

RASSAL ORMAN ALGORİTMASI:

Rassal Orman (Random Forest) algoritması, birçok karar ağacını birlikte kullanan bir makine öğrenme yöntemidir. Hem sınıflandırma hem de regresyon problemlerinde kullanılır. Temel özellikleri şunlardır:

- **Topluluk Öğrenme:** Birden çok karar ağacının birleşimiyle çalışır.
- **Bagging:** Her bir ağaç, orijinal veri setinden rastgele seçilen örneklerle eğitilir.
- **Rastgele Özellik Seçimi:** Her ağaç, her düğümde rastgele seçilen özelliklerle eğitilir.
- **Tahminlerin Birleştirilmesi:** Sınıflandırmada oylama, regresyonda ortalama alınarak nihai sonuç elde edilir.
- **Aşırı Uydurmaya Karşı Dayanıklı:** Birden çok modelin birleşimi sayesinde aşırı uydurmaya karşı dirençlidir.

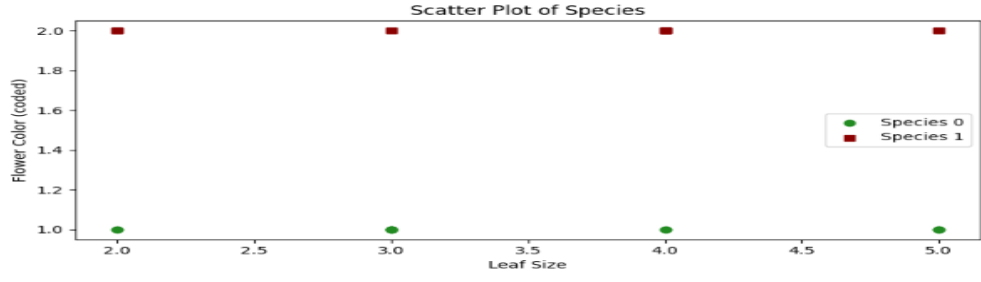
Avantajları arasında yüksek doğruluk, aşırı uydurmaya karşı dayanıklılık ve özellik önemini belirleyebilme yer alır. Dezavantajları ise yüksek hesaplama ve bellek gereksinimleri ile modelin yorumlanabilirliğinin düşük olmasıdır.

Noah, çeşitli bitki türleri ve bunların yaprak boyutu ve çiçek rengi gibi özellikleri hakkında veri toplayan bir botanikçidir. Noah, bu özelliklere dayanarak bir bitkinin türünü tahmin edip edemeyeceğini merak ediyor.

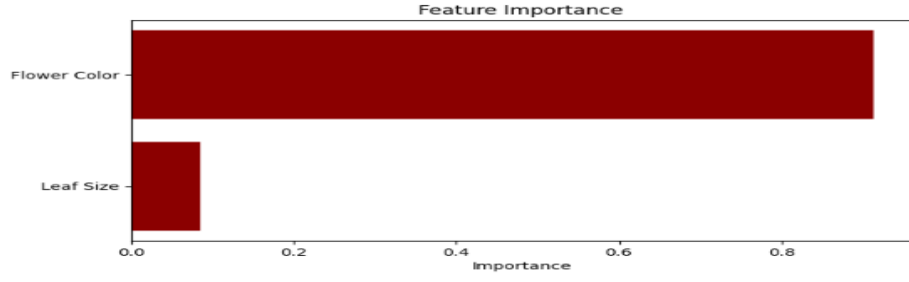
Burada, bitkileri sınıflandırmasına yardımcı olmak için bir topluluk öğrenme yöntemi olan Rastgele Orman'ı kullanacağız.

Teknik olarak, Rastgele Orman modelini kullanarak bitki türlerini belirli belirleyici değişkenlere göre sınıflandırmayı hedefliyoruz.

- **Örnek Veriler**`plants_featuresplants_species`: yaprak boyutunu ve çiçek rengini içerir, ilgili bitkinin türünü belirtir.
- **Train-Test Split:** Verileri eğitim ve test setlerine ayırıyoruz.
- **Model Oluşturma ve Eğitim**`RandomForestClassifier.fit()`: Belirlenen sayıda ağaçla (bu durumda 10) örnek oluşturuyoruz ve onu eğitim verilerimizi kullanarak eğitiyoruz.
- **Tahmin ve Değerlendirme:** Test verileri için türleri tahmin ediyoruz ve tahminleri kesinlik, hatırlama, fl puanı ve destek sağlayan bir sınıflandırma raporu kullanarak değerlendiriyoruz.
- **Özellik Önemlerinin Görselleştirilmesi:** Bitki türlerini tahmin etmede her özelliğin önemini göstermek için yatay bir çubuk grafik kullanıyoruz. Rastgele Orman, burada görselleştirdiğimiz ağaç oluşturma süreci sırasında özelliklerin kullanılabilirliğini ölçer.



[Resim Kaynağı: Yazar] Türlerin dağılım grafiği



[Resim Kaynağı: Yazar] Çiçek Rengi özelliğinin bitki türlerinin belirlenmesinde en büyük etkiye sahip olduğunu görebilirsiniz.