

# Dhawal Joharapurkar



dhawal@ucsc.edu



+1 (831) 346-8702



<https://dhawaljoh.github.io>

## Education

M.S Computer Science  
M.S Project: User Expertise Detection  
in Online Communities  
UC Santa Cruz | 2016-19

B.Tech. Computer Science  
Bachelor Thesis: Temporal Scoping of  
Facts in a Knowledge Base  
Manipal Institute of Technology |  
2011-15

## Honors

Regents' Fellowship  
UC Santa Cruz | 2016-17

## Publications

Online Adspace Posts' Category  
Classification  
Dhawal Joharapurkar, Vaishak Salin,  
Vishal Krishna  
12<sup>th</sup> International Conference on  
Natural Language Processing | 2015

## Talks

From Big Text to Big Knowledge  
SERC Open Day 2015  
IISc, Bangalore | 2015

## Teaching Exp.

Teaching Assistant [Evaluations]  
CMPS5J, CMPS11, CMPS12A,  
CMPS140, CSE184  
UC Santa Cruz | 2017-19

## Work Experience

- Summer '19 Amazon.com Applied Scientist Intern  
Prediction of behavioral priors for item cold start problem using NN  
Replaced several GBDTs with a single multi-output NN that predicts behavioral priors
- Summer '18 Lam Research Intern  
Standardization of the Python development environment across business units  
Used NLP to augment existing ML systems by ingesting text from Lam Knowledge Documents.
- Summer '17 Chobanian Group Intern  
Expert identification in large delivery organizations  
Statistical Relational Learning approaches for expert identification
- Jun-Jan'16 Xerox Research Centre India Research Intern  
Analysis of clinical notes of ICU patients using Topic Models  
Created a plotting tool to highlight different statistics across classes
- Dec-May'15 Indian Institute of Science Project Trainee  
Temporal scoping and ordering of relations in a knowledge base  
Entity linking and disambiguation in large text corpora
- Summer '14 Indian Institute of Technology Intern  
Implemented DBSCAN algorithm to find traffic stoppage points and segmented roads based on their speed profiles  
Plotted the GPS points on Google Maps
- Summer '13 DataWeave Software Pvt. Ltd. Intern  
Created data crawlers using Python that aggregated and stored content in JSON dumps  
Made content available via APIs

## Projects

- Jun-Dec'18 User Expertise Detection in Online Communities UC Santa Cruz  
Identification of experts based on network topology, collaboration graphs and local features such as post strength, experience information, etc. Used StackExchange datasets and Probabilistic Soft Logic for inference.
- May-Jun'18 Survey On Data Programming UC Santa Cruz  
Survey on the recent literature surrounding Data Programming.
- Spring'17 Effects of Social Influences on Culinary Preferences UC Santa Cruz  
Modelled the influence of a friend network on the culinary choices of individuals. Predicted a person's favorite cuisine from social influences from their network and evaluated our model by comparing our predictions to the gold labels that are extracted from the text of the reviews in the Yelp dataset.
- Winter'17 Ideology-Backed Stance Classification UC Santa Cruz  
Improved an approach on jointly-modeling disagreement and stance in online debate forums by integrating ideology as a latent variable in the joint model. Improvement of a few percentage points over the baselines set by prior works. Additionally, we also created a new ideology corpus of around 1.5 million tweets by 645 official members of the U.S. Congress.
- Past Older Projects  
<https://dhawaljoh.github.io/research/projects/>