

ZONGULDAK BÜLENT ECEVİT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SEÇİLEN SİĞİL TEDAVİ YÖNTEMİNİN FARKLI MAKİNE ÖĞRENMESİ
ALGORİTMALARI İLE BAŞARIMININ TAHMİN EDİLMESİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

MUALLA SAKARYA

OCAK 2019

ZONGULDAK BÜLENT ECEVİT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

SEÇİLEN SİĞİL TEDAVİ YÖNTEMİNİN FARKLI MAKİNE ÖĞRENMESİ
ALGORİTMALARI İLE BAŞARIMININ TAHMİN EDİLMESİ

ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

Mualla SAKARYA

DANIŞMAN: Dr. Öğr. Üyesi Rukiye UZUN

İKİNCİ DANIŞMAN: Dr. Öğr. Üyesi YALÇIN İŞLER

ZONGULDAK

Ocak 2019

KABUL:

Mualla SAKARYA tarafından hazırlanan “Seçilen Siğil Tedavi Yönteminin Farklı Makine Öğrenmesi Algoritmaları İle Başarımının Tahmin Edilmesi” başlıklı bu çalışma jürimiz tarafından değerlendirilerek Zonguldak Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalında Yüksek Lisans Tezi olarak oybirliğiyle kabul edilmiştir.18 / 01 / 2019

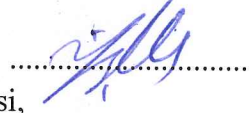
Danışman: Dr. Öğr. Üyesi Rukiye UZUN

Zonguldak Bülent Ecevit Üniversitesi, Mühendislik Fakültesi,
Elektrik-Elektronik Mühendisliği Bölümü



İkinci Danışman: Dr. Öğr. Üyesi Yalçın İŞLER

İzmir Katip Çelebi Üniversitesi, Mühendislik-Mimarlık Fakültesi,
Biyomedikal Mühendisliği Bölümü



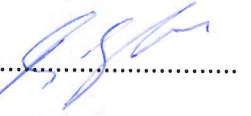
Üye: Dr. Öğr. Üyesi Hakan KAYA

Zonguldak Bülent Ecevit Üniversitesi, Mühendislik Fakültesi,
Elektrik-Elektronik Mühendisliği Bölümü



Üye: Dr. Öğr. Üyesi Okan ERKAYMAZ

Zonguldak Bülent Ecevit Üniversitesi, Mühendislik Fakültesi,
Bilgisayar Mühendisliği Bölümü



Üye: Dr. Öğr. Üyesi Mustafa YENİAD

Ankara Yıldırım Beyazıt Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi,
Bilgisayar Mühendisliği Bölümü



ONAY:

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım. / / 20....



Prof. Dr. Ahmet ÖZARSLAN
Fen Bilimleri Enstitüsü Müdürü

“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”



Mualla SAKARYA

ÖZET

Yüksek Lisans Tezi

SEÇİLEN SİĞİL TEDAVİ YÖNTEMİNİN FARKLI MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE BAŞARIMININ TAHMİN EDİLMESİ

Mualla SAKARYA

Zonguldak Bülent Ecevit Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Rukiye UZUN

İkinci Danışman: Dr. Öğr. Üyesi Yalçın İŞLER

Ocak 2019, 39 sayfa

Siğiller virüs temelli olup en sık karşılaşılan dermatolojik sorunların başında gelmektedir. Siğil tedavisinde çeşitli tedavi yöntemleri geliştirilmiştir. Son zamanlarda yaygın ve ayak taban siğili olan hastalara kriyoterapi ve immünoterapi tedavi yöntemleri uygulanmaya başlanmıştır. Bununla birlikte hangi tedavi yönteminin başarılı olacağına dair bir kanıt bulunmamaktadır.

Bu çalışmada, bu iki yöntemin siğil tedavisinde başarılı olup olmayacağı literatürde yaygın kullanıma sahip makine öğrenmesi yöntemleri ile tedavi uygulanmadan önce tahmin edilmeye çalışılmıştır. Bu algoritmaların girişlerine açık erişime sahip olan UCI veri tabanındaki veriler uygulanarak algoritmalar koşturulmuştur. Bu veriler yaygın ve / veya ayak tabanı siğil türlerine sahip 180 hastaya ait iki veri setinin birleştirilmesiyle elde edilmiştir. Hastaların yarısına kriyoterapi ve diğer yarısına immünoterapi yöntemleri uygulanmıştır.

ÖZET (devam ediyor)

Sonuç olarak, seçilen siğil tedavi yönteminin başarısı duyarlık, belirleyicilik ve genel başarımlar ölçütlerine göre belirlenmiştir. Bu değerler kullanılan yöntemlere göre sırasıyla Naive Bayes yöntemi için %68,43, %67,61 ve %67,78, Lojistik Regresyon yöntemi için %26,32, %94,37 ve %80,00, Karar Ağacı yöntemi için %52,63, %94,37 ve %85,56, 7-Yakın Komşuluk yöntemi için %15,79, %97,19 ve %80,00, Destek Vektör Makineleri yöntemi için %47,37, %95,78 ve %85,46, Aşırı Öğrenme Makinesi yöntemi için %36,85, %90,15 ve %78,89, Çok Katmanlı Algılayıcı yöntemi için %78,95, %98,60 ve %94,45 olarak tespit edilmiştir. Bu değerler içinde en yüksek genel başarımları veren Çok Katmanlı Algılayıcı yöntemi ile literatürdeki mevcut az sayıdaki çalışmadan çok daha yüksek bir doğrulukla tercih edilen siğil tedavi yönteminin başarılı olup olamayacağını tahmin edilebilmiştir.

Anahtar Kelimeler: Siğil, kriyoterapi, imünoterapi, makine öğrenmesi, naive Bayes, lojistik regresyon, karar ağacı, k-yakın komşu, destek vektör makineleleri, aşırı öğrenme makinesi, çok katmanlı algılayıcı.

Bilim Kodu: 608.03.04

ABSTRACT

M.Sc. Thesis

PREDICTION OF THE SUCCESS OF THE SELECTED WART TREATMENT METHOD USING VARIOUS MACHINE LEARNING ALGORITHMS

Mualla SAKARYA

**Zonguldak Bülent Ecevit University
Graduate School of Natural and Applied Sciences
Department of Electrical and Electronics Engineering**

Thesis Advisor: Assist. Prof. Dr. Rukiye UZUN

Co-Advisor: Assist. Prof. Dr. Yalçın İŞLER

January 2019, 39 pages

Warts, which are virus-based, are the most common dermatoses in society. Several treatment methods have been developed in the wart treatment. Recently, treatment methods of cryotherapy and immunotherapy have been got off the ground to the patients with common and plantar warts. On the other hand, there is no proof of which treatment method will be successful, yet.

In this study, it was predicted if these two methods to be applied in the treatment of warts will success or not before starting the treatment using commonly-used machine learning algorithms. Algorithms were run by applying the two online and freely available UCI data sets to the inputs. These data were combined from two data sets from 180 patients with common and / or plantar warts. The cryotherapy was applied to the half of patients and the immunotherapy was applied to the other half.

ABSTRACT (continued)

As a result, the success of the selected wart treatment method was predicted with the criteria of sensitivity, specificity, and accuracy. These values were obtained for the corresponding algorithms used, respectively: 68.43, 67.61 and 67.78 for Naive Bayes, 26.32%, 94.37% and 80.00% for Logistic Regression, 52.63%, 94.37% and 85.56% for Decision Tree, 15.79%, 97.19% and 80.00% for 7-Nearest Neighbors, 47.37%, 95.78% and 85.46% for Support Vector Machines, 36.85%, 90.15% and 78.89% for Extreme Learning Machines, and 78.95%, 98.60% and 94.45% for Multi-Layer Perceptron. Multi-Layer Perceptron, which is resulted in the highest general accuracy among these methods, can predict whether the selected wart treatment will succeed or not more accurately than the few studies presented in the literature.

Keywords: Wart, cryotherapy, immunotherapy, machine learning, naive Bayes, decision tree, logistic regression, k-nearest neighbors, support vector machines, extreme learning machines, multi-layer perceptron.

Science Code: 608.03.04

TEŐEKKÜR

Çalıőmamın her anında verdikleri katkılardan ve desteklerden ötürü tez çalıőmamın danıőmanı olan Dr. Öđr. Üyesi Rukiye UZUN'a ve bana yol gösteren, her konuda bana yardımcı olan tez çalıőmamın ikinci danıőmanı Dr. Öđr. Üyesi Yalçın İŐLER'e verdiđi katkı ve desteklerinden dolayı teőekkür ederim.

Son olarak, ama asla daha az deđil, beni koőulsuz destekleyen, hayatım boyunca yanımda olan, sabırla eđiten ve büyüten anneme ve babama teőekkürü bir borç bilirim.



İÇİNDEKİLER

| | <u>Sayfa</u> |
|--|--------------|
| KABUL | ii |
| ÖZET | iii |
| ABSTRACT | v |
| TEŞEKKÜR | vii |
| İÇİNDEKİLER..... | ix |
| ŞEKİLLER DİZİNİ..... | xi |
| ÇİZELGELER DİZİNİ | xiii |
| SİMGELER VE KISALTMALAR DİZİNİ..... | xv |
| | |
| BÖLÜM 1 GİRİŞ | 1 |
| | |
| 1.1. SİĞİL | 1 |
| 1.2. SİĞİL ÇEŞİTLERİ | 2 |
| 1.2.1. Yaygın Siğiller | 3 |
| 1.2.2. Ayak Tabanı Siğilleri | 3 |
| 1.2.3. Düz Siğiller | 4 |
| 1.2.4. Etek Siğilleri | 5 |
| 1.3. SİĞİL TEDAVİSİ..... | 5 |
| 1.4. MAKİNE ÖĞRENMESİ..... | 7 |
| 1.5. AMAÇ | 8 |
| | |
| BÖLÜM 2 YÖNTEM | 11 |
| | |
| 2.1. VERİ SETİ | 11 |
| 2.2. MAKİNE ÖĞRENMESİ ALGORİTMALARI..... | 12 |
| 2.2.1. Naive Bayes | 13 |
| 2.2.2. Lojistik Regresyon | 13 |

İÇİNDEKİLER (devam ediyor)

| | <u>Sayfa</u> |
|--|--------------|
| 2.2.3. Karar Ağacı..... | 14 |
| 2.2.4. En Yakın K Komşu..... | 15 |
| 2.2.5. Destek Vektör Makineleri..... | 16 |
| 2.2.6. Aşırı Öğrenme Makinesi..... | 17 |
| 2.2.7. Çok Katmanlı Algılayıcı..... | 17 |
| 2.3. ALGORİTMALARIN DEĞERLENDİRİLMESİ..... | 18 |
| 2.3.1. Çapraz Doğrulama..... | 19 |
| 2.4. WEKA YAZILIMI..... | 19 |
| | |
| BÖLÜM 3 SONUÇLAR..... | 23 |
| | |
| 3.1. VERİLERİN HAZIRLANMASI..... | 23 |
| 3.2. SINIFLANDIRMA BAŞARIMLARININ DEĞERLENDİRİLMESİ..... | 24 |
| 3.3. SINIFLANDIRMA ALGORİTMALARI..... | 24 |
| 3.2.1. Naive Bayes..... | 25 |
| 3.2.2. Lojistik Regresyon..... | 25 |
| 3.2.3. Karar Ağacı..... | 26 |
| 3.2.4. En Yakın K Komşu..... | 26 |
| 3.2.5. Destek Vektör Makineleri..... | 27 |
| 3.2.6. Aşırı Öğrenme Makinesi..... | 27 |
| 3.2.7. Çok Katmanlı Algılayıcı..... | 28 |
| | |
| BÖLÜM 4 TARTIŞMA..... | 31 |
| | |
| KAYNAKLAR..... | 33 |
| ÖZGEÇMİŞ..... | 39 |

ŞEKİLLER DİZİNİ

| <u>No</u> | <u>Sayfa</u> |
|---|--------------|
| Şekil 1.1 Toplumda sık karşılaşılan siğil çeşitlerinden Yaygın siğil. | 4 |
| Şekil 1.2 Toplumda sık karşılaşılan siğil çeşitlerinden Ayak tabanı siğili..... | 4 |
| Şekil 1.3 Kriyoterapi tedavisi uygulanişı. | 7 |
| Şekil 2.1 X ve Y nitelikleri üzerine uygulanan testleri içere basit bir karar ağacı..... | 16 |
| Şekil 2.2 Destek Vektör Makineleri destek vektörleri ve ayırım düzlemi. | 17 |
| Şekil 2.3 Çok katmanlı algılayıcı ağ yapısı için tek nöron yapısı. | 19 |
| Şekil 2.4 WEKA kullanıcı ara yüzü. | 22 |
| Şekil 3.1 Bu çalışmada gerçekleştirilen sınıflandırıcı giriş ve çıkışları. | 23 |
| Şekil 3.2 Aşırı Öğrenme Makinesi yapısındaki yapar sinir ağları modeli ile siğil tedavi yönteminin başarı olup olmayacağıının belirlenmesi için sınıflandırıcı performansları..... | 25 |
| Şekil 3.3 Çok Katmanlı Algılayıcı yapısındaki yapar sinir ağları modeli ile siğil tedavi yönteminin başarı olup olmayacağıının belirlenmesi için sınıflandırıcı performansları..... | 29 |



ÇİZELGELER DİZİNİ

| <u>No</u> | <u>Sayfa</u> |
|--|--------------|
| Çizelge 2.1 Kriyoterapi tedavi yönteminde kullanılan öznitelikler. | 11 |
| Çizelge 2.2 Kriyoterapi tedavi yönteminde kullanılan öznitelikler. | 12 |
| Çizelge 3.1 Lojistik Regresyon algoritması ile siğil tedavi yönteminin başarı olup olmayacağıın belirlenmesi için sınıflandırıcı performansları..... | 25 |
| Çizelge 3.2 Karar Ağacı algoritması ile siğil tedavi yönteminin başarı olup olmayacağıın belirlenmesi için sınıflandırıcı performansları. | 26 |
| Çizelge 3.3 K-Yakın Komşu algoritması ile siğil tedavi yönteminin başarı olup olmayacağıın belirlenmesi için sınıflandırıcı performansları..... | 26 |
| Çizelge 3.4 Destek Vektör Makineleri algoritması ile siğil tedavi yönteminin başarı olup olmayacağıın belirlenmesi için sınıflandırıcı performansları..... | 27 |
| Çizelge 4.1 Seçilen siğil tedavi yönteminin başarılı olup olmayacağıın belirlenmesi için kullanılan sınıflandırıcı performansları. | 31 |



SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

- γ : Destek vektör makinelerinde hata payı parametresi
P : Gözlemlenen olayın görülme sıklığı
c : Destek vektör makinelerindeki genişlik parametresi

KISALTMALAR

- ACC** : Doğruluk (Accuracy)
BEÜ : Bülent Ecevit Üniversitesi
ELM : Aşırı Öğrenme Makinesi (Extreme Learning Machine)
FBE : Fen Bilimleri Enstitüsü
FN : Yanlış Negatif (False Negative)
FP : Yanlış Pozitif (False Positive)
HPV : İnsan Papilloma Virüsü
KA : Karar Ağacı (Decision Tree)
KNN : En Yakın K Komşu (K-Nearest Neighbors)
MLP : Çok Katmanlı Algılayıcı (Multi-Layer Perceptron)
SD : Standart sapma (Standard deviation)
SEN : Duyarlılık (Sensitivity)
SPE : Belirleyicilik (Specificity)
SVM : Destek Vektör Makineleri (Support Vector Machines)
TN : Doğru Negatif (True Negative)
TP : Doğru Pozitif (True Positive)



BÖLÜM 1

GİRİŞ

1.1. SİĞİL

Siğil; insan papilloma virüsü (HPV) kaynaklı bir cilt rahatsızlığıdır. Oldukça bulaşıcı bir rahatsızlık olup, doğrudan temas edilmesi veya virüsün bulaştığı nesnelere temas edilmesiyle dolaylı olarak bulaşabilmektedir (Çarman vd. 2013, James et al. 2006). Siğiller özellikle eller, el parmakları ve ayaklar başta olmak üzere insan vücudunun her hangi bir noktasında görülebilir (Aytekin 1994, Öcal 2009, Hoşrik 2010, Sever 2005). Siğiller çoğunlukla ciltle aynı renkte olup sert ve pürüzlü bir görünüme sahiptirler. Bazı siğiller daha koyu renkte (kahverengi veya gri-siyah), düz ve yumuşak bir görünüme sahiptirler. Mesela ayak tabanında bulunan siğiller düz görünümlü olup sıklıkla bir nasır ile karıştırılabilmektedir. Bu gibi siğillere sahip hastalar yürüdükçe, siğil alttaki dokuya doğru itildiği için ağrı oluşmaktadır.

En sık görüldüğü yaş aralığı dikkate alındığında, özellikle genç yetişkinler başta olmak üzere her yaş grubunda rastlanabilmektedir (Aytekin 1994, Öcal 2009, Hoşrik 2010, Sever 2005).

HPV insanlarda hücre içerisine yerleşmekte ve çeşitli enfeksiyonların oluşmasına neden olmaktadır. Virüsün klinik geçerliği tam olarak ispatlanmış bir tedavisi olmayıp kolay bulaşıcı bir yapısı bulunmaktadır. Sağlıklı deri dokusu üzerinden siğillerin insanlara bulaşması çok zordur. Öte yandan, açık yaralardan ve cinsel ilişki yoluyla bulaşması ve böylece tüm vücuda yayılması mümkün olmaktadır.

Herhangi bir sebepten ötürü ciltte bir kesik veya yara oluştuğunda, siğillerin bulaşması olasılığı artar. Çoğunlukla, siğiller zararsızdır. Bununla birlikte virüs kaynaklı oldukları için son derece bulaşıcıdır. Siğile doğrudan temas etmek ya da siğile değen başka bir şeye temas etmek, virüsün yayılmasına neden olabilir. Ellerde, parmaklarda ve ayaklarda sık görülür. Ayrıca, çalışma temposunun yüksekliği, bozuk ve sağlıksız beslenme düzeni, çok eşlilik, korunmasız

cinsel ilişki, kişisel bakım malzemelerinin ortak kullanımı ve havuz gibi sıhhi olmayan yerlerde siğil bulaşma oranı artmaktadır.

Herkesin siğile sebep olan HPV virüsünü kapma olasılığı bulunmasına rağmen, bazı gruplar bu hastalığa yakalanmaya daha yatkındırlar: çocuklar, ergenlik çağındakiler, tırnaklarını yiyenler, tırnak etlerini koparanlar ve bağışıklık sistemi (vücudun savunma mekanizması) zayıf olan kişiler. Siğile sebep olan virüsün vücuda girişine olanak sağlayan durumlar şu şekilde sıralanabilir:

- Alerjik bir bünyeye sahip olmak,
- Bazı bağışıklık sistemini baskılayan hastalıklara yakalanmış olmak veya ilaçları kullanmak,
- Cinsel ilişki kurmak (özellikle etek siğilleri için),
- Deride kesik veya yara bulunması,
- Sıhhi olmayan havuz kenarı, hamam, kaplıca, banyo, duş ve tuvalet gibi ortak kullanıma sahip sıcak ve nemli ortamlarda bulunmak,
- Sıhhi koşullara özen gösterilmeyen kuaförler, güzellik salonlarına gitmek,
- Ortak kullanılan ve siğile temas etmiş olabilecek havlu, bornoz, terlik, ayakkabı vb. eşyaları kullanmak,
- Cildin sık tıraş edilmesi,
- Siğile doğrudan temas etmek,
- Tırnak yeme alışkanlığına sahip olmak,
- Yorgunluk, uykusuzluk, stres, kötü beslenme, diyet, kansızlık benzeri nedenlerle vücudun zayıf düşmesine sebep olmak.

Siğil, cildin üst tabakasına insan papilloma virüsünün (HPV) yerleşmesiyle oluşan bir cilt hastalığıdır (Çarman vd. 2013). Vücudun hemen her yerinde (özellikle el ve ayaklarda) görülen siğillerin çoğu iyi huylu olarak tanımlanmaktadır. Bununla birlikte virüs kaynaklı oldukları için bulaşıcıdırlar.

1.2. SİĞİL ÇEŞİTLERİ

Sığiller; görünüm, renk ve büyüklük açısından oluştukları bölgeye göre yapısal farklılıklar gösterirler. Bununla birlikte, genelleme açısından dört başlıkta toplanabilirler (Hoşrik 2010): (1) Yaygın sığiller, (2) ayak taban sığilleri, (3) düz sığiller ve (4) etek sığilleri. Bunlar arasında en yaygın olan sığil türleri yaygın sığiller ve ayak tabanı sığilleridir. Bu nedenle, bu çalışmada bu sığil türlerinden en az birisine sahip olan hastalar dikkate alınmıştır.

1.2.1. Yaygın Sığiller

Elin çeşitli bileşenlerinde (avuç içinde, elin üstünde, el parmaklarında, el parmaklarının aralarında, el parmağı tırnaklarının çevresinde) sık karşılaşılan bir sığil çeşididir (Şekil 1.1). Genellikle, deri dokusunun bütünlüğü bozulduğu yerin etrafında daha çok karşılaşılr. Özellikle tırnak yeme alışkanlığı bulunan kişilerde yaygın olarak görülmektedir (Alp Avcı ve Bozdayı 2013, Hoşrik 2010). Bu nedenle, çocuklar arasında en sık görülen sığil türü olarak kabul edilir.

1.2.2. Ayak Tabanı Sığilleri

Bu sığil türü tek veya grup şeklinde ayakların alt yüzeylerinde görülebilir (Şekil 1.2). Genellikle nasır gibi göründükleri ve deri içerisine gömülü olduklarından nasırla karıştırılırlar. Bütün vücut ağırlığına maruz kalan bu sığiller hastada şiddetli bir ağrıya sebep olabilirler (Alp Avcı ve Bozdayı 2013, Hoşrik 2010). Bu nedenle genellikle sadece kozmetik sorunu olarak tanımlanamazlar.



Şekil 1.1 Toplumda sık karşılaşılan siğil türlerinden Yaygın siğil.



Şekil 1.2 Toplumda sık karşılaşılan siğil türlerinden Ayak tabanı siğili.

1.2.3. Düz Siğiller

Diğer siğil türlerinin aksine, düz siğiller küçük, yumuşak ve düz bir yapıya sahiptir. Vücutta herhangi bir yerde görülebilirler. Bazen saç içinde veya yüzün çevresinde dahi görülebilmektedir.

1.2.4. Etek Siğilleri

Cinsel organların bulunduğu genital bölgede yer alan siğil türü olduğundan, daha çok erişkinlerde görülür. Kansere sebep olabildiği için kesinlikle tedavi edilmesi gerekmektedir. Cinsel ilişki yoluyla bulaşma olasılığı oldukça yüksektir.

1.3. SİĞİL TEDAVİSİ

Çocukluk çağında bir siğille karşılaşıldığında, tedavi uygulanmaksızın siğilin kendiliğinden iyileştiği görülmektedir. Bununla birlikte, eğer siğil acı vermekte ve çocuk hastayı rahatsız etmekte ise, kesinlikle bir cildiye hekimine gidilerek gerekli tedavinin alınması gerekmektedir. Yetişkinlerde ise, siğilin kendiliğinden iyileşmesine pek rastlanmamaktadır.

Siğiller toplumda çok yaygındır ve siğillerin bulaşması oldukça kolay ve hızlı olduğundan siğillerin tedavileri hususunda titizlikle çalışılması gerekmektedir (Varol 1998, Arıcan 2004). Siğil tedavisine yönelik birçok alternatif yöntem bulunmaktadır. Bu yöntemlerin seçimi genellikle siğilin çeşidine ve sahip olunan maddi olanaklara göre değişmektedir. Bu tedavilerden bir kısmı sadece uzman hekimler tarafından (kriyoterapi, lazer ve cerrahi eksizyon gibi), bir kısmı da hastaların kendi başlarına uygulayabileceği (salisilik asit, podpphyline gibi) yöntemlerdir. Bunların haricinde günümüzde imünoterapi ve interferon gibi alternatif yöntemler geliştirilmiştir.

Bu şekilde, siğile sebep olan virüsün yerleştiği dokunun bozulması prensibine dayanan çeşitli fiziksel ve kimyasal uygulamalar bulunmaktadır. Bunlar arasında kantaridin, koter, cerrahi, lazer ve çeşitli kimyasal solüsyon kullanımı gibi uygulamalar bulunmaktadır. Bu yöntemlere ek olarak hastanın bağışıklık sisteminin güçlendirilmesi şeklinde ortaya çıkan immünoterapi ve interferon gibi göreceli olarak yeni tedavi yöntemleri de geliştirilmiştir (Khozeimeh et al. 2017a, Khozeimeh et al. 2017b). Bu çalışmalarda, siğil tedavisinde yaygın kullanılmaya başlayan kriyoterapi ve immünoterapi tedavi yöntemlerinin uygulandığı hastaların kayıt edildiği veri setleri kullanılmıştır.

Siğil tedavisinde sıkça kullanılmayan başlanan bir yöntem olan Kriyoterapi yönteminde, hastalığa sahip bölge çok düşük sıcaklıkta dondurularak öldürülmektedir. Donma tedavisi olarak da bilinen bu yöntemle, siğilin yerleştiği kokunun altında kalan yapı bozulmaktadır. Bu

uygulama için açık sprej cihazı ve $-195,6$ °C sıcaklığına sahip sıvı azot kullanılmaktadır. Bu tedavi türünün avantajları arasında tedaviye başlamadan önce uygulama yapılacak bölgenin uyuşturulmasının gerekmemesi ve oldukça basit bir şekilde uygulanabilmesi bulunmaktadır. Sıvı azot sprej cihazı kullanılarak siğilin bulunduğu yerde bir donma halkası görülene kadar siğilin tam üstüne uygulanır (Şekil 1.3). Siğil geçinceye kadar bu uygulama 2 ila 4 hafta aralıklarla tekrar edilir. Uygulama sırasında veya sonrasında tedavinin uygulandığı bölgede ağrı oluşması genellikle karşılaşılan bir yan etkidir. Bunun yanı sıra, uygulamadan sonra tedavi edilen bölgede şişme veya su toplanması ile tedavi tamamlandıktan sonra uygulama yapılan bölgede beyaz veya kahverengi renkte bir iz kalması gibi yan etkiler de görülebilmektedir. Ayrıca tedavi süreci de uzun sürmektedir (Khozeimeh et al. 2017a, Khozeimeh et al. 2017b).

Siğil tedavisinde sıkça kullanılmayan başlanan bir diğer yöntem olan İmmünoterapi yönteminde, kişinin kendi bağışıklık sistemi harekete geçirilerek vücudun kendisinin tedaviyi sağlaması amaçlanmaktadır. Farklı türlerde İmmünoterapi yöntemleri bulunmaktadır. Bu çalışmada kullanılan verilerin elde edildiği çalışmada, hastanın alerjik reaksiyon gösterdiği bilinen bir madde siğil dokusuna enjekte edilmekte ve böylece vücudun siğili alerji kaynağı olarak görmesiyle siğili yok etmesi sağlanmaktadır (Khozeimeh et al. 2017a, Khozeimeh et al. 2017b).



Şekil 1.3 Kriyoterapi tedavisi uygulanışı.

Görüldüğü üzere, siğil tedavisinde birçok farklı yöntem olmasına rağmen halen %100 tedavi başarımları elde edilememiştir (Alp Avcı ve Bozdayı 2013, Lipke 2006). Bunun yanı sıra

hastaların hangi tedavi yöntemine daha iyi yanıt (tepki) vereceği uzman hekimler tarafından da tam olarak bilinmemektedir (McGibbon 2006, Khozeimeh et al. 2017a, Khozeimeh et al. 2017b). Siğil tedavisinde kriyoterapi ve immünoterapi tedavi yöntemleri en sık kullanılanlardandır. Literatürde bu iki yöntemin etkilerinin araştırıldığı birçok çalışma vardır (Silverberg et al. 2000, Clifton et al. 2003, Horn et al. 2005, Maronn et al. 2008, Gamil et al. 2010, Nofal and Nofal 2010, Russell et al. 2010). Ancak hastadan toplanan bilgilere bakılarak, tedavi uygulanmasına başlamadan önce seçilen siğil tedavi yönteminin hasta üzerinde başarılı olup olmayacağını kestirilmesine yönelik bir çalışma henüz ortaya konulamamıştır (Khozeimeh et al. 2017a).

1.4. MAKİNE ÖĞRENMESİ

Kavramsal olarak, aşırı miktardaki veri içinde saklı bulunan bilgiyi bulup çıkarma ve ileride olası olan durumların kestirilmesiyle ilgili modeller bulmaya olanak sağlayan ilişkilerin bir bilgisayar programı aracılığıyla araştırılmasına makine öğrenmesi veya örüntü tanıma denilmektedir (Özkan 2016, Duda et al. 2000). Makine öğrenmesi için yapılan çalışmaların en başında sınıflandırma yer almaktadır. Böylece; riskli müşteri tespiti gibi uygulamalarla bankacılık ve sigortacılık sektörden hastalık teşhisi gibi karar destek uygulamalarıyla tıp sektörüne kadar çok çeşitli alanlarda sınıflandırma problemi çözümlerinde makine öğrenmesi algoritmaları kullanılmaktadır (Kayan 2016). Tüm uygulamalarda en yüksek doğruluğu veren genel geçer bir algoritma olmayıp, uygulama alanına özel bilinen algoritmalar uygulanarak, en yüksek genel başarıyı veren algoritma veya algoritmalar tespit edilmeye çalışılmaktadır. Bir algoritmanın başarıyı, algoritmaya uygulanan eğitim verisinin (ve özniteliklerin) sayısı ve yapısı, uygulanan modelin ve parametrelerin seçimlerine bağlı olmaktadır (Ayhan ve Erdoğan 2014).

Cilt hastalıklarının teşhisinde kullanılan birçok makine öğrenmesi tabanlı akademik çalışmaya rastlanmaya başlanmıştır (Lamminen et al. 2001, Bunte et al. 2011, Giotis et al. 2015, Jain et al. 2015, Shrivastava et al. 2015, Mirzaalian et al. 2016, Vasudevan et al. 2016, Oliveira et al. 2016, Flores and Scharcanski 2016). Literatürde imünoterapi ve kriyoterapi tedavi yöntemlerinin etkilerinin araştırıldığı birçok çalışma bulunmaktadır (Silverberg et al. 2000, Clifton et al. 2003, Horn et al. 2005, Maronn et al. 2008, Khurshid et al. 2009, Gamil et al. 2010, Nofal and Nofal 2010, Russell et al. 2010, Khozeimeh 2017a, Khozeimeh 2017b).

Bununla birlikte, siğil tedavi yöntemlerinin karşılaştırıldığı ve hangi yöntemin kullanılması halinde tedavinin daha başarılı sonuç verebileceğine yönelik fazla çalışma bulunmamaktadır. Bu amaca yönelik olarak literatürde karşılaşılan sadece üç adet çalışma bulunmaktadır (Khozeimeh et al. 2017a, Khozeimeh et al. 2017b, Tanyıldızı vd. 2018). Bunlardan ilkinde, immünoterapi yönteminin siğil tedavisinde başarılı bir şekilde uygulanabileceği istatistiksel olarak ortaya konulmuştur (Khozeimeh et al. 2017b). İkinci çalışmada ise, Bulanık Mantık kullanılarak kural tabloları oluşturulmuş ve bu şekilde %80,00-%83,33 (ortalaması %81,67) oranında kriyoterapi ve immünoterapi siğil tedavi yöntemlerinin başarılı olup olmayacakları tahmin edilebilmiştir (Khozeimeh et al. 2017a). Son çalışmada ise bu iki grup hasta için ayrı ayrı başarı tahmin edilmeye çalışılmıştır. Bu çalışmada Naive Bayes, Çok Katmanlı Ağ, Destek Vektör Makineleri, En Yakın K Komşu, K-Star, Karar Ağacı ve Random Forest algoritmaları denenmiştir. Sonuç olarak, kriyoterapi ile tedavi edilen hastalarda uygulanan tedavinin uygun olup olmadığına yönelik yapılan tahminlerde K-Star algoritması ile %96,66 ve immünoterapi ile tedavi edilen hastalarda ise Random Forest algoritması ile %85,55 başarı elde edildiği rapor edilmiştir (Tanyıldızı vd. 2018). Bu çalışmaya göre ortalama başarı Random Forest algoritması ile %89,44 olarak elde edilmiştir.

1.5. AMAÇ

Bu tez çalışmasında siğil tedavi yönteminin başarılı olup olmayacağı tahmin edilmesinde farklı yapay zeka ve makine öğrenmesi algoritmalarının kullanımı araştırılmıştır. Böylece uzman hekimlerin hastalarından elde edecekleri ön bulgulara göre hangi yöntemin işe yarayabileceğinin önceden kestirilmesi amaçlanmıştır. Böylece işe yaraması en muhtemel tedavi yönteminin uygulanması ile hastanın daha az acı çekmesi ve daha kısa sürede tamamlanan bir tedaviye ulaşması sağlanabilecektir.

Bu çalışmada, İran'ın Meşhed kentindeki Ghaem Hastanesi dermatoloji polikliniğine başvuran 180 hasta ve plantar ve yaygın siğiller için gerçekleştirilen tedavi yöntemleri sırasında kayıt edilen veriler kullanılmıştır. Kayıt edilen verilerde 90 hasta sıvı azot ile kriyoterapi yöntemi ile 90'ı immünoterapi yöntemi ile tedavi edilmiştir. Tedavi yönteminin seçimi rastgele yapılmıştır. Literatürde mevcut durumda sadece istatistiksel karşılaştırma (Khozeimeh et al. 2017b), bulanık mantık kuralına dayalı bir uzman sistem ile tedavi başarılarının tahmin edilmesi (Khozeimeh et al. 2017a) ve Naive Bayes, Çok Katmanlı Ağ, Destek Vektör Makineleri, En Yakın K Komşu, K-Star, Karar Ağacı ve Random Forest algoritmalarının denendiği (Tanyıldızı

vd. 2018) alıřmaları mevcuttur. Bu alıřmalarda, imünoterapi ve kriyoterapi yöntemlerinin öngörme doęruluęu sırasıyla en yüksek olarak %85,55 ve %96,66 olarak belirlenebilmiştir.

Elde edilen mevcut sonuçlara göre, hekime en iyi tedavi yöntemini seçmede yardımcı olacak, hastalar için zamandan tasarruf etmeyi sağlayacak, tedavi maliyetini düşürecek ve tedavinin kalitesini arttıracak bir uzman sistem geliştirilmesine ihtiyaç bulunmaktadır. Takip eden bölümlerde sırasıyla alıřmada kullanılan verilerin ve yöntemlerin kısa açıklaması ile sonuçlar ve tartışma bölümleri yer almaktadır.





BÖLÜM 2

YÖNTEM

2.1. VERİ SETİ

Bu çalışmada, internet üzerinden yayımlanmakta olan UCI Machine Learning Repository veri tabanında Immunotherapy Dataset ve Cryotherapy Dataset başlıkları altında yer alan iki ayrı veri seti kullanılmıştır. Bu veriler, siğil tedavisi için Ocak 2013- Şubat 2015 tarihleri arasında Mashhad-Ghaem Hastanesi'ne giriş yapan 180 hastadan alınmıştır. Giriş yaptıran hastalarda yaygın ve ayak taban siğil türlerinden en az bir tanesi bulunmaktadır (Khozeimeh et al. 2017a, Khozeimeh 2017b).

Çalışmada kullanılan veri setlerinden biri kriyoterapi yöntemi uygulandığında göz önünde bulundurulan yedi adet özniteliği içermektedir (Çizelge 2.1). Diğeri ise immünoterapi yöntemi uygulandığı durumda kayıt edilen sekiz adet öznitelikten oluşmaktadır (Çizelge 2.2). Her iki veri setinde yöntemin başarı olup olmadığını gösteren öznitelik olan Tedaviye Yanıt Verme sınıf karşılığı olarak belirlenmiştir.

Çizelge 2.1 Kriyoterapi tedavi yönteminde kullanılan öznitelikler. Burada SD standart sapmayı göstermektedir.

| Öznitelik | Değerleri | Ortalama \pm SD |
|--------------------------------------|--|--------------------|
| Tedaviye yanıt verme | Evet / Hayır | |
| Cinsiyet | 47 Erkek 43 Kadın | |
| Yaş (yıl) | 15 – 67 | 28,60 \pm 13,36 |
| Tedaviden önce geçen zaman (ay) | 0 – 12 | 7,66 \pm 3,40 |
| Siğil sayısı (adet) | 1 – 12 | 5,51 \pm 3,57 |
| Siğil çeşidi | 1: Yaygın siğil (54) 2: Ayak tabanı siğili (9) 3: İkisi de mevcut (27) | |
| Siğil yüzey alanı (mm ²) | 4 – 750 | 85,83 \pm 131,73 |

Çizelge 2.2 Kriyoterapi tedavi yönteminde kullanılan öznitelikler. Burada SD standart sapmayı göstermektedir.

| Öznitelik | Değerleri | Ortalama \pm SD |
|--------------------------------------|----------------------------|--------------------|
| Tedaviye yanıt verme | Evet / Hayır | |
| Cinsiyet | 41 Erkek | |
| | 49 Kadın | |
| Yaş (yıl) | 15 – 56 | 31,04 \pm 12,23 |
| Tedaviden önce geçen zaman (ay) | 0 – 12 | 7,23 \pm 3,10 |
| Siğil sayısı (adet) | 1 – 12 | 6,14 \pm 4,20 |
| Siğil çeşidi | 1: Yaygın siğil (47) | |
| | 2: Ayak tabanı siğili (22) | |
| | 3: İkisi de mevcut (21) | |
| Siğil yüzey alanı (mm ²) | 6 – 900 | 95,70 \pm 136,61 |
| Test öncesi siğilin çapı (mm) | 5 – 70 | 14,33 \pm 17,22 |

Verilerin elde edildiği orijinal çalışmada bu iki veri seti birleştirilmiş ve sonrasında her birinde 90' ar kişi olacak şekilde rasgele iki ayrı gruba ayrılmıştır. Bu grupların ilkindeki hastaların tedavisinde kandida antijen enjeksiyonlu immünoterapi tedavi yöntemi kullanılmıştır. İntralezyonel olarak uygulanan bu yöntemde hastaların tedavisine üç haftalık aralıklarla üç seansla devam edilmiştir. Hastanın tedaviye erken tepki vermesi halinde süreç erkenden sonlandırılmıştır.

İkinci grupta yer alan hastalara sıvı azotlu kriyoterapi yöntemi uygulanmıştır. Hastaların tedavi süreci siğillerin iyileşme durumuna göre birer hafta arayla on seans veya daha az devam edilmiştir. Farklı tedavi süreçleri ve yöntemleri kullanılmasına rağmen hastanın tedaviye yanıt vermemesi durumunda farklı bir siğil tedavi yönteminin uygulanması gerekmektedir (McGibbon 2006).

2.2. MAKİNE ÖĞRENMESİ ALGORİTMALARI

Bu tez çalışmasında yedi farklı makine öğrenmesi algoritması kullanılmıştır: Naive Bayes, Lojistik Regresyon, Karar Ağacı, K-Yakın Komşu, Destek Vektör Makineleri, Aşırı Öğrenme Makinesi ve Çok Katmanlı Algılayıcı. Sadece lojistik regresyon ve karar ağacı algoritmalarının

gerçekleştirilmesi için WEKA yazılımı kullanılmış olup diğer algoritmaların gerçekleştirilmesinde Matlab 2018a programlama ortamı kullanılmıştır. Aşağıdaki alt bölümlerde sınıflandırıcı algoritmalarının kısa açıklamalarına yer verilmiş olup ayrıntılı bilgi için literatürdeki makine öğrenmesi ve örüntü tanıma kitaplarına bakılabilir (Duda et al. 2000).

2.2.1 Naive Bayes

Bu sınıflandırıcı algoritmasının temelinde kayıt edilen her özneliğin diğer özneliklerden istatistiksel olarak bağımsız olduğu varsayımı bulunmaktadır. Buna göre, bir özneliğin oluşmasında diğer özneliklerin oluşmasının herhangi bir etkisi veya katkısı bulunmamaktadır. Gerçek dünyada gerçekleşme olasılığı zayıf olan bu varsayımına rağmen, Naive Bayes algoritması gerçek dünya problemlerinin çözümünde yüksek oranda başarılı olmuştur (Duda et al. 2000).

Algoritmanın teorik temellerini ortaya koyan ünlü İngiliz matematikçisinin adı ile anılan Bayes Teoremi'ne göre, herhangi bir sınıflandırma probleminde gerçekleşebilecek olası bağımlı durumların olasılıkları şu şekilde hesaplanabilir:

$$P(w|X) = \frac{P(X|w)P(w)}{P(X)} \quad (2.1)$$

Burada $P(w|X)$; X olayı gerçekleşmişse w olayının da gerçekleşmesi olasılığını verir. Benzer şekilde $P(X|w)$ ise w olayı gerçekleştiği bilinirken X olayının meydana gelme olasılığıdır. Bu iki değere koşullu olasılık adı verilmektedir. $P(w)$ ve $P(X)$ ise sırasıyla w ve X olaylarının tek başlarına koşulsuz bir şekilde gerçekleşme olasılıklarını göstermektedir.

Bu sınıflandırıcı algoritması, Matlab 2018a yazılımı üzerinde *classify* hazır komutu kullanılarak çalıştırılmıştır.

2.2.2. Lojistik Regresyon

Lojistik regresyon, sınıflandırma ve atamada sıklıkla kullanılan çok değişkenli veri madenciliği tekniklerindedir. Bu algoritma ile bağımsız değişkenlerin bağımlı değişken üzerindeki etkilerinin bir olasılık değeri şeklinde hesaplanması sağlanır. Lojistik regresyon analizinin

amacı; bağımlı deęişkenlerle bağımsız deęişkenler arasında var olan bağıntıyı en az sayıda deęişken kullanarak en yüksek doğrulukla ifade edebilen geçerli bir matematiksel model ile oluşturmaktadır (Bircan 2004).

Lojistik regresyon modeli;

$$P = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n}} \quad (2.2)$$

biçiminde yazılır. Burada P incelenen olayın gözlenme olasılığı; β_0 sabit; $\beta_1, \beta_2, \dots, \beta_n$ bağımsız deęişkenlerin regresyon katsayıları; x_1, x_2, \dots, x_n bağımsız deęişkenleri ve n bağımsız deęişkenlerinin sayısını ifade etmektedir (Özdamar 2009).

Lojistik regresyon algoritması şu adımlarda yapılır (Şekeroęlu 2010):

1. Öncü grup üyelikler ve deęişkenler belirlenir,
2. Modelde parametreleri tahmin edilerek anlamlılığı incelenir,
3. Modelden kestirilen parametreler kullanılarak her bir gözlemin ait olduęu sınıf bulunur,
4. Modelin uyumluluęu sınıflandırma yüzdesi ve yapay olarak oluşturulan R2 ölçütü kullanılarak test edilir.

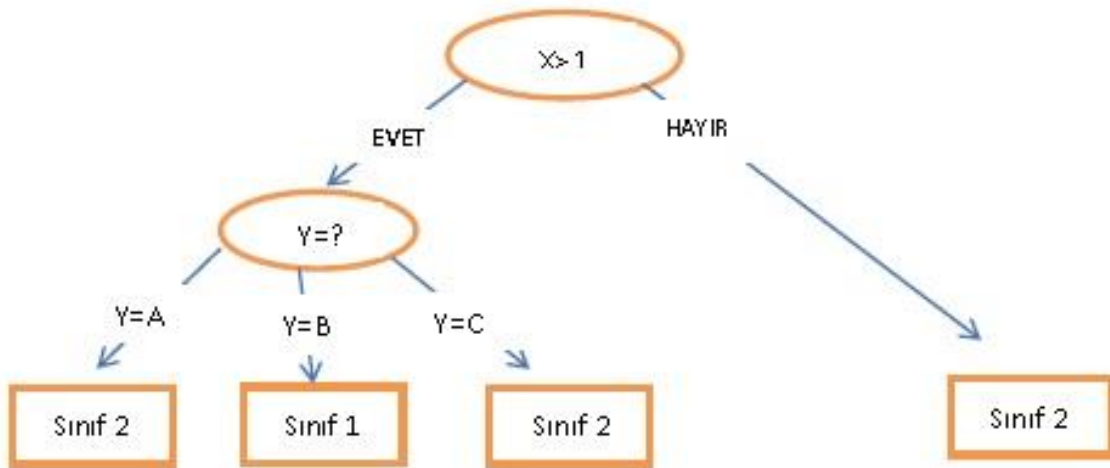
2.2.3. Karar Aęacı

Karar aęacı (KA), sınıfları bilinen veri setinden tümevarım yöntemiyle öğrenilen aęaç şekilli bir karar yapısına sahip veri madencilięi yaklaşımıdır (Albayrak ve Yılmaz 2009). KA, bağımlı deęişkendeki bilgi kazancına göre veri setini sıralı bir şekilde bölmektedir (Sun and Li 2008). Her başarılı bölme işlemi ile gruplanan örnekler kendi içlerinde gittikçe daha benzer hale getirilmektedir. Tahmin edici veya tanımlayıcı olarak KA, yorumlanmasının kolay, maliyetinin az, veri tabanı sistemleri ile kolayca uyum sağlayabilmesi nedeniyle yaygın bir kullanıma sahip olmuştur (Albayrak ve Yılmaz 2009). Son zamanlarda çok sayıda farklı karar aęacı öğrenme teknikleri geliştirilmiştir. Bunlardan en popüler olanları ID3, C4.5 ve C5 algoritmalarıdır (Özarslan ve Barışçı 2014). Bu çalışmada WEKA yazılımı kullanılarak karar aęacı algoritmalarından J48 uygulanmıştır.

Yapısı bir akış şemasına benzeyen KA' da her öznitelik bir düğüm ile gösterilir. Ağaç yapısını dallar ve yapraklar oluşturur. Bu yapının en üst ve en son kısmını sırasıyla "yaprak" ve "kök", bunlar arasında kalan kısımlar da "dal" olarak isimlendirilir. Şekil 2.1' de iki nitelikten oluşan basit bir karar ağacı gösterilmiştir. Bu yapıda "Sınıf 1" de $X > 1$ ve $Y = B$ değerini taşıyan örnekler, "Sınıf 2" de $Y = A$ ve $Y = C$ değerini taşıyan örnekler yer almaktadır. Y' nin değerinin göz önüne alınmadığı $X \leq 1$ şartını sağlayan örnekler "Sınıf 1" de yer alır (Silahtaroglu 2008, Özkan 2016).

2.2.4. En Yakın K Komşu

En Yakın K Komşu (KNN) algoritması basit bir yapıya sahip bir sınıflandırma yöntemidir. Test edilmek istenilen bir örneğin veri setinde mevcut olan diğer tüm örnekler ile arasındaki uzaklıklar hesaplanır. Bu şekilde en düşük uzaklığa sahip en yakın k adet örneğin ait olduğu sınıflardan hangisi daha fazla ise, test edilen örneğin sınıfı o gruba dahil olarak belirlenir (Duda et al. 2000, Cover ve Hart 1967). Bu algoritmanın basit yapısı nedeniyle literatürdeki makine öğrenmesi çalışmaları arasında oldukça popüler olduğu görülmektedir (İşler ve Kuntalp 2007, İşler ve Kuntalp 2010, Narin 2013, Narin et al. 2014a, Narin et al. 2014b, İşler et al. 2015, Narin et al. 2018, İşler et al. 2019).

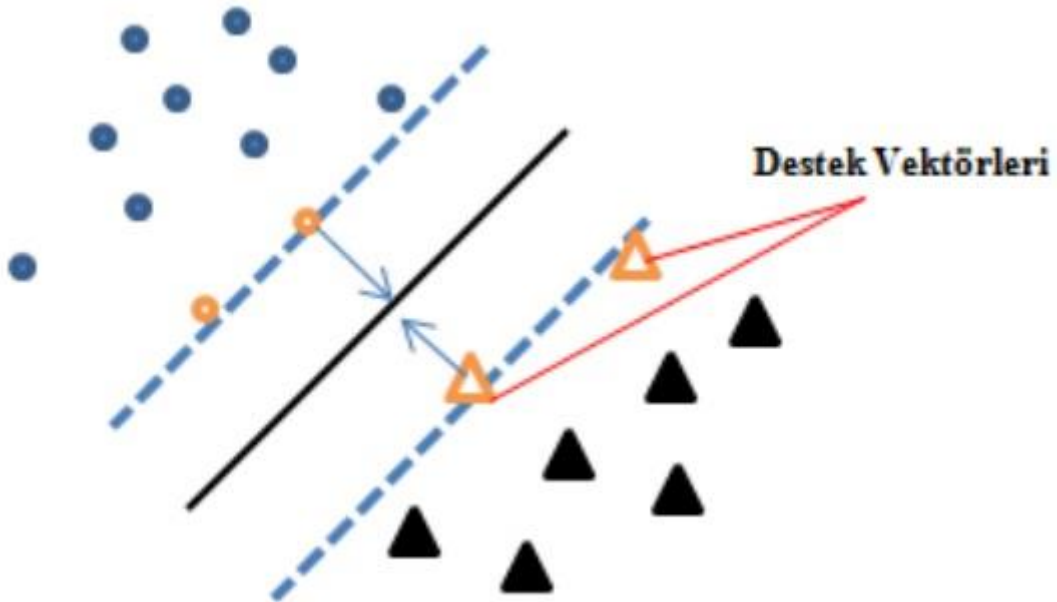


Şekil 2.1 X ve Y nitelikleri üzerine uygulanan testleri içere basit bir karar ağacı (Özkan 2016).

Bu sınıflandırıcı algoritmasının uygulanmasında, Matlab 2018a yazılımındaki hazır kütüphane komutlarından olan *knnclassify* kullanılmıştır. Komşuluk sayısı (k) olarak 1, 3, 5, 7, 9, 11 ve 13 gibi tek sayı olan değerler sırasıyla denenmiştir.

2.2.5. Destek Vektör Makineleri

Destek vektör makinaları (SVM) yönteminde temel amaç iki veri grubunu daha yüksek boyutlu bir uzaya taşıyarak birbirlerinden ayrılabilirliklerini sağlamaktır. Veri grupları doğrusal olabilecekleri gibi doğrusal olmayan yapıya da sahip olabilirler (Duda et al. 2000, Byun and Lee 2003). Bu amaçla destek vektörleri olarak bilinen ve veri gruplarının her ikisine birden yakın olan birbirine paralel iki sınır çizgisi belirlenir. Bu sınıflandırma yapısında, elde edilen bu sınır çizgilerine en uygun bir uzaklıkta konumlandırılan bir ayırıcı düzlem veya çizgi bulunur (Şekil 2.2). Çoğunlukla verileri birbirinden ayırt edebilen doğrusal bir yapı yeterli olmadığından, veriler çekirdek fonksiyonları olarak bilinen doğrusal olmayan fonksiyonlar aracılığıyla daha yüksek boyutlu bir uzaya aktarılır. Bu dönüşüm sonrası, veri grupları bir doğrusal düzlem ile birbirlerinden ayrılabilir bir biçime dönüştürülmektedirler (Narin 2013, Ayhan ve Erdoğan 2014).



Şekil 2.2 Destek Vektör Makineleri destek vektörleri ve ayırım düzlemi.

Bu sınıflandırıcı algoritmasının uygulanmasında, Matlab 2018a yazılımındaki hazır kütüphane komutlarından olan *svmclassify* kullanılmıştır.

2.2.6. Aşırı Öğrenme Makinesi

Aşırı öğrenme makinesi (ELM) son zamanlarda giderek popüler hale gelen bir sınıflandırıcı algoritması haline gelmiştir. Bu algoritma tek gizli katman bulunan ileri yönlü hesaplamalı ağlar arasında yer almaktadır. ELM algoritması daha başarılı regresyon ve sınıflandırma başarımları vermekte ve bu sonuçlara çoğu geleneksel yapay sinir ağları algoritmalarından daha hızlı ulaşmaktadır (Huang et al. 2012). Algoritmanın giriş ağırlıkları rasgele olarak seçilmekte ve çıkış ağırlıkları analitik olarak elde edilmektedir (Huang et al. 2004). ELM aşırı derecede hızlı öğrenme yeteneğiyle iyi bir genelleme başarımı vermektedir (Huang et al. 2006). Üstelik iyi hazırlanmış kaynak dokümanlar ve farklı programlama dillerinde hazırlanmış kaynak kodları barındıran bir web sitesine de sahiptir. Algoritma ile her türlü bilgi ve program örneklerine bu web sitesinden ulaşılabilir (URL-1).

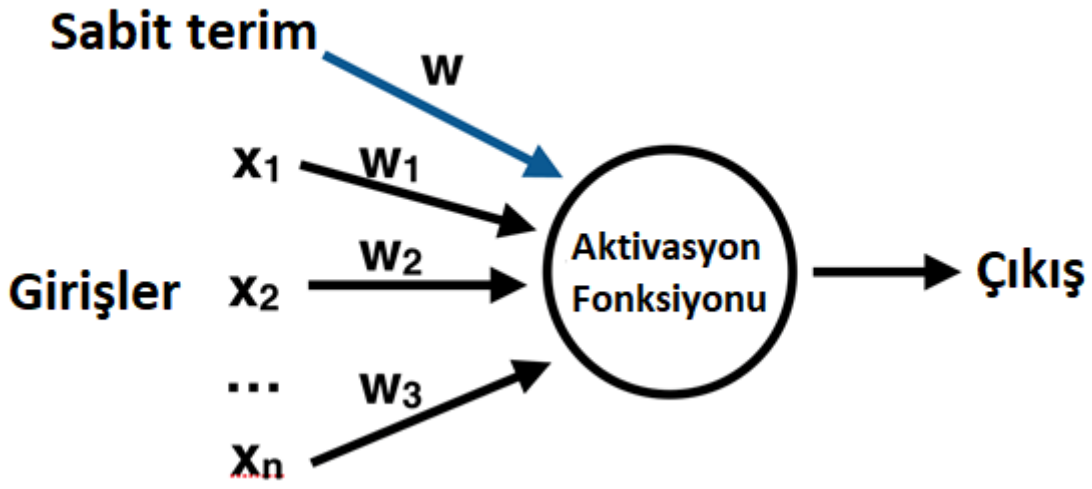
2.2.7. Çok Katmanlı Algılayıcı

Çok Katmanlı Algılayıcı (MLP) ağ yapısı yapay sinir ağı modelleri arasında en çok tercih edileni olmuştur. Klinikte birçok rahatsızlığın teşhisini de içeren çok geniş bir yelpazede uygulamada sık kullanılır hale gelmiştir. Üstelik diğer algoritmaların başarımlarının kıyaslanmasında ortak bir zemin oluşturmaktadır (Koçer ve Canal 2011, Pan et al. 2012, Narin et al. 2018, İşler et al. 2019).

MLP tabanlı çalışmalarında hemen hepsinde üç katmanlı bir yapı tercih edilmektedir: giriş katmanı, gizli katman ve çıkış katmanı. Bazı çalışmalarda birden fazla gizli katmana yer verildiği de görülmektedir. Bu tez çalışmasında kullanılan veri seti nedeniyle 7 adet öznitelik bulunduğundan giriş katmanındaki nöron sayısı 7 olarak belirlenmektedir. Çıkış olarak sadece seçilen tedavi yönteminin başarılı olup olmayacağı şeklinde bir cevap üretileceğinden çıkış katmanında 1 nöron bulunmaktadır. Gizli katmandaki nöron sayısını belirlemeyle ilgili literatürde çalışmalar ve öneriler bulunsa da, genellikle uygulanan yöntem farklı sayılarda nöronlar uygulanarak sınıflandırıcı başarımını en yüksek veren sayı tercih edilmektedir.

Katmanları birbirine bağlayan ağ ağırlıkları hatanın geri yayılımı yöntemi ile tespit edilmektedir. Bu yöntemde, ağ tarafından üretilen çıkış ile beklenen gerçek çıkış arasındaki hataların karelerinin ortalaması, ağdaki her bir nöronun hataya verdikleri katkı oranında ağırlıklarının değiştirilmesi temeli esas alınmaktadır. Bu süreç ortalama karesel hata belli bir eşik değerinin altına düşünceye kadar devam eder (Duda et al. 2000).

Her bir nöronun yapısı aşağıdaki gibi modellenir (Şekil 2.3):



Şekil 2.3 Çok katmanlı algılayıcı ağ yapısı için tek nöron yapısı.

Burada tek nöronun çıkışı şu giriş değerleri (öznitelikler) ile ağırlıkların (weights) çarpımının toplamı bir aktivasyon fonksiyonun (genellikle tanh seçilir) geçirilerek çıkış (output) bulunmuş olur. Bu işlem ağdaki tüm katmanlara ait tüm nöronlar için hesaplanır. Algoritma ayrıntıları için literatürdeki kitaplara bakılabilir (Duda et al. 2000, Özkan 2016).

2.3 ALGORİTMALARIN DEĞERLENDİRİLMESİ

Sınıflandırıcı algoritmalarının başarımlarının ölçümünde en popüler ve basit yöntem olan duyarlılık (SEN), belirleyicilik (SPE) ve doğruluk (ACC) oranları kullanılmıştır. Bu değerlendirme ölçütleri ise sınıflandırıcının test verilerine verdiği yanıtlardan çıkarılmaktadır. Şöyle ki, TP (kayıtlı cevap doğru iken doğru olarak cevap verilenlerin sayısı), TN (kayıtlı cevap yanlış iken yanlış olarak cevap verilenlerin sayısı), FP (kayıtlı cevap yanlış iken doğru olarak cevap verilenlerin sayısı), FN (kayıtlı cevap doğru iken yanlış olarak cevap verilenlerin sayısı) parametrelerine göre hesap yapılmaktadır (Duda et al. 2000):

$$SEN = \frac{TP}{TP + FN} \quad (2.2)$$

$$SPE = \frac{TN}{TN + FP} \quad (2.3)$$

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \quad (2.4)$$

2.3.1. Çapraz Doğrulama

Sınıflandırıcı algoritmalarının eğitimlerinde ve test edilmelerinde farklı verilerin kullanılması, böylece sınıflandırıcının daha önce karşılaşmadığı girdiye verdiği cevaplara göre başarımlarının hesaplanması tercih edilmektedir (Duda et al. 2000, Özkan 2016). Bu amaçla en sık kullanılan doğrulama yöntemleri arasında K-Parçalı Doğrulama ve Birisi-Dışarıda Doğrulama yöntemleri bulunmaktadır. K-Parçalı Doğrulama yönteminde veri seti K parçaya bölünür. Bu parçalardan K-1 adedi sınıflandırıcının eğitiminde ve 1 tanesi sınıflandırıcının test edilmesinde kullanılır. Bu şekilde tüm veri seti parçaları test içi kullanılıncaya kadar bu işlem tekrar edilir. Her bir parça için elde edilen TP, TN, FP ve FN ölçümlerinden genel duyarlılık (SEN), belirleyicilik (SPE) ve doğruluk oranı (ACC) performansları hesaplanır (Duda et al. 2000).

2.4. WEKA YAZILIMI

Günümüzde gelişen teknolojiyle birlikte ölçüm cihazlarının sayısı her geçen gün artmakta ve bu da beraberinde veri sayılarında ve türlerinde artış meydana getirmektedir. Veri toplama araçlarında ve veri tabanı teknolojisinde yaşanan gelişmeler, daha fazla bilginin toplanmasına ve çözümlenmesine olanak sağlamaktadır. Farklı kaynaklardan elde edilen veri yığınları içerisinde bilgiye ulaşılmasına ve gelecekle ilgili tahminlerde bulunmaya olanak sağlayabilecek bağlantıların bilgisayar programı kullanarak aranmasına veri madenciliği denilmektedir (Kudyba 2004, Özkan 2016). Bu bağlamda günümüzde açık kaynaklı ve ticari amaçlı birçok yazılım geliştirilmiştir (İşler ve Narin 2012, Dener vd. 2009). Örnek vermek gerekirse; WEKA açık kaynak kodlu; MATLAB ise ticari amaçlı geliştirilen yazılımlardır (Danacı vd. 2010).

"ScienceDirect" veritabanında "WEKA" anahtar kelimesiyle yapılan aramada 4000' den fazla makale, 200' e yakın kitap bölümü ve 113 tane referans çalışmasının olduğu görülmüştür. Son

üç yıl göz önünde bulundurulduğunda, 2018 yılı ilk 4 ayı için 300 den fazla makale ve 11 kitap bölümü; 2017 yılı için 500' den fazla makale ve 36 kitap bölümü; 2016 yılı için 500' den fazla makale ve 18 kitap bulunduğu görülmektedir. Yapılan bu istatistiksel analiz WEKA yazılımının açık kaynak kodlu olması nedeniyle günümüzde araştırmalarda oldukça sık kullanıldığını göstermektedir. Tıp alanında hastalıkların teşhis ve tahminde WEKA' nın kullanımıyla ilgili de birçok örnek çalışma bulunmaktadır. Bunlar arasında meme kanseri hücrelerinin teşhisi (Danacı vd. 2010), göğüs kanserinin teşhisi (Coşkun ve Baykal 2011), Parkinson hastalığının teşhisi ve tahmini (Özçift 2011) ve Konjestif kalp yetmezliğine sahip hastaların teşhisi (İşler ve Narin 2012) gibi çalışmalar sayılabilir.

WEKA; Waikato Üniversitesi tarafından geliştirilen, makine öğrenmesi algoritmalarının ve veri ön işleme gibi gereksinimlerin birlikte sunulduğu, açık kaynak kodlu Java tabanlı bir veri madenciliği yazılımıdır (Witten et al. 2011). WEKA sayesinde veri ön işleme, sınıflandırma, gruplandırma, özellik seçimi veya çıkarımı gibi makine öğrenmesi ve istatistik ile ilgili birçok kütüphane hazır olarak gelmektedir. Bunun yanı sıra içerisinde görselleştirme araçlarını barındırmaktadır. WEKA birçok farklı dosya tipini (names, metin tabanlı arff., csv., dat., bsi. gibi) desteklemektedir. Bu program kullanıcılarına farklı uygulama (Explorer, Experimenter, Knowledge Flow ve Simple CLI) seçenekleri sunmaktadır (Şekil 2.4).

Bu çalışmada amaç, WEKA' nın siğil tedavisinde sıkça kullanılan kriyoterapi ve immünoterapi tedavi yöntemlerinin hangisinin hastanın tedaviye yanıt verme süresinde daha başarılı olacağını belirlenmesidir. Bu nedenle makine öğrenmesi algoritmalarından bir kısmı WEKA yazılımı kullanılarak gerçekleştirilmiştir (Uzun vd. 2018c).



Şekil 2.4. WEKA kullanıcı ara yüzü.



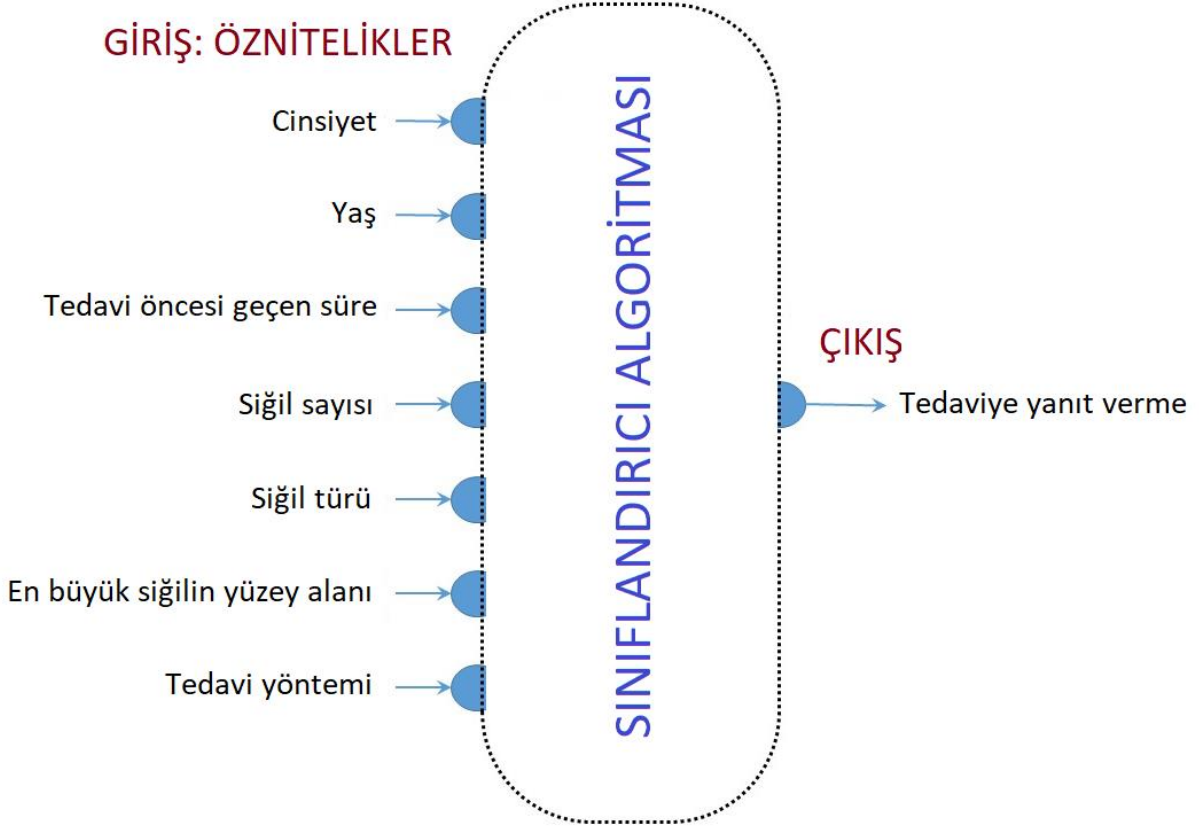
BÖLÜM 3

SONUÇLAR

3.1. VERİLERİN HAZIRLANMASI

Yapılan çalışmada, kriyoterapi ve immünoterapi yöntemleriyle siğil tedavi yapılmış 90'ar hastadan elde edilen öznitelikler kullanılarak bu siğil tedavisi yöntemlerinin hastaya fayda sağlayıp sağlamayacağı tedavi uygulamadan önce tespit edilmesine çalışılmıştır.

Çalışmada bu iki veri seti birleştirilmiş, uygulanan tedavi yöntemi de öznitelik olarak yazılmıştır. Bununla birlikte Tedaviye yanıt verme özneliği ise çıkış değeri olarak düzenlenmiştir. Böylece cinsiyet, yaş, tedavi öncesi geçen süre, siğil sayısı, siğil türü, en büyük siğilin yüzey alanı ve seçilen tedavi yöntemi koşturulacak sınıflandırıcı algoritmasının girişlerine uygulanmıştır. Sınıflandırıcının çıkışında ise tedaviye yanıt verme özneliğinin tahmin edilmesi beklenmiştir (Şekil 3.1).



Şekil 3.1 Bu çalışmada gerçekleştirilen sınıflandırıcı giriş ve çıkışları.

3.2. SINIFLANDIRMA BAŞARIMLARININ DEĞERLENDİRİLMESİ

Sınıflandırıcı başarımlarının değerlendirilmesinde 10 parçalı doğrulama yöntemi kullanılmıştır. Böylece veri seti 10 parçaya bölünmüş ve tüm parçalar test için kullanılmaya kadar 9 parça makine öğrenmesi algoritmasının eğitiminde ve 1 parça eğitilen sınıflandırıcının test edilmesinde kullanılmıştır. Sınıflandırıcı başarımları olan duyarlılık (SEN), belirleyicilik (SPE) ve doğruluk (ACC) değerleri bu parçaların test edilmeleri sırasında oluşan TP, TN, FP ve FN değerleri ile belirlenmiştir. Aşağıdaki sınıflandırma algoritmaları başlığının altındaki alt başlıklar halinde kullanılan sınıflandırıcı algoritmaları sonucunda elde edilen başarımlar şekiller veya çizelgeler halinde verilmektedir.

3.3. SINIFLANDIRMA ALGORİTMALARI

Bu çalışmada literatürde sık kullanılan Naive Bayes, Lojistik Regresyon, Karar Ağacı, K-Yakın Komşu, Destek Vektör Makineleri, Aşırı Öğrenme Makinesi ve Çok Katmanlı Algılayıcı

sınıflandırıcıları sırasıyla denenmiştir. Her bir sınıflandırıcı algoritması algoritmaya özgü farklı parametreler kullanılarak koşturulmuştur.

Bu çalışmada, sadece lojistik regresyon ve karar ağaçları algoritmalarının gerçekleştirilmesi için WEKA yazılımı kullanılmıştır. Diğer algoritmaların gerçekleştirilmesinde Matlab 2018a programlama ortamı kullanılmıştır.

3.2.1. Naive Bayes

Bu sınıflandırıcıların başarımı Matlab programlama ortamında hazır olarak bulunan komutlar kullanılarak bulunmuştur. Buna göre, duyarlılık için %68,43, belirleyicilik için %67,61 ve doğruluk için %67,78 başarımları elde edilmiştir (Uzun vd. 2018a). Bu değerler literatürde bulunan az sayıdaki çalışmadan bile daha düşük sonuçlardır.

3.2.2. Lojistik Regresyon

Lojistik Regresyon algoritmasının hesaplanmasında dikkate alınacak ara katman değeri 2'den 20'ye kadar olan çift sayılar denenmiştir. Bu katman sayılarına karşılık gelen duyarlılık, belirleyicilik ve doğruluk değerleri çizelge halinde düzenlenmiştir (Çizelge 3.1). Buna göre en yüksek sınıflandırıcı başarımına 10 ve 14 ara katmanın kullanıldığı yapılarda ulaşılabilmektedir. En yüksek doğruluk değerini veren sınıflandırıcı başarımında 10 ara katman değeri için %26,32 duyarlılık, %94,37 belirleyicilik ve %80,00 doğruluk değerleri ve 14 ara katman değeri için %15,79 duyarlılık, %97,18 belirleyicilik ve %80,00 doğruluk değerleri elde edilmiştir (Uzun vd. 2018c).

Çizelge 3.1 Lojistik Regresyon algoritması ile siğil tedavi yönteminin başarı olup olmayacağıın belirlenmesi için sınıflandırıcı performansları. Burada SEN duyarlılık, SPE belirleyicilik ve ACC doğruluk değerlerini vermektedir.

| Ara katman sayısı | SEN (%) | SPE (%) | ACC (%) |
|-------------------|--------------|--------------|--------------|
| 8 | 15,79 | 94,37 | 77,78 |
| 10 | 26,32 | 94,37 | 80,00 |
| 12 | 21,05 | 94,37 | 78,89 |
| 14 | 15,79 | 97,18 | 80,00 |
| 16 | 15,79 | 94,37 | 77,78 |
| 18 | 21,05 | 94,37 | 78,89 |
| 20 | 15,79 | 94,37 | 77,78 |

3.2.3. Karar Ağacı

Karar Ağacı algoritmasının hesaplanmasında dikkate alınacak ara katman değeri 2' den 20' ye kadar olan çift sayılar denenmiştir. Bu katman sayılarına karşılık gelen duyarlılık, belirleyicilik ve doğruluk değerleri çizelge halinde düzenlenmiştir (Çizelge 3.2). Buna göre en yüksek sınıflandırıcı başarımına 14 ara katmanın kullanıldığı yapıda ulaşılabilmektedir. En yüksek doğruluk değerini veren sınıflandırıcı başarımında %52,63 duyarlılık, %94,37 belirleyicilik ve %85,56 doğruluk değerleri elde edilmiştir (Uzun vd. 2018c).

Çizelge 3.2 Karar Ağacı algoritması ile siğil tedavi yönteminin başarı olup olmayacağını belirlemek için sınıflandırıcı performansları. Burada SEN duyarlılık, SPE belirleyicilik ve ACC doğruluk değerlerini vermektedir.

| Ara katman sayısı | SEN (%) | SPE (%) | ACC (%) |
|-------------------|--------------|--------------|--------------|
| 8 | 57,89 | 91,55 | 84,44 |
| 10 | 47,37 | 91,55 | 82,22 |
| 12 | 52,63 | 91,55 | 83,33 |
| 14 | 52,63 | 94,37 | 85,56 |
| 16 | 47,37 | 94,37 | 84,44 |
| 18 | 52,63 | 91,55 | 83,33 |
| 20 | 52,63 | 90,14 | 82,22 |

3.2.4. En Yakın K Komşu

En Yakın K Komşu (KNN) algoritmasının hesaplanmasında dikkate alınan komşu sayısı olan k parametresi 1 ile 13 arasındaki tek sayılar kullanılarak denenmiştir. Bu komşuluk değerlerine karşılık gelen duyarlılık, belirleyicilik ve doğruluk değerleri çizelge halinde düzenlenmiştir (Çizelge 3.3). Buna göre en yüksek sınıflandırıcı başarımına k=7 komşuluk değerinde ulaşılabilmektedir. En yüksek doğruluk değerini veren sınıflandırıcı başarımında %15,79 duyarlılık, %97,19 belirleyicilik ve %80,00 doğruluk değerleri elde edilmiştir (Uzun vd. 2018a).

Çizelge 3.3 K-Yakın Komşu algoritması ile siğil tedavi yönteminin başarı olup olmayacağını belirlemek için sınıflandırıcı performansları. Burada SEN duyarlılık, SPE belirleyicilik ve ACC doğruluk değerlerini vermektedir.

| K komşuluk sayısı | SEN (%) | SPE (%) | ACC (%) |
|-------------------|--------------|---------------|--------------|
| 1 | 31,58 | 84,51 | 73,34 |
| 3 | 21,06 | 91,55 | 76,67 |
| 5 | 10,53 | 92,96 | 75,56 |
| 7 | 15,79 | 97,19 | 80,00 |
| 9 | 5,27 | 97,19 | 77,78 |
| 11 | 0,00 | 100,00 | 78,89 |
| 13 | 0,00 | 100,00 | 78,89 |

3.2.5. Destek Vektör Makineleri

Destek Vektör Makineleri (SVM) kullanılan çalışmada, γ yaygınlık değeri 0,05'lik adımlarla 0,1 ile 1,0 arasındaki tüm değerler ve c merkez değeri -3,0 ile +3,0 arasındaki tüm değerler için sırasıyla test edilmiştir. Böylece 19 farklı yaygınlık ve 61 farklı merkez parametre kombinasyonlarıyla toplam 1159 kez bu algoritma koşturulmuştur. Bu denemeler sonucunda elde edilen sınıflandırıcı başarımları en yüksek doğruluk değerine göre duyarlılık için %47,37, belirleyicilik için %95,78 ve doğruluk için %85,46 olarak elde edilmiştir (Uzun vd. 2018b). Bu başarımlara ait iki farklı parametre kombinasyonu çizelge olarak verilmiştir (Çizelge 3.4).

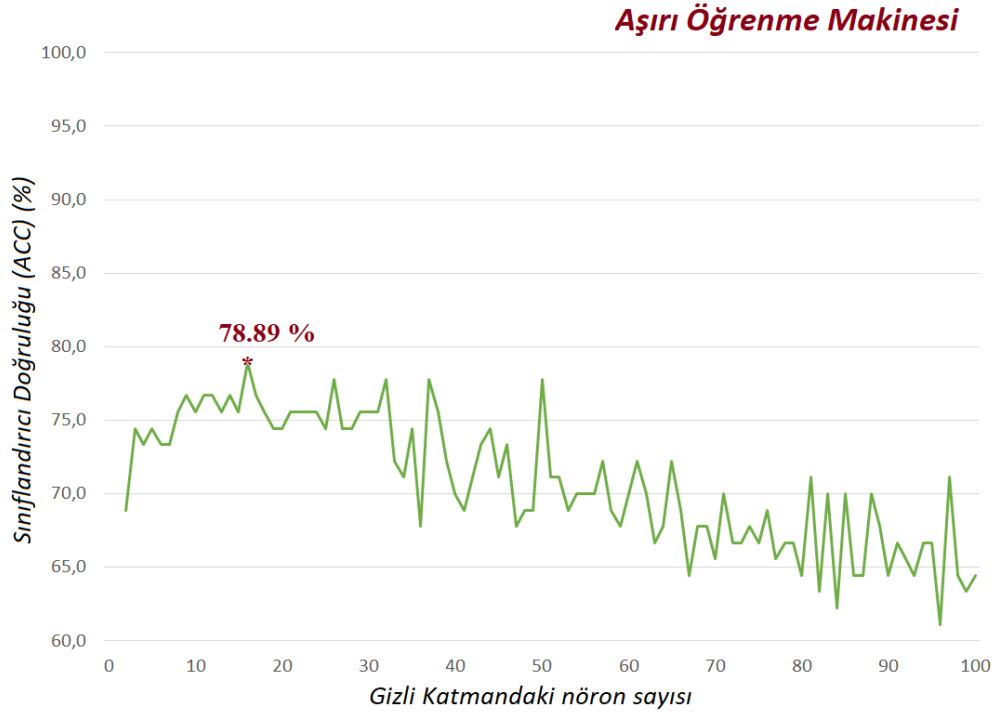
Çizelge 3.4 Destek Vektör Makineleri algoritması ile siğil tedavi yönteminin başarı olup olmayacağını belirlemek için sınıflandırıcı performansları. Burada SEN duyarlılık, SPE belirleyicilik ve ACC doğruluk değerlerini vermektedir.

| γ | c | SEN (%) | SPE (%) | ACC (%) |
|----------|------|---------|---------|---------|
| 0,60 | -2,0 | 47,37 | 95,78 | 85,46 |
| 0,75 | -2,4 | 47,37 | 95,78 | 85,46 |

3.2.6. Aşırı Öğrenme Makinesi

Aşırı Öğrenme Makinesi (ELM) algoritmasının gizli katmanında yer alan nöron sayısı 2 ile 100 arasındaki tam sayılar halinde sırayla denenmiştir. Ayrıca her bir nöron sayısı için algoritma 10 kez tekrarlanarak her bir gizli katmandaki nöron sayısına karşılık gelen en yüksek doğruluk değerleri grafik olarak düzenlenmiştir (Şekil 3.2). Böylece algoritma toplam 990 kez koşturulmuştur. En yüksek doğruluk değerini veren sınıflandırıcı başarımlarında %36,85

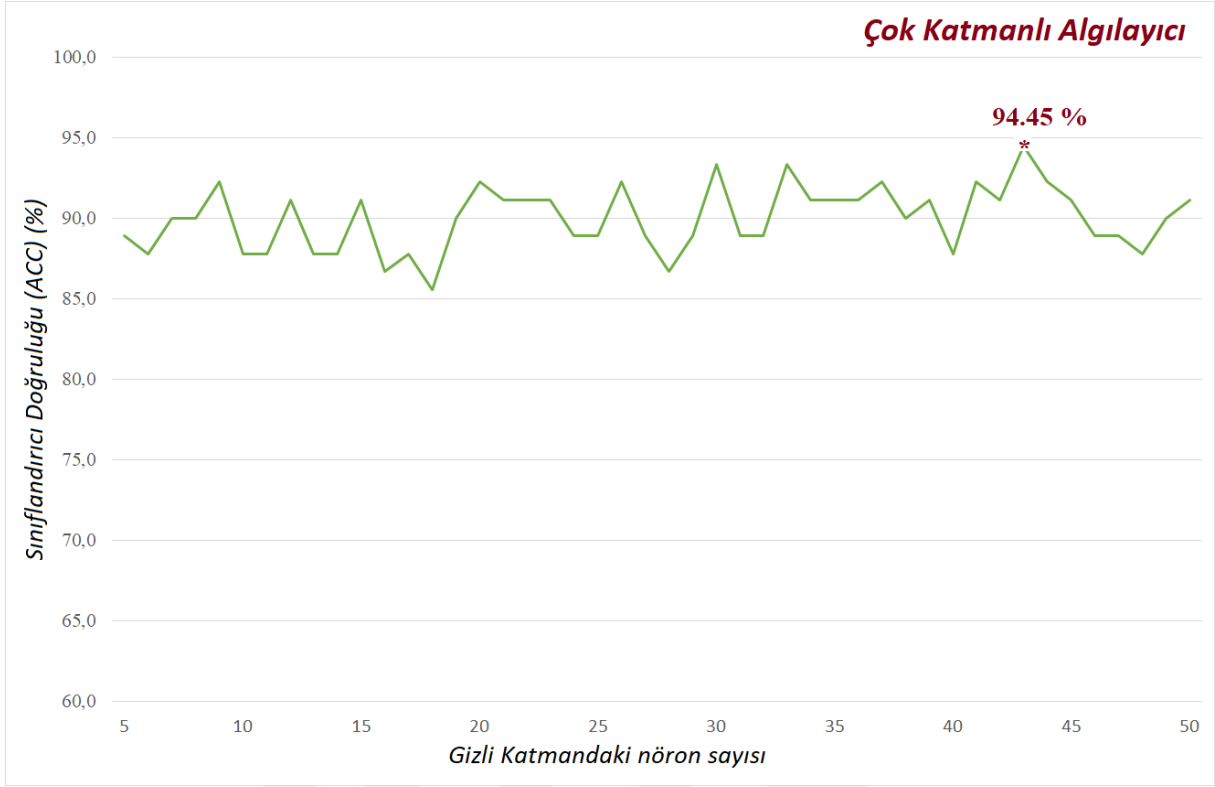
duyarlılık, %90,15 belirleyicilik ve %78,89 doğruluk değerleri elde edilmiştir (Uzun vd. 2018d).



Şekil 3.2 Aşırı Öğrenme Makinesi yapısındaki yapar sinir ağı modeli ile siğil tedavi yönteminin başarı olup olmayacağını belirlemek için sınıflandırıcı performansları. Gizli katman değerlerine göre elde edilen sınıflandırıcı performansları verilmektedir.

3.2.7. Çok Katmanlı Algılayıcı

Çok Katmanlı Algılayıcı (MLP) ağı yapısındaki yapay sinir ağı algoritmasının gizli katmanında yer alan nöron sayısı 5 ile 50 arasındaki tam sayılar halinde sırayla denenmiştir. Ayrıca her bir nöron sayısı için algoritma 10 kez tekrarlanarak her bir gizli katmandaki nöron sayısına karşılık gelen en yüksek doğruluk değerleri grafik olarak düzenlenmiştir (Şekil 3.3). Böylece algoritma toplam 460 kez çalıştırılmıştır. En yüksek doğruluk değerini veren sınıflandırıcı başarımında %78,95 duyarlılık, %98,60 belirleyicilik ve %94,45 doğruluk değerleri elde edilmiştir (Uzun vd. 2018d). Bu değerler hem literatürdeki hem de bu tez çalışmasındaki en yüksek değerlerdir.



Şekil 3.3 Çok Katmanlı Algılayıcı yapısındaki yapar sinir ağı modeli ile siğil tedavi yönteminin başarı olup olmayacağını belirlemek için sınıflandırıcı performansları. Gizli katman değerlerine göre elde edilen sınıflandırıcı performansları verilmektedir.



BÖLÜM 4

TARTIŞMA

Elde edilen sonuçların literatürdeki diğer benzer çalışmalar ile kıyaslamasının yapılması için literatür karşılaştırma tablosu oluşturulmuştur (Çizelge 4.1). Bu tez çalışması dışında yayımlanmış ve sınıflandırıcı başarımlarını içeren sadece iki adet çalışma Khozeimeh et al. (2017a), Tanyıldızı vd. (2018) bulunduğu için, tablo yazarın ve danışmanlarının bu tez kapsamında çıkan yayınlarının Uzun vd. (2018a), Uzun vd. (2018b), Uzun vd. (2018c), Uzun vd. (2018d) karşılaştırılması şeklinde görünmektedir. Bu tabloya giren sınıflandırıcıların kendi algoritmalarında yer alan farklı parametrelere göre sınıflandırıcı başarımları bulunmasına rağmen, o sınıflandırıcı elde edilen en yüksek sınıflandırıcı başarımları listelenmiştir. Ayrıca, eğer ilgili çalışmada belirtilmişse, sınıflandırıcıların bu başarımları elde edilirken kullanılan parametreleri algoritma isminin hemen yanında parantez içinde verilmiştir.

Çizelge 4.1 Seçilen siğil tedavi yönteminin başarılı olup olmayacağını belirlemek için kullanılan sınıflandırıcı performansları. Burada SEN duyarlılık, SPE belirleyicilik ve ACC doğruluk değerlerini vermektedir. Ayrıca * ile işaretli çalışmalar iki veri seti ayrı olarak gerçekleştirildiklerinden ortalamaları alınarak verilmiştir.

| Çalışma | Yöntem | SEN (%) | SPE (%) | ACC (%) |
|-------------------------|-------------------------------------|--------------|--------------|--------------|
| Uzun vd. 2018a | Naive Bayes | 68,43 | 67,61 | 67,78 |
| Uzun vd. 2018d | Aşırı öğrenme makinesi (16) | 36,85 | 90,15 | 78,89 |
| Uzun vd. 2018a | En yakın komşu (7) | 15,79 | 97,19 | 80,00 |
| Uzun vd. 2018c | Lojistik regresyon (10 ve 14) | 26,32 | 94,37 | 80,00 |
| Tanyıldızı vd. 2018* | Naive Bayes | 48,50 | 91,00 | 80,00 |
| Tanyıldızı vd. 2018* | En yakın komşu | 56,00 | 86,50 | 80,00 |
| Khozeimeh et al. 2017a* | Fuzzy kural tabanlı öğrenme | 87,00 | 74,00 | 81,67 |
| Tanyıldızı vd. 2018* | Destek vektör makineleri | 45,50 | 92,50 | 83,33 |
| Tanyıldızı vd. 2018* | Çok katmanlı algılayıcı | 66,50 | 88,50 | 84,44 |
| Uzun vd. 2018b | Destek vektör makineleri | 47,37 | 95,78 | 85,46 |
| Tanyıldızı vd. 2018* | K-Star | 65,00 | 91,00 | 85,55 |
| Uzun vd. 2018c | Karar ağacı (14) | 52,63 | 94,37 | 85,56 |
| Tanyıldızı vd. 2018* | Karar ağacı | 72,50 | 91,00 | 87,78 |
| Tanyıldızı vd. 2018* | Random Forest | 72,50 | 96,00 | 89,44 |
| Uzun vd. 2018d | Çok Katmanlı Algılayıcı (43) | 78,95 | 98,60 | 94,45 |

Bu listeye göre, literatürdeki en yüksek sınıflandırıcı başarısına bu tez çalışmasında kullanılan çok katmanlı algılayıcı yapısındaki gizli katmanında 43 hesaplama biriminin kullanıldığı yapay sinir ağı modelinde ulaşılabilmektedir. Buna göre, hastaların siğil tedavisinde kullanılması için seçilen immünoterapi ve kriyoterapi yöntemlerinin başarılı olup olmayacağı yüzde 94,45 doğruluk oranıyla tespit dileyebilmiştir. Sonuç olarak, bu çalışma mevcut çalışmalardan daha yüksek bir sınıflandırıcı duyarlılığına, belirleyiciliğine ve doğruluğuna sahiptir. Böylece seçilecek siğil tedavi yönteminin başarılı olup olmayacağını çok yüksek bir doğrulukla tahmin eden bir yöntem önerilmiştir.

Bu tez çalışmasında herhangi bir öznitelik seçimi algoritması kullanılmamıştır. İlerleyen çalışmalarda literatürde mevcut olan öznitelik seçimi yöntemleri kullanılabilir. Sadece 7 öznitelik mevcut olduğu için genetik algoritma gibi sezgisel yöntemler kullanılması yerine tüm olasılıkların denemesi tercih edilebilir.

Ayrıca literatürde mevcut farklı sınıflandırıcı modellerinin bulunmasına rağmen belirli sayıdaki yüksek başarımlar verdiği bilinen ve literatürde sık kullanıma sahip algoritmalar denenmiştir. Diğer mevcut sınıflandırıcı algoritmalarının denemesi ile daha yüksek başarımlara ulaşılabilir.

KAYNAKLAR

- Albayrak A S ve Koltan Yılmaz Ş** (2009) Veri Madenciliği: Karar Ağacı Algoritmaları ve İMKB Verileri Üzerine Bir Uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 14 (1): 31-52.
- Alp Avcı G ve Bozdayı G** (2013) İnsan Papilloma Virüsü. *Kafkas Tıp Bilimleri Dergisi*, 3 (3): 136-144.
- Arıcan Ö** (2004) Verrukalar'da Güncel Tedavi. *Dermatose*, 3 (3): 153-159.
- Ayhan S ve Erdoğan Ş** (2014) Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü için Çekirdek Fonksiyonu Seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 9 (1): 175-198.
- Aytekin S** (1994) Human Papilloma Virüs Enfeksiyonlarında İnterferon Alfa 2A Etkinliğinin Araştırılması. *Tıpta Uzmanlık Tezi*, Dicle Üniversitesi, Tıp Fakültesi, Dermatoloji Anabilim Dalı, Diyarbakır, Türkiye, 41 s.
- Bircan H** (2004) Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2: 185-208.
- Bunte K, Biehl M, Jonkman M F and Petkov N** (2011) Learning Effective Color Features for Content Based Image Retrieval in Dermatology. *Pattern Recognition*, 44: 1892-1902.
- Byun H and Lee S W** (2003) A Survey of Pattern Recognition Applications of Support Vector Machines. *International Journal of Pattern Recognition and Artificial Intelligence*, 17 (3): 459-486.
- Clifton M M, Johnson S M, Roberson P K, Kincannon J and Horn T D** (2003) Immunotherapy for Recalcitrant Warts in Children Using Intralesional Mumps or Candida Antigens. *Pediatric Dermatology*, 20 (3): 268-271.
- Coşkun C ve Baykal A** (2011) Veri Madenciliğinde Sınıflandırma Algoritmalarının Bir Örnek Üzerinde Karşılaştırılması, *XIII. Akademik Bilişim Konferansı (AB'11)*, 2-4 Şubat 2011, Malatya, Türkiye, 51-58.
- Cover T M and Hart P E** (1967) Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, 13 (1): 21-27.
- Çarman K B, Sağlam H, Ekici A ve Arslantaş D** (2013) Eskişehir'de Bir Grup Okul Çocuğunda Siğil Sıklığı. *Türkiye Klinikleri Journal of Pediatrics*, 22 (2): 66-69.

KAYNAKLAR (devam ediyor)

- Danacı M, Çelik M ve Akaya A E** (2010) Veri Madenciliği Yöntemleri Kullanılarak Meme Kanseri Hücrelerinin Tahmin Ve Teşhisi, *ASYU'2010: Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu*, 21-24 Haziran 2010, Kayseri, Türkiye, 9-12
- Dener M, Dörterler M ve Orman A** (2009) Açık Kaynak Kodlu Veri Madenciliği Programları: WEKA'da Örnek Uygulama, *XI. Akademik Bilişim Konferansı*, 11-13 Şubat 2009, Şanlıurfa, Türkiye, 1-11.
- Duda R O, Hart P E and Stork D G** (2000) *Pattern Classification*. 2nd edition, Wiley, New York, USA, 688 pp.
- Flores E and Scharcanski** (2016) Segmentation of Melanocytic Skin Lesions Using Feature Learning and Dictionaries. *Expert Systems with Applications*, 56: 300-309.
- Gamil H, Elgharib I, Nofal A and Abd-Elaziz T** (2010) Intralesional Immunotherapy of Plantar Warts: Report of a New Antigen Combination. *Journal of the American Academy of Dermatology*, 63 (1): 40-43.
- Giotis I, Molders N, Land S, Biehl M, Jonkman M F and Petkov N** (2015) MED-NODE: A Computer-Assisted Melanoma Diagnosis System Using Non-Dermoscopic Images. *Expert Systems with Applications*, 42: 6578-6585.
- Horn T D, Johnson S M, Helm R M and Roberson P K** (2005) Intralesional Immunotherapy of Warts with Mumps, Candida, and Trichophyton Skin Test Antigens: A Single-blinded, Randomized, and Controlled Trial. *Archives of Dermatology*, 141 (5): 589-594.
- Hoşrik M E** (2010) Dua ve Plasebonun Siğiller Üzerindeki Etkisi. *Yüksek Lisans Tezi*, Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Felsefe ve Din Bilimleri Anabilim Dalı, Ankara, Türkiye, 205 s.
- Huang G B, Zhoui H, Ding X and Zhang R** (2004) Extreme Learning Machine For Regression And Multiclass Classification, *2004 IEEE International Joint Conference on Neural Networks*, Budapest, Hungary, 42 (2): 513-529
- Huang G B, Zhu Q Y and Siew C K** (2006) Extreme Learning Machine: Theory and Applications. *Neurocomputing*, 1-3: 489-501.
- Huang G B, Zhu Q Y and Siew C K** (2012) Extreme Learning Machine: A New Learning Scheme Of Feedforward Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42 (2): 513-529.
- İşler Y and Kuntalp M** (2007) Combining Classical HRV Indices With Wavelet Entropy Measures Improves to Performance in Diagnosing Congestive Heart Failure. *Computers in Biology and Medicine*, 37 (10): 1502-1510.

KAYNAKLAR (devam ediyor)

- İşler Y and Kuntalp M** (2010) Heart Rate Normalization in the Analysis of Heart Rate Variability in Congestive Heart Failure. *Proceedings of the IMechE Part H: Journal of Engineering in Medicine*, 224 (3): 453-463.
- İşler Y ve Narin A** (2012) WEKA Yazılımında k-Ortalama Algoritması Kullanılarak Konjestif Kalp Yetmezliği Hastalarının Teşhisi. *Süleyman Demirel Üniversitesi Teknik Bilimler Dergisi*, 2 (4): 21-29.
- İşler Y, Narin A and Özer M** (2015) Comparison of the Effects of Cross-validation Methods on Determining Performances of Classifiers Used in Diagnosing Congestive Heart Failure. *Measurement Science Review*, 15 (4): 196-201.
- İşler Y, Narin A, Özer M and Perc M** (2019) Multi-Stage Classification of Congestive Heart Failure based on Short-Term Heart Rate Variability. *Chaos, Solitons & Fractals*, 118: 145-151.
- Jain S, Jagtap V and Pise N** (2015) Computer Aided Melanoma Skin Cancer Detection Using Image Processing. *Procedia Computer Science*, 48: 735-740.
- James W D, Berger T G and Elston D M** (2011) *Viral Diseases. Andrews's Disases of the Skin: Clinical Dermatology*. 11th edition, Saunders Elsevier, Toronto, 968 pp.
- Kaya T** (2015) Makine Öğrenme Yöntemleri ile Trafik Kazaları için Risk Tahmini Yapabilen Web Tabanlı Bir Yazılım. *Yüksek Lisans Tezi*, Gazi Üniversitesi, Bilişim Üniversitesi, Bilgisayar Eğitimi Anabilim Dalı, Ankara, 90 s.
- Khozeimeh F, Alizadehsani R, Roshanzamir M, Khosravi A, Layegh P and Navahandi S** (2017a) An Expert System for Selecting Wart Treatment Method. *Computers in Biology and Medicine*, 81: 167-175.
- Khozeimeh F, Jabbari Azad F, Mahboubi Oskouei Y, Jafari M, Tehranian S, Alizadehsani R and Layegh P** (2017b) Intralesional Immunotherapy Compared to Cryotherapy in the Treatment of Warts. *International Journal of Dermatology*, 56 (4): 474-478.
- Khurshid K, Ali U and Pal S S** (2009) Role of Candida Antigen in Treatment of Viral Warts: A Placebo-controlled Study. *Journal of Pakistan Association of Dermatologists*, 19 (1): 146-150.
- Koçer S and Canal M R** (2011) Classifying Epilepsy Diseases Using Artificial Neural Networks And Genetic Algorithm. *Journal of Medical Systems*, 35 (4): 489-498.
- Kudyba S** (2004) *Managing Data Mining: Advice from Experts*. CyberTech Publishing, Hershey PA, USA 260 s.
- Lamminen H, Ruohonen K and Uusitalo H** (2001) Visual Tests for Measuring the Picture Quality of Teleconsultations for Medical Purposes. *Computer Methods and Programs in Biomedicine*, 65 (2): 95-110.

KAYNAKLAR (devam ediyor)

- Lipke M M** (2006) An Armamentarium of Wart Treatments. *Clinical Medicine & Research*, 4 (4): 273-293.
- Maronn M, Salm C, Lyon V and Galbraith S** (2008) One-year Experience with Candida Antigen Immunotherapy for Warts and Molluscum. *Journal Pediatrics Dermatology*, 25 (2): 189-192.
- McGibbon D** (2006) Rook's Textbook of Dermatology (7th edition). *Clinical and Experimental Dermatology*, 31 (1): 178-179.
- Mirzaalian H, Lee T K and Hamarneh G** (2016) Skin Lesion Tracking Using Structured Graphical Models. *Medical Image Analysis*, 27: 84-92.
- Narin A** (2013) Çoklu Sınıflandırıcı Sistemleri ile Konjestif Kalp Yetmezliği Hastalarının Teşhisi. *Yüksek Lisans Tezi*, Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Zonguldak, Türkiye, 100 s.
- Narin A, İşler Y and Özer M** (2014a) Investigating the Performance Improvement of HRV Indices in CHF using Feature Selection Methods based on Backward Elimination and Statistical Significance. *Computers in Biology and Medicine*, 45: 72-79.
- Narin A, İşler Y and Özer M** (2014b) Konjestif Kalp Yetmezliği Teşhisinde Kullanılan Çapraz Doğrulama Yöntemlerinin Sınıflandırıcı Performanslarının Belirlenmesine Olan Etkilerinin Karşılaştırılması. *Dokuz Eylül Üniversitesi Mühendislik Bilimleri Dergisi*, 16 (48): 1-8.
- Narin A, İşler Y, Özer M and Perc M** (2018) Early Prediction of Paroxysmal Atrial Fibrillation based on Short-Term Heart Rate Variability. *Physica A*, 509: 56-65.
- Nofal A and Nofal E** (2010) Intralesional Immunotherapy of Common Warts: Successful Treatment with Mumps Measles and Rubella Vaccine. *Journal of the European Academy of Dermatology and Venereology*, 24 (10): 1166-1170.
- Oliveira R B, Marranghello N, Pereira A S and Tavares J M R S** (2016) A Computational Approach for Detecting Pigmented Skin Lesions in Macroscopic Images. *Expert Systems with Applications*, 61: 53-63.
- Öcal S H** (2009) Viral Siğillerde Kriyoterapi ile %0,7 Kantaridin Solüsyon Tedavisinin Etki ve Yan Etkilerinin Karşılaştırılması. *Tıpta Uzmanlık Tezi*, Ondokuz Mayıs Üniversitesi, Tıp Fakültesi, Dermatoloji Anabilim Dalı, Samsun, Türkiye, 76 s.
- Özarıslan S ve Barışçı N** (2014) Öğrenci Performansının Veri Madenciliği ile Belirlenmesi, *ISITES'2014: 2nd International Symposium on Innovative Technologies in Engineering and Science*, 18-20 Haziran 2014, Karabük, Türkiye, 1958-1965.
- Özçift A** (2011) Biyomedikal Verilerin Akıllı Sistemler ile Sınıflandırma Başarımlarının Analizi. *Doktora Tezi*, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Elazığ, Türkiye, 105 s.

KAYNAKLAR (devam ediyor)

- Özdamar K** (2009) *Paket Programlar ve İstatistiksel Veri Analizi*. Kaan Kitapevi, Eskişehir, Türkiye, 522 s.
- Özkan Y** (2016) *Veri Madenciliği Yöntemleri*. 3. Baskı, ISBN: 978-975-6797-82-2, Papatya Yayıncılık, İstanbul, Türkiye, 240 s.
- Pan S, İplikçi S, Warwick K and Aziz T Z** (2012) Parkinsons Disease Tremor Classification- A Comparison Between Support Vector Machines And Neural Networks. *Expert Systems with Applications*, 39 (12): 10764-10771.
- Russel D W, McCann W A and Maroncelli H E** (2010) Candida and Subsequent Cell Mediated-panel Driven Intralesional Immunotherapy of Common Warts in Children and Adults. *Journal of Allergy and Clinical Immunology*, 125 (2): AB204.
- Sever S** (2005) Tedaviye Dirençli Verruka Vulgaris ve Verruka Plantaris Olgularında Skuarik Asit Dibütil Ester Tedavisinin Etkinliği. *Tıpta Uzmanlık Tezi*, Haydarpaşa Numune Eğitim ve Araştırma Hastanesi, Deri ve Zührevi Hastalıklar Kliniği Anabilim Dalı, İstanbul, Türkiye, 51 s.
- Shrivastava V K, Londhe N D, Sonawane R S and Suri J S** (2015) Exploring the Color Feature Power for Psoriasis Risk Stratification and Classification: A Data Mining Paradigm. *Computers in Biology and Medicine*, 55: 54-68.
- Silahtaroglu G** (2008) *Kavram ve Algoritmalarıyla Veri Madenciliği*. Papatya Yayıncılık, İstanbul, Türkiye, 304 s.
- Silverberg N B, Lim J K, Paller A S and Mancini A J** (2000) Squaric Acid Immunotherapy for Warts in Children. *Journal of the American Academy of Dermatology*, 42 (5 Pt 1): 803-808.
- Sun J and Li H** (2008) Data Mining Method for Listed Companies, Financial Distress Prediction. *Knowledge-Based Systems*, 21 (1): 1-5.
- Şekeroğlu S** (2010) Hizmet Sektöründe Bir Veri Madenciliği Uygulaması. *Yüksek Lisans Tezi*, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Mühendislik Yönetimi Anabilim Dalı, İstanbul, Türkiye, 135 s.
- Tanyıldızı E, Karabatak M, Yıldırım G ve Özpolat Z** (2018) Siğil Tedavisinde Sınıflandırma Algoritmalarının Performans Analizi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 30 (2): 249-256.
- URL-1** <http://www.ntu.edu.sg/home/egbhuang/elm_codes.html>, Ziyaret tarihi: 14.12.2018.
- Uzun R, İşler Y and Toksan M** (2018d) Multi-Layer Perceptron Based Prediction of the Success of Wart Treatment Methods. *Proceedings of IMechE Part H: Journal of Engineering in Medicine*, under review, 78-91.

KAYNAKLAR (devam ediyor)

- Uzun R, İşler Y ve Toksan M** (2018a) Naive Bayes ve En Yakın k Komşu Sınıflandırıcıları ile Siğil Tedavi Yöntemi Seçimi, *SIU'2018: 26. IEEE Sinyal İşleme ve İletişim Uygulamaları Kurultayı*, 2–5 Mayıs 2018, Çeşme, İzmir, Türkiye, 1-4.
- Uzun R, İşler Y ve Toksan M** (2018b) Siğil Tedavi Yöntemlerinin Başarısının Tahmininde Destek Vektör Makineleri Kullanımı, *ASYU'2018: Akıllı Sistemlerde Yenilikler ve Uygulamaları Konferansı*, 4–6 Ekim 2018, Adana, Türkiye, 62-71.
- Uzun R, Toksan M ve İşler Y** (2018c) WEKA Yazılım Paketinin Siğil Tedavi Yöntemlerinin Başarısının Tahmininde Kullanımı. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 699-708.
- Varol G** (1998) Yaygın Verruca Vulgaris ve Verruca Plantaris Tedavisinde Oral Rantidin'in Etkinliği. *Tıpta Uzmanlık Tezi*, Gülhane Askeri Tıp Akademisi, Deri ve Zührevi Hastalıklar Anabilim Dalı, Ankara, Türkiye, 72 s.
- Vasudevan H, Joshi A R, Shekokar N M, Sumithra R, Suhil M and Guru D S** (2016) Segmentation and Classification of Skin Lesions for Disease Diagnosis. *Procedia Computer Science*, 45: 2076-2085.
- Witten I H, Frank E and Hall M A** (2011) *Data Mining: Practical Machine Learning Tools And Techniques*. Elsevier, London, UK, 644 pp.

ÖZGEÇMİŞ

Mualla SAKARYA Aksaray'da doğdu; ilk ve orta öğretimini Aksaray'da tamamladı. 2012 yılında girdiği Kırıkkale Üniversitesi Elektrik Elektronik Mühendisliği bölümünden 2016 yılında mezun oldu. 2016-2018 yılları arasında Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik Elektronik Mühendisliği Anabilim dalında yüksek lisansını tamamladı. Şu anda özel bir firmada Elektrik-Elektronik Mühendisi olarak çalışmaktadır.

ADRES BİLGİLERİ:

Adres : Çiğdemtepe Mah. 1084 Sok. Saray Apt. No: 8/16 Yenimahalle/ANKARA.

Tel : (+90) 372 257 40 10

E-posta : muallatoksan@gmail.com