

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
12th 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

Mentorship

Mentor [CruzHacks, 2020]

UC Santa Cruz | Jan, 2020

Recent Experience

Summer '19 Amazon.com | Applied Scientist Intern

Palo Alto, CA

- Created model to rank new products in accordance to their projected activity and not rank lower because of lack of data
- Reduced deployment effort by 12.5% by substituting existing models with a single multi-output neural network

Summer '18 Lam Research | Intern

Fremont, CA

- Standardized Python development environment across business units, reducing codebase integration costs by ~30%
- Created word-embeddings for Lam Knowledge Documents to provide data to ML/NLP models

Summer '17 Chobanian Group | Intern

San Jose, CA

- Deployed a relational joint-model to identify experts in large delivery organizations for customer support case assignment
- Improved case resolution times by 45% by reducing case re-routing in a sample subset of 10K cases

Jun-Jan '16 Xerox Research Centre | Research Intern

Bangalore, IND

- Analyzed clinical notes of ICU patients using Topic Models
- Created a plotting tool to monitor topic distribution across output classes over time

Dec-May '15 Indian Institute of Science | Project Trainee

Bangalore, IND

- Performed temporal scoping and ordering of relations extracted from a news dataset
- Worked on Entity Resolution of Named Entities in large text corpora by using their temporal information to disambiguate

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 dataset with ~5M users and Probabilistic Soft Logic for inference. Predicted "accepted answer" with an accuracy of 87%.

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 using collective classification. 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. Achieved an accuracy of 94% and recall of 97%.

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. Improved existing baseline results by 2.5%. Created a new ideology corpus of around 1.5 million tweets by 645 official members of the U.S. Congress.

Other

Other Projects

<https://dhawaljoh.github.io/2019/projects/>

Skills

- Programming Languages: Python, Java
- Operating Systems: Unix-based OS, Microsoft Windows
- Specializations: Machine Learning, Data Science, Analytics