

BÜYÜK VERİ

1	Ders Adı:	BÜYÜK VERİ
2	Ders Kodu:	IYS4213
3	Ders Türü:	Seçmeli
4	Ders Seviyesi	Lisans
5	Dersin Verildiği Yıl:	4
6	Dersin Verildiği Yarıyıl	7
7	Dersin AKTS Kredisi:	6.00
8	Teorik Ders Saati (saat/Hafta)	3.00
9	Uygulama Ders Saati(saat/Hafta)	0.00
10	Laboratuvar Ders Saati (saat/hafta) :	0
11	Dersin Önkoşulu:	
12	Dersin Dili:	Türkçe
13	Dersin Veriliş Şekli:	Yüz yüze
14	Dersin Koordinatörü:	Doç. Dr. MELİH ENGİN
15	Dersi Veren Diğer Öğretim Elemanları:	Doç.Dr. Melih ENGİN
16	Koordinatör İletişim Bilgileri:	Doç.Dr. Melih ENGİN 0224 294 26 95 melihengin@uludag.edu.tr
17	Dersin WEB adresi:	
18	Dersin Amacı:	Bu ders Büyük Verideki mevcut teknolojileri, araçları, mimarileri ve sistemlerin kullanımını kapsıyarak analitik veri üretimi, depolama, yönetim, transfer, gelen büyük verilerde derinlemesine analizi kapsar. bir kapsama alanı sağlar yüksek performanslı ağlarda veri işlem çözümleri. Çeşitli alanlarda ortaya çıkan bigdata uygulamalarını inceleyerek, uygulama, geliştirme konularında kapsayarak yaygın olarak kullanılan büyük veri uygulamalarını test eder. Ayrıca büyük veriyi analiz etmek için veri madenciliği ve makine öğrenme algoritmaları üzerinde durulacak.
19	Dersin Mesleki Gelişime Katkısı:	Bir işletme için gerekli olan sistemleri tasarlama ve iders bağlamındaki ihtiyaçlarına çözüm üretebilme.
20	Dersin Öğrenme Kazanımları:	
	1	"Teori" ve "uygulama" dengeli bir şekilde götürülerek öğrencinin büyük veri analizini ve yönetimini anlaması, kullanması ve pratik analizini öğretmeye yöneliktir. Ders öğrencinin büyük on-line depo sistemlerinde yer alan konuların ve sorunların anlaşılmasını sağlamak için tasarlanmıştır, böyle bir sisteminin ihtiyaçlarını karşılamak için günümüz pratik teknikleri, bilgi ve yarının araştırma yaklaşımlarının çözümleri için bir temel oluşturmayı planlamaktadır.
	2	Büyük veri kavramlarını, terminolojiyi, veri analitiği özelliklerini, 5V-yapısal-yapısal olmayan-metadada gibi büyük veri tiplerini öğreneceklerdir.
	3	Nitel - nicel veri madenciliği, istatistiksel analiz, A/B testi, korelasyon, regresyon analizi gibi analiz tekniklerini kavrarlar.
	4	Kümeleme, dağıtık dosya sistemleri, ilişkisel veritabanı sistemleri, NoSQL, hafıza içi depolama gibi depolama kavramlarına ve paralel, dağıtık, toplu veri işleme gibi büyük veri işleme kavramlarına hakim olurlar.

	5	Büyük veri işlenmesinde Cloudera sanal makina, HDFS (Hadoop Dağıtık Dosya Sistemi), YARN (Yet Another Resource Negotiator and Hue) gibi paralel işleme ve diğer dizayn kalıplarını kavrarlar.
	6	
	7	
	8	
	9	
	10	
21	Dersin İçeriği:	
Hafta	DERS İÇERİKLERİ	
	Teorik	Uygulama
1	Büyük Veriye Giriş: Kavramlar, terminoloji, özellikleri ve 5V, yapılandırılmış, yapılandırılmamış, yarı yapılandırılmış ve metadata gibi büyük Veri türlerini kapsar. İş ve araştırma motivasyonlarını kapsar.	
2	Büyük Veride Depolama ve Analiz: Kümeler gibi depolama kavramları kapsar, dağıtık dosya sistemleri, RDBMS, NoSQL, sharding, bellek depolama, aynı zamanda paralel, dağıtık, toplu veri işleme ve Hadoop gibi büyük veri işleme konseptlerini kapsar.	
3	Büyük Veri Analizi Teknikleri: Nicel, nitel veri madenciliği, istatistiksel analiz, Machine Learning, anlamsal analiz ve görsel veri analizi gibi analiz teknikleri ve konularını kapsar.	
4	MapReduce Çerçevesi ve Hadoop: Paralel süreçler ve büyük veri işleme için diğer tasarım desenleri kapsar. Cloudera sanal makine. HDFS, YARN ve Hue.	
5	Java ile mapreduce API ve Basic Programlama: Hadoop MapReduce Java API'yi inceleyeceğiz.	
6	Hive kullanımı: "veri ambarı" HDFS'ler ve Hadoop. Depolanan veriler üzerinde SQL sorguları.	
7	Spark Kullanımı: MapReduce çerçevesinin bellek tabanlı veri analizi ve RDD-Spark.	
8	Flume ve Kafka eğitimi.	
9	Spark veri akışı, Kafka ve Cassandra: Hızlı verilerin işlenmesi için standart haline gelmiş yığın kullanımı.	
10	Spark ile MLLib, Machine Learning: Veriler üzerinde öngöründe bulunmak için öğrenme algoritmalarını gözden geçireceğiz.	
11	Büyük Veri Kümelerinde görselleştirme: Büyük verinin içeriğini ve özelliklerini grafiksel olarak incelenmesini göreceğiz.	
12	Büyük veride ileri konular ve uygulamalar.	
13	Büyük veride ileri konular ve uygulamalar.	
14	Tekrar haftası.	

22	Ders Kitabı, Referanslar ve/veya Diğer Kaynaklar:	1. Big Data Fundamentals: Concepts, Drivers & Techniques (1st ed.). Thomas Erl, Wajid Khattak, and Paul Buhler. Prentice Hall Press, Upper Saddle River, NJ, USA. 2016. 2. Big Data, Principles and Best Practices of Scalable Realtime Data Systems, Nathan Marz and James Warren, Manning Publications 2015. 3. Hadoop: The Definitive Guide, Tom White, O'Reilly, 2015.
----	---	--

23	Değerlendirme
----	---------------

YARIYIL İÇİ ÇALIŞMALARI	SAYISI	KATKI YÜZDESİ
Ara Sınav	1	40.00
Kısa Sınav	0	0.00
Ödev	0	0.00
Yıl Sonu Sınavı	1	60.00
Toplam	2	100.00
Yıl içi çalışmalarının Başarıya Oranı		40.00
Finalin Başarıya Oranı		60.00
Toplam		100.00
Kullanılan Ölçme ve Değerlendirme Yaklaşımları		Ölçme ve değerlendirme, Bursa Uludağ Üniversitesi Önlisans ve Lisans Eğitim Öğretim Yönetmeliği ilkelerine göre yapılmaktadır.

24	AKTS / İŞ YÜKÜ TABLOSU
----	------------------------

ETKİNLİK	SAYISI	Süresi (Saat)	Toplam İş Yüğü (Saat)
Teorik Dersler	14	3.00	42.00
Uygulamalı Dersler	0	0.00	0.00
Sınıf Dışı Ders Çalışma Süresi (Ön çalışma, pekiştirme)	0	0.00	0.00
Ödevler	0	0.00	0.00
Projeler	0	0.00	0.00
Arazi Çalışmaları	0	0.00	0.00
Arasınavlar	1	60.00	60.00
Diğer	0	0.00	0.00
Yarıyıl Sonu Sınavı	1	75.00	75.00
Toplam İş Yüğü			177.00
Toplam İş Yüğü / 30 saat			5.90
Dersin AKTS Kredisi			6.00

25	PROGRAM YETERLİLİKLERİ İLE DERS ÖĞRETİM KAZANIMLARI İLİŞKİSİ TABLOSU															
	PY1	PY2	PY3	PY4	PY5	PY6	PY7	PY8	PY9	PY10	PY11	PY12	PY13	PY14	PY15	PY16
ÖK1	0	3	3	3	3	3	2	2	1	0	1	0	0	0	0	0
ÖK2	0	3	3	3	3	3	2	2	1	0	1	0	0	0	0	0
ÖK3	0	3	3	3	3	3	2	2	1	0	1	0	0	0	0	0
ÖK4	0	3	3	3	3	3	2	2	1	0	1	0	0	0	0	0

ÖK5	0	3	3	3	3	3	2	2	1	0	1	0	0	0	0	0
ÖK: Öğrenme kazanımlar PY: Program yeterlilikleri																
Katkı Düzeyi:	1 çok düşük		2 Düşük		3 Orta		4 Yüksek		5 Çok Yüksek							