VERİ MADENCİLİĞİ

Prof.Dr. Bülent ALTUNKAYNAK

Kaynaklar

- Ders slaytları kaynak kitaba dayalı olarak hazırlanmıştır.
- https://rpubs.com/DrAltunkaynak



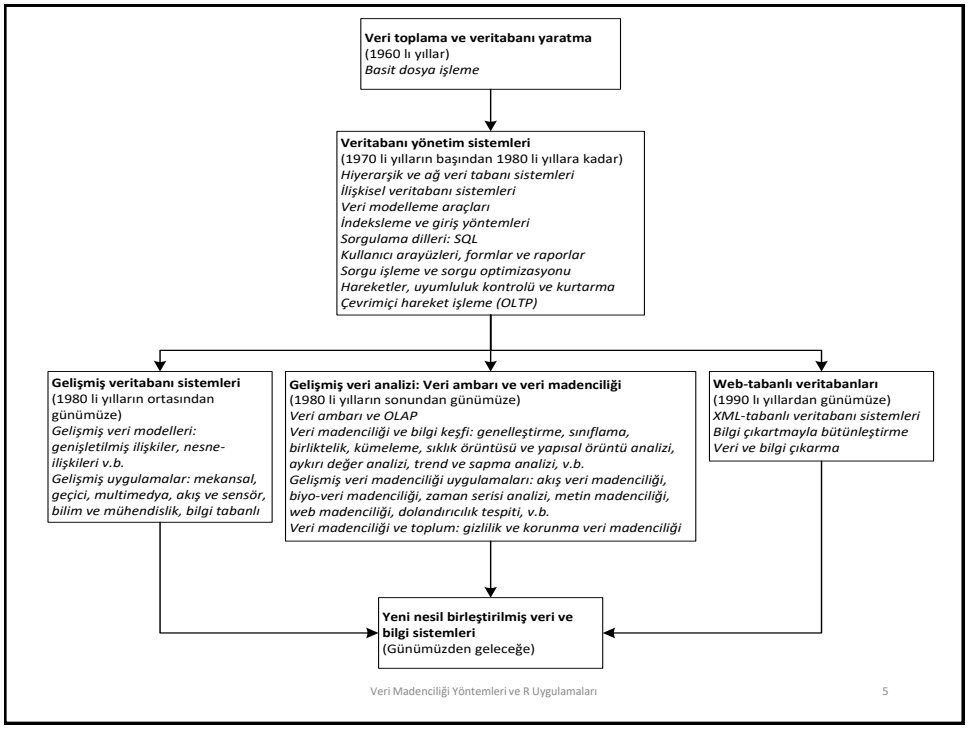
Ders İçeriği

1. Giriş
2. Verinin hazırlanması ve incelenmesi
3. Sınıflandırma Teknikleri
4. Birliktelik Kuralları
5. Kümeleme Teknikleri
6. Özellik seçimi

- 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



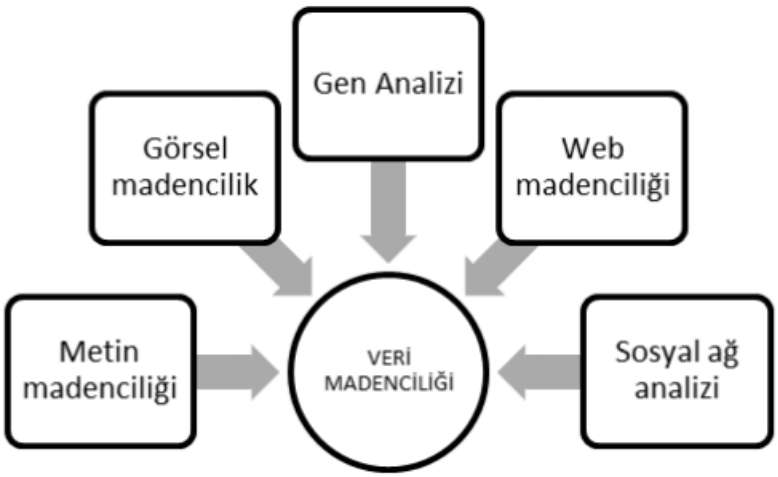
- Uygulama alanları:
  - Müşterilerin satın alma alışkanlıklarının incelenmesi
  - Mevcut müşterilerin elde tutulması ve yeni müşterilerin kazanılması
  - Dolandırıcılık tespiti
  - Satış beklentilerinin tahmin edilmesi
  - İnternet ortamındaki ilişkilerin belirlenmesi ve takip edilmesi
  - Birimlerin farklı özelliklere göre gruplandırılması
  - Birliktelik kurallarının oluşturulması

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

Kullanım Alanları



Şekil 1.3. Veri madenciliğine ilişkin alt alanlar

Kullanılan Yöntemler

