



# MARMARA ÜNİVERSİTESİ MÜHENDİSLİK FAKÜLTESİ

2022-2023 Bahar

CSE4078 Introduction to Natural Language Processing

## DERS TANITIM FORMU

Dersin Açan Bölüm	Bilgisayar Mühendisliği Bölümü	Teknik Seçmeli						
Ders Kodu	CSE4078							
Ders Adı	Introduction to Natural Language Processing (Doğal Dil İşlemeye Giriş)							
Öğretim Dili	İngilizce							
AKTS	5							
Haftalık Ders Saati	Teorik(T): 3	Uygulama(U): - Laboratuvar(L): -						
Ön Koşul(lar)								
Öğretim Elemanı	İsim	Doç. Dr. Murat Can Ganiz						
	E-mail	murat.ganiz@marmara.edu.tr						
Ders Materyalleri	Zorunlu	Speech and Language Processing (3rd ed. 2023 January Draft ) by Dan Jurafsky and James H. Martin						
	Önerilen	Chris Manning and Hinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA: May 1999.						
Dersin Amacı	Bu ders, öğrencilere doğal dil işleme (NLP) alanında güçlü bir temel sağlamayı ve büyük dil modelleri gibi güncel gelişmeleri tanıtmayı amaçlamaktadır. Dil modelleme, sözdizimi, anlambilim gibi temel kavramlar ve NLP'de kullanılan teknikler öğretilecektir. Ayrıca metin sınıflandırma, bilgi çıkarma, özetleme gibi alt problemlere de değinilecektir. Dersin sonunda, öğrenciler bu teknikleri gerçek dünya NLP görevlerinde uygulayabilecek yeteneklere sahip olacaklardır. Ayrıca, öğrenciler NLP alanındaki güncel teknikleri keşfedecek ve dönem projelerinde gerçek dünya veri setleriyle çalışma fırsatı bulacaklardır.							
Ders İçeriği	Dilbilimsel Temeller ve Korpus Tabanlı Çalışma, Collocations, n-gram dil modelleri, Naive Bayes ve Duygu Sınıflandırması, Lojistik Regresyon, Gradyan İnişi ve Geriye Yayılım, Vektör Semantiği ve Kelime Gömmeleri, Sinir Ağları ve Sinirsel Dil Modelleri, Sözcük Türleri ve İsimlendirilmiş Varlıklar için Sıralama Etiketleme, Yinelemeli Sinir Ağları ve Uzun Kısa Süreli Bellek (LSTM), Kodlayıcı-Çözücü Modelleri, Kendi Dikkat Ağı, Transformer mimarileri, Büyük Dil Modelleri ve Sohbet Tabanlı Yapay Zeka, NLP'nin Aşağı Akış Uygulamaları							
Öğrenim Çıktıları	ÖÇ1	Doğal Dil İşleme (NLP) alanındaki zorlukları, insan dilinin karmaşıklığını, temel kavramları ve prensipleri anlamak						
	ÖÇ2	n-gram dil modelleri ve kolokasyon analizi gibi çeşitli teknikleri uygulayarak dil desenlerini ve yapısını analiz etmek ve modellemek						
	ÖÇ3	DDİ alanında kullanılan temel makine öğrenimi algoritmaları ile gradyan inişi ve geriye yayılım algoritmalarını anlamak, uygulamak, gözetimli modellerin değerlendirme ölçütlerini hesaplayabilmek.						
	ÖÇ4	Vektör semantiği ve kelime gömme gibi DDİ'deki ileri düzey konuları keşfetmek						
	ÖÇ5	Yinelemeli sinir ağları, dönüştürücü mimarileri ve büyük dil modelleri gibi DDİ'deki güncel teknikler ve sohbet tabanlı yapay zeka alanında temel bilgi sahibi olmak.						
Program Çıktıları		ÖÇ1	ÖÇ2	ÖÇ3	ÖÇ4	ÖÇ5		
PÇ1	Matematik, fen bilimleri (a) ve bilgisayar mühendisliğine (b) özgü konularda yeterli bilgi birikimi (1); bu alanlardaki kuramsal ve uygulamalı bilgileri, karmaşık mühendislik problemlerinde kullanabilme becerisi (2).	1b	1b	1b	1b	1b		
PÇ4	Mühendislik uygulamalarında karşılaşılan karmaşık problemlerin analizi ve çözümü için gerekli olan modern teknik ve araçları geliştirme (a), seçme ve kullanma (b) becerisi (1); bilişim teknolojilerini etkin bir şekilde kullanma becerisi (2).		1b	1b	1b			
PÇ5	Karmaşık mühendislik problemlerinin veya bilgisayar mühendisliği araştırma konularının incelenmesi için deney tasarlama (a), deney yapma, veri toplama (b), sonuçları analiz etme ve yorumlama becerisi (c).		abc	abc				
Öğretilen Konular, Konuların Öğrenim Çıktılarına Katkıları ve Öğrenim Değerlendirme Metotları	No	Hafta	Konular	ÖÇ1	ÖÇ2	ÖÇ3	ÖÇ4	ÖÇ5
	K1	1	Doğal Dil İşlemeye giriş					
	K2	2	Dilbilgisel Temeller ve Korpus Tabanlı Çalışmalar					
	K3	3	Kolokasyonlar					
	K4	4	n-gram dil modelleri					
	K5	5	Naive Bayes ve Duygu Sınıflandırması					
	K6	6	Lojistik Regresyon, Gradyan İnişi ve Geri Yayılım					
	K7	7	Vektör Semantiği ve Kelime Gömme					
	K8	8	Sinir Ağları ve Sinirsel Dil Modelleri					
	K9	9	Sekans etiketleme uygulamaları (POS ve NER)					
K10	10	Yinelemeli Sinir Ağları ve Uzun Kısa Süreli Bellekler (LSTM)						

	K11	11	Kodlayıcı-Çözücü Modeller							
	K12	12	Öz-Dikkat Ağları, Dönüştürücü Mimarileri, Ön-eğitim, İnce-ayarlı ile aktarım öğrenme							
	K13	13	Büyük Dil Modelleri ve Sohbet Tabanlı Yapay Zeka							
	K14	14	Proje Sunumları							
<b>Öğrenim Değerlendirme Metotları, Ders Notuna Etki Ağırlıkları ve Uygulama Kuralları</b>	<b>No</b>	<b>Tür</b>	<b>Ağırlık</b>	<b>Uygulama Kuralı</b>		<b>Telafi Kuralı</b>				
	MF	Sınav	%60	Sınavlar, kitaplar ve tüm ders malzemeleri kapalı yapılır. Sınavlar sırasında hiçbir hesaplama ve iletişim aracına izin verilmez.		Üniversite prosedürüne uygun bir sağlık raporu veya görevlendirme yazısı getirildiğinde en fazla bir ara sınav için telafi sınavı verilir.				
	Q	Kısa Sınav (Quiz)	-	-		-				
	H	Ödev	-	-		-				
	P	Proje	%40	Disiplin içi takımlar şeklinde DDİ konularının gerçek hayat veya akademik veri kümeleri üzerinde uygulaması yapılır		-				
<b>TOPLAM</b>			%100							
<b>Harf Notu Belirleme Metodu</b>	<ul style="list-style-type: none"> <li>Bir ara sınav, bir final sınavı, ödev ve proje notlandırması yapılarak harf notu verilecektir.</li> <li>Öğrencilerin toplam ortalama notlarına göre çan eğrisi referans alınarak veya öğretim üyesi takdirinde katalog yapılarak harf notu belirlenecektir.</li> <li>Dersten geçmek için final notunun ve toplam ortalama notunun en az 35 olması gereklidir.</li> <li>Marmara Üniversitesi Lisans yönetmeliğine göre lisans derslerinde final sınavının puanı %40'tır.</li> </ul>									
	Değerlendirme		Ara Sınav	Proje	Ödev	Final	TOPLAM			
Puan		20	40	-	40	100				
<b>Öğretim Metotları ve Harcanan Tahmini Saat</b>	<b>Öğretim elemanı tarafından uygulanan süre</b>									
	<b>No</b>	<b>Tür</b>	<b>Açıklama</b>						<b>Saat</b>	
	1	Sınıf Dersi	Ders konuları sunum ile ve tahtaya yazarak anlatılır. Öğrenimi pekiştirmek için ders sırasında örnek sorular çözülür. Birinci ara sınav sınav haftasında, ikinci ara sınav ise sınıfta yapılır.						14x3=42	
	2	Problem Dersi	Problem setleri tahtaya yazarak çözülür.							
	3	Laboratuvar	Laboratuvarda deney yapılır ve rapor hazırlanır, ya da derste öğrenilen teorik konular ile ilgili bilgisayar ortamında uygulama yapılır.							
	4	Etkileşimli Ders	Ders sırasında öğrencilere belli sorular yöneltilir, cevapları tahmin etmeleri teşvik edilir. (Not: Akran eğitimi bu kategoriye giriyor.)						5	
	5	Saha Çalışması	Okul dışı etkinliklere katılır.							
	6	Ara Sınav	Ara sınav haftasında ara sınav uygulanır.						2	
	7	Final	Final haftalarında final sınavı uygulanır.						2	
	<b>Öğrencinin ayırması beklenen tahmini süre</b>									
	8	Proje	Proje olarak verilen problem konusu araştırılır, tasarım ve gerçekleştirme yapılır ve rapor hazırlanır.						40	
	9	Ödev	Ödev olarak verilen soruların cevapları hazırlanır.							
	10	Derse Hazırlık	Yeni konular sınıfta işlenmeden önce öğrenilir (ders materyallerinden).							
11	Ders Tekrarı	Sınavlar ve ödevlere hazırlık için konular tekrar edilir (ders materyallerinden).						36		
12	Ofis Saati	Öğretim elemanı veya ders asistanına birebir sorular sorulur.						2		
<b>TOPLAM</b>									129	
<b>Akademik Dürüstlük</b>	<p>Akademik dürüstlüğü ihlali; intihal yapmayı, kopya çekmeyi ve kopya çekmeye teşebbüs etmeyi, sahte bilgi veya alıntı göstermeyi, referans göstermeden daha önce kendisi tarafından yapılan bir çalışmayı kullanmayı, akademik çalışmayı başka öğrenciler yerine yapmayı, haksız avantaj elde etmeye yönelik davranışlarda bulunmayı, sahte belge düzenlemeyi ve kullanmayı, sınavları izinsiz elde etmeyi, başkasının yerine sınava girmeyi ya da başkasını kendi yerine sınava sokmayı içermekle birlikte, bu eylemlerle sınırlı değildir.</p> <p>Akademik dürüstlüğü herhangi bir biçimde ihlal edilmesi durumunda Yükseköğretim Kurumları Öğrenci Disiplin Yönetmeliği hükümlerince işlem uygulanır.</p>									