

# Harsha Vardhan Reddy Tamatam

• [harsha.vardhan@rutgers.edu](mailto:harsha.vardhan@rutgers.edu) • +1 617-642-4707 • [linkedin.com/in/harsharshavardhantamatam/](https://www.linkedin.com/in/harsharshavardhantamatam/)

## EDUCATION

- RUTGERS UNIVERSITY**, NEW BRUNSWICK, NJ, USA 09/2023 - 04/2025  
Master of Science in Data Science, Dept of Statistics
- MAHINDRA ECOLE CENTRALE**, HYDERABAD, INDIA 08/2016 - 05/2020  
Bachelor of Technology in Computer Science

## WORK EXPERIENCE

- DATA ANALYST INTERN - LS DIRECT**, NY, USA 06/2024 - 12/2024
- Predicted customer relocation using an imbalanced classification model, increasing targeted ad efficiency by 11%.
  - Designed vector data stores for 15M+ users using LangChain & LlamaIndex, improving search precision and retrievals.
  - Built a Python module for automating SQL to BigQuery migrations with validation checks to ensure data integrity.
- MACHINE LEARNING ENGINEER - VEDANTU LEARNING**, INDIA 11/2021 - 06/2023
- Developed a deep ranking-based recommendation system, increasing CTR by 19% and purchases by 2%.
  - Built BERT-based classifiers, boosting NER and sentiment accuracy by 12%, enhancing adaptive learning.
  - Worked in a cross-functional team with Analysts and Product Managers to build XGBoost models and dashboards to track and reduce churn rate by identifying high-risk users.
- FOUNDING ENGINEER - KHABRI AUDIO**(YC STARTUP), INDIA 12/2020 - 11/2021
- Worked on ML based audio search engine, using MFCC extraction, deep learning embeddings, and contrastive learning.
  - Built data infrastructure from scratch, logging user interactions, storing data, and enabling analytics-driven decisions.
  - Conducted A/B testing and statistical modeling to cluster behavioral insights for targeted user acquisition and retention.
- FOUNDER - AUTODUB**, INDIA 04/2018 - 12/2020
- Founded and led Autodub.in, developing AI-powered speech translation software to enable multilingual accessibility.
  - Scaled AI infrastructure from scratch, enabling rapid video dubbing at 90% lower cost and near-human accuracy.
  - Autodub's technology was acquired by Khabri Audio after national media recognition for dubbing COVID-19 awareness content in Indian languages.

## SKILLS

- **Programming:** Python, Pytorch, FastAPI, Scikit-learn, Transformers, LangChain, Flask, Django, Streamlit
- **AI & ML:** Gen-AI, Machine Learning, Deep learning, NLP, LLMs, RAG, ML System Design, Data Modeling
- **Databases & Cloud:** SQL, NoSQL, VectorDB, BigQuery, Redis, Airflow, GCP, AWS, Kubernetes, CI/CD

## RESEARCH EXPERIENCE & PROJECTS

- FOUNDATION MODELS EVALUATION & USE-CASE MAPPING** 01/2025 - Present  
*AI ETHICS LAB, RUTGERS UNIVERSITY* Advisor - Dr. Mhamed Bettaieb
- Co-Authoring a research study on existing foundation models, comparing architectures and applications across domains.
  - Developing a structured model selection framework, guiding optimal LLM choices for different real-world scenarios.
- CODE-MIXED LANGUAGE TRANSLATION USING DILATED-CNN** 02/2017 - 01/2018  
*SPEECH & SIGNAL PROCESSING LAB, IIIT HYDERABAD* Advisor - Dr. Achal Agarwal
- Designed a Dilated-CNN based machine translation model replacing RNNs, increasing training speed by 30%.
  - Trained Word2Vec embeddings for code-mixed Hinglish text, outperforming English-only text semantics.
- LLM-BASED FINANCIAL ANALYSIS SYSTEM**
- Deployed DeepSeek R1-14B model to analyze credit card statements, identifying essential vs discretionary spending.
  - Built an AI-Agent for financial analysis, integrating PDF parsing, LLM inference, and structured storage for insights.
- INCARCERATION-CRIME DYNAMICS: A MACHINE LEARNING APPROACH**
- Built an XGBoost & Linear models to assess prison cycling's impact on crime rates using socio-economic data.
  - Optimized model interpretability via feature selection, multicollinearity reduction, validated with A/B testing on cohorts.