PRIYANSHU SINHA ⬈

Data Science and Machine Learning Professional with 3+ Years' Experience

Seeking Full Time Data Scientist / ML Engineer Role

📞 +1-317-982-4563 ✉ priyanshu_sinha@outlook.com in linkedin.com/in/pri2si17 ○ github.com/pri2si17-1997

Education

Indiana University-Purdue University Indianapolis   **Indianapolis, Indiana**

Master of Science in Health Informatics   *May 2023*

Jaypee Institute of Information Technology   **Noida, India**

Bachelor of Technology in Computer Science and Engineering   *May 2018*

Experience

Labcorp Drug Development   **Durham, North Carolina**

Student Intern - Data Science   *September 2022 - Present*

- Creating a tool to assess the deep neural network black box models explanability in order to comprehend the aspects influencing prediction and to meet stakeholder needs using Learning Interpretability Tool (LIT).
- Making a survival analysis forecasting model to forecast kits in time for clinical kits sent out.

Labcorp Drug Development   **Princeton, New Jersey**

Data Science Intern   *May 2022 - August 2022*

- Performed time-series data analysis, created a deep LSTM network for weekly clinical screening rate forecasting, and examined the effect of COVID-19 on screening rate using TensorFlow and AWS Sagemaker.
- Developed a prototype for augmented annotation to categorize patient emails with a few samples using the BERT model and Hugging Face framework, which resulted in a 10x reduction in manual labeling work without noticeably affecting the model's performance.

Indiana University-Purdue University   **Indianapolis, Indiana**

Graduate Research Assistant   *August 2021 - Present*

- Developing a simulation tool to calculate departmental costs, faculty requirements, and student intake to reduce manual efforts by 40%.
- Made a Virtual Reality teaching tool prototype in collaboration with Emory University researchers and surgeons to scale immersive learning during surgical procedures. By enabling VR annotations of surgical process scenes via a voice and hand controller, the prototype supports live interaction with students and doctors during procedures.
- Optimized deep learning imaging models using INT-8 quantization, achieving a 75% reduction in size, resulting in reduced memory consumption and 4x gain in inference speed for edge devices such as edge devices using TensorFlow Model Optimization Toolkit.
- Made tutorials for students and healthcare professionals to understand the application of machine learning and deep learning algorithms on healthcare data using MIMIC-IV dataset, Postgress database and TensorFlow.

Mentor Graphics India Pvt. Ltd. (Siemens DISW)   **Noida, India**

Senior Member Technical Staff   *June 2018 - August 2021*

- Created LoLa (Mentor's Proprietary Low Latency Framework) components to cluster Lidar's point cloud and project it on image for classification and detction using Python and C++.
- Created a semi-automated annotation tool to label data for autonomous cars, including classification, detection, and segmentation tasks of scenes leveraging deep learning models to generate pseudo-labels, reducing manual labeling efforts by 60%.
- Compared evaluation performance of SOTA object detection models like YOLO V3, RetinaNet with proprietary hierarchical classifier on KITTI and simulated data from Simcenter Prescan.
- Made web-based tool to scan for security vulnerabilities (CVEs) in packages, used in Linux based operating systems for automobiles, and checked if security patch is applied or not, reducing manual efforts by 75%.

Skills

Technologies : Machine Learning, Deep Learning, Data Structures and Algorithms, Computer Vision, Natural Language Processing, Time Series Analysis, Database Management System, REST APIs, Object Oriented Programming
Languages: Python, C++, C, R, JavaScript
Databases: SQL, MySQL, PostgreSQL, MongoDB
Libraries: Tensorflow, PyTorch, Keras, Numpy, Pandas, Scikit-Learn, Scipy, OpenCV, Matplotlib, Fastai, NLTK, Plotly, Dask, Flask, Django, Apache Spark, Simpy, Pytest, React
Operating System: Linux, Windows
Developer Tools: VS Code, Git, Docker, Jupyter Notebook, MATLAB(Basic), RStudio

Publications

1. **Priyanshu Sinha**, Sai Sreya Tummala, Saptarshi Purkayastha, and Judy Gichoya. Energy efficiency of quantized neural networks in medical imaging. In *Medical Imaging with Deep Learning*, 2022

2. **Sinha, Priyanshu**, Judy W. Gichoya, and Saptarshi Purkayastha. Leapfrogging medical ai in low-resource contexts using edge tensor processing unit. In *2022 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*, pages 67–70, 2022

3. Judy W. Gichoya, **Priyanshu Sinha**, Melissa Davis, Jeffrey W. Dunkle, Scott A. Hamlin, Keith D. Herr, Carrie N. Hoff, Haley P. Letter, Christopher R. McAdams, Gregory D. Puthoff, Kevin L. Smith, Scott D. Steenburg, Imon Banerjee, and Hari Trivedi. Multireader evaluation of radiologist performance for COVID-19 detection on emergency department chest radiographs. *Clinical Imaging*, 82:77–82, 2021

4. Areeba Abid, **Sinha, Priyanshu**, Aishwarya Harpale, Judy Gichoya, and Saptarshi Purkayastha. Optimizing Medical Image Classification Models for Edge Devices. In Kenji Matsui, Sigeru Omatu, Tan Yigitcanlar, and Sara Rodríguez González, editors, *Distributed Computing and Artificial Intelligence, Volume 1: 18th International Conference*, pages 77–87, Cham, 2021. Springer International Publishing

5. Pradeeban Kathiravelu, Puneet Sharma, Ashish Sharma, Imon Banerjee, Hari Trivedi, Saptarshi Purkayastha, **Sinha, Priyanshu**, Alexandre Cadrin-Chenevert, Nabile Safdar, and Judy Wawira Gichoya. A DICOM Framework for Machine Learning and Processing Pipelines Against Real-time Radiology Images. *Journal of Digital Imaging*, 34(4):1005–1013, 2021

6. Ananth Bhimireddy, **Sinha, Priyanshu**, Bolu Oluwalade, Judy W Gichoya, and Saptarshi Purkayastha. Blood Glucose Level Prediction as Time-Series Modeling using Sequence-to-Sequence Neural Networks. In *CEUR workshop proceedings*, 2020

7. Saptarshi Purkayastha, Ananth R Bhimireddy, **Sinha, Priyanshu**, and Judy W Gichoya. Using ImageBERT to improve performance of multi-class Chest Xray classification. 2020

8. Imon Banerjee, **Sinha, Priyanshu**, Saptarshi Purkayastha, Nazanin Mashhaditafreshi, Amara Tariq, Jiwoong Jeong, Hari Trivedi, and Judy W Gichoya. Was there COVID-19 back in 2012? Challenge for AI in diagnosis with similar indications. *arXiv preprint arXiv:2006.13262*, 2020

9. **Sinha, Priyanshu**, Saptarshi Purkayastha, and Judy Gichoya. Full training versus fine tuning for radiology images concept detection task for the ImageCLEF 2019 challenge. In *CEUR workshop proceedings*, 2019

Open-Source Contributions

Google Summer of Code (GSoC) Mentor @ LibreHealth | *Python, ML, Unity, VR*   **May 2018 - Current**
- Mentored GSoC students to develop UI for Virtual-Reality application and integrate machine learning models for classification, detection, and segmentation of disease in a VR environment for immersive learning during medical procedures. This app includes features such as voice chat, live streaming and video playback.
- Mentored a GSoC student to migrate the LibreHealth EHR application from native PHP to Django for modular code and ease in test-driven development.

Awards and Achievements

- Our team (of 3 members) won **1**st place in the AT&T 5G Sports hackathon in <u>Fan Engagement</u> category (held in Indianapolis, USA) where more than 25 teams participated.
- Our team (of 2 members) was selected in **top-50** in Hackerearth IndiaHacks hackathon (2016) out of 1500 teams globally.

Projects

Explainable M5 Accuracy Forecasting | *Python, Tensorflow, Time Series*   **December 2022**
- Created LSTM model to forecast 1 day and 28 day unit sales on M5 Forecasting - Accuracy dataset and explained the prediction of model using what-if-tool via counterfactual analysis.

Explainable AI in Radiology Image Analysis Pipeline | *Python, Tensorflow, Computer Vision*   **December 2022**
- Created prototype for explainable AI in radiology image analysis using NIH-14 chest x-ray dataset. Densenet121 model is used to classify the disease and SHAP values are used to explain the prediction. The explainable AI will help radiologist comprehend the prediction and gain trust in black box AI.