Project Planning Phase
Technology Stack (Architecture & Stack)

Date	27 September 2023
Team ID	TEAM -592869
Project Name	Lip Reading using Deep Learning
Maximum Marks	5 Marks
Team	Madhav Sonu Kumar Gyaneshwer Jha Jodgudri Pratik Shivamurti

Technical Architecture:

"Technological Architecture is the process of organising software and network systems to achieve particular technological goals. Technical architects supervise the design, guaranteeing effectiveness, safety, and conformity to project objectives. As a guide for intricate technological solutions, this entails choosing hardware, software, data flow, and communication protocols."

Table-1 : Components & Technologies:

S.No	Components	Description	Technology
1.	Data Collection	Gather a dataset of video clips featuring people speaking and simultaneously showing their lip movements.	Digital cameras, webcams, and video recording software for capturing training data.
2.	Data Preprocessing	Cleaning, normalizing, and transforming collected data into a suitable format for modeling.	Python (Pandas), data cleaning techniques
3.	Feature Engineering	Feature engineering for lip reading involves extracting and preprocessing visual lip movement cues from video frames .	Feature selection techniques, domain knowledge
4.	Exploratory Data Analysis (EDA)	Visualizing and analyzing data to uncover patterns and relationships among variables.	Data visualization tools (Matplotlib, Seaborn)
5.	Model Selection	Choose the appropriate deep learning architecture for lip reading accuracy and performance optimization.	Scikit-Learn, TensorFlow or PyTorch
6.	Model Architecture:	Design the neural network architecture to map video frames to corresponding text transcriptions. Implement model to capture temporal dependencies in the lip movements and predict the accurate output text and to to extract features from the video frames.	Recurrent Neural Network (RNN) , Long Short-Term Memory (LSTM) Convolutional Neural Networks (CNNs)
7.	Training and Testing	Splitting the dataset into training and testing sets to train and evaluate model performance.	Cross-validation, model evaluation metrics
8.	Hyperparameter	Optimizing model hyperparameters to improve predictive	Grid search, random search, hyperparameter

	Tuning	accuracy.	optimization tools
9.	Model Evaluation	Assessing model performance using metrics like Mean Absolute Error (MAE) and R-squared.	Scikit-Learn, custom evaluation scripts
10.	Deployment	Implementing the model for real-time predictions, potentially through APIs or within hotel management systems.	Flask, python, HTML, CSS, JS, Bootstrap
11.	Monitoring and Maintenance	Continuously monitoring the model's performance and making updates to maintain accuracy.	Logging, alerting systems, automated pipelines via AWS CloudWatch
12.	Visualization and Reporting	Creating dashboards and reports for interpreting model results and environmental impact.	Data visualization tools (AWS QuickSight, Power BI)

Table-2: Application Characteristics:

S.No.	Characteristics	Description	Technology
1.	Real-time Lip Reading	detect and transcribe spoken words as they are being spoken. This characteristic requires low latency and high processing speed.	CNN , RNN , LSTM
2.	Noise Robustness	Filter out noise and improve transcription accuracy	Noise reduction techniques and robust audio processing methods
3.	Accessibility Features	Offers text-to-speech (TTS) for hearing-impaired users.	Text-to-speech engines (e.g., Google Text-to-Speech) for audio feedback.
4.	Adaptive Learning	The application continuously adapts and improves its accuracy over time by learning from user interactions and feedback	Reinforcement learning algorithms, user profiling, and online learning techniques
5.	Continuous Monitoring (Real-time Data)	Ensures real-time monitoring of data for quick adaptation.	Real-time data streaming and processing (e.g., Apache Kafka).
6.	Interpretability (Model Explanations)	Enhances model interpretability for predictions.	Model explainability techniques (e.g., SHAP values, feature importance scores).

==========END==========