

YAZILIM LABORATUVARI 1

1. PROJE

Bilgisayar Mühendisliği Bölümü
Kocaeli Üniversitesi

Ali Murat Ekmekçi *
200202114

Melih Turaneri †
200202075

EKİM 2022

1 Özet

Bu doküman Yazılım Laboratuvarı 1 dersi 1. Projesi için çözümümüzü açıklamaya yönelik oluşturulmuştur. Dökümanda projenin tanımı, çözüme yönelik yapılan araştırmalar, kullanılan yöntemler, proje hazırlanırken kullanılan geliştirme ortamı ve kod bilgisi gibi programın oluşumunu açıklayan başlıklara yer verilmiştir. Doküman sonunda projemizi hazırlarken kullandığımız kaynaklar bulunmaktadır.

2 Proje Tanımı

2.1 Proje Tanımı

Web scraping, yapılandırılmış web verilerini otomatik bir şekilde toplama işlemidir. Aynı zamanda web veri çıkarma olarak da adlandırılır. Web kazımının ana kullanım du-

rumlarından bazıları, diğerleri arasında fiyat izleme, fiyat istihbaratı , haber izleme, olası satış oluşturma ve pazar araştırmasını içerir. Genel olarak, web verilerinin çıkarılması, daha akıllı kararlar almak için halka açık büyük miktarda web verisinden yararlanmak isteyen kişiler ve işletmeler tarafından kullanılır. Web scraping işlemi 2 aşamadan oluşur,

1. Web Taraması: Genelde “örümcek” olarak adlandırdığımız bir web tarayıcısı, aradığımız içeriği araştırmak için interneti tarayan ve linkleri toplayan bir sistemdir. Bu bir insan da olabilir, bir yapay zeka da.
2. Web Kazıması: Bir web kazıyıcı, bir web sayfasından verileri doğru ve hızlı bir şekilde çıkarmak için tasarlanmış özel bir araçtır. Web kazıyıcıları, projeye bağlı olarak tasarım ve karmaşıklık bakımından geniş ölçüde farklılık gösterir.

Peki web scraping ile neler yapılabilir?

*alimuratekmekci@gmail.com

†turaneri45@gmail.com

1. Fiyat İzleme: Çeşitli e-ticaret sitelerinden veriler toplanarak, ürününüz veya ürünleriniz için, Dinamik Fiyatlandırma, Rakip Analizi, Yatırım Kararı verme gibi projelerde kullanılabilir.
2. Haber ve İçerik Takibi: Sadece sosyal medyada değil modern medya üzerinde de hakkınızda neler söylendiğini, şirketiniz hakkında ne gibi haberler yapıldığını anlık olarak takip edip buna bağlı olarak aksiyonlar alabilir, ayrıca yatırım yapacağınız şirket hakkındaki verileri analiz ederek yatırımınızı nasıl veya nerelere yapacağınız kararınızı netleştirilebilir.
3. Duygu Analizi: Şirketiniz hakkında sosyal medya da nasıl bir algı oluştuğunu öğrenmek için duygu analizi modeli kurup sosyal medya verisini sürekli olarak kazıyarak ve analiz ederek aktif olarak müşteri beklentinizi ve deneyiminizi sürekli olarak arttırılabilir.

2.2 Amaç

1. Web scraping ile farklı alışveriş sitelerinden istenilen ürün hakkında bilgi elde etmek.
2. Bir uygulama içerisinde istenilen özellikteki ürünlerin filtrelenmesi ve sıralanması becerisini geliştirmek.
3. Dinamik özelliklere sahip bir program geliştirmek.
4. Web kodlama hakkında bilgi ve beceriye sahip olmak.

2.3 İsterler

1. Kendimize ait bir e-ticaret sitesi oluşturmalıyız ve bu siteyi Admin olarak güncelleyebilmeliyiz. Sitede yapılacak değişikliklerde
 - Bir notebook ürünün fiyatının değiştirilmesi,
 - Notebook ürünü ait bir kaydın kaldırılması,
 - Yeni notebook ürünün eklenmesi,
 - Notebook ürünün puanında değişiklik yapılması,
 - Ürün bilgilerinin güncellenmesi,

bilgilerin web scraping ile çekilmesi gerekmektedir. Alınan tüm bu değişiklikler farklı sitelerdeki ürünlerin listelendiği diğer web sitesinde anlık olarak göstermeliyiz.

2. Ürün bilgisinin gösterildiği web sayfasında tüm notebook ürünlerinin listelenmesi ve ürün bilgilerinin yer aldığı ana başlıklarının oluşturmamız gerekmektedir. Bu başlığa tıklayınca ürün hakkında bilgilerin (fiyat bilgileri dahil) olduğu sayfa açılmalıdır. Ayrıca ürünlerin listelendiği sayfadaki E-ticaret linki üzerine tıklayınca ürünün bu E-ticaret sitesindeki sayfasına kullanıcı yönlendirmeliyiz.
3. Dinamik arama ile modele göre ve E-ticaret sitelerine göre arama işlemini yapabilmeliyiz.
4.) Ürün bilgilerinin gösterildiği web sayfası üzerinde Filtreleme ve Fiyat bilgisine göre küçükten büyüğe veya büyük-

ten küçüğü sıralama ayrıca puanı yüksek olan ürüne göre sıralama işlemini yapabilmeliyiz.

5. Oluşturduğumuz E- Ticaret sitesine kayıtlı olan bir ürünün aynı bilgileri ile kayıt edilmeli ve veritabanında Duplicate Kontrolü yapabilmeliyiz.

3 Araştırmalar

Projeye başlamadan önce web scraping nedir, nasıl yapılır sorularının cevaplarını bulduk. Sonrasında kazıdığımız verileri MongoDB içerisine nasıl çekebiliriz, çektikten sonra kendi e-ticaret sitemiz ile MongoDB bağlantısını nasıl yapabiliriz bunu araştırdık. Son olarakta kendi e-ticaret sitemizi yapabilmek için Django hakkında çeşitli araştırmalar yaparak bilgi sahibi olduk ve işe koyulduk.

4 Yöntem

Projemize ilk olarak web scraping işlemleri ile başladık. Sırasıyla N11, Hepsiburada ve Trendyol sitelerinden uzun uğraşlarımız sonucu ürünlerin modelleri, fiyatları, genel özellikleri, linkleri, kullanıcı puanları ve resimlerini elde etmiş olduk. Sonrasında ilk olarak Amazon.com'u denemeye başladık ancak response 503 ve 403 hataları ile karşılaştığımızdan dolayı veri elde edemedik. Sonrasında Vatan.com'u denedik ancak Amazon'da aldığımız hataların aynısını aldık. Son çare olarak Teknosa.com'dan veri çekmeye başladık. Ürünlerin model, fiyat, numara, link ve resimlerini çekmekte sıkıntı yoktu ancak genel özelliklerinin bulunduğu liste bu özellikleri çekmemize izin

vermiyordu. Bu sebeple Teknosa.com'u da es geçerek projeye devam etme kararı aldık. Sonrasında ise çektiğimiz verileri bir veritabanında saklamamız lazımdı. Biz veritabanı olarak MongoDB kullanmaya karar verdik ve verileri içine aktardık. Sonrasında Django ile web sitemizi yapmaya başladık. İlk olarak N11'den çektiğimiz verileri MongoDB'yi Django'ya bağlayarak aktardık. Bu veriler ile sitemizi şekillendirmeye başladık. Ürünleri gösterme, sıralama ve arama özelliklerini başarıyla yaptık ve Hepsiburada ve Trendyol'dan çektiğimiz verileri de siteye entegre ederek sıralama ve arama işlemlerimizin başarısını test ettik ve başarılı olduğumuzu gördük. Sonrasında veritabanı içerisinde ürünlerin modellerini eşitleyerek ortak olan notebookları bulduk. Bunların olduğu bir web sayfasını da web sitemize ekledik. Sonrasında ortak olmuş olan ürünlerin hangi sitede fiyatı ne ve aldığı puanını da göstererek projemizi başarıyla sonlandırmış olduk.

4.1 Kazanımlar

Web scraping, MongoDB ve Django hakkında bilgi sahibi olma ve bildiklerimizi deneyimleme fırsatı elde etmiş olduk.

5 Geliştirme Ortamı

Web scraping işlemleri için Python, veritabanı işlemleri için MongoDB, e-ticaret sitesi oluşturmak için ise Python Django kullandık.

6 Kod Bilgisi

6.1 Akış Diyagramı

Kısım ektedir. [1](Akış Diyagramı)

- <https://app.diagrams.net/> -Akış di-yagramı oluşturmak için.

<https://ceaksan.com/tr/veri-kazima-data-scraping-nedir>

7 Sonuç

Projenin gerektirdiği tüm isterleri sağladık.

8 Kaynakça

- <https://ceaksan.com/tr/veri-kazima-data-scraping-nedir>
- https://tr.wikipedia.org/wiki/Web_kazıma
- <https://medium.com/kaveai/web-scraping-453e96a86195>
- <https://www.sisasoft.com.tr/web-scraping/>
- <https://www.gtech.com.tr/mongodb-nedir-nerelerde-kullanilir/>
- <https://tr.wikipedia.org/wiki/MongoDB>
- <https://www.yusufsezer.com.tr/mongodb-nedir/>
- <https://medium.com/@berkekurnaz/nedir-bu-mongodb-994a94a9d1df>
- <https://halilozel1903.medium.com/django-nedir-52b29ebb7298>
- <https://tr.wikipedia.org/wiki/Django>
- <https://tutorial.djangogirls.org/tr/django/>
- <https://www.hosting.com.tr/blog/django/>
- item <https://www.texdizgi.com/category/genel/>