



**T.C.  
FIRAT ÜNİVERSİTESİ  
TEKNOLOJİ FAKÜLTESİ**

**YAZILIM MÜHENDİSLİĞİ  
YMH 418 GÜNCEL  
KONULAR DERSİ PROJE  
DOSYASI  
20.04.2020 – 24.04.2020**

**BÖLÜMÜ : YAZILIM MÜHENDİSLİĞİ**

**NUMARASI : 15542525**

**ADI ve SOYADI : ÇAĞDAŞ KARACA**

### III. Aşama - Verinin Karakteristiklerinin Analizi:

#### **SimpleKmeans Algoritması:**

K-means algoritması bir kümeleme algoritmasıdır. Kümeleme algoritmaları otomatik olarak verileri daha küçük kümelerle yada alt kümelerle ayırmaya yarayan algoritmalarlardır. Algoritma istatistiksel olarak benzer nitelikteki kayıtları aynı gruba sokar. Bir elemanın yalnızca bir kümeye ait olmasına izin verilir. Küme merkezi kümeyi temsil eden değerdir.

Algoritmanın amacı:

1- Küme içindeki değerler birbirlerine en çok benzemeli,

2- Kümeler birbirine mümkün olduğunca benzememeli

Bu istekleri gerçekleştirmek için algoritma tarafında sırasıyla şu adımlar gerçekleştirilmelidir:

1- Sınıf merkezlerinin belirlenmesi

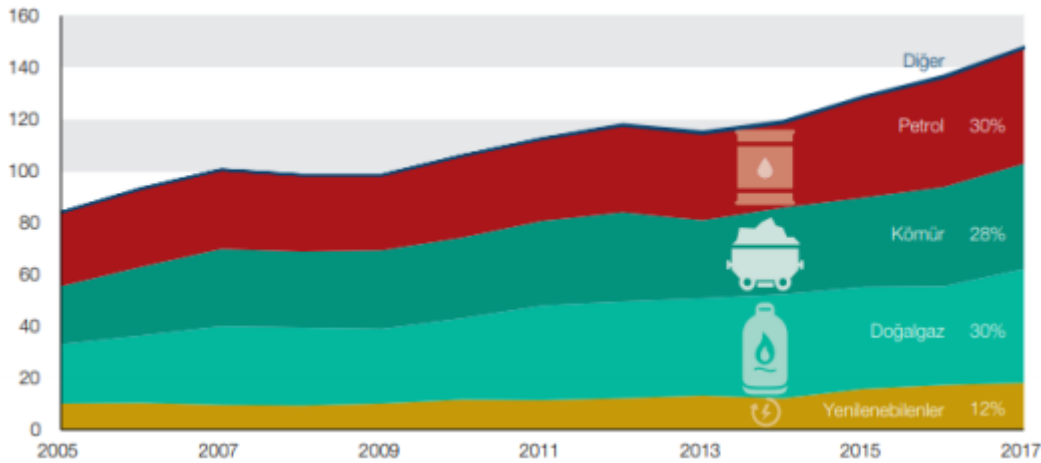
2- Örneklerin mesafelere göre sınıflandırılması

3- Yapılan sınıflandırma sonrasında yeni merkezlerin belirlenmesi

4- İstenilen hale gelinceye kadar 2. ve 3. adımların algoritmik olarak tekrarlanması.

### **HAVA KİRLİLİĞİ**

Hava kirliliği, insan sağlığı, canlı hayatına zarar verecek, eşya ve malzemelerin bozulmalarına, ekolojik dengenin bozulmasına neden olabilecek düzeyde katı, sıvı veya gaz formundaki kirleticilerin atmosferde bulunmasıdır. Hava kirliliği, volkanik aktiviteler, orman yangınları, depremler ve bataklıklar gibi doğal nedenlerle oluşabildiği gibi, sanayileşme, ısınma, ulaşım ve enerji üretimi gibi insan aktivitelerine bağlı da oluşabilmektedir. Nüfus artışı, ve buna bağlı olarak artan şehirleşme ve sanayileşme eğilimi hava kirliliği sorununu beraberinde getirmektedir. Kuraklık, topoğrafik koşullar ve iklimsel özellikler de hava kirliliğinin düzeyini etkilemektedir. Endüstriler ve enerji santralleri, yüksek miktarda partikül madde, azot oksitler, kükürt oksitler, karbon monoksit, hidrokarbon, organik bileşik ve diğer kimyasalların atmosfere salınmasına ve hava kalitesinin azalmasına neden olur. Bu tesisler sadece yerel düzeyde insan ve çevre sağlığına zarar vermekle kalmaz, aynı zamanda bölgesel ölçekte asit yağmurları, ozon oluşumuna, ya küresel ölçekte ise iklim değişikliğine neden olarak dünyanın geleceği için yıkıcı sonuçlara sebep olabilir. Endüstriyel tesislerin enerji üretimi veya ürün prosesi sırasında hava kirliliğine neden olabilen kirletici emisyonları oluşabilmektedir. Enerji üretimindeki hava kirliliğinin en büyük aktörü fosil yakıt kullanımıdır. 19. yüzyıldan beri kullanılmakta olan fosil yakıtlar insan ve çevre sağlığı açısından büyük tahribata neden olmuştur. Enerji ihtiyacı ulaşım sektöründe de kendini göstermektedir. Hava, deniz ve kara ulaşımı için gereken enerjinin çok büyük kısmı halen fosil yakıtlardan karşılanmaktadır.



Görsel 1. Yıllara ve enerji kaynaklarına göre Türkiye'deki kullanımı

Proje de hava kirliliğinin ölçülmesinde kullanılacak olan maddeler aşağıda verilmiştir.

### **Azot dioksit (NO<sub>2</sub>)**

Havanın içersinde ne zaman bir şey yansa, Azot oksitleri oluşacaktır. Bunun nedeni de, solumakta olduğumuz havanın temel olarak Azot (78%) ve Oksijen'den (21%) oluşması ve bunların da ortamda enerji (yanan maddelerden) bulunması durumunda birleşmeleri söz konusu olmaktadır.

En yaygın azot oksitler (genel olarak NO<sub>x</sub> tanımlanmaktadır) azot oksit (NO) ve Azot dioksit (NO<sub>2</sub>)'dir. Azot oksit (NO) kokusuz, renksiz bir gaz olup, içerdeki yakıtın yüksek sıcaklıklarda yakılması sonucu elde edilir, örneğin otomobiller ve diğer karayolu araçları, ısıtıcılar ve pişiriciler. Hava ile temas geçtiklerinde, derhal oksijen ile birleşmeye geçmekte ve sonucunda azot dioksit (NO<sub>2</sub>) oluşturmaktadırlar.

Bu kırmızı kahverengi renkte, alev almayan ve ayırt edici kokuya sahip bir gazdır. Önemli konsantrasyonlarda oldukça zehirli olup sonucunda ciddi gecikmeli etkisi görülen akciğer hasarları ortaya çıkmaktadır. Azot gazına mazur kalmanın diğer sağlık etkileri, nefessiz kalma ve göğüs ağrılarıdır. Azot dioksit kuvvetli bir oksidasyon ajanı olup, hava ile su buharı vasıtası ile reaksiyona girerek korozif nitrik asit oluşturdıkları gibi, aynı zaman da toksik organik nitratlar da oluşturmaktadırlar. Bu durumda, ağaçların, balıkların ve hayvansal yaşamın ölümüne neden olan asit yağmurlarının oluşmasına neden olmaktadır. NO<sub>2</sub> aynı zamanda toprak seviyesi ozon ve dumanlı sis oluşmasını sağlayan atmosferik reaksiyonlarda temel rol oynamaktadırlar. Azot dioksit trafik-alakalı bir kirleticisi olup, konsantrasyonları genel olarak kırsal kesimlere oranla kentsel kesimlerde daha yüksektir.

### **Partiküler Maddeler (PM<sub>10</sub> ve PM<sub>2.5</sub>)**

Partiküller aerodinamik çapları uyarınca tanımlanmaktadır, örneğin PM<sub>10</sub> (aerodinamik çapları 10 µm den daha küçük olan partiküller) veya PM<sub>2.5</sub> (aerodinamik çapları 2,5 µm den daha küçük olan partiküller). PM'in temel komponentleri sülfatlar, nitratlar, ammonia, sodium klorit, karbon, mineral tozları ve sudur. Madde, havada askıda bulunan organik ve inorganik maddelerin katı ve sıvı partiküllerinin bir kompleks karışımıdır. PM<sub>10</sub> genellikle toprak kabuklu malzemeleri ile karayolu aracı ve endüstriyel tesis tozu içermektedir. PM<sub>2.5</sub> ise ikincil derecede oluşmuş aerosoller, yanma partikülleri ve tekrar kondense olmuş organik ve metalik buharlardan oluşmuştur. Partiküler maddelerin asit komponentleri genellikle ince partiküller şeklinde oluşmaktadır. Daha ileri bir ayrıştırma ise partiküllerin orijinleri üzerinden birincil veya ikincil olarak sınıflandırılmaları olacaktır. Birincil partiküller atmosfere direkt olarak yayılırken, ikincil partiküller ise diğer kirleticilerin reaksiyonları sonucunda oluşanlardır. Kırsal çevrede oluşan ikincil partiküller genellikle, içersinde kükürt dioksit ve azot oksitler bulunan reaksiyonlar sonucunda oluşan sülfatlar ve nitratlar şeklindedir.

### **Troposferik (Atmosferin Alt Tabakalarında) Ozon (O<sub>3</sub>)**

Ozon (O<sub>3</sub>), moleküler oksijen'in (O<sub>2</sub>) bir tri-atomik formudur. Madde toksik, uçuk mavi renkte, istikrarsız bir gaz olup, keskin bir kokuya sahiptir. Ozon özellikle stratosferde olmak üzere dünya yüzeyinden 19 ila 30 km. yukarıda doğal olarak bir tabaka şeklinde bulunmaktadır. Söz konusu yükseltilerde ozon, yer yüzeyine inmekte olan ultraviyole (UV) radyasyonunu filtre etmektedir. Yeryüzü seviyesinde ozon, insan sağlığına önemli bir tehdit arz etmektedir. Ozon güçlü bir oksitleyicidir. Troposferik Ozon'un atmosferdeki ömrü yaklaşık 22 gündür.

Kirletici	Ana Kaynağı	Etkisi
Kükürtdioksit(SO <sub>2</sub> )	Fosil Yakıt Yanması, Taşıt Emisyonları	Solunum Yolu Hastalıkları, Asit Yağmurları
Azotoksitler(NO <sub>x</sub> )	Taşıt Emisyonları, Yüksek Sıcaklıkta Yakma Prosesler	Göz Ve Solunum Yolu Hastalıkları, Asit Yağmurları
Partikül Madde (PM)	Sanayi, Taşıt Emisyonları, Fosil Yakıt Yanması, Tarım Ve İkincil Kimyasal Reaksiyonlar	Kanser, Kalp Problemleri, Solunum Yolu Hastalıkları, Bebek Ölüm Oranlarında Artış
Karbonmonooksit(CO)	Eksik Yanma Ürünü, Taşıt Emisyonları	Kandaki Hemeogloblin İle Birleşerek Oksijen Taşınma Kapasitesinde Azalma, Ölüm
Ozon (O <sub>3</sub> )	Trafikten Kaynaklanan Azot Oksitler ve Uçucu Organik Bileşiklerin (VOC) Güneş Işığıyla Değişimi	Solunum Sistemi Problemleri, Göz Ve Burunda İritasyon, Astım, Vücut Direncinde Azalma

Tablo 1.Hava da tespit edilen maddeler, ana kaynakları ve maddelerin etkileri

İSTASYONLAR	Ulusal Sınır Değeri Aşan Saat Sayısı	DSÖ Sınır Değeri Aşan Saat Sayısı	Ölçüm Yapılmayan Saat Sayısı	Veri Alımı (%)	Yıllık Ortalama (µg/m <sup>3</sup> )
İzmir - Alsancak İBB	-	-	8760	0,00	-
İzmir - Bornova İBB	0	0	237	97,29	17
İzmir -Çigli İBB	0	0	1417	83,82	40
İzmir - Güzelyalı İBB	0	0	371	95,76	9
İzmir - Karşıyaka İBB	-	-	8760	0,00	-
İzmir -Sirinyer İBB	0	0	181	97,93	59

Tablo 2. İstasyonlardaki ölçümlere göre NO<sub>2</sub> sınır değerlerini aşan, ölçüm yapılmayan gün sayıları, veri alımı yüzdeleri ve yıllık ortalamalar

Tablo 2 de bulunan veriler Çevre Mühendisleri Odasının 2018 yılı Hava Kirliliği Raporuna bakılarak elde edilmiştir. Alsancak ilçesi için projede kullanılmak üzere alınan veriler Meteoroloji Genel Müdürlüğünden alınmıştır.

Weka programı kullanılarak veri seti üzerinde sınıflandırma işlemi yapılması için veriseti programda çalıştırılmıştır. Elde edilen veriler bölge bölge ayrılarak tüm ilçeler için SimpleKmeans Algoritması kullanılarak aşağıdaki veriler elde edilmiştir.

**Alsancak ilçesi için elde edilen veriler:**

İlk başlangıç noktaları (rastgele):

Küme 0: 2019/5 / 7,17,6.918762,5,98.287879

Küme 1: 2018/3 / 18,26,6.918762,4,98.287879

Attribute	Cluster#		
	Full Data	0	1
	(2200.0)	(273.0)	(1927.0)
=====			
date	2020/4/1	2020/4/10	2020/4/1
pm10	31.753	37.9653	30.8729
no2	6.9188	10.3269	6.4359
so2	4.5847	5.0981	4.512
co	98.2879	7.9969	111.0795

**Kümelenmiş Örnekler**

0 273 (% 12)  
1 1927 (% 88)

0. Küme değeri 2019/5/7,17,6.918762,5,98.287879 1. Küme değeri 2018/3/18,26,6.918762,4,98.287879 olarak alınmış ve elde edilen verilere göre 2020/4/10 tarihinden 2020/4/1 tarihi arasındaki pm10,no2,so2,co değerleri kullanılarak oluşturulan modele göre küme değerleri %88 oranla 1. Kümeye ait olduğu anlaşıyor. Yani 2020/4/1 tarihindeki elde edilen veriler 2018/3/18 tarihinde elde edilen verilerle aynı kümede bulunmaktadır.

**Bayraklı ilçesi için elde edilen veriler:**

Küme 0: 2019/1 / 11,48,15.285714,3,4.465116

Küme 1: 2015/8 / 14,54,15.285714,1,4.465116

Final cluster centroids:

Attribute	Cluster#		
	Full Data	0	1
	(2216.0)	(2198.0)	(18.0)
=====			
date	2020/4/1	2020/4/1	2020/3/15
pm10	52.6083	48.2597	583.6111
no2	15.2857	15.2857	15.2857
so2	4.4861	4.4348	10.7492
co	4.4651	4.4651	4.4651

**Kümelenmiş Örnekler**

0 2198 (% 99)  
1 18 (% 1)

Burada elde edilen verilere bakıldığında 2019/1/11 tarihinde elde edilen veriler 2020/04/1 tarihinde elde edilen verilerin sınıfına ait olduğu anlaşıyor.

### Bornova ilçesi için elde edilen veriler:

İlk başlangıç noktaları (rastgele):

Küme 0: 2018/10 / 1.125,1,3,1

Küme 1: 2016/2 / 11,29,22.371163,1,5

Final cluster centroids:

Attribute	Full Data	Cluster#	
		0	1
	(2170.0)	(1938.0)	(232.0)
=====			
date	2020/4/1	2020/4/1	2020/4/9
pm10	44.7906	44.3716	48.2909
no2	22.3712	14.8117	85.5188
so2	4.6589	4.8514	3.0507
co	5.5005	5.3481	6.7737

Kümelenmiş Örnekler

0 1938 (% 89)

1 232 (% 11)

Rastgele seçilen küme değerleri arasından 2020/4/1 tarihindeki değerler ile 2018/10/1 tarihinde elde edilen veriler ile aynı kümede bulunmaktadır.

### Gazimir ilçesi için elde edilen veriler:

Küme 0: 2018/4 / 13,46,2,16.847695,8,3.068966

Küme 1: 2017/2 / 6,44,2,16.847695,17,3.068966

Final cluster centroids:

Attribute	Full Data	Cluster#	
		0	1
	(2083.0)	(1753.0)	(330.0)
=====			
date	2020/4/1	2020/4/1	2020/2/16
pm10	38.8363	36.557	50.9442
o3	2	2	2
no2	16.8477	15.5844	23.5583
so2	6.7219	5.0677	15.5091
co	3.069	2.8647	4.154

Kümelenmiş Örnekler

0 1753 (% 84)

1 330 (% 16)

2018/4/13 tarihinde elde edilen veriler ile 2020/4/1 tarihinde elde edilen veriler kümesiyle %84 benzerlik gösterdiği gözlemlenmiştir.

### Çiğli ilçesi için elde edilen veriler:

Küme 0: 2014/9 / 25,22,10.516611,7,3.134775

Küme 1: 2019/5 / 29,18,11,5,3.134775

Attribute	Full Data (2206.0)	Cluster#	
		0 (330.0)	1 (1876.0)
date	2020/4/1	2020/4/1	2020/4/2
pm10	35.3831	31.3464	36.0932
no2	10.5166	3.2371	11.7971
so2	6.8696	6.6152	6.9144
co	3.1348	1.8893	3.3539

#### Kümelenmiş Örnekler

0 330 (% 15)  
1 1876 (% 85)

Seçilen 1. Küme değeri 2019/5/29 tarihindeki veriler ile 2020/4/2 tarihinde elde edilen veriler %85 oranında benzerlik göstermektedir ve bu değere bakılarak aynı kümede oldukları belirlenmiştir.

### Güzelyalı ilçesi için elde edilen veriler:

Küme 0: 2014/4 / 25,36,7,1,1

Küme 1: 2019/6 / 10,28,4,4,1

Attribute	Full Data (2206.0)	Cluster#	
		0 (1494.0)	1 (712.0)
date	2020/4/1	2020/4/2	2020/4/1
pm10	37.7799	32.0301	49.8448
no2	7.1364	6.6955	8.0616
so2	5.3797	4.2641	7.7206
co	2.9478	1.8716	5.2061

#### Kümelenmiş Örnekler

0 1494 (% 68)  
1 712 (% 32)

2014/4/25 tarihinde elde edilen veriler ile 2020/4/2 tarihinde elde edilen veriler %68 oranında aynı kümede bulunmaktadır.

### Karşıyaka ilçesi için elde edilen veriler:

Küme 0: 2018/10 / 3,20,8.529511,9,6.053908

Küme 1: 2019/4 / 1,16,8.529511,3,6.053908

Attribute	Full Data (2042.0)	Cluster#	
		0 (546.0)	1 (1496.0)
date	2019/10/1	2019/4/8	2019/10/1
pm10	26.3601	31.754	24.3915
no2	8.5295	9.4159	8.206
so2	5.2303	9.2857	3.7501
co	6.0539	7.7754	5.4256

### Kümelenmiş Örnekler

0 546 (% 27)  
1 1496 (% 73)

2019/4/1 tarihinde elde edilen veriler ile 2019/10/1 tarihinde elde edilen veriler %73 oranında benzerlik gösterdiklerinden aynı kümede bulunmaktadır.

Veri Madenciliği algoritmaları kullanılarak veri setleri ile yapılan çalışma 27/04/2020-01/05/2020 tarihlerinde planan iş akışına göre daha detaylı bir şekilde aktarılacaktır.

Saygılarımla,  
15542525 Çağdaş Karaca