

Bugünkü çalışmamda, projenin gerçek zamanlı performansını artırmak amacıyla önemli adımlar attım. İlk olarak Google Colab üzerinde, GPU desteğinden faydalanarak bir prototip geliştirdim. Bu prototip, mikrofona verisini eş zamanlı olarak işleyerek anlık transkripti ekrana yazdırıyor ve konuşma bittiğinde tüm ses kaydını tek parça halinde işleyerek nihai ve daha doğru bir metin çıktısı üretiyordu. Elde edilen bu çıktıyı, sistemin diğer bileşenleriyle uyumlu olması için JSON formatına dönüştürdüm. Colab üzerinde çalışan bu servisi test edebilmek ve dışarıdan erişilebilir bir arayüze taşımak için

ngrok aracını kullandım. Ngrok, yerel makinede (bu durumda Colab ortamında) çalışan bir servisi, internet üzerinden erişilebilir geçici bir genel URL adresi oluşturarak dış dünyaya açan bir araçtır. Bu sayede, henüz sunucuya kurulmamış uygulamaları kolayca test etme ve sunma imkânı buldum.

Hazırladığım bu tasarımı sorumlu mühendisime sunduğumda, bir önceki görüşmemizde konuştuğumuz gibi sistemin anlık çıktı verebilmesi için WebSocket protokolüne geçmemin daha doğru olacağını belirtti. Bunun üzerine WebSocket teknolojisini detaylıca araştırmaya başladım.

WebSocket protokolü, web tarayıcısı (istemci) ile sunucu arasında sürekli, çift yönlü (full-duplex) ve gerçek zamanlı bir iletişim kanalı kuran bir teknolojidir. Bu yapı, sürekli istek-cevap döngüsüne dayanan HTTP protokolünün aksine, veri akışını kesintisiz hale getirerek gecikmeyi en aza indirir.

Yaptığım araştırmalar sonucunda, mevcut mimarimde kullandığım Flask ve Waitress-serve kombinasyonunun WebSocket ile tam uyumlu çalışmadığını öğrendim. Bu nedenle, asenkron yapısı sayesinde WebSocket desteğini doğal olarak sunan

FastAPI kütüphanesini kullanmaya karar verdim. Servisimi FastAPI tabanlı olarak yeniden yazdım ve WebSocket bağlantısını başarıyla kurdum. Front-end tarafından ses verisinin blob formatında geleceğini öğrendiğim için, kodumda bu formattaki veriyi işleyecek gerekli dönüşümleri de ekledim.

Günün sonunda, FastAPI ve WebSocket tabanlı yeni mimarim ilk haliyle çalışır duruma geldi. Ancak şu anki haliyle sistemde bazı eksiklikler ve hatalar mevcut. Ayrıca sesin parça parça işlenmesi, Colab'daki GPU destekli prototip kadar verimli ve hızlı olmadı.

ÖZET => Colab üzerinde ngrok ile test edilen bir prototip geliştirildi, ardından mühendis geri bildirimiyle WebSocket'e geçiş için FastAPI tabanlı yeni bir servis yazıldı ve front-end'den gelen blob verisi işlendi.