

Proje Dokümanı

Hazırlayanlar

Gökalp Coşgun

Muhammed Berkay Şan

Proje kodu, “General_Models.ipynb” dosyasında görülebileceği üzere Machine Learning modellerinin preprocess edilmiş veri seti üzerinde denenmesi ile Rain, Fog ve Thunderstorm etiketlerinin her biri için ayrı olmak üzere en iyi modellerin belirlenmesi ve kaydedilmesi ile oluşmuştur. Exploratory Data Analysis ve raporda bahsedilen veya denenmiş ancak dahil edilmeyen tüm kodlar “Notebooks” klasörü içerisinde Jupyter dosyası olarak bulunabilir.

Deployment için “Scaler” ve modeller tüm veri üzerinde eğitilip, “Preprocessing/all_scaler.pkl” ve “Models/best*” olarak kaydedilmiştir. Bu kodlara “Preprocessing/Preprocessing.py” ve “Notebooks/General_Models.ipynb” içerisinden ulaşılabilir.

Klasör Açıklamaları :

- Ara Rapor : Projenin ara raporunu içerir.
- Data : Projede kullanılan veriyi içerir.
- Models : Elde edilen en iyi modelleri içerir.
- Notebooks : Deney yapılan kodları içerir.
- Preprocessing : Kaydedilen “Scaler”ı ve Preprocess.py(Modelde kullanılan tüm preprocessing adımları) içerir.
- SHAP : Raporda kullanmak üzere kaydedilen model açıklamalarını içerir.

Preprocessing adımları :

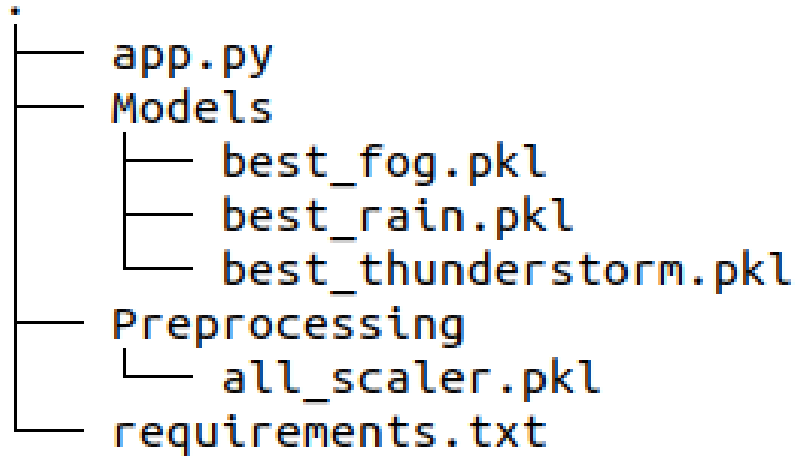
- “PrecipitationSumInches” özniteliğindeki T(Trace), çok az yağış anlamına geldiği için bu değerler 0 ile değiştirildi.
- Verinin geri kalanında “-” NaN yerine kullanıldığı için bu değerler NaN ile değiştirildi.
- NaN değerleri ortalama ile değiştirmek yerine “Date” özniteliği kullanılarak her NaN değeri o ayın ortalama değeri ile değiştirildi.
- “Date” özniteliğini makine öğrenmesi algoritmalarında kullanmayacağımız için bu özniteliği kaldırdık.
- MinMaxScaler ile veriyi 0-1 arasına scale ettik.

Deployment Adımları :

- General_Models.ipynb dosyasında bulunan ve kaydedilen en iyi modeller “Pickle” kütüphanesi kullanılarak yüklenir.
- Daha önceden tüm veri için fit edilmiş MinMaxScaler “Pickle” kütüphanesi kullanılarak yüklenir.
- Kullanıcıdan alınan değerler ile ekrana prediction ve probability bastırılır.

Projenin Ayağa Kaldırılması :

- Projenin çalıştırılması için figürde gösterilen dosya yapısı elde edilmelidir.



- Klasör içerisinde “streamlit run app.py” komutu çalıştırılarak ile proje ayağa kaldırılır.

Projenin Deploylanması:

Projeyi Streamlit cloud üzerinden deployladık. Yukarıdaki adımda gösterilen klasör yapısına sahip bir github reposuna streamlit erişimi verildiğinde projeyi public erişime açık bir şekilde yayınlıyoruz.