

**T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**K-MEANS ALGORİTMASI İLE  
ELEKTROMANYETİK KİRLİLİK  
ANALİZİ**

**YÜKSEK LİSANS TEZİ**

**Feyza ÇEREZCİ**

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM  
MÜHENDİSLİĞİ**  
**Tez Danışmanı : Yrd.Doç.Dr. Serap KAZAN**

**Haziran 2015**

T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**K-MEANS ALGORİTMASI İLE  
ELEKTROMANYETİK KİRLİLİK  
ANALİZİ**

**YÜKSEK LİSANS TEZİ**

**Feyza ÇEREZCİ**

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM  
MÜHENDİSLİĞİ**

**Bu tez 15 / 06 /2015 tarihinde aşağıdaki jüri tarafından oybirliği/oyçokluğu ile kabul edilmiştir.**

.....  
**Jüri Başkanı**

.....  
**Üye**

.....  
**Üye**

## **BEYAN**

Tez içindeki tüm verilerin akademik kurallar çerçevesinde tarafımdan elde edildiğini, görsel ve yazılı tüm bilgi ve sonuçların akademik ve etik kurallara uygun şekilde sunulduğunu, kullanılan verilerde herhangi bir tahrifat yapılmadığını, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezde yer alan verilerin bu üniversite veya başka bir üniversitede herhangi bir tez çalışmasında kullanılmadığını beyan ederim.

Feyza ÇEREZCİ

15.06.2015

## **TEŐEKKÖR**

Çalıőmalarım boyunca deęerli yardım ve katkılarıyla beni yönlendiren hocam Yrd. Doç. Dr. Serap KAZAN'a ve manevi desteęiyle beni hiçbir zaman yalnız bırakmayan sevgili Anne ve Babama, teőekkÖrü bir borç bilirim.

## İÇİNDEKİLER

TEŞEKKÜR.....	i
İÇİNDEKİLER.....	ii
SİMGELER VE KISALTMALAR LİSTESİ.....	iv
ŞEKİLLER LİSTESİ.....	v
ÖZET .....	vi
SUMMARY.....	vii
BÖLÜM.1.	
GİRİŞ .....	1
BÖLÜM.2.	
ELEKTROMANYETİK RADYASYON VE GENEL TANIMLAR.....	4
2.1. Elektromanyetik Spektrum.....	5
2.2. Elektromanyetik Alan ve Dalga Kavramı .....	6
2.2.1. Reaktif yakın alan.....	7
2.2.2. Işıyan yakın alan.....	7
2.2.3. Uzak alan.....	7
2.2.4. Elektromanyetik güç yoğunluğu .....	8
2.3. Elektromanyetik Alanların İnsan Vücudu İle Etkileşimi.....	9
2.4. Yasal Düzenlemeler ve Sınır Değerler .....	10
2.5. Elektromanyetik Alan İnceleme ve Ölçüm Yöntemleri .....	14
2.6. Ölçüm Cihazları .....	16
2.6.1. Ölçüm prosedürü .....	16
2.6.2. Belirsizlik kestrimi .....	17

### BÖLÜM.3.

VERİ MADENCİLİĞİ .....	20
3.1. Veri Madenciliğinin Tanımı.....	20
3.2. Veri Madenciliğinin Tarihi .....	21
3.3. Veri Madenciliği Uygulama Alanları .....	23
3.4. Veri Madenciliğinde Karşılaşılan Problemler .....	24
3.5. Veri Madenciliği Teknikleri.....	27
3.5.1. Sınıflama .....	27
3.5.2. Kümeleme analizi.....	28
3.5.2.1. K-means algoritması .....	31
3.5.3. Birliktelik kuralları .....	32

### BÖLÜM.4.

UYGULAMA.....	33
---------------	----

### BÖLÜM.5.

SONUÇLAR.....	37
---------------	----

KAYNAKLAR .....	39
-----------------	----

ÖZGEÇMİŞ.....	41
---------------	----

## SİMGELER VE KISALTMALAR LİSTESİ

BBB	: Kan beyin bariyeri
CW	: Sürekli dalga
ELF	: Düşük frekanslı dalga
EMR	: Elektromanyetik radyasyon
IARC	: Uluslararası Kanser Ajansı
ICNIRP	: İyonize olmayan radyasyondan korunma komitesi
RF	: Radyo frekanslı dalga
SAR	: Özgün soğurma oranı
ICNIRP	: İyonize olmayan radyasyondan korunma komitesi
WHO	: Dünya sağlık örgütü
UV	: Morötesi ışın

## ŞEKİLLER LİSTESİ

Şekil 2.1. Elektromanyetik radyasyonun frekans dağılımı.....	6
Şekil 2.2. 10m yüksekliğinde bir binaya yerleştirilmiş baz istasyonu anteninin yönlü yaptığı ana ışıma lobu .....	12
Şekil 2.3. Radyo link antenleri .....	13
Şekil 2.4. Ölçüm konfigürasyonu. ....	15
Şekil 2.5. Uzun süreli baz istasyonu EMR ölçümü . ....	19
Şekil 2.6. Baz istasyonu karşısındaki bir evde elektromanyetik radyasyonu frekanslara göre dağılımı . ....	19
Şekil 3.1. Veri madenciliğinin birlikte bulunduğu disiplinler. ....	20
Şekil 3.2. Veri madenciliğinin tarihsel gelişim süreci.....	22
Şekil 3.3. Bilgi keşfi sürecinde veri madenciliği.....	26
Şekil 3.4. Hiyerarşik ve hiyerarşik olmayan kümeleme.....	30
Şekil 4.1. Weka programına kullanılan verilerin yerleştirilmesi.....	35
Şekil 4.2. Antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin kümelenmesi.....	36
Şekil 4.3. Antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin risk durumu .....	36
Şekil 4.4. Elektromanyetik alan değerlerinin iç ve dış mekana bağlı olarak kümeleme sonuçları .....	37



## TABLolar LİSTESİ

Tablo 2.1. Türkiye’de geçerli elektromanyetik radyasyon sınır değerleri .....	11
Tablo 2.2. Baz istasyonları için İtalya, İsviçre ve Türkiye’de uygulanan limitlerin karşılaştırılması .....	12
Tablo 2.3. Zamanla değişen alanlara maruziyet için ICNIRP Referans seviyeleri...	13
Tablo 2.4. Standard belirsizlik tablosu .....	17
Tablo 2.5. Elektromanyetik alan ölçümünde güvenlik analizi akış şeması.....	18
Tablo 4.1. Veri tabanının özellikleri .....	34
Tablo 4.2. Elektrik Alan Şiddet Seviyelerine Göre Doz Dağılımlarına Göre Risk Kriterleri.....	35

## ÖZET

Anahtar kelimeler: Veri Madenciliği, Örnek Uygulama, Elektromanyetik Kirlilik, GSM antenleri

İletişim teknolojisinin hızla yaygınlaştığı günümüzde, telekominikasyon antenlerinden yayınlanan elektromanyetik alanlar kentsel bölgedeki elektromanyetik kirliliğin ana sebebidir. Özellikle baz antenleri ev, okul, hastane gibi duyarlı alanlara yakın konumlandırılmışsa elektromanyetik radyasyon maruziyeti açısından risk oluşturabilmektedir. Yapılan bu veri madenciliği uygulaması çalışmasında Rize şehir merkezinde değişik lokasyonlara yerleştirilmiş olan baz istasyonlarının elektromanyetik kirlilik ölçüm değerleri 3 değişik kümede analiz edilmiştir. Kullanılan veriler SEMAM (Sakarya Üniversitesi Elektromanyetik Araştırma Merkezi) tarafından yayımlanan "Rize ili Elektromanyetik Kirlilik 2012 Raporu" 'ndan elde edilmiştir. Kümeleme opsiyonları baz istasyonlarının evlere uzaklıkları, elektromanyetik ölçüm verilerinin indoor/outdoor olması durumları ve üç farklı risk parametreleri itibarıyla belirlenmiştir. Elde edilen sonuçlar sürekli yaşam alanlarında elektromanyetik maruziyet seviyelerinin kontrolü ve azaltılması açısından önemli bilgiler sunmaktadır.

# **ANALYZING OF ELECTROMAGNETIC EXPOSURE FROM GSM ANTENNAS USING K-MEANS ALGORITHM**

## **SUMMARY**

Keywords: Data Mining, Sample Application, electromagnetic pollution, GSM antennas

Nowadays, overall the world the Electromagnetic Fields radiated from telecommunication antennas are main source of electromagnetic pollution in urban, especially when base antennas located adjacent the living areas such as homes, schools, hospitals in city. In this study, electromagnetic pollution measurements were analyzed with 3 different cluster for base stations which is placed at different locations in a city center. K-means algorithm is used for clustering. The data used are obtained from the Report named "Electromagnetic pollution 2012 study in Rize" which is published by the SEMAM (The electromagnetic research center of Sakarya University). Clustering option is determined by the parameters such as the distance from homes to base stations, measurements of electromagnetic field data's risk status.

## **BÖLÜM 1. GİRİŞ**

Radyo frekanslı mikro dalga enerjisinin yoğun kullanıldığı kablolu, kablosuz iletişim teknolojisinin sürekli gelişimi ve bu teknolojinin piyasaya sürdüğü ürünlerinin kullanımının hızla yaygınlaştığı günümüzde sokak, ev ve ofislerimizde elektromanyetik radyasyon birikmesine sebep olmaktadır. İletişim teknolojisinin günlük kullanıcılara haberleşmede sağladığı piyasaya sunduğu her yeni ürünlerle birlikte çevremizde exponansiyel artan bir elektromanyetik radyasyon oluşmaktadır. İnsan ve çevre sağlığı bakımından istenmeyen bu radyasyon birikimine bilimsel literatüre “elektromanyetik kirlilik” isimli yeni bir kirlilik tanımlaması ortaya çıkartmıştır (Balmori,2009). Günümüzde şehir yaşamımızda bu kirliliğin en önemli bileşeni çevremize baktığımız zaman bazen baca içerisine gizlenmiş ve bazende açıktan GSM frekanslarında elektromanyetik radyasyon yayan baz istasyonlarıdır.

Elektromanyetik radyasyonun canlı yaşamına yapacağı olumsuz etkisine dair, kamu oyunda tartışmalar yapılmakla birlikte, yoğun şiddette ve sürekli elektromanyetik enerji etkisinde kalmanın önemli sağlık sorunlarına yol açacağı konusunda herhangi bir tereddüt mevcut değildir. Ancak bu konuda araştırma yapan dünyaca saygın bilim adamlarının yayınladığı Bioinitiative 2012 raporunda belirtildiği gibi mevcut limitlerin çok altında bile insan vücudunu elektromanyetik radyasyonla biyolojik etkileşime girmesi nedeniyle ciddi sağlık sorunlarının ortaya çıkacağı ifade edilmektedir.

Son zamanlarda yapılan araştırmalardan AB ülkelerinde 12 araştırma merkezinin yaptığı Reflex projesine göre cep telefonlarının ve baz istasyonlarının kullandığı frekans bölgesindeki radyasyonun DNA kırılmalarına, gen değişikliklerine sebep olabildiği rapor edilmektedir. ABD de yayınlanan ve uluslararası saygın 14 bilim adamının yaptıkları çalışmalara dayanan Bioinitiative Raporunda; elektromanyetik

alanlar ile ilgili standartların yetersiz olduđu, kanser yaptığının delilleri, bağıřıklık sistemine zarar verdiđi açıklanmaktadır (Bioinitiative , 2012).

Bunlardan bazıları;

1. Kan Beyin Bariyeri (BBB, Blood Brain Barrier) sızıntısı (Alzheimer's, Parkinson hastalığı)
2. Nöronların ölümü ve beyin hücrelerinin zarar görmesi
3. Kalsiyum akışı, kalp ritim bozukluđu ve kalp durması
4. Habis ve kötü huylu beyin tümörleri
5. Sperm hücrelerinin zarar görmesi ve ölümü
6. Büyümei etkilemek, kusurlu doğum ve kısırlık, Akustik neuroma
7. Alzheimer's, Bunama, Leukemia ve Lymphoma

olarak belirtilmektedir. Bu nedenle Uluslararası Kanser Ajansı (IARC) 2011 yılında elektromanyetik alanları (glioma) etkisi yapması nedeniyle 2B sınıfı kanserojen olarak listesine almıştır.

Ülkemizde, sayıları 90.000 civarında olan şehirlerimizde faaliyet gösteren baz istasyonlarından yayılan elektromanyetik radyasyon maruziyetinin oluşturduđu risk seviyelerinin belirlenmesi; yüksek riskli durumların önlenmesi açısından önemlidir.

Elektromanyetik enerjinin iletişim, sađlık, eğitim, askeri gibi her alanda yoğun ve kaçınılmaz kullanımı, belli bir yarar ve zarar boyutunun birlikte tartıřılmasını da gerektirmektedir. Her aktivitede olduđu gibi modern yaşamın zorunlu kıldıđı elektromanyetik radyasyon maruziyetinde de az veya çok seviyelerde zarar durumu mevcuttur. İnsan yaşamında bazı hasarlara yol açacak tüm etkenler risk olarak adlandırılır ve sonucu belli olmayan her işin şüphesiz bir riski vardır.

Risk, elektromanyetik radyasyon maruziyetinde olduđu gibi kişilere bađlı veya kişilerden bağımsız etkenler ile ortaya çıkabilecek, istatistik olarak kalıplandırılabilir olacak olayların tümünü oluşturur. Kontrol edilebilmeleri deđişkenlik gösteren ve kötü sonuçlar doğurabilecek risklerin önüne geçilmesi için gerekli

arařtırmaların yapılması kaçınılmazdır. İnsan hayatının her türlü durumda korunması için gerekli önlemlerin alınması, bir ülkenin gelişmişliğini gösteren en büyük parametredir. Bu konuda örneğin İsviçre, İtalya, Polonya, Çin gibi ülkeler elektromanyetik kirlilik kontrolü için; ALARA olabildiğince en düşük doz (as low as achieve) ilkesini yıllardır uygulamaktadırlar (Cerezci,2015).

Bu çalışmada; Rize Şehir merkezinde 2012 yılında iç mekan ve dış mekan mekanlarda yapılan baz istasyonlarından yayılan elektromanyetik radyasyon ölçüm değerleri veri madenciliği kümeleme tekniğinin K means algoritması kullanılarak analiz edildi ve üç değişik risk kümelemesi ile elde edilen sonuçlar baz istasyonlarına yakınlık, ölçüm lokasyonlarının iç mekan – dış mekan olması parametrelerine göre tartışılmıştır.

İnsanların yaşam ortamlarında evlerinde ve çalıştıkları mekanlarda sağlıklı ve güvenli olması beklenir. Bu amaç doğrultusunda risk analizi yaptırılır. Ancak risk analizi sonucunda risk oluşturan etkenler tespit edildikten sonra her bir durum kimlerin ne şekilde ve hangi şiddette zarar göreceği belirlenir. Oluşturulan veriler doğrultusunda ulusal veya uluslararası standartlar için risk azaltıcı yöntemler geliştirilebilir. Bu tezde yapılan çalışmalar konsept itibariyle ülkemizde elektromanyetik radyasyon kirlilik risk kontrolünde yapılması düşünülecek yöntem tasarımlarına katkı sağlayacağı umulmaktadır.

## **BÖLÜM 2. ELEKTROMANYETİK RADYASYON VE GENEL TANIMLAR**

Alan yani radyasyon; belirli bir bölgeye dağılmış bulunan ve o bölgedeki herhangi bir cismi etkileyebilen fiziksel bir nesnedir. Elektrikle çalışan tüm cihazlar, akım taşıyan tüm iletkenler, herhangi bir yük olmaksızın akım taşımaya bile üzerinde belli bir gerilim bulunan sistemler potansiyel birer elektromanyetik (EM) alan kaynağıdır. Elektromanyetik alan veya elektromanyetik radyasyon denilince dalgalar halinde yayılmış bulunan, çevremizi bizimle paylaşan, elektrik ve manyetik özellikleri ile vücudumuzla yayılan biyo-etkileşimler anlatılmak istenir.

Son yıllarda bilim ve teknolojiadaki gelişmelere bağlı olarak, kişisel, endüstriyel ve ticari amaçlı, televizyon, bilgisayar, radyo vericileri, cep telefonu ve baz istasyonları, mikro dalga fırınlar, yüksek gerilim hatları, endüstriyel ve tıbbi tanı için kullanılan cihazlar vb. elektromanyetik dalga ve alan oluşturduğu bilinen sistemlerin, çevre ve insan sağlığı açısından bazı risklere yol açtığı bilinmektedir. Kablosuz ev telefonları, internet ve cep telefonu kullanımının tüm dünyada hızla çoğalması, endüstri ve tıp alanındaki çeşitli uygulamalar nedeniyle çevremizde radyo frekanslı elektromanyetik alan seviyeleri çok hızlı artmaktadır.

Çevremizde oluşan yoğun elektromanyetik radyasyon yakın çevrede bulunan bina dışındaki baz istasyonlarından, yüksek gerilim hatlarından, radyo-TV verici antenlerinden kaynaklanabilir. Bina içinde ise evin elektrik tesisatından, eğer tavan yüksekliği düşükse, tasarruflu ampulden, saç kurutma makineleri, TV ekranları, mutfaktaki mikrodalga fırından, Wi-Fi kablosuz erişimlerden kısacası elektrikle çalışan her cihazdan sızıntı şeklinde kaynaklanabilir. Yaşam alanlarında oluşan bu alan seviyeleri hatta bazı durumlarda ciddi sağlık sorunlarına neden olabilecek yüksek seviyelere ulaşabilir.

Büyük iş merkezlerinin kat tavanlarına yerleştirilmiş bina içi haberleşme baz (indoor baz) antenleri de bina içinde elektromanyetik radyasyon kaynağıdır. Gözle görünmeyen elektromanyetik radyasyonlar binalardan hatta kurşundan da geçebilmektedir. Elektromanyetik radyasyonun önlenmesi için ekranlama ismi ile tanımlanan elektromanyetik radyasyonun frekansına göre enerjisini absorbe edici uygulamalar yapılır.

## 2.1. Elektromanyetik Spektrum

Radyasyon, iyonlaştırıcı radyasyon (yani nükleer radyasyon) ve iyonlaştırmayan radyasyon (yani elektromanyetik radyasyon-EMR) olmak üzere iki sınıfta incelenir. İyonlaştıran ve iyonlaştırmayan ayrımının gerisindeki temel yaklaşım şu fiziksel olgudan anlaşılabilir. Bir atomu iyonlaştırmak için son yörüngesinden bir elektronu kopartacak düzeyde enerji aktarılması gerekmektedir. 12 eV' dan daha yüksek enerjili dalgalar atomu iyonlaştıracak enerjiye sahiptir. Bu şekilde enerjisi olan radyasyonlar iyonlaştırıcı radyasyon olarak isimlendirilir.

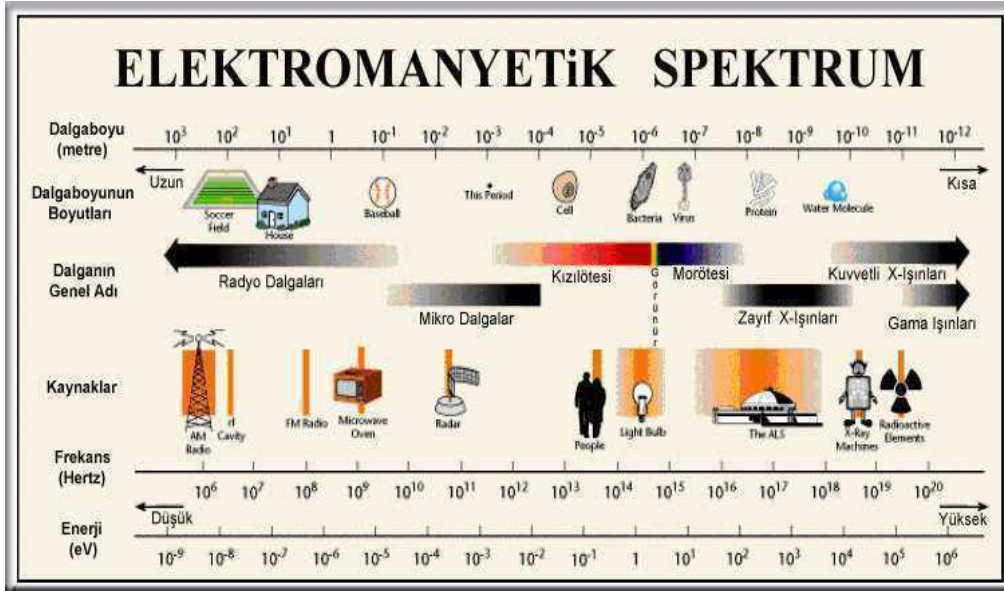
Elektromanyetik radyasyon için genel kabul frekans aralığı 0 Hz-300 GHz arası dalgalar anlaşılır. Şekil 2.1.'de verilen spektrumun 300 GHz'den bölünmesi, 300 GHz'in altındaki RF dalgalarının iyonlaştırmayan (non-ionizing) üzerindeki de iyonlaştıran (ionizing) olarak tanımlanmasıdır. Elektromanyetik radyasyon genel olarak mikrodalgalar, kızılötesi (IR) ışık, morötesi (UV) ışık, x ışınları ve gamma ışınlarıdır. Mikro dalgalar 1mm ile 100 cm arasında (frekansı, 1 GHz-300 GHz) dalga boyuna sahiptirler. Görünür ışık frekansları  $10^{14}$ Hz' lerdir. Bu frekansların üstü iyonlaştırıcı radyasyon bölgesidir.

Elektromanyetik spektrumda;

1. 3 Hz – 3 kHz arası çok çok düşük frekans bölgesi
2. 3 kHz – 30 kHz arası çok düşük frekans bölgesi
3.  $10^{12}$  Hz'lere kadar olan bölge radyo dalgaları diye anılır.
4.  $10^{12}$  Hz'ler kızıl ötesi ışımaya bölgesidir.



Bütün nesnelar bu frekanslarda ışınım yaparlar. Örneğın vücudumuzdaki ısınnın yaklaşık %60'ı kızıl ötesi ışınım ile dışarı atılır. Işıl kamera denen cihaz kızıl ötesi frekanslara duyarlı bir alıcıdır.



Şekil 2.1. Elektromanyetik Radyasyonun Frekans Dağılımı

## 2.2. Elektromanyetik Alan ve Dalga Kavramı

Elektromanyetik alanlar, sağlık etkisi bakımından yüksek frekanslı ve düşük frekanslı olarak iki ayrı grupta toplanabilir. Çok düşük frekanslı alanlar ELF bandı olarak incelenir. Yüksek frekanslı alanlar ise RF veya mikrodalga bantlı alanlar olarak incelenir. Elektromanyetik alanın oluşabilmesi için bir elektrik yükünün olması gerekmektedir. Elektromanyetik alan, elektrik alan ve manyetik alan olmak üzere iki bileşenden oluşur. Elektrik alan şiddetinin birimi V/m, manyetik alan şiddetinin birimi ise A/m'dir. Elektromanyetik Radyasyon (veya enerji) elektrik ve manyetik bileşenler halinde (örneğin; radyo dalgaları, GSM sinyalleri gibi) enerjinin yayılmasıdır.

Antenden yayılan elektromanyetik alanlar boşlukta ışık hızıyla yayılır. Saniyede oluşan elektrik veya manyetik alan darbe sayısına dalganın frekansı ve Hertz birimi ile gösterilir. Radar gibi bazı uygulamalarda genliğin ve frekansın sabit olduğu sürekli dalga (CW) kullanılır. Bazı uygulamalarda da genlik veya frekansın modüle edildiği dalgalar data transferi için kullanılır. Dalganın bir  $1/f$  sürede yani bir

periyotta aldığı yola dalga boyu denir. Radyo frekanslı dalgaların (RF) dalga boyları 1MHz frekansı için 300m olurken mm seviyelerine kadar değişebilmektedir. (Görünür ışığın dalga boyu  $4.10^{-4}$  mm seviyesindedir.) Elektromanyetik dalgalar düşük frekanslı olduğunda elektromanyetik alan ismiyle ve yüksek frekanslı ise elektromanyetik radyasyon ismiyle tanımlanır. Radyo frekanslı (RF) alanlar; Elektromanyetik spektrumun genellikle 100kHz - 60GHz bandını kapsayan alanlardır. İletişim, baz istasyonları ve radio-TV antenlerinden yayılan dalgalar bu band içindedir. Çok çok düşük frekanslı (ELF) alanlar elektromanyetik spektrumun genellikle 20Hz - 2kHz bandını kapsar. Ev içi elektriksel cihazlardan, yüksek gerilim hatları, trafolar, tomografi cihazları gibi bazı hastane cihazlarından yayılan alanlar bu band sınıfına girer. Yüksek frekanslı (RF) elektromanyetik alanlar (EM) kaynaklarından olan mesafeye göre reaktif yakın EM alan, ışılan yakın EM alan ve uzak alan olmak üzere üç ayrı durumda incelenir.

### **2.2.1. Reaktif yakın alan**

Anteni çevreleyen en yakın bölgedir. Bu bölgede bulunmak tehlikelidir. Özel koruyucu giysilerle bu bölgede çalışma yapılır. Kaynağa  $2D^2/\lambda$  (D anten boyu) ye kadar olan mesafeyi ifade eder.

### **2.2.2. Işıyan yakın alan**

Reaktif yakın alan ile uzak alan arasında kalan bölgedir. Bu bölgede radyasyon alanlarının ayrı ayrı etkili olduğu ve açılal dağılımlarının antenden olan mesafeye bağlı bölgedir.

### **2.2.3. Uzak alan**

Yayılan Elektromanyetik dalganın düzlem dalga özelliği gösterdiği ve D anten boyutunu göstermek üzere antenden  $2D^2/\lambda$  dan daha uzak olduğu mesafedir. Uzak alanda elektrik alan ve manyetik alanlar ortamın empedansı ile birbirlerine bağlanır (Blake, 2009).

#### 2.2.4. Elektromanyetik güç yoğunluğu

Yayılan elektromanyetik radyasyonun birim yüzeyden geçirdiği güç akısı elektromanyetik güç yoğunluğu olarak tanımlanır. Ve birimi Watt/m<sup>2</sup> (W/m<sup>2</sup>) ile gösterilir. Genellikle alt birimleri olan (mW/m<sup>2</sup>) veya (μW/m<sup>2</sup>) kullanılır. Kaynağın uzak alanında EM dalga, düzlem dalga olarak kabul edilir. Bu durumda E elektrik alanı ve H manyetik alanları birbirine dik olup dalganın yayıldığı doğrultuya dik düzlemde olup, E ile H alanları birbirlerine ortam empedansı (Z<sub>0</sub>=120Ω=377 Ω) ile bağlanırlar. Uzak alanda düzlem dalga bileşenleri arasında (Denklem 2.1).

$$S = \frac{E^2}{377} \text{ (W/m}^2\text{)} \text{ veya } S = 377 \times H^2 \text{ (W/m}^2\text{)} \quad (2.1)$$

bağıntısı geçerli olup, burada S (W/m<sup>2</sup>) güç yoğunluğu, E (V/m) elektrik alan şiddeti, H (A/m) manyetik alan şiddetidir.

Elektrik ve manyetik alanların özellikleri birbirlerinden farklı olup, canlı üzerindeki etkileri de farklıdır. Manyetik alanların etki alanlarının daha geniş olduğu varsayılmaktadır. Bu varsayımın nedenlerinden birisi de elektrik alanlarının duvarlardan geçememesi, insan derisinden geçerken şiddetini kaybetmesi, manyetik alanların ise özel birtakım maddeler dışında sınır tanımamasıdır. Ortak özellikleri ise kaynaktan uzaklaştıkça azalmalarıdır. Elektrik alanlar, elektrik yüklerinin birbirini etkilemesi ile oluşmaktadır. Manyetik alanlar ise elektrik akımının akışı oluşturur. Bir lamba fişe sokulduğunda şebeke frekanslı bir elektriksel alan meydana gelir, lamba yakıldığında ise kordonda akan elektriğe bağlı olarak bir manyetik alan oluşur. Manyetik alan birimi Tesla (T) ya da Gauss' dur (G). Aralarındaki dönüşüm bağıntısı 1 T = 10000 G şeklindedir.

### 2.3. Elektromanyetik Alanların İnsan Vücudu İle Etkileşimi

Elektromanyetik alanların dokular içindeki iyonlara olan etkileri neticesinde onların hareketlerini arttırmaları neticesinde şiddetlerine bağlı olarak bir ısı enerjisi de ortaya çıkar. Bunun sonucunda da dokular içerisinde sıcaklık artışı görülür. İnsan vücudunda herhangi bir dokunun kendi iç sıcaklığının 0.5C den daha fazla artması o dokunun tolere edemeyeceği bir değer olarak alınmıştır. Bu değeri temel alarak geliştirilen bir sınır değer tüm vücut ortalama özgül soğurma değeri olarak kabul edilmiştir. 4W/kg olarak verilen bu limitin 10 kat düşük değeri (0.4W/kg) ihtiyat ilkesi ışığında Dünya Sağlık Örgütü, Elektrik-Elektronik Mühendisleri Enstitüsü (IEEE), Milletlerarası İyone Olmayan Radyasyondan Korunma Komitesi (ICNIRP) tarafından insan vücudunun RF ve mikrodalga etkilerinin hissedilmeyeceği sınır termal etkilerin başladığı değer olarak şimdiye dek kabul edilmiştir. İnsan vücudu bünyesindeki serbest yükler, çevredeki elektrik alandan etkilenir. Manyetik Alan ise bir iletken olarak düşünebileceğimiz insan bedeninde zayıf akımlar oluşturur. Elektromanyetik dalga vücut yüzeyine çarptığında bir kısmı yansır, bir kısmı vücut içine girerek soğurulur. Elektromanyetik radyasyonun canlı vücuduna soğurulması, doku ısınmasına (ısıl etki) sebep olur. Isıl etkinin biyolojik sonuçları ile ilgili ulusal ve uluslar arası standartlar belirlenmiştir. İnsanda cep telefonu veya baz istasyonu gibi cihazlardan yayılan alanların insan vücudunda oluşturacakları enerji birikiminin vücut sıcaklığını 1<sup>0</sup>C arttıracak miktarı organizmadaki mekanizmaları harekete geçirmekte ve bu nedenle zararlı olarak kabul edilmektedir. Bu değer birim canlı kütlede saniyede biriken enerji olan özgül soğurma oranı (Specific Absorbtion Rate-SAR) olarak belirlenmiştir. SAR değerini oluşturan ölçülebilir alan olarak elektrik alan birimi Volt/m kullanılır. Elektrik alanın üst sınır değeri baz istasyonlarında 900 MHz frekanslı vericiler için 42 Volt/m, 1800 MHz frekanslı vericiler için 57 Volt/m ve 2100 frekanslı vericiler için ise 61 V/m dir. Cep telefonu ile konuşurken cihazın kulağa yakın tutulması nedeniyle bu elektrik alan şiddetlerine ve bazen daha yüksek değerlere ulaşılabilir. Isıl etkiler yanında, ısıl olmayan etkiler (*non-thermal effects*) olduğuna ve hücrelerin, dokuların olumsuz etkilenebileceğini savunan çalışmalar da görülmeye başlanmıştır (Furse, 2009).

## 2.4. Yasal Düzenlemeler ve Sınır Değerler

Günümüzde elektromanyetik alanların yaşama getirmiş olduğu kolaylıklarla birlikte toplumun her kesimi az veya çok dozda elektromanyetik dalga maruziyeti altındadır. Kaçınılmaz maruziyetler elektromanyetik doz limitlerinin doğru belirlenmesi halk sağlığının korunmasında çok önemli bir kontrol yöntemidir. Ayrıca bu konuda izlenecek olan politikalara da yön verir.

Elektromanyetik radyasyon konusunda her ülke kendi standartlarına göre limit değerler belirlemiştir. Avrupa Birliği'ne üye ülkeler ve ABD dâhil olmak üzere birçok Dünya ülkesinde ortak olarak kabul gören ve uygulanan sınır değerler bulunmaktadır.

Bu sınır değerler Dünya Sağlık Örgütü (WHO) tarafından da tanınan ve uluslararası bir komisyon olan ICNIRP (International Commission on Non-Ionizing Radiation Protection-İyonize olmayan radyasyondan koruma komisyonu) tarafından genel halk için günde 24 saat maruz kalındığı kabulüyle belirlenmiştir. Sınır değerler yayılan radyasyonun frekansına bağlı olarak değişmekte olup, her frekans için farklıdır.

Ülkemizde Bilgi Teknolojileri ve İletişim Kurumu tarafından 21.04.2011 tarih ve 279 sayılı Resmi Gazetede yayınlanan, “Elektronik Haberleşme Cihazlarından kaynaklanan elektromanyetik alan şiddetinin uluslararası standartlara göre maruziyet limit değerlerinin belirlenmesi, kontrolü ve denetimi hakkında Yönetmelik” çerçevesinde uygulanacak olan mevzuat belirlenmiştir. Diğer bir yönetmelikte ELF frekanslı EMR ile ilgili düzenlemeler olup 24.07.2010 tarihli resmi gazetede yayınlanan Çevre ve Orman Bakanlığının yönetmeliği ile yayınlanmıştır. Bu iki yönetmelikte ICNIRP standartlardan faydalanılarak hazırlanmıştır.

Bilgi Teknolojileri ve İletişim Kurumu yönetmeliğinde ortamın toplamı için olan sınır değerler ile tek bir cihaz için olan sınır değerler birbirinden ayrılmıştır ve ortamın toplamı için olan sınır değerinin %25'i alınarak tek bir cihaz için olan sınır değeri belirlenmiştir. Buna göre baz istasyonlarının çalıştığı frekanslar için ülkemizde geçerli sınır değerler Tablo 2.1'deki gibidir

Tablo 2.1. Türkiye’de geçerli elektromanyetik radyasyon sınır değerleri

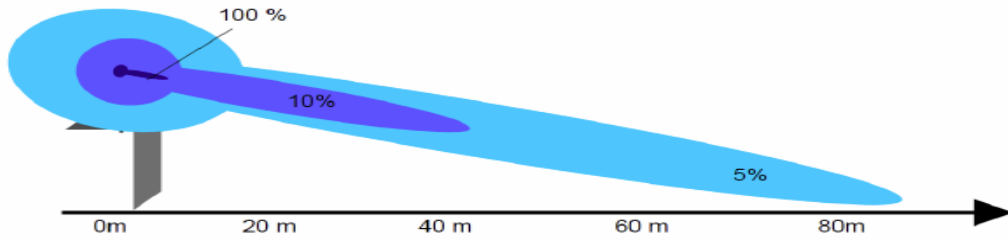
GSM Operatörü	Frekans Bandı	Elektrik Alan Şiddeti (V/m)		Manyetik Alan Şiddeti (A/m)	
		Tek bir cihaz için	Ortamın toplamı için	Tek bir cihaz için	Ortamın toplamı için
VODAFONE	900 MHz	10,23	41,25	0,027	0,111
TURKCELL	900 MHz	10,23	41,25	0,027	0,111
AVEA	1800MHz	14,47	58,34	0,038	0,157
3G (Her Üç Operatör)	2100 MHz	15	61	0,04	0,16

Böylece, çalışma frekansı 900 MHz olan sabit bir telekomünikasyon cihazının tek başına yaymakta olduğu elektrik alan şiddetinin 10,23 V/m, manyetik alan şiddetinin ise 0,027 A/m değerini aşmaması gerekmektedir. 1800MHz için ise bu değerler sırasıyla 14,47 V/m ve 0,038 A/m’dir. Ortamın toplamı için ise, diğer ICNIRP tarafından belirlenen sınır değerler ülkemizde de en üst seviyede geçerlidir. Ancak ülkemizde 42 birim olan bu limitler İtalya’da 6 V/m ve İsviçre’de ise 5 V/m olarak uygulanmaktadır.

ABD ve bazı Avrupa ülkeleri ICNIRP ’ın oluşturduğu sınır değerleri uygularken, İsviçre, İtalya gibi bazı Avrupa ülkeleri ise sınır değerler olarak ICNIRP güvenlik limitlerinin 1/10’nunu (onda biri) uygulamaktadır. Örneğin İsviçre’de baz istasyonu kaynaklı EM Radyasyonun olası olumsuz etkisi için öncelikli olarak İhtiyat İlkesi’ni benimsemesi dolayısıyla ev, ofis, hastane, okul ve çocuk oyun alanları gibi hassasiyet gösteren sürekli yaşam yerlerinde 5 (V/m)’yi sınır değer olarak kabul etmiş olup bunu uygulamaktadır. ICNIRP sınır değerleri elektromanyetik radyasyonu insan üzerinde oluşturacağı biyolojik ve termal etki eşiği nedeni ile oluşturulmasına karşılık İsviçre’de uygulanan ICNIRP’in önerdiği sınır değerlerin 1/10’unu oluşturan seviye bilimsel bulgulara dayanmayan fakat insanları psikolojik olarak rahatlatan ve İhtiyat İlkesi’ne dayanarak insanların yaşam kalitesini bozulmasına engel olmak amacıyla kullanılmaktadır.

Tablo 2.2. Baz istasyonları için İtalya, İsviçre ve Türkiye’de uygulanan limitlerin karşılaştırılması

ÜLKE İSMİ	Elektrik Alan Sınır Değeri (V/m)
İtalya (Duyarlı Mekânlar)	6
İsviçre (Duyarlı Mekânlar için ortamın toplamı)	5
Türkiye (Ortamın Toplamı )	41.25



Şekil 2.2. 10m yüksekliğinde bir binaya yerleştirilmiş baz istasyonu anteninin yönlü yaptığı ana ışın lobu

Baz istasyonlarından yayılan EM radyasyon maruziyeti sonucu ortaya çıkan termal etkiden insanları korumak amacı ile baz istasyonu anteni ile en yakın çevrede yaşayanlar arasında güvenlik limitleri denilen kurallar zinciri ile bir bariyer konulmaya çalışılır. Bunun için güvenlik mesafesi tanımı vardır. Bu mesafe; baz anteninin yakınına girilmemesi gereken bir yaşam alanını ifade eder.

Ülkemizde şehir içlerinde bina ve evlere konulan baz istasyonlarında genellikle bu güvenlik mesafesi antenin tam bakış yönünde 8.5 m -13 m arasındadır. Şekil 9’da görüldüğü gibi bu mesafeler antenin hemen ucundaki EM radyasyonun yaklaşık %25’ine karşı gelmektedir. Güvenlik mesafesi hesabı baz istasyonunun gücüne ve anten özelliğine bağlı olarak belirlenir Bu, genellikle 10 metre, 12-13 metre civarında oluyor. Aydınlatma direklerine yerleştirilen istasyonlarda ise 5-6 metredir.



Şekil 2.3. Radyo Link Antenleri

Elektromanyetik radyasyonun artık göz ardı edilemeyecek diğer bir etkisi de biyolojik etkilerdir. AB ülkelerinde 12 araştırma merkezinin yaptığı REFLEX projesine göre cep telefonları frekanslarındaki elektromanyetik radyasyonun DNA da kırılmalar ile genetik zararlar yaptığı belirlenmiştir. ABD’de yayınlanan uluslararası saygın 14 bilim adamının ortak Bioinitiative Raporunda standartların yetersiz olduğu, kanser yaptığıının delilleri, bağışıklık sistemine ve benzeri 21 konuda en son bilgiler verilmiştir. (Seker,2010)

Tablo 2.3. Zamanla değişen alanlara maruziyet için ICNIRP Referans seviyeleri

Frekans Aralığı	Elektrik Alan Şiddeti E (V/m)	Güç Yoğunluğu S ( $W/m^2$ )
10-400 MHz	28.00	2
400-2000 MHz	$1.375 f^{1/2}$	$f/200$
2-300 GHz	61.00	10



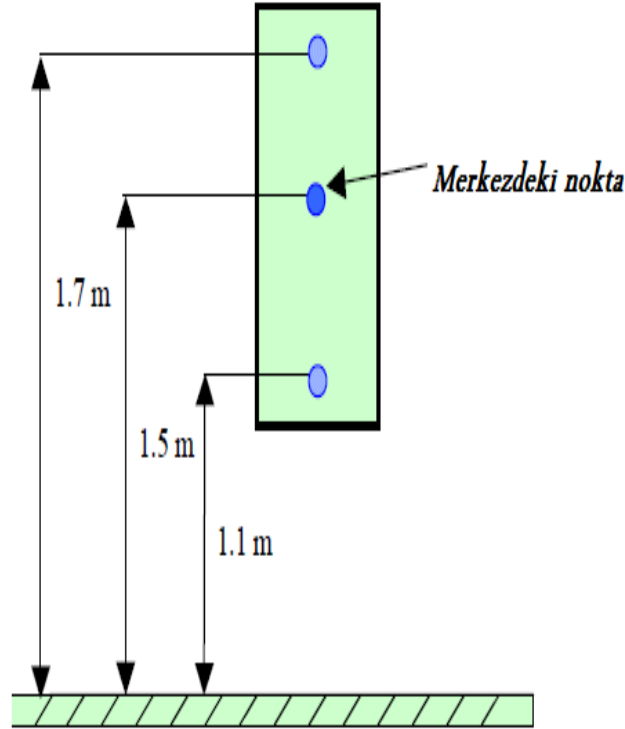
## 2.5. Elektromanyetik Alan İnceleme ve Ölçüm Yöntemleri

Ölçüm çalışmalarında elektromanyetik radyasyonu (EMR) oluşturan elektrik alan ve manyetik alan bileşenleri ayrı ayrı ölçülebilir. Elektromanyetik radyasyon (veya elektromanyetik kirlilik : EMK) ifadeleri kullanıldığında bu bileşenlerin her ikisi birlikte kastedilmektedir. Yüksek frekanslarda dalga kaynaktan uzaklara doğru ilerlerken elektrik ve manyetik alan birbirine bağlı olarak hareket eder ve ortamda bir tanesi (özellikle elektrik alan) ölçülür ve diğeri de ilgili bağıntı yardımıyla belirlenir.

Elektromanyetik kirlilik düzeyi belirlenecek olan mekanın önce çevresinde gözlem yapılır. Bu ön inceleme çalışması sonunda bina içindeki ve bina dışındaki elektromanyetik alan yayıcı kaynaklar varsa özellikleri tespit edilir. Kroki üzerine işlenir. Kaynakların özellikleri (frekans bandı) dikkate alınarak ölçümde hangi cihazların kullanılacağı ve ölçümlerin uzak alanda mı yoksa yakın alanda mı yapılacağı tespit edilir. Daha sonra da yapılacak olan ölçümler için planlama yapılır. Ölçüm yapılacak olan mekânlar bölümlere ayrılarak plana işlenir. Ölçüm yapılacak kritik noktalar tespit edilir. Çalışma için fiziksel ve lojistik şartların sağlanması ve ölçüme ne zaman başlanacağı ve süresi belirlenir. İnceleme yapılacak olan her bir mekânda maksimum sinyal seviyesinin olduğu noktalar tespit edilerek o bölge içerisinde bulunabilecek kişinin maksimum maruziyet seviyesinin belirlenmesi için ölçümler bu noktalara göre yapılır. Ölçümlerde spektrum analizörü kullanılacaksa ekranda max.hold ayarı yapılır ve ortamın sinyal seviyesine duyarlı olan uygun antenler kullanılır. Ölçümler tripod kullanılarak yerden 1.5m yüksekte yapılır. Kontrol ölçümleri ise sırasıyla 1.1 m, 1.5m ve 1.7 m olmak üzere üç ayrı yükseklikte yaparak ortalaması alınır. Ölçüm konfigürasyonu Şekil 2.4’de gösterilmiştir.

Baz istasyonlarının ölçümü elektrik alana duyarlı prob antenlerle ölçülür ve manyetik alan karşılığı bu ölçüm değerinden faydalanılarak hesaplanır. Baz istasyonunun çok yakınında yapılacak ölçümlerde manyetik alan probu kullanılır. Belirli bir uzaklıktan ( $\geq D^2/2\lambda$ ) sonra elektrik alan probu ile ölçüm yapılır. Ölçüm yapılan mekânda birden çok baz istasyonu varsa her birinin etkisini anlamak için spektrum analizörü denilen frekansa bağlı cihazla ölçüm yapılır. Belirli bir sürede örneğin 6 dakikalık

ölçüm sürelerinin yanı sıra günlük, haftalık ölçümler yapılarak ölçüm değerlerinin günün saatlerine göre gösterdiği değişim izlenmelidir. Ölçüm yapılan noktalarda genel olarak geniş bant izotropik elektromanyetik alan ölçer kullanılmıştır. Ölçüm değerleri yüksek olduğunda spektrum analizör kullanılarak farklı frekanslar için selektif ölçümler yapılmıştır.



Şekil 2.4. Ölçüm konfigürasyonu

Ölçüm yapılan mekândaki geniş bant sinyal seviyesi, ölçüm cihazının algılayamayacağı kadar çok düşük ise spektrum analizörü ile bant aralıkları ayrı ayrı kontrol edilir. Ölçüm değerlerinden bazıları teorik hesaplamalarla kontrol edilir. Varsa benzetim programı kullanılır. Sonuçlar karşılaştırılır. Her bir konum için yapılan ölçüm için en az bir adet fotoğraf çekimi yapılır.

Ölçüm değerleri tablo ve grafiklerle aşağıda verilen plan dâhilinde yorumlanır. Şekil 2.4'de görülen her bir üç noktada ölçülen değerlerin ortalaması ile ortamdaki ortalama elektrik ve manyetik alan şiddetleri hesaplanabilir. (Denklem 2.2).

$$E_{spatial\_ortalama} = \sqrt{\frac{\sum_{i=1}^3 E_i^2}{3}} \quad \text{ve} \quad H_{spatial\_ortalama} = \sqrt{\frac{\sum_{i=1}^3 H_i^2}{3}} \quad (2.2)$$

Belirlenen iç ve dış kaynaklara bağlı olarak, ev ortamından yapılacak ölçümler, çok düşük frekanslar bölgesi ile mikrodalga frekanslar bölgesini kapsayacak şekilde yapılmalıdır. Temel olarak ölçümler, elektromanyetik alan şiddeti veya güç yoğunluğu ölçümlerini içerir. Ölçümler geniş bantlı cihaz ile yapıldığında, yakın alanda elektrik alan ve manyetik alan karmaşık bir yapı gösterir. Bu nedenle, elektrik alan ve/veya manyetik alan problemleri ayrı ayrı kullanılarak ölçüm yapılır. Ölçümler, ölçümü yapılacak kaynağın çalışır durumu için yapılır, yani mesken içerisindeki kaynakların toplam etkisi için bu kaynakların çalışır durum olmaları gerekir.

## 2.6. Ölçüm Cihazları

Ölçümlerde geniş bant için elektromanyetik radyasyon ölçüm cihazları izotrop problemleri ile birlikte kullanılır. Bu cihazlar ortamdaki alanın anlık, etkin, ve maksimum değerlerini göstermeli ve zamana göre ortalama alabilme yeteneğine sahip olmalıdırlar. Elektromanyetik alan ölçüm cihazı; ortalama güç yoğunluğunu ( $\text{W/m}^2$ ,  $\text{mW/cm}^2$ ), ortalama elektrik alanını ( $\text{V/m}$ ) veya elektrik alanının karesel ortalamasını, ortalama manyetik alanını ( $\text{A/m}$ ,  $\mu\text{T}$ ) veya manyetik alanının karesel ortalamasını ölçebilecek özellikte olup, bu parametrelerden bir veya birkaçını gösterebilmelidir. Frekansa bağlı ölçümlerde ise spektrum analizörü ya da alıcılar uygun antenlerle birlikte kullanılır.

### 2.6.1. Ölçüm prosedürü

Ölçümlerde kullanılan problemler ortamdaki sinyal seviyesini kapsayacak bant genişliğini sağlamalıdır. Evlerde ve çevredeki kaynak özelliklerine bağlı olarak, gerekli durumlarda birden fazla prob kullanılır. Ölçüm süresi 6 dakika olup, gerektiğinde farklı zaman aralıklarında tekrarlanabilir. Daha uzun süreli ölçümler saha monitörleri kullanılarak yapılır. Ölçüm cihazları tripoda yerleştirilerek ortamdaki alanı bozmadan uzaktan izleme ile yapılır.

### 2.6.2. Belirsizlik kestirimi

Ölçümler sonunda aşağıdaki hususlar dikkate alınarak ölçüm belirsizliği bulunur. Bu işlem için standart belirsizlik  $u(x_i)$  ve duyarlılık katsayısı  $c_i$ , her bir  $x_i$  ölçüm değeri üzerinden tahmin edilir ve standart birleşik ölçüm belirsizliği olan  $u_c(y)$  fonksiyonu aşağıdaki bağıntıdan hesaplanır.(Denklem 2.3).

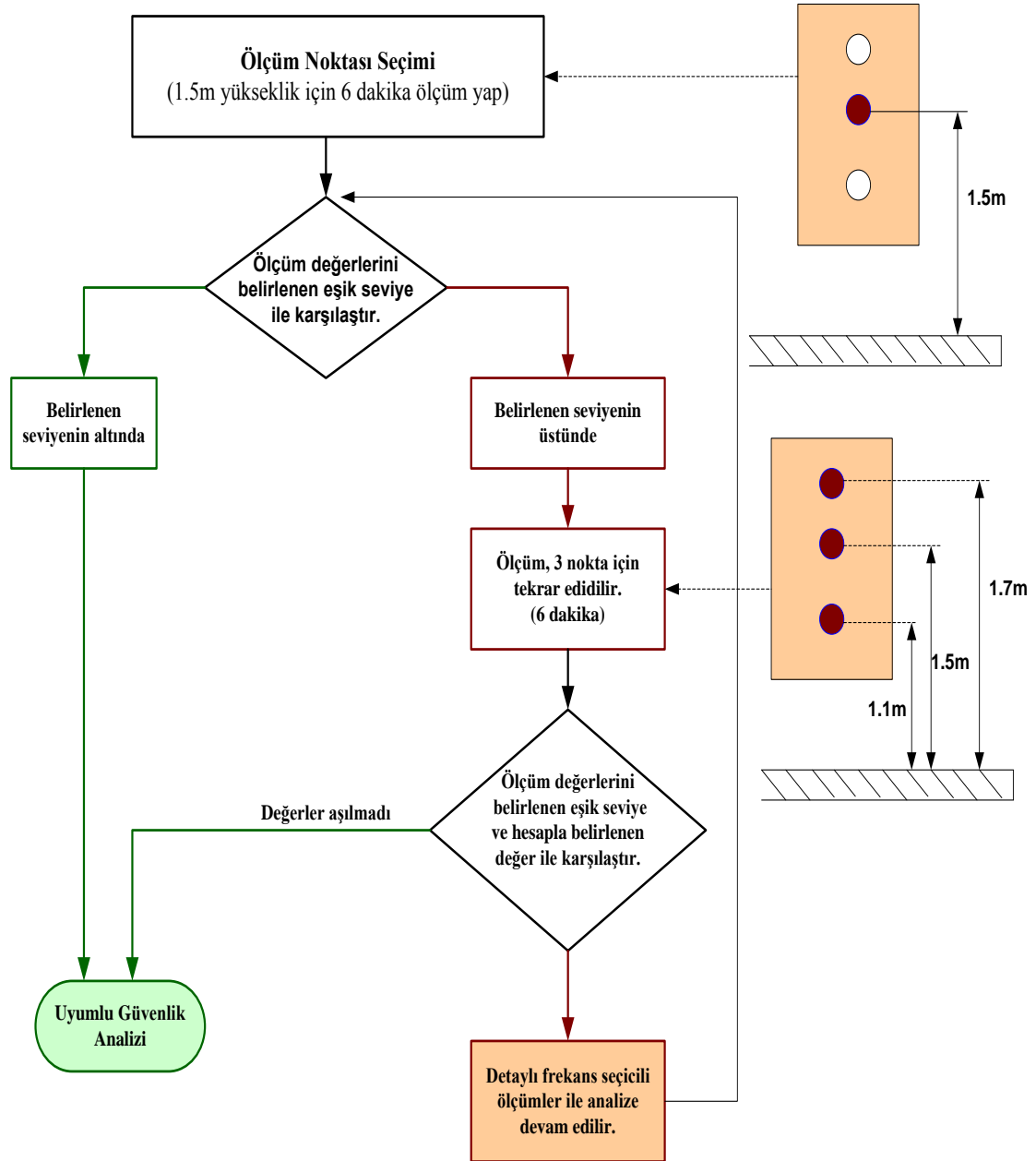
$$u_c(y) = \sqrt{\sum_{i=1}^n (c_i \times u_{(x_i)})^2} \quad (2.3)$$

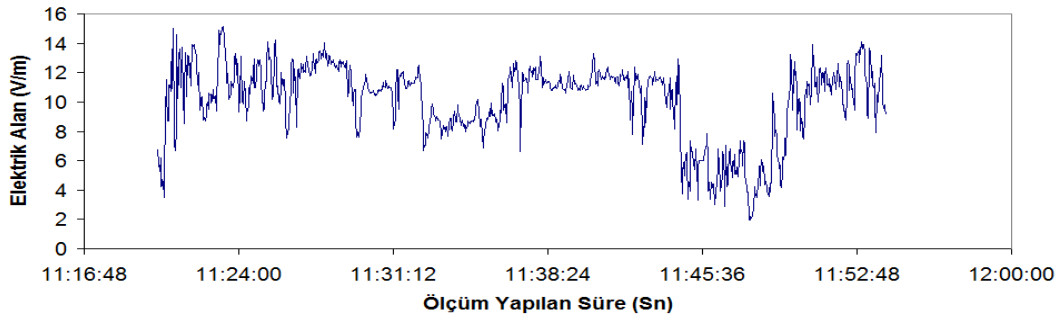
Burada  $u_e$  genişletilmiş ölçüm belirsizliği için  $u_e=1.96u_c$  olarak alınır ve raporda belirtilir. Burada  $x_i$  belirsizliği için olasılık dağılım bölüni giriş değeri isotrop için  $k = \sqrt{3}$  ve  $c_i=1$  alınır.Örneğin elektromanyetik alan ölçerler için problemleri ile birlikte Tablo1'e göre toplam belirsizlik  $u_c(y) = 1.045$  dB ve genişletilmiş belirsizlik  $u_e=1.94$ dB dir.

Tablo 2.4. Standard Belirsizlik Tablosu

Giriş Parametresi	Belirsizlik (dB)	Belirsizlik(sayı)	$u(x_i)$ Standart Belirsizlik(sayı)
İsotrop	1.5	0.19	0.095
Lineer	1.0	0.12	0.06
Düz	1.0	0.12	0.06

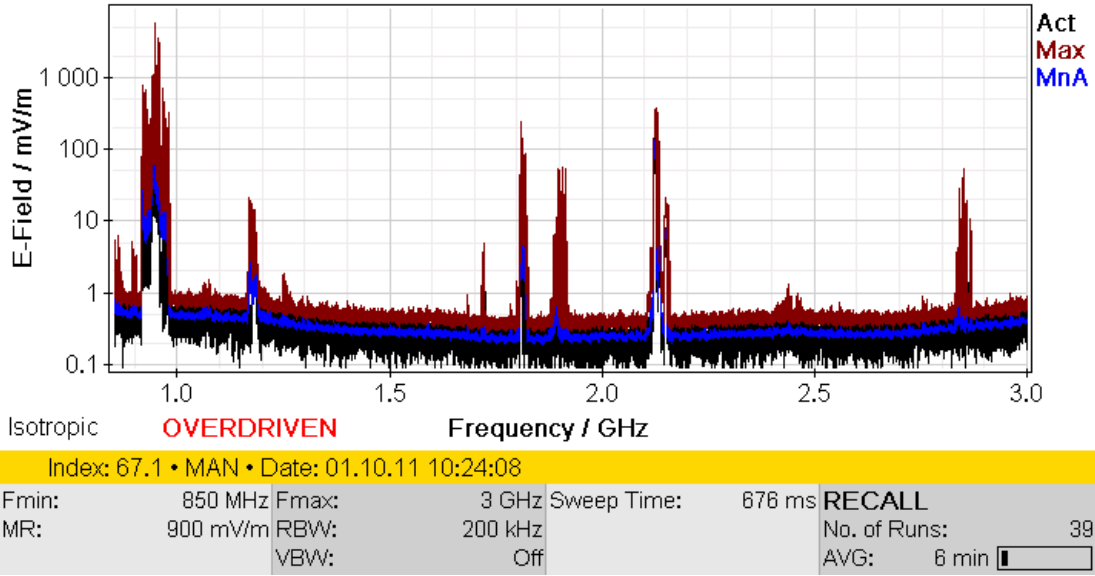
Tablo 2.5. Elektromanyetik Alan Ölçümünde Güvenlik Analizi Akış Şeması





Şekil 2.5. Uzun süreli baz istasyonu EMR ölçümü

Battery: 01.10.11 10:24:08 GPS: 41°1'19.8" N Ant: 3AX 0.4-8G SrvTbl: EU Full Band  
40°31'05.8" E Cable: --- Stnd: ICNIRP GP



Şekil 2.6. Baz istasyonu karşısındaki bir evde Elektromanyetik Radyasyonu Frekanslara göre dağılımı

## BÖLÜM 3. VERİ MADENCİLİĞİ

### 3.1. Veri Madenciliği Tanımı

Veri madenciliği, büyük miktarlardaki verinin içinden geleceği tahmin edilmesinde yardımcı olacak anlamlı ve yararlı bağlantı ve kuralların bilgisayar programlarının aracılığıyla aranması ve analizidir. Ayrıca veri madenciliği, çok büyük miktardaki verilerin içindeki ilişkileri inceleyerek aralarındaki bağlantıyı bulmaya yardımcı olan ve veri tabanı sistemleri içerisinde gizli kalmış bilgilerin çekilmesini sağlayan veri analizi tekniğidir (Kalikov, 2006). Bu işlemlerin uygulama alanı oldukça geniştir. Bu alanlar içerisinde Şekil 3.1’de gösterildiği gibi, Veri Tabanı Sistemleri, Veri Görselliği, Yapay Sinir Ağları, İstatistik, Yapay Öğrenme, vb. gibi disiplinler bulunmaktadır.



Şekil 3.1. Veri Madenciliğinin birlikte bulunduğu disiplinler

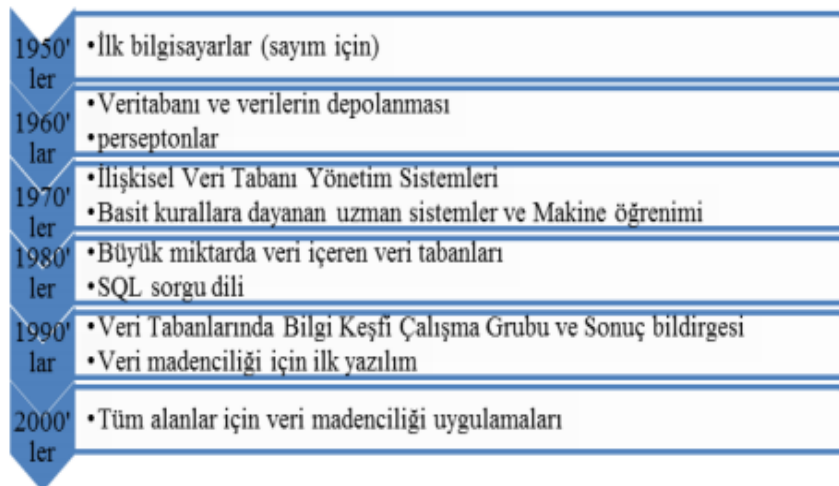
Veri madenciliği araçları kullanılarak, işletmelerin daha etkin kararlar almasına yönelik karar destek sistemlerinde gerekli olan eğilimlerin ve davranış kalıplarının ortaya çıkarılması mümkün olmaktadır. Geçmişteki klasik karar destek sistemlerinin kullanıldığı araçlardan farklı olarak, veri madenciliğinde çok daha kapsamlı ve otomatize edilmiş analizler yapmaya yönelik, birçok farklı özellik bulunmaktadır (İnan, 2003). Veri madenciliğinin işletmelere sunduğu en önemli özellik, veri grupları arasındaki benzer eğilimlerin ve davranış kalıplarının belirlenmesidir. Bu süreç aynı zamanda otomatize edilmiş bir biçimde hayata geçirilebilmektedir. Bu fonksiyon özellikle hedef pazarlara yönelik pazarlama faaliyetlerinde yoğun olarak kullanılmaktadır (İnan, 2003). Başka bir özelliği ise daha önceden bilinmeyen, veri ambarları içerisinde bulunan ancak ilk etapta görülemeyen bilgilerin ortaya çıkarılabilmesidir. Örneğin bir firma sattığı ürünleri analiz ederek, ilerideki kampanyalarını şekillendirebilir ya da sattığı ürünler arasındaki bağları keşfedebilir. Burada amaç daha önceden fark edilmeyen veri kümelerinin bulunabilmesidir. Günümüzün ekonomik koşulları ve yaşanan hızlı değişim ortamlarında, iş deneyimi ve önseziyle dayanarak alınan kararlarda yanlış karar alma riski çok yüksektir. Riski azaltmanın tek yolu bilgiye dayalı yönetimi öngören karar destek çözümleridir. Veri madenciliği teknikleri gerçek anlamda bir karar destek sistemi oluşturmadıkça olmaz araçlardır. Bu noktada bilgi teknolojilerinden yararlanmak kaçınılmaz olmuştur.

### **3.2. Veri Madenciliğinin Tarihi**

Günümüzde neredeyse her eve bilgisayar girmiştir ve internet erişimi hemen hemen her yerden sağlanmaktadır. Disk kapasitelerinin artması, her yerden bilgiye ulaşma olasılığı, bilgisayarların çok büyük miktarlarda veri saklamasına ve daha kısa sürede işlem yapmasına olanak sağlamıştır. Geçmişten günümüze veriler her zaman yorumlanmış, bilgi elde etmek istenmiştir ve bunun için de donanımlar oluşturulmuştur. Bu sayede bilgi, geçmişten günümüze taşınır hale gelmiştir. 1950'li yıllarda ilk bilgisayarlar sayımlar için kullanılmaya başlamıştır. 1960'larda ise veri tabanı ve verilerin depolanması kavramı teknoloji dünyasında yerini almıştır. 1960'ların sonunda bilim adamları basit öğrenmeli bilgisayarlar geliştirebilmişlerdir.



Minsky ve Papert, günümüzde sinir ağı olarak bilinen perseptron'ların sadece çok basit olan kuralları öğrenebileceğini göstermişlerdir (Adriaans ve Zantinge, 1997). 1970'lerde İlişkisel Veri Tabanı Yönetim Sistemleri uygulamaları kullanılmaya başlanmıştır. Bilgisayar uzmanları bununla beraber basit kurallara dayanan uzman sistemler geliştirmişler ve basit anlamda makine öğrenimini sağlamışlardır. 1980'lerde veri tabanı yönetim sistemleri yaygınlaşmış ve bilimsel alanlarda, mühendisliklerde vb. alanlarda uygulanmaya başlanmıştır. Bu yıllarda şirketler, müşterileri, rakipleri ve ürünleri ile ilgili verilerden oluşan veri tabanları oluşturmuşlardır. Bu veri tabanlarının içerisinde çok büyük miktarlarda veri bulunmaktadır ve bunlara SQL veri tabanı sorgulama dili ya da benzeri diller kullanarak ulaşılabilir. 1990'larda artık içindeki veri miktarı katlanarak artan veri tabanlarından, faydalı bilgilerin nasıl bulunabileceği düşünölmeye başlanmıştır. Bunun üzerine çalışmalara ve yayınlara başlanmıştır. 1989, KDD (IJCAI)-89 Veri Tabanlarında Bilgi Keşfi Çalışma Grubu toplantısı ve 1991, KDD (IJCAI)-89'un sonuç bildirgesi sayılabilecek "Knowledge Discovery in Real Databases: A Report on the IJCAI-89 Workshop" makalesinin KDD (Knowledge Discovery and Data Mining) ile ilgili temel tanım ve kavramları ortaya koyması ile süreç daha da hızlanmış ve nihayet 1992 yılında veri madenciliği için ilk yazılım gerçekleştirilmiştir. 2000'li yıllarda veri madenciliği sürekli gelişmiş ve hemen hemen tüm alanlara uygulanmaya başlanmıştır. Alınan sonuçların faydaları göröldükçe, bu alana ilgi artmıştır. Veri madenciliğinin tarihsel gelişim süreci, Şekil 3.2'de gösterilmiştir.



Şekil.3.2. Veri madenciliğinin tarihsel gelişim süreci

### 3.3. Veri Madenciliği Uygulama Alanları

Günümüzde veri madenciliği farklı farklı alanlarda uygulanmaya başlanmıştır. Astronomi, biyoloji, finans, pazarlama, sigorta, tıp, kimya, sosyal bilimler, web madenciliği ve birçok farklı alanlarda uygulanmaktadır.

**Pazarlama:**

Bu alana baktığımız zaman en önemli uygulamaların, farklı müşteri gruplarını belirlemek ve bu müşteri gruplarının davranışlarını tahmin etmeye yönelik olarak analiz çalışmaları yapmak olduğunu görüyoruz. Örneğin American Express veri madenciliği tekniklerini kullanarak %10 - %15 arasında kredi kartı kullanımını arttırmıştır. Müşterilerin satın alma özelliklerinin belirlenmesi, müşterilerin demografik özellikleri, mevcut müşterilerin elde tutulması, pazar sepet analizi (Market Basket Analysis), müşteri ilişkileri yönetimi (CRM Customer Relationship Management), müşteri değerlendirme, şans tahmini, hedef pazar analizi, müşteriler arası benzerliklerin saptanması gibi alanlarda da veri madenciliği teknikleri uygulanmaktadır.

**Biyoloji, Tıp ve Genetik:**

DNA içerisindeki genlerin tespiti, gen haritalarının analizi, kanserli hücrelerin tespiti, genetik hastalıkların tespiti gibi alanlarda veri madenciliği uygulanmaktadır.

**Banka ve Sigorta :**

Finansal tablolar arasında korelasyon tespiti, kredi kartı dolandırıcılarının tespiti, kredi kartı taleplerinin değerlendirilmesi, sigorta dolandırıcılarının tespiti, riskli müşterilerin belirlenmesi, yeni poliçe talep edebilecek müşterilerin tahmini gibi alanlarda veri madenciliği teknikleri uygulanmaktadır. Kredi kartı ödemelerini aksatan, gecikmeli yapan veya hile yapmayan insanların özellikleri incelenerek bu tip sonuçlan verebilecek muhtemel kişilerin belirlenmesinde kullanılmaktadır.

Web Madenciliği (Web Mining) :

İnternet ve web üzerindeki dosyalar hem hacim hem de karmaşıklık olarak hızla artmaktadır. Web madenciliği veri madenciliği tekniklerini kullanarak World Wide Web 'de bulunan dosya ve servislerden otomatik olarak örüntüler bulur ve öngörülemeyen bilgilere ulaşır, böylece veriye ulaşım süresinin azalması amaçlanır.

Yüzey Analizi ve Coğrafi Bilgi Sistemleri :

Bölgelerin coğrafi özelliklerine göre sınıflandırılması, kentlerde yerleşim yerlerinin belirlenmesi, kentlerde su oranının tespiti, otobüs duraklarının yerlerinin belirlenmesi gibi durumlar için veri madenciliği uygulanmaktadır

### **3.4. Veri Madenciliğinde Karşılaşılan Problemler**

Büyük hacimli verilerin bulunduğu veri ortamlarında büyük sorunlar ortaya çıkabilir. Bu nedenle küçük veri kümelerinde, benzetim ortamlarında hazırlanmış veri madenciliği sistemleri, büyük hacimli, eksik, gürültülü, boş, atık, aykırı veya belirsiz veri kümelerinin bulunduğu ortamlarda yanlış çalışabilir. Bu nedenle veri madenciliği sistemleri hazırlanırken bu sorunların çözülmesi gerekmektedir. Veri madenciliği uygulamalarında karşılaşılabilecek sorunlar şunlardır:

**Artık veri:** Artık veri, problemde istenilen sonucu elde etmek için kullanılan örneklem kümesindeki gereksiz niteliklerdir. Bu durum pek çok işlem sırasında karşımıza çıkabilir.

**Belirsizlik:** Yanlılıkların şiddeti ve verideki gürültünün derecesi ile ilgilidir.

**Boş veri:** Bir veri tabanında boş değer, birincil anahtarlar yer almayan herhangi bir niteliğin değeri olabilir. Boş değer, tanımı gereği kendisi de dâhil olmak üzere hiçbir değere eşit olmayan değerdir.

**Dinamik veri:** Kurumsal çevrim içi veri tabanları dinamiktir ve içeriği sürekli olarak değişir. Bu durum, bilgi keşfi metodları için önemli sakıncalar doğurmaktadır.

Eksik veri: Veri kümesinin büyüklüğünden ya da doğasından kaynaklanmaktadır.

Eksik veriler olduğunda yapılması gerekenler şunlardır:

1. Eksik veri içeren kayıt veya kayıtlar çıkarılabilir.
2. Değişkenin ortalaması eksik verilerin yerine kullanılabilir.
3. Var olan verilere dayalı olarak en uygun değer kullanılabilir.

Eksik veriler, yapılacak olan istatistiksel analizlerde önemli problemler yaratmaktadır. Çünkü istatistiksel analizler ve bu analizlerin yapılmasına olanak veren ilgili paket programlar, verilerin tümünün var olduğu durumlar için geliştirilmiştir (Albayrak, 2008).

Farklı tipteki verileri ele alma: Gerçek hayattaki uygulamalar makine öğreniminde olduğu gibi yalnızca sembolik veya kategorik veri türleri değil, fakat aynı zamanda tamsayı, kesirli sayılar, çoklu ortam verisi, coğrafi bilgi içeren veri gibi farklı tipteki veriler üzerinde işlem yapılmasını gerektirir.

Gürültülü ve kayıp değerler: Veri girişi veya veri toplanması esnasında oluşan sistem dışı hatalara gürültü denir. Büyük veri tabanlarında pek çok niteliğin değeri yanlış olabilir. Veri toplanması esnasında oluşan hatalara ölçümden kaynaklanan hatalar da dahil olmaktadır. Bu hataların sonucu olarak birçok niteliğin değeri yanlış olabilir ve bu yanlışlardan dolayı veri madenciliği amacına tam olarak ulaşmayabilir.

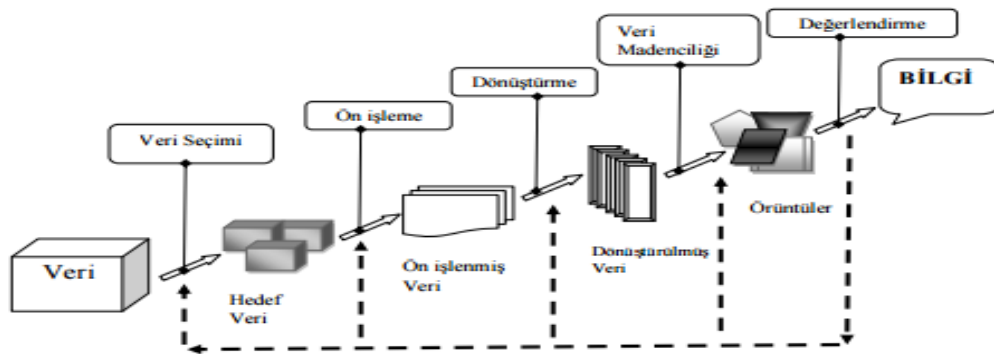
Sınırlı bilgi: Veri tabanları genel olarak basit öğrenme işlerini sağlayan özellik veya nitelikleri sunmak gibi veri madenciliği dışındaki amaçlar için hazırlanmışlardır. Bu yüzden, öğrenme görevini kolaylaştıracak bazı özellikler bulunmayabilir.

Veri tabanı boyutu: Veri tabanı boyutları büyük bir hızla artmaktadır. Veri tabanı algoritması çok sayıda küçük örnekleme ele alabilecek biçimde geliştirilmiştir. Aynı algoritmaların yüzlerce kat büyük örneklemlerde kullanılabilmesi için çok dikkat gerekmektedir. Veri madenciliği, aynı zamanda bir süreçtir. Veri yığınları arasında,

soyut kazılar yaparak veriyi ortaya çıkarmanın yanı sıra, bilgi keşfi sürecinde örüntüleri ayrıştırarak süzmek ve bir sonraki adıma hazır hale getirmek de bu sürecin bir parçasıdır. Bu süreç Şekil 3.3’de gösterilmiştir. Üzerinde inceleme yapılan işin ve verilerin özelliklerinin bilinmemesi durumunda ne kadar etkin olursa olsun hiç bir veri madenciliği algoritmasının fayda sağlaması mümkün değildir. Bu sebeple, veri madenciliği sürecine girilmeden önce, başarının ilk şartı, iş ve veri özelliklerinin detaylı analiz edilmesidir.

Veri madenciliği sürecinde izlenen adımlar genellikle aşağıdaki şekildedir (Shearer, 2000):

1. Problemin tanımlanması,
2. Verilerin hazırlanması,
3. Modelin kurulması ve değerlendirilmesi,
4. Modelin kullanılması,
5. Modelin izlenmesi.



Şekil 3.3. Bilgi keşfi sürecinde veri madenciliği

**Problemin tanımlanması:** Veri madenciliği çalışmalarında başarılı olmanın en önemli şartı, projenin hangi işletme amacı için yapılacağını ve elde edilecek sonuçların başarı düzeylerinin nasıl ölçüleceğinin tanımlanmasıdır.

**Verilerin hazırlanması:** Modelin kurulması aşamasında ortaya çıkacak sorunlar, bu aşamaya sık sık geri dönülmesine ve verilerin yeniden düzenlenmesine neden olacaktır. Bu durum verilerin hazırlanması ve modelin kurulması aşamaları için, bir

analistin veri keşfi sürecinin toplamı içerisinde enerji ve zamanının %50 - %85'ini harcamasına neden olmaktadır (Piramuthu,1998). Verilerin hazırlanması, “toplama”, “değer biçme”, “birleştirme ve temizleme”, “örneklem seçimi” ve “dönüştürme” aşamalarından oluşmaktadır.

**Modelin kurulması ve değerlendirilmesi:** Tanımlanan problem için en uygun modelin bulunabilmesi, olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir.

**Modelin kullanılması:** Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi, bir başka uygulamanın alt parçası olarak kullanılabilir.

**Modelin izlenmesi:** Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve yeniden düzenlenmesini gerektirecektir

### **3.5. Veri Madenciliği Teknikleri**

#### **3.5.1. Sınıflama**

Sınıflandırma veri madenciliğinin en çok kullanıldığı alandır. Var olan veri tabanının bir kısmı eğitim olarak kullanılarak sınıflandırma kuralları oluşturulur. Bu kurallar yardımıyla yeni bir durum ortaya çıktığında nasıl karar verileceği belirlenir. Veri madenciliğinin sınıflandırma grubu içerisinde en sık kullandığı teknik karar ağaçlarıdır. Aynı zamanda lojistik regresyon, diskriminant analizi, sinir ağları ve fuzzy setleri de kullanılmaktadır. İnsanlar verileri daima sınıflandırdıkları, kategorize ettikleri ve derecelendirdikleri için sınıflandırma, hem veri madenciliğinin temeli olarak hem de veri hazırlama aracı olarak da kullanılabilir.

### 3.5.2. Kümeleme analizi

Kümeleme analizi, sınıflama analizinden farklı olarak denetimsizdir. Öngörülecek alanların belirlenmesini ve birbirine benzeyen verilerin altkümelere ayrılmasını hedefler. Kümeleme analizinin hedefi, veri setinde doğal olarak meydana gelen altsınıfları bulmaktır.

Kümeleme, verilerin kendi aralarındaki benzerliklerin göz önüne alınarak gruplandırılması işlemidir ve kümeleme yöntemlerinin çoğu veri arasındaki uzaklıkları kullanır. Hiyerarşik Kümeleme yöntemleri en yakın komşu algoritması ve en uzak komşu algoritmasıdır. Hiyerarşik olmayan kümeleme yöntemleri arasında k-ortalama yöntemini sayılabilir.

Kümeleme teknikleri; uzaklık matrisini kullanarak nesnelere veya değişkenlere kendi içinde homojen ve kendi aralarında heterojen gruplar oluşturmaya imkan sağlamaktadır. Kümeleme analizi için birçok algoritma öne sürülmüştür. Ancak literatürde bu algoritmalar iki başlık altında toplanmıştır: Hiyerarşik kümeleme teknikleri, Hiyerarşik olmayan kümeleme tekniği (Ketchen and Shook, 1996: 444). Her iki teknikte de ortak amaç kümeler arasındaki farklılıkları ve kümeler içi benzerlikleri en yüksek düzeye çıkarmaktır. Yani, küme içi homojenlik artırılırken kümeler arası homojenlik ise azaltılmaktadır.

Aralık ve oransal ölçekteki değişkenleri bulunan birimleri kümelerken, değişik uzaklık ölçüleri kullanılır. Bu ölçülerden bazıları ;

Öklidyen: Öklidyen uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık aşağıdaki formülle hesaplanır.

$$d(i, j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2} \quad (3.1)$$

Manhattan: Manhattan uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık aşağıdaki formülle hesaplanır.

$$d(i, j) = (|x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|) \quad (3.2)$$

Minkowski Uzaklık Ölçüsü: Minkowski uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık aşağıdaki formülle hesaplanır

$$d(i, j) = \left[ |x_{i1} - x_{j1}|^m + |x_{i2} - x_{j2}|^m + \dots + |x_{ip} - x_{jp}|^m \right]^{1/m} \quad (3.3)$$

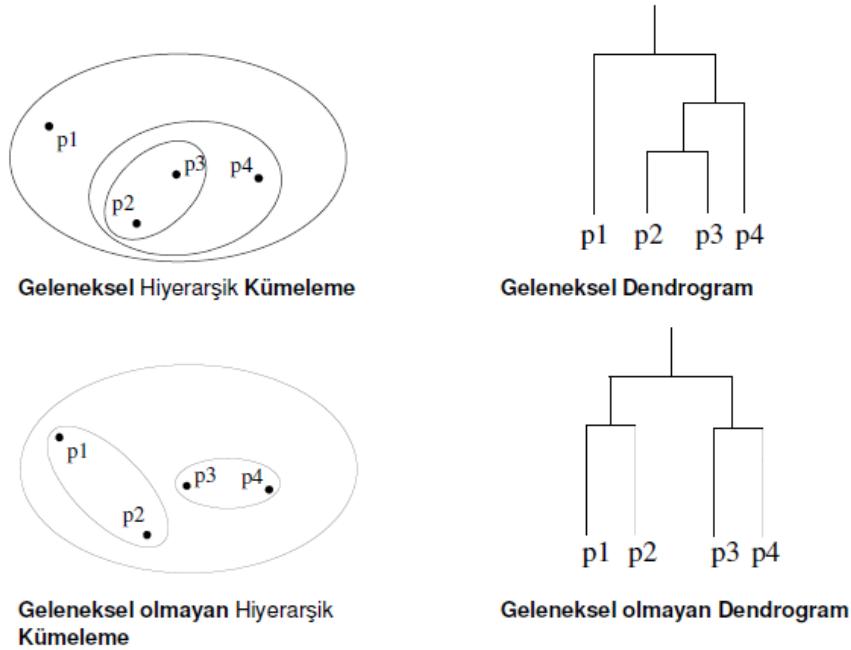
Pearson Uzaklık Ölçüsü: Pearson uzaklık ölçüsü kullanılarak iki birim arasındaki uzaklık aşağıdaki formülle hesaplanır. Bu formülde kullanılan  $S_p$ , uzaklığın hesaplandığı değişkene ait varyanstır.

$$d(i, j) = \sqrt{\frac{(x_{i1} - x_{j1})^2}{S_1^2} + \frac{(x_{i2} - x_{j2})^2}{S_2^2} + \dots + \frac{(x_{ip} - x_{jp})^2}{S_p^2}} \quad (3.4)$$

Kümeleme yöntemleri hiyerarşik ve hiyerarşik olmayan kümeleme olarak iki gruba ayrılır. Hiyerarşik kümeleme yöntemi kendi içinde birleştirici (agglomerative) ve ayrıştırıcı (divisive) kümeleme yöntemleri olarak ikiye ayrılmaktadır. Birleştirici yöntemlerde, tüm gözlemler başlangıç düzeyinde ayrı birer küme olarak ele alınmakta ve daha sonra uzaklık veya benzerlik ölçülerine göre en yakın veya en çok benzeyen gözlemler birleşerek bir küme oluşturmaktadır. Bu doğrultudaki her aşamada gözlem sayısı bir azalarak tüm gözlemler tek bir kümede birleşinceye kadar kümeleme işlemi devam etmektedir. Bu çerçevede ve farklı uzaklık veya benzerlik ölçüleri ile farklı küme bağlantı tekniklerine göre bir araya gelen gözlemlerin oluşturduğu kümeler, ağaç diyagramları (dendogram) veya buz saçaklarına benzeyen (icicle plot) grafikler ile gözlemlenebilmektedir. Ayrıştırıcı kümeleme yönteminde ise, birleştirici yöntemdeki süreç tersine işleyerek başlangıçta tüm gözlemlerin oluşturduğu tek bir küme, benzemezlik veya uzaklık ölçüleri temelinde her gözlem tek bir kümeyi temsil edene kadar devam etmektedir. Hiyerarşik kümeleme yöntemlerinde küme sayısının ne olacağına ilişkin önsel bir bilgiye ihtiyaç duyulmazken, hiyerarşik olmayan kümeleme yöntemlerinde ise, oluşturulacak küme



sayısının önceden bilinmesi gerekmektedir. Hiyerarşik olmayan kümeleme yöntemlerindeki bu kısıtlamaya karşın, her gözlem için uzaklık veya benzerlik matrislerinin hesaplanmasına gerek duyulmadığı için daha büyük ( $n > 300-400$  gibi) veri setlerine uygulanabilmesindeki kolaylık ve aykırı (outlier) değerlere daha az duyarlı olmaları bu yöntemleri avantajlı kılmaktadır (Alpar, 2011).



Şekil 3.4. Hiyerarşik ve Hiyerarşik Olmayan Kümeleme

Küme sayısı konusunda ön bilgi var ise ya da araştırmacı anlamlı olacak küme sayısına karar vermiş ise bu durumda hiyerarşik olmayan kümeleme yöntemi kullanılabilir. Bu kümeleme yönteminde birimlerin kümelere parçalanması rastgele yapılabilir. Birimlerin ayrılacakları küme sayısı belirlendikten sonra, küme belirleme kriterine göre birimlerin hangi kümelere gireceklerine karar verilir ve atama işlemleri yapılır.

Hiyerarşik olmayan kümeleme algoritmalarından en çok kullanılanı K-Means'tir. K-means'ın atama mekanizması her verinin sadece bir kümeye ait olabilmesine izin verir (Evans, 2005). Bu nedenle, keskin bir kümeleme algoritmasıdır (bölümlemeli kümeleme). K-means algoritmasının genel mantığı  $n$  adet veri nesnesinden oluşan bir veri kümesini, giriş parametresi olarak verilen  $k$  adet kümeye bölümlemektir. Amaç,

gerçekleştirilen bölümlenme işlemi sonunda elde edilen kümelerin, küme içi benzerliklerinin maksimum ve kümeler arası benzerliklerinin minimum olmasını sağlamaktır.

### 3.5.2.1. K - means algoritması

K-means algoritması en iyi bilinen ve yaygın kullanılan bir kümeleme algoritması ve bölümlenme tekniğidir İlk olarak J. MacQueen tarafından 1967 yılında tanıtılmıştır. Bu yöntem yıllardır bilimsel ve endüstriyel uygulamalarda en yoğun kullanılan kümeleme algoritması haline gelmiştir. Algoritmaya k-means adı verilmesinin nedeni, algoritmanın çalışmasından önce sabit bir küme sayısına ihtiyaç duyulmasıdır. Küme sayısı  $k$  ile gösterilir ve elemanlarının birbirlerine olan yakınlıklarına göre oluşacak grup sayısını ifade eder. Buna göre  $k$  önceden bilinen ve kümeleme işlemi bitene kadar değeri değişmeyen sabit bir pozitif tamsayıdır. Kümeleme işlemi, verilerin en yakın veya benzer oldukları küme merkezleri (centroid) etrafına yerleştirilmesi ile gerçekleştirilir. Çalışma yönteminde genellikle, öklit bağıntısı temel alınarak kümeleme yapılmaktadır. Algoritmanın başında  $k$  sayısı giriş parametresi olarak verilir. Eğer küme sayısı belirli değil ise deneme yoluyla en uygun sayı bulunur veya bu değer algoritmaya dışardan verilir.  $K$  adet rastgele küme merkezi belirlenir veya ilk  $k$  eleman merkez olabilir. Elemanların merkezlere yakınlıkları hesaplanarak, yakın oldukları merkezlere göre kümeleme yapılır. Oluşan kümelerin ortalamaları hesaplanarak yeni küme merkezleri belirlenir. Bu işlem kümelenecek eleman kalmayınca kadar sürer. Çok yaygın kullanımı olan bu algoritmanın zayıf yanları da bulunmaktadır: öncelikle algoritmanın başında giriş parametresi olarak bir  $k$  sayısının verilmesine gerek vardır. Elde edilecek sonuçlar  $k$  sayısına göre değişkenlik gösterebilir. Eğer küme sayısı belirli değil ise deneme yoluyla en uygun sayı bulunur. Aşırı gürültü ve istisna veriler algoritmayla hesaplanan ortalamayı değiştirdiği için k-means algoritması gürültü ve istisnaya karşı çok duyarlıdır. Algoritma çakışan kümelerde iyi sonuç vermemektedir ve sadece sayısal veriler ile kullanılabilir.

### 3.5.3. Birliktelik kuralları

Birliktelik kuralları, birbiriyle ilişkili olan deęişkenlerin ortaya ıkarılması ve aralarındaki baęlantının büyüklüğünün tespit edilmesine yöneliktir. Birliktelik kuralları belirli türlerdeki veri yapıları arasındaki ilişkileri tanımlamaya alışan bir yöntemdir.

Baęıntı analizleriyle cinsiyet ile eğitim durumu gibi çeşitli deęişkenler arasında anlamlı ve kuvvetli bir baęıntı kurulabilir. Müşteri yaşı ve gelir seviyesi ile satın alma tutum ve davranışları arasında da bir baęıntı kurulabilir.

Müşteriye sunulacak herhangi bir ürün-hizmet teklifinin müşteri tarafından kabul edilip edilmemesi, eęer kişi hakkındaki bir demografik veri veya onun bir dięer tutum ve davranışıyla ilişkilendirilebilirse pazarlama faaliyetleri bakımından önemli bir bilgi birikimi elde edilmiş olur.

Örneęin; bankadan hizmet alan müşterilerin maaş hesabı sahip olmalarıyla ile özel emeklilik sigortasına sahip olmaları arasında kuvvetli bir ilişki bulunmuşsa, maaş hesabı bulunan dięer müşterilere özel emeklilik sigortası konusunda bir teklif sunulabilir.

Baęıntı analizi esasına dayanan ve veri madencilięi uygulamalarında ok kullanılan yöntemlerden birisi sepet analizidir. Sepet analizi, özellikle işlemsel veriyi ilişkilendirir.

Örneęin; A hizmetinin talep edilmesiyle B hizmetinin veya C hizmetinin talep edilmesi arasında bir baęlantı olup olmadığı, varsa, bu baęlantının kuvvet ve önem derecesi sepet analizleriyle ortaya ıkarılmaya alışılır. Amaç bu analizin sonucunda A hizmeti talebiyle B hizmeti talebi arasında kuvvetli bir baęıntı bulunması durumunda A hizmeti talep eden müşteriye B hizmetini de sunmaktır. Bu şekilde apraz satış ve üst seviye satış imkânı doğmaktadır.

## BÖLÜM 4. UYGULAMA

Bu çalışmada veri madenciliğinin kümeleme tekniği kullanılarak Rize merkezinde yapılmış olan baz istasyonları antenlerinden yayılan elektromanyetik radyasyon ölçüm değerlerinin baz istasyonu yakınlarında bulunan insanlar için maruziyet seviyelerinin analiz ve değerlendirilmesi yapılmıştır. Çalışmada 87 noktada yapılan elektromanyetik alan ölçüm değerleri , yaşam alanlarına uzaklıklarına göre GSM frekans bandlarına göre ve sürekli yaşam alanları (indoor ortamlar) ile dış ortamlar(outdoor) için 3 ayrı kümeye göre risk ayrışımı yapılmıştır.

Çalışma gerçekleştirilirken WEKA Yazılımının Hiyeraşik olmayan kümeleme algoritmalarından K-Means algoritması kullanıldı.Ve Euclidean Distance, Chebyshev Distance, Edit Distance, Manhattan Distance uzaklık fonksiyonlarından Manhattan seçilerek 3 ayrı kümeleme yapıldı.

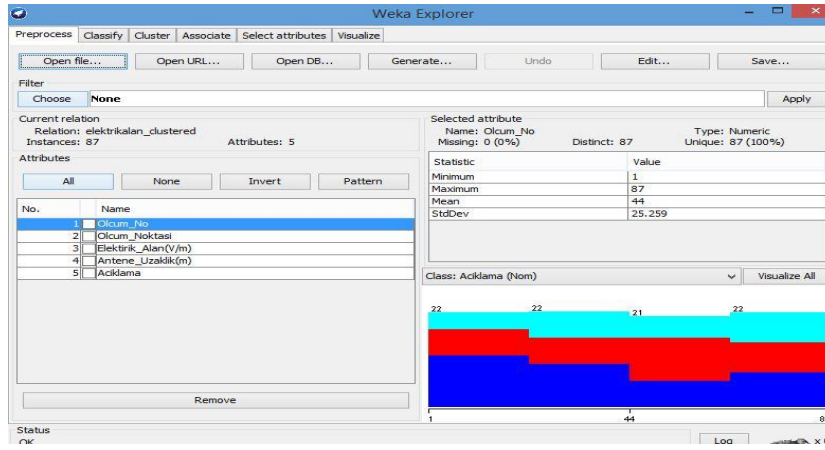
Veriler hazırlandıktan sonra gerekli düzenlemeler Tablo.4.1’de görüldüğü gibi hazırlandı ve csv uzantılı formata dönüştürülerek yazılımda kullanılabilir hale getirildi. Hiyeraşik olmayan k means algoritması kullanılarak elde edilen sonuçlar Şekil 4.2, Şekil 4.3 ve Şekil 4.4 ‘de gösterilmiştir.

Tablo 4.1 Veri tabanının özellikleri

VERİ ALANININ ADI	ÖZELLİKLERİ
Ölçüm No	Sayısal
Ölçüm Noktası	Metin(indoor,outdoor)
Elektrik Alan	Sayısal
Antene Uzaklık	Sayısal
Açıklama	Metin(ND,KED,RD)

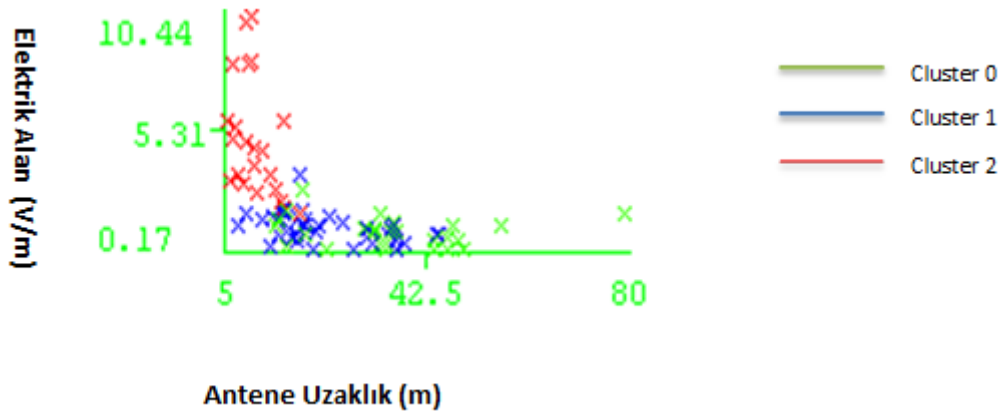
Tablo 4.2 Elektrik Alan Şiddet Seviyelerine Göre Doz Dağılımlarına Göre Risk Kriterleri

Elektromanyetik Dozimetre	Elektrik Alan Şiddeti (V/m)
ND	$0 < E \leq 0,75$
KED	$0,75 < E \leq 2$
RD	$E > 2$



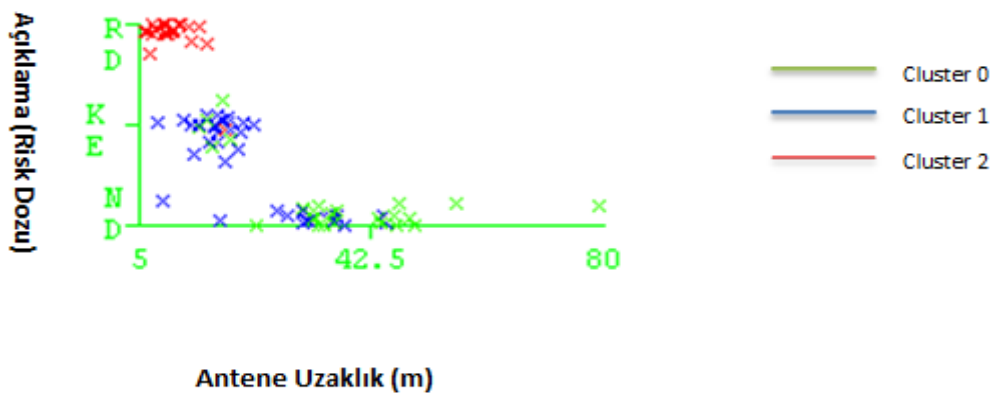
Şekil 4.1. Weka programına kullanılan verilerin yerleştirilmesi

Şekil 4.2.'de antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin kümelenmesi, Şekil 4.3'de Antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin risk durumu, Şekil 4.4.'de ise elektrik alan değerlerinin iç ve dış mekana bağlı olarak kümeleme sonuçlarının ekran çıktıları görülmektedir.



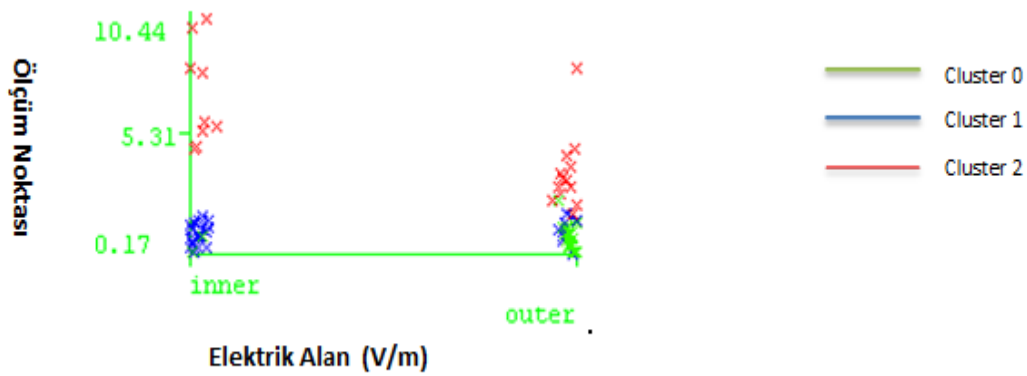
Şekil 4.2. Antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin kümelmesi

Şekil 4.2 görüldüğü gibi tüm ölçüm değerleri frekans bandlarına, iç ve dış mekan ölçümlerine bakılmaksızın Cluster0, Cluster1, Cluster2 olarak 3 ayrı kümeye ayrıldı. Şekil 4.2’de Cluster 0 , Cluster 1, Cluster 2 içerisindeki veriler; elektrik alan - antenden uzaklığına göre dağılımını göstermektedir. Grafikten görüldüğü gibi Cluster 0 da düşük seviyedeki elektrik alan değerleri antenden daha uzak noktalarda yoğunlaşmaktadır, Cluster 2’de ise yüksek seviyeli elektrik alan değerleri antene yakın lokasyonlarda yoğunluk kazanmakta iken Cluster 1 de ise her iki küme arasında kalmaktadır.



Şekil 4.3. Antene uzaklığa bağlı olarak elektromanyetik alan değerlerinin risk durumu

Şekil 4.3’de Antene olan uzaklığa göre belirlenen 3 ayrı risk kriterine göre dağılımları gösterilmiştir. Risk değerlendirme kriteri; V/m birimindeki elektrik alan şiddet seviyelerine göre Tablo 4.2’de kullanılmıştır. Şekilden görüldüğü gibi Normal Doz (ND) ve Kabul Edilebilir Doz (KED) dağılımı içerisinde Cluster0 ve Cluster 1 birlikte yer almaktadır. Ancak Cluster 0 da çoğunluklu olarak ND dominant ve Cluster 1 de ise KED dominanttır. Cluster 2 de ise tümüyle Riski Dozlar(RD) antene çok yakın pozisyonda görülmektedir.



Şekil 4.4. E.A değerlerinin iç ve dış mekana bağlı olarak kümeleme sonuçları

Şekil.4.4’de ise Elektrik Alan Değerlerinin iç ve dış mekana bağlı olarak kümelenmesi gösterilmiştir. İç mekanlarda ve dış mekanlarda yapılan ölçümlerde Cluster0 , Cluster1 ve Cluster2 e ait kümeler birlikte bulunmaktadır. İç mekanlarda daha çok Cluster1 gözükürken, dış mekanlarda ise Cluster2 çoğunluktadır. Ancak dış mekanlarda Cluster0 riski de göze çarpmaktadır.

## BÖLÜM 5. SONUÇ ve ÖNERİLER

Günümüzde iletişim teknolojisinin hızla gelişmesi sonucu piyasaya her çıkan yeni ürün, yeni bir elektromanyetik kirlilik kaynağı olmaktadır. Her teknolojik haberleşme araçları modern hayatın vazgeçilmez bir unsuru olarak eğitim, tıp, askeri vs alanlarda yoğun ve kaçınılmaz kullanımı, elektromanyetik radyasyonun yararlı özelliği ile birlikte biyoelektromanyetik etkileşim dolayısıyla zarar boyutunun tartışılmasını gündeme getirmiştir.

Her faaliyette olduğu gibi modern yaşamda vazgeçilemez olan ve cep telefonu kullanımı gibi bağımlılık veren elektromanyetik enerjinin sürekli kullanımı insan sağlığı açısından değişik seviyelerde zarar oluşturma potansiyeline sahiptir. İnsan yaşamında bazı sağlık sorunlarına sebep olacak tüm etkenler risk olarak tanımlanır. Neticesi belirsiz olan her işin bir riski olabilir. Elektromanyetik radyasyon maruziyeti açısından risk olduğu bazen kişisel bazda kişilere bağlı veya kişilerden bağımsız toplumun tümüne yönelik etkenlerden kaynaklanabilir.

Örneğin cep telefonu kullanan bir kişi kullandığı süre içinde elektromanyetik etkiye bilerek ve farkında olarak kabullenerek maruz kalır. Ancak yaşamış olduğu evinin yakınında bulunan ve uygun olmayan bir şekilde konumlandırılmış baz istasyonundan yayılan elektromanyetik radyasyon maruziyetinden farkındasız ve istemdişi etkilenir. Bu durumdaki maruziyetlerin elektromanyetik alan ölçümleri yapılarak risk analizlerinin yapılması ve bu suretle sağlıklı yaşam ortamlarının temini açısından önemlidir.

Yapılan bu elektromanyetik kirlilik analiz çalışmasında şehir yaşamında mevcut elektromanyetik risk görüntüsü ortaya çıkarılmıştır. Veri madenciliği kümeleme metodlarından K- means algoritması kullanılarak analiz edilmiştir. Yapılan analizde üç değişik risk kümelemesi ile elde edilen sonuçlar baz istasyonlarına yakınlık, ölçüm lokasyonlarının iç mekan ve dış mekan olması parametrelerine göre



tartışılmıştır. Elde edilen sonuçlar yapılacak yeni düzenlemelere katkı sağlayacak sonuçlar içermektedir. Bu çalışmaların benzerleri diğer şehir merkezleri için de yapılarak elektromanyetik risk azaltılması ve elektromanyetik kirlilik kontrolü hususunda önemli sonuçlar elde edilebilir.

Bu tezde yapılan çalışmalar veri madenciliği konsepti itibariyle ülkemizde elektromanyetik radyasyon kirlilik risk kontrolünde yapılması düşünülecek yöntem tasarımlarına katkı sağlayacağı umulmaktadır.

## KAYNAKLAR

Adriaans, P. ve Zantinge, D., Data Mining, , Boston, MA, USA Addison Wesley Longman Publishing,

Tantug, A.,Veri Madenciliği ve Demetleme, ITO, Yüksek Lisans Tezi, 2002.

Albayrak, M., EEG Sinyallerindeki Epileptiform Aktivitenin Veri Madenciliği Süreci ile Tespiti, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 2008.

Alper Vahaplar, Dr. Mustafa Inceoglu, "Veri Madenciliği ve Elektronik Ticaret", Kasım, 2003.

Balmori Alfonso, "Electromagnetic pollution from phone masts. Effects on wildlife, Spain, 2009.

Baltrenas, P. ; Buckus, R.; Vasarevicius, S., Research and evaluation of the intensity parameters of electromagnetic fields produced by mobile communication antennas, Journal of Environmental Engineering and Landscape Management 20(4): 273-284., 2012.

Cerezci O. , Kanberoglu B. , Yener Ç.,Analysis on trending electromagnetic exposure levels at homes and proximity next to base stations along three years in a city, Journal of Environmental Engineering and Landscape Management, 23:1, 71-81, 2015.

G. Linoff, M.J.A. Berry,Data Mining Techniques for Marketing Sales and Customer Support, Wiley Computer Publishing, New York, USA , 1997.

<http://www.google.com> , Erişim Tarihi: 15.04.2015.

Hua Zhu,On-Line Analytical of Association Rules, University of Science and Technology of China, 1995.

İnan, A., Privacy Preserving Distributed Spatio-Temporal Data Mining, ,Sabancı University, Computer Science and Engineering, Yüksek Lisans Tezi, 2006.

İnan, O., Veri Madenciliği, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 2003.

Kalıkov, A., Veri Madenciliği ve Bir E-Ticaret Uygulaması, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 2006.

Piramuthu, S. , Evaluating Feature Selection Methods For Learning in Data Mining Applications, Thirty-First Annual Hawaii International Conference on System Sciences, IEEE Computer Society, Kohala Coast Hawaii USA, 1998.

Robert L. Grossman, Chandrika Karnath, Vipin Kumar, Data Mining For Scientific And Engineering Approach, Kluwer Academic Publishers, 2001.

Seker, S. ,Determining of electromagnetic pollution in Bursa Nilufer district and recommending a sample model to decrease exposure levels, in National Conference on Electrical, Electronics and Computer Engineering, 2010.

Shearer, C., The Crisp-DM Model: The New Blueprint for Data Mining Journal of Data Warehousing, Cilt 5 No 4, 13-2., 2000.

The Bioinitiative Report (A Rationale for Biologically-based Public Exposure Standards for Electromagnetic Fields ) 2012.

Fayyad, U., Gregory, P., Padhraic Smyth, From Data Mining to Knowledge Discovery in Databases, USA, AAAI Pres., 1996.

Fayyad, U., Gregory, P., R. Uthurusamy, Advances in Data Mining and Knowledge Discovery, MIT Pres, 1994.

## ÖZGEÇMİŞ

Feyza ÇEREZCİ, 29.09.1991'da Sakarya'da doğdu. 2009 yılında Figen Sakallıođlu Anadolu Lisesi'nden mezun oldu. 2009 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nü 2013 yılında bitirdi. 2013 yılında Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nde yüksek lisans eğitimine başladı. 2015 yılında Sakarya Üniversitesi'nde araştırma görevlisi olarak çalışmaya başladı. Ve halen Sakarya Üniversitesi Fen Bilimleri Enstitüsünde araştırma görevlisi olarak görev yapmaktadır.