Sanskriti Singh

530-364-8327 | ssingh7@mit.edu | http://www.linkedin.com/in/sanskritisingh7

OVERVIEW

Massachusetts Institute of Technology **Aug. 2024 - May 2027**

Coursework: Intro to Algorithms, Deep Learning, Intro to Machine Learning, Math for CS, Probability and Random Variables
Lab Assistant: Intro to Machine Learning

TECHNICAL EXPERIENCE

News-to-Arxiv pipeline (Independent Project) Jun. 2025 – Curr.

- Built a full-stack Python pipeline transforming >500 news queries into arXiv search tasks, leveraging LLMs and heuristic scoring to retrieve top 10 relevant papers per query. Implemented JSON-based storage and retrieval across 50,000+ arXiv articles for matching.

MIT CSAIL Undergraduate Research - Professor Amar Gupta Nov. 2024 – Curr.

- Advancing credit card fraud detection by designing and training LSTM and Transformer models on >1M transaction records, incorporating synthetic data generation techniques to enhance training diversity and improve overall model F1 score by ~12%.
- Developing O-Health symptom extractor to process >50,000 patient records, automatically classifying symptoms and accurately identifying relevant medical specializations with >60% accuracy via unsupervised learning and clustering mechanisms.

Engineering Internship at Kognitos Jun. 2025 – Aug. 2025

- Engineered automated pipelines leveraging LLMs and heuristics to enhance product error messaging by 89% and performed regex structural matching with 99% accuracy, significantly reducing customer wait times and improving overall issue resolution efficiency.

MIT CSAIL Undergraduate Research - Professor Manolis Kellis Aug. 2024 – Dec. 2024

- Developing latent space embeddings with personalzied metrics on > 6000 patient records and external data for cardiovascular disease

Water Consumption via Unsupervised ML and Distance Algorithms (Independent Project) Sep. 2023 – Sep. 2024

- Developing model to detect the amount of water in gallons being used by each water appliance (averaging around 9 appliances) within a household after single use appliance calibration > 60 seconds - using kmeans, DTW distance, mathematical equations, and clustering

RESEARCH EXPERIENCE

Co-Founder and President – Andromeda AI (https://andromedaaigalaxy.wixsite.com/website) Jan. 2020 – Aug. 2024

- Reached 5 countries and 10 states, received $5,000 in awards for innovative projects, and actively mentors students through personalized one-on-one sessions and collaborative group meetings, fostering a strong interest in computer science and digital design.

Yale Interventional Oncology Lab Research Intern June. 2023 – Oct. 2023

- Led a cutting-edge project designing an innovative GAN by developing a novel teacher-student model architecture, significantly improving performance by 30% and DICE coefficient by 23% on Eovist MRI data, advancing automated medical imaging analysis.

CheX-Nomaly: Segmenting Lung Abnormalities from Chest Radiographs using Machine Learning
(1st Author: In review/Accepted - arXiv) Sep. 2022 – Jan. 2024

- Developed a siamese binary U-net model for localized detection of chest X-ray abnormalities, using transfer learning and contrastive learning techniques to improve generalizability across various thoracic diseases; Mentors: Chenyu You (Stanford)

Automated Coronary Calcium Scoring using U-Net Models through Semi-supervised Learning on Non-Gated CT Scans
(1st Author: Accepted - IEEE, MIT URTC) Sep. 2021 – Oct. 2022

- Predicted heart attack risk from non-gated CT scans by developing novel mathematical equations to adapt and crop to mimic gated scans, achieving a 91% improvement in mean absolute error and a 32% improvement in F1 score; Mentor: Weicheng Dai (NYU)

A Novel Mask R-CNN Model to Segment Heterogeneous Brain Tumors through Image Subtraction
(1st Author: Accepted - arXiv) Aug. 2020 – May 2021

- Enhanced brain tumor segmentation in MRI scans using Mask R-CNN with a ResNet backbone and image subtraction, achieving DICE coefficient of 0.75, and compared to state-of-the-art models; Mentors: Dr. Sussane Soin, consultant radiologist in London

PneumoXttention: A CNN compensating for Human Fallibility when Detecting Pneumonia through CXR images with Attention
(1st Author, Accepted - IEEE, ISPA) Nov. 2019 – Sep. 2021

- Developed a machine learning algorithm utilizing a dual ensemble of 13-layer CNNs and heatmap techniques to detect pneumonia in chest X-rays. Validated with high F1 against radiologists across diverse datasets; Mentor: Dr. Colin-Whitby Strevens (Cambridge)

Emphasis on the Minimization of False Negatives or False Positives in Binary Classification
(1st Author: Accepted - arXiv) Sep. 2018 – Oct. 2019

- Reduced false negatives or false positives in binary classification by adjusting input values post-pre-training, demonstrating improved recall or precision across various datasets and model architectures without significantly affecting the overall F1 score

HONORS AND AWARDS

- Top 300 Scholars Regeneron Science Talent Search - 2024
- SCVSEFA 1st place winner - 2019, 2020, 2021, 2022, 2023
- BioGENEius (HM, 3rd, HM) - 2021, 2022, 2023
- NCWIT National Winner - 2023

SKILLS

Python (Advanced), Java (Intermediate), PyTorch, TensorFlow, Machine Learning, Deep Learning, Computer Vision, Git, Data Analysis

PUBLICATIONS

- **S. Singh,** "PneumoXttention: A CNN compensating for Human Fallibility when Detecting Pneumonia through CXR images with Attention," 2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA), Zagreb, Croatia, 2021, pp. 61-66, doi: 10.1109/ISPA52656.2021.9552171.
- **S. Singh**, "Automated Coronary Calcium Scoring using U-Net Models through Semi-supervised Learning on Non-Gated CT Scans," 2022 IEEE MIT Undergraduate Research Technology Conference (URTC), Cambridge, MA, USA, 2022, pp. 1-5, doi: 10.1109/URTC56832.2022.10002228.
- **Singh, S.** (2023, November 3). CheX-Nomaly: Segmenting Lung Abnormalities from Chest Radiographs using Machine Learning. arXiv.org. https://arxiv.org/abs/2311.01777
- **Singh, S.** (2022a, April 6). Emphasis on the minimization of false negatives or false positives in binary classification. arXiv.org. https://arxiv.org/abs/2204.02526
- **Singh, S.** (2022a, April 4). A Novel Mask R-CNN Model to Segment Heterogeneous Brain Tumors through Image Subtraction. arXiv.org. https://arxiv.org/abs/2204.01201