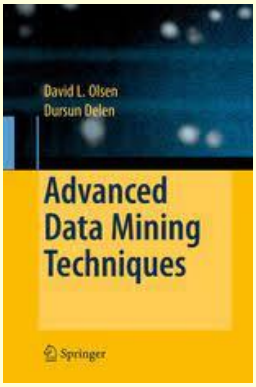
İleri Veri Madenciliği

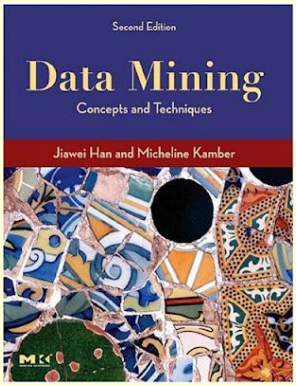
Prof.Dr. Bülent Altunkaynak

Yarıyıl	Kodu	Adı	T+U	Kredi	AKTS
1	VB5062	İLERİ VERİ MADENCİLİĞİ	3+0	3	7

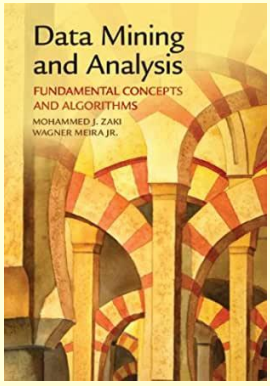
Dersin Detayları	
Dersin Dili	Türkçe
Dersin Düzeyi	Yüksek Lisans
Bölümü / Programı	VERİ BİLİMİ (YL)
Öğrenim Türü	Örgün Öğretim
Dersin Türü	Zorunlu
Dersin Amacı	İleri seviye veri madenciliği algoritmalarını öğrenmek.
Dersin İçeriği	Öğrencilere veri madenciliği ile ilgili temel uygulamaları, güçlü ve zayıf yanlarını, temel kavramları ve teknikleri tanıtmak.
Ön Koşulları	
Dersin Koordinatörü	Yok
Dersi Verenler	Prof.Dr. BÜLENT ALTUNKAYNAK



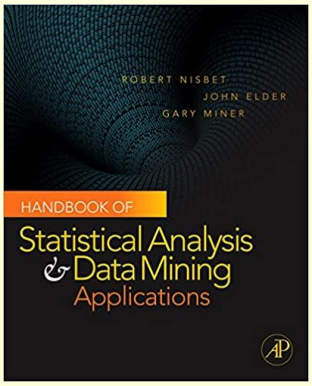
Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Springer Science & Business Media.



Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.



Zaki, M. J., Meira Jr, W., & Meira, W. (2014). *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press.



Nisbet, R., Elder, J., & Miner, G. (2009). *Handbook of statistical analysis and data mining applications*. Academic press.

Yarıyıl Çalışmaları	Sayısı	Katkı
Ara Sınav	1	% 30
Ödev	3	% 30
Uygulama	0	% 0
Projeler	0	% 0
Pratik	0	% 0
Kısa Sınav	0	% 0
Yarıyıl Sonu Sınavı	0	% 40
Toplam :	**4**	**% 100**

Hafta	Konu	Ön Hazırlık	Dökümanlar
1	Veri Madenciliği Süreci		
2	Bellek Tabanlı Akıl Yürütme		
3	Birliktelik Kuralları		
4	Veri Madenciliğinde Bulanık Kümeleme		
5	Veri Madenciliğinde Bulanık Kümeleme		
6	Kaba (Rough) Kümeler		
7	Destek Vektör Makineleri		
8	Destek Vektör Makineleri		
9	Sınıflandırma		
10	Sınıflandırma		
11	Kümeleme		
12	Kümeleme		
13	Web Ve Metin Madenciliği		
14	Zamansal Ve Mekânsal Veri Madenciliği		

- Veri madenciliği (data mining) bilgisayar ortamında depolanmış büyük hacimli verilerden değerli bilgilerin çıkartılmasında kullanılan bir bilgi keşif (knowledge discovery) süreci/yöntemi olarak tanımlanabilir.



Şekil 1.2. Bilgi keşfi yöntemleri

Veri Madenciliği <-> Klasik İstatistik

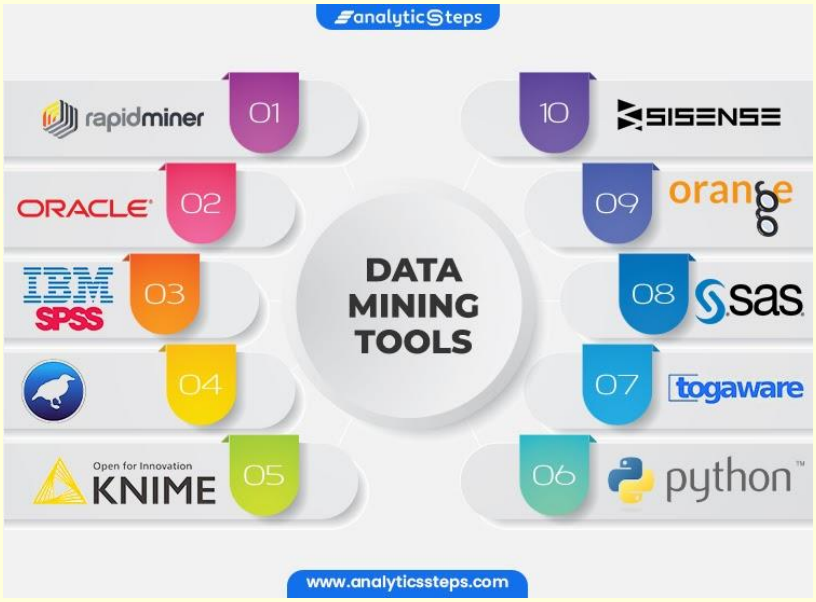
- Veri madenciliği yöntemlerini klasik istatistiksel yöntemlerden ayıran bazı özellikler söz konusudur. Bunlardan bazıları,
  - veri madenciliği yöntemlerinin nispeten daha büyük veri yapılarında kullanılıyor olması,
  - klasik istatistiksel yöntemlerin ihtiyaç duyduğu bir çok varsayımın (dağılım varsayımları, varyans homojenliği varsayımı, dağılımlar arasındaki yakınsamalar için gerekli olan varsayımlar v.b.) çoğu zaman gerekli olmaması,
  - hipotezler oluşturup bu hipotezi test etmeye yönelik verileri toplamak yerine elde bulunan verilerden yararlı sonuçlara ulaşmaya çalışması,
  - yeni veriler elde edildikçe elde edilen modellerin veya kuralların eğitilmesine devam edilmesi

şeklinde verilebilir

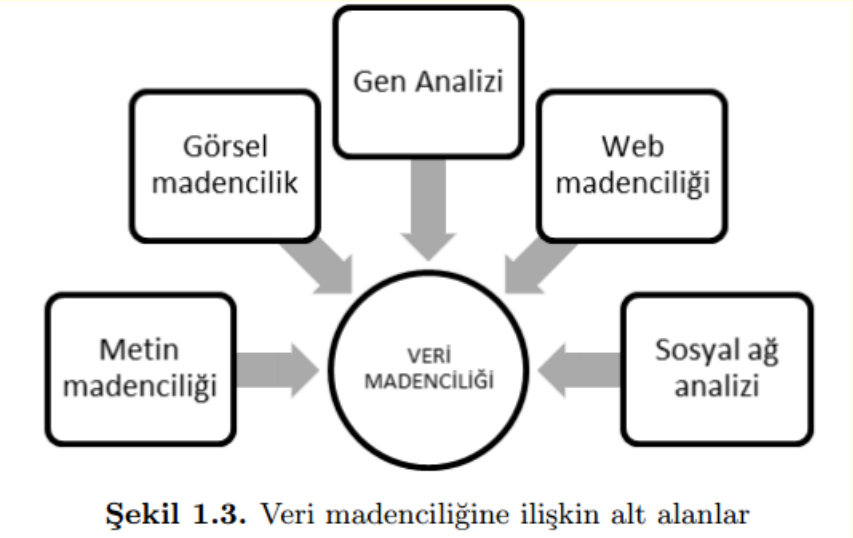
Uygulama Alanları



Yazılımlar

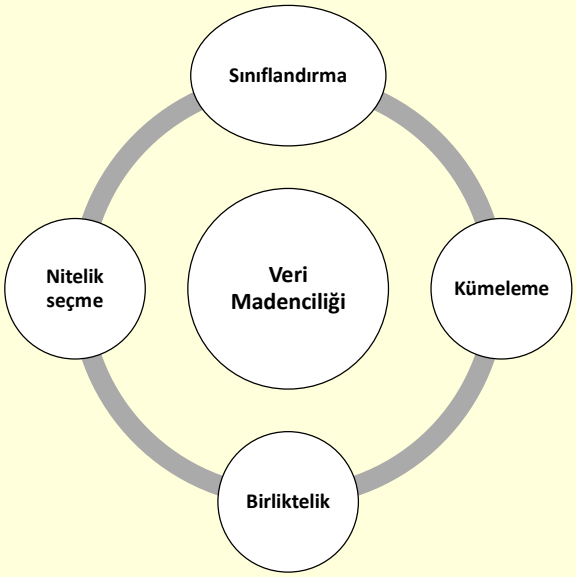


Kullanım Alanları



Şekil 1.3. Veri madenciliğine ilişkin alt alanlar

Kullanılan Yöntemler



Yöntemler

Yaygın Kullanılanlar

- Regresyon (normal, lojistik)
- Yapay sinir ağları
- Karar ağaçları

Daha Az Kullanılanlar

- Birliktelik kuralları
- Bulanık veri madenciliği
- Kaba küme modelleri
- Destek vektör makineleri
- Genetik algoritmalar

Çok yönlü (versatile) ve ölçeklenebilir (scalable)