# Prathmesh Vaidya

Navi Mumbai, India ,400614| 7666342411 | prathmeshvaidya20@gmail.com | Linkedin | Github |Portfolio

## Introduction

I'm a Data Engineer & AI/ML Engineer passionate about turning complex data challenges into intelligent, scalable solutions. With hands-on experience in Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), I specialize in building AI-driven applications that combine cutting-edge machine learning with robust data engineering.

## Skills & Abilities

- **Programming Languages:** Python, SQL,R ,Java.
- **Machine Learning:**
  - **Algorithms:** Regression, Classification (Random Forest), Clustering (K-Means, DBSCAN), Sequence Modeling (BiLSTM)
  - **NLP:** Retrieval-Augmented Generation (RAG), Large Language Models (LLM), Transformers
- **Python Libraries:** NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn
- **Web Development:** Flask, HTML, CSS

  **Data Visualization**: Power BI, Tableau, MS Excel, Word, and PowerPoint

## Project

**THERAPY CONNECT |** Gemini 2.5 Pro **|** Python, Rag, Llm, Lstm, Whisper, Chromadb, Flask

- Architected an end-to-end mental health support system that provides personalized, empathetic analysis of user's spoken thoughts.
- Engineered a BiLSTM model to classify six core human emotions and implemented a RAG pipeline using MiniLM and ChromaDB to source insights from a corpus of 31 psychology books.
- Integrated Google's Gemini 2.5 Pro to generate reflective, therapy-aligned questions and recommendations based on user input.
- Developed the application using Flask for the backend and JavaScript for the frontend.

**US TRAFFIC ACCIDENT SEVERITY** | Machine Learning | Python, Google colab

- Engineere**d** a predictive model to classify the severity of US traffic accidents, achieving **82.4%** accuracy and a **70.9%** macro average precision.
- Executed comprehensive feature engineering, including handling class imbalance and encoding categorical variables, to enhance model robustness.
- Utilized DBSCAN and K-Means clustering to identify high-risk geographical zones and accident density patterns from the data.

**UNDERSTANDING HUMAN EMOTIONS THROUGH SPEECH |** Machine Learning **|** Python,

- Developed a deep learning model to classify emotions from human voice recordings using the RAVDESS dataset**.**
- Transformed raw audio into Mel-spectrograms and trained a Bidirectional LSTM model, achieving **81.02%** accuracy with a test loss of **0.6589**.
- The project integrates techniques from audio signal processing**,** computer vision, and sequence modeling to identify emotions like happiness, sadness, anger, and fear, showcasing a multidisciplinary approach to emotion recognition through speech.

## Education

**Post Graduate Diploma in Big Data Analytics (PG-DBDA)** | C-DAC, Mumbai | *Expected Aug 2025*

**Bachelor of Engineering in Electronics and Telecommunication** | Terna Engineering College, Mumbai University | *2021 - 2024*