DATA INSTITUTE CONFERENCE SCHEDULE

San Francisco | March 12-14, 2023 | #DSC023

SUNDAY, MARCH 12

8:00am-9:00am: **Coffee and Registration**, 5th Floor
9:00am-11:30am: **Short Course Session I**

BAYESIAN METHODS IN A/B TESTING
Nathaniel Stevens, University of Waterloo, Rm 450

INTRODUCTION TO DEEP LEARNING WITH PYTORCH
Carlos Garcia, University of San Francisco, Rm 452

12:30pm-3:00pm: **Short Course Session II**

ADAPTING PRETRAINED MODELS FOR DOCUMENT CLASSIFICATION
David Guy Brizan, University of San Francisco, Rm 450

INTRODUCTION TO GRAPH NEURAL NETWORKS AND THEIR GENERALIZATIONS
Mustafa Hajij, University of San Francisco, Rm 452

9:00am-3:00pm: **Datathon**, Rm 527

Hosted by Amazon Web Services & Amazon SageMaker Studio Lab

3:30pm: **Opening Remarks**, Rm 150

Henry Humadi and Shan Wang, Conference Co-Chairs
Jeff Hamrick, Senior Director, Data Institute

4:00pm-5:00pm: **Plenary**, Rm 150

Anima Anandkumar, Bren Professor, Caltech & Senior Director ML Research, Nvidia

5:00pm-6:00pm: **Opening Reception**, 5th Floor
Light Refreshments and Hors d'oeuvres

MONDAY, MARCH 13

8:00am-9:00am: **Coffee and Registration**, 5th Floor
9:00am-11:00am: **Concurrent Sessions**

PRACTICAL ISSUES AND ADVANCES IN A/B TESTING, Rm 527
Chair: **Nathaniel Stevens**, University of Waterloo

Nicholas Larsen, North Carolina State University
HODOR: A Two-Stage Hold-Out Design for Online Controlled Experiments on Networks

Jiannan Lu, Apple
Towards Trustworthy, Efficient and Private A/B Tests: Learnings and Insights

Lo-Hua Yuan, Airbnb
How Airbnb Learns From Product Releases Without A/B Testing

Cindy Zhang, Pinterest
Long-term Holdouts as Ground Truth Measurements

DATA SCIENCE IN SPORTS, Rm 529
Chair: **Steve Devlin**, University of San Francisco

Justin Jacobs, Squared2020, NBA
A Brief Tour into Computer Vision in Sports

Jake Toffler, New York Mets
Model Benchmarking in a Competitive Industry

Thomas Treloar, Hillsdale College
Global and Iterative Ranking Models Based on Network Diffusion

David Uminsky, University of Chicago
NBA Lineup Attribution: A Fourier Transform Approach

TOPICS IN MACHINE LEARNING, Rm 150
Chair: **Jeff Hamrick**, University of San Francisco

Oliver Zeigermann, Open Knowledge
Resilient Machine Learning

Phil Mui, Salesforce
An Embedding Analysis on the Impact of Diversity Awareness on Diversity of Ideas

Yuchi Cho, University of Southern California
Flexible and Robust Real-Time Intrusion Detection Systems to Network Dynamics

Connor Gibbs, Colorado State University
ECoHeN: A Hypothesis Testing Framework to Extract Communities from Heterogeneous Networks

MONDAY, MARCH 13

11:15am-12:15pm: Mentor Mentee Lunch, 5th Floor

12:30pm-1:30pm: Plenary, Rm 150

DJ Patil, Former U.S. Chief Data Scientist

1:30pm-2:00pm: Coffee Break
2:00pm-4:00pm: Concurrent Sessions

DATA PRIVACY AND SECURITY RISKS EXPOSED BY AI, Rm 527

Chair: **Hyrum Anderson**, Robust Intelligence

Daniel Zielaski, Salesforce
The importance of AI to Business, Marketing, and Digital Products

Naveen Jain, Salesforce
Privacy by Design: Building Privacy into your Data Systems from the Ground Up

Hyrum Anderson, Robust Intelligence
Data and Model Supply Chain Risk in the AI Development Lifecycle

Erwin Quiring, International Computer Science Institute (ICSI) and Ruhr University Bochum
Dos and Don'ts of Machine Learning in Computer Security

MACHINE LEARNING APPLICATIONS IN INDUSTRY, Rm 529

Co-Chairs: **Jennifer Zhu** and **Vidya Sagar Ravipati**, Amazon Web Services

Sean McCurdy, Pinterest
Intro to Pinterest Ads and ML Systems

Saman Sarraf, Johnson & Johnson
Computer Vision Applications Across Industries: Wins and Challenges

Ruixuan Zhang, Airbnb
Machine Learning Driven User Growth and Lifecycle Engagement Marketing

April Liu, Intuit
AI Application in Fighting Synthetic Account Fraud

STATISTICAL THEORY AND APPLICATIONS, Rm 150

Chair: **Daniel Jerison**

Moinak Bhaduri, Bentley University
Distribution-free, Online Change Detection in Multidimensional Point Processes Through Repeated Testing

Wiranthe Bandara Herath, Drake University
Dimension Reduction for Vector Autoregressive Models

Mustafa Hajij, University of San Francisco
What is Topological Deep Learning?

Greeshma Balabhadra, Stony Brook University
High-Frequency Risk Estimators Using Change Point Detection Methods

4:30pm-5:30pm: Data Science as an Equalizer, Rm 150

Mark Freeman, On the Mark Data
Carly Villareal, Nextdoor

Chair: **Dr. Bushra Anjum**, Doximity

Megan Yahya, Google Cloud
Elaine Zhou, Change.org

5:00pm-6:00pm: Opening Reception, 5th Floor
Light Refreshments and Hors d'oeuvres

TUESDAY, MARCH 14

8:00am-9:00am: Coffee and Registration, 5th Floor
9:00am-11:00am: Concurrent Sessions

AI FOR PRECISION MEDICINE AND LEARNING HEALTHCARE SYSTEMS, Rm 527

Chair: **William Bosl**, University of San Francisco

Conrad Yiu, Trident.ai
AI in Medicine: What Are We Really Trying to Do?

Vincent Liu, Kaiser Permante Research
Augmented Intelligence: The Future of AI in Healthcare

Sophia Wang, Stanford University
Envisioning the Future: Artificial Intelligence for Preserving Eyesight

William Bosl, University of San Francisco
AI Meets Cura Personalis: Computational Biomedical Knowledge for Integrated, Precision Medicine

PRACTICAL ISSUES AND ADVANCES IN A/B TESTING (II), Rm 529

Chair: **Nathaniel Stevens**, University of Waterloo

Wenjing Zheng, Netflix
Double Robust Causal Effect Generalization and Transportation under Covariate Shifts

Nick Ross, University of Chicago
Hidden Integration Costs in Online Controlled Experimentation Platforms

Nathaniel Stevens, University of Waterloo
General Additive Network Effect Models: A Framework for the Design and Analysis of Experiments on Networks

Steve Howard, The Voleon Group,
Augmented Inverse Propensity Weighting for Randomized Experiments

MACHING LEARNING IN INDUSTRY, Rm 150

Session Chair: **Daniel O'Connor**, University of San Francisco

Dili Ezeme, Ab-InBev
New Assortment Recommendation Approach for Differentiated Products

Rushil Manglik, University of San Francisco
Fine-Tuning Layout Parser Deep Learning Model for Document Image Analysis and Table Extraction

Diane Woodbridge, University of San Francisco
Bundle Recommender for Complimentary Menus

Sanghamitra Deb, Chegg Inc.
Computer Vision Landscape at Chegg: Present and Future

11:30am-12:20pm: Lunch Break

TUESDAY, MARCH 14

12:30pm-1:30pm: Plenary

Eric Berlow, CEO Vibrant Data Labs

1:30pm-2:00pm: Coffee Break
2:00pm-4:00pm: Concurrent Sessions

PRIVACY AND PERSONALIZATION IN MARKETING, Rm 527

Chair: **Daniel Zielaski**, Salesforce

Carson Forter, Twitch (Amazon)
Putting Personalization to the Test: A Causal Inference Approach

Sherry Guo, Salesforce
From 0 to 1: Leveraging Data Science Algorithms for Personalization

Chad Kimner, Meta Reality Labs
Navigating the Intersection of Personalization and Privacy: A Marketer's Perspective

Maura Tuohy, Discord
The Fine Line: Balancing Personalization and Legal Compliance

REINFORCEMENT LEARNING FOR WEALTH MANAGEMENT, Rm 529

Chair: **Matthew Dixon**, Illinois Institute of Technology

Karén Chaltikian, Money Lion
Wealth Management Benchmarks

Sanjiv Das, Santa Clara University
Optimizing The Probabilities Of Obtaining Single And Multiple Financial Goals

Matthew Dixon, *Illinois Institute of Technology*
Time Consistent Risk-Aware Q-Learning of Optimal Consumption under Epstein-Zin Preferences

MACHINE LEARNING AND STATISTICS IN MEDICINE, Rm 150

Chair: **Shan Wang**

Zirui Zhang, University of California, Irvine
Parameter Inference in Diffusion-Reaction Models of Glioblastoma Using Physics-Informed Neural Networks

Souradipto Ghosh Dastidar, University of Minnesota Twin Cities
Correcting for Spatial Correlation in Functional Connectivity from task-based fMRI

Vignesh Ravindranath, University of California, San Francisco
Optimizing Treatment Selection in Crohn's Disease Using Patient-Specific Features: An Individual Participant Data Meta-Analysis of Fifteen Randomized Controlled Trials

4:30pm-5:30pm: Ethical Perspectives of AI-Generated Art, Rm 150

Quinn Keck, Disney Streaming
Chris Brooks, University of San Francisco

Chair: **Robert Clements**, University of San Francisco

Paul Kim, University of California, Santa Barbara
David Aughenbaugh, Factory VFX & EyeTriplmages.com

5:30pm-6:30pm: Closing Reception, 5th Floor
Light Refreshments and Hors d'oeuvres

Thank You — Sponsors & Members

Craig Newmark Philanthropies
aws
liftoff
VeNoM Arc Technologies
University of San Francisco | Master of Science in Applied Economics
Pacific Gas and Electric Company
BlackRock
Meta
The Nature Conservancy
GORE
Association of Bay Area Governments | Metropolitan Transportation Commission
pendulum

IN COOPERATION WITH:

siam
NSF
ASA

ABOUT THE DATA INSTITUTE

MISSION

- Foster new partnerships between industry and academia to tackle industrial data science problems
- Build an inclusive community of data scientists
- Support data science research
- Create innovative curriculum to support the training of the next generation of ethical data scientists
- Partner with nonprofit and civic organizations to seek data-driven solutions to address pressing social, economic and environmental challenges

Founded in 2016 at the University of San Francsico, the Data Institute serves as the umbrella organization for data science education, activities and programming at USF. The Data Institute houses interdisciplinary data science research, the Bachelors and Masters in Data Science degree programs, and continuing education certificates. Through our strong industrial-academic partnerships, we provide unparalleled experiential learning opportunities for our students.

Affiliated Data Institute faculty, researchers and postdoctoral fellows explore the latest theoretical advances and applications in data science. Key research initiatives focus on medical applications of data science, the environment, and ethics.

The Data Institute corporate, nonprofit and civic organizations membership provides engagement opportunities such as faculty consulting, practicum placements, talent acquisition and on-site training. Through our innovative certificate programs, members can provide employee development and education. Membership can also include access to our industrial-academic postdoctoral fellows program.

ABOUT THE UNIVERSITY OF SAN FRANCISCO

UNIVERSITY OF SAN FRANCISCO

CHANGE THE WORLD FROM HERE

The University of San Francisco, the city's first university, was established by the Jesuits in October 1855. Jesuit tradition defines USF's approach to learning and our commitment to welcoming students of every faith and no faith. Our vision and mission are the foundations of our university, and reflect the shared views of our institution.

The core mission of the university is to promote learning in the Jesuit Catholic tradition. The university offers undergraduate, graduate, and professional students the knowledge and skills needed to succeed as persons and professionals, and the values and sensitivity necessary to be people for others.

The university distinguishes itself as a diverse, socially responsible learning community of high quality scholarship and academic rigor sustained by a faith that does justice. The university draws from the cultural, intellectual, and economic resources of the San Francisco Bay Area and its location on the Pacific Rim to enrich and strengthen its educational programs.

OMC12487