

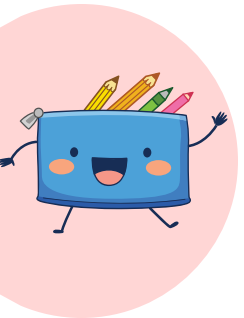
# Veri Bilimi

Common Voice kullanılarak cinsiyet-yaş-aksan analizi ve sınıflandırması



# Common Voice

Veri Seti Google Colab Ortamında Eğitilmiştir.  
Bu eğitim süreci boyunca kullandığı kaynak bilgisi;



## SİSTEM BİLGİLERİ (Google Colab)

- ◆ CPU Çekirdek Sayısı: 12
- ◆ Toplam RAM: 167.1 GB
- ◆ Kullanılabilir RAM: 164.4 GB
- ◆ Toplam Disk: 235.7 GB
- ◆ Boş Disk: 196.9 GB
- ◆ GPU: NVIDIA A100-SXM4-80GB, 81920 MiB

📌 Tahmin: Colab Pro+ (High-RAM)

Veri Seti aşağıdaki kolonlardan oluşmaktadır.

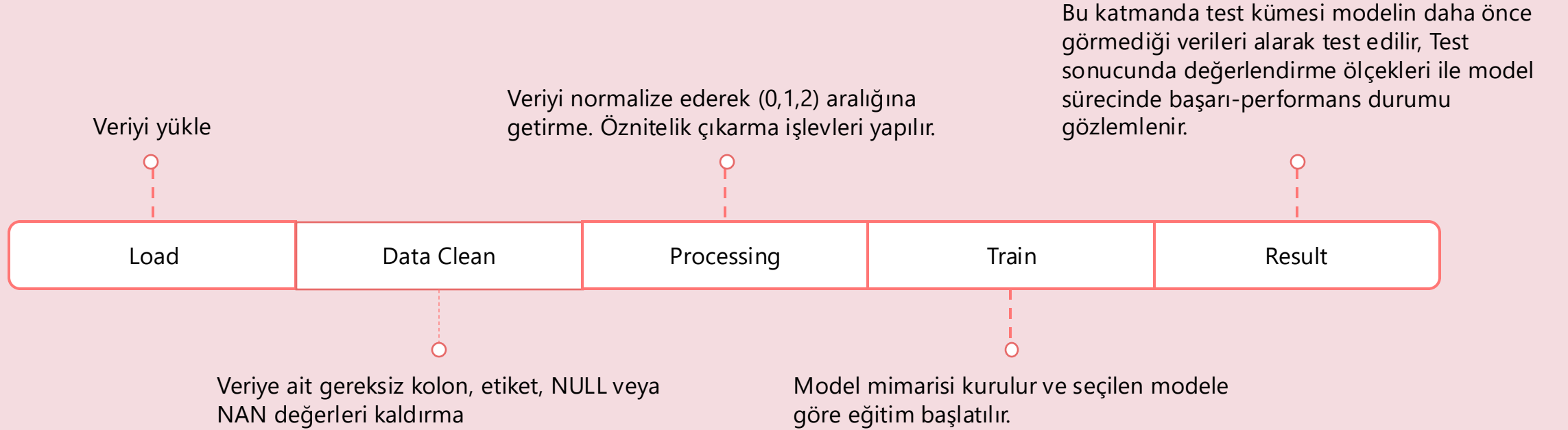
- Text
- Up\_votes
- Down\_votes
- Age
- Gender
- Accent
- Duration

# Verileri Yapılandırma(Preprocessing) Adımları

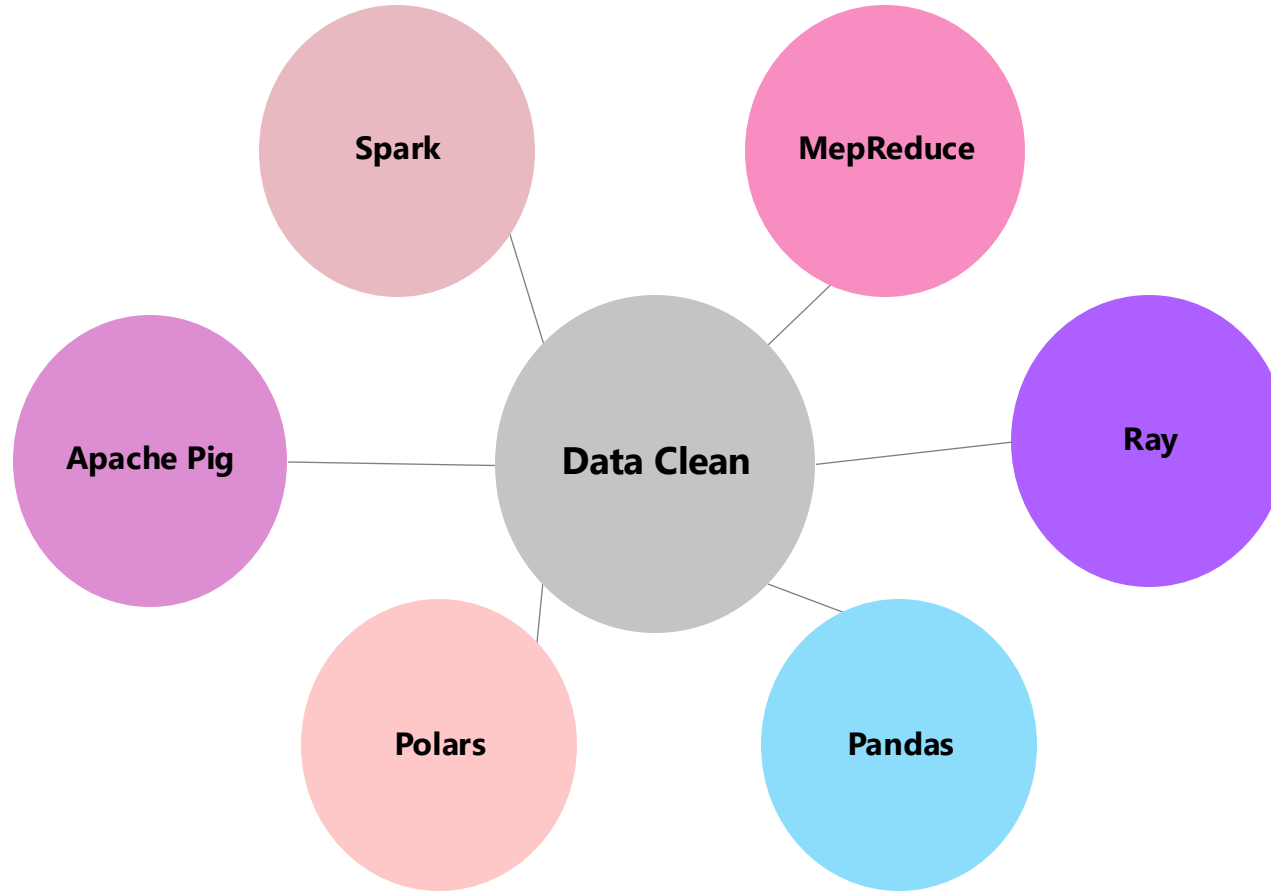
1. Veriyi Yükle
2. Gereksiz kolonları kaldır
3. İşlevsiz verileri kaldır
4. Veri Az ise alt etiketlerde topla



# Süreç



# Veri Analizi İçin Kullanılan Araçların Bazıları



# Feature Extraction - Öznitelik Çıkarımı

Mel-Frekans Kepstral Katsayıları Mel-Frekans Kepstral Katsayıları (Mel-Frequency Cepstral Coefficients) Mel frekans ölçeği, insan kulağının ses frekanslarındaki değişimi algılayışını gösteren bir ölçektir. MFCC (Mel-Frequency Cepstral Coefficients), ses sinyalinin kısa zamanlı güç spektrumunun Mel ölçeği üzerindeki ifadesidir.

Frekanstan Mel ölçeğine dönüşüm formülü;

$$M = 1125 \times \ln(1 + (f \div 700))$$

$M \rightarrow$  Mel ölçeği

$f \rightarrow$  Frekans(Hz)



# Spark ve Ray Kullanımı

Veri üzerinde Spark ve Ray ile veride yapılan işlemler

- Veri yüklendi ve analiz için head ile kontrol edildi
- Gereksiz kolonlar, NULL ve NAN değerleri kaldırıldı.
- Normalizasyon yapıldı
- StringIndexer veya Label Encoding uygulandı
- Öznitelikler çıkartıldı
- Spark'a özgü model ile eğitim başlatıldı
- Ray için Sklearn kullanıldı



# MODEL EĞİTİMİ

Model Eğitim aşamasında 2'li ve daha fazla sınıflarda eğitim performansı yüksek olan ve ses verileri için kullanılan modeller

## 1. SVM

## 2. Random Forest

### GENDER (CİNSİYET) SINIFLANDIRMA

📌 RandomForest eğitiliyor...

✅ RandomForest - Accuracy: 0.8590, Süre: 5.5s

📌 SVM RBF eğitiliyor...

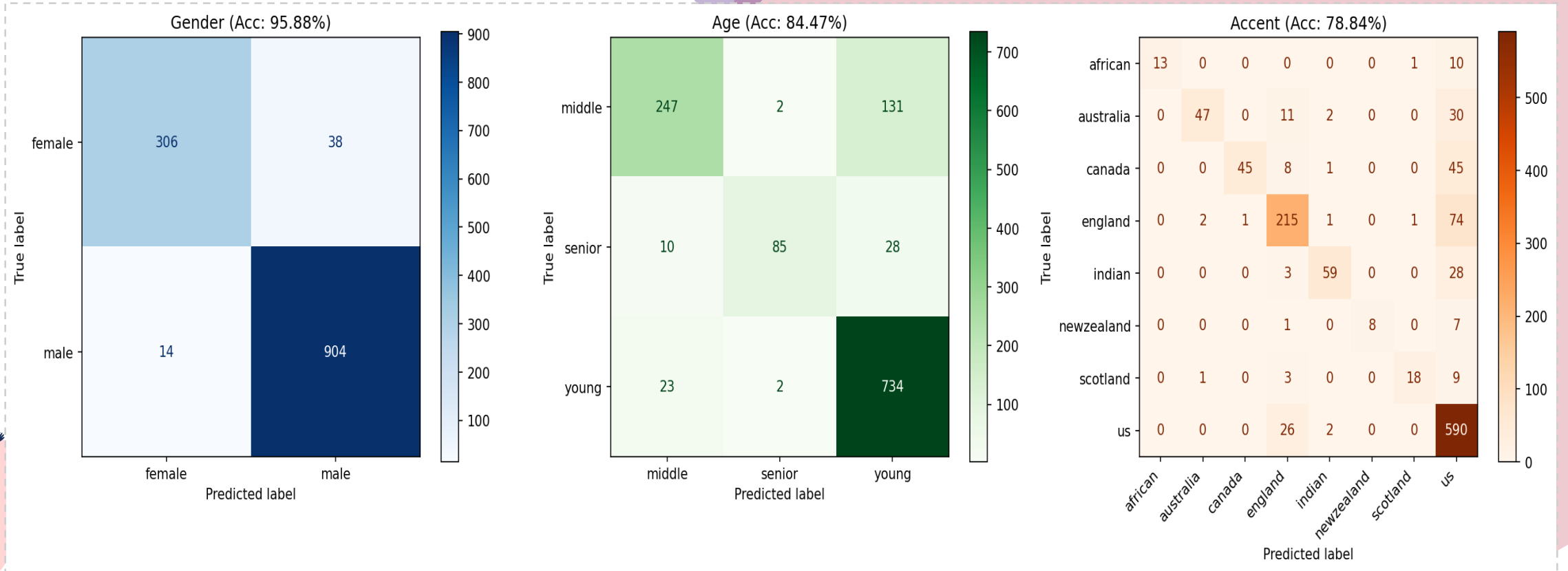
✅ SVM RBF - Accuracy: 0.9588, Süre: 102.6s

### 📊 SVM RBF Classification Report (Gender):

	precision	recall	f1-score	support
male	0.96	0.98	0.97	918
female	0.96	0.89	0.92	344
accuracy			0.96	1262
macro avg	0.96	0.94	0.95	1262
weighted avg	0.96	0.96	0.96	1262



# Result



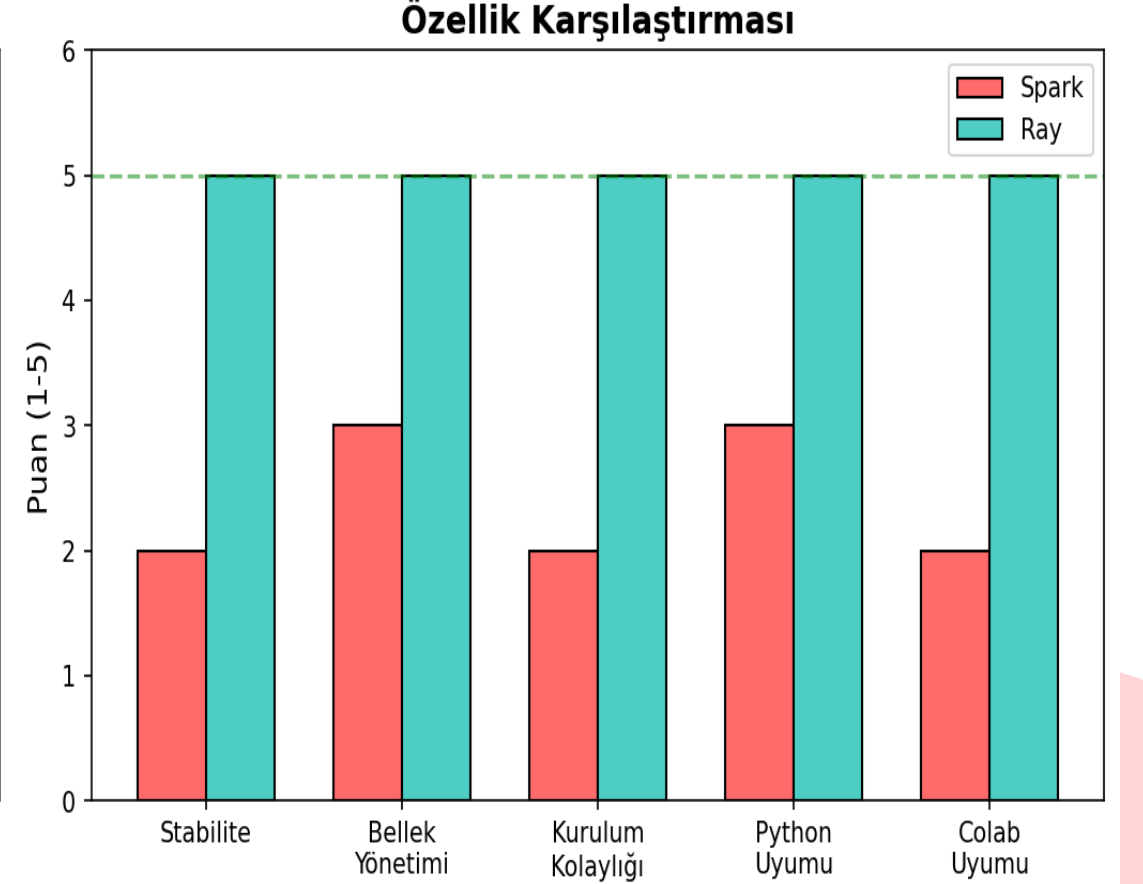
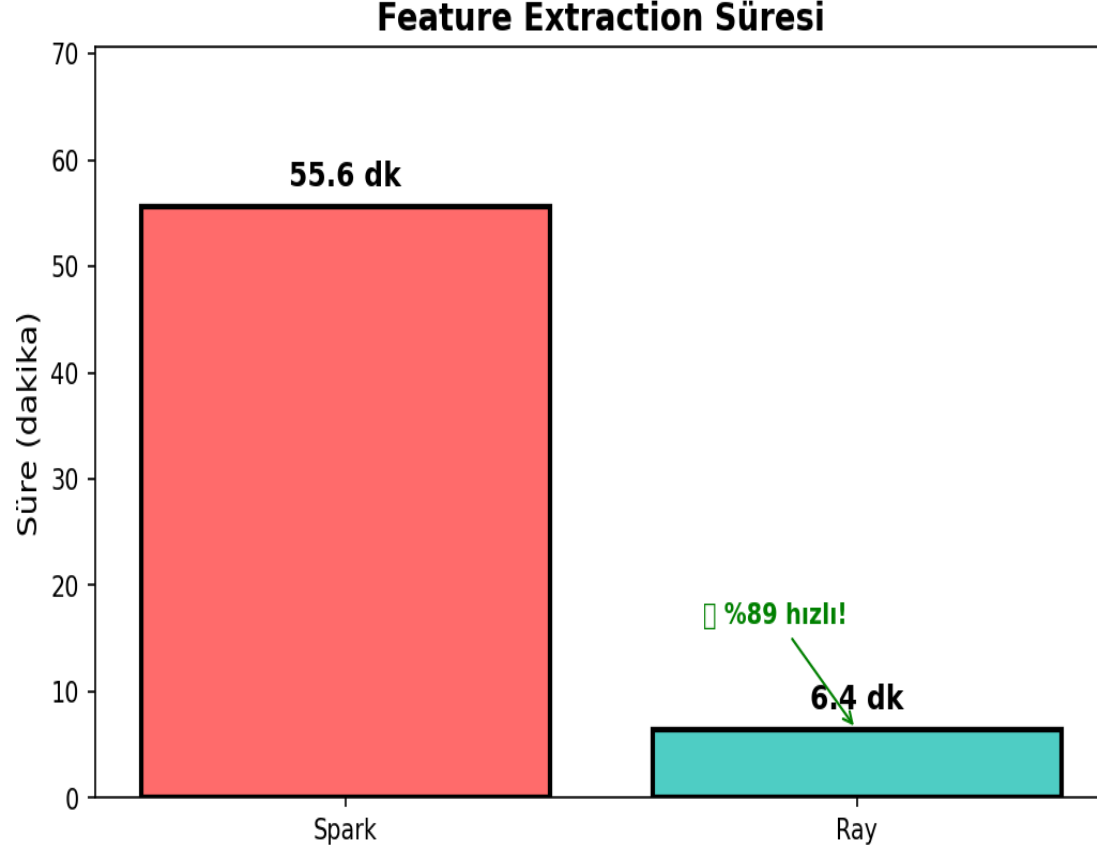
Veriye ait her bir sınıfın bireysel performansı

# Her Sınıfı Ortak Bir Çıktıda Birleştirme

- Veriye ait modellerde birden fazla etiket olduğu için hesaplama yapılırken her sınıfın doğruluk oranı birlikte doğru olma ve sonucundan ortak bir ağırlıklı performans değeri hesaplanır
- Multi-label Accuracy (Gender-Age-Accent):
  - 3 ayrı model ait etiketlerin hepsinin doğru olma ihtimali atanmış olur
  - True True True ortak ağırlık performansı alınır.

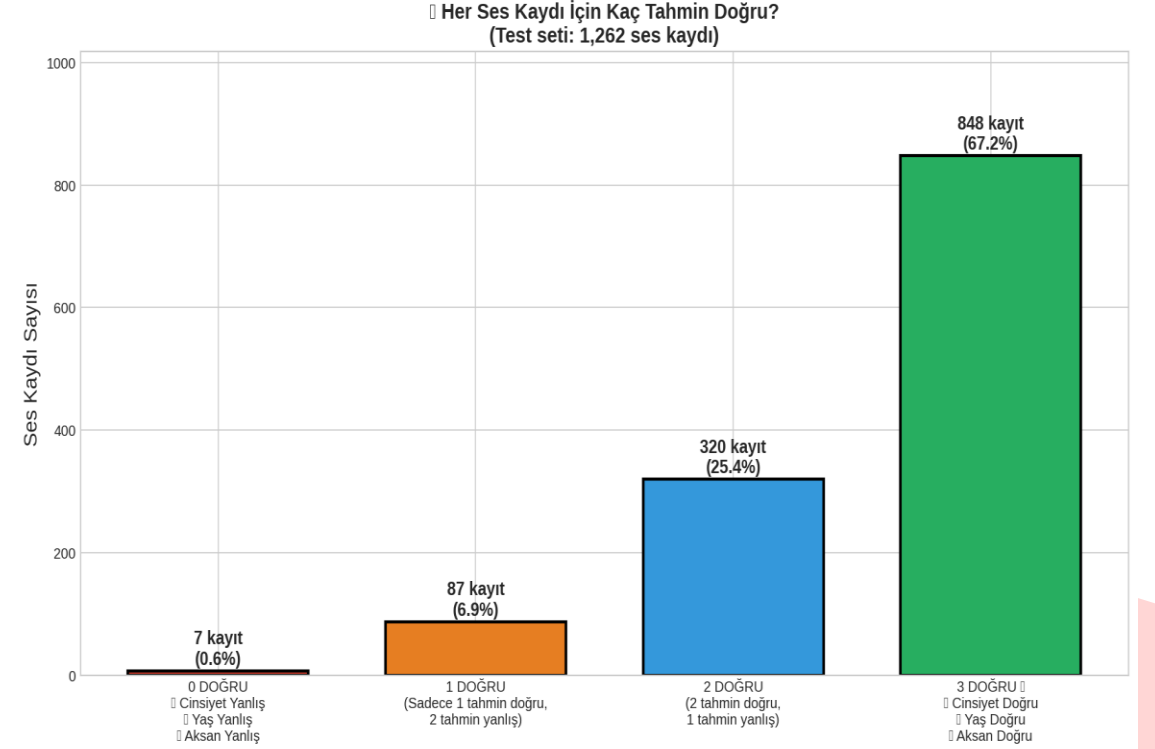
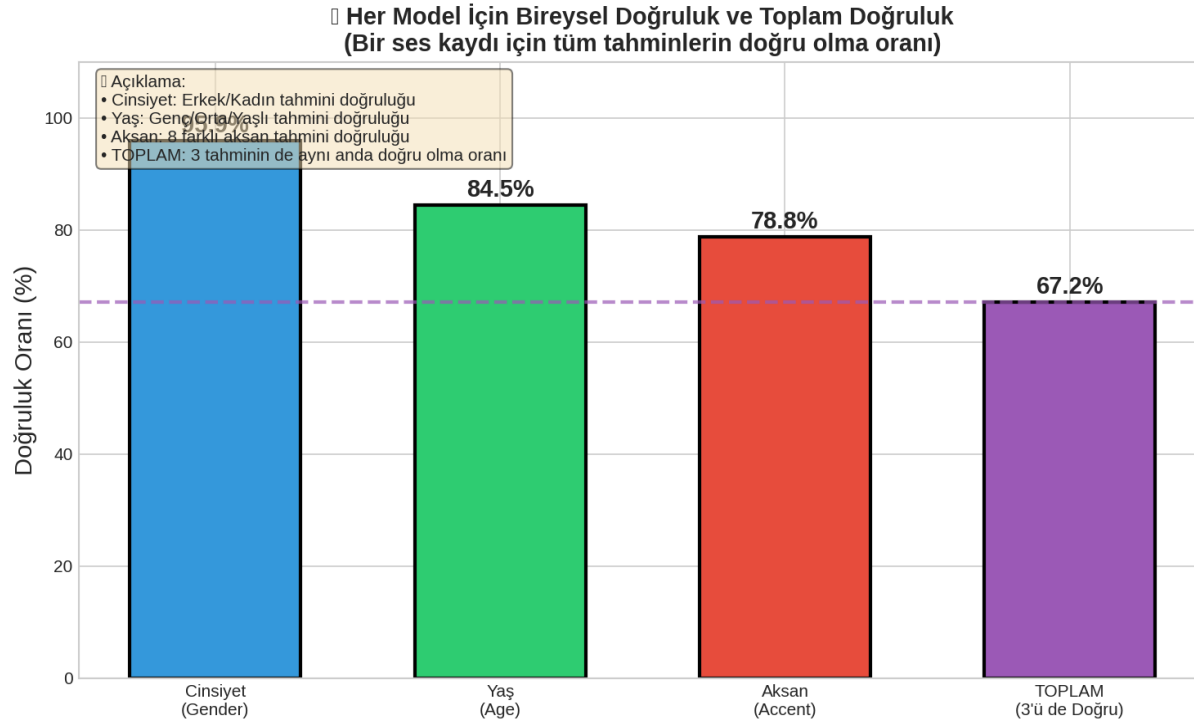
# Benchmark Göre Performans Analizi

Spark vs Ray Performans Karşılaştırması (Colab Pro+ (High-RAM), 12 CPU, 167GB RAM)



Tablodan da anlaşılacağı üzere ray hem veri seti üzerinden hem de genel kabul olarak daha performanslı olmuştur.

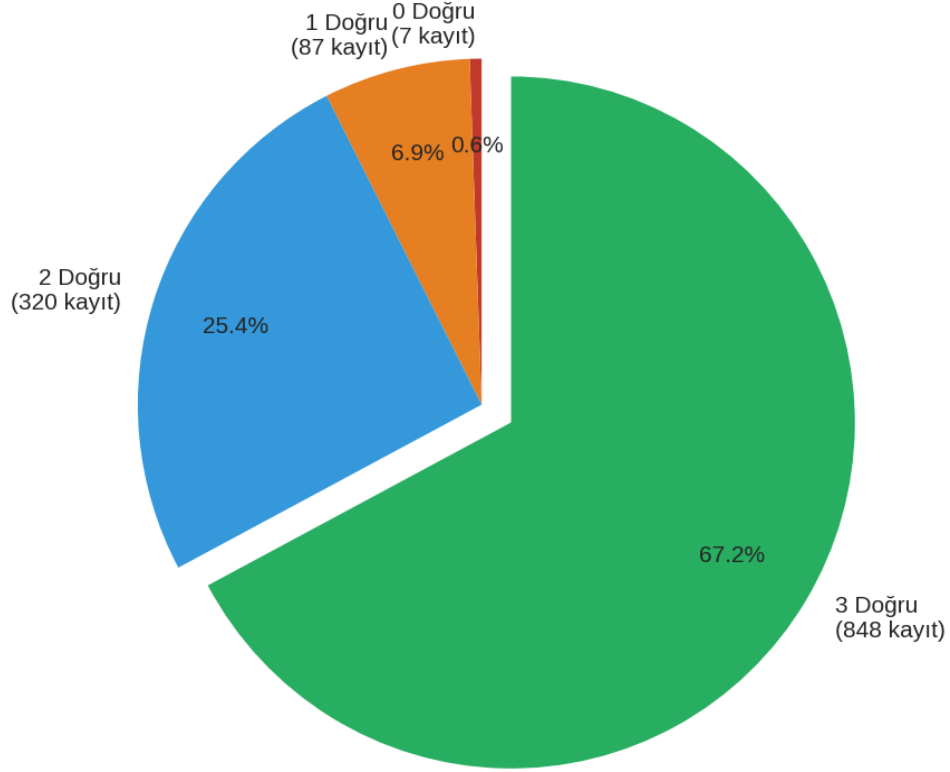
# Modelin Bireysel ve Toplam Doğruluk Grafikleri



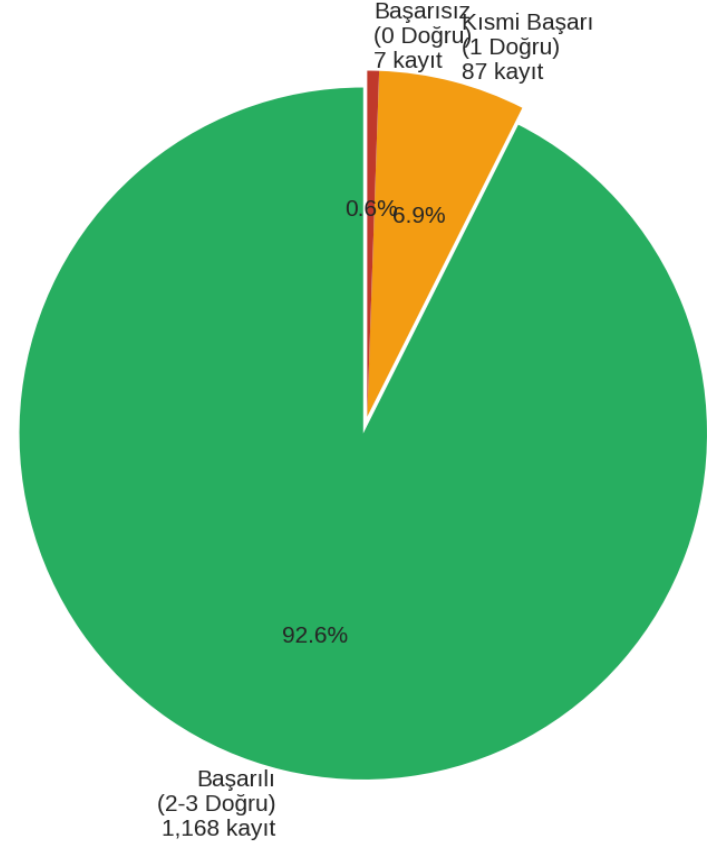
# Modelin Bireysel ve Toplam Doğruluk Grafikleri

## Multi-Label Performans Özeti (Test: 1,262 ses kaydı)

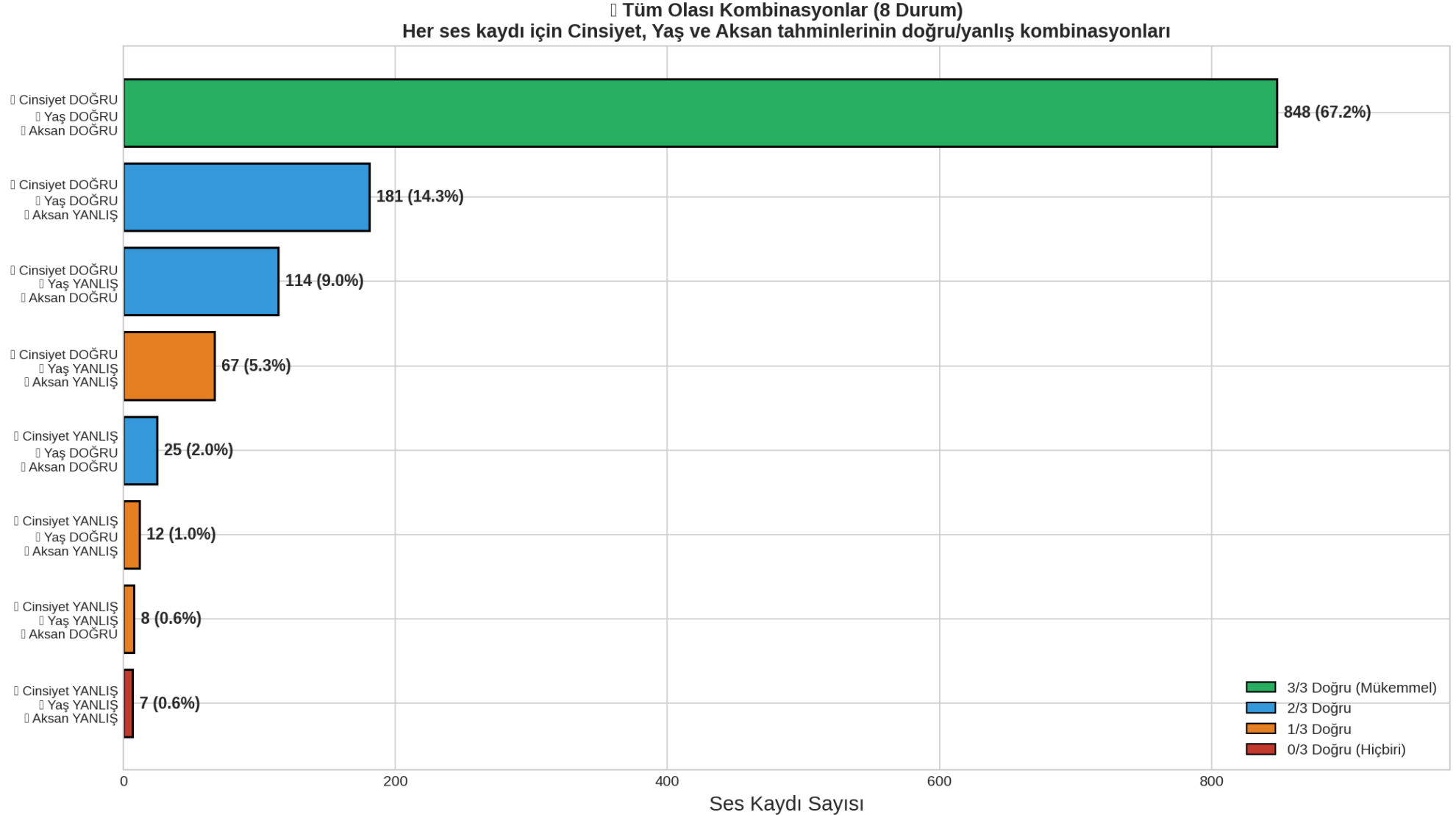
Doğru Tahmin Sayısına Göre Dağılım



Genel Başarı Durumu



# Modelin Bireysel ve Toplam Doğruluk Grafikleri



# Model Test

## Modeli Test Etmek

Huggenface üzerinden yayınlanmış hali üzerinden test edebilirsiniz.

<https://huggingface.co/spaces/elifksz/common-voice-demo>