
Ömer Faruk Bulut

Hava Durumu Verileri

GENEL BAKIŞ

Bu projede, Türkiye'deki hava durumu bilgileri üç farklı web sitesinden toplanıp MongoDB veritabanına kaydedilmektedir. Kullanılan web siteleri şunlardır:

1. Met Office (<https://www.metoffice.gov.uk/weather/world/turkey>)
2. Weather.com (<https://weather.com/tr-TR/>)
3. HavaDurumuX (<https://www.havadurumux.net/>)

Projede Kullanılan Kütüphaneler

- requests
- BeautifulSoup
- pymongo

Projede Database İsmi

- omer_faruk_bulut

AŞAMALAR

Linklerin Toplanması

- İlk adım olarak, yukarıda belirtilen web sitelerinden Türkiye'deki şehirlerin hava durumu sayfalarının linkleri toplanıyor.
- Toplanan bu linkler, MongoDB'deki "links" adlı koleksiyona ekleniyor.
- Weather.com sitesi, şehirlerin posta kodlarını içeriyor. Bu posta kodları kullanılarak şehirlerin plaka numaraları elde ediliyor ve veritabanına kaydediliyor.
- Diğer sitelerde yer alan şehir isimleri, plaka numaraları ile eşleştirilerek bu bilgiler de veritabanına ekleniyor.

Hava Durumu Bilgilerinin Toplanması

Veri Toplama Süreci

- Projede, daha önce toplanan ve MongoDB de “links” koleksiyonunda saklanan linkler üzerinden işlem yapılıyor.
- Bu linkler kullanılarak, ilgili web sitelerinden bir haftalık hava durumu bilgileri toplanıyor.

Süreklilik ve Dayanıklılık Özelliği

- Projede önemli bir özellik, kodun kapatılması veya herhangi bir sorunla karşılaşılması durumunda, işlemlerin son durdukları yerden devam edebilmesidir.
- Bu özellik, links koleksiyonundaki her link için bir “last_activity” alanı kullanılarak sağlanıyor.
- Bir link üzerinden veri toplama işlemi tamamlandığında, bu linkin “last_activity” alanı güncelleniyor.
- Program tekrar başlatıldığında veya devam ettirildiğinde, “last_activity” alanı bugünün tarihinden eski olan linkler taranıyor.
- Bu yaklaşım, veri toplama sürecinin kesintisiz ve verimli bir şekilde devam etmesini sağlıyor ve olası hatalar veya kesintiler durumunda veri kaybını önlemeye yardımcı oluyor.
- Bu şekilde, hava durumu verilerinin toplanması ve işlenmesi süreci, hem verimli hem de kesintilere karşı dayanıklı bir şekilde yürütülmektedir.

Projede Oluşan Sorunlar

Türkiye'deki İllerle İlgili Sorunlar

- 81 İlin Tamamının Bulunmaması: Met Office web sitesinde Türkiye'nin 81 ilinin tamamının verisi mevcut değil.
- 79 Bölge Mevcut: Sitede Türkiye için sadece 79 bölge bulunmakta. Bu bölgeler içerisinde sadece şehir statüsündeki bölgeler için hava durumu verisi toplanıyor.

Alınan Yaklaşım ve Çözümler

- Projede, Met Office web sitesinin bu sınırlamaları göz önünde bulundurularak, mümkün olan en fazla verinin toplanması hedeflenmiş.
- Bu sorunlar, projenin kapsamı ve hedefleri doğrultusunda yönetilmiş, veri toplama sürecinde mevcut kaynakların en iyi şekilde kullanılmasına odaklanılmıştır.