

Kocaeli Üniversitesi Bilgisayar Mühendisliği

Bölümü Yazılım Laboratuvarı I - Proje I

İbrahim İnce
Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü
2. Öğretim
200202102

I. ÖZET

Trendyol, Hepsiburada, Vatan Bilgisayar ve Teknosa e-ticaret sitelerinden istenilen ürüne ait bilgilerin yer aldığı bir veritabanı oluşturulacak ve bu veritabanından istenilen bilgileri web sayfası üzerinde gösterilecektir.

Veri çekilecek olan siteler:

- hepsiburada.com,
- teknosa.com,
- trendyol.com,
- vatanbilgisayar.com’idir.

II. GİRİŞ

Proje Python programlama dili ile hazırlanmıştır. Projede BeautifulSoup ve request kütüphanelerinden yararlanılmıştır.

E-ticaret sitelerinden çekilen veriler öncelikle listelerde tutulup düzenlenmiş, ardından MongoDB Atlas veri tabanına aktarılmıştır.

III. YÖNTEM

Öncelikle veri çekilecek olan siteler incelenmiştir. İstenilen ürünler "Notebooks" cihazları olduğu için sitelerin notebooks sayfalarına ulaşılmıştır. Ardından ilk sayfanın html kodları incelenmiştir. Diğer ürünleri görmek için diğer sayfalara geçildiğinde url yapısının değişimi gözlenmiştir.

A. Sitelerden "html" Bilgilerinin Çekilmesi

"request" kütüphanesi kullanılarak sitelerin html bilgileri çekilmiştir. BS4 kütüphanesinden BeautifulSoup modülü kullanılarak çekilen html bilgileri parçalanıp düzenlenmiştir.

Bu düzenleme

- Ürünün linki şeklindedir.

Erişilen sayfadaki her bir ürünün detay linki alındıktan sonra diğer sayfalardaki ürünlere erişebilmek için o sayfaların urlleri ziyaret edilip diğer ürünlerin linkleri de toplanmıştır.

B. Ürün Sayfalarından Bilgilerin Toplanması

Her bir ürün sayfasından tek tek html bilgileri çekilmiştir. Çekilen html bilgisinden şu veriler elde edilmiştir:

- Ürün fiyatı,
- ürün puanı,
- İşletim Sistemi bilgisi,
- İşlemci Tipi bilgisi,
- İşlemci Nesli bilgisi,
- Ram miktarı bilgisi,
- SSD Boyutu bilgisi,
- Ekran boyutu bilgisi bilgisi,
- Ürünün fotoğraf urlsi.

Bu elde edilen bilgiler "tempDict" adındaki geçiçi sözlük yapısına aktarılmıştır. Ardından bu sözlük doğrudan veritabanındaki ilgili sitenin tablosuna aktarılır.

Veritabanı olarak MongoDB Atlas kullanılmıştır. MongoDB bir noSQL veritabanıdır. MongoDB Atlas ise "bulut" tabanlı bir sistemdir. Yani veritabanındaki bilgiler kendi cihazımda değil, MongoDB sunucularında saklanmaktadır.

C. Her Bir Siteden Çekilen Ürün Bilgilerinin Depolanması

Her bir siteden çekilen ürünlerin bilgileri, ziyaret edilen sitenin adı verilen tablolarda saklanmaktadır.

D. Ürünlerin Karşılaştırılması

Son olarak sitelerden çekiliğ veritabanındaki tablolara aktarılan ürünler tablolardan çekilip birbiriyle karşılaştırılır. Aynı olan ürünler "Final" adındaki tabloya aktarılır.

Aynı ürünleri bulmak için gerekli olan bilgiler, yukarıda belirtilen e-ticaret sitelerinde açıkça belirtilmemektedir. Bu sebeple verimli bir karşılaştırma

yapılabilmesi için ürün markası ve ürün özellikleri eşleşiyor ise bu iki ürün aynı cihaz kabul edilir ve fiyat karşılaştırılması buna göre yapılır.

IV. DENEYSEL SONUÇLAR

Django frameworkü ile Python tabanlı web site tasarımı sırasında; veritabanından "pymongo" kütüphanesindeki metodlar ile çekilen sözlük yapılarının içeriklerine ulaşılması sırasında sıkıntı yaşanıldı. Django'da Python daki syntaxtan farklı bir syntax kullanılması gerekti. Bu nedenle "Final" tablosundaki verilerin "key"leri yeniden düzenlendi.

V. SONUÇ

Python dili, Django Frameworkü, Html, CSS ve Bootstrap kullanılarak bir web sitesi

oluřturuldu. Bu web sitesine eřitli sitelerden ekilen rnlerden aynı olanlar yerleřtirildi. Her bir rnn her bir sitedeki fiyatları belirtilerek fiyat karřılařtırılması yapılmıř oldu.

KAYNAKA

- [1] w3schools.com. Eriřim Tarihi: 21.10.2022, https://www.w3schools.com/python/python_mongodb_query.asp/diagrams.net. Akıřřeması 27.02.2022, <https://app.diagrams.net/>

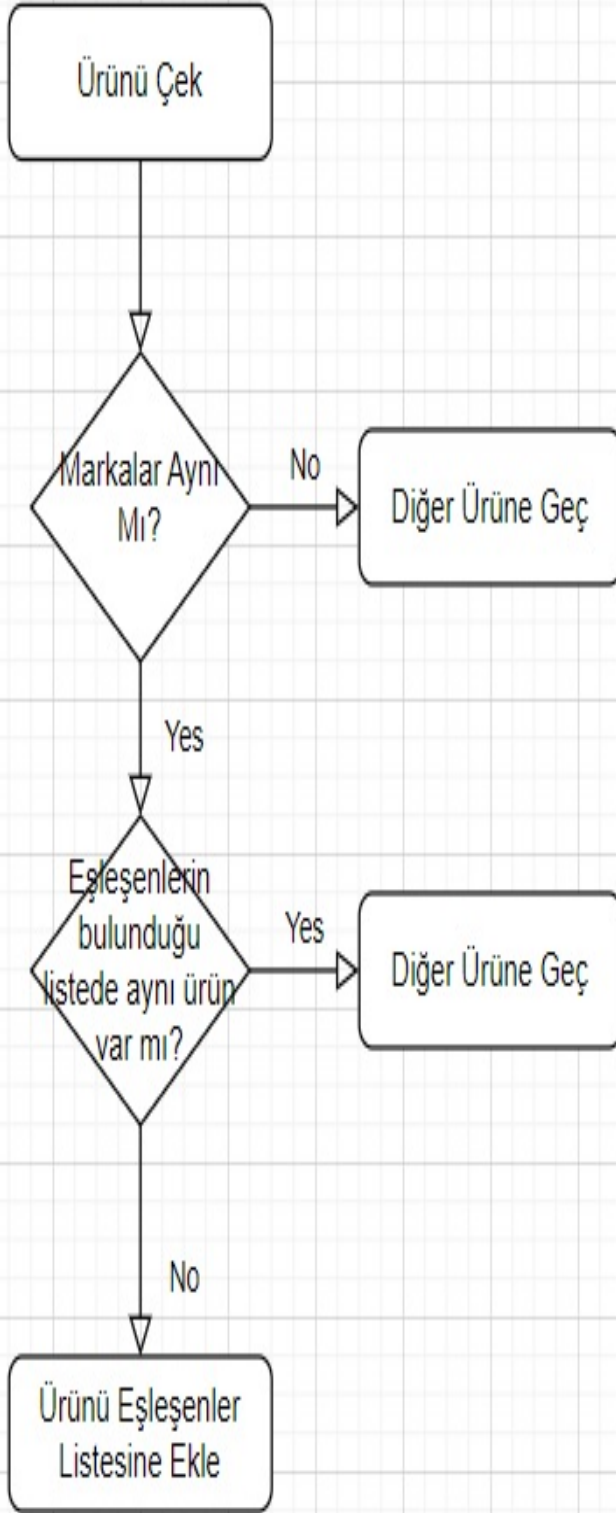


Fig. 1. rn Karřılařtırma Akıř řeması