

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

Skills and Technologies

Languages and tools: Java; Go; Python; MySQL; 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 Few-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.
- Implemented and dockerized modules (to deploy on Kubernetes cluster) to train, deploy and onboard data/models from/to client.
- Implemented 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:- Developed toolkit to analyze OS performances, reducing work load from weeks to days, using deep learning and Computer Vision [Python (Tensorflow, OpenCV)]
- PRODUCT INTELLIGENCE:- Maintained video search API; Generated 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 performance 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>
- Robust UAV Object Tracking [Masters Project]** Python; PyTorch
- As a part of Masters project, developing an online tracking algorithm for tracking objects through UAV video sequences.
- Implement online tracker by updating query template 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** 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** 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** 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.
- Covid-19 Vizualization** 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.

Other Experiences And Achievements

- **Google Kickstart '20 Round B [Rank 430/10k]** 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.