşadınız mı?		
	<i>Evet</i>	4320	12,3
	<i>Hayır</i>	30675	87,7
ODEMEGUCLUGU_RUHTEDV	Son 12 ay içerisinde, ruhsal tedaviye (Psikiyatrist vb. uzmanlar tarafından) ihtiyacınız olduğu halde ödeme gücünüz nedeniyle karşılayamadığınız oldu mu?		
	<i>Evet</i>	1063	3,0
	<i>Hayır</i>	22339	63,8
	<i>İhtiyacım olmadı</i>	11593	33,1

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

ODEMEGUCLUGU_DISBAKIM	Son 12 ay içerisinde, dış bakımına ihtiyacınız olduğu halde karşılayamadığınız oldu mu?		
	<i>Evet</i>	4087	11,7
	<i>Hayır</i>	24801	70,9
	<i>İhtiyacım olmadı</i>	6107	17,5
SON12AY_6VEUZERI	Son 12 ay içerisinde, (herhangi bir fırsatta) ne sıklıkta 6 adet veya daha fazla alkollü içecek içtiniz? (Örneğin; partide, yemekte, evde tek başına iken)		
	-	30297	86,6
	<i>Her gün/Hemen hemen her gün</i>	58	0,2
	<i>Haftada 5-6 gün</i>	24	0,1
	<i>Haftada 3-4 gün</i>	44	0,1
	<i>Haftada 1-2 gün</i>	190	0,5
	<i>Ayda 2-3 gün</i>	305	0,9
	<i>Ayda bir kere</i>	510	1,5
	<i>Ayda bir kereden az</i>	1748	5,0
	<i>Son 12 ay içerisinde içmedim</i>	771	2,2
	<i>Hayatım boyunca hiçbir zaman 6 adet veya daha fazla içmedim</i>	1048	3,0
TEDAVI_MASRAF_GSS	Tedavi masraflarınız GSS tarafından mı karşılanıyor?		
	<i>Evet</i>	3677	10,5
	<i>Hayır</i>	31318	89,5
SAGKURUM_TERCIH	En son gittiğiniz sağlık kuruluşunu tercih etme sebebiniz nedir?		
	<i>Sosyal güvencesi nedeniyle</i>	13951	39,9
	<i>Ucuz olduğu için</i>	1032	2,9
	<i>Yakın olduğu için</i>	8915	25,5
	<i>Başka seçeneği yok</i>	2196	6,3
	<i>İyi davrandıkları için</i>	727	2,1
	<i>Güvenilir olduğu için</i>	2196	6,3
	<i>Bekletilmediği için</i>	577	1,6
	<i>Daha iyi sağlık hizmeti verdiği için</i>	2704	7,7
	<i>Tanıdık personel olduğu için</i>	880	2,5
	<i>Sevk edildiği için</i>	279	0,8
	<i>Tavsiye edildiği için</i>	766	2,2
	<i>Rahatsızlığı ciddi olmadığı için</i>	235	0,7
	<i>İlaç temini için</i>	332	0,9
	<i>Diğer</i>	205	0,6

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

TUTUNKULLANIM_DURUM	Tütün mamulü kullanıyor musunuz?		
	<i>Evet, her gün</i>	9004	25,7
	<i>Evet, ara sıra</i>	1660	4,7
	<i>Hiçbir zaman</i>	18494	52,8
	<i>Bıraktım</i>	5837	16,7
TUTUN_BASLAMANEDEN	Tütün mamulü kullanmaya başlama nedeniniz nedir?		
	-	18840	53,8
	<i>Merak</i>	4851	13,9
	<i>Özenti</i>	3601	10,3
	<i>Aile sorunları</i>	725	2,1
	<i>Kişisel sorunlar</i>	926	2,6
	<i>Arkadaş etkisi</i>	4701	13,4
	<i>Eğlence amaçlı</i>	480	1,4
	<i>Hiçbir özel neden yok</i>	468	1,3
	<i>Diğer(Belirtiniz)</i>	403	1,2
DIS_HEK_HIZALMAZAMAN	Kendiniz için en son ne zaman bir dış hekiminden hizmet aldınız?		
	<i>Son 6 ay içinde</i>	7259	20,7
	<i>6 aydan fazla, 12 aydan az süre önce</i>	4406	12,6
	<i>12 ay ve daha fazla süre önce</i>	19437	55,5
	<i>Hiçbir zaman</i>	3893	11,1
TEDAVI_MASRAF_SGK	Tedavi masraflarınız SGK tarafından mı karşılanıyor?		
	<i>Evet</i>	29018	82,9
	<i>Hayır</i>	5977	17,1
KILO	Elbise ve ayakkabılarınız olmadan vücut ağırlığınız (kg) ne kadardır?		
	<i>Sürekli değişken</i>		
ALKOL_KULLANMANEDEN	Alkol kullanmanızın ya da kullanmayı denemenizin en önemli nedeni nedir?		
	-	28036	80,1
	<i>Merak</i>	2501	7,1
	<i>Özenti</i>	595	1,7
	<i>Aile sorunları</i>	72	0,2
	<i>Kişisel sorunlar</i>	128	0,4
	<i>Arkadaş etkisi</i>	1657	4,7
	<i>Eğlence amaçlı</i>	1555	4,4
	<i>Hiçbir özel neden yok</i>	170	0,5
	<i>Diğer(Belirtiniz)</i>	281	0,8

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

S_BIRAKMAYONTEM	Sigarayı bırakmak için en son hangi yöntemi kullandınız?		
	-	25365	72,5
	<i>Kendi kendime bıraktım</i>	8950	25,6
	<i>Alo 171 Sigara Bırakma Danışma hattını arayarak bıraktım.</i>	112	0,3
	<i>Hekim yardımıyla bıraktım.</i>	231	0,7
	<i>Hekim yardımı ve ilaçla bıraktım.</i>	114	0,3
	<i>İlaçla bıraktım</i>	147	0,4
	<i>Diğer</i>	76	0,2
SPOR_ZAMAN_SAAT	Normal bir hafta içerisinde spor, fitness veya boş zaman aktivitesi gerçekleştirirken toplamda ne kadar zaman harcarsınız?		
	<i>Sürekli değişken</i>		
MEYVE_YEMESIKLIK	Ne sıklıkta meyve yersiniz?		
	<i>Günde bir kere ya da daha fazla</i>	17851	51,0
	<i>Haftada 4-6 kere</i>	5899	16,9
	<i>Haftada 1-3 kere</i>	8236	23,5
	<i>Haftada bir kereden az</i>	2319	6,6
	<i>Hiç</i>	690	2,0
GRIPASISOLMA_ZAMAN	En son ne zaman grip aşısı oldunuz?		
	<i>Son bir yıl içinde</i>	1179	3,4
	<i>Uzun zaman önce (Geçen yıldan önce)</i>	4732	13,5
	<i>Hiçbir zaman</i>	29084	83,1
SEBZE_SALATA_YEMESIKLIK	Ne sıklıkta sebze ya da salata yersiniz(Patates hariç)?		
	<i>Günde bir kere ya da daha fazla</i>	21532	61,5
	<i>Haftada 4-6 kere</i>	6751	19,3
	<i>Haftada 1-3 kere</i>	5543	15,8
	<i>Haftada bir kereden az</i>	929	2,7
	<i>Hiç</i>	240	0,7
KENDI_ILAC	Son 2 hafta içerisinde kendi kararınız/tercihinizle ilaç kullandınız mı?		
	<i>Evet</i>	10486	30,0
	<i>Hayır</i>	24509	70,0

EK-2. (devam) Değişkenler frekans yüzde dağılımı (%)

SIGARA_ADET	Ortalama olarak her gün kaç adet sigara kullanıyorsunuz? <i>Sürekli değişken</i>		
TUTUNKUL_DUZENLI_DURUM	Düzenli olarak (yılda en az 100 adet olmak üzere) hiç tütün mamulü kullandınız mı? - <i>Evet</i> <i>Hayır</i>	29730 3016 2249	84,955 8,618 6,427
ALKOLKUL_DUZENLI_YAS	Düzenli olarak kaç yaşında alkollü içecek kullanmaya başladınız? <i>Sürekli değişken</i>		
TUTUNDUMAN_MARUZ	Kapalı alanda, tütün dumanına hangi sıklıkta maruz kalıyorsunuz? <i>Hiç/Hemen hemen hiç</i> <i>Günde bir saatten az</i> <i>Günde bir saatten fazla</i>	26118 4568 4309	74,6 13,1 12,3
TUTUNKUL_DUZENLI_YAS	Düzenli olarak kaç yaşında tütün mamulü kullanmaya başladınız? <i>Sürekli değişken</i>		
TEDAVI_MASRAFI_KENDI	Tedavi masraflarınızı kendiniz mi karşılıyorsunuz? <i>Evet</i> <i>Hayır</i>	7589 27406	21,7 78,3

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : YAMANTAŞ, Özge
 Uyruğu : T.C.
 Doğum tarihi ve yeri : 08.03.1991, Ankara
 Medeni hali : Evli
 Telefon : + 534 625 45 35
 e-mail : ozge.erkilic1@gazi.edu.tr



Eğitim

Yüksek lisans
 Lisans
 Lise

Eğitim Birimi

Gazi Üniversitesi / İstatistik
 Ankara Üniversitesi / İstatistik
 Uluğbey Anadolu Lisesi

Mezuniyet Tarihi

Devam Ediyor
 2013
 2009

İş Deneyimi

Yıl

2016-Halen

Yer

Türkiye İstatistik Kurumu

Görev

TÜİK Uzm. Yrd.

Yabancı Dil

İngilizce

Yayınlar

1. Yamantaş, Ö. (2018, 27-29 Haziran). *Modelling of Turkey health survey data with artificial neural networks*. International Conference on Mathematics and Mathematics Education, Ordu.

Hobiler

Tenis, Kayak



GAZİ GELECEKTİR..