A multimodal evaluation model for improving pedagogical outcomes

Md Shafiqul Islam

Supervisor **Dr Md. Mizanur Rahoman**

26 May , 2025

Introduction

- Importance of evaluating teaching performance in higher education

- Traditional reliance on student feedback — subjective and limited

- Emergence of multimodal systems using AI/ML

Objective: Compare reliability and validity
between traditional and multimodal methods

Research problem and objectives

Problem

Student feedback suffers
from bias and lacks depth

Objectives

- Evaluate reliability and
  validity of both methods

- Analyze correlation between student
  feedback and multimodal evaluation

Literary review

Student feedback

Common but biased by
non-academic factors
Subjective & biased

Auditory Evaluation

Captures vocal tone,
emotion, prosody
Sensitive to Noise & Accent

Vision-Based
Systems

Track movement, gestures,
classroom interaction
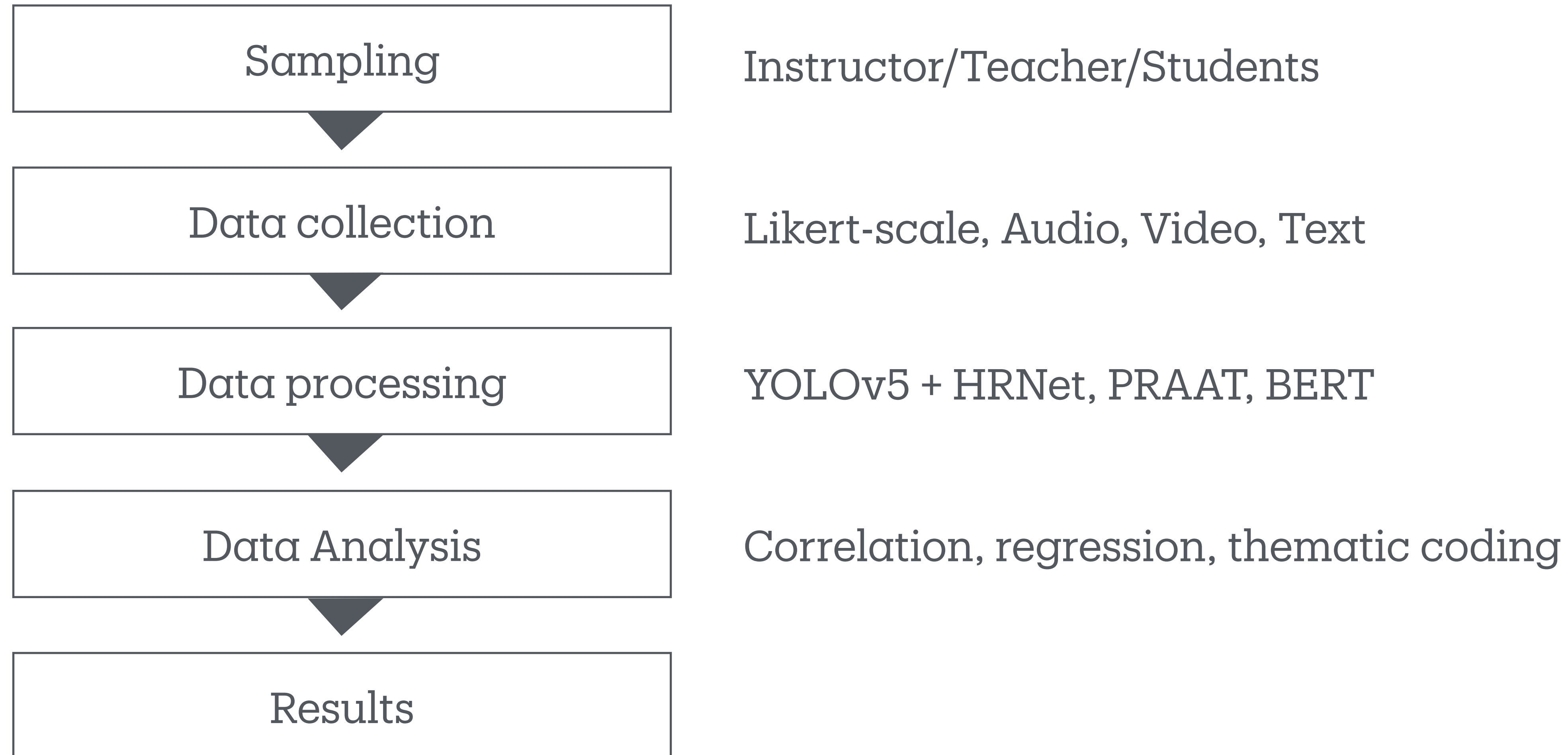Lacks Context

Linguistic/NLP

Analyze instructional clarity,
sentiment, question quality
Misses Non-Verbal Cues

Methodology

Sampling	Instructor/Teacher/Students
Data collection	Likert-scale, Audio, Video, Text
Data processing	YOLOv5 + HRNet, PRAAT, BERT
Data Analysis	Correlation, regression, thematic coding
Results	

Methodology

1. Voice prosody
2. Pitch variation
3. Speech rate
4. Pause patterns

1. Facial expressions
2. Gestures
3. Proxemics
4. Class movement

1. Language complexity
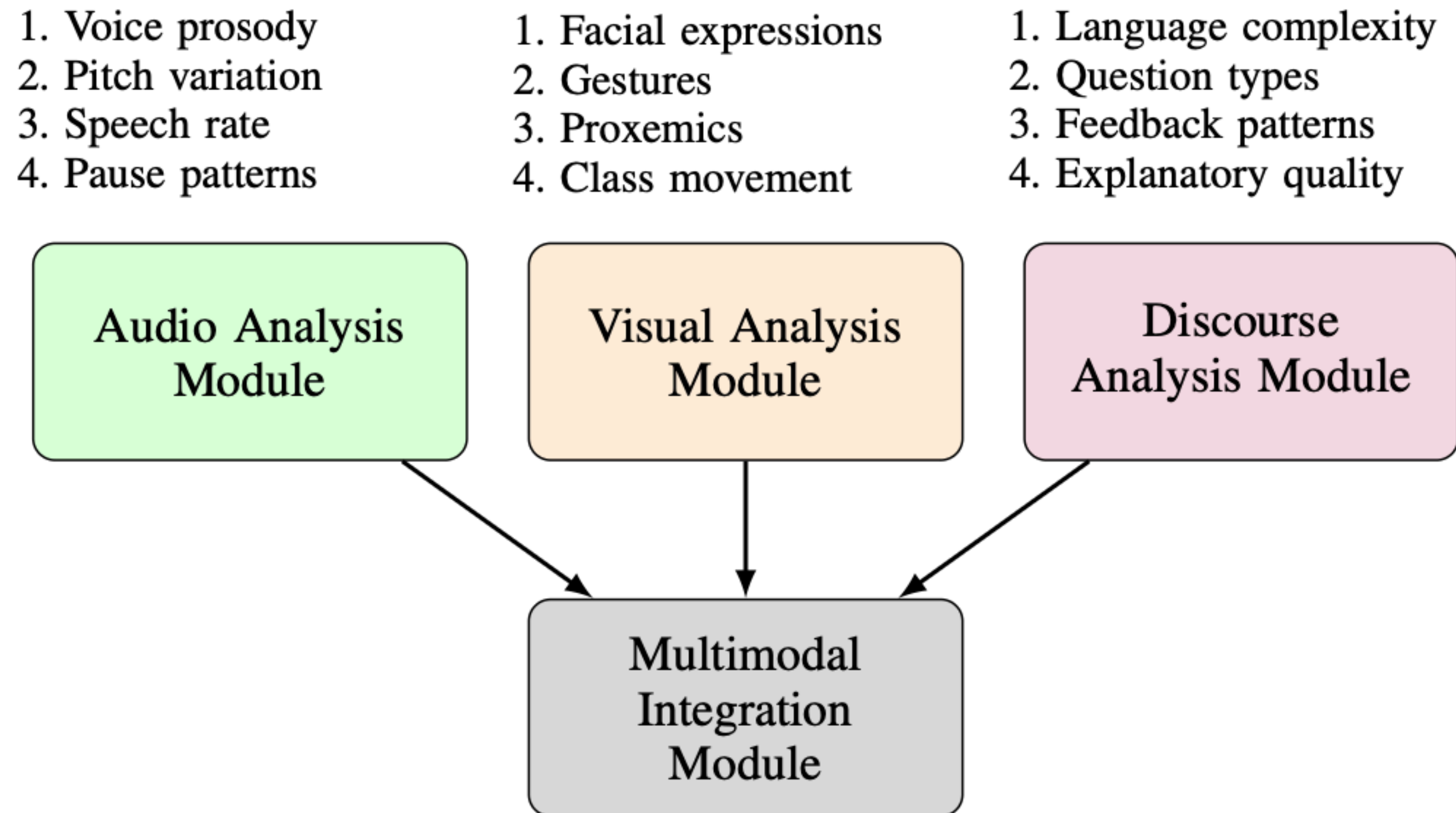2. Question types
3. Feedback patterns
4. Explanatory quality

Fig. 2: Multimodal system components and feature extraction modules.

Experimental setup

Hardwares

- *RTX 3060 + 32 GB RAM*
  For faster processing

Softwares

- *PRAAT*
  For audio processing

- *OpenPose/HRNet*
  For pose estimation

- *Python + hugginface*
  For analysis and modeling

Data collection summery

- *Audio recordings*
  WAV, 44.1Hz

- *Video recordings*
  MP4, 1080p

- *Transcripts*
 TXT

- *Student feedback*
  CSV

TABLE IV: Summary of Collected Dataset

Data Type	Sessions	Total Size
Audio	90	18 hours (12 GB)
Video	90	18 hours (90 GB)
Transcripts	90	1.2M words (8 MB)
Student Feedback	90	1,350 responses (0.5 MB)

Multimodal evaluation system

- *Audio*
  Speech rate, pitch,
  emotion using Wav2Vec2

- *Video*
  Pose estimation, gesture
  tracking (YOLOv5, HRNet)

- *Text feedback*
  Sentiment analysis, question
  type (BERT/DistilBERT)
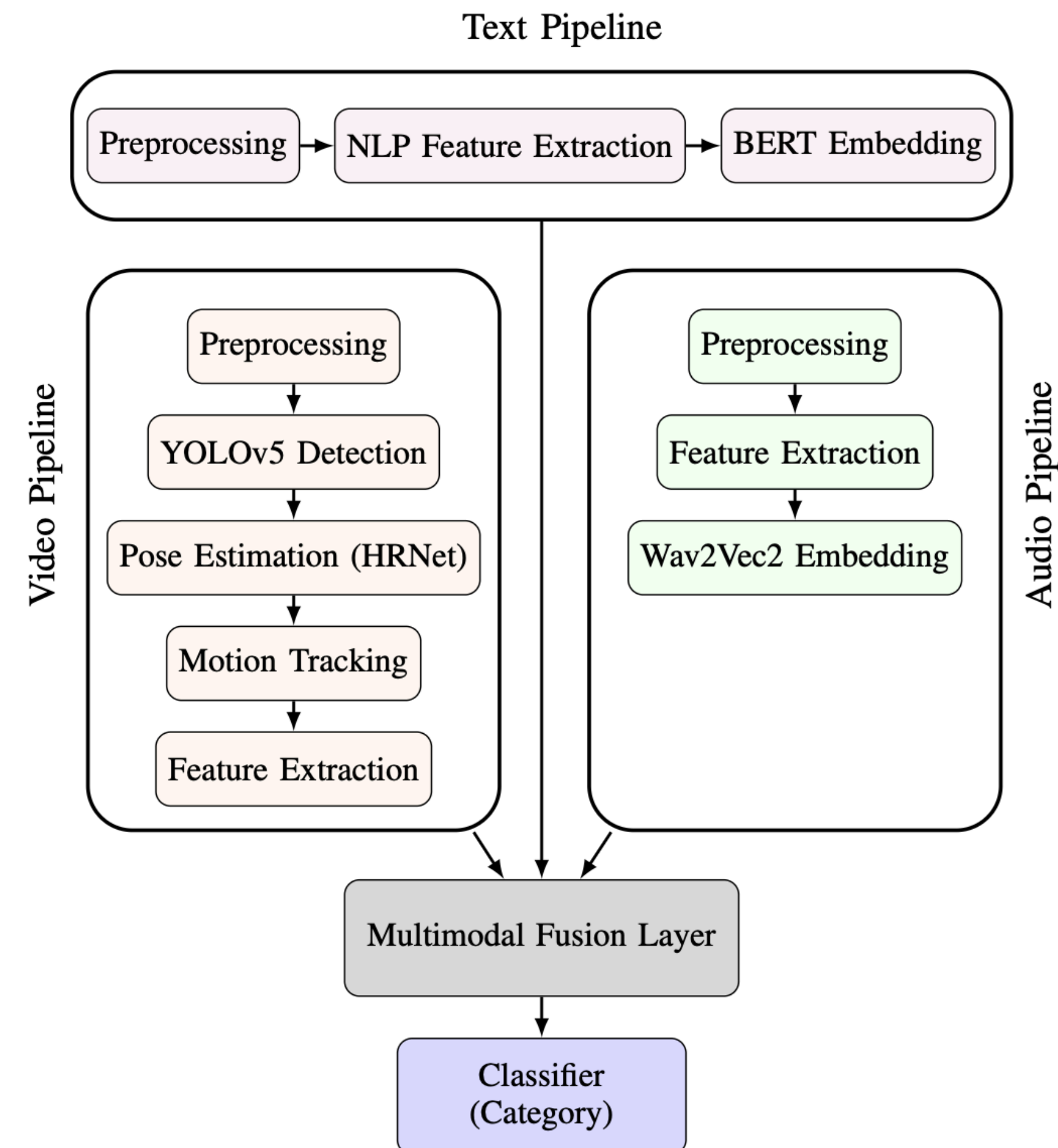
Evaluation metrics

We evaluated engagement, clarity, responsiveness, and overall effectiveness by comparing student feedback with multimodal features such as gestures, speech patterns, response timing, and combined audio-visual-text data.

TABLE III: Evaluation Dimensions and Corresponding Metrics

Dimension	Student Feedback Metric	Multimodal Metric
Engagement	Likert rating (1-5)	Interaction frequency + Voice animation index
Clarity	Likert rating (1-5)	Speech rate + Pause ratio + Example frequency
Organization	Likert rating (1-5)	Topic coherence score + Transition clarity index
Responsiveness	Likert rating (1-5)	Response time + Student engagement rate

Results

To be continued ...

Comparison and analysis

To be continued ...

Discussion and future works

To be continued ...

THANK YOU