

Project Report

Offline Hindi Voice Assistant using Raspberry Pi and Vosk ASR

Methodology:

The system follows a modular pipeline architecture:

Step 1: Audio Capture

- Microphone input via PyAudio
- 44100 Hz mono sampling
- Frame-based streaming

Step 2: Automatic Speech Recognition

- Vosk Hindi acoustic model (small)
- KaldiRecognizer processes audio frames
- Converts speech → Hindi text

Step 3: Text Processing

- Normalization of input
- Phrase-based intent detection
- Command mapping using dictionary-based matching

Step 4: Intent Handling

- System commands
- Weather cache reading (JSON-based)
- Conversational responses

Step 5: Speech Output

- espeak-ng Hindi voice
- Adjustable speed, pitch, and voice variant
- Audio routed via PipeWire/Bluetooth speaker

Hardware Utilization:

1] Raspberry Pi 5

- ARM-based processing for ASR
- Handles large Hindi model decoding
- System-level command execution

2] Microphone

- USB mic preferred for cleaner input
- 44100 Hz sampling for optimal model compatibility

3] Bluetooth Speaker

- PipeWire audio routing
- pw-play for proper audio output

Optimization Techniques Applied

TTS Optimization:

- Used hi voice variant
- Adjusted speed and pitch
- Used PipeWire routing to fix Bluetooth issues
- Reduced latency by avoiding ALSA direct output

NLP Optimization:

- Text normalization
- Removal of filler words
- Phrase matching instead of exact match
- Global LAST_RESPONSE for repeat command

Memory Optimization:

- Used JSON cache for weather instead of API

- Lightweight intent mapping (no heavy ML)

Results:

- Accurate Hindi speech recognition sometimes makes mistakes in recognising Hindi words
- Fully offline operation
- Stable Bluetooth audio output
- Multi-intent support
- Repeat command functionality