Amir Aeiny

AI/ML Developer | Python - NLP - Audio Transcription

GitHub: https://github.com/DarkOracle10

LinkedIn: https://www.linkedin.com/in/amir-aeiny-dev

Project: Persian Audio Transcriber

AI/ML Production Tool - Audio Transcription and NLP System

Technology Stack

Backend	Python 3.11+, AsyncIO
AI/ML	OpenAI Whisper, Faster-Whisper, Google Speech, PyTorch
NLP	Hazm (Persian), NLTK
GPU	CUDA, torch GPU acceleration
Audio	FFmpeg, librosa, pydub
Data	pandas, numpy
CLI	argparse, tqdm, colorama
Formats	MP3, WAV, M4A, FLAC, MP4, OGG, WEBM, AAC, WMA, AIFF
Testing	pytest, unittest
Version Control	Git, GitHub

Project Statistics

Lines of Code	2,500+
Transcription Engines	3
Supported Audio Formats	10+
Output Formats	3
Supported Languages	90+
GPU Acceleration	Yes, with CPU fallback
Batch Processing	Parallel execution
Status	Production Ready

Project Overview

A production-ready AI-powered audio transcription system combining Faster-Whisper, OpenAI Whisper, and Google Speech with intelligent fallback. Features GPU acceleration with automatic CPU fallback, Persian text normalization via Hazm, batch processing with live progress, and multiple outputs including timestamped JSON and SRT. Built for large-scale transcription with robust error handling, logging, and scalable architecture.

Key Features

1. Multi-engine transcription with automatic fallback

2. CUDA acceleration with CPU fallback

3. Persian NLP via Hazm for normalization

4. Batch processing with real-time progress

5. Multiple outputs: TXT, JSON with timestamps, SRT

6. Universal format support with auto conversion

7. Comprehensive error handling and logging

8. Live progress indicators and status updates

9. Quality tuning: language detection and model selection

10. Production-ready architecture and tooling

Skills Demonstrated

AI and Machine Learning

- OpenAI Whisper integration and optimization
- Faster-Whisper acceleration and tuning
- Google Speech Recognition API integration
- Multi-engine fallback orchestration
- GPU acceleration with PyTorch and CUDA
- Model performance benchmarking

Natural Language Processing

- Persian text normalization with Hazm
- Tokenization and lemmatization
- Text cleaning and standardization
- Multi-language support (90+ languages)
- Character encoding handling for Persian script

Backend Development

- Python 3.11+ with AsyncIO
- Multi-threaded and parallel pipelines
- REST-ready architecture and patterns

- Robust error handling and logging

Audio Processing

- FFmpeg format conversion and extraction
- Audio quality analysis and optimization
- Waveform analysis and metadata extraction
- Real-time audio streaming support

Performance Optimization

- CUDA-accelerated execution with CPU fallback
- Batch processing for throughput
- Memory-aware handling for large files
- Resource usage monitoring and tuning

DevOps and Production

- Modular architecture and clean code
- Production-grade logging and monitoring
- Cross-platform compatibility (Win, Mac, Linux)
- Dependency and virtual environment management

Similar Projects I Can Build

Audio and Video Processing

- Podcast transcription with diarization
- Video subtitle generation and sync
- Audio content indexing and search
- Multi-language transcription platforms
- Real-time transcription APIs

NLP and Language Processing

- Persian text analysis and summarization
- Multi-language translation systems
- Sentiment analysis for Persian and Arabic
- Named entity recognition for Persian
- Text classification pipelines

AI and ML Systems

- Custom GPT-based content tools
- AI document processing systems
- Image recognition and classification
- Chatbots with NLP capabilities
- Recommendation systems

Automation and Data Processing

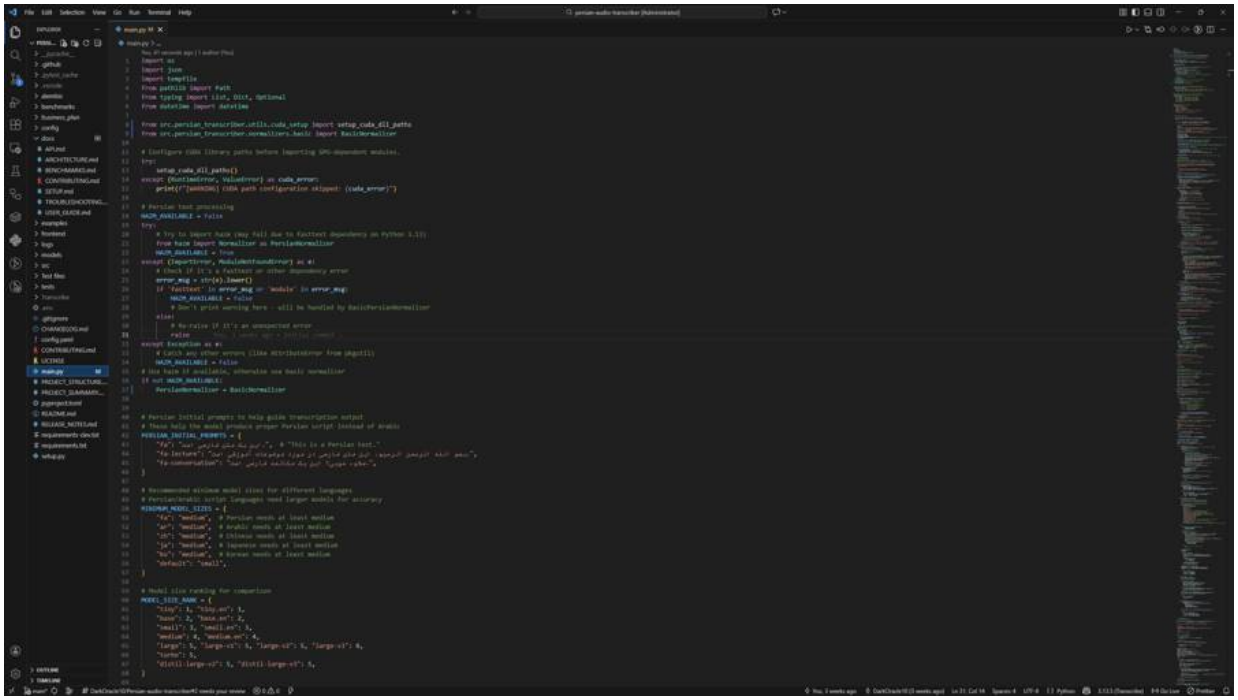
- Batch file processing systems
- Data pipeline automation (ETL)
- Report generation automation
- Workflow automation tools
- API-driven data collection

API Development

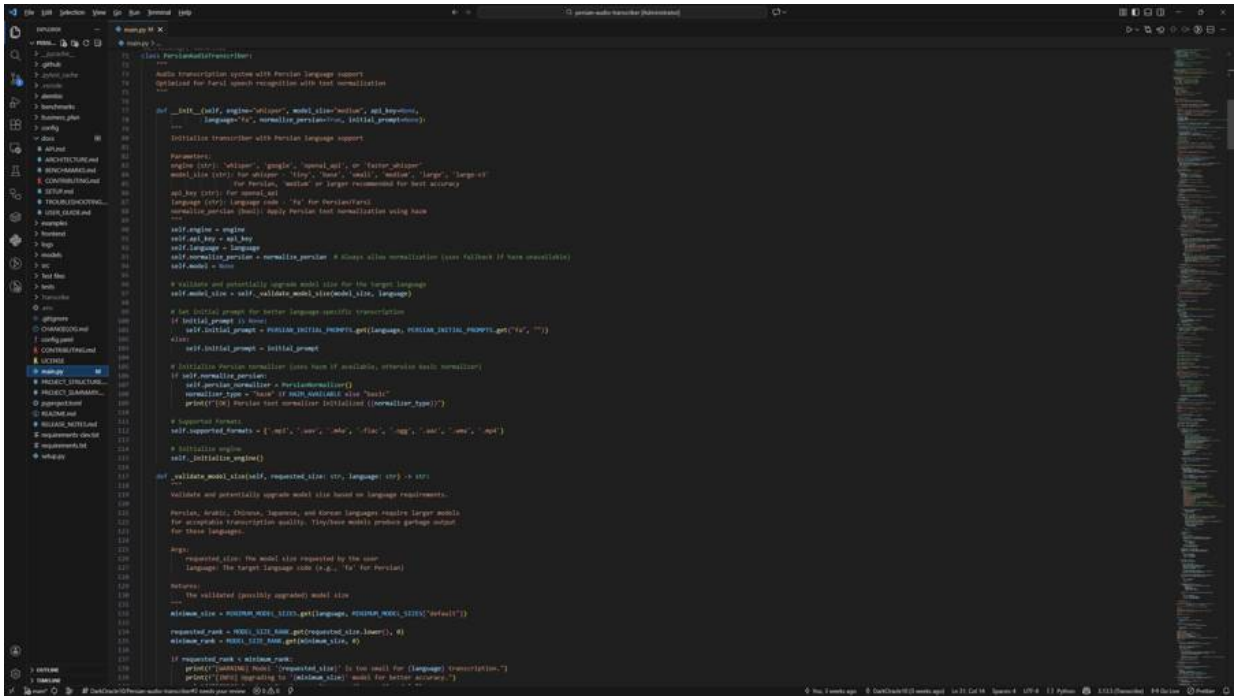
- RESTful wrappers for AI services
- Webhook integration systems
- Third-party API integration
- Microservices-ready designs
- Rate limiting and caching strategies

Screenshots

Code Structure and Engine Logic



GPU Acceleration and Fallback



Generated Outputs (TXT, JSON, SRT)

Sample Persian Transcription Output