

**T.C.**  
**FIRAT ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**TEKNOLOJİ VE BİLGİ YÖNETİMİ ANABİLİM DALI**



**TIP'TA VERİ MADENCİLİĞİ UYGULAMALARI:**  
**YENİDOĞAN SEPSİSİ VERİ SETİ ANALİZİ**

**YÜKSEK LİSANS TEZİ**

**DANIŞMAN**  
**Dr. Öğr. Üyesi Mustafa ULAŞ**

**HAZIRLAYAN**  
**Aytaç TEKİN**

**ELAZIĞ-2018**

**T.C.**  
**FIRAT ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**TEKNOLOJİ VE BİLGİ YÖNETİMİ ANABİLİM DALI**

**TIP'TA VERİ MADENCİLİĞİ UYGULAMALARI:**  
**YENİDOĞAN SEPSİSİ VERİ SETİ ANALİZİ**

**YÜKSEK LİSANS TEZİ**

**DANIŞMAN**

**Dr. Öğr. Üyesi Mustafa ULAŞ**

**HAZIRLAYAN**

**Aytaç TEKİN**

Jürimiz, 13/07/2018 tarihinde yapılan tez savunma sınavı sonunda bu yüksek lisans tezini oy birliği / oy çokluğu ile başarılı saymıştır.

Jüri Üyeleri:

1. **Dr. Öğr. Üyesi Mustafa ULAŞ**
2. **Dr. Öğr. Üyesi Aytağ BOYACI**
3. **Dr. Öğr. Üyesi Musa ÇIBUK**

F. Ü. Sosyal Bilimler Enstitüsü Yönetim Kurulunun ..... tarih ve .....sayılı kararıyla bu tezin kabulü onaylanmıştır.

**Prof. Dr. Osman Ömer UMAR**  
**Sosyal Bilimler Enstitüsü Müdürü**

**ÖZET****Yüksek Lisans Tezi****Tıp'ta Veri Madenciliği Uygulamaları: Yenidoğan Sepsisi Veri Seti Analizi****Aytaç TEKİN****Fırat Üniversitesi****Sosyal Bilimleri Enstitüsü****Teknoloji ve Bilgi Yönetimi Anabilim Dalı****Elazığ – 2018, Sayfa: IX+47**

Gelişen Teknolojiler ile birlikte eskiden ölçülemeyen veya ölçülebilir olup da kayıt altına alınmayan birçok veri günümüzde kayıt altına alınarak veri tabanlarında saklanmaktadır. Klasik istatistik bilimi az sayıda veri içerisinde anlamlı veriyi çıkarmamıza imkân tanıyordu fakat modern teknoloji ile klasik istatistik metotlarının birleşmesi ile veri madenciliği yöntemleri ortaya çıkmıştır. Veri madenciliği yöntemleri ile artık birçok veri yığını arasından anlamlı veriyi tespit etmek, elde edilen anlamlı veriler ile karar destek sistemleri oluşturmak daha kolay bir hal almıştır.

Bu çalışmanın amacı veri madenciliği yöntemlerinin sağlık hizmetleri ile olan ilişkisine değinmektir. Hekimler adına karar verecek bir sistem, metot veya yöntem geliştirmek değildir. Sonuçta en iyi kararı hekimin kişisel gözlemleri ve tecrübesi oluşturmaktadır. Yoğun iş temposu ve veri karmaşası arasından hekime karar verme noktasında destek olacak akıllı bir sistem geliştirmek anlamında çalışmalar olacaktır. Bundan hareketle Yenidoğan Sepsis veri setine KNN Algoritması uygulanmış, toplam 128 Örnek içerisinde 121 örnek doğru sınıflandırılmıştır. KNN algoritmasının doğruluk oranı bu veri seti için %94.53 'dir. Yenidoğan Sepsis veri setine Naive Bayesian algoritması uygulanmış ve doğruluk oranı %93.73 olarak hesaplanmıştır.

**Anahtar kelimeler:** Tıp'ta Veri Madenciliği, Veri Madenciliği, Sepsis Veri Madenciliği, KNN, Naive Bayesian

**ABSTRACT****Master Thesis****Data Mining Applications in Medicine: Newborn Sepsis Data Set Analysis****Aytaç TEKİN****The University of Firat****The Institute of Social Science****The Department of Technology and Information Management****Elazığ – 2018, Page: IX+47**

With the developing technologies, many data that are not measured or measurable in the past but not recorded are now recorded and stored in the databases. Classical statistical knowledge has made it impossible to extract meaningful data from a small number of data, but data mining methods have emerged by combining classical statistical methods with modern technology. With data mining methods it is now easier to establish decision support systems with meaningful data obtained from meaningful data among many data stacks

The purpose of this study is to address the relationship between data mining methods and health services. It is not a system, method or method to decide on behalf of physicians. Ultimately, the best decision is the personal observation and experience of the physician. The intensive business will work in the sense of developing an intelligent system that will support the tempo of decision making and the complexity of data among the tempo. From this, the KNN algorithm was applied to the Sepsis data set and 121 samples were classified correctly in a total of 128 samples. The accuracy of the CNN algorithm is 94.53% for this data set. Naive Bayesian algorithm were applied to the Sepsis data set and the accuracy rate was calculated as 93.73.

**Keywords:** Data Mining in Medicine, Data Mining, Sepsis Data Mining, KNN, Naive Bayesian

## İÇİNDEKİLER

<b>ÖZET .....</b>	<b>II</b>
<b>ABSTRACT.....</b>	<b>III</b>
<b>İÇİNDEKİLER.....</b>	<b>IV</b>
<b>TABLolar LİSTESİ .....</b>	<b>VI</b>
<b>ŞEKİLLER LİSTESİ .....</b>	<b>VII</b>
<b>ÖNSÖZ .....</b>	<b>VIII</b>
<b>KISALTMALAR ve TERİMLER.....</b>	<b>IX</b>
<b>GİRİŞ .....</b>	<b>1</b>

### BİRİNCİ BÖLÜM

<b>1. VERİ, BİLGİ, VERİ MADENCİLİĞİ KAVRAMI .....</b>	<b>4</b>
1.1. Veri, Veri Ambarı, Veri tabanı Kavramları .....	4
1.2. Veri Madenciliği Kavramı .....	5
1.3. Veri Madenciliği Modelleri.....	5
1.4. Veri Madenciliği Kullanım Alanları .....	6

### İKİNCİ BÖLÜM

<b>2. VERİ MADENCİLİĞİ YÖNTEMLERİ .....</b>	<b>7</b>
2.1. Regrasyon Analizi .....	7
2.2. K-En Yakın Komşuluk (K Nearest Neighborhood).....	8
2.3. Kümeleme (Clustering).....	9
2.4. Karar Ağaçları .....	10
2.5. Birliktelik Kuralı .....	11
2.6. Bayes Teoremi ve Ağları .....	12
2.7. Veri Madenciliği Algoritmalarının Doğruluk Analizi İçin Kullanılan Yöntemler.....	14
2.8. Veri Madenciliği Analiz Yazılımları .....	15

### ÜÇÜNCÜ BÖLÜM

<b>3. VERİ MADENCİLİĞİ METODLARININ TIP ALANINDA KULLANILMASI.....</b>	<b>17</b>
3.1. Sağlık Hizmetlerinde Veri Kavramı.....	17
3.2. Tıpta Veri Madenciliği .....	17
3.3. Tıpta Veri Madenciliği Çözümleri .....	17
3.4. Tıpta Veri Madenciliği Yöntemleri Kullanılarak Yapılmış Çalışmalar.....	18

**DÖRÜNCÜ BÖLÜM**

<b>4. YENİDOĞAN SEPSİSİ VERİ SETİNİN TANIMI VE ELDE EDİLEMSİ .....</b>	<b>21</b>
4.1. Sepsis Kavramı.....	21
4.2. Sepsis Tanısı.....	21
4.3. Yenidoğan Sepsisi Veri Seti.....	22
4.4. Sepsis Veri Seti Analizi .....	22
4.4.1. Genel Veriler .....	22
4.4.2. Bebekte Tespit Edilen Mevcut Hastalıklar ve Riskler İlişkin Veriler .....	24
4.4.3. Kan Kültürü ve Üreme Sonucu İle İlişkili Veriler .....	28
4.4.4. Tanı, Skorlama Sistemleri İle İlgili Bulgular ve Testlere İlişkin Veriler .....	28
4.4.5. Bebeğe Ait Risk Faktörleri İle İlgili Veriler.....	29
4.4.6. Beslenme Bozukluğu İle İlgili Veriler .....	29
4.4.7. Anneye Ait Risk Faktörleri İle İlgili Veriler .....	29
4.4.7.1. Annenin Önceden Var Olan Hastalıkları İle İlgili Veriler.....	29
4.4.7.2. İntrapartum (Doğum Sırasında) Komplikasyonlar İlişkin Veriler.....	30
4.4.7.3. Obstetrik Komplikasyonlar İlişkin Veriler .....	31
4.4.8. Gelişmiş Laboratuvar Testleri İle İlgili Veriler.....	32
4.5. Veri Setindeki Sayısal Değerler .....	33

**BEŞİNCİ BÖLÜM**

<b>5. VERİ SETİNE UYGULANACAK OLAN VERİ MADENCİLİĞİ</b>	
<b>METOTLARI.....</b>	<b>35</b>
<b>SONUÇ VE ÖNERİLER.....</b>	<b>39</b>
<b>KAYNAKÇA.....</b>	<b>40</b>
<b>EKLER .....</b>	<b>46</b>
Ek 1. Orjinallik Raporu .....	46
<b>ÖZGEÇMİŞ .....</b>	<b>47</b>

**TABLolar LİSTESİ**

<b>Tablo 1.</b> Regrasyon Analizi Sebep-Sonuç İlişkisi.....	7
<b>Tablo 2.</b> Değişik zamanlar satın alınmış alışveriş sepetleri .....	11
<b>Tablo 3.</b> Alışveriş Sepetlerinin içerikleri kullanılarak Destek Sayısının Belirlenmesi..	12
<b>Tablo 4.</b> Karışıklık Matrisi .....	14
<b>Tablo 5.</b> Veri Seti Çalışma Grubu.....	22
<b>Tablo 6.</b> Veri Seti içerisindeki Laboratuvar Testleri.....	32
<b>Tablo 7.</b> Veri Seti içerisindeki Genel verilere ilişkin bazı sayısal değerler .....	33
<b>Tablo 8.</b> Veri Seti içerisindeki bazı hastalık verilerine ilişkin bazı sayısal değerler .....	34
<b>Tablo 9.</b> KNN ve Naive Bayesian Tabakalı çapraz doğrulama Özeti.....	36
<b>Tablo 10.</b> KNN ve Naive Bayesian Karışıklık Matrisi Değerleri .....	37
<b>Tablo 11.</b> KNN ve Naive Bayesian Sınıf Tarafından Detaylı Doğruluk .....	38

**ŞEKİLLER LİSTESİ**

<b>Şekil 1.</b> Veri ambarından, veri madenciliği ile bilgiye erişim.....	4
<b>Şekil 2.</b> A K-En Yakın Komşu Örneği.....	8
<b>Şekil 3.</b> B K-En Yakın Komşu Örneği.....	8
<b>Şekil 4.</b> C K-En Yakın Komşu Örneği.....	9
<b>Şekil 5.</b> Kümeleme Algoritması.....	9
<b>Şekil 6.</b> Karar Ağacı Örnek Uygulaması.....	10





## ÖNSÖZ

Eğitimim ve çalışmalarım süresince bilgi ve sabrıyla bana yol gösteren, zamanını ve desteğini esirgemeyen Sayın Danışmanım Dr. Öğr. Üyesi Mustafa ULAŞ' a, onlara ayıracağım zamanlarından feragat eden sevgili eşim ve çocuklarıma, Tez çalışmamızda kullanılmış olduğumuz verileri bizlerle paylaşan Uz. Dr. Fatma UZUN ve onun nezrinde danışmanı Prof. Dr. Erdal TAŞKIN 'a

Teşekkür ederim...

**ELAZIĞ-2018**

**Aytaç TEKİN**



**KISALTMALAR ve TERİMLER**

<b>EKG</b>	: Elektro Kardiyo Grafi
<b>GNU</b>	: Genel Kamu Lisansı
<b>IP</b>	: İntrapartum (Doğum Sırasında)
<b>KNN</b>	: K Nearest Neighborhood (En Yakın Komşu)
<b>OBS</b>	: Obstetrik Tıp'ta Doğum, Gebelik ve Loğusalık dönemini inceleyen bilim dalıdır.
<b>VERİ SETİ</b>	: Bir Evreni Temsil Eden Deneklerin Değişebilen Özelliği
<b>VM</b>	: Veri Madenciliği
<b>VTYS</b>	: Veri Tabanı Yönetim Sistemi
<b>GB</b>	: Gigabyte
<b>PC</b>	: Personel Computer

## GİRİŞ

Günümüzde teknolojinin gelişmesi ile birlikte kurum ve kuruluşlarda devasa veri ambarları ortaya çıkmıştır. Bu veri ambarlarının büyümesi içerisinde doğru ve anlaşılır bilgiyi bulmayı zorlaştırmıştır. İstatistik bilim dalı, veri ambarlarında biriken verilerin boyutunun artması ile yeni arayışlar ortaya çıkarmıştır. Gelişen bilişim teknolojileri ve klasik istatistik metotlarının birleşmesi ile artık milyarlarca veri arasından değerli bilgiye erişmek daha da kolaylaşmıştır. Bu anlamda veri madenciliği kavramı ortaya çıkmıştır. Veri madenciliği, anlamlı veri oluşturmak için istatistik, veri tabanı yöntemleri, yapay zeka, örüntü tanıma, uzman sistemler gibi bir çok alandan yararlanmaktadır.

Gelişen teknolojiler ve veri analizi yöntemleri sağlık alanını önemli ölçüde etkilemiştir. Daha önceden sınırlı sayıda parametre ile çalışan laboratuvar cihazları günümüzde çok sayıda parametre 'ye, radyolojide tek boyuttan birkaç boyuta, 8-16 kesit görüntülerden 128-256 kesit görüntülere ulaşılmıştır. Ölçülemeyen birçok durum artık ölçülebilir hale gelmiş, daha önemlisi geçmiş veriler üzerinden hareketle yeni verileri mukayese etmek daha çok önem kazanmıştır. Durumun böyle olması daha çok verinin oluşmasına sebep olurken artık sağlık profesyonellerine yardımcı olacak karar destek sistemlerinden faydalanmayı zorunlu kılmıştır.

Yenidoğan Sepsisi yaşamın ilk otuz gününde enfeksiyon bulguları ile ortaya çıkan bir durumdur. Bulguların görülmeye başlandığı zamana göre erken başlangıçlı, geç başlangıçlı ve çok geç başlangıçlı sepsis olarak ifade edilebilir. Erken yenidoğan sepsisi genellikle doğum kanalında veya doğum sırasında kazanılırken, Geç yenidoğan sepsisi ise doğum kanalında veya hastane, çevre, toplum, ev ortamından kazanılabilmektedir. Çok geç başlangıçlı sepsis ise genellikle düşük doğum ağırlığı ile dünyaya gelmiş bebeklerin uzun süre hastanede tedavi görmesine ve nadiren toplumdan kazanıldığı görülmüştür. (CENGİZ A.B.,2009)

Yenidoğan sepsisi epidemiyolojine bakılacak olursa Gelişmiş ülkelerde her canlı 1000 doğumdan 1-10 arasında görülürken bu sayı gelişmemiş olan ülkelerde 49-170 arasında olabilmektedir. Sepsis oluşumuna neden olan bakteriler coğrafi konum, ülke ve hastaneye göre farklılık göstermektedir. Erken yenidoğan sepsisine neden olan birçok risk faktörü bulunmaktadır. Fakat bu risk faktörlerinin birden fazla olması yenidoğan sepsisi olma olasılığını arttırmaktadır. Prematüre doğum, düşük doğum ağırlığı, annenin

doğum sularının erken gelmesi, Bebeği çevreleyen zarın enfeksiyon kapması, çoğul gebelik ve doğum sürecinde fazla vajinal muayenenin yapılması durumları, başlıca risk faktörleri arasında yer almaktadır. (CENGİZ A.B.,2009)

Tıpta veri madenciliği kullanımı birçok soruna çözüm getirmeye çalışmaktadır. Hacettepe Üniversitesi hastanesindeki sağlık kayıtları içerisinde gizli kalmış verileri aramak için, Acil servise başvuran hasta verilerini kullanarak senaryo belirlemek, hastane enfeksiyonların takibi ve kaynağını bulmak için, kalp sinyallerini inceleyerek (EKG) fazla gürültülü veriyi veri madenciliği yöntemleri ile temizleyerek daha net bir sonuç elde etmek veya akciğer hastalığının tanısını koymada patoloji yapılmadan yada patoloji sonuçlarını beklemeden önden bir tahmin modeli oluşturmak için kullanılabilir. (GÜLTEPE E. ve Ark.,2013)

Amerika'nın Vanderbilt Üniversitesi, Monroe Carell Jr. Çocuk Hastanesi yenidoğan yoğun bakım ünitesine 1 Ocak 2006 tarihinden başlamak kaydıyla 18 aylık süre içerisinde başvuran 299 bebek geç başlangıçlı sepsis konusunda değerlendirilmiştir. Tüm Klinik, Laboratuvar ve Mikrobiyoloji verileri toplanarak dokuz farklı makine öğrenmesi algoritması bu veriler üzerine uygulanmıştır. (S.MANI ve Ark.,2013)

Amerika'nın Kaliforniya eyaletindeki Davis Sağlık Sistemleri Üniversitesinde Sistematik inflamatuvar yanıt sendromu kriterlerinkinden en az ikisini taşıyan 741 erişkin hastanın hastane bilgi sistemindeki kayıtları, bulguları, laboratuvar testlerinin sepsis hastalığına bağlı ölüm ile olan ilişkisi araştırılmıştır. Bayes, Makrov Modeller, Destek Vektör Makinası, Gauss Karışı algoritmaları kullanılarak ölüm tahmini, laktat seviyesinin tahmini üzeninde çalışmalar yapılmıştır. (GÜLTEPE E. ve Ark.,2013)

Portekizdeki Centro Hospitalar Tâmega e Sousa hastanesinin doğum bölümünde yapılan çalışmada veri madenciliği algoritmaları ile yenidoğan bebeklerdeki sarılığa yakalanan 231 hastadan klinik verileri ve 8 saatlik zaman aralıkları ile Sarılık Ölçer cihazı kullanılarak toplanan verilerden veri seti oluşturulmuştur.. WEKA programı yardımıyla J48, C4.5, Naive Bayes, Sıralı minimum optimizasyon algoritmaları uygulanarak bu algoritmalar doğruluk, duyarlılık ve özgünlük açısından bir birleriyle karşılaştırılmıştır. (OLIVEIRA A. ve Ark., 2012)

Veri madenciliği yöntemleri Elektronik Ticaret, Toptan ve Perakende Pazarlama sektörlerinde başarılı uygulamaları olduğu gibi diğer sektörlerde de uygulamalar geliştirilebilmektedir. Bu sektörlerden biride sağlık sektörüdür. Tıbbi veriler içeriği

bakımında çok zengin ve keşfedilmeyi bekleyen birçok bilgiyi barındırmaktadır. KNN ve Bayes sınıflandırma metotları kullanılarak 909 hasta verisi üzerinde 13 nitelik değerlendirilerek Akıllı Kalp Hastalığı Tahmin Sistemi geliştirmek için kullanılmıştır. Weka 3.6.0 versiyonunda veri tutarsızlıkları giderilerek yapılan çalışmada Naive Bayesian'da doğruluk oranı %96.5 iken KNN'de doğruluk oranı %88.3 olarak saptanmıştır (ANSARİ U.,2011). Günümüz sağlık kuruluşlarında üretilen ve tıbbi cihazlardan toplanan verinin büyük miktardaki artışı bu verilerden çıkarım yapılması için otomatik bir yol gerekliliğini ortaya çıkarmıştır. Veri madenciliği teknikleri kullanarak veri yığınları arasında ilginç ve yararlı bilgiye erişmek mümkündür. Bu şekilde elde edilen değerli bilgiler iş verimliği ve karar verme süreçlerinin kalitesi arttırmak ve hızlandırmak için kullanılabilir. Veri madenciliği teknikleri kullanılarak sağlık hizmetlerinde tahmin, karar verme, sağlık hizmetlerinin maliyetini azaltmak bulaşıcı hastalıkları erkenden tespit etmek, hastaların eğilimleri incelenerek davranışların tahmini, uygulanan tanı ve tedavilerin modellenmesi veri madenciliği teknikleri ile olaylara farklı bakış açılarından bakmamıza gizli kalmış değerli bilgilere erişmemize olanak sağlayacaktır. (MILOVIC B., 2012)

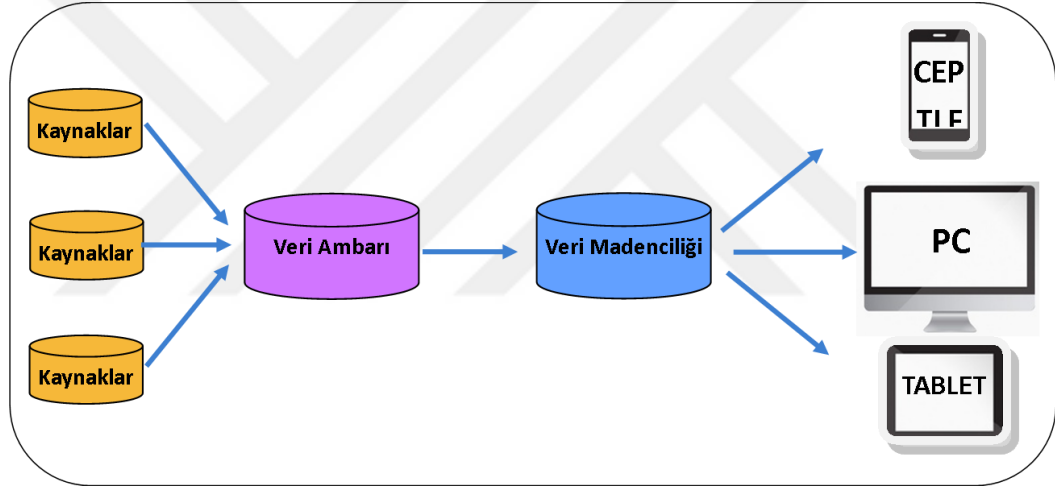
Bu çalışmanın amacı veri madenciliği yöntemlerinin sağlık hizmetleri ile olan ilişkisine değinmektir. Hekimler adına karar verecek bir sistem, metot veya yöntem geliştirmek değildir. Sonuçta en iyi kararı hekimin kişisel gözlemleri ve tecrübesi oluşturmaktadır. Yoğun iş temposu ve veri karmaşası arasından hekime karar verme noktasında destek olacak akıllı bir sistem geliştirmek anlamında çalışmalar olacaktır. Bundan hareketle Sepsis veri setine KNN Algoritması uygulanmış, toplam 128 Örnek içerisinde 121 örnek doğru sınıflandırılmıştır. KNN algoritmasının doğruluk oranı bu veri seti için %94.53 'dir. Sepsis veri setine Naive Bayesian algoritmaları uygulanmıştır ve doğruluk oranı 93.73 olarak hesaplanmıştır.

# BİRİNCİ BÖLÜM

## 1. VERİ, BİLGİ, VERİ MADENCİLİĞİ KAVRAMI

### 1.1. Veri, Veri Ambarı, Veri tabanı Kavramları

Genelde bu kavramlar bir birine karıştırılsa da veri için Bilgi'nin işlenmemiş ham hali diyebiliriz. Yığınlar halinde duran veri blokları çeşitli aşamalardan geçerek anlaşılır hale dönüşmektedir. İşte bu anlaşılır durum bilgi olarak ifade edilmektedir. Verinin anlam kazanabilmesi için bazı enformasyon teknikleriyle işlenerek bilgiye dönüşmesi gerekmektedir. Veri ambarı ise gün geçtikçe artan veri topluluklarının birikmesiyle oluşturdukları yığınlardır.



Şekil 1. Veri ambarından, veri madenciliği ile bilgiye erişim

Şekil-1'deki gibi veri ambarı belirli kaynaklardan sürekli olarak beslenmektedir. Veri ambarında biriken verilere çeşitli teknikler ve algoritmalar uygulanarak anlamlı bilgi halinde teknolojik cihazlara yansımaları gözlemleyebiliriz.

Veri tabanı bir birleriyle ilişki içerisinde olan verilerin belirli bir düzene göre depolandığı fiziksel ve mantıksal olarak var olan sistemlerin genel ismidir. Veri tabanı, 1980'li yıllarla birlikte bilişim dünyasına girmiş bir kavram olup çeşitli veri tabanı programları olmasına rağmen çalışma mantıkları bir birlerine yakındır. Günümüzde kullandığımız hemen hemen tüm uygulamalar birer veri tabanına sahiptir. Basit bir web sayfası, uluslararası bir banka, çok uluslu şirketler çoğu veri tabanı uygulamasına

ihtiyaç duymaktadır. (ŞENTÜRK A.,2006:6) Verilerin fiziksel olarak hafızadaki durumlarını gözlemek, yeni veri tabanı oluşturmak, kullanıcı erişimlerini yetkilendirmek, yedekleme işlemlerinin uygulanması işlemlerinin yürütüldüğü sistemlere Veri Tabanı Yönetim Sistemleri (VTYS) denir. Günümüzde yaygın olarak kullanılan VTYS'leri arasında MS SQL, My SQL, Sybase gelmektedir. (ŞENTÜRK A.,2006:7)

### **1.2. Veri Madenciliği Kavramı**

Yaşanan teknolojik gelişmeler ile birlikte dünyadaki veri miktarında hızla bir büyüme görülmektedir. Artık yapılan tüm alışveriş, bankacılık, yazışma, mesajlaşma, otelcilik, sağlık hizmetleri v.b. tüm işlemler gelişen ve maliyetleri düşen depolama üniteleri sayesinde elektronik olarak kayıt altına alınmaktadır. Bu kayıt altına alma işlemleri çok büyük veri yığınları oluşturmaktadır. Bu büyük veri yığınları beraberinde yeni bir sorunla birlikte gelmektedir bu samanlıktaki işi nasıl bulacağız?

Bilişim teknolojilerinin bu kadar gelişmediği dönemlerde karar verici yöneticiler bilgi kıtlığı arasında karar vermemeye çalışırken şuanda bu kıtlık yerini bolluğa bırakmış ve yöneticiler karar noktasında binlerce seçenekle karşı karşıyadır. (ŞENTÜRK A.,2006:1) Bu çalışmanın konusu olmamasına rağmen bu aşırı veri yoğunluğu beraberinde veri güvenliği kavramını da ortaya çıkarmış ve saklanan bu verilerin kötü amaçlar için kullanılmaması için güvenlik tedbirleri de geliştirmek ortaya çıkmıştır.

İşte burada veri madenciliği kavramı ortaya çıkmakta bu veri yığınları ve veri bolluğu içerisinde değerli, anlamlı bilgiyi bulmak için veri madenciliği (data mining) yöntemleri devreye girmektedir. Veri madenciliği yöntemleri ile veriler arasındaki ilişki ortaya çıkarılarak geleceğe yön vermeye çalışılmaktadır. Veri Madenciliği, büyük veri yığınları içerisinde gelecek ile ilgili tahminler yapabilmemize yardımcı olmaktadır. (PEKTAŞ, A.O.,2013:100)

### **1.3. Veri Madenciliği Modelleri**

Veri madenciliğinde kullanılan modeller ikiye ayrılmaktadır. Bunlardan ilki Tanımlayıcı Modellerdir. Tanımlayıcı modellerde amaç, karar vermeye rehberlik etmede kullanılacak olan verilerdeki örüntüyü yakalayabilmektir Birliktelik Kuralı, Ardışık zamanlı Örüntüler bu model kapsamında yer almaktadır..(GÜR SOY U.,2012:5)

Tahmin edici modellerde sonuçlardan yola çıkarak bilinen verileri işleyerek bir model geliştirilir bundan sonra gelecek olan veriler için karar verebilecek, sonucunu tahmin edebilecek bir model geliştirmeyi hedeflemektedir. Yapay Sinir Ağları, Genetik Algoritmaları, Karar Ağaçları bu model kapsamında yer almaktadır.

#### **1.4. Veri Madenciliği Kullanım Alanları**

Genel olarak kullanım alanları; analiz yapma, karar verme noktasında destek, çapraz inceleme, Risk analizi yapma, müşteriler arasındaki tercih benzerliklerinin tespiti, öngörü, sahtekârlıkların tespiti, belgeler arası benzerlik incelemesi, kurum ve kuruluşların kaynaklarının en iyi şekilde yönetilmesi ve kullanılması, Geçmiş ve mevcut verilerin analizini yaparak geleceğe yönelim çıkarım ve tahminlerde bulunulmasıdır.

Pazarlama alanında ise müşterilerin satın alma alışkanlıklarının tespit edilmeye çalışılması, Müşterilerin yaşadıkları çevreye, nüfus'a göre bir birleri arasındaki bağlantıların bulunması, Mevcut müşterileri elde tutmak için politika geliştirmekte, yeni müşterilerin kazanılması, Pazar sepeti analizi (Market Basket Analysis), Müşteri ilişkileri yönetimi (Customer Relationship Management), Müşteri değerlendirme (Customer Value Analysis), Satış tahmini (Sales Forecasting).

Bankacılık ve Sigortacılık alanında ise Farklı finansal göstergeler arasında gizli korelasyonların bulunması, kredi kartı dolandırıcılıklarının tespiti, kredi kartı harcamalarına göre müşteri gruplarının belirlenmesi, kredi taleplerinin değerlendirilmesi. Yeni poliçe talep edecek müşterilerin tahmin edilmesi, Sigorta dolandırıcılıklarının tespiti, Riskli müşteri örüntülerinin belirlenmesi.

Tıp alanında ise belirli hastalığa yakalanmış olan hastaların ortak özelliklerinin belirlenmesi, hastalara uygulanan tedavilerin sonuçlarının tahmin edilmesi, Hastanenin finansal analizi yapılarak maliyetlerin tahmin edilmesi, salgın hastalıkların ve ölüm oranlarının tahmin edilmesi, Geçmiş verilerden yararlanarak hastalar hakkında tahminlerde bulunma Bunun yanı sıra veri madenciliği yöntemleri eğitim alanında da çok sayıda kullanım örneği bulunmaktadır. Öğrenci ve Akademisyenlerin başarımlarını, Öğrenci profillerini tespit etmekte, öğrenci başarısını etkileyen faktörlerin analizinde veri madenciliği yöntemlerinden yararlanılmaktadır. (BAYKAL A.,2006:95-107)



## İKİNCİ BÖLÜM

### 2. VERİ MADENCİLİĞİ YÖNTEMLERİ

Veri madenciliğinde kullanılmakta yöntemlere her geçen gün yeni bir algoritma veya yeni bir yöntem eklenmektedir. Veri madenciliğinin temel yöntemleri çoğu istatistiksel yöntemlere dayanmaktadır. Veri madenciliğinde kullanılan bazı klasik yöntemler aşağıdaki gibidir.

- Regrasyon Analizi
- Sınıflandırma
- K-En Yakın komşuluk
- Kümeleme
- Karar Ağaçları
- Birliktelik Kuralı
- Sınır Ağları
- Bayes Teoremi
- Destek Vektör Makinası

#### 2.1. Regrasyon Analizi

Aralarında veri girişine göre sebep-sonuç ilişkisi bulunan bir veya birden fazla değişken arasındaki ilişkiyi belirlemek ve onlarla ilgili tahminlerde bulunmak amacıyla yapılır. (ORHUNBİLGE N.,2017:20) Şimdi bu durumu basit bir Sebep-Sonuç tablosu ile açıklamaya çalışalım.

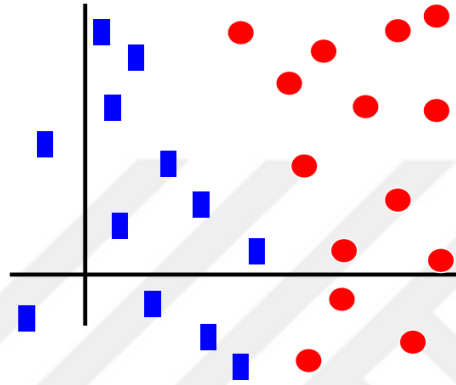
**Tablo 1.** Regrasyon Analizi Sebep-Sonuç İlişkisi

Sebep	Sonuç
Performans	Maaş
Alışveriş	Para
Gübre	Verim
Gelir	Harcama
Yaş	Boy

Tablodan anlaşılacağı üzere durumların sebepleri incelenerek muhtemel olması beklenen sonuçları tahmin edebiliriz.

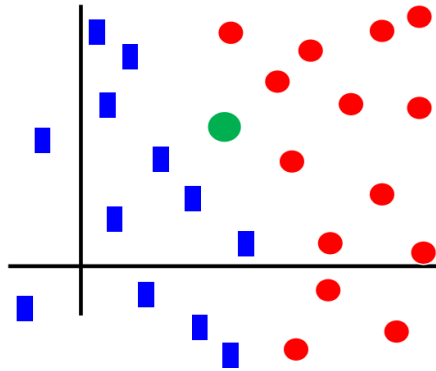
## 2.2. K-En Yakın Komşuluk (K Nearest Neighborhood)

Büyük veri tabanlarında kullanılan bir sınıflandırma tekniği olup gruplandırılmak istenen yeni bireyin daha önceki bireylerden k tanesine olan yakınlığına ve benzerliğine bakılarak sınıflandırma yapılmasıdır. (ŞEKER Ş.,2013:161-163)



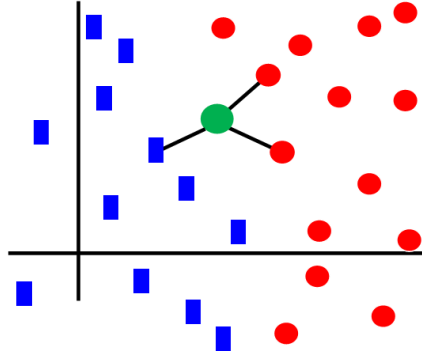
Şekil 2. A K-En Yakın Komşu Örneği

Şekil 3-A daki Görüntüde sınıflandırılmış halde bulunan elemanlar görülmektedir.



Şekil 3. B K-En Yakın Komşu Örneği

Şekil 4-B deki görüntüde olduğu gibi sınıflandırma grubuna yeni bir eleman katılması halinde KNN yöntemine göre bu gelen yeni üyenin en yakın olduğu 3 üye (veya 3'den büyük başka bir tek sayı) tespit edilir.

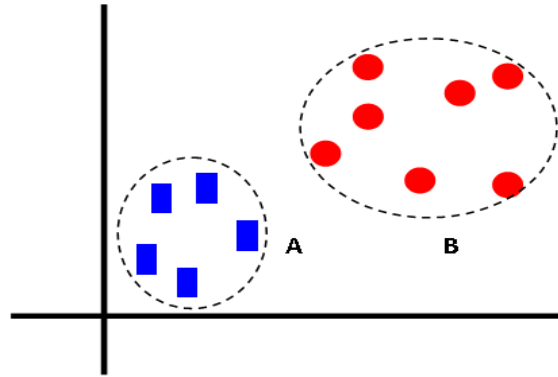


Şekil 4. C K-En Yakın Komşu Örneği

Şekil 5-C deki görüntü 'den de anlaşılacağı üzere yeni gelen elemana en yakın iki elamanın kırmızı yuvarlak grubu elemanları olduğu görülür ve yeni eleman bu sınıfta değerlendirilir.

### 2.3. Kümeleme (Clustering)

Kümeleme birbirine benzeyen veri parçalarının ayırma işlemidir ve kümeleme yöntemlerinin çoğu veriler arasındaki uzaklıkları kullanarak hesaplama yapar. Kümeleme yöntemi aklımıza en yakın komşu algoritmasını getirmektedir.(ÖZKAN Y.,2016:101)



Şekil 5. Kümeleme Algoritması

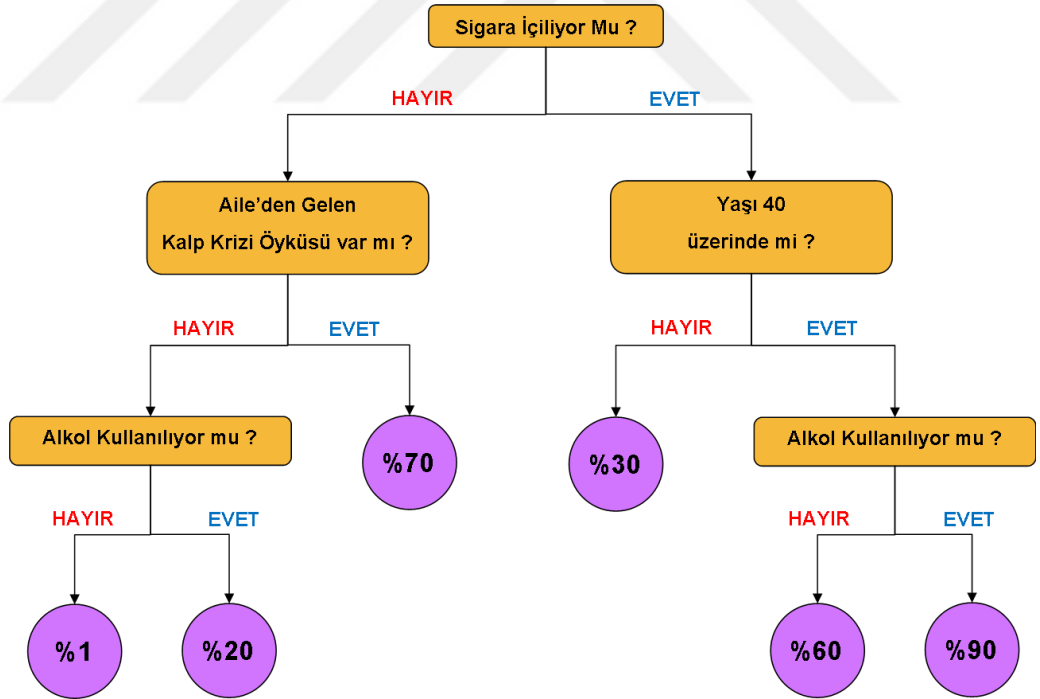
Kümeleme (Clustering) Algoritması veri seti içerisinde yer alan elemanları bir birlerine olan yakınlıklarına göre kümelere ayırarak sınıflandırabilmektedir.

## 2.4. Karar Ağaçları

Karar ağaçlarında amaç verileri parçalayarak bir ağaç yapısı oluşturmak ve bu yeni yapıda daha rahat karar vermeyi sağlamaktır. Karar ağacı öğrenmesinin temeli istatistikte makine öğrenmesine dayanmaktadır. Günümüzde en sık kullanılan güvenilir veri madenciliği tekniklerinden olan karar ağaçları; karar düğümleri, yapraklar ve dallardan meydana gelmektedir.

Karar düğümü yapılacak testi belirler, bu test sonrası karar ağacında veri kaybı yaşanmadan dallara ayrılır. Her yeni dalda sınıflandırma işlemi yapılmaktadır şayet sınıflandırma işlemi gerçekleşmiyorsa bu noktada karar düğümü meydana gelmektedir. Sınıflandırma işlemi yapılabiliriyorsa bu durum o dalın sonunda yaprak olduğunu işaret etmektedir. (ŞENTÜRK A.,2016:34) Bu durum ağacın tepesinden yapraklara ulaşıncaya kadar devam etmektedir.

Başlıca kullanılan karar düğümü algoritmaları C4.5, CHAID, CART, ID3'tür. Karar ağaçlarının görsel olarak ifadesi programcılıkta kullanılan algoritmalara benzemektedir. Aşağıdaki şekilde örnek bir karar ağacı yapısı görülmektedir.



Şekil 6. Karar Ağacı Örnek Uygulaması

Şekil-7 de yer alan yer alan sigara kullanımına bağlı olarak yaşanaabilecek kalp krizi risk durumunu sınıflandırma işlemi sonrası karar düğümü oluşturmak üzere ortaya koyulmuştur.

## 2.5. Birliktelik Kuralı

Genelde Ticari veri tabanlarında sıkça kullanılan bir yöntemdir. Olayların birlikte gerçekleşme ihtimaline dayalı olarak çalışan bir algoritmadır. Bu kuralın en yaygın kullanıldığı durum perakende satışlarda müşteri profilinin belirlenerek ürünlerin sunumunun veya mağaza içi dizilimlerinin çıkan sonuçlara göre düzenlenmesidir. (ÖZKAN Y.,2016:157)

Mağazada mangal satın alan müşterinin mangal kömürü satın alması tahmin edilebilen bir öngörüdür burada önemli olan öngörülemeyen durumların önceden tespitine Pazar sepet analizi adı verilmektedir. Birliktelik Kuralında bazı temel kavramları açıklayacak olursak;

**-Öğeler Kümesi:** Alışveriş sepetinde bulunan nesnelere dir.

**-Destek Sayısı:** Alışveriş sepetindeki ürünlerin veriler içerisinde görülme sayısıdır.

**-Destek:** Alışverişteki nesnelere içinde bulunduğu birliktelik sayısının toplam birliktelik sayısına olan oranını göstermektedir.

**-Güven:** Nesnelere arasındaki birlikteliğin doğruluğunu ifade etmekte kullanılır.

**-Yaygın Öğeler:** Destek değeri minimum destek değerinden büyük yada eşit olan nesnelere kümesidir.

**Tablo 2.** Değişik zamanlar satın alınmış alışveriş sepetleri

Alışveriş-1	Süt, Ekmek
Alışveriş-2	Ekmek, Yumurta
Alışveriş-3	Süt, Peynir
Alışveriş-4	Yumurta, Ekmek, Peynir, Süt
Alışveriş-5	Peynir, Yumurta, Süt

Pazar sepet analizi ile incelenecek olan Tablo-2 deki veriler değişik zamanlarda yapılmış alışverişlerin içeriğini göstermektedir.

**Tablo 3.** Alışveriş Sepetlerinin içerikleri kullanılarak Destek Sayısının Belirlenmesi

Küme Nesne Sayısı	Nesne	Destek Sayısı
1	Süt	4
	Ekmek	3
	Yumurta	3
	Peynir	3
2	Süt, Ekmek	2
	Ekmek, Yumurta	2
	Süt, Peynir	3
	Yumurta, Peynir	2
	Yumurta, Süt	2
3	Yumurta, Ekmek, Peynir	1
	Yumurta, Ekmek, Süt	1
	Ekmek, Peynir, Süt	1
	Peynir, Yumurta, Süt	2
4	Yumurta, Ekmek, Peynir, Süt	1

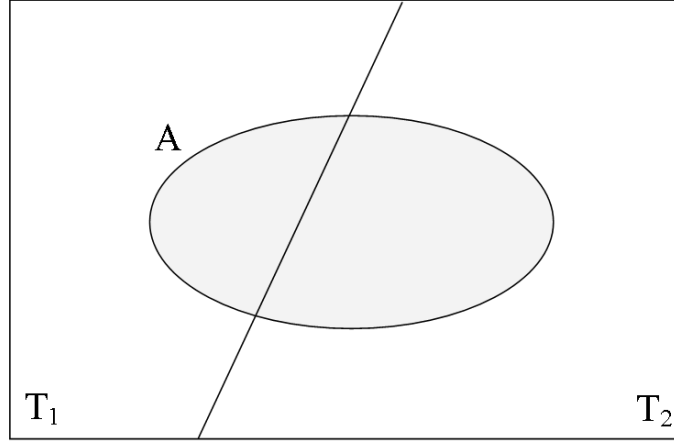
Tablo-3'den anlaşılacağı üzere alışveriş sepetleri incelenerek ürünlerin alınma ve birlikte alınma durumlarından destek sayısı tespiti yapılır.

**Destek** ({Yumurta} $\Rightarrow$ {Peynir})= Destek Sayısı(Yumurta, Peynir)/Toplam Alışveriş Sayısı=2/5

**Güven** ({Peynir, Süt} $\Rightarrow$ {Yumurta})=Destek Sayısı(Peynir, Yumurta, Süt)/Destek Sayısı(Peynir, Süt)=2/3

### 2.6. Bayes Teoremi ve Ağları

Bayes Teoremi olasılıklar hesabında önemli bir yere sahiptir. Bayes Teoremine dayanarak sınıflandırma yapmak mümkündür. Bu teoremin esasını ortaya koymak üzere  $T_1$  ve  $T_2$  gibi iki olayı göz önüne alalım. Bu iki olayın bağdaşmaz olaylar olduğunu yani  $T_1 \cap T_2 = \Phi$  olduğunu varsayalım.



Şekil 7. Bayes Teoremi T1-T2 olayları

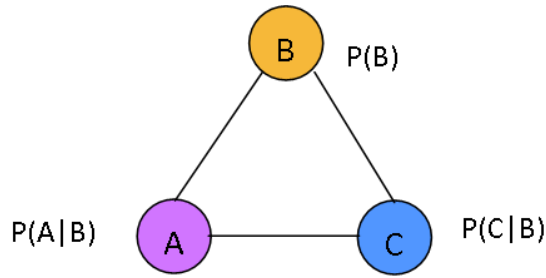
Bir olayın  $T_1$  ve  $T_2$  olayları cinsinden ifade etmeniz mümkündür. Bunun için koşullu olasılık bağlantısından yararlanır. Yani  $P(T_1 \setminus A)$  olasılığı

$$P(T_1 \setminus A) = \frac{P(A \setminus T_1)}{P(A)}$$

Biçimindedir. Şekil 7 üzerinde görüldüğü gibi A olayı  $T_1$  ve  $T_2$  olaylarından birinde gerçekleşmektedir. A olayı için şu bağıntı yazılabilir.

$$A = (T_1 \cap A) \cup (T_2 \cap A)$$

Bayes ağları, değişkenlere ait koşullu olasılık dağılımlarını ortaya koymak ve değişkenlere ait alt kümeler arasındaki koşullu bağımsızlıkları tanımlamak üzere kullanılır. (ÖZKAN Y.,2016:185)



Şekil 8. Bayes ağlar sınıflandırması

## 2.7. Veri Madenciliği Algoritmalarının Doğruluk Analizi İçin Kullanılan Yöntemler

Veri madenciliği algoritmaları içerisinde algoritma sonuçlarını daha kolay anlayabilmemiz ve yorumlayabilmemiz için birçok analiz ve istatistik yöntemi bulunmaktadır.

**Karışıklık Matrisi (Confusion Matrix):** Gerçek değerlerin belirli olduğu bir dizi test verisi üzerinde sınıflandırma algoritmasının performansını değerlendirmede kullanılan bir tablodur. Tabloya bakarak sonuçları yorumlamak daha kolay olabilir.

**Tablo 4.** Karışıklık Matrisi

TAHMİN EDİLEN			Toplam
GERÇEK	TP Gerçek Pozitif (True Pozitif)	FN Yanlış Negatif (False Negatif)	N <sup>+</sup> Gerçek Pozitif Sayısı
	FP Yanlış Pozitif (False Pozitif)	TN Gerçek Negatif (True Negatif)	N Gerçek Negatif Sayısı
Toplam	N <sup>+</sup> Tahmini Pozitif Sayısı	N <sup>-</sup> Tahmini Negatif Sayısı	N Toplam Örnek Sayısı

Sınıflandırma modelini içinde İkili bir tahmin durumu oluşması halinde sınıflandırma sonucu yukarıdaki şekildeki gibi ifade edilmesi durumunda daha anlaşılır bir hal alacaktır. Karışıklık matrisi verilerine dayanarak yapılabilecek bazı hesaplamalar ve çıkarımlar aşağıdaki gibi hesaplanmıştır.

**Sınıf (Class):** Tahmin edilen sınıflar a=Hastalık var ve b=Hastalık Yok olarak gruplandırılmıştır.

**Gerçek Pozitif (True Positive-TP)** Gerçek pozitif tahminlerin sayısı, veri seti içerisinde hasta olarak tahmin ettiklerimizdir.

**Yanlış Pozitif(False Positive-FP)** : Yanlış pozitif tahminlerin sayısı, veri seti içerisinde hasta olarak tahmin edilen fakat gerçekte hasta olmayanların sayısıdır. (Tip 1 Hatası olarak adlandırılmaktadır.)

**Gerçek Negatif (True Negative-TN):** Veri seti içerisinde hastalık öngörülme, hastalığı olmayanların sayısıdır.

**Yanlış Negatif (False Negatice-FN):** Veri seti içerisinde hastalığa sahip değil diye öngördüğümüz fakat normalde hasta olanların sayısıdır. (Tip 2 Hatası olarak



adlandırılmaktadır.)

**Doğruluk (Accuracy Rate):** Sınıflandırıcı Algoritmanın başarı oranı yüze olarak ifade edilmesidir. Aşağıdaki formüle göre hesaplanır.

$$AC = \frac{TP + TN}{\text{Toplam}}$$

**Keskinlik (Precision) :** Keskinlik değeri, Gerçek Pozitif değerinin Gerçek Pozitif ve Yanlış pozitif değerinin bölünmesi ile bulunmaktadır.

$$PPV = \frac{TP}{TP + FP}$$

**Hassasiyet (Recall) :** Elde edilen doğru sonuçların, olması gereken doğru sonuçlara oranı olarak hesaplanır.

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN}$$

**F-Skoru (F-Measure) :** İkili sınıflandırmanın istatistiksel analizinde, bir testin doğruluğunun denemede  $F_1$  skoru bir ölçütüdür. Skoru hesaplamak için keskinlik ve hassasiyet harmonik ortalaması alınmalıdır. (POWERS D.,2011)

$$F_1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$$

## 2.8. Veri Madenciliği Analiz Yazılımları

Veri madenciliği analiz yazılımı WEKA ücretsiz olması nedeniyle çok tercih edilmektedir. WEKA programı Yeni Zelanda'nın Waikato Üniversitesinde makine öğrenme grubu araştırmacıları tarafından geliştirilmiş bir veri madenciliği yazılımıdır. Adını Yeni Zelanda adalarında bulunan doğuştan uçamayan bir kuş olan Weka'dan alan yazılım, GNU lisansı ile yayınlanan açık kaynak kodlu bir yazılımdır.

İçerisinde yer alan Algoritmalar direk olarak bir veri setine uygulanabileceği gibi kendi java kodunuz içerisinden de çağırabilirsiniz. Weka içerisinde veri ön işleme, sınıflandırma, gerileme, kümeleme, ilişkilendirme kurulları ve görselleştirme araçları içerir. (WEKA)

Matlab, MathWorks tarafından geliştirilmiş çok yönlü sayısal hesaplama, bilimsel ve mühendislik grafiksel veri gösterimi, programlamanın yapıldığı bir yazılımdır. Matlab yardımcı ile veri madenciliğine ilişkin birçok hesaplamalar yaptırılabilir.

Python, Guido Van Rossum tarafından geliştirilmiş olan açık kaynak kodlu bir programlama dilidir. Python dilini diğer programlama dillerinden ayıran özelliklerinden birkaçı ise bilimsel yöntemlere yönelik program yazılabilmesi ve çok hızlı olmasıdır. Python programlama dili makine öğrenmesi teknikleriyle çok ilişkili bir şekilde çalışmakta ve bu konuda geliştirilmiş çok sayıda araştırma ve proje yapılmıştır.

Bu programlama dili ile makine öğrenme sistemini bir arada kullanarak verilerin edinimini, veri içerisindeki gürültünün temizlenmesi, vektörle hale getirilerek normal verinin elde edilmesini, Model oluşturmayı, Oluşturulmuş olan modelin test edilmesini, Modelin performansının denenmesini yapılabilmektedir.



## ÜÇÜNCÜ BÖLÜM

### 3. VERİ MADENCİLİĞİ METODLARININ TIP ALANINDA KULLANILMASI

Bu bölümde veri madenciliği yöntemlerinin tıp alanında nasıl kullanılacağını ve daha önce kullanılmış örneklerden bahsedilecektir.

#### 3.1. Sağlık Hizmetlerinde Veri Kavramı

Günümüz teknolojileri ile birlikte Sağlık hizmetlerinin sunumu yoğun bir bilişim trafiği içerisinde geçmektedir. Kaliteli hasta bakımı ve hizmetlerinin de ön plana çıktığı günümüzde hastanın tıbbi geçmişi çok önem arz etmektedir. Hekimlerin ve yöneticilerin karar alabilmeleri için sürekli olarak sağlık bilgi sistemi kayıtlarına ihtiyaç duyarlar. Toplanan tüm veriler bu sistem aracılığı ile anlamlı bilgiler haline dönüştürülerek sunulmaktadır. Sağlık Bilgi Sistemi'nin veri girişini oluşturan bazı unsurları şöyledir. Hastalara ait önemli Klinik bilgiler, Gözlem, Bulgular, Laboratuvar Teşhis Sistemleri, tıbbi Görüntüleme Sistemleri, Biyolojik Sinyal izleme Sistemleri, Optik Tıbbi Cihazlardır.

#### 3.2. Tıpta Veri Madenciliği

Sağlık sektörü bilginin doğruluğunun ve hızlı erişiminin en önemli olduğu sektörlerden birisidir. Profesyonellerin güncel ve doğru bilgiye en hızlı şekilde erişmeleri gerekmektedir. Bu anlamda Tıpta veri madenciliği karar desteği sağlamak, sağlık kurumlarının yönetimi ve sağlık politikalarının oluşturulmasında kullanılmaktadır. (KOYUNCUGİL S.,2009)

#### 3.3. Tıpta Veri Madenciliği Çözümleri

Ülkemizde Sağlık bakanlığının politika geliştirmesi için verilerin toplanması ve veri madenciliği yöntemleri kullanılarak analiz edilmesi hayati önem taşımaktadır. Sağlık alanında yaşanacak sorunlara veri madenciliği yöntemleri ile getirilebilecek bazı çözümler şöyledir. (EDWARD H.,2001)

Hastanelerde farklı veri tabanlarında tutulmakta olan veriler tek bir veri tabanında toplanarak veri madenciliği yöntemleri ile temiz ve analiz edilebilir veriye

erişim sağlanabilir. Elektronik hasta kayıtlarının kronolojik<sup>1</sup> sıralaması yapılarak hastanın teşhis, tedavi, ilaç, laboratuvar, radyoloji v.b. verileri tek dosyada tutularak hekime karar-destek sağlayıcı veriye erişimi sağlanabilir. Yaşam süresine bağlı olarak ortaya çıkabilecek süreğen hastalıklar için veri madenciliği yöntemleri hastalığa yönelik sosyal, demografik<sup>2</sup>, coğrafi, ekonomi v.b. değişkenler baz alınarak hastalık için erken uyarı sistemi geliştirilebilir. Laboratuvar testleri veri madenciliği yöntemleri ile incelenerek sağlık hizmetlerinin sunumunda ortaya çıkabilecek hataların tespiti yapılır. Risklerin minimize edilmesi sağlanabilir.

Hekimlere, hastanın geçmiş verilerini analiz ederek veri karmaşası yerine grafiksel bir ara yüz ile karar destek sistemi sunulabilir. Ülkemizde sağlık hizmetlerinin büyük bir bölümü kamu finansmanına bağlı olduğundan fatura yolsuzluğuna ve suiistimale açık durumdadır. Bu durum veri madenciliği yöntemleri kullanılarak takip edilebilir. Maliyete etki eden faktörlerin tespiti ve Maliyetleri düşürmek için uygulanacak politikalar veri madenciliği yöntemleri ile tespit edilebilir. Veri madenciliği yöntemleri ile finansal risk analiz uyarı sistemleri ve yönetsel karar destek sistemleri geliştirilerek yol haritaları çizilebilir. Elektronik hasta kayıtlarına tek merkezden erişim sağlanarak ülke içerisinde veri bütünlüğü sağlanabilir. Böylelikle hasta verilerine ülkenin her tarafından erişim sağlanır. Gereksiz test maliyetlerinin önüne geçilmiş olunur.

### **3.4. Tıpta Veri Madenciliği Yöntemleri Kullanılarak Yapılmış Çalışmalar**

Bu bölümde tıpta veri madenciliği yöntemlerini kullanarak yapılmış olan makale ve tez çalışmalarının çalışma konuları, çalışmalar sonucunda elde edilen çıkarımlara ve önerilere yer verilecektir. Hacettepe Üniversitesi Hastanesine başvurmuş hasta verileri üzerinden veri madenciliği teknikleri kullanılarak sağlık kayıtları içerisinde gizli kalmış kritik bilgileri ön plana çıkarmak amacıyla yapılmış bu çalışma ile aynı hastalığı taşıyan kişilerin ortak profillerinin çıkarılması, tedavi sonrası durumların tahmini, tedavi süreçlerinin maliyeti ve hastalıklarla ilgili bir tahmin sistemi geliştirilmesi öngörülmektedir. (YILDIRIM P.,2008)

Hastanelerin Acil Servisleri başvuran vakaların çeşitliği ve kesintisiz hizmet vermesi münasebetiyle hizmet sunumun yoğun olduğu ve kusursuz planlanması gereken

<sup>1</sup> Kronoloji: Olayların tarihsel sıralanması ile ilgili bilim dalı.

<sup>2</sup> Demografi: Nüfusun yapısını, durumunu, dinamik özelliklerini inceleyen bilim dalı.

birimleridir. Bu anlam Ege bölgesinde faaliyet göstermekte olan bir Araştırma ve Uygulama hastanesin 'in Acil Servis birimine başvuran hasta verileri ele alınarak veri madenciliği teknikleri ile başvuran hastaların yaşı, cinsiyeti, tanısı, fiziki ve coğrafi şartları, çalışan sayısı, hastaların acil serviste kalma süreleri gibi verileri kullanarak farklı sunum senaryoları geliştirilmiştir. (ÖZDAĞOĞLU A, 2009)

Görülme sıklığı her geçen gün artan ve mücadelesi oldukça güçlü bir süreç olan hastane enfeksiyonları önemli derecede ölüm sebebidir. Tedavisinin zorlu olması ve finansal maliyetleri göz önünde bulundurulduğunda ilgilenilmesi gereken bir konudur. Bu anlamda Eskişehir Osmangazi Üniversitesi hastanesi Klinik Mikrobiyoloji ve Enfeksiyon hastalıkları bilim dalı tarafından toparlanmış veri setine sınıflandırma, karar ağacı, yapay sinir ağları ve veri doğrulama yöntemleri uygulanarak tedavi süresince kullanılan bazı girişimsel metotların sürelerinin enfeksiyon ile olan oranı veri madenciliği yöntemleri ile başarılı bir şekilde ortaya koyulmuştur. (AKSOY İ.,2010)

Veri, gelişen ve ucuzlayan teknolojiye bağlı olarak her geçen gün artmaktadır. Veri madenciliği yöntemleri ise her geçen gün devasa bir şekilde artan veri ambarlarının en etkili şekilde kullanılması için kaçınılmaz yegane çözümdür. Her geçen gün bayanlar arasında görülme sıklığı artış gösteren Meme Kanseri Wisconsin veri setine temeli ID3, C4.5 algoritmalarına dayanan J48 ve Bayes Sınıflandırma, Kstar algoritmalarına dayanarak modeller oluşturulmuş ve oluşturulan modellerin başarımlar oranları bir birleriyle karşılaştırılarak Meme Kanseri Veri setine veri madenciliği metotlarının uygulanmasının başarımlarını tespit edilmiştir. (POYRAZ O.,2012)

Kullanım alanı gün geçtikçe artan doğru sınıflandırma başarımları ile ön plana çıkan lojistik regresyon analizi ile laboratuvar bölümünden hemogram ve biyokimya testlerinin sonuçlarını analiz ederek kalp hastalığının olma olasılığı üzerinde durulmuştur. 2011 yılı içerisinde Mayıs ve Kasım ayları arasında Kardiyoloji ve diğer kliniklere yatan hastaların testleri lojistik regresyon analizi ile değerlendirilmiştir. Kardiyoloji kliniğindeki hastaların değerlendirilmesi kardiyoloji uzmanları tarafından yapılmıştır. (AKYOL K.,2012)

Pamukkale Üniversitesi Hastanesi'ne 2008, 2009, 2010 ve 2011 yılı içerisinde ayaktan başvuran hastalara ilişkin veriler hastanenin SQL veri tabanından alınarak veri madenciliği yöntemleri ile incelenerek hasta profillerini tespit etmek üzere çalışma yapılmıştır. Bu çalışmanın sonucunda; aylara göre hasta dağılımı, cinsiyet, doğum yeri,

hasta türü, başvuru polikliniklere göre incelemeler yapılmış ve hasta profil yapısı ortaya koyulmuştur. (ERTUĞRUL İ.,2013)

Elektro Kardiyo Grafi (EKG) Kalp kasını ve iletim sisteminin çalışmasını incelenmesine yardımcı olan bir sistemdir. EKG sinyallerini işlemek üzere Kaba Kümeleme yardımıyla eksik verilerden yararlanılarak olası veya kesin veriyi elde etme yöntemi ve veri madenciliği yöntemleri kullanılarak EKG sinyallerin tutarsız verilerden temizlenerek sinyalleri sınıflandırma yaparak az sayıdaki parametre ile sağlık profesyonellerinin karar vermesine yardımcı olmak amacıyla bir model ortaya koyulmuştur. (ÇEKİÇ R., 2014)

Akciğer kanserinin kesin tanısını koymak için hastalığın klinik seyri ve hastadan alınan biyopsi<sup>3</sup>siye yapılan patolojik inceleme sonucu hastalığın evresi belirlenebilmektedir. Veri madenciliği yöntemleri ile hastalığın klinik bilgileri ile patoloji sonuçları arasındaki ilişkiyi bulmak amacıyla kullanılmıştır. (HAOFAN Y.,2015)

---

<sup>3</sup> Hastadan patoloji biriminde incelenmek ve hastalığın evresini tespit etmek üzere parça alma işlemi

## DÖRÜNCÜ BÖLÜM

### 4. YENİDOĞAN SEPSİSİ VERİ SETİNİN TANIMI VE ELDE EDİLEMSİ

Bu bölümde Yenidoğan sepsisi hastalığı ve veri madenciliği yöntemleri kullanarak yeni doğan bebeklerde görülen sepsis hastalığının teşhisinde uygulanan skorlama sistemleri, testler ve gözlemlerin düzeylerinin önemi incelenecektir.

#### 4.1. Sepsis Kavramı

Sepsis, vücuttaki bütün organları ve sistemleri tutan ağır bir enfeksiyondur. Yenidoğan sepsisi bebeklerde görülen hayati tehlikeye sahip bir hastalıktır. Sepsis yenidoğan bebekler için önemli bir morbidite<sup>4</sup> ve mortalite<sup>5</sup> sebebidir. Bundan ötürü zamanında tanı koyulması ve tedavi edilmesi büyük önem taşımaktadır. (CENGİZ A.B.,2007). Gelişmiş ülkelerde Yenidoğan sepsis'in nüfus içerisindeki oranı 1000 canlı doğumda 1-10 arasındaki iken gelişmekte olan ülkelerde ise bu oran 1000 canlı doğumda 49 ile 170 arasında olmaktadır. (CENGİZ A.B.,2009)

#### 4.2. Sepsis Tanısı

Yenidoğan Sepsis'i tanısı canlı doğumdan sonra 28 gün içerisinde ortaya çıkan bir belirdir. Kesin tanısı kan kültüründe bakteri üremesi takip edilerek konulmaktadır. Fakat hastalığa özgü belirtilerin ve bulguların güç anlaşılıyor olması, kan kültür sonuçlarının zaman alması, kan kültürü sonuçlarının yanlış sonuçlanabilmesi tanı koymayı güçleştirmektedir. Kesin tanıyı hızlı koymak için Laboratuvar çalışmaları ve bilimsel çalışımlar henüz daha devam etmektedir. Yenidoğan sepsisi tespit edileme zamanlarına göre Tıp otoriteleri tarafından üç farklı sınıfa ayrılmıştır. (UZUN F.,2015)

- Erken Başlangıçlı Yenidoğan Sepsisi
- Geç Başlangıçlı Yenidoğan Sepsisi
- Çok Geç Başlangıçlı Yenidoğan Sepsisi

<sup>4</sup> Morbidite: Belirlenmiş bir zaman diliminde belli bir hastalığa yakalanan ve tanı koyulan hasta sayısıdır.

<sup>5</sup> Mortalite: Genel popülasyon içerisinde belli bir hastalığa bağlı ölüm oranıdır.

### 4.3. Yenidoğan Sepsisi Veri Seti

Bu çalışmada kullanılan veri setindeki veriler Uzm. Dr. Fatma UZUN tarafından Fırat Üniversitesi Hastanesi Çocuk sağlığı ve Hastalıkları Anabilim dalı Yenidoğan Yoğun Bakım Ünitesi'nde Mayıs 2013 ile Ocak 2014 tarihleri arasında veri bilgi formları doldurulmak suretiyle toplanmıştır. Veri seti çalışma grubu içerisindeki hastaların sayıları ve dağılımı aşağıdaki gibidir.

**Tablo 5.** Veri Seti Çalışma Grubu

Tanı Durumu	Hasta Sayısı
Kanıtlanmış Yenidoğan Sepsisi	40
Klinik Yenidoğan Sepsisi	48
Kontrol Grubu	40

Kanıtlanmış Yenidoğan Sepsisinde 40 hastanın Kan Kültürü sonuçlarında üreme olmuştur. Klinik Yenidoğan Sepsisindeki 48 hastanın Kan Kültüründe üreme olup olmamasına bakılmaksızın klinik seyirlerine göre ve diğer bulgulara dayanarak sepsis tanısı koyulmuştur. Kontrol Grubundaki 40 hasta ise sağlıklı bir anneden doğmuş, kan Kültürü Testlerinde üreme veya Diğer Klinik seyirlerinde sepsis tanı koymak için bir emareye rastlanmamıştır. (UZUN F.,2015)

### 4.4. Sepsis Veri Seti Analizi

Veri Seti içerisindeki birçok alanda kısaltma isimleri ve latince terimler kullanıldığından çalışmanın daha anlamlı ve anlaşılabilir olması açısından veri setindeki alanların ve bu alandaki verilerin nasıl, niçin toplandığı bu bölüm içerisinde açıklanacaktır.

#### 4.4.1. Genel Veriler

**Cinsiyet:** 1-Erkek, 2-Kız

**Doğum Şekli:** 1- Sezaryen Doğum, 2-Normal Doğum

**Doğum Kilosu:** 1-1500 gram altı, 2-1500-2500 gram arasında, 3-2500-4000 gram arasında, 4-4000 gram üstü

**Doğum Haftası:** 1- 38 Hafta ve Üstü, 2- 32-37 Hafta arasında, 3- 28-31 Hafta arasında, 4- 28 Haftanın altında



**Gerçek Doğum Haftası:** Doğumun gerçekleştiği hafta değeri

**APGAR Skor 1DK ve APGAR Skor 5DK:** APGAR Skor Sistemi; Amerikalı Anestezi Doktoru Virginia APGAR tarafından katıldığı binlerce doğum sonrası 1952 geliştirdiği ve dünyaya tanıttığı skor sistemidir. Bu sistem; Kan Tonusu, Kalp Hızı, Uyarılara Cevap, Cilt Rengi, Solunum değerlerine doğumun 1.dakikasına ve 5.dakikasına bakarak skora yapılır. (MUMCU A.)

Skora göre;

- 7 ile 10 arasında olanlar normal
- 4-6 arasında olanlar birkaç solunum desteği yapılarak bu dönemi atlatır.
- 0 ile 3 arasında olan bebekler ise acil müdahale gerektirmektedir.

Veri Seti içerisinde Apgar Skorlama sisteminde birinci ve beşinci dakikalarında 7 ve altında puanlama alan bebekler 1 değeri ile gösterilirken. 8 ve üzerinde değer alan bebekler 2 değeri ile gösterilmiştir.

**Kan Grup Uyuşmazlığı:** Ana ve Bebek kan Grupları arasındaki uyumsuzluk durumuna bakılır. 1-Yok, 2-Var

**Celestone Uygulaması:** Celestone Ampül halinde etken maddesi betametazon olan iltahap giderici anneye enjeksiyon yoluyla uygulanan ilaç.

**Rhogam Uygulaması:** Gebeliğin sekizinci haftasından sonra düşük tehdidi olan Kan Grubu Rh- olan annelere yapılan bir ilaç uygulamasıdır.(TOPALOĞLU S.)

**Sepsis Tanı Zamanı:** Sepsis tanısı konulma zamanına göre 3 bölüme ayrılmıştır.

- 72 Saatten önce olanlar Erken Sepsis (Değer:1)
- 72 Saatten sonra olanlar Geç Sepsis (Değer:2)
- 30 Gün ve sonrasında olanlar Çok Geç Sepsis (Değer:3)

**Sepsis Hikâye:** Daha önceden sepsis hikâyesi var mı? Doğum sonrası Yenidoğan daha sepsis tanısı almış mı? (1:Yok, 2:Var)

**Antibiyotik Kullanımı:** Önceden antibiyotik kullanımı var mı? (1:Yok, 2:Var)

**RDS:** RDS (Respiratuvar distress sendromu) , sürfaktan<sup>6</sup> eksikliği ve akciğerlerin olgunlaşmamasına bağlı olarak doğumdan sonraki ilk 4 saat içinde ortaya çıkan, 24 saatten fazla süren, Morarma, Hızlı ve inlemeli solunum, ile kendini gösteren bir klinik tablodur. (ALDEMİR B.)

<sup>6</sup> Yetişkinlerde ve normal bebeklerin akciğerinde bulunan sabunsu bir maddedir. Sürfaktan olmadığı zaman hava kesecikleri nefes verildiği zaman yapışma eğilimi gösterirler.

**Surfactan:** İlaç uygulaması (1:Yapıldı, 2:Yapılmadı)

#### 4.4.2. Bebekte Tespit Edilen Mevcut Hastalıklar ve Riskler İlişkin Veriler

Gebelik Sürecinde veya doğum sonrası bebeğe eşlik eden hastalıklara ilişkin veriler bu bölümde incelenecektir. Veri Seti içerisinde aşağıdaki hastalıkların olması durumu 1 ile ifade edilirse olmaması durumu 2 ile ifade edilmiştir.

**Prematüre:** Hamileliğin 28 ile 37. Haftasında olan doğumlar.

**Yenidoğan Sarılığı:** Sağlıklı ve Doğum Haftasına göre normal olan bebeklerin %60'da ve süresinden önce gerçekleşmiş doğumlarda ise %80'inde kanda yer alan bilirubin adlı maddenin oranının yükselerek bebeğin gözlerinin ve cildinin sarıya çalmasıyla ortaya çıkar. (MUMCU Y.,2015)

**Plasenta Previa<sup>7</sup> Anne Bebeği:** Plesenta<sup>7</sup>'nin doğum öncesi bebeğin önüne geçtiği ve doğumun sezaryen ile gerçekleşmesinin gerektiren bir durumdur. (MUMCU A.)

**EMR'li Anne Bebeği:** EMR(Erken Membran Ruptürü) Suların Erken Gelmesi durumudur.

**Oligohidramniyoz:** Bebeğin çevresindeki Gebelik suyu miktarının anormal bir şekilde azalmasıdır. (ULUDAĞ S.)

**Yenidoğan'ın Geçici Takipnesi (YDGT):** Zamanında doğum sonrası bebeğin geçici bir süreyle solunum sayısının normalden daha fazla olması durumudur. Yaş akciğer ismi verilir. (ÖRS R.,2013)

**IUGR:** Rahim içi gelişme geriliği, anne karnındaki bebeğin doğum ağırlığının 2,5 kg altında olması olarak tanımlanır. (URMAN B.)

**Koryoamniyot:** Fetus<sup>8</sup>'ün içerisinde bulunduğu sıvıyı çevreleyen zarın enfeksiyon olması durumudur. (UZUN F.,2015)

**Opr intestinalatrezi:** Yenidoğan bebeklerde ince bağırsak tıkanıklığı olarak bilinen 2700 canlı doğumda görülen bir hastalıktır. Bebeğin bu hastalıktan operasyon geçirip geçirmediği durumudur. (YAĞMURLU A.)

**Direkt Bilirubinemi:** Karaciğer, Safra Kesesi ve Yolları hastalıklarının tespitinde kullanılan bir testtir.

<sup>7</sup> Gebeliğin ile haftalarından itibaren anne ile bebek arasında besin alışverişini sağlayan organdır.

<sup>8</sup> Üçüncü gebelik ayının başından doğuma kadarki süre içerisinde ana rahmindeki canlıya verilen isimdir.

**Opr Pilorstenozu:** Midenin bağırsaklara ile geçiş bölümünün daralması ile ortaya çıkan ve besinlerin mideye geçişini engelleyen 500 ile 1000 arasında canlı doğumlara görülen bir hastalıktır. Bebeğin bu hastalıktan operasyon geçirip geçirmediği durumudur. (OĞUZKURT P.)

**Metabolik Hastalık:** Protein, yağ ve karbonhidrat asitlerinin yapısının bozulması ile ortaya çıkan yeni doğan bebeklerde doğumda normal olup doğum sonrası ortaya çıkabilen, uygun tedavi başlanmaz ise ölümlü sonuçlanabilen bir hastalık durumudur. (ERBAŞ O.)

**PDA:** Anne karnındaki bebeğin kalbinden biri vücuda diğeri de akciğere giden iki büyük damar vardır. Bu atardamarlar arasında açıklık vardır bu açıklık doğum sonrası kapanır. Kalp 'den çıkan bu damarların açıklığının doğum sonrası kapanmaması durumuna PDA (Patent Duktus Arteriosus) denir. (US M.)

**Böbrek Yetmezliği:** Böbreğin normal fonksiyonların giderek azalması veya görevini kaybetmesi durumudur.

**Asfiksi-Hipoksik Doğum:** Organ veya dokulara yeterince oksijen ulaşmaması, yeterince kan ulaşmaması durumunda gelişen durumdur. (SAY A.,2011)

**Polikistik Böbrek:** Böbreklerde kist<sup>9</sup> gelişmesi durumudur.

**Vezikoüreteral Reflü (VUR):** İdrarın mesaneden ureter<sup>10</sup>lere geri kaçması durumunda ortaya çıkan hastalıktır.(SEVER L,2005)

**Opr özofagusatrezisi:** Özofagus yemek borusu anlamındadır. Atrezisi ise doğumdan olması gereken yeterince gelişmemesi anlamındadır. Yeni yemek borusunun gelişmemesi, kısmen olmaması veya yokluğu durumudur. Bebeğin bu hastalıktan operasyon geçirip geçirmediği durumudur. (OĞUZKURT P.)

**Ambigusgenitalya:** Bebeklerde Cinsiyet belirsizliği hastalığıdır.

**Pnömotoraks:** Göğüs boşluğu içerisine herhangi bir sebepten ötürü hava toplanması durumudur. (HALEZEROĞLU S.)

**VSD, ASD:** Kalpteki sağ ve sol karıncık arasında delik oluşması durumudur. Bu duruma bağlı olarak akciğer atardamar kan basıncının artmasıdır. (EREK. E.)

**Preekleptik Anne Bebeği:** Gebelik sırasında annenin tansiyonun yüksek düzeyde seyretmesi, el, ayak ve yüzde şişlik, idrarda fazla miktarda protein atımı ile ilgili bir hastalıktır. (YAKUT Y.)

<sup>9</sup> Sıvı yâda yarı sıvı bir maddeyle dolu torba

<sup>10</sup> Böbrek ile idrar torbası arasında bulunan ve meseneye kadar uzanan kanallardır.

**İdrar yolu enfeksiyonu (IYE):** İdrar yolundan mesaneye kadar ulaşan bakterilerin neden olduğu bir enfeksiyon hastalığı durumudur

**Hipoksik İskemik Ensefalopati (HIE):** Böbrekler, karaciğer, akciğerler ve solunum sistemi, kalp ve dolaşım sistemi gibi pek çok organ sistemi yetersizliği sonucunda beyin fonksiyonlarında ortaya çıkan bir bozulma durumudur. (KARAMAN ÖZİŞİK H., 2007)

**Bronkopnömoni (BP):** Çocuklarda görülen solunum yolu iltihabı durumudur. (CANTILAV N. Ve Ark. 2017)

**Nekrotizan Enterokolit (NEC):** Yenidoğan bebeklerde yaşamın ilk haftalarında görülen bağırsak duvarından iltihap, doku ölümü durumudur. (TÜRK E.,2016)

**ASA-PFO:** Kalibin sağ ve sol kulakçığı arasındaki duvarda oluşan ve doğumdan sonra kapanması gereken deliğin kapanmamasına bağlı gelişen durumdur. (KILIÇ T.,2014)

**Opr Hidrosefali:** Hidro Su anlamında, Sefali baş anlamındadır. İki tanımın birleşiminden oluşur. Beyinde aşırı su birikmesiyle suyun beyinin bazı odacıklarına baskının yükselmesi olarak tanımlanabilir. Bebeğin bu hastalıktan operasyon geçirip geçirmediği durumudur. (ŞEN. O)

**Opr Meningomyelosele:** Bebeğin bel ya da sırt bölgesinde omuriliğin veya omurilik sıvısının dışarıya doğru kese şeklinde fitikleşmesi durumudur. Bebeğin bu hastalıktan operasyon geçirip geçirmediği durumudur. (MUMCU.A)

**Spinabifida:** Bebeğin omurgası ilk oluştuğunda kapalı değildir. Döllenme sonrası 29. Günde birleşir veya kapalı bir oluşum haline gelir. Spina Bifida da bu durum gerçekleşmez veya omurlar arası kapanmaz. (MUMCU.A)

**Chiarymalformasyon:** Beyin ve Omuriliğin birleştiği kafa bölgesinde oluşan doğumsal bir anomalidir<sup>11</sup>.

**Kernikterus:** Yenidoğan sarılığındaki yüksek bilirubin seviyesine bağlı olarak bebeklerde ortaya çıkabilecek kalıcı nörolojik hasarlar durumu olarak açıklanabilir. (GÜRTUNA A.,2017)

**Bronko Plmoner Dispilazi (BPDP):** Yenidoğan bakımında Oksijen ve pozitif basınçlı ventilasyon ile tedavi edilen prematüre bebeklerde en sık gelişen uzun dönemli bir akciğer hastalığıdır. (ÖZKAN H. ve Ark.,2008)

<sup>11</sup> Tıp Dilinde normalin dışında olan anlamındadır.

**Kalp Kapak Hastalığı:** Kalp kapakları kanın kalpte tek yönde ilerlemesine yardımcı olan bir anahtar görevindedir. Kalbin her atışında kapaklar açılarak kanın bir bölümden diğerine geçmesini sağlamaktadır. (CİVELEK A.,2014)

**Yenidoğan Konvizyonu:** Yenidoğan bebeklerde sıklıkla 0 ile 28 gün arasında nöbet geçirme şeklinde görülen nörolojik<sup>12</sup> bir hastalıktır.

**Diyabetik Anne Bebeği:** Şeker hastalığı olan gebe annenin bebeği

**Sendromik Bebek:** Yenidoğan bebekte özel bir bozukluğu gösteren tanıyı kolaylaştıran belirtilerin veya bulguların tümü ve buna bağlı hastalıklardır.

**Kas Hastalığı:** Liflerin hastalığına bağlı olarak güçsüzlük, ağrı, incelme veya krampların kaslarda görüldüğü durumdur.

**Diafragmahernisi:** Fetüs<sup>13</sup>'ün oluşumu ile birlikte diyaframın<sup>14</sup> normal dışı gelişimi ile ortaya çıkan genelde doğum öncesi ultrason ile teşhisi koyulabilen oldukça ciddi bir hastalıktır. (YAĞMURLU A.)

**Non immün hidrops:** Fetal dokularda veya vücut boşluklarında sıvı toplanmasıyla ortaya çıkan bir durumdur.(UÇAR B.A.,2011)

**Hematolojik Hastalık (Nötropeni/Trombositopeni):** Hematoloji latince bir terim olup Kan hastalıkları anlamına gelmektedir.

**Omfolit:** Göbek deliğinin iltihaplanması anlamına gelir. (BEYATLI E.,2015)

**Hipotiroidi:** Trioit<sup>15</sup> bezin normal dışına fazla çalışmasıyla vücutta aşırı miktarda tiroid hormonu üretmesi durumudur. (KABASAKAL L.)

**Mekoyumaspirasyon Sendromu (MAS):** Mekanyum, gebelik esnasında bebeğin bağırsaklarının içerisinde biriken bir maddedir. Mekonyum ile boyalı aminyo sıvısı bebeklerde etkili solunum hastalıklarına yol açmaktadır. (ŞEN V. ve Ark.,2014)

**ROP:** Prematüre, düşük doğum ağırlığı ve erken doğan bebeklerde görülen çocukluk döneminde körlüğe yol açan bir göz hastalığıdır. (KÖKSAL N.,2005)

**Anal atrezi:** Bebeğin doğuştan makatının olmaması veya kapalı olması durumudur. (SOYSAL F.)

<sup>12</sup> Sinir Sistemi ve onunla ilgili hastalıklarla uğraşan bir tıp bilim dalıdır.

<sup>13</sup> Anne kardındaki bebeğe 8.haftadan sonra verilen isimdir.

<sup>14</sup> Karın boşluğunu göğüs boşluğundan ayıran ince kas yapısına verilen isimdir.

<sup>15</sup> Boyunda bulunan kelebek şeklinde hormon üreten bir bezdir.

#### 4.4.3. Kan Kültürü ve Üreme Sonucu İle İlişkili Veriler

**Kan Kültür Üreme Durum:** Normalde kan hiçbir bakteri veya mantar içermez. Kanda enfeksiyon tespiti için yapılan testtir. Yaklaşık 7 gün içerisinde sonuçlanmaktadır.

**Üreme Sonucu:** Kan Kültüründe Üremenin durumu

#### 4.4.4. Tanı, Skorlama Sistemleri İle İlgili Bulgular ve Testlere İlişkin Veriler

**SIRS:** Sistemik İnflamatuvar Response Sendromu, aşağıdaki kriterlerin iki tanesini olmasına verilen kısa isimdir. (TARHAN. Ö.R.)

- Vücut Sıcaklığının 36 Derecenin altında olması veya 38 derecenin üstünde olması
- Kalp Hızının dakikada 90'ın üstünde olması
- Solunum Hızının dakikada 20'nin altında olması veya PaCO<sub>2</sub><sup>16</sup> değerinin 32'nin altında olması
- Kandaki WBC (Lökosit) değerinin 12.000 mm<sup>3</sup>'den yüksek olması veya 4.000 mm<sup>3</sup>'den düşük olması

**Platelet (Troboosit):** Kanın elemanlarından biridir. Kanamayı durdurma özelliği vardır. Azlığı kanamaya, fazlalığı ise damar içi pıhtı oluşumuna sebep vermektedir. (UÇAR A.B,2011)

**WBC:** Kanda bulunurlar, Beyaz Kan hücrelerinin görevi vücuttaki enfeksiyonlarla savaşmaktır. Diğer adıda Lökosit'tir.

**Kan Gazı Asidoz:** Hastaların metabolik ve solunumsal durumu hakkında güvenilir bilgi veren biz dizi önemli laboratuvar testlerinden oluşmaktadır. (SARIHAN A.,2016)

**TOLLNER Skoru:** Yeni doğan sepsis tespitinde Gözlemlere ve yukarıda çalışılmış tarama testleri sonuçlarına dayanarak hastalık tespitinde kullanılan skorlama sistemidir. (GÖZÜBÜYÜK A.,2016)

---

<sup>16</sup> PaCO<sub>2</sub>: Kanda bulunan oksijen ve karbondioksit gibi gazların basınçlarının ölçümü

#### 4.4.5. Bebeğe Ait Risk Faktörleri İle İlgili Veriler

**BEBEK İnvazif:** İnvazif<sup>17</sup> girişimlerin sık yapılması

**BEBEK Deri Bütünlük:** Deri bütünlüğünün yetersiz oluşu

**BEBEK Tekrarlayan Antibiyotik:** Tekrarlayan antibiyotik kullanımı

**BEBEK Vasküler Katater:** Damarda içerisinde uzun süreli kalan ve her an örnek alınması hazırda duran tıbbi bir yöntemdir.

**BEBEK Mekanik Ventilasyon:** Solunum yetmezliği çeken hastalarda uzun süreli olarak solunum desteğini cihaz aracılığı ile sağlanması durumu.

**BEBEK Glikokortikoid Tedavi:** İnsan sağlığında önemli olan kortizol salgısının olmaması buna bağlı olarak böbreklerde, böbrek üstü bezlerinde, oluşabilecek problem durumudur.

**BEBEK YB Salgın:** Yenidoğan yoğun bakım ortaya çıkabilecek salgınlar

**BEBEK YB Personel Yetersizlik:** Yenidoğan yoğun bakım personelinin yetersizliği

**BEBEK El Yıkama Azlığı:** Yenidoğan yoğun bakım personelinin El yıkamayı olayını az yapması

#### 4.4.6. Beslenme Bozukluğu İle İlgili Veriler

**BEBEK Enteral Beslenme:** Beslenme işlevine sahip olmasına rağmen besinleri ağız yoluyla alamayan hastalar için farklı beslenme metodudur.

**BEBEK Tam Enteral Beslenme:** Beslenme işlevine sahip olmadığında besinleri mideye kadar ulaştıran farklı beslenme metodudur.

**BEBEK Doğum Ağırlığına Ulaşma:** Bebek normalde olması gereken doğum ağırlığına ulaşmış mı?

#### 4.4.7. Anneye Ait Risk Faktörleri İle İlgili Veriler

##### 4.4.7.1. Annenin Önceden Var Olan Hastalıkları İle İlgili Veriler

**ANNE Kronik Hastalık:** Vücudun herhangi bir yerinde ortaya çıkan ve uzun süre tedavi edilmeyen veya ömür boyu kalıcı olan hastalıklardır.

**ANNE Diyabet:** Diyabet (Şeker Hastalığı), Vücuttaki insülin hormonunun eksikliği veya fazlalığı sonucu ortaya çıkan ömür boyu süren bir hastalıktır. (TEMĐ)

<sup>17</sup> Vücut bütünlüğüne en az şekilde zarar verecek şekilde yapılan girişimler

**ANNE HipoTiroidi:** Trioit<sup>18</sup> bezin normal dışına fazla çalışmasıyla vücutta aşırı miktarda tiroid hormonu üretmesi durumudur. (KABASAKAL L.)

**ANNE Tirotoksikoz:** Vücuttaki tiroid hormonlarının artışı sebebi ile ortaya çıkan durumdur. (BALKAN F.,2017)

**ANNE Böbrek:** Böbreğin normal fonksiyonların giderek azalması veya görevini kaybetmesi durumudur.

**ANNE Nöbet Geçirme:** Beynin beklenmedik bir anda, anormal bir şekilde davranış bulunması durumudur. (KIYLIOĞLU N.,2013)

**ANNE SLE:** Sistemik Lupus eritematozus (SLE), Bağışıklık sisteminde sebebi bilinmeyen birçok organı ve sistemi etkileyen bir bağ dokusu hastalığıdır. (DÜZGÜN N.)

**ANNE Kalp Hastalık:** Annenin Kalp Hastalıkları ile ilgili durumu

**ANNE Astım:** Annenin Astım<sup>19</sup> Hastalığı ile ilgili durumu

**ANNE Kistik Fibroz:** Kistik Fibrozis kalıtsal bir hastalık olup, doğum sonrası salgı bezlerinin fonksiyonlarının bozulması ve buna bağlı olarak bir çok organın etkilenmesiyle ortaya çıkan bir durumdur.

**ANNE Zayıflık:** Annenin zayıf olması durumu

**ANNE Üreme Sistemi:** Annenin üreme sistemi ile ilgili hastalık durumu

**ANNE Yaş:** Annenin gebelikteki yaşı

**ANNE RIA:** Rahim İçi Araçlar (RIA), Anne rahmine doğum kontrol amaçlı yerleştirilen küçük plastik cisimdir.

#### 4.4.7.2. İntrapartum (Doğum Sırasında) Komplikasyonlar İlişkin Veriler

**IP Premature:** Hamileliğin 28 ile 37. Haftasında olan doğumlar.

**IP EMR:** EMR(Erken Membran Ruptürü) Suların Erken Gelmesi durumudur. (1:Yok, 2:Var)

**IP Ateş Enfeksiyon:** Ateş ve Enfeksiyon Durumu

**IP FD Hipoksi:** Organ veya dokulara yeterince oksijen ulaşmaması, yeterince kan ulaşmaması durumunda gelişen durumdur. (SAY A.,2011)

**IP Müdahaleli Doğum:** Doğum esnasında beklenmedik bir durum karşısında müdahale edilmesi

<sup>18</sup> Boyunda bulunan kelebek şeklinde hormon üreten bir bezdir.

<sup>19</sup> Nefes Darlığı Hastalığı



**IP Serklaj:** Rahim ağzı normal bir hamilelikte kapalıdır. Fakat bazı durumlarda Rahim ağzı ilerleyen bir gebeliği ve doğum sancılarını taşıyamayacak kadar güçsüz olabilir. Bu durumda serklaj adı verilen rahim ağzına dikiş geçirilerek gebelik sonuna kadar rahim ağzının kapalı tutulması durumudur. (MUMCU A.,)

**IP Fetal Taşikardi:** Anne karnındaki bebeğin kalbinin normalden farklı olarak hızlı bir şekilde çarpması şeklinde görülen bir kalp rahatsızlığı durumudur. (YAYLA M.)

**IP Doğum Travma:** Annenin doğum esnasında kendisinin veya bebeğinin öleceği, yaralanacağı hissiyatına kapılması durumudur. Bu durum anneyi, bebeği ve ailesini ruhsal olarak etkiler. (İSBİR GÖKÇE G.,2014)

#### 4.4.7.3. Obstetrik Komplikasyonlar İlişkin Veriler

Obstetrik Tıp'ta Doğum, Gebelik ve Loğusalık dönemini inceleyen bilim dalıdır.

**OBS Antepartum Kanama:** Gebeliğin 28. Haftasından doğum başlangıcına kadar geçen süre içerisinde anne üreme organında kanama olması olarak tanımlanmaktadır. (KÜÇÜKÖZKAN T.,2011)

**OBS Kronik Hipertansiyon:** Ortalama her 10 gebelikten 1'inde görülen yüksek kan basıncı hastalığı olarak tanımlanabilen ve anne, bebek ölümleri ile sonuçlanabilen bir durumdur. (LAÇIN S.)

**OBS\_Preeklamsi:** Gebelik sırasında annenin tansiyonun yüksek düzeyde seyretmesi, el, ayak ve yüzde şişlik, idrarda fazla miktarda protein atımı ile ilgili bir hastalıktır. (YAKUT Y.)

**OBS HELLP:** Gebelikte annede görülen Hipertansiyonu ileri boyutu Kırımızı kan hücrelerinin tahribatı, Karaciğer enzim artışı, Kanın pıhtılaşma görevini sağlayan hücrelerin azalmasıyla ortaya çıkan HELLP Sendromu olarak adlandırılan durumdur. (MUMCU A.)

**OBS Anne Enf:** Gebelik esnasında annede enfeksiyon görüme durumudur.

**OBS\_Izo\_imminuzasyon:** Gebe ile fetüs arasında kan grubu uyumsuzluğuna bağlı olarak ortaya çıkan hastalık veya hastalıklar durumudur. (DANIŞMAN N.,2011)

**OBS\_EMR:** EMR(Erken Membran Rüptürü) Suların Erken Gelmesi durumudur. (1:Yok, 2:Var)

**OBS Çoğul Gebelik:** İkiz veya daha fazla sayıda gebelik olması durumudur.

**OBS\_Polihidramniyoz:** Gebelik sırasında bebeğin içerisinde bulunduğu sıvının normalden çok fazla olması durumudur. (KOYUNCUOĞLU A.O.)

**OBS\_R\_Preterm\_eylem:** Gebeliğin 20. Haftası ile 37. Haftası arasında gerçekleşen doğumlardır. (İLBEYİ P.)

**OBS\_Oligohidramniyoz:** Gebelik sırasında bebeğin içerisinde bulunduğu sıvının normalden az olması durumudur. (KOYUNCUOĞLU A.O.)

#### 4.4.8. Gelişmiş Laboratuvar Testleri İle İlgili Veriler

Yenidoğan yoğun bakıma sepsis şüphesi ile yatırılmış olan hastalara 0 Saatte tedavi başlamadan önce alınmış kan değerleri ve tedavi başladıktan sonra 24 saat sonra tedavinin ne durumda olduğunu görmek için aynı testler 24 saat sonra tekrar alınmıştır.

**Tablo 6.** Veri Seti içerisindeki Laboratuvar Testleri

0 Saatte	24 Saate
CRP0H	CRP24H
PROCal0H	PROCal24H
SSLinFSLinA0H_Lenfosit	SSLinFSLinA24H_Lenfosit
SSLinFSLinC0H_Monosit	SSLinFSLinC24H_Monosit
SSLinFSLinD0H_Granülosit	SSLinFSLinD24H_Granülosit
CD14_CD45_E4_0H	CD14_CD45_E4_24H
CD14_0H_Monosit	CD14_24H_Monosit
CD45_0H_Lenfosit	CD45_24H_Lenfosit
CD11B_0H_Lenfosit	CD11B_24H_Lenfosit
CD64_0H_Lenfosit	CD64_24H_Lenfosit
CD11B_0H_Monosit	CD11B_24H_Monosit
CD640H_Monosit	CD64_24H_Monosit
CD11B_0H_Granülosit	CD11B_24H_Granülosit
CD64_0H_Granülosit	CD64_24H_Granülosit
CD14_CD45_0H_Lenfosit	CD14_CD45_24H_Lenfosit
CD11B_CD64_0H_Lenfosit	CD11B_CD64_24H_Lenfosit
CD11B_CD64_0H_Monosit	CD11B_CD64_24H_Monosit
CD11B_CD64_0H_Granulosit	CD11B_CD64_24H_Granulosit

CRP<sup>20</sup> Düzeyleri hastalardan alınan 2-3 cc'lik kanların BEHRING BN2 cihazında CRP düzeyini ölçen laboratuvar kitleri yardımıyla düzeyleri belirlenmiş ve veri seti içerisine kayıt edilmiştir. Hastaların Prokalsitonin<sup>21</sup> değerleri ise ROCHE COBAS 601 DİAGNOSTIN cihazı ile çalışılarak sonuçlar veri setine işlenmiştir. CD11b-CD64 değerleri ise Bechman Coulter marka CYTOMICS FC500 model cihaz ile çalışılarak sonuçlar veri setine kayıt edilmiştir. (UZUN F.,2015)

#### 4.5. Veri Setindeki Sayısal Değerler

Veri seti içerisindeki 128 hastaya ilişkin bazı sayısal değerler aşağıdaki tabloda gösterilmiştir.

**Tablo 7.** Veri Seti içerisindeki Genel verilere ilişkin bazı sayısal değerler

		Erkek		Kız	
Cinsiyet		77	60%	51	40%
Doğum Şekli	Sezeryen (C/S)	58	62%	36	38%
	Normal	19	56%	15	44%
Doğum Ağırlığı	<1500	14	48%	15	52%
	1500-2500	20	61%	13	39%
	2500-4000	42	65%	23	35%
	>4000	1	100%	0	0%
Doğum Haftası	>=38	37	60%	25	40%
	32-37	23	70%	10	30%
	28-31	6	35%	11	65%
	<28	11	69%	5	31%
Apgar Skoru	1.Dakikası <=7 olanlar	71	59%	50	41%
	1.Dakikası >=8 olanlar	6	86%	1	14%
	5.Dakikası <=7 olanlar	15	48%	16	52%
	5.Dakikası >=8 olanlar	62	64%	35	36%
Kan grubu Uyuşmazlığı	Yok	67	59%	46	41%
	Var	10	67%	5	33%
Celestone Uygulaması	Yok	69	60%	46	40%
	Var	8	62%	5	38%
Rhoam Uygulaması	Yok	2	67%	1	33%
	Var	75	60%	50	40%
Sepsis Tanı Zamanı	<72 Erken Sepsis	2	67%	1	33%
	>72 Geç Sepsis	37	51%	36	49%
	>30 Gün Çok Geç Sepsis	11	92%	1	8%
	Tanı almamış hastalar	27	68%	13	33%
Çocuk Sepsis ile ilgili Hikâyesi varmı?	Yok	57	55%	47	45%
	Var	20	83%	4	17%
Antibiyotik Kullanımı	Yok	29	66%	15	34%
	Var	48	57%	36	43%
RDS	Yok	64	63%	37	37%
	Var	13	48%	14	52%
Surfactan Uygulaması	Yapıldı	13	50%	13	50%
	Yapılmadı	64	63%	38	37%

<sup>20</sup> Vücudunuzda bulunan iltihap ve enfeksiyon ile ilgili bilgi veren bir değerdir.

<sup>21</sup> Tiroid bezinde, akciğerde ve bağırsaklarda üretilen, kalsitonin hormonunun düzeyinin belirlenmesi

Veri seti içerisinde erkek bebeklerin oranı %60 iken kız bebeklerin oranı %40 düzeyindedir. Doğumu gerçekleşen erkek çocuklarının %75 Sezeryen yöntemi ile dünyaya gelmişken %25 normal yolla dünyaya gelmiştir. Kız çocuklarında ise oranı %70 sezeryen %30 normal yollarla doğum şeklinde görülmektedir. Doğum ağırlıkları bakımından Kız çocuklarının doğum ağırlıklarının erkek çocuklara göre düşük olduğu saptanmıştır. Kız çocuklarının doğumunun Erkek çocuklara göre doğum haftalarının daha erken olduğu saptanmıştır.

**Tablo 8.** Veri Seti içerisindeki bazı hastalık verilerine ilişkin bazı sayısal değerler

	<b>Erkek</b>		<b>Kız</b>	
	Prematüre Doğum	38	%49	24
EMR'li Anne Bebeği	6		7	
Koryoamniyoit	-		3	
Bebeğe Doğum öncesi, doğum sonrası veya doğum esnasında eşlik eden hastalıkları	202		135	

Veri Seti içerisindeki 128 bebekten 62 bebek prematüre doğmuştur. Bunlardan 38 bebek erkek, 24 bebeğin ise cinsiyeti kızdır. Doğumu Prematüre olmayan 16 erkek ve 15 kız bebek sepsis tanısı almıştır. Prematüre doğduktan sonra sepsis tanısı koyulmayan 35 bebek bulunmaktadır. Bunların 23'ü erkek 12'si ise kız bebeğidir. Veri seti içerisindeki prematüre doğum yüzdeleri iki cinsiyet içinde birbirlerine çok yakındır. Veri Seti içerisinde 13 Erkek bebeğin 5'ine 2 tane hastalık eşlik ederken 8 tanesine 1 hastalık eşlik etmesine rağmen bu bebekler sepsis tanısı almamıştır.

## BEŞİNCİ BÖLÜM

### 5. VERİ SETİNE UYGULANACAK OLAN VERİ MADENCİLİĞİ METOTLARI

Veri seti içerisindeki genel veriler bölümündeki Cinsiyet, Doğum Şekli, Doğum Kilosu, Doğum Haftası, Gerçek Doğum Haftası, Kan Grup Uyuşmazlığı, Celestone Uygulaması, Rhogam Uygulaması, Antibiyotik Kullanımı, RDS, Surfactan alanlarına Mesafeye dayalı sınıflandırma yöntemi olan K-En yakın komşu (k-Nearest Neighbors - KNN) ve İstatistiğe dayalı sınıflandırma için ise Naive Bayesian algoritması uygulanarak veri seti içerisinde sepsis tanı zamanı alanı bulunmaya çalışılmış ve bir erken uyarı karar destek sistemi ortaya koymak hedeflenmiştir.

K Nearest Neighborhood algoritması (KNN) içerisinde farklı sınıflarda bulunan elemanları en yakın benzerlik algoritmasına göre sınıflandırarak bundan sonra gelecek yeni elemanı daha önceki sınıflarla karşılaştırarak sınıfını belirlemeye veya en yakın sınıfını tahmin etmeye yarayan algoritmadır.(KAŞIKÇI T.,2014) Karışıklık matrisi makine öğrenimi alanında çalışan bir algoritmanın görselleşmesini, anlaşılabilir olmasını sağlayan tablo yapısıdır. KNN algoritmasının sepsis veri seti için üretmiş olduğu çıktı değerleri Tablo-8, Tablo-9'da açıklanmaya çalışılacaktır. KNN algoritmasında N varsayılan değeri 3 olarak hesaplama yapılmıştır. (POWERS D.,2011)

Naive (Saf) Bayes Algoritması istatistiğe dayalı Bayes sınıflandırma algoritmalarından birisidir. Olasılık tabanlıdır. Bu algoritmaya göre kelimeler sınıftan bağımsızdır. Uygulanabilirliği kolaydır ve performans bakımından başarılıdır. Naive Bayesian uygulanması esnasında Çapraz Doğrulama (Cross-validation) yöntemi çok önemli rol oynamaktadır. Temel olarak hedefin öngörüldüğü ortamlarda kullanılır ve bir tahmin modelinin pratikte ne kadar doğru bir performans göstereceğini tahmin etmek ister. Bir tahmin probleminde, bir modele genellikle eğitimin çalıştırıldığı bilinen bir veri kümesi ( eğitim veri kümesi ) ve modelin test edildiği ( veri doğrulama veri seti veya test olarak adlandırılır) bilinmeyen veri (veya ilk görülen veriler) veri kümesi verilir. Çapraz geçerliliğin amacı, modelin tahmin edilmesinde kullanılmayan yeni verileri tahmin etme yeteneğini test etmek, aşırı takmak gibi sorunları işaretlemek ve modelin nasıl geliştirileceğine dair bir fikir vermektir.

Bir çapraz doğrulama, bir veri numunesinin tamamlayıcı alt kümelere ayrılmasını, analizi bir alt kümede ( eğitim seti olarak adlandırılır) gerçekleştirmeyi ve diğer alt kümedeki ( doğrulama kümesi veya test kümesi olarak adlandırılan) analizini doğrulama işlemini içerir. Değişkenliği azaltmak için, çoğu yöntemde çapraz bölümlenimin birden fazla turu farklı bölümler kullanılarak gerçekleştirilir ve validasyon<sup>22</sup> sonuçları, modelin tahmini performansının bir tahminini vermek için turlar üzerinde birleştirilir. (PIERRE A.,1982)

**Tablo 9.** KNN ve Naive Bayesian Tabakalı çapraz doğrulama Özeti

	<b>KNN</b>	<b>Naive Bayesian</b>
<b>Doğruluk Oranı</b>	<b>% 94.53</b>	<b>% 93.75</b>
Toplam Örnek Sayısı (Total Number of Instances)	128	128
Doğru Sınıflandırılmış Örnekler (Correctly Classified Instances)	121	120
Ortalama Mutlak Hata (Mean absolute error)	0.0935	0.0919
Karekök Ortalama Hata (Root mean squared error)	0.2283	0.2255
Algoritmanın Çalışma Süresi (Operation Time of Algorithm)	0.8769	0.8603

Toplam 128 hastadan oluşan veri seti içerisinde 121 hastanın verisi KNN algoritmasına göre %94,53 'lük bir doğruluk oranı sınıflandırılırken, Naive Bayesian algoritması aynı hasta sayısına göre 120 hasta verisini %93.75 oranında doğru sınıflandırılmıştır. KNN algoritmasında ortalama mutlak hata oranı 0.0935 oranında iken Naive Bayesian'da durum 0.0919'dur. KNN algoritmasında Karekök Ortalama hata oranı 0.2283 iken Naive Bayesian' da bu oran 0.2255 hesaplanmıştır. Algoritmaların çalışma süreleri göz önüne alındığında Intel Core I5, 4 GB Ram, 2 GB Ekran Kartına sahip PC'de KNN 0.8769 oranında iken Naive Bayesian algoritmasına göre kappa istatistik değeri 0.8603 olarak hesaplanmıştır.

**Karışıklık Matrisi:** Karışıklık matrisi ile algoritma sonuçları arasında anlaşılabilir birçok veri tablo halinde ifade edilerek anlamlı bir hal almaktadır. Aşağıda

<sup>22</sup> Bir ürünün, bir sürecin veya bir sistemin işlevini, önceden belirlenmiş gerekliliklere uygun olarak yerine getirdiğine emin olmak için gerçekleştirilen çalışmalardır.

yeni doğan veri setine uygulanmış olan KNN ve Naive Bayesian Algoritmalarına ilişkin sonuçlar tek tabloda verilmiştir.

**Tablo 10.** KNN ve Naive Bayesian Karışıklık Matrisi Değerleri

	KNN		Naive Bayesian	
	P	N	P	N
Sepsis Pozitif	82	6	81	7
Sepsis Negatif	1	39	1	39

KNN algoritmasına göre Gerçek pozitiflerin sayısı 82 iken Gerçek negatiflerin sayısı 39'dur. Yanlış negatiflerin sayısı 6 iken yanlış pozitiflerin sayısı 1 olarak tespit edilmiştir. Naive Bayesian algoritmasına göre gerçek pozitif sayısı 81 iken gerçek negatiflerin sayısı 39'dur. Yanlış negatiflerin sayısı 7 iken, yanlış pozitiflerin sayısı 1 olarak tespit edilmiştir. Her iki algoritmada yanlış pozitif sayısını 1 bulurken, Gerçek Negatif sayısını 39 bulmuştur. İki algoritma arasındaki farklılık KNN algoritması Gerçek pozitif Naive bayesian algoritmasından 1 fazla bulmuştur. KNN algoritması 88 sepsis tanısı almış hastanın 82 'sini direk tespit ederken 40 adet kontrol sepsis hastasının 39'unu doğru tespit etmiş ve 1 adet sağlam olan hastanın aslında sepsis hastası olduğunu tespit etmiştir. Naive Bayesian algoritması 88 sepsis tanısı almış hastanın 81 'sini doğrudan tespit ederken 40 adet kontrol sepsis hastasının 39'unu doğru tespit etmiş ve 1 adet sağlam olan hastanın aslında sepsis hastası olduğunu tespit etmiştir.

**Tablo 11.** KNN ve Naive Bayesian Sınıf Tarafından Detaylı Doğruluk

Sınıf (Class)	KNN			Naive Bayesian		
	Sepsis Pozitif	Sepsis Negatif	Ağırlıklı Ort. (Weighted Avg.)	Sepsis Pozitif	Sepsis Negatif	Ağırlıklı Ort. (Weighted Avg.)
Gerçek Pozitif (TP Rate)	0,932	0,975	0,945	0,920	0,975	0,938
Yanlış Pozitif (FP Rate)	0,025	0,068	0,038	0,025	0,080	0,042
Keskinlik (Precision)	0,988	0,867	0,950	0,988	0,848	0,944
Hassasiyet (Recall)	0,932	0,975	0,945	0,920	0,975	0,938
F-Skoru (F-Measure)	0,959	0,918	0,946	0,953	0,907	0,939

KNN algoritması Gerçek pozitif sepsis tanısı almış hastaların ortalamasını %0.945 bulmuştur. Yanlış pozitifler sepsis tanısı almamış hastaların ortalamasını %0.038 olarak tespit etmiştir. Naive Bayesian algoritması ise gerçek pozitif sepsis tanısı almış hastaların ortalamasını %0.938 bulunmuştur. Yanlış pozitifler sepsis tanısı almamış hastaların ortalaması ise %0.042'dir. Her iki algoritma için keskinlik değerlerini karşılaştıracak olursak KNN için ortalama %0,950 olan oran Naive Bayesian algoritması için ortalama %0.944'dür. Hassasiyet değerleri ise KNN algoritmasında ortalama %0.945 iken Naive Bayesian algoritmasında bu değer %0.938'dir. F-Skoru karşılaştırması KNN algoritmasında ortalama %0.946 iken Naive Bayesian algoritmasında ortalama %0.938 olarak hesaplanmıştır. Her iki algoritmanın da yeni doğan veri setinde üzerindeki etkileri analiz edilmiş ve çıkarımları ortaya koyulmuştur.



## SONUÇ VE ÖNERİLER

Veri madenciliği yöntemlerinin tıp alanında aktif olarak kullanılması, veri madenciliği yöntemlerinin bilimsel makaleler dışında Hastane Bilgi Yönetim Sistemi yazılımı geliştiren firmalar tarafından sistemlerine kazandırılması bazı hastalıkların tespitinden önden bir uyarı niteliğinde olacak ve bazı hastalıklara karar verme noktasında hekime yardımcı karar destek sistem olacaktır. Hastane bilgi sistemleri üzerine yazılım geliştirmekte olan ekiplerin veri madenciliği metotlarını özümseyerek bu metotları Hastane bilgi yönetim sistemlerinin tüm süreçlerinde geliştirilen sistemlerin içerisine gömülü şekilde yerleştirmeleri faydalı olacaktır. Ülkemizde sağlık sisteminde olabildiğinden fazla veri toplanarak, bu verilerin veri madenciliği uzmanları tarafından incelenmesi ülkemizde geliştirilecek sağlık politikalarına yön vermede etkili olacaktır. Hastalıktan bölgeye, bölgeden hastalığa çok yönlü yapılacak incelemeler sayesinde yapılacak olan sağlık yatırımlarına ve ilaç sanayisi sektörüne fikir verecektir. Bunun yanı sıra veri tabanlarında zamanla biriken hastalık verileri incelenerek karar verme süreçlerini destekleyecek klinik karar destek sistemlerinin sayısı artırılmalıdır.

Bu tez çalışmasında tıpta kullanılan veri madenciliği metotları, tıpta yapılmış çalışmalar, yeni doğan sepsisi veri seti irdelenmiş, veri madenciliği algoritmalarından KNN ve Naive Bayesian algoritmaları açıklanmış ve bu algoritmalar yeni doğan veri sepsisi veri setinin bir bölümüne uygulanarak bir erken tanı tespit sistemi oluşturulmaya çalışılmıştır. Bu çalışmanın amacı otomatik bir tanı sistemi yapmaktansa erken teşhis için hekimlere bir karar destek sistemi ortaya koymaktır. Çalışma kapsamında elde edilen sonuçlara göre KNN algoritması % 94.53 doğruluk oranı ile veri tahmini yaparken Naive Bayesian algoritması % 93.75'lik bir doğruluk oranı üretmiştir. Bu doğruluk oranlarda göstermektedir ki; tıp alanında henüz daha keşfedilmeyi bekleyen içerisinden çok değerli bilgiler çıkarılabilecek veri yığınlarının var olduğu görülmektedir. Hiçbir olayın insan hayatından değerli olmadığı fikriyle birlikte İşletme açısından da bu tarz karar destek sistemlerinin sayısının artması Maliyet-Performans ve Maliyet-Zaman bakımından tasarruf sağlayacaktır. Ülke politikaları gereği oluşturulmuş öncelikler içerisinde var olan tıpta karar destek sistemlerinin oluşturulması bu gibi çalışmaları sayısının artırılması ile artış sağlayacaktır.

## KAYNAKÇA

- AKSOY İ. Ve arkd. “Hastane enfeksiyonlarının gizli örüntülerinin bulunması: Bir veri madenciliği yaklaşımı ”,İstanbul Üniversitesi İşletme Fakültesi Dergisi, 2010
- ALDEMİR B. “Respiratuvar distress sendromu(RDS)”, <http://www.drbulentaldemir.com/rds/> , Erişim Tarihi: 28.04.2018
- ANSARİ U. ve Ark. “Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction”, International Journal of Computer Applications (0975 – 8887) Volume 17– No.8, March 2011
- BALKAN F., “Tirotoksikoz ve Graves Hastalığı”, <http://endokrin.com.tr/tirotoksikoz-ve-graves-hastaligi/>, Erişim Tarihi:30.04.2018
- BAYKAL A. Veri Madenciliği Uygulama alanları, D.Ü.Ziya Gökalp Eğitim Fakültesi Dergisi 7, 95-107, 2006
- BEYATLI E., “Göbek iltihabı (Omfalit) – Sebepleri, tanı ve tedavisi”, <http://www.ertanbeyatli.com/gobek-iltihabi/>, Erişim Tarihi:30.04.2018
- CANTİLEV N. ve Ark. “Bronkopnömoni Nedeniyle Hastaneye Yatan Çocukların Annelerinin Öz-yeterliliklerinin İncelenmesi”, İzmir Kâtip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi, 2017
- CENGİZ A.B. , “Yenidoğan Sepsisi”, Çocuk Enfeksiyon Dergisi, 2009
- CENGİZ A.B. , “Yenidoğan Sepsisinde Değerlendirme ve Yönetim”, Güncel Pediatri Dergisi, 2007
- CİVELEK A., ”Kalp Kapak Hastalıkları”, <http://www.alicivelek.com/kalp-kapak-hastaliklari/>, Erişim Tarihi:30.04.3018
- ÇEKİÇ R. , TELÇEKEN S., “Ekg Sinyallerinin Kaba Kümeler Teorisi Kullanılarak Sınıflandırılması”, Anadolu Üniversitesi Bilim ve Teknoloji Dergisi A-Uygulamalı Bilimler ve Mühendislik, 2014
- DANIŞMAN N.,” Rh İzoimmünizasyonu”, Türkiye Klinikleri Dergisi, 2011
- DÜZGÜN N., ” Sistemik Lupus Eritematozus”, <http://ichastaliklariromatoloji.medicine.ankara.edu.tr/files/2014/02/Sistemik-Lupus-Eritematozus.pdf>, Erişim Tarihi:30.04.2018
- EDWARD H. SHORTLIFFE, JAMES J. CIMINO, Biyomedikal Bilişimi, Sağlık Hizmetleri ve Biyotıpta Bilgisayar Uygulamaları Kitabı

- ERBAŞ O., “Metabolik Hastalıklar”, <http://www.oytunerbas.com.tr/metabolik-hastaliklar/>, Erişim Tarihi: 30.04.2018
- ERTUĞRUL İ. Ve arkd. “Veri Madenciliği Uygulamasına İlişkin Paü Hastanesinde Hasta Profiline Belirlenmesi”, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 2013
- GÜLTEPE E. ve Ark. “From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system”, Journal of the American Medical Informatics Association, 2013
- GÜRSOY U., Uygulamalı Veri Madenciliği Sektörel Analizler Kitabı, Pagem Yayınevi, Ankara, 2012, s5.
- GÜRTUNA A.,” Bebeklerde Sarılık & Bilirubin Yüksekliği (Kernikterus)”, <https://www.aligurtuna.com/bebeklerde-sarilik-bilirubin-yuksekligi-kernikterus-1365.html>, Erişim Tarihi:30.04.2018
- HALEZEROĞLU S.,”Pnömotoraks Nedir?” , <http://www.akcigercerrahisi.com/pnomotoraks/>, Erişim Tarihi:30.04.3018
- HAOFAN YANG VE Yİ-PİNG PHOEBE CHEN , “ Data mining in lung cancer pathologic staging diagnosis: Correlation between clinical and pathology information” 2015
- İLBEYİ P.,” Erken Doğum (Preterm Eylem)”, <http://www.pelinilbeyi.com/hamilelik/komplikasyonlar/erken-dogum/>, Erişim Tarihi:05.05.2018
- İSBİR GÖZDE G.,” Travmatik Doğum Ve Hemşirelik Yaklaşımları” KAHSED, 2014
- KABASAKAL L.,” Hipotiroid”, <http://www.leventkabasakal.com/hastalik-turleri/tiroit/hipotiroit/>, Erişim Tarihi:30.04.2018
- KARAMAN ÖZİŞİK H., “İskemik-Hipoksik (İskemik-Anoksik) Ensefalopati”, Türkiye Klinikleri Cerrahi Tıp Bilimleri Dergisi, 2007
- KAŞIKÇI T., GÖKÇEN H.,” Metin Madenciliği ile E-Ticaret Sitelerinin Belirlenmesi”, Bilişim Teknolojileri Dergisi, 2014
- KILIÇ S., “İstatistikî İfadeyle kappa testi”, Journal of Mood Disorders Volume: 5, 2015
- KOYUNCUGİL S., ÖZGÜLBAŞ N., Veri Madenciliği: Tıp ve Sağlık hizmetlerinde kullanımı ve Uygulamaları, Bilişim Teknolojileri Dergisi Cilt:2 Sayı:2 Mayıs 2009

- KOYUNCUOĞLU A.O.,” Gebelikte Amniyon Sıvısı Ölçümü Nasıl Yapılır?”,  
<http://www.aliosmankoyuncuoglu.com/gebelikte-amniyon-sivisi-olcumu-nasil-yapilir>, Erişim Tarihi: 05.05.2018
- KÖKSAL N., “Prematüre Retinopatisi – Derleme”, Güncel Pediatri Dergisi, 2005
- KÜÇÜKÖZKAN T.,” Doğum Öncesi Kanama (Antepartum)”, <http://doktor-7.blogspot.com.tr/2011/06/dogum-oncesi-kanama-antepartum-hemoraji.html>,  
 Erişim Tarihi:05.05.2018
- LAÇİN S.” Gebelik ve Tansiyon Yükselmesi (Hipertansiyon)”,  
<https://www.selmanlacin.com/jinekoloji/gebelik/sorunlar-problemler/285-gebelik-ve-tansiyon-yukselmesi-hipertansiyon.html>, Erişim Tarihi:05.05.2018
- MANI S. ve Akd. “Medical decision support using machine learning for early detection of late-onset neonatal sepsis”, Journal of the American Medical Informatics Association, 2013
- MILOVIC B.,” Prediction And Decision Making In Health Care Using Data Mining”,  
 Kuwait Chapter of Arabian Journal of Business and Management Review, 2012
- MUMCU A. “Apgar Skoru”, <http://www.mumcu.com/apgar-skoru/> Erişim Tarihi:  
 28.04.2018
- MUMCU A. “Meningomyelose ve Spina bifida”, <http://www.mumcu.com/spina-bifida/>,  
 Erişim Tarihi: 28.04.2018
- MUMCU A.,”HELLP Sendromu”, <http://www.mumcu.com/hellp-sendromu/>, Erişim  
 Tarihi: 05.05.2018
- MUMCU Y. “Yenidoğan Sarılığı”, <https://tobbetuhastanesi.com.tr/bulletin/54-yenidogan-sariligi> , Erişim Tarihi: 28.04.2018
- OĞUZKURT P. “Özofagus atrezisi”, <http://www.pelinoguzkurt.com/ozefagus-atrezisi>,  
 Erişim Tarihi:30.04.2018
- OĞUZKURT P., “Pilor Stenozu”, Çocuk Cerrahisi Uzmanı,  
<http://www.pelinoguzkurt.com/pilor-stenozu> , Erişim Tarihi:30.04.2018
- OLIVEIRA A. ve Ark.,” Applying data mining techniques to improve diagnosis in neonatal jaundice.” BMC Medical Informatics and Decision Making, 2012
- ORHUNBİLGE N. Uygulamalı Regrasyon ve Korelasyon Analizi Kitabı 3.Baskı,  
 Nobel Akademik Yayıncılık, İstanbul, 2017, s20
- ÖRS R., “Yenidoğanın Geçici Takipnesi”, Türkiye Klinikleri J Pediatr Sci, 2013

- ÖZDAĞOĞLU A. ve arkd. “Ege Bölgesi’ndeki Bir Araştırma Ve Uygulama Hastanesinin Acil Hasta Verilerinin Simüle Edilerek Analizi” , İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 2009
- ÖZKAN H. ve ark. “Bronkopulmoner Displazide Risk Faktörleri - Özgün Araştırma”, Güncel Pediatri Dergisi, 2008
- ÖZKAN, Y., Veri Madenciliği Yöntemleri Kitabı, Papatya Yayınları Baskı 4,Ankara, 2016, s131
- ÖZKAN, Y., Veri Madenciliği Yöntemleri Kitabı, Papatya Yayınları Baskı 4,Ankara, 2016, s157
- ÖZKAN, Y., Veri Madenciliği Yöntemleri Kitabı, Papatya Yayınları Baskı 4,Ankara, 2016, s185
- PEKTAŞ, A.O. , SPSS ile Veri Madenciliği Kitabı, Dikey Eksen yayın evi, Ankara, 2013, s100.
- PIERRE A.,” Pattern Recognition: A Statistical Approach”,Londra, GB: Prentice-Hall.
- POWERS D.,” Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation”, Journal of Machine Learning Technologies.
- POYRAZ O., “Tıp’da veri madenciliği Uygulamaları: Meme kanseri veri seti analizi”, Trakya Üniverfsitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 2012
- SARIHAN A.,” Kan Gazı Nedir, Nasıl Yorumlanır ?”, <http://xn--aciltp-t9a.com/kan-gazi>, Erişim Tarihi:30.04.2018
- SAY A., “Asfiksi ve Hipoksi”, <https://www.diyadinnet.com/sa%C4%9F1%C4%B1k-4398-asfiksi-ve-hipoksi>, Erişim Tarihi: 30.04.2018
- SEVER L., “Vezikoüreteral Reflüde Tedavi ve İzlem Protokolü”, Güncel Pediatri Dergisi, 2005
- SOYSAL F.,”Anal Atrezi”, <http://www.feryalgunsoysal.com/tr/article/desc/37885/anal-atrezi.html>, Erişim Tarihi:30.04.2018
- ŞEKER Ş. , İş Zekası ve Veri Manenciliği Weka İle Kitabı, Net Kitap Yayınları, Ankara, 2013, s161-163
- ŞEN O., “Hidrocefali”, <http://www.orhansen.com/hidrocefali>, Erişim Tarihi:30.04.2018
- ŞEN V. ve Ark. “Yenidoğan Yoğun Bakım Ünitesinde Mekonyum Aspirasyon Sendromlu Yenidoğanların Değerlendirilmesi”, Van Tıp Dergisi, 2014

- ŞENTÜRK A., Veri Madenciliği Kavram ve Teknikleri Kitabı, Ekin Yayın Evi, İstanbul, 2006, s6.
- ŞENTÜRK A., Veri Madenciliği Kavram ve Teknikleri Kitabı, Ekin Yayın Evi, İstanbul, 2006, s7.
- ŞENTÜRK A., Veri Madenciliği Kavram ve Teknikleri Kitabı, Ekin Yayın Evi, İstanbul, 2006, s1.
- ŞENTÜRK A., Veri Madenciliği Kavram ve Teknikleri Kitabı, Ekin Yayın Evi, Ankara, 2006, s34
- TARHAN Ö.R. , “SIRS, Sepsis, Septik Şok, MODS, MOF”, Türk Cerrahi, <http://www.turkcerrahi.com/makaleler/sok/sirs-sepsis-septik-sok-mods-mof/>, Erişim Tarihi: 28.04.2018
- TEMĐ, Türkiye Endokronoloji ve Metabolizma Derneđi Diyabetik Çalışma Grubu Eğitim Broşürü, <http://www.turkendokrin.org/files/pdf/DiyabetNedir.pdf>, Erişim:30.04.2018
- TOPALOĞLU S.,” Gebelik ve rh (kan) uyumsuzluğu”, <http://www.suheylatopaloglu.com/gebelik/19-site-icerigi/makaleler/gebelik/146-gebelik-ve-rh-kan-uyusmazl-g>, Erişim Tarihi:28.04.2018
- TÜRK E., “Nekrotizan Enterokolit (NEC, NEK)”, <https://www.erdalturk.com.tr/hastaliklar/nekrotizan-enterokolit.html>, Erişim Tarihi:30.04.2018
- UÇAR A.B.,” Trombosit Nedir? Trombosit Düşüklüğü ve Yüksekliğinin Nedenleri Nelerdir?”, <http://www.tibbiyardim.com/trombosit-nedir-trombosit-dusuklugu-ve-yuksekliginin-nedenleri-nelerdir.html>, Erişim Tarihi:30.04.2018
- ULUDAĞ S.,” Gebelik Suyunun Azalması Oligohidramnios”, <http://www.seyfetinuludag.com.tr/Content.Asp?ContentID=9&SubContentID=20>, Erişim Tarihi: 28.04.2018
- URMAN B., “Intrauterin Gelişme Geriliđi – IUGR”, <http://www.bulenturman.com/gebelik-konulari/gebelik-komplikasyonlari/intrauterin-gelisme-geriligi-iugr> , Erişim Tarihi:30.04.2018
- US M., “Patent Duktus Arteriozus (PDA)”, [http://www.melihus.com.tr/Patent-Duktus-Arteriozus-\(PDA\).asp](http://www.melihus.com.tr/Patent-Duktus-Arteriozus-(PDA).asp), Erişim:30.04.2018

- UZUN F., “Yenidoğan Sepsisli Bebeklerde Erken Tanıda Serum Crp, Prokalsitonin Ve Cd64 Düzeylerinin Önemi”, Fırat Üniversitesi, Tıp Fakültesi-Çocuk Sağlığı ve Hastalıkları Anabilim Dalı, Tıpta Uzmanlık Tezi, 2015
- WEKA, “Data Mining Software in Java”, <https://www.cs.waikato.ac.nz/ml/weka/>
- YAĞMURLU A.,” Diyafragma Hernisi”, Çocuk Cerrahisi ve Çocuk Ürolojisi Uzmanı  
<http://www.aydinyagmurlu.com/tr/icerik/60/diyafragma-hernisi> Erişim Tarihi:30.04.2018
- YAĞMURLU A.,” İntestinal Atreziler”, Çocuk Cerrahisi ve Çocuk Ürolojisi Uzmanı  
<http://www.aydinyagmurlu.com/tr/icerik/65/intestinal-atreziler>, Erişim Tarihi:30.04.2018
- YAKUT Y., “Preeklampsi (gebelik zehirlenmesi) nedir? Belirtileri, nedenleri ve tedavisi”, <https://www.medikalakademi.com.tr/preeklampsi-gebelik-zehirlenmesi-belirtileri-nedenleri-ve-tedavisi/>, Erişim Tarihi:30.04.2018
- YAYLA M.,” Fetal Medikal Tedavi” ,  
<http://www.muratyayla.com/default.asp?CatId=55> , Erişim Tarihi:30.04.2018
- YILDIRIM P. ve arkd. , “Hastane Bilgi Sistemlerinde Veri Madenciliği”, Akademik Bilişim, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, 30 Ocak - 01 Şubat 2008

## EKLER

## Ek 1. Orjinallik Raporu



SOSYAL BİLİMLER ENSTİTÜSÜ  
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

## ÖĞRENCİ BİLGİLERİ

Adı-Soyadı	Aytaç TEKİN
Öğrenci Numarası	141217113
Enstitü Anabilim Dalı	Teknoloji ve Bilgi Yönetimi
Programı	Yüksek Lisans
Danışmanın Unvanı, Adı-Soyadı	Dr. Öğr. Üyesi Mustafa ULAŞ
Tez Başlığı (Türkçe)	Tıp'da Veri Madenciliği Uygulamaları: Yenidoğan Sepsisi Veri Seti Analizi

## SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

Yukarıda başlığı belirtilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 41 sayfalık kısmına ilişkin, 25/06/2018 tarihinde şahsım/tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orjinallik raporuna göre, tezimin benzerlik oranı % 4'tür.

Uygulanan filtrelemeler:

- 1- Kabul/Onay ve Bildirim sayfaları hariç,
- 2- Kaynakça hariç
- 3- Alıntılar hariç/dâhil
- 4- 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Yukarıda bilgileri verilen öğrencinin doktora tezi Sosyal Bilimler Enstitüsü Yönetim Kurulu tarafından belirlenen azami benzerlik oranlarını aşmadığını ve tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Dr. Öğr. Üyesi Mustafa ULAŞ  
Danışmanın Adı-Soyadı  
(İmzası)

Dr. Öğr. Üyesi Cem AYDIN  
Anabilim Dalı Başkanı  
(İmzası)

Lisansüstü tezler, savunma öncesinde intihal program raporu ile birlikte enstitüye teslim edilir.

İntihal raporu ile ilgili olarak etik kurallar dâhilindeki benzerlik oranları ilgili Enstitü Yönetim Kurulu tarafından belirlenir. (Enstitü Yönetim Kurulu tarafından tezin, intihal kapsamı dışında değerlendirilmesi için TURNITIN'den alınan raporda "benzerlik oranı"nın, "alıntılar hariç" en fazla %10, "alıntılar dâhil" % 30'u geçmemesi şeklinde kabul edilmiştir).



## ÖZGEÇMİŞ

1983 yılında Elazığ'da doğdum. İlköğrenimimi Ziya Gökalp İlkokulunda, Orta Öğrenimimi Atatürk Ortaokulunda, Lise Eğitimimi ise Gazi Endüstri Meslek Lisesi, Bilgisayar Yazılım bölümünü birincilikle bitirerek tamamladım. Üniversite Eğitimi, Fırat Üniversitesi Bilgisayar Teknolojileri ve Programcılığı Bölümü ve Anadolu Üniversitesi İşletme Fakültesini bitirerek tamamladım.

2013 yılında Gazi Üniversitesi, Bilişim Enstitüsü, Sağlık Bilişimi Anabilim dalında iki yıl süresince eğitim aldım. 2015 Yılında Fırat Üniversitesi, Sosyal Bilimleri Enstitüsü, Teknoloji ve Bilgi Yönetimi Anabilim dalında öğrenime devam etmekteyim.

Bilişim Teknolojileri alanında yirmi yıldır çalışmakta olup, Evli ve iki çocuk babasıyım.