

Yashraj Shrivastava

+91-8839664523

yashraj170803@gmail.com

linkedin.com/yashraj

github.com/yashraj

Education

Indian Institute of Technology, Madras

B.S. in Data Science And Application

Expected May 2027

Chennai, Tamil Nadu

- **Relevant Coursework (till now):** Machine Learning, Probability & Statistics, Linear Algebra, SQL for Data Science

Projects

TruthTell: Real-time Fake News Detector | Flask, PassiveAggressiveClassifier, NLP, Speech Recognition, Deep Learning

- Designed text-based AI solutions for real-time misinformation detection, achieving an **F1-score of 87.6%**.
- Reduced processing time by **65%** with optimized Flask deployment (latency cut from **8s to 4.5s**).
- Processed over **5000 news articles** to uncover key misinformation patterns.
- Currently expanding prototype to support audio, image, and video verification.

Video Person Tracker | Computer Vision, YOLOv8, DeepFace, OpenCV

- Engineered a real-time facial recognition pipeline with **92% accuracy** using YOLOv8 and DeepFace.
- Increased efficiency by **40%** through frame sampling (0.5s intervals) and GPU inference.
- Reduced false positives by **35%** via custom temporal clustering.
- Achieved **3.2s detection time per video minute** with multi-threading and model quantization.

News Sentiment Analyzer & Hindi Audio Summarizer | FastAPI, VADER, RAKE, MarianMT, gTTS

- Built a real-time NLP pipeline that fetched and processed over **100+ live news articles daily** using **NewsAPI**, generating extractive summaries and sentiment insights.
- Achieved **85%+ accuracy** in sentiment classification using **VADER** and zero-shot learning for contextual validation.
- Automated keyword extraction using **RAKE** and translated summaries to Hindi using **MarianMT**; produced human-like speech with **gTTS**.
- Deployed production-ready REST APIs via **FastAPI** with **CORS**, enabling seamless integration with frontend apps.

Technical Skills

Languages: Python, SQL, C++, HTML, CSS

Mathematics: Probability, Statistics, Linear Algebra

Machine Learning: Supervised and Unsupervised Learning

Deep Learning: Neural Networks, CNNs, TensorFlow, Model Training and Evaluation

Natural Language Processing (NLP): Text Preprocessing, Sentiment Analysis, Named Entity Recognition (NER)

Libraries and Tools: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, Flask, Power BI, Git, GitHub

Data and Deployment: Data Cleaning, Feature Engineering, PCA, Time Series, SQLite, MySQL, Flask API, Streamlit

Certifications

- Pursuing a 4-year certification from PW Institute of Innovation.
- **Data Analytics and Visualization Job Simulation** - Accenture (Jul 2024):
- **Analyzing and Visualizing Data with Microsoft Power BI** - Great Learning (Jan 2024):
- **Pandas and NumPy Libraries** - Infosys Springboard (Feb 2024, valid till Dec 2034):
- **Matplotlib** - Infosys Springboard (Apr 2024, valid till Apr 2030):
- **MySQL 8** - Great Learning (Jan 2024):
- **GitHub** - Infosys Springboard: