Web Madenciliği
(Web Mining)

Hazırlayan: M. Ali Akcayol
Gazi Üniversitesi
Bilgisayar Mühendisliği Bölümü

Bu dersin sunumları, "Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, Bing Liu, Springer, 2011." kitabı kullanılarak hazırlanmıştır.

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

Giriş

- Günümüzdeki arama algoritmaları kullanıcı sorgusu ile ilgililik düzeyini hesaplamak için farklı yöntemleri kullanır.
- **Bir arama motorunun en önemli bileşeni** kullanıcı sorgusuyla ilgili olduğu belirlenen sayfaların **rank değerlerinin belirlenmesidir.**
- Web arama motorları crawler yazılımları ile orijinal sunuculardaki Web sayfalarını belirli aralıklarla güncellemesi gereklidir.
- Crawl yapılan sayfalar arama işlemlerinde etkinliği artırmak için indekslenmiş bir şekilde saklanır.

3

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

4

Metin Ön İşlemleri

- Bir doküman erişim için kullanılmadan önce bazı önişlemler yapılır.
- Klasik dokümanlar için (HTML hariç) yapılan önişlemler, **stopword silme**, **stemming**, **rakamların silinmesi**, **tire işaretlerinin silinmesi**, **noktalama işaretlerinin silinmesi** ve **büyük/küçük harf ayırımının giderilmesi** işleridir.
- **Stopword silme** ile dokümanın anlamına katkı sağlamayan dilin yapısal özelliklerinden dolayı eklenen kelimeler silinir (a, about, an, by, for, who, ben, ki, de/da, için, ...).
- **Stemming** ile bir kelimenin çoğul yapılması, zaman ekleri gibi öneklerinin ve soneklerinin silinmesi yapılır ("walks", "walking" ve "walker" kelimeleri "walk" yapılır).

5

Metin Ön İşlemleri

- **Rakamların silinmesi** ile klasik IR sistemlerinde doküman içerisindeki rakamlar silinir (Web arama motorları rakamları da indeksler).
- **Tire işaretlerinin silinmesi** ile ifadeler arasındaki tutarsızlıklar giderilir ("state-of-the-art" ifadesi "state of the art" yapılır).
- **Noktalama işaretlerinin silinmesi** ile tire işaretlerinde olduğu gibi tutarsızlıklar giderilir.
- **Büyük/küçük harf ayırımının giderilmesi** ile doküman ya tümü küçük harf ya da tümü büyük harf yapılır.

6

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

7

Web Sayfası Ön İşlemleri

- **HTML dokümanında farklı text alanları vardır.** Her text alanı farklı öneme sahiptir. Bu farklı text alanlarının ayırt edilmesi gereklidir.
- **Anchor text'ler** arama motorları tarafından yüksek öneme sahiptir. Anchor text'lerin doğru tanımlanması gereklidir.
- **HTML etiketlerinin** indeks yapılmadan önce silinmesi gereklidir.
- **Temel içerik bloklarının** belirlenmesi gereklidir.

8

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

9

Web Arama

- Günümüzdeki arama algoritmaları vector space model ve term matching yöntemlerini kullanır.
- Bir arama motoru Web sayfalarını crawl yaparak başlar.
- Crawl edilen sayfalar **parse edilir**, **indekslenir** ve **saklanır.**
- Sorgu sırasında, bilgiye erişim için indeks etkin bir şekilde kullanılır.

Parsing

- **Parser** bir HTML dokümanını giriş olarak alır, indekslenmek üzere **token veya terim kümesi üretir.**
- Bir parser YACC (Yet Another Compiler Compiler) veya Flex (Fast Lexical Analyzer) gibi açık kaynak kodlu lexical analyzer kullanılarak yapılandırılabilir.

10

Web Arama

Indexing

- **Arama motorları arama etkinliğini artırmak için birden fazla indeks oluşturabilir.**
- Örneğin title ve anchor text sayfa hakkında daha doğru bilgiyi sağlayabilir. Bunlar için ayrıca indeks oluşturulabilir.
- Sayfa metinlerinin tümünü içeren full indeks te ayrıca oluşturulur.
- Arama algoritması öncelikle title ve anchor text'leri içeren indeks'te arama yapar ardından full indeks'te arama yapar.
- **Yeterli sayıda sayfa ilk aramada elde edilirse full indeks üzerinde arama yapılmaz.**

11

Web Arama

Searching and Ranking

- Girilen bir kullanıcı sorgusuna göre arama aşağıdaki adımlardan oluşur:
  - Sorgu terimleri üzerinde **önişlemler**
  - İndeks üzerinde **sorgudaki tüm terimleri bulunduran sayfaların bulunması**
  - Kullanıcya sunulmak üzere sayfaların **rank değerlerinin belirlenmesi**
- **Bir arama motorunun en önemli kısmı ranking algoritmasıdır.**
- Ticari arama motorlarında ranking algoritmalarıyla ilgili çok az bilgi bulunmaktadır.

12

Web Arama

Searching and Ranking

- Klasik IR sistemlerinde dokümanların rank değerlerinin belirlenmesi için **cosine similarity** gibi yöntemler kullanılır.
- **Cosine similarity gibi yöntemler içerik tabanlıdır ve Web aramada yeterli değildir.**
- Kullanıcı sorgusu için çok sayıda ilgili doküman elde edilebilir.
- **Örneğin, Google arama motorunda "data mining" şeklinde sorguya 41.900.000 adet sonuç bulunmaktadır.**
- Bu sayfaların kullanıcıya hangi sırada sunulduğu ve üstte hangilerinin yer alacağı önemlidir.
- **Web üzerindeki sayfaların kalitesi ve güvenilir olması ranking için çok önemlidir.**

13

Web Arama

Searching and Ranking

- Bir Web sayfasının kaliteli olup olmadığı içeriğe göre belirlenemez.
- **Hyperlink'ler sayfaların kalitesinin belirlenmesinde en önemli bileşenlerdir.**
- X sayfasında Y sayfasına link verildiğinde, **X sayfasının yazarı Y sayfasının içeriğine güvendiğini dolaylı olarak gösterir.**
- **Bir sayfa ne kadar çok link alırsa (in-links) o kadar kaliteli olduğu varsayılır.**
- PageRank algoritmasının temelinde in-links sayısı yer almaktadır.
- **Web sayfaları kendi içeriklerine ve aldıkları linklere göre değerlendirilebilir.**

14

Web Arama

Searching and Ranking

- **İçerik tabanlı (content based) değerlendirme** aşağıdaki bilgilere dayanır:
  - Occurence type
  - Count
  - Position

Occurence type

- Sorgu terimlerinin bulunduğu yerler:
  - **Title:** Sayfa başlığında yer alabilir
  - **Anchor text:** Linke ait metinde yer alabilir
  - **URL:** http://www.domain.edu/Webmining
  - **Body:** Sayfa içerisinde herhangi bir yer olabilir. (bold, font size gibi değerlerine göre değerlendirilir.)

15

Web Arama

Searching and Ranking

Count

- Her terimin sayfa içerisinde bulunduğu **tekrar sayısıdır.**

Position

- Her terimin **sayfada bulunduğu pozisyonu gösterir.**
- Birden fazla terime sahip proximity sorgularda terimlerin birbirine yakınlığını belirlemek için kullanılır.
- **İçerik tabanlı skor hesabında terimlerin bulunduğu yerlere göre ağırlıklandırma yapılır (title, URL, body, …).**

16

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

17

Meta-Arama ve Sonuçların Birleştirilmesi

- **Kullanıcıların %75'i** sorgu sonucunda gelen **ilk sayfadaki sonuçlara bakmaktadır.** Sadece %25'i ikinci sayfaya geçmektedir.
- **Kullanıcıların %70'i** doğrudan **organik sonuçları seçmektedir** ve reklam içerikli sonuçlara bakmamaktadır.
- **Şirketlerin %81'i** kendileri için **yazılan blog bilgilerini dikkate almaktadırlar.**
- Google arama motorunun sunduğu organik listedeki sonuçlardan **1.sıradakini seçen kullanıcı oranı %18**, **2.sıradakini seçen %10** ve **3.sıradakini seçen %7** oranındadır.
- Bing arama motoru organik sonuçlarından **1.sıradakini seçen kullanıcı oranı %9.7**, **ikinci sıradakini seçen %5.5** ve **3.sıradakini seçen %2.7** oranındadır.

18

Meta-Arama ve Sonuçların Birleştirilmesi

- Google, Yahoo, Live ve Ask arama motorlarında yapılan araştırmaya göre,
  - **İki** farklı arama motorunun **ilk sayfa sonuçlarındaki çakışma oranı %8.9**,
  - **Üç** farklı arama motorunun **ilk sayfa sonuçları arasında çakışma oranı %2.2**
  - **Dört** farklı arama motorunun **ilk sayfa sonuçları arasındaki çakışma oranı ise %0.6** çıkmıştır.

	Percentage % of **G-Y-L-A** Total Results
Shared by all 4 engines	0.6%
Shared by any 3 engines	2.2%
Shared by any 2 engines	8.9%
Unique to 1 engine	88.3%

19

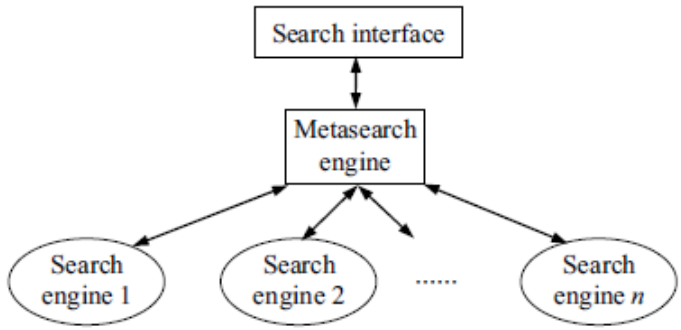
Meta-Arama ve Sonuçların Birleştirilmesi

- Kullanıcı sadece **Google** arama motorunu kullanırsa sorgusuyla ilgili Web'teki **%72.7 en iyi sonucu ilk sayfada görememektedir.**
- Kullanıcı sadece **Yahoo** arama motorunu kullanırsa **%69.2** ve sadece **Live** kullanırsa **%69.9** en iyi sonucu **ilk sayfada görememektedir.**
- Meta arama motorları, normal arama motorlarına göre Web'teki bilgilerin daha büyük bir kısmında arama yapar.
- Kullanıcıların farklı arama motorlarına sorgu girerek gelen sonuçları incelemeleri yerine tek bir sorgu ile birleştirilmiş tek bir liste halindeki sonuçları almaları sağlanır.
- Meta arama motorları kullanıcının girdiği sorguya hangi arama motorunun daha uygun olduğuna çok hızlı ve doğru bir şekilde karar verebilirler.

20

Meta-Arama ve Sonuçların Birleştirilmesi

- Meta-arama motoru, kullanıcı sorugusunu **birden fazla arama motoruna gönderir ve gelen sonuçları birleştirerek kullanır.**
- **Meta-arama motoru kendisine ait veritabanına sahip değildir.**



21

Meta-Arama ve Sonuçların Birleştirilmesi

- Arama arayüzünden alınan kullanıcı sorgusu seçilen arama motorlarına gönderilir.
- Arama motorlarından gelen sonuçlar birleştirilerek kullanıcıya sunulur.
- **Meta arama motorları Web'teki arama bölgesini artırır.**
- Web çok büyük bir bilgi kaynağıdır ve **arama motorları çok küçük bir kısmına erişebilir.**
- Eğer bir arama motoru kullanılırsa erişemediği bölgedeki ilgili dokümanlar elde edilemez.
- **Her arama motoru elde ettiği ilgili sayfalar için rank değeri hesaplar.**

22

Meta-Arama ve Sonuçların Birleştirilmesi

- **Meta aramada en önemli işlem, arama motorlarından alınan sonuçların birleştirilmesidir.**
- Farklı arama motorlarından gelen sonuçlardan aynı olanların belirlenmesi gereklidir.
- Elde edilen sonuçlardaki aynı sayfaların teke indirilmesi gereklidir.
- Meta arama motorunun elde edeceği liste için farklı arama motorlarının rank değerlerinin birleştirilmesi gereklidir.
- Elde edilen **sonuçların benzerliğine göre** veya **rank pozisyonlarına göre birleştirme işlemi yapılır:**
  - **Borda ranking**
  - **Condorcet ranking**
  - **Reciprocal ranking**

23

Meta-Arama ve Sonuçların Birleştirilmesi

Borda ranking

- Aday sayfaların oylanmasına dayanan birleştirme yapar.
- Toplam n tane aday varsa, bir oylama sonucunda **birinci sırada yer alan n puan**, **ikinci sırada yer alan n-1 puan**, ... şeklinde devam eder.
- Eğer **oylanmayan** adaylar varsa **kalan puan oylanmayan adaylar arasında eşit dağıtılır.**
- En yüksek puan alan aday kazanan adaydır.
- **Sıralama puanı yüksek olandan düşük olana doğru yapılır.**

24

Meta-Arama ve Sonuçların Birleştirilmesi

Condorcet ranking

- Her adayın **ikili olarak diğerleriyle sıralamadaki yerinin karşılaştırılmasına dayanır.**
- Bir oylamada **bir aday diğerine göre daha yüksek rank değerine sahipse kazanmış olur (win).**
- Bir oylamada bir aday diğerine göre daha düşük rank değerine sahipse kazanmış olur **(lose).**
- Bir oylamada **iki adayda oylanmadıysa birbirlerine bağlanır (tie).**

25

Meta-Arama ve Sonuçların Birleştirilmesi

Reciprocal ranking

- Her oylamada **birinci sıradaki 1 puan**, **ikinci sıradaki 1/2 puan**, **üçüncü sıradaki 1/3 puan**, ... şeklinde puan alır.
- Eğer bir aday oylanmazsa boş geçilir.
- Sonuç rank değerleri elde edilen skorların toplamı alınarak elde edilir.
- **Skor değeri yüksek olan ilk sırada yer alır.**
- Üst sıradaki adaylara Borda ranking yöntemine göre daha yüksek puan verir.

26

Meta-Arama ve Sonuçların Birleştirilmesi

Örnek

- Bir meta arama 5 arama motorundan aşağıdaki listeleri alsın.

system 1: a, b, c, d
system 2: b, a, d, c
system 3: c, b, a, d
system 4: c, b, d
system 5: c, b

Borda Ranking

$\text{Score}(a) = 4 + 3 + 2 + 1 + 1.5 = 11.5$
$\text{Score}(b) = 3 + 4 + 3 + 3 + 3 = 16$
$\text{Score}(c) = 2 + 1 + 4 + 4 + 4 = 15$
$\text{Score}(d) = 1 + 2 + 1 + 2 + 1.5 = 7.5$

- Sonuç ranking: **b, c, a, d**

27

Meta-Arama ve Sonuçların Birleştirilmesi

Örnek

Condorcet Ranking

	a	*b*	*c*	*d*
a	-	1:4:0	2:3:0	3:1:1
b	4:1:0	-	2:3:0	5:0:0
c	3:2:0	3:2:0	-	4:1:0
d	1:3:1	0:5:0	1:4:0	-

	win	*lose*	*tie*
a	1	2	0
b	2	1	0
c	3	0	0
d	0	3	0

system 1: a, b, c, d
system 2: b, a, d, c
system 3: c, b, a, d
system 4: c, b, d
system 5: c, b

- Sonuç ranking: **c, b, a, d**

28

Meta-Arama ve Sonuçların Birleştirilmesi

Örnek

Reciprocal Ranking

Score(a) = 1 + 1/2 + 1/3 = 1.83
Score(b) = 1/2 + 1 + 1/2 + 1/2 + 1/2 = 3
Score(c) = 1/3 + 1/4 + 1 + 1 + 1 = 3.55
Score(d) = 1/4 + 1/3 + 1/4 + 1/3= 1.17

- Sonuç ranking: **c, b, a, d**

29

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

30

Web Spamming

- Bir rank algoritmasının oluşturduğu Web sayfa listesi ziyaret edilme sıklığı açısından oldukça önemlidir.
- Gerçekte **kullanıcı sorgusuyla çok ilgili olduğu halde listede üst sıralarda yer almayan** Web sayfaları kullanıcı tarafından görülmeyecektir.
- Gerçekte **kullanıcı sorgusuyla çok ilgili olmayan Web sayfalarının üst sıralarda yer alması ise** kullanıcı tarafından istediği bilgiye ulaşamamasına yol açacaktır.
- **Spamming,** arama motorlarını yanıltmayı ve kullanıcı sorgusuyla çok ilgili olmadığı halde üst sıralarda yer almayı sağlayan tüm düzenlemeleri ifade eder.

31

Web Spamming

- Bir Web sayfasının içerdiği bilginin değeri artmadan arama motoru yanıltılarak rank değeri artırılırsa kullanıcıya uygun liste sunulamaz.
- **Arama motorları Web sayfalarının içeriğini anlamaz**, yapısal özelliklerini inceleyerek Web sayfasının içerdiği bilgiyi değerlendirir.
- **Arama motorlarının eksik yönlerinden faydalanarak bir Web sayfasının rank değeri artırılabilir.**
- Web sayfalarının rank değerini artırmak için ticari **firmalar (SEO – Search Engine Optimization)** bulunmaktadır.
- **SEO firmalarının** bir kısmı etik davranır ancak **bir kısmı spam oluşturur.**

32

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

33

Content Spamming

- **Çoğu arama motoru** kullanıcı sorgusuyla Web sayfalarının ilgililik düzeyini **TF-IDF tabanlı yöntemler kullanılarak hesaplar.**
- Arama motorlarının yüksek ağırlık verdiği Web sayfası bileşenlerinde term spamming yapılabilir.
- Rank değerini artırmak için **sayfaya eklenen terimler önemli kısımlara eklenebilir.**
- Bu bileşenler:
  - **Title:** Web sayfasının başlığına spam terim eklenebilir.
  - **Meta-tags:** Web sayfasının yazar adı, keywords, content language, abstract gibi bilgilerinin içerisine spam terim eklenebilir.
  - **Body:** Body içerisindeki herhangi bir yere spam terim eklenebilir.
  - **Anchor text:** Hyperlink'lerin anchor text'leri arama motorları için çok önemlidir. Spam terimler bu kısma yerleştirilebilir.
  - **URL:** Spam terimler Web sayfasına ait URL içerisine yerleştirilebilir.

34

Content Spamming

- Temel olarak iki spam yöntemi vardır: **önemli terimlerin tekrar edilmesi** ve **ilgisiz terimlerin eklenmesi.**
- Önemli **terimlerin tekrar edilmesi TF skor değerini artırır.**
- Önemli terimlerin tekrar edilmesi sonucunda dokümanın ilgilik düzeyi artırılmış olur.
- **İlgisiz terimlerin eklenmesi** ile dokümanın **çok sayıdaki sorguyla ilgili olmasına yol açar.**
- **İlgisiz terimlerin eklenmesi** yönteminde genellikle **popüler terimler Web sayfası içerisine yeleştirilir.**

35

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

36

Link Spamming

- Hyperlink'ler Web sayfalarının güvenilir olmasına yönelik değerlendirmede çok önemli rol oynar.
- **Out-link spamming** ile bazı **otorite** sayfalara (in-links çok fazla olan sayfalar) linkler verilir.
- **Out-link spamming** ile Web sayfası, **hub** sayfa (out-links çok fazla) haline getirilir.
- **In-link spamming** oluşturmak out-link spamming oluşturmaya göre **çok daha zordur.**
- **In-link spamming** oluşturmak için başka safalardan kendi sayfasına link oluşturması gereklidir.

37

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

38

Spam Gizleme Teknikleri

- Spam terimler genellikle kullanıcıların görmemesi için gizlenir.
- **Content hiding** ile eklenen terimler kullanıcının görmeyeceği şekilde renklendirilebilir. Örneğin font rengi background rengiyle aynı yapılabilir.

```
<body background = white>
   <font color = white> spam items</font>
   …
</body>
```

- Eklenen bir hyperlink'i gizlemek için çok küçük bir resim kullanılarak link oluşturulur.

```
<a href = target.html"><img src="blank.gif"> </a>
```

39

Spam Gizleme Teknikleri

- **Cloaking** yönteminde ise Web sunucu **kullanıcıya bir HTML gönderir**, **Web crawler'a farklı bir doküman gönderir.**
- Web sunucular, **Web crawler yazılımlarını ayırt etmek için arama motorlarının IP listesini bulundurur.**
- Web sunucular, **Web browser'ları ise user-agent field'i ile ayırt eder.**

```
GET /pub/WWW/TheProject.html HTTP/1.1
Host: www.w3.org
User-Agent: Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1)
```

- **Redirection** ile spam sayfa **redirect edilerek kullanıcıdan gizlenebilir.**
- **Redirection ile kullanıcı** spam sayfadan başka bir sayfaya **script kullanılarak yönlendirilir.**

40

Konular

- Giriş
- Metin Ön İşlemleri
- Web Sayfası Ön İşlemleri
- Web Arama
- Meta-Arama ve Sonuçların Birleştirilmesi
- Web Spamming
- Content Spamming
- Link Spamming
- Spam Gizleme Teknikleri
- Spam ile Mücadele Yöntemleri

41

Spam ile Mücadele Yöntemleri

- **Redirection** yapan script arama motoru crawler yazılımıyla algılanamaz.
- Arama motorunun normal browser olarak tanımlaması gereklidir.
- Her spam ile ayrı ayrı uğraşmak yerine **TrustRank yöntemi tüm spam türleriyle mücadele etmek için önerilmiştir.**
- TrustRank yöntemi güvenilir ve spam olmayan sayfaları ayırt edebilir.
- **Bir spam sayfa güvenilir sayfalara çok sayıda link verebilir.**
- Ancak, **güvenilir Web sayfası spam sayfalara link vermez.**

42

Spam ile Mücadele Yöntemleri

- Spamming ile mücadele **sınıflandırma problemidir** ve **supervised learning yöntemleri kullanılarak spam sayfa denetimi yapılabilir.**
- Supervised learning için kullanılan özellikler:
  - **Sayfadaki popüler kelime sayısı:** Spam sayfalar çok sayıda popüler kelimeye sahiptir.
  - **Ortalama kelime uzunluğu:** İngilizcede bir kelime uzunluğu ortalama olarak 5 harftir. Sentetik içeriklerde kelime uzunluğu ortalaması daha yüksektir.
  - **Sayfa başlığındaki kelime sayısı:** Sayfa başlığı yüksek ağırlıklı öneme sahip olduğundan başlıktaki kelime sayısı popüler kelimeler ile artırılır.
  - **Sayfadaki gizli içerik kısmının miktarı:** Spam sayfalar kullanıcıdan spam içeriği gizler. Bunu sonucunda daha çok gizli alana sahpitirler.

43