🎯 Project Proposal — Idea #1: Voice-Based Disease Detection Hardware

---

🧭 1. What is the Problem?

Many serious diseases show **early physiological changes in voice** before clinical symptoms appear. Traditional diagnosis often occurs **late, is costly, invasive, and limited to clinical settings**. There's no widely accessible screening tool that uses simple, non-invasive data to detect disease early at scale.

Voice is a biological signal — not just speech — that encodes subtle biomarker changes. Research shows that motor and neurological impairments affecting voice occur early in several diseases, including Parkinson's disease (PD), depression, and metabolic disorders. (Nature)

📊 2. Why is This Important? (Need)

Global Disease Burden

- Neurological conditions are now the **leading cause of ill-health worldwide**, affecting ~43% of the global population (~3.4 billion people) and leading to 11.1 million deaths in 2021. (The Guardian)
- Parkinson's disease cases are projected to **double to ~25.2 million by 2050** due to aging populations and other factors. (The Washington Post)
- Depression is a top cause of disability globally, affecting ~1 in 10 people; it often co-occurs with neurodegenerative disorders. (PubMed)

Early Diagnosis Matters

- Voices show measurable acoustic changes (e.g., jitter, tremor, spectral anomalies) in early disease stages that aren't noticeable to the human ear but detectable with machine learning. (Nature)
- Studies report voice-based models achieving **>90% accuracy** in distinguishing early Parkinson's cases from healthy controls in experimental settings. (PubMed)
- Voice biomarkers have also shown promise in depression classification with AUC up to ~0.93. (PubMed)

Why this matters: A simple voice test expands early detection to **millions who lack access to specialized diagnostics**.

👥 3. Who Is Affected? (Target Audience)

Primary Users

- **Individuals in at-risk populations** (elderly, history of neurological/mood disorders)
- **Remote & underserved communities** lacking access to clinical diagnostics
- **Telemedicine patients** wanting low-cost, frequent monitoring

Secondary Users

- **Clinics & neurologists** for screening support
- **Healthcare systems** aiming to reduce late-stage care costs
- **Insurance providers** seeking preventive tools

📍 4. Where Does This Matter?

- **Globally**, especially in regions with poor access to traditional diagnostic infrastructure (LMICs).
- Can be deployed via **mobile phones + inexpensive hardware**, not limited to clinics.

📅 5. When Does This Project Fit?

- The technology can be prototyped and validated in **~6–12 months** with consumer hardware (mic + DSP + phone app).
- The platform can continue to evolve via **continual ML training with new voice datasets**.

🧠 Scientific Core

The core idea leverages the fact that **neurological and systemic diseases affect the motor control of speech**. This manifests in:

- Acoustic variations (e.g., jitter, shimmer, spectral patterns)
- Prosodic changes (e.g., pauses, pitch variability)
- Altered articulatory dynamics

Recent research demonstrates:

- Machine learning models can classify Parkinson's disease from voice with **>90% accuracy** using MFCCs and spectral features. (PubMed)
- Robotic deep learning models achieve similar or better performance (>92–98% accuracy) in identifying early Parkinson's markers. (arXiv)
- Voice biomarkers correlate with mood disorders like depression with moderate accuracy in large samples (~78–96%). (PubMed)

Key insight: These acoustic markers appear *before* overt clinical symptoms, making voice a powerful **digital biomarker**.

🎯 EDP (Engineering Design Process) Breakdown

📌 A. Problem Definition

- Late diagnosis of debilitating diseases increases morbidity, healthcare cost, and mortality.
- Existing diagnostics require clinical visits, lab tests, and specialists.

Problem Statement:
"Develop an embedded, low-cost voice monitoring system that identifies early disease biomarkers to enable pre-symptomatic health assessments outside clinical settings."

📌 B. Requirements & Constraints

Must-have:

- Real-time or near real-time processing
- Low-cost microphone + DSP circuitry
- Smartphone interface for data collection and feedback
- Robust to recording noise (environmental, language variation)

Constraints:

- Battery life of embedded hardware
- Privacy/security of health data

📌 C. Concept Generation

Core modules:

1. **Microphone + embedded signal conditioning**

2. **DSP for feature extraction**
3. **Machine learning classifier (edge + cloud hybrid)**
4. **Mobile app for UI + personalized feedback**

📌 D. Prototype & Architecture

Hardware

- MEMS microphone (e.g., Knowles)
- Low-power MCU (e.g., Cortex-M series)
- BLE for phone connectivity

Software

- Voice feature extraction (MFCC, jitter, shimmer)
- ML model (TensorFlow Lite / Edge ML)
- Mobile app dashboard

📌 E. Testing & Validation

Validation Plan

- Recruit voice samples from healthy & disease cohorts
- Split dataset (train/validation/test)
- Evaluate accuracy, sensitivity, specificity

Metrics

- Accuracy > 90% target for initial DN (e.g., Parkinson's)
- False positive rate < 5% desirable
- Cross-language robustness testing

🧪 Expected Impact

Impact	Expected Outcome
Healthcare accessibility	Early detection without clinical visits
Patient empowerment	Self-monitoring + preventive action
Cost reduction	Reduced hospital admissions + late-stage care

Global reach	Scalable via smartphones

📌 Five Ws Summary

Question	Answer
What	Voice-based early disease detection embedded system
Why	Enable early, accessible diagnosis; reduce disease burden
Who	Patients, healthcare systems, telemedicine users
Where	Global, especially underserved areas
When	Prototype in 6–12 months; iterative improvement ongoing

🎓 Conclusion

Voice-based disease detection presents a **novel, high-impact**, and **feasible embedded health solution**, validated by emerging research showing strong predictive performance of voice biomarkers for neurological disorders. This project fits a clear need for early, non-invasive screening and has strong potential for **real-world deployment, social impact, and competition success**.

IDEA 1: Voice-Based Disease Detection Hardware

Problem: Neurological, respiratory, and metabolic diseases alter voice *before diagnosis*.

Novelty: Voice is a **biomarker**, not just speech.

System

- Embedded mic + DSP
- Voice stress, jitter, tremor extraction
- App analyzes Parkinson's, depression, COVID-like illness, ALS risk

Why it's big

- Non-invasive
- Zero consumables
- Massive population-scale screening

CATEGORY 1 — NEW BIO-SIGNALS

🔥 Idea 1: Voice-Based Disease Detection

Metric	Score
Novelty	⭐⭐⭐⭐ (4)
Risk	⚠️⚠️⚠️ (3)
Sustainability	🌱🌱🌱🌱🌱 (5)
Profitability	💰💰💰💰 (4)

Pros

- No hardware complexity
- Extremely scalable
- Preventive healthcare
- No wet lab

Cons

- Needs large datasets
- Risk of false positives

1. Voice-Based Disease Detection

Problem Solved: Early detection of neurological, metabolic, and respiratory diseases before clinical symptoms.

Novelty Edge:

- Uses **subtle voice biomarkers** (tremor, jitter, pitch, stress patterns) instead of just speech recognition.
- Non-invasive, continuous, and scalable; can detect **multiple diseases from a single modality**.
- Moves medicine from reactive diagnosis to predictive monitoring.

Competitors:

- **Cognoa**, **Winterlight Labs**, **VoiceSense**, **Sonde Health**
- Mostly focused on **specific diseases** (e.g., Parkinson's or depression) rather than multi-disease detection.

Differentiators:

- Multi-disease platform from **one embedded mic**.

- Embedded DSP + AI on-device → real-time feedback, low latency.
- Designed for **population-scale screening** rather than individual clinical trials.