Tanaya Pawar

New York, USA | tp2623@nyu.edu | (929)-702-3832 | LinkedIn | GitHub | Portfolio

EDUCATION

New York University New York, USA
Masters in Computer Science, GPA: 3.88/4 Aug 2024 - May 2026
Coursework: Machine Learning, Open Source, Software Engineering, Big Data, Deep Learning, Algorithms

Savitribai Phule Pune University Pune, India
Bachelor Of Engineering in Computer Engineering *with distinction*, GPA: 8.64/10 Aug 2020 - June 2024
Coursework: Data Structures, Object Oriented Programming, Operating Systems, Cloud Computing

EXPERIENCE

New York University, New York, USA |*Grader/Tutor* Starting Sep 2025
- Evaluate assignments & projects for graduate level course programming languages covering OOP, structures
- Will provide feedback on functional programming code, ensuring conceptual understanding and code correctness

Referrio, New York, USA |*Full Stack Developer Intern* June 2025 - August 2025
- Engineered a scalable full-stack application for a fast growing startup using **TypeScript**, **Next.js** and **React.js**, developing reusable UI components with TailwindCSS and ShadCN UI to optimizations dramatically boosted user experience and system performance, reduced data latency, and were deployed via a CI/CD pipeline
- Built and validated **Node.js** and **Express.js** API endpoints with Postman, achieving 100% data consistency
- Applied Agile methodologies & developed a secure, real-time support ticketing system using **Node Express** & **PostgreSQL**, which centralized issue tracking and accelerated internal support resolution by approximately 20%

Wilo Mather And Platt, Pune, India |*Data Analyst Intern* Oct 2022 - Dec 2022
- Engineered data solutions to monitor **financial risk**s and expenses and optimize **$23M** annual budget using **Power BI and SQL APIs** to streamline financial analytics for **cross-functional** teams
- Prepared monthly status reports and presented findings to cross-functional teams along with alert to identify high cost areas & potential cost savings opportunities
- Achieved a **38%** reduction in expenses using data analysis over **2022** by analyzing employee behavior patterns

PROJECTS

CrisisCast |GitHub |*Apache Kafka, Pyspark, MongoDB, Docker, FastAPI, Qdrant*
- Developed real-time crisis detection pipeline using **Kafka**, **Spark**, **Docker**, ingestion from Reddit & Bluesky
- Integrated a locally hosted LLM via FastAPI for crisis classification, achieved low latency inference & filtering
- Engineered a modular architecture with **MongoDB** and **Qdrant** for enriched storage and fast vector search, powering a real-time dashboard with semantic querying and trend visualizations
- Tuned Kafka pipeline for resilience & scalability achieved 800–1000 msg/s throughput with fault recovery

Parameter-Efficient Transformer |GitHub |*PyTorch, Hugging Face, LoRA, Python*
- Fine-tuned **RoBERTa**-base on AG News using **LoRA**, training just 0.4% of weights for parameter efficiency
- Achieved **92.3%** test accuracy in 3 epochs, outperforming full fine-tuning with 99.6% fewer parameters
- Reduced GPU memory by **50%** and training time to under 5 min/epoch, enabling fast, scalable deployment

Comparative Analysis of LLMs |GitHub | *Hugging Face, PyTorch, LLM's*
- Explored BERT and GPT by implementing both architectures from scratch and fine-tuning them for NLP tasks
- Trained scaled-down BERT, GPT on WikiText & fine-tuned on SQuAD, CNN/DailyMail using transfer learning
- Achieved losses of **1.67 (BERT) & 2.97 (GPT)**, highlighting strengths in question answering & summarization
- Implemented a modular **inference pipeline** for **real-time** predictions and model performance

PUBLICATIONS

"Precision Farming Methods Using Machine Learning", **Agro-Advice** |*ICESTM* Feb 3 2024
- Conducted systematic analysis of 24 peer-reviewed studies on ML-driven agricultural recommendation systems
- Identified critical gaps in existing models,insufficient data and limited adaptability to microclimate variations
- Proposed hybrid approach combining Random Forest and XGBoost algorithms, improving prediction accuracy by 17.3% introduced paired cropping optimization demonstrating 22% yield improvement in experimental validation

SKILLS

Programming: Python, JavaScript, C++, R, HTML, CSS, Typescript, SQL, MongoDB
Frameworks: Django, Django Rest Framework, React, NextJs, Nodejs, Expressjs, TailwindCSS, ShadCN UI
Libraries: Streamlit, Tensorflow, Keras, Sci-kit Learn, Pandas, NumPy, NLTK, PyTorch, OpenCV, Matplotlib
APIs and Tools: REST APIs, Apache Kafka, Power BI, Tableau, Microsoft Office, Hadoop, Spark, Linux
Cloud: AWS, Kubernetes, Docker, Snowflake