Sahil Garg

Google Scholar, sahil.garg.cs@gmail.com, +1-3236372603, add: Mount Sinai, 1399 Park Avenue, NY-10029

Research Interest

I am interested in *problem-driven* research for real-world impact, relying upon my foundations in theoretical and algorithmic concepts from *computer science*.

Prior to the PhD program, I had worked upon the problem of monitoring environmental dynamics using mobile sensors. This helped me build my foundations in *machine learning*, *information theory*, *and robotics*.

For the PhD thesis, I worked to infer hypotheses on *cancerous* bio-pathways via *information extraction* from the existing bio-medical literature. In the ongoing struggle to solve this challenging *natural language processing* problem, I have built models, which are *computationally scalable* and *robust* to overfitting training data sets, relying upon theoretical concepts, such as *locality sensitive hashing*, *nonstationary kernel functions*, *semantic parsing*, *ensemble classification*, *semi-supervision*, randomized algorithms, etc.

From research summer internships and active collaborations, I have also developed an interest in *psychiatry* and *computational neuroscience*. For instance, the modeling of brain functions that are relevant to psychiatric diseases; in this problem, we explore *dictionary learning* a.k.a *sparse coding* as a computational model (*autoencoder*) for the phenomenon of *adult neurogenesis* in human brains. In special relevance to psychiatry, we have developed a framework for assisting a human psychiatrist with our hashing based *dialog generation* model, optimized maximizing novel *info-theoretic bounds*, that has potential applications, such as training/evaluating human therapist, analyzing therapeutic sessions, etc.

Education

University of Southern California (USC) Aug 2013 -Aug 2019
PhD in Computer Science Advisor: Aram Galstyan
Dissertation Committee: Aram Galstyan (chair), Kevin Knight, Greg Ver Steeg, Roger Ghanem, Irina Rish.

Thapar Institute of Engineering and Technology July 2005 - July 2009
Bachelor of Engineering (B.E.) in Computer Science

Selected Publications

(Tier-1) Conference Proceedings

Modeling Psychotherapy Dialogues with Kernelized Hashcode Representations: A Nonparametric Information-Theoretic Approach. Sahil Garg*, Irina Rish, Guillermo Cecchi, Palash Goyal, Shuyang Gao, Sarik Ghazarian, Greg Ver Steeg, Aram Galstyan. In submission.

Nearly-Unsupervised Hashcode Representations for Relation Extraction. Sahil Garg*, Aram Galstyan, Greg Ver Steeg, Guillermo Cecchi. To appear in proceedings of the Conference on Empirical Methods in Natural Language Processing (**EMNLP-19**).

Kernelized Hashcode Representations for Biomedical Relation Extraction. Sahil Garg*, Aram Galstyan, Irina Rish, Guillermo Cecchi, Shuyang Gao. Proceed-

ings of Thirty-Third AAAI Conference on Artificial Intelligence (**AAAI-19**).

Neurogenesis-Inspired Dictionary Learning: Online Model Adaption in a Changing World. Sahil Garg*, Irina Rish, Guillermo Cecchi, Aurelie Lozano. Proceedings of the Twenty-sixth International Joint Conference on Artificial Intelligence (**IJCAI-17**).

Extracting Biomolecular Interactions Using Semantic Parsing of Biomedical Text. Sahil Garg*, Aram Galstyan, Ulf Hermjakob, Daniel Marcu. Proceedings of the Thirtieth *AAAI* Conference on Artificial Intelligence (**AAAI-16**).

Persistent Monitoring of Stochastic Spatio-temporal Phenomena with a Small Team of Robots. Sahil Garg*, Nora Ayanian. Proceedings of Robotics: Science and Systems (**RSS-14**).

Learning Nonstationary Space-Time Models for Environmental Monitoring. Sahil Garg*, Amarjeet Singh, Fabio Ramos. Proceedings of the Twenty-Sixth *AAAI* Conference on Artificial Intelligence (**AAAI-12**).

Workshop Proceedings (co-located with tier-1 conferences)

Therapeutic Dialogue Modeling via Locality Sensitive Hashing. Sahil Garg*, Guillermo Cecchi, Irina Rish, Shuyang Gao, Greg Ver Steeg, Palash Goyal, Aram Galstyan. Presented in *ICML* 2018 Workshop, AI and Computational Psychology: Theories, Algorithms and Applications (CompPsy 2018).

Dialogue Modeling via Hashing Functions. Sahil Garg*, Guillermo Cecchi, Irina Rish, Shuyang Gao, Greg Ver Steeg, Sarik Ghazarian, Palash Goyal, Aram Galstyan. Proceedings of *IJCAI* 2018 Workshop, Linguistic and Cognitive Approaches to Dialog Agents (LaCATODA 2018).

Neurogenesis-Inspired Dictionary Learning: Online Model Adaption in a Changing World. *Sahil Garg**, Irina Rish, Guillermo Cecchi, Aurelie Lozano. *ICLR* 2017 - Workshop Track.

Efficient Space-Time Modeling for Informative Sensing. *Sahil Garg**, Amarjeet Singh, Fabio Ramos. Proceedings of KDD 2012 Workshop, Knowledge Discovery from Sensor Data (SensorKDD 2012).

Journal Articles

Computational Modeling of Adult Neurogenesis with Online Sparse Autoencoders. Sahil Garg*, Irina Rish, Guillermo Cecchi, Aurelie Lozano. In preparation for a neuroscience journal.

Therapeutic Dialog Modeling via Hash Functions: An info-theoretic approach. Sahil Garg*, Guillermo Cecchi, Irina Rish, Aram Galstyan. In preparation for a psychiatry journal.

Professional Services

Program Committee Member for (Tier-1) Conferences in CS

The AAAI Conference on Artificial Intelligence (AAAI)	2017, 2018, 2020
Neural Information Processing Systems (NeurIPS)	2017, 2018, 2019
International Conference on Machine Learning (ICML)	2018, 2019
International Conference on AI and Statistics (AISTATS)	2019, 2020
International Conference on Learning Representations (ICLR)	2019, 2020

Research Experience

Icahn School of Medicine at Mount Sinai **Postdoctoral Fellow**
Sept 2019 - present Advisor: Cheryl Corcoran
Developing interpretable machine learning models for research problems in computational psychiatry.

University of Southern California (USC) **Research Assistant**
Mar 2015 - Aug 2019 Advisor: Aram Galstyan
Developing machine learning models which are computationally scalable, trainable in a robust manner on small sets of labeled examples, applicable to natural language processing for healthcare.

May 2015 - Mar 2016 Advisor: Aram Galstyan
Info-theoretic modeling of brain fMRI dynamics using CorEx.

June 2014 - Oct 2015 Advisor: Aram Galstyan
Phase transitions in community detection using CorEx.

April 2014 - Sept 2014 Advisor: Aram Galstyan
Generative modeling of a complex network and its structural properties like clustering, power law degree distribution, degree correlation, etc.

Oct 2013 - April 2014 Advisor: Nora Ayanian
Persistent sensing of environmental phenomena with a team of robotic sensors.

June 2013 - Sept 2013 Advisor: Milind Tambe
Developing computationally scalable game theoretic algorithms for securing natural resources.

IIIT Delhi **Undergraduate Research Assistant**
April 2011 - May 2013 Advisor: Amarjeet Singh and Fabio Ramos
Learning non-stationary models efficiently for sensing environment dynamics.

Research Internship Experience

IBM T. J. Watson Research Center Computational Biology Center
Summers of 2016, 2017 Mentors: Irina Rish & Guillermo Cecchi

June 2016 - present Collab: Irina Rish, Guillermo Cecchi & Aurelie Lozano
We investigated computational plausibility of adult neurogenesis *phenomenon.*

June 2017 - present Collab: Irina Rish & Guillermo Cecchi
We developed an info-theoretic framework for modeling therapeutic dialogues via hash functions.

June 2017 - Aug 2017 Collab: Elif K Eyigoz & Guillermo Cecchi
Language based Discrimination for Parkinson's Disease.

July 2016- Aug 2016 Collab: Stephen J. Heisig
We investigated into the spatio-temporal modeling of pressure Mat data dynamics for an early detection of Parkinson's disease.

Teaching Experience

University of Southern California (USC)
Coordinated Mobile Robotics, Spring 2014 Teaching Advisor: Nora Ayanian

Software Engineering Experience

Snowpal Software Services Co-founder
June 2010 - Dec 2011 Manager: Harman Singh & Krish Palaniappan
Developed a server side application in education domain including database design and a RESTful API.

Commdel, India Software Engineer
Aug 2009 - June 2010 Manager: Srinivasareddy Chennareddy
Developed a component to parse the data packets, as per the configurable ISO8583 format, into business objects for financial transactions.

Global Logic, India Software Intern
Feb 2009 - Aug 2009 Manager: Atul Srivastava
Developed a component for subscribing RSS feeds in a user friendly manner with an efficient search utility.

Graduate Coursework

Artificial Intelligence, Database Systems, Coordinated Mobile Robotics, Machine Learning, Applied Linear Algebra, Estimation Theory, Advanced Analysis of Algorithms, Randomized Algorithms (A), Digital Geometry Processing, Scientific Computing and Visualization.

Educational programs

2015 Complex Systems Summer School, Santa Fe Institute.

Other Accomplishments

99% percentile secured in all India entrance exams IIT-JEE-05 (200k participants) and AIEEE-05 (600k participants).
1st rank secured in C++ skill exams (for online placements in undergrad school) conducted by companies Informatica Business Solutions (CS batch of 80 students), and Global Logic (220 students).

Research References

Aram Galstyan galstyan@isi.edu
Research Associate Professor Univ. of Southern California
Director of AI Division ISI USC

Irina Rish rish@us.ibm.com
Research Staff Member IBM T. J. Watson Research Center

Guillermo A. Cecchi gcecchi@us.ibm.com
Principal Research Staff Member IBM T. J. Watson Research Center

Greg Ver Steeg gregv@isi.edu
Research Associate Professor Univ. of Southern California

Amarjeet Singh amarjeet@iiitd.ac.in
Assistant Professor IIIT Delhi

Co-Founder & Chief Technology Officer — Zenatix

Fabio Ramos — fabio.ramos@sydney.edu.au
Associate Professor — Univ. of Sydney

Daniel Marcu — marcu@isi.edu
Research Associate Professor — Univ. of Southern California
Director of MT/NLP — Amazon

Kevin Knight — kevin.crawford.knight@gmail.com
Professor — Univ. of Southern California
Chief Scientist for NLP — Didi Chuxing

Engineering References

Nitin Gupta — nitin@commdel.net
Co-Founder & Managing Partner — Commdel Consulting Services Pvt Ltd
Co-Founder & Director- Product & Strategy — Agility MobileForce Solutions
Co-Founder & Director — Core Doc2Info Services Pvt. Ltd.

Amit K Verma — amit@commdel.net
Co-Founder & Director — Commdel Consulting Services Pvt Ltd
Co-Founder & Director — Agility MobileForce Solutions
Co-Founder — Core Doc2Info Services Pvt. Ltd.

Srinivasareddy Chennareddy — srinivasa.chennareddy@gmail.com
VP Products, Digital Marketing & Sales — CG Parivar Group
Co-Founder (exited in 2015) — Agility MobileForce Solutions

Harman Singh — hpssahni@gmail.com
Senior Software Engineer — Amazon
Co-Founder (exited in 2011) — Snowpal Software Services