1. Product Features (MVP Scope)

We will use a P0/P1 system for ruthless prioritization. **We only build P0 features.**

Priority	Feature ID	Feature Name	Description & Justification
P0	**F1**	**Web-Based Video/Audio Capture**	A simple frontend component using react-webcam or similar to record video and audio. **Must be reliable.**
P0	**F2**	**Speech-to-Text Transcription**	Use the **Whisper API**. Do not try to build our own. The output is a clean text transcript.
P0	**F3**	**LLM-Powered Content Analysis**	Use the **Gemini API**. We will create a robust prompt that takes the transcript from Whisper and returns a JSON object with: 1) A list of filler words ("um", "ah", "like"), 2) A confidence score (1-10), 3) A clarity score (1-10), and 4) A 1-sentence summary of the key message with customized feedback for the user.
P0	**F4**	**CV-Powered Eye Contact Metric**	We will focus on ONE reliable CV metric: **Eye Contact**. We'll use a library like **MediaPipe Face Mesh** to detect facial landmarks. We can calculate the percentage of time the user's head is oriented towards the camera. This is achievable and provides a powerful, easy-to-understand metric. Note that this stat will also be delivered to the user ALONG with the content analysis stats.

P0	**F5**	**Results Dashboard**	A single, non-scrolling page built in React/Next.js. It will display the results from F3 and F4 in visually distinct cards (e.g., "Filler Words: 7", "Eye Contact: 85%", "Clarity Score: 8/10"). Use a component library like Chakra UI or MUI for speed. **This page must look professional.**
P0	**F6**	**Actionable Feedback Card**	For each metric on the dashboard, we will generate ONE "Next Step". Do this with a second call to the Gemini API with the results to generate a personalized tip. This delivers the "personalized plan" value in a bite-sized, achievable way.
		DEFERRED TO FUTURE:	
P1	**---**	**User Accounts & Profiles**	**OUT OF SCOPE.** The demo will be a single, ephemeral session. No login, no database for users.
P1	**---**	**Complex Gesture/Body Analysis**	**OUT OF SCOPE.** Tracking hand gestures, posture, etc., is a time sink with a high risk of failure. We stick to eye contact.
P1	**---**	**Payment & Subscription**	**OUT OF SCOPE.**
P1	**---**	**Practice History**	**OUT OF SCOPE.** The dashboard shows results for the one session only.

2. Technical Stack & Architecture

Ofc use whatever you're comfortable, feel free to discuss.

- **Frontend: Next.js (or Vite + React)**. Use a component library like **Chakra UI, MUI, or Tailwind UI** to build the UI quickly.
- **Computer Vision: MediaPipe.js** (Face Mesh) running directly in the browser for eye contact analysis. This avoids complex backend CV infrastructure.
- **Backend: Next.js API Routes or a simple Express/FastAPI server.** Its only job is to be a proxy that securely calls the Whisper and Gemini APIs.
- **AI - Transcription: OpenAI Whisper API**.
- **AI - Analysis: Google Gemini API** (specifically the model tuned for JSON output if available).
- **Deployment: Vercel** or **Netlify**. One-click deployment from a GitHub repo.

3. Team Roles & Responsibilities

Clear roles are essential to avoid blocking each other.

- **Mikhai 1 (Frontend Lead):**
    - **Owns:** The Results Dashboard (F5) and the overall UI/UX.
    - **Tasks:** Set up the Next.js project, choose and integrate the UI library, build the static components for the landing and dashboard pages, and integrate the data from the backend.
- **Jaitra (CV & Frontend Integration):**
    - **Owns:** Video Capture (F1) and Eye Contact Metric (F4).
    - **Tasks:** Build the video recording component. Integrate MediaPipe.js to run analysis on the captured video blob. Pass the resulting metric (eye_contact_percentage) to the dashboard.
- **Daniel (Backend & AI Lead):**
    - **Owns:** The backend API server and all external API integrations (F2, F3).

    - **Tasks:** Set up the API endpoint(s). Write the logic to call the Whisper API for transcription and the Gemini API for content analysis. Ensure API keys are secure. Create the data contract (the JSON structure) that the frontend will receive.

- **Eric (Data & Logic Lead):**
    - **Owns:** The "brains" connecting the analyses to the final output (F6).
    - **Tasks:** Design the logic for the Actionable Feedback Card. This might involve creating a "prompt chain" where the output of the first Gemini call is used to generate a more specific prompt for the feedback tip. Work closely with the Backend Lead to structure the final API response.
- Pitching will be done all together, after we are knowledgeable about our own features, we will come together to draft a pitch.