

Dersin Tanımı					
Adı	Kodu	Yarıyıl	T+U Saat	Kredi	AKTS
VERİ MADENCİLİĞİ VE İŞ ZEKASI	MIS4112149	Güz Dönemi	3+0	3	5
Ön Koşul Dersleri					
Önerilen Seçmeli Dersler					
Dersin Dili	İngilizce				
Dersin Seviyesi	Lisans				
Dersin Türü	Zorunlu				
Dersin Koordinatörü	Prof.Dr. Gökhan SİLAHTAROĞLU				
Dersi Verenler	Öğr.Gör. Nada A. M. MİSK, Prof.Dr. Gökhan SİLAHTAROĞLU				
Dersin Yardımcıları					
Dersin Amacı	Öğrencilere veri tabanlarından veri ambarı yaratarak, bu veri ambarları üzerinde OLAP ve veri madenciliği modüllerini kullanarak araştırma becerisi kazandırmak ve veri madenciliği algoritmalarını kod düzeyinde uygulayabilecek bilgi düzeyine getirmek.				
Dersin İçeriği	Bu ders; Veri Madenciliğine Giriş,Veri Ambarı –OLAP – ETL süreci – Veri Ambarı ve Veri Tabanı arasında farklılıklar,Veri Madenciliğinde Kullanılan Araçlar: Python, R, KNIME, pivoting.(KNIME Kurulumu, Arayüz, Veri okuma ve görüntüleme),Veri önışlemleri (Uc veriler, Kayıp veriler, Normalizasyon, Değişken Dönüşümü, Binning Histogram),Birlikte Analizi (Alışveriş Sepeti Analizi),Denetimli Öğrenme, Sınıflandırma (Classification) Ve Karar Ağaçları, Gini / Entropy Kısa tanıtım (Uygulama: BEARS),Karar Ağaçları (Devam) (İLAC ve Wine uygulamaları ; , Java Snippet, filtreleme, , train-test, validation, accuracy, Color Manager),Random Forest, Booststrapping, Ensemble Learning (Parfüm Uygulaması + Telco Uygulaması),Regresyon ve Lojistik Regresyon,Yapay sinir Ağları (çalışma prensibi, Hyperparameters),Denetimsiz Öğrenme / Kümeleme (Uygulama: Wholesale Customer Data, fuzzy c-means ve kalite ölçümleri, Entropy Scorer),Boyut İndirgeme, Sentetik Veri üretimi , PCA, Scatter Plot ile kümeleme (DBSCAN),Büyük Veri Konseptleri – HADOOP, Spark, MongoDB, NOSQL,Proje Sunumları - Değerlendirme; konularını içermektedir.				
Dersin Öğrenme Kazanımları	Öğretim Yöntemleri	Ölçme Yöntemleri			
1. Veri Tabanından veri ambarı üretebilecektir.	10, 16, 9	A			
1.1. Veri Madenciliğini açıklar.	16, 9				
1.2. Veri Ambarını tanımlar.	16, 9				
1.3. Yıldız veri ambarı modelini tasarlar.	10, 9				
1.4. Ana tablolar birliği veri ambarı modelini tasarlar.	16, 9				
2. Veri Madenciliği Modellerini birbirleriyle ilişkilendirebilecektir.	16, 9	A, F			
2.1. Veri Madenciliği Modellerini tanımlar.	16, 9				
2.2. Sınıflandırma Kavramını tanımlar.	16, 9				
2.3. Kümeleme kavramını tanımlar.	16, 9				
2.4. Bağlantı analizi kavramını tanımlar.	16, 9				
3. Sınıflandırma modelini uygulayabilecektir.	16, 9	A, F			
3.1. Denetimli öğrenmeyi tanımlar.	16, 9				
3.2. Sınıf kavramını belirler.	9	A			
3.3. İstatistiksel algoritmaları sıralar.	16, 9				
3.4. Karar ağaçlarını uygular.	16, 9	F			
3.5. Karar ağacı algoritmalarını tanımlar.	16, 6, 9				
3.6. Budama ve saflık değerlerini tanımlar.	9				
4. Kümeleme modelini uygulayabilecektir.	14, 6, 9	F			
4.1. Denetimsiz Öğrenmeyi tanımlar.	6, 9				
4.2. Kümeleme kavramını açıklar.	16, 9				
4.3. Kümeleme algoritmalarını sıralar.	6, 9	F			
4.4. K-means algoritmasını uygular.	14, 16, 9				
4.5. Genetik algoritmaları tanımlar.	10, 8, 9	F			
5. Bağlantı analizi modelini uygulayabilecektir.	14, 16, 6, 9	A			
5.1. Bağlantı analizi kurallarını yorumlar.	14, 6, 9				
5.2. Kaldıraç kavramını açıklar.	16, 9	A			
5.3. İlişki analizi yöntemini uygular.	14, 9				
5.4 Bağlantı analizi ile kümelemeyi birleştirir.	2	F			
6. Veri madenciliği Algoritmalarını uygulayabilecektir.	2	F			
6.1. Sınıflandırma algoritmalarını veriler üzerinde uygular.	2	F			
6.2. Kümeleme algoritmalarını veriler üzerinde uygular.	2	F			
6.3. Bağlantı analizi algoritmalarını veriler üzerinde uygular.	2	F			
6.4. Veri madenciliği uygulama çıktılarını yorumlar.	2	F			
Öğretim Yöntemleri	10: Tartışma Yöntemi, 14: Bireysel Çalışma Yöntemi, 16: Soru - Cevap Tekniği , 2: Proje Temelli Öğrenme Modeli, 6: Deneyimle Öğrenme Modeli, 8: Ters-yüz Edilmiş Sınıf Modeli, 9: Anlatım Yöntemi				
Ölçme Yöntemleri	A: Klasik Yazılı Sınav, F: Proje Görevi				
Ders Akışı					
Sıra	Konular	Ön Hazırlık			
1	Veri Madenciliğine Giriş				
2	Veri Ambarı –OLAP – ETL süreci – Veri Ambarı ve Veri Tabanı arasında farklılıklar				
3	Veri Madenciliğinde Kullanılan Araçlar: Python, R, KNIME, pivoting.(KNIME Kurulumu, Arayüz, Veri okuma ve görüntüleme)				
4	Veri önışlemleri (Uc veriler, Kayıp veriler, Normalizasyon, Değişken Dönüşümü, Binning Histogram)				
5	Birlikte Analizi (Alışveriş Sepeti Analizi)				

Ders Akışı		
Sıra	Konular	Ön Hazırlık
6	Denetimli Öğrenme, Sınıflandırma (Classification) Ve Karar Ağaçları, Gini / Entropy Kısa tanıtım (Uygulama: BEARS)	
7	Karar Ağaçları (Devam) (İLAC ve Wine uygulamaları ; Java Snippet, filtreleme, , train-test, validation, accuracy, Color Manager)	
8	Random Forest, Booststrapping, Ensemble Learning (Parfüm Uygulaması + Telco Uygulaması)	
9	Regresyon ve Lojistik Regresyon	
10	Yapay sinir Ağları (çalışma prensibi, Hyperparameters)	
11	Denetimsiz Öğrenme / Kümeleme (Uygulama: Wholesale Customer Data, fuzzy c-means ve kalite ölçümleri, Entropy Scorer)	
12	Boyut İndirgeme, Sentetik Veri üretimi , PCA, Scatter Plot ile kümeleme (DBSCAN)	
13	Büyük Veri Konseptleri – HADOOP, Spark, MongoDB, NOSQL	
14	Proje Sunumları - Değerlendirme	
Değerlendirme Yöntemleri		Sınava Katkısı
Ara Sınav		40
(Genel Sınav) Proje (Genel Sınav olarak)		60
Genel Sınav		60

Kaynaklar
1. Data Mining Introductory and Advanced Topics, Margaret H. Dunham, Prentice Hall. Krnime Application: https://docs.knime.com/ 1. Data Mining Concepts and Techniques , J. Han & M. Kamber, Morgan Kaufman. 2. Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results , Bernard Marr, Wiley, 2016 3. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, Cathy O'Neil ,2017 4. Naked Statistics: Stripping the Dread from the Data, Charles Wheelan, 2013 5. Kavram ve Algoritmalarıyla Temel Veri Madenciliği, Gökhan Silahtaroğlu, Papatya Yayıncılık.