SOUNAK RAY

+91 · 99579 · 85181 ◇ sounakray1997@gmail.com ◇ LinkedIn ◇ Homepage ◇ Google Scholar

Education

- **Columbia University** — New York City, NY
  Master of Science in Computer Science; Specialization: Machine Learning — *Aug 2021 – Dec 2022*
- **Indian Institute of Technology Guwahati** — Guwahati, India
  B. Tech. in Electronics and Communication Engg., Minor: Computer Science; GPA: 8.44/10.0 — *Jul 2016 – Jun 2020*

Experience

- **Sprinklr** — Gurgaon, India
  Data Scientist, Machine Learning Team — *August 2020 - Present*
  - **Customer Feedback**: Worked on incorporating client feedback in our training datasets using similarity search libraries like Faiss and baseline models. This helped in improving our training sets and achieve an improvement of 10% F1-Score on our validation datasets.
  - **Liveness Verification**: Finetuned an Attention Convolutional LSTM network with only 100 training videos for detecting whether a person in a video is alive through hand wave detection and eye blink detection.
  - **Insights**: Worked on generating product insights from millions of consumer feedback and social media conversations. Used Transformers-based architectures to perform Named Entity Recognition and Text Classification.
  - **Training, Deployment and Maintenance**: Responsible for understanding the data and the requirements of various client specific requests, training of the model and deployment using Jenkins, Docker and Kubernetes. Also responsible for maintaining the deployments and optimize the resource utilization.
- **Sprinklr** — Gurgaon, India
  Data Science Intern, Machine Learning Team — *June 2020 - August 2020*
  - **Sentiment Classification**: Created a word embedding model to find embeddings of emojis along with other words in a message based on the Unicode of an emoji. Used the Hierarchical Attention Network with pre-trained word2vec embeddings to train the model. Improved the sentiment classification F1-score on messages with emojis by 11%.
- **Indian Statistical Institute, Kolkata** — May 2018 - July 2019
  Research Intern, CVPR Unit — *Supervisor : Dr. Utpal Garain*
  - Developed a Residual Capsule Network for Skin Lesion Classification which was fine-tuned with ImageNet weights and compared the results with ResNet50. Achieved an improvement of **7%** on **F1-score** over the baseline models.
  - Explored the use of Attention Mechanism to generate attention-aware features for classification.

Projects

- **Online Writer Identification** — July '19 - June '20
  Bachelor's Thesis, Dr. Suresh Sundaram, Associate Professor, IIT Guwahati — *Thesis Link*
  - Developed histogram based features using concepts of gravitation and momentum to represent handwriting characteristics of a writer and obtained an improvement of **7%** on the BIT Casia Dataset.
  - Encoded the temporal information using supervised LSTM autoencoders. The encoded vectors were classified using SVM classifier.
- **Graphical User Interface for Computer-Aided Diagnosis of Psoriasis** — May '18 - June '18
  Dr. Utpal Garain, CVPR Unit, ISI Kolkata — *Project Link—Demonstration Link*
  - Developed a Graphical User Interface (GUI) for computer-assisted medical image analysis and processing diseased images of Psoriasis using PyQt, Keras and OpenCV.
  - The GUI was designed to perform severity assessment of Psoriasis disease using ResNet50. It was also designed to perform segmentation of Psoriasis histopathology images using U-Net architecture.

Publications and Preprints

Sounak Ray, Addrish Roy, Suresh Sundaram. **A Deep Learning Framework with Histogram Features for Online Writer Identification,** International Conference on Frontiers in Handwriting Recognition (ICFHR), 2020. Link to paper

Anabik Pal, **Sounak Ray** and Utpal Garain; **Skin disease identification from dermoscopy images using deep convolutional neural network**; arXiv:1807.09163; ISIC 2018 Challenge

Technical Skills

- **Programming Languages:**: Python, C++, C, MATLAB, MySQL
- **Tools/Frameworks/IDE:**: Keras, TensorFlow, PyCharm, Git, OpenCV, spaCy, Docker, Kubernetes, Faiss
- **Miscellaneous**: LaTeX, MS-Excel
- **Operating Systems**: Windows, Linux, MacOS