Graduate Student
MS in Computer Science
Stony Brook University, NY

HARSHIT

Contact: +1(631)7108849
hharshit@cs.stonybrook.edu
https://harshit13.github.io/

Passionate developer, aspiring to work on infrastructure, systems and apps using Containers, Kubernetes. Linkedin: https://bit.ly/2B8UvLT

Education

M. S. Computer Science, **Stony Brook University**, New York; GPA: 3.87/4.0 [2019-Dec 20]

- *Distributed Systems, Big Data, Computer Vision, Visualization, Probability and Statistics, Analysis of Algorithms*

B. Tech. Computer Science and Engineering, **Indian Institute of Technology**, Patna, India; CPI: 8.84/10 [2013-17]

- *Operating Systems, Deep Learning, Databases, Algorithms, Data Structures, Object-oriented Programming*

Skills and Technologies

Languages and tools: *Java; Go; Python; Spark; C; C++; JavaScript; MySQL; HTML; CSS; RaspberryPi; Arduino; Latex; Linux*
Frameworks: Py*spark; HDFS; Google Cloud; Spring, Hibernate; D3.js; Node.js, Flask; Django; React Native; PyTorch; Tensorflow; OpenCV*

Work Experience

SWE Intern - Machine Learning **VisioLab** *Summer 2020*

- Develop object recognition pipeline using One-Shot Learning; Similarity/Metric Learning via DNN
- Reduced training time by 5x using cloud resources and set-up train-cycle with the world largest food caterer Aramark.
- Implement and dockerize modules (to deploy on Kubernetes cluster) to train, deploy and onboard data/models from/to client.
- Implement library for feature extractor training on cloud GPUs. [*Docker, Google Cloud Storage, Flask; Python; PyTorch; CoreML*]

Software Engineer **Samsung R&D** *2017-2019*

- OPERATING TIME PERFORMANCE for Samsung VD:- Develop toolkit to analyze OS performances, reducing work load from weeks to days, using deep learning and Computer Vision [*Python (Tensorflow, OpenCV)*]
- PRODUCT INTELLIGENCE:- Maintain video search API; Generate statistics and analytics influencing proactive product decisions; monitoring timelines of various products via Machine Learning models. [*Java-Spring-SOLR, Python, Splunk*]
- VERTICAL OPTIMIZATION of memory and peformance issues for Samsung TVs ensuring stable software released. [*C, C++*]

Research Internship **Nanyang Technological University, Singapore** *Summer 2016*

- Collected and processed EEG responses [using EMOTIV Epoc headset] on audio and visual stimuli.[C#, MATLAB]
- Implemented the authentication system, achieving 80% accuracy. Published at IECBES 2016. https://bit.ly/2m2WKII

- Graduate Teaching Assistant, Data Structures (CSE214), UI Development (CSE333) **Stony Brook University**

Selected Publications

- LaSOT: A High-quality Large-scale Single Object Tracking Benchmark, IJCV 2020. https://bit.ly/3kAAKhA
- Road Congestion Sensing via Crowdsourcing and MapReduce, IPSN 2015. https://bit.ly/2kQQP9B

Projects

Mining School Surveys for Quality Education *PySpark; HDFS; Tensorflow*

- Conducted Multi-Hypothesis test, to find significance of feedback information (20 GB survey data)
- Found clusters of area codes based on feedbacks via **Locality Sensitive Hashing** using Spark+HDFS. Train models on feedbacks.
- Cross nation area codes belonged to same clusters with Jaccard Similarity of 0.80. https://bit.ly/3d0CrS9

Sharded Replicated KeyValue Store *Go, Distributed Systems*

- Implemented a key value store replicated across multiple machines with RAFT consensus for fault tolerance.
- The system also used sharding and snapshots for performance. Done as a part of in-course project in Distributed Systems.

Chord: Distributed Hash Table *Java (javafx)*

- Implement a peer to peer distributed hash table using chord protocol and algorithm.
- Chord adapts efficiently as new keys join the system, and can answer queries even if the system is continuously changing.

Pintos: Operating System *C; Bash; Qemu*

- OS courswork project. Used Pintos OS available from Stanford and on Qemu VM emulator. Implemented FCFS, priority schedulling algorithm. Completed virtual memory library. Used semaphores and locks for synchronization tasks.
- Modified file system to allow directory entries to point (links) to files or to other directories.

Robust UAV Object Tracking [Masters project] *OpenCV, PyTorch*

- As a part of Masters project, developing an online tracking algorithm for tracking objects through UAV video sequences.
- Implement template update for online tracker via alignment matching and updateNet for higher accuracies.
- Worked on largest Single Object Tracking Dataset, LaSOT, used as a standard for benchmarking. https://bit.ly/2S0z6du

Copter QL: The Q-Learning Helicopter Game *Deep-learning[Tensorflow], Pygame*

- Aimed to make agent learn to play copter using a Deep QNetwork (DQN) for learning Q-values for approximate state-action pairs.
- Agent balanced exploration and exploitation using experience replay and update delay, achieving the best strategy to score after 3000 attempts at gameplay. https://bit.ly/2AK4qqO

Other Experiences And Achievements

- Google Kickstart '20 Round B **[Rank 430; 96 percentile]** Hashcode '20 **[US Rank 90]; Competition Expert @ Kaggle** [97 percentile]

- **Bronze medal** in IoT Innovation at Inter-IIT tech meet (2016). Runner-up in **IEEE ISED Grand Challenge** December 2016.