İçerik

Ders Kodu	Dersin Adı	Yarıyıl	Teori	Uygulama	Lab	Kredisi	AKTS
INF 511	Veri Bilimi	1	3	0	0	3	6

Ön Koşul	
Derse Kabul Koşulları	

Dersin Dili	İngilizce
Türü	Seçmeli
Dersin Düzeyi	Yüksek Lisans
Dersin Amacı	Bu ders, öğrencilere veri madenciliği sürecini tanıtmayı amaçlamaktadır. Dersin temel amaçları arasında veri hazırlama ve ön işlemenin, çeşitli veri madenciliği algoritmalarının ve bunların sonuçlarını değerlendirmek için mevcut araçların anlaşılması ve kullanılabilmesi vardır. Ders, birliktelik kuralları madenciliği, denetimli sınıflandırma ve denetimsiz sınıflandırma (kümeleme) ile ilgili standart yaklaşımlara odaklanır. Madencilik algoritmalarını ve kalite değerlendirme araçlarını anlamak için temel istatistiksel bilgi gereklidir. Böylece öğrencinin veri analizi alanında pratik çözümler üretebilmesi hedeflenmektedir.
İçerik	1. Veri madenciliği ve kestirimci analitiğe giriş 2. Veri ön işleme, keşifsel veri analizi 3. Boyut indirgeme yöntemleri, tek değişkenli istatistiksel analiz 4. Çok değişkenli istatistik, veriyi modellemeye hazırlama 5. Basit doğrusal regresyon, çoklu regresyon 6. Model oluşturma 7. k-en yakın komşu algoritması, karar ağaçları 8. Lojistik regresyon, Naive Bayes ve Bayes ağları 9. Vize sınavı 10. Model değerlendirme teknikleri 11. Sınıflandırma modellerinin grafiksel değerlendirilmesi 12. Hiyerarşik ve k-means kümeleme, küme kalitesinin ölçülmesi 13. Birliktelik kuralları, topluluk (ensemble) yöntemleri 14. Öğrenci sunumları
Kaynaklar	1. Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005. 2. Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall,1999. 3. Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000. 4. Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Willey & sons, 2006. 5. 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	Veri madenciliği ve kestirimci analitiğe giriş
2	Veri ön işleme, keşifsel veri analizi
3	Boyut indirgeme yöntemleri, tek değişkenli istatistiksel analiz
4	Çok değişkenli istatistik, veriyi modellemeye hazırlama
5	Basit doğrusal regresyon, çoklu regresyon
6	Model oluşturma
7	k-en yakın komşu algoritması, karar ağaçları
8	Lojistik regresyon, Naive Bayes ve Bayes ağları
9	Vize sınavı
10	Model değerlendirme teknikleri
11	Sınıflandırma modellerinin grafiksel değerlendirilmesi

Hafta	Konu Başlıkları
12	Hiyerarşik ve k-means kümeleme, küme kalitesinin ölçülmesi
13	Birliktelik kuralları, topluluk (ensemble) yöntemleri
14	Öğrenci sunumları