

KOCAELİ ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ YAZILIM LABORATUVARI 1 PROJE 1

WEB SCRAPING YAPISI VE PYTHON DİLİNDE UYGULANIŞI

1st Yunus Emre Kırıcı
Bilgisayar Mühendisliği 2. Öğretim
200202056
200202056@kocaeli.edu.tr

2nd Mustafa Can Öncü
Bilgisayar Mühendisliği 2. Öğretim
200202100
2002202100@kocaeli.edu.tr

I. PROJENİN ÖZETİ

Projemizin amacı "Python" dilinde "web scraping" algoritmasının gerçekleştirilmesidir. Verilen fonksiyonları gerçekleştirerek ödev isterlerini yerine getirmemiz beklenmektedir. Bu fonksiyon isterlerinden ilki input dosyasının "web scraping" yapısına uygunluğunun kontrolüdür. Bu isterlerden ikincisi ise kullanıcıdan alınacak olan (txt formatındaki input dosyasının içersine koyarak) sıralamanın gerçekleştirilmesidir. Bu isterlerden üçüncüsü ise sıralanmış verilerin id, saat, bekleme süresi, istenen iniş süresi gibi özelliklerinin istenen veri yapısı ile istenilen formata sokulmasını sağlamaktadır. Bu yapı hem linked list hem de array ile çözülebilecekken biz array kullanmayı tercih ettik ve tüm projeyi buna göre şekillendirdik. Sonuncu fonksiyonumuz ise bizden verilen tüm şartların sonucunu output dosyasına aktarmamızdır. Konsol sürekli "enter" tuşu yardımıyla kendini yenilemektedir ve böylece proje tamamlanmaktadır. Raporumuz L^AT_EX formatında yazılmıştır.

II. GİRİŞ

A. Asıl projeye ait bilgilerin detaylarına giriş kısmı

Projenin yazıldığı IDE süreç boyunca farklılık göstermektedir. "VSCODE" ve "Py-Charm" temel olarak kullanılan IDE'ler olarak belirlendi. Bunun yanı sıra web üzerinden online editörler de kullanılmıştır. Temel amacımız algoritmayı efektif bir şekilde kurup, bu mantığı konsola aktarma aşamasında başarılı olabilmektir.

III. TEMEL BİLGİLER

Projemizde pek çok array yapısı tanımlandı ve kullanıldı. Bu sayede "web scraping" oluşturacağımız yapı mevcut hale geldi. "int eklenecek olanlar urunadi;" fonksiyonunu etkileyecek işlemler yapabilmek için şu fonksiyonları tanımladık: yazdır(), güncelle(), ertele(), add(). Otomatik çalışma sayesinde yukarıda verilen fonksiyonlara erişim sağlanarak kullanıcının isteğine göre çalıştırılması sağlandı.

IV. YÖNTEMLER

Projemizde veri yapıları ve algoritmaları dersini baz olarak çeşitli yöntemler uyguladık. Fonksiyonlarımızın hepsi tam olarak

çalışmaktadır. Bu sayede istenilen yapı gerçekleştirilmiştir.

A. Projenin çalışma prensibini gösteren kısım

Bu bölümde fonksiyonlar tek tek ele alınacaktır. void guncelle():kalkış saatlerinin güncellenmesini sağlar ve "web scraping" yapısını oluşturur. void yazdir():Son olarak artık yazılması gereken verilerini outputa yazdırır.void ertele(int oncelik,int ucakid,int inisaat) girilen verilere göre erteleme işlemini gerçekleştirir.void add(int oncelik,int ucakid,int inisaat): Parantez içinde verilen değerleri alarak ekleme yapmayı sağlayan fonksiyonumuzdur.

V. DENEYSEL SONUÇLAR

Kullanıcı programı çalıştırdığında input dosyasından çekilmiş tek bir verinin konsol ekranında yazması ile karşılaşmaktadır. Priority queue mantığı işlemler tanımlanınca görülebilir. Yapılmak istenen işlemlerin yanında numaraların olduğu menü ile karşılaşırız. Bu sayede işlemlerin gerçekleşeceği menüyü görüyoruz. "ENTER" tuşu ile sonuçları ekrana bastırılır. Veriler tam anlamıyla bitince son "enter" programdan çıkışı sağlar.

A. Projenin deneysel sonucunu gösteren kısım

VI. SONUÇLAR

A. Projenin sonucunu gösteren kısım

Projemiz başarılı bir şekilde sonuca ulaştı. Bizden istenen fonksiyonlar istenen yapılar kullanılarak gerçekleştirildi. Bu proje sayesinde pythonun yapısını ve web geliştirmenin yapısını çok daha detaylı inceleme ve öğrenme fırsatı bulduk. Bu algoritma yapıları dışında bir library(kütüphane) dosyası indirerek bunu kodumuza entegre etmeyi başardık. Bunun

sonucu olarak öncelikli BeautifulSoup yapısını for döngüsü yoluyla yapmayı kavradık ve projemizde kullanabildik.

VII. KAYNAKÇA

A. Projede yararlandığımız kaynaklar

- www.cs.bham.ac.uk/~jxb/DSA/dsa.pdf
- Çölkesen, Rifat, Veri Yapıları Ve Algoritmaları
- <https://www.youtube.com/watch?v=aPRqocoBsFQ>
- <https://www.geeksforgeeks.org/advanced-data-structures/>
- <https://medium.com/analytics-vidhya/web-scraping-with-scrapy-and-django-94a77386ac1b>
- <https://www.youtube.com/watch?v=GK2Zl5DG0VM>
- <https://www.youtube.com/watch?v=aUKoTGnzLic>
- <https://www.youtube.com/watch?v=XDxLEUgVDMm>

VIII. YALANCI KOD

- BAŞLA
- BeautifulSoup library import edilir
- siteden request alınır
- alınan request değere atanır
- atanan değer kullanılarak alınabilen değerler alınır
- alınmayan değerler için link değeri alınır
- link değeri için ayrı request oluşturulur
- değerlerin html kodda karşılığı aranır
- alınan değerlerin "div" ya da "li" içerisinde olup olmadığına bakılır
- class'a göre değerler alınır
- liste ya da table olması durumunda for döngüsü devreye girer
- değer if bloğu ile kontrol edilir
- çekilen data adı markaya eşitse
- değeri markaya atar çağırır
- çekilen data adı fiyata eşitse
- değeri fiyata atar çağırır
- çekilen data adı işlemciye eşitse
- değeri işlemciye atar çağırır
- çekilen data adı işletim sistemine eşitse
- değeri işletim sistemine atar çağırır
- çekilen data adı rame eşitse
- değeri rame atar çağırır
- Değerler database yazar
- Değerler databaseden silinir
- Databaseden çıkılır
- BİTİR

isim	Tip	Şema
> django_content_type		CREATE TAB
> django_migrations		CREATE TAB
> django_session		CREATE TAB
▼ product_category		CREATE TAB
id	integer	"id" integer
title	varchar(30)	"title" varchar(30)
keywords	varchar(200)	"keywords" varchar(200)
description	varchar(200)	"description" varchar(200)
image	varchar(100)	"image" varchar(100)
status	varchar(10)	"status" varchar(10)
slug	varchar(50)	"slug" varchar(50)
create_at	datetime	"create_at" datetime
update_at	datetime	"update_at" datetime
parent_id	bigint	"parent_id" bigint
▼ product_product		CREATE TAB
id	integer	"id" integer
title	varchar(30)	"title" varchar(30)
keywords	varchar(200)	"keywords" varchar(200)
description	varchar(200)	"description" varchar(200)
image	varchar(100)	"image" varchar(100)
price	real	"price" real
amount	varchar(100)	"amount" varchar(100)
detail	text	"detail" text
status	varchar(10)	"status" varchar(10)
create_at	datetime	"create_at" datetime
update_at	datetime	"update_at" datetime

isim	Tip	Şema
▼ Tablolar (1)		
▼ c		CREATE TABLE c(marka TE
marka	TEXT	"marka" TEXT
urun_adi	TEXT	"urun_adi" TEXT
model_no	TEXT	"model_no" TEXT
isletim_sistemi	TEXT	"isletim_sistemi" TEXT
islemci_tipi	TEXT	"islemci_tipi" TEXT
islemci_nesli	TEXT	"islemci_nesli" TEXT
ram	TEXT	"ram" TEXT
disk	TEXT	"disk" TEXT
ekran_boyutu	TEXT	"ekran_boyutu" TEXT
puani	TEXT	"puani" TEXT
urun_fiyati	TEXT	"urun_fiyati" TEXT
site_ismi	TEXT	"site_ismi" TEXT
link	TEXT	"link" TEXT
İndisler (0)		
Görünümler (0)		
Tetikleyiciler (0)		

Veritabanı Yapısı

Veriyi Görüntüle

Pragmaları Düzenle

SQL kodunu yürüt

Tablo:

c



Filter in any column

	marka	urun_adi	model_no	isletim_sistemi	islemci_tipi	islemci_nesli	ram
	Filtre	Filtre	Filtre	Filtre	Filtre	Filtre	Filtre
1	Lenovo	Lenovo IdeaPad 3 15ITL6 82H802F6TX i5-1135G7 8 GB 1 ...	IdeaPad 3 15ITL6 82H802F6TX	Freedos	Intel Core i5	Intel i5-1135G7	8 GB
2	Asus	Asus VivoBook 15 Oled D513UA-L1346 Ryzen 5 5500U 8 ...	D513UA-L1346	Freedos	Amd Ryzen 5	Amd Ryzen 5-5500U	8 GB
3	Asus	Asus X415EA-EK977 i5-1135G7 8 GB 256 GB SSD 14" Fre...	X415EA-EK977	Freedos	Intel Core i5	Intel i5-1135G7	8 GB
4	Dell	Dell Inspiron 15 3511 I35111015U i5-1135G7 8 GB 512 GB ...	Inspiron 15 3511 I35111015U	Ubuntu	Intel Core i5	Intel i5-1135G7	8 GB
5	HP	HP Victus 16-D1021NT 6G0E9EA i5-12500H 16 GB 512 G...	6G0E9EA	Freedos	Intel Core i5	Intel i5-12500H	16 GB
6	Dell	Dell G15 5511 G55112306U i5-11260H 16 GB 512 GB SSD ...	G15 5511 G55112306U	Ubuntu	Intel Core i5	Intel i5-11260H	16 GB
7	Asus	Asus X515JA-EJ2120A3 i7-1065G7 16 GB 512 GB SSD ...	X515JA-EJ2120A3	Freedos	Intel Core i7	Intel i7-1065G7	16 GB
8	Lenovo	Lenovo Legion 5 82JM0013TX i7-11800H 16 GB 512 GB ...	Legion 5 82JM0013TX	Freedos	Intel Core i7	Intel i7-11800H	16 GB
9	Acer	Acer Extensa 15 EX215-54 NX.EGJEY.005 i5-1135G7 8 GB...	NX.EGJEY.005	Freedos	Intel Core i5	Intel i5-1135G7	8 GB
10	Acer	Acer Gaming Aspire 7 A715-42G NH.QE5EY.006 R5-5500...	NH.QE5EY.006	Freedos	Amd Ryzen 5	Amd Ryzen 5-5500U	8 GB
11	Dell	Dell Inspiron 15 5515 CYBORG N515AMD2201_102 ...	Inspiron 15 5515 CYBORG N515AMD2201_102	Windows 11 Home	Amd Ryzen 5	Amd Ryzen 5-5500U	8 GB
12	HP	HP 250 G8 45Q24ES1 i3-1005G1 8 GB 256 GB SSD MX13...	45Q24ES1	Windows 11 Pro	Intel Core i3	Intel i3-1005G1	8 GB
13	Acer	Acer Nitro 5 AN517 NH.QFCEY.002 i7-11800H 16 GB 512 ...	NH.QFCEY.002	Linux	Intel Core i7	Intel i7-11800H	16 GB
14	Monster	Monster Abra A5 V19.1.1 i5-12500H 8 GB 500 GB SSD 4 ...	Abra A5 V19.1.1	Windows 11 Home	Intel Core i5	Intel i5-12500H	8 GB
15	HP	HP 255 G8 4P3M4ES Ryzen 7 5700U 8 GB 512 GB SSD ...	4P3M4ES	Windows 10 Home	Amd Ryzen 7	Amd Ryzen 7-5700U	8 GB
16	Asus	Asus D515DA-BR3169 Ryzen 3 3250U 4 GB 256 GB SSD ...	D515DA-BR3169	Freedos	Amd Ryzen 3	Amd Ryzen 3-3250U	4 GB
17	Dell	Dell Inspiron 5415 I54151012WH Ryzen 5 5500U 8 GB 25...	Inspiron 5415 I54151012WH	Windows 11 Home	Amd Ryzen 5	Amd Ryzen 5-5500U	8 GB
18	Asus	Asus Rog Strix G15 G513RM-HF266 R7-6800H 16 GB 512...	Rog Strix G15 G513RM-HF266	Freedos	Amd Ryzen 7	Amd Ryzen 7-6800H	16 GB
19	Asus	Asus Tuf Gaming F15 FX506LHB-HN348 i5-10300H 8 GB ...	FX506LHB-HN348	Freedos	Intel Core i5	Intel i5-10300H	8 GB
20	HP	HP 15S-EQ3004NT 68N36EA R7-5825U 8 GB 256 GB SS...	68N36EA	Freedos	Amd Ryzen 7	Amd Ryzen 7-5825U	8 GB
21	Acer	Acer Aspire 7 A715-42G NH.QBFEY.007 R5-5500U 8 GB ...	NH.QBFEY.007	Freedos	Amd Ryzen 5	Amd Ryzen 5-5500U	8 GB
22	Lenovo	Lenovo V14 G2 ALC 82KC0005TX Ryzen 3 5300U 4 GB ...	V14 G2 ALC 82KC0005TX	Freedos	Amd Ryzen 3	Amd Ryzen 3-5300U	4 GB