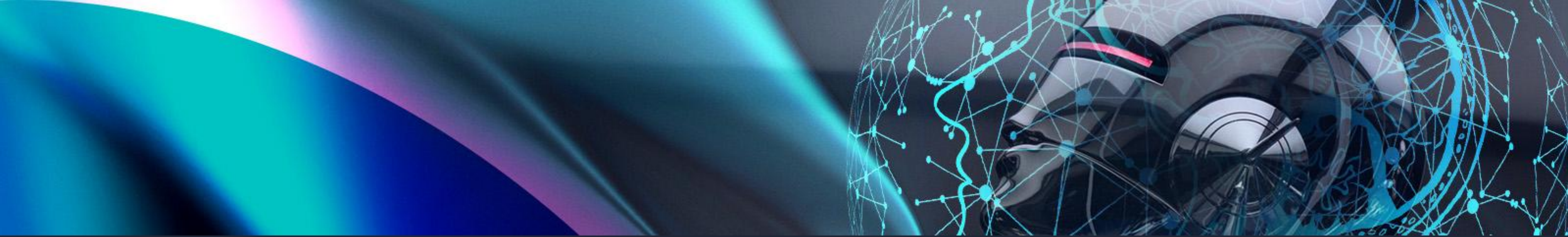
CardioBeat AI: Heart Sound Classification with Deep Learning

By:
Hemil Shah
Niharika Das

Contents

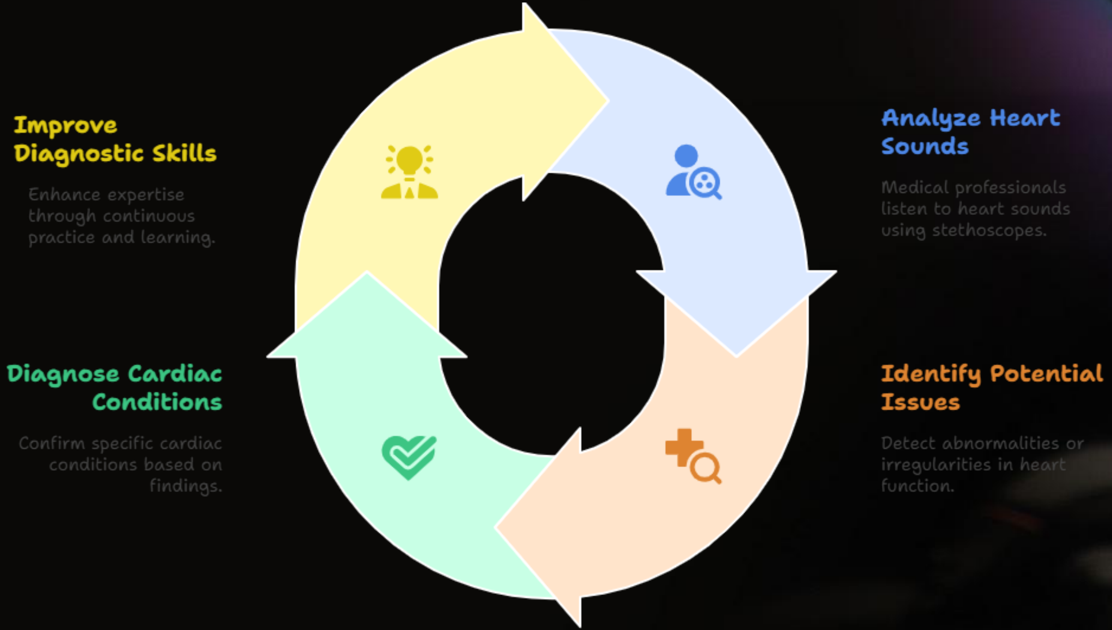


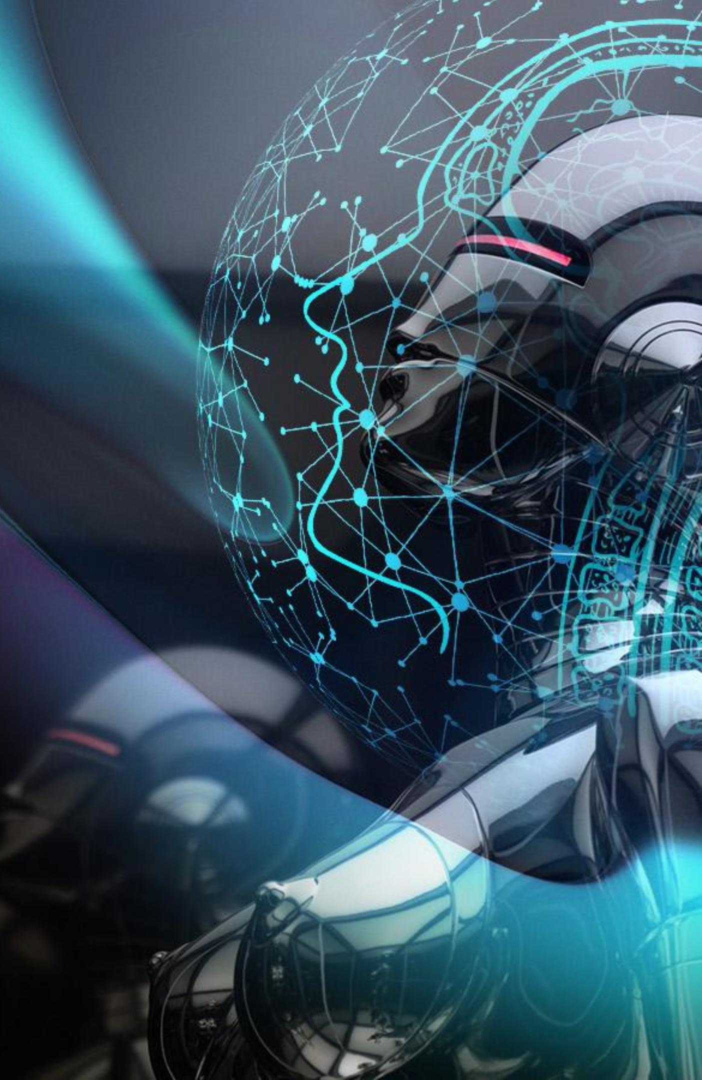
Part-01: Introduction to Heart Sound Analysis

Diagnosing with Heart Sounds

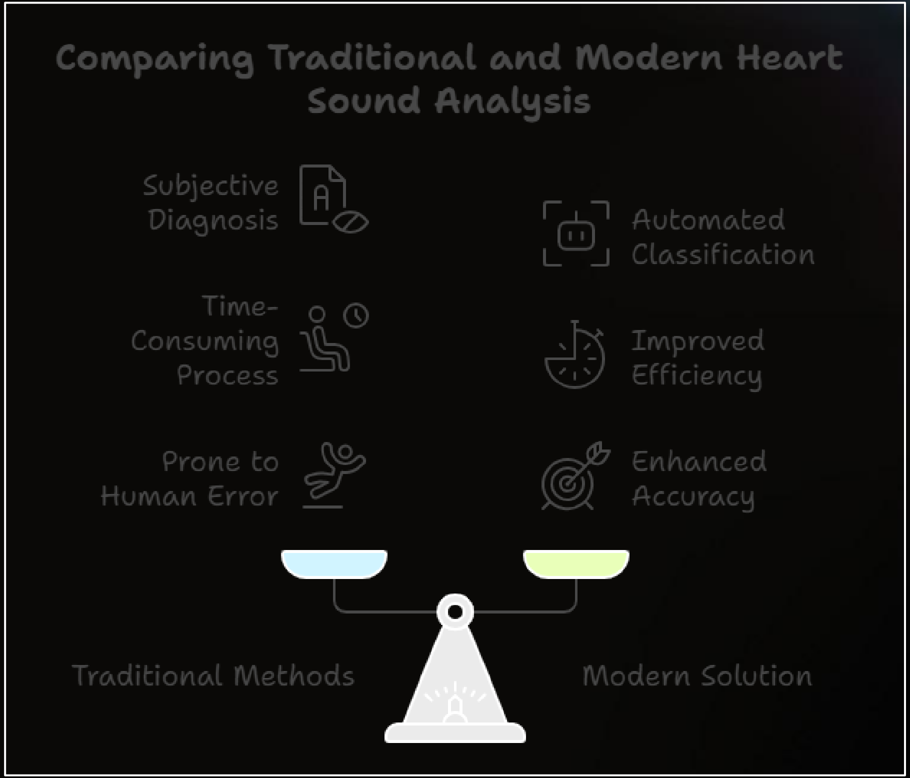
Traditional Methods:

- Analyzing heart sounds is key in diagnosing cardiac conditions, heavily relying on the expertise of medical professionals using stethoscopes.





Diagnosing with Heart Sounds



Diagnosing with Heart Sounds

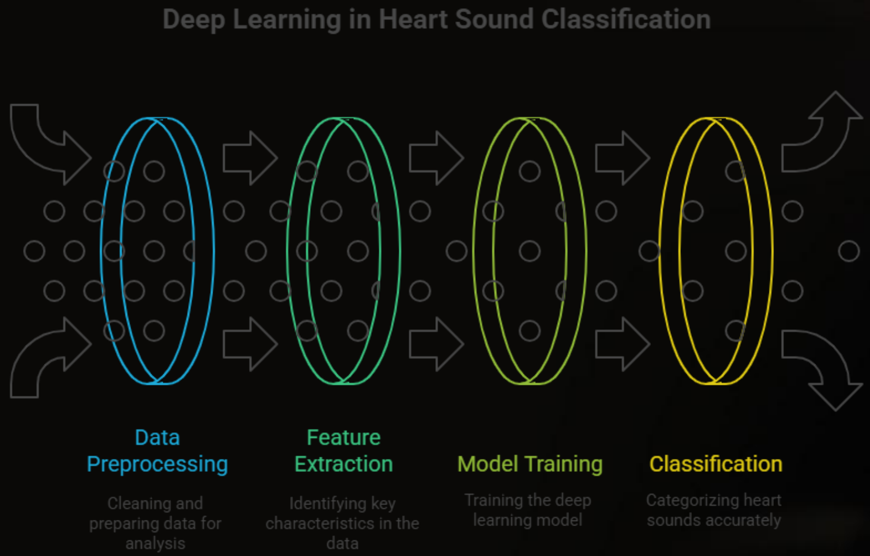
The Challenges:

- This process is subjective, requires time, and is prone to human error, presenting limitations. We aim to address.

Diagnosing with Heart Sounds

Modern Solution:

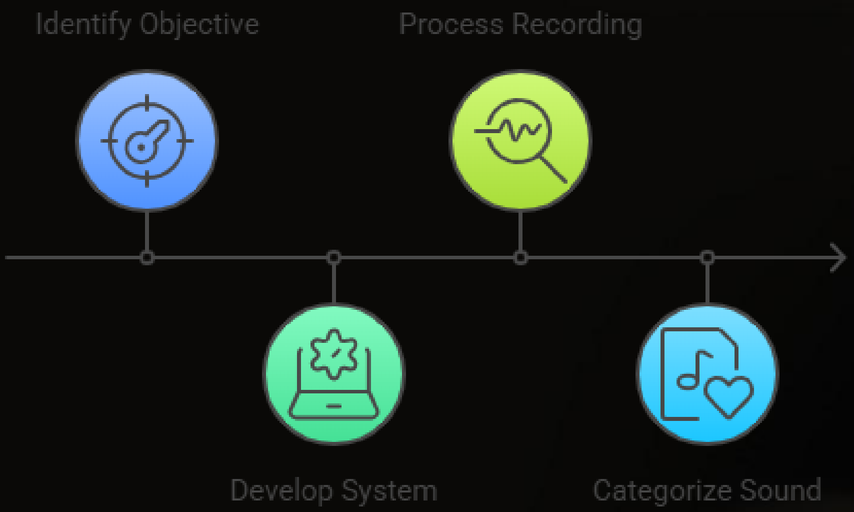
- Deep learning automates heart sound classification, providing a potent tool for preliminary diagnosis, improving efficiency and accuracy.

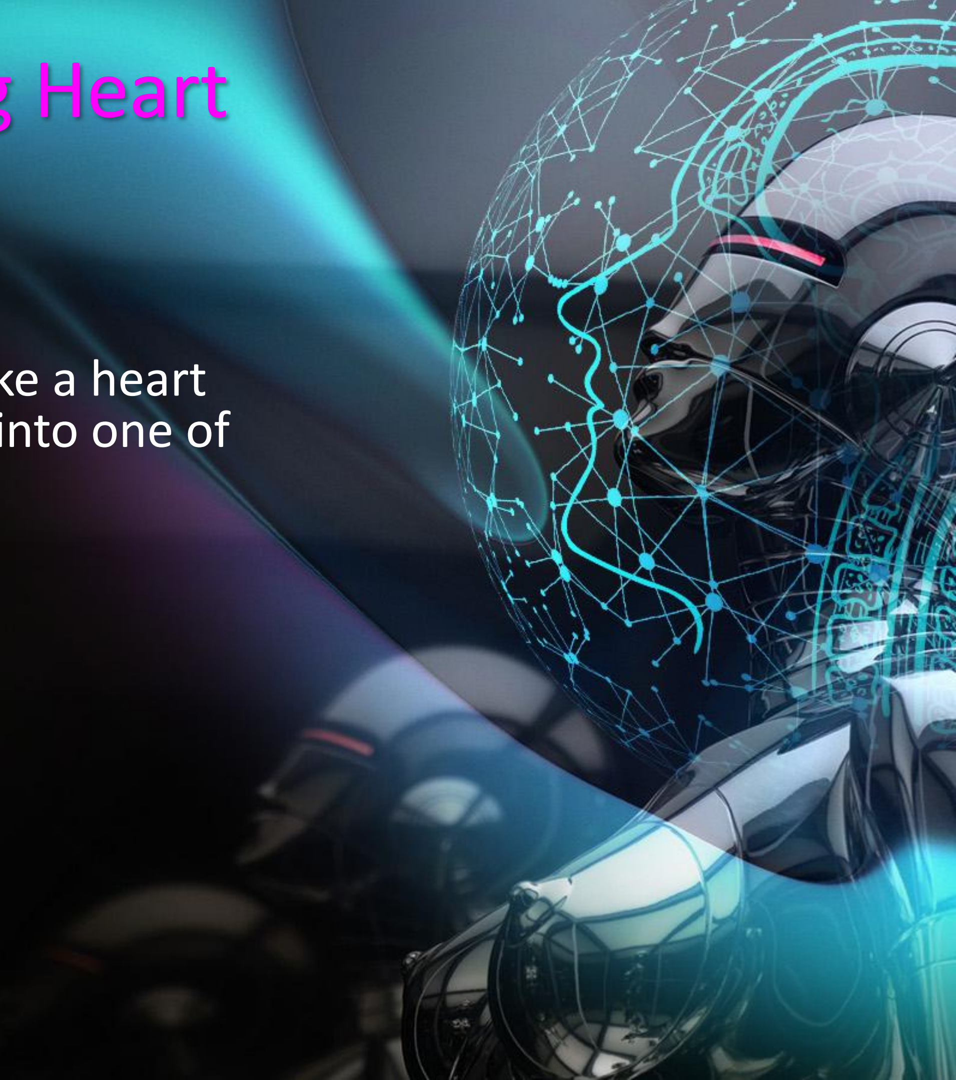


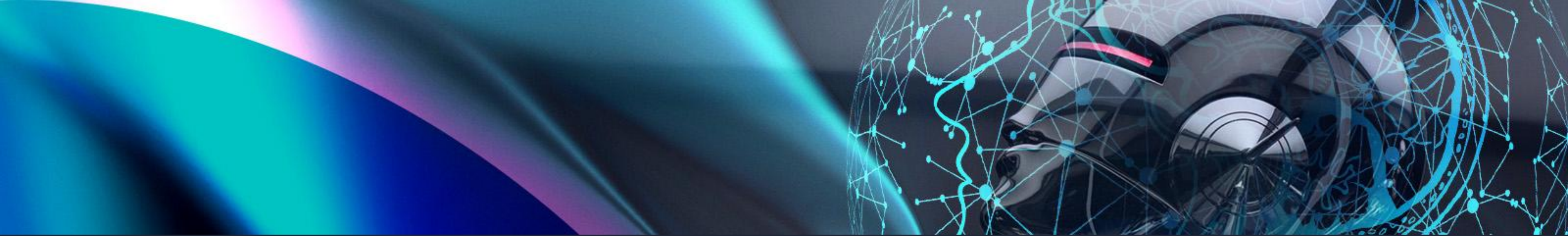
The Problem: Diagnosing Heart Sounds

Our Goal:

- To build an end-to-end system that can take a heart sound recording and accurately classify it into one of several predefined categories.







Part-02: Dataset Overview

About the Dataset

Data Source:

- The heart sound data is provided as a compressed RP1Heart Data final.zip file, encompassing a collection of heart sound recordings for analysis.

Structure:

- The project utilizes a dataset of heart sounds that's provided as a compressed RP1Heart Data final.zip file. The application is designed to automatically unzip this file and organize the audio into distinct folders. Each folder represents a specific class, allowing the model to be trained on labeled data.
- The model is trained on five specific categories of heart sounds:
    1. AS (Mitral Valve prolapse)
    2. MR (Mitral Regurgitation)
    3. MS (Mitral Stenosis)
    4. MVP (Mitral Valve Prolapse)
    5. N (Normal)
- The system then processes the individual audio files within these folders, extracting the necessary features to train the deep learning model.

The Data Processing Sequence

Download Data

The compressed data file is downloaded from the provided link.

Unzip Data

The downloaded file is automatically unzipped to access the data.

Organize Folders

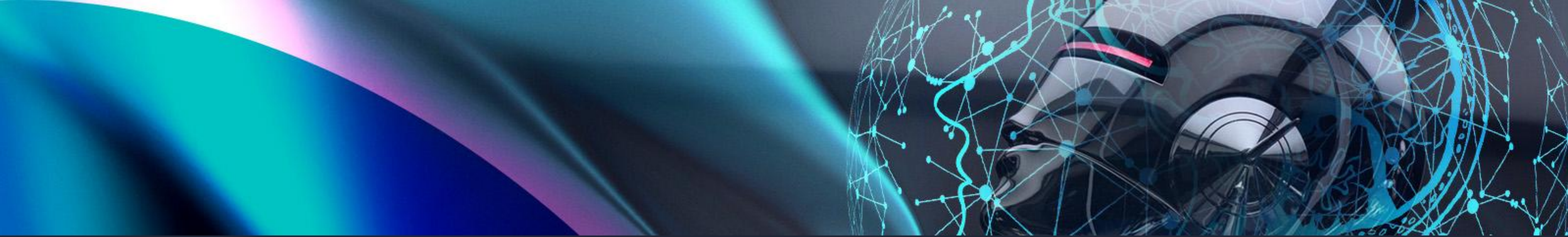
The unzipped data is organized into folders representing different heart sound classes.

Train Model

The model is trained using the organized data to recognize heart sound classes.

Process Audio Files

Individual audio files are processed to extract features for training.



Part-03: The Process Funnel: Sound to Diagnosis

Process Funnel

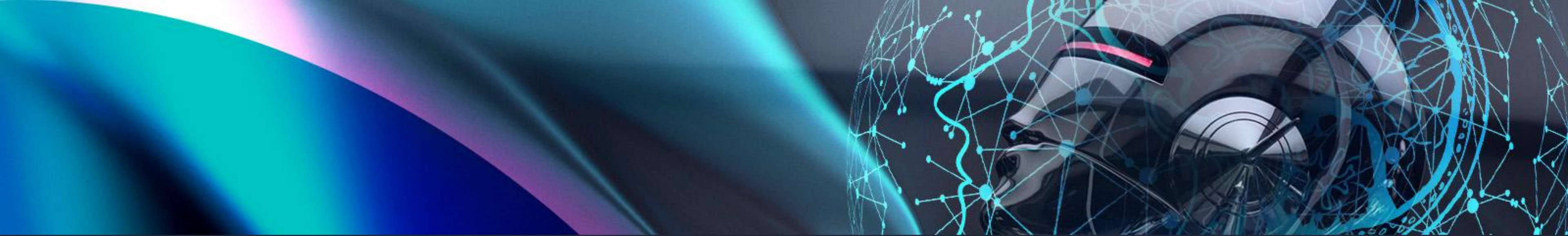
Raw Audio Input
- It all begins when a user uploads a heart sound recording in formats like WAV, MP3, or MP4.

Audio Preprocessing & Feature Extraction
- The raw audio is first cleaned by trimming out silent sections.
- The file is then processed to a standardized format, and the key features are extracted. Our model uses Mel-spectrograms, which are visual representations of the audio's frequency and time.

Deep Learning Model
- The Mel-spectrogram is fed into our pre-trained Convolutional Neural Network (CNN).
- The CNN analyzes the image-like data to find patterns it learned during training.

Final Prediction
- The model outputs a prediction of the heart sound's class.
- The app displays the result with a confidence score and a breakdown of the probabilities for each category, giving you clear insights into the classification.

Process Funnel



Part-04: Solution Features: CardioBeat: A User-Friendly System

Features:

- ***A Comprehensive & User-Friendly System:***

Navigation through this system is very easy. The process is very easy, and the user can perform the predictions very easily, on a very good looking UI.

- ***Seamless Data Pipeline:***

The application automates the entire process, from unzipping the raw data to training the model, making it a self-contained system.

- ***Deep Learning at the Core:***

Utilizes a custom-built, multi-layered Convolutional Neural Network (CNN), specifically designed to classify complex audio patterns with high accuracy.

- ***Feature Engineering:***

Transforms raw audio signals into Mel-spectrograms , converting time-series data into a visual format that the CNN can analyze effectively.

Features:

- ***Interactive Streamlit UI:***

Features a simple, elegant web interface that allows users to easily upload audio files and receive instant predictions.

- ***Multi-Format Support:***

Goes beyond basic WAV files by supporting MP3 and MP4 formats, thanks to the integration of the pydub library.

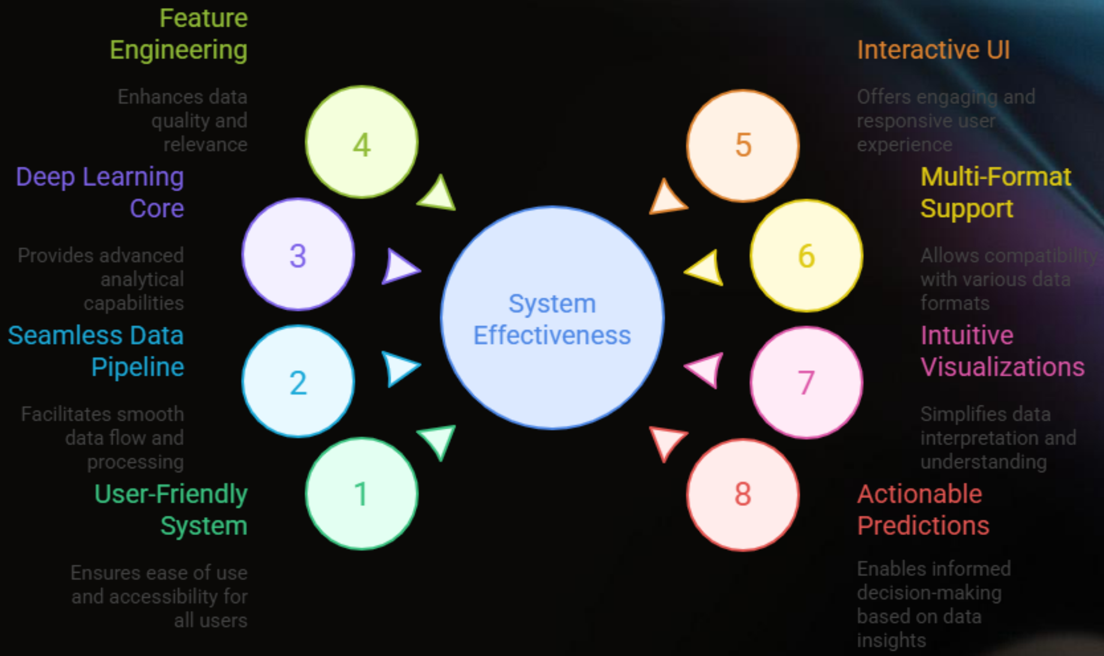
- ***Intuitive Visualizations:***

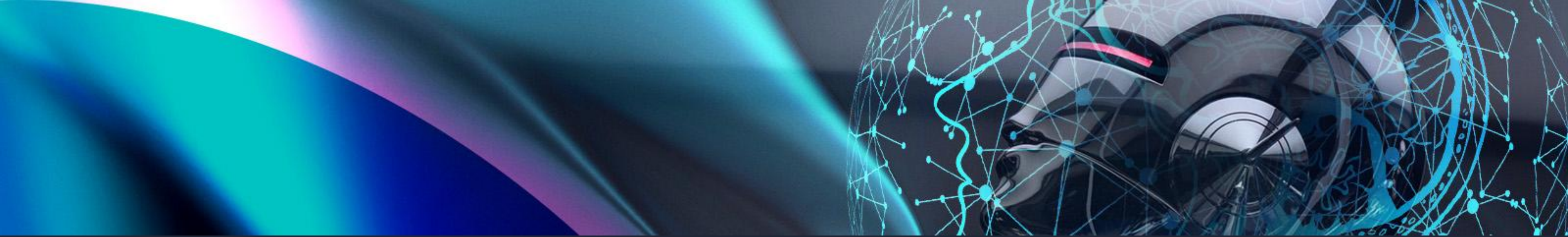
Provides clear visual feedback by displaying both the audio waveform and the Mel-spectrogram of the uploaded file.

- ***Actionable Predictions:***

Presents the classification result with a confidence score and a detailed probability breakdown for all categories, offering valuable insights beyond a simple label.

Features:


System Effectiveness
1
User-Friendly System
Ensures ease of use and accessibility for all users
2
Seamless Data Pipeline
Facilitates smooth data flow and processing
3
Deep Learning Core
Provides advanced analytical capabilities
4
Feature Engineering
Enhances data quality and relevance
5
Interactive UI
Offers engaging and responsive user experience
6
Multi-Format Support
Allows compatibility with various data formats
7
Intuitive Visualizations
Simplifies data interpretation and understanding
8
Actionable Predictions
Enables informed decision-making based on data insights



Part-05: Impact of the Solution: Transforming Diagnostics

Impact:

- ***Democratized Access to Diagnostics:***

Our solution provides a low-cost, accessible tool for preliminary heart sound analysis. This can be particularly impactful in areas with limited access to specialized medical equipment or professionals.

- ***Empowering Users:***

The application doesn't just give a diagnosis; it offers clear visualizations of the heart sound's waveform and Mel-spectrogram. This helps users understand the data and provides valuable educational insight into the analysis process.

- ***Preliminary Screening & Triage:***

While not a replacement for professional medical advice, the app can serve as a valuable initial screening tool. It could help flag potential issues, encouraging users to seek a formal diagnosis from a doctor.

- ***Bridging Technology & Medicine:***

The project demonstrates how cutting-edge deep learning techniques can be applied to real-world medical challenges, paving the way for more AI-powered diagnostic tools in the future.

Impact:

Limited Diagnostics

Access to medical diagnostics is restricted

User Empowerment

Users gain control over their health data

Preliminary Screening

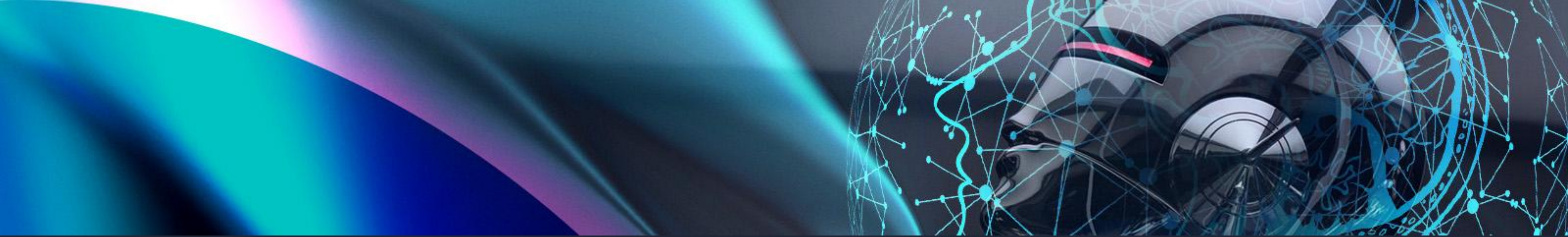
Initial health assessments are conducted

Technology Integration

Combining tech and medical expertise

Democratized Access

Diagnostics available to a wider population



Part-06: Stand-Out Areas: Innovative and Accessible

Stand-Out Areas:

- ***Broad Accessibility:***

Unlike traditional medical devices, this is a web-based, portable solution that can be used on a smartphone, tablet, or computer, making preliminary heart sound analysis available to a wider audience.

- ***Empowering Visuals:***

The application goes beyond a simple prediction by providing live visualizations of the sound wave and Mel-spectrogram. This not only makes the process transparent but also helps in understanding the underlying data.

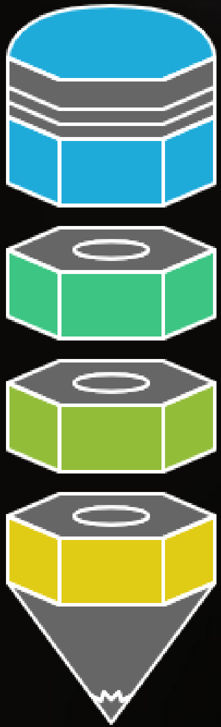
- ***Real-World Utility:***

This project is a practical demonstration of how a sophisticated deep learning model can be deployed in a user-friendly application to address a real-world need, serving as a powerful proof of concept for future AI in medicine.

- ***Flexible Input:***

By supporting a variety of audio and video formats, including MP3 and MP4, the solution is adaptable and easy to use, avoiding the need for manual file conversion.

Stand-Out Areas:



Broad Accessibility

Ensures content is easily accessible to a wide range of users

Empowering Visuals

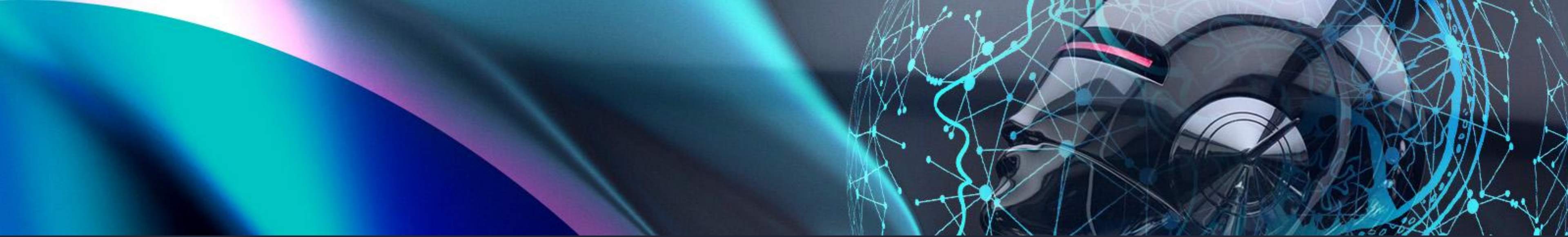
Enhances understanding and retention through visual aids

Real-World Utility

Guarantees concepts are applicable and beneficial in everyday scenarios

Flexible Input

Allows users to interact with the material in a personalized way



Part-07: Architecture Diagram

Details:
Architecture Diagram

The system's architecture can be visualized as a straightforward, linear pipeline.

- ***Audio Input:***

The user uploads a heart sound file (.wav, .mp3, .mp4) via the Streamlit interface.

- ***Preprocessing & Feature Extraction:***

The audio is loaded using librosa and converted to a standardized format. Mel-spectrogram features are then extracted, transforming the raw sound data into a dense, visual representation that the model can interpret.

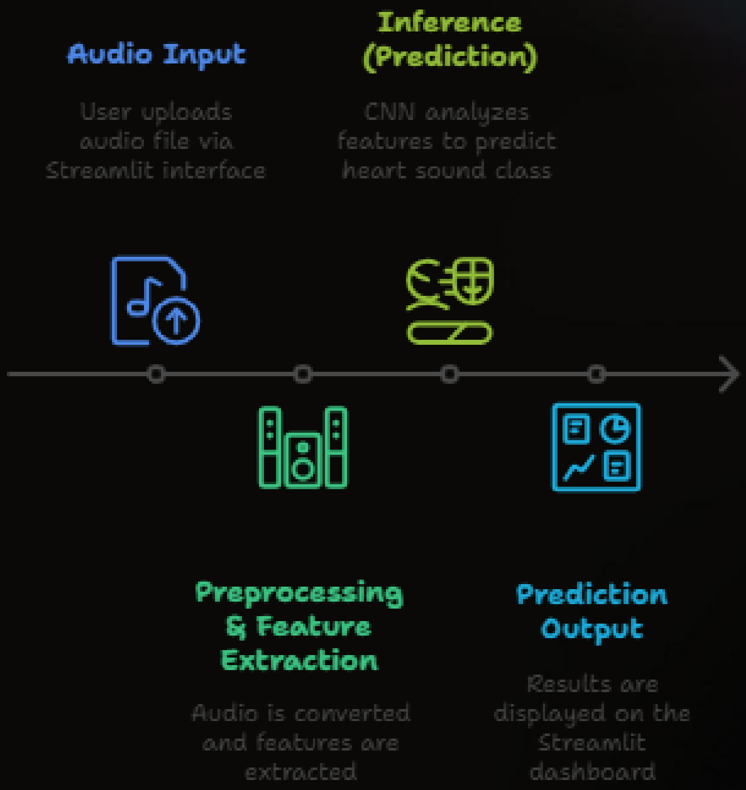
- ***Inference (Prediction):***

The preprocessed features are fed into the trained Convolutional Neural Network (CNN), which analyzes the Mel-spectrogram to predict the heart sound's class.

- ***Prediction Output:***

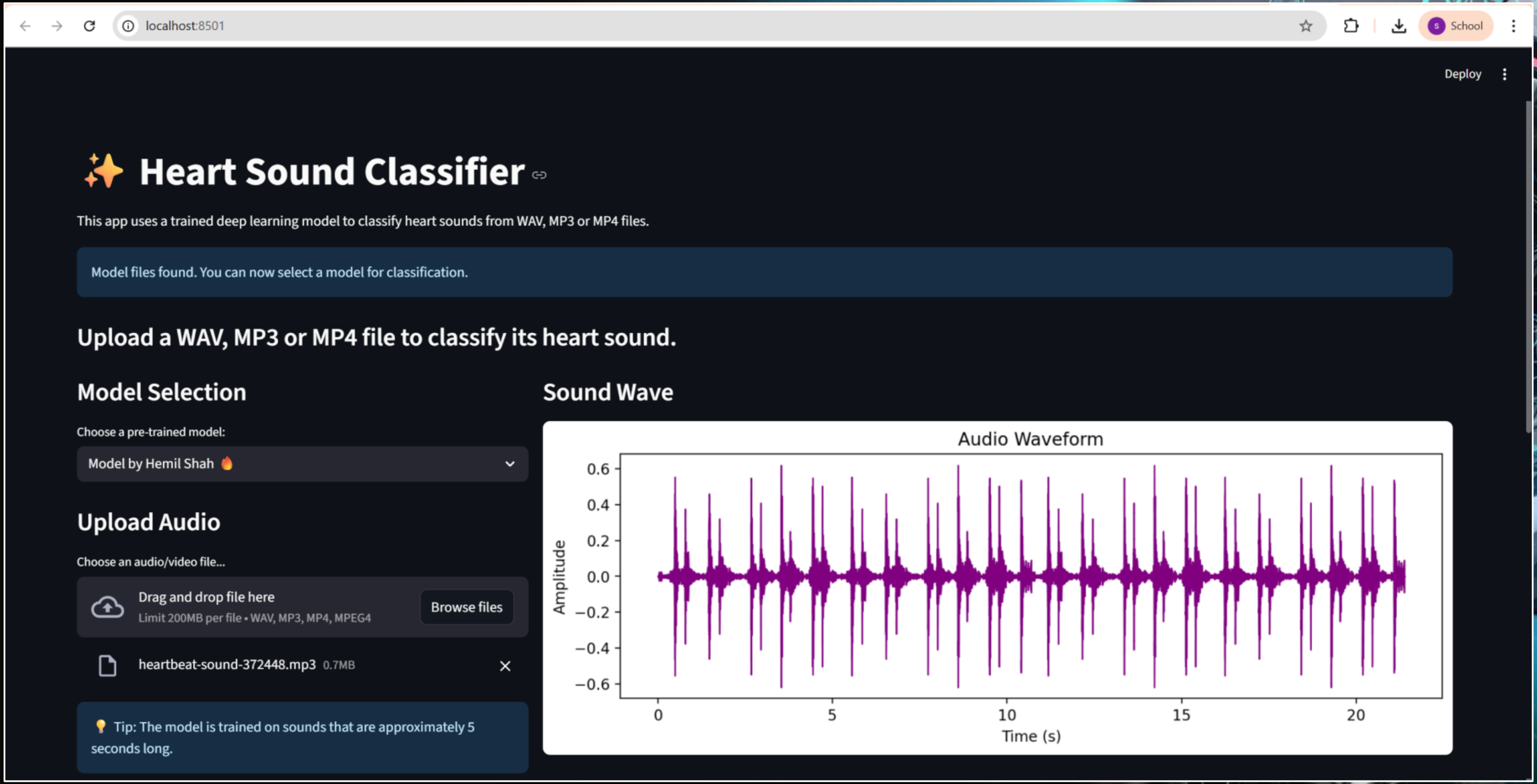
The result is displayed to the user on the Streamlit dashboard. It includes the predicted class, the confidence score, and a breakdown of the probabilities for each class.

Heart Sound Analysis Pipeline:

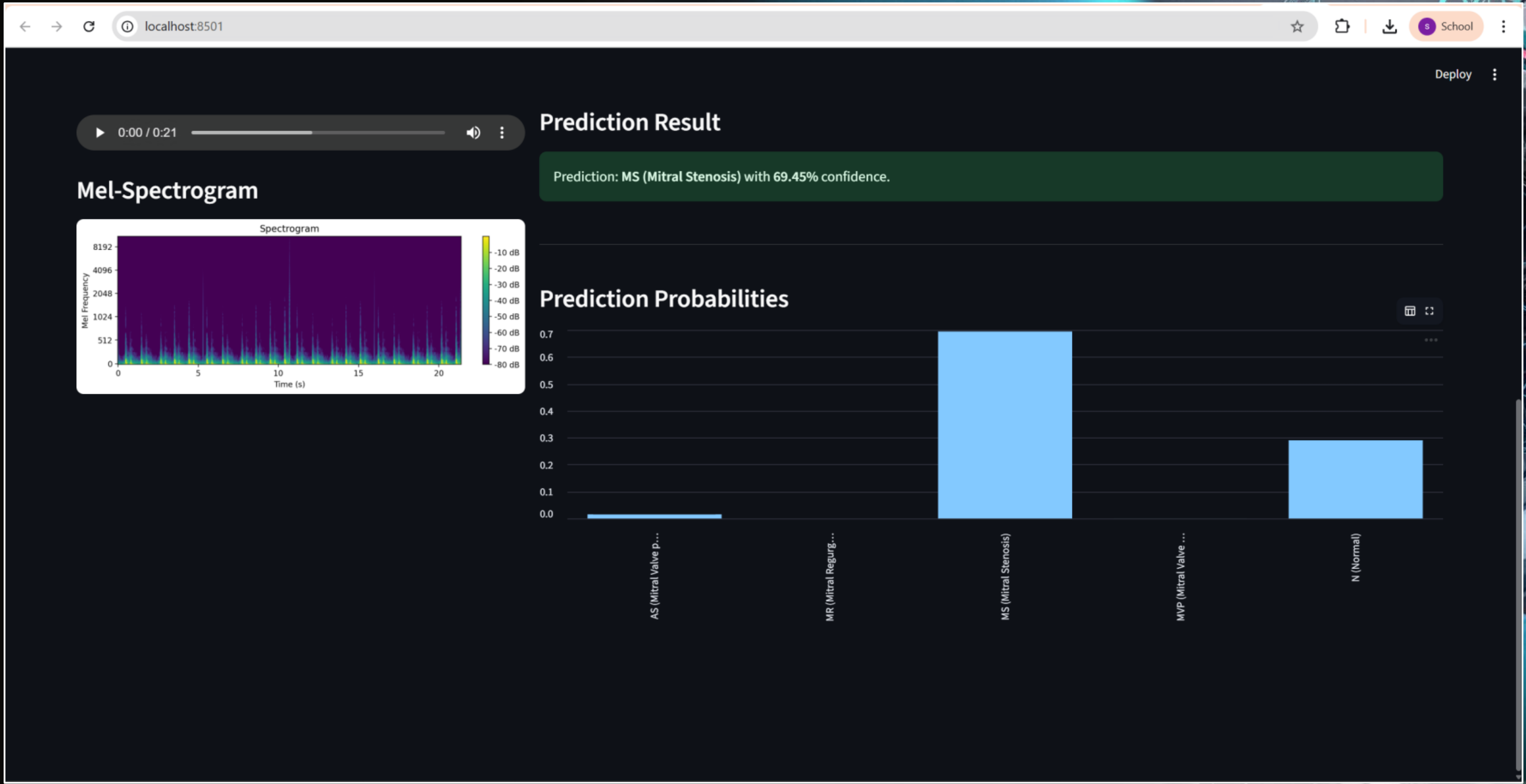


Part-08: Interface Preview

User-Side Interface



User-Side Interface



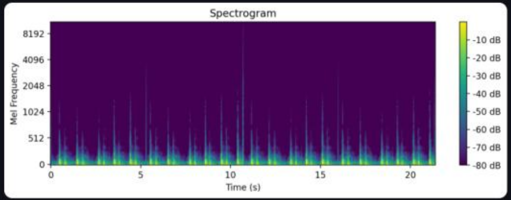
User-Side Interface

localhost:8501

Deploy

0:00 / 0:21

Mel-Spectrogram



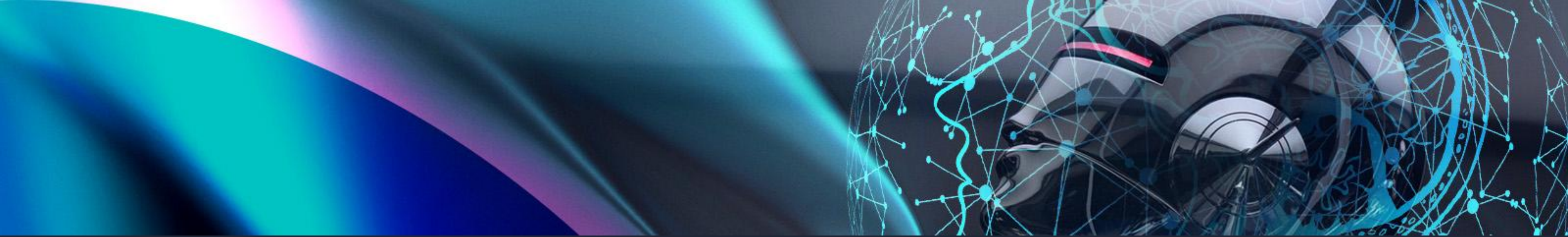
Prediction Result

Prediction: **MS (Mitral Stenosis)** with **69.45%** confidence.

Prediction Probabilities

	index -- streamlit-generated	Probability
0	AS (Mitral Valve prolapse)	0.015
1	MR (Mitral Regurgitation)	0.00006
2	MS (Mitral Stenosis)	0.6945
3	MVP (Mitral Valve Prolapse)	0.0001
4	N (Normal)	0.2903

Admin-Side Interface



Part-09: Future Scope

Future Scope

- ***Live Audio Input:***

Implement real-time classification directly from a device's microphone, allowing for instant analysis without needing to upload a pre-recorded file.

- ***Mobile App Development:***

Port the application to a dedicated mobile platform (iOS/Android) to make the diagnostic tool even more accessible and portable for on-the-go use.

- ***Database Integration:***

Integrate with a database to securely store anonymized data, which could be used to continuously train and improve the model's accuracy over time.

- ***Expanded Classifications:***

Broaden the model's capabilities by training it on a larger, more diverse dataset to recognize a wider range of heart conditions and abnormalities.

Future Scope

1. Live Audio Input

Enhances real-time interaction

2. Mobile App Development

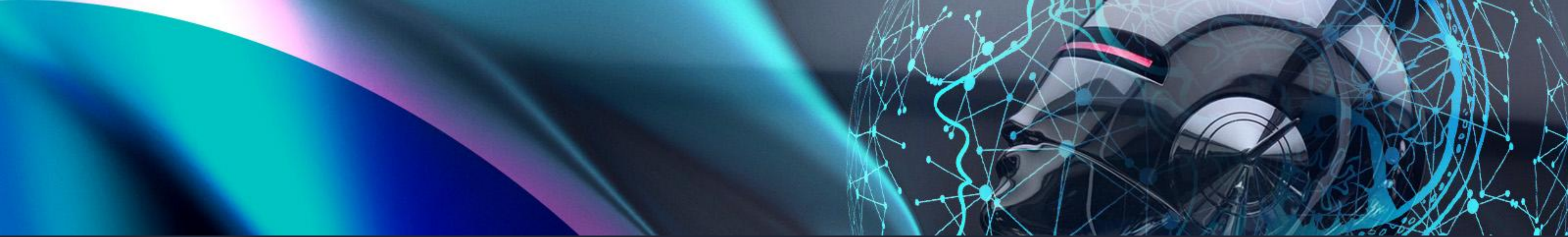
Increases accessibility and user engagement

3. Database Integration

Streamlines data management and improves efficiency

4. Expanded Classifications

Allows for broader categorization and analysis



Thank you