İçerik

Ders Kodu	Dersin Adı	Yarıyıl	Teori	Uygulama	Lab	Kredisi	AKTS
VM 531	Veri Önizleme ve Kümeleme	3	4	0	0	3	8

Ön Koşul	
Derse Kabul Koşulları	

Dersin Dili	İngilizce
Türü	Zorunlu
Dersin Düzeyi	Yüksek Lisans
Dersin Amacı	Bu ders, öğrencilere veri ön işleme ve kümeleme sürecini tanıtmayı amaçlamaktadır. Bu kapsamda veri hazırlama ve ön işleme, çeşitli veri madenciliği algoritmaları ve bunların sonuçlarını değerlendirmek için kullanılan araçlar ele alınır. Ders; veri ön işleme ve gözetimsiz sınıflandırma (kümeleme) konularında standart yaklaşımlara odaklanır. Madencilik algoritmalarını ve kalite değerlendirme araçlarını anlayabilmek için temel istatistik bilgisi gereklidir.
İçerik	Giriş, genel bakış Veri ön işleme 1 (Temizlik, kayıp veri, gürültülü veri işleme vb) Veri ön işleme 2 (Boyut indirgeme, modelleme ile veri ön işleme vb) Kod uygulaması 1 Kümeleme 1 (genel tanımlar, uzaklık tanımları, kümeleme algoritma kategorileri) Kümeleme 2 (parçalama ve hiyerarşik algoritmalar) Kümeleme 3 (yoğunluk temelli algoritmalar ve kümeleme performans ölçümü) Kod uygulaması 2
Kaynaklar	• Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005. • Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall,1999. • Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000. • Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Willey & sons, 2006. • The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, T. Hastie, R. Tibshirani & J. Friedman, Springer, 2009.

Teori Konu Başlıkları

Hafta	Konu Başlıkları
1	Giriş, genel bakış
2	Tanımlayıcı istatistikler
3	Veri ön işleme 1
4	Veri ön işleme 2
5	Kod Uygulaması 1
6	Ara Sınav
7	Kümeleme 1 (Tanımlar, parçalama yöntemleri)
8	Kümeleme 2 (Hiyerarşik yöntemler ve yoğunluk temelli yöntemler)
9	Kümeleme 3 (Başarı ölçümleri)
10	Kod Uygulaması 2
11	Proje sunumları