

GENEL TANIM / GENERAL DESCRIPTION

Ders Adı / Course Name	UNSUPERVISED MACHINE LEARNING METHODS / UNSUPERVISED MACHINE LEARNING METHODS	
Ders Kodu / Course Code	İST431	
Ders Türü / Course Type		
Ders Seviyesi / Course Level	First Cycle / First Cycle	
Ders Akts Kredi / ECTS	5.00	
Haftalık Ders Saati (Kuramsal) / Course Hours For Week (Theoretical)	3.00	
Haftalık Uygulama Saati / Course Hours For Week (Objected)	0.00	
Haftalık Laboratuvar Saati / Course Hours For Week (Laboratory)	0.00	
Dersin Verildiği Yıl / Year	4	
Öğretim Sistemi / Teaching System	Face to Face / Face to Face	
Eğitim Dili / Education Language	English / English	
Ön Koşulu Olan Ders(ler) / Precondition Courses	Yok	None
Amacı / Purpose	Dersin amacı, öğrencilere denetimsiz öğrenme yöntemlerini detaylı bir şekilde öğretmektir. Bu kapsamda, Temel Bileşenler Analizi, k-ortalama ve k-medoid kümeleme, hiyerarşik kümeleme ile birlikte model ve dağılıma dayalı kümeleme yöntemleri gibi temel modeller ve algoritmalar hem klasik, hem de bulanık yöntemler ele alınarak incelenecektir. Bu süreçte R programlama dili ve Weka yazılımı kullanılacak olup, öğrencilerden lineer cebir ve olasılık konularında temel bir bilgiye sahip olmaları beklenmektedir.	The aim of the course is to provide students with a comprehensive understanding of unsupervised learning methods. In this context, fundamental models and algorithms, including Principal Component Analysis, k-means and k-medoid clustering, hierarchical clustering, as well as model and distribution-based clustering methods, will be extensively explored. Both classical and fuzzy methods will be covered. Throughout this process, R programming language and Weka software will be utilized. Students are expected to have a basic knowledge of linear algebra and probability.
İçeriği / Content	Veri bölme, çapraz doğrulama ve bootstrap, Temel Bileşenler Analizi, kümelemede uzaklık ve benzerlik ölçüleri, Bölünmeye dayalı kümeleme yöntemleri (k-ortalamlar, bulanık c-ortalamlar, k-medoid, bulanık c-medoid), hiyerarşik kümeleme yöntemleri, modele dayalı kümeleme yöntemleri, komşuluğa/yoğunluğa dayalı klasik ve bulanık kümeleme yöntemleri (DBSCAN, FN-DBSCAN), küme geçerliliği yöntemleri	Data splitting, cross-validation, and bootstrap; Principal Component Analysis; distance and similarity measures in clustering; Partition-based clustering methods (k-means, fuzzy c-means, k-medoid, fuzzy c-medoid); Hierarchical clustering methods; Model-based clustering methods; Crisp and fuzzy clustering methods based on proximity/density (DBSCAN, FN-DBSCAN); Cluster validity methods
Önerilen Diğer Hususlar / Recommended Other Considerations		
Staj Durumu / Internship Status	Yok	None

Kıtabı / Malzemesi / Önerilen Kaynaklar / Books / Materials / Recommended Reading	1.Tom, M. Mitchell (1997). Machine Learning, McGrawHill, 2.Friedman J., Hastie, T., Tibshirani, R. (2013). The Elements of Statistical Learning Data Mining, Inference and Prediction Preface to the Second Edition. 3.Friedman J., Hastie, T., Tibshirani, R. (2013). The Elements of Statistical Learning Data Mining, Inference and Prediction Preface to the Second Edition. 4.James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer, New York. 5.Albalate, A. & Minker, W. (2011). Semi-Supervised and Unsupervised Machine Learning. John Wiley & Sons, Inc., London. 6.Alpar R. (2017). Uygulamalı Çok Değişkenli İstatistiksel Yöntemler. Detay Yayıncılık.	1.Tom, M. Mitchell (1997). Machine Learning, McGrawHill, 2.Friedman J., Hastie, T., Tibshirani, R. (2013). The Elements of Statistical Learning Data Mining, Inference and Prediction Preface to the Second Edition. 3.Friedman J., Hastie, T., Tibshirani, R. (2013). The Elements of Statistical Learning Data Mining, Inference and Prediction Preface to the Second Edition. 4.James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer, New York. 5.Albalate, A. & Minker, W. (2011). Semi-Supervised and Unsupervised Machine Learning. John Wiley & Sons, Inc., London. 6.Alpar R. (2017). Uygulamalı Çok Değişkenli İstatistiksel Yöntemler. Detay Yayıncılık.
Öğretim Üyesi (Üyeleri) / Faculty Member (Members)	Prof.Dr.Gözde ULUTAGAY	Prof.Dr.Gözde ULUTAGAY

ÖĞRENME ÇIKTILARI / LEARNING OUTCOMES

1	Veri odaklı öğrenme konusunda bir anlayış geliştirme.	Develop an understanding of data-driven learning.
2	Veri boyutunu azaltma yeteneği kazanma.	Acquire the ability to reduce data dimensionality.
3	Çeşitli kümeleme yöntemlerini veri üzerinde uygulama becerisi edinme.	Gain the skill to apply various clustering methods to data.
4	Kümelenmiş veriyi etkili bir şekilde görselleştirebilme.	Effectively visualize clustered data.
5	Kümeleme algoritmalarını değerlendirme ve model seçme süreçlerini anlama.	Understand the process of evaluating clustering algorithms and selecting models.
6	Kümeleme yöntemlerini uygulamak için R programlama dilini ve Weka yazılımını kullanabilme.	Use the R programming language and Weka software to implement clustering methods.

HAFTALIK DERS İÇERİĞİ / DETAILED COURSE OUTLINE

Hafta / Week					
1	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Makine Öğrenimi ile İstatistik Arasındaki Fark: Büyük Örneklem Boyutu, Yüksek Boyut, Çok Değişkenli Yanıtlar				
	Statistics versus Machine Learning: Large Sample Size, High Dimensionality, Multivariate Responses				
2	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Veri bölme, çapraz doğrulama ve bootstrap				
	Data splitting, cross-validation, and bootstrap				
3	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Temel Bileşenler Analizi				
	Principal Component Analysis				
4	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Kümelemede uzaklık ölçütleri- uzaklık ve benzerlik ölçme yöntemleri				
	Distance metrics in clustering- methods for measuring distance and similarity				
5	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Kümelemede uzaklık ölçütleri- veri standartlaştırma, uzaklık matrisi oluşturma ve uzaklık matrislerini görselleştirme				
	Distance metrics in clustering- data standardization, creating distance matrices, and visualizing distance matrices				

	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
6	Bölünmeye dayalı kümeleme yöntemleri: k-ortalamalar, bulanık c-ortalamalar (FCM)				
	Partition-based clustering methods: k-means, fuzzy c-means (FCM)				
7	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Bölünmeye dayalı kümeleme yöntemleri: k-medoids bulanık k-medoids				
	Partition-based clustering methods: k-medoids, fuzzy k-medoids				
8	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Hiyerarşik kümeleme yöntemleri				
	Hierarchical clustering methods				
9	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Modele dayalı kümeleme yöntemleri				
	Model-based clustering methods				
10	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Komşuluğa (Yoğunluğa) dayalı kümeleme yöntemleri: DBSCAN				
	Density-based (neighborhood-based) clustering methods: DBSCAN				
11	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Bulanık Komşuluğa (Yoğunluğa) dayalı kümeleme yöntemleri: FN-DBSCAN				
	Fuzzy density-based (fuzzy neighborhood-based) clustering methods: FN-DBSCAN				

	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
12	Küme geçerliliği yöntemleri				
	Cluster validity criteria				
13	Proje sunumları				
	Project presentation				
14	Proje sunumları				
	Project presentation				

DEĞERLENDİRME / EVALUATION

Yarıyıl (Yıl) İçi Etkinlikleri / Term (or Year) Learning Activities	Sayı / Number	Katkı Yüzdesi / Percentage of Contribution (%)
Ara Sınav / Midterm Examination	1	100
Toplam / Total:	1	100
Başarı Notuna Katkı Yüzdesi / Contribution to Success Grade(%):		40

Yarıyıl (Yıl) Sonu Etkinlikleri / End Of Term (or Year) Learning Activities	Sayı / Number	Katkı Yüzdesi / Percentage of Contribution (%)
Final Sınavı / Final Examination	1	100
Toplam / Total:	1	100
Başarı Notuna Katkı Yüzdesi / Contribution to Success Grade(%):		60

Etkinliklerinin Başarı Notuna Katkı Yüzdesi(%) Toplamı / Total Percentage of Contribution (%) to Success Grade:	100
Değerlendirme Tipi / Evaluation Type:	

İŞ YÜKÜ / WORKLOADS

Etkinlikler / Workloads	Sayı / Number	Süresi (Saat) / Duration (Hours)	Toplam İş Yüğü (Saat) / Total Work Load (Hour)
Ara Sınav / Midterm Examination	1	2.00	2.00
Ara Sınav İçin Bireysel Çalışma / Individual Study for Mid term Examination	1	25.00	25.00
Final Sınavı / Final Examination	1	2.00	2.00
Final Sınavı için Bireysel Çalışma / Individual Study for Final Examination	1	30.00	30.00
Performans / Performance	14	1.00	14.00
Proje Hazırlama / Project Preparation	1	30.00	30.00
Proje Sunma / Project Presentation	1	20.00	20.00
Uygulama/Pratik / Practice	13	1.00	13.00
Toplam / Total:	33	111.00	136.00
<p>Dersin AKTS Kredisi = Toplam İş Yüğü (Saat) / 30.00 (Saat/AKTS) = 136.00/30.00 = 4.53 ~ / Course ECTS Credit = Total Workload (Hour) / 30.00 (Hour / ECTS) = 136.00 / 30.00 = 4.53 ~</p>			

PROGRAM VE ÖĞRENME ÇIKTISI / PROGRAM LEARNING OUTCOMES

Öğrenme Çıktıları / Learning Outcomes	Program Çıktıları / Program Outcomes																							
	1.1.1	1.1.2	1.1.3	1.1.4	1.1.5	1.1.6	1.1.7	1.1.8	1.1.9	1.1.10	1.1.11	1.1.12	1.1.13	1.1.14	1.1.15	1.1.16	1.1.17	1.1.18	1.1.19	1.1.20	1.1.21	1.1.22	1.1.23	1.1.24
1. Veri odaklı öğrenme konusunda bir anlayış geliştirme. / Develop an understanding of data-driven learning.	5	5	5			5					3													
2. Veri boyutunu azaltma yeteneği kazanma. / Acquire the ability to reduce data dimensionality.	5	4																						
3. Çeşitli kümeleme yöntemlerini veri üzerinde uygulama becerisi edinme. / Gain the skill to apply various clustering methods to data.	5	5		5																				
4. Kümelenmiş veriyi etkili bir şekilde görselleştirebilme. / Effectively visualize clustered data.	5	4	4																					
5. Kümeleme algoritmalarını değerlendirme ve model seçme süreçlerini anlama. / Understand the process of evaluating clustering algorithms and selecting models.	5	4																						
6. Kümeleme yöntemlerini uygulamak için R programlama dilini ve Weka yazılımını kullanabilme. / Use the R programming language and Weka software to implement clustering methods.	5	4																						

Katkı Düzeyi / Contribution Level : 1-Çok Düşük / Very low, 2-Düşük / Low, 3-Orta / Moderate, 4-Yüksek / High, 5-Çok Yüksek / Very high