

AIM-AHEAD







Georgetown University Innovation Center for Biomedical Informatics

PUBLIC HEALTH

CENTER FOR HEALTH EQUITY



Data Science Training Core (DSTC) Resources

Dr. Legand (Lee) Burge, PhD Dr. Nawar Shara, PhD Dr. Guodong Gao, PhD & Dr. Ritu Agarwal, PhD Dr. Peter McGarvey, PhD



DSTC Purpose

The Data Science Training Core (DSTC) will **identify training needs and gaps** as well as **identify or develop training and workforce development resources** to support AI/ML and health disparities research.

The DSTC will assess training requests for AIM-AHEAD participants and connect requestors with required training and training support via the <u>Help Desk</u>.



The DSTC Roadmap



Virtual Applied Data Science Training Institute, VADSTI: Fall 2022 Training Series

Description:

• To extend the existing Virtual Applied Data Science Training Institute (VADSTI) to engage the next generation of investigators in using data science methodology, and AI/ML to support minority health disparities research and address health inequities.

Objectives

- Continue to provide training in the foundations of data science, advance analytic skills and introduce tools for clinical and genomic research with application to minority health and health disparity research;
- Demonstrate understanding of the foundations of data science;
- Demonstrate knowledge of analytic techniques utilized in biomedical, behavioral and health disparities research;
- Perform data analysis and interpret results from output with health disparities datasets; Understand the concepts of data partitioning and the practice behind supervised and unsupervised learning;
- Understand the elements and use of EHR; Be introduced to tools for applied data science using cloud based platforms.
- Learn from experts on current research topics in data science including application to social and environmental justice.

Lead Investigator: Dr. John Kwagyan, Howard University College of Medicine.

Alignment with North Stars: North Star 1 (Develop a diverse, equitable, and inclusive AI/ML workforce)

North Star 3 (Use AI/ML to address disparities and minority health)

Completion Date: 10/28/2022



[•]

Academic EHR (SimChart) Undergraduate Training (Nursing)

Description:

• This project will integrate academic EHR (SimChart) and data analytic skills and enhance SimChart activities across the DON clinical courses.

Objectives

- To infuse the SimChart across the undergraduate curriculum and increase undergraduate nursing students' interest and retention in research and data analytics for the 2022/2023 cohort using an approved HU IRB.
- To obtain and analyze data in the DON clinical curriculum to enable answer pertinent research questions that would assist in addressing health disparities among the underserved and underprivileged population using SEDoH variables.

Outcome:

 Abstract titled, "<u>Students' Perspective on Academic EHR Usability:</u> <u>SimChart Enhanced</u>" to be presented at the upcoming NYU-HU summit (March 29, 2023).

Lead Investigator: Priscilla Okunji, Ph.D., RN-BC, Howard University College of Nursing and Health Sciences



Academic EHR Introduced to Juniors in Health Assessment Class



Data Science for Wearable Health Monitoring Devices

Description:

This 13-week hybrid course will allow in- person and remote students to gain real world experience in the design of wearable devices that implement machine learning methods for disease monitoring. The course will focus specifically on cardiovascular diseases, the leading cause of death globally. We will tackle racial bias in medical device technology (pulse oximetry) and methods to eliminate bias.

Objectives

- The course will provide an **overview of data science concepts** and engineering design principles required to develop mobile health devices for real-time vital sign monitoring.
- To obtain **data on in-patient cardiometabolic**, clean and analyze such data to answer pertinent research questions that would assist in **addressing health disparities** among the underserved and underprivileged population **using SEDoH** variables.
- The hybrid course is designed as an "interactive learning community" to enhance interaction between students and classmates, the instructor, course material, world-renowned researchers, and technology.

Lead Investigator: Dr. Valencia Koomson, Tufts University



Probability and Statistics with R

Description:

- Intermediate to advance level course.
- Provide comprehensive probability and statistical models, and R statistical tutorials that implement the learned statistical concepts for data analysis.

Objectives

- Statistical learning using a statistical package R
- Use clinical data sets
- Data visualization methods
- T-tests and paired t-tests to compare two groups;
- The analysis of variance (ANOVA) to quantify differences among multiple groups
- Linear regression to investigate the association between two variables and the ability of one variable to predict another

Lead Investigator: Dr. Joon Ha, Howard University



Introduction to R for Data Science

Description:

• This course serves as an introductory course to machine learning principles with an application to cardiovascular diseases and oncology. An emphasis will be placed on how the social determinants of health impact these disease states. As a part of this course, students will learn to prepare and utilize datasets to apply machine learning-based quantitative analysis. Although students will learn about essential theoretical concepts crucial to machine learning, the emphasis will be placed on applied skills as relates to applying various machine learning algorithms and models to datasets.

Objectives

- Understand how to properly prepare datasets for use in computing workspaces.
- Be able to describe the principal models used in machine learning/data science and the benefits and limitations of these approaches
- Gain hands-on experience using data gathered from a national clinical informatics dataset to evaluate the role to which social determinants of health and/or precision medicine contribute to disparities in cancer and cardiovascular-related prevalence and treatment outcomes.

Lead Investigator: Dr. La'Marcus Wingate and Dr. Salome Weaver, Howard University



Omics: Big Data in Biology

Description:

• A data science course to expose underrepresented minority undergraduate and graduate college students to omics data analysis.

Objectives

- Students are expected to learn the genomics, epigenomics, transcriptomics, proteomics and metabolomics
- Understand the concepts of bioprogramming and machine learning.
- Understand the roles of sequence variants in human diversity and diseases.
- Explain the RNAseq for gene expression and gene regulatory network.
- Discuss the omics results in a biological and biomedical context.
- Relate the omics analysis to human diseases and health disparities.
- Apply SWB for programming and machine learning analysis
- Write the small Python scripts for bioinformatics analysis.
- Use R to perform statistical computing and to produce simple graphs.
- Perform basic next generation sequencing data analysis using GATK, IGV, Galaxy.
- Learn how to use biological databases such as NCBI, PDB and UCSC Genome Browser.
- Analyze the OCHIN's RDW and All of Us Research Hub datasets.

Lead Investigator: Dr. Shaolei Teng, Howard University



Open Data Science & AI/ML



Open Data Science & Al/ML

Welcome to the AIM-AHEAD Open Data Science & AI/ML Course

OPEN DATA SCIENCE & AI/ML About this Course This is an asynchronous, self-directed course for beginnerand intermediate-level learners. The main presentations and Jupyter Notebooks in this course were adopted from IBM's OpenDS4AII repository. ...

About this Course

This is an asynchronous, self-directed course for beginner- and intermediate-level learners. The main presentations and Jupyter Notebooks in this course were adopted from <u>IBM's</u> <u>OpenDS4All repository</u>. OpenDS4All is an open-source data science curriculum that provides educators with access to the tools and materials needed to establish data science courses at institutions across the world.

The perspective of the materials largely comes from computer science, with an emphasis on data wrangling and engineering as well as machine learning and validation. Prior versions of the content have been used to teach students ranging from freshmen to PhD students, across a wide range of fields. The emphasis is largely on core concepts and algorithms with grounding in today's technologies and best practices.

Course Structure

This course contains six modules (units):

- 1. Introduction & Data Science Ethics
- 2. Building AI Trust
- 3. Intermediate Data Science
- 4. Unsupervised Machine Learning
- 5. Supervised Machine Learning
- 6. Data Science and Healthcare Use Cases

This course curated for PRIME trainees is available for all AIM-AHEAD members.



AI for Health Equity Symposium 2022



AIHES2022 Workshop Series Modules

- The Basics pt I and pt II
- Ethics in AI/ML
- All of Us Data Browser
- Implementing AI/ML in health equity applications in low-resources settings
- Conducting a project in AI/ML for health equity from start to finish in low-resource institutions
- AI/ML Knowledge and Communication for Leadership in Healthcare
- ML Methods with Healthcare Data
- Tools for Deep Learning Exploring and Understanding DL Use Cases in Biomedicine & Beyond
- Cutting-Edge AI/ML Applications
- Cutting-Edge AI/ML Applications
- Career Development

AI for Health Equity Symposium 4-week Virtual Workshop Series https://www.youtube.com/watch?v=Rm6fcziKGdg&list=PLMdVHrgRTn1XKzUnRTMnkGxnFggkeBNI



AIM-AHEAD Professional Development Program (AA-PDP)

2 Thematic: Healthcare Management & AI Management

9-week virtual training opportunity – Recorded and available on AIM-AHEAD Connect

Module 1: Healthcare Equity & Fairness

- Prioritizing treatment & care based on need
- Representative data

Algorithmic approaches Al-Induced biases

Module 2: Healthcare Finance & Revenue Cycle

- The healthcare ecosystem
- Major data types & sources

- Google Colab and Python
- Fundamentals of AI projects

Module 3: Healthcare Operations & Patient Engagement

- Care delivery process
- Care quality & capacity

Module 4: Health IT & Informatics

- Major health IT systems
- Interoperability & security

Module 5: Leading AI Transformation

- Leadership Essentials
- Manages Changes & Risks



- Clinical decision support
- Deep learning
- Cloud Computing
- Nurture AI culture
- High performance AI teams









Al Essentials for Healthcare

AI Essentials for Healthcare

Objectives

- Understand the role of AI in addressing critical health disparities issues
- Demystify AI "buzzwords" like machine learning, deep learning, computer vision, and natural language processing (NLP)
- Identify the right AI tool for different challenges and opportunities
- Manage and launch AI projects in healthcare.

Course Structure

This course contains nine modules, each composed of 4-6 lessons. Each lesson contains a video, typically 10 minutes in length or shorter. At the end of each module, there is a quiz for self-assessment.



AI Essentials for Healthcare



Al Essentials for Healthcare

Welcome to AI Essentials for Healthcare AI Essentials for Healthcare is a self-directed course designed to support professional development on the application of AI in healthcare. Course Objectives Understand the

Enroll

Coming Soon to AIM-AHEAD Connect

Training Materials to be made available on AIM-AHEAD Connect:

Online Mini-Course of curriculum focused on Fundamental <u>AI/ML in Health Equity for multiple target audiences</u>, revised from a curriculum co-created with a HS student (From GUMC Gateway Exploration Program pilot for Washington, DC Public School HS scholars)

Al for Healthcare Applications training materials in the form of asynchronous lectures with Python notebooks for an interactive set of modules on AIM AHEAD Connect

Online workshops investigating Social Determinants of Health using SDoH indices and interactive **Tableau dashboards**.

