Santosh Patapati

sapatapatiwork@gmail.com | linkedin.com/in/spatapati | github.com/Soontosh | spatapati.vercel.app

Experience

Machine Learning Research Intern | Part-Time May 2024 – Present
University of Washington *Seattle, WA*

- Explored ML x Nanotech for low-resource SARS-CoV-2 strand classification from electronic biomolecular data
- Surpassed state-of-the-art across all metrics with extremely high statistical significance
- Experimented with various ML techniques, including GNNs & attention mechanisms, optimizing hyperparameters
- Created compiler-level optimizations to improve model speed five-fold, allowing for real-time use
- Fine-tuning open-source LLMs for materials science applications using scraped data
- Experimenting with: Different LM heads, higher quality feature extraction, custom cross-attention layers, explainability, DoRA, curriculum learning, and more through the Google Cloud Platform

Machine Learning Engineer | Contract Jan. 2024 – June 2024
3Blue1Brown

- Built ML pipeline to automate 3Blue1Brown (6M Subscribers) audio translation process for 30+ languages
- Developed, evaluated, and combined several ML models across multiple modalities; Scraped online data
- Created automated tools for 3Blue1Brown employees to translate videos; Made presentations for client meetings

Nanobot Research Assistant | Part-Time June 2024 – July 2024
SMU BAST Lab

- Invited to investigate different propulsion systems' impact on nanobot maneuverability in simulated vascular networks
- Compiled experimental data to create visualizations in Python that facilitated comparative analysis
- Authored a brief research report detailing methdology and findings of my research

Machine Learning Engineer | Contract July 2024 – December 2024
Genloop

- Developed multilingual culturally adaptive WhatsApp-based AI health advisor for under-resourced communities, integrating Ayurvedic medical knowledge and low-resource language NLP
- Implemented backend infrastructure using FastAPI and SQL/GraphQL to support scalable deployment
- Led technical research on dataset creation and model fine-tuning for dialect-specific healthcare advice
- Designed and iterated on a custom RAG pipeline to ground AI responses in culturally-relevant healthcare data

Full-Stack Developer Intern | Full-Time August 2023 – December 2023
AI Camp

- Built automated fraud detection with Django and custom RESTful API, almost $30,000 in savings in one month
- Developed secure and highly-documented JWT-encrypted API to retrieve invoice and customer information
- Created weekly web-based report system using Redis message broker to constantly notify team of potential fraud
- Designed DB schema and ERD for storing customer and invoice information w/ PostgreSQL; Deployed to cloud

Instructor & Course Developer | Part-Time October 2023 – Present
AI Camp

- Instructed & presented to 120+ students, ranging from middle schoolers to undergraduates
- Instructed students on AI ethics and basic deployment, resulting in capstone projects combining several tools
- Developed course for undergrads and high school students on the use of Autoencoders and Transformer architectures

Notable Projects

ML-Based Approach to MDD Diagnosis | *Python, Tensorflow, Pytorch, Django* August 2023 – August 2024

- Developed a novel approach to automated depression diagnosis - 91% accuracy with Leave-One-Out cross-validation
- Pre-processed and augmented hundreds of clinical interviews for training: Extracted biomarkers such as movement of facial muscles, gaze, head rotation, changes in pitch, text sentiment, and more
- Built trimodal, BiLSTM hybrid fusion architecture for diagnosis w/ resource-efficient hyperparameter tuning
- Deployed model to web application using Django, HTML/SCSS, and Zoom API OAuth.

VisualKeras Python Package | *Keras, PIL, Git, Open-Source Software* June 2024 - Present

- Sole maintainer of ML visualization library with 500+ Github stars, 200k+ downloads, and 120+ uses in papers
- Maintaining detailed documentation and changelog across Pypi versions; Following best practices
- Reviewing pull requests & user issues

Publications and Conference Presentations

WebNav: An Intelligent Agent for Voice-Controlled Web Navigation: Wrote preprint presenting a novel voice-controlled navigation agent using a ReAct-inspired architecture, offering improved accessibility for the visually impaired.

Integrating Large Language Models into a Tri-Modal Architecture for Automated Depression Classification: Wrote preprint detailing methodology and results for research project involving multimodal diagnosis of MDD. Featured on "The AI Timeline" and Arxiv "30 Most Popular [ML] Papers in the Past 30 Days".

GenECA: A Generalizable Framework for Real-Time Multimodal Embodied Conversational Agents: Research introducing a robust framework for multimodal interactions with embodied conversational agents, emphasizing emotion-sensitive interaction; Accepted at CVPR 2025 Demo Track

VIZ: Virtual Physical Navigation System for the Visually Impaired: Research project addressing the challenges faced by the visually impaired by utilizing generative AI to mimic human behavior for complex digital tasks and physical navigation; Accepted at CVPR 2025 Demo Track

IEEE MIT URTC: Accepted to deliver a talk on computational psychiatry to attendees at the top undergraduate-level technology research conference; Additionally presented at poster session on computational nanotechnology.

Rice AI in Health Conference, Poster Presenter: Poster presentation on computational psychiatry research at a conference hosted by the Rice Ken Kennedy Institute in Houston, Texas. 50% acceptance rate

The Future of Biology Conference, Poster Presenter: Poster presentation on computational psychiatry research at a conference hosted by the Harvard OpenBio Lab and MIT Biomakers in Cambridge, Massachusets.

Honors and Awards

2024, GripTape Learning Challenge PLUS Scholar: Mentored by industry psychiatry professional for human-computer interaction research project; Top 5% of applicants selected for PLUS scholar

2024, U.S. Innovator Convention, National Finalist: Qualified by winning at the Independent Innovator Challenge; Won additional awards from IFIA and awarded for Most Rigorous Research.

2024, 2nd in Category TXSEF & 1st in Category DRSEF: Placed first at the Dallas regional science fair and placed second at Texas state science fair (top 0.42%).

2024, SMU Texas Workforce Commission Grant: Awarded $2k grant to pursue further research at SMU.

2024, American Psychological Association Achievement in Research in Psychological Science: Formally recognized by the APA for my research in computational psychiatry; Aditionally awarded $300

Technical Skills

Languages: Python, Typescript, HTML/CSS/SCSS, C#, Java, C++, Google App Scripts
Frameworks: Django, React, React Native, Node.js, Flask
Developer Tools: Git, Docker, Redis, Postgres, Firebase, MongoDB, Linux, React, Anaconda, Jupyter, VS Code