

Graduate Student
MS in Computer Science
Stony Brook University, NY

HARSHIT

Linkedin: <https://bit.ly/2B8UvLT>

Contact: +1(631)7108849
hst.kmr@gmail.com
<https://harshit13.github.io/>

Education

- M. S. Computer Science, Stony Brook University, New York; GPA: 3.87/4.0 [2019-Dec 20]
- *Computer Vision, Big Data, Distributed Systems, 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]
- *Deep Learning, Network Science, Algorithms, Data Structures, Object-oriented Programming, Operating Systems*

Skills and Technologies

Languages and tools: Java; Python; MySQL; Go; Spark; C; C++; JavaScript; HTML; CSS; RaspberryPi; Arduino; Latex; Linux
Frameworks: Pyspark; HDFS; Google Cloud; Spring, Hibernate; D3.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 [Python; PyTorch; GCP, CoreML]
- Implement library for feature extractor training on cloud and APIs to deploy models to client and onboard data from client.
- Dockerized the module for scalability. This work helped the team setup active learning cycle with Aramark (world largest food service).
- Engineer I** **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(lightgbm, sklearn), Splunk]
- Setup of MEMORY and PERFORMANCE TASK FORCE for profiling Tizen OS ensuring stable software. [C++, Python, Jenkins]
- 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), Benevolent Computing (ISE339), **Stony Brook University**, NY, US
- **Research Internship**, CNeRG lab, under Prof Niloy Ganguly, **IIT Kharagpur**, India [Summer 2015]

Conference Publications

- Online Electroencephalogram (EEG) based biometric authentication using visual and audio stimuli, IECBES 2016
- Road Congestion Sensing via Crowdsourcing and MapReduce, IPSN 2015. <https://bit.ly/2kQQP9B>

Projects

- Mining School Surveys for Quality Education (2020)** *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>
- Covid-19 Vizualization (2020)** *D3.js; Flask; Bootstrap; jQuery*
- Developed a dashboard with Map visualization to show COVID-19 statistics for each country using color-maps. <https://bit.ly/2zlkZJw>
- Used parallel coordinate and radar chart to show trends between health expenses, population density and covid-19 stats.
- Robust UAV Object Tracking (2020 - Present) [Masters Project]** *Python; 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>
- Sharded Replicated KeyValue Store (2019)** *Go*
- Implemented a scalable 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**.
- Copter QL: The Q-Learning Helicopter Game (2019)** *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>
- Chord: Distributed Hash Table (2016)** *Java*
- Implemented a peer to peer distributed hash table simulator using chord protocol and algorithm.
- Chord adapts efficiently as new keys join, and can read/write in sublinear time efficiently.

Other Experiences And Achievements

- **Google Kickstart '20 Round B [Rank 430] Hashcode '20 [US Rank 90]; Competition Expert @ Kaggle** [Currently ranked - 3547]
- **Bronze medal** in IoT Innovation at Inter-IIT tech meet (2016). Runner-up in **IEEE ISGD Grand Challenge** December 2016.