

Kocaeli Üniversitesi

Bilgisayar Mühendisliği Bölümü

Yazılım Laboratuvarı I

Web Scraping E-Ticaret

Bünyamin Ekşici Oğün Bingöl

170201014@kocaeli.edu.tr 170201089@kocaeli.edu.tr

Web scraping ile belirlenmiş en az 4 siteden dizüstü bilgisayar kategorisindeki ürünlerin özellikleri, başlık, resim ve daha bir çok özelliğinin çekilip bir web sitesinde gösterilmesi

Anahtar Kelimeler: web scraping, e-ticaret, kategorileştirme, filtreleme

Projenin Özeti

Yazılım laboratuvarı 1 projesi olarak bizden “Web Scraping E-Ticaret ” adındaki bir uygulama yapılması istendi. Biz projenin isterlerinden biri olan admin paneli kısmını yapamadık ve tabi ki buna neden olan en büyük faktörlerinden birisi de zamandı. O yüzden raporumuzda admin panelden bahsetmeyeceğiz.

Biz bu proje için bizden istenenler doğrultusunda programlama dillerinden JAVA, PHP, HTML, CSS dillerini database olarakta MYSQL kullanmayı uygun gördük.

Projemizin isterleri üzerine web scraping ile aldığımız verileri bir database ekleyip daha sonrasında bir sitede göstermekti. Bizde n11, trendyol, teknosa, vatan sitelerinden aldığımız verileri pcdb1 veri tabanına ekledik. Bu verileri kullanarak localde yer alan bir web sitesi tasarladık.

I. GİRİŞ

Projenin konusu Web scraping ile belirlenmiş en az 4 siteden dizüstü bilgisayar kategorisindeki ürünlerin özellikleri, başlık, resim ve daha bir çok özelliğinin çekilip bir web sitesinde gösterilmesi Biz bu proje için bizden istenenler doğrultusunda programlama dillerinden JAVA, PHP, HTML, CSS dillerini database olarakta MYSQL kullanmayı uygun gördük.

II. YÖNTEM

Bu projede kullanılan yöntem 4 aşamada anlatılacaktır.

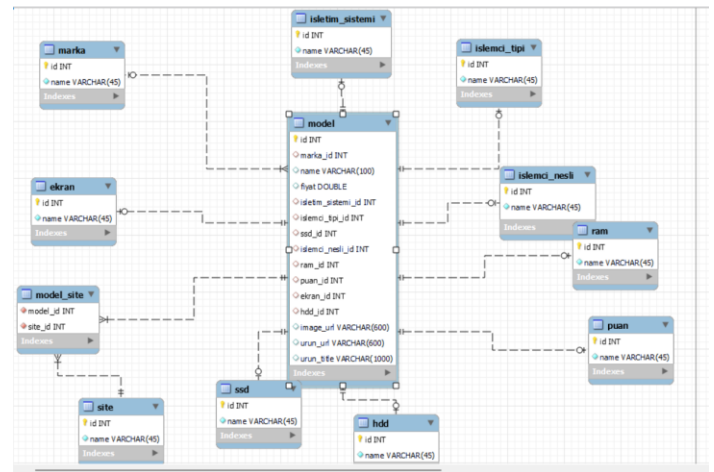
1.Aşama: Öncelikle 4 site belirlendi. Bu siteler N11, Trendyol, Vatan, Teknosa. Bu sitelerin yapısı incelenip web scraping için gerekli kütüphaneler araştırıldı. Java da iki alternatifimiz bulunmaktaydı. HtmlUnit ve JavaSoup. Kullanışlılık ve hız konusunda daha iyi olduğunu düşündüğümüz JavaSoup jar'ını kullandık. Kütüphanenin içinde barınan fonksiyonlarla

urun_title, urun_image, fiyat ve diğer özellikler için de arraylist ile bu bilgileri kayıt ettik.



Ürünün bütün özelliklerini Jsoup ile aldıktan sonra bunları bir veritabanına aktarmak gerekiyordu. Dünyada açık kaynaklı ve aynı zamanda çok kullanılan Mysql veritabanına aktarmalıydık.

2.Aşama: 1.aşamada aldığımız verileri bir veritabanına kayıt etmek için bir veri tabanı oluşturduk. Adı pcdb1 olan bu veritabanına verilerimizi kaydettik. Fakat öncelikle pcdb1 veritabanını oluşturmadan önce bir EER diyagramını oluşturduk. Çünkü bu projede isterler arasında olan normalizasyon işlemlerini daha kolay yapabiliyor ve görüntü olarakta genel şablonu görmek için kullandık. Oluşturduğumuz EER diyagram ; 1nf, 2nf, 3nf yani çokaçok(many to many) ilişkisine de uymaktadır. Oluşturulan veritabanına INSERT ettikten ve oluşan hataları düzelttikten sonra 3.aşamaya geçildi.



3.Aşama: Bu aşamada 2.aşamadaki verileri insert edip düzenledikten sonra bu verilerin gösterileceği site yapımına

başlandı. Öncelikle HTML bilgilerimizi tazeledikten sonra ana şablon hazırlandı. CSS ile bu oluşturduğumuz temel yapıyı görüntü olarak düzenlemiş olduk. Ve geriye sadece PHP ile mantıksal aynı zamanda veritabanı bağlantısını yaptık. Ve ürünleri bir tabloda göstermek için gerekli kodlar ve algoritmalar kuruldu. Son aşamaya geçildi

4.Aşama: Bu aşamada elimizde bir web sitesi ve bu web sitesinde bine yakın ürün bilgisi bulunmakta.

İsterler arasında filtreleme yapmamız gerektiği de vardı bu yüzden filtreleme yapmamız gerekti. Öncelikle veri tabanından bu özelliklerin dinamik olarak eklenmesi ve filtreleme sonucunu tabloda görmemiz gerekli. Dinamik olucak şekilde yapılan filtreleme yeni bir sekmede görüntülendi. Aynı zamanda Artan fiyat ve Azalan fiyat adında iki buton eklendi. Bu butonlarda görevini yerine getirmektedir.

Oluşturulan Dosyalar :

- anasayfa.php
- filtrenmiş.php
- urun_ozellikleri.php
- Teknosa.java
- N11.java
- Trendyol.java
- Vatan.java

Kullanılan IDE'ler: Visual Studio Code ,Eclipse

SÖZDE KOD

1-BAŞLA

2-Bağlan(n11,teknosa,vatan,trendyol)

3-Marka verisini al

4-Ssd verisini al

5-Model verisini al

6-İşlemci tipi verisini al

7- İşlemci nesli verisini al

8-İşletim sistemi verisini al

9-Ram verisini al

10- Puan verisini al

11-Site verisini al

12-Fiyat verisini al

13-Ürün resmi url'sini al

14-Ürünün url'sini al

15-Ürün başlığını al

16-Alınan verileri veritabanına yaz

17-Site oluştur(Html)

18-Site Tasarla(Css)

19-Tablo oluştur

20-Veritabanına bağlan

21-Veritabanından ürünlerin verilerini al

22-Ürünleri listele(göster)

23-Ürünlerin özelliklerini veritabanından çek

24-Dinamik filtreleme sidebar oluştur

25-Filtelenmiş ürünleri listele

26-Ürünlerin özelliklerini veritabanından çek

III. DENEYSEL SONUÇLAR



Ürün Özellikleri										
Acer Aspire 3 3.Nesil Ryzen 5 3500U-8Gb-256Gb Ssd-15.6inc-Freedom										
Marka	Model	Fiyat	İşletim Sistemi	İşlemci Tipi	İşlemci Nesli	SSD	HDD	Ram	Ekran Boyutu	Puan
ACER	Aspire 3	8.299	FreeDOS	Ryzen5	3	256	-	8	15,6	0



Ürün Listesi		
<div> <div>RAM</div> <div> <div>RAM</div> <div>8</div> <div>16</div> <div>32</div> <div>64</div> <div>128</div> <div>256</div> </div> </div> <div> <div>MARKA</div> <div> <div>ACER</div> <div>LENOVO</div> <div>CHUAWI</div> <div>ASUS</div> <div>CHOMETECH</div> <div>DELL</div> <div>HP</div> <div>MONSTER</div> <div>GAME</div> <div>SMS</div> <div>APPLE</div> <div>EVEREST</div> <div>GAOPER</div> </div> </div> <div> <div>İŞLETİM SİSTEMİ</div> <div> <div>LinuxOS</div> <div>Windows</div> </div> </div>	<div> <div>Resim</div> <div> <div>Ürün Başlığı</div> <div> <div>Acer Aspire 3 3.Nesil Ryzen 5 3500U, 8GB, 256GB, Ssd, 15.6inc, Freedos</div> <div>Lenovo Ideapad 3 11.Nesil Core i3 1115G4, 8GB, 256GB, Ssd, 15.6inc, W11</div> <div>Lenovo Ideapad 3 11.Nesil Core i3 1115G4, 8GB, 256GB, Ssd, 15.6inc, W11</div> <div>Huawei Matebook D15 11.Nesil Core i5 1135G7, 8GB, 256GB, Ssd, 15.6inc, W11</div> </div> </div> </div>	<div> <div>Site</div> <div> <div>Home</div> <div>Home</div> <div>Home</div> <div>Home</div> </div> </div>

Filtrelenmiş Ürün Listesi		
Resim	Ürün Başlığı	Site
	Acer Aspire 3 3.Nesil Ryzen 5 3500U, 8GB, 256Gb, Ssd, 15.6inc, Freedos	Home
	Acer TravelMate P2 10.Nesil Core i5 10210U, 8GB, 512Gb, 15.6inc, Mx230 20p, Freedos	Home
	Acer Extensa 15 EX215-54 NX EGUEV 005 0, 1135G7 8 GB 512 GB SSD 15.6" Free Dos FHD Display Bıyıkayır	Home
	Acer Gaming Aspire 7 A715-42G N61CPSEY 006 RS 5500U 8 GB 256 GB SSD RTX3050 144Hz 15.6" Dos FHD Display Bıyıkayır	Home

V. SONUÇ

Bu proje sayesinde web scraping ve site oluşturma konusunda kendime değer kattım özellikle PHP konusunda kendimi iyileştirdiğimi düşünüyorum. Aynı zamanda bir kütüphane nasıl indirilir, kurulur ve fonksiyonları nasıl araştırılır öğrendim. Son olarak bir web sitesinin yapımının hangi aşamalardan geçtiğini ve ne kadar emek harcandığını gördüm. Teşekkür ederim.

KAYNAKLAR

- 1-Intro to Java Programming, Comprehensive Version (10th Edition) Y. Daniel Liang
- 2- [PHP Tutorial \(w3schools.com\)](https://www.w3schools.com/)
- 3- Veritabanı Mantığı Kerem Köseoğlu
- 4- Udemy -Komple Uygulamalı Web Geliştirme Eğitimi "Sadık Turan
- 5-Google images –
Vatan.png,Teknosa.png,N11.png,Trendyol.png