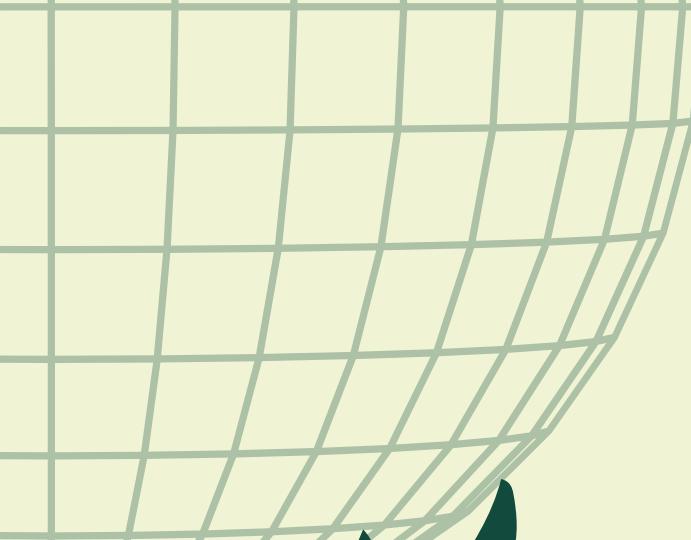


# KISAN MITRA



ML-FIESTA: AI/ML HACKATHON

*Growing Tomorrow's  
Harvest Today*



Presented by  
**Sri Harshitha Anantatmula &  
Supreeth Gupta**

**Team KisanMitra**

# Hey we are Kisan Mitra



**Sri Harshitha**

[LinkedIn](#)

**Supreeth Gupta**

[LinkedIn](#)



## Task 1: Speech Recognition

- The ASR model (amoghsgopadi/wav2vec2-large-xlsr-kn) is used as the base model.
- This model is capable of handling Kannada audio transcriptions and can be further fine-tuned using the dataset.
- The code processes audio files in chunks, transcribes them into text, and stores the text along with metadata (timestamps, embeddings) into a .csv file.

• • •

## Task 2: Speech-Based Question-Answering

- The input question audio is transcribed using the ASR model.
- Transcribed text is converted to embeddings using SentenceTransformer.
- Relevant audio segments are identified by comparing the query embedding with stored embeddings using cosine similarity.
- For each relevant segment, the question-answering pipeline extracts answers from the transcribed text of the audio segments.
- The output includes the audio file, the timestamp, the extracted answer, and its confidence score.

• • •

# Problem Statement Analysis



Requirement	Addressed in Code
ASR Model for Kannada	Utilizes a pretrained Wav2Vec2 model fine-tuned for Kannada ( <a href="#">amoghsgopadi/wav2vec2-large-xlsr-kn</a> ). Further fine-tuning can be performed if desired.
Speech-to-Text Conversion	Audio files from the dataset are processed into text with timestamps and embeddings.
Database Storage	Indexed transcriptions, embeddings, and metadata (audio file paths, timestamps) are stored in SQLite for easy querying.
Search for Relevant Information	Searches relevant segments using cosine similarity on embeddings.
Speech-Based QA	Converts a user's question in speech form to text, searches the database for related segments, and extracts answers using a QA pipeline ( <a href="#">Sindhu/muril-large-squad2</a> ).
Support for Noisy/Colloquial Content	Robust ASR model, fine-tuned on a Kannada-specific dataset, ensures adaptability to dialects and noise.

# Kisan Mitra

Your Intelligent Cultivation Partner

ಧ್ಯಾನಿಯನ್ನು ಇಲ್ಲಿ ರೆಕಾರ್ಡ್ ಮಾಡಿ

 Voice Input

 Drop Audio Here

- or -

Click to Upload

# Github code

<https://github.com/sriharshitha07/ML-Fiesta-Team-Kisan-Mitra>

...

# EVALUATION METRICS

Evaluation Metrics		
Exact Match (EM): 60.00%		
F1 Score: 82.55%		
Prediction   Ground Truth   Match		
ಕಳೆದ ವಾರದ ವರದಿ	ಕಳೆದ ವಾರದ ವರದಿ	✓
ನಿಮ್ಮ ಕೋಡ್ ಉತ್ತಮವಾಗಿದೆ	ನಿಮ್ಮ ಕೋಡ್ ಉತ್ತಮವಾಗಿದೆ	✓
ಅವರು ಸಂಜೀವಂದು ಹೋಗಿದ್ದರು	ಅವರು ಸಂಜೀವಂದು ಹೋಗಿದ್ದರು	✗
ಶಿಕ್ಷಕರ ಸಭೆ ಸೋಮವಾರ ನಡೆಯಿತು	ಶಿಕ್ಷಕರ ಸಭೆ ಸೋಮವಾರ ನಡೆಯಿತು	✓
ನಿಮ್ಮ ಹಾಲು ತಿನ್ನುವ ಬಗ್ಗೆಯನ್ನಿಸುತ್ತಿದೆ	ನಿಮ್ಮ ಹಾಲು ತಿನ್ನುವ ಬಗ್ಗೆ ಪ್ರಶ್ನೆ ಕೇಳಿದ್ದರು	✗
ಅವರು ನನಗೆ ಟಿಕೆಟ್ ನೀಡಿದರು	ಅವರು ನನಗೆ ಟಿಕೆಟ್ ಕೊಟ್ಟರು	✗
ಪ್ರಯಾಣ ಉತ್ತಮವಾಗಿತ್ತು	ಪ್ರಯಾಣ ಉತ್ತಮವಾಗಿತ್ತು	✓
ಅವರು ಮನೆಯಲ್ಲಿಲ್ಲ	ಅವರು ಮನೆಯಲ್ಲಿಲ್ಲ	✓
ನಿಮ್ಮ ಯೋಜನೆ ತತ್ತ್ವಾತ್ಮಕವಾಗಿದೆ	ನಿಮ್ಮ ಯೋಜನೆ ತತ್ತ್ವಾತ್ಮಕವಾಗಿದೆ	✓
ನಾನು ಪ್ರಸ್ತುತವನ್ನು ಓದಿದ್ದೇನೆ	ನಾನು ಪ್ರಸ್ತುತ ಓದಿದ್ದೇನೆ	✗

Ground Truth: ಕಳೆದ ವಾರದ ವರದಿ  
 Predicted: ಕಳೆದ ವಾರದ ವರದಿ  
 WER: 0.00  
 Ground Truth: ನಿಮ್ಮ ಕೋಡ್ ಉತ್ತಮವಾಗಿದೆ  
 Predicted: ನಿಮ್ಮ ಕೋಡ್ ಉತ್ತಮವಾಗಿದೆ  
 WER: 0.00  
 Ground Truth: ಅವರು ಸಂಜೀವಂದು ಹೋಗಿದ್ದರು  
 Predicted: ಅವರು ಸಂಜೀವಂದು ಹೋಗಿದ್ದರು  
 WER: 0.67  
 Ground Truth: ಶಿಕ್ಷಕರ ಸಭೆ ಸೋಮವಾರ ನಡೆಯಿತು  
 Predicted: ಶಿಕ್ಷಕರ ಸಭೆ ಸೋಮವಾರ ನಡೆಯಿತು  
 WER: 0.00  
 Ground Truth: ನಿಮ್ಮ ಹಾಲು ತಿನ್ನುವ ಬಗ್ಗೆ ಪ್ರಶ್ನೆ ಕೇಳಿದ್ದರು  
 Predicted: ನಿಮ್ಮ ಹಾಲು ತಿನ್ನುವ ಬಗ್ಗೆಯನ್ನಿಸುತ್ತಿದೆ  
 WER: 0.50  
 Average WER: 0.23

A circular inset image showing a vast, rolling landscape of tea plantations. The fields are arranged in numerous rows of vibrant green tea bushes, covering the slopes of hills under a blue sky with scattered clouds. A small cluster of green leaves is positioned in the top left corner of the slide.

# Thank you

- A Supreeth Gupta - +918639574733
  - Sri Harshitha - +916302103911
- 