

Bir Soru Cevaplama Sistemi: BayBilmiş

M. Fatih Amasyalı, Banu Diri
Yıldız Teknik Üniversitesi Bilgisayar Müh. Bölümü, İstanbul
{mfatih,banu}@ce.yildiz.edu.tr

Özetçe

Bilgiye erişimin en yaygın ve en kolay yollarından biri Internet, Internet'teki bilgiye erişimin en kolay yolu ise arama motorlarıdır. Ancak arama motorlarından istenilen bilgiye erişmek için önce uygun anahtar kelimeler seçilmeli, daha sonra sonuç sayfaları ya da sonuç sayfalarındaki bağlantılardaki sayfalar kullanıcı tarafından incelenmelidir. Oysa kullanıcıların belirli tipteki soruları için tüm bu işlemleri otomatik hale dönüştürmek mümkündür. İngilizce dili için bu tür birçok sistem geliştirilmiştir. Kullanıcısından sorusunu doğal dillerde kabul eden ve cevabı bulması için sayfa adresleri listesi yerine cevabın kendisini veren sistemlere soru cevaplama sistemleri denmektedir. Bu çalışmada; Türkçe için doğal dille çalışan bir soru cevaplama sistemi gerçekleştirilmiştir. Sistem öncelikle kullanıcısının doğal dille sorduğu sorusunu arama motoru sorgusuna çevirmekte ve arama motorunun sonuç sayfasından ya da bağlantılarındaki sayfalardan olası cevap cümlelerini seçmektedir. Olası cevap cümlelerini çeşitli kriterlere göre puanlandırıp en yüksek puanı alan ilk beş cümle kullanıcıya iletilmektedir. Sistem 524 adet soruyla değerlendirilmiş ve arama motorunun sonuç sayfası kullanıldığında soruların yaklaşık %43'üne, sonuç sayfasının bağlantılarındaki sayfalar kullanıldığında %60'ına ilk beş cümle içinde doğru cevap verebilmiştir.

Abstract

The easiest and common way to achieve information is Internet. The easiest way to achieve information on Internet is using a search engine. But user must do extra processes

(selection of keywords, research on result pages and web sites) while using a search engine. These extra processes can be done automatically for some kind of queries. The systems, that accepts user queries in natural language and results in only answer not a document list, are named Question-Answering Systems. The instances of these systems exist for English.

In this study, a Question-Answering system is built for Turkish. The system converts user question in natural language to search engine query and sends these queries to a search engine and receives result page. In result page and the web pages which web addresses are in result page, the answer candidate sentences are selected. The candidate answers are ranked according to four criteria. The five candidates which have the highest scores are the answers. The system is examined with 524 questions. 43% of questions are answers correctly in first five sentences when the result page of search engine was used. 60% of questions are answers correctly in first five sentences when the real pages(at the result page's links) was used.

1. Giriş

Son yıllarda Internet arama motorları ve bilgi dizinleri konularında önemli adımlar atılmış olsa da, bilginin Internet'e verimli bir şekilde yerleştirilmesi problemi hala çözülememiştir. Bu çalışmada Internet üzerindeki bilgiye, bir kişiye soru sorar gibi (doğal dille) sorular sorarak erişimi sağlayan bir sistem geliştirilmiştir.

Sistemin odaklandığı soru türleri, "Dünyadaki en yüksek ikinci dağ hangisidir?", "Uzaya giden ilk insan kimdir?" gibi gerçek sorulardır. Her ne kadar bu çeşit sorular, arama motorlarında

bulunan sınırlı sorgulama alt kümesini temsil etseler de, cevabı bulabilmek için kullanıcı tarafından harcanacak zamanın önemli derecede azaltılması olanağını içermektedirler. Çalışmanın amacı, kullanıcıyı bir dizi arama motoru sonuç sayfalarındaki sorunun cevabını potansiyel olarak içeren bir çok web sitesini okumaya zorlamak yerine, kullanıcıya cevabı içeren cümleler veren bir sistem oluşturmaktır.

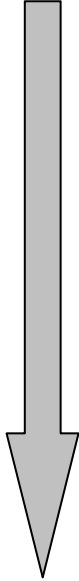
Kullanıcı sorusunun kesin bir cevabı olduğu durumlarda (ör: Plüton'u kim keşfetti? Polonya'nın başkenti neresidir?) puanlanmış bir cevap listesi, puanlanmış bir doküman adres listesinden çok daha kullanışlıdır.

Sistemin amacını gerçekleştirmek üzere kullanılan modüller, değerlendirmenin daha etkili bir biçimde değerlendirilebilmesi için ayrı ayrı geliştirilmiştir. Bu sayede her bir modülün sistemin performansına etkileri daha kolay anlaşılmıştır.

Birinci bölümün devamında konu ile ilgili diğer alanlara ait giriş niteliğinde bilgiler verilmiş, ikinci bölümde ise bu konuda daha önceden yapılmış olan çalışmalardan bir kısmı anlatılmıştır. Üçüncü bölümde, tasarlanan sistemin çalışma akışı ve modülleri ayrıntılı bir biçimde sunulmuş, dördüncü bölümde ise sistemin farklı puanlama türlerine göre sistemin performansı sunulmuş ve değerlendirilmiştir. Beşinci bölümde ise yapılan çalışmadan çıkarılan sonuçlar verilmiş ve son bölümde gelecek çalışmalar için önerilerde bulunulmuştur.

Doğal dilleri işlemek için birçok araç geliştirilmiştir. Doğal dil işleyen yazılım araçlarıyla, uygulamada nerelerde kullanıldıkları Tablo-1'de gösterilmiştir. Yukarıdan aşağıya doğru kelimelerin anlamlarına olan ihtiyaç artmaktadır.

Tablo- 1 Doğal dil işleme araçları ve uygulamaları arasındaki ilişki.
(Daelemans,2002)

NLP araçları	METİN	Uygulamalar
Morfolojik Analiz		Yazım Denetleme
Kelimelerin Türlerini Bulma		Grammer Denetleme
Birliktelikleri Bulma(Tamlama vb.)		Doküman Erişim
Sözdizimsel Analiz		Doküman Sınıflandırma
Kelimeler Arası İlişkileri (öğeleri) Bulma		Bilgi Çıkarımı
Etiketleme		Özet Çıkarımı
Şüpheli Durumların Çözümlemesi		Soru Cevaplama
Anlamsal Analiz		Diyalog Sistemleri
Zamir Çözümleme		Otomatik Tercüme
Söylem Analizi		ANLAM

1.1 Arama Motorları

Arama motorlarının bir an için yok olduğu düşünüldüğünde arama motorlarının Internet'e erişim için(dolayısıyla bilgiye erişim) kullanılan en önemli kaynak olduğu anlaşılır. Arama motorları olmadığında web sayfalarının adresleri ancak etraftaki kişilerden öğrenilebilirdi. Bu durumda Internet'teki bilginin ne kadar az bir kısmına erişilebileceği açıkça ortaya çıkar.

Arama motorlarının çoğunun içerisinde her kelime için hangi dokümanda ve kaç kere geçtiğinin tutulduğu bir veritabanı mevcuttur.

Böylesi bir yapı kullanıcının arama sorgusuyla ilgili metinlere kolayca erişmesini sağlamaktadır. Bir kelimenin bir doküman içerisinde kaç kere

geçtiği, dokümanların sorguyla olan ilgisinin derecesini de bulmakta kullanılır.

Bu yapıya bakılarak genelde bir sitenin arama motoruna kaydının aşağıdaki adımları içerdiği söylenebilir.

- Sitenin sayfasındaki kelimelerin ve kaçar defa geçtiklerinin bulunması.
- Veritabanında eğer bir kelime varsa geçtiği adreslere yeni sitenin adresinin ve frekansının eklenmesi
- Yoksa veritabanına bu kelimenin, yeni sitenin adresteki frekansının ve site adresinin eklenmesi

Daha karmaşık arama motoru veritabanlarında kelimelerin bir site içinde nerede geçtikleri de tutulur. Kullanıcı sorgusundaki kelimelerin site içinde birbirine yakınlığı, sitelerin birbiriyle karşılaştırılmasında kullanılan bir yöntemdir.

1.2 Soru Cevaplama Sistemleri

İnternet'teki bilgi patlamasının sonucu olarak doğal dille çalışan soru cevaplama sistemleri büyük bir potansiyel olarak görülmeye başlandı. Soru cevaplama sistemlerinin ilk geliştirilen örnekleri genelde, sınırlı konularda çalışmak üzere tasarlanmışlardır.

TREC-8(Text REtrieval Conference) ile birlikte, “gerçek dünya verileriyle”, “konu bağımsız soru cevaplama sistemleri” birbiriyle karşılaşma olanağı bulmuşlardır.

Günümüzde kullanılan arama motorları, aranan bilgiyi içeren ya da ilgili olabilecek dokümanların kendilerini veya yerlerini kullanıcılarına sunmaktadırlar. Kullanıcı karşısına sunulan puanlanmış bir listeden ihtiyaç duyduğu bilgiyi çıkarma işlemini kendi başına yapmak zorundadır. Bu işlem yine (göreceli olarak) büyük miktarda dokümandan bilgi çıkarımı işlemine dönüşmektedir.

Kullanıcı sorusunun kesin bir cevabı olduğu durumlarda; bu işlemde kullanıcıyı kurtarmak mümkün gözükmemektedir. Günümüzde bu işlemi(kullanıcıya sadece ihtiyacı olan bilgiyi verme) gerçekleştirecek araçlara ihtiyaç

duyulmakta olup İnternet'teki bilgi miktarı arttıkça bu ihtiyaç da artacaktır. Soru cevaplama sistemlerinin yakından ilgili olduğu diğer alanlar ve kısa açıklamaları Çizelge 1.3'te verilmiştir.

Tablo-2 Soru cevaplama sistemlerinin ilgili olduğu alanlar

Alan	Yapısı belli bir ilişkisel veritabanı üzerinde SQL yerine doğal dille sorgulanma yapılmasına imkan veren sistemlerdir.
Doğal Dille İlişkisel Veritabanları Sorulama	Kullanıcının istediği konudaki dokümanları ya da adreslerini bulan sistemlerdir. Tüm arama motorları bu tür sistemlere örnektir.
Bilgiye Erişim Sistemleri	Kullanıcının önceden tanımladığı şablonlar ile dokümanları tarayarak dolduran sistemlerdir. Şablonlar bir veri tabanı da olabilir. Örneğin firmaların web sitelerinden firma profillerini bulup daha önceden yapısı verilmiş olan bir veri tabanı tablosuna yazma işlemi bu tür bir işlemdir.
Bilgi Çıkarım Sistemleri	Sisteme tek bir metin verilir; yine bu metinle ilgili sorulara cevap verilmesi istenmektedir.
Okuduğunu Anlama	Soru-cevap ikilileri içeren veritabanlarında(ör: sıkça sorulan sorular), web sitelerinde(ör: yardım masası) kullanıcının sorduğu soruya en yakın olan sorunun bulunup oradaki cevabın kullanıcıya iletilmesi işlemidir.
Soru-Cevap İkili İçeren Veritabanları Cevap Arama	Yapısı belli bir ilişkisel veritabanı üzerinde SQL yerine doğal dille sorgulanma yapılmasına imkan veren sistemlerdir.

2. Önceki Çalışmalar

Türkçe ile yapılmış bir çalışma olmamasına karşın İngilizce dili için yapılmış birçok soru cevaplama sistemi mevcuttur.

WEBCLOPEDIA

Webclopedia (Hovy, 2001) soru cevaplama sistemi, sonuçlarının doğruluğunu(kesinliğini) geliştirmek için yazımsal ve anlamsal bilginin kullanımını sağlamaktadır. Daha önceki soru cevaplama sistemleri de bu şekilde dış

bilginin(dünya bilgisinin) kullanılması gereğini açıkça ortaya koymuştur. Örneğin, “Bant genişliği nedir? ” gibi tanım türü sorularda, sistem cevap yığnında tanımları aramaya başlamadan önce, terimsel açıklamada kullanılan kelimeleri seçmek için WordNet(Miller, 1990) kullanmaktadır.

Webclopedia soruları cevaplamak için iki temel kaynak kullanmaktadır. İlki dilbilgisi, ikincisi de dünya bilgisidir. İlki doğal dil işleme araçlarından oluşmakta; ikincisi ise 140 adet soru tipi ve bunlara ait şablonlardan, ayrıca 10,000 düğümden oluşan kavramlar hiyerarşisinden oluşmaktadır.

Soru türleri Internet’ten indirilmiş yaklaşık 17,000 adet soru-cevap ikililerinin analizinden elde edilmiştir. Soru tiplerine ait şablonlar ise yine bu soru-cevap ikililerinden otomatik olarak çıkarılmış ve belli ağırlık değerleri verilmiştir.

Şablonların otomatik olarak elde edilmesi işlemi aşağıdaki adımlardan oluşmaktadır:

Örnek soru-cevap ikilileri alınmış ((Mozart-1756)(Gandhi-1869)...) ve Altavista’ya sorgu olarak gönderilmiştir.

İlk 1000 sonuç sayfasındaki bu ikililerin beraber geçtiği cevap cümlelerinde en uzun eşleşen parçalar bulunmuştur. Örneğin ‘A B C D E F’, ‘B C D N’, ‘S L B C D’ cümleleri için ‘B C D’ ifadesi en uzun eşleşen parçadır.

Cevap cümlelerindeki en uzun eşleşen parçadaki soru kelimesi ;<NAME>, cevap kelimesi ise <Answer> ile etiketlenmiştir.

Örneğin doğum yılı soran yukarıdaki ikilliler için

1.00 <NAME> (<Answer> -)

0.87 <NAME> was born on < Answer > , in

0.87 <NAME> was born on < Answer >

0.86 <NAME> was born on < Answer > in

0.86 <NAME> was born on < Answer >

şeklinde birçok şablon elde edilmiştir. Şablonların başlarındaki ağırlık değerleri ise şablonun geçtiği doğru cevabı içeren cümle sayısının, şablonun geçtiği cevap cümlesi sayısına bölümünden elde edilmiştir.

Soru türleri ve bunlara ait şablonların tümüne

http://www.isi.edu/natural-language/projects/webclopedia/Taxonomy/taxonomy_toplevel.html adresinden erişilebilir.

ANSWERBUS

Zhiping Zheng tarafından 2002 yılında gerçekleştirilen sistem kullanıcıdan İngilizce, Almanca, Fransızca, İtalyanca ve Portekizce sorular kabul edebilmektedir. AnswerBus ilkönce kullanıcının dilini basit bir modülle bulmakta, eğer İngilizce’den farklı bir dil ise Altavista’nın BabelFish(<http://babel.altavista.com>) adlı çeviri modülünü kullanarak İngilizce’ye çevirmektedir. Daha sonra soru türünü belirlemekte ve her soru türüne göre farklı bir arama motoru seçerek ona uygun sorgular hazırlamaktadır. Örneğin günlük olaylar için YahooNews’un, Google’dan daha iyi bir tercih olacağı belirtilmiştir.

Seçilen arama motorlarına gönderilen sorguya cevap olarak geri dönen sayfalar içerisinde cevabın olması muhtemel cümleleri seçmekte ve puanlamaktadır.

Hangi soruya hangi arama motorunun daha iyi cevap vereceğini bulmak için önceden cevaplanmış 2000 TREC sorusu alınmış ve değişik arama motorlarına gönderilmiştir. Bu sayede arama motorlarının hangi sorulara doğru cevap verdiği bulunmuştur. 2000 sorudaki her kelime indekslenerek kelimelerin geçtiği soruların, kaç tanesine, hangi arama motorundan doğru cevap alındığı çıkarılır. Örneğin *nasıl* kelimesi için Google 20 doğru cevap bulurken Yahoo 12 doğru cevap bulmuşsa *nasıl* kelimesini içeren sorulara Google’ın daha iyi cevap bulduğu söylenebilir.

QUASM

Amerika’daki istatistiksel veritabanlarını sorgulamak için geliştirilmiş bir sistemdir(Pinto, 2000). Sistem www.fedstats.gov adresindeki yaklaşık 170,000 web sitesinde arama yapmaktadır. İşlemin hızlı olması için tüm web siteleri yerel bir veritabanına indirilmiştir. Veritabanında sayfalar orjinal halleriyle değil; sistemce etiketlenmiş halleriyle saklanmaktadır. Her sayfa işlenerek içinde konu bütünlüğü olduğu düşünülen parçalara ayrılmaktadır. Bu yapılırken HTML etiketlerinden yararlanılmıştır. Örneğin

kalın ya da büyük puntolarla yazılmış kelimelerin başlık belirttiği düşünülerek bir sonraki başlığa kadarki bölüm, başlıktaki konuyla ilgilidir şeklindeki kurallarla konulara ayırma işlemi gerçekleştirilmiştir. QuASM'nin çalışması diğer soru cevaplama sistemlerine benzemektedir. Kullanıcının girdiği soru incelenerek türü belirlenir. Bu işlem Naive Bayes metodu kullanılarak yapılmıştır. Soru ve türlerinden oluşan bir veritabanında kelimelerin ve kelime ikililerinin hangi tür sorularda bulunduklarının olasılıkları($P(w_i|C)$) ile çıkarılmıştır. Sorunun hangi sınıfa ait olduğu ise eşitlik 2.6'teki formülle bulunmuştur.

$$P(Q|C) = P(w_1|C) * P(w_2|C) * ... * P(w_n|C) \quad (2.6)$$

ASKMSR

Eric Brill, Susan Dumains ve Michelle Banko tarafından 2002 yılında geliştirilen AskMSR kullanıcının girdiği soruları yeniden biçimlendirerek Google'a göndermekte ve Google'ın sonuç sayfasındaki sayfa özetlerinden cevabı bulmaya çalışmaktadır (Brill, 2002). Soruların yeniden biçimlendirilme kuralları ve ağırlıkları elle elde edilmiştir. Web sitelerinin içeriğinde arama yapmak yerine Google'ın ürettiği sayfa özetinde arama yapmak sistemin hızını büyük ölçüde arttırmaktadır. Sistemde öncelikle sayfa özetlerindeki birli, ikili ve üçlü n-gramlar çıkarılmaktadır. N-gramların frekansları ve sayfayı getiren sorgunun ağırlığına bakılarak n-gramlar puanlandırılmaktadır.

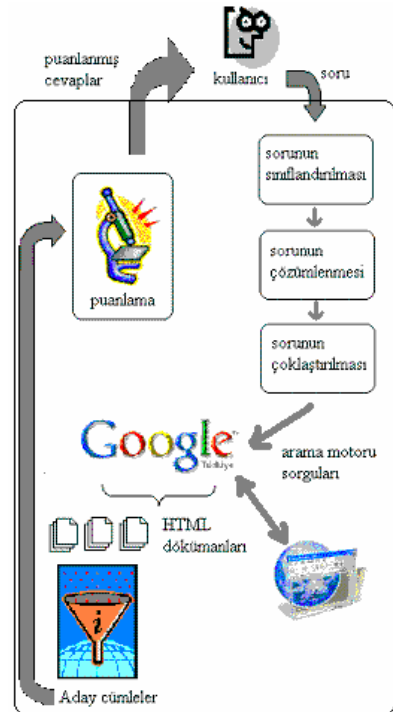
3. Tasarlanan Sistem

Türkçe için ilk olan bu çalışmada İngilizce için yapılan çalışmalar ayrıntılı olarak incelenmiş ve Türkçe içinde geçerli olabilecek kısımları belirlenmiştir. Birçok sistemin İngilizce için çok sayıda mevcut olan doğal dil işleme araçlarından faydalandığı görülmüştür. Maalesef Türkçe için bu tür araçlar henüz mevcut olmadığından gerekli olan modüller (Türkçe için mevcut olan kelime ayrıştırıcı dışındaki) sistemle birlikte geliştirilmiştir. Ancak diğer sistemlerin kullandıkları bazı modüllerin Türkçe için geliştirilmesi (örneğin en önemli araçlardan biri

olan WordNet) başlı başına bir proje olacağından dolayı gerçekleştirilmemiştir.

Geliştirilen sistem her bir modülün performans üzerindeki etkisinin görülebilmesi için istendiği zaman eklenip çıkartılabilecek bir şekilde tasarlanmıştır. Ayrıca sistemin daha sonraki gelişmelere (örneğin Türkçe için bir WordNet) ayak uydurabilmesi için yeni modüller eklenebilir bir yapıda tasarlanmıştır.

Tasarlanan sistem diğer soru cevaplama sistemleriyle benzer bir mimariye sahiptir. Sistemin çalışma akışı Şekil 1'de gösterilmiştir. Sistem kullanıcının sorduğu soruyu ilk önce sınıflandırmakta ve çözümlemekte; daha sonra çoklaştırmak arama motoru sorgularına dönüştürmektedir. Arama motorundan gelen sonuç sayfalarındaki cümleler bir filtreden geçirilerek aday cümleler belirlenmektedir. Son bölümde de aday cümlelere puan verilerek en yüksek puana sahip ilk 5 cevap kullanıcıya iletilmektedir.



Şekil-4: Sistemin çalışma akışı

3.1 Sorunun Tipinin Belirlenmesi

Kullanıcının sorduğu sorunun ne tür soru olduğunu anlamak için, içerdiği soru kelimeleri incelenmektedir. Bunun için Çizelge 3.1'deki gibi bir çizelge kullanılmış ve hangi soru kelimelerini içeriyorsa buna karşılık gelen soru türü belirlenmiştir. Sistemde 53 adet soru türü bulunmaktadır.

Tablo-2 Soru kelimelerinden soru türünün belirlenmesinde kullanılan tablodan bir kesit

sorukelime	sorutürü	silinecekkelime
yüzde kaç	oran	kaç
yıl nedir?	yıl	nedir?
yıl kaçtır?	yıl	kaçtır?
yapışları nelerdir?	bileşen	nelerdir?
tarih?	zaman	tarih?
tarih nedir?	zaman	tarih nedir?

3.2 Kullanıcı Sorusunun Çözülmesi

“Bir Türkçe Metindeki Kelimelerin Cümle İçindeki Durumlarına Bakılarak Eklerine Ayrılması” isimli çalışmada (İşler, 2000) geliştirilen ayrıştırıcı kullanılarak kelimelerin gövdeleri, ekleri belirlenmiştir. Ayrıca, ayrıştırıcıya kelimelerin etiketlenmesini sağlayan bir modül eklenmiştir. Etiketler ve örnekleri Çizelge 3.2'de verilmiştir.

Çizelge 3.2 Cümlelerdeki kelimelerin ayrıştırılması ve etiketlenmesi örnekleri

Orijinal kelime	Eklerine ayrılmış kelime	Gövde ve eklerin tanımları	Kelimenin görevi
13	13	i[i]	gagun
Aralık'ta	aralık*da	i[i]*bul	gaay
okullar	okul*lar	i[i]*ç	
açıldı	açıl*dı	f[f]*dgz	
13 Aralık'ta okullar açıldı.			
29	29	i[i]	ttgun
mayıs	mayıs	i[i]	ttay
1453'te	1453*te	i[i]*blnmyr	ttyil
okullar	okul*lar	i[i]*ç	
açıldı	açıl*dı	f[f]*dgz	
29 mayıs 1453'te okullar açıldı.			
29	29	i[i]	sayı
Km	km	i[i]	uz

yürüdüm	yürü*dü*m	f[f]*dgz*setl	
29 km yürüdüm.			
29	29	i[i]	sayı
Km'lik	km*lik	i[i]*blnmyr	uz
Yolu	yol*u	i[i]*bli	
3	3	i[i]	sayı
Saate	saat*e	i[i]*yon	zaman
Aldık	al*dı*k	f[f]*dgz*seçl	
29 km'lik yolu 3 saate aldık.			

Çizelge 3.2'de verilen örneklerden ‘13 Aralık’ta okullar açıldı.’ ifadesindeki ‘13’; ‘13 Aralık’ gün-ay tamlamasının günü belirten kısmı, ‘Aralık’; tamlamanın ayı belirten kısmıdır. Bu görevleri kelimenin görevleri kısmında ‘gagun’ ve ‘gaay’ ile ifade edilmiştir. ‘okullar’ kelimesi ise *okul* isminden sonra çoğul eki almıştır ve çözümlenmiş halinde ‘i[i]*ç’ ile gösterilmiştir. ‘açıldı’ kelimesi; *açıl* fiiline di’li geçmiş zaman eki gelmesiyle oluşmuştur. ‘29 km’lik yolu 3 saate aldık.’ ifadesindeki ‘29 km’lik’ tamlamasındaki 29 sayı; *km’lik* uzunluk(uz) olarak belirlenmiştir. Yine ‘3 saate’ ifadesindeki 3 sayı, *saate* ifadesi ise zaman olarak belirlenmiştir. ‘2500 kilodur.’ ifadesinde 2500 sayı; *kilodur* ağırlık(ag) olarak belirlenmiştir.

Çizelge 3.3'te yer, zaman, ağırlık ve uzunluk etiketlerinin belirlenmesinde kullanılan anahtar kelimeler verilmiştir.

Çizelge 3.3 Kelimelerin etiketlemesinde kullanılan anahtar kelimelere örnekler

[yer]	[zaman]	[ag]	[uz]
ili	dakika	gram	metre
barajı	saat	kilo	km
dağı	gün	kg	kilometre
gölü	hafta	kilogram	santim
şehri	ay	ton	cm
yöresi	yıl	litre	santimetre
bölgesi	asır		mil
kenti			

Kullanıcının sorduğu sorudan soru kelimeleri silinerek ve çekim eklerinden ayrılarak iki sorgu

oluşturulmuştur. Çizelge 3.4'te bu işleme örnekler verilmiştir.

Çizelge 3.4 Kullanıcı sorusunun eklerinden ve soru kelimesinden arındırılması örnekleri

Kullanıcı sorusu	Asya ile Avrupa'yı ayıran boğazın adı nedir?
1. sorgu	Asya ile Avrupa ayıran boğazın
2. sorgu	Asya ile Avrupa ayır boğaz ad
Kullanıcı sorusu	enzimlerin vücuttaki görevi nedir?
1. sorgu	enzimlerin vücuttaki görevi
2. sorgu	enzim vücut görev
Kullanıcı sorusu	Deli Dana hastalığının belirtileri nelerdir?
1. sorgu	Deli Dana hastalığının belirtileri
2. sorgu	Deli Dana hastalık belirti
Kullanıcı sorusu	dünyanın en uzun boylu insanı kimdir?
1. sorgu	dünyanın en uzun boylu insanı
2. sorgu	dünya en uzun boy insan
Kullanıcı sorusu	su hangi atomlardan oluşur?
1. sorgu	su atomlardan oluşur
2. sorgu	su atom oluş

Çizelge 3.4'te 1.sorgu sadece soru kelimesinin çıkarılmış, 2.sorgu eklerinden arındırılmış halidir.

3.3 Arama Motorundan Sonuçların Alınması

Kullanıcı sorusundan elde edilen iki sorgu, arama motoruna gönderilmiştir. Aday cevap cümlelerin elde edilmesinde, iki farklı yol kullanılmıştır. Birincisi sadece arama motorunun sonuç sayfasındaki sayfa özetlerindeki cümleleri incelemek; ikincisi ise sonuç sayfasındaki bağlantıların gösterdiği sayfaları incelemektir. Her iki yaklaşımın avantaj ve dezavantajları Çizelge 3.5'te verilmiştir.

Çizelge 3.5 Cevapların nerede aranacağına dair karşılaştırma

Cevabın nerede aranacağı	Avantajları	Dezavantajları
Sadece sayfa özetlerinde arama	Hızlı	Cümleler kesik ve anlamsız dolayısıyla daha az doğru sonuçlar (%43.9)
Sayfaların kendilerinde arama	Cümleler tam ve anlamlı ve dolayısıyla daha doğru cevaplar (%43.9 + %16.6)	Çok yavaş

Sayfaların kendilerinden cevap üretilirken sayfaların orijinal hallerinin yerine Google'ın önbelleğindeki halleri kullanılmıştır. Bunun iki önemli avantajı vardır. Birincisi sayfalardaki bilgilerin indirilmesi daha az zaman alacaktır; ikincisi ise sayfanın arama yapılırkenki (belki değiştirilmiş ya da yayından kaldırılmış) içeriği yerine Google'da indekslendiği hali kullanılarak daha doğru bilgiye erişim sağlanacaktır.

3.4 Sonuç sayfalarından aday cümlelerin seçilmesi

Sitelerdeki cümlelerin aday cümleler olabilmesi için bazı filtrelemeler yapılmıştır. Bir cümlemin cevap cümlelerinden olabilmesi için aşağıdaki şartları sağlaması gerekmektedir.

- İçinde sorgu kelimelerinin sayısının en az yarısı kadar sorgu kelimesi içermeli
- Sorgudaki özel isim olarak etiketlenmiş en son kelimeyi içermeli
- Soru türü cevapta sayısal bir değer beklenen türdense(yıl, mesafe, oran, ağırlık vb.) içinde sayı bulundurmalıdır.

Yukarıdaki her üç şarta uygun cümleler aday cevap cümleleri olarak belirlenmiş ve puanlama modülüne gönderilmiştir. Puanlama işlemi zaman alıcı bir işlem olduğundan; bu filtreleme sayesinde, cevap verme süresi kısaltılmıştır.

Çizelge 3.6'da "Atatürk hangi yıl dünyaya geldi?" sorusu için aday cümlelerin seçiminden örnekler verilmiştir.

Çizelge 3.6 “Atatürk hangi yıl dünyaya geldi?” sorusu için sonuç sayfalarından aday cümlelerin seçilmesi örnekleri

Cümle	Aday cümle mi	Sebebi
6 Eylül 1944, Pink Floyd grubunun üyelerinden Roger Waters dünyaya geldi.	Hayır	2. Şarta uymuyor
Kemal Atatürk,1881(Rumi 1296) yılında Selanik'te Koca Kasım Paşa Mahallesi Islahhane Caddesi'nde bugün müze olan üç katlı bir evde dünyaya geldi.	Evet	Tüm şartları sağlıyor
1964 yılında İstanbul'da dünyaya geldi.	Hayır	2. Şarta uymuyor
Muhtar Cem Karaca 5 Nisan 1945 de İstanbul da dünyaya geldi.	Hayır	2. Şarta uymuyor
Atatürk'ün babası Ali Rıza Efendi de bu kocacık nahiyesinde dünyaya geldi.	Hayır	3. Şarta uymuyor
Nilüfer Hatun Ortaokulu ve Taksim Atatürk Lisesi ni bitirdi.	Hayır	1. Şarta uymuyor
Atatürk'ün babası Ali Rıza Efendi'de bu nahiye de dünyaya geldi.	Hayır	3. Şarta uymuyor
1881 - Selanik' te dünyaya geldi.	Hayır	2. Şarta uymuyor

3.5 Aday cümlelerin puanlanması

Her cümle için eşleştirme, sıralama ve şablon puanları olmak üzere 3 farklı puan hesaplanmaktadır.

Eşleştirme puanının belirlenmesi: Cümlemin içerdiği sorgu kelimesi sayısı, o cümlemin eşleştirme puanı olarak belirlenir. Çizelge 3.7'de “Boğaz Köprüsü’nün yüksekliği ne kadardır?” sorusu için aday cümlelerin eşleştirme puanları verilmiştir.

Çizelge 3.7 Aday cümlelerin eşleştirme puanlarından örnekler

Cevap cümlesi	Eşleştirme puanı
Boğaz Köprüsü'nün yüksekliğinin yaklaşık 3 katına eşdeğerdir .	3
64 metre yüksekliğindeki Boğaz Köprüsü'yle, Saipem 7000 .	3
Boğaz Köprüsü'nün denizden yüksekliği ise 60 metre.	3
Mesela Boğaziçi Köprüsü'nün deniz seviyesinden yüksekliği sadece 64 .	3
boyu 165 metre olan Birinci Boğaz Köprüsü'nün .	2

Sıralama puanının hesaplanması: Cümle içindeki sorgu kelimelerinin birbirine yakınlığı, cevabın o cümlede olma olasılığını artırmaktadır. Bu yüzden sorgudaki kelimelerinin sayısına, cümledeki sorgu kelimelerinin sayısına ve yakınlığına bağlı bir formülle 0-10 arasında bir sıralama puanı hesaplanmıştır. Hesaplama formülü Eşitlik 3.1'de verilmiştir.

$$SP = \left[\frac{10 * C_k * (C_k - 1)}{Tp * S_k} \right] \quad (3.1)$$

Eşitlik 3.1'deki C_K cümlemin içerdiği sorgu kelimesi sayısını, S_K sorgudaki kelime sayısını göstermektedir. Tp ise cümledeki sorgu kelimelerinin birbirlerine kelime bazındaki uzaklıklarının toplamıdır.

Örneğin "A B C D" sorgu cümlesi için "A E F G H B C" cevap cümlesinin sıralama puanı bulunurken öncelikle cevap cümlesindeki sorgu kelimelerinin cümledeki kaçınıcı kelime oldukları bulunmuştur. Örnek için, 1 – 6 – 7 olarak bulunur. Bu dizi sıralı değilse ise sıralanmaktadır. Sonuç olarak eşitlikteki Tp ifadesi; bu dizideki her rakamın bir sonrakinden farkı bulunduktan sonra bu farkların mutlak değerleri toplanarak elde edilmiştir. Örnek için $Tp = |(1-6)| + |(6-7)| = 6$ olarak bulunur. Örneğin sıralama puanı ise $SP = [10 * 3 * (3-1)] / (6 * 4) = 2.5$ dir.

Bir başka örnek vermek gerekirse; "A B" sorgu cümlesi için "A B C" cevap cümlesinin sıralama

puanı $SP=[10*2*(2-1)]/(1*2)=10$ olarak bulunmuştur.

Çizelge 3.8'de “**Boğaz Köprüsü’nün yüksekliği ne kadardır?**” sorusu için aday cümlelerin sıralama puanları verilmiştir.

Çizelge 3.8 Aday cümlelerin sıralama puanlarından örnekler

Cevap cümlesi	Sıralama puanı
64 metre yüksekliğindeki Boğaz Köprüsü ’yle, Saipem 7000 arasında yaklaşık 8 metre mesafe kalacak.	10
Fatih Sultan Mehmet Köprüsü’nün güvenli geçiş yüksekliği 64 metre olduğu için bu köprünün altından geçişte bir sorun görülüyor.	2
Gemi, denizden 64 metre yüksekliğinde olan Boğaziçi Köprüsü’nün altından yaklaşık 10 dakikada geçişini tamamladı.	6

Şablon puanının belirlenmesi: Kullanıcının girdiği soru türüne ait şablonlar veritabanından çekildikten sonra cümlede bu şablonların olup olmadığı araştırılır. Eğer varsa o var olan şablona ait ağırlık değeri cümlelerin şablon puanı olarak belirlenir.

Örnek olarak “plastik neden elde edilir?” sorusunun türü “bileşen” olarak belirlendikten sonra; şablonlar veritabanından “bileşen” sorusuna cevap olabilecek şablonlar seçilmiştir. Çizelge 3.9’da bu şablonlar verilmiştir.

Çizelge 3.9 Bileşen sorularında aranan şablonların listesi

Şablonlar		
sorutipi	cevaptipi	ağırlık
bileşen	orj=ana%maddesi	5
bileşen	orj=elde%edil	6
bileşen	orj=hammadde	5
bileşen	orj=içerir	5
bileşen	orj=içinde	5
bileşen	orj=kullanılarak	5
bileşen	orj=yapılır	7
bileşen	tnm=ayr#orj=elde%edil	9

Şablonlar		
sorutipi	cevaptipi	ağırlık
bileşen	tnm=ayr#orj=oluş	7
bileşen	tnm=ayr#orj=üretilir	8
bileşen	tnm=ayr#orj=yapılır	9

Cevap cümlesinin içinde Çizelge 3.9’daki yapılardan biri yer alıyorsa; yer alan şablonun ağırlık değeri cümlelerin şablon puanı olarak belirlenmiştir.

Üç farklı şekilde puanlandırılan soruların aldıkları puanlar toplanarak o cümlelerin puanı elde edilmiştir.

Çizelge 3.10’da örnek aday cümleler, cevabı oldukları sorularındaki anahtar kelimeler, soru türleri ve aldıkları puanlar verilmiştir.

Çizelge 3.10 Aday cümlelerin aldıkları puanlar

Aday cümle	Sorudaki anahtar kelimeler	Soru türü	Eşleştirme	Sıralama	Şablon
yeni GSM telefonu M50’yi ilk kez tüketicilerle buluşturdu .	telefon ilk kez kullanıldı	zaman	3	5	0
"@",e-mail adresinin bir parçası olarak ilk kez 1971 yılında Ray Tomlinson tarafından kullanıldı.	telefon ilk kez kullanıldı	zaman	3	2	1
Probleme dayalı öğrenme (PDÖ) kavramı ilk kez 1976 yılında Kanada McMaster Üniversitesi’nde Howard Borrows tarafından kullanıldı.	telefon ilk kez kullanıldı	zaman	3	1	0
Amerika, 1867’de, Alaska’yı Rusya’dan 7.2 milyon \$ altına satın aldı.	Amerika Alaska aldı	zaman	3	2	8
Amerika Alaska’yı da, 1867’de Ruslar’dan satın aldı.	Amerika Alaska aldı	zaman	3	3	8

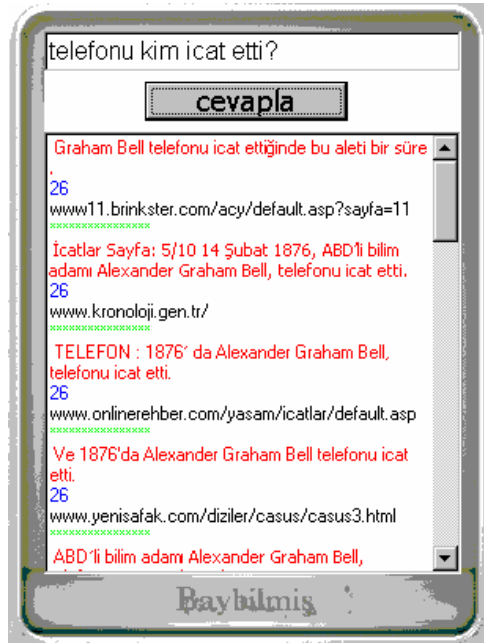
Amerika'nın Bağımsızlık Günü'nden banane; 4 Temmuz, oğlum Berent'in doğumgünü :) Geçtiğimiz günlerde, Berent, doğumgünü Paris Purple	Amerika bağımsızlık günü	zaman	2	0	5
Eylül 2001 Çarşamba Bilindiği üzere 11 Eylül 2001 Salı günü Amerika'da çok .	Amerika bağımsızlık günü	zaman	2	6	5
4'ün Today (Bugün) adlı programında Amerika'nın bağımsızlık .	Amerika bağımsızlık günü	zaman	2	6	0
Hürriyet, Amerika'daki Bağımsızlık günü kutlamalarına ilişkin 4 Temmuz tarihli haberinde .	Amerika bağımsızlık günü	zaman	2	3	5
1 Kasım 1927, TBMM'nin üçüncü dönemi açıldı.	TBMM açıldı	zaman	2	3	8
Teklif, bugün (24 Şubat 2000) koalisyon partileri genel başkanlarının düzenledikleri zirveden hemen sonra TBMM'de imzaya açıldı.	TBMM açıldı	zaman	2	5	8
Sonuçta 23 Nisan 1920 Tarihinde Ankara'da TBMM açıldı ve Türk .	TBMM açıldı	zaman	2	0	8
TBMM Genel Kurulu saat 15.00'te açıldı.	TBMM açıldı	zaman	2	2	0
25 Aralık'ta Gaziantep'in kurtuluş gününde açacağımız sergiye gelecek insanlardan böyle bir grup çıkarabileceğimizi ve gelecek sene beş .	Gaziantep kurtuluş	zaman	2	0	5
Şöyle ki, bugün, 25 Aralık 1999, Gaziantep İlimizin kurtuluş günüdür.	Gaziantep kurtuluş	zaman	2	5	0

Körfez Başkent'te vuruldu:2-0 Ligden çok kupayı düşünen Kocaelispor, Ankaragücü deplasmanında Kennedy'nin gollerine boyun eğdi Ligde UEFA Kupası .	Kennedy vuruldu	zaman	2	1	0
O zaman hoppala diyeceksiniz; Kennedy 1963'de vuruldu biliyorsunuz, bunların çıkması ise 1980'den sonra.	Kennedy vuruldu	zaman	2	5	8
Senatör Kennedy vuruldu (6 Haziran 1968) Eski ABD Başkanı JF.	Kennedy vuruldu	zaman	2	0	8
Sadri Alışık Sadri Alışık 5 Mart 1925 yılında İstanbul'da doğdu.	Sadri Alışık doğdu	doğum zaman	3	2	0
Sadri Alışık, 5 Mart 1925, 18 Mart 1995.	Sadri Alışık doğdu	doğum zaman	2	6	0
Sadri Alışık 5 Mart 1925 yılında İstanbul'da doğdu.	Sadri Alışık doğdu	doğum zaman	3	2	0
Sadri Alışık 1925 yılında İstanbul'da doğdu.	Sadri Alışık doğdu	doğum zaman	3	4	0
Kerem Alışık, 1963, İstanbul Sadri Alışık ve Çolpan İlhan'ın oğlu olan Kerem Alışık, 1963 yılında İstanbul'da doğdu.	Sadri Alışık doğdu	doğum zaman	3	1	0
Türkiye Cumhuriyeti'nin ilk nüfusu 12.4 milyon olarak kabul edildi.	Türkiye ilk nüfus sayımı yapıldı	zaman	3	4	0
Türkiye'nin ilk nüfus sayımı ise 1927 yılında yapıldı.	Türkiye ilk nüfus sayımı yapıldı	zaman	5	5	0
28 Ekim 1927, Türkiye'de birinci genel nüfus sayımı yapıldı.	Türkiye ilk nüfus sayımı yapıldı	zaman	4	4	8
28 Ekim 1927 Türkiye'de ilk nüfus sayımı yapıldı.	Türkiye ilk nüfus sayımı yapıldı	zaman	5	0	8

3.6 Cevapların kullanıcıya verilmesi

Puanlanmış cümlelerden, birbirinden farklı olan ve en yüksek puanı almış ilk 5 cümle kullanıcıya cevap olarak iletilmektedir.

Şekil 3.2'de sistemin ara yüzü ve verdiği cevap örnekleri gösterilmiştir.



şekil 4- Sistemin arabirimi ve verdiği cevap örnekleri

3.7 Şablonlar Veritabanı

Çizelge 3.12'de görüleceği gibi şablonlar veritabanında hangi soru türünün şablonu oldukları, şablonların kendileri ve ağırlık değerleri bulunmaktadır.

Sistemde şablonlar üç türlü bulunmaktadır. Birincisi kelimelerin orijinal hallerini içeren şablonlar; ikincisi kelimelerin ayrıştırılmış hallerini içeren şablonlar; üçüncü ise kelimelerin görevlerini içeren şablonlardır.

Bir sorunun cevap şablonlarında yukarıda bahsedilmiş olan üç tür şablonla yer alabilir. Çizelge 3.12'de şablonlar veritabanından bir kesit verilmiştir.

Çizelge 3.12 Şablonlar veritabanından bir kesit

Şablonlar		
sorutipi	Cevaptipi	ağırlık
eser	tnm=ö[i]#orj=eseridir	8
eser	tnm=ö[i]#orj=tarafından	8
eser	tnm=ö[i]#orj=yapıtlarından	8
eser	tnm=ö[i]#orj=yaptı	7
eser	tnm=ö[i]*iet2#orj=başyapıtı	8
eser	tnm=ö[i]*iet2#orj=eseri	8
eser	tnm=ö[i]*iet2%ö[i]*iet3	6
eser	tnm=ö[i]*tm#orj=başyapıtı	10

Çizelge 3.12'de yer alan şablonlardaki 'orj=' ifadesi kelimelerin orijinal hallerini, 'grv=' ifadesi kelimelerin görevlerini, 'tnm=' ifadesi kelimelerin gövde eklerinin tanımlarını belirtmektedir. Şablonlar veritabanındaki '#' işareti 've', '%' işareti ise 'ardından' anlamındadır.

Örnek olarak 'tnm=ö[i]*tm#orj=icadı' ifadesi cümlede 'XXX'in icadı' şeklinde bir yapıyı ifade etmektedir.

Şablonların Oluşturulma Metodu

Daha önce gerçekleştirilmiş olan birçok soru cevaplama sisteminde olduğu gibi cevap şablonları ve ağırlık değerleri elle oluşturulmuştur. Bu işlem oldukça zaman isteyen bir modüldür. Sistemin performansını arttırmak için gerçekten çok sayıda şablona ihtiyaç duyulmaktadır. Şu an mevcut olan sistem için 279 adet şablon oluşturulmuştur. Şablonların elde edilmesinde "nedir soruları" için ve "nedir haricindeki sorular" için farklı yöntemler uygulanmıştır.

Nedir sorularının şablonlarının oluşturulması

"Nedir soruları" bir kelimenin tanımını sorduğu için Internet'teki bir sözlükten yararlanılmıştır. Yaklaşık 4000 Türkçe kelimededen oluşan bir listedeki kelimeler Türk Dil Kurumu'nun Internet sitesindeki sözlüğü (<http://www.tdk.gov.tr/tdksozluk/sozara.htm>) gönderilip sonuç sayfalarındaki kelime anlamları bir dosyaya kaydedilmiştir.

Bir sözlükteki en çok tekrar edilen kelimeler Türkçe'deki tanım ifadelerini içereceğinden;

kelime anlamlarını içeren dosyadaki kelimelerden en çok tekrarlanan 50 kelime(Türkçe'de en sık tekrarlanan kelimeler çıkarıldıktan sonra) “nedir sorularının” şablonları olarak belirlenmiştir.

Nedir haricindeki soruların şablonlarının oluşturulması

Her soru türü için soru cevap çiftleri belirlenmiş ve İnternet'te soru ve cevabın birlikte geçtiği cümleler bulunmuştur. Bu cümleler incelenerek, her soru türü için sıkça kullanılan kelimeler ve yapılar elle belirlenmiştir.

Çizelge 3.13'te zaman soran sorular için belirlenen şablonlar ve bu şablonların elde edildiği örnek cümleler verilmiştir.

Çizelge 3.13 Zaman şablonlarının elde edilmesinde kullanılan cümlelerden örnekler

X (Y -	ERZURUM KONGRESİ (23 Temmuz -
Y X	23 TEMMUZ ERZURUM KONGRESİ
X Y de	Erzurum Kongresi 23 Temmuz 1919'da
Y günü ... X	4 Eylül 1919 günü toplanan Sivas Kongresi
Y de ... X	4 Eylül 1919'da toplanan Sivas Kongresi'nin
X, Y günü	Sivas Kongresi, 4 Eylül 1919 günü
Y tarihinde ... X	17 Ağustos 1999 tarihinde saat 03.02 te meydana gelen
Y X (Y)	13 Kasım. Napolyon Viyana'ya girdi (1805).
Y tarihli X	3 Kasım 1839 tarihli Gülhane Hattı
Y: X	19 Temmuz: Paris'te metro hizmete girdi.
Y yılında ... X	25 Haziran 1950 yılında çıkan Kore Savaşı

4- SİSTEMİN UYGULAMA SONUÇLARI

Soru cevaplama sistemlerinin değerlendirilmesi oldukça güç bir işlemdir. Birbirinden farklı sorularla test edilen iki sistemin birbirleriyle karşılaştırılması pek bilimsel bir yöntem değildir. Bu yüzden geliştirilen sistemin diğer sistemlerle karşılaştırılabilmesi için aynı sorularla olmasa

bile aynı tür sorularla test edilebilmesi için öncelikle TREC-9 ve TREC-10 soruları yarışmanın sitesinden toplanmış ve daha sonra Türkçe'ye çevrilmiştir. Türkçe web sitelerinde cevabı olması mümkün gözükmeyen bazı özel sorular elendikten sonra 4 kişiye bu soruları başka türlü nasıl sorabilecekleri sorulmuştur. Kişilerden gelen sorular birleştirilerek 524 adet test sorusu elde edilmiştir. Cevapları sadece Google'ın sonuç sayfasında arayan sisteme, 524 soru verilmiş ve cevapları aldıkları puanlarla birlikte bir veritabanına kaydedilmiştir.

Sistemin cevap verdiği sorularlar 5 farklı puanlama tekniğine (1. eşleme 2. eşleme+sıralama 3. eşleme+şablon 4. sıralama+şablon 5. eşleme+sıralama+şablon) göre sıralanıp doğru cevabın kaçınıcı sırada olduğu bir tabloya kaydedilmiştir. Bu tablo kullanılarak soru türlerine göre doğru cevap sayıları elde edilmiştir. Çizelge 4.1'te bu sayılar verilmiştir.

Çizelge 4.1 Soru ve doğru cevap sayılarının soru türlerine göre dağılımı

Soru türü	Soru Sayısı	Eşleme	Eşleme+Sıralama	Eşleme+Şablon	Sıralama+Şablon	Eşleme+Sıralama+Şablon
Ağırlık	1	1	1	1	1	1
Başkent	6	6	6	6	6	6
bileşen	8	2	1	1	1	2
Belirtili nesne	6	4	4	4	4	4
derece	9	4	4	3	3	3
Dil	4	2	2	2	2	2
Doğum zaman	7	3	2	3	3	4
hangi	13	5	4	5	4	4
hangisi	32	14	12	14	12	12
Hız	5	3	3	3	3	3
icat	10	6	6	6	6	6
il	19	7	9	7	9	8
kim	49	21	18	22	16	18

TREC yarışmasında sistemlerin performansları MRR adını verdikleri bir puanlama tekniğiyle ölçülmektedir. Bu puanlama tekniğinde önce her sorunun puanı hesaplanmakta daha sonra tüm soruların ortalama puanı sistemin MRR puanı olarak bulunmaktadır. Bir sorunun MRR puanı, sistemin büyükten küçüğe puanlayarak verdiği cevaplar listesinde doğru cevabın kaçınıcı sırada olduğuna bağlı olarak bulunmaktadır. Doğru cevabın sırası X olursa sorunun MRR puanı $1/X$ olarak hesaplanmaktadır.

Çizelge 4.2'de 5 farklı puanlama tekniğinin soru türlerine bağımlı olarak aldıkları MRR puanları verilmiştir. En yüksek MRR puanını 0.91 'le 6 adet başkent sorusunun 6'sına birden ilk sıralarda doğru cevap veren sıralama+şablon puanlama türü almıştır.

Çizelge 4.2 Soru türlerine göre MRR puanlarının dağılımı

Soru türü	Soru Sayısı	Eşleme	Eşleme+Sıralama	Eşleme+Şablon	Sıralama+Şablon	Eşleme+Sıralama+Şablon
ağırlık	1	0.5	0.5	0.5	0.5	0.5
başkent	6	0.86	0.83	0.87	0.91	0.83
bileşen	8	0.15	0.12	0.12	0.06	0.09
Belirtili nesne	6	0.37	0.58	0.38	0.58	0.58
Derece	9	0.38	0.38	0.25	0.33	0.33
Dil	4	0.31	0.5	0.31	0.5	0.5
Doğum zaman	7	0.35	0.17	0.42	0.31	0.45
Hangi	13	0.17	0.16	0.17	0.16	0.16
hangisi	32	0.24	0.23	0.23	0.27	0.23
Hız	5	0.36	0.46	0.5	0.6	0.6
İcat	10	0.6	0.45	0.31	0.29	0.3
İl	19	0.33	0.37	0.32	0.29	0.35
Kim	49	0.27	0.24	0.31	0.21	0.26

TREC-9 yarışmasına katılan sistemlerin ortalama MRR puanı 0.234 olduğu ve sistemin sadece Google'ın sonuç sayfasında cevapları aradığı ve yarışmaya katılan diğer sistemlerin sahip oldukları doğal dil işleme araçları kullanılmadığı

düşünüldüğünde sistemin ulaştığı 0.313 'lük MRR puanı iyi bir sonuç olarak değerlendirilebilir. Daha yüksek MRR puanı elde etmek için yapılabilecekler öneriler bölümünde belirtilmiştir. Çizelge 4.3'te TREC yarışmasının birinci ve ikincilerinin MRR puanları verilmiştir.

Çizelge 4.3 TREC yarışmasının en iyileri(MRR puanları)

	TREC-8	TREC-9	TREC-10
Birinci	0.66	0.58	0.68
ikinci	0.56	0.32	0.59

Çizelge 4.4'te sisteme sorulan sorulardan ve sistemin verdiği cevaplardan örnekler verilmiştir. Puanlama türü eşleme+sıralama+şablon'dur.

Çizelge 4.4 Sisteme sorulan sorulara ve sistemin cevaplarına örnekler

soru		Cevaplar
TBMM ne zaman açıldı?	1	İşte hemen sonra da Ankara'da 23 Nisan 1920 tarihinde TBMM açıldı.
	2	Sonuçta 23 Nisan 1920 Tarihinde Ankara'da TBMM açıldı ve Türk .
	3	23 Nisan 1920'de geniş bir katılımıla TBMM açıldı.
	4	Teklif, bugün (24 Şubat 2000) koalisyon partileri genel başkanlarının düzenledikleri zirveden hemen sonra TBMM'de imzaya açıldı.
	5	1 Kasım 1927, TBMM'nin üçüncü dönemi açıldı.
Gaziantep'in kurtuluş günü ne zamandır?	1	Şöyle ki, bugün, 25 Aralık 1999, Gaziantep İlimizin kurtuluş günüdür.
	2	25 Aralık'ta Gaziantep'in kurtuluş gününde açacağımız sergiye gelecek insanlardan böyle bir grup çıkarabileceğimizi ve gelecek sene beş .
	3	-
	4	-
	5	-

Tarkan ne zaman doğdu?	1	Tarkan Süalp 1968 yılında İstanbul'da doğdu.
	2	TARKAN GÖZÜBÜYÜK Tarkan Gözübüyük Bass Gitar & Vokal 1970 yılında Erzurum'da doğdu.
	3	Tarkan Gözübüyük: Bass, Vocal 1970 yılında Erzurum'da doğdu.
	4	-
	5	-
Türkiye İnternet ile ne zaman tanıştı?	1	12 Nisan 1993'de Ankara-Washington 64K'lık bağlantısı ile Türkiye İnternet'le tanıştı.
	2	Türkiye İnternet ile 12 nisan 1993'de tanıştı ve İnternetimiz 6.
	3	12 Nisan 1993'te Ankara-Washington arasında kiralık hatla kurulan bağlantı ile Türkiye İnternetle tanıştı.
	4	Türkiye İnternet'le 12 nisan 1993'de ODTÜ'den gerçekleştirilen Ankara-Washington bağlantısı ile tanıştı.
	5	-
Boğaz Köprüsü'nün yüksekliği ne kadardır?	1	Boğaz Köprüsü'nün yüksekliğinin yaklaşık 3 katına eşdeğerdir .
	2	64 metre yüksekliğindeki Boğaz Köprüsü'yle, Saipem 7000 .
	3	Boğaz Köprüsü'nün denizden yüksekliği ise 60 metre.
	4	Mesela Boğaziçi Köprüsü'nün deniz seviyesinden yüksekliği sadece 64 .
	5	boyu 165 metre olan Birinci Boğaz Köprüsü'nün .
Ay'ın Dünya'dan uzaklığı ne kadardır?	1	Pembe-mavi benekli bu uydunun Neptün'e uzaklığı, yaklaşık olarak Dünya-Ay arası uzaklığı kadardır (354.200 Km.
	2	Ay'ın Yer yüzeyine uzaklığı 350000 km ile 400000 km .
	3	Ay'ın Dünya'ya olan ortalama uzaklığı 384000 km dir.
	4	Yer çevresinde yaklaşık olarak her 48 saatte bir dönmekte ve bu hareketi sırasında kimi zaman Yer'e Ay uzaklığı kadar yaklaşmakta, kimi zaman da Ay .
	5	Bu uzaklık, Ay-Dünya arasındaki uzaklığın (6) katıdır.

Çizelge 4.5'da sistemin puanlama türlerinin MRR puanları ve doğru cevap oranları karşılaştırılmıştır.

Çizelge 4.5 MRR ve doğru cevap oranlarının karşılaştırılması

	MRR	Doğru cevap oranı
eşleme	0.298	42,70%
eşleme+ sıralama	0.289	40,00%
eşleme+ şablon	0.308	43,90%
sıralama+ şablon	0.301	41,60%
eşleme+sıralama +şablon	0.313	41,98%

Çizelge 4.5 incelendiğinde sistemin; sadece cevap cümlesindeki soru kelimelerinin sayısına göre hesaplanan eşleme puanı ile çalıştırıldığında MRR puanının düşük ancak buna karşın doğru cevap oranının yüksek olduğu görülmektedir. Bunun sonucu olarak kullanıcı emeğinin artması durumunda(düşük MRR) doğru cevaba erişme ihtimalinin arttığı söylenebilir.

Puanlamaya soru kelimelerinin birbirine yakınlığı katıldığında MRR ve doğru cevaplama oranının düşük olduğu görülmektedir. Bunun Türkçe'de cümlelerdeki kelime sıralamasının önemsiz olmasından kaynaklandığı düşünülmektedir.

Eşleme puanına şablon puanı katıldığında MRR puanının ve doğru cevap oranının daha yüksek değerlere ulaştığı görülmektedir. Bu şablonların sisteme katılmasının önemli olduğunun bir göstergesidir.

Sistemde mevcut olan yaklaşık 250 şablonun sayısının artırılmasının sistemin performansına yapacağı olumlu etki puanlamaya eşlemenin katıldığı her puanlama türü için sonuçların yüksek olmasından anlaşılmaktadır. İçinde eşleme puanı içermeyen puanlama türünde bile sistemin performansı yüksektir.

Sistemin performansına en büyük etkiyi eşleme ve şablon puanlarının yaptığı görülmektedir.

Sistemin performansı seçilen puanlama metotlarının hepsinin birden kullanıldığı puanlama türünde en yüksek değere ulaşmıştır. Buna dayanılarak sistemi oluşturan parçaların

birbirini tamamlayıcı nitelikte oldukları söylenebilir.

5-SONUÇ ve ÖZET

Bilgiye erişimde Internet'in, Internet'teki bilgiye erişimde ise arama motorlarının önemi tartışılmazdır. Ancak arama motorlarından sadece bilginin adresi öğrenilebilmektedir. Ayrıca bu cevaba erişmek için yine bir site içinde arama yapmamız gerekmektedir. Buna ek olarak aradığımız bilgiyi belirtmek için kendi dilimizi(doğal dilimizi) değil aradığımız şeyi ifade eden anahtar kelimeler seçilmelidir. İşte tüm bu işlemleri(anahtar kelime seçme, arama motorundan gelen sonuç sayfasını inceleme, sonuç sayfasındaki linklerdeki sayfalara gitme ve ihtiyaç duyulan bilgiyi arayıp bulma) bizim yerimize yapacak sistemlere olan ihtiyaç artmıştır. Bu işlemin "niçin", "nasıl" ve "nelerdir" soruları haricindeki sorularda (nerede, ne zaman, kim, nedir vb.) başarıyla gerçekleştirilmesi mümkün olmuştur.

İşte bu noktadan hareketle, bu çalışmada İngilizce dili için birçok örneği bulunan ancak Türkçe için bir ilk olan Internet üzerinden soru cevaplama sistemi gerçekleştirilmiştir.

Sistem kullanıcının günlük diliyle sorduğu sorusunu, önce arama motoru sorgularına dönüştürüp daha sonra Google'a göndermekte ve cevabı sonuç sayfası üzerinde ya da sonuç sayfasındaki bağlantılardaki sayfalar üzerinde aramaktadır. Cevap aranırken cevap olabilecek cümleler, sayfalardan seçilip sorunun türüne göre çeşitli yöntemlerle puan verilip en yüksek puanı alan cevap ya da cevaplar kullanıcıya iletilmektedir.

Sistem değerlendirilirken bu konuda yapılmış olan diğer sistemlerle karşılaştırılabilmesi için ortak soru bankasından (TREC) çeşitli sorular alınmış ve Türkçe'ye çevrilip sisteme sorulmuştur (524 adet soru). Sistemin (sadece sonuç sayfası üzerinde arama yapan hali) soruların %43.9'una(230 adet); sonuç sayfalarındaki adreslerde bulunan sayfalar üzerinde arama yapan hali ise %60'ına ilk 5 cevap içinde doğru cevap verdiği görülmüştür. Ayrıca sistemin MRR puanı 0.313 olarak

ölçülmüştür. TREC-9 yarışmasına katılan sistemlerin ortalama MRR puanı 0.234 olduğu ve sistemin sadece Google'ın sonuç sayfasında cevapları aradığı ve yarışmaya katılan diğer sistemlerin sahip oldukları doğal dil işleme araçları kullanılmadığı ve Türkçe Internet içeriğinin İngilizce'ye göre daha az olduğu düşünüldüğünde sistemin ulaştığı MRR puanı iyi bir sonuç olarak değerlendirilebilir.

Sistemin başarısızlık nedenleri olarak cevabın farklı kelimelerle ifade edilmiş olması, cevabın çeşitli ekler almış olması gösterilebilir. Bu nedenleri ortadan kaldırmak için yapılabilecekler son bölümde önerilmiştir.

6-GELECEKTEKİ ÇALIŞMALAR İÇİN ÖNERİLER

Soru cevaplama sistemlerinin kullanıcılarına belli soru türleri için sağladığı avantajlar düşünüldüğünde üzerinde çalışılması gereken bir alan olarak karşımıza çıkmaktadır. Bu çalışmada geliştirilen sistemin elde ettiği performansın daha da yükseltilebilmesi için sistem geliştirilirken elde edilen deneyimler ışığında şunlar söylenebilir:

Sistemde kullanılan şablonlar veritabanı daha da geliştirildiğinde sistemin performansının artacağı kuşkusuzdur.

Sisteme yeni kelime etiketlerinin (görevlerinin) eklenmesi, şablonların ve filtrelerin dolayısıyla tüm sistemin performansını arttıracaktır.

Aday cevapların puanlanmasında ve kullanıcı sorusunun arama motoru sorgularına dönüştürülmesinde, eşanlamlı kelimelerin ya da yapıların kullanılması sistemin başarısına en büyük katkıyı sağlayacaktır.

Google yerine Türkçe için tasarlanmış olan arama motorları içerik olarak daha da zenginleştiği takdirde başarıya olumlu bir katkısı olacaktır. Örneğin Google'da 'yüzük' ifadesi aratıldığında içinde 'yüzüklerin' olan sayfalar döndürülmemektedir. Bu durum Türkçe için tasarlanmış arama motorlarında söz konusu değildir. Örneğin www.arama.com'da bulunan arama motoru içinde 'yüzüklerin', 'yüzüklerdir',

‘yüzükleri’ kelimelerini içeren sayfaları da döndürmektedir. Sistemin bu tür arama motorlarında kullanılabilmesi için sadece arama sorgusundaki kelimelerin çekim eklerinden arındırılmış hallerini göndermek yeterli olacaktır.

Kaynakça

- [1] **Abney S., Collins M., ve Singhal A.**, (2000), “Answer Extraction”, www.vinartus.net/spa/99f.pdf.
- [2] **Akmajian A., Heny F.** (1975), “An Introduction to the Principle of Transformational Syntax”, MIT Press, Cambridge.
- [3] **Antword E. L.**, (1990), “PC-KIMMO: a two level processor for morphological analysis”..
- [4] **Brill E., Dumais S. ve Banko M.**, (2002), “An Analysis of the AskMSR Question-Answering System”, <http://research.microsoft.com/~brill/Pubs/EMNLP2002.pdf>.
- [6] **Buchholz S.**, “Open-Domain Question Answering on the World Wide Web”, 2002, <http://ilk.kub.nl/~antalb/abvi/week3/>.
- [7] **Charniak E.**, (1999), “A Maximum-Entropy-Inspired Parser”
- [8] **Daelemans R.**, (2002), “Learning Text Mining”
- [9] **Erkman F. ve Özil S.**(1998), “Türkçe'de Niteleme”, Simurg yayınları, İstanbul
- [10] **Erkman F.**, (2000), “Dile genel bir bakış”, Multilingual Yayınları
- [11] **Fromkin, Victoria, Rodman ve Robert** (1993), “An Introduction to Language”, 17
- [12] **Henzinger M. R., Heydon A., Mitzenmacher M., Najork M.**, (1999), “Measuring index quality using random walks on the Web”
- [13] **Hovy E., Hermjakob U. ve Ravichandran D.**, (2001), “The Use of External Knowledge in Factoid QA”, www.isi.edu/~cyl/papers/TREC10-webclopedia.pdf
- [14] **İşler S., Amasyalı M. F., Tatlı E.**, (2000), “Bir Türkçe Metindeki Kelimelerin Cümle İçindeki Durumlarına Bakılarak Eklerine Ayrılması”, Bilişim Zirvesi'01, İstanbul
- [15] **Kwok C., Etzioni O. ve Weld D. S.**, (2000), “Scaling Question Answering to the Web”, <http://www10.org/cdrom/papers/120/>
- [16] **Plamondon, L. ve Kosseim, L.**, (2002), “QUANTUM: A Function-Based Question Answering System”
- [17] **Pinto D, Branstein M., Coleman R., Croft W. B. ve King M.**, (2000), “QuASM: A System for Question Answering Using Semi-Structured Data”, <http://ciir.cs.umass.edu/pubfiles/ir-244.pdf>
- [18] **Robertson S.E. ve Walker S.**, (1998), “Okapi/Keenbow at TREC-8”, In Proceedings of The Eighth Text Retrieval Conferanse, 151-162.
- [19] **Zheng, Z.** 2002, “Developing a Web-based Question Answering System”, <http://www2002.org/CDROM/poster/203/>.