



Submitted by: Archit Gupta, Shivi Mehrotra

Submitted to: Dr Bhawna Saxena

Concept Ideas

1

A Speech Emotion Recognition (SER) extracts speech features, classifies emotions via ML, integrates with voice interfaces, and needs annotated data. It detects happiness, sadness, or anger in real-time for empathetic human-computer communication.

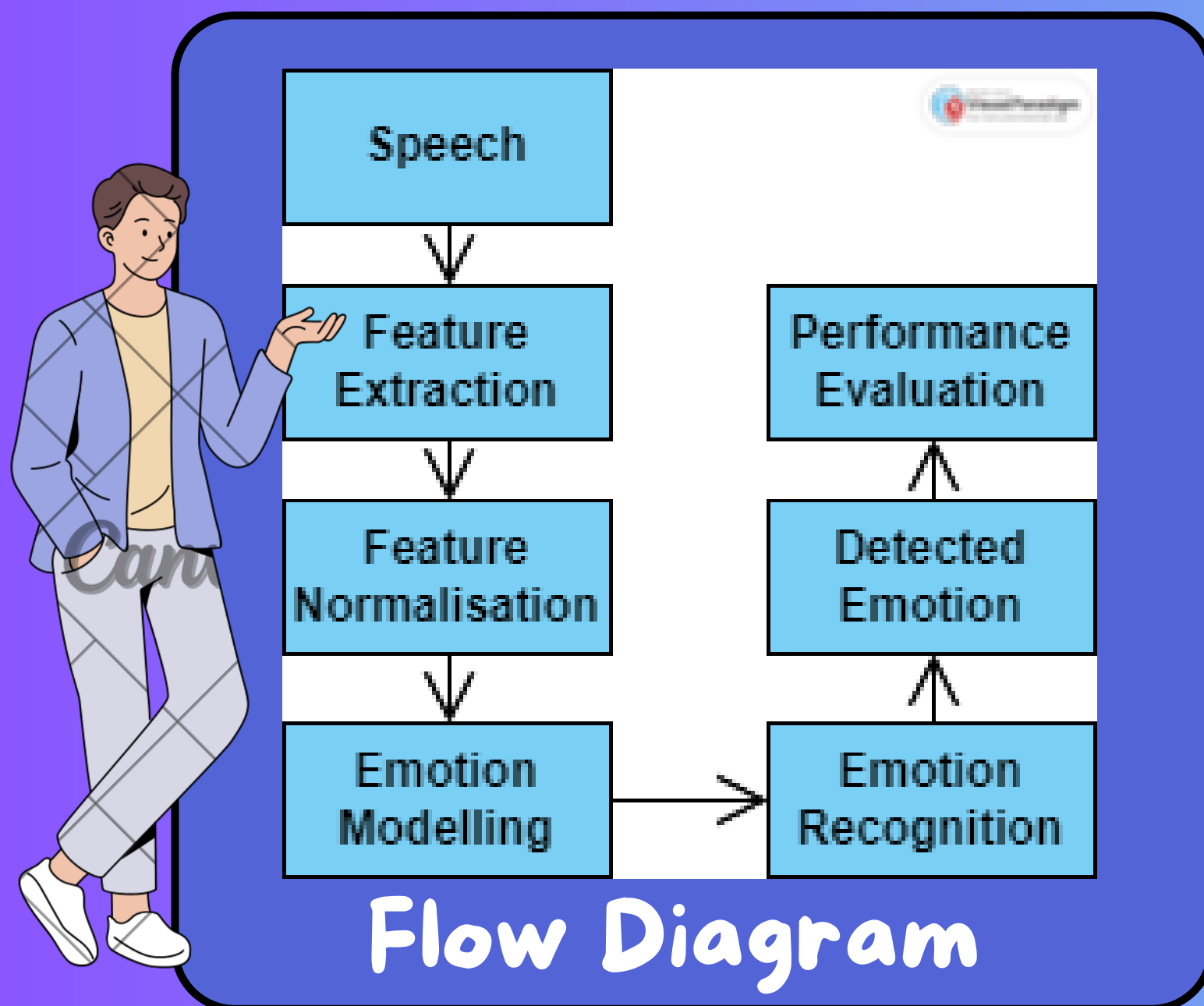


Solution Approach

2



- Dataset loading and preprocessing: Extract features from audio datasets.
- Data augmentation: Enhance dataset diversity with noise addition, stretching, shifting, and pitch changes.
- Model building: Construct a CNN-based model with Keras, incorporating residual blocks and a softmax layer for multi-class classification.
- Training: Utilize Adam optimizer, categorical cross-entropy loss, & callbacks for model optimization & saving.
- Evaluation: Assess model accuracy and performance on training, validation, and test sets.
- Backend: Create a Flask backend for audio data processing and emotion prediction.
- Integration: Connect frontend & backend to enable data exchange & emotion label display.



RESULTS

3

The SER system demonstrates high accuracy and performance in detecting and classifying emotions from speech signals. It shows robustness across different speech styles and accents, indicating its potential for real-world applications.



CONCLUSION

4

The real-time speech emotion recognition project utilized a CNN model integrated with a React frontend to accurately detect emotions from speech audio. Challenges included dataset diversity and model training. Project demonstrates feasibility for practical applications, suggesting future enhancements for broader impact and robustness.