

# Dhawal Joharapurkar

---

Contact Information 1056, 10th Main  
Mahalaxmipuram  
Bangalore - 560086

**Phone:** +91 9035 69 7023  
**email:** dmjan21@gmail.com  
**web:** <https://dhawaljoh.github.io>

Work Experience **Xerox Research Centre India**, Bangalore June '15– Present  
Research Intern, Machine Learning and Statistics Group

**Indian Institute of Science**, Bangalore December '14– May '15  
Project Trainee, Supercomputer Education and Research Centre (SERC)

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

**Indian Institute of Technology**, Kharagpur May '14 – June '14  
Summer Research Intern, Dept. of Computer Science & Engineering

- Automatic profiling of Driver Behaviour on a GPS dataset provided by MHRD.
- Implemented DBSCAN algorithm to find traffic stoppage points and segmented roads based on their speed profiles
- 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 here

Education **Manipal Institute of Technology**, Manipal 2011 - Present  
B.Tech in Computer Science & Engineering  
**Bachelor Thesis:** Temporal ordering and scoping of facts in a knowledge base  
**GPA:** 7.01

**National Public School, Rajajinagar**, Bangalore 1997-2011  
Central Board of Secondary Education

Online Courses **The Data Scientist's Toolkit** June '14  
coursera.org, 100% Johns Hopkins University

**Design and Analysis of Algorithms** May '14  
Massively Empowered Classrooms, 100% Microsoft Research

**Algorithms: Design and Analysis, Part 1** July '13  
coursera.org, 98% Stanford University

**Machine Learning** April '14  
coursera.org, 100% Stanford University

Projects **Detecting Fibrous Regions in Protein Sequences** November '13 – May '14  
**Guide:** Dr. Smitha Nair Manipal Institute of Technology, Manipal  
Worked on the detection of fibrous regions in protein sequences using Support Vector Machines and Bee Colony Optimization for PCA.

	<b>Photo Tagger: Multi-class classification</b> <b>Rank:</b> 84 out of 644 Used SVM to classify photos into various classes (people, cars, shoes, buildings, flowers). The parameters of the SVM were optimized using GridSearchCV. The features were extracted using the SIFT algorithm.	March '14 CSA, IISc, Bangalore
	<b>Craigslist Post Classification</b> <b>Accuracy:</b> 81% Used bag of words model, tf-idf and SVM to classify posts on Craigslist into sections based on the product description. The open dataset was available on HackerRank	October '13 Manipal Institute of Technology, Manipal
Conference Presentations	<ol style="list-style-type: none"> <li>1. <b>Craigslist Post Classifier: Identifying the category of a Craigslist post</b>", IEEE Student Branch, Manipal - February, 2014 (Adjudged Best Paper)</li> <li>2. <b>Craigslist Post Classifier: Identifying the category of a Craigslist post</b>", ICCMEH 2014, December, 2014</li> <li>3. <b>Enhancing a Financial Service organizations cross-sell strategy using Artificial Neural Networks</b>", ICCMEH 2014, December, 2014</li> </ol>	
Talks	<b>From Big Text to Big Knowledge</b> SERC Open Day 2015	Feb '15 IISc, Bangalore
Programming Competitions	<b>IEEEExtreme Programming Competition 7.0</b> University Rank: 1, Country Rank: 265	October '13 IEEE
Positions Held	<b>IEEE Student Branch, Manipal</b> Technical Secretary Organized several events that saw a participation of 80+ teams each.	August '13 – May '14 Manipal Institute of Technology, Manipal
Skills	<ul style="list-style-type: none"> <li>• <b>Languages:</b> C/C++, Python, Octave, SQL, L<sup>A</sup>T<sub>E</sub>X</li> <li>• <b>Operating Systems:</b> Linux(various distributions), Microsoft Windows</li> <li>• <b>Tools:</b> Emacs, Sublime Text, Enthought Canopy, IPython</li> <li>• <b>Version Control System:</b> git</li> </ul>	
Curriculum Courses	Neural Networks and Fuzzy Systems; Data Mining and Data Warehousing; Software Testing; Distributed Computing Systems; Computer Communication and Networks; Operating Systems; Graph Theory; Cryptography and Network Security; Parallel Computing; Language Processors; Relational Database Management Systems; Discrete Mathematics; Switching Theory and Logic Design; Formal Languages and Automata Theory; Data Structures; Design and Analysis of Algorithms; Design and Implementation of Programming Languages; Computer Organization and Design	