

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

TÜRKÇE ARAMA MOTORU SONUCU KÜMELEME ÇALIŞMALARI

BURAK DURAL

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
DOÇ. DR. BANU DİRİ**

İSTANBUL, 2013

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TÜRKÇE ARAMA MOTORU SONUCU KÜMELEME ÇALIŞMALARI

Burak DURAL tarafından hazırlanan tez çalışması 18.02.2013 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Banu DİRİ

Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Banu DİRİ

Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. M. Fatih AMASYALI

Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Cüneyd TANTUĞ

İstanbul Teknik Üniversitesi

ÖNSÖZ

Öncelikle tezin yazımı aşamasında hiçbir desteğini esirgemeyen, sabırla bana yardımcı olan Sayın hocam Doç. Dr. Banu DİRİ'ye, bu süreçte ve tüm eğitim hayatım boyunca her zaman yanımda olan ve bana destek veren aileme teşekkürlerimi bir borç bilirim.

Ocak, 2013

Burak DURAL

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ.....	ix
ÇİZELGE LİSTESİ	x
ÖZET	xii
ABSTRACT	xiv
BÖLÜM 1	
GİRİŞ.....	1
1.1. Literatür Özeti	3
1.2. Tezin Amacı	3
1.3. Hipotez	4
BÖLÜM 2	
ÖNCEKİ ÇALIŞMALAR	6
2.1. Scatter/Gather	6
2.2. Son Ek Ağacı Kümeleme.....	7
2.3. Son Ek Ağacı Kümeleme Üzerinde Yapılan Geliştirmeler	7
2.4. SHOC: Semantic Hiearchical Online Clustering.....	9
2.5. Lingo.....	10
2.6. Wordnet Kullanılan Çalışmalar	10
BÖLÜM 3	
GENEL BİLGİLER	12
3.1. İnternet Arama Motoru	12
3.1.1 Arama Motorları Tarihçesi	13
3.1.2 Arama Motorlarının Yapısı	15
3.2. Kümeleme	17

3.2.1 Doküman Kümeleme	18
3.2.1.1 Hiyerarşik Kümeleme	18
3.2.1.2 Bölümlemeli Kümeleme	19
3.2.2 Arama Sonucu Kümeleme	19
3.2.3 Kümeleme Arayüzü	21
3.3 Doküman Benzerliği Hesaplanması	22
3.3.1 TF-IDF	22
3.3.2 Kosinüs Benzerliği	23
3.4 Son Ek Ağacı Kümeleme.....	24
3.5 Zemberek	28
3.6 Carrot2	29
3.7 Google Arama Motoru	31
3.7.1 Google Custom Search API	31
3.8 Doküman Kümelemede Ön İşleme Adımları	32
BÖLÜM 4	
ÖNERİLEN YÖNTEM	33
4.1 Kullanılan Veriler.....	33
4.2 Son Ek Ağacı Kümeleme.....	34
4.2.1 Ön İşleme Adımları.....	34
4.2.2 Klasik SAK - KSAK.....	36
4.2.3 Geliştirilmiş SAK - GSAK	38
4.3 Distance Based Clustering.....	39
4.3.1 Ön İşleme Adımları.....	39
4.3.2 Yöntem.....	40
4.4 Melez GSAK Yöntemi (M-GSAK).....	42
BÖLÜM 5	
DEĞERLENDİRME VE DENEYLER	44
5.1 Değerlendirme Yöntemleri	44
5.2 Yapılan Deneyler	45
5.2.1 Balık Sorgusu Sonuçları	46
5.2.1.1 Klasik SAK	47
5.2.1.2 Geliştirilmiş SAK	48
5.2.1.3 DBC	49
5.2.1.4 M-GSAK	50
5.2.2 Dünya Sorgusu Sonuçları	51
5.2.2.1 Klasik SAK	52
5.2.2.2 Geliştirilmiş SAK	53
5.2.2.3 DBC	53

5.2.2.4	M-GSAK	54
5.2.3	Ocak Sorgusu Sonuçları.....	55
5.2.3.1	Klasik SAK	55
5.2.3.2	Geliştirilmiş SAK	56
5.2.3.3	DBC.....	57
5.2.3.4	M-GSAK	57
5.2.4	Takvim Sorgusu Sonuçları	59
5.2.4.1	Klasik SAK	59
5.2.4.2	Geliştirilmiş SAK	61
5.2.4.3	DBC.....	61
5.2.4.4	M-GSAK	62
5.2.5	Yüz Sorgusu Sonuçları	64
5.2.5.1	Klasik SAK	64
5.2.5.2	Geliştirilmiş SAK	65
5.2.5.3	DBC.....	66
5.2.5.4	M-GSAK	67
5.3	Deney Sonuçlarının Değerlendirilmesi	68
BÖLÜM 6		
SONUÇLAR VE ÖNERİLER		75
KAYNAKLAR.....		78
EK-A		
BALIK SORGU SONUÇLARI.....		81
EK-B		
DÜNYA SORGU SONUÇLARI		85
EK-C		
OCAK SORGU SONUÇLARI.....		90
EK-D		
TAKVİM SORGU SONUÇLARI.....		94
EK-E		
YÜZ SORGU SONUÇLARI		100
ÖZGEÇMİŞ.....		105

SİMGE LİSTESİ

ϕ	Boş küme
tf	Terim Frekansı
idf	Ters Doküman Frekansı
tf-idf	TF-IDF
$\cos(\theta)$	Kosinüs Benzerliği
S	Karakter katarı
s	S karakter katarı'nın son eki
s(B)	Temel küme skoru

KISALTMA LİSTESİ

DBC	Distance Based Clustering
GSAK	Geliştirilmiş Son Ek Ağacı Kümeleme
KSAK	Klasik Son Ek ağacı Kümeleme
M-GSAK	Melez Geliştirilmiş Son Ek Ağacı Kümeleme
SAK	Son Ek Ağacı Kümeleme
URL	Uniform Resource Locator

ŞEKİL LİSTESİ

	Sayfa
Şekil 5.1 Ort. F-ölçüm değerleri.....	70
Şekil 5.2 Ortalama Ort. F-ölçüm değerleri.....	70
Şekil 5.3 Çalışma Veri Seti Ort. F-Ölçüm Değerleri	71
Şekil 5.4 Çalışma Veri Set Ortalama Ort. F-Ölçüm Değerleri.....	71
Şekil 5.5 Ort. Tutturma değerleri	72
Şekil 5.6 Ortalama Ort. Tutturma Değerleri	73
Şekil 5.7 Ort. Bulma Değerleri	74
Şekil 5.8 Ortalama Ort. Bulma Değer	74

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 5.1	Balık sorgusu için yöntemlerin çıkardığı etiketler 46
Çizelge 5.2	Balık sorgusu için elle çıkarılan sonuçlar 47
Çizelge 5.3	Balık sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları 47
Çizelge 5.4	Balık sorgusu için KSAK yöntemi F-ölçüm sonuçları..... 48
Çizelge 5.5	Balık sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları 49
Çizelge 5.6	Balık sorgusu için GSAK yöntemi F-ölçüm sonuçları 49
Çizelge 5.7	Balık sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları..... 49
Çizelge 5.8	Balık sorgusu için DBC yöntemi F-ölçüm sonuçları 50
Çizelge 5.9	Yöntemler ve elle yapılan etiketlemede çıkan adreslerin sayısı..... 50
Çizelge 5.10	<i>Restaurant</i> etiketi için adres birleştirmeleri 51
Çizelge 5.11	Dünya sorgusu için yöntemlerin çıkardığı etiketler 51
Çizelge 5.12	Dünya kelimesi için elle çıkarılan sonuçlar..... 52
Çizelge 5.13	Dünya sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları.. 52
Çizelge 5.14	Dünya sorgusu için KSAK yöntemi F-ölçüm sonuçları 52
Çizelge 5.15	Dünya sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları . 53
Çizelge 5.16	Dünya sorgusu için GSAK yöntemi F-ölçüm sonuçları..... 53
Çizelge 5.17	Dünya sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları ... 53
Çizelge 5.18	Dünya sorgusu için DBC yöntemi F-ölçüm sonuçları 54
Çizelge 5.20	<i>Büyük</i> etiketi için adres birleştirmeleri 54
Çizelge 5.21	Ocak sorgusu için yöntemlerin çıkardığı etiketler 55
Çizelge 5.22	Ocak kelimesi için elle çıkarılan sonuçlar..... 55
Çizelge 5.23	Ocak sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları.... 56
Çizelge 5.24	Ocak sorgusu için KSAK yöntemi F-ölçüm sonuçları 56
Çizelge 5.25	Ocak sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları ... 56
Çizelge 5.26	Ocak sorgusu için GSAK yöntemi F-ölçüm sonuçları..... 56
Çizelge 5.27	Ocak sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları 57
Çizelge 5.28	Ocak sorgusu için DBC yöntemi F-ölçüm sonuçları..... 57
Çizelge 5.30	<i>Fiyat</i> etiketi için adres birleştirmeleri 58
Çizelge 5.31	Ocak sorgusu M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları 58
Çizelge 5.32	Ocak sorgusu için M-GSAK yöntemi F-ölçüm sonuçları..... 58
Çizelge 5.33	Takvim sorgusu için yöntemlerin çıkardığı etiketler 59
Çizelge 5.34	Takvim kelimesi için elle çıkarılan sonuçlar 59
Çizelge 5.35	Takvim sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları 60

Çizelge 5.36	Takvim sorgusu için KSAK yöntemi F-ölçüm sonuçları	60
Çizelge 5.37	Takvim sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları	61
Çizelge 5.38	Takvim sorgusu için GSAK yöntemi F-ölçüm sonuçları	61
Çizelge 5.39	Takvim sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları..	61
Çizelge 5.40	Takvim sorgusu için DBC yöntemi F-ölçüm sonuçları	62
Çizelge 5.42	<i>Öğretim</i> etiketi için adres birleştirmeleri.....	63
Çizelge 5.43	<i>Güz</i> etiketi için adres birleştirmeleri.....	63
Çizelge 5.44	Takvim sorgusu M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları..	63
Çizelge 5.45	Takvim sorgusu için M-GSAK yöntemi F-ölçüm sonuçları	64
Çizelge 5.46	Yüz kelimesi için yöntemlerin çıkardığı etiketler	64
Çizelge 5.47	Yüz sorgusu için elle çıkarılan sonuçlar	64
Çizelge 5.48	Yüz sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları	65
Çizelge 5.49	Yüz sorgusu için KSAK yöntemi F-ölçüm sonuçları.....	65
Çizelge 5.50	Yüz sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları.....	65
Çizelge 5.51	Yüz sorgusu için GSAK yöntemi F-ölçüm sonuçları	66
Çizelge 5.52	Yüz sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları.....	66
Çizelge 5.53	Yüz sorgusu için DBC yöntemi F-ölçüm sonuçları	66
Çizelge 5.54	Yöntemler ve elle yapılan etiketlemede çakışan adreslerin sayısı	67
Çizelge 5.55	<i>Wikipedi</i> etiketi için adres birleştirmeleri	67
Çizelge 5.56	Yüz sorgusu için M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları	67
Çizelge 5.57	Yüz sorgusu için M-GSAK yöntemi F-ölçüm sonuçları	68
Çizelge EK-A.1	Balık Sorgu Sonuçları	81
Çizelge EK-B.1	Dünya Sorgu Sonuçları	85
Çizelge EK-C.1	Ocak Sorgu Sonuçları.....	90
Çizelge EK-D.1	Takvim Sorgu Sonuçları.....	95
Çizelge EK-E.1	Yüz Sorgu Sonuçları	100

TÜRKÇE ARAMA MOTORU SONUCU KÜMELEME ÇALIŞMALARI

Burak DURAL

BilgisayarMühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Doç. Dr. Banu DİRİ

İnternetteki bilgiler gün geçtikçe artmakta ve bu geniş bilgi ağında istenilen bilgiye ulaşmak büyük bir problem oluşturmaktadır. Bu problemi çözmek adına geliştirilen arama motorları, internet sitelerini olabildiğince indeksleyerek, kullanıcıların bu veriler üzerinde istedikleri bilgileri aramalarına imkan sağlamaktadır.

İnternet gibi geniş bir ağ düşünüldüğünde, yapılan aramalarda –seçilen sorguya göre– yeri geldiğinde milyonlarca internet sayfası bulunabilmektedir. Bulunan bu kadar sonuç arasından kullanıcının ulaşmak istediği bilgiye en yakın sayfayı bulması ise, ayrı bir problem oluşturmaktadır. Bu problemi çözmek için günümüz arama motorları bir çok yöntem sunmaktadır. Arama sonuçlarının belirli kurallara göre derecelendirilerek sıralandırılması için sonuçların kümelenmesi ya da gruplanmasına dayanan yöntemler geliştirilmiştir.

Arama sonucunun kümelenmesi de, bu probleme çözüm getirmek için uygulanan yöntemlerden biridir. Arama motoru tarafından getirilen sonuçların, çeşitli bilgi çıkarımı yöntemleriyle içeriğine göre kümelere ayrılmasına dayanır. Arama sonuçları, kullanıcıya açıklayıcı etiketler içeren kümeler halinde gösterilirler. Böylece kullanıcı, kümelere aradığı bilgiye en yakın olanını seçerek, aradığına daha çabuk ulaşabilir.

Arama sonucu kümeleme için bir çok çalışma yapılmıştır. Bu alanda kullanılan en yaygın ve hızlı algoritmalarından biri olan Son Ek Ağacı Kümeleme (Suffix Tree Clustering) algoritmasıdır.

Bu tez'deki amacımız, Son Ek Ağacı Kümeleme algoritmasıyla ve doküman benzerliğine dayanan kendi geliştirdiğimiz bir algoritmayla Türkçe sayfalar üzerinde arama sonucu kümeleme çalışmaları yapmak ve mümkün olduğunca bu algoritmaları Türkçe için daha iyi sonuçlar verecek şekilde düzenleyerek başarısını ölçmektir.

Çalışmamız çerçevesinde, doküman benzerliğine dayanan DBC, klasik SAK (KSAK), KSAK üzerinde geliştirmeler yaparak oluşturduğumuz GSAK ve GSAK'den elde edilen sonuçların DBC sonuçlarıyla karşılaştırılarak geliştirilmesine dayanan M-GSAK yöntemleri kullanılmıştır. Bu yöntemler kullanılarak yaptığımız deneyler neticesinde, Türkçe arama motoru sonuçları üzerinde, GSAK ve M-GSAK ile yaptığımız kümeleme işlemlerinde, KSAK'ye göre daha başarılı sonuçlar alınmıştır. F-Ölçüm sonuçlarına göre GSAK, KSAK'den %77, M-GSAK de, GSAK'den %13 oranında daha başarılı olmuştur. Ayrıca üzerinde geliştirmeler yapılan SAK tabanlı yöntemler ile doküman benzerliği tabanlı yöntemlerden daha başarılı sonuçlar alındığı da görülmüştür. GSAK, DBC'den %9 oranında daha başarılı sonuç elde etmiştir.

Anahtar Kelimeler: Son Ek Ağacı Kümeleme Algoritması, Arama Motoru Sonucu Kümeleme, Doküman kümeleme

SEARCH RESULT CLUSTERING STUDIES IN TURKISH

Burak DURAL

Department of Computer Engineering

MSc. Thesis

Advisor: Assoc. Prof. Banu DİRİ

Information in the Internet increasing day by day and reaching desired information in this huge network is a major problem. In order to solve this problem, web search engines developed. Web search engines can index web sites to allow users search for the information they need on indexed data.

If a large network such as internet is concerned, depends on the search query, a web search can return millions of web pages. Finding the most relevant information in this huge result set is a separate problem. Today's search engines provide several ways to solve this problem. Generally; methods relies on sorting search results according to ranking with certain rules or clustering and grouping results developed.

Search result clustering as well, is one of the methods to bring solution to this problem. It depends on applying various information extraction techniques to search results for clustering them by their content. The search results are displayed to users in clusters that contain descriptive labels. Thus, user selects one of the closest clusters to the information he/she need and he/she can access the information he need quicker.

Many studies have been conducted for search result clustering. Suffix Tree Clustering (STC), one of the most widely used and fast algorithm in this field.

Our aim in this thesis, performing search result clustering operations with Suffix Tree Clustering and another algorithm based on document similarity on Turkish web pages.

Then, trying to improve these algorithms to get better results and measuring success of these algorithms.

In our work, we used DBC which is based on document similarity calculation, classical STC, GSAK which is an improved STC that we prepared and M-GSAK, which tries to improve GSAK results with using DBC results. On our tests, GSAK and M-GSAK get better results than classical SAK algorithm on Turkish web results. Depends on the F-Score calculations; GSAK scored 77% better performance than KSAK and M-GSAK scored 13% more than GSAK. Also we can say that; improved STC techniques get better results than document similarity calculation based DBC method. GSAK scored 9% better than DBC.

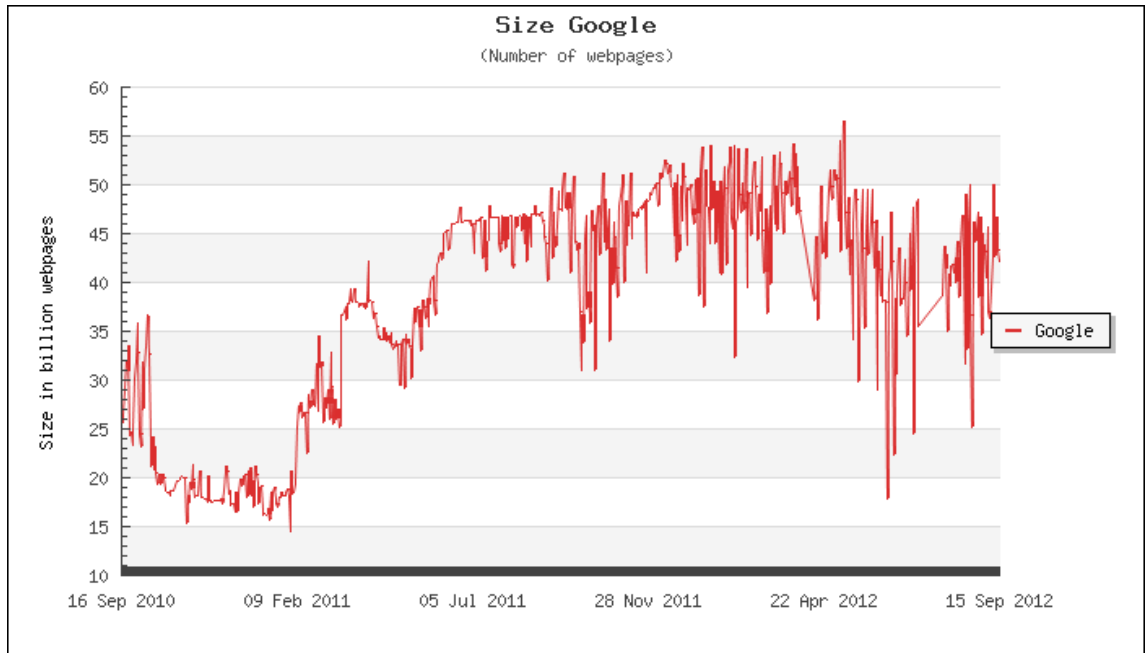
Key words: Suffix Tree Clustering, Search Engine Result Clustering, Document Clustering

BÖLÜM 1

GİRİŞ

İnternetin hızla büyümesi ile birlikte interneti indeksleme çalışmaları ve internet arama motorları da yavaş yavaş ortaya çıkmaya başlamıştır. Böylece insanlar hem internette istediği içeriklere kolayca ulaşabilmişler, hem de adresleri bile bilinmeyen internet siteleri ulaşılabilir hale gelmiştir. İnternet arama motorları, hem internet sitelerinin hem de kullanıcılarının artmasına olumlu yönde katkı sağlamıştır.

Günümüze bakıldığında, en popüler internet arama motorlarından olan Google (<http://www.google.com>)'da 50 milyar civarında internet sayfası indekslenmektedir [1]. İki yıl önce bu sayının yaklaşık 35 milyar civarında olması, internetin büyüme hızı hakkında bize bir fikir vermektedir.



Şekil 1.1 Son iki yılda Google'ın indekslediği internet sayfası sayıları

İnternette içeriğin hızla büyümesi ile birlikte internet arama problemi de ön plana çıkmış, bu konuda yapılan incelemeler ve çalışmalar da artmıştır. İnternet arama problemi, kullanıcının aradağı sorguyla eşleşen az sayıda doküman ile ilgilenir. Kullanıcı da, arama motorundan dönen bu sınırlı sayıda dokümanların çok az bir kısmını gerçekten inceleyebilir. İlk zamanlarından bu yana, arama motorlarının çoğu, sonuçları aranan sorguyla ilgi derecesine göre bir algoritmayla sıralamakta ve beraberinde sayfa hakkındaki URL (Uniform Resource Locator), başlık ve sayfa içeriği hakkında kısa bir açıklama(snippet) bilgisi vermektedirler.

Broder yaptığı çalışmada[2], internet arama motorlarına gönderilen sorguları üç gruba ayırmıştır. Navigasyonel sorgular, kullanıcının aradığı siteyi bildiği fakat adresini bilmediği ya da hatırlayamadığı siteler için yaptığı sorgulardır (ör. “compaq”). Bilgisel sorgular, bir veya daha fazla internet sitesinde var olduğu varsayılan bir bilgiye ulaşmak için yapılan sorgulardır(ör. “San Fransisco”). Üçüncü grup olan işlemsel sorgular ise, internet tabanlı bir aktiviteyi gerçekleştirebilmek için girilen sorgulardır(ör: “pdf dosyası indirmek”). Broder'ın sorgu kayıtları analizinde görülmüştür ki, navigasyonel sorguların tüm sorgulara oranı yaklaşık %20 civarındadır. Bilgisel sorgular %48, işlemsel sorgular ise %30 civarındadır. Standart arama motorlarının yaptığı sıralı liste şeklinde gösterme şekli navigasyonel sorgular için uygundur. Yapılan bazı çalışmalar [3] ayrıca göstermektedir ki, kullanıcılar genelde kısa ve kötü tanımlanmış sorgular kullanmaktadırlar ve yalnızca ilk sıralardaki arama sonuçlarını incelemektedirler. Böylece geleneksel arama sonucu gösteriminde kullanıcı çok anlamlı bir kavram üzerinde arama yaptığında, sadece aradığı şeyin en popüler anlamıyla ilgili içeriklere bakabilir. Çok popüler olmayan alternatif anlamlar ise, genelde kullanıcıların hiç bakmadığı kadar ileriki sayfalarda çıkabilirler.Bu da özellikle bilgisel sorgularda kullanıcın aradığını bulamamasına sebep olabilir. Aranan sorguyla ilgili alternatif anlamlara daha kolay ulaşabilmek için farklı arama sonucu gösterim yöntemleri üzerinde çalışmalar yapılmıştır. Bunlar arasında en popülerlerinden biri de, arama sonucu kümeleme yöntemidir.

1.1. Literatür Özeti

Metin kümeleme işlemlerinin tarihi daha eskiye dayansa da, arama sonucu kümeleme işlemleri üzerindeki çalışmalar, arama motorlarının ortaya çıkmasıyla yakın tarihlere denk gelmektedir. İlk arama motoru olan Archie'nin [4] yayınlanmasından yalnızca iki yıl sonra 1992'de Hearst vd.[5] tarafından 1992 yılında ortaya atılan "Scatter/Gather göz atma paradigması"nda arama sonucu kümeleme işlemleri ilk defa ele alınmıştır. Daha sonraki yıllarda bu yöntemi temel alan bir çok çalışma da yapılmıştır. Arama sonucu kümeleme ile arama sonuçları anlamlı etiketler altında toplandıklarında; kullanıcılar, büyük boyutlu verilerle daha kolay başa çıkmışlardır. İlk arama sonucu kümeleme işlemlerinde, *Bag of words* yöntemi ile frekans vektörleri oluşturularak klasik doküman kümeleme işlemlerine benzer teknikler kullanılmıştır. Daha sonra, arama sonuçlarında geçen kelime grupları tekrarlarını ele alan yaklaşımlar ortaya çıkmış ve bunların daha başarılı olduğu görülmüştür. Bunlardan biri ve en çok bilineni de 1998 yılında yayınlanan Son Ek Ağacı Kümeleme (SAK) [6] algoritmasıdır. SAK yöntemi yayınlandıktan sonra, arama sonucu kümeleme alanında oldukça popüler olmuş ve üzerinde birçok farklı geliştirme çalışmaları yapılmış [7, 8, 9], farklı bakış açıları değerlendirilerek SAK üzerinde değişiklikler yapılmıştır.

Arama sonucu kümeleme alanında, SAK yöntemi dışında bir çok çalışma gerçekleştirilmiştir. 2003 yılında Lingo [10] diğer yöntemlerde kullanılan aksine, öncelikle dokümanlar arasında ayrımı sağlayacak etiketler tespit etmiş, daha sonra da belirlenen etiketlere ait kümelere dokümanları atamıştır.

1.2. Tezin Amacı

İnternette içerik hızla büyümektedir. Ülkemizde internet kullanımı hızla arttığından, internet üzerindeki Türkçe içerik de aynı hızda artış göstermektedir. Güncel arama motorlarının çoğunda Türkçe veriler için de başarılı sonuçlar alınmasına rağmen, diğer dillerde olduğu gibi, Türkçe için de arama sonucu olarak dönen bazen milyonlara erişen sayfalardan aranılanı bulabilmek çoğu zaman güç olmaktadır. Bu sorunu çözmek için kullanılabilecek yöntemlerden biri de arama sonucu kümeleme işlemleridir. Yaptığımız literatür tarama araştırmaları sonucunda görülmüştür ki; Türkçe dilindeki arama sonuçları için yapılan arama sonucu kümeleme çalışmaları oldukça sınırlı sayıdadır.

İngilizce gibi diller için kullanılan yöntemlerin hiç değiştirilmeden Türkçe için kullanılmasında da yeterli başarı elde edilememektedir. Bu sebeple çalışmamızda Türkçe dili için daha başarılı arama sonucu kümeleme işlemlerinin yapılabilmesi için araştırmalar yaptık. Yaptığımız çalışmada, en popüler arama sonucu kümeleme yöntemlerinden biri olan SAK'yi Türkçe dili için daha verimli ve başarılı çalışacak hâle getirmek istedik. Bunun için de, algoritmayı Türkçe dilinin yapısına daha uygun olacak şekilde düzenlemeye ve anlamsal bilgileri de kullanarak geliştirip daha başarılı olmasını sağlamaya çalıştık. Ayrıca, klasik doküman kümeleme çalışmalarında, doküman benzerlik ölçümlerinde sıkça kullanılan Kosinüs Benzerlik temeline dayanan basit bir kümeleme algoritması olan Distance Based Clustering (DBC) algoritmasını geliştirip, kullandık. Böylece klasik doküman kümelemeye dayanan yöntemlerle, arama sonucu kümeleme işlemi için özel geliştirilmiş bir yöntem olan SAK'nin başarımını karşılaştırabilecek sonuçlara ulaşmayı hedefledik. Daha sonra da; DBC'den elde ettiğimiz sonuçları da kullanarak SAK algoritmasının performansını arttırmayı hedefleyen melez bir yöntem geliştirdik.

1.3. Hipotez

Kümeleme, verilen girdiden birbirine yakın olan objeleri bir grupta toplama işlemine denir. Arama sonuçlarına uygulandığında ise; bu işleme, arama sonuçlarının kolaylıkla göz gezdirilebilen tematik gruplara yerleştirilmesi diyebiliriz.

Arama sonucu kümeleme problemi, doküman kümeleme problemine benzetilse de içerisinde bir çok farklılıklar taşımaktadır. Klasik doküman kümeleme işleminin performansını arttıracak ilkeler çoğunlukla arama sonucu kümeleme işleminde işe yaramazlar. Arama sonucu kümeleme işleminde, kümelemek için kullanılacak veri, internet sitesinin başlığı ve site hakkındaki bir iki cümlelik kısa bir açıklama yazısından (snippet) ibarettir. Bu kısıtlı veri, algoritmaların verideki parazitlenmeye olan toleransını azaltır ve klasik doküman kümeleme işleminde kullanılan algoritmaların başarısız olmasına sebep olur.

Bu kısıtları göz önünde bulundurarak geliştirilen arama sonucu kümeleme algoritmaları; kısa zamanda, az veriyle, başarılı kümeleme işlemlerini gerçekleştirmeyi hedeflemektedirler.

En bilinen ve en çok kullanılan arama sonucu kümeleme yöntemlerinden SAK algoritması, arama sonucu kümeleme alanında, özellikle İngilizce dilinde en başarılı algoritmalarından biridir. Fakat, yalın haliyle kullanıldığında Türkçe’de başarısı yetersiz kalmaktadır.

Tez kapsamında yaptığımız çalışmada; SAK algoritmasına yeni bir birleştirme adımı eklediğimizde ve hem kümelerin oluşturulması, hem de etiketlerin belirlenmesi aşamasında Zemberek [11] Türkçe dil işleme kütüphanesi kullanılması ile geliştirdiğimiz GSAK (Geliştirilmiş Son Ek Ağacı Kümeleme) yöntemi ile Türkçe arama sonuçlarında yalın haldeki SAK algortimasından daha başarılı sonuçlar elde edilmiştir.

Ayrıca, DBC yöntemiyle elde edilen kümeler ile GSAK sonuçlarının karşılaştırılmasıyla elde edilen melez yöntem ile bu başarı biraz daha da arttırılmıştır.

ÖNCEKİ ÇALIŞMALAR

Arama sonucu kümeleme, doküman kümeleme konusunun bir alt dalıdır. Doküman kümelemede kullanılan yöntemlerin burada da kullanılabileceği düşünülse de, arama sonucu kümeleme problemi birçok farklılıklar içermektedir. İnternet sürekli büyümekte olan geniş bir bilgi hazinesidir. Bu sebeple arama sonucu kümeleme işleminin, arama yapıldığı anda başlayıp, arama sonucu ekranda gösterilene kadar verimli bir şekilde, gerçek zamanlı olarak gerçekleştirilip, kullanıcıya yansıtılması gerekmektedir. İlk arama sonucu kümeleme çalışmaları, neredeyse ilk arama motorları kadar eskidir.

2.1. Scatter/Gather

Scatter/Gather algoritması [5] arama sonrası uygulanabilen ilk arama sonucu kümeleme algoritmasıdır. Scatter/Gather temel olarak kümeleme tabanlı doküman göz atma (browsing) uygulamasıdır. Scatter/Gather algoritması dokümanları verilen sayıda kümeye ayrılmaktadır. Girdi olarak, doküman vektörlerinin birbirleri arasındaki benzerlik ölçüleri - kosinüs benzerliği kullanılmıştır - kümeleme algoritmasına verilir. Burada *Fractionation* adı verilen algoritma devreye girer ve bu algoritma parçalı kümeleme için gerekli olan ilk k başlangıç merkezini belirler. Daha sonra; dokümanlar oluşturulan merkezlere yakınlıklarına göre, bir kümeleme alt yordamları ile kümelere aktarılırlar.

Scatter/Gather yönteminde işlemlerin interaktif bir şekilde kullanıcıya bir an önce yansıtılması hedeflendiğinden, kümeleme işlemi yapılırken doğruluktan daha çok işlemin hızlı bir şekilde yerine getirilmesi üzerinde durulmuştur. Kümelere ait olan

dokümanlar belirlendikten sonra, bu kümelerde en çok geçen kelimeler bulunarak kümenin başlığı/etiketi olarak kullanılmıştır.

2.2 Son Ek Ağacı Kümeleme

Zamir ve Etzioni'nin geliştirdiği Son Ek Ağacı Kümeleme (SAK) algoritması [6], daha sonra da birçok arama sonucu kümeleme çalışmasında kullanılmıştır. SAK, lineer zamanlı, artırımsal (incremental) bir algoritmadır. Zamir ve Etzioni SAK'nin bu alanda, standart kümeleme algoritmalarından daha hızlı çalıştığını, daha yararlı ve kullanabilir olduğunu göstermiştir.

Son Ek Ağacı Kümeleme iki aşamadan oluşmaktadır. Bunlar: temel kümelerin oluşturulması ve temel kümelerin birleştirilerek sonuç kümelerinin elde edilmesidir. İlk aşamada genelleştirilmiş son ek ağacı kullanılır. Bu son ek ağacına tüm dokümanlardaki bütün cümleler, kelimeleri temel elemanlar olarak yerleştirilir. Veri setindeki dokümanların tamamı bu şekilde ağaç yapısı üzerinde gösterildikten sonra, son ek ağacı'nın düğümleri hangi kelimelerin ya da kelime gruplarının, hangi dokümanlarda kullanıldığı konusunda bilgi verebilecek hale getirilir. Bu bilgi kullanılarak aynı kelime gruplarını kullanan dokümanlar aynı küme altına alınarak temel kümeler oluşturulur. İkinci aşamada birbirine benzer temel kümelerin birleştirilmesi hedeflenmektedir. Bu aşamada, ortak doküman sayıları belirlenen bir eşik değerden fazla olan temel kümeler birleştirilir. Zamir ve Etzioni bir sene sonra yayınladıkları Grouper [12] makalesiyle, SAK algoritması üzerinde geliştirme yapmışlardır. Geliştirilen SAK algoritmasıyla, birbiriyle çok yakın anlamlarda olan sözcük grupları elenmiş, çok bilgi vermeyen, genel sözcük grupları da filtrelenmiştir. Ayrıca, durak (stop word) kelimelerinin çıkarılması, az kullanılan sözcük gruplarının derlemde (corpora) çıkarılması gibi yöntemler de uygulanarak iş yükü %50 azaltılmıştır.

2.3 Son Ek Ağacı Kümeleme Üzerinde Yapılan Geliştirmeler

Wang [13], geleneksel SAK algoritmasında kullanılan benzerlik hesaplama yönteminin iki dezavantajı olduğunu tespit etmiştir. Karşılaştırılan iki küme arasındaki doküman oranı dengesiz olduğunda, bir küme diğerinin alt kümesi olsa dahi, bu iki küme birleştirilememektedir. Ayrıca, birbirine benzer dokümanlar, birbiriyle çakışmayan ayrı

kümelerde oldukları zaman, algoritma sadece kümeler arasındaki çakışan dokümanlara baktığı için bu kümeler yine birleştirilemiyordu. Wang, bu problemleri çözmek amacıyla, temel kümelerin birleştirilmesi üzerine bir algoritma geliştirmiştir. Bu algorithmada iki çeşit kümeler arası benzerlik yöntemi kullanılmıştır. İlkinde geleneksel SAK'de olduğu gibi kümeler arasındaki çakışan dokümanlardan yararlanılır. Fakat burada çakışma için, geleneksel SAK'de kullanılan farklı bir formül kullanılmış böylece dengesiz kümeler ile ilgili problem çözülmüştür. İkincisinde ise; çakışan dokümanları olmayan kümeler için, metinsel benzerliğe bakılarak, bu durumdaki kümelerin problemleri çözülmeye çalışılmıştır.

Chim vd. [8] yaptığı çalışmada, SAK yönteminde, standarttan farklı bir benzerlik ölçümü geliştirmiştir. Son Ek Ağacı oluşturulurken, her düğüm için TF ve IDF değerleri de hesaplanır. Böylece son ek ağacı oluşturulduğunda, aynı anda tüm dokümanlardaki düğümlerin, TF-IDF (terim frekans-ters doküman frekans) değerleri de hesaplanmış olur. Bu bilgiler de Vector Space Document (VSD) [14] üzerinde tutulur. Böylece; model yardımıyla, geleneksel benzerlik ölçme yöntemleri kullanılabilir. Yapılan çalışmada Kosinüs benzerlik hesaplama yöntemi kullanılmıştır. Tüm dokümanlar için, Kosinüs benzerlik yöntemi kullanılarak ikili benzerlik değerleri hesaplanır ve bu değerlerle de Group-average Hierarchical Clustering (GAHC) [15] algoritması kullanılarak kümelerin son halleri oluşturulur.

2009 yılında Kale vd. [7] yaptıkları çalışmada, geleneksel SAK'nin temel kümeleri birleştirme ve kümeleri etiketleme mekanizmaları üzerinde değişiklik yapmışlardır. Bu çalışmada Son Ek Ağacı n-gram'lar kullanılarak ($n \leq 3$) oluşturulmuştur. Geleneksel yöntemde iki temel küme birleştirilirken, oluşturulan yeni kümenin etiketi olarak, etiketi daha uzun olan temel kümenin etiketi seçilmekteydi. Bu da diğer temel küme açısından bilgi kaybı yaratabiliyordu. Bu sebeple geliştirilen yeni birleştirme algoritmasıyla, iki küme birleştirildikten sonra, oluşturulan yeni kümede yer almaması gereken dokümanların çıkarılması hedeflenmiştir.

Kale vd. [7] ayrıca, kümelerin etiketlenmesinde de geleneksel SAK tekniğinden farklı bir yöntem kullanmıştır. Küme içerisindeki dokümanlarda geçen, n-gram ile oluşturulmuş

tüm sözcük grupları çıkartılarak, en yüksek frekansa sahip sözcük grubu kümenin etiketi olarak atanmıştır.

Wen vd. [9] yaptığı çalışmada, geleneksel SAK'den farklı olarak, kümelerin etiketlemesinin yapılmasında Anlamsal İçerik İndeksleme (Latent Semantic Indexing) yöntemini kullanmıştır.

Gizli Anlamsal İndeksleme (Latent Semantic Indexing) bir bilgi çıkarımı yöntemidir. Terim tabanlı matrisin derecelerini düşürmeye çalışarak, kirlilik yaratan kelimelerden ve eş anlamlı kelimelerden kurtarmayı amaçlar. Böylece dokümanın içerisindeki gizli kalmış anlamı açığa çıkarmaya çalışır. Bunu yaparken de Tekil Değer Ayrıştırması (Singular Value Decomposition) tekniğini kullanır.

SAK algoritmasının oluşturduğu sonuç kümelerinin etiketleri her zaman kümenin içerisindeki dokümanların içeriğiyle doğru orantılı olmayabilir. Eş anlamlılık ve çok anlamlılık durumları da işin içerisine girdiğinden, kümelerin etiketlenmesinde hatalı sonuçlar doğabilir. Wen vd. [9] bu sebeplerle daha başarılı küme etiketlendirilmesinin anlamsal boyutta ele alınması gerektiğini savunmaktadır. Yapılan çalışmada, temel kümeleri temsil edecek etiketleri hesaplamak için, aday sözcük gruplarının, terim-doküman matrisinin anlamsal içerik uzayı ve aday sözcük gruplarının uzunluğuna dayanan bir formül yardımıyla puanları hesaplanmış en yüksek puana sahip sözcük ya da sözcük grubu temel kümenin etiketi olarak seçilmiştir.

2.4 SHOC: Semantic Hiearchical Online Clustering

Anahtar sözcük grubu bulma ve ortogonal kümeleme işlemlerinin bir arada kullanılmasıyla oluşturulmuştur. Sadece İngilizce için değil Çince gibi doğu dillerinde de başarılı bir şekilde çalışması sağlanmıştır. Temel olarak SAK'nin işlem akışını almasına rağmen, SAK doğu dillerinde verimli olmadığı için birçok değişiklik yapılmış, farklı algoritmalar kullanılmıştır. Anahtar kelime grupları, yani etiketler oluşturulurken SAK'nin aksine doğu dillerinin çok harfli ve kelime bitimlerinde İngilizce gibi boşluk kullanmayan yapısına uygun bir metot kullanılmıştır. Dokümanları kümelere yerleştirme işleminde, ilk etapta çıkarılan anahtar kelime grupları terim olarak kullanılarak bir terim-doküman matrisi oluşturulur. Latent Semantic Indexing

algoritmasından yola çıkılarak, aynı dokümanda geçen etiket kelime grupları ve aynı etiketi içeren dokümanlar dikkate alınır. Böylece yoğun olarak aynı dokümanları işaret eden etiketler bir araya getirilir ve böylece de kümeler oluşturulur. Bir araya getirilen terimlerde, bu terimleri içeren dokümanların sayısı incelenerek bir terimin diğer terimin alt kümesini oluşturabilecek nitelikte olup olmadığına da karar verilir. Böylece arama sonuçlarında etiketlerin ağaç yapısı şeklinde, alt kümeleriyle gösterilmesine de olanak sağlanmış olunur [16].

2.5 Lingo

Genelde çoğu arama sonucu kümeleme algoritmasında öncelikli olarak kümeler oluşturulur. Daha sonra da bu kümeleri tanımlayıcı etiketler oluşturulur. Fakat bu yaklaşımda, çoğu zaman dokümanlar arasındaki karmaşık benzerlik hesapları küme içerisindeki dokümanları bir arada tutan kavramın ne olduğu konusunda insanlara fikir vermeyebilir. Bu sebeple Lingo [10] bu yaklaşımı tersten ele almıştır. Önce kümeleri tanımlayabilecek anlamlı, insanlara daha açıklayıcı gelecek etiketler oluşturulur. Bunun için veri setindeki dokümanlar içerisinde en çok kullanılan sözcük grupları bulunur. Daha sonra orijinal terim-doküman matrisi üzerinde Singular Value Decomposition kullanılarak azaltma işlemi uygulanır. Böyle yapılarak arama sonuçlarındaki çeşitli konuların gizli kalmış yapıları keşfedilmeye çalışılır. Son olarak da çıkarılan konular, grup açıklamaları ile eşleştirilir ve dokümanlar ilgililiklerine göre bu gruplara atanırlar. Dokümanların gruplara atanma işleminde VSM (Vector Space Model) kullanılır. Bu atama işlemi VSM tabanlı doküman getirme(document retrieval) işlemini andırmaktadır. Bu aşamada, oluşturulan konular ile doküman veri seti üzerinde tekrar bir sorgulama yapıldığı da söylenebilir. Bir konu ile dokümanın benzerlik değeri hesaplanır. Bu değer önceden belirlenen sınır değerden yüksekse doküman konunun temsil ettiği gruba dahil edilir.

2.6 Wordnet Kullanılan Çalışmalar

Hotho vd. [17] çalışmasında, parçalı kümeleme üzerinde WordNet'i [18] kullanarak başarı elde etmeye çalışmıştır. Kullanılan terimlerin (yalnızca isimler) WordNet synset'leri bulunarak, her doküman için terim vektörünün haricinde bir de

synset'lerden oluşan kavram vektörleri oluşturulmuştur. Kavram vektörleri ile terim vektörlerini birleştirilerek, yalnızca kavram vektörlerini kullanarak ya da yalnızca belirli bir sayıda terimi kavram karşılıklarıyla değiştirerek kullanılacak esas vektörleri oluşturmuşlardır. Bu vektörler kullanılarak, Bi-Section, K-Means algoritmasıyla kümeleme işlemi yapılmıştır. Bu çalışmayla, özellikle terim vektörü ve kavram vektörünün birleştirildiği durumlarda %8 civarı bir başarı elde edilmiştir.

Hu vd. Son Ek Ağacı Kümeleme'de oluşturulan küme etiketlerinin bazen birbiriyle aynı ya da benzer olduğunu gözlemlemiştir [19]. Bu da kullanıcının istediği bilgiyi kolayca bulmasını engelleyecek ek bir yük getirmektedir. Bu problemi gidermek için yapılan çalışmada, anlamsal açıdan birbirine benzeyen kümelerin bir araya getirilmesi ve kümelerin etiketlerine göre hiyerarşik bir biçimde gösterilmesi için bir çalışma yapılmıştır. Oluşturulan sistemde öncelikle SAK yöntemi standart şekliyle uygulanır ve sonuç kümeleri elde edilir. Küme etiketleri incelenerek, anlamsal olarak birbirine denk olan kavramlar varsa bu kümeler birleştirilir. Örneğin *edebiyat* ve *literatür* etiketlerine sahip iki ayrı küme varsa bu kümeler birleştirilir. Hiyerarşi oluşturmak için ise yine küme etiketleri karşılaştırılarak, bir etiketin diğerini anlamsal ya da yapısal olarak kapsaması durumunda, kısa olan diğerinin üst hiyerarşisinde gösterilir. Örneğin; *İngiliz* ve '*İngiliz Edebiyatı*' etiketli iki küme varsa, '*İngiliz*' etiketli küme, '*İngiliz Edebiyatı*' etiketli kümenin üst hiyerarşisine yerleştirilir.

GENEL BİLGİLER

Yaptığımız çalışmanın ve amacının daha iyi kavranabilmesi için, bu bölümde tez içerisinde kullandığımız birçok kavram ve konu hakkında genel bilgileri bir araya getirdik.

3.1 Internet Arama Motoru

Arama motorları, internette bilgi aramak için geliştirilmiş araçlardır. Arama motorlarının öncelikli amacı, internet kullanıcılarının ilgilendikleri konuları internette bulmalarını sağlamaktır. Günümüz arama motorlarının çok büyük bir kısmı, bu işlemleri HTTP (Hyper Text Transfer Protocol) üzerinden, web uygulamaları ile yaparlar. Örneğin bir arama yapmak için, kullanıcı web tarayıcısını kullanarak arama motorunun web sayfasına girer. Burada kullanıcı sorgusunu web sayfasında bulunan arama kutucuğuna yazar. Genelde bu sorgu, kullanıcının aradığı bilgiyi tanımlayan bir veya daha fazla anahtar kelimelerden oluşan sözcüklerdir. Sorgu sözcükleri girildikten sonra kullanıcı arama işlemini başlatacak düğmeye basarak işlemi tetikler (Şekil 3.1). Arama sonucu olarak, kullanıcının sorguladığı konuları içeren web sayfalarının linkleri liste halinde aranan sözcükle en çok ilgili sayfalardan en az ilgililere doğru sıralanacak bir şekilde getirilir. Genelde bu bağlantıların yanında, sayfayı kısaca açıklayan ufak bir yazı (snippet) bulunur. Kullanıcı bu listeden, açıklama yazılarından da yararlanarak arama yaptığı sözcükle en çok alakalı olduğunu düşündüğü web sayfalarına ilgili linki tıklayarak geçiş yapabilir.



Şekil 3.1 Günümüzde en çok kullanılan arama motoru olan Google'ın ekran görüntüsü

3.1.1 Arama Motorları Tarihçesi

İnternet üzerindeki web sitelerinin hızla artması ve internetin de hızlanmasıyla, internet, insanların dünya çapında hemen hemen her konu hakkında bilgiler bulabileceği bir yer haline gelmiştir. Web siteleri arttıkça, bu kadar çok bilgi arasından istenilene ulaşmanın zorluğu da aynı hızla artmıştır. Bu sebeple arama motorlarına olan ihtiyaç da artmıştır. İnternet büyüdükçe, buna paralel olarak, arama motorları da artmaya ve daha yeni teknolojilerle daha başarılı sonuçlar elde etmeye uğraşmaktadırlar. Bu sebeple internetin doğuşundan bu yana birçok arama motoru geliştirilmiştir ve halen arama motorları konusunda yenilikler ve geliştirmeler hızla devam etmektedir [20].

İnternette arama yapmak için kullanılan en eski araç Archie'dir [4]. 1990 yılında geliştirilen Archie, genel anonim FTP klasörlerinde bulunan dizinlerin dosya isimlerini veritabanında saklayarak bunlar üzerinden arama yapabilmekteydi. 1991 yılında Gopher protokolünün [21] geliştirilmesi, metin dosya içeriklerinin de indekslenmesini mümkün kılmıştır. Bu protokolü kullanan **Veronica** ve **Jughead** arama motorları geliştirilmiştir. Bunlar da Archie gibi FTP'de arama yapabiliyorlardı. Çoğu üniversitelerin olmak üzere, ilk birkaç yüz web sitesinin çıkması da 1993 yılını bulmaktadır. Birazda bu sebeple 1993 yılına gelinene kadar, henüz web siteleri üzerinde arama yapabilen bir

arama motoru yoktu. Bu tarihte geliştirilen **W3Catalog**, kaynağını o dönem internette bulunan siteyi hazırlayanlar tarafından oluşturulan katalog niteliğindeki web sitelerinden alınmaktadır ve yazılan bir kodla katalog sitelerdeki web site listeleri çekilmekte ve standart bir formatta kullanıcıya sunulmaktaydı. Listeler, düzenli aralıklarla yine bu sitelerden yararlanılarak güncelleniyordu. Yine 1993 yılında geliştirilen **World-Wide Web Wanderer** ilk kez bir **web-robot** kullanan arama motoru oldu. Bu web-robot internette dolaşarak, Wide Web Wanderer'ın veritabanına URL'leri ekliyordu. 1993 yılının sonlarında geliştirilen **AliWeb** bir web-robot içermiyordu. Sitelerini AliWeb'de aranabilir kılmak isteyen web site yöneticilerinin, siteleri hakkında kısa bir açıklama ve anahtar kelimeler içeren bir index dosyası hazırlamaları ve bunu da AliWeb'e bildirmeleri gerekmektedir. Aralık 1993 yılında ortaya çıkan **JumpStation**; crawler, indexer ve arama olmak üzere üç temel arama motoru özelliğini de içermektedir. JumpStation, internet üzerinden sorgulama yapmak için bir web arayüzü de içeriyordu. 1994 yılında geliştirilen **WebCrawler**; ilk “full text”, crawler tabanlı arama motoru oldu. Sayfaların sadece başlığı ya da açıklamaları üzerinden değil, tüm sayfa üzerinde arama yapabilmekteydi.

Bu zamandan sonra birçok arama motoru geliştirilmeye başlandı. **Magellan, Excite, InfoSeek, Inktomi, vs. AltaVista** da bu dönemlerde (1995) ortaya çıkmıştır. Diğer arama motorlarından farklı olarak, çok kanallı (multithread) bir crawler kullandığı için, kısa zamanda daha fazla sayıda site Altavista tarafından indekslemiş, oluşan bu geniş bilgi kaynağı da arama sonuçlarının daha başarılı olmasını sağlamış, bu da o dönem, Altavista'nın en popüler arama motoru olmasına sebep olmuştu.

1998 yılında yayınlanan **Google** [22], 2000'li yılların başından itibaren en popüler arama motoru haline geldi. Bu unvanını halen daha sürdürmektedir. Google kullandığı PageRank [23] algoritmasıyla diğer arama motorlarından çok daha başarılı bir arama sonucu sıralaması gerçekleştirmektedir. Google bu başarısını sürdürerek ve sürekli arama teknolojisini geliştirerek, günümüze dek başarılı olmuş ve dünya çapında en çok kullanılan arama motoru olmayı başarmıştır. 2004 yılına kadar başka arama motorlarını kendi web sitesinde kullanan Microsoft da, bu tarihte kendi crawler'ı **MsnBot**'u piyasaya sürerek, arama işlemleri teknolojisini kendisi geliştirmeye başlamıştır.

Microsoft 2009 yılında da, semantik arama özelliklerine sahip olan Bing arama motorunu tanıtmıştır.

3.1.2 Arama Motorlarının Yapısı

Arama motorunun oluşturulmasında farklı fiziksel bileşenler kullanılabilir. Tipik bir arama motorunun yapısını mantıksal olarak ele aldığımızda ise: Web Crawler, Doküman ön belleği, Doküman indeksi, Doküman sıralama, Sorgu işleme ve Sunum arayüzü [20] bileşenlerinden oluşmaktadır.

Web Crawler: WebCrawler, WebSpider veya robot olarak da adlandırılır. Bu bileşenin amacı internetteki yeni kaynakları keşfederek, bunları arama motorunun veri tabanına eklemek; bu sayede bu kaynakların da arama sonuçlarına yansıtılabilmesini sağlamaktır. Bu işlemi de gezdikleri web siteleri üzerindeki URL'leri öz yinelemeli bir şekilde tarayarak yaparlar. Böylece kendisine bir başlangıç adresi verilen webcrawler bu internet adresi sayfasının içerisindeki URL'lerden başlayarak, sayfada bulunduğu her linke dallanarak gider. Gittiği URL arama motorunun veri tabanında kayıtlı değilse, ekler.

Doküman İndeksi: Doküman indeksinin amacı, verilen kelimelerin hangi internet sayfalarında bulunduğunu indekslemektir. Bunun için genellikle arama motorları ters indeks (inverted index) veri yapısının çeşitlerini kullanırlar.

Doküman Ön Belleği: Bir çok arama motoru indekslediği internet sayfalarının orijinal hallerini de saklarlar. Bazı arama motorlarında internet sayfasının bu ön belleklenmiş hali de kullanıcıya gösterilebilmektedir. Bu sayede eğer internet sitesi geçici olarak ulaşılamaz durumdaysa, sayfanın ön bellekteki haline bakılabilir.

Doküman Sıralama: Arama motoruna gönderdiğimiz sorgu ne kadar açıkça belirtilmiş de olsa, sorgu ile ilgili binlerce ve bazen milyonlarca internet sayfası bulunacaktır. Doküman sıralama algoritmaları burada devreye girmektedir. Arama sonucunda dönen dokümanları, aranan sözcükle en alakalı ve en ilgi çekici olanı en üstte gelecek şekilde sıralayıp, kullanıcıya döndürmeye çalışırlar. Bunun için arama motorları farklı algoritmalar kullanmaktadırlar.

PageRank: Bu sıralama algoritmalarından en çok bilineni şüphesiz günümüzde dünyanın en çok kullanılan arama motoru olan Google'ın PageRank sistemidir [22]. PageRank, Google arama motorunda ilk zamanlarından beri kullanılan bir sayfa derecelendirme sistemidir. Bir sayfanın PageRank değeri hesaplanırken; bu sayfadan diğer sayfalara giden linkler ve bu sayfayı işaret eden başka sayfalardaki linkler önem teşkil etmektedir. Bir sayfanın PageRank değeri hesaplanırken bu sayfaya gelen ve bu sayfadan işaret edilen sayfaların PageRank değerleri sayfanın PageRank değerinin hesaplanmasında önemli bir etkidir. Hesaba ilk başlandığında bu diğer sayfaların PageRank değerleri bilinmiyor olabilir. Bu durumda, diğer sayfalar için tahmini bir başlangıç PageRank değeri verilerek hesaplama başlatılır. Buradan elde edilen PageRank değerleri kullanılarak işlem sürekli tekrarlanır. Yeterince tekrar yapıldıktan sonra PageRank değeri bir noktada sabitlendiğinde işlem tamamlanmış olur.

Google tarafından indekslenen her sayfa için bir PageRank değeri hesaplanır. Bir arama yapıldığında sonuçlar listelenirken PageRank değeri en yüksek olan sayfa en yukarıda gelecek şekilde listeleme yapılır.

Günümüzde Google'ın sayfaları derecelendirmek için kullandığı sıralama algoritması burada anlatılandan daha karmaşık ve hızlı çalışan, yeni eklenen sayfaları daha çabuk sıralamaya dahil eden bir sistemdir. Örneğin: Google Caffeine [24]

Sorgu İşlemcisi: Bu bileşenin görevi kullanıcının girdiği sorguyu işlemektir. Bu işlem yapılırken, sorgu işlemcisi; doküman indeksi, önbellek ve sıralama bileşenlerini kullanarak sorguyla ilgili sonuçları hazırlar. Hazırlanan sonuç doküman listesi, sunum arayüzüne iletilerek, istenilen şekilde kullanıcıya sunulur.

Sunum Arayüzü: Arama motorunun döndürdüğü sonuçları görüntülemek için bir sunum arayüzüne ihtiyaç vardır. Bunun için yıllar içerisinde birçok farklı arayüz geliştirilmiştir. En popüler olanlarından bazıları şunlardır:

* **Sıralı liste sunumu:** Günümüz arama motorlarında en çok kullanılan sunum arayüzüdür. Bu gösterimde, döndürülen internet sayfaları, bir algoritma yardımıyla sorgu ile olan ilgilerine göre sıralanırlar. Girilen sorguyla en çok ilgili olabilecek sayfalar sıralamada daha yukarılarda yer alırlar.

* **İlinti Geribildirimi:** Bu görünüm arayüzünde, arama motoru kullanıcıya; sorgusunu nasıl değiştirerek, aradığı konu ile ilgili daha iyi bilgi almasını sağlayacak hale getirebileceği konusunda öneriler sunar. Bu öneri de, sorgunun içeriği ve dönen dokümanlar incelenerek yapılır.

* **El Yapımı Web Dizinleri:** Bu yapıda, web siteleri konularına göre, hiyerarşik bir kataloğa eklenirler. Ancak bu işlem yalnızca el ile yapılabilir. Bu gösterim arayüzü sayesinde oluşturulan dizinler konuyla ilgisiz dokümanlardan arındırılmış olur, fakat diğer taraftan işlem el ile yapıldığı için çok yavaş ilerlemektedir ve yine bu yavaşlık sebebiyle bazı internet sayfaları geçerliliğini veya güncelliğini çoktan yitirmiş olabilirler.

* **Kümeleme Arayüzü:** Arama sonucu kümeleme işlemi; arama sonuçlarının, konularına göre el yapımı web dizinleri yönteminde olduğu gibi organize etme işlemidir. Yalnız buradaki gruplara ayırma işlemi el ile değil arama motoru tarafından yazılımsal olarak sağlanmaktadır. Buradaki gruplar da daha önceden elle oluşturulan gruplar değil, arama motoru tarafından arama sonuçlarına göre dinamik bir biçimde oluşturulan gruplardır. Bu arayüzün detaylarına bölüm 3.2.3’de değinilmiştir.

3.2 Kümeleme

Kümeleme bir grup nesneyi ayırma işlemidir. Böylece aynı kümedeki nesnelerin, diğer nesnelere göre birbirine daha benzer olması hedeflenmektedir. Kümeleme işlemi bir araştırımsal veri madenciliği işidir ve yaygın kullanılan bir istatistiksel veri analizi tekniğidir. Birbirine benzeyen özellikteki nesnelerin bir araya getirilmesi temeline dayanır. Aynı kümedeki nesneler, benzerlik oranında olabildiğince birbirine yakın, farklı kümelerdeki nesneler ise olabildiğince birbirine uzak olmalıdır. Makine öğrenmesi, örüntü tanıma, görüntü analizi, bilgi çıkarımı vb. birçok alanlarda kümeleme işlemi yaygın bir şekilde kullanılmaktadır. Doküman kümeleme ve arama sonucu kümeleme de bilgi çıkarımı alanındaki kümeleme uygulamalarındandır.

3.2.1 Doküman Kümeleme

Doküman (ya da metin) kümeleme, veri kümeleme işleminin bir alt dalıdır. Doküman kümelemenin, bilgi çıkarımı, doğal dil işleme (natural language processing) ve makine öğrenmesinin de bazı kavramlarını kullanır. Doküman kümeleme işlemi, bir doküman koleksiyonu içerisindeki doğal grupları tespit etmek ve bu grupların içeriğiyle ilgili bir tanıtıcı metin ya da başlık oluşturma işlemini kapsar.

Doküman kümeleme işlemi, sayısal verileri kümeleme işlemlerinden daha zorlu bir görevdir. Dokümanların sayısal olarak gösterilme biçimi genelde yalnızca kelime frekanslarıyla oluşturulur. Çoğu doğal diller için, içerikte en çok kullanılan kelimeler de çoğunlukla işleme yardımcı olmaktan çok veri kirliliği yaratırlar. Bu da doküman kümeleme işlemini daha da zor bir hale getirmektedir.

Kümeleme işlemi, temelde sınıflandırmaya benzese de birçok yönden sınıflandırma probleminden ayrılır. Bir sınıflandırma probleminde oluşacak sınıflar, gruplar ve bunların özellikleri ön bilgi olarak verilir. Dokümanlar sınıflandırma işlemi uygulanarak bu sınıflara atanırlar. Kümeleme işleminde ise, işlem başlamadan önce oluşturulacak sınıfların sayısı, özellikleri, bileşenleri belirli değildir. Bütün bu saydıklarımız kümeleme işlemi sırasında dinamik olarak oluşturulurlar. Bir kümelemenin başarılı olmasının şartları, aynı kümeye alınan dokümanların birbirine benzer ve diğer dokümanlardan da uzak olması ile sağlanır. Gerekğinde ayrık ya da çakışan bölümler yapılabilir. Bir doküman birden fazla kümenin içerisinde yer alabilir. Doküman kümeleme işlemlerinde en çok kullanılan algoritma türleri Hiyerarşik kümeleme ve bölümlenmeli kümeleme algoritmalarıdır.

3.2.1.1 Hiyerarşik Kümeleme

Hiyerarşik kümelemede, her doküman başlangıçta kendi kümesindedir. Hiyerarşik algoritmalar, bu kümeleri birbirlerine yakınlıklarına göre başarılı bir şekilde birleştirerek ilerlerler ve dokümanları hiyerarşik bir yapıda göstermeye çalışırlar. Toplaşım (Agglomerative) ve Bölünür (Divisive) olmak üzere iki türe ayrılırlar.

Toplaşım: Aşağıdan yukarıya giden bir yaklaşımdır. Her doküman kendi kümesinde başlar. Uygun küme ikilileri birleştirilerek hiyerarşik olarak bir üst seviyeye çıkılır.

Bölünür: Yukarıdan aşağıya doğru giden bir yaklaşımdır. Başlangıçta tüm dokümanlar tek bir kümenin içerisinde. Özyinelemeli bir şekilde, kümeler bölünerek, hiyerarşide altlara doğru inilir.

Hiyerarşik kümenin avantajları, ilerlemek için bir kaç küme verilmesine gerek olmaması ve sonuçta oluşan kümelerin, gözlemlenebilir bir yapıda olmasıdır. Dezavantajları ise, hiyerarşik algoritmaların ölçümlendirilebilir olmamasıdır. Bu yüzden gerçek zamanlı uygulamalar veya büyük derlem (corpus)'ler için uygun değildir.

3.2.1.2 Bölümlemeli Kümeleme

Bölümlemeli metotların en bilinen örneği k-Means algoritmasıdır. Klasik k-Means algoritması işe ilk önce rastgele k adet dokümanı başlangıç kümeleri olarak seçerek başlar. Tekrarlamalı olarak, dokümanları bu kümelere atarken, bu kümelerin ağırlık merkezlerini de günceller. Yakınsama oluştuğunda ve artık kümelerin ağırlık merkezleri çok fazla değişmemeye başladığında işlem tamamlanır. Yukarıdan aşağıya doğru bir yaklaşım ile hareket edilerek, kümeler bölünerek işlemlere devam ettirilir. k-Means algoritmasının birçok değişik versiyonu bulunmaktadır.

3.2.2 Arama Sonucu Kümeleme

Arama sonucu kümeleme işlemi, doküman kümeleme işleminin alt başlıklarından birisidir. Arama sonucu kümeleme algoritmalarından, sınırlı girdileri olan sezgisel açıklamaları kullanarak hızlı bir şekilde kümeleme yapmaları beklenmektedir. Arama sonucu kümeleme algoritmalarına verilen girdiler genelde yalnızca bir iki cümle uzunluğunda olurlar. Diğer kümeleme işlemlerinden farklı olarak, girdi olarak küçük metin parçalarının kullanılması, arama sonucu kümeleme algoritmasının başarısına etki edecek en önemli faktördür.

Arama sonucu kümeleme problemini tanımlamak istersek, bunu daha iyi bir biçimde maddeler halinde şu şekilde anlatabiliriz [25]:

- Her arama sonucu, internette bir dokümanı temsil etmektedir. Bir arama sonucu, dokümanın URL'si, başlığı ve dokümanın içeriğiyle ilgili kısa bir yazı barındırır. Bu kısa içerik parçası, dokümanın içeriğinde geçmiyor olsa bile,

dokümanın konusuyla ilgili bir içerik barındırır. Bazı dokümanlar için bu kısa yazı parçası, arama motorlarından döndürülmeyebilir ya da hiç bir cümlesel yapı içermeyen çok kısa yazılar olabilir.

- Bir doküman, birden fazla başlığa ait olabilir. Başlıklar anlamsal gruplardır. Başlıklar birbirleri arasında ilişkiler barındırabilirler ve başlıklar arasında bir hiyerarşi yapısı da olabilir.

Genel olarak bir arama sonucu kümeleme algoritmasından beklenenleri şöyle sıralayabiliriz [25]:

- Algoritma hızlı olmalıdır. Doğrusal zamanda çalışması ve artırımsal olması da tercih edilir.
- Algoritma küçük metin parçalarıyla bile iyi sonuçlar üretebilmelidir.
- Algoritma, çakışan kümelere olanak tanınmalıdır. Bir arama sonucu birden fazla kümede yer alabilmelidir ve genelde birden fazla başlık içerebilmelidir.
- Küme etiketleri sezgisel olmalıdır. Etiket, kullanıcının kümenin içeriğinde neler olduğunu anlamasını sağlamalıdır.

Klasik doküman kümeleme ve sınıflandırma problemleriyle karşılaştırdığımızda, arama sonucu kümeleme işlemleri arasındaki en önemli farklılıklardan birisi de, algoritmanın başarısı ölçülmek istendiğinde ortaya çıkar. Klasik metotlar genelde; test verilerinin doğruluğunun matematiksel ispatı, Tutturma ve Bulma gibi yöntemlerle kalite ölçümü yapabilirler. Arama sonucu kümeleme işleme ise tamamıyla kullanıcı odaklı bir alandır. Oluşan kümelerin başarılı olup olmadığı subjektif gözlemlerle elde edilebilir. Arama sonucu kümeleme işleminin başarılı olup olmadığı sadece matematiksel yöntemler kullanılarak doğrulanamaz. Başarıyı test edebilmek için insan gözlemleri de işin içine girmelidir.

Arama sonucu kümeleme algoritmaları, problemi genelde kelimelerin ya da kelime gruplarının frekansları gibi istatistiksel bilgilerle çözmeye çalışırlar. Fakat insan dilleri, bu istatistiksel dağılımı saptıracak birçok yapı içerirler. Eş anlamlı kelimeler, amaçlı bir şekilde yapılmış kelime sırası değişimleri, zamir kullanımları vb. durumları buna örnek

olarak verebiliriz. Bu yapılar dilden dile farklı olabilir ya da farklı etkiler yaratabilirler. Bu nedenle, kümeleme işleminde, kullanılacak dile özgü dil bilgisi bilgilerini kullanabilen algoritmalar geliştirilmelidir.

3.2.3 Kümeleme Arayüzü

Arama sonuçlarını kümeleme işlemi, arama sonuçlarını konularına göre web dizinlerinin yaptığına benzer bir şekilde organize etme işidir. Fakat bu işlem insan yapımı web dizinlerinden birçok noktada ayrılır. Öncelikle, kümeleme işleminde sadece arama sonucundan dönen veriler konularına göre gruplanır. Kümeleme işleminde oluşturulan konusal gruplar, önceden belirlenmiş sabit konular değildir. Konusal gruplar, dinamik bir şekilde, arama sonucundan elde edilen sonuçlara göre oluşturulurlar. Ayrıca, kümeleme arayüzü arama motorunun bir parçası olarak çalışmalıdır. Kümeleme işlemi; arama sonuçları elde edilir edilmez, online, hızlı ve verimli bir biçimde yapılmalıdır. Bu sebeple, kümelenecek dokümanların tamamının indirilip, işlenmesi kabul edilemez. Bunun yerine arama motorunun sonuçlarında döndürülen internet sayfasının ufak parçaları (snippet) kullanılmalıdır. Kümeleme arayüzü, yeni bir algoritma sınıfının ortaya çıkmasına neden olmuştur. Bu da bilgi getirme sonrası doküman sınıflandırma algoritmalarıdır. Bu sınıftaki algoritmaların başarılı olmasına etki edecek faktörler şunlardır:

Uygunluk: Algoritma'nın dokümanları gruplayarak oluşturacağı kümeler, kullanıcının sorgusuna uygun olmalıdır.

Gözlemlenebilir Özetler: Kullanıcı sadece göz gezdirerek, oluşturulan kümenin ilgi alanına uyup uymadığını görebilmelidir. Algoritma kümelerin kısa ve doğru açıklamalarını içermelidir.

Çakışan Kümeler: Bir dokümanın birden fazla başlığı olabileceği için, bir doküman birden fazla kümeye dahil olabileceğini de göz önünde bulundurmak gerekir.

Özet Metin Toleransı: Yalnızca arama motorundan döndürülen özet metinlerin (snippet) kullanmasına rağmen algoritma kaliteli kümeler oluşturmalıdır. Çünkü, kullanıcıların çoğu tüm dokümanların tek tek indirilip işlenmesini beklemek istemez.

Hız: Algoritma, arama sonuçlarının döndürülmesi sonrasında online bir işlem olarak çalışacağı için, kullanıcıya fark edilebilir bir gecikme yaşatılmamalıdır. Kümelemenin bir amacı da standart sıralı listeli görünüme göre, kullanıcıya daha kısa zamanda daha çok dokümana göz gezdirme imkânı sağlamasıdır.

Arttırımsal İşleme: Zaman kazanmak için, algoritma bir snipet'ı internetten çeker çekmez diğerlerini beklemeden işlemeye başlamalıdır.

3.3 Doküman Benzerliği Hesaplanması

Dokümanların benzerliklerinin hesaplanabilmesi için dokümanların Vektör Uzay Model (Vector Space Model) üzerinde gösterilmesi gerekmektedir. Böylece dokümanlar, üzerlerinde hesaplama işlemleri yapılabilecek cebirsel bir temele oturtulmuş olurlar. Vektör Uzay Modeli, dokümanlar ve benzeri objeleri belirleyici özellikleri üzerinden vektörel düzlemde göstermeye yarar. Dokümanlar için örnek vermek gerekirse bu özellik, doküman içerisinde geçen terimlerin frekansları ya da indeksleri olabilir.

Vektör Uzay Modelde her doküman bir vektör ile gösterilir. Doküman setindeki her ayrı kelime vektörlerin boyut sayısını belirler. Eğer bir kelime, dokümanda geçiyorsa bu kelimeye karşılık gelen boyut, doküman vektöründe genelde sıfırdan farklıdır. Sadece terimin doküman içerisinde geçip geçmemesine veya kelimenin doküman içerisindeki frekansına dayanan birçok farklı ağırlıklandırma yöntemi vardır. Bu yöntemler ile doküman vektörleri üzerinde puanlama yapılır. Bu konuda en çok kullanılan yöntemlerden biri TF-IDF ağırlıklandırma yöntemidir.

3.3.1 TF-IDF

TF-IDF iki istatistik değerin birlikte kullanılarak yapılan hesaplamalar sonucu ortaya çıkar. Bu iki istatistik değer: Term Frequency (Terim Frekansı), ve Inverse Document Frequency (Ters Doküman Frekansı)'dır.

Terim Frekansı: Terim Frekansı bir çok farklı şekilde kullanılmasına rağmen en çok kullanılan haliyle;

t terim, d Doküman olmak üzere, terim frekansı (3.1)'deki gibi hesaplanır.

$$tf(t, d) = f(t, d) \quad (3.1)$$

Burada $f(t,d)$ verilen d dokümanında, t teriminin kaç kere geçtiğini göstermektedir. Diğer bir gösterimde ise; eğer terim, dokümanın içerisinde en az bir kere geçiyorsa 1, aksi takdirde 0 döndürülür.

Ters Doküman Frekansı: Ters Doküman Frekansı bir terimin, doküman setinde genel kullanılan kelime mi yoksa sadece az sayıda dokümanda mı yer aldığını göstermektedir. Bu değer, veri setindeki toplam doküman sayısının terimi içeren doküman sayısına bölümünün logaritmasının alınmasıyla hesaplanmaktadır. t terim, d doküman ve D doküman seti olmak üzere, Ters Doküman Frekansı (3.2)'de gösterildiği şekilde hesaplanır.

$$idf(t, D) = \log \frac{|D|}{|\{d \in D: t \in d\}|} \quad (3.2)$$

TF-IDF: TF-IDF değeri ise TF ve IDF değerlerinin çarpılmasıyla elde edilir. TF-IDF değeri yüksek olan bir terim, bakıldığı doküman içerisinde yüksek frekansta, tüm doküman seti içerisinde de düşük frekansta olmalıdır. Böylece TF-IDF ile, doküman için belirleyiciliği yüksek olan kelimelere daha yüksek değerler verilmiş olur (3.3).

$$tf - idf(t, D) = tf(t, d) * idf(t, D) \quad (3.3)$$

3.3.2 Kosinüs Benzerliği

Bir doküman setindeki dokümanlar Vektör Uzay Model'de gösterildikten sonra, artık bu değerler yardımıyla doküman benzerlik algoritmalarına gönderilebilirler.

Vektörel benzerlik için en çok kullanılan algoritmalarından biri Kosinüs Benzerlik (Cosine Similarity) ölçüm algoritmasıdır. Kosinüs benzerliği kabaca, iki vektör arasındaki Kosinüs değerinin ölçümüdür. A ve B iki vektör olmak üzere, Kosinüs benzerliğini (3.4) eşitliğinde gösterildiği gibi formüle edebiliriz.

$$similarity = \cos(\theta) = \frac{A.B}{||A|| ||B||} = \frac{\sum_{i=1}^n A_i x B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} x \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (3.4)$$

3.4 Son Ek Ağacı Kümeleme

Son Ek Ağacı Kümeleme, Zamir vd. çalışmasıyla [6] ortaya çıkan; hızlı, artırımsal işlemeye olanak veren, çakışan kümeleri oluşturma yeteneğine sahip, kısa doküman parçalarıyla çalışmaya elverişli bir algoritmadır. Bu sebeple, ilk geliştirilmeye başlandığı 1997 yılından itibaren, arama sonucu kümeleme çalışmalarında sıklıkla kullanılan bir yöntem olmayı başarabilmiştir. Son Ek Ağacı Kümeleme algoritması, merkezinde genelleştirilmiş son ek ağacı kullanan bir algoritmadır. Son Ek Ağacı doğrusal zamanda oluşturulabilmektedir. Oluşturulduktan sonra artırımsal bir biçimde de kullanılabilir. Ayrıca, karakter katarı işlemlerinin hızlı bir şekilde yapılmasına da imkan tanır.

Geleneksel Son Ek Ağacı Kümeleme algoritması üç aşamada çalışmaktadır:

İlk aşama; dokümanların temizlenmesi aşamasıdır. Bu aşamada, dokümanları temsil eden metin parçaları basit bir kök bulma algoritmasından geçirilir. Kelimenin ön/son ekleri silinir, çoğul olan kelimeler tekil hale çevrilir. Noktalama işaretlerinden yararlanılarak cümlelerin sınırları belirlenir. Cümle sonlarında yer almayan noktalama işaretleri kaldırılır.

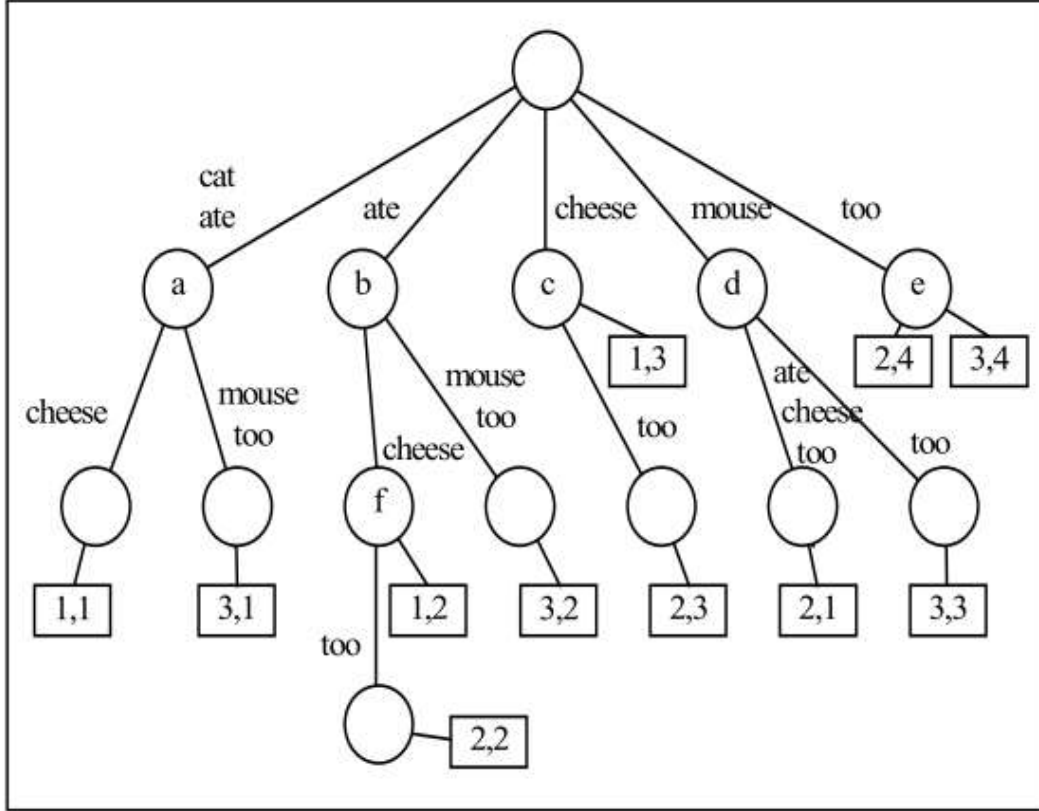
İkinci aşama; temel kümelerin belirlenmesi aşamasıdır. Bunun için de dokümanlardaki kelimeleri kapsayan bir Son Ek Ağacı oluşturulması gerekmektedir. Son Ek Ağacı'nın tanımını şöyle yapabiliriz: *"Bir S karakter katarının son ek ağacı S'nin tüm son eklerini içeren bir compact trie'dir"* [6].

Zamir vd. [6]'nin SAK için kullandığı son ek ağacında; son ekler yerine kelimeler, karakter katarları yerine de dokümanlar temsil edilmektedir. Böylece son ek ağacı, karakter katarı üzerinde harf veya harf dizilerine ulaşmak yerine doküman içerisinde geçen kelime veya kelime gruplarına ulaşmak için kullanılabilir hale getirilmiştir.

1. Son Ek Ağacı köklü, yönlendirilmiş bir ağaçtır.
2. Her iç düğüm, en az iki çocuk düğüme sahiptir.
3. Her *kenar*, S'nin boş olmayan bir alt kümesiyle etiketlenmelidir.
4. Aynı düğümden çıkan iki *kenar*, aynı kelimeyle başlayan etiketlere sahip

olmamalıdır.

5. S'nin her s son eki için, etiketi s'ye eşit olan bir son ek düğümü bulunmalıdır.



Şekil 3.2 “cat ate cheese”, “mouse ate cheese too” ve “cat ate mouse too” dokümanları için oluşturulan son ek ağacı [6]

Şekil 3.2’de verilen dokümanlarla oluşturulan Son Ek Ağacı gösterilmektedir. Burada daire ile düğümler gösterilmektedir. Her sonek düğümüne bir ve birden fazla kutu iliştilmiştir. Bu kutulardaki ilk sayı düğümün hangi dokümanda bulunduğunu; ikinci sayı ise, düğüm etiketinin dokümanın kaçınıcı kelimesinden başladığını göstermektedir. Son Ek Ağacındaki her düğüm, aynı kelime ya da kelime grubunu paylaşan bir grup dokümanı temsil etmektedir. Böylece aynı kelime grubunu içeren dokümanlar bir araya getirilerek temel kümeler tanımlanmış olur. Şekil 3.2’de gösterilen Son Ek Ağacından oluşturulan temel kümeler Çizelge 3.1’de görülebilir.

Çizelge 3.1 Şekil 3.2'de gösterilen Son Ek Ağacı için bulunan temel kümeler.

Düğüm	Etiket	Dokümanlar
A	cat ate	1,3
B	ate	1,2,3
C	cheese	1,2
D	mouse	2,3
E	too	2,3
F	ate cheese	1,2

Oluşan her temel küme için Eşitlik 3.5'deki gibi bir skor hesaplanır.

$$s(B) = |B| \cdot f(|P|)$$

(3.5)

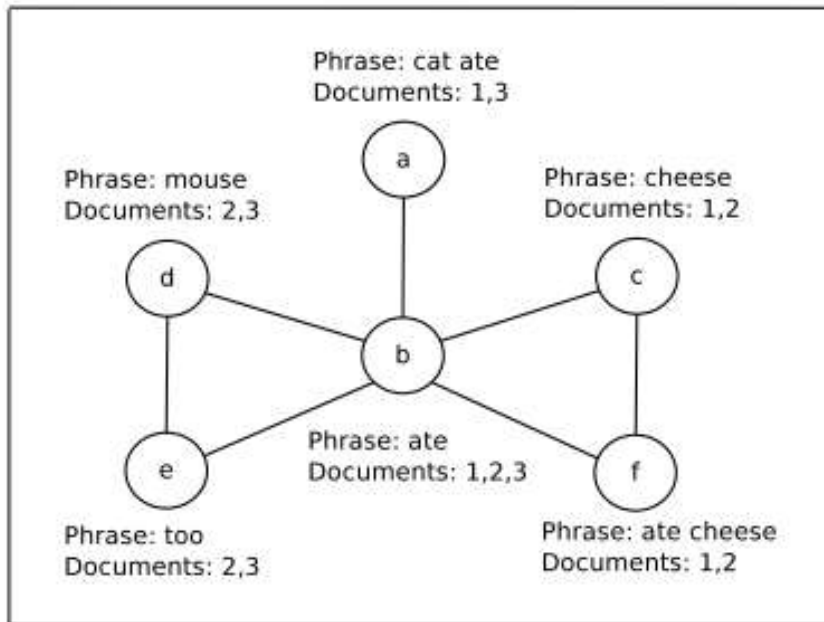
Burada $|B|$, temel kümedeki doküman sayısı, $|P|$ ise kümenin etiketindeki kelime sayısıdır. Çok az sayıda dokümanda ya da çok fazla sayıda dokümanda bulunan kelimeler için ya da önceden belirlenen durak listesinde yer alan kelimeler için $|P|$ değeri daha düşük tutulur.

Üçüncü aşama; temel kümelerin birleştirilerek sonuç kümelerinin oluşturulması işlemidir. Dokümanlar birden fazla kelime grubu içerebilirler. Bunun bir sonucu olarak da, birbirinden farklı temel kümelerde bulunan doküman setleri kesişebilir. Hatta, birbirine eş temel kümeler bile oluşmuş olabilir. Bu durumun yaygın bir şekilde görülmesini engellemek için üçüncü aşama, yani temel kümelerin birleştirilmesi aşaması uygulanır. Temel kümelerin birbirleriyle ne kadar kesiştiğini öğrenebilmek için, ikili bir benzerlik ölçümü yapılır. Verilen iki temel küme B_m ve B_n olsun. Bunlardaki doküman sayıları da $|B_m|$ ve $|B_n|$ olarak gösterilsin. $|B_n \cap B_m|$ de iki kümede birden olan Doküman sayıları belirtilir. Eğer:

$$|B_m \cap B_n| / |B_m| > 0.5 \text{ ve}$$

$|B_m \cap B_n| / |B_n| > 0.5$ ise, iki temel küme arasındaki benzerlik değeri 1 olur. Aksi takdirde bu değer 0 olur.

Tüm temel kümeleri arasındaki benzerlik ölçüleri hesaplandıktan sonra, temel kümeler çizgesi oluşturulur. Burada birbiri arasındaki benzerlik ölçümü 1 olan temel kümeler birbirine bağlı gösterilirler. Şekil 3.3’de örnek olarak kullandığımız veri seti için temel kümeler çizgesi gösterilmektedir. Birbirine bağlantısı olan temel kümeler birleştirilerek, sonuç kümeleri elde edilmiş olur. Bağlı temel küme skorlarının toplamı da sonuç kümelerinin skorunu belirler. Genellikle, oluşturulan sonuç kümelerinden en çok 10 tanesi kullanıcılar için ilgi çekici başlıklardır. Bu sebeple, kullanıcıya tüm sonuç kümeleri değil, en yüksek skora sahip sınırlı sayıdaki sonuç kümesi gösterilir.



Şekil 3.3 Oluşturulan temel küme graph'ı [6].

Eğer burada *ate* kelimesi durak kelimeler listemizde olsaydı, bu temel kümenin skoru 0 olacağından, b temel kümesi, temel kümeler listesinde olmayacaktı. Bu durumda, elimizde üç sonuç kümesini gösteren üç adet bağlı eleman olacaktı.

SAK algoritması artırimsal bir algoritmadır. İnternette dokümanlar alındıkça, ilk aşamadan geçirilerek Son Ek Ağacına yerleştirilirler. Yeni düğüm bilgileri de bu değişikliğe göre değişir. Düğümlerdeki değişikliklerden dolayı, temel kümeler de

değişebilir. Eğer değişen kümeler varsa, bunların benzerlik değerleri yeniden hesaplanır. Son olarak da, bu değişikliklerin sonuç kümeleri etkileyip etkilemediğine bakılarak, gerektiğinde sonuç kümelerinde de değişiklik yapılır. Bu işlemler, yeni dokümanlar sisteme eklendikçe tekrarlanır. Benzerlikler hesaplanırken, işlemin daha verimli ve hızlı bir şekilde çalışabilmesi için değişen temel kümeler, diğer tüm temel kümelerle değil de sadece en yüksek skora sahip k temel küme ile karşılaştırılır. Daha sonra yayınlanan bir makaleyle [12], Zamir vd. *Grouper* adını verdikleri sistemi tanıtmış, burada SAK'nin üçüncü aşaması olan temel kümeleri birleştirme aşamasında geliştirmeler yapmıştır ve böylece daha iyi küme etiketlerinin oluşturulması sağlanmıştır. Bundan sonra da SAK tekniği üzerinde birçok farklı makalede iyileştirme çalışmaları üzerine araştırmalar yapılmıştır. Özellikle algoritmanın üçüncü aşamasına yenilikler getirilmeye çalışılmıştır.

3.5 Zemberek

Zemberek [11], açık kaynak kodlu, tüm Türki diller için kullanılması hedeflenen, Türkçe doğal dil işleme kütüphanesidir. Zemberek, Türkçe için aşağıda listelediğimiz temel doğal dil işleme işlemlerini gerçekleştirebilme yeteneğine sahiptir:

- Yazım denetimi
- Hatalı kelimeler için öneri sistemi
- Kelimeyi hecelerine ayırma
- Kelime oluşturma
- Biçimbirimsel (Morfolojik) ayrıştırma
- Kelime kökü bulma
- Sadece ASCII olarak yazılmış kelimelerin dönüştürülmesi (deasciifier)

Zemberek kütüphanesi iki temel kısımdan oluşmaktadır. Bunlar: dil yapısı bilgisi ve doğal dil işleme işlemleridir. Farklı diller için dil yapısı bilgileri sonradan sisteme eklenerek, bu diller için de kütüphane çalışır hale getirilebilir. Fakat özellikle Türki diller

iin bir kütüphane oluşturulmak istendiğinden, farklı kökenli dillerin kullanılması problem yaratabilir.

Zemberek kütüphanesi, Java programlama diliyle yazılmıştır. Türke ile ilgili bir ok akademik alıřmada ve OpenOffice (<http://www.openoffice.org>), Firefox (<http://www.mozilla.org/firefox>), Pardus (<http://www.pardus.org.tr/>) gibi birok aık kaynak kodlu yazılımlarda kullanılmaktadır. Zemberek kütüphanesinin kendisi de aık kaynak kodlu bir uygulamadır.

3.6 Carrot2

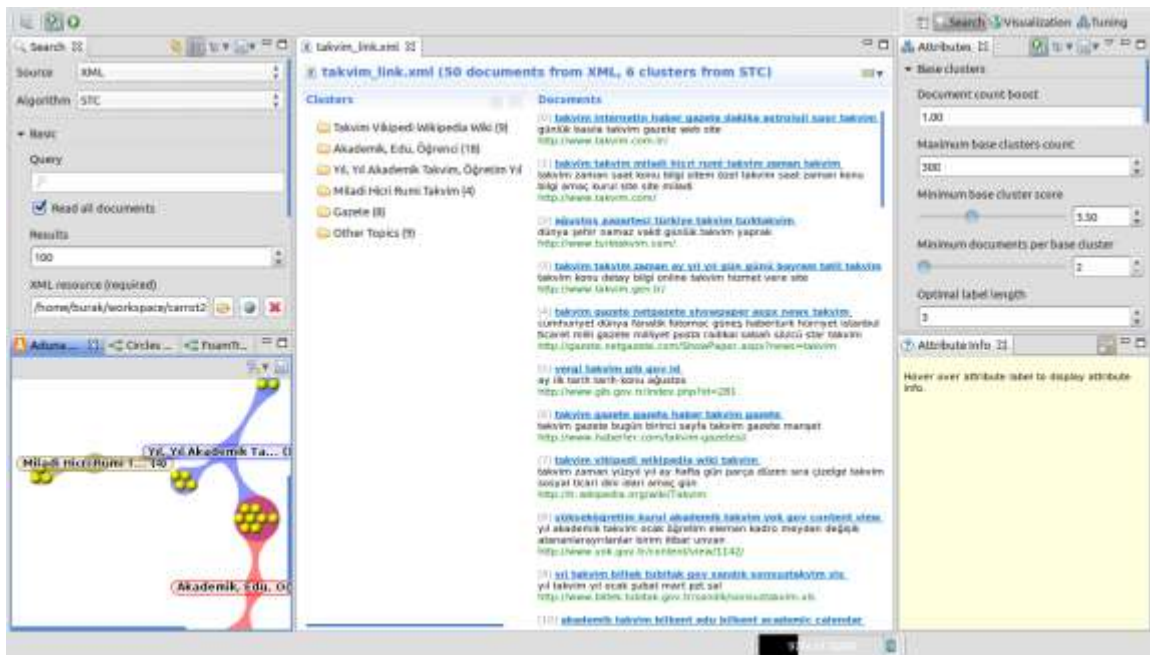
Carrot2, aık kaynak kodlu arama sonucu kümeleme problemine yoğunlaşmış bir web bilgi toplama ve web madencilik çatısıdır (framework) [26]. Carrot2, arama sonucu kümeleme problemi iin kullanılan popüler yöntemlerden, k-Means, SAK ve Lingo'nun uygulamalarına sahiptir. Ayrıca URL'ye ve kaynağā göre kümeleme seçeneklerine sahiptir. Carrot2 kullanacağı veriyi bir ok farklı kaynaktan alabilmektedir. Veriler etoolsh.ch (<http://www.etools.ch/>), Wikipedia (<http://www.wikipedia.org/>), Bing (<http://www.bing.com>), PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) gibi internet üzerinde bulunan farklı kaynaklardan, sağlanan arayüz sayesinde aranarak elde edilebilir. Ayrıca, Carrot2 sistemi tarafından detayları belirlenen bir XML formatı halinde de kümeleme işlemi yapılacak veriler sisteme aktarılabilir.

Carrot2, metin kümeleme alıřmalarında kullanılan birok işlemin uygulamasını da kullanıcılara sağlamaktadır (Şekil 3.3). Bunlar:

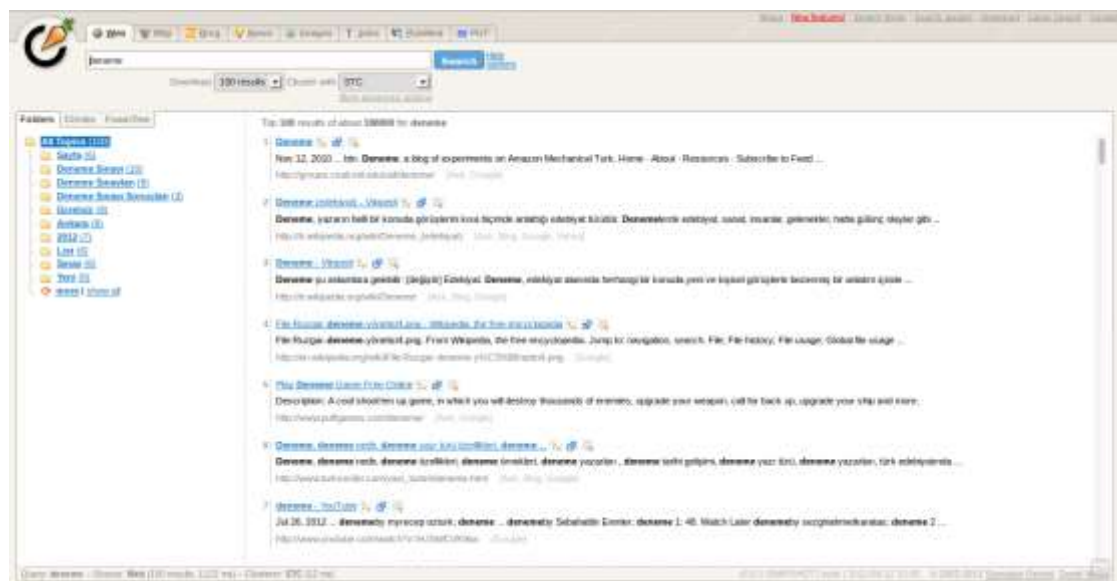
- Verimli, JFlex tabanlı (<http://jflex.de>) metin ayrıştırıcı
- Trigram tabanlı dil tanıma
- Bir ok dil iin durak kelimeleri silme ve kök bulma mekanizması
- Bazı test veri tabanlarına erişim. Ör: Open Directory Project (<http://dmoz.org>)
- Sonuçların HTML tabanlı gösterimi, otomatik kalite ölçümü.

Araştırma projelerinde kullanılmak üzere tasarlanmış olan Carrot2'in, bu amaca hizmet etmesi iin bazı özellikleri sağlaması gerekmektedir. Carrot2'nin sağladığı bu özellikleri şöyle sıralayabiliriz:

Carrot2, Eclipse (<http://www.eclipse.org>) tabanlı bir uygulama ve bir web uygulaması içermektedir (Şekil 3.4).



Şekil 3.4 Carrot 2 workbench



3.7 Google Arama Motoru

Google, günümüzde dünya çapında en çok kullanılan internet arama motorudur. Her gün Google üzerinde dünya genelinde milyarlarca sorgulama yapılmaktadır ve yine milyarlarca internet sayfası Google'ın veritabanında indekslenmektedir. Klasik Google arama motorunun amacı, internet üzerindeki internet sayfalarının içerisindeki metinleri indekslemek ve daha sonra bunlar üzerinden arama yapılmasını sağlamaktır. Daha sonra, internetteki resimler üzerinden arama yapmayı sağlayan “Google görsel arama” gibi farklı alanlarda da indeksleme yapabilen özel alt birimleri de oluşturulmuştur.

Google arama motoru ilk olarak 1997 yılında, Larry Page ve Sergey Brin tarafından bir tez çalışması olarak geliştirilmeye başlanmıştır. Arama sonuçlarının sıralanması için bir puanlama yöntemi olan PageRank [23] algoritması ile Google, diğer internet arama motorlarına kıyasla, kullanıcın sorgusuyla isteyebileceği sonuçları daha kolay erişebilecekleri yerlerde gösterebilmiş, böylece hızlı bir şekilde popülerliğini arttırarak, ezici bir farkla en popüler internet arama motoru olmayı başarmıştır.

3.7.1 Google Custom Search API

Google, projelerinin farklı ortamlarda rahatça kullanılması ve kendi araçlarını kullanan bir çok uygulamanın geliştirebilmesini kolaylaştırmak için bir çok API(Application Programming Interface) sunmaktadır. Custom Search API, geliştiricilere Google arama motoru üzerinde özelleştirilebilir aramalar yapabilmeleri için sağladıkları bir API'dir[27]. Geliştiriciler, bu API'yi kullanarak isterlerse sadece tanımladıkları internet siteleri üzerinde arama yapabilirler. Aramayı sadece istedikleri dilde yapabilirler. Dönen arama sonucu sayısını sınırlayabilirler.

Google Custom Search'ün sağladığı JSON/Atom API'si ile geliştiriciler arama sonuçlarını birçok programlama dilini kullanarak, programlanabilir bir şekilde elde edebilirler. RESTful [28] yöntemiyle kullanıcıların API'ye yaptıkları isteklerden, seçime göre JSON (<http://json.org>) ya da Atom (<http://www.atomenabled.org/>) formatında sonuçlar

döndürülür. API tarafında kullanıcıya, her arama sonucu ile ilgili detaylı bilgiler döndürülür. Web sayfasının başlığı, bağlantının adresi, sayfanın içeriği hakkında bilgi veren kısa bir metin, sayfanın Google sunucularında tutulan önbellekli halinin *id*'si, dosya formatı gibi bilgiler kullanıcıya döndürülür.

3.8 Doküman Kümelemede Ön İşleme Adımları

Klasik doküman kümeleme işlemleri yapılırken, düz yazının belirli bir ön işlemten geçilerek gereksiz kısımların çıkarılması ve verinin sayısal gösterimlere daha uygun bir hale getirilmesi gerekmektedir. Bu sayede hem işlemdeki gereksiz bilgiler çıkarılarak kümeleme işlemi daha hızlı hale getirilecektir, hem de bilgi kirliliği yaratacak verilerin elenmesiyle daha doğru sonuçlar elde edilecektir. Bu sebeple doküman kümeleme işlemlerinde bazı ön işleme adımları uygulanmaktadır. Bunlardan en çok kullanılanları şöyle sıralayabiliriz:

Filtreleme: Düz yazı içerisinde özel karakterlerin ve noktalama işaretlerinin çıkarılması işlemidir. Bu tip karakterler hiç bir ayrıştırıcı değer teşkil etmezler. İnternet sayfaları gibi belirli şekilde formatlanmış yazılarda bu işlem daha da önemlidir.

Ayrıştırma(Tokenization): Cümleler kendisini oluşturan kelimelere ayrıştırılır. Daha karmaşık yöntemler uygulanarak belirli dilbilgisi kurallarına göre kelime grupları da alınabilir.

Kök Bulma: Kelimelerin kök hallerinin alınması işlemidir. Kelimelerin sonundaki takılar, ekler kaldırılarak, kelimelerin kökleri üzerinden işlem yapılması sağlanır. Böylece sistemde aynı kökene sahip sözcükler, aynı kelimeler olarak gösterileceği için, sistemin genel başarımı arttırılabilir. Bu işlemi gerçekleştirebilmek için dillere özgü programlar ya da API'ler kullanılmaktadır. İngilizce için, diğer birçok dili de destekleyen Porter Stemmer [29] kullanılmaktadır. Türkçe için ise bu işlemleri en doğru bir biçimde gerçekleştiren, bölüm 3.6'da da bahsettiğimiz Zemberek kütüphanesidir.

Budama: Tüm dokümanlar içerisinde (derlem), en az geçen kelimelerin, kümeleme işlemine çok az etki ettiği için çıkarılması işlemidir. Genelde bu işlem için bir eşik değeri kullanılır. Derlem içerisinde çok az sayıda geçen kelimeler ya da çok fazla sayıda geçen kelimeler, belirleyici nitelikleri düşük olduğu için derlemden çıkartılırlar.

ÖNERİLEN YÖNTEM**4.1 Kullanılan Veriler**

Yaptığımız çalışmada, kendi oluşturduğumuz veri setini kullandık. Veri setimizi her biri rastgele seçilen tek kelimelerden oluşan 5 sorgu oluşturmaktadır. Seçtiğimiz kelimeler Çizelge 4.1’de verilmektedir. Sorgulardan *Balık* ve *Yüz*, geliştirme aşamasında yapılan çalışmalar sırasında kullanılmış Geliştirme Veri Setini oluşturmaktadır. *Takvim*, *Dünya* ve *Ocak* sorguları ise deney veri seti olarak kullanılmıştır.

Çizelge 4.1: Seçilen Kelimeler

Kelime	Doküman Sayısı	Etiket Sayısı	Etiket İsimleri
Balık	50	7	Oyun, Wiki, Avcılık, Burç, Restaurant, Diğer, Hobi
Yüz	50	8	Oyun, Wikipedia, Bakım Ürünleri, Yüz Nakli, Uzman TV, Diğer, Estetik, Facebook
Takvim	50	4	Akademik Takvim, Gazete, Wikipedia, Zaman
Dünya	50	5	Haber, Wikipedia, Dünya Mirası, Dizi, Şarkı
Ocak	50	4	Wikipedia, Pişirme, Döviz, Maden Ocağı

İnternet araması yapmak için dünya genelinde ve ülkemizde en popüler internet arama motoru konumunda olduğu için Google arama motorunu tercih ettik. Google, Türkçe veriler için de başarılı sonuçlar döndürülebilir konumdadır. Çalışmamızı Türkçe veriler üzerinde yapacağımızdan, Google’ın sunduğu verileri çalışmamız için yeterli kabul ettik. Ayrıca, Google’ın geliştiriciler için sunduğu araçlar ve API’lerinin de başarılı ve çalışmamızı kolaylaştıracak nitelikte olması sebebiyle çalışmamızda kullanılacak verileri elde etmek için Google arama motorunu tercih ettik.

Kullanacağımız verileri Google’dan çekebilmek için Google Custom Search API’sini kullandık. API’ye dil parametresi olarak ‘tr’ vererek, sadece Türkçe içeriğe sahip

sayfaların döndürülmesini sağladık. Sorgu parametresine de seçtiğimiz kelimeleri sırası ile girdik. Her sorgu için, arama motorundan dönen ilk 50 arama sonucunu aldık. API'den dönen bilgilerden, her arama sonucunun URL'sini, sayfa başlığını ve sayfa hakkındaki kısa açıklamayı (snippet) çektik. Topladığımız bu sonuçları, kümeleme çalışmaları sırasında kullanacağımız veri setimizin özelliklerini oluşturmaktadır. Veri setimizde böylece her sorgu kelimesi için 50 URL bilgisi, sayfa başlığı ve kısa özet bilgisi tutuldu.

4.2 Son Ek Ağacı Kümeleme

Bölüm 3.4'de anlatılan SAK algoritmasının uygulamasını gerçekleştirebilmek için temel olarak Carrot2'de sunulan SAK algoritması kullanıldı. Ön işleme çalışmalarında ve yöntemin kendisinde, Türkçe diline daha uygun olacak bir yapı sunabilmek için bir takım değişiklikler gerçekleştirildi.

4.2.1 Ön İşleme Adımları

Doküman kümelemede uygulanan benzer ön işleme adımları SAK algoritmasında da uygulanmaktadır. Fakat SAK'de uygulanan yöntem daha basittir. SAK'nin işlediği veriler üzerinde noktalama işaretleri, kelime grupları gibi yazı özelliklerini kullanması sebebiyle bu basit yöntem uygulandı. Carrot2'de hali hazırda ön işleme adımları bulunmasına rağmen, kendi geliştirdiğimizi tercih etmemizin sebebi, Zemberek'i de kullanarak bu işlemi Türkçe dili için daha başarılı yapabileceğimizi düşünmemizdir. Aldığımız verileri algoritmaya vermeden önce sırasıyla; filtreleme, ayrıştırma, budama ve kelime hatalarının düzeltilmesi işlemlerinden geçirdik.

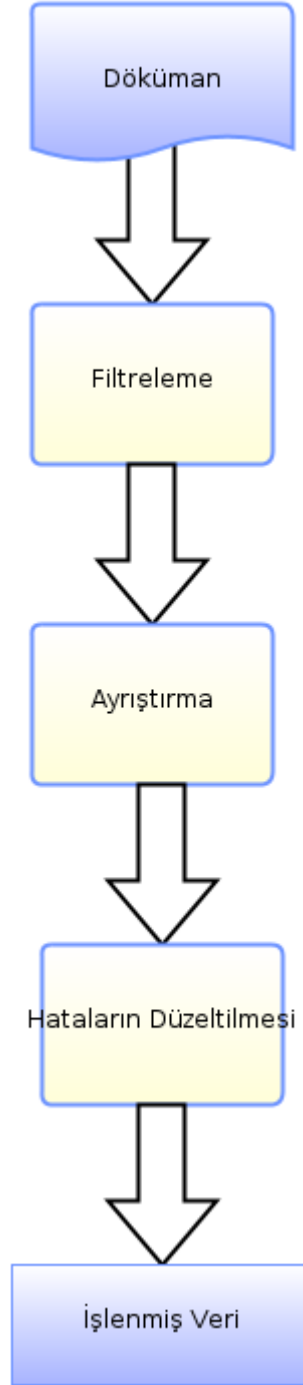
Filtreleme: Filtreleme aşamasında, cümlelerin sonuna gelenler (nokta, soru işareti, ünlem, üç nokta) dışındaki tüm noktalama işaretlerini sildik. Cümlelerin sonlarına gelen noktalama işaretlerini bırakmamızın nedeni, SAK algoritmasında cümle sınırlarının önem teşkil etmesinden dolayıdır.

Ayrıştırma: Ayrıştırma aşamasında; cümleleri, kelimelerine ayırdık. Eğer varsa, kelime sonlarındaki çoğul eklerini (-lar, -ler) çıkarttık. Çoğu kümeleme yönteminin ön işleme adımlarında genellikle kelimeler köklerine kadar ayrıştırılırlar. SAK algoritması, küme

etiketlerini doküman içerisinde ard arda geçen kelimeler ve kelime grupları ile oluşturmaktadır. Kelimelerin sadece köklerini almamız küme etiketleri için anlamsal kayıplara neden olabilirdi. Örneğin, '*balık avı teknikleri*' isimli küme etiketindeki kelimelerde kök bulma işlemi uygularsak, '*balık av teknik*' haline gelir. Bu da ilk haline göre daha az açıklayıcı olacaktır. Bu nedenden ötürü veri setimizdeki kelimeler üzerinde kök bulma işlemi uygulanmadı.

Budama: Budama işlemi de, yine SAK'deki kelime sırasına verilen önemden dolayı uygulanmadı. Eğer az kullanılan kelimeler ya da bağlaçlar çıkarılsaydı, olası kelime gruplarının yapısı bozulabilirdi. Bu da SAK'nin başarısını düşüren bir etken olacaktı. Bunun yerine, bu tip kelimelerin yarattığı gürültü puanlama aşamasında değerlendirildi.

Hataların Düzeltilmesi: Veri setinden elde ettiğimiz verilerde, kelimelerin hatalı yazılması (örneğin veri setinin bazı yerlerinde restaurant, bazı yerlerinde restoran şeklinde) ya da internet üzerinde sıkça karşılaşılan bir durum olan Türkçe karakter içeren bir kelimenin, sadece İngilizce karakterle yazılması (balık yerine balik yazılması) gibi hatalar mevcuttu. Bu tip hatalar algoritmanın başarımını da olumsuz yönde etkilemektedir. Bu tip hataları olabildiğince en aza indirmek için Zemberek'in yazım denetimi ve deasciifier kütüphaneleri kullanıldı. Eğer Zemberek üzerinden kök ve eklerine ayırtıramadığımız kelimeler varsa, bunları önce yazım hatası var mı diye yazım denetimi kütüphanesi ile, eğer yoksa deasciifier'dan geçirerek tekrar ayırtırmaya çalıştık. Eğer kelime hâlâ ayırtıramıyorsa bunun Türkçe kökenli olmayan bir kelime olduğuna karar verildi ve bunları olduğu şekilde ağaca ekledik. Bu aşamada kelimeleri düzeltebilmişsek, düzeltilmiş hallerini kullandık.

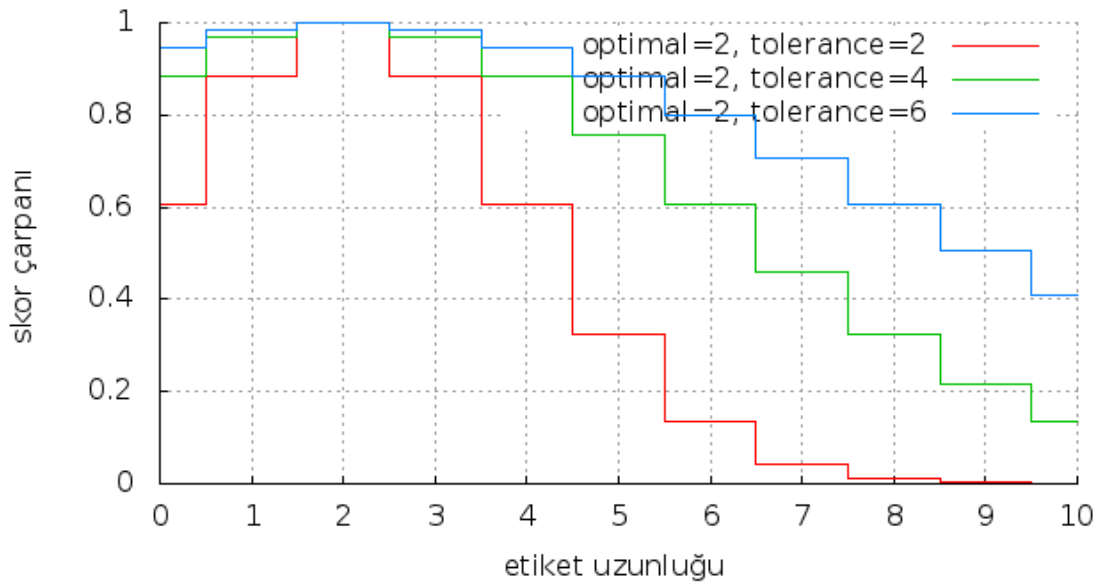


Şekil 4.1 Ön işleme adımları

4.2.2 Klasik SAK - KSAK

İlk önce üzerinde fazlaca değişiklik yapmadığımız bir haliyle SAK algoritmasının performansını görmek ve başarısını ölçmek istedik. Bu sebeple, Carrot2'de bulunan SAK gerçekleştirmesini kullandık. Bu gerçekleştirmede büyük ölçüde bölüm 3.4'de anlattığımızla aynı olmakla birlikte, bazı ufak farklılıklar da taşımaktadır.

Carrot2'nin SAK uygulamasında temel kümelerin skorlarının hesaplanması Zamir vd. [6]'nin geliştirdiği standart yöntemden farklılıklar içermektedir. Carrot2'de etiket etkisini hesaplayan formül farklı bir biçimde ele alınmıştır. Zamir vd. [6]'nin yönteminde etiket olarak seçilen kelime grubunun uzunluğuna ve etikette geçen kelimelerin durak listelerinde olup olmamasına göre, sabit bir değer döndürülmektedir. Carrot2 uygulamasında ise ideal etiket uzunluğu ve etiket bu ideal uzunluğa olan toleransın yoğunluğu üzerine kurulan bir Gaussian fonksiyonu kullanılmaktadır.



Şekil 4.2 İdeal etiket uzunluğu fonksiyonu

İdeal uzunluk ve tolerans parametreleri önceden tanımlanmaktadır. Bu fonksiyona göre, temel kümenin etiketi ideal değerdeyken en yüksek skoru alır. Tolerans değerine göre de ideal değerden uzaklaştıkça skor azalır Şekil 4.2'de ideal değer 2 olarak alınmıştır. Tolerans değerinin 2, 4 ve 6 olduğu durumlardaki değişim Şekil 4.2'den görülebilir. Tolerans değeri daha yüksek alınsaydı, skor daha geç düşüşe geçecekti. Carrot2'de ideal uzunluk ve tolerans için verilen varsayılan değerler, sırasıyla 3 ve 2 idi. Klasik Son Ek Ağacı Kümeleme adı altında yaptığımız çalışmada bu değerleri olduğu gibi bıraktık. Ön işleme kısmında, Carrot2'de uygulanan ön işleme kısmı yerine, Türkçe için

daha uygun olduğunu düşündüğümüz kendi geliştirdiğimiz ön işleme adımlarını 4.1’de belirttiğimiz şekilde kullandık. Ön işleme çalışmaları kısmında da belirtildiği gibi bu aşamada Zemberek kütüphanesinden yararlandık.

Bunun dışındaki kısımlarda, SAK uygulamamız Zamir vd. [6]'nin çalışmasında anlatıldığı biçimde çalışmaktadır.

4.2.3 Geliştirilmiş SAK - GSAK

Klasik SAK yönteminde olduğu gibi, geliştirilmiş SAK'de de Carrot2 içerisinde bulunan SAK gerçekleştirimi temel alınmıştır.

Ön işleme kısmında, klasik SAK'de olduğu gibi, Türkçe için daha uygun olduğunu düşündüğümüz, kendi geliştirdiğimiz ön işleme adımlarını kullandık. Bu adımlar ile ilgili detaylar bölüm 4.2.1’de anlatılmaktadır.

Klasik SAK algoritması üzerinde çeşitli deneyler yaparak, bu algoritmanın eksik kalan kısımlarını düzelterek ve algoritma adımlarına yeni eklemeler yaparak başarıyı arttırmak istedik.

Klasik SAK algoritması üzerinde yaptığımız deneylerde gördük ki; etiketleri ya da içerisindeki elemanları birbirine çok benzeyen farklı kümeler oluşmakta. Bu sebeple, klasik SAK yöntemiyle elde ettiğimiz sonuç kümeleri üzerinde işlem yapmak için yeni bir adım ekleyerek bu kümeleri birleştirmeye çalıştık. Eklediğimiz bu adımda, üç aşamada sonuç kümelerini inceleyerek gereksiz olanları çıkarmaya ve birbirine benzer kümeleri birleştirmeye çalıştık.

İlk aşamada, en az %70 oranında aynı elemanları taşıyan kümeleri tespit ederek bunları birleştirdik.

İkinci aşamada, skoru 5'in altında olan kümeleri tek tek inceledik. Eğer bu kümelerin elemanlarının %50'si diğer kümeler içerisinde yer alıyorsa, bu kümeyi bozduk. Kümenin, diğer kümelerde yer almayan elemanlarını, *Diğer* ile etiketlenmiş kümeye ekledik.

Son aşamada ise elimizdeki küme etiketlerini tekrar gözden geçirdik. Aynı kelimeleri içeren etiketler varsa, bu kümeler arasındaki eleman benzerliğine baktık. Eğer benzerlik

%50'nin üzerindeyse bu kümeleri de birleştirdik. Birleştirdiğimiz bu kümenin etiketine de, birleştirilen kümelerden skoru yüksek olanını atadık.

4.3 Distance Based Clustering

Distance Based Clustering (DBC) yöntemi, doküman kümeleme yöntemlerine benzer; fakat bunlardan daha basit bir yapıda olacak şekilde hazırlanmıştır. Bu yöntemin amacı, doküman kümeleme çalışmalarına uygun ve basitleştirilmiş bir uygulamanın arama sonucu kümeleme işlemlerinde başarısını ölçmektir. Bu yöntemin geliştirilmesi ile ilgili detaylardan, 4.3.1'de ön işleme adımları, 4.3.2'de yöntem olmak üzere bahsedilmiştir.

4.3.1 Ön İşleme Adımları

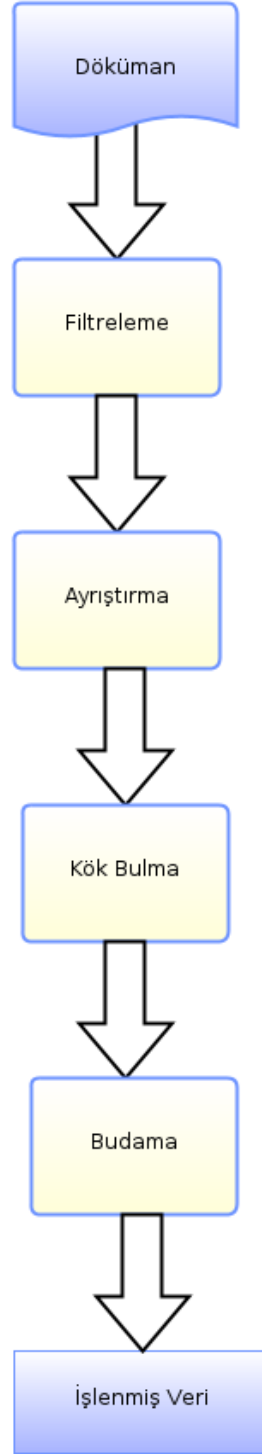
Uygulanan DBC yöntemi için, doküman kümeleme işlemlerinde genel olarak kullanılan tüm ön işleme adımlarını kullandık. Diğer yöntemlerde olduğu gibi veri seti, sayfa başlıklarından ve snippet'lardan oluşmaktadır. Bu iki kısmın bir araya getirilmesiyle ilgili sayfanın, kümeleme işleminde kullanacağımız tüm yazısı elde edilmiş olur. Şekil 4.3'de DBC algoritması için ön işleme adımları gösterilmektedir.

Filtreleme: Filtreleme aşamasında, yazı içerisindeki tüm noktalama işaretleri ve rakamlar çıkartıldı.

Ayrıştırma: Ayrıştırma aşamasında yazı tek tek kelimelere ayrıştırıldı.

Kök Bulma: Bu aşamada tüm kelimeler, Zemberek kütüphanesi kullanılarak, tek tek köklerine ayrıldı. Tüm kelimelerin kök halleri alındı.

Budama: Tüm derlem içerisinde, çok az sayıda geçen kelime varsa, bunların başarıma olan etkileri düşük olacakları için çıkartıldı. Ayrıca, önceden belirlenen durak kelimeleri-stop words (bağlaçlar, internet arama sonuçlarında sık kullanılan fakat bir açıklayıcılığı olmayan kelimeler: www, internet, vb.) çıkartıldı. Aynı şekilde tüm derlem içerisinde, dokümanların çoğunda geçen kelimeler varsa, bunların da belirleyicilikleri düşük oldukları için çıkarıldı.



Şekil 4.3 DBC Ön işleme adımları

4.3.2 Yöntem

Döküman kümeleme algoritmalarında kullanılan dokümanlar arası benzerlik yöntemlerine dayanan, basit bir kümeleme yöntemi geliştirmek istedik. Bununla

amacımız, arama sonucu kümeleme işlemleri için uygulanan SAK algoritmasıyla böyle bir algoritmanın başarısını kıyaslamaktır.

Arama sonucunda elde ettiğimiz bilgiler üzerinde benzerlik yöntemleri uygulamadan önce, bu veriler üzerinde 4.3.1'de anlatıldığı şekilde ön işleme çalışması uyguladık. Buradan elde ettiğimiz işlenmiş verileri, Vektör Uzay Model üzerinde göstermek için TF-IDF yöntemini kullandık. Her sayfanın TF-IDF değerlerini (3.3) eşitliğinde gösterildiği şekilde hesapladık. TF-IDF değerlerini kullanarak her bir sayfanın, diğer sayfalarla olan benzerliğini kosinüs benzerlik (3.4) yöntemi ile hesapladık. Hesaplanan bu sayfalar arası benzerlik değerlerini bir matris üzerine kaydettik.

Kümeleri oluşturma aşamasında, veri setimizdeki tüm sayfaları tek tek ele aldık. İncelediğimiz her sayfa için diğer sayfalarla arasındaki benzerlik değerlerine bakarak, benzerlik oranı belirlenen eşik değerden (0.20) daha yüksek olan sayfa varsa, incelediğimiz sayfayı, benzeşen sayfaların kümelerine ekledik. Burada belirlediğimiz eşik değeri, yaptığımız denemeler sonucunda en verimli olan değeri seçilmiştir. Eğer benzeşen sayfalar henüz herhangi bir kümeye dahil edilmemişse, incelenen sayfayı, benzeşen sayfalarla birlikte yeni bir küme oluşturup, buna dahil ettik. Bu şekilde tüm sayfaları sırayla tarayarak, benzerlik oranı birbirine yakın sayfaları aynı kümeler içerisine dahil ettik. Yapılan bu işlem sonucunda, eğer bir kümede verilen üç gibi (3) sabit bir değerden az sayıda sayfa varsa, bu küme bozulur. Buradaki sabit değeri olarak belirlenen 3, yaptığımız denemeler neticesinde ortaya çıkmıştır. Genellikle bu sabit değerin altındaki eleman sayılarıyla oluşturulan kümeler, ayrıştırıcı niteliğe sahip değildirler. Bozulan kümeler içerisindeki sayfalar, herhangi bir kümeye dahil edilemeyen sayfaların bulunduğu *Diğer* kümesine eklenir.

Bu şekilde tüm sayfaları kümelere yerleştirdikten sonraki aşamada, *Diğer* kümesi haricindeki kümelere, küme etiketlerini atadık. Bu işlem için, bir küme içerisindeki tüm sayfalarda en çok kullanılan kelimeyi tespit ederek, bu kelimeyi küme etiketi olarak verdik. En çok kullanılan kelimeyi tespit etmek için, küme içerisindeki sayfaların içerdiği kelimelerin TF değerlerini inceledik. En yüksek TF değerine sahip kelimeyi etiket olarak seçtik. Böylece etiket için, SAK yönteminin aksine, her zaman tek kelimeden oluşan bir etiket seçmiş olduk.

Etiket in de belirlenmesiyle, DBC kümeleme işlemi tamamlanmış oldu.

4.4 Melez GSAK Yöntemi (M-GSAK)

Sorgu karşılığında, sonuç olarak dönen sayfaları içeriklerine göre sınıflandırarak kullanıcıya sunmak için geliştirdiğimiz sistem için, en çok kullanılan arama sonucu kümeleme yöntemlerinden olan SAK'yi ve dokümanlar arası benzerlik ölçümlerine dayanan bir yöntem olan DBC'yi kullandık. SAK yönteminden elde ettiğimiz sonuçların başarısını daha yukarılara çekebilmek için, SAK'nin sonuçlarını DBC'nin sonuçlarıyla karşılaştırıp bu sonuçlardan yararlanabileceğimiz melez bir yöntem oluşturmaya çalıştık.

Bu melez yöntem için aşağıdaki çalışma adımları izlenmiştir:

Adım-1 Her iki yöntemde çıkarılan etiketler karşılaştırılır. GSAK'de olup, DBC'de olmayan etiketler belirlenir. Eğer iki yöntem tarafından da bulunmuş ortak etiketler varsa, bu etiketler melez işlem sonrası oluşturulan sonuç kümelerine aynen aktarılırlar. Sadece GSAK tarafından çıkarılmış etiketler ve bu etiketlerin gösterdiği kümeler, üzerlerinde işlem yapmak için seçilir.

Adım-2 Sadece GSAK yöntemi ile çıkarılmış her etiket için:

- Bu etiket içerisinde yer alan sayfalardan, GSAK'nin bulduğu diğer kümeler içerisinde olan var mı?
- Bu etiket içerisinde yer alan sayfalardan DBC'nin bulduğu diğer kümeler içerisinde olan var mı?

Kontrolü yapılır.

Adım-3 Adım-2'de değerlendirilen her sayfa için :

- Eğer, iki algoritmadan da aynı etiket geliyorsa, sayfanın yeni etiketi olarak bu atanır.
- İncelenen sayfa, eğer hem GSAK, hem de DBC'nin diğer kümelerinden biri içerisinde yer alıyorsa ve bu iki etiket birbirinden farklıysa, yeni etiket olarak, GSAK'nin genel başarıımı oranı DBC'den fazla olduğu için, GSAK'nin etiketi atanır.

- Eğer adres, sadece bir algoritmanın **diğer** kümelerinden birinde yer alıyorsa, sayfanın etiketi olarak bu küme atanır.

Sayfalara yeni etiketler belirlenirken; eğer bir sayfaya, iki algoritmanın sonucunda birden bulunmayan, bu yüzden de tekrar incelenmesi gereken, bir etiket atanmışsa, Adım-4 devreye girer.

Adım-4 Üzerinde işlem yapılan etiket ile işaretlenmiş sayfalar için yöntemlerden biri iptal edilecek aday etiketini veya boş etiketi ikinci etiket olarak veriyorsa, üzerinde işlem yapılan etiket kalıcı etikete çevrilir.

Uygulanan bu dört adım sonunda birleştirme işlemi tamamlanarak, M-GSAK sonuçları elde edilmiş olur.

DEĞERLENDİRME VE DENEYLER

5.1 Değerlendirme Yöntemleri

Uygulanan yöntemlerin başarısını ölçmek için F-ölçüm (F-score, F_1 -score) yöntemi kullanıldı. Yöntemlerin başarısını ölçmek için, her bir kümenin F-ölçüm değeri ayrı ayrı hesaplandı. Daha sonra her kümenin F-ölçüm değeri kümedeki sayfa sayısı ile çarpıldı ve bu değerler toplandı. Toplanan değerler getirilen toplam sayfa sayısına bölünerek, çalışmamızda en önemli değerlendirme kriteri olarak kullandığımız ortalama F-ölçüm değeri elde edilmiş olur.

F-ölçüm hesaplamasında Tutturma-Precision, Bulma-Recall değerleri kullanılmıştır.

Tutturma; yöntemden getirilen verinin ne kadarının gerçekte getirilmesi gereken veri arasında olduğunu gösterir. Tutturma (5.1) eşitliğinde gösterilmektedir.

Bulma ise; getirilmesi gereken verinin ne kadarının yöntem tarafından getirildiğini gösterir. Bulma (5.2) eşitliğinde gösterildiği gibi hesaplanır.

F-ölçüm; Tutturma ve Bulma değerlerinden yararlanılarak (5.3) eşitliğinde gösterildiği gibi hesaplanır.

$$Tutturma = P = \frac{\{getirilen sayfa\} \cap \{ilgili sayfa\}}{getirilen sayfa} \quad (5.1)$$

$$Bulma = R = \frac{\{getirilen sayfa\} \cap \{ilgili sayfa\}}{ilgili sayfa} \quad (5.2)$$

$$F - \text{Ölçüm} = \frac{2 * Tutturma * Bulma}{Tutturma + Bulma} \quad (5.3)$$

getirilen sayfa: Yöntemin, incelenen etiket için getirdiği sayfalar

ilgili sayfa: İncelenen etiketin, elle sınıflandırma ile elde edilen sayfalar

T , tüm kümeler setini belirtmek üzere, her c kümesi için F_c F-ölçüm, P_c Tutturma, R_c Bulma değerlerini göstermek üzere; Ortalama F-ölçüm değeri (5.4)'de, Ortalama Tutturma değeri (5.5)'de, Ortalama Bulma değeri (5.6)'da gösterildiği şekilde hesaplanır.

$$Ort.F - ölçüm = \frac{\sum_{c \in T} (F_c * |c|)}{\sum_{c \in T} |c|} \quad (5.4)$$

$$Ort.F - Tutturma = \frac{\sum_{c \in T} (P_c * |c|)}{\sum_{c \in T} |c|} \quad (5.5)$$

$$Ort.F - Bulma = \frac{\sum_{c \in T} (R_c * |c|)}{\sum_{c \in T} |c|} \quad (5.6)$$

Örneğin, klasik SAK yöntemi ile *balık* sorgusundan elde edilen *burç* etiketini inceleyelim. SAK yöntemi, *burç* etiketine 5 sayfayı atmıştır. Elle yapılan etiketlendirmede ise burç ile etiketlenmiş 4 adet sayfa bulunmaktadır. Hem SAK ile, hem de elle yapılan etiketlendirmede ortak bulunan sayfa sayısı, yani getirilen sayfa ve ilgili sayfa kesişimi 3'tür. Bu durumda, burç etiketi için F-ölçüm değeri şu şekilde hesaplanır:

$$Tutturma=P = 3 / 5 = 0.6$$

$$Bulma = R = 3 / 4 = 0.75$$

$$F - Ölçüm = 2 * (0.6 * 0.75) / (0.6 + 0.75) = 0.67$$

5.2 Yapılan Deneyler

Geliştirdiğimiz yöntemler üzerinde ölçümler yaparken, 4.1'de anlatıldığı gibi, seçtiğimiz anahtar kelimeleri arama motoruna vererek veri setimizi oluşturduk. Oluşturduğumuz veri seti üzerinde, yöntemlerin başarısını ölçebilmek için öncelikle elle kümeleme yaptık. Yaptığımız elle kümeleme sonuçlarını, başarılı kümeleme işlemi elde edilmesi gereken sonuçlar olarak kabul ettik. Veri seti üzerinde KSAK, GSAK ve DBC yöntemlerini uyguladık. Bu yöntemlerle elde ettiğimiz sonuçları, el ile etiketlendirdiğimizde elde ettiğimiz sonuçlarla karşılaştırarak, yöntemlerin başarımlarını bölüm 5.1'de anlatılan

yöntemlerle ölçtük. Daha sonra, 4.4’de anlatılan melez yöntem ile KSAK algoritmasının performansını geliştirmeye çalıştık ve elde ettiğimiz bu sonuçların başarısını da aynı şekilde ölçtük. Böylece SAK algoritmasının varyasyonları arasındaki performans farklarını, ayrıca SAK algoritmaları ile DBC arasındaki performans farklarını ölçebilmiş olduk.

Deneylerimizde kullandığımız veri setleri kullanılarak, tüm yöntemlerce oluşturulan kümeler ve ölçüm oranları:

- Bölüm 5.2.1’de *balık* sorgusu için
- Bölüm 5.2.2’de *dünya* sorgusu için
- Bölüm 5.2.3’te *ocak* sorgusu için
- Bölüm 5.2.4’te *takvim* sorgusu için
- Bölüm 5.2.5’te *yüz* sorgusu için

olacak şekilde anlatılmaktadır.

5.2.1 Balık Sorgusu Sonuçları

Veri setimizde **Balık** sorgusu için 50 adet internet sayfası bulunmaktadır. Bu veriler kullanılarak, KSAK, GSAK, DBC, M-GSAK algoritmalarıyla kümeleme işlemi yapıldı. Çizelge 5.1’de her bir yöntemin ve elle etiketleme sonucu elde ettiğimiz kümelerin, *Balık* sorgusu için çıkardığı etiket bilgileri gösterilmektedir. Çizelge 5.2’de elle yapılan etiketleme sonucu oluşan kümeler ve bu kümeler içerisindeki sayfalar gösterilmektedir.

Çizelge 5.1 **Balık** sorgusu için yöntemlerin çıkardığı etiketler

Yöntem	Etiket Sayısı	Çıkarılan Etiket İsimleri
KSAK	9	Beceri Oyunu, Wikipedia Avcılık, Burç, Restaurant, Deniz, Oyun, Av Teknikleri, Diğer
GSAK	6	Oyun, Wikipedia, Avcılık, Burç, Restaurant, Diğer
DBC	5	Oyun, Avcılık, Burç, Diğer, Et
Elle Etiketleme	6	Oyun, Wikipedia, Avcılık, Burç, Restaurant, Diğer

Çizelge 5.2 **Balık** sorgusu için elle çıkarılan sonuçlar

Etiketler	Sayfalar
Oyun	15,16,20,36,44,45
Wikipedia	0,21,27,29,31,32,35,39,40,47
Avcılık	1,3,5,8,9,11,33
Burç	6,12,23,49
Restaurant	4,17,19,24,25,26,28,34,38,41,42,43
Diğer	2,7,10,13,14,18,22,30,33,37,46,48

5.2.1.1 Klasik SAK

Veri setimiz üzerinde KSAK algoritmasını çalıştırdığımızda Çizelge 5.3’de gösterilen sonuçlar elde edilmiştir. KSAK, 15 sayfaya birden fazla etiket atamaktadır. (0, 1, 19, 20, 21, 24, 27, 29, 39, 40, 44, 45) numaralı sayfalara iki etiket atarken, (5 ve 9) sayfalara ise üç kez etiket ataması yapmıştır.

Çizelge 5.3 **Balık** sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Beceri Oyunu	20,44,45
Wikipedia	0,21,27,29,31,32,35,39,40,47
Avcılık	1,3,5,9,11
Burç	6,12,23, 31 ,49
Restaurant	19,24 ,28,34,38
Deniz	4, 5 ,8, 9 , 19,24,39 ,46
Oyun	15,16, 20,36,44,45
Av Teknikleri	1,5,9
Diğer	2,7,10,13,14,17,18,22,25,26,30,33,37,41,42,43,48

Çizelge 5.4’de **Balık** sorgusu için elde edilen sonuçlar gösterilmektedir. Buradan her kümenin ayrı ayrı *Tutturma*, *Bulma* ve *F-Ölçüm* değerleri ile *Ort. Tutturma*, *Ort. Bulma* ve *Ort. F-Ölçüm* değerleri gösterilmektedir.

Çizelge 5.4 **Balık** sorgusu için KSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Beceri Oyunu	0	0	0
Wikipedia	1	1	1
Avcılık	1	0,71	0,83
Burç	0,8	1	0,89
Restaurant	1	0,42	0,59
Deniz	0	0	0
Oyun	1	1	1
Av Teknikleri	0	0	0
Diğer	0,65	0,92	0,76
Ortalamalar	0,66	0,68	0,65

5.2.1.2 Geliştirilmiş SAK

Çizelge 5.3'deki KSAK yöntemi sonuçlarından da görülebileceği üzere; bazı kümeler, etiketleri açısından ya da içerdikleri sayfalar açısından büyük oranda benzerlik taşımaktadırlar. 4.2.3'de belirttiğimiz aşamaları KSAK'den aldığımız sonuç kümeleri üzerinde uyguladık. Birbirine %70 oranından fazla benzerlik taşıyan kümeler belirlendi. Çizelge 5.3'den de rahatlıkla görülebileceği üzere: *Oyun - Beceri Oyunu, Avcılık – Av Teknikleri, Wikipedia– Balık Vikipedi* etiketli kümeler bu koşula uymaktadırlar. Koşula uyan bu kümeler, GSAK'de birleştirilmiştir. Geriye kalan kümelere baktığımızda, *Deniz* etiketli kümenin skorunun belirlenen eşik değer olan 5'in altında olduğunu gördük. Bu kümenin elemanlarını da inceleyecek olursak, bu kümeye ait 8 sayfanın 5 tanesinin diğer kümelerde de yer aldığı görülebilir. Bu 4.2.3'de belirlediğimiz %50 oranının üzerinde bir değerdir. Bu sebeple *Deniz* etiketli küme silinir. Kümenin diğer kümeler içerisinde yer almayan 3 elemanı (4, 8 ve 46 nolu sayfalar), belirli bir kümeye eklenememiş sayfaların bulunduğu *Diğer* kümesine dahil edilir. Yapılan bu geliştirmeler sonrasında oluşan kümeler ve bu kümelerin içerdikleri sayfalar Çizelge 5.5'de gösterilmektedir.

Çizelge 5.5 **Balık** sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Oyun	15,16,20,36,44,45
Wikipedia	0,21,27,29, 31 ,32,35,39,40,47
Avcılık	1,3,5,9,11
Burç	6,12,23, 31 ,49
Restaurant	19,24,28,34,38
Diğer	2,4,7,8,10,13,14,17,18,22,25,26,30,33,37,41,42,43,46,48

Çizelge 5.6 **Balık** sorgusu için GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Oyun	1	1	1
Wikipedia	1	1	1
Avcılık	1	0,71	0,83
Burç	0,8	1	0,89
Restaurant	1	0,42	0,59
Diğer	0,65	0,92	0,76
Ortalamalar	0,85	0,88	0,84

5.2.1.3 DBC

DBC yöntemini uygulayarak yaptığımız birbirlerine benzer dokümanları bulma ve bu benzer dokümanları aynı küme altında toplama işlemi neticesinde elde ettiğimiz sonuçlar Çizelge 5.7’de gösterilmektedir.

Çizelge 5.7 **Balık** sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Oyun	2,14,15,16,20,26,33,36,44,45,48
Avcılık	1,3,5,8,9,11,13,18,19,38
Burç	6,12,21,23,25,29,31,49
Et	7,10,28,30,32,39
Diğer	0,4,17,22,24,27,34,35,37,40,41,42,43,46,47

DBC yönteminde her bir sayfaya sadece tek bir etiket atanmaktadır.

Çizelge 5.8 **Balık** sorgusu için DBC yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Oyun	0,55	1	0,71
Avcılık	0,6	0,86	0,71
Burç	0,5	1	0,67
Et	0	0	0
Diğer	0,2	0,25	0,22
Ortalamalar	0,38	0,63	0,47

5.2.1.4 M-GSAK

İki yöntemin birleştirilmesi çalışmasında 4.4’de anlatıldığı şekilde bir çalışma yapılmıştır. GSAK ve DBC algoritmalarının sonuçları karşılaştırılarak, SAK yönteminin başarımı arttırılmaya çalışılmıştır.

Ek-A’da **Balık** için arama motorunda yapılan sorgu sonucu dönen ilk 50 internet sayfası numaralandırılmış olarak verilmektedir.

Çizelge 5.9’ da bu 50 sayfanın elle etiketlenme sonucunun diğer yöntemler ile yapılan kümeleme sonuçları ile ne kadar uyduğu verilmektedir. Çizelge 5.9’da (a, b) formunda belirttiğimiz kullanımda a , çakışan adres sayısını, b de, etiket içerisine dahil edilen sayfa sayısını göstermektedir. Birinci satırda *Oyun* kelimesi için alınan değerleri yorumlayacak olursak; (6,6) değeri bize, GSAK yöntemi kullanılarak yapılan etiketlemede 6 sayfanın bu küme altına dahil edildiğini ve 6’sının da elle yapılan ve *Oyun* etiketli küme altında yer alan 6 sayfa değeri ile aynı olduğunu söylemektedir. (6, 11) ise, DBC ile yapılan etiketlemede 11 sayfanın *Oyun* etiketli küme altında olduğunu ve bunlardan 6 tanesinin elle yapılan etiketleme ile aynı olduğunu söylemektedir.

Çizelge 5.9 Yöntemler ve elle yapılan etiketlemede çakışan adreslerin sayısı

Etiketler	Yöntemler		
	GSAK	DBC	Elle Sınıflama
Oyun	(6,6)	(6,11)	6
Wikipedia	(10,10)	-	10
Avcılık	(5,5)	(6,10)	7
Burç	(4,5)	(4,8)	4
Restaurant	(5,5)	-	12
Et	-	(0,6)	-
Diğer	(11,17)	(3,15)	12

4.4'te anlattığımız iki yöntemin birleştirilmesi işleminden yararlanmak istersek, GSAK yönteminde bulunup, DBC yönteminde bulunmayan kümeleri incelememiz gerekir. Çizelge 5.9'dan da rahatça görülebileceği üzere, incelenmesi gereken kümeler, *Restaurant* ve *Wikipedia* etiketli kümelerdir.

Bölüm 4.4'te anlatılan adımlar doğrultusunda, *Restaurant* etiketinin GSAK ve DBC içerisindeki sayfalar incelendiğinde, Çizelge 5.10'daki sonuç elde edilir. Sayfalarının tümünün diğer GSAK etiketi boş küme olduğundan, *Restaurant* etiketli küme geçerli bir küme olarak kabul edilir.

Çizelge 5.10 *Restaurant* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
19	(ϕ , Avcılık)	Restaurant
24	(ϕ , Diğer)	Restaurant
28	(ϕ , Et)	Restaurant
34	(ϕ , Diğer)	Restaurant
38	(ϕ , Avcılık)	Restaurant

Aynı işlemler *Wikipedia* etiketi üzerinde de tekrarlayacak olursak, yine SAK'den başka bir küme sonucu dönmediği görülür. Bu sebeple *Wikipedia* etiketli küme de aynen kalır.

Bu kurallar işletildiğinde Çizelge 5.5'deki sonuçlar üzerinde hiçbir değişiklik olmadığı görülür. Sonuçlar üzerinde bir değişiklik olmadığından M-GSAK'nin başarısı da GSAK ile aynı kalmıştır.

5.2.2 Dünya Sorgusu Sonuçları

Çizelge 5.11'de **Dünya** sorgusu için tüm yöntemlerle elde edilen küme etiketlerinin sayısı ve isimleri verilmektedir.

Çizelge 5.11 **Dünya** sorgusu için yöntemlerin çıkardığı etiketler

Yöntem	Etiket Sayısı	Çıkarılan Etiket İsimleri
KSAK	6	Büyük, Türkiye, Haber, Güncel, Vikipedi, Diğer
GSAK	4	Büyük, Haber, Vikipedi, Diğer
DBC	4	Haber, Dünya Mirası, Dizi, Diğer
Elle Etiketleme	6	Haber, Vikipedi, Dünya Mirası, Dizi, Şarkı, Diğer

Dünya sorgusu üzerinde yapılan elle etiketleme sonucu oluşan kümeler ve bu kümelere dahil edilen sayfalar Çizelge 5.12’de gösterilmektedir.

Çizelge 5.12 **Dünya** kelimesi için elle çıkarılan sonuçlar

Etiketler	Sayfalar
Haber	0,2,4,5,6,7,8,10,13,14,15,18,19,20,26,27,33,37
Vikipedi	1,16,23,30,38,45
Dünya Mirası	12,39
Dizi	24,46,48
Şarkı	42,44
Diğer	3,9,11,17,21,22,25,28,29,31,32,34,35,36,40,41,43,47,49

5.2.2.1 Klasik SAK

Dünya kelimesi KSAK yöntemi ile sorgulandığında 6 adet küme elde edilmiştir. 39 sayfa yalnızca bir küme içerisinde yer almaktadır. 9 sayfa (0, 2, 4, 5, 16, 18, 19, 23, 38) iki kümede birden bulunmaktadır. 2 sayfa ise 3 küme (7, 14) içerisinde yer almaktadır. Çizelge 5.13’da elde edilen kümeler ve bu kümeleri oluşturan sayfalar gösterilmektedir.

Çizelge 5.13 **Dünya** sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Büyük	6,7,14,22,23,25
Türkiye	4,7,12,14,16,19,20,24,29,37,47
Haber	0,2,4,5,7,10,13,14,15,18,19,33
Güncel	0,2,5,8,18,38
Vikipedi	1,16,23,30,38,45
Diğer	3,9,11,17,21,26,27,31,32,34,35,36,39,40,41,42,43,44,46,48,49

Dünya sorgusu için oluşturulan F-ölçüm sonuçları da Çizelge 5.14’de verilmektedir.

Çizelge 5.14 **Dünya** sorgusu için KSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Büyük	0	0	0
Türkiye	0	0	0
Haber	1	0,67	0,8
Güncel	0	0	0
Vikipedi	1	1	1
Diğer	0,67	0,74	0,7
Ortalamalar	0,52	0,48	0,49

5.2.2.2 Geliştirilmiş SAK

Dünya sorgusu için GSAK yöntemi ile yapılan etiketlendirmede 47 sayfa tek bir küme içerisinde yer almaktadır. Geri kalan 3 sayfa (7,14,23) iki kümede birden yer almaktadır. Çizelge 5.15’de GSAK yöntemiyle elde edilen kümeler ve bu kümelere yerleştirilen sayfalar görülmektedir.

Çizelge 5.15 **Dünya** sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Büyük	6,7,14,22,23,25
Haber	0,2,4,5,7,8,10,13,14,15,18,19,33
Vikipedi	1,16,23,30,38,45
Diğer	3,9,11,12,17,20,21,24,26,27,28,29,31,32,34,35,36,37,39,40,41,42,43,44,46,47,48,49

Dünya sorgusu için GSAK yöntemiyle elde edilen sonuçlar Çizelge 5.16’da gösterilmektedir.

Çizelge 5.16 **Dünya** sorgusu için GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Büyük	0	0	0
Haber	1	0,73	0,84
Vikipedi	1	1	1
Diğer	0,61	0,90	0,73
Ortalamalar	0,68	0,77	0,70

5.2.2.3 DBC

DBC yöntemiyle oluşturulan kümeler Çizelge 5.17’de verilmektedir. Bu kümeleme sonucu değerlendirme yöntemlerimizle yaptığımız ölçümlerin sonuçları ise Çizelge 5.18’de gösterilmektedir.

Çizelge 5.17 **Dünya** sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Haber	2,5,7,13,15,18,0,1,3,4,6,8,9,10,11,14,16,17,19,20,21,22,23,25,26,30,31,32,33,35,36,38,40,41,42,43,44,45,47,49
Dizi	24,28,46,48
Dünya Mirası	12,37,39
Diğer	27,29,34

Çizelge 5.18 **Dünya** sorgusu için DBC yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Haber	0,4	0,89	0,55
Dizi	0,75	1	0,86
Dünya Mirası	0,67	1	0,8
Diğer	0,67	0,11	0,18
Ortalamalar	0,46	0,86	0,57

5.2.2.4 M-GSAK

Çizelge 5.19’da GSAK ve DBC yöntemlerinin çakışan etiketleri, bu etiketlere ait kümelerdeki getirilen ve doğru bilinen sayfa sayıları verilmektedir.

Çizelge 5.19 Yöntemler ve elle yapılan etiketlemede çakışan adreslerin sayısı

Etiketler	Yöntemler		
	GSAK	DBC	Elle Sınıflama
Büyük	(0,5)	-	-
Haber	(13,13)	(16,40)	18
Vikipedi	(6,6)	-	6
Diğer	(17,28)	(2,3)	19
Dizi	-	(3,4)	3
Dünya Mirası	-	(2,3)	2
Şarkı	-	-	2

Çizelge 5.19’dan görülebileceği üzere, GSAK kümelerinden *Büyük* ve *Vikipedi* kümelerinin DBC ile yapılan kümelemede bir karşılığı yoktur. Bu sebeple bu iki küme elemanlarını birleştirme aşamasında tekrar gözden geçirmemiz gerekmektedir.

Büyük etiketli kümenin elemanlarını bu doğrultuda 4.4’te anlattığımız şekilde incelediğimizde Çizelge 5.20’deki durum oluşmaktadır.

Çizelge 5.20 *Büyük* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
6	(ϕ , Haber)	Büyük
7	(ϕ ,Haber)	Büyük
14	(ϕ ,Haber)	Büyük
22	(ϕ ,Haber)	Büyük
23	(ϕ ,Haber)	Büyük
25	(ϕ ,Haber)	Büyük

Çizelge 5.20'deki gibi değerlendirme yapıldığında, 4.4'te belirttiğimiz gibi, tüm sayfalar için GSAK tarafında boş küme bulunduğundan, *Büyük* etiketi sonuç etiketi olarak kabul edilir. Aynı işlemleri *Vikipedi* etiketli kümeye uyguladığımızda, burada da aynı durumla karşılaştığımızdan *Vikipedi* etiketi de GSAK yönteminde olduğu şekliyle kalır.

Bu durumda *Büyük* ve *Vikipedi* etiketli kümelerin sayfalarına yeni etiket ataması yapılmaz. Bu kümeler birleştirme işleminden sonra da aynı kalırlar.

Herhangi bir değişiklik yapılmadığından, birleştirme sonrasındaki başarımlar. GSAK yönteminin başarısıyla aynı olur. Bir gelişme olmaz.

5.2.3 Ocak Sorgusu Sonuçları

Bu alt bölümden itibaren **Ocak**, **Takvim** ve **Yüz** sorgu kelimeleri için detay bilgi verilmeden sadece sonuçlar çizelgeler yardımıyla aktarılacaktır.

Çizelge 5.21 **Ocak** sorgusu için yöntemlerin çıkardığı etiketler

Yöntem	Etiket Sayısı	Çıkarılan Etiket İsimleri
KSAK	8	Ürün, Pişirme, Türkiye, Pazartesi, Fiyat, Döviz, Vikipedi, Diğer
GSAK	5	Vikipedi, Döviz, Fiyat, Pişirme, Diğer
DBC	5	Ada, Pişirme, Vikipedi, Döviz, Diğer
Elle Etiketleme	5	Vikipedi, Pişirme, Döviz, Maden Ocağı, Diğer

Çizelge 5.22 **Ocak** kelimesi için elle çıkarılan sonuçlar

Etiketler	Sayfalar
Vikipedi	0,12,21,29,46
Pişirme	3,4,5,19,20,32,34,44
Döviz	6,8,10,18,22,25,27
Maden Ocağı	13,26
Diğer	1,2,7,9,11,17,14,15,16,23,24,28,30,31,33,35,36,37,38,39,40,41,42,43,45,47,48,49

5.2.3.1 Klasik SAK

KSAK ile yapılan kümeleme işleminde; 2 sayfa (19, 32) iki kümede, 4 sayfa (3, 6, 8, 18) ise 3 kümede, 44 sayfa da tek bir kümede yer almaktadır.

Çizelge 5.23 **Ocak** sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Ürün	3,5,19,20,26
Piştirme	3,19,32
Türkiye	1,6,8,10,11,18,22,23,25,27,31,34
Pazartesi	6,8,14,18,37
Fiyat	3,4,13,32,42,48
Döviz	6,8,18
Vikipedi	0,12,15,21,29,30,36,41,46,47
Diğer	2,7,9,16,17,24,28,33,35,38,39,40,43,44,45,49

Çizelge 5.24 **Ocak** sorgusu için KSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Ürün	0	0	0
Piştirme	1	0,38	0,55
Türkiye	0	0	0
Pazartesi	0	0	0
Fiyat	0	0	0
Döviz	1	0,43	0,6
Vikipedi	0,5	1	0,67
Diğer	0,94	0,54	0,68
Ortalamalar	0,43	0,35	0,35

5.2.3.2 Geliştirilmiş SAK

44 sayfa bir kümede, 2 sayfa (3) iki küme içerisinde yer almaktadır.

Çizelge 5.25 **Ocak** sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Piştirme	3,5,19,20,26
Fiyat	3,4,13,32,42,48
Döviz	6,8,10,18,22,25,27,34
Vikipedi	0,12,15,21,29,30,36,41,46,47
Diğer	1,2,7,9,11,14,16,17,23,24,28,31,33,35,38,39,40,43,44,45,49

Çizelge 5.26 **Ocak** sorgusu için GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Piştirme	0,8	0,5	0,62
Fiyat	0	0	0
Döviz	0,88	1	0,93
Vikipedi	0,5	1	0,67
Diğer	0,95	0,71	0,82
Ortalamalar	0,72	0,71	0,69

5.2.3.3 DBC

Çizelge 5.27 **Ocak** sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Ada	11,17,1,7,13,14,35
Pişirme	3,19,32,4,5,20,26,40,42,44,48
Döviz	6,8,10,18,22,27,23,25,31,34,37
Vikipedi	12,21,29,46,0,15,30,33,36,38,39,41,43,47
Diğer	2,9,16,24,28,45,49

Çizelge 5.28 **Ocak** sorgusu için DBC yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Ada	0	0	0
Pişirme	0,64	0,88	0,74
Döviz	0,64	1	0,78
Vikipedi	0,36	1	0,53
Diğer	1	0,25	0,4
Ortalamalar	0,52	0,73	0,54

5.2.3.4 M-GSAK

Çizelge 5.29’da GSAK ve DBC yöntemleriyle elde edilen sonuçlar karşılaştırılmalı olarak gösterilmektedir.

Çizelge 5.29 Yöntemler ve elle yapılan etiketlemede çakışan adreslerin sayısı

Etiketler	Yöntemler		
	GSAK	DBC	Elle Sınıflama
Pişirme	(4,5)	(7,11)	8
Fiyat	(0,6)	-	-
Döviz	(7,8)	(7,11)	7
Vikipedi	(5,10)	(5,14)	5
Diğer	(20,21)	(7,7)	28
Ada	-	(0,7)	-
Maden Ocağı	-	-	2

Çizelge 5.29’daki sonuçlar incelenerek, GSAK yöntemi sonuçlarında olup, DBC’de olmayan *Fiyat* etiketi incelenir.

Çizelge 5.30 *Fiyat* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
3	(Pişirme, Pişirme)	Pişirme
4	(ϕ ,Pişirme)	Pişirme
13	(ϕ ,Ada)	Ada
32	(ϕ ,Pişirme)	Pişirme
42	(ϕ ,Pişirme)	Pişirme
48	(ϕ ,Pişirme)	Pişirme

Çizelge 5.30’da görüldüğü gibi *Fiyat* etiketli küme bozularak, içerisindeki sayfalar Çizelge 5.30’da görüldüğü gibi yeni kümelere aktarılır. Bu durumda, M-GSAK işlemi sonrasında kümeler Çizelge 5.31’deki gibi olur.

Çizelge 5.31 **Ocak** sorgusu M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Pişirme	3,4,5,19,20,26,32,42,48
Ada	13
Döviz	6,8,10,18,22,25,27,34
Vikipedi	0,12,15,21,29,30,36,41,46,47
Diğer	1,2,7,9,11,14,16,17,23,24,28,31,33,35,38,39,40,43,44,45,49

Çizelge 5.31’deki kümeler doğrultusunda, F-ölçüm değerlerini yeniden hesaplayacak olursak Çizelge 5.32’deki sonuçları elde ederiz.

Çizelge 5.32 **Ocak** sorgusu için M-GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Pişirme	0,67	0,75	0,71
Ada	0	0	0
Döviz	0,88	1	0,93
Vikipedi	0,5	1	0,67
Diğer	0,95	0,71	0,82
Ortalamalar	0,78	0,81	0,77

5.2.4 Takvim Sorgusu Sonuçları

Çizelge 5.33 **Takvim** sorgusu için yöntemlerin çıkardığı etiketler

Yöntem	Etiket Sayısı	Çıkarılan Etiket İsimleri
KSAK	15	Üniversite, Güz, Öğretim Yılı, Takvim, Tıp Fakültesi, Akademik Takvim, Takvim Önlisans lisans, Güz Dönemi, Eğitim, Zaman, Öğretim Yılı Akademik, Eğitim Öğretim Yılı, Lisans, Eylül, Diğer
GSAK	5	Zaman, Güz, Akademik Takvim, Öğretim, Diğer
DBC	6	Akademik Takvim, Gazete, Zaman, Türktakvim, Lisans, Diğer
Elle Etiketleme	4	Akademik Takvim, Gazete, Vikipedi, Zaman

Çizelge 5.34 **Takvim** kelimesi için elle çıkarılan sonuçlar

Etiketler	Sayfalar
Akademik Takvim	10,11,14,15,16,17,18,19,24,26,27,29,31,32,35,36,37,38,39, 40,41,42,43,44,47,48,49
Gazete	0,3,6,9,20,21,22,23,33
Vikipedi	5,7,8,25,30,46
Zaman	1,2,4,12,23,28,34,45

5.2.4.1 Klasik SAK

Takvim sorgusunda KSAK ile yapılan kümeleme işleminde; 19 sayfa tek bir kümede, 12 sayfa (1, 2, 10, 11, 18, 19, 28, 29, 31, 34, 39, 44) 2 kümede, 3 sayfa (16, 17, 40) 3 kümede, 3 sayfa (37, 38, 47) 4 kümede 7 sayfa (14, 24, 27, 32, 35, 41, 49) 5 kümede, 2 sayfa (36, 42) 6 kümede, 3 sayfa (15, 26, 48) 7 kümede, 1 sayfa (43) da 8 kümede birden bulunmaktadır.

Çizelge 5.35 **Takvim** sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Üniversite	14,17,18,26,27,29,37,41,48
Güz	11,14,15,35,36,41,43,47
Öğretim Yılı	14,15,16,17,24,27,36,38,42,43,44,49
Takvim	1,2,4,7,8,12,14,15,16,17,23,24,26,27,28,32,34,35,36,38,39, 40,42,43,46,48,49
Tıp Fakültesi	24,32,38,42,48
Akademik Takvim	10,11,14,15,16,18,19,24,26,27,29,31,32,35,36,37,38,39,40,41, 42,43,44,47,48,49
Takvim Önlisans Lisans	32,43,48
Güz Dönemi	26,35,36,41,47
Eğitim	15,19,27,37,42,43
Zaman	1,28,34
ÖğretimYılı Akademik	15,24,36,42
Eğitim Öğretim Yılı	26,35,40,48,49
Lisans	26,31,32,37,43,47,48
Eylül	2,10,15,26,41,43,49
Diğer	0,3,5,6,9,13,20,21,22,25,30,33,45

Çizelge 5.36 **Takvim** sorgusu için KSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Üniveriste	0	0	0
Güz	0	0	0
Öğretim Yılı	0	0	0
Takvim	0	0	0
Tıp Fakültesi	0	0	0
Akademik Takvim	1	0,97	0,98
Takvim Önlisans Lisans	0	0	0
Güz Dönemi	0	0	0
Eğitim	0	0	0
Zaman	1	0,38	0,55
Öğretim Yılı Akademik	0	0	0
Eğitim Öğretim Yılı	0	0	0
Lisans	0	0	0
Eylül	0	0	0
Diğer	0	0	0
Ortalamalar	0,21	0,19	0,19

5.2.4.2 Geliştirilmiş SAK

GSAK ile yapılan kümeleme işleminde, 31 sayfa (0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 18, 19, 20, 21, 22, 23, 25, 26, 29, 30, 31, 32, 33, 37, 39, 40, 45, 46, 48) tek bir kümede, 15 sayfa (1, 11, 16, 17, 24, 27, 28, 34, 35, 38, 41, 42, 44, 47, 49) iki kümede birden, 4 sayfa (14, 15, 36, 43) da üç kümede birden yer almıştır.

Çizelge 5.37 **Takvim** sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Zaman	1,28,34
Güz	11,14,15,35,36,41,43,47
Akademik Takvim	1,2,4,7,8,10,11,12,14,15,16,17,18,19,23,24,26,27,28,29,31,32,34,35,36,37,38,39,40,41,42,43,44,46,47,48,49
Öğretim	14,15,16,17,24,27,36,38,42,43,44,49
Diğer	0,3,5,6,9,13,20,21,22,25,30,33,45

Çizelge 5.38 **Takvim** sorgusu için GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Zaman	1	0,38	0,55
Güz	0	0	0
Akademik Takvim	0,73	1	0,84
Öğretim	0	0	0
Diğer	0	0	0
Ortalamalar	0,41	0,52	0,45

5.2.4.3 DBC

Çizelge 5.39 **Takvim** sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Zaman	1,23,28,34,4,5,12,25,30,33
Gazete	0,3,6,9,21,22
Akademik Takvim	14,17,24,27,42,11,15,16,18,19,29,32,35,36,38,39,40,41,43,44,48
Türktakvim	13,46,2,7,8
Lisans	26,31,37,47
Diğer	10,20,45,49

Çizelge 5.40 **Takvim** sorgusu için DBC yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Zaman	0,6	0,75	0,67
Gazete	1	0,67	0,8
Akademik Takvim	1	0,78	0,88
Türktakvim	0	0	0
Lisans	0	0	0
Diğer	0	0	0
Ortalamalar	0,66	0,56	0,60

5.2.4.4 M-GSAK

Çizelge 5.41’de GSAK ve DBC yöntemleriyle elde edilen sonuçlar karşılaştırılmalı olarak gösterilmektedir.

Çizelge 5.41 Yöntemler ve elle yapılan etiketlemede çıkan adreslerin sayısı

Etiketler	Yöntemler		
	GSAK	DBC	Elle Sınıflama
Akademik Takvim	(27,37)	(21,21)	27
Öğretim	(0,12)	-	-
Güz	(0,8)	-	-
Zaman	(3,3)	(6,10)	8
Diğer	(0,13)	(0,4)	-
Gazete	-	(6,6)	9
TürkTakvim	-	(0,5)	-
Lisans	-	(0,4)	-
Vikipedi	-	-	6

Çizelge 5.41’i incelediğimizde, *Öğretim* ve *Güz* etiketli kümeleri tekrar incelememiz gerektiği gözükmemektedir. *Öğretim* etiketli kümenin sayfalarının yeni etiketleri Çizelge 5.42’de gösterilmektedir.

Çizelge 5.42 *Öğretim* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
14	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
15	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
16	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
17	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
24	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
27	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
36	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
38	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
42	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
43	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
44	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
49	(Akademik Takvim, Diğer)	Akademik Takvim

Güz etiketli kümenin sayfalarını da aynı şekilde incelediğimizde, Çizelge 5.43'deki sonuçları elde ederiz.

Çizelge 5.43 *Güz* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
11	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
14	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
15	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
35	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
36	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
41	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
43	(Akademik Takvim, Akademik Takvim)	Akademik Takvim
47	(Akademik Takvim, Akademik Takvim)	Akademik Takvim

Öğretim ve *Güz* etiketli kümeler üzerinde değişiklikler yapıldıktan sonra, oluşan yeni kümeler Çizelge 5.44'te gösterilmektedir.

Çizelge 5.44 **Takvim** sorgusu M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Zaman	1,28,34
Akademik Takvim	1,2,4,7,8,10,11,12,14,15,16,17,18,19,23,24,26,27, 28 ,29,31,32, 34 ,35,36,37,38,39,40,41,42,43,44,46,47,48,49
Diğer	0,3,5,6,9,13,20,21,22,25,30,33,45

Yeni kümeler doğrultusunda değerlendirme yöntemleri uygulandığında elde edilen sonuçlar Çizelge 5.45'te gösterilmektedir.

Çizelge 5.45 **Takvim** sorgusu için M-GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Zaman	1	0,38	0,55
Akademik Takvim	0,73	1	0,84
Diğer	0	0	0
Ortalamalar	0,57	0,72	0,62

5.2.5 Yüz Sorgusu Sonuçları

Çizelge 5.46 **Yüz** kelimesi için yöntemlerin çıkardığı etiketler

Yöntem	Etiket Sayısı	Çıkarılan Etiket İsimleri
KSAK	9	Bakım, Türkiye, Vikipedi, Yüz Nakli, Oyun, Bakım Oyunu, Uzman TV, Bakım Ürünleri, Diğer
GSAK	6	Bakım Ürünleri, Oyun, Yüz Nakli, Vikipedi, Uzman TV, Diğer
DBC	10	İş Bildirim, Estetik, Kız, Yüz Nakli, Oyun, Facebook, Eğitim, Uzman TV, Bakım Ürünleri, Diğer
Elle Etiketleme	8	Oyun, Vikipedi, Bakım Ürünleri, Yüz Nakli, Uzman TV, Estetik, Facebook, Diğer

Çizelge 5.47 **Yüz** sorgusu için elle çıkarılan sonuçlar

Etiketler	Sayfalar
Oyun	8,10,11,14,16,24,25,33,40,42,43,46
Wikipedi	1,13,19,23,28,30,39
Bakım Ürünleri	37,41,47
Yüz Nakli	13,34,38,45,49
Uzman TV	3,4,29,35,44
Estetik	6,7,29,31,35
Facebook	21,27
Diğer	0,4,5,9,12,15,17,18,19,20,21,22,26,27,32,36,48

5.2.5.1 Klasik SAK

Yüz etiketi kullanılarak yapılan KSAK kümeleme çalışmasında kullanılan 50 sayfadan 37'si tek bir kümede yer almaktadır. 10 sayfa (4, 8, 13, 24, 33, 34, 37, 38, 42, 49) 2 kümede birden, 3 sayfa (16, 40, 43) da 3 kümede birden yer almaktadır.

Çizelge 5.48 **Yüz** sorgusu için KSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Bakım	16,24,37,40,43
Türkiye	4,8,33,34,38,42,49
Wikipedi	1,13,19,23,28,30,39
Yüz Nakli	13,34,38,45,49
Oyun	8,10,11,14,16,24,25,33,40,42,43,46
Bakım Oyunu	16,40,43
Uzman TV	3,4,29,35,44
Bakım Ürünleri	37,41,47
Diğer	0,2,5,6,7,9,12,15,17,18,20,21,22,26,27,31,32,36,48

Çizelge 5.49 **Yüz** sorgusu için KSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Bakım	0	0	0
Türkiye	0	0	0
Wikipedi	1	1	1
Yüz Nakli	1	1	1
Oyun	1	1	1
Bakım Oyunu	0	0	0
Uzman TV	1	1	1
Bakım Ürünleri	1	1	1
Diğer	0,79	0,89	0,83
Ortalamalar	0,71	0,74	0,72

5.2.5.2 Geliştirilmiş SAK

Yüz kelimesi kullanılarak yapılan GSAK kümeleme çalışmasında 50 sayfadan 48'i tek bir kümede yer almaktadır. 2 sayfa (13, 37) ise 2 kümede yer almaktadır.

Çizelge 5.50 **Yüz** sorgusu için GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Bakım Ürünleri	37,41,47
Oyun	8,10,11,14,16,24,25,33,37,40,42,43,46
Yüz Nakli	13,34,38,45,49
Wikipedi	1,13,19,23,28,30,39
Uzman TV	3,4,29,35,44
Diğer	0,2,5,6,7,9,12,15,17,18,20,21,22,26,27,31,32,36,48

Çizelge 5.51 **Yüz** sorgusu için GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Bakım Ürünleri	1	1	1
Oyun	0,92	1	0,96
Yüz Nakli	1	1	1
Wikipedi	1	1	1
Uzman TV	1	1	1
Diğer	0,79	0,88	0,83
Ortalamalar	0,90	0,96	0,93

5.2.5.3 DBC

Çizelge 5.52 **Yüz** sorgusu için DBC yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
İş Bildirim	36
Estetik	8,10,11,14,16,24,25,33,37,40,42,43,46
Kız	7,18,31
Yüz Nakli	2,8,9,10,14,16,17,24,32,33,40,42,43,46
Oyun	3,4,29,35,44
Facebook	21,26,27
Eğitim	5,12,15,20
Uzman TV	3,4,29,35,44
Bakım Ürünleri	37,41
Diğer	1,6,28,30,39,47,48

Çizelge 5.53 **Yüz** sorgusu için DBC yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
İş Bildirim	0	0	0
Estetik	0,67	0,4	0,5
Kız	0	0	0
Yüz Nakli	0,56	1	0,71
Oyun	0,71	0,83	0,77
Facebook	0,67	1	0,8
Eğitim	0	0	0
Uzman TV	1	1	1
Bakım Ürünleri	1	0,67	0,9
Diğer	0,14	0,59	0,83
Ortalamalar	0,54	0,63	0,57

5.2.5.4 M-GSAK

Çizelge 5.54’te GSAK ve DBC yöntemleriyle elde edilen sonuçlar karşılaştırılmalı olarak gösterilmektedir.

Çizelge 5.54 Yöntemler ve elle yapılan etiketlemede çıkan adreslerin sayısı

Etiketler	Yöntemler		
	GSAK	DBC	Elle Sınıflama
Uzman TV	(5,5)	(5,5)	5
Wikipedi	(7,7)	-	7
Yüz Nakli	(5,5)	(5,9)	5
Oyun	(12,13)	(10,14)	12
Bakım Ürünleri	(3,3)	(2,2)	3
Diğer	(15,19)	(1,7)	17
İş Bildirimi	-	(0,1)	-
Estetik	-	(2,3)	5
Kız	-	(0,2)	-
Facebook	-	(2,3)	2
Eğitim	-	(0,4)	-

Çizelge 5.55’te görüldüğü gibi *Wikipedi* etiketli kümenin incelenmesi gerekmektedir.

Çizelge 5.55 *Wikipedi* etiketi için adres birleştirmeleri

Adres No	(GSAK etiket, DBC etiket)	Yeni Etiket
1	(ϕ , Diğer)	Diğer
13	(Yüz Nakli, Yüz Nakli)	Yüz Nakli
19	(ϕ , Yüz Nakli)	Yüz Nakli
23	(ϕ , Yüz Nakli)	Yüz Nakli
28	(ϕ , Diğer)	Diğer
30	(ϕ , Diğer)	Diğer
39	(ϕ , Diğer)	Diğer

Çizelge 5.55’teki değişiklikler uygulandığında oluşan yeni kümeler Çizelge 5.56’daki gibi olmaktadır.

Çizelge 5.56 Yüz sorgusu için M-GSAK yöntemiyle oluşturulan etiketler ve sayfaları

Etiketler	Sayfalar
Bakım Ürünleri	37,41,47
Oyun	8,10,11,14,16,19,23,24,25,33,37,40,42,43,46
Yüz Nakli	13,34,38,45,49
Uzman TV	3,4,29,35,44
Diğer	0,1,2,5,6,7,9,12,15,17,18,20,21,22,26,27,28,30,31,32,36,39,48

Çizelge 5.56'daki yeni kümeler doğrultusunda ölçümler tekrar hesaplandığında elde edeceğimiz sonuç Çizelge 5.57'deki gibi olmaktadır.

Çizelge 5.57 **Yüz** sorgusu için M-GSAK yöntemi F-ölçüm sonuçları

Etiketler	Tutturma	Bulma	F-Ölçüm
Bakım Ürünleri	1	1	1
Oyun	0,92	1	0,96
Yüz Nakli	0,71	1	0,83
Uzman TV	1	1	1
Diğer	0,65	0,88	0,75
Ortalamalar	0,78	0,95	0,85

5.3 Deney Sonuçlarının Değerlendirilmesi

Bölüm 5.2'de yaptığımız deneylerle, oluşturduğumuz tüm veri seti için incelediğimiz DBC, KSAK, GSAK ve M-GSAK yöntemlerini kullanarak sonuçlar elde ettik. Bölüm 5.1'de anlattığımız F-ölçüm, Tutturma ve Bulma ölçümlerini kullanarak da etiket bazlı ortalama değerler elde edildi. Elde ettiğimiz bu sonuçların daha çok anlam kazanabilmesi için, genel ve özel olarak incelenmesi gerekmektedir.

Şekil 5.1'de tüm yöntemler için ortalama F-ölçüm değerleri verilmektedir. Bu şekilden ve Bölüm 5.2'de anlattığımız deneylerden de görülebileceği üzere, GSAK algoritması genelde DBC'den daha iyi sonuç vermektedir.

KSAK ile DBC'yi karşılaştırsak DBC'nin daha iyi sonuç verdiği gözükmemektedir. KSAK algoritması, bir dokümanı birden fazla kümeye dahil edebilmektedir. DBC'de ise bir doküman yalnızca bir küme içerisinde yer alabilmektedir. Deneyde kullandığımız verileri incelediğimizde ve elle yaptığımız kümelendirmede de görmekteyiz ki; genelde bir doküman, bir küme içerisinde yer almaktadır. KSAK ile yaptığımız deneylerde, oluşan küme sayısı çoktur ve birçok doküman birden fazla küme içerisinde yer almaktadır. Bu sebeple, KSAK'nin *Tutturma* değeri yüksek olsa bile, *Bulma* değeri düşük olduğundan, algoritmanın genel başarısı da düşmektedir.

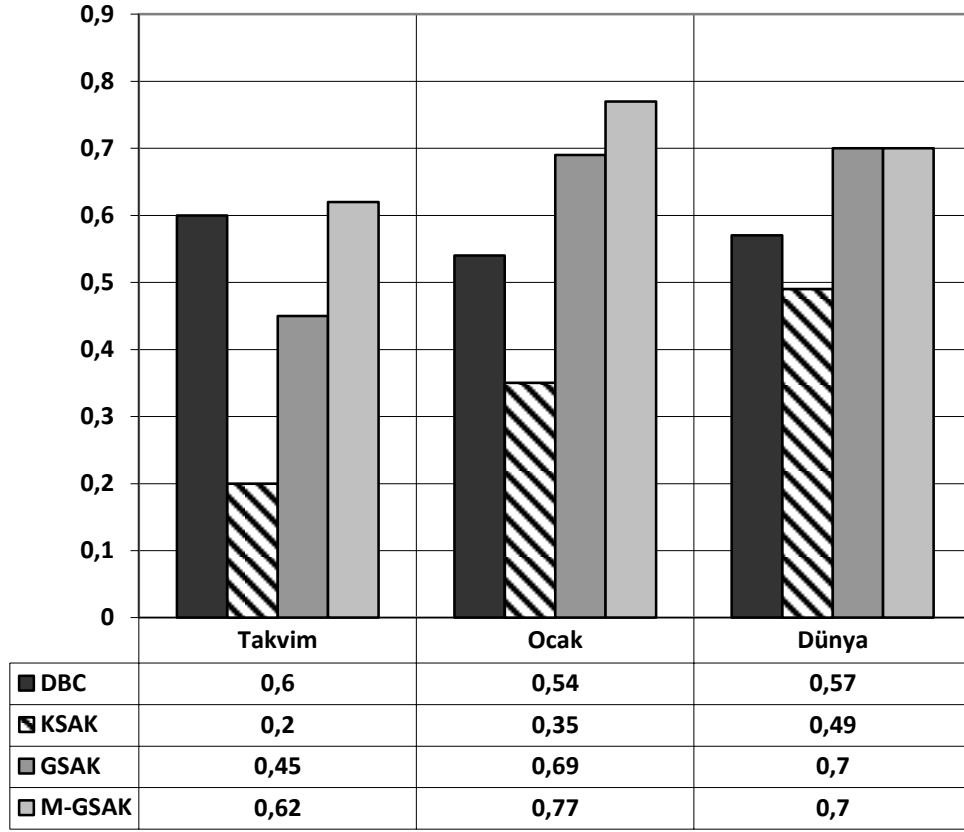
DBC algoritmasında her doküman tek bir kümeye dahil edilmektedir. Hatalı kümeleme işlemi sayısı, KSAK'ye göre azdır. Fakat, bir dokümanı tek kümeye yerleştirme işlemi *Tutturma* oranını da düşürmektedir.

Şekil 5.2’de her algoritmanın ortalama başarı oranları gözükmemektedir. Burada DBC ve KSAK’nin sonuçlarını incelediğimizde, KSAK’nin başarısının 0.35, DBC’nin başarısının da 0.57 olduğunu görebiliriz. Bu değerleri karşılaştırsak görmekteyiz ki DBC, KSAK’dan daha başarılı bir sonuç elde etmektedir.

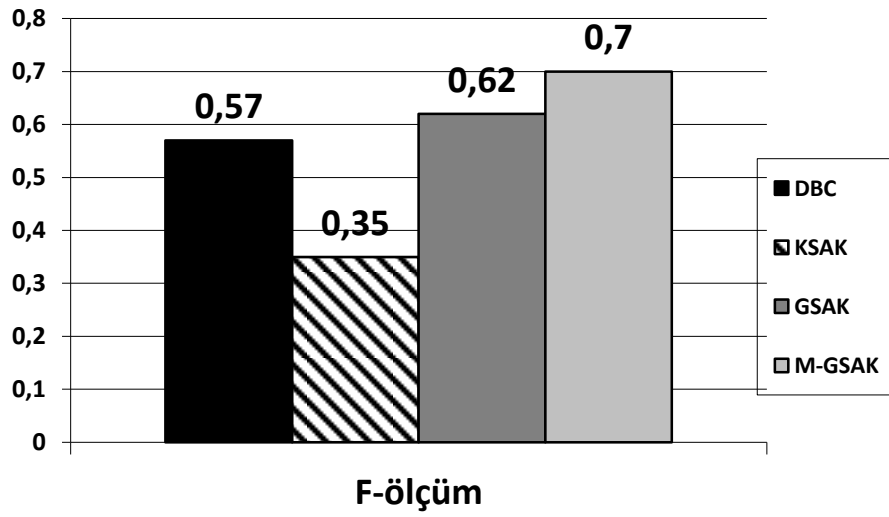
SAK yöntemi geliştirilip, GSAK haline getirildiğinde bu algoritmanın başarı artmaktadır. Yapılan çalışmalarla KSAK’ye göre daha az sayıda küme döndürülmüş ve dokümanların birbirine çok benzeyen birçok küme içerisinde birden fazla yer alması da KSAK’ye göre azaltılmıştır. Bu da GSAK’nin daha başarılı bir yöntem olmasını sağlamıştır. Şekil 5.2’deki ortalama sonuçlar incelendiğinde, GSAK’nin KSAK’dan %77 gibi önemli bir oranda daha başarılı olduğu gözlemlenmektedir. GSAK, DBC’den de %9 oranında daha başarılı olmuştur.

Uyguladığımız melez yöntem ile GSAK sonuçlarını, DBC’nin sonuçları ile karşılaştırıp, sağlamasını yaparak başarısını arttırmaya çalıştık. Şekil 5.1 ve Şekil 5.2’den de görülebileceği üzere, melez yöntem sadece *ocak* ve takvim sorgularında başarının artmasını sağlarken, *yüz* sorgusunda GSAK’nin başarısının altında kalmış, diğer sorgularda ise bir değişiklik meydana getirmemiştir. Tüm veri setimizin ortalama değerlerine bakacak olursak, Şekil 5.2’den görüleceği üzere, M-GSAK ile GSAK’nin başarıları yaklaşık %13 oranında artmıştır.

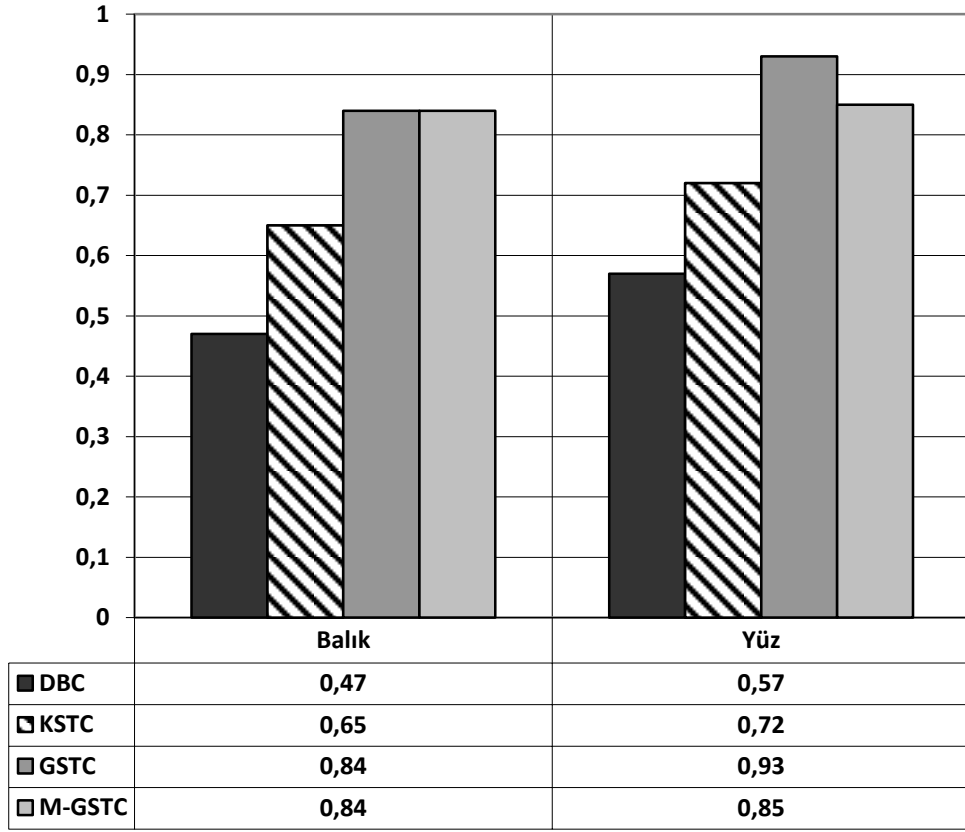
Şekil 5.3 ve Şekil 5.4’te Geliştirme Veri Seti’nde elde edilen F-Ölçüm değerleri görülebilir.



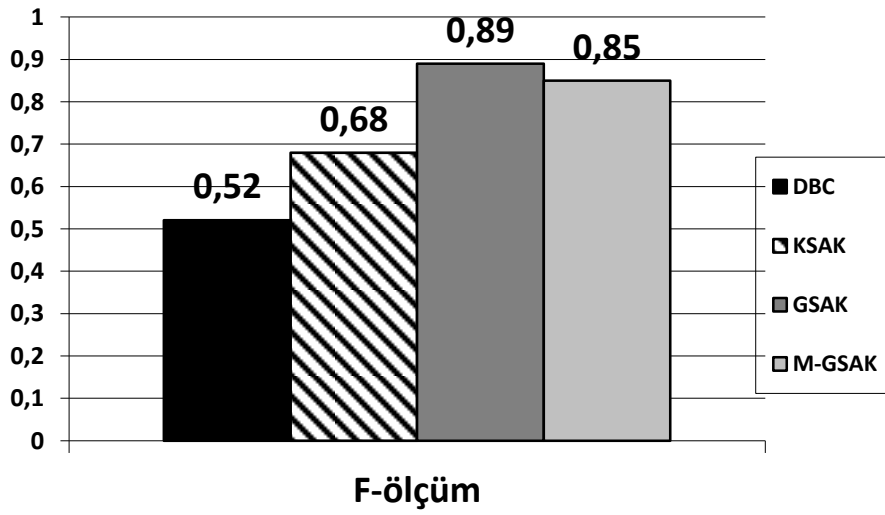
Şekil 5.1 Ort. F-ölçüm değerleri



Şekil 5.2 Ortalama Ort. F-ölçüm değerleri



Şekil 5.3 Çalışma Veri Seti Ort. F-Ölçüm Değerleri

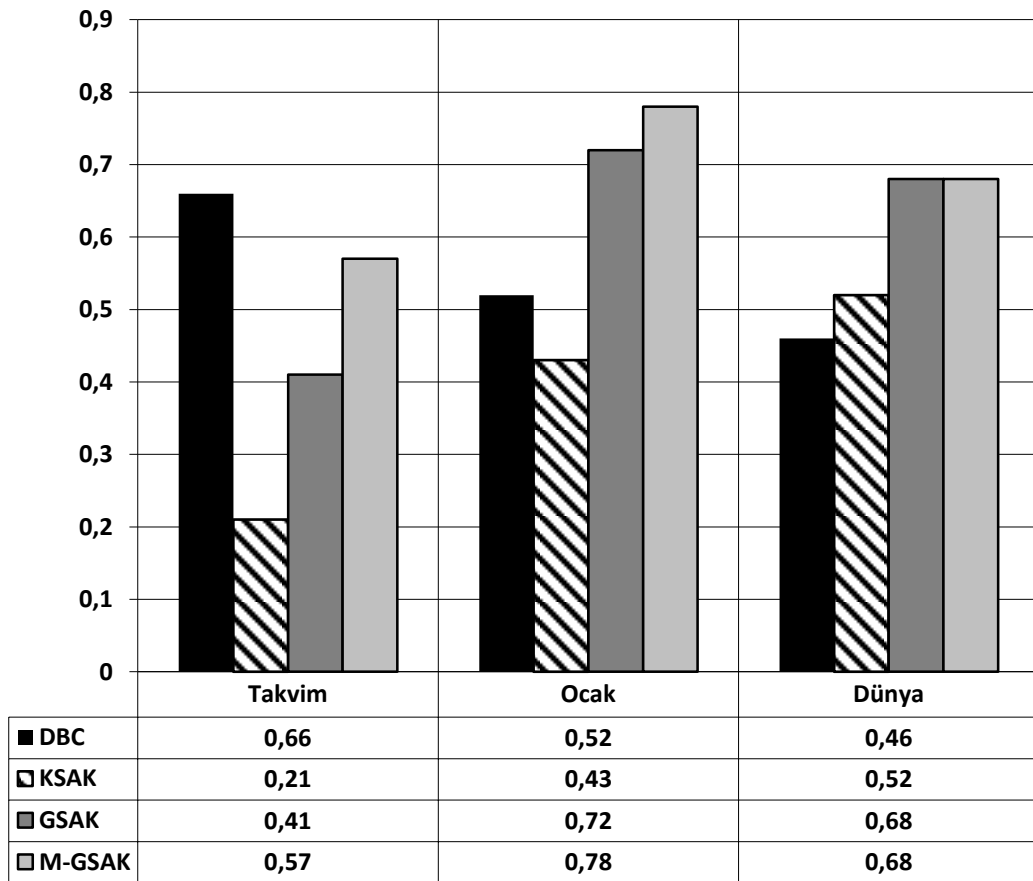


Şekil 5.4 Çalışma Veri Seti Ortalama Ort. F-Ölçüm Değerleri

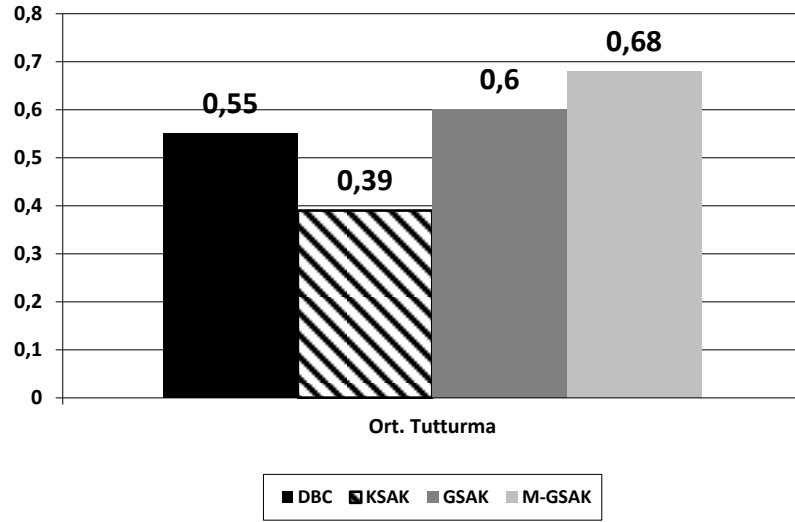
Sonuçları, sadece *Tutturma* değerlerini ele alacak şekilde inceleyecek olursak, tüm veri setleri için Şekil 5.5'teki sonuçları elde ederiz. Bu sonuçların ortalamaları da Şekil 5.6'da gösterilmektedir.

Bu sonuçları incelediğimizde; KSAK ve DBC arasındaki farkın burada azaldığı görülebilir. Şekil 5.6'dan ortalama Tutturma sonuçları oranı DBC için 0,55, KSAK için 0.39'dur. Buradan anlaşılacağı üzere, KSAK'de yapılan bir doküman için birden fazla kümeleme yapılması nedeniyle doğru kümelemenin yanında hatalı kümeleme sayısı da fazla olduğundan, KSAK'nin F-ölçüm başarıları daha düşüktür.

GSAK'nin *Tutturma* için başarı sonuçlarına baktığımızda, GSAK'nin *Tutturma* değerlerinde de, DBC ve KSAK'den daha başarılı olduğu görülebilir. Şekil 5.6'da çıkartılacağı üzere, *Tutturma* oranlarında GSAK; KSAK'den ve DBC'den daha başarılı bir sonuç elde etmektedir.



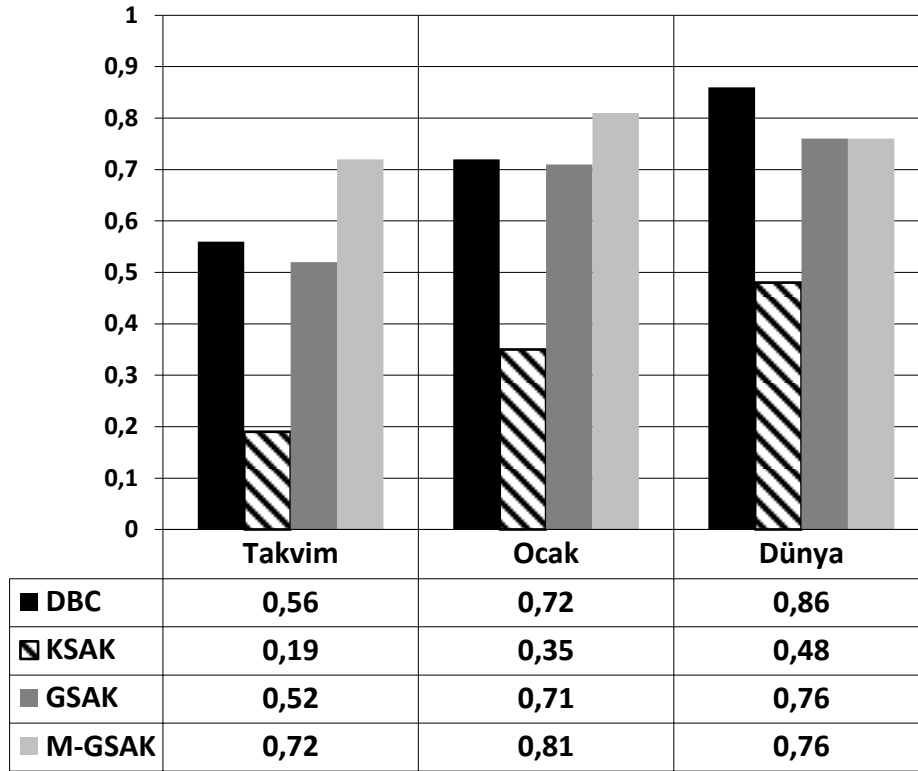
Şekil 5.5 Ort. Tutturma değerleri



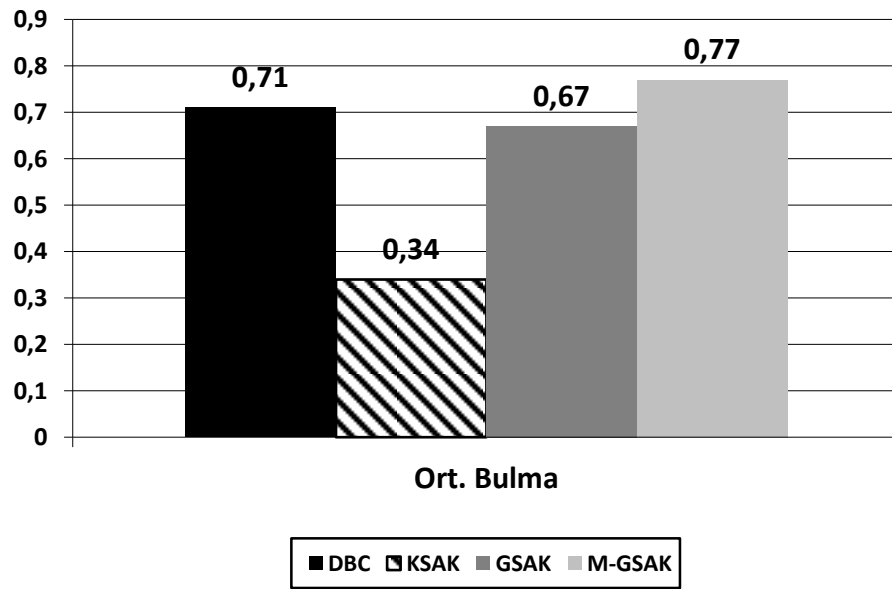
Şekil 5.6 Ortalama Ort. Tutturma Değerleri

Bulma değeri, bir kümede ilgili sayfaların ne kadarının getirildiğini göstermektedir. Şekil 5.7 ve Şekil 5.8’den *Bulma* değerleri ve Ortalama *Bulma* değerlerine ulaşabiliriz.

Bu sonuçlardan görülebileceği üzere DBC’nin *bulma* değeri, F-ölçüm ve Tutturma değerlerinden daha yüksektir. DBC’nin *bulma* değeri GSAK’dan bile fazladır. Şekil 5.6’daki sonuçları değerlendirirsek; DBC, GSAK ve KSAK’dan daha başarılı olmuştur. M-GSAK ise tüm yöntemler arasında en başarılı olanıdır.



Şekil 5.7 Ort. Bulma Değerleri



Şekil 5.8 Ortalama Ort. Bulma Değer

SONUÇLAR VE ÖNERİLER

Arama sonucu kümeleme işlemleri, arama motorlarının kullanılmaya başlanması ve yaygınlaşmasıyla ortaya çıkmıştır. İnternette içeriğin hızla büyümesi ve aranılan sorguya göre çoğu zaman döndürülen milyonlarca sonuç içerisinde, istenilen sonuca ulaşmanın zorlaşmasıyla önemini gittikçe arttırmaktadır.

Çalışmalarımızda hedefimiz, Türkçe içerikli internet sayfaları için arama sonucu kümeleme işleminin başarılı bir şekilde yapılmasını sağlamaktır. Bunun için öncelikle, bu konuda yapılan çalışmalar ve kullanılan algoritmalar hakkında literatür araştırması yaptık. Yaptığımız bu araştırma sonucu, bu alanda çok popüler olan ve performansını da başarılı gördüğümüz bir algoritma olan SAK'yi kendi çalışmalarımızda kullanmaya karar verdik. Ayrıca, doküman benzerliği ölçümlerinde sıkça kullanılan Kosinüs benzerliğini kullanarak hazırladığımız DBC yöntemiyle de, doküman benzerliği tabanlı yöntemlerin arama sonucu kümeleme işleminde nasıl sonuçlar vereceğini gözlemlemek istedik.

Bölüm 4.1'de detaylarından bahsettiğimiz, seçtiğimiz tek kelimeli sorgular sonucunda elde ettiğimiz Türkçe sonuçlardan oluşan veri setini yaptığımız çalışmada kullandık. Bu arama sonuçlarının içerikleri Ekler kısmında bulunmaktadır. Bir değerlendirme ve karşılaştırma yapabilmek için, aşağıdaki algoritmalar deneylerimizde kullanılmıştır:

- SAK'nin Zamir vd. [6]'nin geliştirdiği, bölüm 4.2.1'de detaylarını belirttiğimiz hali Klasik SAK (KSAK) olarak

- SAK'nin üzerinde bizim geliştirmeler yaptığımız hali (bölüm 4.2.2) Geliştirilmiş SAK (GSAK) olarak
- Bölüm 4.3'te bahsettiğimiz doküman benzerliği tabanlı Distance Based Clustering (DBC) algoritması
- Bölüm 4.4'te bahsedilen, DBC sonuçlarından yararlanılarak GSAK'nin başarısını arttırmayı hedefleyen Melez yöntem (M-GSAK)

Deneylerde kullandığımız bu algoritmaları bölüm 5.1'de anlatılan değerlendirme yöntemlerini kullanarak değerlendirdik. Veri seti olarak Türkçe veriler kullanarak, algoritmaların özellikle Türkçe arama sonuçları üzerindeki etkilerini gözlemlemek istedik.

Yaptığımız deneylerde elde ettiğimiz ortalama F-ölçüm sonuçlarına göre, DBC 0.57, KSAK 0.35, GSAK 0.62 ve M-GSAK 0.70 olarak başarılı sonuç vermiştir. Ulaştığımız sonuçlarda, SAK'nin üzerinde hiçbir düzenleme ve ayar yapılmamış halinin DBC'den daha başarısız olduğu ya da zaman zaman birbirlerine yakın sonuçlar aldığı gözlemlenmektedir. Üzerinde geliştirmeler yapılmış SAK algoritması olan GSAK, KSAK'ye göre Ort. F-ölçüm değerleri karşılaştırıldığında önemli oranda daha başarılı olmuştur. GSAK'nin başarısının DBC'den de yukarıda olduğu yine elde ettiğimiz sonuçlardan gözükmemektedir. Buna dayanarak, özellikle arama sonucu kümeleme işlemi için tasarlanmış bir algoritmanın, doküman kümeleme işlemleri için tasarlanmış bir yöntemle göre daha başarılı olduğu sonucunu çıkartabiliriz. Bunun en büyük sebebi, arama sonucu kümeleme işlemlerinde kullanılan verilerin, normal dokümanlara göre kısıtlı olmasındandır.

Özellikle arama sonucu kümeleme işlemleri için geliştirilmiş bir algoritma olan SAK ile daha olumlu sonuçlar alınmaktadır. Algoritma üzerinde yapılacak bir takım değişikliklerle, üzerinde deneyler yaptığımız Türkçe veri seti üzerinde başarı kolayca arttırılabilmektedir. DBC sonuçlarıyla GSAK'nin geliştirilmesine dayanan M-GSAK yöntemiyle de görüldüğü üzere başarı az da olsa artmaktadır. Buna karşın, M-GSAK'de eklenen yöntemlerin sonuçlarını karşılaştırma aşaması, işlemin daha uzun sürede çalışmasına sebep olacaktır. Bu da sonuç döndürme ve çalışma zamanını arttıracaktır.

Arama sonucu kümeleme işlemlerinde, sonuçlar kullanıcıya anında yansıtılmak istendiğinden, M-GSAK'deki bu ek süre olumsuz bir özellik olarak gösterilebilir.

Hazırladığımız sistemde, her ne kadar kullandığımız algoritmalarda yaptığımız geliştirmelerle başarıda belirgin bir artış sağladıysak da, çoğunlukla tezin hazırlanma aşamasındaki süre kısıtlarından dolayı gerçekleştiremediğimiz, ileride yapılabilecek geliştirmeler ve iyileştirmeler mevcuttur.

Hazırladığımız sistemle elde ettiğimiz kümeleme işlemlerinde, küme etiketleri ve içerikleri için olabildiğince genel kavramlar atamaya çalıştık. Kullanılacak daha geniş veri setlerinden de yararlanılarak, daha ayrıntılı kümeler oluşturulabilir. Küme etiketleri daha özel kavramları temsil eder hale getirilebilir. Böylece birbirini kapsayan kümelerin de dikkate alınması sağlanarak, kümeler arasında bir hiyerarşi oluşturulabilir. Kümelerin alt-üst kümeleri bu şekilde bir hiyerarşik ağaç yapısında gösterilebilir. Bu şekilde bir kullanım, kullanıcının aradığına daha da kolay ulaşmasını sağlayacaktır.

Hazırlanan sistem, bir arama motoruna entegre edilerek özelleştirilmiş, gerçek zamanlı arama sonucu kümeleme işlemleri yapan bir internet sayfası oluşturulabilir. Böylece hazırladığımız yöntemler, daha çok ve geniş test verileriyle sınanacağı için, buradan çıkacak sonuçlarla, algoritmanın olası başka eksiklikleri de tespit edilebilir ve bu eksiklikler çözülebilir.

KAYNAKLAR

-
- [1] Worldwebsitesize, The size of the World Wide Web (The Internet), <http://www.worldwidewebsitesize.com/>, 16 Mayıs 2012.
 - [2] Broder, A., (2002). "Ataxonomy of web search". In ACM Sigir Forum, 36(2):3-10.
 - [3] Ferragina, P. ve Gulli, A., (2005). "A personalized search engine based on web-snippet hierarchical clustering", In Special interest tracks and posters of the 14th international conference on World Wide Web, 801-810, ACM New York, NY, USA.
 - [4] Chapman, archie - An Electronic Directory Service for the Internet, http://www1.chapman.edu/gopherdata/archives/Internet%20Information/w_hatis.archie, 3 Kasım 2012
 - [5] Cutting, D.R. , Karger, D.R. , Pedersen, J.O. ve Tukey, J.W., (1992). "Scatter/gather: A cluster-based approach to browsing large document collections", In Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval, (1): 318-329 ACM New York, NY, USA.
 - [6] Zamir, O. ve Etzioni, O., (1998). "Web document clustering: A feasibility demonstration", In Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, 46-54. ACM New York, NY, USA.
 - [7] Kale, A., Bharambe, U. ve Sashikumar, M., (2009). "A New Suffix Tree Similarity Measure and Labeling for Web Search Results Clustering", 2009 Second International Conference on Emerging Trends in Engineering & Technology: 856-861.
 - [8] Chim, H. ve Deng, X., (2007). "A New Suffix Tree Similarity Measure for Document Clustering", Proceedings of the 16th international conference on World Wide Web - WWW '07, 121-130.
 - [9] Wen, H. ve Huang, G., (2009). "Clustering web search results using semantic information", Machine Learning and Cybernetics, 2009 International Conference on, 3:1504-1509.
 - [10] Osinski, S., (2003). An Algorithm For Clustering of Web Search Results, Master Thesis, Poznan University of Technology, Polonya

- [11] Google Code, Zemberek 2 is an open source NLP library for Turkic languages. <http://code.google.com/p/zemberek/>, 9 Nisan 2012.
- [12] Zamir, O. ve Etzioni, O., (1998). "Grouper: a dynamic clustering interface to Web search results", Computer Networks-the International Journal of Computer and Telecommunications Networkin, 31(11):1361–1374.
- [13] Wang, J.ve Li, R. (2006). "A New Cluster Merging Algorithm Of Suffix Tree Clustering", In Intelligent information processing III: IFIP TC12 International Conference on Intelligent Information Processing (IIP 2006), September 20-23, Adelaide, Australia, 228:197-203.
- [14] Salton, G., Wong, A. ve Yang, C.S., (1975). "A vector space model for automatic indexing", Communication of ACM, 18(11):613–620.
- [15] Willet, P., (1988). "Recent trends in hierarchical document clustering: a critical review", Information Processing and Management, 24:577-97.
- [16] Zhang, D. ve Dong, Y., (2003). "Semantic, hierarchical, online clustering of web search results". Lecture notes in computer science, 69–78.
- [17] Hotho, A, Staab, S. ve Stumme, G., (2003). "Wordnet improves text document clustering", In Proc. of the SIGIR 2003 Semantic WebWorkshop, 541–544.
- [18] Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. ve Miller, K., (1990). "Introduction to Wordnet: An On-line Lexical Database", International Journal of Lexicography, 3(4):235-244
- [19] Hu, G., Zuo, W., He, F., ve Wang, Y. (2009). "Semantic-Based Hierarchicalize the Result of Suffix Tree Clustering". 2009 Second International Symposium on Knowledge Acquisition and Modeling, 3:221–224.
- [20] Search Engine Watch, Anatomy of a Search Engine: Inside Google, <http://searchenginewatch.com/article/2064446/Anatomy-of-a-Search-Engine-Inside-Google>, 12 Ekim 2012.
- [21] Anklesaria, M., McCahill, M., Lindner, P., Johnson, D. ve Torrey, D., (1993). "The Internet Gopher Protocol (a distributed document search and retrieval protocol)", Internet Engineering Task Force, Albert.
- [22] Brin, S. ve Page L., (1998). "The anatomy of a large-scale hypertextual Web search engine", Computer Networks and ISDN Systems, 30(1-7):107-117, Lawrence.
- [23] Brin S., (1999). "The PageRank Citation Ranking: Bringing Order to the Web", Stanford InfoLab, 66.
- [24] Google Blog, Our new search index: Caffeine, <http://googleblog.blogspot.com/2010/06/our-new-search-indexcaffeine.html>, 25 Nisan 2012
- [25] D. Weiss, (2002). "Introduction to search results clustering", In Proceedings of the 6th International Conference on Soft Computing and Distributed Processing, Rzeszów. Citeseer.

- [26] Osiriski, S. ve Weiss, D., (2005).“Carrot2: Design of a flexible and efficient web information retrieval framework”, In Advances in web intelligence: Third International Atlantic Web Intelligence Conference, AWIC 2005, Lodz, Poland, June 6-9, 2005: proceedings, 439. Springer-Verlag New York Inc.
- [27] Google Developers, Google Custom Search <https://developers.google.com/custom-search/>, 21 Kasım 2011.
- [28] Wikipedia, REST http://en.wikipedia.org/wiki/Representational_state_transfer, 29 Aralık 2011.
- [29] Porter, M.F., (1980).“An algorithm for suffix stripping”, Program, 14(3): 130–137.

BALIK SORGU SONUÇLARI

Çizelge EK-A.1 Balık Sorgu Sonuçları

0	http://tr.wikipedia.org/wiki/Bal%C4%B1k Balık - Vikipedi--Balıklar (Latince: Pisces) poikloterm olan, neredeyse sadece suda yaşayan ve solungaçları ile solunum yapan, soğuk kanlı, yürekleri çift gözlü, çoğunun vücudu ...
1	http://www.balikavi.net/ BALIK AVI--Balık Avı ve Amatör Balıkçılık, balık avı teknikleri, balık türleri ve olta takımları hakkında herşey.
2	http://search.izlesene.com/?vse=balik Balık Videoları İzlesene.com--Balık ile ilgili bütün videoları burada izleyebilirsiniz.
3	http://www.youtube.com/watch?v=I93gW2Ahu8Y BALIK AVI - YouTube--6 Nis 2007 ... arkadaşlar balık tam 26 kg geldi. dikkatli izlerseniz sudan çıktığı zaman ağzının hala acilip kapandığını görebilirsiniz. bu sene devami gelecek
4	http://www.nepisirsem.com/yemekkategori.aspx?k=Bal%FDk+ve+Deniz Balık ve Deniz - Resimli yemek tarifleri - ne pişirsem?--Balık ve Deniz - Resimli yemek tarifleri. ... Balık Çorbası (3) (8 KİŞİLİK) · Balık Çorbası · Balık Çorbası Terbiyesi · Balık Graten Sebzeli · Balık Köftesi · Balık Köftesi ...
5	http://www.baliksevdam.com/ Balık Avı Amatör ve Sportif Balık Avcılarının Buluşma Noktası ...--Balık Avı Sitemiz Deniz Balıkçılığı ve Tatlı Su Balıkçılığı ile ilgili Balık Avı Teknikleri, Balık Avı Malzemeleri, Misina, Olta , Kamış, Makine, İğne, Fırdöndü, ...
6	http://www.balikburcu.net/ Balık Burcu Burç Balık Burçlar Balık Burcu Özellikleri Günlük Balık ...--Balık Burcu - Birebir ilişkiler adına sizi kısıtlayan ve zorlayan konular üzerinde düşünmeniz gerek. Eşinizle olan ilişkinizde eleştirilere açık davranmalı ve ılımlı
7	http://www.ebk.gov.tr/ ET ve BALIK KURUMU GENEL MÜDÜRLÜĞÜ--ET ve BALIK KURUMU GENEL MÜDÜRLÜĞÜ. ... 14.08.2012 - Et ve Balık Kurumu Genel Müdürlüğü tarafından kamuoyuna duyurulur ... DEVAMI »
8	http://www.balikmarketim.com/ Balık Marketim.com Olta ve Balık Av Malzemesi Dalış Malzemeleri--Deniz malzemeleri ve Balık av malzemelerinin online satışının yapıldığı alışveriş sitesi.
9	http://www.balikdostlari.com/ Amatör Balıkçılık Ve Balık Dostları Platformu--Amatör Balıkçının Açılış Sayfası. Deniz, tatlısu balıkları, av araçları, balık avı teknikleri, balık yemekleri, av raporları, organizasyonları paylaşım alanı.

Çizelge EK-A.1 Balık Sorgu Sonuçları (Devam)

10	http://www.akvaryumculuk.biz/baliklar.html Balıklar ve Akvaryum--Balık, tatlı ve tuzlu suda yaşayan, evrimleşme çizgileri farklı, soğukkanlı omurgalıların genel adıdır. Bu terim, bir sınıflandırmadan çok bir yaşam biçimini tanımlar
11	http://www.amatorbalikavi.com/ Amatör Balık Avı--Ana Sayfa · Balık Cinsleri · Forum · Forum Resim Sayfası · Haberler · JİGGİNG BALIK AVI TURLARI ... Aileleri ile Beraber Yaptığımız Çeşme Balık Avı Turları
12	12. http://astroloji.ivillage.mynet.com/fal/gunluk/balik MYNET Astroloji - Balık Burcu 09 Eylül 2012 Yorumu--Sevgili Balıklar saat 03:13 Merkür-Jüpiter görünümü ile Pazar günümüz başlıyor. Bizim iş ortağından veya hayatımızı birlikte sürdüreceğimiz bir eşten neler ...
13	http://www.sadikoglu.info/balik_tur.htm Balık Türleri--Pervane balığı da denir. Okyanuslar ve Akdeniz'in açıklarında uzun göçler yapmadan, bazen derinlerde, bazen de su yüzeyinde yan yatarak dolaşır.
14	http://www.balikhotel.com/ Balık Otel--Balık Otel Denize sıfır, Antalya hava limanına 120 km, Şehir merkezine 1 km, uzaklıktadır. Doğanın güzellikleri yeryüzünde dağıtılırken, Alanya bundan fazlasıyla ...
15	http://www.tipeez.com/ch/oyunlar/Tipeez-Oyunlari/oyna/2750/balik-istifi.aspx Balık İstifi ve tüm Tipeez Oyunları Tipeez'de--Balık İstifi oyununda sizi çeşit çeşit suşi yapmaya yönlendiren 3 mod var. Aynı renkleri bir araya getirerek balıklardan suşi yap. Ancak hızlı olmayı unutma ve
16	http://www.oyungemisi.com/turkce-oyun/bir-balik-efsanesi/ Bir Balık Efsanesi Oyunu - Beceri Oyunları, Oyun, Oyunlar--Bir Balık Efsanesi Oyun. Küçük balıkları yiyerek büyü ve tehlikelerle dolu okyanusta hayatta kal. Bir Balık Efsanesi oyunu. En Güzel Beceri Oyunları. Oyun.
17	http://www.setbaliklokantasi.com/ Set Balık--Log in to Parallels Plesk Panel 9.5. Enter the login name into "Login" and password into the "Password" fields respectively. Then click "Log In".
18	http://video.cnnturk.com/actions/video/forward/177760 Kuşun ilginç balık yakalama tekniği CNN TÜRK Video--27 Ağu 2012 ... Bu kuşun avlanırken kullandığı teknik, balıkçıları kışkırtacak cinsten... Bir ekmek parçasıyla avını tuzağa düşüren kuş, izleyenleri şaşırtıyor...
19	http://www.yukselbalik.com/ Yüksel Balık Restaurant – Yeşilköy--Yüksel Balık Yüksel Balık, bugüne dek modern ve geleneksel servis anlayışımızla yüksek kalitede deniz ürünlerini oltanın ucundan sofranıza ulaştırmayı ilke ...
20	http://www.oyunlar1.com/kucukbalik.html Küçük Balık oyunları - Beceri oyunu--OYUNLAR 1 - Türkiye'nin en çok ziyaret edilen oyun sitesi : oyunlar1.com
21	http://tr.wikipedia.org/wiki/Beta_(balık) Beta (balık) - Vikipedi--Betta (Betta splendens), doğal yaşam alanı Asya (Tayland, Malezya, Kamboçya ve Vietnam) olan, yaklaşık 4-6 cm boyunda bir balık türü. Özellikle erkeklerin ...
22	http://www.youtube.com/watch?v=VS5es5LmcEQ Keloğlan Masalları Altın Balık Hikayesi [Tam,HD] - YouTube--3 Mar 2012 ... Keloğlan Masalları Altın Balık Hikayesi [Tam,HD]. KeloglanTV. Subscribe Subscribed Unsubscribe. Loading... 15 videos. Loading... Alert icon

Çizelge EK-A.1 Balık Sorgu Sonuçları (Devam)

23	http://www.posta.com.tr/astroloji/balik Balık - Posta Astroloji-+-Balık Burcu, balık burcunun özellikleri, burçlar ve özellikleri, burçlar, burç, astroloji, Balık Burcu, balık burcunun özellikleri, burçlar ve özellikleri, burçlar, burc, ...
24	http://www.surbalik.com/ Sur Balık Restaurant-+-Sur Balık restoranları, alanlarında uzmanlaşmış şef ve aşçılarıyla, Akdeniz mutfağı ve dünya mutfaklarının en seçkin deniz ürünleri ve mezelerini, sağlıklı, lezzetli ...
25	http://www.berkbalik.com/ Berk Balık, Bodrum-+-Berk Balık, egenin özel lezzetlerinin harmanlandığı balıkçı restoranıdır. Burayı diğer balıkçılardan ayıran, alıştığımız ve de her yerde bulabildiğimiz mezelerin ...
26	http://takanik.com.tr/ Takanik Balık-+-Takanik balık hizmete 1998 yılında küçük teknesiyle Yeniköy limanında başladı. lezzeti ve uygun fiyatlarıyla kısa zamanda müşteri kitlesini genişleten Takanik, ...
27	http://tr.wikipedia.org/wiki/Akci%C4%9Ferli_bal%C4%B1klar Akciğerli balıklar - Vikipedi-+-Akciğerli balıklar (Dipnoi), Sarcopharyngii sınıfından solungaç solunumu yapmakla beraber ihtiyaç duyulduğunda hava solunumu da yapabilen tatlı sularda ...
28	http://catana.com.tr/ ÇATANA balık restaurant-+-İstanbul Anadolu yakasında, Suadiye sahilinde bulunan; balık ve meze kültürüne yeni bir soluk getiren, Çatana Balık 2001 yılında hizmete açılmıştır. Açıldığından ...
29	http://tr.wikipedia.org/wiki/Wanda_Ad%C4%B1nda_Bir_Bal%C4%B1k Wanda Adında Bir Balık - Vikipedi-+-Wanda Adında Bir Balık 1988 ABD - İngiltere ortak yapımı kara komedi filmidir. Özgün adı A Fish Called Wanda olan film çevrildikten bir yıl sonra, 1 Aralık ...
30	http://www.ebk.gov.tr/magazalarimiz-icerik-227-1.htm ET ve BALIK KURUMU GENEL MÜDÜRLÜĞÜ - Mağazalarımız-+-ET ve BALIK KURUMU GENEL MÜDÜRLÜĞÜ - ... 10.04.2012. ET VE BALIK KURUMU BALIK SATIŞINA BAŞLAMIŞTIR. 29.11.2011. EBK 2010 YILI ...
31	http://tr.wikipedia.org/wiki/Bal%C4%B1k_burcu Balık burcu - Vikipedi-+-Balık(00) Zodyağın son astrolojik burcudur. Batı astrolojisine göre 20 Şubat ile 20 Mart arasında doğanlar bu burca mensup olur. Astrolojide, Balık negatif bir burç ...
32	http://tr.wikipedia.org/wiki/Kalkan_bal%C4%B1%C4%9F%C4%B1 Kalkan balığı - Vikipedi-+-Kalkan balığı (Psetta maxima), Scophthalmidae familyasına ait, gözleri vücudunun sol tarafında bulunan ve sağ tarafı ile denizin tabanında yatan bir yassı balık ...
33	http://www.balikpesinde.com/ Balikpesinde.Com Balık Peşinde Yaşanacaklara-+-Balık peşinde koşan bir grup avcının günlüğü ve incelemeleri.
34	http://www.poyrazbalik.com/ Poyraz Balık Restaurant-+-Poyraz Balık Restaurant'a Hoşgeldiniz. Mavi ve yeşilin iç içe girdiği, natürel bahçe mimarisi, her türlü ayrıntının düşünüldüğü, çağdaş dekorasyonu, maksimum ...
35	http://tr.wikipedia.org/wiki/Astronot_bal%C4%B1%C4%9F%C4%B1 Astronot balığı - Vikipedi-+-Astronot balığı ya da Oskar (Astronotus ocellatus), doğal yaşam alanı Güney Amerika olan, yaklaşık 20-30 cm boyutlarında, yutabileceği tüm küçük balıklara ...

Çizelge EK-A.1 Balık Sorgu Sonuçları (Devam)

36	http://www.balikoyunlari.net/ Balık Oyunları, Balık Oyunu-Balikoyunlari.Net-- Balık oyunları, balık tutma, balık besleme, akvaryum oyunları hergün güncellenmektedir.
37	http://www.serminbalik.com/ Sermin BALIK--Aslen Elazığ ili nin Merkeze bağlı YALAVUZ köyünden olan Sermin Balık anne ve babasının memuriyetinden dolayı Tunceli ilinde 1972 yılında dünyaya geldi.
38	http://www.misinabalik.com/ Misina Balık Restaurant--Misina Balık Restaurant. ... İletişim & Rezervasyon · Misina Galeri. Yükleniyor.. Misinada balık yemek ayrıcalıktır. Yükleniyor.. Misinada balık yemek ayrıcalıktır.
39	http://tr.wikipedia.org/wiki/Bal%C4%B1k_eti Balık eti - Vikipedi--Balık eti deniz veya tatlı sularda yakalanan soğuk kanlı solungaçları ile solunum yapan hayvanların besin olarak kullanılan etleridir. Midye, yengeç, istakoz gibi .
40	http://tr.wikipedia.org/wiki/Palamut_(bal%C4%B1k) Palamut (balık) - Vikipedi-- "http://tr.wikipedia.org/w/index.php?title=Palamut_(balık)&oldid=11336452" adresinden alındı. Kategoriler: Balık taslakları · Balıkçılık ...
41	http://www.sandalbalikyalova.com/ Yalova Sandal Balık Restorant--Yalova Sandal Balık Restorant - Yalova Sandal Balık - Yalova Sandal - Sandal Balık - Balık Restorant - Balık Restoranı - Yalova Restorant - Yalova Balık - Balık ...
42	https://foursquare.com/v/revma-balik-arnavutkoy/4f80a49fe4b01e69d72b8943 Revma Balık Arnavutkoy --+See 37 photos and 19 tips from 147 visitors to Revma Balık Arnavutkoy. "Gerek yediğimiz yemeğin çeşit ve lezzetinden gerekse servis ve ağırlamadan..."
43	http://www.adembaba.com/ Adem Baba Balık Lokantası--Mevsimin tüm balıklarını en taze ve en ucuz şekilde sunmayı amaç edinmiş, gösterişten uzak dekore edilmiş, sınırsız keyifli mi keyifli bir alkolsüz balık lokantası .
44	http://www.oyunlar1.com/balikarci.html Balık Arçı oyunları - Beceri oyunu--OYUNLAR 1 - Türkiye'nin en çok ziyaret edilen oyun sitesi : oyunlar1.com.
45	. http://www.oyunskor.com/savaşanbalik.htm Savaşan Balık oyunları - Beceri oyunu oyna--Savaşan Balık Oyunu, Oyun Skor en güzel oyunlar Beceri Oyunları.
46	. http://www.metro-tr.com/public/taze-balik Metro Toptancı Marketiniz Taze Balık-- Her gün en az 40, sezonunda 100 çeşit ve üzeri taze ve hesaplı deniz ürününü bulabilmeniz için çalışıyoruz!
47	http://tr.wikipedia.org/wiki/Bal%C4%B1k_kartal%C4%B1 Balık kartalı - Vikipedi--Balık kartalı (Pandion haliaetus), balık kartalığıllar (Pandionidae) familyasından balıkla beslenen yırtıcı bir kuş türü. İri, uzun kanatlı, suya dalarak avlanan bir ...
48	http://balikfosilleri.com/ Balık Fosilleri.com Balık Fosilleri, Fosil, Fosil Nedir, Fosiller--Bu sitede yer alan balık fosilleri, Yaratılış gerçeğini ispatlayan yüz milyonlarca fosilin sadece birkaç örneğidir. Ve sadece bu örnek bile, evrim teorisinin bilim ...
49	http://www.astroloji.org/balik.asp Balık Burcu--Balık burcu kadını hırslı,tuttuğunu koparan, çekici, sempatik ve iyi huylu tavırlarıyla dikkati çeker. Duygusal ve romantiktir. Evcimen ve evine bağlı bir yapısı vardır ...

DÜNYA SORGU SONUÇLARI

Çizelge EK-B.1 Dünya Sorgu Sonuçları

0	http://www.dunya.com/ Dünya Gazetesi - Ekonomi Haber Portalı--Güncel ekonomi, finans ve politika haberleri, emtia verileri, imkb, döviz ve yurt dışı piyasa verileri ve daha fazlası.
1	http://tr.wikipedia.org/wiki/D%C3%BCnya Dünya - Vikipedi--Dünya (Yeryuvarı, eski: Küre-i Arz), Güneş Sistemi'nin Güneş'e uzaklık açısından üçüncü sıradaki gezegenidir. Üzerinde zeki yaşam formları barındırdığından ...
2	http://dunya.milliyet.com.tr/ Dünya HABER Son Dakika Dünya Haberleri Milliyet Haber--Milliyet Haber'de en son dünya haberleri sizleri bekliyor. En güncel son dakika dünya haberleri için hemen tıklayın.
3	http://www.dunya.com.tr/ Dünya Radyo - Ana sayfa--Ulusal Yayın yapan radyo kuruluşunun yayın akışı ve programları hakkında bilgi veren resmi sitesi.
4	http://www.hurriyet.com.tr/dunya/ Dünya - Hürriyet--Türkiye'den ve dünyadan tüm son dakika gelişmeler burada, haber spor yaşam ifestyle interaktif forum bilgi piyasalardaki son dakika gelişmeler ...
5	http://www.haberler.com/dunya/ dünya DÜNYA dünya DUNYA dünya haberleri haber haberi--dünya DÜNYA dünya DUNYA dünya ile ilgili en yeni güncel haberler, son dakika haberleri. dünya.
6	http://www.stargazete.com/dunya/ Dünya--Şırnak'ta PKK'ya büyük darbe Şırnak'ın Beytüşşebap ilçesinde yapılan operasyon kapsamında 3 terörist etkisiz hale getirildi. Cudi Dağı'nda dün yapılan ...
7	http://www.trthaber.com/haber/dunya/ Dünya - TRT Haber - Türkiye'nin Haber Kaynağı--Dünya Dünya haberler... TRT Haber DD Kasım Sayısını İncelediniz mi? t.co/ rRcjdsPg · trt_gonultas Türkiye dünyada 7. büyük tarım ülkesi imiş. Hedef ise ilk 5 .
8	http://www.cnnturk.com/dunya/default.aspx Dünya Gündemi Avrupa, Amerika, Afrika, Asya, Ortadoğu, ABD, AB ...--Dünya Gündemi Avrupa, Amerika, Afrika, Asya, Ortadoğu, ABD, AB Gündemi, Uluslararası Gelişmeler ve Politika , Son Dakika ve Güncel Gelişmeler, Fotoğraf ...

Çizelge EK-B.1 Dünya Sorgu Sonuçları (Devam)

9	http://www.butundunya.com/ Bütün Dünya--+--blah blah. Kapat. Kapak. Kapak. Doğum Yılı (*). Mesleğiniz (*). Bulunduğunuz Kent (*). Bulunduğunuz Ülke (*). E-Posta. Eleştiri ve Yorumlarınız ...
10	http://www.ntvmsnbc.com/id/24927413/ Dünya Haberleri- ntvmsnbc.com--+-- NTVMSNBC Dünya Haberleri. ... Dünya sendikalarından Gül'e çağrı · AB: Açlık grevlerinden endişe duyuyoruz. Video - Foto Galeri. Başlıklar. Almanya · Fransa .
11	http://www.worldometers.info/tr/ Worldometers - Gerçek zamanlı dünya istatistikleri--+--Dünya nüfusunun anlık ölçümleri, karbon dioksit co2 emisyonu, dünyadaki açlık, kamusal harcamalar, üretim rakamları, tüketim verileri ve daha bir çok konuda ...
12	http://www.goturkey.com/dunya_mirasi_tr.php?lng=tr UNESCO Dünya Mirası Türkiye--+--Dünya Mirası Listesi Nedir? İnsanoğlu tarih sahnesine çıktığından bu yana yaşadığı alanları değerli kılmış ve kültürünü gelecek nesillere taşıyacak şaheserler ...
13	http://www.sondakika.com/dunya/ Son Dakika Dünya Haberleri--+--Son dakika Dünya haberleri, Dünya DÜNYA dünya Dunya DUNYA dünya.
14	http://www.haberturk.com/dunya Dünya-HABERTÜRK-TÜRKİYE'NİN EN BÜYÜK İNTERNET GAZETESİ--+--Haber Türkiye'de Habertürk'ten takip edilir, haberin kaynağı Haberturk.com. En son haberler Habertürk'te.
15	http://www.internethaber.com/dunya-haberleri-5.htm DÜNYA haberleri - İnternet Haber - Son Dakika Haberler - Haber ...--+--DÜNYA haberleri DÜNYA ile ilgili haberler. ... Dünya'nın takip ettiği ABD Başkanlık seçimlerinde ABD'nin seçimi yine Obama oldu. Obama 303 delege çıkardı ...
16	http://tr.wikipedia.org/wiki/D%C3%BCnya_(anlam_ayr%C4%B1m%C4%B1) Dünya (anlam ayrımı) - Vikipedi--+... ile ilgili tanım bulabilirsiniz. Dünya, aşağıdaki anlamlara gelebilir: ... yayınlanmış olan gazete. Dünya (gazete, 1981) - Türkiye'de yayımlanan ekonomi gazete ...
17	http://www.namazvakti.com/ Dünya Namaz Vakitleri--+--Dünya Şehirleri için namaz vakitleri (Prayer times for cities the worldover)
18	http://www.dunya.com.au/ Dünya Gazetesi Dünya, Avustralya, Türkiye, Güncel Haberleri ...--+--Dünya Gazetesi Dünya, Avustralya, Türkiye, Güncel Haberleri, Gazete Haberleri , Avustralya Yerel Haberleri, Spor Haberleri, Magazin Haberleri.
19	http://www.dunyabulteni.net/ Dünya Bülteni--+--Dünya ve Türkiye gündemini meşgul eden haberler ve makaleler yer alıyor.
20	http://gazete.netgazete.com/ShowPaper.aspx?news=dunya dünya--+--Dünya · Fanatik · Fotomac · Güneş · Haberturk · Hürriyet · İstanbul Ticaret · Milli Gazete · Milliyet · Posta · Radikal · Sabah · Sözcü · Star · Takvim · Taraf · Türkiye ...
21	http://www.dunyatelekom.com/ DÜNYA TELEKOM ucuz telefon görüşmesi İnternet üzerinden ...--+--internet telefon görüşmesi, internet üzerinden telefon görüşmesi, internetten telefon görüşmesi, telefon görüşmesi, ucuz telefongörüşmesi, uluslar arası telefon ...

Çizelge EK-B.1 Dünya Sorgu Sonuçları (Devam)

22	http://www.wtcistanbul.net/ İstanbul Dünya Ticaret Merkezi--İstanbul Dünya Ticaret Merkezi. ... Bugün dünya üzerinde faaliyet gösteren yaklaşık 300 Dünya Ticaret Merkezinin büyük bir kısmı gayrimenkul ... İFM ...
23	http://tr.wikipedia.org/wiki/Orta_D%C3%BCnya Orta Dünya - Vikipedi--Orta Dünya kıtası, J. R. R. Tolkien'in kurguladığı Orta Dünya evreninde, yazarın eserlerinin büyük bir kısmına mekân olan hayalî kıtadır. Bu eserlerin en ...
24	http://www.kanald.com.tr/YalanDunya Yalan Dünya / Dizi / Kanal D / İzle--12 saat önce ... Yalan Dünya dizi sayfası izle. ... Yalan Dünya / Dizi / Kanal D / İzle. Kanal D - Türkiye'nin Kanalı. Yalan Dünya. Yalan Dünya. Sevdiğin dizinin ...
25	http://www.dunyagoz.com/ Dünya Göz Hastanesi--Dünya'nın en büyük göz hastanesi, en yüksek standartlarda LASIK, lazer tedavisi , katarakt, keratokonus, retina, glokom, kornea ve diğer göz sağlığı hizmetleri ...
26	http://www.sporx.com/futbol/dunya/ Dünya Futbolu - Anasayfa--FIFA dünya ülkeler sıralamasını açıkladı. ... 07.11.12 FIFA Dünya Sıralaması açıklandı (Dünya Futbolu) ... 07.11.12 Güiza: "Her şey para değil" (Dünya Futbolu) ...
27	http://tr.euronews.com/hava/ world Hava durumu euronews: world 10 günlük hava durumu tahmini--Hava durumu tahmin world euronews, hava durumu tahminleri world (sıcaklık, rüzgâr, yağış miktarı...). 10 günlük hava durumu tahmini world.
28	http://www.dunyastore.com/ Dünya Store--Dünya Grubu Yayınları · Dünya Gazetesi · Dünya Grubu Dergileri · e-Dünya · Kitaplar. Akkor · Albüm Dizisi · Amatörler İçin Serisi · Anı · Anlatı Ustaları · Antoloji ...
29	http://www.dunyatv.com.tr/ Dünya TV - Türkiye'nin ilk özel kürtçe televizyonu--Di heqê me de; Filmê Danasînê · Katalog; Dünya TV Logo. PSD · AI · PNG · JPG · Bername · ZİMANÊ ŞÎRÎN · TARÇIN Û HEVALÊN WÊ · RÛ BÎ RÛ · ROJA KU ...
30	http://tr.wikipedia.org/wiki/D%C3%BCnya_Bankas%C4%B1 Dünya Bankası - Vikipedi--Dünya Bankası, II. Dünya Savaşı'nın ardından 1945 yılında Uluslararası Yeniden Yapılanma ve Kalkınma Bankası (IBRD=International Bank for Reconstruction ...
31	http://www.dunyagida.com.tr/ Dünya Gıda--Dünya Gıda Derigisi Logo. Reklam ... Dünya Gıda Dergisi - 2012/Ağustos ... Dünya gıda kaynaklarına erişimde önemli bir sektör; su ürünleri yetiştiriciliği ...
32	http://www.bigpara.com/borsa/dunya-borsa-endeksleri/2/kasim/2012/ Dünya Borsa Endeksleri Borsa Bigpara.com--6 gün önce ... Kudret AYYILDIR Altın'da kritik viraja geldik... ○ Figen ÖZAVCI Notunu artıramadıklarımızdan mısınız? ○ Oğuz AKDENİZ Merak etmiyor ...
33	http://www.sabah.com.tr/dunya/ Sabah Dünya Haberleri - 06 Kasım 2012--Dünya. Liderlerin kanser laneti Liderlerin kanser lanetiLatin Amerikalı liderlerin teker teker kansere yakalanmaları şüpheleri ve komplo teorilerini de beraberinde ...

Çizelge EK-B.1 Dünya Sorgu Sonuçları (Devam)

34	http://history1900s.about.com/od/worldwarII/World_War_II.htm World War II--A truly world war which lasted from 1939 to 1945. The countries of the world aligned with either the Axis powers or the Allies and battled in a total war.
35	http://www.facebook.com/DunyaGerceklerim Dünya Gerçekleri Realities of the World Facebook--To connect with Dünya Gerçekleri Realities of the World, sign up for Facebook today Kuralları yada yöneticileri olmayan sınır yada engel tanımayan bir dünya , ...
36	http://egazete.dunya.com/ e-Dünya--Diğer Sitelerimiz. kobiden.com · dunyastore.com · pcworld.com.tr · computerworld .com.tr · byte.com.tr · dunyagida.com.tr · dunyainsaat.com.tr · makinamagazin.
37	http://www.sabah.com.tr/fotohaber/dunya Sabah - HaberPlus - Dünya - 05 Kasım 2012--En kötü 50 icat · Boynuz taşıyan insanlar · Dünyanın en gelişmiş savaş uçakları · Dünyanın en tuhaf ağaçları · Rakamlarla Türkiye ve Suriye orduları · 1700 kişiyi ...
38	http://tr.wikipedia.org/wiki/D%C3%BCnya_n%C3%BCfusu Dünya nüfusu - Vikipedi--Bu terim, dünya üzerinde yaşayan insan sayısını belirtmek amacıyla kullanılır. En güncel tahminlere göre dünya nüfusu 6,6 milyara 2007 Haziran'ında ulaşmıştır ...
39	http://www.kulturvarliklari.gov.tr/TR,44423/dunya-miras-listesi.html Dünya Miras Listesi--Bütün insanlığın ortak mirası olarak kabul edilen evrensel değerlere sahip kültürel ve doğal varlıkları dünyaya tanıtmak, toplumda söz konusu evrensel mirasa ...
40	http://www.yenidunyadergisi.com/ Yenidünya Dergisi {Gönül Dünyanıza Açılan Kapı}--Yaratılanların en şerefli olan insanın, kendisini yaratan Rabbinin istediği yolda, gönderdiği kitap ve elçilerinin yolunda gidebilmesi, sonuçta dünya imtihanının ...
41	http://www.dektmk.org.tr/ DEK-TMK Dünya Enerji Konseyi--Dünya Enerji Konseyi Türk Milli Komitesi (DEKTMK)'nın websitesinden kuruluşun , tanım, amaç, tüzük, üye ve yönetim bilgisi gibi idari bilgilerine erişebilirsiniz.
42	http://www.youtube.com/watch?v=jjDKGtNswtE Dünya (Yasemin Mori) - YouTube--4 Tem 2012 ... Beni Severmiş O (Asfalt Dünya)by muyap1,024,882 views; Nil Karaibrahimgil - Ben Buraya Çıplak Geldim 4:06. Watch Later Nil Karaibrahimgil ...
43	http://www.dunyaekonomi.tv/ DÜNYA TV » Ana Sayfa--Her hafta ekonomik gidişatın ve iş hayatının nabzını tutan Volkan Özsoy'la İş DÜNYA'sı Turu, İş dünyasının her kesimine uzanan konuk ve program yelpazesıyla ...

Çizelge EK-B.1 Dünya Sorgu Sonuçları (Devam)

44	http://www.youtube.com/watch?v=IDD5uuHATRU Koro - Batsın Bu Dünya / 2012 Orijinal Şarkı / "Orhan Gencebay İle ...-+-20 Eyl 2012 ... Müzik devi Orhan Gencebay'ın 20 Eylül'de yayınlanan Tribute albümü "Orhan Gencebay İle Bir Ömür"den albümde yer alan tüm şarkıcıların ...
45	http://tr.wikipedia.org/wiki/D%C3%BCnyan%C4%B1n_Yedi_Harikas%C4%B1 Dünyanın Yedi Harikası - Vikipedi-+-Dünyanın Yedi Harikası, tamamı insanoğlu tarafından inşa edilmiş, olağanüstü antik yapı ve yapıtlardır. Ayrıca Antik Dönemin Yedi Harikası adıyla da anılırlar.
46	http://www.tekdiziizle.com/diziler/agir-roman-yeni-dunya Ağır Roman Yeni Dünya-+-Ağır Roman Yeni Dünya 5.Bölüm tek parça izle Ağır Roman Yeni Dünya 5.Bölüm izle, Ağır roman yeni dünya dizisinin 5.bölüm 1 kısım 2012 perşembe tarihli ...
47	http://www.wwf.org.tr/ WWF Türkiye - Yaşayan bir dünya için-+-Türkiye'de su kaynakları, ormanlar, denizler ve kıyıların korunması ile ilgili programlar uygulayan kuruluş.
48	http://www.hddiziseyretizle.com/diziseyret/diziler/yalan-dunya Yalan Dünya-+-Dizi izle, Son Bölüm izle, Dizi Seyret, Yeni Bölüm izle, Full Dizi, HD Dizi izle. HDDiziseyretizle.com Hos Geldiniz. Felekten Bir Gece 2 Türkçe Dublaj izle ...
49	http://dunya.nedir.com/ Dünya - Nedir-+-Dünya nedir? Güneş sistemindeki gezegenlerden biri. Dünyanın uzaydan görünüşü mavi olduğu için uzay dilinde dünya mavi gezegen ismi ile de...

OCAK SORU SONUÇLARI

Çizelge EK-C.1 Ocak Sorgu Sonuçları

0	http://tr.wikipedia.org/wiki/Ocak Ocak - Vikipedi--Sözcüğün, ocak >>> "ocak" olduğu savlanırken, öte yandan Eski Türkçedeki "oc -mak, uc-mak" "ateş tutuşturmak" kökünden de olabilir. Ocak, ateş yakılan yer, ...
1	http://www.ocak.com.tr/ Ocak Elektronik - Anasayfa fotoğraf makinası, video kamera, baskı ...--Sony, Kodak ve daha birçok markanın Türkiye dağıtıcısı.
2	http:// bayi.ocak.com.tr/ OCAK ELEKTRONİK - BAYİ SİSTEMİ--Şifre paylaşımından doğabilecek zararlardan Ocak Elektronik sorumlu tutulamaz. Bu site Ocak Elektronik bayileri içindir. Eğer henüz bayimiz değilseniz, iletişim
3	http://www.ucuzu.com/ocak-c4172/ Set üstü ocak--Set üstü ocak - Fiyatları, ürün yorumları, resimleri ve en ucuz ...
4	http://www.akakce.com/ocak.html en ucuz Ocak / Setüstü Ocak fiyatları akakce.com'da--Onlarca alışveriş sitesinden en ucuz Ocak / Setüstü Ocak ...
5	http://www.siemens-home.com/tr/%C3%BCr%C3%BCnler/pi%C5%9Firme-grubu/ocaklar.html Siemens Ev Aletleri Ocaklar - Ürünler - Pişirme Grubu - Ocaklar - List--70 cm cam seramik ocak , 4G + 1W , Facette Tasarım , flameTronic: otomatik yeniden ateşleme , MainSwitch: Tek tuş ile cihazı açma ve durdurma , 6 saat sonra ...
6	http://www.tcmb.gov.tr/kurlar/201001/Jan.html ocak - 2010 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 04 OCAK PAZARTESİ / 04 JANUARY MONDAY ...
7	http://www.5ocaknews.com/ 5 Ocak Gazetesi 5 Ocak Çukurovanın En İyi Gazetesi--Beş Ocak Gazetesi, 5 ocak gazetesi, adana, adana haber, çukurova, adana medya, adana spor, demir spor.
8	http://www.tcmb.gov.tr/kurlar/201201/Jan.html ocak - 2012 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 02 OCAK PAZARTESİ / 02 JANUARY MONDAY ...
9	http://www.ocakder.org.tr/ Ocak Köyü Derneği - Kemaliye Erzincan--OCAK KÖYÜ TANITIMI · TARİHÇE · COĞRAFI KONUM · NASIL GİDİLİR · BAĞLI KÖYLER · AMAÇ · YÖNETİM KURULU · OCAK KÖYÜ TANITIMI · TARİHÇE ...

Çizelge EK-C.1 Ocak Sorgu Sonuçları (Devam)

10	http://www.tcmb.gov.tr/kurlar/200401/Jan.html ocak - 2004 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 02 OCAK CUMA / 02 JANUARY FRIDAY · 05 OCAK ...
11	http://www.turkpatent.gov.tr/dosyalar/istatistik/marka/Marka_tescil_ulkelere_gore_dagilim_aylik.pdf Ocak Şubat Mart Nisan Mayıs Haziran Temmuz Ağustos Eylül Ekim ...--Ocak Şubat Mart Nisan Mayıs Haziran Temmuz Ağustos Eylül Ekim Kasım Aralık. Ülke kodu Ülke Adı. TR TÜRKİYE. 3777 3796. 4983. 2856. 3847. 2179. 3405 ...
12	http://tr.wikipedia.org/wiki/15_Ocak 15 Ocak - Vikipedi--15 Ocak, Gregoryen Takvimi'ne göre yılın 15. günüdür. ... Yaser Arafat'ın 11 Kasım 2004'te ölmesinin ardından 9 Ocak'ta devlet başkanı seçilen Mahmud Abbas, ...
13	http://www.ocakhaddecilik.com.tr/ Ocak Haddecilik Metal San. Ltd. Şti.--Ocak Haddecilik Süper Kalite ve uygun fiyatları ile demir, sıcak izli demir, ferforje, kare lama ve kolon boyu inşaat demirinde tercih edilen marka.
14	http://www.turktakvim.com/6/Aylik_Namaz_Vakitleri/Ocak/1/16741/%C4%B0st_anbul.html İstanbul İçin 2012 Ocak Ayı Namaz Vakitleri--Tarih, İmsâk, Güneş, İşrak, Öğle, İkinci, Akşam, Yatsı. 1 Pazar, 5:33, 7:22, 8:17, 12:19, 14:39, 16:54, 18:31. 2 Pazartesi, 5:33, 7:22, 8:17, 12:19, 14:40, 16:55, 18: ...
15	http://www.iletisim.com.tr/kitap/ocak-ve-zanaat-1276.aspx Ocak ve Zanaat / Güney Çeğin, Emrah Göker, Alim Arlı, Ümit ...--Günümüz sosyolojisini ve kuramlarını en derinden etkileyen isimlerin başında kuşkusuz Fransız sosyolog Pierre Bourdieu gelmektedir. 2002'de hayatını ...
16	http://www.ocakmobilya.com/ Ocak Mobilya--Ocak Mobilya. Your image title Hayatın her alanında yeniliklere açıksan, yaşamına başka bir değer katmak istiyorsan, ideallerinle öndeysen, alışılmışı değil, yeni ...
17	http://www.turkpatent.gov.tr/dosyalar/istatistik/marka/Marka_tescil_illere_gore_dagilim_aylik.pdf Ocak Şubat Mart Nisan Mayıs Haziran Temmuz Ağustos Eylül Ekim ...--Ocak Şubat Mart Nisan Mayıs Haziran Temmuz Ağustos Eylül Ekim Kasım Aralık. İl Kodu İller. 01 ADANA. 56. 51. 81. 52. 87. 31. 49. 59. 65. 02 ADIYAMAN. 5. 1 ...
18	http://www.tcmb.gov.tr/kurlar/201101/Jan.html ocak - 2011 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 03 OCAK PAZARTESİ / 03 JANUARY MONDAY ...
19	http://www.arcelik.com.tr/set-ustu-ocak.html Arçelik - Set Üstü Ocak--Arçelik set üstü ocak ürünlerinin özelliklerini ve kampanyalarını bu sayfada bulabilirsiniz.
20	http://www.naturelocak.com/ :::NATUREL OCAK::: Mutfağınıza Özel--NATUREL OCAK::: Mutfağınıza Özel. ... GÜVENİLİR. KALİTELİ. EKONOMİK. AR- GE. ÜRÜNLER. HAKKIMIZDA. İLETİŞİM. Copyright © 2012 Naturel Ocak.
21	http://tr.wikipedia.org/wiki/1_Ocak 1 Ocak - Vikipedi--1 Ocak, Gregoryen takvimine göre yılın 1. günü, yılbaşı. öğrencisi Adem Doğan, başına isabet eden kurşunla yaralandı. Doğan 4 Ocak'ta hastanede öldü.
22	http://www.tcmb.gov.tr/kurlar/200801/Jan.html ocak - 2008 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 02 OCAK ÇARŞAMBA / 02 JANUARY ...

Çizelge EK-C.1 Ocak Sorgu Sonuçları (Devam)

23	http://www.ocak-koy.com/ Ocak koy holiday village fethiye turkey--Love at first sight: Those who have been to Ocakköy once admit that this really exists. To be in love with a place where nature is still allowed to remain nature, ...
24	http://www.idefix.com/kitap/ahmet-yasar-ocak/urun_liste.asp?kid=327 Ahmet Yaşar Ocak - Kitap idefix.com--Ahmet Yaşar Ocak - Kitap. ... kultur-tarihi-kaynagi-olarak-menakibnameler-ahmet -yasar-ocak. kitap ... osmanli-toplumunda-tasavvuf-ve-sufiler-ahmet-yasar-ocak ...
25	http://kur.doviz.com/arsiv/serbest-piyasa/euro/2012/01/ 01-2012 Euro Serbest Piyasa Kur Arşivi - Doviz.com--01/Ocak/2012, 2,442, 2,462. 02/Ocak/2012, 2,438, 2,458. 03/Ocak/2012, 2,435, 2,455. 04/Ocak/2012, 2,426, 2,444. 05/Ocak/2012, 2,407, 2,424. 06/Ocak/2012 ...
26	http://www.madenveocak.com.tr/ Maden ve Ocak Teknolojileri Dergisi--Maden ve Ocak Teknolojileri ürünleri ve ekipmanları sektörünün internet gazetesidir.
27	http://www.tcmb.gov.tr/kurlar/200701/Jan.html ocak - 2007 / january--Türkiye Cumhuriyet Merkez Bankası Döviz Kurları / Central Bank of the Republic of Turkey Exchange Rates. 04 OCAK PERŞEMBE / 04 JANUARY THURSDAY ...
28	http://www.diyadinnet.com/RuyaTabiri-7493&Tabir=ocak 'rüyada Ocak görmek' ne demektir tabiri 7493 ? Ocak...--Rüyada boş ocak görmek henüz hayata geçirmediğiniz plan halinde olan bir iştir. İçinde kömür bulunan ocak, iş yaşamınızda verdiğiniz emekler sonrasında ...
29	http://tr.wikipedia.org/wiki/8_Ocak 8 Ocak - Vikipedi--8 Ocak, Gregoryen takvimine göre yılın 8. günü. ... bir mezarda bulundu ve çalanlar yakalandı. Koç'un naaşı, 10 Ocak'ta ikinci kez toprağa verildi. 2003 - THY'nin ...
30	http://www.wien.gv.at/tr/entegrasyon/takvim/ocak.htm Ocak 2012 - En büyük dini ve etnik grupların önemli bayramları--6 Ocak: Ortodoksların Noel Gecesi (Julian Takvimine göre) ... 6 Ocak ayrıca Epifani bayramı olarak da kutlanır, Epifani Yunanca bir sözcük olup, "ortaya çıkma, ...
31	http://www.tff.org/Default.aspx?pageID=72&hakemID=19703 ÖMER FARUK OCAK - Hakem Bilgileri TFF--ÖMER FARUK OCAK - Türkiye Futbol Federasyonu Resmi İnternet Sitesi - Milli Takımlar, Spor Toto Süper Lig, Puan cetveli, Fikstür, Bankasya 1. Lig, Spor Toto 2 ...
32	http://www.gittigidiyor.com/arama/?k=set+%FCst%FC+ocak Set üstü ocak - GittiGidiyor--Set üstü ocak için arama sonuçları. Set üstü ocak için benzersiz seçenekler, marka ve uygun fiyatları ile sizi bekliyor. Set üstü ocak, GittiGidiyor'da!
33	http://www.kadinlarkulubu.com/ocak-anneleri-f315/ Ocak Anneleri--Ocak Anneleri - ... Aşağıda "Kadınlar Kulübü, Kadın Sosyal Ağı" kategorisinde " Ocak Anneleri" ... Anket: Ocak 2013 'te ANNE OLUCAKLAR burada toplanalım ...
34	http://www.teknosa.com/kategori/beyaz_esya/ocak/1483 OCAK - Teknosa.com--Akbank, x 12. Garanti Bankası, x 12. İş Bankası, x 5. HSBC, x 5. Finansbank, x 5. Yapı Kredi Bankası, x 5. Taksit Seçenekleri ...

Çizelge EK-C.1 Ocak Sorgu Sonuçları (Devam)

35	http://www.history.hacettepe.edu.tr/ocak.html Prof. Dr. Ahmet Yaşar OCAK - Hacettepe Üniversitesi Tarih--Adı Soyadı, Prof. Dr. Ahmet Yaşar OCAK. Görevi, Öğretim Üyesi. Öğrenim. Lisans : İst. Yüksek İslam Enstitüsü (M.Ü. İlahiyat Fak.). (1963-1967). İ.Ü.Edebiyat Fak.
36	http://www.milliyetemlak.com/ocakemlakgayrimenkulcom OCAK GAYRİMENKUL İNŞ.TURZ.PAZR.TİC.VE LTD.ŞTİ Emlak Ofisi ...--OCAK GAYRİMENKUL ALTI YILLIK TECRÜBE KONUSUNA HAKİMİYET VE GÜVEN İN ... Ocak gayrimenkul öveçler 4 ncü cadde yakını satılık 7+1 dubleks daire
37	http://www.accuweather.com/tr/it/rome/213490/january-weather/213490 Roma İçin Ocak Ayı Hava Durumu 2012 - İtalya İçin AccuWeather ...--Pazar, Pazartesi, Salı, Çarşamba, Perşembe, Cuma, Cumartesi. Oca 1. Gerçek Sıcaklık. 11° Dü 1°. Geç. Ort. 12° Dü 6°. 2. Gerçek Sıcaklık. 15° Dü 9°. Geç. Ort.
38	http://cekerekcpl.meb.k12.tr/ YOZGAT-ÇEKEREK (Çekerek ?ehit Özgür Ocak Çok Programl? Lise)--Okulun tarihçesi, personel ve öğrenci bilgileri, etkinlikleri, duyuruları, rehberlik bölümü ve fotoğraflar yer alıyor.
39	http://www.ademocak.com/ Adem Ocak El Yapımı Mobilya - Modoko--Adem Ocak El Yapımı Mobilya. Yüksek kalite ve işçilik ile Adem Ocak Mobilya Modoko'da.
40	http://www.oyun2d.com/ocak.asp Ocak Oyunları - Ocak Oyunları oyna - Oyun 2D--Ocak Oyunları,Ocak Oyunu oyna,Ocak oyunları oyna. Tüm Ocak oyunları ve benzer Ocak oyunu çeşitleri.
41	http://www.muneccim.gen.tr/363-1-ocak-oglak-hangi-burc-oluyor.html 1 Ocak Oğlak Hangi Burç Oluyor Muneccim – Fal, Burçlar, Rüya --1 Ocak Oğlak Hangi Burç Oluyor. Burcunuz : Oğlak. Gurubunuz : Toprak. Uğurlu gününüz : Cumartesi. Uğurlu sayınız : 8. Uğurlu taşınız : Lal. Uğurlu renkleriniz
42	http://www.emrullahocak.com/ Emrullah OCAK - Can'dan Birisi--1 Nis 2012 ... Dijital yaşam koçu Bilkom, iPad 2 almak isteyenlerin yüzüne güldü ve son kullanıcı fiyatlarında KDV'si dahil 118 Dolar indirim yapınca iPad ...
43	http://www.habibocak.com/ Habib OCAK'la Bilişim Dünyasına Giriş Yapın. - Anasayfa--Design : Habib OCAK 2012 @ Tüm Hakları Saklıdır. RSS Kaynağı Yazar Girişi Yazarlık Başvurusu. Altyapı: MyDesign Haber Sistemi.
44	http://www.icf.com.tr/tr/urunler/ankastre-ocak/camli-ocak Camlı Ocak ICF Ankastre--Ankastre Ocak. Ankastre Fırınlar · Mikrodalga · Bulaşık Makinesi ... 2332 ICF 485 HGG6044SP2 C. Isıya dayanıklı cam gazlı ocak. Otomatik ateşleme. Otomatik ...
45	http://www.istanbul.edu.tr/sosyaltesisler/index.php/tesis/restorantlarmz/nerede/capa-ocak-restorant.html Çapa Ocak Restorant--You are here Restorantlar ve Tesisler Restorantlarımız Nerede Çapa Ocak Restorant. Top. Skip to content. Telif Hakkı © 2012 Open Source Matters. Tüm Hakları ...
46	http://tr.wikipedia.org/wiki/3_Ocak 3 Ocak - Vikipedi--3 Ocak, Gregoryen takvimine göre yılın 3. günüdür. Wikimedia Commons'ta. 3 Ocak ile ilgili çoklu ortam belgeleri bulunur. Dünya'nın Güneş'e en yakın ...

Çizelge EK-C.1 Ocak Sorgu Sonuçları (Devam)

47	http://personel.saglik.gov.tr/2012-yili-ocak-donemi-kurum-ici-naklen-atama-kurasi--id3417-60.html 2012 Yılı Ocak Dönemi Kurum İçi Naklen Atama Kurası--T.C. Sağlık Bakanlığı Yönetim Hizmetleri Genel Müdürlüğü.
48	http://www.tatilbudur.com/tags/ocak-ayi-otel-fiyatlari.html Ocak Ayı Otel Fiyatları--Ocak Ayı Otel Fiyatları, Ocak Ayında Ucuz Oteller, Ocak Ayı Ucuz Oteller ve fiyatları 2011 erken rezervasyonda %50 ye varan indirimlerle.5 yıldızlı balayı otelleri ...
49	http://www.19ocak.org/ 19 OCAK • Geniş katılımlı organize bir suikast--Bu site, iki amaçla hazırlandı. İlki, Hrant Dink'in aramızdan çekilip alınmasına giden süreci ve sonrasını unutturmamak. Bu, resmî görevlilerin, kışkırtıcıların, ...

TAKVİM SORGU SONUÇLARI

Çizelge EK-D.1 Takvim Sorgu Sonuçları

0	http://www.takvim.com.tr/ Takvim - İnternetin Haber Gazetesi, Son Dakika, Astroloji, Spor ...-+Günlük olarak basılan Takvim gazetesinin web sitesi.
1	http://www.takvim.com/ Takvim.com - 2012 Takvimi - Miladi, Hicri, Rumi Takvim, Zaman ...-+Takvim, zaman, saat konularında bilgiler. ... Sitemiz özellikle takvim, saat ve zaman konusunda bilgi sunmak amacıyla kurulmuş olan bir sitedir. Sitemizde; Miladi ...
2	http://www.turktakvim.com/ 9 Eylül Pazar - Türkiye Takvimi-+-Dünya şehirleri için namaz vakitleri ve günlük takvim yaprağı.
3	http://gazete.netgazete.com/ShowPaper.aspx?news=takvim takvim-+-Cumhuriyet · Dünya · Fanatik · Fotomac · Güneş · Haberturk · Hürriyet · İstanbul Ticaret · Milli Gazete · Milliyet · Posta · Radikal · Sabah · Sözcü · Star · Takvim ...
4	http://www.takvim.gen.tr/ takvim - 2012 takvimi,2013 takvimi,yıl,yılı,gün,günü,bayram tatili ...-+Takvim konusunda detaylı bilgiler bulabileceğiniz online takvim hizmeti veren bir site.
5	http://tr.wikipedia.org/wiki/Takvim Takvim - Vikipedi-+-Takvim,zamanın yüzyıl,yıl,ay,hafta ve gün gibi parçalara bölünüp düzenli bir sırayla gösterildiği çizelgedir. Takvim, sosyal, ticari, dini ya da idari amaçla gün ...
6	http://www.haberler.com/takvim-gazetesi/ Takvim Gazetesi, Gazeteler-+-Takvim gazetesinin bugünkü birinci sayfası ve Takvim gazetesini manşetleri için tıklayın.
7	http://tr.wikipedia.org/wiki/J%C3%BClyen_takvimi Jülyen takvimi - Vikipedi-+-Jülyen takvimi, Jül Sezar tarafından M.Ö. 46 yılında kabul edilen ve batı dünyasında 16. yüzyıla kadar kullanılan takvim. Artık yıl hesaplamasındaki ufak bir fark ...
8	http://tr.wikipedia.org/wiki/Hicri_takvim Hicri takvim - Vikipedi-+-Hicri takvim, Müslüman takvimi ya da İslami takvim (Arapça: التقويم الهجري; at- taqwīm al-hijrī; Farsça: تقویم هجری قمری ... :ye qamari; Urduca-jrie he-taqwīm
9	http://twitter.com/takvim Takvim (takvim) on Twitter-+-Takvim Gazetesi resmi twitter hesabı. Son dakika haberleri, spor, magazin, ekonomi, siyaset, teknoloji, televizyon dünyasından gelişmeler anında @takvim ' de.

Çizelge EK-D.1 Takvim Sorgu Sonuçları (Devam)

10	http://pmyo.pa.edu.tr/index.php?pid=23 Polis Meslek Yüksek Okulları Resmi İnternet Sitesi - Akademik Takvim--08 Eylül 2012, Cumartesi. Giriş Yap. Başvuru Koşulları. P.M.Y.O. Hakkında. Okutulan Dersler. Akademik Takvim. İletişim Bilgileri ...
11	http://www.bilkent.edu.tr/bilkent-tr/academic/calendar/ Akademik Takvim--1 Ağustos 2012, Çarşamba, Yatay geçiş ve bölümlerarası geçiş başvuruları için son gün. 2 Ağustos 2012, Perşembe, Güz yarıyılında açılacak derslerin ...
12	http://www.nesiltakvim.com/ Nesil Takvim--Nesil 2012 Takvim,Gemici Takvimi,Takvim,Ajanda,İmsakiyeleri İnternet Sitemizden Bulabilirsiniz.
13	http://www.takvim.com/takvim_donusum.php Takvim Dönüşümleri--Hicri Takvim Hz. Muhammed'in Mekke'den Medine'ye hicretini başlangıç kabul eden ve ayın dünya çevresinde dolanımını esas alan bir takvim sistemidir.
14	http://www.anadolu.edu.tr/ogrenci_isleri/akademik_takvim.aspx Akademik Takvim - Anadolu Üniversitesi--Öğrenci İşleri > Akademik Takvim. Akademik Takvim. 2012-2013 Öğretim Yılı ... 2012-2013 Açıköğretim Kayıt ve Sınav Takvimi · 2011-2012 Öğretim Yılı - Güz ...
15	http://odb.kocaeli.edu.tr/akademik_takvim.php Akademik Takvim--2012-2013 EĞİTİM - ÖĞRETİM YILI AKADEMİK TAKVİMİ. GÜZ YARIYILI. 03-07 Eylül 2012, Yeni Öğrenci Kayıtları. 07 Eylül 2012, Seçmeli Ders Kontenjanlarının ...
16	http://www.yok.gov.tr/content/view/1142/ Yükseköğretim Kurulu - Akademik Takvim 2012--2012 YILI AKADEMİK TAKVİMİ 06 Ocak • Öğretim elemanı kadrolarında meydana gelen değişiklikler (atananlar/ayrılanlar), birimler itibarıyla unvan bazında ...
17	https://www.anadolu.edu.tr/ogrenci_isleri/akademik_takvim/ao_takvim.aspx 2011-2012 Açıköğretim Kayıt ve Sınav Takvimi - Anadolu Üniversitesi--2011-2012 ÖĞRETİM YILI AÇIKÖĞRETİM DÖNEMLİK-KREDİLİ PROGRAMLARI KAYIT, KAYIT YENİLEME TARİHLERİ. Kayıt Türü, Kayıt ve Kayıt Yenileme ...
18	http://takvim.marmara.edu.tr/ Marmara Üniversitesi Takvim Sistemi--Marmara Üniversitesi'ne bağlı bütün bölümlerin tüm akademik takvimleri yayımlanmaktadır.
19	http://www.gau.edu.tr/tr/330/akademik-takvim Akademik Takvim--Akademik Takvim. ... Akademik Takvim. 2010-2011 akademik takvim. Alt Sayfalar . Öğrenci İşleri · Kayıt Koşulları · Yatay ve Dikey Geçişler · Burslar · Eğitim Harcı ...
20	http://www.takvim.com.tr/Ekonomi/ EKONOMİ HABERLERİ - Takvim--Harcıyo'rus Rus zenginler Antalya'dan sonra İstanbul'u keşfetti. Özel uçaklarla gelip... Güle güle IMF Türkiye ile IMF'nin borç ilişkisi önümüzdeki Nisan'da sona ...
21	http://www.takvim.com.tr/Saklambac/ SAKLAMBAÇ HABERLERİ - Takvim--Aile boyu çalışan Başarılı turizmci Özgür Ardil, bu sezon Alaçatı Casa Oliva'da ailesiyle... Sevgilisine derin mesaj Şarkıcı Tuğba Özerk, Çeşme'de bir ...
22	http://www.takvim.com.tr/Spor/ SPOR HABERLERİ - Takvim--TAKVİM. Canlı Skor; Puan Durumu; Fikstür; TV Programı. Süper Lig; PTT 1.Lig; İngiltere; İspanya; Almanya; İtalya; Fransa. Tüm Canlı Maçlar ». Takımlar O Av P ...

Çizelge EK-D.1 Takvim Sorgu Sonuçları (Devam)

23	http://www.takvim.com/yillik_takvim.php Takvim.com - 2012 Takvimi - Miladi, Hicri, Rumi Takvim, Zaman ...-+Yıllık takvimleri açıp, Ctrl+P tuşlarıyla yazıcınızdan bastırabilirsiniz. 2007 Takvimi ... Takvim sayfaları aynı sitemizde olduğu gibi "pop-up" şeklinde açılacaktır. <!
24	http://www.bayar.edu.tr/ogrenci/takvim.php Akademik Takvim-+-2011-2012 Öğretim Yılı Akademik Takvimi. 2012 - 2013 ÖĞRETİM YILI AKADEMİK TAKVİMİ. Akademik Takvim · Tıp Fakültesi Akademik Takvimi ...
25	http://tr.wikipedia.org/wiki/Takvim,_saat_ve_%C3%B6l%C3%A7%C3%BClerde_de_e%C4%9Fi%C5%9Fiklik Takvim, saat ve ölçülerde değişiklik - Vikipedi-+-Takvim, saat ve ölçülerde değişiklik ; 1925-1935 arasında gerçekleştirilen yasal değişikliklerle Türkiye Cumhuriyeti'nde kullanılan takvim, saat, rakam sistemleri.
26	http://www.khas.edu.tr/egitim/akademik-takvim/akademik-takvim-2011-2012.html Akademik Takvim 2011-2012 Özet Kadir Has Üniversitesi-+-2011-2012 EĞİTİM-ÖĞRETİM YILI AKADEMİK TAKVİMİ (Lisans ve Ön Lisans Düzeyi) SONBAHAR DÖNEMİ: : 19 Eylül 2011- 28 Aralık 2011 İlk Kez Kayıt ...
27	http://ogrenci.cu.edu.tr/takvim.asp Akademik Takvim-+-2012- 2013 Akademik Takvim. Çukurova Üniversitesi 2012-2013 Eğitim Öğretim Yılı Yıl Esasına Göre Eğitim Veren Fakülteler İle Devlet Konservatuarı ...
28	http://www.takvim.com/2014_takvimi.html Takvim.com - 2014 Yılı Takvimi-+-2014 Yılı Takvimi Takvim, zaman, saat konularında bilgiler.
29	http://www.itu.edu.tr/?akademik/akademik-takvim Akademik Takvim-+-AKADEMİK TAKVİM. Güncel akademik takvimler Öğrenci İşler Daire Başkanlığı' nca duyurulmaktadır. Akademik Takvimler. İstanbul Teknik Üniversitesi ...
30	http://tr.wikipedia.org/wiki/Milad%C3%AE_takvim Miladî takvim - Vikipedi-+-Miladi takvimya da Gregoryen takvim, Jülyen takviminin yerine Papa XIII. Gregory tarafından yaptırılan takvim. Milad'ı tarih başlangıcı ve Dünya'nın Güneş ...
31	http://www.ktu.edu.tr/index.php?pid=akadtak Akademik Takvim-+-AKADEMİK TAKVİM. 2012/2013 AKADEMİK TAKVİM (Lisans ve Önlisans) 2011/ 2012 AKADEMİK TAKVİM (Lisans ve Önlisans) 2010/2011 AKADEMİK TAKVİM ...
32	http://www.gantep.edu.tr/takvim/ Akademik Takvim-+-2012 - 2013 Akademik Takvimi. Önlisans, Lisans ve Lisansüstü, Tıp Fakültesi, Diş Hekimliği Fakültesi, Naci Topçuoğlu Meslek Yüksekokulu ...
33	http://www.takvim.com.tr/Yazarlar Yazarlar - Takvim-+-hakki.yalcin@takvim.com.tr HAKKI YALÇIN · HAKKI YALÇIN Normal hayat! Beyefendi Afyon Valisi... Şehrinde 25 şehit verilmiş. Derin bir yas tutmasını ...
34	http://www.takvim.com/2013_takvimi.html Takvim.com - 2013 Yılı Takvimi-+-2013 Yılı Takvimi Takvim, zaman, saat konularında bilgiler.
35	http://ogris.aksaray.edu.tr/001.html Akademik Takvimler-+-2012-2013 Eğitim-Öğretim Yılı. Güz Dönemi · Bahar Dönemi · Uluslararası Öğrenci Program. 2012-2013 Eğitim-Öğretim Yılı - Enstitü Akademik Takvimleri ...
36	http://www.cankaya.edu.tr/akademik_takvim/index.php Akademik Takvim-+-2012-2013 ÖĞRETİM YILI AKADEMİK TAKVİMİ. GÜZ DÖNEMİ.

Çizelge EK-D.1 Takvim Sorgu Sonuçları (Devam)

37	http://www.okan.edu.tr/sayfa/lisans-2011-2012-akademik-takvim Lisans - 2011 - 2012 Akademik Takvim Okan Üniversitesi--Uluslararası standartlardaki eğitim ve öğretimi ile ülkenin ve iş yaşamının sorunlarının çözümüne bilimsel destek sağlayan, sosyal sorumluluk ve çevre bilinciyle ...
38	http://oidb.nku.edu.tr/index.php?option=com_content&view=category&id=20&layout=blog&Itemid=66 Akademik Takvim--2012 2013 Akademik Takvim, PDF · Yazdır ... 2012-2013 ÖĞRETİM YILI TIP FAKÜLTESİ AKADEMİK TAKVİMİ(pdf). (Tıp Fakültesi dışındaki akademik birimler) ...
39	http://www.trakya.edu.tr/anasayfa/?i_id=100 Akademik Takvim--5 Tem 2012 ... 2012-2013 YILI AKADEMİK TAKVİMİ · 2011-2012 YILI AKADEMİK TAKVİMİ · 2010-2011 YILI AKADEMİK TAKVİMİ · 2009-2010 YILI AKADEMİK ...
40	http://www.gop.edu.tr/akademiktakvim.aspx AKADEMİK TAKVİM--2012-2013 Eğitim-Öğretim Yılı Akademik Takvimi · 2011-2012 Eğitim-Öğretim Yılı Akademik Takvimi · 2010-2011 Eğitim-Öğretim Yılı Akademik Takvimi ...
41	http://www.neu.edu.tr/node/77 Akademik Takvim 2012/2013 Yakın Doğu Üniversitesi--GÜZ DÖNEMİ. Eylül 10-14, 2012, Ders Alma. Eylül 17, 2012, Derslerin Başlaması. Ekim 01, 2012, Ders Ekleme/Bırakma İçin Son Günü. Ekim 05, 2012, Geç ...
42	http://web.karaelmas.edu.tr/ogrenci/online-islemler/akademik-takvim/ Akademik Takvim Öğrenci İşleri Daire Başkanlığı--2012-2013 ÖĞRETİM YILI AKADEMİK TAKVİM •2012-2013 EĞİTİM ÖĞRETİM YILI (Tıp Fakültesi, Diş Hekimliği Fakültesi ve Devlet Konservatuvarı hariç.) ...
43	http://www.iku.edu.tr/TR/akademik_takvim.php Akademik Takvim--2011 - 2012 Akademik Takvimi (ÖNLİSANS ve LİSANS). Güz Yarıyılı. Eğitim - Öğretim Başlangıcı, 19 Eylül 2011. Güz Yarıyılı Bitiş Tarihi, 23 Aralık 2011 ...
44	http://www.msgsu.edu.tr/msu/pages/50.aspx MSGSÜ - Akademik Takvim-- Akademik Takvim · Öğretim Mevzuatı · BURS KOORDİNATÖRLÜĞÜ · Uluslararası İlişkiler (ULİK) · Gerekli Belgeler · Sunuşlar/Raporlar · Ortak Üniversiteler ...
45	http://www.soylutaakvim.com/ Soylu Takvim Matbaacılık Ltd. Şti. Takvim Ajanda İmsakiye Masa ...--Firmamız takvim, imsakiye, ajanda üretimi ve tüm matbaacılık faaliyetlerini sunmaktadır.
46	http://tr.wikipedia.org/wiki/Rumi_takvim Rumi takvim - Vikipedi--Rumi Takvim, Hicret'i (Miladi 622) başlangıç kabul eden güneş yılı esasına dayalı bir takvim. Dünya'nın Güneş etrafında dolanımını esas alan Şemsi Takvim ...
47	http://www.ieu.edu.tr/tr/akademik-takvim Ana Sayfa - Akademik Takvim--Güz Dönemi. 26-29 Haziran 2012. Lisansüstü Programlara Başvurular (İlk Başvurular) (Fen Bilimleri Enstitüsü). 3 Temmuz 2012. Yüksek Lisans Mülakat Sınavı ...
48	http://rize.edu.tr/alt/tr/?page_id=359 Akademik Takvim T.C. Recep Tayyip Erdoğan Üniversitesi--2012-2013 Eğitim-Öğretim Yılı Akademik Takvimi (Önlisans ve Lisans) · 2012- 2013 Eğitim-Öğretim Yılı Akademik Takvimi (Tıp Fakültesi) · 2012-2013 ...

Çizelge EK-D.1 Takvim Sorgu Sonuçları (Devam)

49	http://oidb.sdu.edu.tr/assets/uploads/sites/73/files/2012-2013-akademik-takvim-13072012.pdf 2012-2013 EĞİTİM-ÖĞRETİM YILI AKADEMİK TAKVİMİ--13 Tem 2012 ... birimlerinde Normal Öğretim ve İkinci Öğretim için 2012-2013 Eğitim-Öğretim yılında uygulanacak Akademik Takvim aşağıdaki gibidir. 03 Eylül
----	---

YÜZ SORU SONUÇLARI

Çizelge EK-E.1 Yüz Soru Sonuçları

0	http://www.mehmetauf.com/index2.html Yüz Okuma Sanatı--8 Tem 2001 ... Kitap : İnsan tanımada yüz okuma sanatı. Mehmet Auf.
1	http://tr.wikipedia.org/wiki/Y%C3%BCz_(organ) Yüz (organ)-Vikipedi--Yüz, canlıların başlarının ön bölümüne verilen isim: alın, gözler, burun, yanaklar, ağız ve çenenin oluşturduğu bütün. Surat, sima, çehre kelimeleri ile
2	http://www.youtube.com/watch?v=WVIGFA7Y3nE 4 Yuz - Salla videoclip 4yüz - YouTube--29 Haz 2007 ... 4 Yuz Salla ! yepyeni videosu . 4yuz grubu'nun ilk albumleri için çıktıkları 3. video klip Salla www.4yuz.com.
3	http://www.uzmantv.com/etiket/y%C3%BCz Yüz - UZMANTV--UZMANTV'de yüz hakkındaki sorulara uzmanların verdiği görüntülü yanıtlar.
4	http://www.uzmantv.com/konu/yuz-yogasi Yüz Yogası - UZMANTV--Türkiye'yi yüz yogası ile tanıştıran Lourdes Çabuk gençlik sırlarını UZMANTV ile paylaştı. Tüm kırışıklardan kurtulmak artık hayal değil. Yüz yogasında hangi alın ...
5	http://yuzturkedebiyatcisi.meb.gov.tr/ Yüz Türk Edebiyatçısı--Yüz Türk Edebiyatçısının hayatı ve eserleri konusunda sesli medyaların yer aldığı web sitesi. T.C. Milli Eğitim Bakanlığı Eğitim Teknolojileri Genel Müdürlüğü.
6	http://vshape.org/ Vshape Ameliyatsız Yüz Germe Kırışıklık Tedavisi Yüz ...- +-VShape, ağrısız ve cerrahi olmayan bir yüz germe (face lifting), yüz gençleştirme boyun germe, yüz gerdirmeye, göz kaldırma ve yanak kaldırma uygulamasıdır.
7	http://www.yuzgermeameliyati.org/ Estetik Yüz Germe Ameliyatı Yüz Gerdirmeye Estetiği Fiyatları Yüz ...--7 May 2012 ... 1919 yılında Paris'te bir bayan hastanın cerrah Dr. Passot'ya gidip, "Ensemde derimi gerebilirseniz yüzüm de gerilir" demesiyle yüz cildini
8	http://www.kraloyun.com/Oyun/Yuz_Guzelligim Yüz Güzelliğim Oyunu - KralOyun'da Oyna--27 Haz 2012 ... Yüz Güzelliğim Oyunu Kral Oyun 'da oyna. Yüz Güzelliğim Oyunları Türkiye'nin lider Oyun sitesinde.

Çizelge EK-E.1 Yüz Sorgu Sonuçları (Devam)

9	http://www.sinemalar.com/film/875/yuz-yuze Yüz Yüze - Face/Off - Sinemalar.com--2 gün önce ... Yüz Yüze - 27 Haziran 1997 - FBI ajanı Sean Archer, sadist bir terörist olan Castor Troy'un Los Angeles'ta sakladığı biyolojik silahın peşine ...
10	http://www.oyunskor.com/yuz_yilin_savasi.htm Yüz Yılın Savaşı oyunları - Savaş oyunu oyna--Yüz Yılın Savaşı Oyunu, Oyun Skor en güzel oyunlar Savaş Oyunları.
11	http://www.kiz-oyun.net/oyun//1149Yuz_temizleme Yüz temizleme oyunu - Temizlik Oyunları--Aslında çok güzel ve tatlı olan kız blüğ çağına girdiğinden dolayı yüzünün çeşitli yerlerinde ve alnında çeşitli lekeler ve sivilceler çıkmış, sizde öncelikle yüzünü .
12	http://www.meb.gov.tr/haberler/html_haberler/EgitimeDestek.htm Eğitime Yüzde Yüz Destek kampanyası--Eğitime Yüzde Yüz Destek kampanyası. Milli Eğitim Bakanlığı'nın "Eğitime Yüzde Yüz Destek kampanyası" Milli Eğitim Bakanlığı Şura Salonu'nda düzenlenen ...
13	http://tr.wikipedia.org/wiki/Y%C3%BCz_nakli Yüz nakli - Vikipedi--Yüz nakli, bir insanın yüzünün bir kısmının ya da tamamının tekrar yerleştirilmesine dair medikal prosedürdür. Dünyanın ilk tam yüz nakli 2010 senesinde ...
14	http://www.oyunskor.com/kumlu-yuz-maskesi.htm Kumlu yüz maskesi oyunları - Makyaj oyunu oyna--Kumlu yüz maskesi Oyunu, Oyun Skor en güzel oyunlar Makyaj Oyunları.
15	http://www.dr.com.tr/Kitap/Bin-Yuz-Bir-Insan/Aret-Vartanyan/Egitim-Basvuru/Kisisel-Gelisim/urunno=0000000365010 D&R - Kitap Aret Vartanyan Bin Yüz Bir İnsan 14.99 TL Kişisel ...--Arka Kapak Birkaç kitabın arkasını çevirip bakıyorsun ve yine kendini arıyorsun. Yaşamında yapmak isteyip de yapmadığın ya da yapamadığın ne kadar çok şey
16	http://www.yuzbakimioyunlari.com/ Yüz Bakımı Oyunları, Makyaj Oyunları--En güzel yüz bakımı oyunları ve makyaj yapma oyunları sitemizdedir.Yüzlerce yüz bakımı oyunu arasından dilediğinizi oynayın.
17	http://www.sinemalar.com/film/619/yarali-yuz Yaralı Yüz - Scarface - Sinemalar.com--1 gün önce ... Yaralı Yüz - 09 Aralık 1983 - Tony Montana isimli Kübalı suçlu, Miami'ye gelip uyuşturucu lordu Loggia'nın emrinde çalışmaya başlar.
18	http://www.kitapyurdu.com/kitap/default.asp?id=114501 kitapyurdu: kitap - Dokuz Yüz Katlı İnsan - Dr. Mustafa Merter, ,--Dokuz Yüz Katlı İnsan, Dr. Mustafa Merter, , DOKUZ YÜZ KATLI İNSAN, DR. ... O senin muazzam varlığın, belki dokuz yüz kattır; dibi, kıyısı olmayan bir denizdir.
19	http://tr.wikipedia.org/wiki/Gizli_Y%C3%BCz_(film) Gizli Yüz (film) - Vikipedi--Gizli Yüz, senaryosunu Orhan Pamuk'un yazdığı, yönetmenliğini Ömer Kavur'un yaptığı 1991 yapımı Türk filmidir. Pamuk senaryoyu Kara Kitap'taki "Karlı ...
20	http://maol.meb.gov.tr/html_files/yuzyuze_egitim.html Yüz yüze Eğitim--Yüz yüze eğitim alabileceğiniz okullar her dönem değişmektedir. Bu nedenle kayıt yaptırma esnasında yüz yüze eğitim alabileceğiniz okul ve kurumları öncelikle ...

Çizelge EK-E.1 Yüz Sorgu Sonuçları (Devam)

21	http://www.facebook.com/fceyuzifdleri Face Yüz İfadeleri Facebook--Face Yüz İfadeleri is on Facebook. To connect with Face Yüz İfadeleri, sign up for Facebook today. Sign UpLog In · Cover Photo · Like · Create a Page · Privacy ...
22	http://www.yuzbinyuz.com/ Yuz bin Yuz--YÜZ BİN YÜZ. YÜZ BİN YÜZ PAYLAŞIMI. İLETİŞİM. İNSAN KENDİNİ YALNIZCA İNSANDA TANIR. GOETHE. YÜZ BİN YÜZ. YÜZ BİN YÜZ PAYLAŞIMI. İLETİŞİM ...
23	http://tr.wikipedia.org/wiki/Y%C3%BCz Yüz - Vikipedi--Yüz - organ; Yüz - sayı; Yüz - yüzey, satıh (Suyun yüzü); Yüz - kesici araçlarda ağız (Bıçağın keskin yüzü); Yüz - kumaşın dikiş sırasında dışa getirilen gösterişli ...
24	http://www.cerezoyun.com/makyaj-oyunlari/yuz-bakimi-ve-kas-alma.html Yüz bakımı ve kaş alma oyunları makyaj oyunları oyunu oyna--Yüz bakımı ve kaş alma ; tatlı kızımız jessy bu akşam büyük bir partiye gidecek o yüzden güzelce hazırlanıp partinin en gözde kızı olması gerekiyor. ona bu ...
25	http://kiz.oyunuoyna.com/Yuz-Maskesi-Yap-4925.html Yüz Maskesi Yap Oyunu Oyna - Kız--Oyun adı: Yüz Maskesi Yap Oyunu Oyna - Kategori : Kız - Nasıl oynanır: Yüzünde çiller bulunan kıza çeşitli yüz maskeleri yapacak ve daha güzel görünmesini ...
26	http://www.bilgimanya.com/facebook-sekilleri-ifadeleri-kullanilan-yuz-ifadeleri/128 Facebook Şekilleri İfadeleri Kullanılan Yüz İfadeleri Bilgimanya.Com--Windows Live Messenger, eski adıyla MSN'den bildiğimiz ve yazışmalarımızda sürekli kullandığımız ifadeleri artık facebookda sohbet ederken.
27	http://www.facebook.com/pages/Y%C3%BCz-Kitab%C4%B1/230589994927 Yüz Kitabı Facebook--Highlights. Highlights · Posts by Page · Posts by Others. Highlights, Highlights, Posts by Page, Posts by Others. Recent Posts by Others on Yüz KitabıSee All .
28	http://tr.wikipedia.org/wiki/Y%C3%BCz_G%C3%BCn Yüz Gün - Vikipedi--Yüz Gün, Birinci Fransa İmparatorluğu imparatoru Napolyon Bonapart'ın 1814 Fontainebleau Antlaşması gereğince sürgün edildiği Elba Adasındaki sürgünden ..
29	http://www.uzmantv.com/yuz-masaji-nasil-yapilir Yüz masajı nasıl yapılır? - UZMANTV--Yüz masajı nasıl yapılır sorusunu Masaj Uzmanı İpek Çaldemir Çalışkan görüntülü yanıtıyor
30	http://tr.wikipedia.org/wiki/Bin_Dokuz_Y%C3%BCz_Seksen_D%C3%B6rt Bin Dokuz Yüz Seksen Dört - Vikipedi--Bin Dokuz Yüz Seksen Dört, George Orwell tarafından kaleme alınmış alegorik bir politik romandır. Hikayesi distopik bir dünyada geçer. Distopya romanlarının ...
31	http://www.gurkankayabasoglu.com/ Yüz Germe - Burun Estetiği Op. Dr. Gürkan Kayabaşoğlu--yüz germe,burun estetiği ameliyat bilgilerini Estetik plastik cerrah Op. Dr. Gürkan Kayabaşoğlu.
32	http://www.yckm.org/ Yüz Çiçek Açsın Kültür Merkezi--2 gün önce ... Yüz Çiçek Açsın Kültür Merkezi (YÇKM) yeni dönem etkinliklerine başlıyor. Özgün yorumuyla dinleyicilerin büyük beğenisini toplayan müzisyen ...

Çizelge EK-E.1 Yüz Sorgu Sonuçları (Devam)

33	http://www.oyunlar1.com/yeniyuz.html Yeni Yüz oyunları - Makyaj oyunu-- OYUNLAR 1 - Türkiye'nin en çok ziyaret edilen oyun sitesi : oyunlar1.com.
34	http://video.cnntrk.com/2012/saglik/3/22/3-yuz-nakli-ameliyatini-anlattilar 3. yüz nakli ameliyatını anlattılar CNN TÜRK Video--22 Mar 2012 ... Türkiye'nin 3. Yüz naklinin gerçekleştirildiği hasta Hatice Nergis'in durumu iyi. 17 Mart'ta Gazi Üniversitesi Hastanesi'nde yapılan naklin ...
35	http://www.uzmantv.com/konu/yuz-germe-ameliyati Yüz Germe Ameliyatı - UZMANTV--Yüz germe ameliyatları hakkında merak edilenler... Hangi yaşlar yüz germe ameliyatı için uygundur? Kaç yaşından önce tavsiye edilmez? Yüzümüzde az ...
36	http://www.gib.gov.tr/dilekce/01iseBaslamaBirakmaBildirimi/iseBaslamaBirakmaBildirimiOnYuz.html İşe Başlama / Bırakma Bildirimi - ÖN YÜZ--İşe Başlama / Bırakma Bildirimi - ÖN YÜZ. İl seçiniz. İSTANBUL, İZMİR, ANKARA , ADANA, ADIYAMAN, AFYON, AĞRI, AKSARAY, AMASYA, ANTALYA ...
37	http://www.eczaonline.com/kategori/Y%C3%BCz-Bak%C4%B1m-%C3%9Cr%C3%BCnleri/ Yüz Bakım Ürünleri - EczaOnline--Zayıflama, Diyet, Cinsel Sağlık, Saç Bakımı, Cilt Bakımı ve çok daha fazlası EczaOnline farkı ile kapınızda.
38	http://www.sabah.com.tr/Yasam/2012/05/15/turkiyenin-dorduncu-yuz-nakli-yapiliyor Türkiye'nin dördüncü yüz nakli yapıldı - Sabah - 15 Mayıs 2012--15 May 2012 ... Türkiye'nin dördüncü yüz nakli, Antalya'daki Akdeniz Üniversitesi ... Türkiye'de ilk yüz naklinin gerçekleşmesinde Uşak büyük rol oynamıştı.
39	http://tr.wikipedia.org/wiki/%C4%B0ki_Y%C3%BCz İki Yüz - Vikipedi--Two-Face, (Çift yüz)Amerikançizgiromanşirketi DC Comics'in Batmaniğin yarattığı hayali bir karakterdir. Karakter ilk olarak Detective Comics'in 66. sayısında ...
40	http://www.oyunskor.com/yuzbakimi.htm Yüz Bakımı oyunları - Makyaj oyunu oyna--Yüz Bakımı Oyunu, Oyun Skor en güzel oyunlar Makyaj Oyunları.
41	http://www.nivea.com.tr/Urunler/Yuz-Bakim Yüz Bakım - NIVEA--NIVEA Visage Yüz Bakım Ürünleri. ... Hakkımızda · Beiersdorf · Tarihçe · Kariyer · Sürdürülebilirlik · İletişim · E-Bülten · Ana Sayfa · Ürünler; Yüz Bakım ...
42	http://www.oyunlar1.com/kediyuzyildiz.html Kedi yüz yıldız oyunları - Macera oyunu--OYUNLAR 1 - Türkiye'nin en çok ziyaret edilen oyun sitesi : oyunlar1.com.
43	http://www.sunoyun.com/Yuz-Bakimi-Oyunu-4fab5498ab23.html Yüz Bakımı Oyunu, Oyunları Bedava Türkçe 3d Oyunlar Oyna--Yüz Bakımı: Kız Oyunları kategorimizde yer alan Yüz Bakımı oyunu oynayarak en güzel kız oyunlarından biri ile eğleneceksiniz. Yüz bakımı evde nas.

Çizelge EK-E.1 Yüz Sorgu Sonuçları (Devam)

44	http://www.uzmantv.com/konu/yuz-felci-tedavisi Yüz Felci Tedavisi - UZMANTV-+-Yüz felci tedavisi ile ilgili bilmedikleriniz UZMANTV'de! Yüz felci nasıl teşhis edilir ? Çamaşırcı kadın ya da şoför hastalığı olarak da bilinen yüz felcinin tedavisi ...
45	http://www.haberturk.com/saglik/haber/743408-yuzde-yuz-basari Turan Çolak'ın yüz nakli ameliyatından sonraki ilk fotoğrafı ...-+-18 May 2012 ... Turan Çolak'ın yüz nakli ameliyatından sonraki ilk fotoğrafı yayınlandı-En Son Nakil Haberleri.
46	http://www.oyunskor.com/yuztasarim.htm Yüz tasarım oyunları - Makyaj oyunu oyna-+-Yüz tasarım Oyunu, Oyun Skor en güzel oyunlar Makyaj Oyunları.
47	http://www.lorealparis.com.tr/cilt-bakimi/yuz-bakimi.aspx YÜZ BAKIMI-+-L'Oréal Paris'ten tüm cilt tiplerine uygun yüz bakım ürünleri hakkında herşey: tonik, nemlendirici, temizleyici, göz makyajı temizleyicisi, yüz temizleyicileri ve ...
48	http://www.sinemadafilmizle.com/yuz-numarali-adam-sinema-film-izle.html YÜZ NUMARALI ADAM FİLMİNİ İZLE sinema izle - film izle - film ...-+-Kemal sunal ,oya aydoğan, Yeşilçam, filmini, filmleri, izle, seyret.
49	http://www.sabah.com.tr/multimedya/galeri/yasam/3-yuz-nakli-ameliyatindan-fotograflar 3. yüz nakli ameliyatından fotoğraflar - Sayfa 1 - Sabah - Galeri ...-+-1 gün önce ... Gazi Üniversitesi Tıp Fakültesi'nde Türkiye'nin 3. yüz nakli Hatice Nergiz'e yapıldı . Ameliyatı gerçekleştiren doktorlar, vericinin burun kökünden ...

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Burak DURAL
Doğum Tarihi ve Yeri : 1986, İstanbul
Yabancı Dili : İngilizce
E-posta : burakdural@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Bilgisayar Mühendisliği	Trakya Üniversitesi	2009
Lise	Pendik Lisesi (Y.D.A.)		2004

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2011	OBSS	Yazılım Uzmanı
2009	Mint BT	Yazılım Uzmanı