SOUNAK RAY

(995)–798–5181 ⋄ sr3846@columbia.edu ⋄ linkedin.com/sounak-ray ⋄ Homepage ⋄ GitHub

EDUCATION

Columbia University   New York City, NY
M.S. in Computer Science; Specialization: Machine Learning   Sep 2021 – Dec 2022
- Relevant Courses: Computational Learning Theory, User Interface Design, Representation Learning.

Indian Institute of Technology Guwahati   Guwahati, IN
B. Tech. in Electronics and Communication Engineering, Minor: Computer Science; GPA: 8.44/10.0   Jul 2016 – Jun 2020
- Relevant Courses: Data Structures and Algorithms, Computer Vision, Artificial Intelligence, Machine Learning.

TECHNICAL SKILLS

Programming Languages/Web Technologies: Python, C++, C, MATLAB, MySQL, HTML, CSS.
Tools/Frameworks/IDE: Keras, TensorFlow, PyCharm, Git, OpenCV, spaCy, Docker, Kubernetes, scikit-learn, numpy, pandas.
Miscellaneous: LaTeX, MS-Excel.
Operating Systems: Windows, Linux, MacOS.

EXPERIENCE

Sprinklr   Gurgaon, IN
Data Scientist, Machine Learning Team   Aug 2020 – Aug 2021
- **Customer Feedback**: Devised methods to automate incorporation of client feedback in datasets by leveraging similarity search with Faiss and baseline models. Attained an improvement of 4% F1-Score on validation set for various text classification tasks.
- **Liveness Verification**: Designed and finetuned an Attention ConvLSTM network with 100 videos for liveness detection through hand waving and eye blinking with Python and TensorFlow. Achieved an F1-Score of 93% on validation set.
- **Insights**: Generated insights from millions of consumer feedback and social media conversations with Hugging Face library by leveraging Transformers-based architectures to implement Named Entity Recognition and Text Classification across 12 languages.
- **Training and Deployment**: Responsible for understanding data requirements of deep learning projects through interaction with product managers. Developed, deployed and maintained sentiment classification pipelines using Jenkins, Docker and Kubernetes.

Sprinklr   Gurgaon, IN
Data Science Intern, Machine Learning Team   Jun 2020 – Aug 2020
- **Sentiment Classification**: Created an embedding model for emojis based on the Unicode of an emoji using Keras. Utilized Hierarchical Attention Network with word2vec and emoji embeddings to improve the F1-score on messages with emojis by 9.7%.

Indian Statistical Institute, Kolkata   Kolkata, IN
Research Intern, Dr. Utpal Garain, CVPR Unit   May 2019 – Jul 2019
- Developed a Residual Capsule Network for Skin Lesion Classification finetuned with ImageNet weights. Achieved an improvement of 3% F1-score over baseline ResNet models.
- Explored efficacy of Attention Mechanism by generating attention-aware features for classification.

PROJECTS

Residents COVID-19 Vaccination Status Portal   Jun 2021 – Jul 2021
- Created a platform to maintain an inventory of all people vaccinated in a neighborhood by applying HTML, CSS, Node.js and MySQL. Users can enter, update or delete vaccination details of beneficiaries and retrieve a list of vaccinated people.

Online Writer Identification   Jul 2019 – Jun 2020
Bachelor's Thesis, Dr. Suresh Sundaram, Associate Professor, IIT Guwahati
- Constructed histogram based features using concepts of gravitation and momentum to represent handwriting characteristics of a writer and obtained an improvement of 7% over state-of-the-art results on BIT Casia Dataset at text-line level.
- Encoded temporal information using supervised LSTM autoencoders. Encoded vectors were classified with SVM. Presented a paper (oral) published at ICFHR 2020.

Graphical User Interface for Computer-Aided Diagnosis of Psoriasis   May 2018 – Jun 2018
- Constructed a Graphical User Interface (GUI) for computer-assisted medical image analysis of Psoriasis using PyQt, Keras and OpenCV. Users can apply image processing functions filtering, rotating and cropping on diseased images.
- Users can perform severity assessment of Psoriasis disease using ResNet50 and MobileNet and segmentation of Psoriasis histopathology images with the help of 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.

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