Sinjoy Saha

sinjoysaha555@gmail.com
linkedin.com/in/sinjoysaha github.com/sinjoysaha sinjoysaha.github.io Google Scholar

EDUCATION

University of Calcutta Aug 2016 - Sep 2020
B. Tech. in Electronics and Communication Engineering CGPA - 9.3/10
Courses include *Analog & Digital Circuits, Control Theory, Digital Signal Processing, Computer Organization & Architecture, Data Structures & Algorithms, Microprocessors and Communication Theory.*

WORK EXPERIENCE

Siemens Healthineers **Bangalore, Karnataka**
Senior Engineer - Research and Technology Mar 2023 - Present

- Developed a semantic segmentation pipeline for calculating **tumor-stroma ratio (TSR)** of histopathological whole slide images (WSI) for breast cancer based on **DeepLabV3+** and **ResNet50** backbone. Optimized pipeline using ONNX, pruning, quantization, normalization, and background separation. Achieved Dice coefficient of **0.85** and IoU of **0.63**.
- Introduced a novel algorithm for detection of **semantically uncertain regions** in WSIs based on ensembled model divergence, enabling minimal-input feedback loop for active model learning.
- Led a team to migrate the MI service analytics tool from full-text SQL search to a language model-based **semantic search**. Developed custom API using Python, BERT, Milvus VectorDB, and fastAPI, and deployed backend on AWS cloud, leading to savings of **$20,000** per annum and onboarding of new modalities.
- Developed an unsupervised **topic modeling** method based on BERT and HDBSCAN and created a Power BI dashboard to identify semantic and domain-relevant failure modes, leading to **redesign of autoQC** sub-component in SPECT.

Engineer - Research and Technology Dec 2021 - Feb 2023

- Developed a novel set of image features for an ML classifier and an ROI-agnostic artifact segmentation model for **automated grading** of SPECT Gamma Camera QC images. Improved F1-score from **0.46** to **0.75** compared to baseline uniformity metrics, resulting in a successful patent filing.
- Developed an unsupervised method based on Word2Vec embeddings and DBSCAN for **clustering domain-relevant words**. Increased cluster count from **15** to **346**, improving baseline full-text SQL search for PET/SPECT service tickets.
- Developed an end-to-end **cause and action phrase** extraction pipeline from service tickets using roBERTa and NER. Increased average ticket coverage in search results from **14%** to **52%** compared to earlier POS tagging model.

Cognizant Technology Solutions **Kolkata, West Bengal**
Programmer Analyst Trainee Nov 2020 - Sep 2021

- Developed and executed automated test scripts using Java, Selenium, REST API, HTML, CSS, and JavaScript to ensure quality of API and web applications. Contributed to development of the automation testing framework.
- Actively participated in the test development process for web and mobile applications, including requirements analysis, test case design, and test execution. Worked closely with developers to identify and resolve defects.

INTERNSHIPS

Plant Vision Lab, University of Nebraska-Lincoln **Remote**
Research Intern *under Dr. Sruti Das Choudhury, Prof. Ashok Samal* Jul 2021 – Dec 2021

- Developed two novel algorithms to predict onset of drought stress in cotton plants using **dynamic time warping (DTW)** on computed phenotypes and leveraged **1D-CNN** and **LSTM**-based network for predicting propagation of temporal stress using hyperspectral imagery. Predicted stressed pixels showed F1-score of **0.98** and average Pearson correlation coefficient of **-0.85** with soil water content. The work led to a publication in a peer-reviewed journal.

Sensordrops Networks Pvt. Ltd., IIT Kharagpur **Remote**
Research Intern *under Dr. Rituparna Saha, Prof. Sudip Misra* Jul 2021 – Dec 2021

- Introduced a novel algorithm for Federated Learning in **non-IID** datasets based on K-Means clustering of client data statistics. Surpassed accuracy of baseline Federated Averaging (FedAvg) by **2%** and reduced aggregation time by **67%** compared to baseline clustering on local weights. Currently working towards a publication.

AI Lab, Jadavpur University **Kolkata, West Bengal**
Summer Research Intern *under Prof. Amit Konar* May 2019 - Jul 2019

- Project conducted under **INAE** Student Mentoring Program under INAE Fellows. Programmed a 6-DOF Robotic Arm from Kinova (JACO) in C#. Developed an **Obstacle Avoidance and Trajectory Planning system** for 3-DOF Robotic Arms based on the A* Heuristic Search algorithm and kinematics. Verified the algorithm using simulations on MATLAB and submitted a technical report to INAE, upon which the stipend was disbursed.

IoT Lab, University of Calcutta **Kolkata, West Bengal**

Summer Research Intern *under S. Basu, Dr. S. Pandit, Dr. S. Barman (Mandal)* May 2019 - Jul 2019

- Implemented an **IoT-based health monitoring system** using Arduino UNO, different non-invasive sensors, Wi-Fi module, and Thingspeak IoT platform to read health parameters, display it on an LCD, and transmit data to cloud for secure storage and future analysis. The work resulted in a conference paper that was presented at the International Conference on Computational Intelligence in Pattern Recognition (CIPR), 2019, IIEST Shibpur.

PATENT & PUBLICATIONS

- Daga S., Saha S., Crawford T. E., Khan K., Morris B. Patent filed in April 2023 (USPTO). Artifact Segmentation and/or Uniformity Assessment of a Gamma Camera. Under review.
- Das Choudhury S.†, Saha S.†, Samal A., Mazis A. and Awada T. (2023) Drought stress prediction and propagation using time series modeling on multimodal plant image sequences. *Front. Plant Sci.* 14:1003150. †Equal contribution. https://doi.org/10.3389/fpls.2023.1003150
- Basu S., Saha S., Pandit S., Barman (Mandal) S. (2020) Smart Health Monitoring System for Temperature, Blood Oxygen Saturation, and Heart Rate Sensing with Embedded Processing and Transmission Using IoT Platform. In: Das A., Nayak J., Naik B., Pati S., Pelusi D. (eds) Computational Intelligence in Pattern Recognition. Advances in Intelligent Systems and Computing, vol 999. Springer, Singapore. https://doi.org/10.1007/978-981-13-9042-5_8

PROJECTS

Brain Tumor Segmentation from MRI using PSPNet | *Python, CNN, TensorFlow/Keras* | [Kaggle] Apr 2021

- Motivation was to show the efficacy of Pyramid Scene Parsing Network (PSPNet) for segmentation of brain tumors from MRI data. Built using TensorFlow/Keras in Python. Achieved Dice coefficient of 0.66 and IoU of 0.552.

Car Price Prediction ML WebApp | *Python, Random Forest Regressor, Scikit-Learn* | [GitHub] Mar 2021

- Built a machine learning web application for predicting price of used cars using Random Forest Regressor model. Implemented using Python, Flask, HTML, Bootstrap and hosted on Heroku cloud.

Quadrone, IIT Roorkee | *Arduino UNO* Mar 2018

- Built a radio-controlled quadcopter by implementing the flight controller using Arduino UNO, Flysky CT6B 2.4GHz 6CH Transmitter, and other off-the-shelf components. Implemented PID controller and calibration system for the drone. Participated in Quadrone Event at the national tech-fest of IIT, Roorkee, Cognizance 2018.

CERTIFICATIONS

- **Structuring Machine Learning Projects** - Coursera (Jun 2021) [Link]
- **Neural Networks and Deep Learning** - Coursera (May 2021) [Link]
- **Sequence Models** - Coursera (Apr 2021) [Link]
- **SQL for Beginners: Learn SQL using MySQL and Database Design** - Udemy (Oct 2020) [Link]
- **Machine Learning** by Prof. Andrew Ng, Stanford - Coursera (Aug 2018) (audited)
- **Programming, Data Structures and Algorithms using Python** by IIT Madras - NPTEL (Mar 2017) *Roll No: NPTEL17CS1026450003AN*

TECHNICAL SKILLS

Domains: Data Science, Medical Image Processing, Machine Learning, Deep Learning, Natural Language Processing
Languages: Python, Java, C/C++, MATLAB, SQL, HTML/CSS, JavaScript
Software & Tools: Linux, SQL Server, Milvus DB, Git/GitHub, VS Code, LaTeX, Arduino
Libraries/Frameworks: TensorFlow/Keras, PyTorch, ONNX, nltk, spaCy, OpenSlide

ACHIEVEMENTS

SCM Hackathon, DX, Siemens Healthineers Feb 2023

- Achieved 2nd position out of 44 participating teams from two countries (India and Slovakia) in the 24-hour Supply Chain Management Data Science Hackathon organized by Diagnostics (DX) department within Siemens Healthineers.

JBNSTS Senior Scholar, Govt. of West Bengal Dec 2016

- Qualified in the Jagadis Bose National Science Talent Search (JBSTNS) Examination.
- Received a monthly scholarship and annual book grant throughout undergraduate studies.

EXTRACURRICULARS

- Virtual invited talk on "Introduction to AI and Data Science: Navigating the vector space" (May 2023) for first and second-year undergraduates, organized by the AI & Robotics Club and IEEE Student Branch, University of Calcutta.
- Organized Annual Chess Tournament (Dec. 2019) at Institute of Radio Physics and Electronics, University of Calcutta.
- Volunteered in the photography team for the 3rd Science and Technology Exhibition and Competition (Oct. 2018) organized by IEEE Student Branch, University of Calcutta, along with IEEE Photonics Society Kolkata Chapter.