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
Current CHOT Industry Members

Texas Children's Hospital
UAB Health System
Lakeshore General Health System
Opelousas General Health System
Palm Healthcare Foundation, Inc.
Seattle Children's Hospital
Alacare Home Health & Hospice
Philips
MULTICARE
BetterConnected
Main Line Health
Well ahead
CARE Coordination Institute
Children's Healthcare of Atlanta
Morehouse School of Medicine
Grady
Penn State Hershey Milton S. Hershey Medical Center
Pennsylvania Office of Rural Health
Highmark
AT&T
Siemens
Quantum Innovations
American Society of Anesthesiologists
Carena
VA
U.S. Department of Veterans Affairs
Last Best Chance
HealthSouth
Research Inspiration

https://www.youtube.com/watch?v=jfSNxVqprvM
Current State of Healthcare?

- Patient
- Culture
- Trust
- etc.
- Physician
Population Distribution: State of Pennsylvania

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

Population per sq. mile
<1
1...10
10...25
25...50
50...100
100...250
250...500
500...1000
1000...2500
2500...5000
>5000
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

- Sense
- Acquire
- Model
- Classify
- Feedback
The Process of Learning

Humans

- Sense
- Acquire
- Model
- Classify
- Feedback

Machines

- Sense
- Acquire
- Model
- Classify
- Feedback

ctucker4@psu.edu

http://www.engr.psu.edu/datalab/
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
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
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.
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
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
**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

Sense

Acquire

Model

Classify

Feedback
Population Health: How do we learn what a patient population needs?
Woke up Wednesday morning with razor blades in my throat and a headache so painful I couldn’t... https://t.co/nDgffw6Z7d from @melissasbinkley at 03/03/2017 13:34
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

Wearable Sensors

Non Wearable Sensors

Healthcare Sensors and Sensing Systems

http://www.engr.psu.edu/datalab/
Healthcare Sensors and Sensing Systems

Sensors and Sensing Systems

Wearable Sensors

Non Wearable Sensors
Capturing of Patients’ Biometric Data

- Heartbeat
- Pain Level
Remote Capturing of Patients’ Biometric Data

Heartbeat

Pain Level

Before Color Amplification

After Color Amplification

*Arrows indicate dominant regions of color change corresponding to change in blood volume.
Remote Capturing of Patients’ Biometric Data

Heartbeat

Pain Level

Before Color Amplification

After Color Amplification

*Arrows indicate dominant regions of color change corresponding to change in blood volume

Heart Rate (B.P.M)

B.P.M.
Remote Capturing of Patients’ Biometric Data

Heartbeat

Pain Level
Remote Capturing of Patients’ Motor Functions

Posture

Gait

ctucker4@psu.edu

http://www.engr.psu.edu/datalab/
Healthcare Sensors and Sensing Systems

Sensors and Sensing Systems

- Wearable Sensors
- Non Wearable Sensors

Video Based
Sensors and Sensing Systems in Healthcare

Research Method

STEP 1: Sensor Data Acquisition

STEP 2: Skeletal Data Preprocessing

STEP 3: Data Mining Knowledge Discovery

STEP 4: Healthcare Decision Support

http://www.engr.psu.edu/datalab/
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?
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.


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