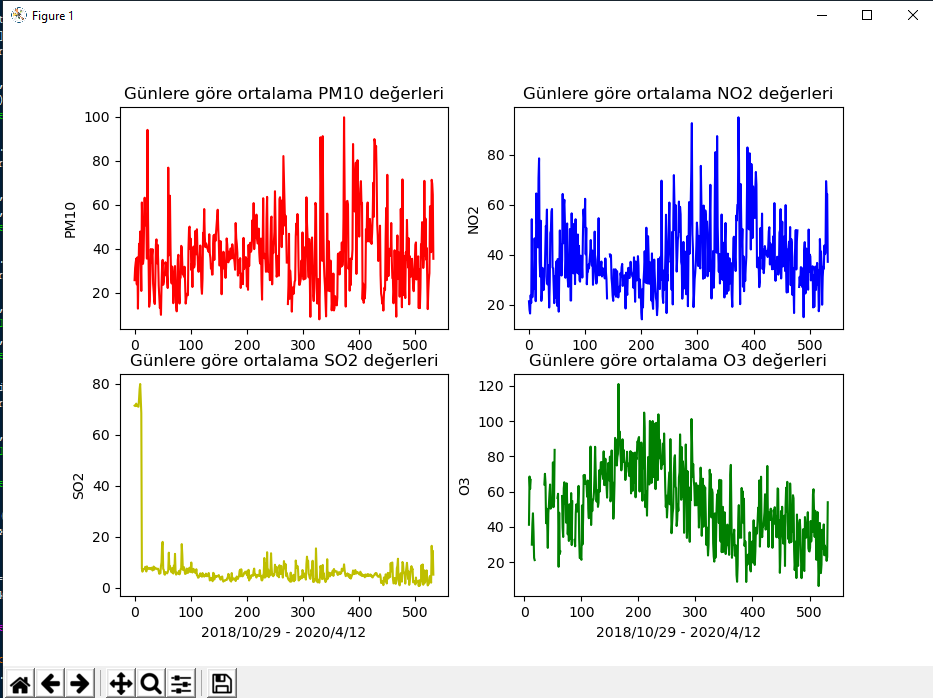
01.05.2020

YAZILIM MÜHENDİSLİĞİ GÜNCEL KONULAR

Çevre ve Şehircilik Bakanlığı’nın Ulusal Hava Kalite İzleme Ağı verilerine bakıldığında 4 adet parametre içeren Yalova ilinin verilerine dayanarak veri seti hazırlandı. Bu 4 parametre NO2, SO2, O3, PM10 şeklinde olmaktadır. Bu parametreler bölgenin hava kalitesi hakkında saat saat bilgi vermektedir. Bu bilgiler ışığında projenin hangi tarih ve saat aralığında hangi değerlere ulaşıldığı bilgisine bakılıp, ortalaması alınıp gün bazında projeye katkıda bulundu.

Projede <https://matplotlib.org/gallery/index.html> kaynağından grafikler projeye dâhil edildi. <https://havaizleme.gov.tr> alınan veriler “hava.csv“ isimli dosya aktarıldı. “hava.csv” isimli dosyadaki tarih verisinin formatı proje içerisinde tarihi sıralamak için formatı uygun olmadığından uygun hale getirildi ve “hava.csv” isimli dosya yeniden düzenlendi. Grafiklere aktarılmak üzere “hava.csv” isimli dosyadan veriler çekildi. Bu veriler saatlik olduğu için tarih verileri gruplandırılıp ortalama değerleri hesaplanarak günlük veriler elde edilmiş oldu. Bu veriler 4 farklı parametre içerdiğinde 4 farklı grafik çizerek veriler daha anlaşılır hale getirildi.

<http://www.havaizleme.gov.tr> ‘nin belirlemiş olduğu hava kalite index değer aralıklarına bakılarak “İyi, Orta, Hassas, Sağlıksız, Kötü, Tehlikeli” şeklinde sınıflandırma işlemi yapıldı. Veri setinde bu sınıflandırma işlemine ait bir sütun yer almadığı için “kalite” başlığına sahip yeni bir sütun oluşturuldu. “Kalite” sütununa, daha önceden gün bazında ortalaması alınmış PM10, SO2, NO2 ve O3 değerlerine bakılarak maksimum değer olan parametrenin hava kalite index değerinin karşılığı olan hava kalitesi eklendi. Eklenen bu değerin hangi parametreye ait olduğunu görmek ve üzerinde işlemler yapmak için “Secilen” isimli yeni bir sütun oluşturuldu. İş günleri ve tatil günleri olmak üzere 2 grubun hava kalitesine bakılarak bir çıkarım yapılması planlanmaktadır. Bu yüzden veri setinden gelen verilerin tatil günümü, yoksa iş günümü olduğunu anlamak için veri setinin ilk günkü verisini pazartesi olacak şekilde ayarlandı. Tatil ve iş günü hava kaliteleri bu şekilde toplandı. Toplanan bu verilere veri madenciliği uygulayarak Yalova ilinin hava kalitesinin tatil ve iş gününe göre değişikliği incelendi. Veri madenciliği için Gini algoritması kullanılmıştır.



**Şekil 1** 2018/10/29-2020/4/12 Tarihleri Arası Günlere Göre Ortalama PM10 Değerlerinin Kodu ve Grafiği



**Şekil 2** İş ve Tatil Günlerine Göre Veri Madenciliğinin Uygulanması ve Değerlerin Elde Edilmesi