

# Keval D. Shah

2300, Arapahoe Avenue, apt 338  
Newton Court, Boulder, CO - 80302

<https://www.linkedin.com/in/kevalds51/>

**Email:** [keval.shah@colorado.edu](mailto:keval.shah@colorado.edu)

**(M):** +1-(720)-917-5943

---

## EDUCATION

### University of Colorado - Boulder

**2018-2020**

MS, Computer Science

**GPA:** 3.86/4.00

- TA for “High Performance Scientific Computing” (Fall’18) - C, Python, OpenMP
- TA for “Introductory Computer Science - Engineering Applications” (Spring’19) - MATLAB and C++
- Courses: Natural Language Processing, Machine Learning, Object Oriented Analysis and Design, Big Data Architecture, Design and Analysis of Algorithms.

### Dhirubhai Ambani Institute of ICT, India

**2014-2018**

B.Tech, ICT (Honors) with minor in Computational Science

**CGPA:** 7.8/10.0

---

## INTERNSHIPS

### Software Engineer Intern, Splunk Inc

**June 2019-Present**

- Built an AWS lambda + Container based health check application that reports and alerts on data quality in the underlying data stores. Coded, tested and deployed the checker framework into production through a CI-CD configured infrastructure.

### System performance & analysis Intern, Robocop.io

**May-July 2016**

- Created plugins for system monitoring daemons in C and Python that were to be installed on client’s machines. These plugins would collect the system usage data and parse it for further processing.
- Developed a package security checker for unix systems using popular vulnerability databases like Common Vulnerabilities and Exposures (CVE) and National Vulnerability Database (NVD)

### Open source contribution to SymPy: <https://github.com/kevalds51>

- Development of the matrix multiplication module for the popular Python library SymPy

---

## PUBLICATIONS

### Analysing search queries to predict commercial success of movies:

- Trained a prediction model using data from search queries ( from Google trends) and movie databases. Predicted the commercial success of movies based on location, actors and internet searches.
- Used Python with scikit-learn and numpy libraries to train a linear-regression models.
- Achieved the highest performance and accuracy (70%) for this application problem.
- Published in International Journal of Control Theory and Application, 9(41) 2016, pp. 629-640 ISSN: 0974-5572. [Document link](#)

### High-Performance Computing applications in simulation of plasma:

- Parallelization of the PIC-MCC algorithm for plasma simulation on Intel Xeon-phi architecture.
- Implemented efficient load balancing and improved cache re-usability using novel sorting trigger.
- Achieved an overall speed-up of 20x with respect to serial implementation on Intel Xeon KNL.
- Presented as a [research poster](#) at 30<sup>th</sup> Supercomputing conference (SC17) at Denver, USA.

---

## TECHNICAL SKILLS

- **Programming Languages:** Python (8000 lines), C (3000 lines), Java (3000 lines), MATLAB
- **Tools:** Splunk, GIT, Latex, AWS tools, numpy, scikit-learn, TensorFlow, SQL, mongoDB, pytest, pipenv, Jenkins, Docker, RESTful APIs, openMP, Unix, Jira (Agile tool)

---

## HONORS

- Third place in Parallel Programming Challenge supported by Intel and NVIDIA at 23rd International Conference on HPC, Data and Analytics (HiPC 2016) and received travel grant to present solution.
- Recipient of “Experiencing HPC for UG” program of 30<sup>th</sup> Supercomputing (SC17) conference.
- Built a slack bot for incoming graduate channel in HackCU’18 (University Hackathon organized by Microsoft/Github)
- Certificate of Merit for securing 100% marks in Informatics Practices in All India Senior School Certificate Examination (AISSCE) 2014.