THE UNIVERSITY OF RHODE ISLAND

Introduction: Lightning Talks

Department of Computer Science and Statistics

2023

Krishna Venkatasubramanian☼
Jing Wu☂

☼ Graduate Director, Compute Science, ☂ Graduate Director, Statistics

THINK BIG WE DO™

What are lightning talks?

- Lightning talks are intended to inform students (both graduate and undergraduate) about the research going on our department

- The idea is for professors in the department to give a quick (5-minute) pitch about their research work/projects

- There won't be much time for questions. So taking notes would be helpful!

How should you use the information you get here?

- As you hear each professor, see if any of the project they are working on, or their general research area interests you

- If it does, feel free to contact the professor(s) and talk to them about the common interests

- For PhD and MS (thesis) students this is particularly important as:
  - you need to **find your advisor** as soon as possible to be able to start with your research work

Major areas of CS research

- In **computer science**, some of the major areas of research include: algorithms, accessibility, computer algebra, computational biology, CS education, computer vision, data science/machine learning, databases, digital forensics, human-computer interaction, programming languages, and simulation.

What if I miss something

- If you miss something that a professors talked about, you can find more information about the various labs and research groups in the department at:

https://web.uri.edu/cs/research/

Statistics information

- Major areas of departmental research include Bayesian inference, big data analysis, biostatistics, high dimensional inference, machine learning, missing data analysis, network data analysis, nonparametric statistics, precision medicine, spatio-temporal data analysis, time series analysis, …

- Several professors also work at the intersection of statistics and computer science.

Statistics Graduate Courses

Spring 2024	Fall 2024
STA 545 Bayesian Statistics	STA 502 Applied Regression Analysis
STA 542 Categorical Data Analysis	STA 500 Analysis of Missing Data
STA 592 Survival Analysis	STA 525 Programming and Data Management in SAS
STA 592 Advanced Inference	STA 585 Network Data Analysis
	STA 560 Time Series Analysis

STA 515 Spatial Data Analysis
STA 536 Applied Longitudinal Analysis
STA 550 Ecological Statistics

AGENDA

Time	Faculty Name
12:30-12:35	Tim Colaneri (CS)
12:35-12:40	Yichi Zhang (Stats)
12:40-12:45	Lisa DiPippo (CS)
12:45-12:50	Gavino Puggioni (Stats)
12:50-12:55	Sarah Brown (CS)
12:55-1:00	Jing Wu (Stats)
1:00-1:05	Mark Yu (Stats)
1:05-1:10	Noah Daniels (CS)
1:10-1:15	Shaun Wallace (CS)
1:15-1:20	Ed Lamagna (CS)
1:20-1:25	Jean-Yves Hervé (CS)
1:25-1:30	Krishna Venkatasubramanian (CS)
1:30-1:35	Guangyu Zhu (Stats)
1:35-1:40	Vic Fay-Wolfe (CS)
1:45-1:50	Abdeltawab Hendawi (CS)
1:50-2:00	GAU Rep

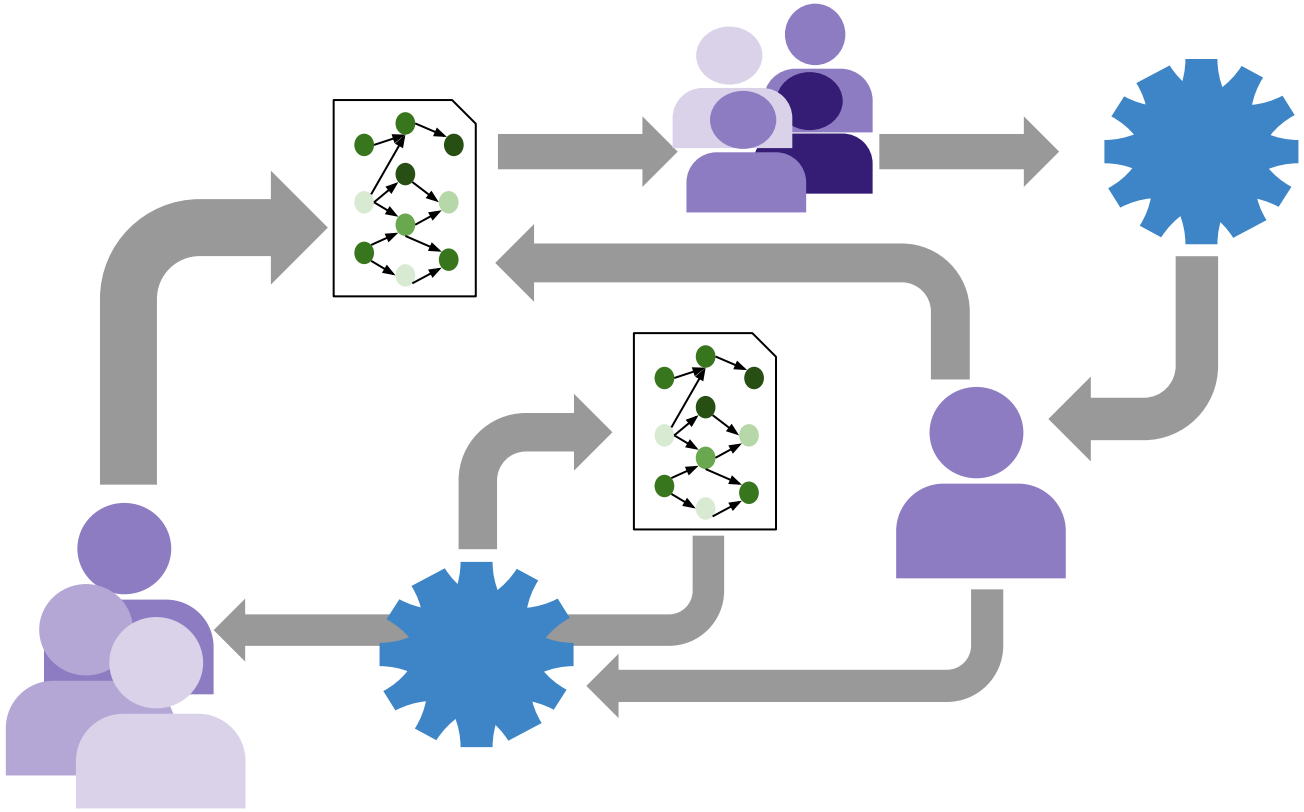
LUNCH!!!

ml4sts.com

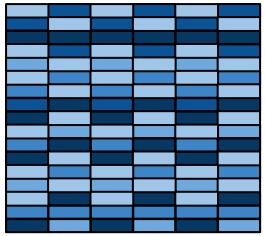
Machine Learning

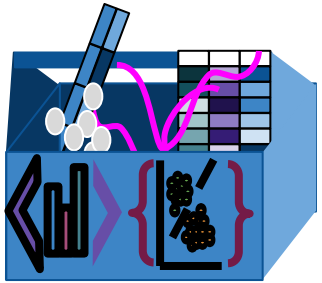
for

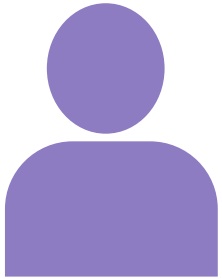
Sociotechnical Systems



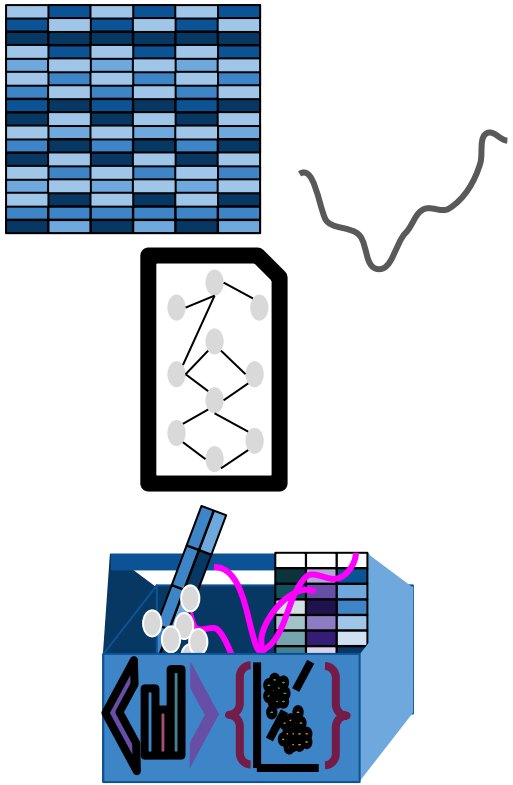
Data Scientist’s Recourse



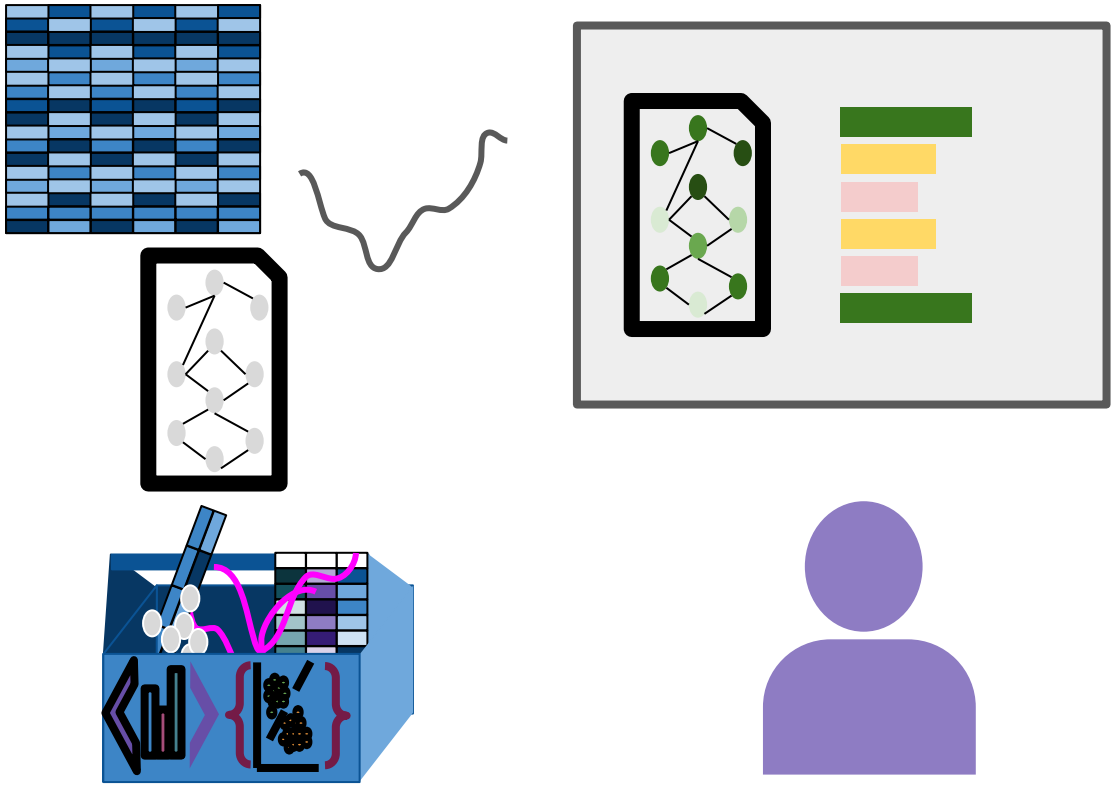




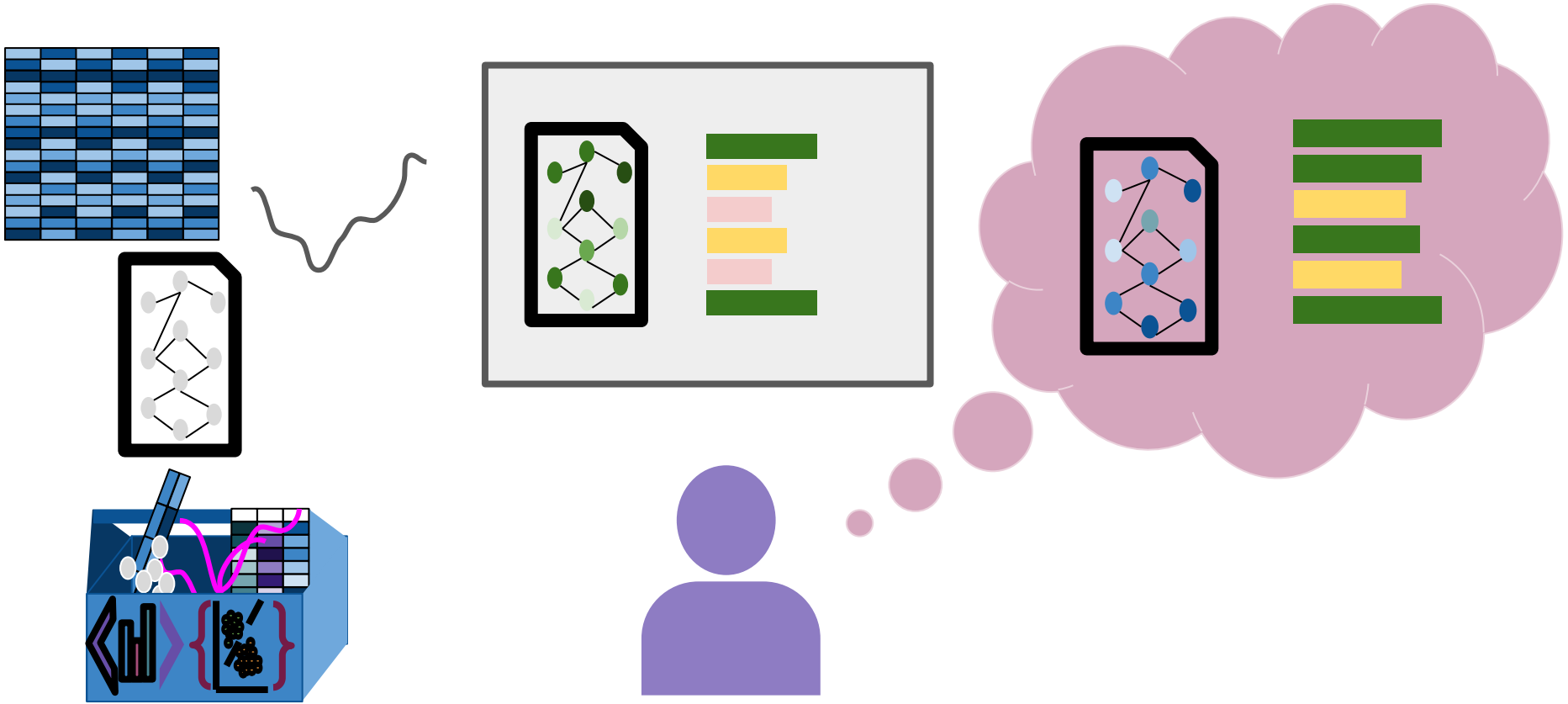
Data Scientist's Recourse



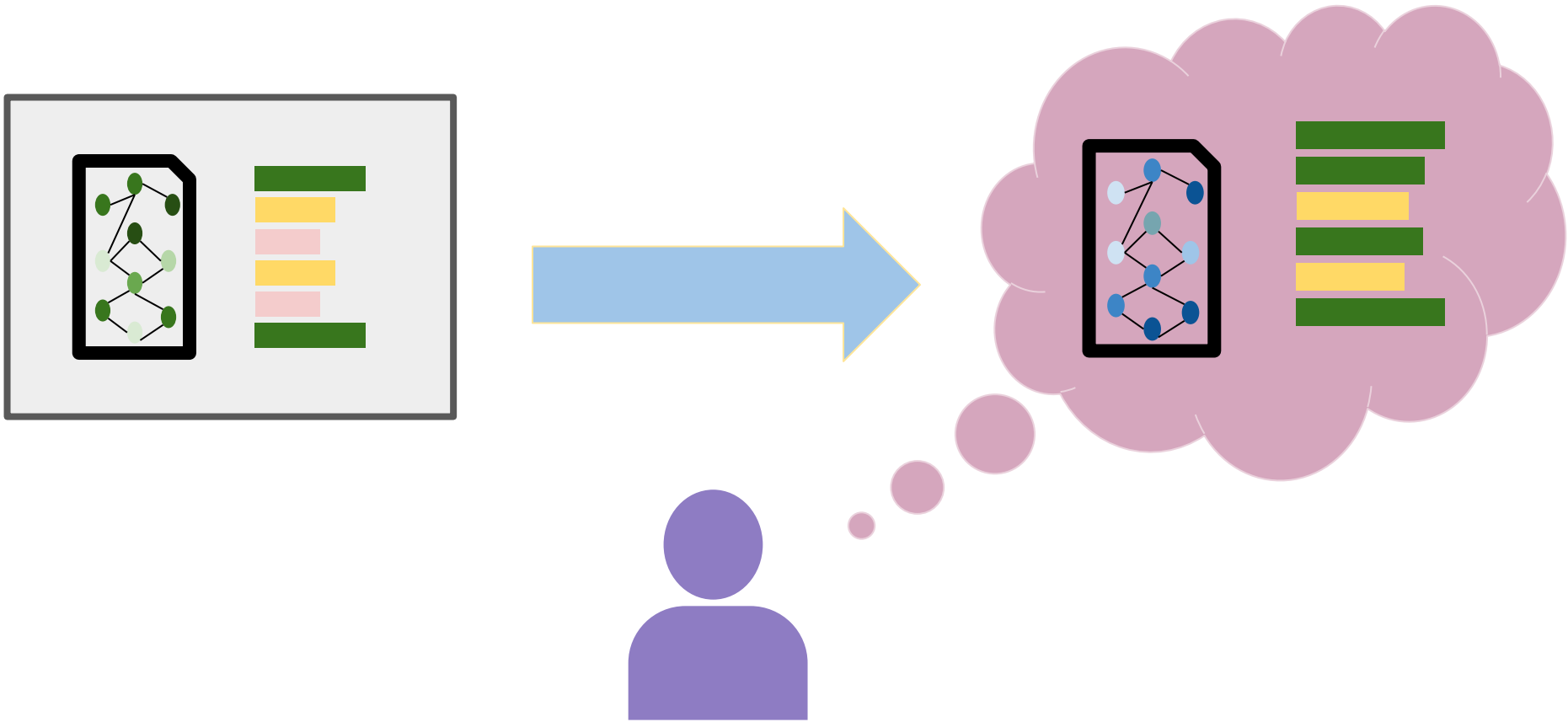
Data Scientist's Recourse



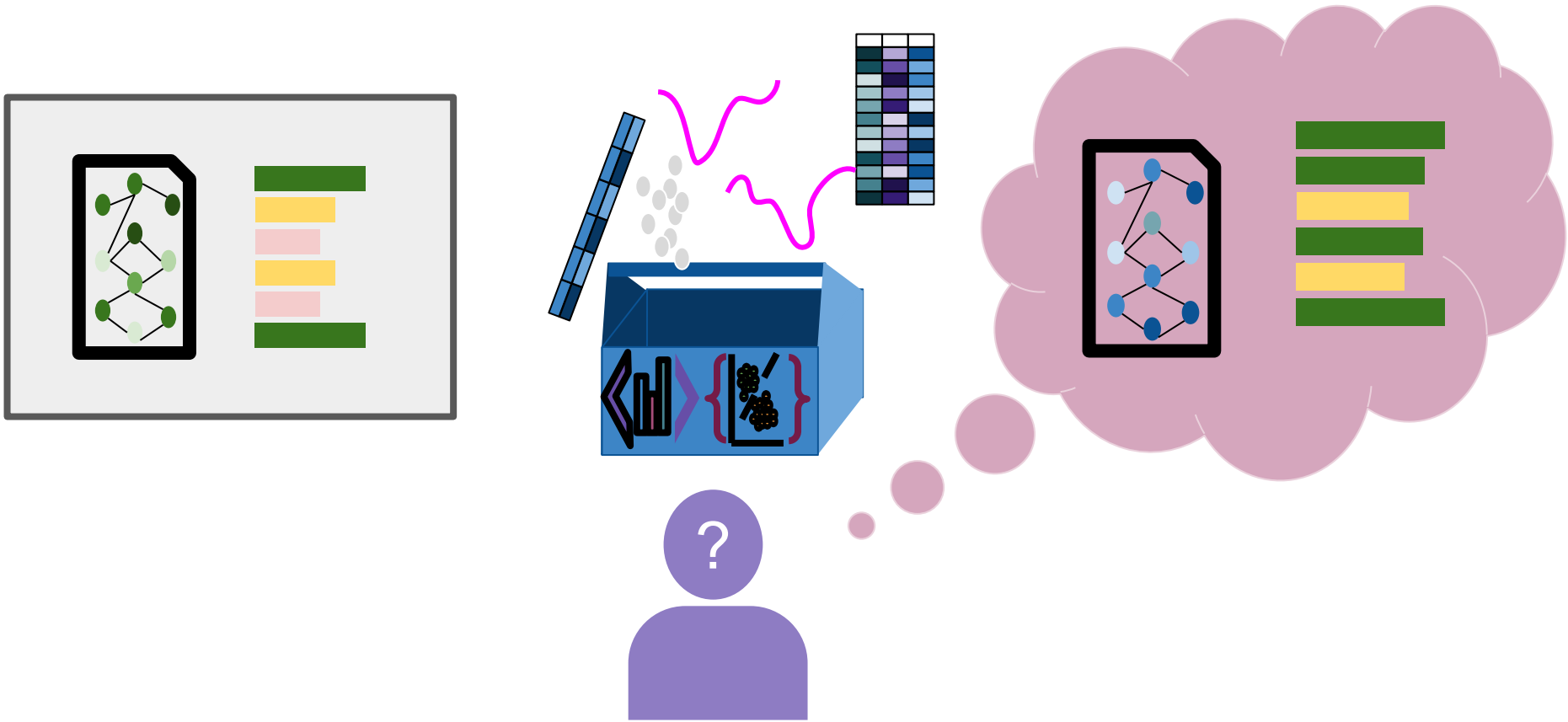
Data Scientist's Recourse



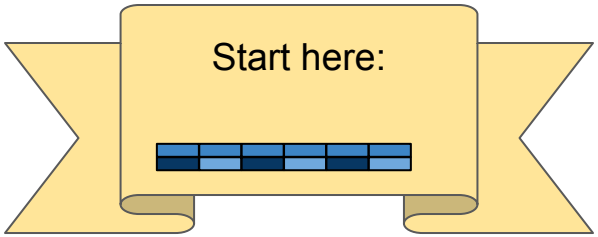
Data Scientist's Recourse

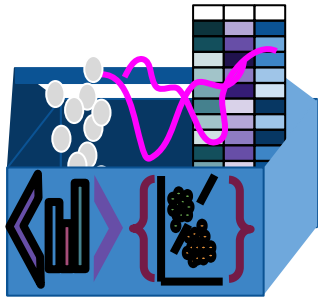


Data Scientist's Recourse

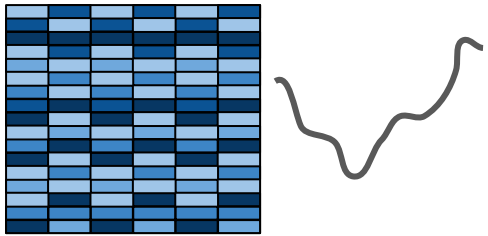


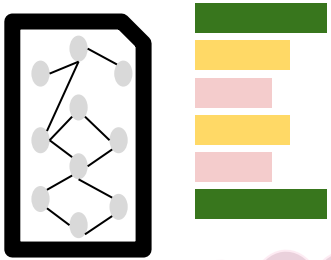
Data Scientist's Recourse

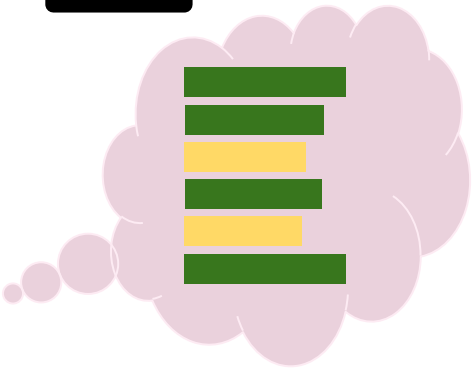


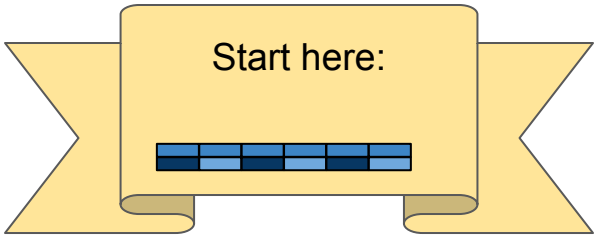


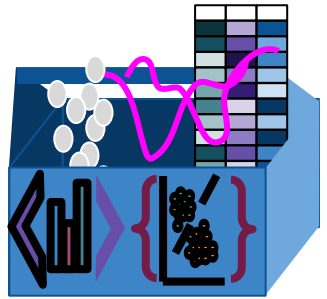
Data Scientist's Recourse



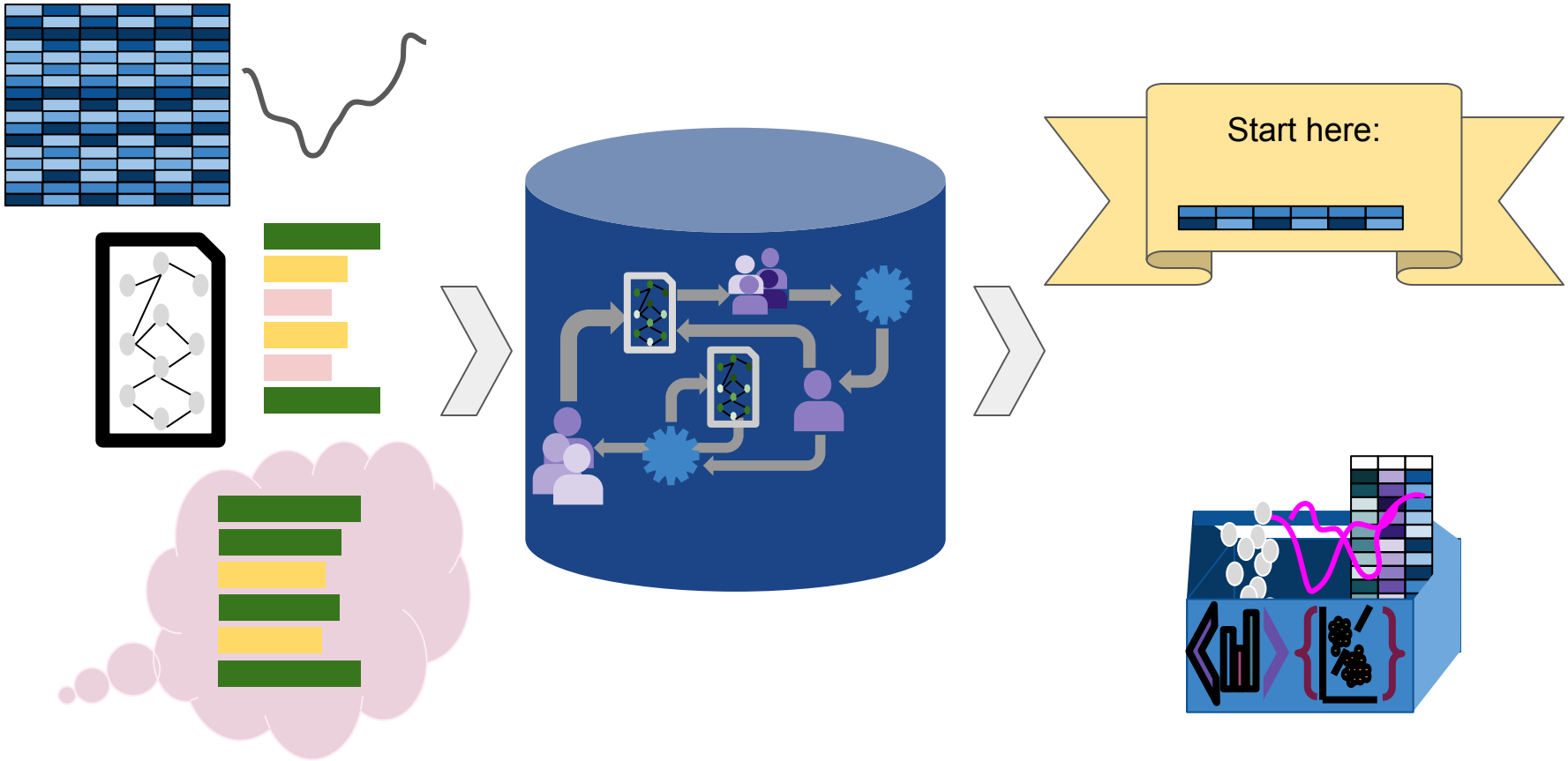






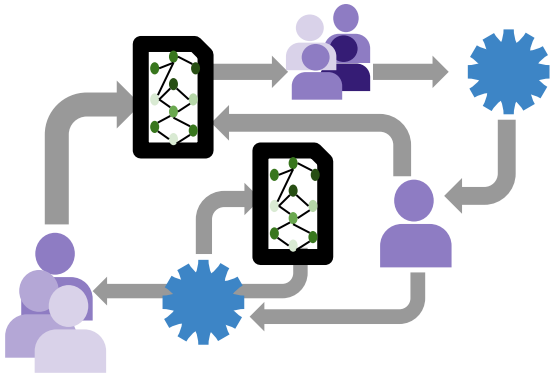


Data Scientist's Recourse



Types of work:

- Statistical modeling of biases
- Deep learning to represent complex data
- Software infrastructure for fair data science



Get Started:

ml4sts.com/join

brownsarahm@uri.edu

ASTEROID
TOWARDS PROGRAMMING WITH FIRST-CLASS PATTERNS

- Asteroid is an open source, dynamically typed, multi-paradigm programming currently under development here, at the University of Rhode Island.
- Asteroid supports patterns as first-class citizens enabling whole new ways of solving problems.
- In this new programming paradigm, patterns and pattern-matching are supported by all major programming language constructs, making programs succinct and robust.

```
let POS_INT = pattern with (n:%integer) %if n > 0.
let NEG_INT = pattern with (n:%integer) %if n < 0.
```

Running Asteroid

```
$ asteroid
Asteroid Version 1.0.0
Run "asteroid -h" for help
Press CTRL+D to exit
> load system io.
> io @println "Hello, World!".
Hello, World!
>
```

```
function fact
  with 0 do
    return 1
  orwith *POS_INT do
    return n * fact (n-1).
  orwith *NEG_INT do
    throw Error("undefined for "+n).
  end.
```

```
function sign
  with 0 do
    return 0.
  orwith *POS_INT do
    return 1.
  orwith *NEG_INT do
    return -1.
  end.
```

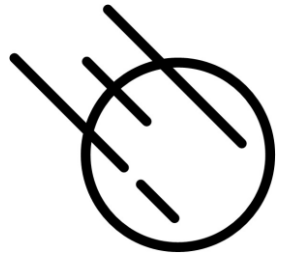
Official Release:
https://asteroid-lang.org

Interested in contributing?
Contact: lutzhamel@uri.edu

Research Team:

Professor Lutz Hamel
Timothy Colaneri, Ariel Finkle, Theodore Henson, Calvin Higgins, Oliver Mclaughlin, Christian Tropeano

Dept. of Computer Science and Statistics
University of Rhode Island
Kingston, Rhode Island, USA
www.uri.edu



Algorithms for Big Data
the cloud isn't enough

CLAM (Clustered Learning of Approximate Manifolds)
CHAODA (Clustered Hierarchical Anomaly and Outlier Detection Algorithms)
CLAM CAKES (CLAM-Accelerated KNN Entropy-scaling Search)

Noah M. Daniels, Ph.D github.com/URI-ABD

MEDFORD

A human- and machine-readable markup language to facilitate FAIR coral metadata

John Freeman, Jill Ashey, Polina Shpilker, Hailey McKelvie, Hollie M. Putnam, Jane Greenberg, Lenore Cowen, Alva Couch, **Noah M. Daniels**

Supported in part by the National Science Foundation under grants 1939795 (to HP), 1940239 (to JG) and 1939263 (to LC)

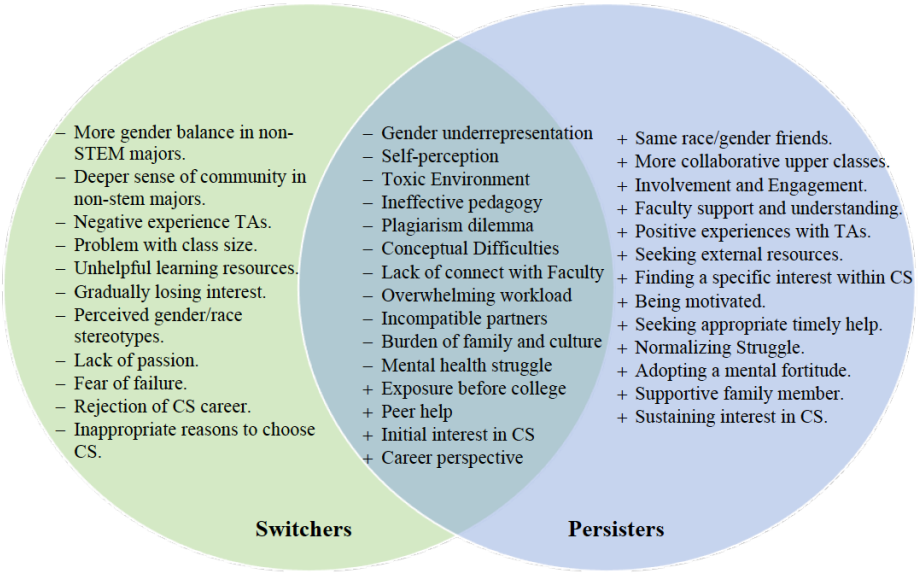
Shpilker, et al. 2022, "MEDFORD: A human and machine readable metadata markup language". Database, OUP

Retention of Underrepresented Students in Computer Science

Dr. Lisa DiPippo

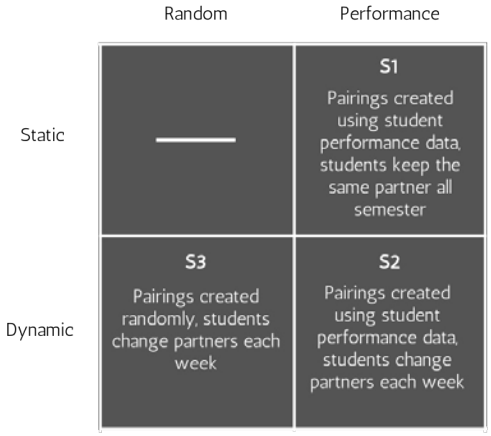
ANALYZING FACTORS THAT CONTRIBUTE TO ATTRITION OF UNDERREPRESENTED UNDERGRADUATE STUDENTS IN COMPUTER SCIENCE

Noura Albarakati



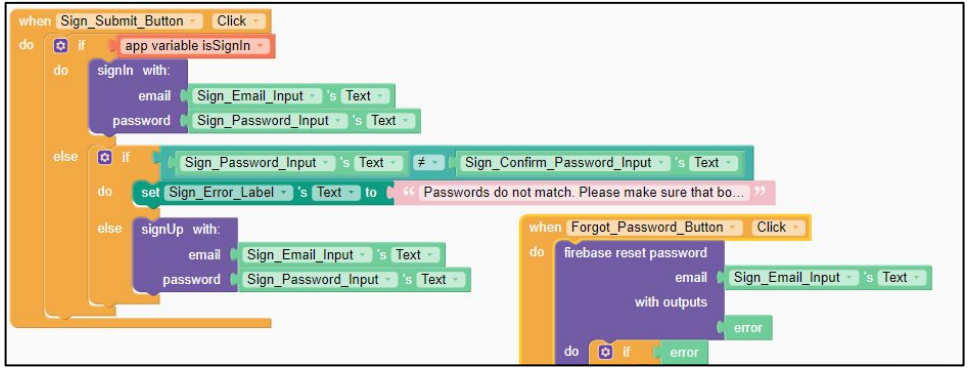
EXAMINATION OF VARYING PAIR PROGRAMMING STRATEGIES IN COMPUTER SCIENCE LABORATORY ENVIRONMENTS

Lily Sisouvong



CS Education Research

Dr. Vic Fay-Wolfe
URI Professor of Computer Science
CSC 305 Course Coordinator
State CS4RI K-12 CS Lead
Academic Co-Director URI Cybersecurity Program

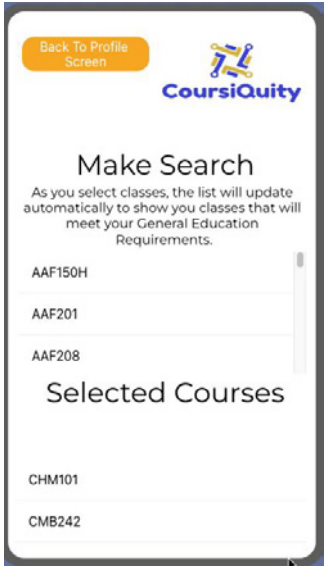


- *Using Low-Code/No-Code to Teach Software Engineering* - currently seeking MS students

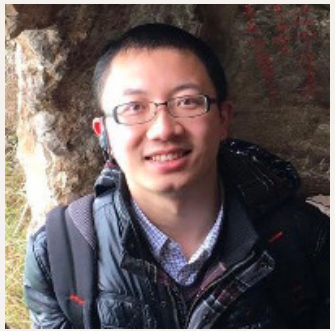
- *Work-based learning in K-12 Computer Science*.
  $3.5M Dept of Education grant
  www.k12.cs.uri.edu/

- *CyberCorps Scholarship For Service*.
  $3.9M NSF Grant
  web.uri.edu/cs/academics/cybercorps-scholarship-for-service/

- *OnRamp Bridge To A Cybersecurity Career*.
  $350K NSF Grant
  web.uri.edu/cs/academics/onramp/

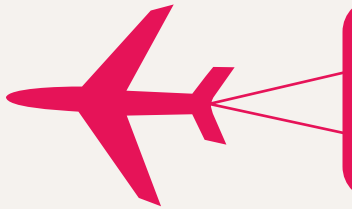


A brief introduction to my research

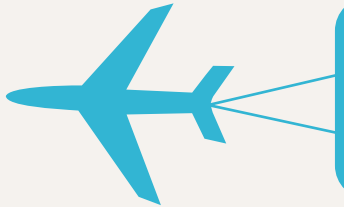


Guangyu Zhu

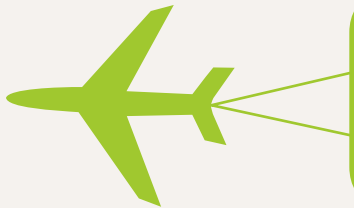
Research Areas



High Dimensional Data Analysis

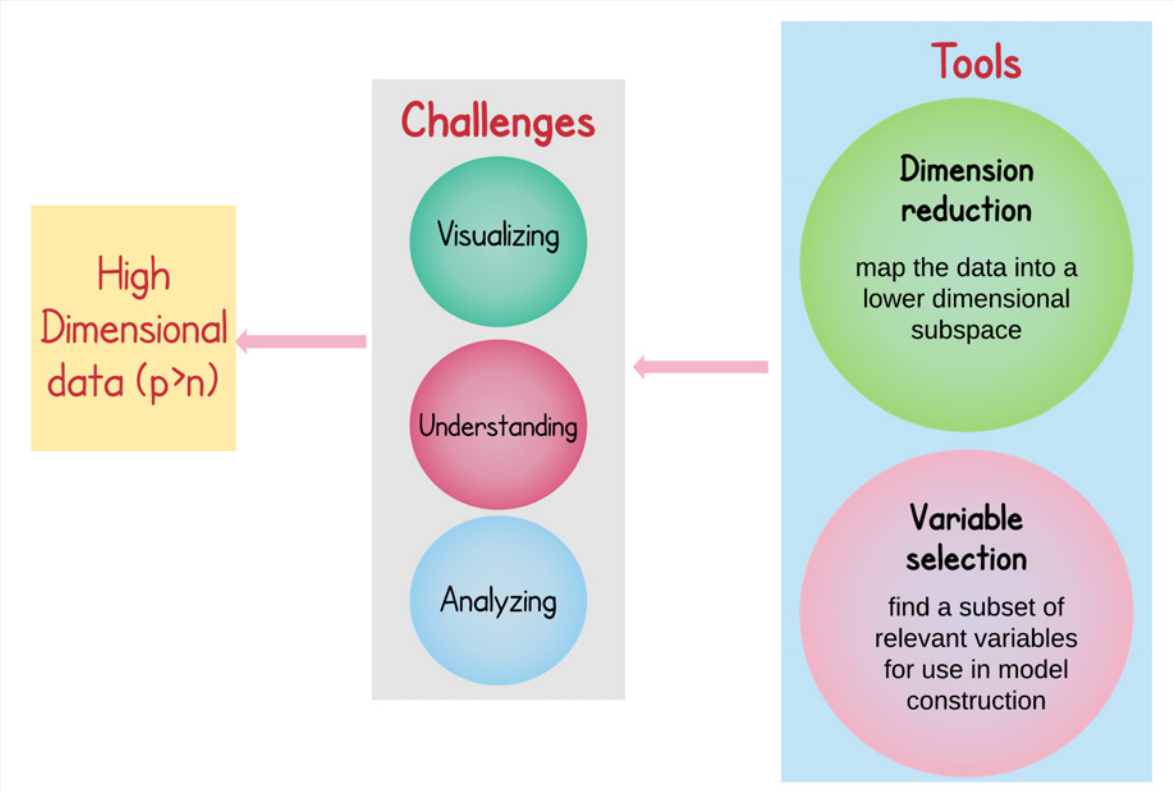


Interpretable Machine Learning Methods

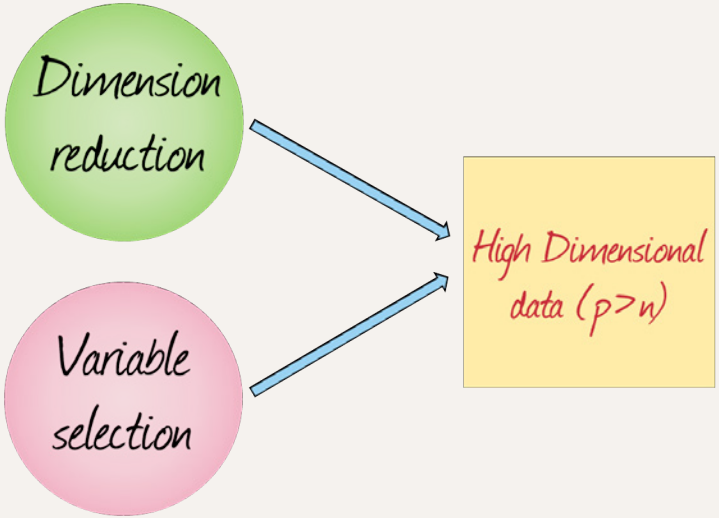


Health Analytics

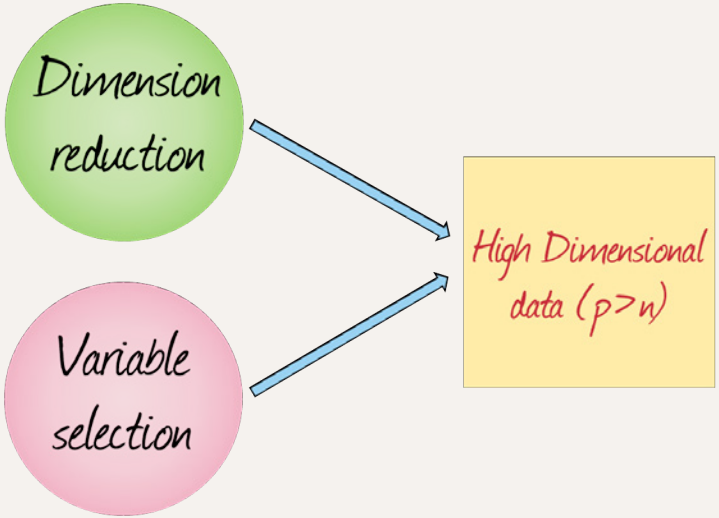
High Dimensional Data Analysis



High Dimensional Data Analysis



High Dimensional Data Analysis



Biometrika (2016), **103**, 3, *pp.* 579–593
Printed in Great Britain
doi: 10.1093/biomet/asw036

Sparse envelope model: efficient estimation and response variable selection in multivariate linear regression

BY Z. SU, G. ZHU

The Annals of Statistics
2020, Vol. 48, No. 1, 161–182
https://doi.org/10.1214/18-AOS1796
© Institute of Mathematical Statistics, 2020

ENVELOPE-BASED SPARSE PARTIAL LEAST SQUARES

BY GUANGYU ZHU[1] AND ZHIHUA SU[2]

[1]*Department of Computer Science and*
[2]*Department of Statist*

Electronic Journal of Statistics
Vol. 15 (2021) 6191–6227
ISSN: 1935-7524
https://doi.org/10.1214/21-EJS1943

Empirical likelihood ratio test on quantiles under a density ratio model

, Guangyu Zhu[2], and Jiahua Chen[1]

epartment of Statistics
rsity of British Columbia
ver, BC, Canada, V6T 1Z4
ng@stat.ubc.ca; jhchen@stat.ubc.ca

f Computer Science and Statistics
iversity of Rhode Island
ngston, RI, USA 02881
ail: guangyuzhu@uri.edu

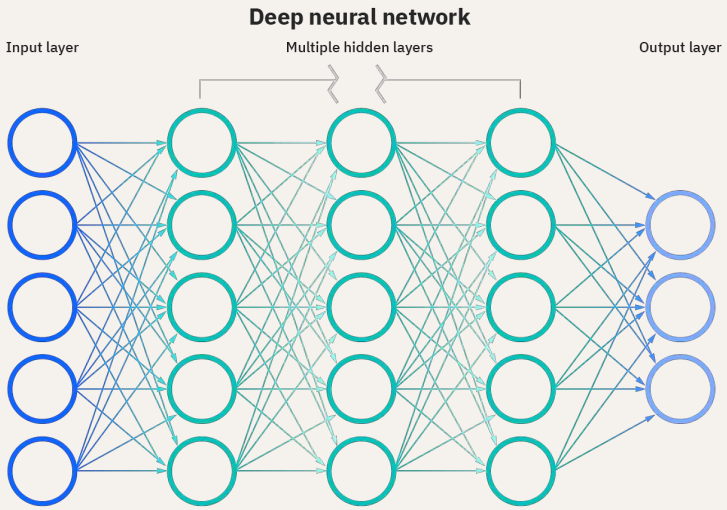
Statistica Sinica **31** (2020), 79-105
doi:https://doi.org/10.5705/ss.202018.0060

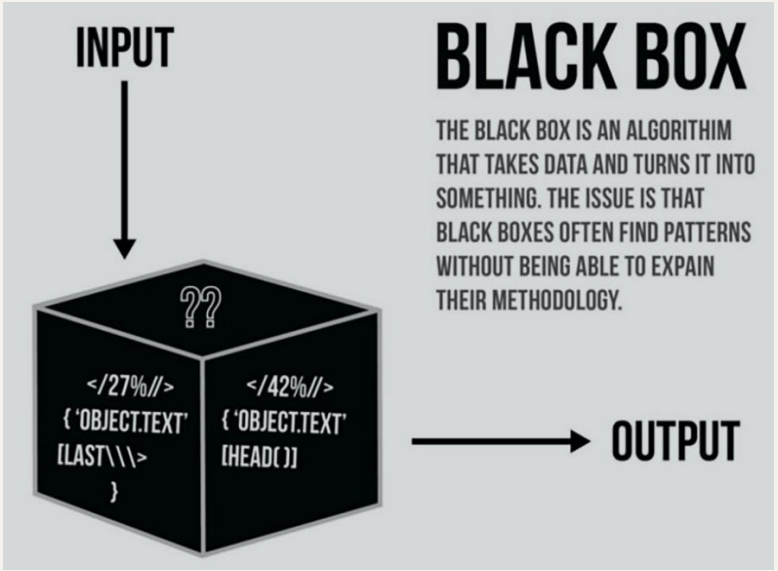
ENVELOPE QUANTILE REGRESSION

Shanshan Ding, Zhihua Su, Guangyu Zhu and Lan Wang

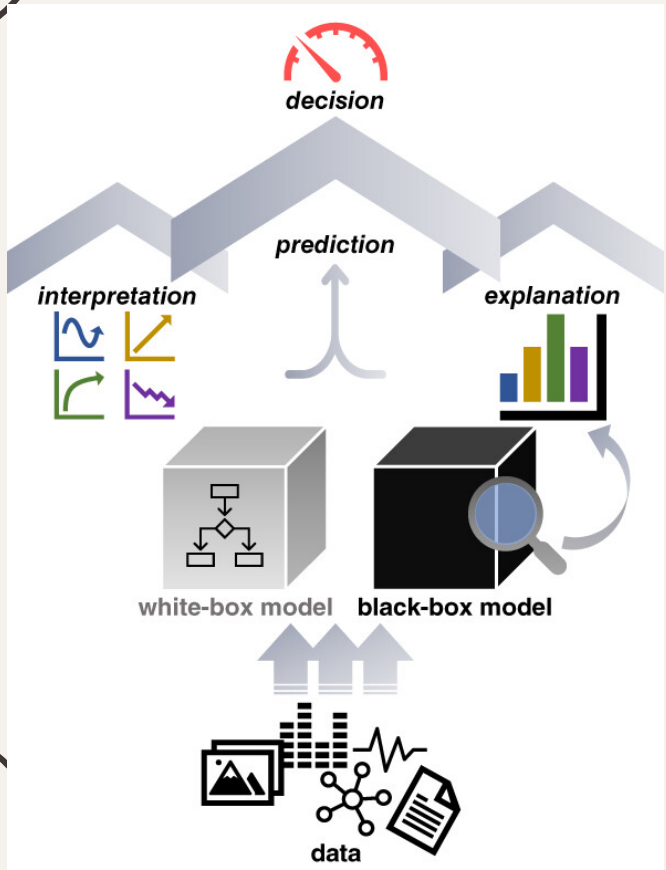
University of Delaware, University of Florida,
University of Rhode Island and University of Minnesota

Interpretable Machine Learning Methods

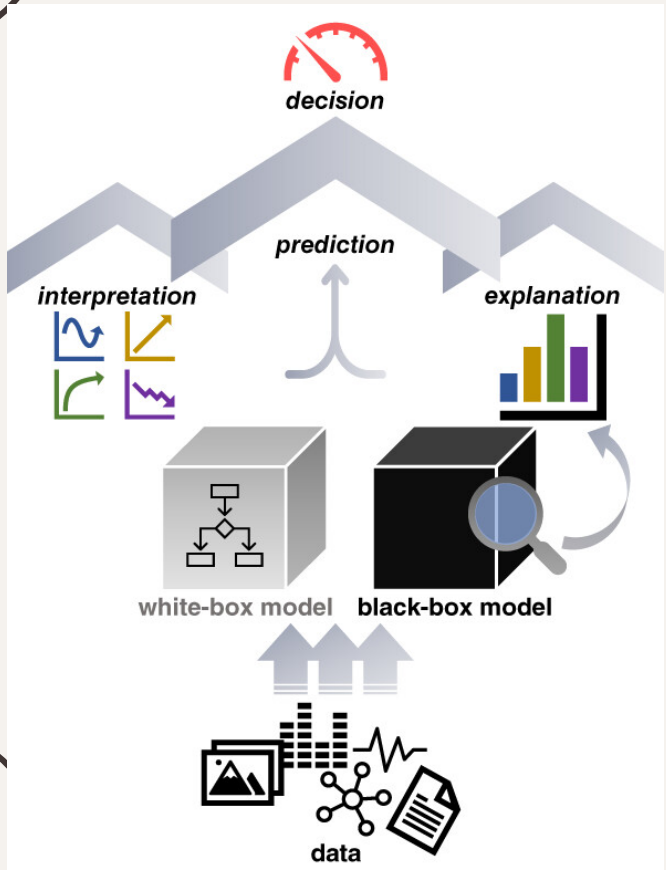




Interpretable Machine Learning Methods



Interpretable Machine Learning Methods



Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Deep-gKnock: Nonlinear group-feature selection with deep neural networks

Guangyu Zhu [a,*], Tingting Zhao [b]

[a] Department of Computer Science and Statistics, University of Rhode Island, United States of America
[b] Department of Electrical and Computer Engineering, Northeastern University, United States of America

OXFORD

Briefings in Bioinformatics, 2023, 1–13
https://doi.org/10.1093/bib/bbad084
Problem Solving Protocol

Identification of significant gene expression changes in multiple perturbation experiments using knockoffs

Tingting Zhao, Guangyu Zhu, Harsh Vardhan Dubey and Patrick Flaherty
Corresponding authors: Tingting Zhao, 1150 Douglas Pike, Smithfield, RI 02917, Phone/Fax: 401-232-6000 (US). E-mail: tzhao@bryant.edu; Patrick Flaherty, 710 N. Pleasant St., Amherst, MA 01003, Phone/Fax: 413-545-1796 (US). E-mail: pflaherty@umass.edu

Health Analytics

Visiting Scientists at Boston Children's Hospital
Until every child is well™
HARVARD MEDICAL SCHOOL

Health Analytics

Visiting Scientists at Boston Children's Hospital — Until every child is well™ — HARVARD MEDICAL SCHOOL

Child Neurology Open
Volume 10: 1-9
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2329048X231203817
journals.sagepub.com/home/cno
Sage

The RITA-T (Rapid Interactive Screening Test for Autism in Toddlers) Community Model to Improve Access and Early Identification of Autism in Young Children

Roula Choueiri, MD[1], William T. Garrison, PhD[2], Valerie Tokatli, BA[3], Naaz Daneshvar, BA[2], Jillian Belgrad, BA[2], Guangyu Zhu, PhD[3,4], and Bo Zhang, PhD[3,5]

Neurol Ther
https://doi.org/10.1007/s40120-022-00382-4

STUDY PROTOCOL

Interrater and Intrarater Agreement of Epileptic Encephalopathy Among Electroencephalographers for Children with Infantile Spasms Using the Burden of Amplitudes and Epileptiform Discharges (BASED) EEG Grading Scale: Study Design and Statistical Considerations

Xinting Liu · Jian Chen · Lin Wan · Zhichao Li · Yan Liang · Huimin Yan · Guangyu Zhu · Bo Zhang · Guang Yang

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 05 October 2022
DOI 10.3389/fmed.2022.913361

Check for updates

OPEN ACCESS

EDITED BY
Redhwan Ahmed Al-Naggar,
National University of Malaysia, Malaysia

REVIEWED BY
Rui Guo,
Capital Medical University, China
Guangqiang Wang,
University of Shanghai for Science and Technology, China

*CORRESPONDENCE
Bo Zhang

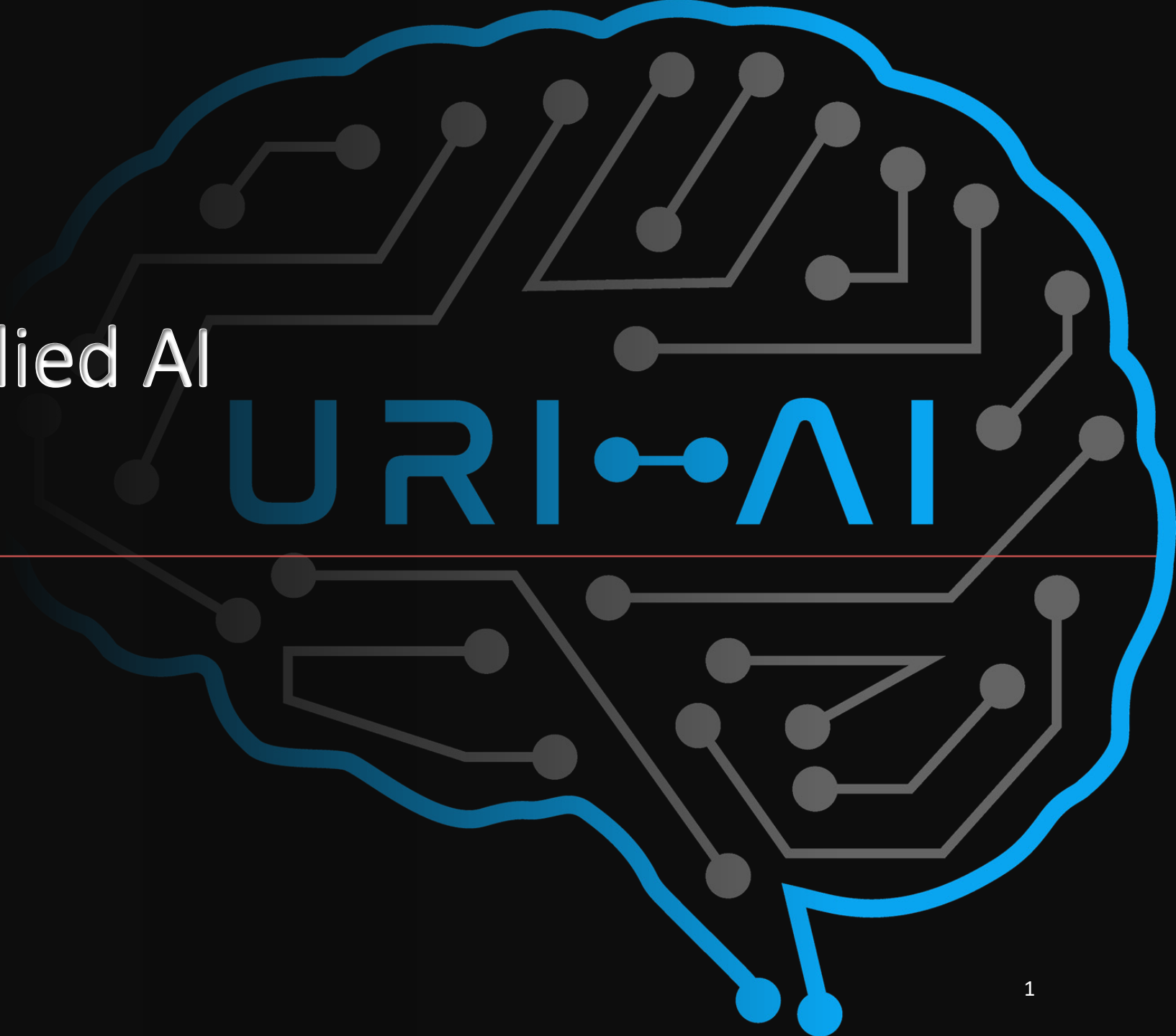
Changes in rate and socioeconomic inequality of cervical cancer screening in northeastern China from 2013 to 2018

Yaqian Liu[1], Jing Guo[1], Guangyu Zhu[2], Bo Zhang[3]* and Xing Lin Feng[1]*

Personal Website

https://guangyuzhu.rbind.io/

Big Data and Applied AI



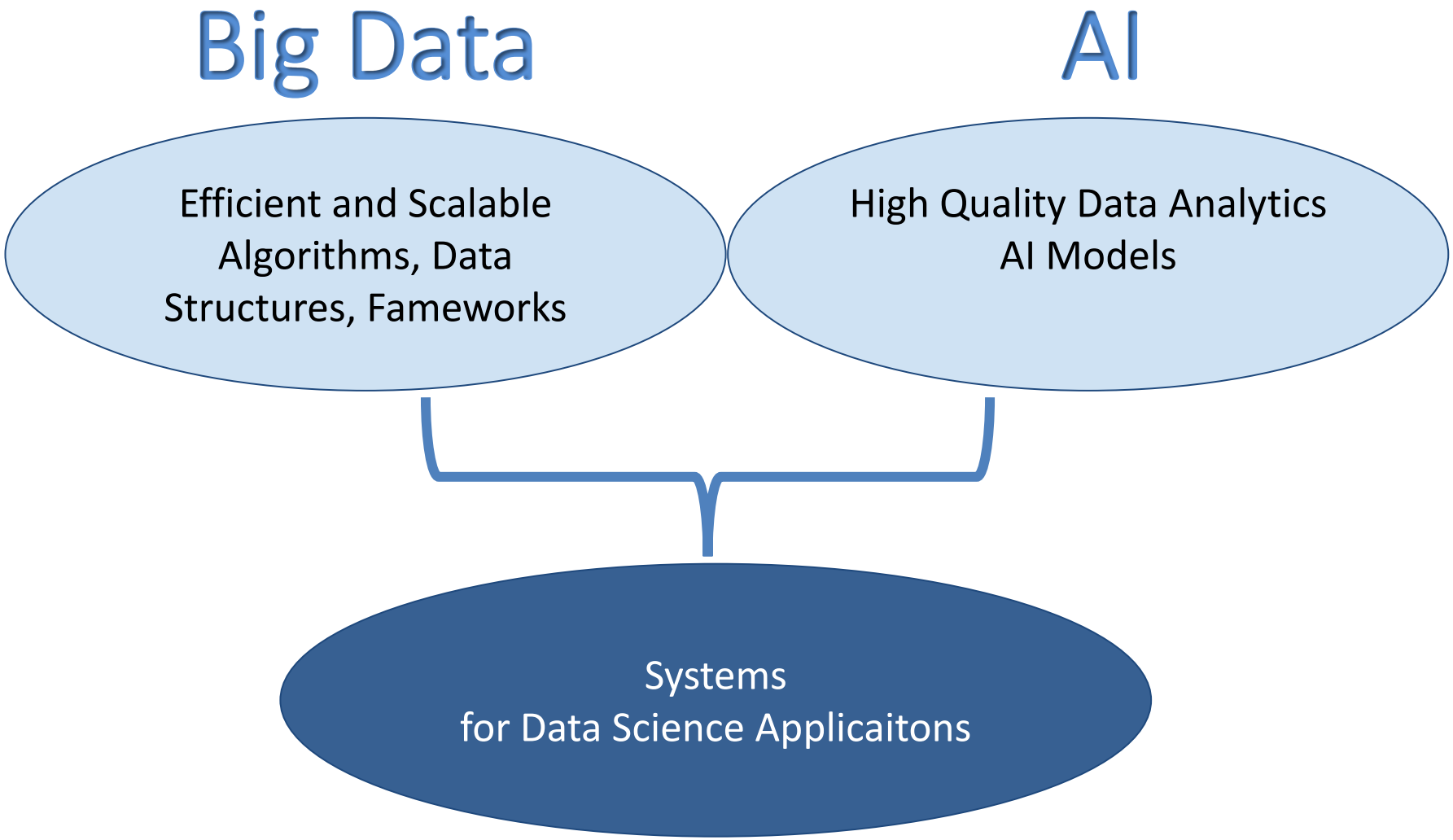
Abdeltawab Hendawi

University of Rhode Island, USA

Co-Director of the AI-Lab at URI

hendawi@uri.edu

Research Contributions



Sample Projects

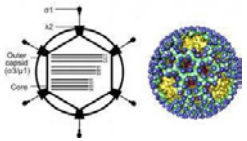
Real World Problems

- **Smart Cities**
  - **Drivers Behavior** — Umass, MIT, Michelin Co.

  - **Satellite Imagery** — UW, Microsoft Seattle

- **Smart Health**
  - **Covid-19 Contact Tracing** — International Collaborators

  - **Reovirus Detection** — UMN, SDSU



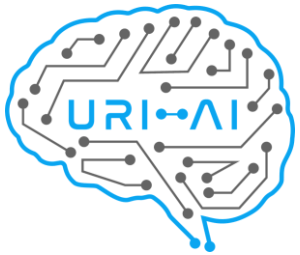
Students and Collaborators

Daniel Bernard

Brennan Rivera

Gavin Hamel

- UW
- UVA
- UMN
- SDSU
- MIT
- UMASS
- Microsoft
- Michelin Co.



THANK YOU !

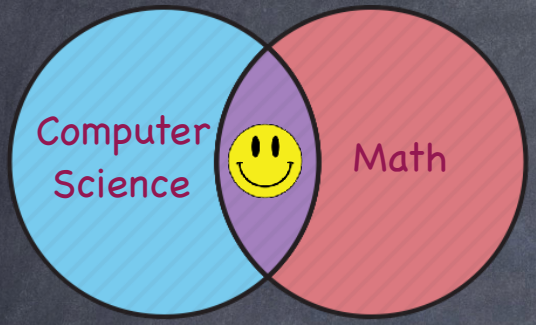
Jean-Yves Hervé

- Modeling and Simulation
  - → Serious Gaming (insidious gaming)
  - → Interactive multimedia for art exhibits/museums
- VR Immersion
- Classical Computer Vision (no ML)
  - → Augmented Reality
  - → Visual navigation (hand-eye coordination)
  - → Functional space representation
  -

Interposition for AR

Compliant Interposition

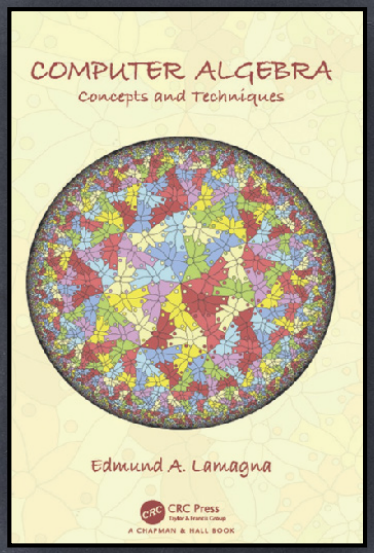
Ed Lamagna



eal@cs.uri.edu

Design & Analysis of Algorithms

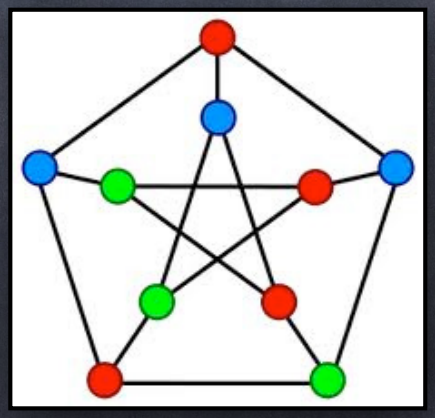
- computer algebra (CAS)
  summation, recurrences
- sorting
- algebra
- graphs/networks
- experimental algorithmics



Elections

- security and auditing
- redistricting/gerrymander

Educational Tools



- cryptography toolkit
- Graph Playground

Puzzles & Games

- algorithmic thinking
- games on graphs
- math card tricks

URI Center for Computational Research (CCR)
Research Computing, ITS, AI lab

For URI researchers:

- HPC support through local clusters and the MGHPCC (Mass. Green High Performance Computing Center); quantum computing support through partnership with IBM
- Expert support and funding for ***interdisciplinary data enabled*** research – all disciplines – all colleges.
- Speed networking and lightning talk events
- Featured speaker series and Fireside Chats - AI, Ethics, Team formation …
- Collaborative with URI AI Lab, and the AI and Business Analytics group in the College of Business

Check us out: https://its.uri.edu/center-for-computational-research/

Few words about my research

Gavino Puggioni, Ph.D.

Dept. of Computer Science and Statistics, and CELS

Current Research

- Innovation in Bayesian modeling, novel stochastic processes
- Space and time data analysis in different areas of applications: ecology and the environment, medical data, spread of infectious diseases, finance
- Hybrid stochastic machine learning models

I would like to supervise…

- One Ph.D. student with the following skills and interests:

1. Strong mathematical and statistical background
2. Proficient in coding (R or Python)
3. An interest in learning new material in different scientific areas
4. Comfortable reading scientific articles
5. Academic career oriented

Accessible & Socially Aware Technologies Lab

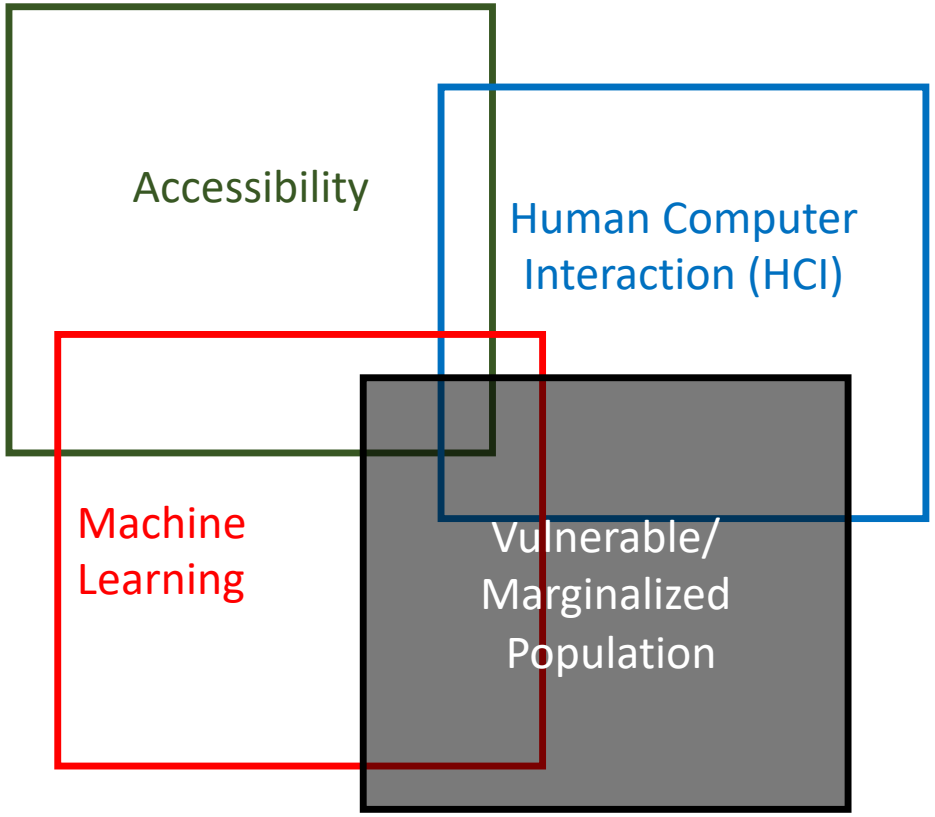
A S S E T

The University of Rhode Island

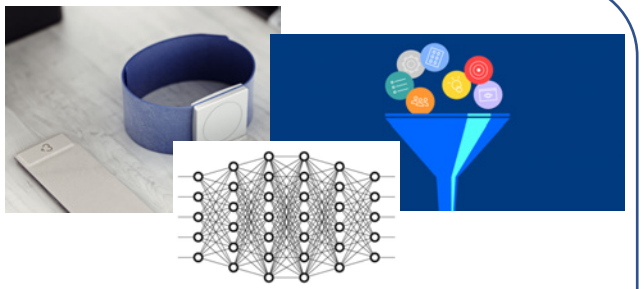
Building **accessible virtual reality (VR)** for **people with** intellectual and developmental disabilities **(I/DD)**

Developing **tools** for enabling **reporting for people with** intellectual & developmental disabilities **(I/DD)**

Developing co-designed **self care tools for people with** intellectual & developmental disabilities **(I/DD)**



Contact Info:
Prof. Krishna Venkatasubramanian, Director, ASSET Group
krish@uri.edu



Data curation and modeling for biosensor-based **medication adherence detection for opioid use disorders**

Digital phenotyping to ensure prophylaxis medication adherence for men to have sex with men with addiction disorders

Designing better **end-user-focused digital tools** for **medievalists** humanities **scholars**

Delivering Personalization Across Tasks and Disciplines

Adobe

HAXLab

THE READABILITY CONSORTIUM

Google

Creating a Personalized Behavior Platform

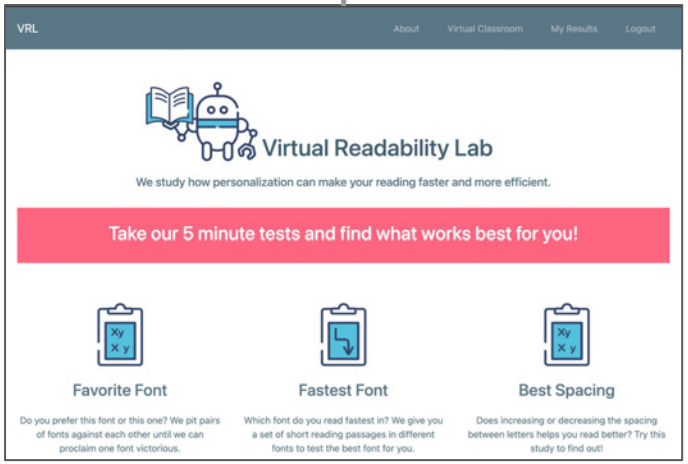
Custom low code environment.

Develop personalized readability tests.

Integrate and develop AI and LLMs.

Introduce activities beyond readability.

Collaborate with K-12 and specific populations.

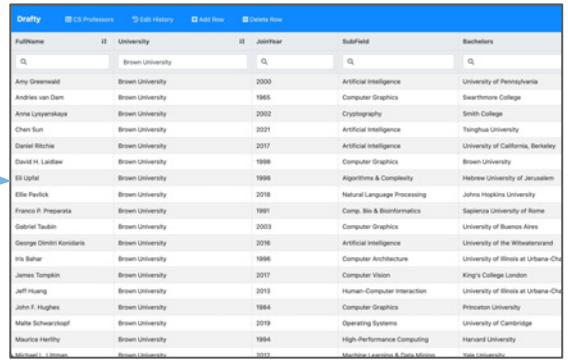


Want to help develop this research in Spring 24? **CSC 592 HCI Research Seminar,** or join the lab :)

Want to help develop this research in Spring 24?
CSC 592 HCI Research Seminar, or join the lab :)

HAXLab

Integrating Intelligent Information Systems with AI Tools and Diverse Populations



+

WIKIDATA

OpenAI
ChatGPT

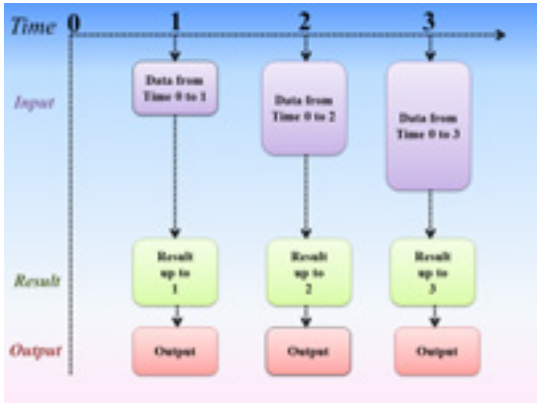
How do we personalize human-information interactions in people's everyday lives?

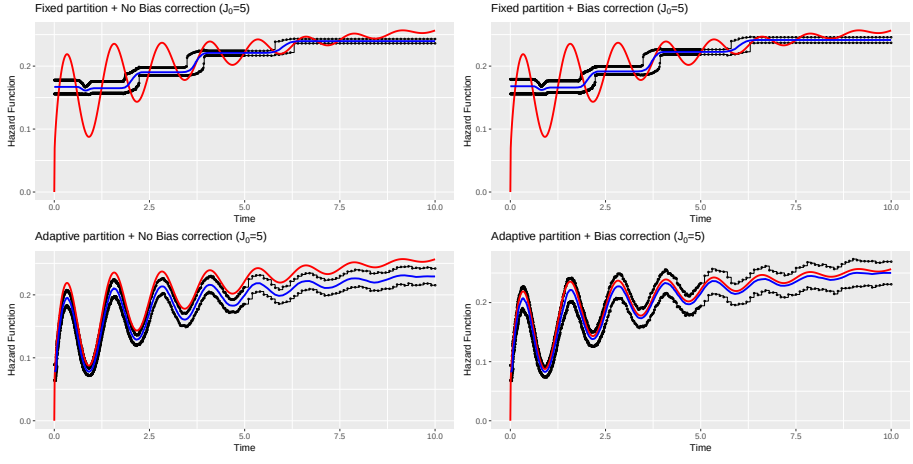
`crowdsourcing + peer production`

Research Interests

Jing Wu

- Missing Data & Bayesian
  - Clinical Trials
  - Survival Analysis
  - Education
  - Social Network
- Big Data
  - Modeling: LM, GEE, GLM, Survival
  - Hypothesis Testing
  - Variable Selection and Regularization
- Interdisciplinary Research
  - Prostate Cancer Research
  - Radiosensitizing Nanoparticles
  - Neurobehavioral Development of Preterm Infants





Statistical tool manufacturer

Mark (Haihan) Yu

October 25, 2023

Manufacture statistical tools

- Invent new statistical tools for challenging data problems, such as
  - construct new estimator or test statistics;
  - provide new inference procedures;
  - ...
- Improve an existing statistical tool from multiple perspectives, including:
  - extend its range of applicability;
  - investigate its mechanism under unconventional settings;
  - ...

More specifically

I have been working on

- Nonparametric inference tools $\Rightarrow$ make inference based on data not on model
  - Bootstrap, resampling, subsampling, empirical likelihood
- Dependent data analysis $\Rightarrow$ time series
  - spectral analysis, change-point problem
- Open to all interesting problems...

My goal is to develop **user-friendly** and **theoretically compact** statistical tools.

Thank you!

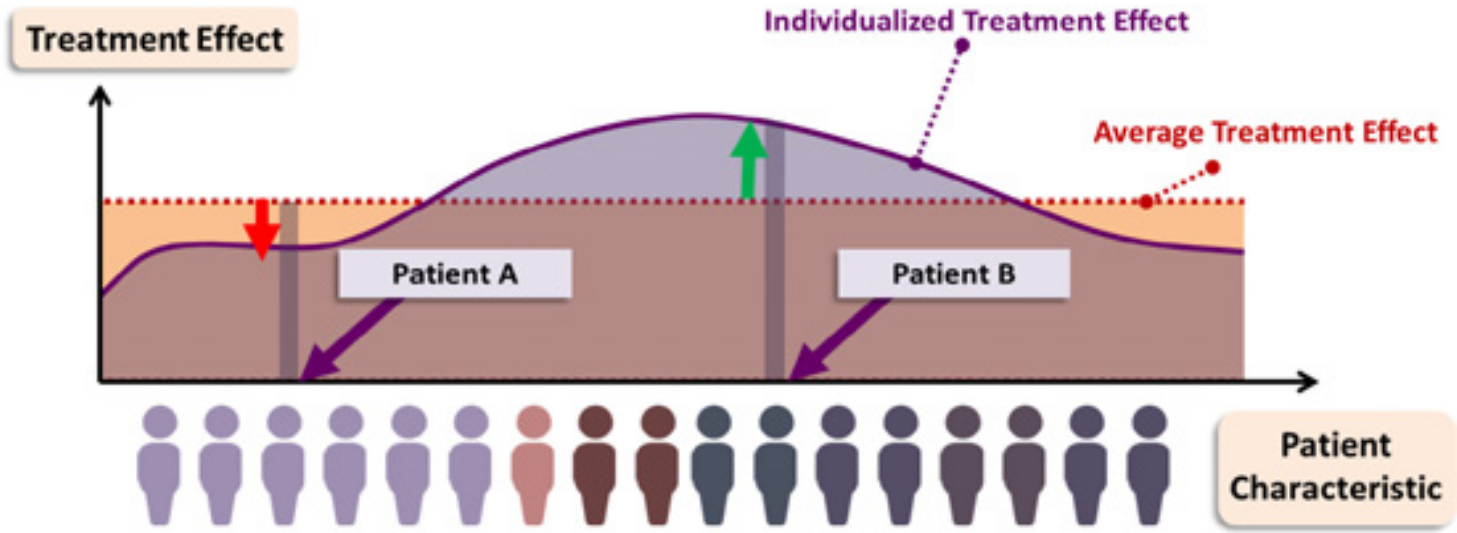
A Brief Overview of My Research

Yichi Zhang
yichizhang@uri.edu

October 27, 2023

Precision Medicine

- Individualized treatment rules: the right treatment for the right patient at the right time.
- Reinforcement learning:
  - Unlike supervised learning, we don't observe the optimal treatment for each patient.
  - Unlike unsupervised learning, we do observe some outcome related to the treatment.



Source: Bica, I., Alaa, A. M., Lambert, C., & Van Der Schaar, M. (2021). From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. Clinical Pharmacology & Therapeutics, 109(1), 87–100.

Precision Medicine

- Longitudinal data: Inference for long-term effects.
- Inference for the cutoff value: suppose an aggressive treatment is recommended for younger patients while a conservative treatment is recommended for elder patients. The estimated cutoff value doesn't follow a normal distribution.
- Design of experiment: can we do better than a randomized study?
- Sample splitting: can we make inference about the treatment effects when some parts of the model uses maching learning techniques?
- Interpretability: can we approximate a rule by a tree or a decision list?