

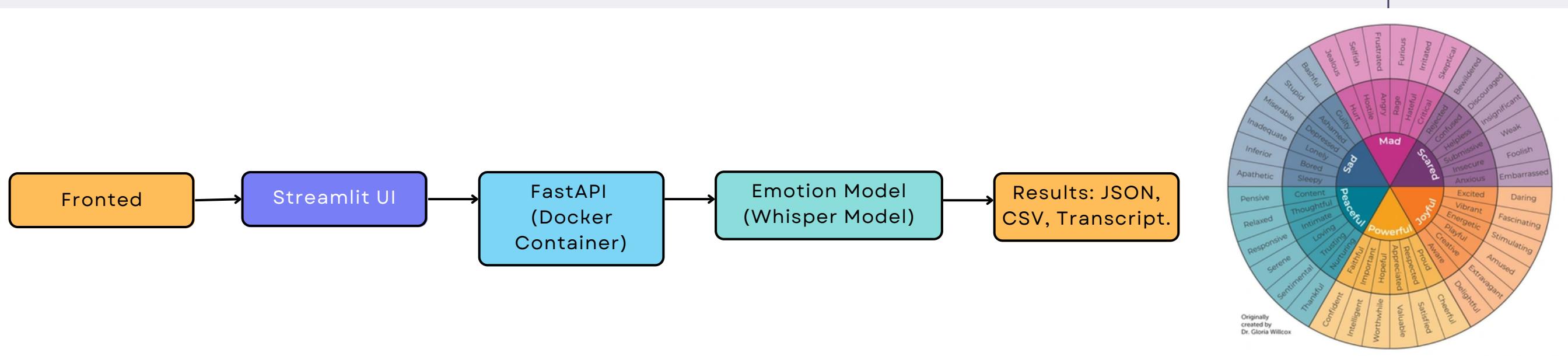
Emotion Recognition API with Dockerized Deployment & Frontend

Authors

Alex Mic, Anastasiia Mokhonko, Artjom Musaelans, Maria Salop, Nikita Orlovs
Block D, Year 2, 2025

Introduction

A scalable API for emotion detection and audio transcription, using FastAPI, Transformers, Whisper, and Docker. Batch and real-time predictions, with an additional Streamlit frontend for interactive demos. Production practices applied: CI, unit testing, logging, stress testing.



Key Features

- Text & Batch Emotion Prediction: via /predict & /upload-csv
 - YouTube & Audio Transcription: audio-to-text, then batch prediction
 - Dockerized for Easy Deployment: runs anywhere, “works on my machine”
 - Logging & Unit Tests: robust and maintainable code
 - Streamlit Frontend: for simple web-based demo (built, attempted Azure deploy)

Technologies Used

Python
FastAPI
Hugging Face Transformers
Whisper
Docker
Streamlit
Locust
Pandas

Evaluation: Performance & Stress

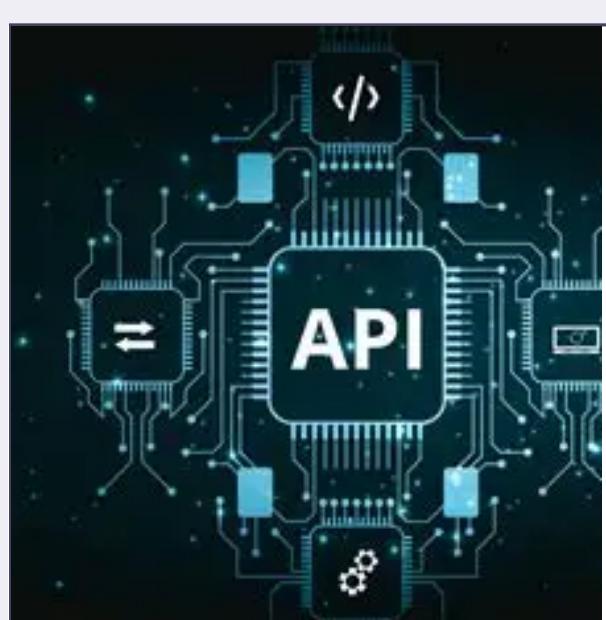
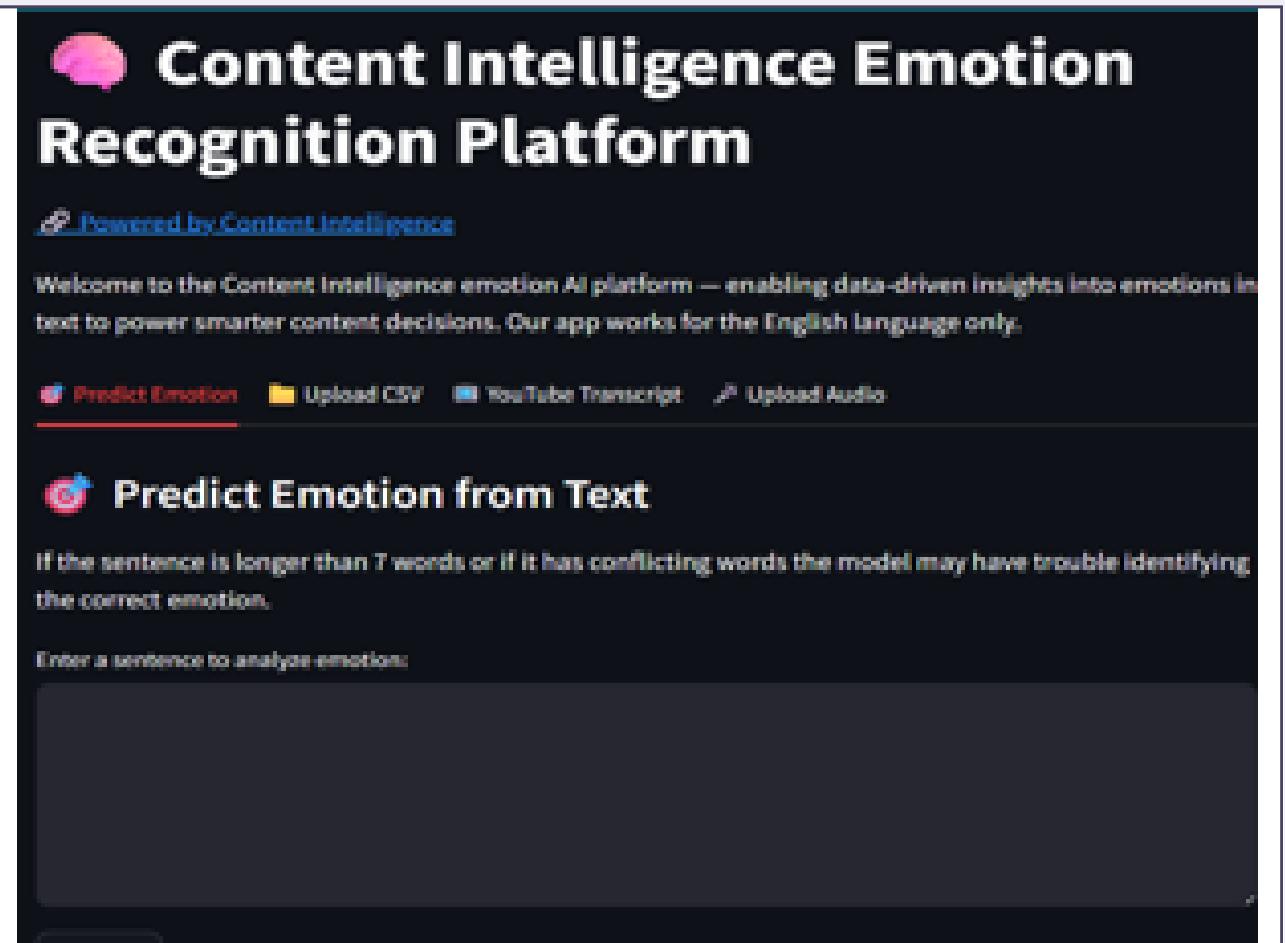
Test

Locust Results:

- 1,978 requests (POST /predict) at up to 44.7 RPS
 - Avg. response: 117 ms | Median: 90 ms | 99%ile: 270 ms
 - Failures: 1 (0.1% error)
 - Conclusion: Stable under high concurrent load; near real-time for typical use.

User Interface (Streamlit Frontend)

- User-friendly interface built in Streamlit for uploading text, audio, or CSV files and viewing predictions.
 - “Designed for ease of use and future deployment. Deployment to Azure not fully functional at demo time, but local interface works.”



Endpoints (API)

Endpoint	Description
/predict	Predict emotion from text
/upload-csv	Batch predict from CSV
/youtube-transcript	Get transcript from Youtube Audio
/upload-audio-transcript	Transcript from audio file

