



VERİ MADENCİLİĞİ

Fırat İsmailođlu, PhD

Giriş – Motivasyon - Genel Bakış



Ders Sorumlusu:

Ar. Gör. Dr. Fırat İsmailođlu

Bilgisayar Mühendisliđi, Oda No:202

fismailoglu@cumhuriyet.edu.tr

Ders Hakkında:

- 12 Hafta Ders
- Kitap 1: Introduction to Data Mining, Pang-Ning Tan, Micheal Steinbach, Vipin Kumar
- Kitap 2 : Data Mining the Textbook, Charu Aggarwal
- Yaz okulu olmayacak
- Deđerlendirme: %40 Vize + %60 Final
- Belki Ödev/Proje?



Veri Madenciliği Nedir?

Veri Madenciliği (VM) (büyük) veri depolarından işe yarar, ilginç ve gizli bilgileri çıkarma sürecine denir.



VM ile ortaya çıkarılan bilgiler kullanılarak, elimizdeki veriyi daha iyi tanıyabilir (tanımlayıcı analiz) ve gelecek hakkında tahminde bulunabiliriz (öngörücü analiz).

- **İnternet Verileri:** Kullanıcıların websitelerine erişimi sonrası oluşan 'web access log' dosyaları. Bu dosyalar serverlarda saklanır. Bu dosyalarda kullanıcıların IP adresleri, websitesine erişim zamanı, toplam erişim süresi ve ziyaret edilen sayfalar gibi bilgiler bulunur. Bu dosyalar üzerinde madencilik yapılarak belirli bir websitesinde geçirilen ortalama zaman, websitesine hangi bölgelerden daha çok erişim olduğu ve diğer kullanıcı alışkanlıkları bulunabilir.
- **Finans Verileri:** Kredi kartı işlemleri, kredi başvurusu dosyaları. Bu alanda veri madenciliği yapılarak olagan dışı işlemler bulunabilir, çalıntı kart kullanımı yakalabilir. Yada bankaya kredi için başvuran kişinin alacağı krediyi ödeyip odememe durumu tahmin edilebilir.
- **Alışveriş Verileri:** Bir süpermarkette gün boyu yapılan alışverişlerin oluşturduğu veriler üzerine veri madenciliği yapılarak hangi ürünlerin beraber satın alındığı; yada müşteri profilleri (yaş, cinsiyet, gelir) ile aldıkları ürünler arasında ilişkiler bulunabilir. Kişiyeye özel kampanya yapılabilir.



- **Sensör Verileri:** Günümüzde sensör teknolojisinin yaygınlaşması ve ucuzlaması ile saatten, çamaşır makinesine birçok ürüne yerleştirilen sensörlerin oluşturduğu gerçek zamanlı veriler. Bu veriler üzerine yapılan veri madenciliği çalışmalarıyla örneğin üzerinde sensör olan ürünün ne zaman bozulabileceği ile ilgili tahmin yapılabilir.
- **Kullanıcı Verileri:** Örneğin cep telefonu ile yaptığımız konuşmalar, ne sıklıkla konuştuğumuz, kimleri aradığımız, ne kadar uzun konuştuğumuzun verileri operatörlerde toplanır. Bu veriler üzerine yapılan veri madenciliği çalışmalarıyla kişinin konuşma alışkanlıkları ortaya çıkarılır ve böylece kişisel kampanyalar oluşturabiliriz.

