

Harsh Pareek

CS PhD student graduating Aug 2017(exp)

CONTACT INFORMATION:

GDC 2.718F, Department of Computer Science,
2317 Speedway, Stop D9500, UT Austin,
Austin, TX 78712

www.cs.utexas.edu/~harshp
+1 (512) 994-5775
harshp@cs.utexas.edu

RESEARCH STATEMENT

I'm interested in the theoretical foundations of Machine Learning but also in applying principled Machine Learning algorithms and approaches at scale for real-world applications.

My PhD work emphasizes the axiomatic and theoretical foundations of Ranking, Rank Aggregation and Human-in-the-loop learning / Crowdsourcing.

I have broad expertise and interest in the following research areas: Ranking and Recommender systems, Large Scale Optimization and Learning, High dimensional methods (compressed sensing), Graphical Models, Social Network Analysis, Time Series Analysis and NLP.

PUBLICATIONS

- *Distributional Rank Aggregation, and an Axiomatic Analysis* A. Prasad*, H. Pareek*, P. Ravikumar. ICML 2015
- *Tracking with ranked signals* Tianyang Li, Harsh Pareek, Pradeep Ravikumar, Dhruv Balwada, Kevin Speer. UAI 2015. (selected for plenary presentation)
- *A Representation Theory for Ranking Functions*. H. Pareek, P. Ravikumar. NIPS 2014
- *Human Boosting*. H. Pareek, P. Ravikumar. ICML 2013
 - *Human Boosting* was honored as one of the “50 best ideas from UTCS”, on the 50th anniversary of the Computer Science department at UT Austin

ACADEMIC BACKGROUND:

- **University of Texas at Austin** 2011 - Summer 2017(*expected*)
Computer Science PhD Program **GPA: 4.0/4.0**
Advisor: **Prof. Pradeep Ravikumar**, Statistical Machine Learning Group
- **Indian Institute of Technology - Bombay (IITB), India** 2007-2011
B.Tech(Equivalent to BSCS) Computer Science with Honours **CPI: 9.58/10**
with Minor in Electrical Engineering
Relevant Coursework: Graduate-level courses on Machine Learning, Web Mining, Convex Optimization and NLP.

OTHER RESEARCH AND INDUSTRY EXPERIENCE:

Clustering of user posts *September-November 2016*
Facebook, Menlo Park CA *SDE Intern, Feed Ranking Team*

- Continued work by a previous intern. Built an online system to cluster incoming posts into clusters. This system is currently in production (to the best of my knowledge).

Ranking for Amazon's mobile homepage *June-August 2016*
Amazon, Austin TX *Research Scientist Intern, Homepage Personalization Team*

- Worked across teams to create a proposal for incorporating diversity and other listwise features into the existing ranking framework for Amazon's mobile homepage. Used a number of internal tools to build an end-to-end system for introducing diversity into the mobile app, which will hopefully be in production soon.
-

Time Series Analysis for eBay transaction data *Summer 2013*
eBay Research Labs *PhD Intern, Data Science Team*

- Performed time series analysis and causality analysis on eBay's historical transaction data using Big Data technologies (Hadoop, MapReduce etc)
-

Transfer Learning Across Multilingual Corpora *June-July 2011*
Microsoft Research, Bangalore, India *Research Intern, Multilingual Systems Group*

- Explored transfer learning of text classifiers across languages to leverage availability of large unlabelled parallel corpora using a novel discriminative CCA technique. This allows using small amounts of labelled data in one language together with large unlabelled corpora to create and improve classifiers for other languages
-

Undergraduate Senior Thesis: Entity Search and Ranking *July 2010-May 2011*
Guide: Prof. Soumen Chakrabarti, IIT Bombay *Intelligent Internet Systems Group*

- Proposed RankSVM-based algorithms for ranking Wikipedia entities using bag of words and proximity based features from text and Random Walk models for entity list completion. Implemented a proof of concept system using Semantic Web technologies
-

Undergraduate Research Internship: Complete Functional Synthesis *May-July 2010*
Guide: Prof. V. Kuncak, EPFL, CH *Lab for Automated Reasoning and Analysis*

- Applied methods from Algebraic Geometry and Convex Analysis to design algorithms for program synthesis from specifications, in particular for programs involving linear integer and rational arithmetic formulae
-

HONOURS AND AWARDS:

- Awarded **MCD Fellowship** by the University of Texas at Austin.
 - Member of Honor Society **Phi Kappa Phi**
-

- **All India Rank 8** in IIT-JEE 2007 among over 500,000 applicants
 - **Silver Medallist** at the 38th **International Physics Olympiad 2007** in Isfahan, Iran
 - Awarded **Gold Medal** in Indian National Physics Olympiad. Indian National Chemistry Olympiad and Indian National Mathematics Olympiad (the **only student** in the history of the National Olympiad Program to be awarded all three)
 - 4th position in online contest and 7th in Onsite in **ACM ICPC Regionals 2009**
 - Ranked **2nd** in Task 2 (among undergraduates) in the **International Predictive Analytics competition** (UC San Diego Data Mining Competition 2010 sponsored by FICO)
-

- Recipient of the prestigious **Aditya Birla Group Fellowship** 2007-2011
 - Awarded **NTSE Scholarship** (National Talent Search Examination) in 2005
 - Ranked 12th in the State Merit List in SSC Examination (10th Grade) 2005
-

SERVICE:

Reviewer for AISTats 2016

OTHER ACTIVITIES:

Participated in startup accelerator **3 Day Startup - Austin** in October 2011

Coordinating TA for CS101, IIT Bombay *July 2010 - December 2010*
One of 4 Coordinating TAs for a team of 92 TAs for 550 freshmen

Overall Coordinator for the Science Club of IIT-Bombay *July 2009 - April 2010*
Led a team of 17 students to manage 6 technical clubs

TECHNICAL SKILLS:

R, MATLAB, Python, C/C++, Java, Weka, Mechanical Turk, Hadoop, MapReduce, Galois