# **Hackathon Project Phases: Audio to Art**

### **Transforming Voice Calls into Visual Creations Using Transformers**

## **Project Title:**

**Audio 2 Art**

## **Team Name:**

Team Codes

## **Team Members:**

* Vigneshwar Reddy
* K.Ritesh
* Rajeev

## **Phase 1: Ideation & Research**

* **Problem Statement: How can we convert voice calls into visually expressive art using AI?**
* **Target Audience: Artists, content creators, designers, and individuals interested in personalized AI-generated visuals.**
* **Research Areas:**
  + **Speech-to-text models (e.g., Whisper, Wav2Vec 2.0)**
  + **Emotional sentiment analysis from audio**
  + **Generative art models (e.g., Stable Diffusion, DALL·E)**
  + **Transformer-based multimodal architectures**

## **Phase 2: Tech Stack Selection**

* **Speech Processing: OpenAI Whisper / Wav2Vec 2.0**
* **Text & Sentiment Analysis: BERT / GPT / LSTMs**
* **Feature Extraction: MFCCs, pitch detection, spectral analysis**
* **Image Generation: Stable Diffusion / ControlNet**
* **Backend: Python (FastAPI, Flask)**
* **Frontend: React.js (for UI to display generated art)**

## **Phase 3: Data Collection & Preprocessing**

* **Voice Data: Use open-source datasets or real-time recorded conversations.**
* **Feature Engineering:**
  + **Extract pitch, tone, frequency, and emotional cues**
  + **Convert speech features into embeddings for art generation**

## **Phase 4: Model Development**

1. **Speech-to-Text & Emotion Analysis**
   * **Use transformers to transcribe audio**
   * **Perform sentiment & tone analysis**
2. **Mapping Audio to Visual Representations**
   * **Develop a latent space mapping from voice features to art styles**
   * **Experiment with GANs / Diffusion Models**
3. **Generating AI Artwork**
   * **Use Stable Diffusion for abstract, surreal, or emotion-based art**
   * **Implement ControlNet for structured outputs**

## **Phase 5: Integration & UI Development**

* **Build an interactive web app**
* **Allow users to upload voice notes or make live calls**
* **Provide real-time visual feedback**

## **Phase 6: Testing & Optimization**

* **Model fine-tuning for better emotion-to-art mappings**
* **Performance optimization for real-time processing**

## **Phase 7: Deployment & Demo**

* **Deploy via Streamlit / Flask on cloud (AWS, GCP)**
* **Live demo for hackathon presentation**