

Advancing Population and Personalized Healthcare through Big Data and Machine Learning

Conrad Tucker, Ph.D., MBA

Associate Professor of Engineering Design, Industrial and Manufacturing
Engineering at Penn State

NSF Center for Health Organization Transformation (CHOT)



www.chot.psu.edu

Current CHOT Industry Members



Research Inspiration

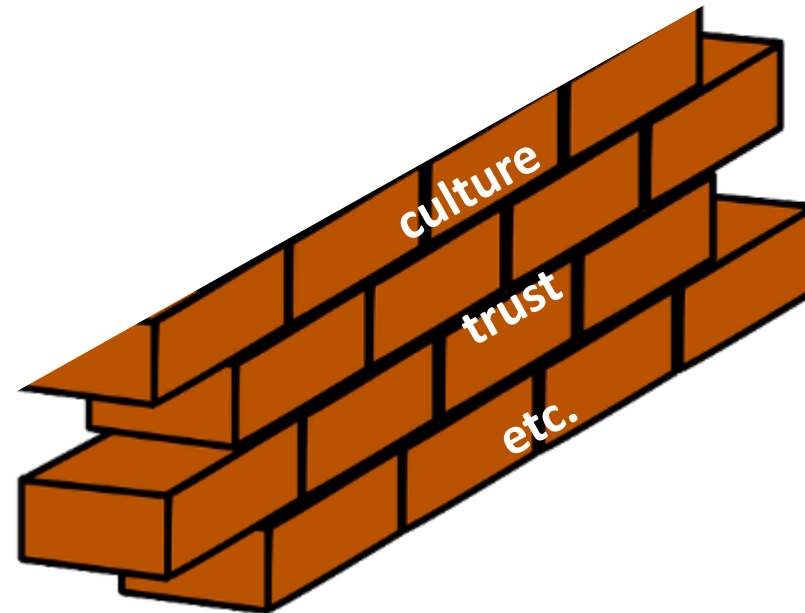
<https://www.youtube.com/watch?v=jfSNxVqprvM>

Current State of Healthcare?

Patient

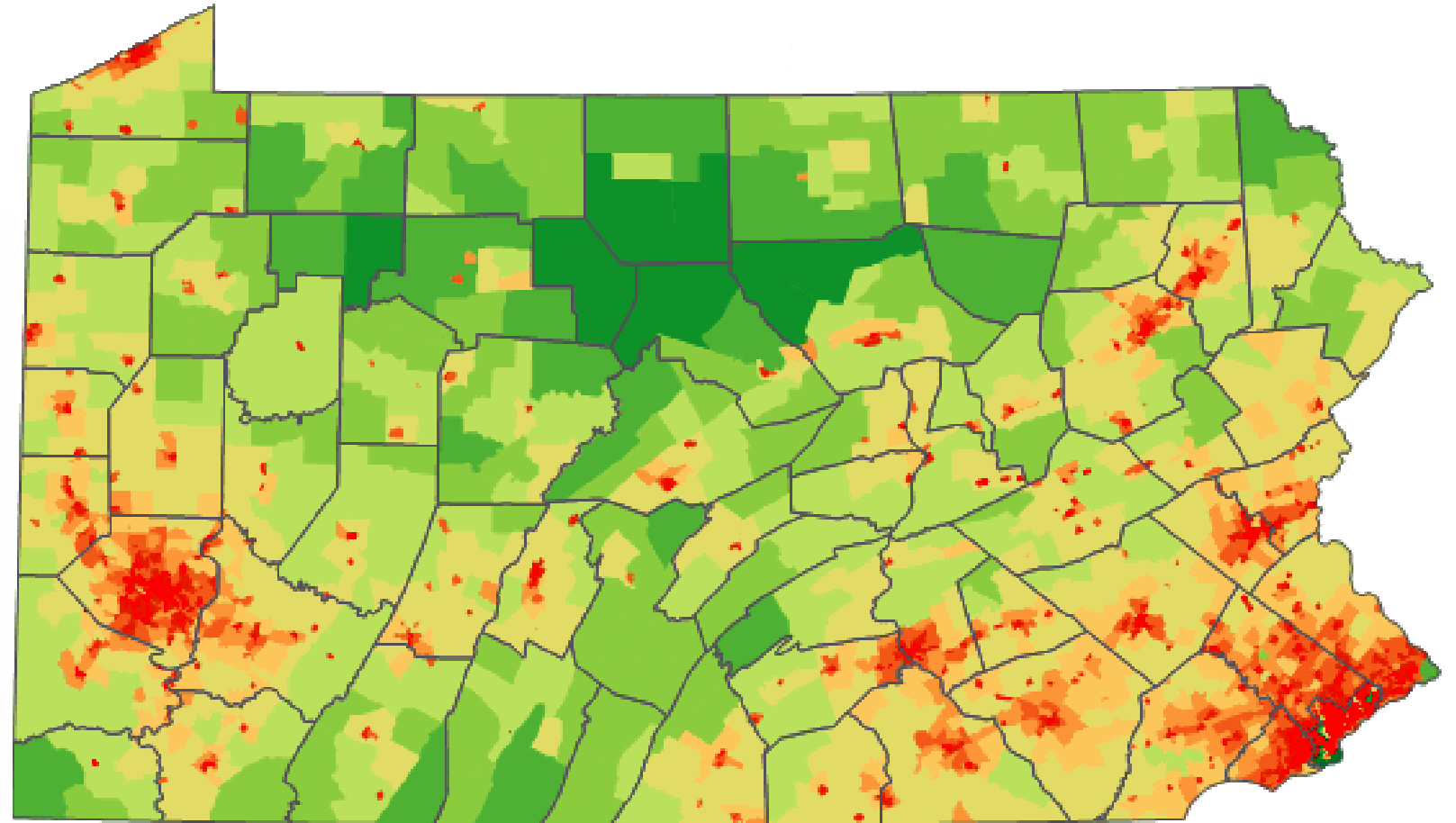
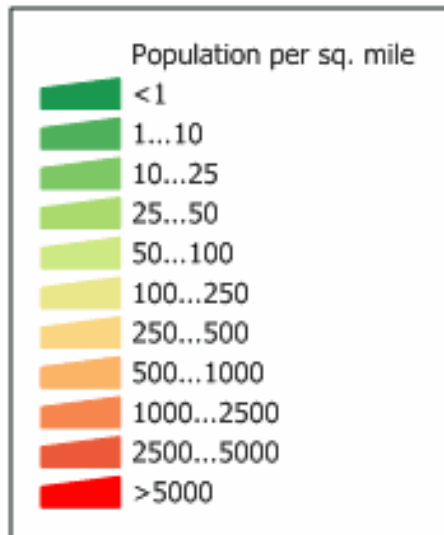


Physician



Population Distribution: State of Pennsylvania

Source: U. S. Census Bureau
Census 2000 Summary File 1
population by census tract.



Research Questions

- How can we learn patients' needs?
- How can we better connect patients with their healthcare providers?
- How can we make the healthcare sector more efficient?

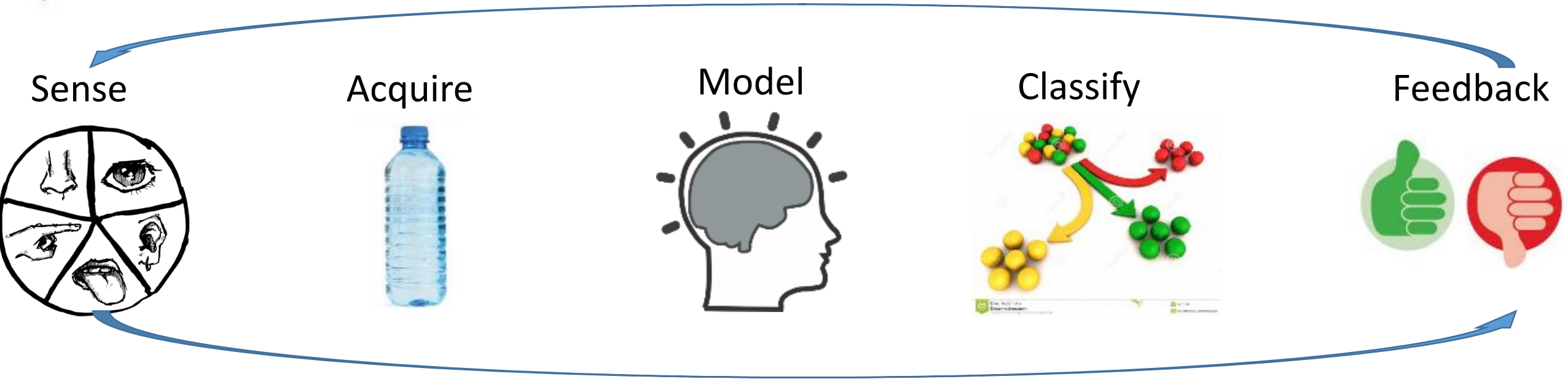
Research Questions

- How can we learn patients' needs?



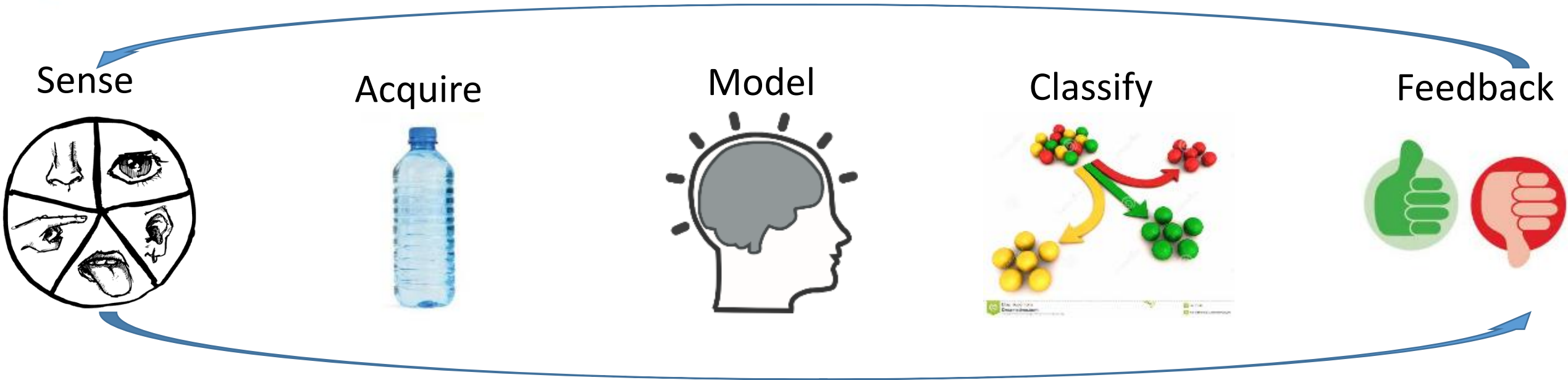
What does it mean “to learn”?

The Process of Learning

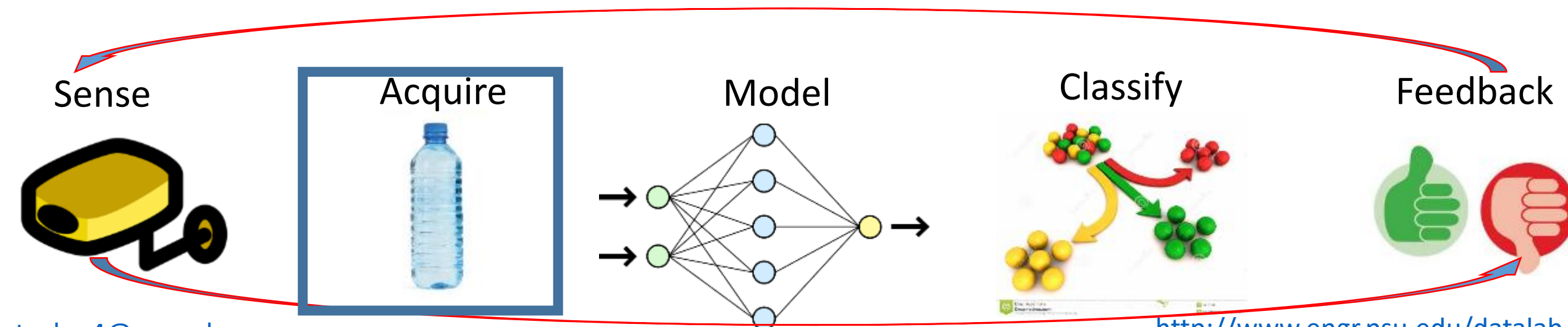


The Process of Learning

Humans



Machines



Machine Learning Through “Big Data”?

Big data “is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” -Gartner

Machine Learning Through “Big Data”?

Big data “is **high-volume**, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” -Gartner



Volume

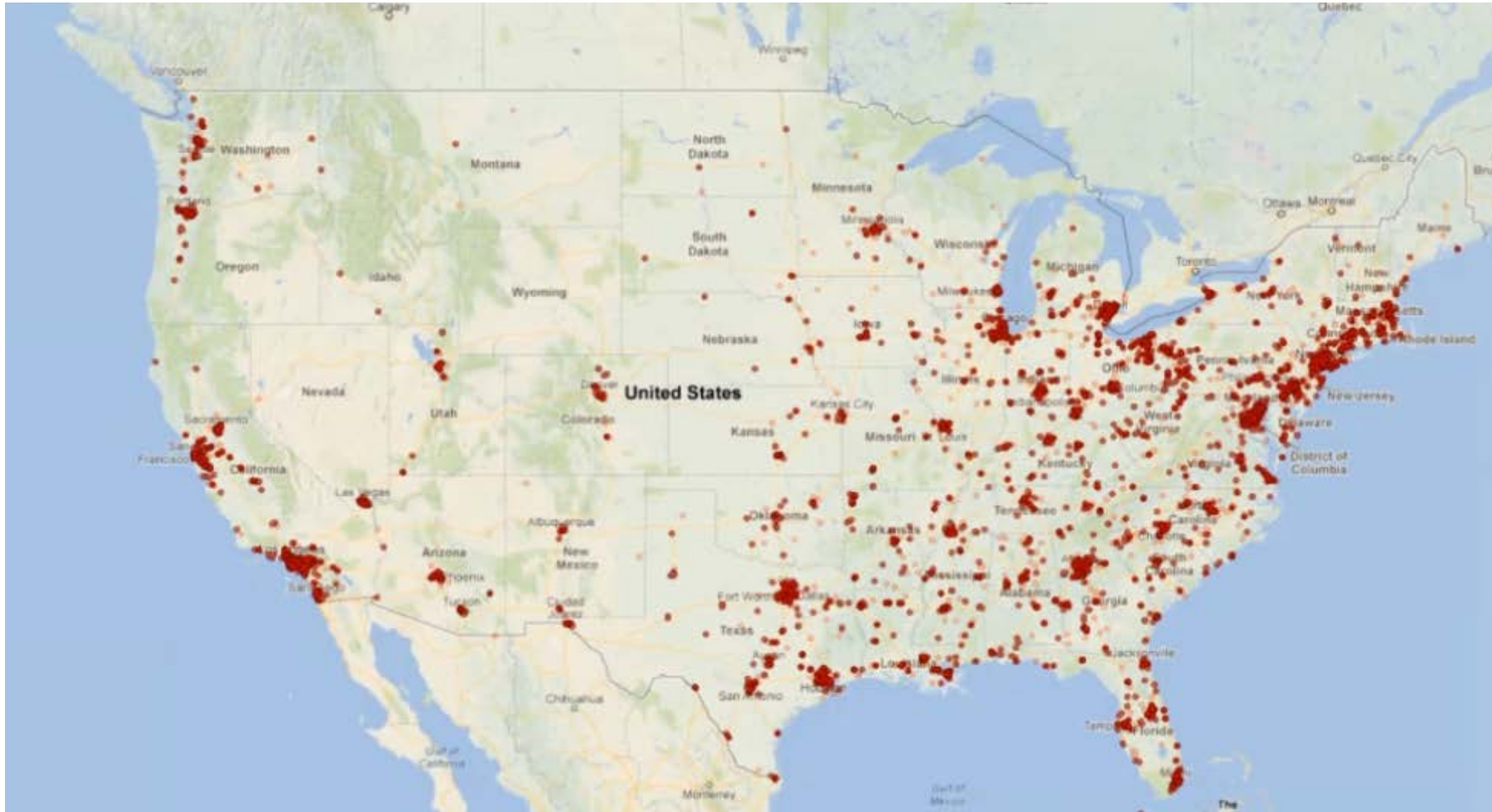
Unstructured (Textual) Data: Social media services such as Twitter[®] and Facebook[®] process more than **500 terabytes** of data each day.

A significant amount of that data is **unstructured textual** that can be used in engineering design to mine user preferences.

Machine Learning Through “Big Data”?

Big data “is high-volume, **high-velocity** and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” -Gartner

500 million tweets each day (one every 173 microseconds)



Velocity

Machine Learning Through “Big Data”?

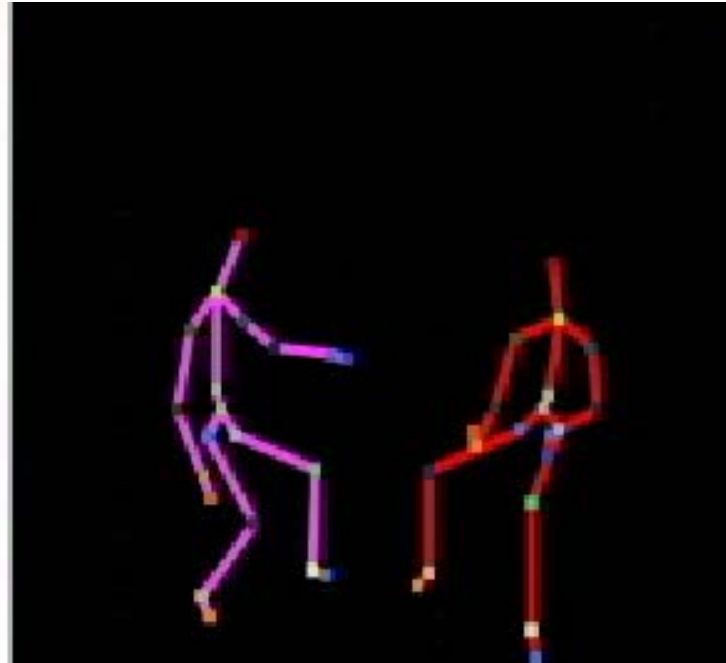
Big data “is high-volume, high-velocity and **high-variety** information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” -Gartner

Structured (Body Posture) Data

Video Data Stream



Skeletal Data Stream



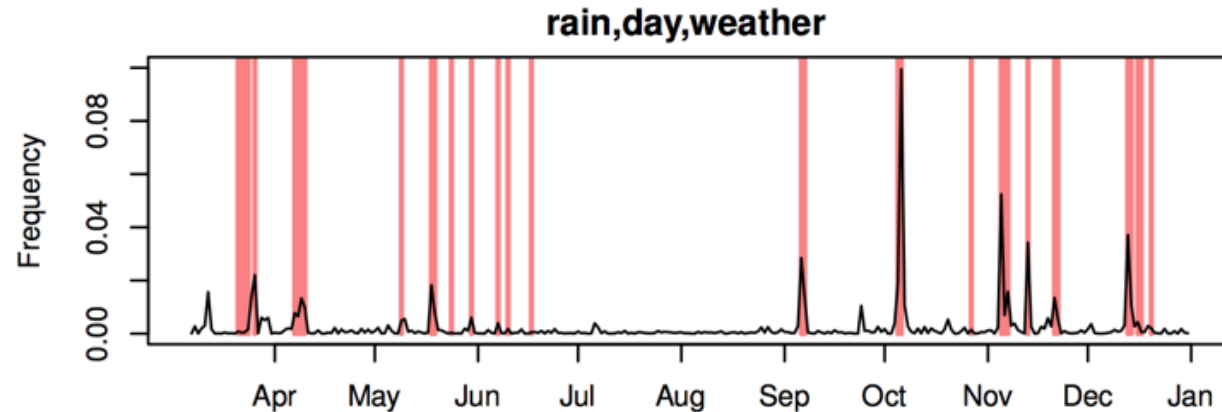
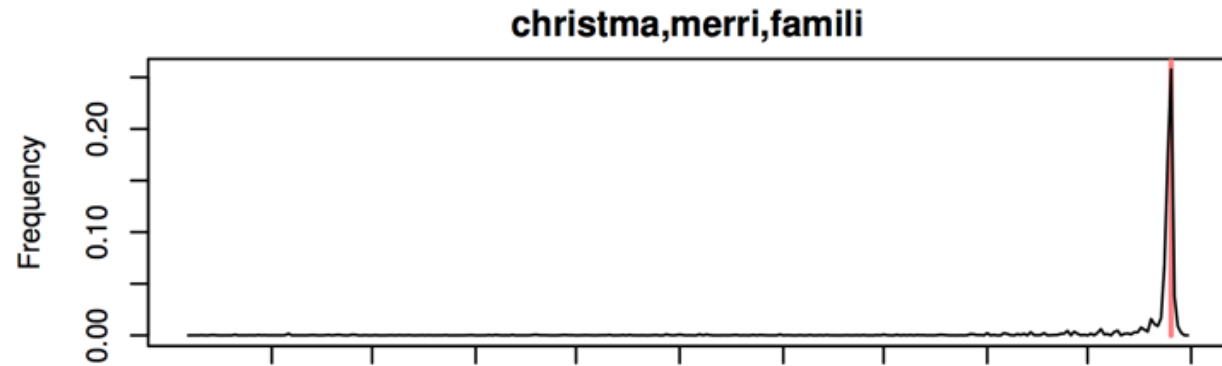
Depth Data Stream



Variety

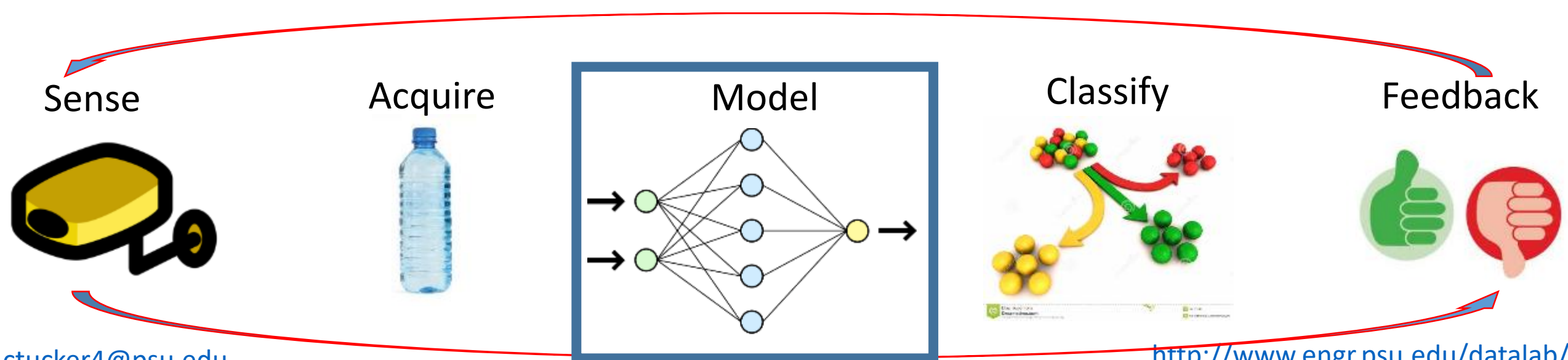
Machine Learning Through “Big Data”?

Big data “is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for **enhanced insight (veracity)** and decision making” - Gartner



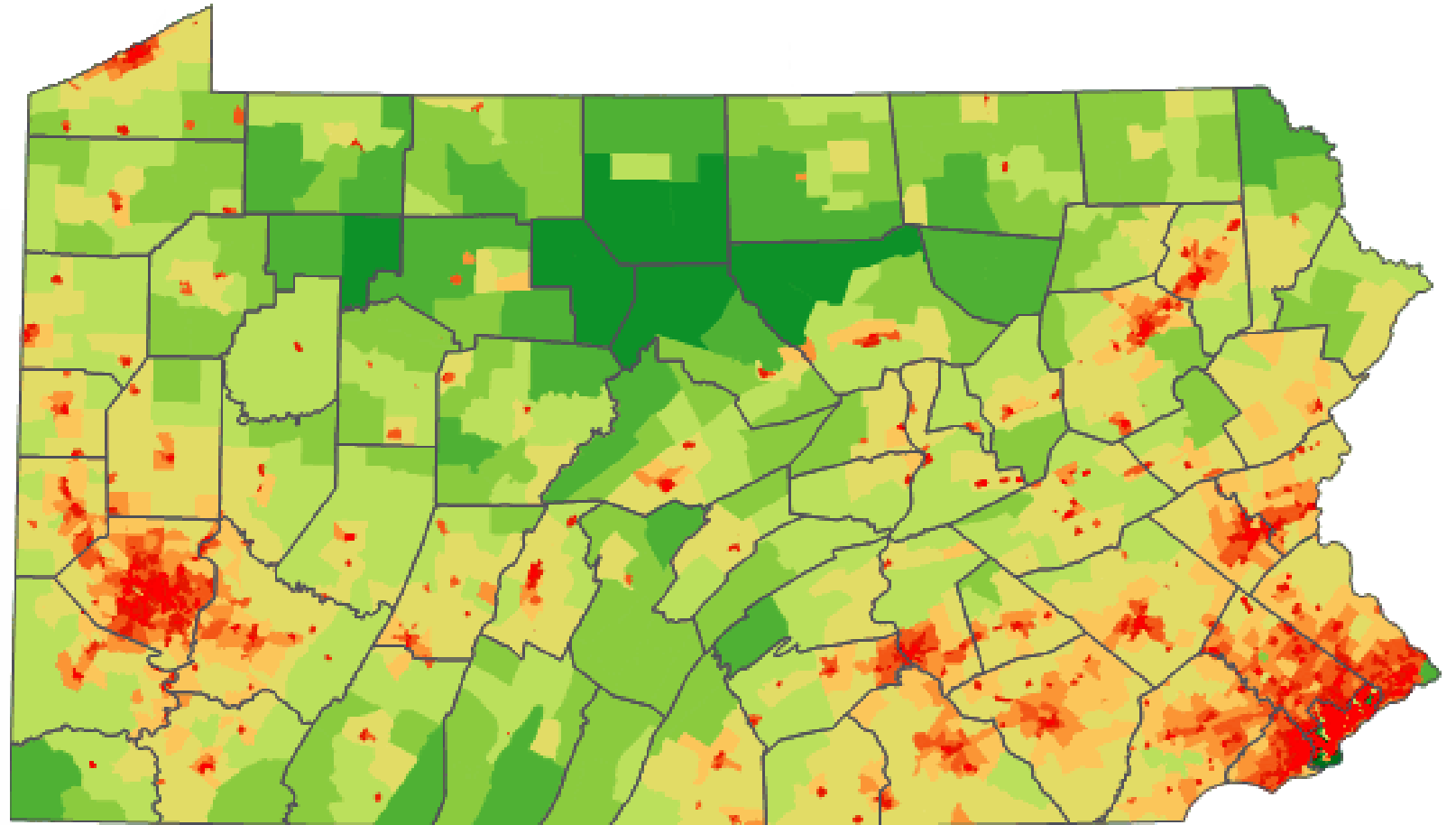
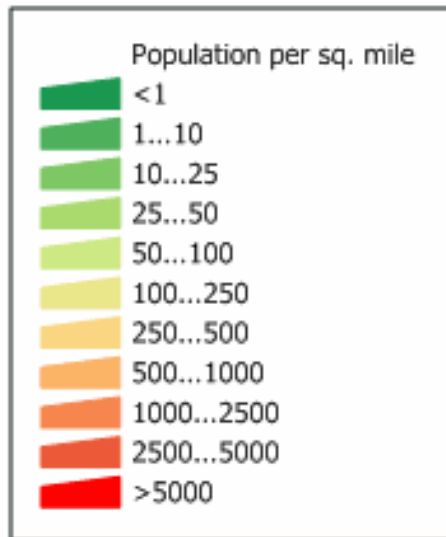
Veracity

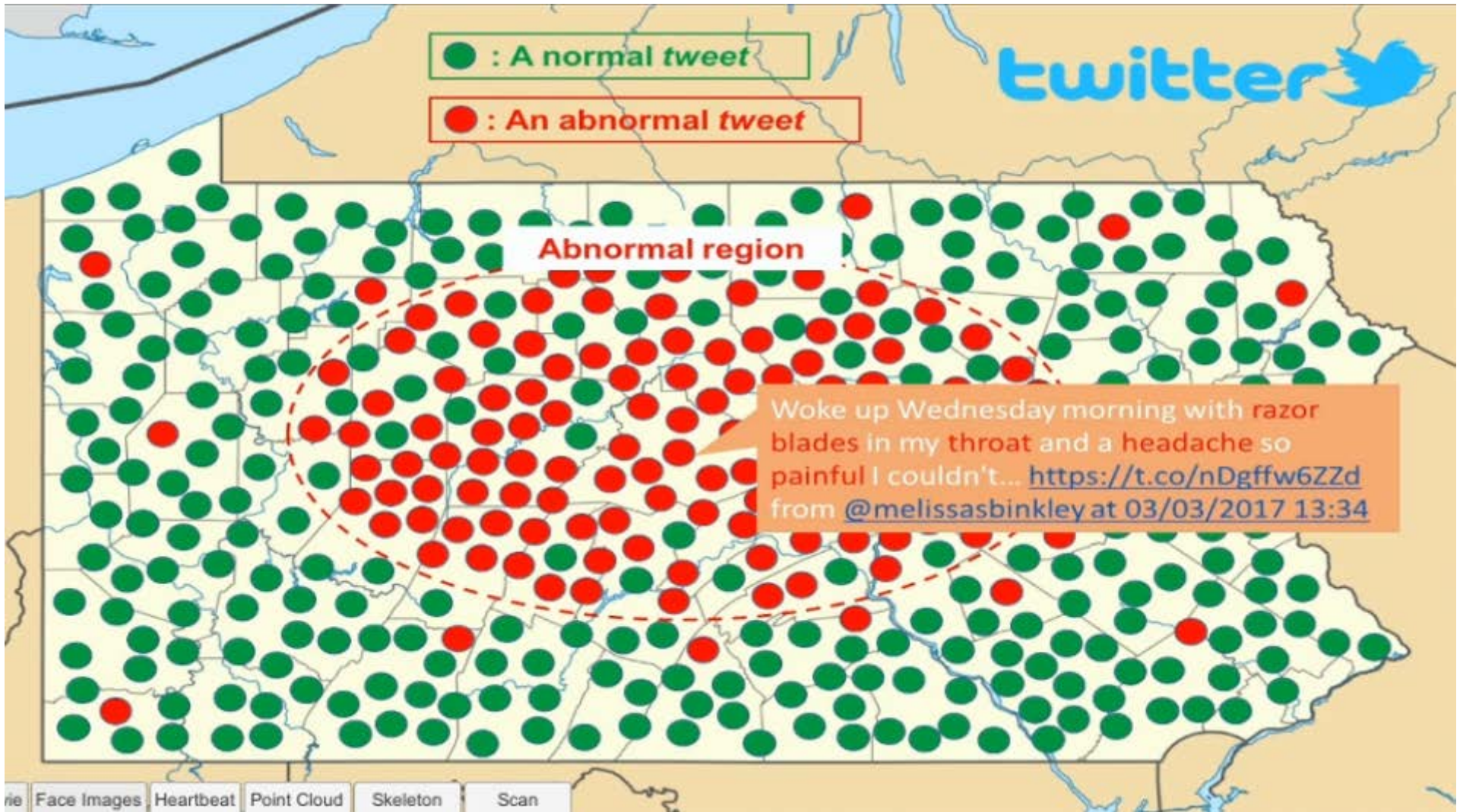
The Process of Machine Learning



Population Health: How do we *learn* what a patient population needs?

Source: U. S. Census Bureau
Census 2000 Summary File 1
population by census tract.





Research Questions

- How can we learn patients' needs?
- How can we better connect patients with their healthcare providers?
- How can we make the healthcare sector more efficient?

Healthcare Sensors and Sensing Systems

Sensors and Sensing Systems

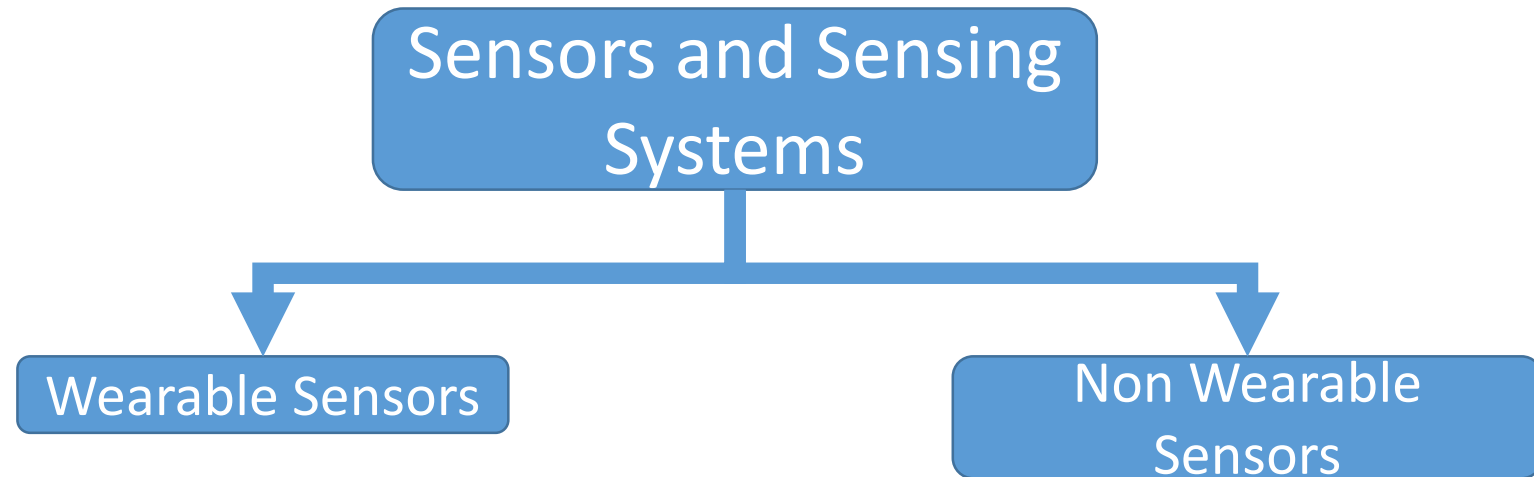
Wearable Sensors



Non Wearable Sensors



Healthcare Sensors and Sensing Systems



Capturing of Patients' Biometric Data

Heartbeat

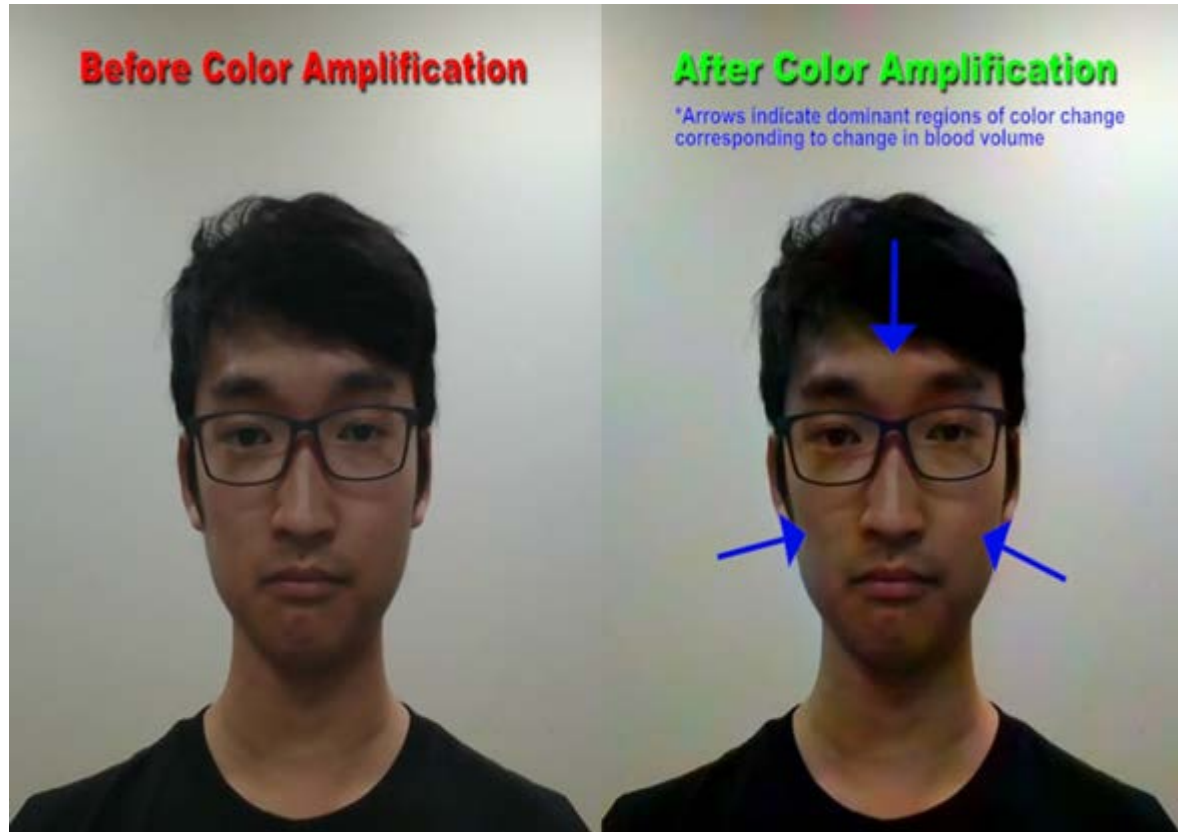


Pain Level



Remote Capturing of Patients' Biometric Data

Heartbeat

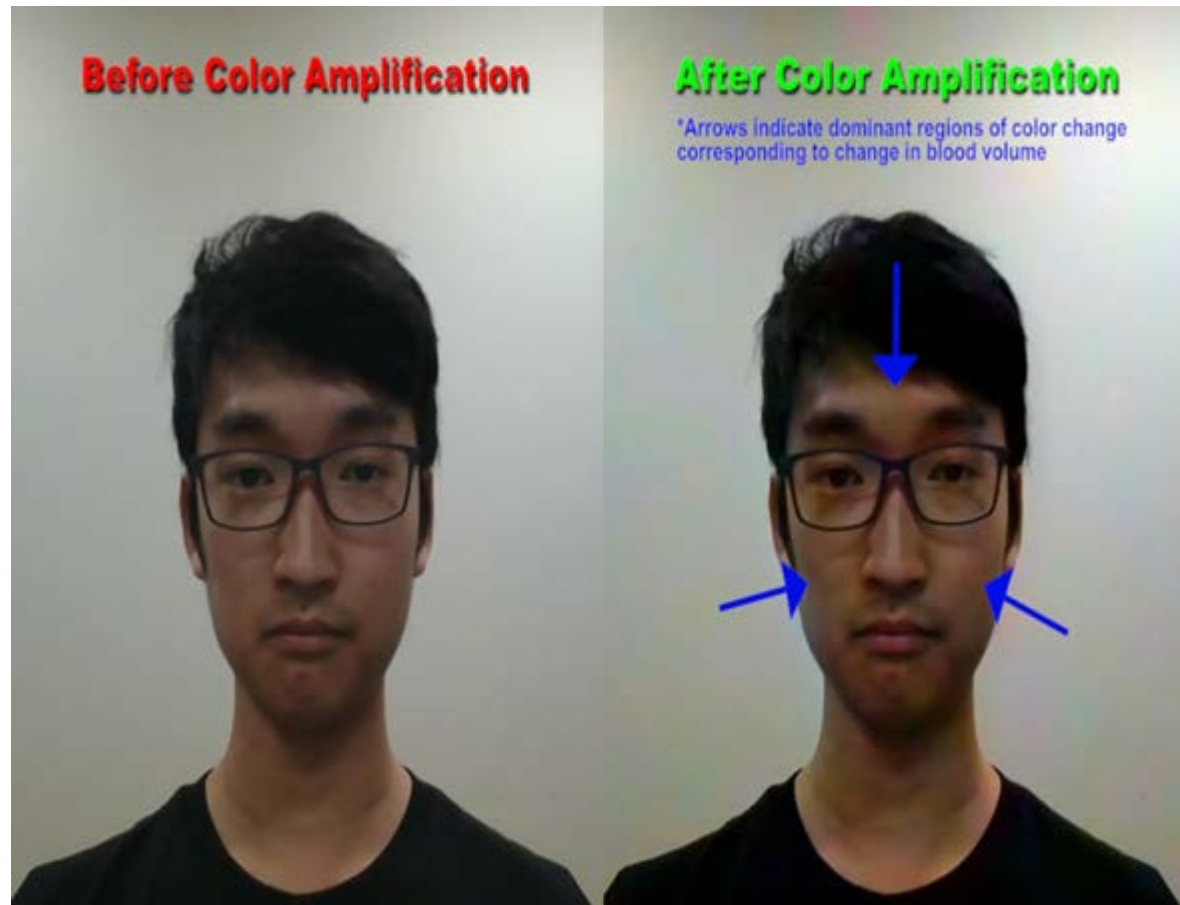


Pain Level

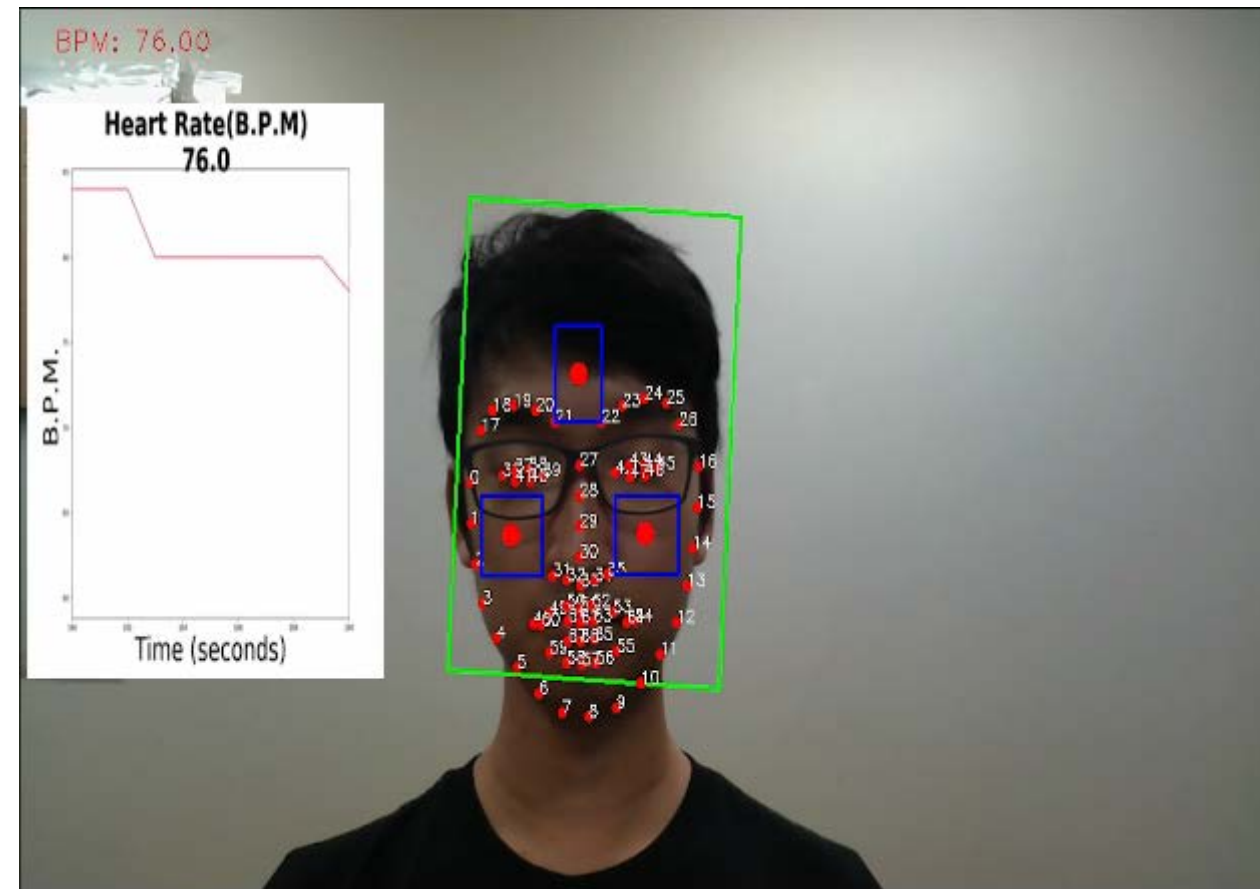


Remote Capturing of Patients' Biometric Data

Heartbeat

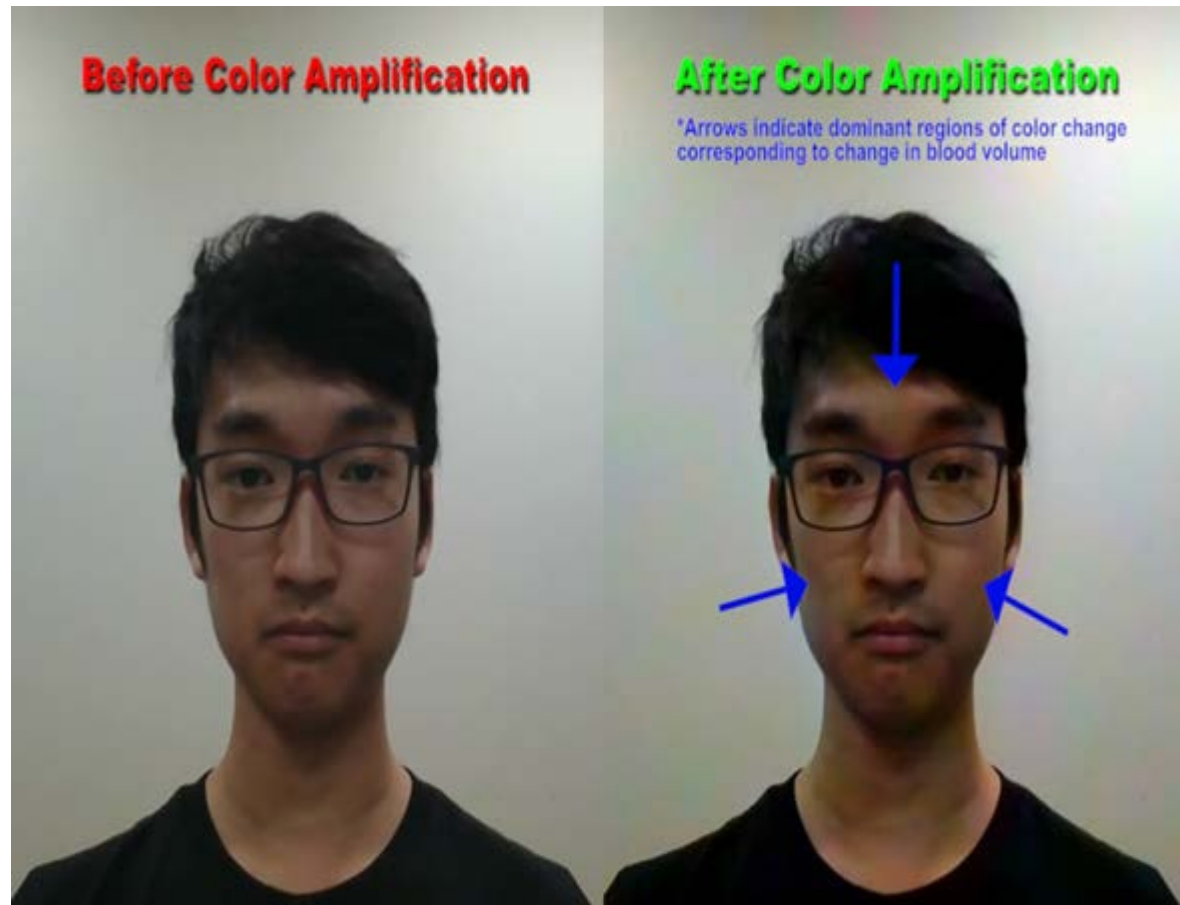


Pain Level

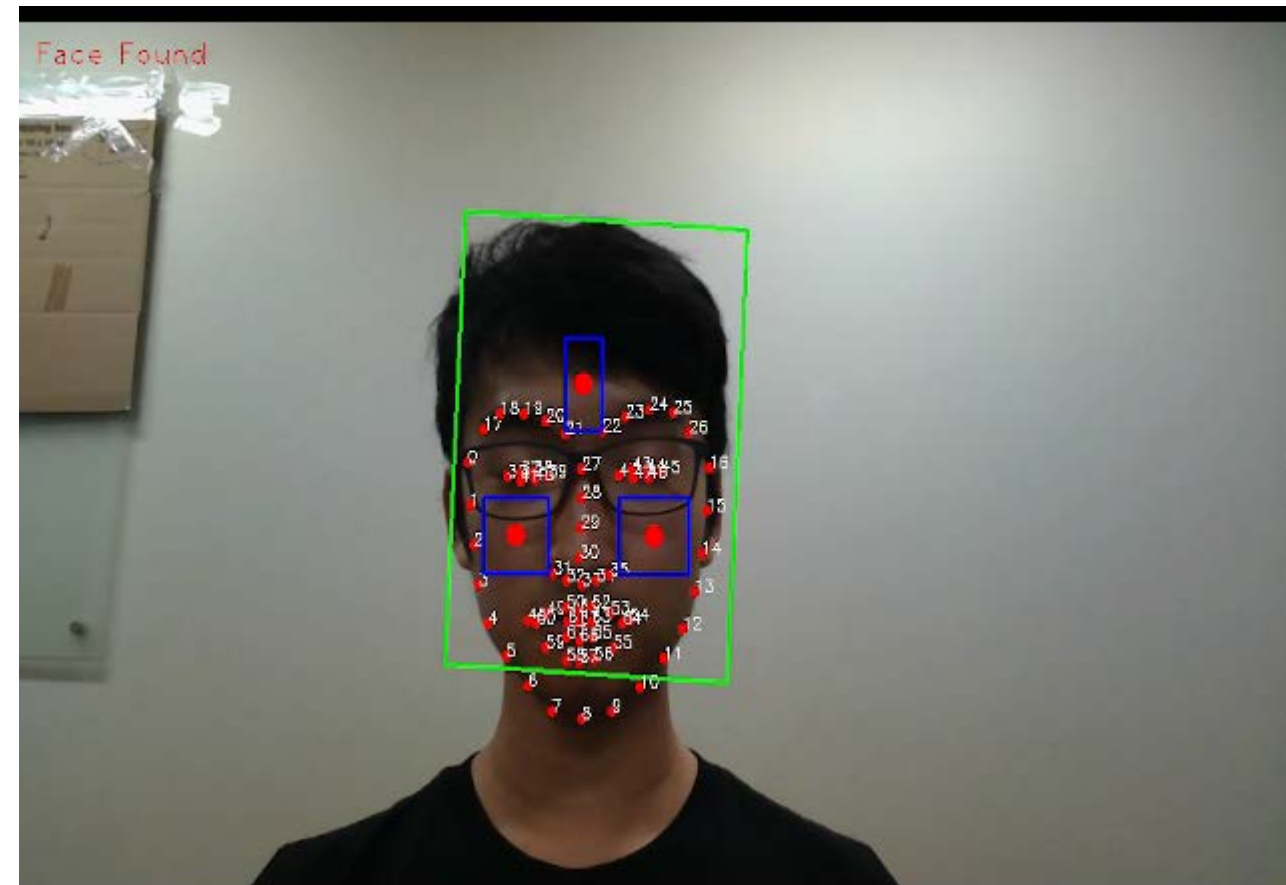


Remote Capturing of Patients' Biometric Data

Heartbeat

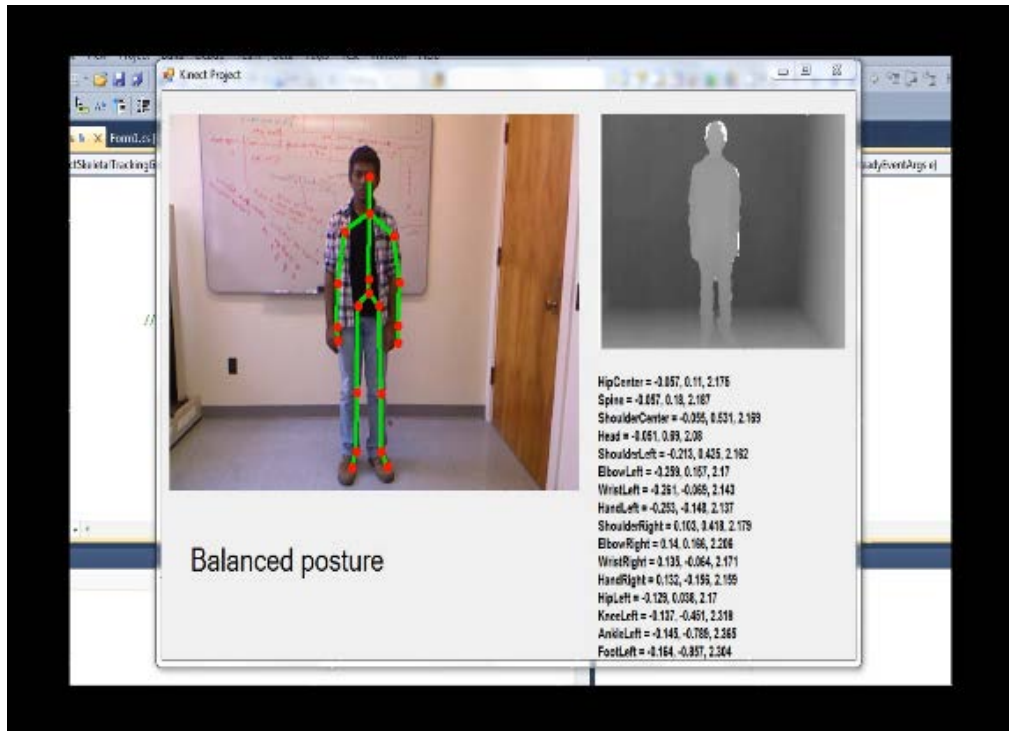


Pain Level

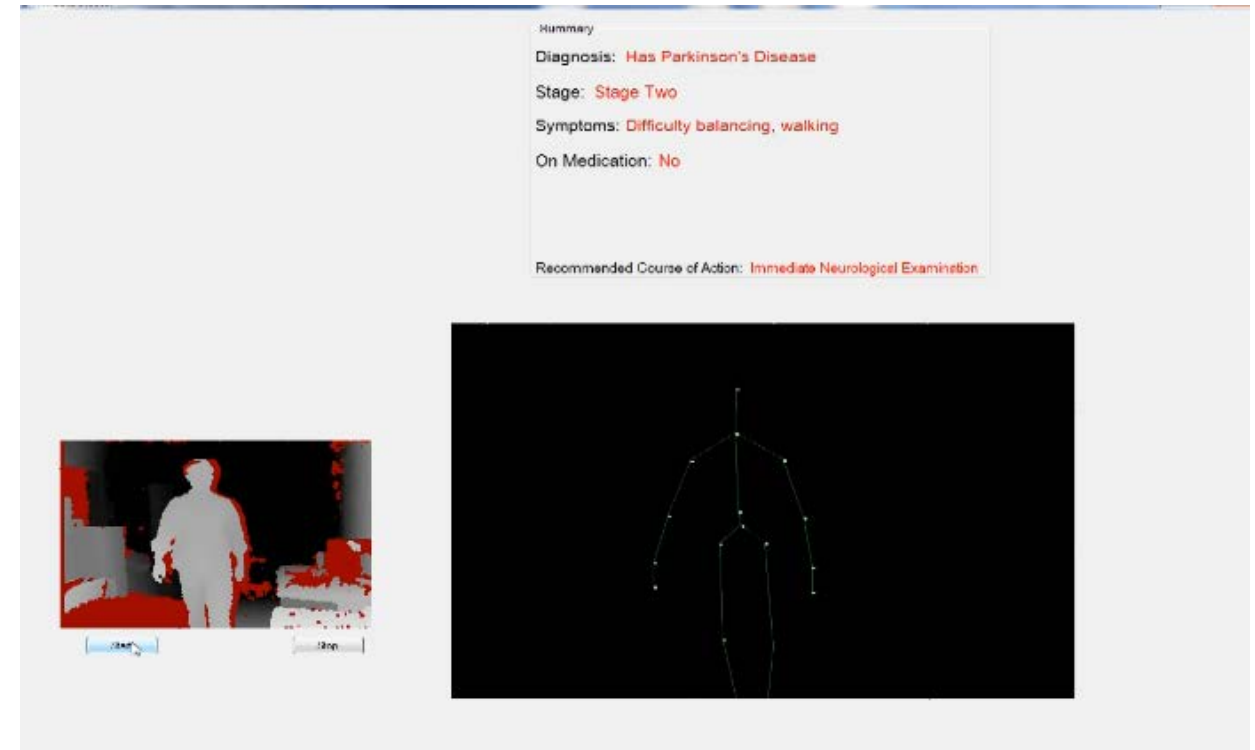


Remote Capturing of Patients' Motor Functions

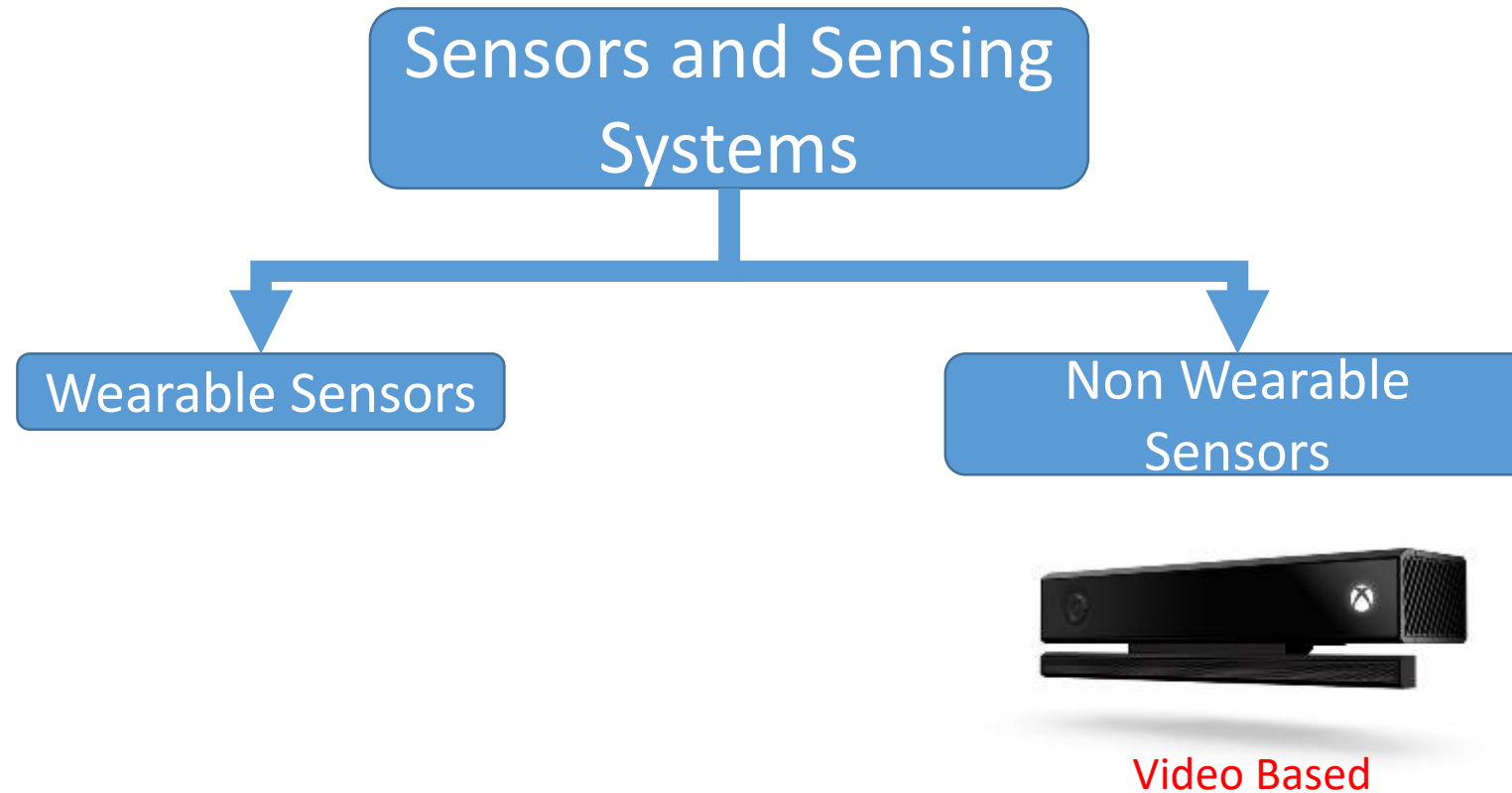
Posture



Gait



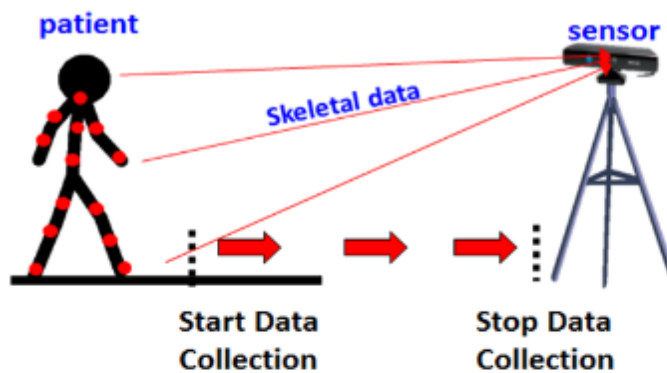
Healthcare Sensors and Sensing Systems



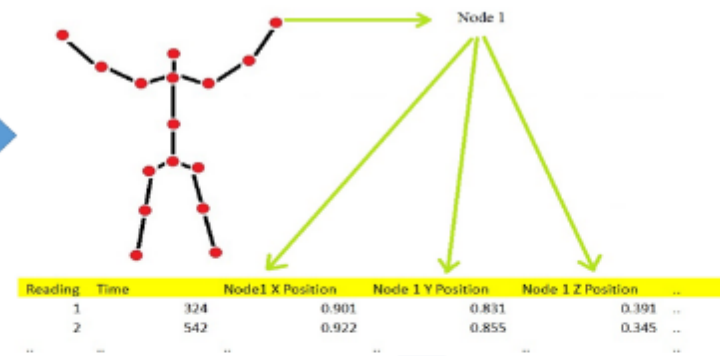
Sensors and Sensing Systems in Healthcare

Research Method

STEP 1: Sensor Data Acquisition



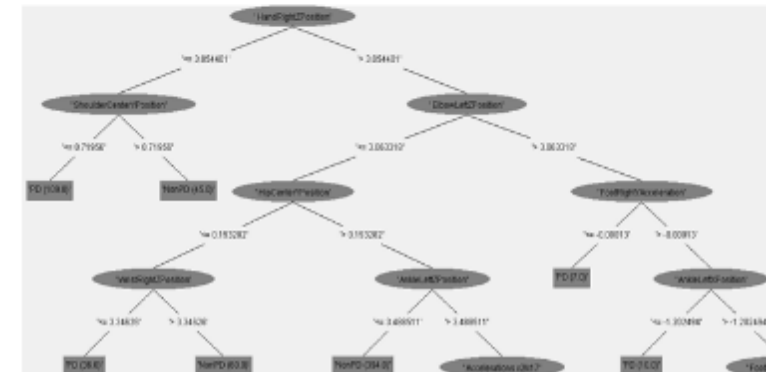
STEP 2: Skeletal Data Preprocessing



STEP 4: Healthcare Decision Support



STEP 3: Data Mining Knowledge Discovery



Research Questions

- How can we learn patients' needs?
- How can we better connect patients with their healthcare providers?
- **How can we make the healthcare sector more efficient?**

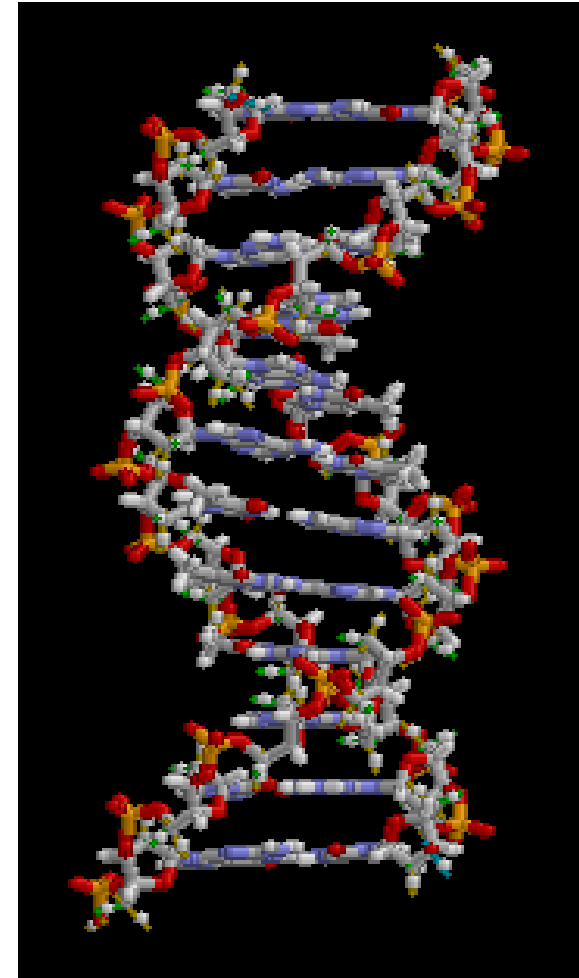
Personalized Healthcare through Precision Medicine

Objective:

- Develop novel machine learning techniques for identifying actionable genomic similarities among patients with chronic immune diseases.

Partners / Relevance:

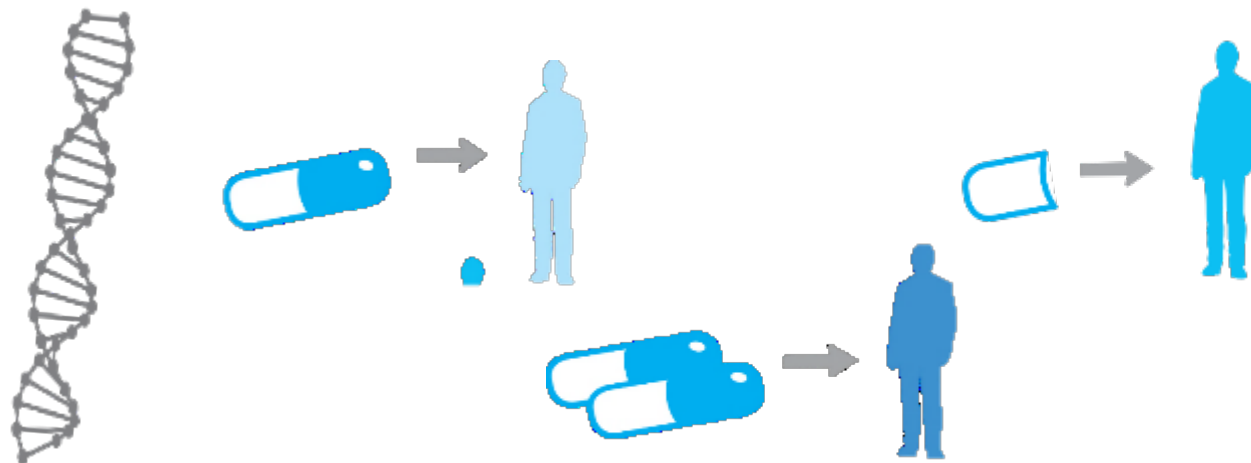
- Reduce trial-and-error treatment costs and improve patient outcomes.



Personalized Healthcare through Precision Medicine

What problem is this project seeking to address?

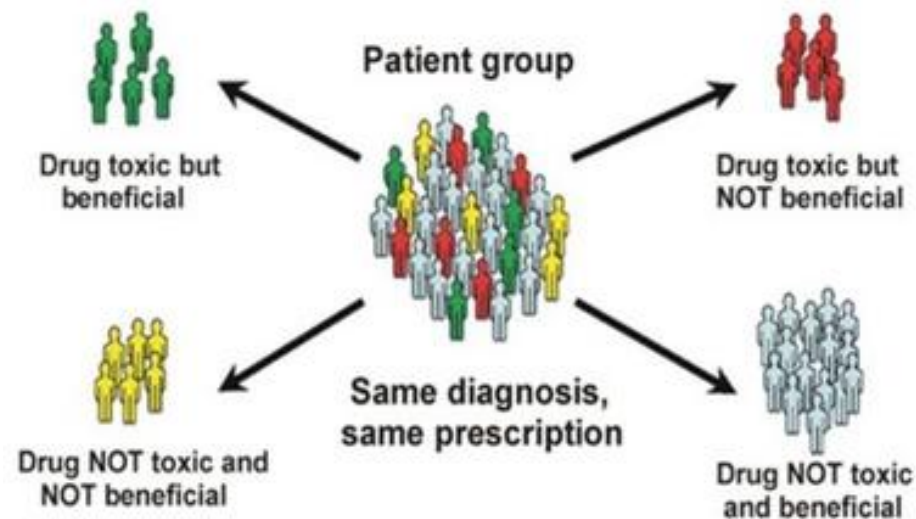
- Chronic immune disorders manifest differently from patient to patient, yet have a genetic etiology.
- Enhanced understanding of the genetic underpinnings can improve disease treatment.



Personalized Healthcare through Precision Medicine

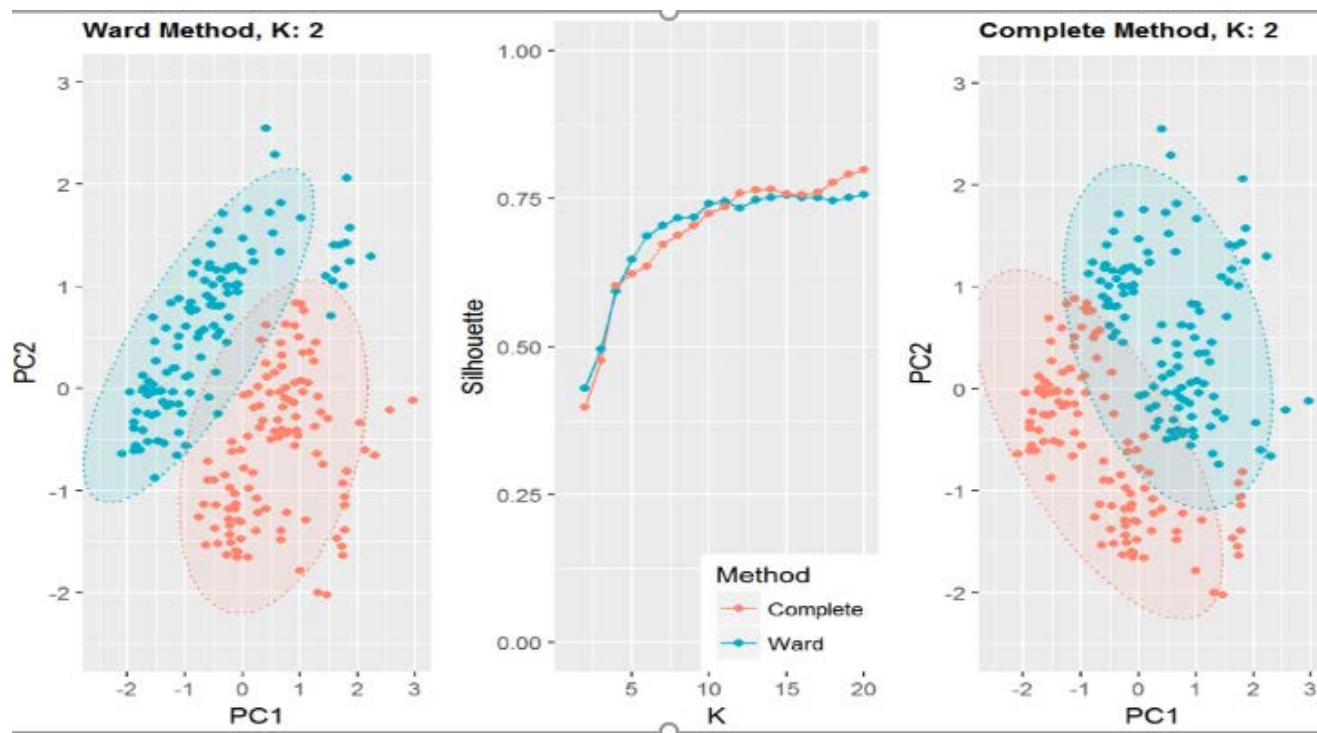
Research Perspective

- Patient – enhanced care
- Provider – informed treatment decisions
- Health system – improved patient outcomes
- Payer – efficient patient-based treatment



Personalized Healthcare through Precision Medicine

- Hierarchical unsupervised machine learning is employed to immunochip data collected from a Multiple Sclerosis patient cohort.
- Significant genetic clusters are defined by the algorithm and patients are assigned to these clusters based on their genetic makeup.





Grounded in Theory: Peer-Reviewed Publications

Prakash, S. K. A., & **Tucker, C. S.** (2018). "Bounded Kalman filter method for motion-robust, non-contact heart rate estimation", Biomedical Optics Express, 9(2), 873-897

Tuarob, S., **Tucker, C.**, Kumara, S., Giles, C. L., Pincus, A. L., Conroy, D. E., & Ram, N. (2017), "How are you feeling?: A personalized methodology for predicting mental states from temporally observable physical and behavioral information", Journal of Biomedical Informatics, 68, 1-19

Lopez, C., and **Tucker, C.** (2017), "A Quantitative Method for Evaluating the Complexity of Implementing and Performing Game Features in Physically-Interactive Gamified Applications", Computers in Human Behavior, DOI: <http://dx.doi.org/10.1016/j.chb.2017.01.036>

Munoz, D. A., Kilinc, M. S., Nembhard, H. B., **Tucker, C.**, & Huang, X. (2017), "Evaluating the Cost-Effectiveness of an Early Detection of Parkinson's Disease through Innovative Technology", The Engineering Economist, DOI: <http://dx.doi.org/10.1080/0013791X.2017.1294718>

Lim, S., **Tucker, C.**, and Kumara, S. (2016), "An unsupervised machine learning model for discovering latent infectious diseases using social media data," Journal of Biomedical Informatics, DOI: <http://dx.doi.org/10.1016/j.jbi.2016.12.007>

Bharathi, A., Singh, A., **Tucker, C.**, and Nembhard, H. (2016), "Knowledge Discovery of Game Design Features By Mining User-Generated Feedback", Computers in Human Behavior, 60, 361-371.

Tucker, C., Behoora, I., Black-Nembhard, H., Lewis, M., Sterling, N., and Huang, X. (2015), "Machine Learning Classification of Medication Adherence in Patients with Movement Disorders Using Non-Wearable Sensors", Computers in Biology and Medicine, 66, 120-134.

Tucker, C., Han, Y., Nembhard-Black, H., Lee, W., Lewis, M., Sterling, N. and Huang, X. (2015), "A Data Mining Methodology For Predicting Early Stage Parkinson'S Disease Using Non-Invasive, High-Dimensional Gait Sensor Data", IIE Transactions on Healthcare Systems Engineering, 5(4), 238-254.

Research Summary

- How can we learn patients' needs?
- How can we better connect patients with their healthcare providers?
- How can we make the healthcare sector more efficient?

Questions?

www.chot.psu.edu



Conrad Tucker, Ph.D.

Associate Professor of Engineering Design and Industrial Engineering

Email: ctucker4@psu.edu