
Final Year Project Proposal

Project Title:

Name should be decide: **Combating Misinformation Through Voice-Driven Fact-Checking for Low-Literacy Populations**

Problem Statement:

In low-literacy populations, misinformation spreads rapidly through **audio voice notes, manipulated images, and deepfake videos**, especially via WhatsApp and social media. Most fact-checking tools are **text-based**, making them **inaccessible to individuals who cannot read** or have limited digital literacy. This results in public health crises, social unrest, and trust erosion.

Project Objective:






To develop a **voice-first, AI-powered mobile system** that can:

- Accept audio, image, and video input from users
 - Detect misinformation using AI (Speech-to-Text, NLP, Computer Vision)
 - Deliver **simple voice-based feedback** in regional languages
 - Provide an interface for journalists to review and enrich the AI knowledge base
-

Key Innovations:

Feature

Uniqueness

 Voice-only Fact-Checking	Delivers synthesized answers via audio in Urdu, Pashto, Punjabi , etc.
 WhatsApp Bot + App	Users can send voice clips to WhatsApp for instant audio response
 Emotion & Dialect Recognition	Detect urgency + understand local accents and slang
 Offline Mode with Edge AI	Allows fact-checking without internet using TFLite models
 Human-in-the-loop System	Journalists can manually verify or reject AI-flagged content

Technical Stack:

Layer	Tools/Technologies
Frontend	React Native (Mobile App), WhatsApp Bot (Twilio API)
Backend	FastAPI (Python), Django (Admin Panel)
AI/ML	Whisper, HuggingFace Transformers, PyTorch, TFLite, OpenCV
CV	Deepfake detection with FaceForensics++ and DFDC models
Database	Neo4j (Knowledge Graph), PostgreSQL (Users/Content)
Deployment	AWS/GCP for scalable model serving

Dataset Plan:

Public Datasets (Pretraining & Model Baselines):


Purpose	Dataset	Link
Speech (Urdu, Pashto, Punjabi)	Common Voice (Mozilla)	https://commonvoice.mozilla.org/en/datasets
Multilingual Audio	OpenSLR SLR64/69	http://openslr.org/resources.php
Fake News (Text)	FakeNewsNet	https://github.com/KaiDMML/FakeNewsNet

Multilingual Misinformation	WELFake	https://huggingface.co/datasets/zeroshot/WELFake
Deepfake Videos	DFDC (Kaggle)	https://www.kaggle.com/c/deepfake-detection-challenge
Image Forgery	FaceForensics++	https://github.com/ondyari/FaceForensics

Custom Dataset Creation (Research Contribution):

Since **no publicly available dataset exists** that contains **real-world regional misinformation in audio, image, and WhatsApp-style formats**, we will:

- Collect **audio clips, social media rumors, videos, and memes** from:
 - WhatsApp public groups
 - Facebook pages
 - YouTube/TikTok
- Label them using:
 - **Fact-checking organizations** (Geo FactCheck, Poynter, AFP)
 - Human annotators (e.g., journalism students, community leaders)
- Build a unique **Misinformation Voice Corpus in Urdu/Pashto**

 This dataset will be **open-sourced** under academic license after anonymization — a contribution to global misinformation research.

Evaluation Criteria:

- Misinformation detection accuracy (audio/image/video)
- User comprehension (via usability testing with low-literacy participants)
- Audio clarity in regional language feedback
- Time taken to generate fact-check



Future Work:

- Integrate with **national media outlets**
- Expand to cover **more regional languages and dialects**
- Add **voice-to-voice interactive conversation** for deeper understanding



Team Members:

- **Yousaf Maaz** – Lead
- **Rehan** – lead
- **Taha** – lead



Final Note:

This project is not just technically ambitious, but socially impactful. It targets a **real, unsolved problem** with a globally unique, **voice-first AI approach**, and contributes to research by building an original dataset.

Rough work for this:

✅ Refined, Unique Version of Your Idea

🚀 Title:

"AI Saathi: An Audio-First AI Companion for Misinformation Detection in Low-Literacy Populations"

🔥 What Makes It Unique?

Existing tools focus on text-based misinformation detection — even multimodal ones rarely prioritize **audio-first**, **low-literacy-friendly UX**, and **regional language synthesis**.

We'll add **5 global-first innovations**:

💡 1. Community-Trained Voice Fact-Checker ("AI Saathi")

Not done yet:

A **voice-only AI agent** that:

- Talks in **regional/local languages** (Pashto, Punjabi, Urdu, etc.).
- Can **understand audio submissions** (WhatsApp voice notes, fake calls).
- Gives **voice-based fact-checks**, like:
"Bhai, ye bat ghalat hai. WHO ke mutabiq, polio drops nuksan nahi dete."

✅ **Bonus:** Enable **voice conversations** for clarification or follow-up — as done in ChatGPT voice mode.

💡 2. Fake Chain Message Detector (Audio + Image + Text)

Unique idea: Build a system that automatically identifies viral **WhatsApp-style chain messages** and checks if they've been **debunked globally** — using a **vector similarity model** + metadata matching + audio fingerprinting.

Example: “Share this message with 10 people to avoid tax fine” → detected and explained as fake.

3. Regional Dialect + Emotion Recognition

Train the AI to:

- Understand **regional dialects/slang** in speech input (e.g., Pashto/Chitrali/Hyderabadi accents).
- Analyze **emotion** to detect urgency or panic (common in fake health/audio rumors).

✅ This hasn't been done before for low-literacy misinformation use cases.

4. Offline Mode Using Edge AI

Your app can **work partially offline** using:

- Pre-downloaded **local models** (on-device TensorFlow Lite).
- Local **voice-based FAQ detection**.
- Enables fact-checking in **rural or low-connectivity** areas.

✅ *This is a technical novelty.*

5. "Fact-Check via WhatsApp" Bot with Audio Response

A low-literate user can **WhatsApp voice** a clip or image to the bot → AI replies **with a voice note**, not text.

This makes your system **platform-independent** and aligns with **real-world use** in Pakistan, India, Africa, etc.

Updated Architecture (Modular View):

Module	Technology Stack
Voice Input & Output (App)	React Native + Google TTS/STT (local language models)
Audio NLP	Whisper + HuggingFace Transformers
Image/Video Analysis	Deepfake & manipulation detection (PyTorch + OpenCV)
Knowledge Graph	Neo4j + Graph embeddings for similarity detection
Bot API (WhatsApp, Web)	Twilio API + FastAPI
Offline Edge AI	TFLite + Quantized BERT models
Admin Panel (Fact-checkers)	Django + Streamlit or Retool

Bonus Research Extension (for SE Report):

Add an **evaluation metric**:

- Accuracy of audio misinformation detection
 - User trust & accessibility (via surveys)
 - Time to correct misinfo
-

SE Final Year Report Sections You'll Nail:

Section	Strong Points with This Idea
Problem Statement	Social crisis, low literacy, dangerous misinfo
Innovation/Novelty	Audio-first, regional dialects, offline edge AI
Technical Feasibility	HuggingFace, Neo4j, TFLite, FastAPI, PyTorch
Evaluation	Field testing + user studies
Future Work	Expand to other countries/languages

Dataset Information :

Great question — **datasets are the backbone** of your FYP, especially since you're dealing with **speech, misinformation, and visual media**. Since you're aiming for a **globally unique project**, you should **mention both public datasets and plans to build your own**, like this:

Datasets — Where You'll Get the Data

✓ 1. Existing Public Datasets (for Pretraining & Baseline Models)

Speech-to-Text (Audio in Regional Languages):

- **Common Voice (Mozilla)**
<https://commonvoice.mozilla.org>
→ Urdu, Punjabi, Hindi datasets for voice recognition.
- **OpenSLR**
<http://openslr.org>
→ Speech datasets (e.g., SLR64 - Urdu, SLR69 - Pashto).

Misinformation / Fake News:

- **FakeNewsNet**
<https://github.com/KaiDMML/FakeNewsNet>
→ Labeled news articles as fake or real.
- **WELFake (Multilingual Misinformation Dataset)**
<https://huggingface.co/datasets/zeroshot/WELFake>
→ News + claims for training veracity detection models.

Image / Video Manipulation:

- **Deepfake Detection Challenge (DFDC)**
<https://www.kaggle.com/c/deepfake-detection-challenge>

→ 100,000+ videos to detect manipulated faces.

- **FaceForensics++**

<https://github.com/ondyari/FaceForensics>

→ Dataset for forgery detection in videos/images.

2. Custom Dataset Creation (Your Innovation)

Since **no dataset exists yet** for **real-world regional misinformation in audio/video form**, **you'll build your own**:

- Collect real-world voice notes, social media clips, and rumors from:
 - **Public WhatsApp groups**
 - **Facebook posts and local pages**
 - **YouTube/TikTok videos spreading rumors**
- Label them manually with help from:
 - Local journalists
 - University language experts
 - Verified fact-checking sites like **Snopes**, **AFP Fact Check**, **Poynter**, **Geo FactCheck**

You can also **run a survey** asking people to **submit suspicious audio messages** they've received. This would become your **contribution to research**.

How You Can Write It in Your SE Proposal:

"To support training and evaluation of our system, we will leverage publicly available datasets such as Mozilla Common Voice, OpenSLR (for Urdu and Pashto), FakeNewsNet (for fake news classification), and DFDC (for deepfake detection).

Additionally, to address the lack of regional and audio misinformation data, we will create a novel custom dataset by collecting voice notes, images, and videos from real-world sources, and labeling them using human-in-the-loop validation through journalist partners and local fact-checkers."
