

HARSHINI SAI DONEPUDI

New York, NY | 702-820-9381 | US Citizen

github.com/HarshiniDonepudi | [Linkedin](#) | hsd39@cornell.edu

EDUCATION

Cornell Tech (Cornell University), New York, NY | Merit Scholar | GPA: 4.0 May 2025

Jacobs Technion – Cornell Dual Master of Science Applied Information Science and Information systems – Health Tech Concentration

Computer Vision, Machine Learning Engineering, Data Science in the Wild, Deep Learning, User Interface Design, Web Development

Andhra University College of Engineering (A), Visakhapatnam, India | Best Outgoing Student '23 | GPA: 3.92 May 2023

Bachelor of Science in Computer Science and Systems Engineering

Machine Learning/AI, Cloud Computing, Data Structures and Algos, Object Oriented Programming, Software Engineering, Operating Systems

TECHNICAL SKILLS

Coding Language: Python, R, SQL, Kotlin, Swift, C++, Javascript, MATLAB, Objective-C, Java, HTML, CSS, SAS
Frameworks: Tensorflow, Pytorch, SkLearn, Plotly, Matplotlib, OpenCV, Pandas, Numpy, Matplotlib, Pyspark, Streamlit, Flask
Tools: Xcode, Android Studio, VScode, Git, Azure, AWS, GCP, Dataiku, Tableau, Figma, MS Office, Latex, Unity, BigQuery

EXPERIENCE

Quest Diagnostics – Data Scientist Intern | Remote Sept 2024 – Present

- Developed boosting models to enhance decision-making, including predictive analytics and clinical predictions for chronic diseases using Python, PySpark, SQL, SAS, and Snowflake to streamline analysis and process large lab and medical coding datasets (ICD, LOINC).
- Collaborating with engineering teams to build scalable AI/ML pipelines in cloud environments (AWS, Azure, GCP), utilizing NLP and deep learning to improve operations and created data visualizations with AWS QuickSight and Snowpark for efficient data management.

Weill Cornell Medicine – Machine Learning Research Intern, New York, NY Sept 2023 – Present

- Developed AI-driven medical imaging solutions (PyTorch: UNet) to improve heart donor-recipient matching and extract MRI reports with NLP, boosting diagnostic accuracy by 10%. Integrated ECG data for enhanced predictive insights.
- Built a generalized DICOM preprocessing pipeline to streamline medical image segmentation, supporting both batch and real-time processing using 3D slicer of visualization.

GE Healthcare – EEDP Software Intern, New York, NY May 2024- August 2024

- Optimized cloud resource usage for GE Healthcare's Portrait mobile using FinOps/DevOps techniques. Automated trend analysis dashboards with the TICK stack, Python, and SQL, integrated into CI/CD pipelines via Jenkins for seamless 12-hour updates.
- Collaborated in cross-functional stakeholder teams to align technical insights with business objectives enhancing efficiency by 15%.

Johnson and Johnson – Innovation Medicine – Commercial Data Science Coop, Titusville, NJ Jan 2024 – May 2024

- Developed a multimodal LLM leveraging GenAI and RAGs to automate insight extraction to generate data stories for drugs in therapeutic areas from SHAP plots, enhancing model interpretability to support decision-making and clinical insights for HCPs.
- Built scalable data pipelines with AWS S3 and Sagemaker for model deployment, reducing manual effort by 30%.
- Enhanced HCP Salesforce recommendations by applying MLOps with Dataiku through real-time monitoring and feedback.

Mayo Clinic, Intern – Biostatistics and Machine Learning, Rochester, MN March 2023 – Jan 2024

- Implemented multi model models with CNNs and VIT using PyTorch, Python and R to analyze EMR, genomics, behavioural and biomarker data for Alzheimer's, Diabetes, and Breast Cancer (TCGA), collaborating with stakeholders for real-world validation.
- Played a central role in fine tuning and refining methodologies to achieve a successful accuracy improvement of 4%

Shibnobi – Software, Technology Developer, Azle, TX August 2022 – Jan 2024

- Developed cross platform applications for projects: Shibnobi Play, Foundation Token, Torqued and Tuvoznov for IOS and Android.
- Led the end-to-end development of AI-driven cross-platform applications, implementing gaming-based predictive analytics and NLP models to enhance user engagement and content recommendations.

PROJECTS

- Wound Whisperer** (Freelance Project for Wound Expert Care Ltd | Python, Azure Databricks, PyTorch) July 2024 – Ongoing
Developed an AI multimodel tool for wound detection and severity assessment using Azure Databricks. Processed 200,000 images for segmentation and classification from EHR Data. Integrated CoreML for real-time iOS deployment, reducing latency by 25%, and added features like wound tracking, body mapping, and healing progress visualization.
- InsuloCarb** (Python, CoreML, GenAI, TensorRT, Swift, HealthKit, ARKit). April 2022 – Ongoing
Developed a computer vision app using ARKit for volume estimation and AI to estimate meal carbohydrates and recommend insulin doses. Enhanced glucose predictions 2-4 hours in advance, factoring in meal sensitivity for precise dosing. Integrated GenAI for tailored meal suggestions and synced with Dexcom systems and Insulin Pumps for real-time compatibility and alerts.
- Adaptive AI Pulse Oximetry Software | Bryant Bionics** (Python, ML, CV, color theory, BigQuery) August 2024 – December 2024
Developed adaptive calibration algorithms using a custom skin tone scale, leveraging EHR data and PPG sensor data to enhance SpO2 accuracy. Collaborated with clinicians to validate performance and meet regulatory standards.

PUBLICATIONS

- A Comprehensive Study on Accident Detection Techniques - DELCON '22(IEEE conference), IEEE Xplore April 2022
- Insulin Usage and Practices in Children and Adolescents with Type 1 Diabetes Endocrinology and Metabolism July 2021