

Education

- 2019- **Ph.D. candidate, Statistics and Data Science, Yale University, U.S..**
2017-2019 **Computer Science, EPFL, Lausanne, Switzerland.**
2015-2016 **M.Tech., Computer Science and Engineering, IIT Kanpur, India.**
2011-2016 **B.Tech., Computer Science and Engineering, IIT Kanpur, India.**

Experience

Publications

- Preprint **Fair Distributions from Biased Samples: A Maximum Entropy Optimization Framework**, *L. Elisa Celis, Vijay Keswani, Ozan Yildiz, Nisheeth K. Vishnoi.*
- Real-world datasets may not always be the true representation of the underlying distribution of the population, and can unintentionally encode bias against certain groups.
 - We develop a fast, novel algorithm, based on max-entropy optimization principles, that can correct such biases, and learn a "fair" distribution from a given dataset.
- Preprint **Implicit Diversity in Image Summarization**, *L. Elisa Celis and Vijay Keswani.*
- For image selection and summarization problems, many tools such as Google Image Search have been shown to be biased with respect to gender, race, etc.
 - We develop a framework that takes as input a visibly diverse control set of images and a query and returns images from the dataset that correspond to the query and is diverse in a similar manner as the control set.
- Preprint **Improved Adversarial Learning for Fair Classification**, *L. Elisa Celis and Vijay Keswani.*
- Classification algorithms are increasingly being used in many societal contexts such as determining criminal recidivism, predictive policing, and job screening. As such, it is becoming increasingly important that these algorithms are unbiased with respect to race, gender, etc.
 - We suggest improved models and algorithms that use the adversarial paradigm to ensure that the returned classifier is fair and accurate.
- ACM-FAT* 2019 **Classification with Fairness Constraints: A Meta-Algorithm with Provable Guarantees**, *L. Elisa Celis, Lingxiao Huang, Vijay Keswani and Nisheeth K. Vishnoi.*
- Once again in the context of classification, most algorithms ensure parity with respect to a particular fairness metric and cannot usually be generalized to other metrics.
 - We developed a meta-algorithm that takes the fairness type and constraint as input and returns a classifier which satisfies the fairness constraint and has good accuracy.
- ICML 2018 **Fair and Diverse DPP-based Data Summarization**, *L. Elisa Celis, Vijay Keswani, Damian Straszak, Amit Deshpande, Tarun Kathuria, Nisheeth K. Vishnoi.*
- Diverse sampling algorithms are used in many real-world scenarios such as data summarization to capture a representative summary of a large dataset. However for a biased dataset, simple sampling techniques can lead to similar or even aggravated bias in the sampled subset.
 - Working in the framework of Determinantal Point Processes (DPPs), used for diverse sampling, we introduced fairness constraints on the sampling procedure to ensure that the sampled summary is fair and provide a simple linear-time algorithm approximate algorithm to sample from DPPs with fairness constraints.

Demo Papers

IJCAI-ECAI 2018 **Balanced News Using Constrained Bandit-based Personalization**, Sayash Kapoor, Vijay Keswani, Nisheeth K. Vishnoi, L. Elisa Celis.

- We created a news-search prototype that de-polarizes the news feed by presenting balanced viewpoints across liberal and conservative articles.
- The prototype uses a novel constrained bandit-sampling algorithm, introduced in the paper *An Algorithmic Framework to Control Bias in Bandit-based Personalization*.

Master's Thesis

01/2015 - **Laplacian Solvers and Graph Sparsification**, Guide: Dr. Rajat Mittal,
06/2016 CSE Department, IIT Kanpur.

- Solving a linear system of equations usually takes *cubic time*. However, for linear systems associated with the graph Laplacian matrix, the graph sub-structure of the matrix can be exploited to obtain nearly linear time solvers.
- These solvers use Spectral Sparsification of graphs as a subroutine in the algorithm. However their usage extends beyond Laplacian solvers, and have been employed in various settings since their discovery.
- Through this thesis, we tried to understand the scope of spectral sparsification algorithms used in Laplacian solvers and the relations between current state-of-the-art Laplacian solvers.

Industry Experience

07/2016 - **Product Engineer**, Sprinklr, Gurgaon, India.

- Sprinklr aimed to provide a single platform for social media management to enterprise.
- As part of the Advertisements team, we worked to ensure that our platform was compatible with current the Ads APIs of all relevant social networks, such as Facebook, Snapchat, etc.
- We also provided several novel analysis tools to the users for them to judge the performance of their ads and to ensure they maximize their returns from these ads.

Internships

05/2014 - **Validation of a Compiler for Critical Embedded Software**, Guide: Dr.
07/2014 Francesco Zappa Nardelli and Dr. Mark Pouzet, INRIA, Paris-Rocquencourt and ENS, Paris.

- *Lustre* is a data synchronous language, designed for description and verification of real-time systems. The compiler for these systems need to fool-proof to prevent any unexpected outcome.
- In this project considered one such compiler for Lustre, *Heptagon*. To verify this compiler, we built a symbolic validation framework in *OCaml*. It generated the symbolic representation for the *Lustre* program and the resulting C program, which were then compared using *Z3 SMT solvers* to authenticate their equivalence.

05/2013 - **Study of Statistical Learning Algorithms**, Guide: Dr. Ambedkar
07/2013 Dukkipatti, CSA Department, Indian Institute of Science, Bangalore.

- Applied information theoretic concepts suggested by Kamal Nigam and Andrew McCallum, to get a better model for **document classification**.

Teaching Experience

2018 **Machine Learning**, Teaching Assistant, EPFL.

2018 **Theory of Computation**, Teaching Assistant, EPFL.

2016 **Algorithms and Data Structures**, Instructor, ACA Summer School, IIT Kanpur.

2016 **Online Learning and Optimization**, Teaching Assistant, IIT Kanpur.

2015 **Theory of Computation**, Teaching Assistant, IIT Kanpur.

2015 **Data Structures and Algorithms**, Teaching Assistant, IIT Kanpur.

2014 **Fundamentals of Computing**, Teaching Assistant, IIT Kanpur.

Skills

Programming *Advanced*:C, Python (Tensorflow, Numpy, Scipy, Sklearn, Flask) R
Familiar:C++, Java, Haskell, Bash