

# 2022 SDSU Data Science Symposium Schedule

## Monday, February 7

Time	Pasque 255	Dakota Room 250 A/C	Phesant
12:30 pm - 5:00 pm	Check-in/Registration (Location : Prairie Lounge)		
1:00 -5:00 pm	<b>Workshop 1</b> <i>Omics Data Science</i> Ali Rahnavard and Himel Mallick	<b>Workshop 2</b> <i>Microsoft Power BI</i> Negassi Tesfay	Workshop 3 Blockchain Analytics Cuneyt Akcora
	<b>Banquet (Performing Art Center)</b>		
6:00 - 8:30 pm	Social time (Cash bar)		
6:30 - 8:00 pm	Dinner		
7:15 - 7:30 pm	<b>Welcome</b> Dr. Kurt Cogswell - Head of Mathematics and Statistics Department		
7:30 - 8:30 pm	<b>Keynote</b> <i>Dissecting Blockchain Analytics: What We Can Learn from Topology and Geometry of Blockchain Transaction Graphs</i> Dr. Yulia R. Gel, University of Texas Dallas		

Tuesday, February 8				
Time	Dakota Room 250 A/C	Pheasant Room 253 A/B	Herold Crest 253 C	Pasque 255
7:30 - noon	Check-in / Luggage Check (Location: Volstorff Lounge)			
7:45 - 8:15 am	Breakfast (Location: Volstorff B)			
8:30-8:45 am	Opening Session Welcome and Introduction: President Barry Dunn (Location: Volstorff B)			
8:45-9:45 am	Keynote: Data Science before, during, (and after!) a global pandemic Speaker: Dr. Justin Smith (Location: Volstorff B)			
9:50-10:50 am	<p>Session 1: Finance Chair: Thomas B. - SDSU</p> <p><i>The Role of Data Science in Fraud Detection</i> Michael Lim – TransUnion</p> <p><i>Modeling Consumer Risk: A Comparison of Logistic Regression, Scorecard, and Machine Learning Models</i> Sommer West – Capital Services</p>	<p>Session 2: Precision agriculture Chair: Dr. Hossein Moradi - SDSU</p> <p><i>A vision guidance system on agricultural sprayers reduces operator stress</i> Travis Burgers – Raven Industries</p> <p><i>Dataset for Precision Agriculture Practices</i> Jiyul Chang– SDSU</p>	<p>Session 3: Cancer Genomics Chair: Dr. Tobias Meissner - AVERA</p> <p><i>Understanding the multi-scale heterogeneity in cancer genomes</i> Anu Amallraja – Avera Cancer Institute</p> <p><i>16 Year Life History and Genomic Evolution of an ER+ HER2- Breast Cancer</i> Bing Xu–Avera Cancer Institute</p> <p><i>Patient-specific analysis and visualization of cancer pathways</i> Priya Swaminathan – Avera Cancer Institute</p>	<p>Session 4 Unsupervised learning methods Chair: Dr. Shuchismita Sarkar</p> <p><i>Conditional mixture modeling and model-based clustering</i> Yang Wang – College of Charleston</p> <p><i>Model-based clustering of directed weighted networks</i> Volodymyr Melnykov – The University Alabama</p>
10:50-11:00 am	Networking break / Exhibitors (Location: Volstorff A)			
11:00 am-12:00 noon	<p>Session 5: Finance Chair: Thomas B. - SDSU</p> <p><i>Equipment Finance Credit Risk Modeling - A Case Study in Creative Model Development &amp; Nimble Data Engineering</i> Ed Krueger, Landon Thomson, and Josh- Channel Partners Finance</p> <p>Discussion</p>	<p>Session 6: Data Science and Startups Chair: TBD</p> <p><i>Decentralized Data Access and Analysis for Cybersecurity Usecases</i> Craig Jorgensen - Customer Success Manager Query AI</p>	<p>Session 7: Genomics Chair: Xijin Ge - SDSU</p> <p><i>Decoding infectious disease omics data: COVID-19 case study</i> Ali Rahnavard – George Washington University</p> <p><i>Tweedie mixed models for spatial transcriptomics and digital pathology</i> Himel Mallick – Merck Research Laboratories</p>	<p>Session 8: Unsupervised learning methods Chair: Dr. Shuchismita Sarkar</p> <p><i>Studying contributions of variables to classification</i> Yana Melnykov – The University of Alabama</p> <p><i>Semi-supervised clustering of time-dependent categorical sequences with application to discovering education-based life patterns</i> Yingying Zhang – University of South Alabama</p>
12:00 - 1:00 pm	Lunch (Volstorff B)			
1:00 - 2:00 pm	Poster session Student poster competition (Location: Volstorff A)			
1:00 - 2:25 pm	Job Fair/ Recruiting Exhibitors (Location: Volstorff A)			
2:30 - 3:25 pm	<p>Session 9: Methods Chair: T. B. zoom session</p> <p><i>Advances in Explanation-friendly Scorecard Technology</i> Gerald Fahner – FICO</p> <p><i>Forecasting During Uncertain Times</i> Lance Cundy -- Federal Reserve Bank</p>	<p>Session 10: Applications Chair: TBD</p> <p><i>Linguistic Cues in Disaster Relief Crowdfunding</i> Kari Sandouka- Dordt College</p> <p><i>Factors of Significance for Graduating at a Private College</i> Jennifer Schon, Northwestern College</p> <p><i>Impacts on Student-Athletes' GPA at Northwestern College</i> Jennifer Schon, Northwestern College</p>	<p>Session 11: The epidemiology of COVID-19 Chair: Emma Spors</p> <p><i>Development and Evaluation of a Rural SARS-Cov-2 Hospital Admissions Predictive Model</i> Kayli Rageth– Avera Health</p> <p><i>Data and Epidemiology: The Need to Educate the Public</i> Bonny Specker– South Dakota State University</p>	<p>Session 12: Forensic Statistics Chair: C. Saunders (SDSU) and A. Carrquiry (Iowa state un)</p> <p><i>Topological data analysis of dynamic Ethereum token networks</i> Yuzhou Chen - Princeton University</p> <p><i>Forensic Handwriting Identification using Random Forests and Score-based Likelihood Ratios</i> Danica Ommen - Iowa State Universty Madeline Q. Johnson, Boston Scientific</p>
3:25 - 3:30 pm	Break			
3:30 - 4:25 pm	<p>Session 13: Student speed session Chair: TBD</p> <p><i>On statistical estimates of the inverted Kumaraswamy Distribution under adaptive type-I progressive hybrid censoring</i> - Qingqing Li and Yuhlong Lio - USD</p> <p><i>Generative Adversarial Networks in Tumor-Reglated Research: A Review and Agenda for Moving Forward</i> - Andrew Behrens - DSU</p> <p><i>Development of strategies for estimating a response surface to characterize a black-box algorithm in terms of a white-box algorithm</i> - Cami Fuglsby- SDSU</p> <p><i>Predicting automobile accident severity and hotspots using multinomial logistic regression</i> Zhuoyu Yang - Minot State University</p>	<p>Session 14: Tools for Data Science Chair: Dr. Zeng - Dakota State University(DSU)</p> <p><i>Deploying Live Dashboard Data API using USDA Data APIs to inform farmers/producers</i> Indira Fuyal and Paurakh Paudel - DSU</p> <p><i>Fine-tuning Transformer-based Natural Language Generation Algorithms for USDA Grains Reports</i> Winston Zeng - University of Arizona</p> <p><i>Towards Long Term Impact of DL Models in Medical Imaging</i> Krutika Sutrave and Rajesh Godasu- DSU</p>	<p>Session 16: Methods Chair: C. Saunders (SDSU) and A. Carrquiry (Iowa state University)</p> <p><i>How Do You Define a Circle? Perception and Computer Vision Diagnostics</i> Susan Vanderplas - University of Nebraska-Lincoln</p>	
4:30 - 5:00 pm	Closing Session Thomas Brandenburger (Location: Volstorff B) Winners of poster announcement			

Tuesday, Feb 8, 2022

Poster presentations - Volstorff A			
No	First name	Last name	Title
1	Loknath	Ambati	A comparative analysis of topic modeling Techniques for short text data
2	Andrew *	Behrens	Generative Adversarial Networks in Tumor-Reglated Research: A Review and Agenda for Moving Forward
3	Dylan	Borchert	An alpha-based prescreening methodology for a common but unknown source likelihood ratio with different subpopulation structures
4	Cami*	Fuglsby	Development of strategies for estimating a response surface to characterize a black-box algorithm in terms of a white-box algorithm
5	Indira	Fuyal	Deploying Live Dashboard Data using USDA Data APIs to inform farmers/producers
6	Laxmi*	Gorugantu	Detection of prostate cancer using machine learning techniques: An exploratory study
7	Janean	Hanka	Atypicality Based Measures for the Identification of Counterfeit Aspirin
8	Jason	Hasse	Multivariate Statistics Applied to Additive Manufacturing of Inconel
9	Hoang	Nguyen	Predicting US wildfire possibility and severity
10	Kosman	Rajapaksha	Wald Type Tests with the Wrong Dispersion Matrix
11	Kari	Sandouka	Antecedents of Success for Proficiency in Core Skills
12	Sadixa	Sanjel	Generalized Estimating Equations (GEE) Approach for Clustered Binary Data with Application to COVID-19 Treatment
13	Andrew	Simpson	Identifying Subpopulations of a Hierarchical Structured Data using a Semi-Supervised Mixture Modeling Approach
14	Emma	Spors	Standardized incidence ratio of the COVID-19 pandemic in a midwestern state
15	Zhuoyu*	Yang	Predicting automobile accident severity and hotspots using multinomial logistic regression
16	Winston	Zeng	Fine-tuning Transformer-based Natural Language Generation Algorithms for USDA Grains Reports for Farmers, Producers, and Small Businesses

\* Students giving 10 min oral presentation from 3:30-4:30 in Session 7 in addition to poster presentattion