Varun Totakura

(848) 667-6729 — varun.totakura@gmail.com — varuntotakura.github.io — Google Scholar @ZbVUfuYAAAAJ

Professional Summary

Passionate Software/ML Engineer with over 4 years of professional experience. Developed ML/Active Learning algorithms as a Research Assistant while pursuing a Master's degree. Excellent team player and individual contributor with strong interpersonal skills, committed to achieving project objectives efficiently.

Technical Skills

- Frameworks: PyTorch, TensorFlow, Scikit-Learn, Keras, LangChain, Hugging Face
- Languages: Python, JavaScript, R, SQL
- Modules: Dask, Numpy, Pandas, PySpark, FastAPI, Ollama, OpenAI, Gemini, Whisper
- MLOps: Jenkins, Docker, Airflow, Kafka, AutoML, Azure AI
- Databases: PostgreSQL, MySQL, BigQuery, BigTable, MongoDB, Cassandra, VectorDB - Chroma
- Clouds: ServiceNow, Google Cloud, Microsoft Azure, Amazon Web Services
- Visualization and Reports: Matplotlib, Seaborn, Plotly, Altreyx, Qlik, Power BI, Tableau
- Version Control: Git, GitHub, GitLab

Professional Experience

AI Engineer - Consultant II, EXL Service Inc., New York, NY (Remote) *October 2024 – Present*

- Drive advanced analytics initiatives in the healthcare sector by leveraging Generative AI techniques, including LLMs, Machine Learning, and Deep Learning.
- Design and develop innovative data analysis, predictive modeling, and AI-powered decision-making solutions to enhance healthcare operations and patient outcomes.
- Developed and deployed code via Jenkins on GCP Cloud Storage Buckets, creating 10+ pipelines with Airflow DAGs to launch DataProc clusters for processing over 1.5 million records and generating predictions, with results stored in BigQuery tables.

AI/ML Engineer - Graduate Research Assistant, Florida State University, Tallahassee, FL *August 2022 – August 2024*

- Engineered a groundbreaking hybrid AI architecture combining TCN, CNN, and GRU to analyze Alzheimer's Disease patterns, achieving a 35% performance boost through domain disparity reduction.
- Developed a state-of-the-art Active Learning Algorithm for Imperfect Oracles with LLMs, optimizing performance through parallel computing with Dask to achieve 4x speedup in model training.
- Pioneered a scalable deep learning pipeline for model correctness across prediction, processing over 300,000 records of 10 diverse datasets with high accuracy.
- Implemented an advanced NLP solution using BERT and GAN networks for precise bug localization in code, analyzing 150,000+ social media posts.

AI Software Developer - Systems Engineer, Tata Consultancy Services, Hyderabad, India *July 2020 – September 2022*

- Spearheaded the development of AI-powered chatbots using Azure Bot Service and Virtual Agent, integrating data from 9+ sources to revolutionize user interactions.
- Orchestrated seamless integrations between ServiceNow and 5 external platforms using REST, SOAP, and Email protocols in an Agile SDLC environment.
- Engineered scalable ServiceNow automations, boosting productivity by over 45% through intelligent process optimization.
- Led a high-stakes data migration project, leveraging AI-driven integrations to transfer 1.5 million records with unprecedented efficiency.
- Architected and maintained 15+ robust applications, incorporating advanced workflows, Service Portal, ACLs, and AI-enhanced scheduled jobs.
- Achieved 95% accuracy in cross-platform data synchronization, ensuring flawless information flow across multiple ServiceNow instances and external systems.
- Pioneered R&D initiatives by designing ML-driven automation solutions with AI Studio, amplifying task efficiency by an impressive 20x.
- Mentored a team of 11 as a peer leader, implementing AI-driven performance tracking strategies that increased individual productivity by 5x.

Research Publications

- Active Learning with Imperfect Oracles using Text Data. *TPAMI - Expected 2024*
- MetaErr: Predicting Error Patterns in Deep Neural Networks using the Meta Model Data. *BMVC - Expected 2024*
- Analysing the behavioural patterns in elderly people suffering with Alzheimer's Disease. *TACCESS - Expected 2024*
- Master's Thesis - A Study on Deep Learning Models for Real World Applications. *August 2024*
- Prediction of Transmittable Diseases Rate in a Location Using ARIMA and GARCH. *November 2021*
- Improved Safety of Self-Driving Car using Voice Recognition through CNN. *October 2021*
- Prediction of Stock Trend for Swing Trades using Long Short-Term Memory Neural Network Model. *March 2020*
- Prediction of Animal Vocal Emotions using Convolutional Neural Network. *February 2020*
- An Integrated Approach to Sentiment Analysis using Machine Learning Algorithms. *April 2020*
- Published another 5 research papers and survey papers. *2019 - 2020*

Awards and Achievements

- Three times Star of the Month, Tata Consultancy Services. *2021 - 2022*
- Five times On the Spot Awards, Tata Consultancy Services. *2021 - 2022*
- Learning Achievement Award, Tata Consultancy Services. *2021 - 2022*
- Ranked 2nd Topper in Academics, Guru Nanak Institutions. *2018 - 2019*
- Peer Reviewed three research papers - ORCID ID - https://orcid.org/0000-0002-5114-5205 *2019 - 2024*
- Google Scholar - https://scholar.google.com/citations?user=ZbVUfuYAAAAJ *2019 - 2024*

Certifications and Trainings

- Completed Deep Learning Specialization taught by Andrew Ng and other instructors. *2020*
- Completed altogether 10 Coursera courses on Machine Learning, Customer Analytics and more. *2020*
- Ethical Hacking, IOT, and Web Development training by Imparta. *June 2019*

Education

Florida State University, Tallahassee, Florida, United States
Master of Science in Computer Science (Thesis) *August 2024*

Jawaharlal Nehru Technological University – Guru Nanak Institutions, Telangana, India
Bachelor of Technology in Computer Science & Engineering *September 2020*