

**Contact Information**      483, Engineering 2      **email:** [dhawal.joharapurkar@gmail.com](mailto:dhawal.joharapurkar@gmail.com)  
 Computer Science Department      **web:** <https://dhawaljoh.github.io>  
 University of California, Santa Cruz      **phone:** +1-831-346-8702  
 Santa Cruz 95060

**Manipal Institute of Technology**, Manipal 2011–2015  
 B.Tech in Computer Science & Engineering  
**Bachelor Thesis:** Temporal Scoping of Facts in a Knowledge Base

**Xerox Research Centre India**, Bangalore Jun '15–Jan '16  
 Research Intern, Machine Learning and Statistics Group  
 Mentors: Dr. Vaibhav Rajan and Sumit Negi

- Analysis of clinical notes of ICU patients using Topic Models
- Created a plotting tool to highlight different statistics across classes

**Indian Institute of Science**, Bangalore  
 Project Trainee, Supercomputer Education and Research Centre (SERC)  
 Mentor: Prof. Partha Talukdar [MALL Lab]

- Temporal scoping and ordering of relations in a knowledge base
- Entity linking and disambiguation in large text corpora

**Indian Institute of Technology**, Kharagpur May '14–Jun '14  
 Summer Research Intern, Dept. of Computer Science & Engineering  
 Mentor: Prof. Sudeshna Sarkar

- Implemented DBSCAN algorithm to find traffic stoppage points and segmented roads based on their speed profiles – automatic profiling of drivers
- Modified “simplekml” Python module to plot the GPS points on Google Maps

**DataWeave Software Pvt. Ltd.**, Bangalore May '13–Jun '13  
Summer Intern

- Created data crawlers using Python that aggregated and stored content in JSON dumps
- Content available via APIs, a few listed at <http://dataweave.in/apis>

Teaching Experience	<b>University of California</b> , Santa Cruz	Winter '17, '18
	Teaching Assistant, CMPS140 Artificial Intelligence	
	<b>University of California</b> , Santa Cruz	Spring '18
	Teaching Assistant, CMPS11 Intermediate Programming in Java	

Projects	<b>Expert Identification</b>	Oct '18-Present
	<b>Mentor:</b> Prof. Lise Getoor	UC Santa Cruz, Santa Cruz
	Identification of experts based on network topology, collaboration graphs and local features such	

as post strength, experience information, etc. We are working on StackExchange datasets and using Probabilistic Soft Logic for inference.

**Question Category Classification**

Jul '15

**Mentor:** Prof. Manish Shrivastava

IASNLP, IIIT, Hyderabad

Question category classification on TREC data using topic composition features generated by LDA.

**Detecting Fibrous Regions in Protein Sequences**

Nov '13–May '14

**Mentor:** Dr. Smitha Nair

Manipal Institute of Technology, Manipal

Detection of fibrous regions in protein sequences using Support Vector Machines and Bee Colony Optimization for PCA.

**Online competition: Photo Tagger – Multi-class classification**

Mar '14

**Leaderboard:** 84 out of 644

CSA, IISc, Bangalore

SVM whose parameters were optimized by GridSearchCV, to classify photos into various classes (people, cars, shoes, buildings, flowers).

**Craigslist Post Classification**

Oct '13

**Accuracy:** 81%

Manipal Institute of Technology, Manipal

Vectors created by bag of words model and then tf-idf, SVM to classify posts on Craigslist into sections based on the product description.

Publications

**Online Adspace Posts' Category Classification**

Dhawal Joharapurkar, Vaishak Salin, Vishal Krishna

12<sup>th</sup> International Conference on Natural Language Processing, December, 2015

Talks

**From Big Text to Big Knowledge**

SERC Open Day 2015

Feb '15

IISc, Bangalore

Honors

**Regents' Fellowship**, UC Santa Cruz

2016–17

Online Courses

**The Data Scientist's Toolkit**

coursera.org, 100%

Jun '14

Johns Hopkins University

**Design and Analysis of Algorithms**

Massively Empowered Classrooms, 100%

May '14

Microsoft Research

**Machine Learning**

coursera.org, 100%

Apr '14

Stanford University

**Algorithms: Design and Analysis, Part 1**

coursera.org, 98%

July '13

Stanford University