

Model Karşılaştırma ve En İyi Model Seçimi

Bu bölümde, eğittiğimiz üç farklı makine öğrenmesi modelinin (Lojistik Regresyon, Naive Bayes ve Karar Ağacı) performanslarını karşılaştıracak ve veri setimiz için en uygun olanı seçeceğiz. Karşılaştırmayı yaparken, önceki adımda elde ettiğimiz değerlendirme metriklerini kullanacağız.

Değerlendirme Metriklerine Genel Bakış

Modelleri karşılaştırmadan önce, kullandığımız temel metrikleri tekrar hatırlayalım:

- **Accuracy (Doğruluk):** Modelin tüm tahminleri içinde doğru yaptığı tahminlerin oranıdır. En genel başarı ölçütüdür, ancak sınıflar arasında dengesizlik varsa yanıltıcı olabilir.
- **Precision (Kesinlik):** Modelin belirli bir sınıf olarak tahmin ettiği örneklerden kaçının gerçekten o sınıfa ait olduğunu gösterir. Örneğin, “gazlı içecek” olarak tahmin edilen ürünlerin ne kadarının gerçekten gazlı içecek olduğunu ölçer. Yüksek kesinlik, modelin yanlış pozitif (yanlış alarm) oranının düşük olduğunu gösterir.
- **Recall (Duyarlılık/Geri Çağırma):** Gerçekte belirli bir sınıfa ait olan örneklerden kaçının model tarafından doğru bir şekilde o sınıf olarak tahmin edildiğini gösterir. Örneğin, gerçekte “gazlı içecek” olan ürünlerin ne kadarının model tarafından doğru tespit edildiğini ölçer. Yüksek duyarlılık, modelin yanlış negatif (gözden kaçırma) oranının düşük olduğunu gösterir.
- **F1-Skoru:** Kesinlik (Precision) ve Duyarlılığın (Recall) harmonik ortalamasıdır. Bu iki metrik arasında bir denge kurar ve özellikle sınıf dağılımları dengesiz olduğunda kullanışlıdır.
- **Weighted Avg (Ağırlıklı Ortalama):** Metriklerin (Precision, Recall, F1-Skoru) hesaplanırken her sınıfın örnek sayısına göre ağırlıklandırılmış ortalamasıdır. Bu, genel model performansını değerlendirirken her sınıfın katkısını dikkate alır ve dengesiz veri setleri için daha adil bir ölçüm sunar.

Modellerin Performans Metrikleri

Aşağıda, her modelin test verisi üzerinde elde ettiği temel metriklerin bir özeti bulunmaktadır:

1. Lojistik Regresyon (Logistic Regression):

- Accuracy: 0.9318
- Weighted Avg Precision: 0.94
- Weighted Avg Recall: 0.93
- Weighted Avg F1-score: 0.93
- *Not:* Model eğitimi sırasında yakınsama uyarısı (ConvergenceWarning) alındı. Bu, modelin optimum sonuca tam olarak ulaşamamış olabileceğini,

ancak yine de yüksek bir performans sergilediğini gösterir. Daha fazla iterasyon veya hiperparametre ayarı ile potansiyel olarak daha da iyileştirilebilir.

2. Naive Bayes (MultinomialNB):

- Accuracy: 0.8525
- Weighted Avg Precision: 0.86
- Weighted Avg Recall: 0.85
- Weighted Avg F1-score: 0.83

3. Karar Ağacı (Decision Tree):

- Accuracy: 0.8725
- Weighted Avg Precision: 0.89
- Weighted Avg Recall: 0.87
- Weighted Avg F1-score: 0.88

Karşılaştırma ve Sonuç

Metrikleri karşılaştırdığımızda aşağıdaki sonuçları görüyoruz:

- **Doğruluk (Accuracy):** Lojistik Regresyon (%93.18) en yüksek doğruluğa sahiptir. Onu Karar Ağacı (%87.25) ve Naive Bayes (%85.25) takip etmektedir.
- **Ağırlıklı Ortalama Metrikler:** Lojistik Regresyon, ağırlıklı ortalama Precision (%94), Recall (%93) ve F1-Skoru (%93) değerlerinde de diğer iki modelden belirgin şekilde daha iyi performans göstermiştir. Bu, modelin genel olarak sınıfları ayırt etmede daha başarılı olduğunu ve sınıf dengesi-zliğinden daha az etkilendiğini gösterir.
- **Naive Bayes ve Karar Ağacı:** Karar Ağacı, Naive Bayes'e göre biraz daha iyi bir performans sergilemiştir, ancak her ikisi de Lojistik Regresyon'un gerisinde kalmıştır.

En İyi Model Seçimi

Elde edilen sonuçlara göre, **Lojistik Regresyon modeli**, test edilen üç model arasında en yüksek performansı göstermiştir. Hem genel doğruluk oranı hem de ağırlıklı ortalama metrikler açısından diğer modellere üstünlük sağlamıştır. Yakınsama uyarısı alınmış olsa da, elde edilen %93'lük doğruluk ve F1 skoru oldukça başarılıdır ve bu modelin ürün isimlerinden kategori tahmini yapma görevinde en uygun model olduğunu göstermektedir.

Bu nedenle, bu analiz için en başarılı model olarak **Lojistik Regresyon** seçilmiştir.