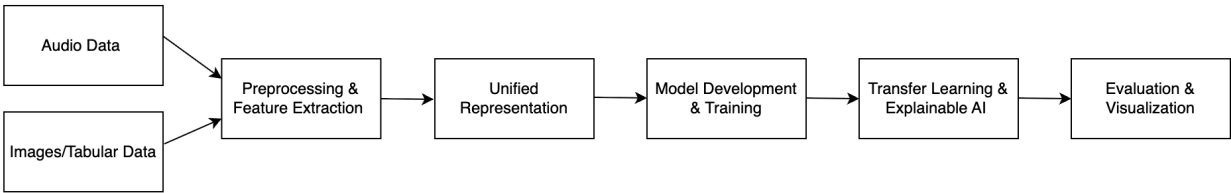# System Design & Technical Implementation

1. System Architecture

The proposed system will be structured as a modular, end-to-end AI framework designed to enable robust, interpretable, and generalizable biomarker analysis in cognitive and mental-health research. The framework will be composed of integrated machine learning (ML), deep learning (DL), and explainable AI (XAI) components to ensure transparency and adaptability across datasets.

Workflow Overview

Publicly available datasets such as *DAIC-WOZ* and *Kaggle* will be utilized, encompassing behavioural, physiological, and speech-based biomarkers. The speech data will be processed to extract acoustic and prosodic features, while tabular and image-derived biomarkers will be standardized and aligned to ensure consistency across modalities. A hybrid AI pipeline will be implemented to integrate machine-learning and deep-learning approaches for comprehensive multimodal analysis. Pre-trained models will be fine-tuned to enhance generalization, and explainable AI (XAI) techniques will be incorporated to maintain transparency and interpretability. The overall model performance will be evaluated, and results will be presented through an interactive visualization dashboard designed to facilitate analysis and insight generation.

Technical Stack

Component	Tools / Frameworks
Programming Language	Python
Frameworks	TensorFlow, PyTorch, Scikit-learn
Processing Libraries	NumPy, Pandas, SciPy, Librosa, OpenSMILE
Visualization	Matplotlib, Plotly, Streamlit
Explainability	SHAP, LIME
Hardware	GPU-enabled environments (Google Colab Pro, institutional servers)

2. Methodology

Preprocessing

All datasets will be subjected to cleaning, normalization, and alignment across speech and tabular modalities. Missing values will be imputed statistically, and acoustic as well as prosodic features will be derived from speech signals. Physiological and behavioural data will be standardized, and dimensionality reduction techniques (e.g., Principal Component Analysis) will be applied to enhance computational efficiency and minimize overfitting.

Modeling Techniques

- A hybrid modelling strategy will be adopted.

- Baseline algorithms, including Logistic Regression, Random Forest, and XGBoost, will be employed to provide interpretable benchmarks for structured data.

- Deep-learning architectures, including CNNs and Transformers, will be utilized to learn spatial and temporal dependencies within the data.

- Pre-trained networks such as Wav2Vec2.0, HuBERT, and BERT will be fine-tuned to enable cross-domain adaptability and improved performance in limited-data scenarios.

Explainability

Explainable AI techniques, including SHAP and LIME, will be integrated into both ML and DL models for global and local interpretation. Counterfactual reasoning will be explored to assess how variations in specific features influence model predictions, ensuring transparency and supporting interpretability in clinical and cognitive research contexts.

Justification

The proposed methodology will be guided by recent research combining structured clinical and speech-based biomarkers for cognitive and mental-health analysis. The approach follows the works of *Ntampakis et al. (2023)* and *Alghowinem et al. (2022)*. By integrating ML, DL, and XAI techniques, a reproducible and interpretable multimodal framework will be developed for future applications in cognitive-health research.

3. Implementation Roadmap

The implementation of the project will be carried out in three structured phases to ensure systematic development, validation, and integration of all components.

Phase	Timeline	Key Activities & Deliverables
Phase 1: Foundation	Month 1	Literature review, dataset selection, preprocessing, and baseline model training.
Phase 2: Modeling & Explainability	Month 2	Development of deep-learning and transfer-learning models; integration of SHAP and LIME.
Phase 3: Evaluation & Integration	Month 3	Model validation, dashboard construction, visualization, and report preparation.

Deliverables:

Trained models, interpretability outputs, and an interactive dashboard will be produced to visualize model performance and feature importance across modalities.

4. Feasibility & Ethical Considerations

Feasibility will be ensured through the utilization of open-access multimodal datasets such as DAIC-WOZ and Kaggle, which will provide sufficient data for model development and validation. Computational requirements will be met through GPU-enabled environments, including Google Colab Pro and institutional servers, allowing efficient training of deep models. The implementation will be conducted entirely using open-source frameworks to ensure scalability, cost efficiency, and reproducibility.

All datasets to be used will remain anonymized and publicly available, guaranteeing compliance with ethical and privacy standards. Model performance will be evaluated across demographic subgroups to identify and mitigate potential bias. Transparency and accountability will be reinforced through the integration of explainable AI methods, promoting responsible and interpretable model development.

5. Evaluation Strategy

Performance Metrics

Model effectiveness will be assessed using metrics such as Accuracy, F1-score, ROC-AUC, and Precision–Recall, depending on the nature of classification tasks.

Validation Techniques

- K-fold cross-validation will be applied to ensure robustness and prevent overfitting, while holdout testing will be conducted to provide unbiased estimates of model generalization.

- Fairness assessments will be performed by comparing subgroup outcomes to confirm equitable model performance.

- Interpretability validation will be conducted by examining the stability of SHAP and LIME explanations across multiple runs.

Through this structured evaluation, the framework will be expected to remain reliable, transparent, and generalizable across multimodal biomarker datasets.

References

[1] Ntampakis, N., Diamantaras, K., Goulianas, K., & Chouvarda, I. (2023). *Predicting the onset of dementia in initially healthy individuals using demographic and clinical data*. IEEE.(*CrossRef*)

[2] Gheorghe, M., Mihalache, S., & Burileanu, D. (2022). *Using deep neural networks for detecting depression from speech*. IEEE. (*CrossRef*)