Pyhton Web Crawler Raporu

Osman Aslancan 170421842

Emirhan Çınar 171419002

Projemiz internet üzerindeki e-ticaret sitelerinden ürün fiyatlarını çekerek kullanıcıya grafik veya tablo şeklinde sunmaktadır. Şu anda hepsiburada.com ve trendyol.com sitelerine uyarlanmıştır fakat başka eklenecek e-ticaret siteleri içinde temel yapılmıştır.

Projemiz sitelerdeki bilgiyi çekebilmek için Scrapy kütüphanesini kullanmaktadır. Scrapy kütüphanesinde ideal olarak her bir site için ayrı bir spider açmamız gerekiyor bu yüzden bizde Hepsiburada ve Trendyol için 2 tane spider yazdık. Fakat bu spider’ların güncel tutulması gerekiyor çünkü eğer ilerde bu sitelerdeki bir değişiklik spider’i bozabilir güncellenmesi gerekir.

Projemizde çekilen bilgileri kullanıcıya göstermek için Pyqt5 kütüphanesini kullandık. Pencerelerimizi Qt Designer ile tasarladık. Pencereleri ui formatından py formatına Visual Studio Code’da bulunan “PYQT Integration” uzantısını kullanarak çevirdik. Projemizde 3 tane form ekranı tasarladık. Bunlar Ana Listeleme, Ürün Ekleme ve Sonuçları Görüntüleme ekranlarıdır.

Database için SQLExpress(MSSQL) kullandık. Database’i pythonda kullanmak için Pyqt5 kütüphanesi içerisinde bulunan QtSql modülünü kullandık. Bunu kullanmamızın nedenleri:

1. Projemize başka bir kütüphane daha ekleyip boyutunu artırmamak.
2. QtSql içerisinde yazılan sorguların direk tabloya yazdırılması özelliğinin (QSqlQueryModel) bulunması.

Ürünler için sitelerden çektiğimiz sonuçları kullanıcıya tablo şeklinde göstermekteyiz. Fakat kullanıcı isterse bu bilgileri grafik şeklinde de görebilir. Bu grafiği çizdirmek için matplotlib kütüphanesini kullanmaktayız.









