

## GENEL TANIM / GENERAL DESCRIPTION

Ders Adı / Course Name	Search Engines / Search Engines	
Ders Kodu / Course Code	9105055462013	
Ders Türü / Course Type		
Ders Seviyesi / Course Level	Second Cycle / Second Cycle	
Ders Akts Kredi / ECTS	8.00	
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	1	
Öğretim Sistemi / Teaching System	Face to Face / Face to Face	
Eğitim Dili / Education Language	Turkish / Turkish	
Ön Koşulu Olan Ders(ler) / Precondition Courses	Yok	None
Amacı / Purpose	Bu dersin amacı öğrencilere arama motorlarına ilişkin temel teknolojilerin ve araçların tanıtılmasıdır. Google, Yahoo, MSN Search gibi genel kullanımı olan büyük ölçekli arama motorları incelenecek ve kıyaslamaları yapılacaktır.	The objective of the course is to give students an introduction to basic technologies and tools associated with the search engines. The large-scale search engines in general-use, such as Google, Yahoo, and MSN Search, will be examined and the comparisons amongs them will be made.
İçeriği / Content	Bu dersin kapsamı: arama motorlarının mimarisi; web örümceklerinin geliştirilmesi; yazılı metinlerin işlenmesi, indekslenmesi, organizasyonu ve sınıflanması; arama motorlarının performans değerlendirilmesi; web madenciliği konularına yoğunlaşmıştır. Ayrıca pratik uygulamalar için Piton dili de anlatılacaktır.	This course will focus on topics such as the architecture of search engines, web crawlers' development; processing, indexing, organizing and classifying of texts; performance evaluation of search engines; and web mining. Python programming language is also explained in this course for the practical applications.
Önerilen Diğer Hususlar / Recommended Other Considerations	Yok	Competency in reading and understanding English and application of statistical methods
Staj Durumu / Internship Status	Yok	None
Kitabı / Malzemesi / Önerilen Kaynaklar / Books / Materials / Recommended Reading	Ders kitabı: Modern Information Retrieval R. Baeza-Yates ve B. Ribeiro-Yeto, 1999 Diğer önemli kaynaklar: Information Retrieval: A survey, E. Greengrass,2000, <a href="http://www.cs.umbc.edu/research/cadip/readings/IR.report.120600.book.pdf">http://www.cs.umbc.edu/research/cadip/readings/IR.report.120600.book.pdf</a> Information retrieval, C.J. van Rijsbergen, <a href="http://www.dcs.gla.ac.uk/Keith/Preface.html">http://www.dcs.gla.ac.uk/Keith/Preface.html</a> Precision and Recall, S.M. Shafi ve R.A. Rather <a href="http://www.webology.ir/2005/v2n2/a12.html">http://www.webology.ir/2005/v2n2/a12.html</a>	TEXTBOOK: Modern Information Retrieval R. Baeza-Yates ve B. Ribeiro-Yeto, 1999 OTHERS: Information Retrieval: A survey, E. Greengrass,2000, <a href="http://www.cs.umbc.edu/research/cadip/readings/IR.report.120600.book.pdf">http://www.cs.umbc.edu/research/cadip/readings/IR.report.120600.book.pdf</a> Information retrieval, C.J. van Rijsbergen, <a href="http://www.dcs.gla.ac.uk/Keith/Preface.html">http://www.dcs.gla.ac.uk/Keith/Preface.html</a> Precision and Recall, S.M. Shafi ve R.A. Rather <a href="http://www.webology.ir/2005/v2n2/a12.html">http://www.webology.ir/2005/v2n2/a12.html</a>

## ÖĞRENME ÇIKTILARI / LEARNING OUTCOMES

1	Arama motorlarının genel yapısı ve çalışma sistemini anlama	To be able to analyze and understand the architecture of search engines.
2	BGG sistemlerinin bileşenlerini analiz edebilme	To be able to analyze the components of information retrieval (IR) systems.
3	Belge temsilinde kullanılan terim ağırlıklandırma süreci için gerekli olan unsurları kavrayabilme	Understand the factors considered for term weighing process in document representation.
4	Sayfa önem derecesinin belirlenmesinde kullanılan ölçütleri kavrayabilme	Understand page ranking measurements
5	Dizinleme yapılarını kavrayabilme	Understand indexing structures
6	Temel metin işleme algoritmalarını yazılımsal olarak gerçekleştirebilme	To be able to code basic text processing algorithms
7	Belge-sorgu alakası hesaplamasını kavrayabilme	Understand document-query relevance computations
8	Arama motorlarının değerlendirmesini yapabilme	To be able to assess search engines
9	Belge kümeleme yöntemlerini anlama	Understand text clustering methods
10	Python dilinde web örümceđi yazabilme	To be able to code search engine in Python.

## HAFTALIK DERS İÇERİĐİ / DETAILED COURSE OUTLINE

Hafta / Week					
1	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Giriş: Arama Motorlarının genel mimarisi, mevcut sistemler ve karşılaştırmaları Introduction: General architecture of Search Engines, available systems and their comparisons				
2	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Web Örümcekleri Web Örümceklerinin çalışması. Piton dilinde Web örümceğinin yazılması için metin işleme komutları Web Crawlers: Operation of Web Crawlers, their development by the text processing commands in Python.	okuma ve tartışma reading and discussion			
3	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Yazılı Metinlerin İşlenmesi Yazılı metinlerde gövdeleme, durma kelimeleri, tamlamaların tespiti, noktalama işaretlerinin çıkarılması, cümle sonu tespiti. Text Processing: elimination of stop-words and punctuation marks, stemming, compound words and end of sentence detection.	Programlama:Piton ile metin işleme programlama ödevi programming: text processing programming assignment with Python.			
4	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Terim Ağırlıklandırma Vektör uzayı modeli.Terim ağırlıklandırma yöntemleri: ikili, terim frekansı, ters belge frekansı Term Weighting: Vector space model, term weighting models: Boolean, term frequency and inverse document frequency.	Programlama:tdm aracı ile verilen bir belgenin terimxbelge matrisini çıkarma programming: obtaining termxdocument matrix of a given document by tdm tool			
5	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Sorguların oluşturulması Anahtar kelime tabanlı sorgulama, şablon eşleştirme ve yapısal sorgular. Query Languages: Keyword-based querying, pattern matching and structural queries.	araştırma ve raporlama research and reporting			

	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
6	İndeksleme ve Arama: Ters dosyalar, Tek geçişli ve iki geçişli indeksleme yöntemleri	Programlama:Verilen bir belge kümesinin ters indis dosyasının oluşturulması			
	Indexing and searching: Introduction to inverted files, single pass and two pass indexing methods.	programming: creating inverted file of a given document collection			
7	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Belge-Sorgu Alaka Sıralaması metin ve sorgu arasındaki benzerlik ölçülmesi	araştırma ve raporlama			
	Relevance ranking of documents on queries: Measuring the similarity between a text and a query.	research and reporting			
8	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Arama Motorlarının Değerlendirilmesi: Anma, Duyarlık ve F ölçütleri, alternatif başarımlar ölçütleri.	araştırma ve raporlama			
	Search Engines Evaluation: Recall and precision measures, alternative performance metrics.	research and reporting			
9	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Ara sınav				
	Midterm Exam				
10	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Bağ Yapılarının analizi Arama analizinde bağ yapılarının kullanımı	araştırma ve raporlama			
	Link Structure Analysis: Usage of link structures in search analysis.	research and reporting			
11	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
	Spam ve kopyaların tespiti Spamlerin tespiti, , filtrelemeler	okuma ve tartışma			
	Spam and replication detection: Spam detection and filtering.	reading and discussion			

	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
12	Hesaplamalı Reklamcılık Reklamların önceliklendirilmesi ve ücretlendirilmesi	araştırma ve raporlama			
	Computational Advertising: Prioritization and pricing of advertisements.	research and reporting			
	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
13	Belge Kümeleme Yöntemleri Bayes, K-nn, Gizli Anlamsal İndisleme	Programlama			
	Document Clustering Methods: Bayes, K-nn, Latent Semantic Indexing	programming			
	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
14	Web'de veri madenciliği Sosyal ağlardan bilgi çıkarımında veri madenciliği yöntemlerinin uygulanması	Okuma ve tartışma			
	Data Mining in WEB: Data mining applications for knowledge extraction from Social Networks.	reading and discussion			
	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
15	Öğrenci Proje Sunumları				
	Students Project Presentations				
	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary
16	Final sınavı				
	Final Exam				
	Teorik Dersler / Theoretical	Uygulama	Lab	Öğretim Yöntem ve Teknikleri/Teaching Methods Techniques	Ön Hazırlık / Preliminary

## 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)
Rapor Hazırlama / Report Preparation	8	8.00	64.00
Ara Sınav İçin Bireysel Çalışma / Individual Study for Mid term Examination	1	20.00	20.00
Proje Hazırlama / Project Preparation	1	25.00	25.00
Proje Sunma / Project Presentation	1	1.00	1.00
Okuma / Reading	5	6.00	30.00
Ara Sınav / Midterm Examination	1	3.00	3.00
Final Sınavı / Final Examination	1	3.00	3.00
Uygulama/Pratik / Practice	6	5.00	30.00
Final Sınavı için Bireysel Çalışma / Individual Study for Final Examination	1	25.00	25.00
Derse Katılım / Attending Lectures	13	3.00	39.00
<b>Toplam / Total:</b>	<b>38</b>	<b>99.00</b>	<b>240.00</b>
Dersin AKTS Kredisi = Toplam İş Yüğü (Saat) / 30.00 (Saat/AKTS) = 240.00/30.00 = 8.00 ~ 8.00 / Course ECTS Credit = Total Workload (Hour) / 30.00 (Hour / ECTS) = 240.00 / 30.00 = 8.00 ~ 8.00			

## 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.Arama motorlarının genel yapısı ve çalışma sistemini anlama / To be able to analyze and understand the architecture of search engines.							3
2.BGG sistemlerinin bileşenlerini analiz edebilme / To be able to analyze the components of information retrieval (IR) systems.			3				5

3.Belge temsilinde kullanılan terim ağırlıklandırma süreci için gerekli olan unsurları kavrayabilme / Understand the factors considered for term weighing process in document representation.		5	4			3	
4.Sayfa önem derecesinin belirlenmesinde kullanılan ölçütleri kavrayabilme / Understand page ranking measurements							5
5.Dizinleme yapılarını kavrayabilme / Understand indexing structures			4				5
6.Temel metin işleme algoritmalarını yazılımsal olarak gerçekleştirebilme / To be able to code basic text processing algorithms	3						
7.Belge-sorgu alaka hesaplamasını kavrayabilme / Understand document-query relevance computations			4				5
8.Arama motorlarının değerlendirmesini yapabilme / To be able to asses search engines		5					
9.Belge kümeleme yöntemlerini anlama / Understand text clustering methods							
10.Piton dilinde web örümceği yazabilme / To be able to code search engine in Python.							

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