Ses Analizi Modelleri Teknik Raporu

**1. Giriş**

Bu teknik rapor, ses dosyalarının duygu, cinsiyet ve yaş gibi özelliklerinin tahmin edilmesi amacıyla geliştirilen bir masaüstü uygulaması olan Ses Tahmini Arayüzü'nün tasarımı ve uygulanmasını detaylandırmaktadır.

**2. Proje Amaçları**

Ses dosyalarından duygu, cinsiyet ve yaş gibi özelliklerin tahmin edilmesini sağlayan bir arayüz tasarlamak.

Kullanıcıların ses dosyalarını yükleyebileceği, analiz edebileceği ve sonuçları görebileceği bir kullanıcı deneyimi sunmak.

Tkinter ve Python kullanarak basit, kullanıcı dostu bir arayüz oluşturmak.

**3. Proje Aşamaları**

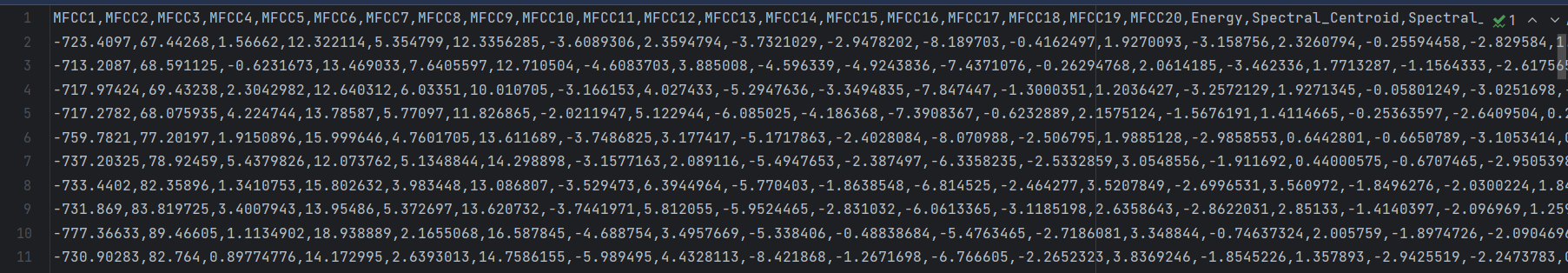
**3.1 Veri Hazırlığı**Bu projede, ses kayıtlarından duygu, cinsiyet ve yaş tahmini yapmak amacıyla bir dizi veri işleme adımı uygulanmıştır. Veri hazırlığı, verilerin temizlenmesi, özelliklerin seçilmesi ve eğitim/test setlerine bölünmesini içermektedir.

Kullanılacak veri setleri (örneğin, RAVDESS veya BVC Voice) indirilir ve işlenir.

Ses dosyalarından MFCC (Mel-Frequency Cepstral Coefficients) gibi ses özellikleri çıkarılır.

**3.1.1 Veriyi Yükleme:**

Veri seti, CSV formatında saklanan bir dosyadan yüklenmiştir.

Bu veri seti, çeşitli ses kayıtlarının özelliklerini ve bu kayıtların etiketlerini içermektedir.

**3.1.2 Özelliklerin Seçimi ve Hedef Değişken:**

Projedeki bazı özellikler, modeller için gerekli görülmediğinden çıkarılmıştır. Bu işlem, veri setini daha yönetilebilir hale getirirken, modelin eğitim süresini de optimize etmeye yardımcı olur.

**Çıkarılacak Gürültülü VerilerSütunlar:**

Energy

Spectral\_Centroid

Spectral\_Bandwidth

Spectral\_Rolloff

Zero\_Crossing\_Rate

Chroma\_STFT

**Hedef Değişken:**Emotion (Duygu)

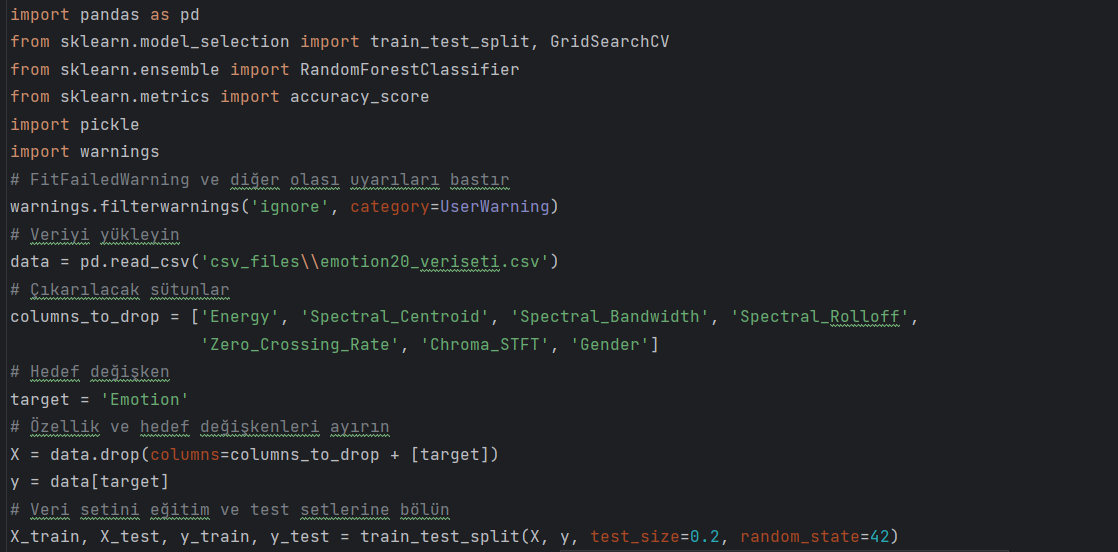
**3.1.3** **Eğitim ve Test Setlerine Bölme:**

Veri seti, modelin performansını değerlendirmek için eğitim ve test setlerine bölünmüştür. Eğitim seti modelin öğrenmesi için kullanılırken, test seti modelin doğruluğunu ve genelleme kabiliyetini değerlendirmek için kullanılmıştır.

**Eğitim ve Test Seti Oranı:**

**Eğitim seti: %80**

**Test seti: %20**

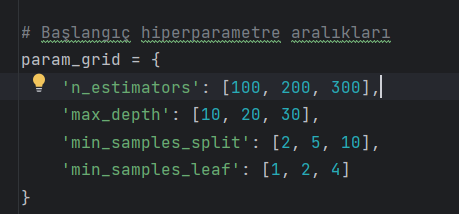


**3.2 Hiper parametre ayarlanması ve model eğitimi:**Bu algoritma, iteratif olarak en iyi hiperparametreleri bulmak ve model performansını artırmak amacıyla kullanılır. Random Forest modeline Grid Search uygulayarak, her iterasyonda hiperparametre aralıklarını daraltır ve modelin doğruluğunu artırmayı hedefler.

**3.2.1Model Tanımlama ve Başlatma:**

**Random Forest Modeli:** İlk olarak, Random Forest sınıflandırıcı modeli tanımlanır.

**Başlangıç Parametreleri:** Grid Search'ün başlangıç hiperparametreleri belirlenir.

****

**3.2.2 Grid Search ile Hiperparametre Optimizasyonu:**

Grid Search Uygulama: Grid Search ile tanımlanan hiperparametre aralıkları içinde en iyi kombinasyonları bulmak için model eğitilir ve değerlendirilir.

En İyi Model ve Parametreleri: Grid Search sonucunda en iyi model ve hiperparametre kombinasyonları belirlenir.

**3.2.3 Performans Değerlendirmesi:**

**Tahmin ve Doğruluk:** En iyi model, test seti üzerinde tahmin yapar ve doğruluk skoru hesaplanır.

Sonuçları Kaydetme: Mevcut iterasyonun hiperparametreleri ve doğruluk skoru kaydedilir.

**3.2.4 Hiperparametre Aralıklarını Daraltma:**

Başarı Kontrolü: Eğer mevcut iterasyonda elde edilen doğruluk skoru önceki en iyi skoru geçerse, en iyi skor ve model güncellenir.

Yeni Hiperparametre Aralıkları: Mevcut en iyi hiperparametrelerin etrafında yeni, daha dar aralıklar belirlenir.

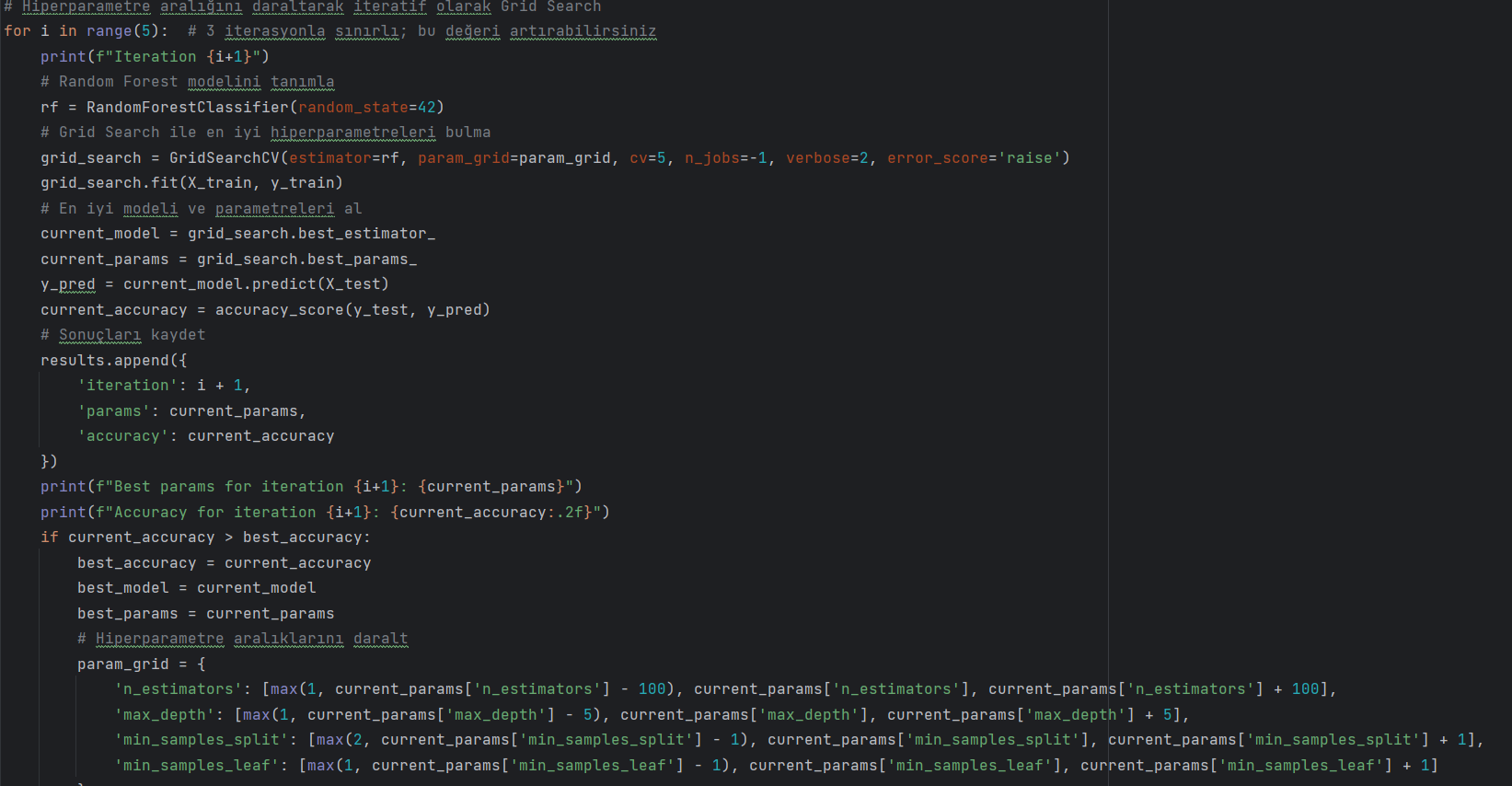
**3.2.5 İteratif İşlem:**

İterasyon Sayısı: Bu süreç belirli sayıda iterasyon (örneğin 5) boyunca tekrarlanır. Eğer bir iterasyonda gelişme olmazsa, işlem durdurulur.

**3.2.6 Sonuç:**

**En İyi Model ve Parametreler**: Sonunda en iyi hiperparametreler ve model belirlenmiş olur.

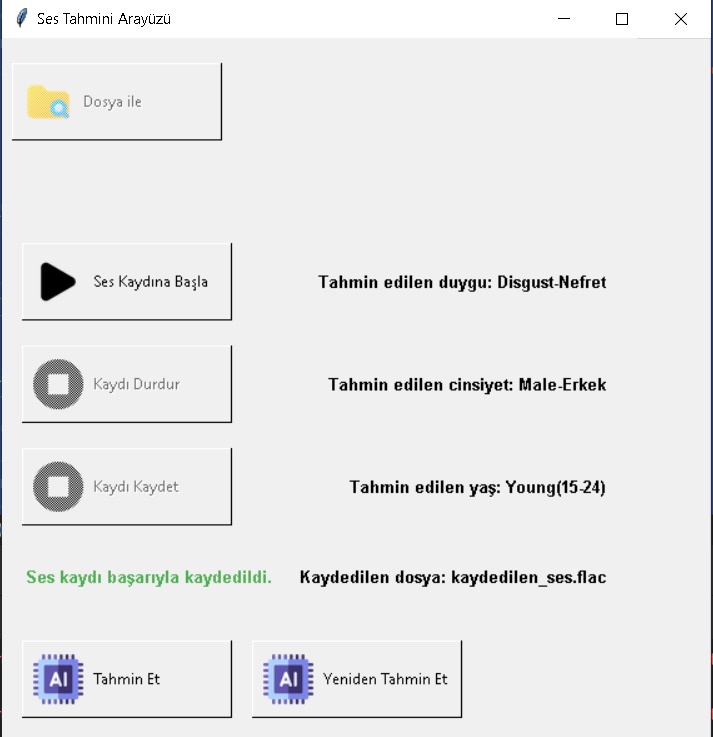
Performans: Modelin doğruluk skoru ve en iyi hiperparametre kombinasyonu raporlanır.

Bu algoritma, hiperparametre aralıklarını her iterasyonda daraltarak, modelin doğruluğunu optimize etmeyi ve en iyi performansı elde etmeyi amaçlar. Bu yöntem, modelin eğitim süresini ve kaynak kullanımını verimli bir şekilde yönetmeye yardımcı olur.  


**4.Arayüz Tasarımı**

Tkinter kullanılarak kullanıcı arayüzü tasarlanır.

Ses dosyası seçme, kaydetme, kayıt başlatma/durdurma, tahmin etme gibi işlemleri yöneten düğmeler ve etiketler eklenir.



**4.1 Arayüz Bileşenleri**

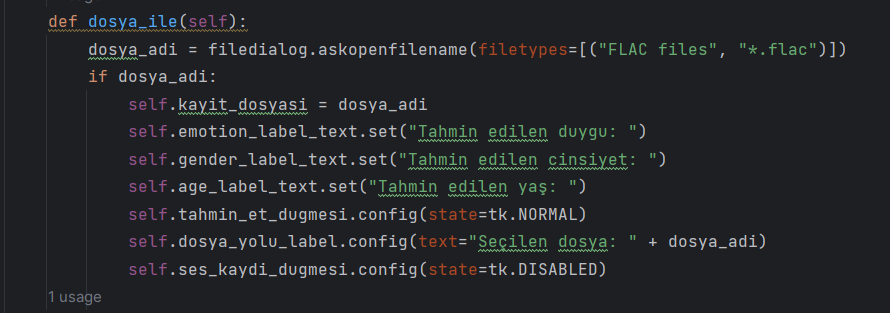
**Dosya Seçme Düğmesi:** Kullanıcının ses dosyasını seçmesine olanak tanır.

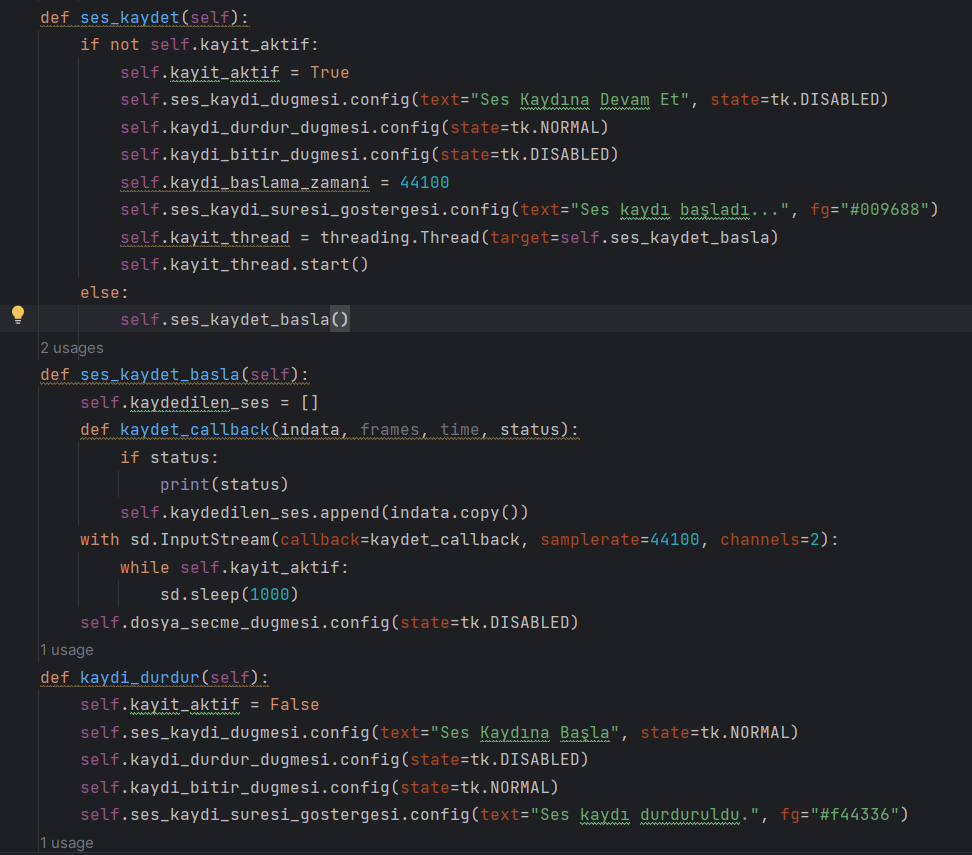
**Kayıt Düğmeleri:** Ses kaydını başlatma, durdurma ve kaydetme işlemlerini yönetir.

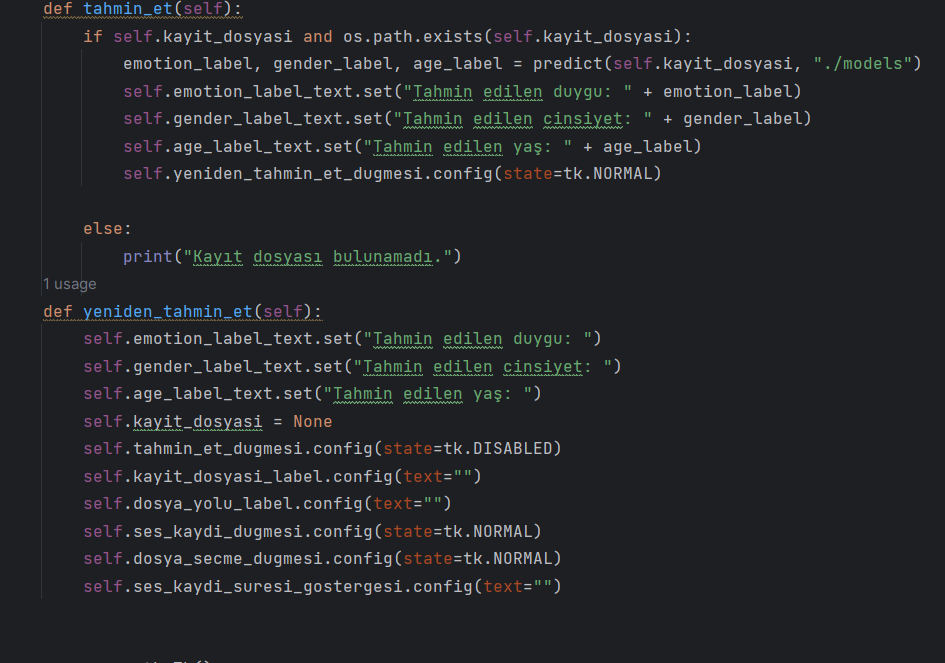
**Tahmin Et Düğmesi:** Seçilen ses dosyasının analiz edilerek sonuçların gösterilmesini sağlar.

**Sonuç Etiketleri:** Tahmin edilen duygu, cinsiyet ve yaş gibi sonuçları gösterir.

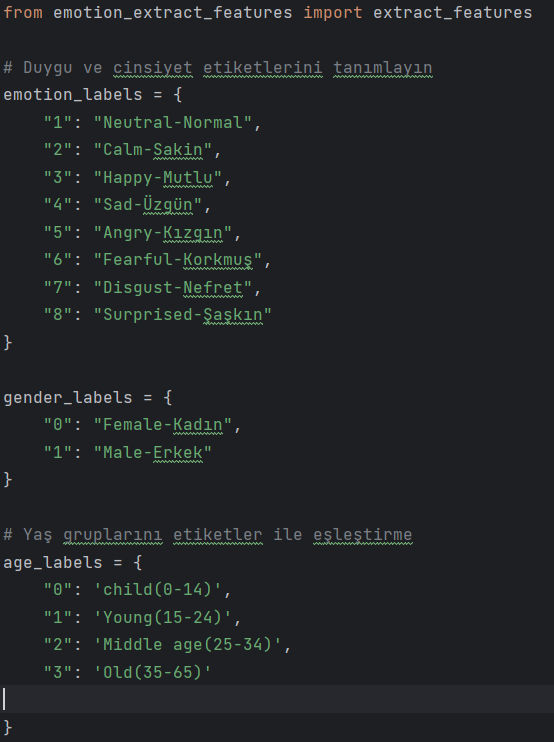
**Dosya Seçme Düğmesi:** Kullanıcının ses dosyasını seçmesine olanak tanır.



**Kayıt Düğmeleri:** Ses kaydını başlatma, durdurma ve kaydetme işlemlerini yönetir.  


**Tahmin Et Düğmesi:** Seçilen ses dosyasının analiz edilerek sonuçların gösterilmesini sağlar.  


**Sonuç Etiketleri:** Tahmin edilen duygu, cinsiyet ve yaş gibi sonuçları gösterir.

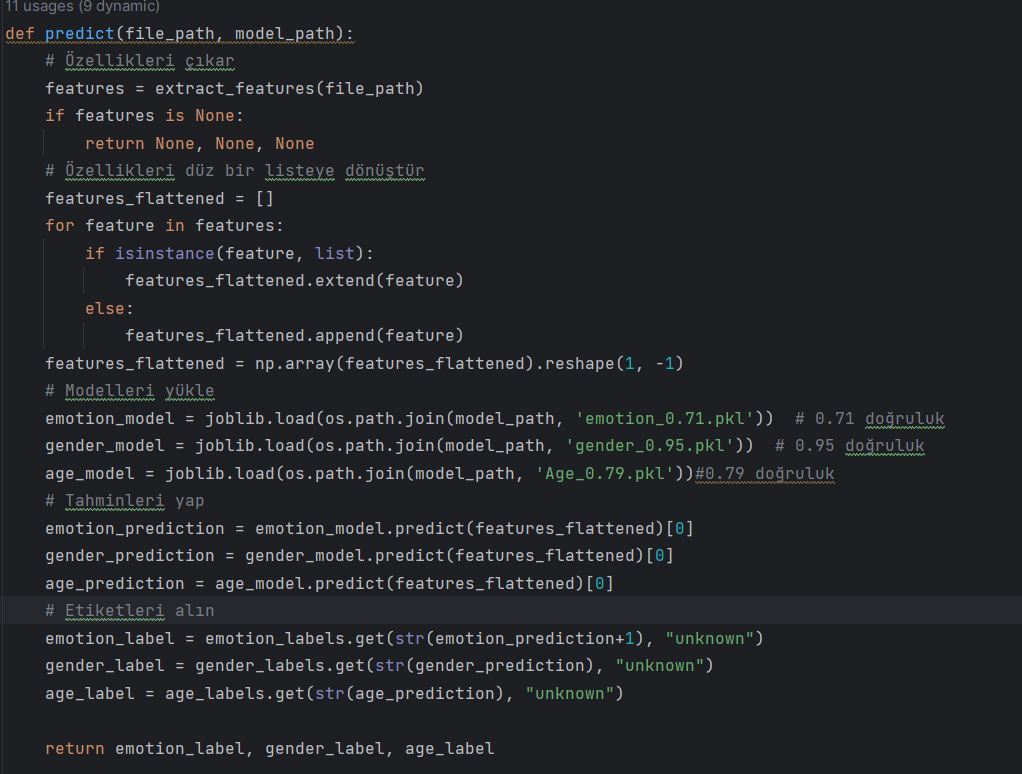


**5. Teknik Detaylar**

**5.1 Back-End İşlemleri**

Ses kaydı sırasında sounddevice kütüphanesi kullanılarak ses verileri toplanır.

Ses dosyaları soundfile kullanılarak kaydedilir.

Model tahminleri main.py dosyasından alınarak arayüze gösterilir. ****

**6. Sonuçlar ve Gelecek Çalışmalar**

Ses Tahmini Arayüzü, kullanıcıların ses dosyalarını kolayca analiz etmelerine olanak tanır.

Gelecekte, arayüzün kullanıcı dostuğunu artırmak ve daha fazla analiz seçeneği eklemek için çalışmalar yapılabilir.

**6.1** **Başarı Oranları:**

**Duygu Tahmini:**

Eğitim ve test veri setleri üzerinde %71 doğruluk oranına ulaşıldı.

**Cinsiyet Tahmini:**

Cinsiyet tahmin modeli, test veri setinde %95 doğruluk oranına ulaştı.

Kadın ve erkek sesleri arasındaki farklar model tarafından başarılı bir şekilde ayrıştırıldı.

**Yaş Tahmini:**

Yaş tahmin modeli, %79 doğruluk oranıyla 10 yıllık yaş gruplarını doğru tahmin edebildi.

Yaş tahmininde en zorlayıcı sınıf çocuk sınıfı oldu.

**7. Referanslar**

RAVDESS Dataset

BVC Voice Dataset