# Sai Vishwanath Venkatesh

saivishy.github.io - linkedin.com/in/saivish - saivishy - saivishw@buffalo.edu

## Education

**University at Buffalo (SUNY)** **September 2021 - January 2023***
*Master of Science in Computer Science — GPA: **3.41*** *Buffalo, New York*
*Courses* : *Parallel & Distributed Programming, Machine Learning, Distributed Systems, ML in Society, Blockchain*

**SRM Institute of Science and Technology** **July. 2016 – May 2020**
*Bachelor of Technology in Computer Science — Percentage: **84.42%*** *Chennai, India*
*Courses* : *Operating Systems, Data Structures, Analysis of Algorithms, Big Data Analytics, Text Mining*

## Technical Skills

**Languages**: C/C++, Python, Bash, R, SQL, Java, JavaScript, HTML, Ruby, LaTeX
**Frameworks and Libraries**: OpenMP, OpenMPI, HPX, Nvidia CUDA, Pandas, Numpy, Keras, EdgeML, Docker, GCP, Tableau
**Concepts**: Parallel & Distributed Programming, Activity Recognition, Natural Language Processing, Anomaly Detection

## Projects

**SCoOL - Scalable Common Optimization Library** | Repository **May 2022**
- Worked under the advise of Dr. Jaroslaw Zola to design and develop a general purpose executor in HPX for search space exploration and optimisation problems (NP-hard problems) by adapting the bulk synchronous parallel (BSP) model of programming.
- Open Source Contributor towards this library.

**Distributed Sorting: Using Open MPI** | Repository **November 2021**
- Implements count sort on a range of short integers distributed across a set of processors in a cluster using Open MPI. Implemented on UB's HPC as well.

**Gaussian KDE: GPU Implementation** | Repository **December 2021**
- Parallel implementation of gaussian kernel density estimation computation for a set of floating point numbers using NVIDIA CUDA.

**Rooting Graph Nodes: Using Spark** | Repository **November 2021**
- Finds the roots of each node in a graph of connected components using Apache Spark. Implemented on UB's HPC.

**Distributed Key Value Store** | Repository **May 2022**
- Developed a fault-tolerant distributed key value store using the RAFT consensus protocol.
- The data is made accessible via REST API endpoints.

**Distracting Message Identification For Instant Messaging Platforms** **April 2020**
- Built a preprocessing pipeline to consolidate and process messages from groups.
- Reconstructed models adapted from existing research to perform context-based filtering of distracting messages.

## Experience

**Solarillion Foundation** **July 2018 - February 2021**
***Researcher*** *Chennai, India*
- Lead a team to instrument efficient multiple instance learning on optical flow by introducing early stopping and intuitive undersampling. This tackled otherwise computationally expensive crime detection for video. Deployed and tested on a resource constrained device: Raspberry Pi3 (ARM v8) our solution performed 2.3x times faster then the benchmark servers (Intel Xeon) using just a third of the number of cores.
- Lead a team towards condensing one of the largest datasets (10 billion+ records) for malware classification in smartphones to 0.1% of its feature set whilst retaining threshold accuracy reported by competing data intensive methods through efficient data preprocessing and online learning.
- Developed a two-stage model to estimate the number of weeks a movie in a multiplex would successfully run based on past occupancy/crowd behaviour. The solution is currently in use by one of India's leading cinema exhibitors.

***Server Administrator***
- Was responsible for Server Automation such as writing shell scripts, linux administration , writing wrappers to automate student performance reports, TA hours and schedule meetings. In use by 100+ users consisting of students and TAs.
- Reconfigured & maintained the network attached storage server as well as the main compute server for the organisations daily use.

## Publications

**Non-intrusive Malware Classification and Data Theft Classification for Smartphones** | Publication Code
$21^{st}$ International Conference on Computational Science (ICCS 2021), Kraków, Poland

**Forecasting the Lifetime of Movies in a Multiplex using deep learning** | Publication Code
Future of Information and Communication Conference (FICC 2020), San Francisco, USA.

**Real-time Crime Detection using Edge Devices** | Publication Code
International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2020)