# DERS TANITIM ve UYGULAMA BİLGİLERİ

| Dersin Adı | Kodu | Yarıyıl | T+U+L (saat/hafta) | Türü (Z / S) | Yerel Kredi | AKTS |
|---|---|---|---|---|---|---|
| Veri Bilimi ve Analitiği | CE 515 | Bahar | 03+00+00 | Zorunlu | 3 | 7.5 |

| | |
|---|---|
| Akademik Birim: | MDBF, Bilgisayar Mühendisliği Bölümü |
| Öğrenim Türü: | Örgün Eğitim |
| Ön Koşullar | yok |
| Öğrenim Dili: | İngilizce |
| Dersin Düzeyi: | Yüksek Lisans |
| Dersin Koordinatörü: | Taner ARSAN |
| Dersin Amacı: | Bu dersin amacı, öğrencilerin büyük veri ve analitik gerektiren projelerde etkin bir şekilde görev alabilmeleri için gerekli bilgi ve pratik deneyime sahip olmalarını sağlamaktır. |
| Dersin İçeriği: | Veri Bilim ve Analitiği dersi, büyük veri ve analitik uygulamaları konularında öğrencileri yetiştirmek üzerine kurulmuştur. Ders, Büyük veri ve veri analizi yaşam döngüsü içerisinde yer alan iş odaklı sorunları çözmek için gerekli ve yeterli bilgi sağlar. Ders içeriği ayrıca, temel ve ileri düzey analitik yöntemler ile MapReduce ve Hadoop'un dahil olduğu büyük veri analitiği teknolojisi ve araçlarını öğrenebimek için gerekli altyapıyı da sağlamaktadır. |
| Dersin Öğrenme Çıktıları (ÖÇ): | • **1-** Dersi alan öğrencilerin büyük veri ve diğer analitik projelerine hemen etkin bir şekilde katılmak için bilgi ve pratik deneyime sahip olmalarını sağlamak. |
| Dersin Öğrenme Yöntem ve Teknikleri | sınıf içi ders |

## HAFTALIK PROGRAM

| Hafta | Konular | Ön Hazırlık |
|---|---|---|
| 1 | Giriş ve Ders İçeriğinin paylaşılması, Modül 1: Büyük Veri ve Analitiğe Giriş | Büyük Veriye Genel Bakış, Analitik Uygulamalarının Durumu |
| 2 | Veri Bilimcisi, Endüstri Sektörlerinde Büyük Veri Analitiği | Bu modül, tipik bir analitik yaşam döngüsünün çeşitli aşamalarını açıklamaya odaklanır - keşif, veri hazırlama, model planlama, model oluşturma, sonuçların ve bulguların iletilmesi ve operasyonel hale getirilmesi. Bu modül ayrıca yaşam döngüsünün her aşamasında meydana gelen kritik faaliyetleri detaylandırır. |
| 3 | Modül 2: Veri Analitiği Yaşam Döngüsü | Keşif, Veri Hazırlama, Model Planlama, Model Oluşturma, Sonuçları Tartışma, Operasyonelleştirme |
| 4 | Modül 3: R / Python Kullanarak Temel Veri Analitik Yöntemlerinin İncelenmesi. Python temelleri veya R Grafik Kullanıcı Arayüzünü kullanma, Genel Bakış: R'ye Veri Alma (ve R'den), R'de Kullanılan Veri Tipleri, Temel R İşlemleri, Temel İstatistikler, Genel Fonksiyonlar | Python temelleri veya R Grafik Kullanıcı Arayüzünü kullanma, Genel Bakış: R'ye Veri Alma (ve R'den), R'de Kullanılan Veri Tipleri, Temel R İşlemleri, Temel İstatistikler, Genel Fonksiyonlar |
| 5 | Modül 4: Gelişmiş Analitik – Teori ve Yöntemlere Genel Bakış: K-means Clustering (Lesson 1) Association Rules (Lesson 2) | • K-means clustering • Association Rules • Linear Regression • Logistic Regression • Naïve Bayesian Classifiers • Decision Trees • Time Series Analysis • Text Analytics |
| 6 | Linear Regression (Lesson 3) Logistic Regression (Lesson 4) | |
| 7 | Naïve Bayesian Classifiers (Lesson 5) | |

| | | |
|---|---|---|
| | Decision Trees (Lesson 6) | |
| 8 | Time Series Analysis (Lesson 7) Text Analytics (Lesson 8) | |
| 9 | Midterm Exam | |
| 10 | Module 5: Advanced Analytics - Technology and Tools Lesson 1: Analytics for Unstructured Data - MapReduce and Hadoop Lesson 2: The Hadoop Ecosystem | |
| 11 | Lesson 3: In-database Analytics SQL essentials Lesson 4: Advanced SQL and MADlib | |
| 12 | Module 6 – The Endgame, or Putting it All Together Lesson 1: Operationalizing an Analytics Project Lesson 2: Creating the Final Deliverables | |
| 13 | Module 6 – The Endgame, or Putting it All Together Lesson 3: Data Visualization | Veri görselleştirme araçları anketi • Sponsorlar ve analistler için farklı görselleştirmeler oluşturma • Kilit noktalarınızı destekleyecek görseller geliştirme • Bir grafik veya görselleştirme nasıl temizlenir |
| 14 | Preparing Project Paper and Presentations | |

Kadir Has Üniversitesi'nde bir dönem 14 haftadır, 15. ve 16. hafta sınav haftalarıdır.

## ZORUNLU ve ÖNERİLEN OKUMALAR

Bütün kaynaklar EMC2 Akademik İşbirliği kapsamında EMC2 tarafından sağlanmaktadır.

## DİĞER KAYNAKLAR

Ele alınan teorilerin uygulamasını içerecek şekilde ek okumalara başvurulabilir. Örneğin;

* Introduction to Machine Learning with Python: A Guide for Data Scientists, Andreas C. Müller and Sarah Guido (for beginners).

## DEĞERLENDİRME SİSTEMİ

| **Yarıyıl İçi Çalışmaları** | **Sayı** | **Katkı Payı (%)** |
|---|---|---|
| Proje | 1 | 30 |
| Sunum/Jüri | 1 | 40 |
| Ara Sınavlar | 1 | 30 |
| **Total:** | **3** | **100** |

## İŞ YÜKÜ HESAPLAMASI

| **Etkinlikler** | **Sayısı** | **Süresi (saat)** | **Toplam İş Yükü (saat)** |
|---|---|---|---|

| | | | |
|---|---|---|---|
| Ders Saati | 14 | 3 | 42 |
| Uygulama | 2 | 3 | 6 |
| Proje | 1 | 49.5 | 49.5 |
| Sunum/Jüriye Hazırlık | 1 | 30 | 30 |
| Sunum | 1 | 30 | 30 |
| Ara Sınavlar | 1 | 30 | 30 |
| **Toplam İş Yükü (saat):** | | | **187.5** |

1 AKTS = 25 saatlik iş yükü

## PROGRAM YETERLİLİKLERİ (PY) ve ÖĞRENME ÇIKTILARI (ÖÇ) İLİŞKİSİ

| # | PY1 | PY2 | PY3 | PY4 | PY5 | PY6 | PY7 | PY8 | PY9 | PY10 | PY11 | PY12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OC1 | | | | | | | | | | | | |

**Katkı Düzeyi:** 1 Düşük, 2 Orta, 3 Yüksek