

YAZILIM LABORATUVARI 1

PROJE 1

TARIK GÖREN
Bilgisayar Mühendisliği
200202022@kocaeli.edu.tr

ÖMER ARAN
Bilgisayar Mühendisliği
190202012@kocaeli.edu.tr

ÖZET

Bu proje ile amacımız web scraping yaparak e-ticaret sitelerinden notebook ürünlerini ve özellikle-rini çekmek ve en ucuz ürünü listelemek, filtrelemek ve sıralamak aynı zamanda kendi e-ticaret sitemizi de oluşturup listelediğimiz ürünlerimizi de scrape etmektir.

I. GİRİŞ

Bu projede backend kısmında java dilini ve java frameworku olan Spring framework kullanmayı tercih ettik. Javanın geniş kullanım alanı yelpazesinden de yararlanarak web scraping için geliştirilmiş olan Jsoup kütüphanesini kullandık. Veritabanı olarak bir RDBMS olan PostgreSQL'i hem kendi e-ticaret sistemimizde hem de scraping yaptığımız sistemde kullandık. Veritabanımızın NoSql kısmında ise Redis kullanmayı tercih ettik bunun sebebi ise Redis'in diğer tüm araçlara göre daha kolay ve güvenilir entegrasyon yapılmasıdır. Projemizin ön yüzünde ise backend geliştiricilerin daha kolay adapte olabileceği bir javascript kütüphanesi olan Angular frameworkü tercih ettik.

Bu proje kapsamında bizden aşağıdaki amaçların gerçekleştirilmesi istenmiştir:

A. Web Scraping

- 4 adet E-Ticaret (Trendyol, Hepsiburada, n11, Amazon) sitesi üzerinden web scraping kullanılarak bilgisayar(notebook) kategorisine ait bilgiler alınacaktır. Ayrıca E-ticaret sitelerine benzer yapıda kendiniz bir web sitesi oluşturacak (sadece bilgisayar(notebook) kategorisi olacak) ve web scraping kullanarak sitedeki bilgiler alınacaktır.

- Web Scraping kullanılırken siteye ait html bilgiler veya siteye request isteği ile bilgiler alınmalıdır. sağlanmalıdır.

B. Veritabanı

- Web scraping ile elde edilen veriler veritabanına kaydedilecektir. sağlanmalıdır.
- Kullanılacak veritabanı proje uygun olarak herhangi bir tanesini kullanabilirsiniz (MongoDB gibi veri tabanları kullanılması artı puan verilecektir.)
- Veritabanında aşağıdaki tabloda belirtilen özellik bilgilerinin kullanılması gerekmektedir. Veritabanındaki tabloların normalizasyon kurallarına uygun olacak şekilde tasarlanması ve organize edilmelidir.

C. Web Sayfası

- Farklı sitelerden çekilen ürün bilgilerinin gösterilmesi için bir web sayfası oluşturmanız beklenmektedir. Örnek web sayfası aşağıdaki resimde görülmektedir.
- Bu sayfa ilk açıldığında veritabanında yer alan tüm kayıtlar listelenmelidir.
- Listeleme işleminde ürün bilgilerinin yer aldığı ana başlık bulunmalıdır. Herhangi bir ürün başlığına tıklandığında ürün hakkındaki bilgilerin (fiyat ve ürün özellikleri) olduğu sayfaya yönlendirilecektir..
- Listeleme işleminde ürüne ait en düşük 3 fiyat ve bunların ait olduğu E- ticaret sitelerinin linkleri isim ve logolarıyla birlikte yer alacaktır. İsim veya logonun üzerine tıklanınca ürünün bulunduğu E-ticaret sitesinin sayfası açılacaktır.
- Web sayfasından herhangi bir ürünle ilgili dinamik arama işlemi yapılabilmelidir. Bu arama alanında bir ürünün isim, model ya

da bulunduđu E-ticaret sitesine göre yapılmalıdır. Ayrıca Arama sırasında yazım yanlış olması durumunda sistem düzeltilmiş öneride bulunmalıdır. Örnek: hepsi burada – yazımı düzeltilmiş sonuçları görüyorsunuz: Hepsiburada şeklinde olmalıdır.

- Web sayfasında dinamik filtreleme işlemi de ayrıca yer almalıdır. Filtreleme işlemi veritabanında yer alan ürünün tüm özellikleri için uygulanabilir olmalıdır. Yukarıdaki görselde e-ticaret sitelerinde bulunan laptop ürünleri markasına göre örnek olarak filtrelenmiştir.
- Web sayfasında en düşük fiyattan en yüksek fiyata, en yüksek fiyattan en düşük fiyata ve en yüksek puandan en düşük fiyata şeklinde sıralama işlemleri yapılabilirdir. Laptop bilgilerinin yer aldığı alanda fiyat bilgisi için de aynı durum geçerli olacaktır

D. İsterler

- Kendinize ait bir E-ticaret sitesi oluşturmalısınız ve bu site Admin tarafından güncellenebilir olmalıdır. 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 listelendiğı diğer web sitesinde anlık olarak gösterilmelidir.
- Ürün bilgisinin gösterildiğı web sayfasında tüm notebook ürünlerinin listelenmesi ve ürün bilgilerinin yer aldığı ana başlıklarının oluşturulması gerekmektedir. Bu başlığa tıklayınca ürün hakkında bilgilerin (fiyat bilgileri dahil) olduğı sayfa açılmalıdır. Ayrıca ürünlerin listelendiğı sayfadaki E-ticaret linki üzerine tıklayınca ürünün bu E-ticaret sitesindeki sayfasına kullanıcı yönlendirilmelidir.
- a) Dinamik arama ile (yukarıda istenilen tüm arama kriterleri dahil olacak şekilde) Asus ..vb modele göre arama işleminin yapılması b) Dinamik arama ile (yukarıda istenilen tüm arama

kriterleri dahil olacak şekilde) hepsiburada ..vb E-ticaret sitelerinin aranması

- a) Ürün bilgilerinin gösterildiğı web sayfası üzerinde Filtreleme işleminin yapılması b) Fiyat bilgisine göre küçükten büyüğe veya büyükten küçüğü sıralama ayrıca puanı yüksek olan ürüne göre sıralama işleminin yapılması
- Sizin oluşturacağınız E- Ticaret sitesine kayıtlı olan bir ürünün aynı bilgileri ile kayıt edilmeli ve veritabanında Duplicate Kontrolü yapılmalıdır.
- Sizin oluşturacağınız E- Ticaret sitesinde iki farklı kayıt işleminin yapılması ve Near Duplicate kontrolünün yapılması

II. YÖNTEM

A. Başlangıç

Öncelikle bizden istenen programları ve programların içerisinde kullanacağımız kütüphaneleri entegre ettik. Programları (SQL araçları, IDE ve web scraping kütüphanesini) yükledikten sonra isterleri gerçekleştirmeye başladık.

Bizden istenen ilk isterden - Web Scraping gerçekleşmesi - başladık. Jsoup dokümanları ve ilgili eğitim videoları ile web scraping yapacağımız siteleri oluşturduğumuz spring projesinin test kısmında denemeler yaparak ürün hakkında birlikte çekmeyi denedik.

B. Geliştirme Ortamı

Projenin backend tarafı Java'nın Spring Boot frameworkü ile geliştirilirken frontend tarafı AngularJS frameworkü ile geliştirilmiştir. Veritabanı olarak PostgreSQL (RDBMS) ve Redis (NoSQL) kullanılmıştır.

Macos-Windows sistemlerde geliştirilip IDE olarak IntelliJ Idea Ultimate Version ve WebStorm kullanılmıştır. Projenin versiyon kontrolü tamamen github üzerinden yapılmıştır. IDE'lerimizde yazdığımız kodları veritabanı bağlantısı sağladıktan sonra localimizdeki portlarda çalıştırıp her türlü olumsuzluğa karşı önlem açısından testler yapılmıştır.

C. Kullandığımız Teknolojiler

• Backend Kısmı

- **Spring Boot:** Spring boot, Java programlama dilinde geliştirilen bir framework'tür. Java dilinde microservice yazılmasına olanak sağlar.
- **Liquibase:** Liquibase veritabanı şeması üzerindeki değişikliklerimizi yönetmemizi sağlayan açık kaynaklı bir Java kütüphanesidir.
- **Hibernate:** Hibernate, Java platformunda yazılmış bir ORM aracıdır. ORM, nesne odaklı dillerdeki nesnelerin, ilişkisel veri tabanlarındaki kayıtlara nasıl karşılık geldiğini yürüten bir teknolojidir.
- **JPA (Java Persistence API):** Verileri kalıcı olarak saklamak, almak, bulmak, silmek ve güncellemek için kullanılan Java ORM şartnamesidir.
- **Swagger:** Web API geliştirmede en önemli ihtiyaçlardan biri dokümantasyon ihtiyacıdır. Çünkü API methodlarının ne işe yaradığı ve nasıl kullanıldığı dokümantasyon içerisinde anlaşılır olması gerekir ve swagger bu dokümantasyonu online ve sürekli güncel olarak kullanmamıza olanak sağlayan araçtır.
- **JSoup:** HTML belgelerinde saklanan verileri ayırtmak, ayıklamak ve değiştirmek için tasarlanmış açık kaynaklı bir Java kütüphanesidir. (Yalnızca HTML getirip parçalamaya yarar. Script execution desteklemez.)
- **Selenium:** Web scraping sırasında javascript yürütmemiz gereken sayfalarda kullandığımız teknolojidir.
- **ModelMapper:** Java'da aynı property isimlerine sahip bir objeden başka bir objeye veri aktarmak isterken kullandığımız tip dönüşüm kütüphanesidir.

– **Lombok:** Spring framework içerisinde class üretirken bu class'lara erişim sağlayıp kullanımı için getter, setter ve constructorlara ihtiyaç duyarız işte bu noktada basmakalıp kodları tekrar etmek yerine basit anotasyonlarla hızlandırmak ve daha az maliyetli hale getirmek amacıyla kullanılan kod üretme amacıyla üretilen bir kütüphanedir.

– **Spring-Security:** Spring framework kullanılarak geliştirilen doğrulama(Authentication), yetkilendirme (Authorization), şifreleme (Password Encoder) ve CSRF gibi güvenlik önlemleri sağlayan, Spring platformunda yer alan bir projedir.

– **JWT (Json Web Token):** Json Web Token bir RFC7519 endüstri standartıdır. JWT, kullanıcının doğrulanması, web servis güvenliği, bilgi güvenliği gibi birçok konuda kullanılabilir. Biz ise projemizde kullanıcının authentication süresini ve kullanıcı adını üretilen token içinde saklayarak kullandık.

– **Validation:** Validation kontrolünün amacı Web uygulamalarında kullanıcıların girmiş olduğu verilerin veritabanına kaydedilmeden önce istenen şartları sağlayıp sağlamadığını kontrol etmektir.

• Frontend Kısmı

– **AngularJS:** Angular, Google tarafından geliştirilen ve desteklenen SPA(Single Page Application) uygulama yapılmasına olanak sağlayan javascript frameworküdür. Projemizin ön yüzünü bu framework ile dizayn ettik.

– **PrimeNG:** PrimeNG, AngularJS tabanlı uygulamalar için açık kaynak kullanıcı arayüzü bileşen kitaplığıdır. Projemizin ön yüzündeki bazı componentlerde PrimeNG kütüphanesine de yer verdik.

- Veritabanı Kısmı

- **PostgreSQL:** PostgreSQL açık kaynak kodlu, SQL destekli bir ilişkisel veritabanı yönetim sistemidir.
- **Redis:** Redis key-value şeklinde tasarlanmış bir NoSQL veritabanıdır. Veriyi bellekte tuttuğu için çok hızlı okuma ve yazma yapılır. Biz ise projemizde kullanıcılar authentication olduğu zaman üretilen tokenın sistem tarafından üretildiğini doğrulamak amaçlı her JWT üretildiğinde üretilen tokenı redis'e key username value JWT olacak şekilde sakladık. HTTP interceptorler aracılığıyla da her istek geldiğinde gelen token'ları redis'te kaydedilen tokenlarla karşılaştırıp doğrulamayı bir üst seviyeye taşıdık.

III. PSUEDO KOD

Uygulama hem backend hem de frontend kısmı başlatıldıktan sonra scrapper sitesi açılıyor ve adminin yetki alanında scrapper yapılabilir. Listeleme kısmında herkes ürünleri görüntüleyebilir olacaktır.

```
login: username password
post -> controller
service -> database check
if(isPresent(user))
  redirect to admin panel
else( give an error)
```

```
add-data-to-db
post->controller
if(!product.isEmpty())
  go db and save
else error: product empty, fill the blanks
delete-data-from-db
http delete
if(processDone)
  print successfully deleted
update-data
if(id and product == full and valid)
  print data updated successfully
```

```
list products
check for cheapest products
if you like
  click product
  send it to scrapper site
  vatan, n11 ...
  product detail page
  pull product id
  product.url -> product detail page
  product details technical details
  go back -> list products
  if(user.role==admin)
    edit-add-delete
    admin can make these processes
  —
  e-commerce web site
  without role -> list products
  if button clicked detail->
    for one product
    detailpage(product.id)
```

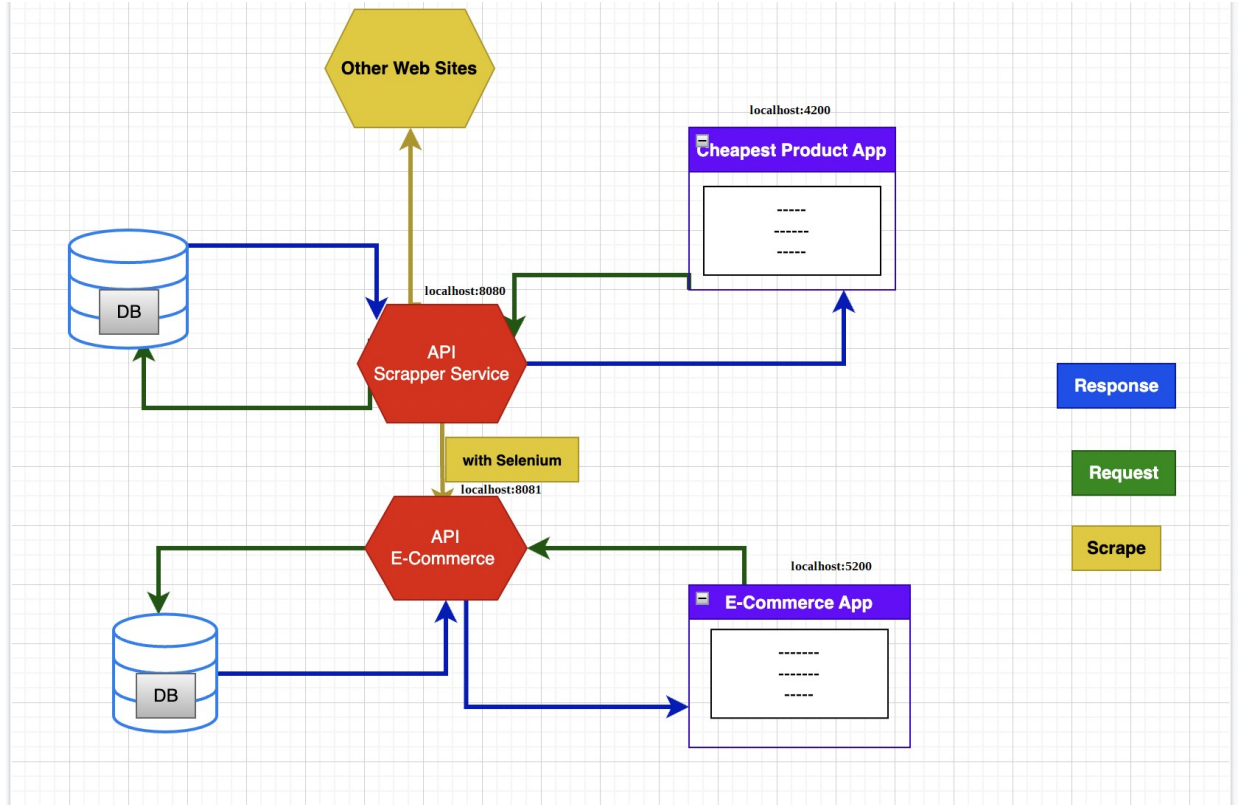
IV. SONUÇLAR

Proje amacı yerine getirilerek web scraping ile kendi e-ticaret sitemiz dahil 11 farklı web sitesinden notebook kategorisinden istediğimiz teknik özellikleri JSoup kütüphanesi ile aldığımız verileri düzenli bir şekilde veritabanımıza kaydettik bu sırada duplicate kontrolü yaparak aynı ürünleri ürün kodlarını karşılaştırarak eğerki aynı ürün birden fazla kez bulunduysa da fiyat karşılaştırması ile fiyatı daha avantajlı olan satıcının ürününü alıyoruz. Veritabanında farklı mağazalara ait aynı ürünlerin listelenmesi ve fiyat karşılaştırması ile Cheapest-Products-App adıyla web scraping sitemizde yayınlıyor ve kullanıcılara aynı ürünü satın alırken avantajlı olan fiyat ile yardımcı oluyoruz. Uygulamamızın ön yüzünde ise google tarafından 2016'da frontend javascript framework olarak sunulan Angular frameworkü ile ürün listeleme, ekleme, silme, düzenleme ve ilgili ürünlerin satıldığı web sitelere yönlendirme gibi özelliklerin dizaynını ve javascript aracılığıyla web servisimize istek atmasını sağladık. Projenin birkaç katmandan oluşması daha fazla iş yükü oluştursa da daha özgür bir şekilde yazılımı geliştirmemize olanak sağladı. Bu proje ile sonuç olarak bir web projesinin yaşam döngüsünün bir resmini görebilme fırsatı yakaladık.

KAYNAKLAR

- [1] Web Scraping yaparken kullandığımız kütüphane ve dokümanları
<https://jsoup.org/>
<https://stackoverflow.com/questions/12361925/html-parsing-with-jsoup>
https://www.youtube.com/watch?v=PF0iyeDmu9E&t=789s&ab_channel=ShaneLee
https://www.youtube.com/watch?v=2naKQHbFQpY&ab_channel=AutomationwithMaksu
<https://www.guru99.com/execute-javascript-selenium-webdriver.html>
- [2] Spring Framework Doküman
<https://spring.io/projects/spring-boot>
<https://www.baeldung.com/learn-spring-course>
- [3] Angular
<https://angular.io/docs>
<https://www.youtube.com/watch?v=Gx4iBLKLVHk>
<https://www.smashingmagazine.com/2018/11/a-complete-guide-to-routing-in-angular/>
<https://www.btkakademi.gov.tr/portal/course/angular-ile-web-programciligi-9914>
- [4] Primefaces - Angular için Frontend UI Kütüphanesi
<http://primefaces.org/primeng/setup>


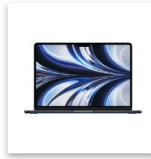
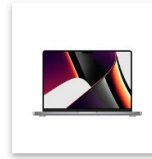

EKRAN GÖRÜNTÜLELİRİ



Şekil 1. System design interview

Cheapest Products App

29 adet ürün bulundu!

<p>Marka</p> <input type="text"/> <input type="checkbox"/> LENOVO <input type="checkbox"/> ACER <input type="checkbox"/> CASPER <input checked="" type="checkbox"/> APPLE	<p>İşlemci</p> <input type="text"/> <input type="checkbox"/> INTEL CORE i7 <input type="checkbox"/> APPLE M2 <input type="checkbox"/> AMD RYZEN™ 7 <input type="checkbox"/> APPLE M1 MAX	<p>Ekran Kartı</p>	
<p>MLY33TU/A</p>  <p>MacBook Air MLY33TU/A M2 8Gb-256Gb Ssd-Liquid Retina- 13.6inc-Gece Yarısı APPLE</p> <p>23575 TL 23999 TL 24704.05 TL</p>	<p>MLY43TU/A</p>  <p>MacBook Air MLY43TU/A M2 8Gb-512Gb Ssd-Liquid Retina- 13.6inc-Gece Yarısı APPLE</p> <p>29069.05 TL 30649 TL 32289 TL</p>	<p>MKGP3TU/A</p>  <p>MacBook Pro MKGP3TU/A M1 Pro 16Gb-512Gb Ssd-Liquid Retina-14inc-Space Grey APPLE</p> <p>36574.05 TL 38039 TL 41999 TL</p>	<p>MLXW3TU/A</p>  <p>MacBook Air MLXW3TU/A M2 8Gb-256Gb Ssd-Liquid Retina- 13.6inc-Uzay Grisi APPLE</p> <p>23602 TL 24604.05 TL 24699 TL</p>

Şekil 2. Cheapest Product App Product List sayfası

Marka

☐

☐ LENOVO
☐ ACER
☐ CASPER
☒ APPLE

İşlemci

☐

☐ AMD RYZEN™ 9
☒ APPLE M1
☐ AMD RYZEN™ 5
☐ INTEL CORE İ3

Ekran Kartı

4 adet ürün bulundu!

MYD82TU/A

MacBook Pro Touch Bar
MYD82TU/A M1 8Gb-256Gb
Ssd-Retina-13inc-Space Grey
APPLE

25599 TL

25649.05 TL

25669.02 TL

MYDC2TU/A

MacBook ProTouch Bar
MYDC2TU/A M1 8Gb-512Gb
Ssd-Retina-13inc-Gümüş
APPLE

25649.05 TL

30999 TL

30999 TL

MGND3TU/A

MacBook Air MGND3TU/A M1
8Gb-256Gb Ssd-Retina-
13.3inc-Gold
APPLE

18999.05 TL

19899 TL

20399.99 TL

MGN63TU/A

MacBook Air MGN63TU/A M1
8Gb-256Gb Ssd-Retina-
13.3inc-Space Grey
APPLE

18999.05 TL

20999 TL

20999 TL

Şekil 3. Cheapest Product App Product List sayfası çoklu filtre

Dizüstü Bilgisayar

MacBook Air MLXW3TU/A M2 8Gb-256Gb Ssd-
Liquid Retina-13.6inc-Uzay Grisi

Genel Bilgiler

- APPLE
- Uzay Grisi
- Ev ve Öğrenci
- 1.24 Kg.

Bulunan Mağazalar

23602 TL ★0/5

24604.05 TL ★5/5

24699 TL ★0/5

25989 TL ★0/5

25999 TL ★5/5

Teknik Özellikler

İşlemci	APPLE M2
Ekran Kartı	Paylaşılmıyor
RAM	8 GB
Disk	256 GB
İşletim Sistemi	macOS Monterey

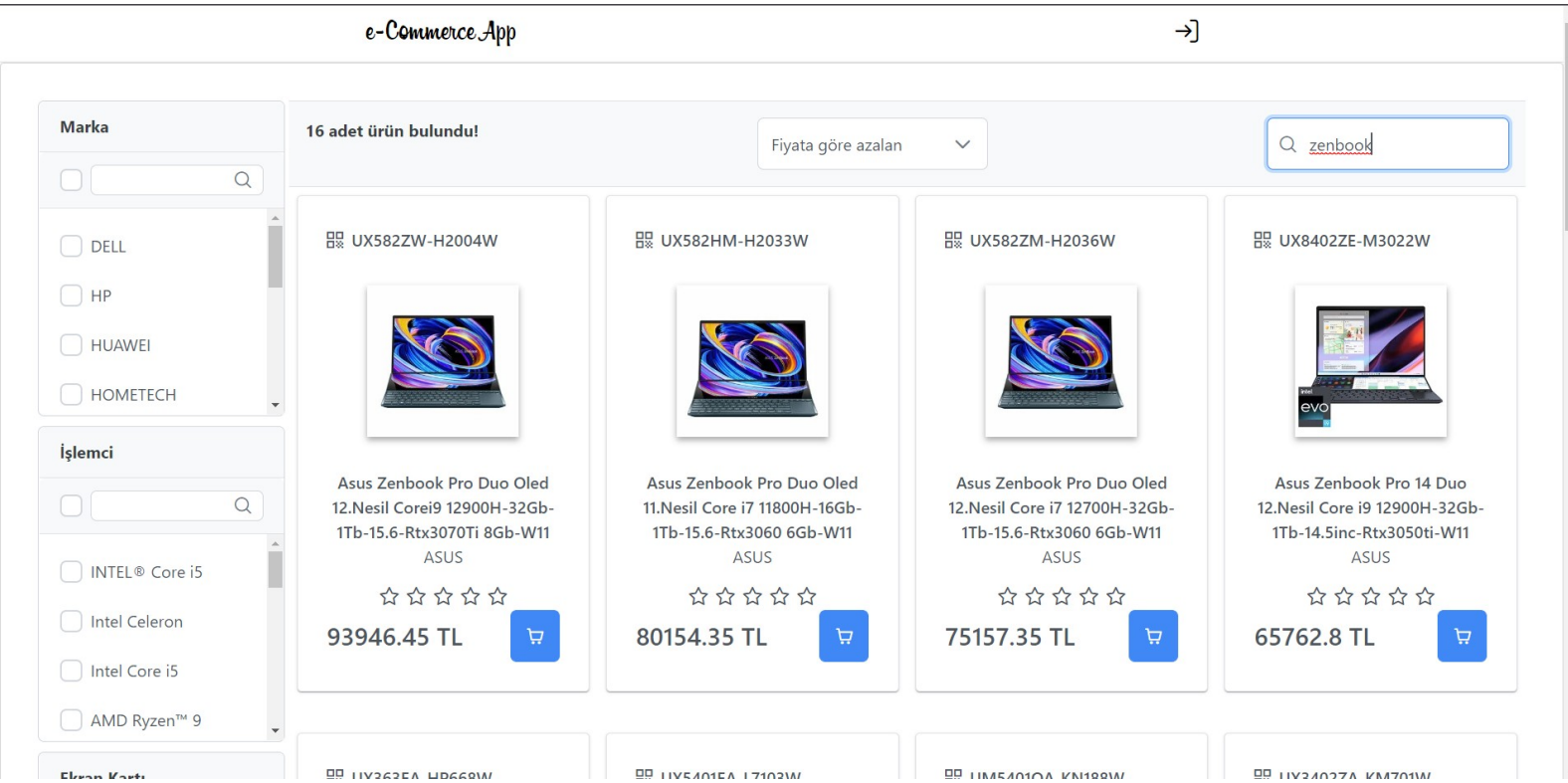
Diğer Özellikler

Ekran Boyutu	13.6 inch
Ağırlık	1.24 Kg.
Renk	Uzay Grisi
Kullanım Tipi	Ev ve Öğrenci
Model	MLXW3TU/A

Şekil 4. Cheapest Product App Product Detail sayfası

```
@Query(nativeQuery = true,
value = "SELECT * " +
" FROM product p\n" +
" WHERE ( :#{filter.brandName.size()} = 0 OR p.product_brand IN (:#{filter.brandName}) )" +
" AND ( :#{filter.cpu.size()} = 0 OR p.product_cpu IN (:#{filter.cpu}) )" +
" AND ( :#{filter.hdd.size()} = 0 OR p.product_hdd IN (:#{filter.hdd}) )" +
" AND ( :#{filter.ram.size()} = 0 OR p.product_ram IN (:#{filter.ram}) )" +
" AND ( :#{filter.color.size()} = 0 OR p.product_color IN (:#{filter.color}) )" +
" AND ( :#{filter.operatingSystem.size()} = 0 OR p.product_operating_system IN (:#{filter.operatingSystem}) )" +
" AND ( :#{filter.gpu.size()} = 0 OR p.product_gpu IN (:#{filter.gpu}) )" +
" AND ( :#{filter.screenSize.size()} = 0 OR p.product_screen_size IN (:#{filter.screenSize}) )" +
" AND ( :#{filter.searchTitle.length()} = 0 OR lower(p.product_title) like concat('%',:#{filter.searchTitle.toLowerCase()}, '%') )"
public List<Product> getProductsByFilter(FilterRequestÜto filter, Pageable pageable);
```

Şekil 5. Filtreleme sorgusu



Şekil 6. E-Commerce App Product List sayfası arama ve sıralama filtreleri

Marka

☐

DELL

HP

HUAWEI

HOMETECH

İşlemci

☐

INTEL® Core i5

Intel Celeron

Intel Core i5

AMD Ryzen™ 9

Ekran Kartı

65 adet ürün bulundu!

Sıralama

4H1T0EA

Hp Spectre X360 14-Ea0008nt
11.Nesil Core i7 1165G7-16Gb-
1Tb Ssd-13.5inc-W10
HP

☆☆☆☆☆

28499.05 TL

6Y7Z4EA

Hp 15-dw4008nt 12.Nesil Core
i5 1235U-8Gb-512Gb Ssd-
15.6inc-W11
HP

★★★★★

15199.05 TL

4H0U5EA

Hp Pavilion 14-Dv0029Nnt
11.Nesil Core i7 1165G7-8Gb-
512Gb Ssd-14-Mx450 2Gb-W10
HP

☆☆☆☆☆

18999.05 TL

5410-F65WP165N

Dell Vostro 14 5410 11.Nesil
Core i7 11370H-16Gb-512Gb
Ssd-14-Mx450 2Gb-W10pro
DELL

★★★★☆

25766.85 TL

3Y4V2EA

Hp Spectre X360 14-Ea0008nt
11.Nesil Core i7 1165G7-16Gb-
1Tb Ssd-13.5inc-W10
HP

☆☆☆☆☆

28499.05 TL

68P09EA

Hp 15-dw4008nt 12.Nesil Core
i5 1235U-8Gb-512Gb Ssd-
15.6inc-W11
HP

★★★★★

15199.05 TL

68P40EA

Hp Pavilion 14-Dv0029Nnt
11.Nesil Core i7 1165G7-8Gb-
512Gb Ssd-14-Mx450 2Gb-W10
HP

☆☆☆☆☆

18999.05 TL

68P30EA

Dell Vostro 14 5410 11.Nesil
Core i7 11370H-16Gb-512Gb
Ssd-14-Mx450 2Gb-W10pro
DELL

★★★★☆

25766.85 TL

Şekil 7. E-Commerce App Product List sayfası çoklu filtre

+ Yeni

Ürün Yönetimi

Marka	Ürün Kodu	Başlık	Resim	Fiyat	Puan	
HUAWEI	MBD15/I511/8/256	Huawei Matebook D15 11.Nesil Core i5 1135G7-8Gb-256Gb Ssd-15.6inc-W11		12073.95 TL	★★★★☆	
LENOVO	82H801AATX	Lenovo Ideapad 3 11.Nesil Core i3 1115G4-8Gb-256Gb Ssd-15.6inc-W11		11548.95 TL	☆☆☆☆☆	
ASUS	M1502IA-EJ129W	Asus Vivobook 15 M1502 4.Nesil Ryzen 5 4600H-8Gb-256Gb Ssd-15.6inc-W11		14016.45 TL	★★★★☆	
LENOVO	82H80252TX	Lenovo Ideapad 3 11.Nesil Core i3 1115G4-8Gb-256Gb Ssd-15.6inc-W11		8923.95 TL	★★★★☆	
HUAWEI	53012TTB	Huawei Matebook D15 11.Nesil Core i5 1135G7-8Gb-512Gb Ssd-15.6inc-W11		14488.95 TL	★★★★☆	

304 üründen 1 ile 5 arası gösteriliyor

1 2 3 4 5

Toplamda 304 ürün mevcut.

Şekil 8. E-Commerce App ürün yönetim sayfası