

Voice Health Analysis Tool

This script extracts basic voice features from a recording in order to analyse a patient's health indicators.

Prerequisites

- Python (tested using Python 3.13.5)
- Pip (tested using Pip 25.1.1)

Installation

In order for the script to run, the modules listed in [requirements.txt](#) must be installed. It is recommended to install the required modules in a virtual environment

```
python -m pip install librosa numpy reducenoise
```

Usage

```
python voice_analysis.py <file_path>
```

Output

The script outputs a formatted JSON object containing the extracted voice feature data, any health indicators, and an analysis of the voice health. An example output is shown below.

```
{
  "timestamp": "2025-07-25T14:07:08.341526",
  "file": "Stressed.wav",
  "features": [
    {
      "name": "pitch",
      "value": 189.2454833984375,
      "unit": "Hz"
    },
    {
      "name": "energy",
      "value": 0.017401184886693954,
      "unit": "normalized"
    },
    {
      "name": "speaking_rate",
      "value": 245.6499488229273,
      "unit": "syllables per minute"
    }
  ]
},
```

```
"health_indicators": [
  "Energy level is low, indicating possible fatigue or lack of engagement.",
  "Speaking rate is above normal, indicating possible excitement or anxiety."
],
"analysis": [
  "The voice indicates a strong possibility of anxiety or stress due to high speaking rate and low energy."
]
}
```

The output includes basic information such as the timestamp of the analysis, and the file path of the sound file that has been analysed. The script extracts three features from the provided sound file - pitch, energy, and speaking rate - shown in the features list. Each feature displays its name, value and a unit for easy understanding. Each feature is explained in more detail below.

Features

Pitch

The pitch value represents the frequency of the patient's voice in Hz. Typical values of pitch for the human voice are in the range of 85-180Hz for an adult male, 165-255Hz for an adult female, and 250Hz and above for children.

Value	Indicator
< 85	Fatigue or Depression
85-255	Normal
>255	Excitement or Anxiety

Energy

The energy value is a representation of the relative amplitude of the audio file, with values ranging between 0 and 1.

Value	Indicator
<0.02	Fatigue or Depression
0.02-0.1	Normal
>0.1	Excitement

Speaking Rate

The speaking rate is the average number of syllables spoken by the patient per minute. A normal speaking rate for a patient speaking in conversational English is typically between 120 and 220 syllables per minute.

Value	Indicator
-------	-----------

Value	Indicator
<120	Fatigue or Depression
120-220	Normal
>220	Excitement or Anxiety

Health Indicators

Anything that has been picked up as abnormal will be displayed in this section, including a short message of possible indications.

In the above example output, the analysis has indicated low energy and high speaking rate and as such two messages are displayed informing the clinician of these observations.

Analysis

This section provides an overall analysis of the patient's voice health indicators. If no abnormal health indicators are found this analysis will determine there are no significant health concerns. If one or more abnormal health indicators are found, the analysis will provide a potential cause for the abnormal indicators as well as describing the confidence of the analysis in the form of "possible" or "strong possibility". There is a strong possibility of a cause if multiple indicators point towards the cause, and only a possibility if there is only one indicator.