

Web Tabanlı Veri Toplama ve Fiyat Karşılaştırma Aracı

DFA Teknoloji Staj Teknik Görev Raporu

1. Proje Tanımı

Bu proje kapsamında, belirli çevrimiçi alışveriş platformlarından (Trendyol, Hepsiburada, N11) ürün bilgilerini çekerek, aynı ürünün farklı satıcılar tarafından listelendiği durumlarda satıcı ismi, satıcı puanı, ürün ismi, ürün paunu ve ürün fiyatı gibi kriterlere göre karşılaştırmalı bir analiz gerçekleştirilmiştir.

Geliştirilen uygulama Python diliyle yazılmıştır ve web kazıma işlemleri için Selenium ve BeautifulSoup kütüphaneleri, görsel arayüz için Tkinter kütüphanesi kullanılmıştır.

2. Kullanılan Teknolojiler

- Python 3.10+
- Selenium
- BeautifulSoup4
- Tkinter (GUI)
- ChromeDriver (WebDriver)
- CSV (veri kaydı için)

3. Uygulama Genel Yapısı

Uygulama çalıştırıldığında kullanıcıdan:

1. Alışveriş sitesini seçmesi (Trendyol, Hepsiburada, N11)
2. Ürünün tam adını girmesi

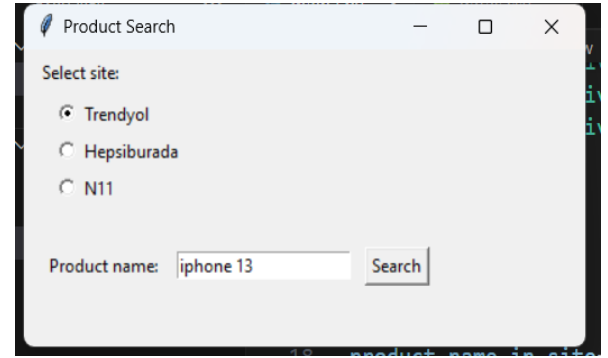
beklenir. Ardından, seçilen sitede arama yapılır ve bulunan ilk ürün sayfası açılır, mevcut

satıcı diğer satıcılar listesinde yoksa ilk mevcut satıcı bilgileri alınır ve csv ye kaydedilir daha sonra diğer satıcılar etiketine tıklanarak diğer satıcılara ulaşılır ve onların bilgileri oradan alınır ve csv ye yazılır. Bir ürünün değerlendirme puanı diğer satıcılar sayfasındaki tüm satıcılarda aynı olduğu için mevcut satıcı sayfasından bir kere çekilir ve diğer satıcılar içinde kullanılır. Tüm veri çekme işlemi bittikten sonra sonuç penceresi açılır ve csv deki veriler okunarak perncereye yansıtılır. Pencereden seçilen en ucuz/ en pahalı 5 ürüne göre tekrar sıralama yapılabilir.

4. Kod Akışı ve Fonksiyonlar

Ana Arayüz

→**create_interface()** fonksiyonunda program kullanıcıdan istediği alışveriş sitesini seçmesini ve aramak istediği ürünün ismini tam şekilde girmesini ister. Bu değerler girildikten sonra kullanıcı “search” butonuna basarak program başlatır.



Web Kazıma Fonksiyonları

→Programda her site için search fonksiyonları bazı yerlerde farklılık göstermektedir. Öncelikle genel search fonksiyonunu açıklayalım.

- Kullanıcı search butonuna basınca chrome webdriver programa bağlanır daha sonra driver kullanılarak site ana sayfasına ulaşılır . Arama çubuğu bulunup arama yapılır . Çıkan ürünlerden ürün ismi ile uyuşan ilk ürün seçilir ve ürüne tıklayarak ürün sayfasına gidilir. Daha sonra ürün sayfasına mevcut satıcı bilgileri alınır ve csv ye kaydedilir. Daha sonra “diğer satıcılar “ gibi bir etiketin varlığı kontrol edildikten sonra öyle bir element yoksa program sonlanır ve mevcut satıcı bilgileri kullanıcıya gösterilir . Eğer “diğer satıcılar” benzeri bir etiket bulunursa ona tıklanır ve diğer satıcılar sayfasına gidilir. Ordaki satıcıların verilei çekilerek sırayka csv ye kaydedilir ve program sonlanır ve bilgileri sonuç penceresinde kullanıcıya gösterir.

→Şimdi her site için farklı olan kodları açıklayalım.

- **Trendyolda** ürün sayfasında “diğer satıcılar” etiketine basınca sayfa aşağı kayıyor ve 4 satıcı ve onların altında “daha fazla göster” butonu bulunuyor. O butona basınca tüm diğer satıcılara ulaşabiliyoruz. Trendyol fonksiyonunda sayfa aşağı kaydıktan sonra önce “daha fazla göster” butonun varlığını try/catch ile kontrol ederiz eğer varsa try içinde o tuşa basılır ve çıkan tüm diğer satıcılarla birlikte tüm satıcı bilgileri çekilir ve csv ye kaydedilir. Eğer “daha fazla göster” tuşu yoksa catch

içinde devam ederek sadece olan satıcı bilgileri alınıp csv ye kaydedilir ve program sonlandıktan sonra veriler kullanıcıya gösterilir.

- **Hepsiburada** da ise diğer kullanıcılar etiketine basınca sağ tarafta bir pencere açılıyor ve diğer satıcılar listeleniyor. Ordan tüm diğer satıcı bilgileri çekilerek csv ye kaydedilir ve sonuç penceresinde gösterilir.
- **N11** de ise diğer satıcılar etiketine basınca ekran aşağı kayıyor ve diğer satıcılar gösteriliyor. Limiti aşarsa diğer satıcıların sayısı 1-2-3 diye sayfalara ayrılıyor. Burada program eğer diğer sayfa “→” butonu varsa önce bulunan sayfadaki veriyi çekip csv ye kaydediyor sonra butona basıyor ve 2. Satıcılar sayfasına geçiyor daha sonraki sayfada yine aynı şekilde buton olmayıncaya kadar devam eder ve en sonunda buton yoksa program sonlanır ve veriler kullanıcıya sonuç penceresinde gösterilir.

Sonuç Görselleştirme

?

→**show_result_window()** fonksiyonu tüm veri çekme ve csv ye yazma işlemleri bittikten sonra oluşturulur . Ürünler ilgili bulunan tüm satıcılar ve bilgileri başta tabloda gösterilir. Aranılan ürünün ismi orta üstte gösterilirken altında 3 adet buton bulunmaktadır. Kullanıcı “Top 5 Cheapest” butonuna basınca program içlerinde en ucuz 5 tanesini gösterir . Eğer kullanıcı “Top 5 Most Expensive” butonuna basarsa program en pahalı 5 ürünü kullanıcıya gösterir. Eğer kullanıcı tüm satıcıları tekrar görmek isterse “Show All” butonuna basması yeterli.

5. Çıktılar



Seller	Seller Rating	Product Rating	Price
Teknosa	9.1	4.7	32.159 TL
MediaMarkt	9.1	4.7	32.169 TL
Elif Ticaret/Balikes...	9.2	4.7	33.089 TL
websinial	9.8	4.7	33.599 TL
HBT TEKNOLOJİ	10	4.7	33.600 TL



Seller	Seller Rating	Product Rating	Price
BALPETEK TEKNOLOJİ	9.8	4.7	44.999 TL
Kaplantiteknoloji	9.7	4.7	42.999 TL
PARK İLETİŞİM - ELEK...	10	4.7	41.999 TL
Kayaalar Plaza	9.4	4.7	40.500 TL
nesiltech	9.3	4.7	40.199 TL



Seller	Seller Rating	Product Rating	Price
Teknosa	9.1	4.7	32.159 TL
MediaMarkt	9.1	4.7	32.169 TL
Elif Ticaret/Balikes...	9.2	4.7	33.089 TL
HBT TEKNOLOJİ	10	4.7	33.600 TL
Koray Shoes	10	4.7	33.601 TL
websinial	9.8	4.7	33.599 TL
TEKNOLOJİ MOBİL STOR...	N/A	4.7	33.602 TL
Glam & Glamorous	9.4	4.7	34.000 TL
VATAN BİLGİSAYAR	8.8	4.7	33.999 TL
Aksaray İletişim	9.6	4.7	34.269 TL
Trendyol	9.3	4.7	34.249 TL
Pratikceli İletişim	9.9	4.7	34.999 TL
Gürgeçler Apple Yet...	9.3	4.7	35.199 TL

6. CSV Veri Örneği

```
main2.py  datas.csv  X
datas.csv > data
1 seller,seller_rating,product_rating,price
2 Teknosa,9.1,4.7,32.159 TL
3 MediaMarkt,9.1,4.7,32.169 TL
4 Elif Ticaret/Balıkes...,9.2,4.7,33.089 TL
5 HBT TEKNOLOJİ,10,4.7,33.600 TL
6 Koray Shoes,10,4.7,33.601 TL
7 websinial,9.8,4.7,33.599 TL
8 TEKNOLOJİ MOBİL STOR...,N/A,4.7,33.602 TL
9 Glam & Glamorous,9.4,4.7,34.000 TL
10 VATAN BİLGİSAYAR,8.8,4.7,33.999 TL
11 Aksaray İletişim,9.6,4.7,34.269 TL
12 Trendyol,9.3,4.7,34.249 TL
13 Pratikcell İletişim,9.9,4.7,34.999 TL
14 Gürgeçler Apple Yet...,9.3,4.7,35.199 TL
15 BittiBitiyor,9.4,4.7,35.249 TL
16 BİRİNCİ TEKNOLOJİ,9.8,4.7,35.999 TL
17 TAHA İLETİŞİM,9.7,4.7,35.999 TL
18 CRATOS TEKNOLOJİ,9.7,4.7,36.000 TL
19 SMTOPTAN,9.5,4.7,35.999 TL
20 ANADOLU TEKNOLOJİ & ...,9.7,4.7,36.000 TL
21 CumbaZ,9.4,4.7,36.299 TL
22 Ceptemax,9.5,4.7,36.399 TL
23 CepHane Teknoloji,9.4,4.7,36.599 TL
24 fdg iletişim bilişim...,9.3,4.7,36.900 TL
25 Alsambunu,9.3,4.7,36.999 TL
26 MAÇİN TEDARİK,9.6,4.7,37.400 TL
27 Nasa İletişim,9.4,4.7,37.499 TL
28 TEKNO.STAR,9.8,4.7,37.999 TL
29 ...
```

7. Analiz Özeti

Kod, CSV'den gelen verileri sıralayarak:

- En ucuz 5 satıcıyı
- En pahalı 5 satıcıyı
- Tüm satıcıları

kullanıcıya gösterir. Bu analiz, fiyat/fayda değerlendirmesi için kullanıcıya fayda sağlar.

8. Karşılaşılan Sorunlar ve Çözümler

Sorun	Çözüm
Web sitelerinin DOM yapısının dinamik olması	Selenium WebDriverWait ile bekleme stratejileri kullanıldı
Ürün bulunamaması durumu	try/except bloklarıyla hata yakalama yapıldı
Site ile etkileşime girmek isterken çerez veya bildirim istekleri yüzünden elemente ulaşamama	çerez ve bildirim isteğinin class bilgilerini alıp reddedildi
Bazı elementlerin aynı class name ' e sahip olmaları	O durumda css_selector ile elemente ulaşım sağlandı

9. Sonuç ve Değerlendirme

Proje kapsamında web tabanlı veri çekme, veri temizleme, görsel arayüz tasarımı ve temel analiz çıktıları başarıyla gerçekleştirilmiştir. Uygulama, Python'un GUI ve web scraping kabiliyetlerini bütünleştiren işlevsel bir araç olarak çalışmaktadır.