DERS TANITIM ve UYGULAMA BİLGİLERİ

Dersin Adı	Kodu	Yarıyıl	T+U+L (saat/hafta)	Türü (Z / S)	Yerel Kredi	AKTS
Doğal Dil İşleme ve Uygulamaları	MIS 458	Bahar	03+00+00	Seçmeli	3	5

Akademik Birim:	Yönetim Bilişim Sistemleri
Öğrenim Türü:	Örgün Eğitim
Ön Koşullar	Bilgisayar Programlama Dilleri (C, Python veya Java)
Öğrenim Dili:	İngilizce
Dersin Düzeyi:	Lisans
Dersin Koordinatörü:	- -
Dersin Amacı:	Bu ders, gerek kuramsal gerekse uygulama seviyesinde metinsel veri türlerini analiz etme konusunda genel ve uygulama alanlarına özel bilgi sağlamayı amaçlamaktadır. Öğrenciler, söz konusu metinden bilgi ve iş zekasını üretebilmek için metin madenciliği tekniklerini kullanarak bir problem üzerinde pratik olarak çalışacaklardır. Bu amaçla herhangi bir programlama dilini kullanabilirler ancak dersin teorik kısmında öğretilen teorik tekniklerin uygulanması için dersler sırasında Python kullanılacaktır. Dönem sonunda öğrenciler metinsel verileri ön işleme tabi tutabilmeli ve bunları yazım düzeltme, POS etiketleme, metin sınıflandırma, semantik analiz, duygu analizi vb. gibi ana göreve hazırlayabilmelidir.
Dersin İçeriği:	Ders, metin ön işleme ve normalleştirme, düzenli ifadeler, yazım düzeltme, POS etiketleme, duygu analizi, metin sınıflandırma ve naif Bayes, sınıflandırma algoritmalarının sınıflandırılması ve değerlendirilmesi, kelime semantiği ve kelime gömme (word embeddings) gibi konuları kapsayacaktır.
Dersin Öğrenme Çıktıları (ÖÇ):	• **1-** Metinsel verileri anlamak için ön işleme ve analiz etme becerisi. • **2-** Bu dersin yazım düzeltme, metin sınıflandırma ve POS etiketleme gibi NLP tekniklerini Python gibi bir programlama dili kullanarak uygulayabilme becerisi. • **3-** Metinsel veriler üzerinde sınıflandırma ve kümeleme yapmak için Python gibi bir programlama dilini ve kütüphanelerini kullanabilme.
Dersin Öğrenme Yöntem ve Teknikleri	Sınıf içi öğrenme. Programlama ödevleri bireysel olarak yapılacaktır. Ancak gerçek dünyadaki bir soruna yönelik bir çözümün tasarlanması ve uygulanması grup çalışması olarak yapılacaktır. Bu derste öğretilen tüm teorik kavramlar kodlama bölümünde uygulanacaktır.

HAFTALIK PROGRAM

Hafta	Konular	Ön Hazırlık
1	NLP ve Python Giriş	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
2	Text ön işleme and Regular ifadeler	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
3	Text (ön) işleme and Regular ifadeler	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
4	Minimum Edit Distance	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
5	N-gram dil model	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
6	Text sınıflandırma and Naïve Bayes	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
7	Yazın düzeltme	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
8	POS etiketleme	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
9	Vector semantics	Ders notları, slaytlar ve diğer çevrim içi kaynaklar

10	Duygu (Sentiment) Analizi	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
11	Affective analysis and Emotions	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
12	Introduction to Word embeddings	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
13	Word Semantic and ilişkiler	Ders notları, slaytlar ve diğer çevrim içi kaynaklar
14	Proje Sunumları	Ders notları, slaytlar ve diğer çevrim içi kaynaklar

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

ZORUNLU ve ÖNERİLEN OKUMALAR

1) Speech and Language Processing, Jurafsky and Martin, Draft chapters in progress, 2021
2) Foundations of Statistical Natural Language Processing, Christopher D. Manning and Henrich Schütze, The MIT Press, Cambridge, Masachusetts , London, England, 2000

DİĞER KAYNAKLAR

Ders Notları, Stanford NLP parser vb NLP araçları

DEĞERLENDİRME SİSTEMİ

Yarıyıl İçi Çalışmaları	Sayı	Katkı Payı (%)
Katılım	1	20
Proje	2	60
Sunum/Jüri	1	20
Total:	**4**	**100**

İŞ YÜKÜ HESAPLAMASI

Etkinlikler	Sayısı	Süresi (saat)	Toplam İş Yükü (saat)
Ders Saati	14	5	70
Proje	2	20	40
Sunum/Jüriye Hazırlık	1	15	15
Toplam İş Yükü (saat):			**125**

1 AKTS = 25 saatlik iş yükü

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

#	PY1	PY2	PY3	PY4	PY5	PY6	PY7	PY8	PY9
OC1									
OC2									
OC3									

Katkı Düzeyi: 1 Düşük, 2 Orta, 3 Yüksek