

Dhawal Joharapurkar

Contact Information 483, Engineering 2 **email:** 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

Education	University of California , Santa Cruz Ph.D. in Computer Science Advisor: Prof. Lise Getoor	2016–Present
-----------	----------------------------------------------------------------------------------------------------------------	--------------

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

Work Experience **Lam Research**, Fremont Jun '18–Present
 Intern, Global Information Systems
 Mentor: Amod Ghangurde

- Initiated the standardization of the Python development environment across business units
- Foundational work on using Natural Language Processing at Lam Research. Used NLP to augment existing ML systems by ingesting text from Lam Knowledge Documents.

Chobanian Group, San Jose Jun '17–Aug '17
 Research Intern, Data Science Group
 Mentor: Dr. Tyler Munger

- Expert identification in large delivery organizations
- Statistical Relational Learning approaches to expert identification

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 Dec '14–May '15
 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	University of California , Santa Cruz	Spring '18
	CMPS11 Intermediate Programming in Java [Evaluations] University of California , Santa Cruz	Winter '17, '18
	CMPS140 Artificial Intelligence [Evaluations - Winter 17'] [Evaluations - Winter 18']	
Current Project	Expert Identification	Oct '18-Present
	Mentor: Prof. Lise Getoor	UC 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.	
	Survey On Data Programming	May '18- Jun '18
	Mentor: Prof. Cormac Flanagan	UC Santa Cruz
	Survey on the recent literature surrounding Data Programming.	
	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.	
	Past Projects	
	Link: https://dhawaljoh.github.io/research/projects/	
Publications	Online Adspace Posts' Category Classification Dhawal Joharapurkar, Vaishak Salin, Vishal Krishna 12 th 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

