Ritik Dutta

+91-9586986574
dutta.ritik@iitgn.ac.in
ritik99.github.io

Education

2016 - 2020 **IIT Gandhinagar**, *B.Tech (Hons.), Computer Science & Engineering*, 9.04/10.0.

2016 **Pace Junior Science College**, *High School*, 87.8%.

2014 **D.A.V. Public School, Thane**, *(CBSE)*, 94.8%.

Research Experience

Apr '18 - Ongoing **ChaLearn** (Remote), Dr. Isabelle Guyon & Dr. Kristin Bennett,
Generating Privacy-Preserving Synthetic Medical Data.

- Goal of this project is to generate synthetic medical data to overcome challenges posed by the use of real patient data
- Tested models such as WGANs, VAEs, random forest imputations and additive noise models to generate medical data
- Developing and exploring metrics to benchmark algorithms based on their utility and ability to preserve privacy
- Results of this work have been presented at ESANN 2019 and AIDR 2019

May '19 - Jul '19 **INRIA**, *Paris-Saclay*, Dr. Isabelle Guyon,
Using Observational Causal Discovery for Synthetic Data Generation.

- Extended the Structural Agnostic Model by Kalainathan et al. to support categorical data
- The modified model uses neural networks to learn the underlying causal graph, and the Gumbel-Softmax trick by Jang et al. for categorical reparametrisation
- Began contributing as a collaborator on the Causal Discovery Toolbox, an open-source Python framework for causal discovery from observational data

May '18 - Jul '18 **Texas A&M University**, *College Station*, Dr. Sunil Chirayath.

- Wrote a Python program to implement a nuclear forensics method. Automated processes to reduce the run time from 30 minutes to 5 minutes

Publications, Extended Abstracts

Apr '19 **Privacy Preserving Synthetic Health Data**,
A Yale, S Dash, **R Dutta**, I Guyon, A Pavao, K Bennett.
ESANN 2019

May '19 **Assessing Privacy and Quality of Synthetic Health Data**,
A Yale, S Dash, **R Dutta**, I Guyon, A Pavao, K Bennett.
AIDR 2019

Dec '19 **Synthetic Event Time Series Health Data Generation**,
S Dash, **R Dutta**, I Guyon, A Pavao, A Yale, K Bennett.
Extended abstract at the ML4H Workshop at **NeurIPS 2019**

Jan '20 **Effect of Feature Hashing on Fair Classification**,
R Dutta, V Gohil, A Jain.
Young Researcher's Symposium at **CoDS-COMAD 2020**

Jan '20 **Causal Discovery Toolbox: Uncover causal relationships in Python**,
D Kalainathan, O Goudet, **R Dutta**.
JMLR, Volume 21

Apr '20 **Generation and Evaluation of Privacy Preserving Synthetic Health Data**, A Yale, S Dash, **R Dutta**, I Guyon, A Pavao, K Bennett.
Neurocomputing

Professional Experience

May '17 - Jun '17 **Humbee.in**, Vivek Nautiyal.
- Wrote ReactJS and python programs to stream data from news and social media sources

Teaching Experience

Jan '20 - Present **Teaching Assistant: Machine Learning (ES 654)**, IIT Gandhinagar.
- Preparing lecture slides, assignments & quizzes, and mentoring student project groups
- Gave a guest lecture on causality and fairness in machine learning

Aug '17 - Nov '17 **Teaching Assistant: Computing (ES112)**, IIT Gandhinagar.
- Supervised lab programming sessions, setting and grading questions for exams

Other Research Projects

Oct '19 - Present **Fairlets for fair regression**, Dr. Anirban Dasgupta.
- Fairlets are minimal sets that satisfy the constraints of fair representation with applications in fair clustering. We are looking into extending the notion of fairlets for a fair regression setting

Aug '19 - Ongoing **Markov Decision Processes and Fair Voting**, Dr. Neeldhara Misra.
- Started working on a parameterized approach for the policy iteration algorithm which is used to solve Markov Decision Processes
- We're also simultaneously exploring various fair voting and committee selection algorithms in terms of transferability of fairness guarantees and parameterised approaches to solve them

Aug '18 - Present **Motif Discovery with Topic Models**, Dr. Anirban Dasgupta.
- Exploring the use of parameterized and non-parameterized topic models to discover binding sites of transcription factors (TF) on DNA
- Engaged in literature review of peak-calling methods and motif disocvery for DNA-TF interactions

Open-Source Projects

Jul '19 - Present **Causal Discovery Toolbox (CDT)**.
- CDT is a Python package for causal inference in graphs. I am contributing to the codebase, managing the documentation and fixing bugs

Major Course Projects

Feb '19 - Apr '19 **Image Hashing as an Adversarial Defense**, Dr. Nipun Batra.
- Evaluated the use of image hashing and SEGAN as defenses against adversarial attacks such as the FGSM and C&W attacks

Feb '19 - Apr '19 **Effect of Feature Hashing on Fair Classification**, Dr. Anirban Dasgupta.
- Evaluated the effect of feature hashing data on fair classification under a multi-task setting. The project was accepted at the Young Researcher's Symposium at CoDS-COMAD 2020

Feb '19 - Apr '19 **Detecting Insults in Social Commentary**, Dr. Mayank Singh.
- Used multiple traditional machine learning methods and ensemble methods to detect insults in social media commentary
- Dataset used for testing was part of a Kaggle competition. Our AUC score was 0.811, while the best score was 0.842

Other Projects

Aug '17 - Nov '17 **Adversarial Learning**, Dr. Dinesh Garg.
- Literature review of GANs and adversarial attacks on machine learning models

Aug '18 - Nov '18 **Estimating defocus blur in images**, 3D Computer Vision Course Project.
- Implemented a paper on estimating defocus blur which uses rank of local patches

Aug '18 - Nov '18 **Implementation of the AES algorithm on an FPGA**, Digital Systems Course Project.
- Implemented the AES algorithm on FPGA

Technical Proficiency

Advanced: Python, Pytorch, Sklearn

Intermediate: TensorFlow, C++ , C, R, MATLAB, NLTK, OpenCV, LaTeX

Basic: Chainer, Numba, Keras, Slurm

Relevant Coursework

Advanced Machine Learning, Artificial Intelligence, Machine Learning, 3D Computer Vision, Natural Language Processing, Nature Inspired Computing, Topology

Extra-Curricular Activities

Winner of TechLeaps 2.0 (intra-college technical innovation challenge) for a facial recognition project, with a funding opportunity of Rs. 1 Lakh for wide-scale implementation

Represented the institute as a part of the debate team at the Inter-IIT Cultural Meet 2016

Part of the executive team of the Coding Club

Interested in competitive programming