

Congratulations!

ICS/DATA 435 and ICS 635

Spring 2023

CES Course Evaluations

- Please fill out to help design future iterations of this course
- +2 points on the final exam for everyone if >95% class participation

Reminder of Deadlines

Final Project Code / Writeup: May 9 at 11:55pm

Final Exam (Laulima): May 15 at 5pm

Outline for Today

Putting It All Together: How ChatGPT Works

Quick Advertisement: Machine Learning for Healthcare Research

Final Exam Review

Course Review

Outline for Today

Putting It All Together: How ChatGPT Works

Quick Advertisement: Machine Learning for Healthcare Research

Final Exam Review

Course Review

Core Ingredients of ChatGPT

1. Attention based Transformer decoder model (GPT)
2. Self-supervised pre-training
3. Reinforcement Learning from Human Feedback (RLHF)

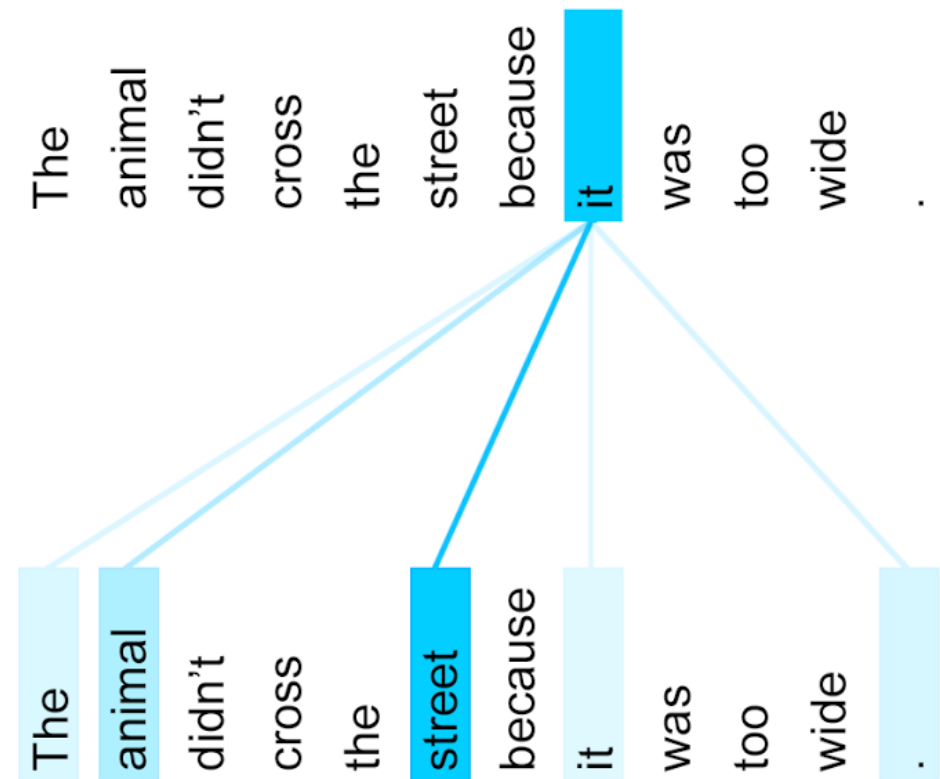
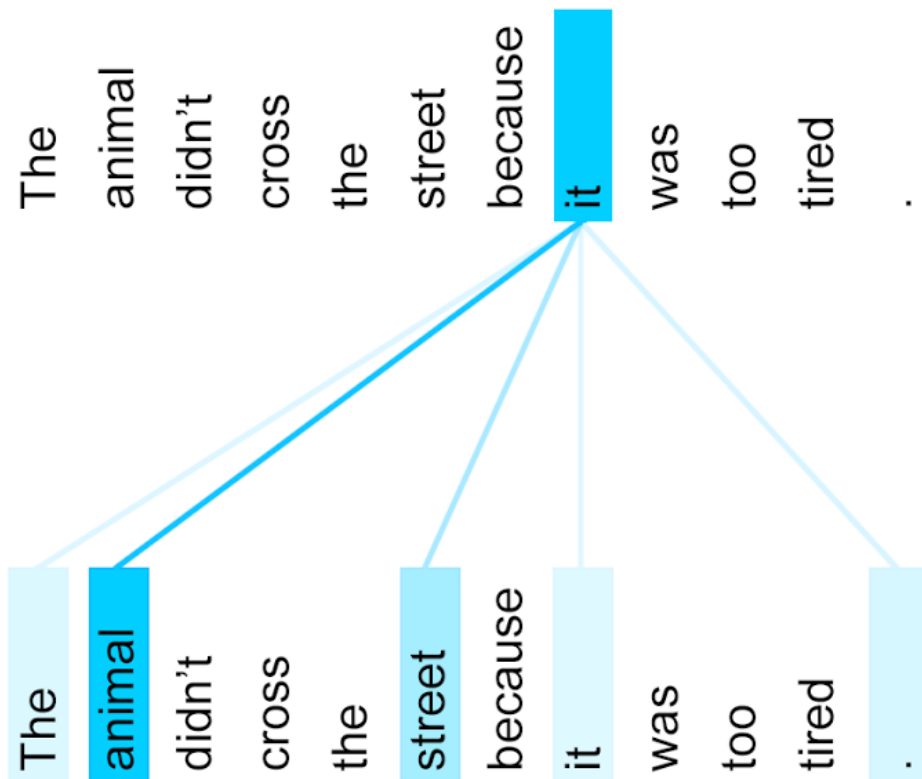
Core Ingredients of ChatGPT

1. Attention based Transformer decoder model (GPT)
2. Self-supervised pre-training
3. Reinforcement Learning from Human Feedback (RLHF)

We've talked about these first two ingredients extensively over the last few classes!

Quick Review: Attention based Transformer decoder model (GPT)

Self Attention:



Quick Review: Self-supervised pre-training

Text Corpus

Nothing is impossible.
Even the word
impossible
says I'm possible



Task: Predict from past

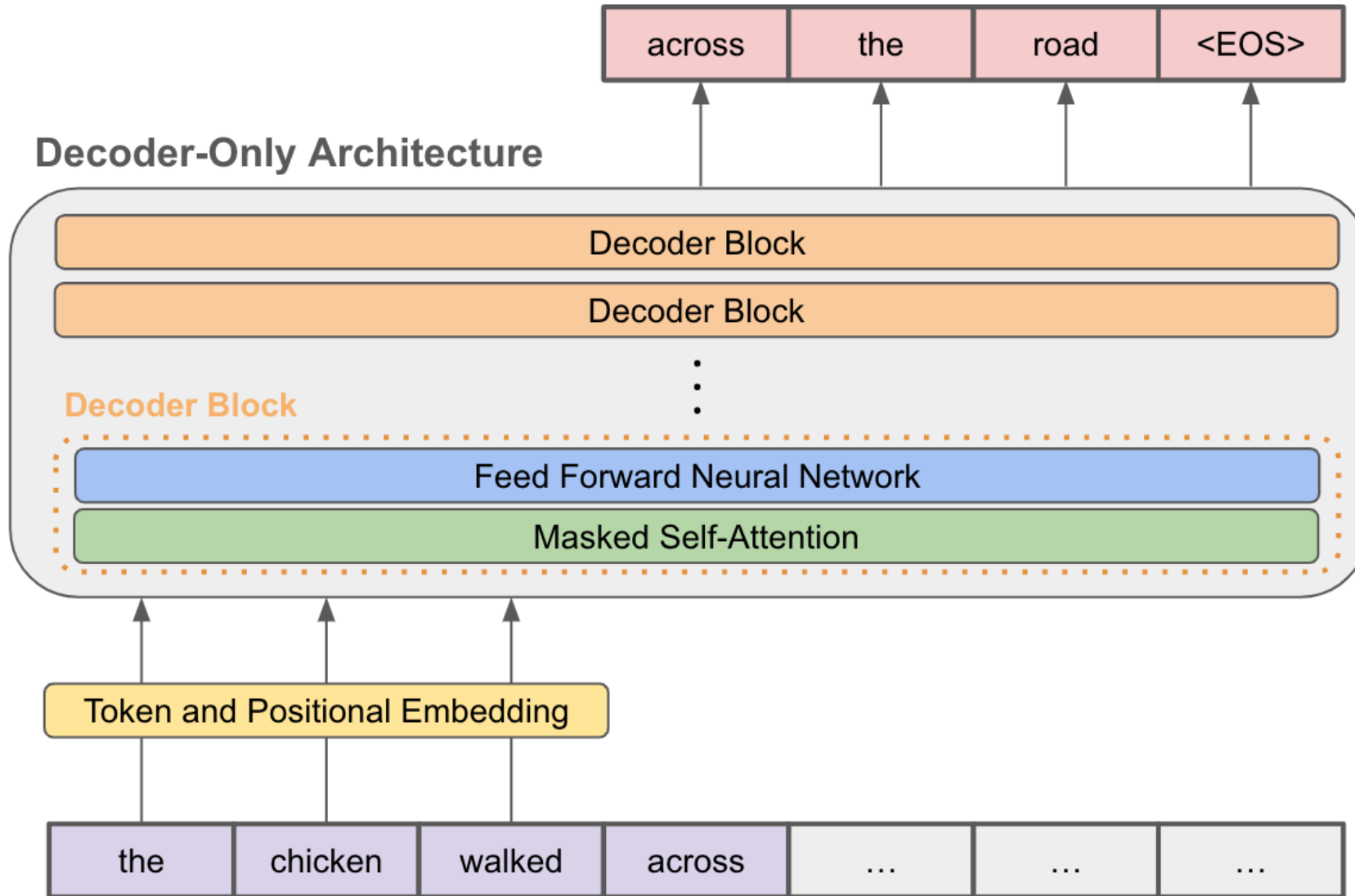
Nothing

Nothing is

Nothing is impossible

...

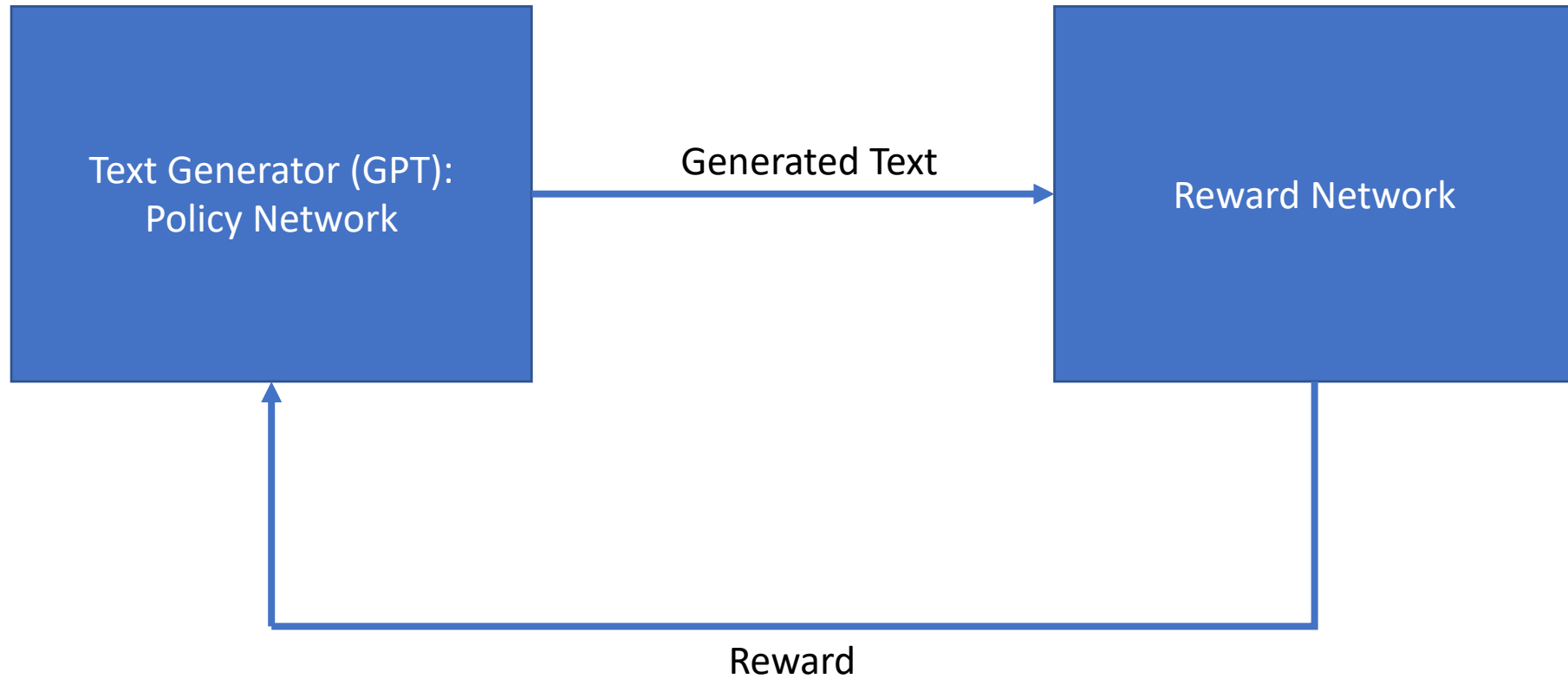
Quick Review: Self-supervised pre-training



Reinforcement Learning from Human Feedback (RLHF)

- Train a "reward model" directly from human feedback
- Use the model as a reward function to optimize an agent's policy (the text generator) using RL through an optimization algorithm
 - Proximal Policy Optimization, a model-free policy estimation method, is used by OpenAI

Reinforcement Learning from Human Feedback (RLHF)

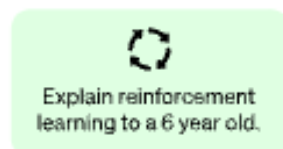


Reinforcement Learning from Human Feedback (RLHF)

Step 1

Collect demonstration data and train a supervised policy.

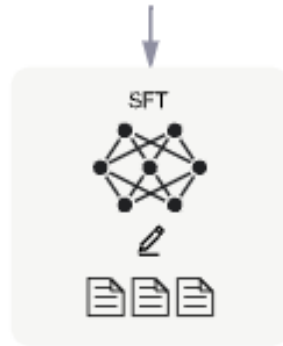
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



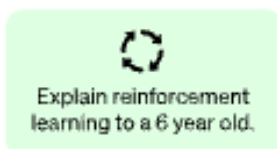
This data is used to fine-tune GPT-3.5 with supervised learning.



Step 2

Collect comparison data and train a reward model.

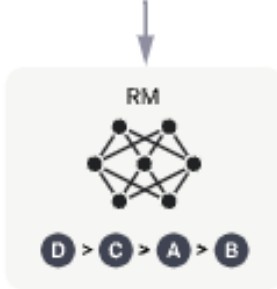
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



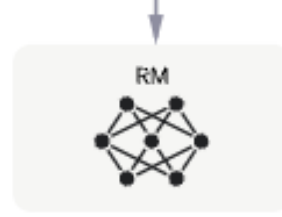
The PPO model is initialized from the supervised policy.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



Core Ingredients of ChatGPT

1. Attention based Transformer decoder model (GPT)
2. Self-supervised pre-training
3. Reinforcement Learning from Human Feedback (RLHF)

ChatGPT Wisdom

P What types of real-world projects can someone who completed a Machine Learning class now accomplish?



Completing a Machine Learning class is an excellent accomplishment, and it equips you with the necessary knowledge to work on a variety of real-world projects. Here are some types of projects you could consider:

1. **Image Classification:** You could work on a project that involves classifying images based on specific categories, such as animals or vehicles. This could involve using deep learning models such as Convolutional Neural Networks (CNNs) to train your algorithm.
2. **Natural Language Processing (NLP):** NLP is an exciting field that involves training models to understand and interpret human language. Projects could include sentiment analysis of social media posts, chatbots that can carry on conversations with humans, or language translation tools.
3. **Fraud Detection:** Machine Learning algorithms can be used to detect fraudulent activities such as credit card fraud or identity theft. You could work on a project that uses supervised learning algorithms to identify patterns in data and flag suspicious transactions.

4. **Predictive Maintenance:** This involves using machine learning algorithms to predict when a machine or device is likely to fail. This could be applied to a variety of industries, such as manufacturing, transportation, and energy.
5. **Recommendation Systems:** Recommendation systems use machine learning algorithms to suggest products, services, or content based on a user's past behavior. Projects in this area could include developing a recommendation engine for an e-commerce site or a music streaming service.
6. **Time Series Analysis:** Time series analysis involves analyzing data that changes over time. This could be applied to a variety of areas such as finance, weather forecasting, and traffic prediction.

These are just a few examples of the types of projects that you could work on after completing a Machine Learning class. The key is to identify a problem or challenge that you are passionate about and apply your newly acquired skills to find a solution.

ChatGPT Wisdom

P What types of jobs can someone get who knows Python, scikit-learn, TensorFlow, and the basics of machine learning theory? What are the typical salaries of such jobs?



There are many job opportunities available for individuals with knowledge of Python, scikit-learn, TensorFlow, and the basics of machine learning theory. Here are some examples:

1. **Machine Learning Engineer:** Machine learning engineers build, test, and deploy machine learning models in production. They also work on designing and optimizing algorithms, as well as maintaining and improving the existing codebase. Salaries for machine learning engineers can range from \$100,000 to \$200,000 per year.
2. **Data Scientist:** Data scientists work on analyzing large datasets to extract insights and inform business decisions. They use statistical and machine learning techniques to build predictive models and identify patterns in the data. Salaries for data scientists can range from \$80,000 to \$150,000 per year.

3. **Research Scientist:** Research scientists work on developing new machine learning algorithms and improving existing ones. They work on cutting-edge research problems in areas such as computer vision, natural language processing, and reinforcement learning. Salaries for research scientists can range from \$100,000 to \$250,000 per year.
4. **Software Engineer:** Software engineers with knowledge of machine learning are in high demand. They work on developing software applications that use machine learning algorithms, such as recommendation systems, fraud detection systems, and chatbots. Salaries for software engineers can range from \$80,000 to \$150,000 per year.
5. **Data Engineer:** Data engineers work on designing, building, and maintaining the infrastructure necessary to store and process large datasets. They work closely with data scientists and machine learning engineers to ensure that the data is available and ready for analysis. Salaries for data engineers can range from \$100,000 to \$200,000 per year.

Note that these salary ranges are based on various factors such as location, industry, experience, and job responsibilities. They are meant to provide a general idea of what to expect, and salaries may vary significantly depending on these factors.

Outline for Today

Putting It All Together: How ChatGPT Works

Quick Advertisement: Machine Learning for Healthcare Research

Final Exam Review

Course Review



HAWAI'I DIGITAL HEALTHCARE LAB

peterwashington.github.io

Peter Yiğitcan Washington

Assistant Professor

Information & Computer Sciences

University of Hawai'i at Mānoa

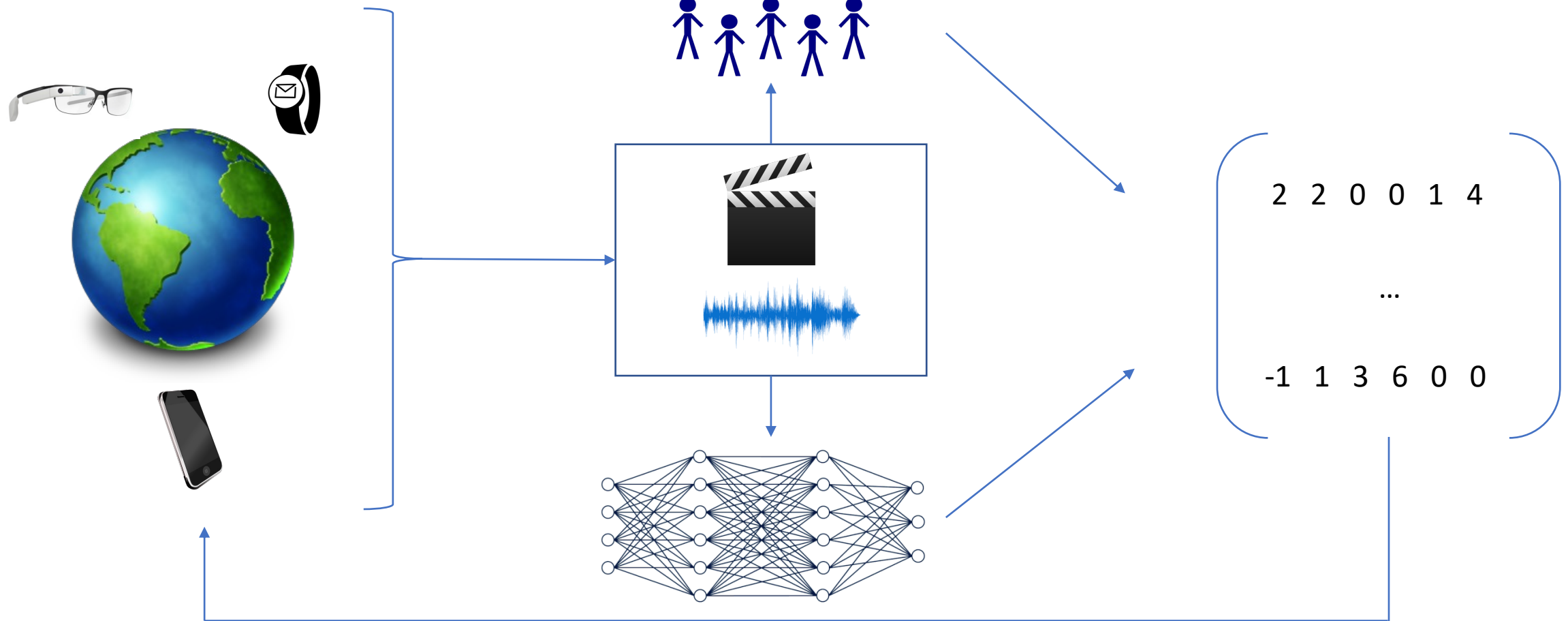
pyw@hawaii.edu

The Digital Phenotyping Process

Digital interventions for health monitoring

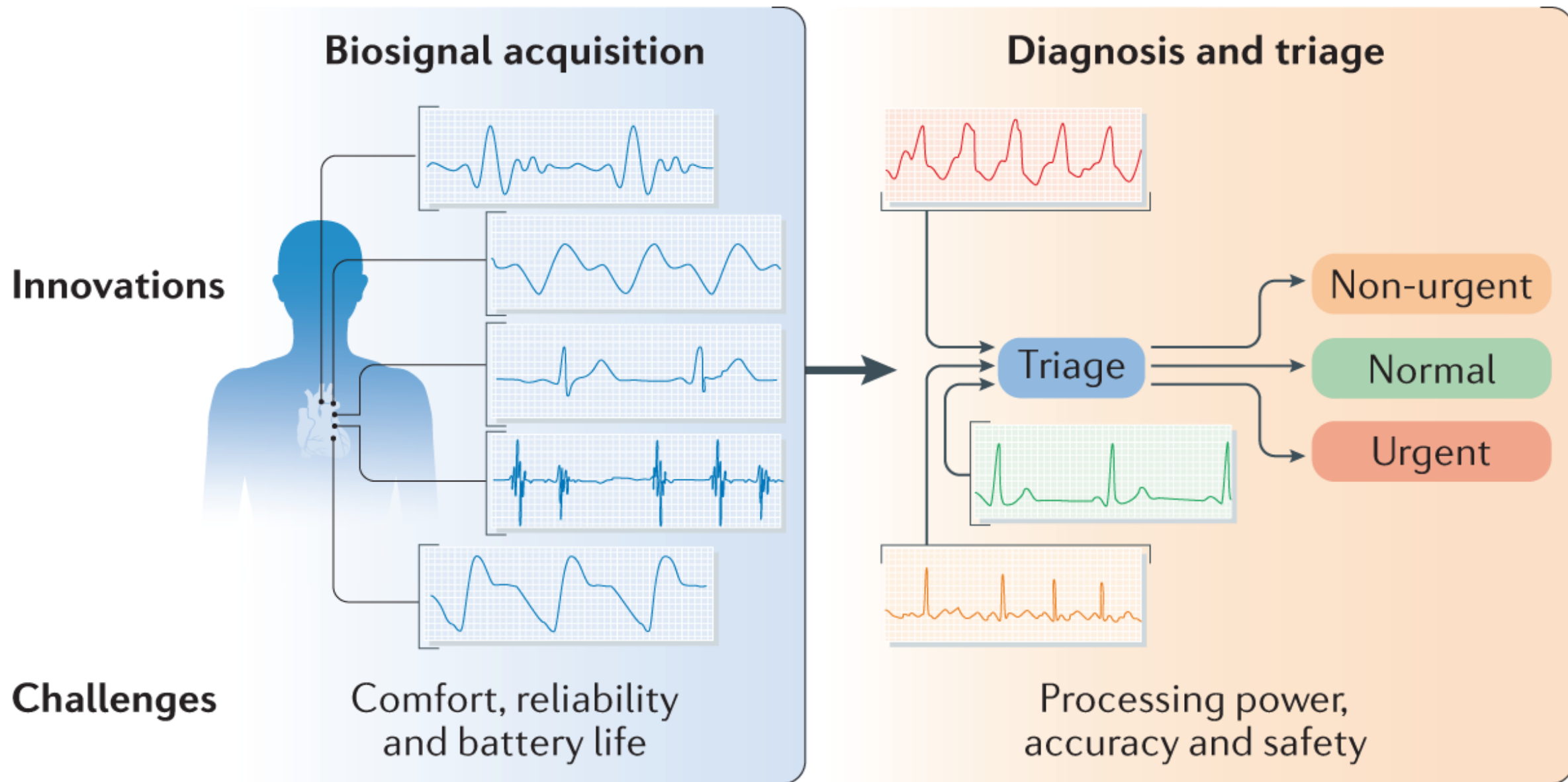
Feature extraction via automated algorithms and crowdsourcing

Digital diagnostics and adaptive therapies

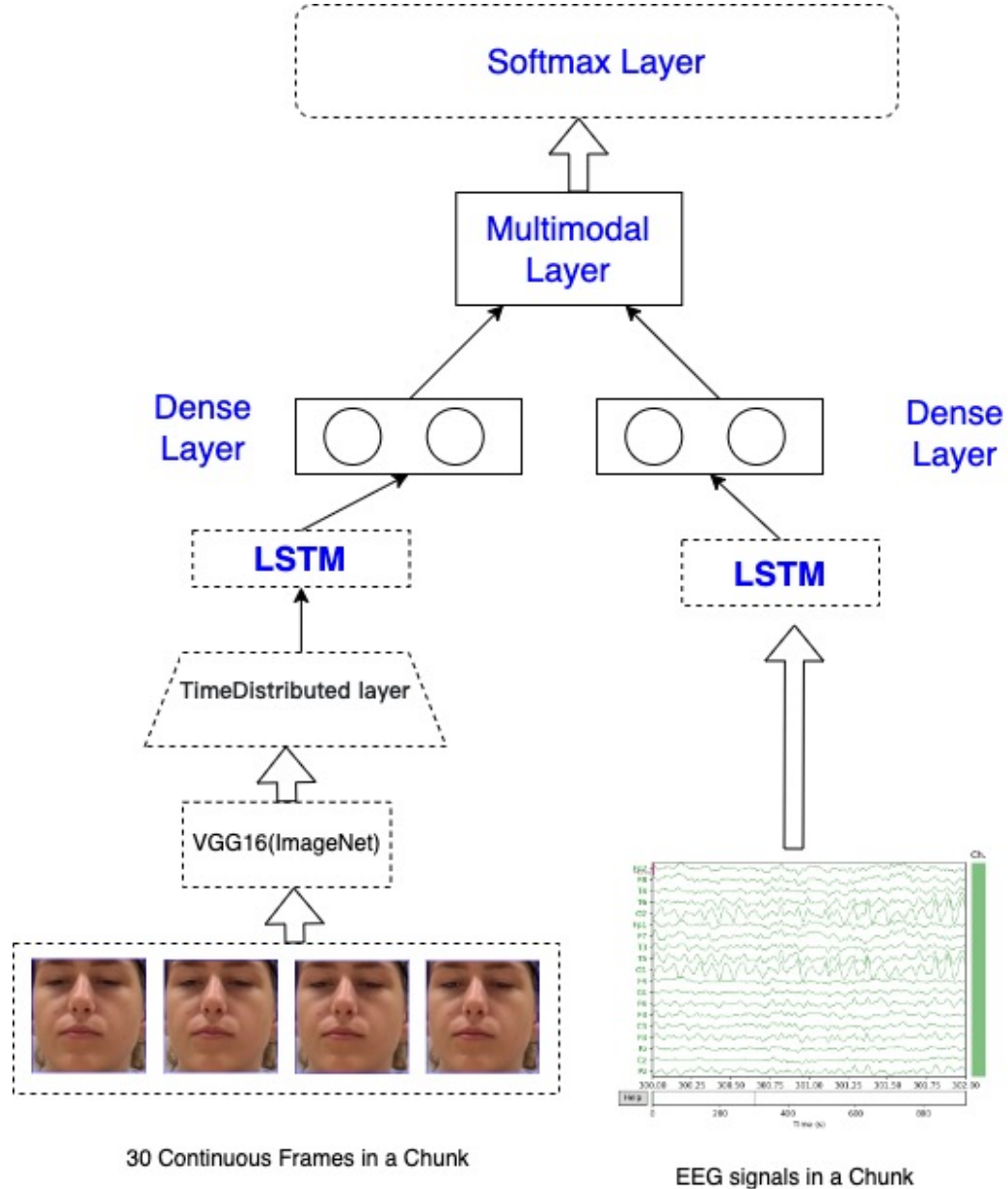


Part 1: Methods

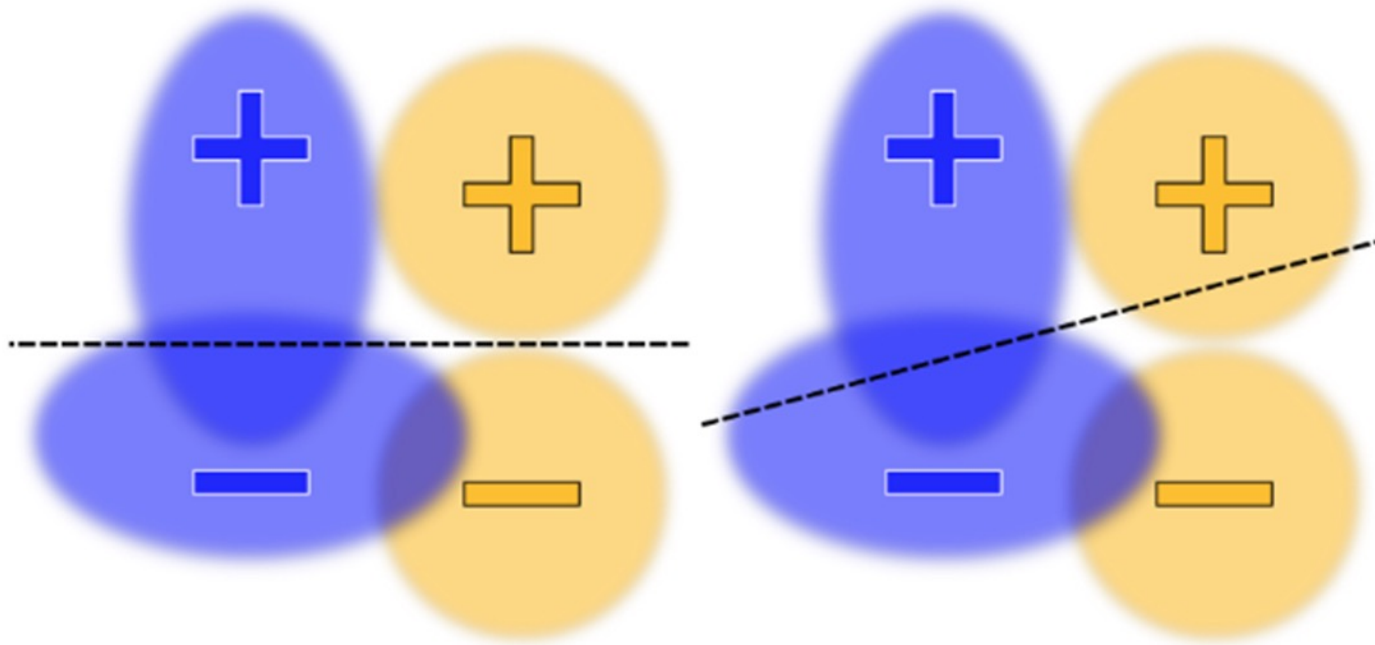
Multimodal Machine Learning



Multimodal Machine Learning

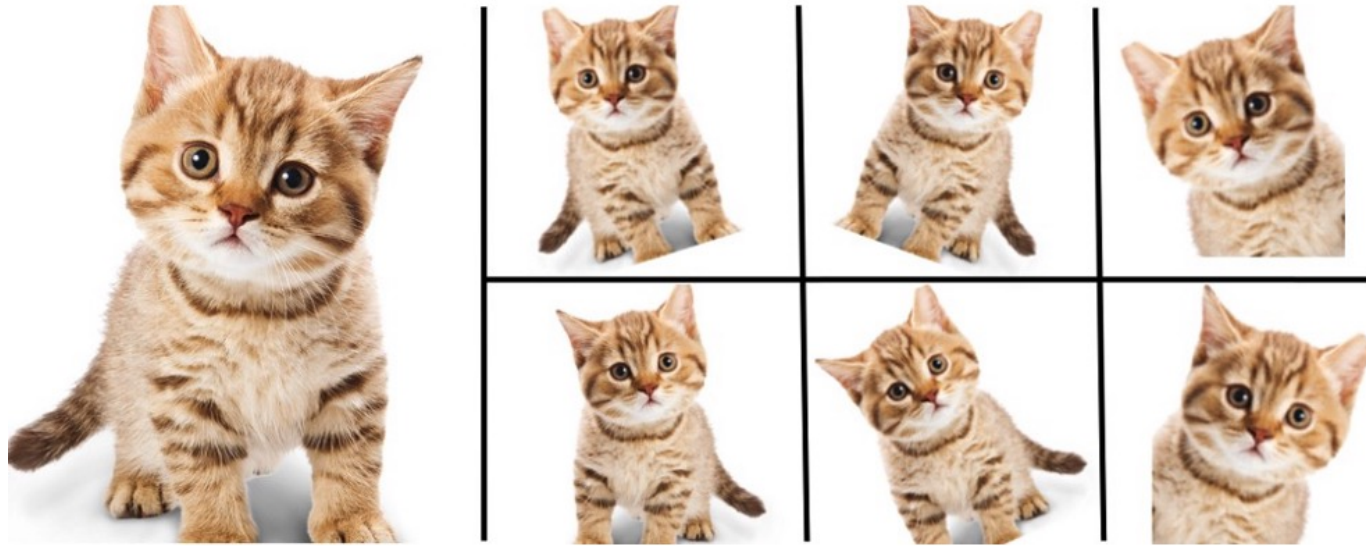


Fair Machine Learning



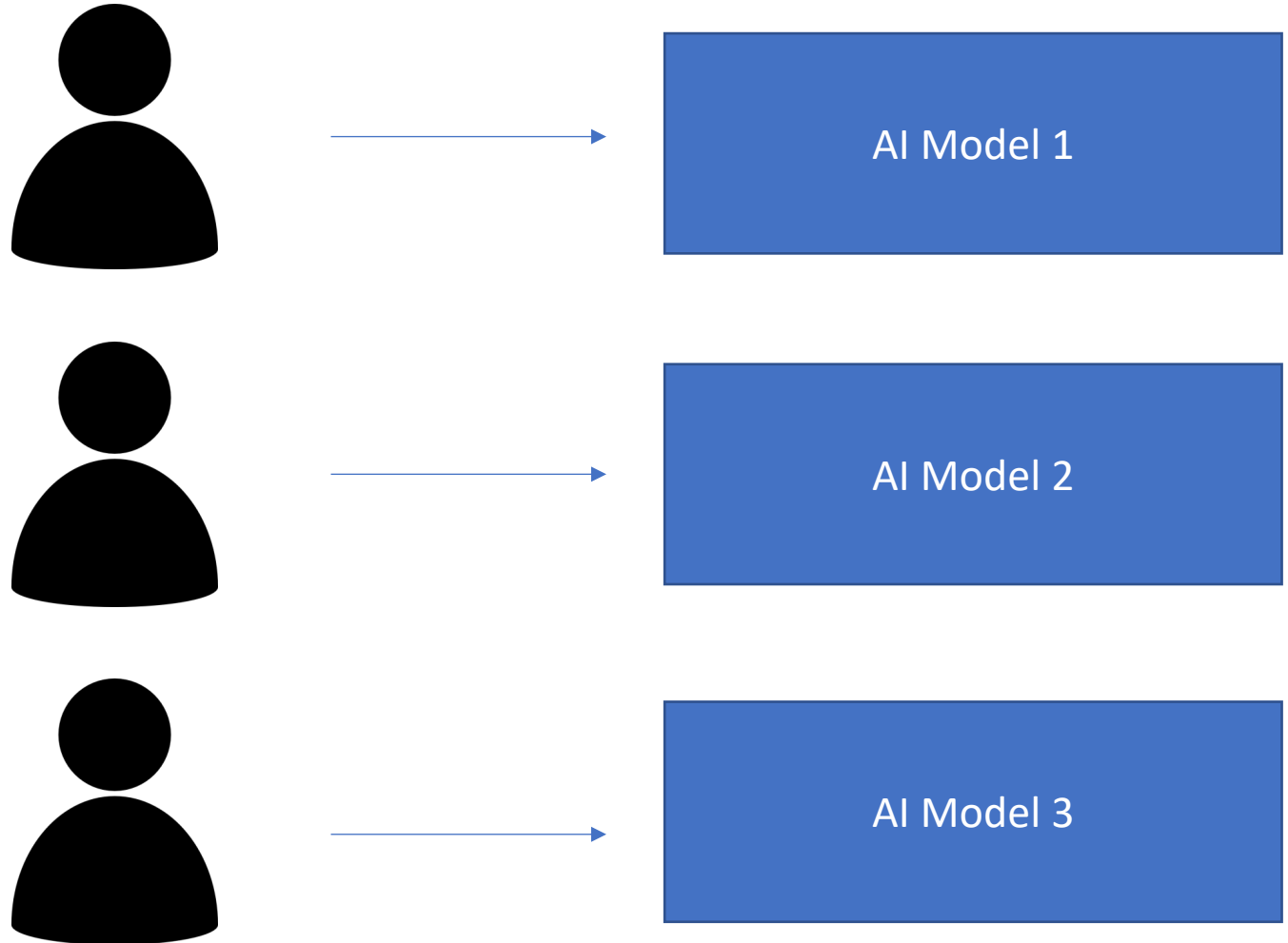
$$\text{MSE} + \frac{1}{n_1 n_2} \sum_{\substack{(x_i, y_i) \in \mathcal{S}_1 \\ (x_j, y_j) \in \mathcal{S}_2}} d(y_i, y_j) (w \cdot x_i - w \cdot x_j)^2.$$

Fair Machine Learning

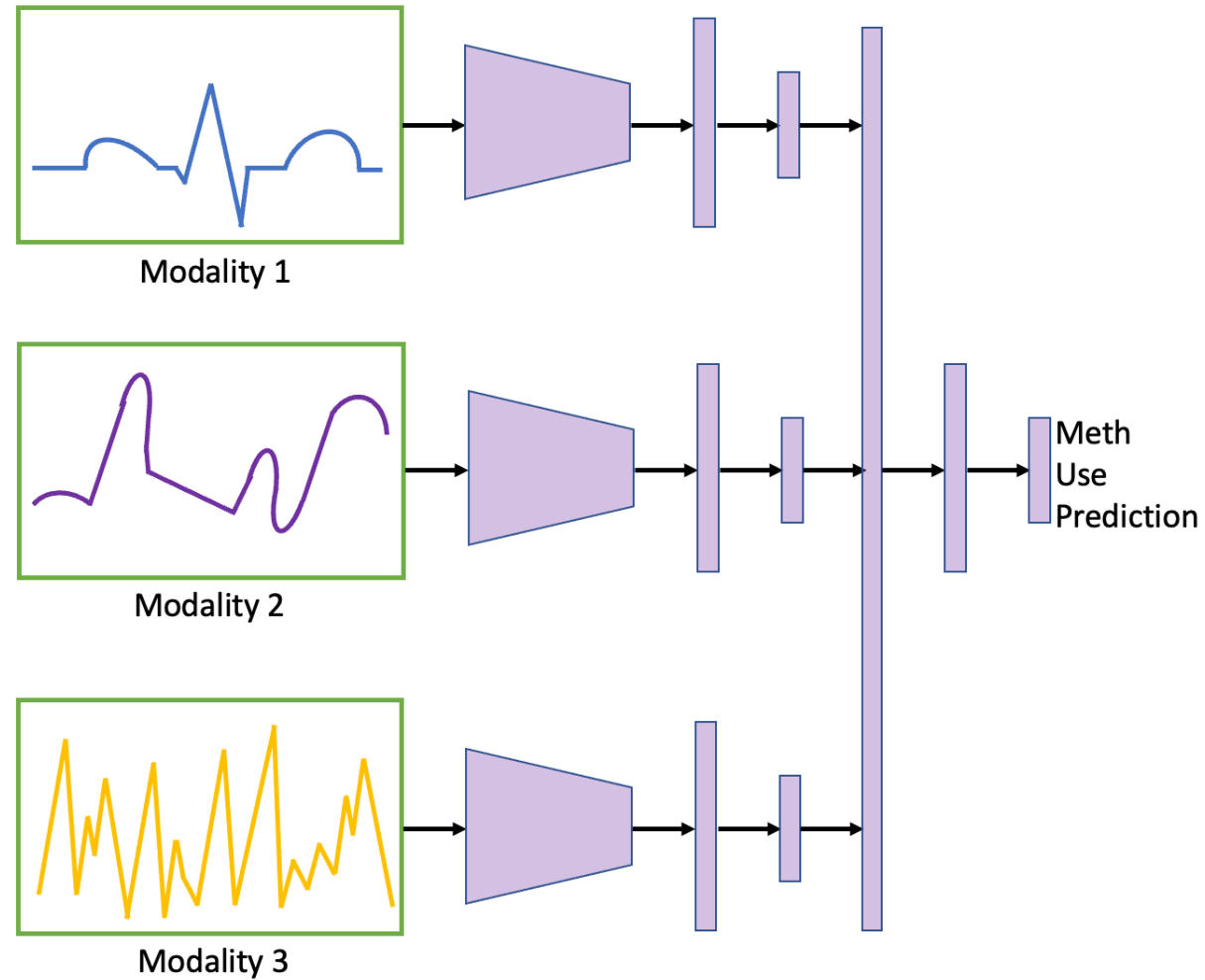


Data Augmentation Upsampling

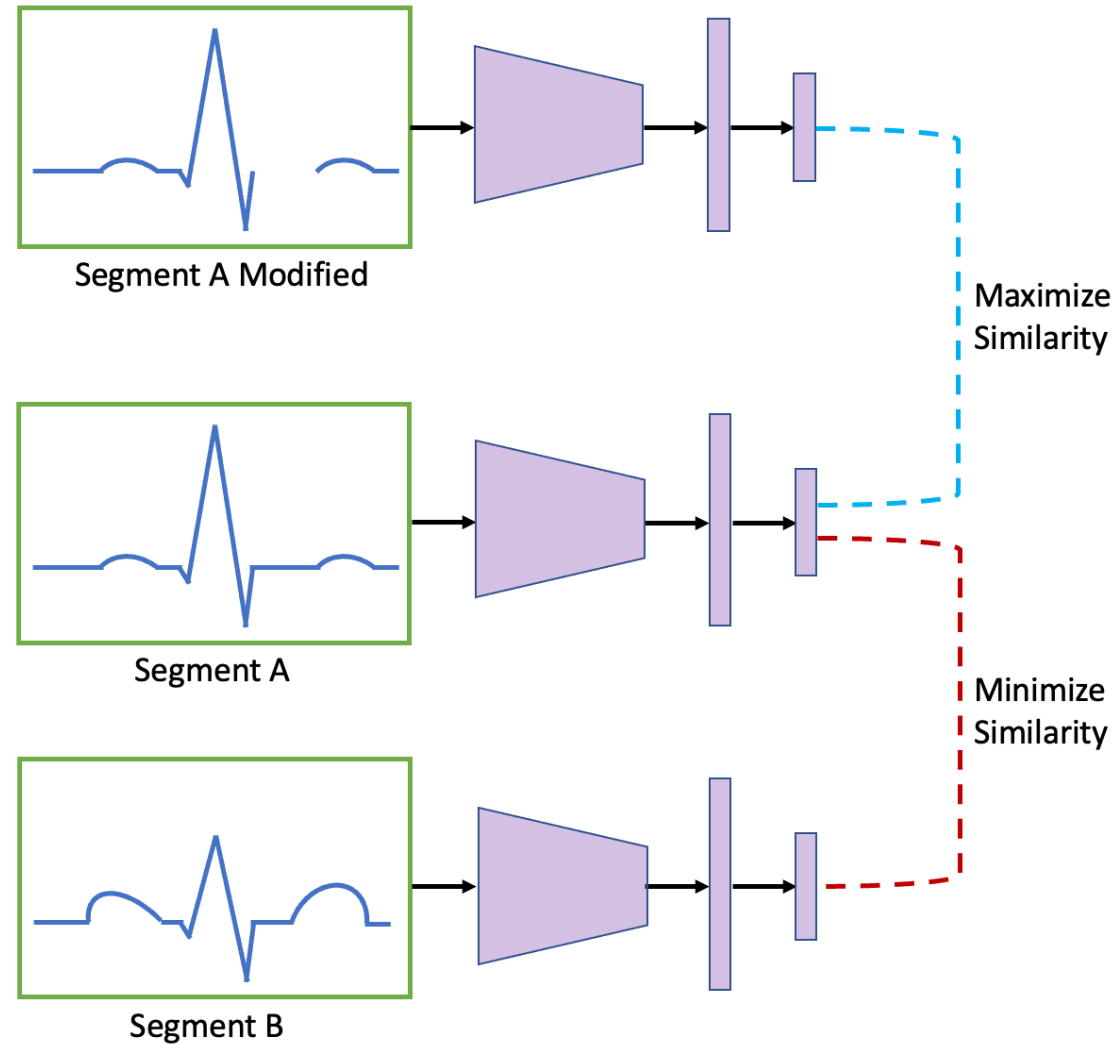
Personalized Machine Learning



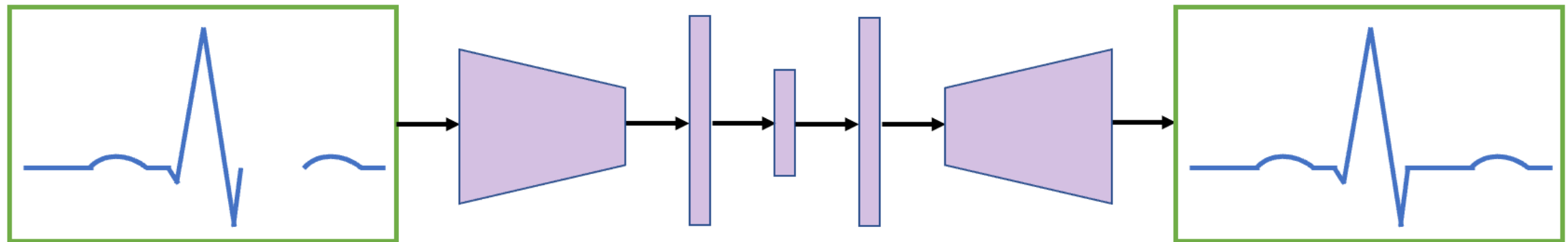
Personalized Machine Learning through Self-Supervised Learning



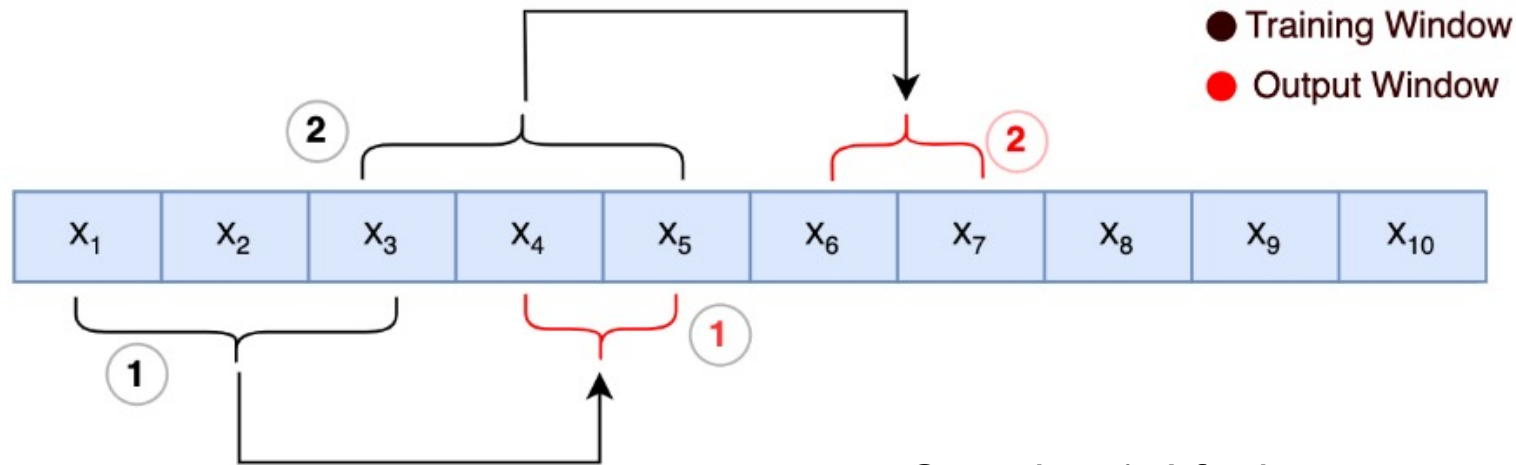
Personalized Machine Learning through Self-Supervised Learning



Personalized Machine Learning through Self-Supervised Learning

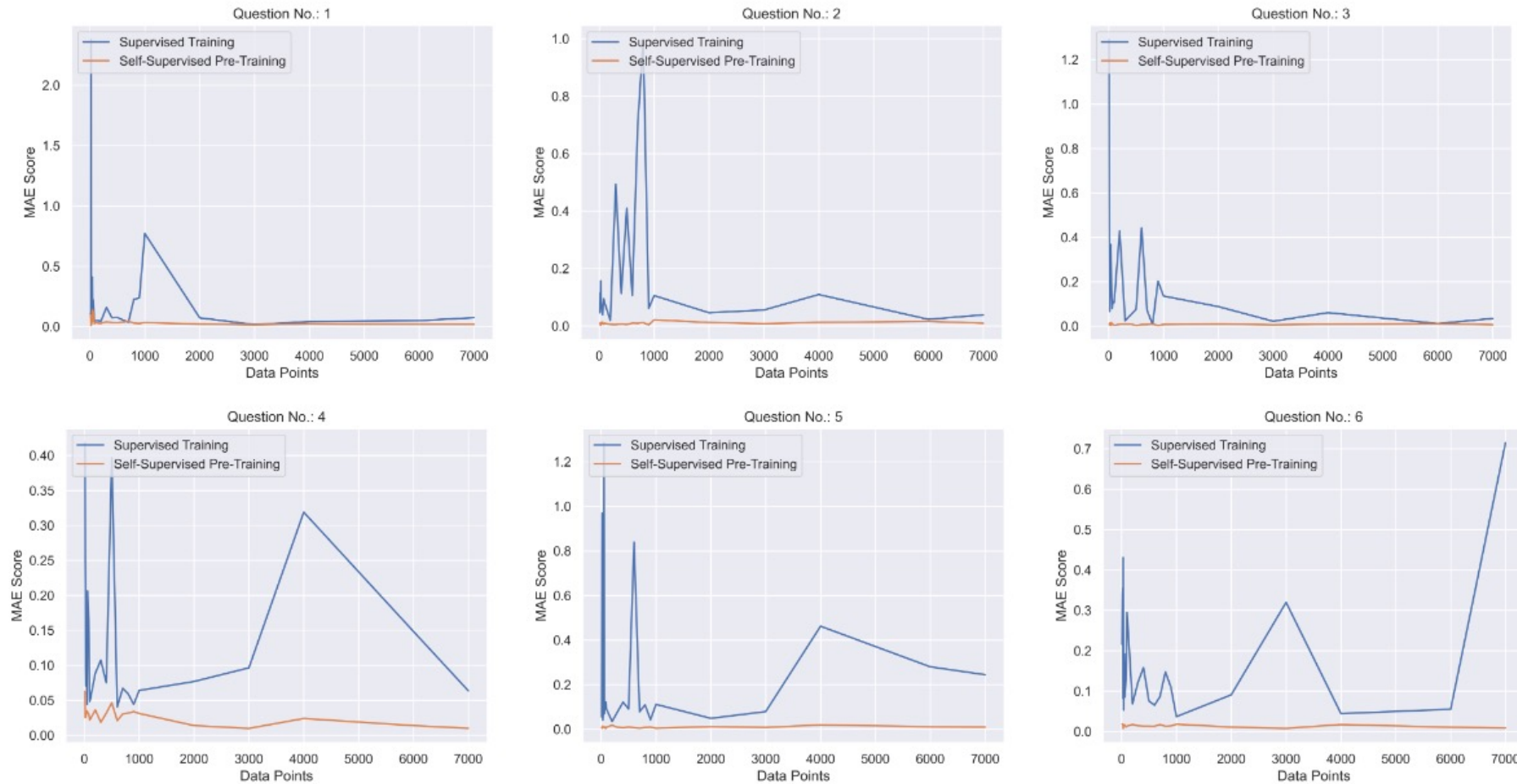


Results: Wearable Biosignal-Based Stress Detection



- Question 1: I feel at ease
- Question 2: I feel nervous
- Question 3: I am jittery
- Question 4: I am relaxed
- Question 5: I am worried
- Question 6: I feel pleasant

Results: Wearable Biosignal-Based Stress Detection



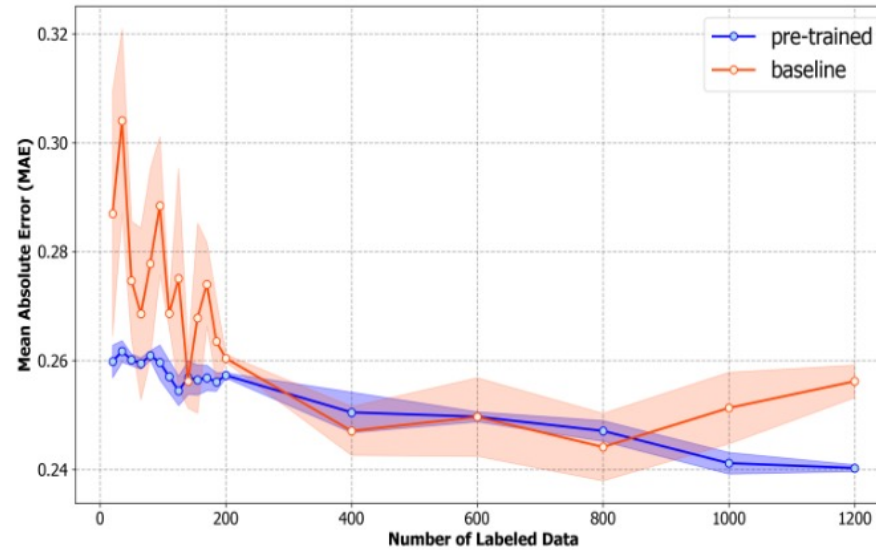
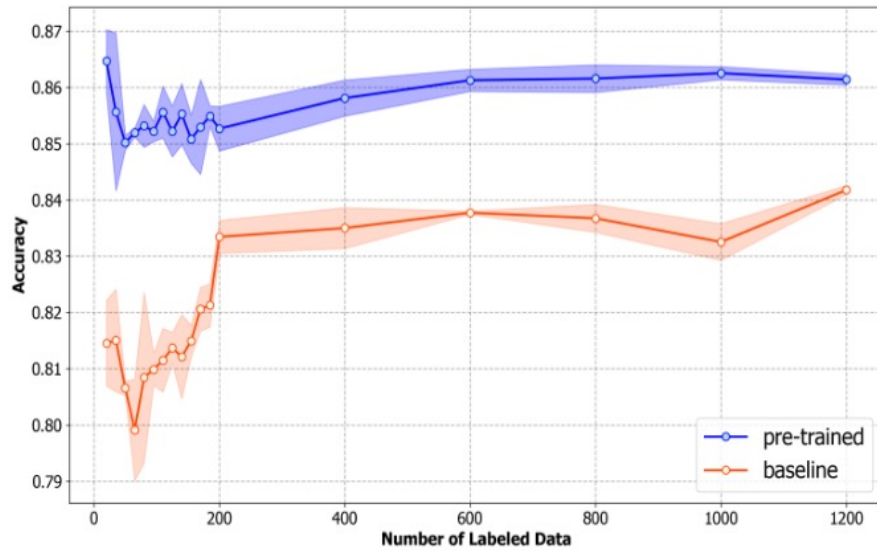
- Question 1: I feel at ease
- Question 2: I feel nervous
- Question 3: I am jittery
- Question 4: I am relaxed
- Question 5: I am worried
- Question 6: I feel pleasant

Results: Wearable Biosignal-Based Stress Detection

| Question | Subject | MAE with pretraining | MAE with supervised training |
|----------|---------|----------------------|------------------------------|
| 1 | S2 | 0.03 ± 0.02 | 0.23 ± 0.25 |
| 1 | S3 | 0.13 ± 0.08 | 0.44 ± 0.7 |
| 1 | S4 | 0.2 ± 0.11 | 0.05 ± 0.03 |
| 1 | S5 | 0.04 ± 0.03 | 0.2 ± 0.13 |
| 1 | S6 | 0.09 ± 0.05 | 0.43 ± 0.35 |
| 1 | S7 | 0.02 ± 0.02 | 0.52 ± 0.36 |
| 1 | S8 | 0.02 ± 0.01 | 0.12 ± 0.1 |
| 1 | S9 | 0.03 ± 0.01 | 0.09 ± 0.05 |
| 1 | S10 | 0.03 ± 0.03 | 0.07 ± 0.02 |
| 1 | S11 | 0.07 ± 0.05 | 0.55 ± 0.65 |
| 1 | S13 | 0.02 ± 0.01 | 0.11 ± 0.1 |
| 1 | S14 | 0.02 ± 0.02 | 0.38 ± 0.45 |
| 1 | S15 | 0.08 ± 0.04 | 0.37 ± 0.2 |
| 1 | S16 | 0.02 ± 0.01 | 0.06 ± 0.03 |
| 1 | S17 | $0.1 \pm .04$ | 00.34 ± 0.26 |

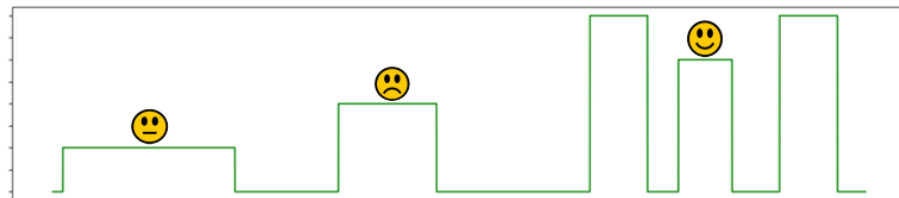
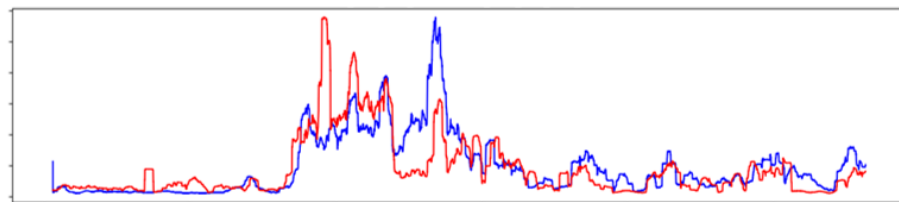
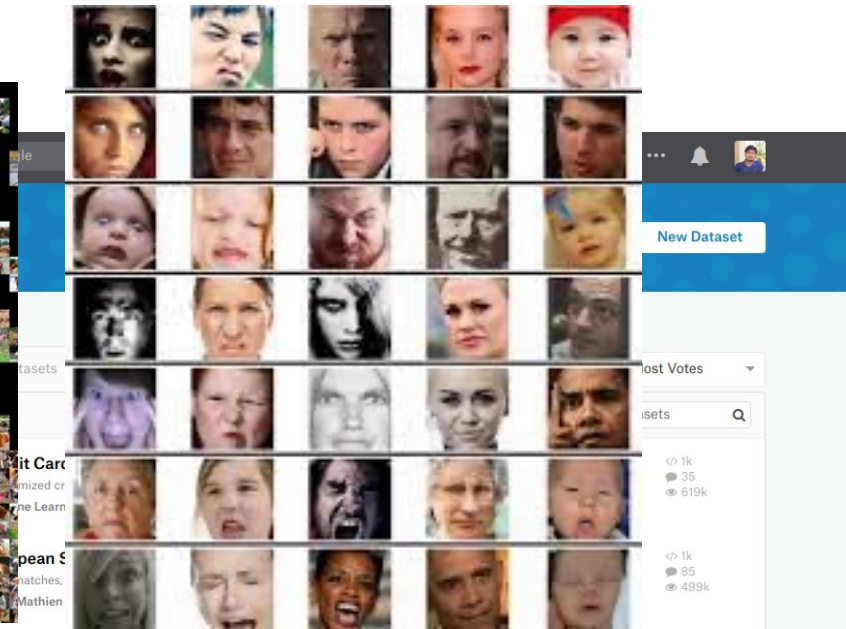
Tanvir Islam and Peter Washington. "Personalization of Stress Detection Models using Self-Supervised Learning." In submission. 2023.

Results: Speech-Based Emotion Recognition



Peranut Nimitsurachat and Peter Washington. "Self-Supervised Learning for Audio-Based Emotion." In preparation. 2023.

Benchmark Dataset Development



| | | | | |
|------|---|---------------------------------------|-------------------------|-------------------|
| 1226 | TMDb 5000 Movie Dataset Metadata on ~5,000 movies from TMDb The Movie Database (TMDb) updated a year ago (Version 2) | film | CSV 9.3 MB Other | 1k 50 516k |
| 992 | Global Terrorism Database More than 180,000 terrorist attacks worldwide, 1970-2017 START Consortium updated 25 days ago (Version 3) | crime terrorism international... | CSV 27.9 MB Other | 634 11 239k |
| 922 | Bitcoin Historical Data Bitcoin data at 1-min intervals from select exchanges, Jan 2012 to July 2018 Zielak updated 3 months ago (Version 14) | history finance | CSV 148.3 MB CC4 | 88 22 226k |
| 765 | Data Science for Good: Kiva Crowdfunding Use Kernels to assess welfare of Kiva borrowers for \$30k in prizes Kiva updated 7 months ago (Version 5) | geography finance lending + 2 more... | CSV 41.9 MB CC0 | 230 47 133k |
| 752 | Wine Reviews 130k wine reviews with variety, location, winery, price, and description zackthoutt updated 10 months ago (Version 4) | critical the... food and dr... | CSV 50.9 MB CC4 | 1k 18 143k |

Benchmark Dataset Development

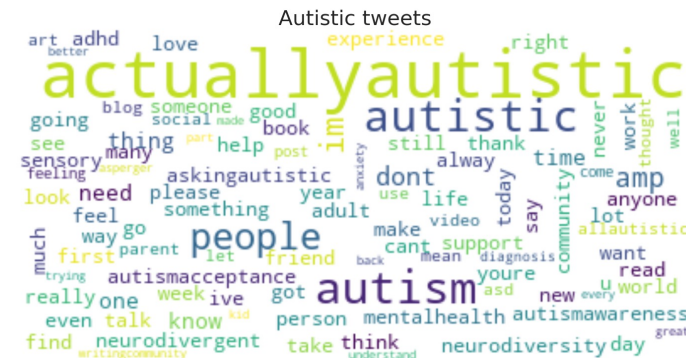
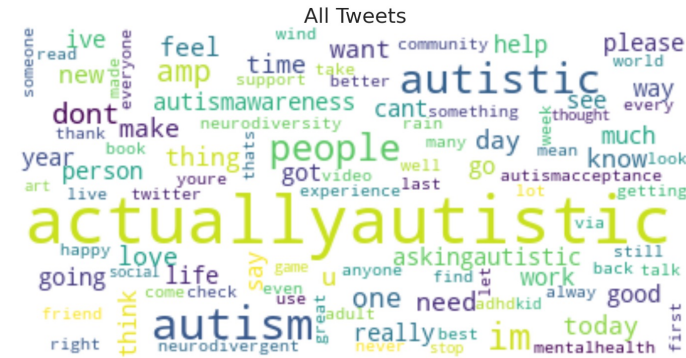
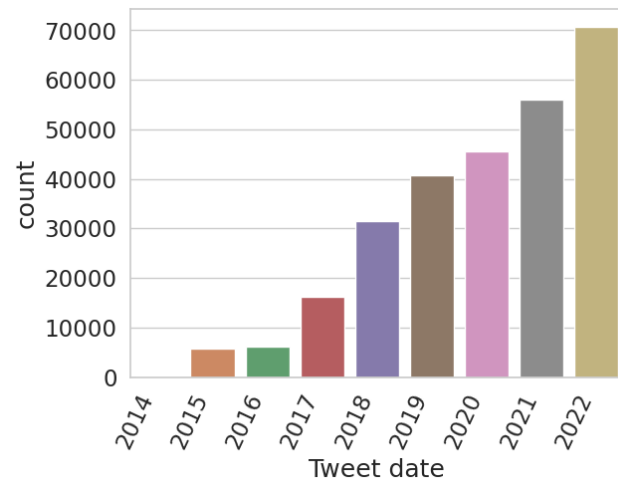
 **Neurodivergent Rebel** 
@NeuroRebel

I know, technically, autism doesn't have "a look" but there are DEFINITELY occasions where I would call myself "visibly autistic" - mostly when my guard is down or if I am suddenly overwhelmed with strong emotions - be them joy, fear, or sorrow. Anyone relate?
#AskingAutistics

4:39 AM - 27 Dec 2018 from Texas, USA

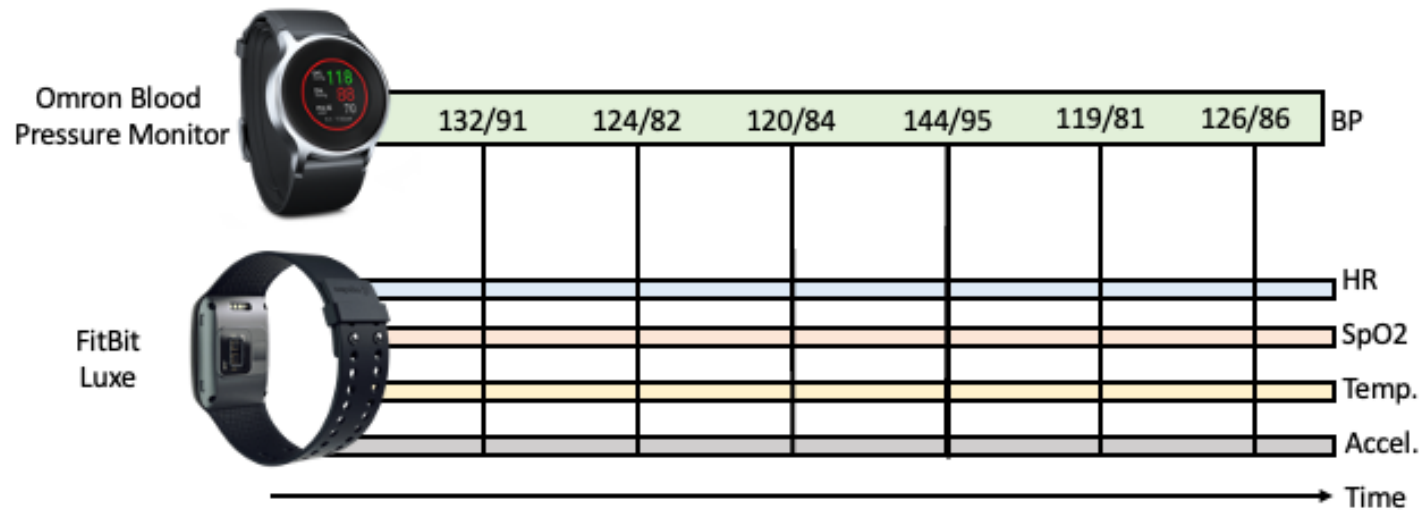
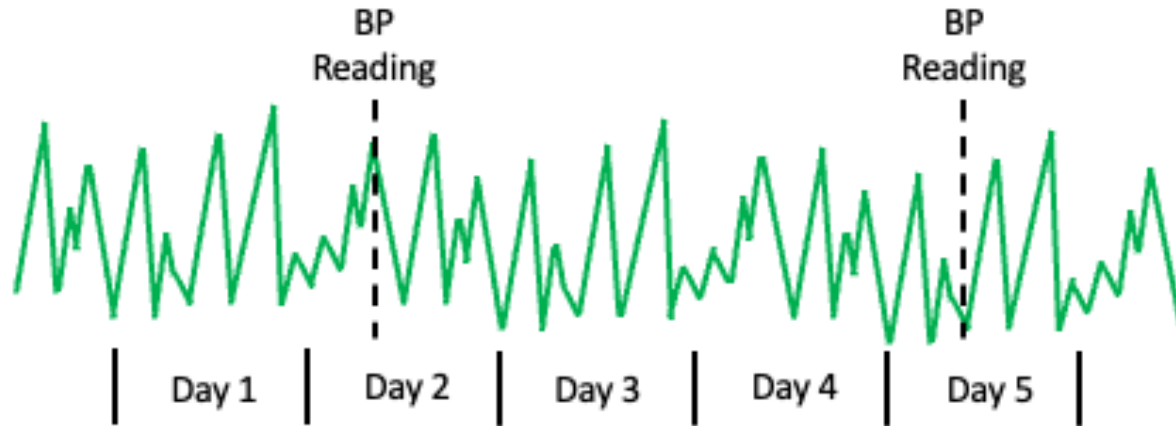
42 Retweets 293 Likes

61 42 293

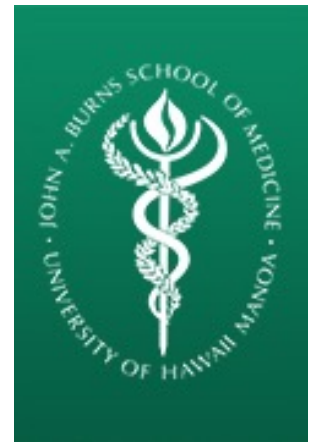


Part 2: Applications

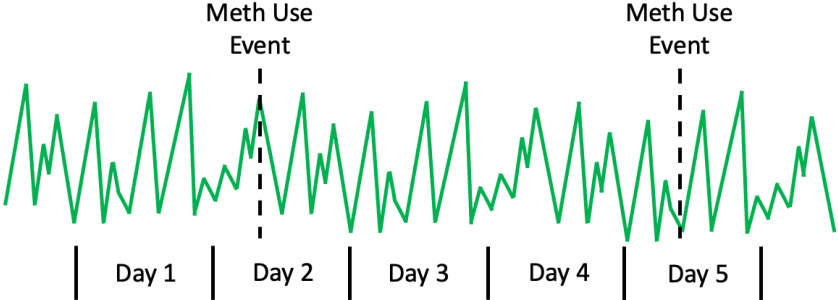
Detecting Blood Pressure Spikes



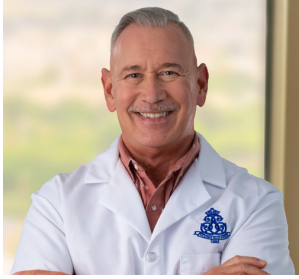
Dr. Kaholokula



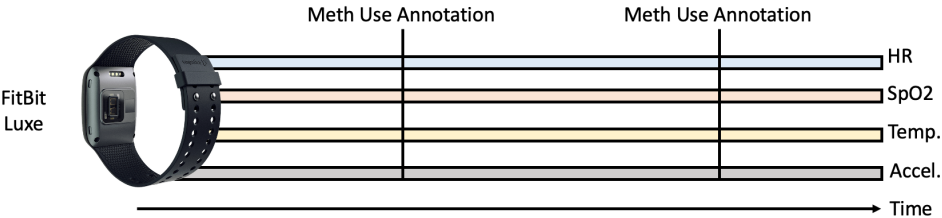
Detecting Meth Craving and Use



Dr. Guerrero



Dr. Busch



Heather Lusk



Dr. Phillips



PIKO



Psychiatry



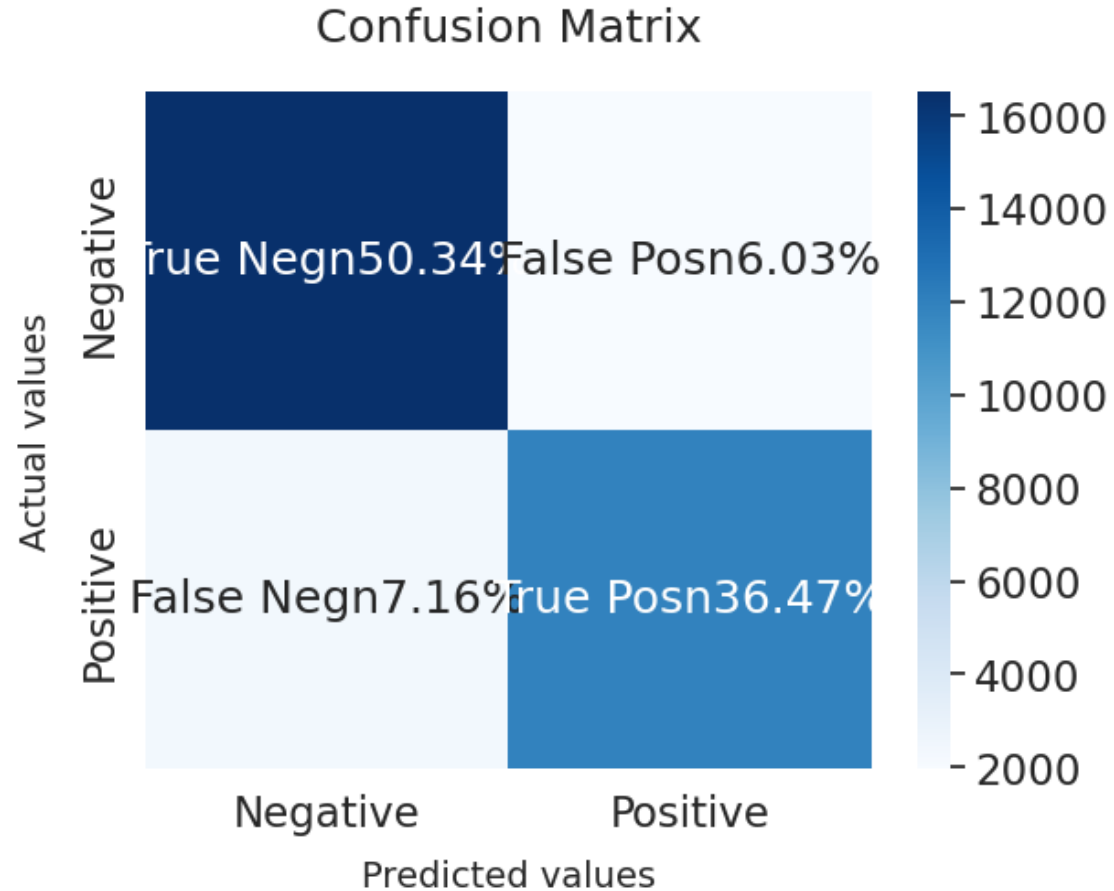
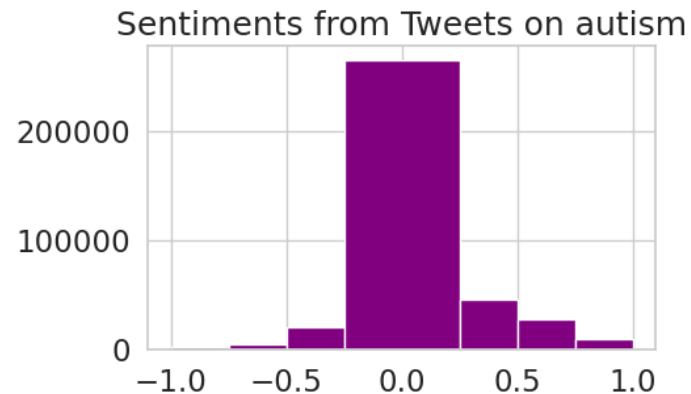
HINAMAUKA
Giving Hope for Recovery



Social Media Data Mining



Autism Analysis for Public Health



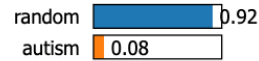
Manuscript in preparation.

Autism Analysis for Public Health

Manuscript in preparation.

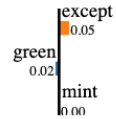
except green mint
True label: random

Prediction probabilities



random

autism



Text with highlighted words

except green mint

actual think slow bit today sparklingicesparklingicecontest
True label: random

Prediction probabilities



random

autism

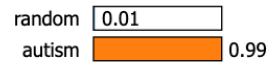


Text with highlighted words

actual think slow bit today sparklingicesparklingicecontest

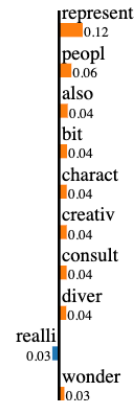
realli appreci diver represent charact goodtroubl realli wonder got consult peopl also anoth white ci male charact realli think couldv bit creativ
True label: autism

Prediction probabilities



random

autism

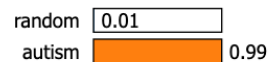


Text with highlighted words

realli appreci diver represent charact goodtroubl realli wonder got consult peopl also anoth white ci male charact realli think couldv bit creativ

attend mental health awar cours today nervou look forward mentalhealthmatt
True label: autism

Prediction probabilities



random

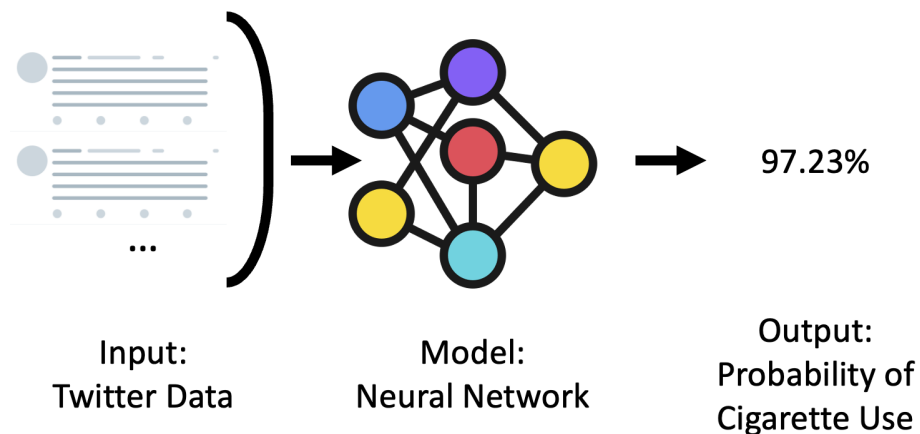
autism



Text with highlighted words

attend mental health awar cours today nervou look forward mentalhealthmatt

Nicotine Addiction Analysis on Social Media



UNIVERSITY OF HAWAI'I
CANCER CENTER



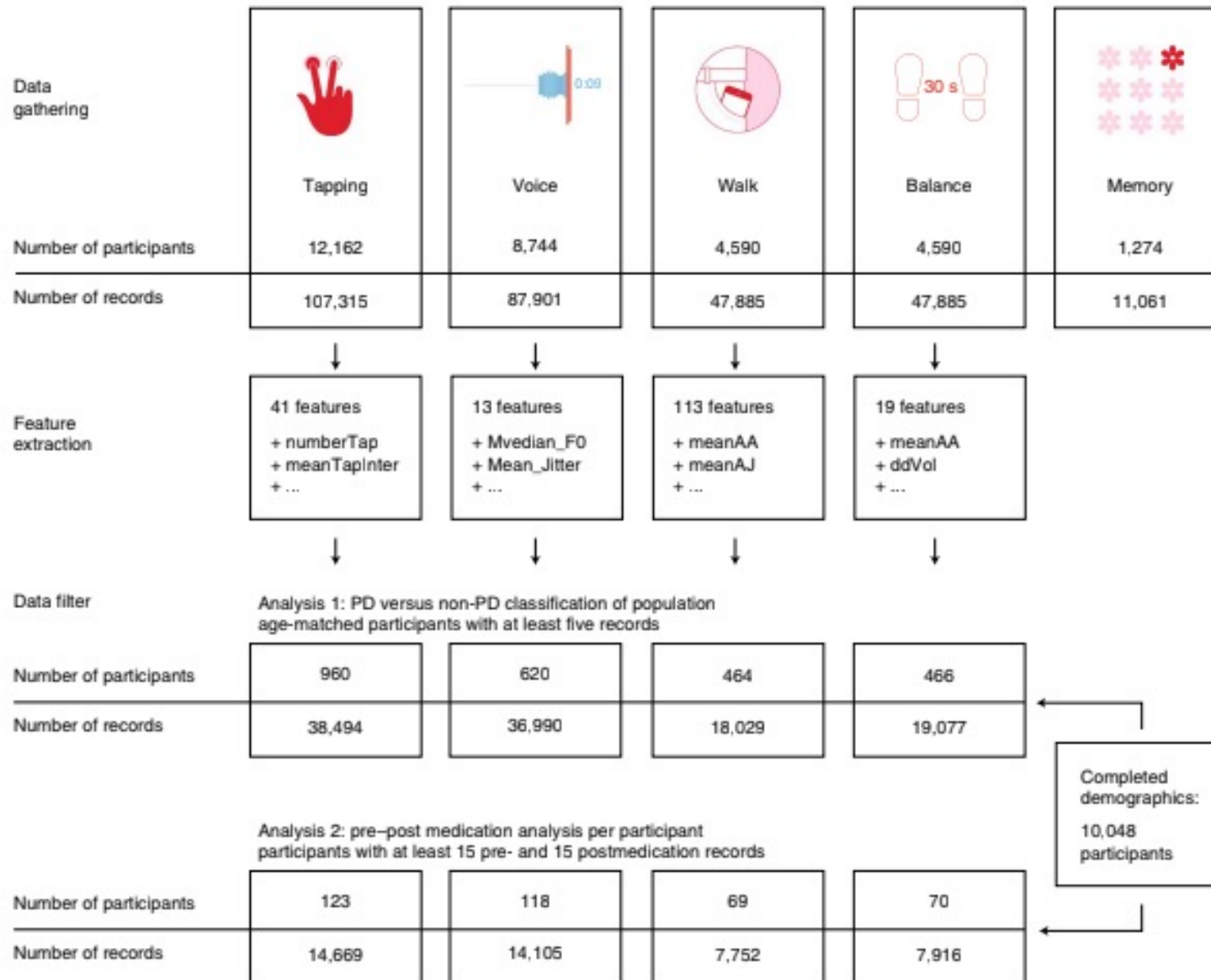
Dr. Pokhrel



Dr. Amin

Manuscript in preparation.

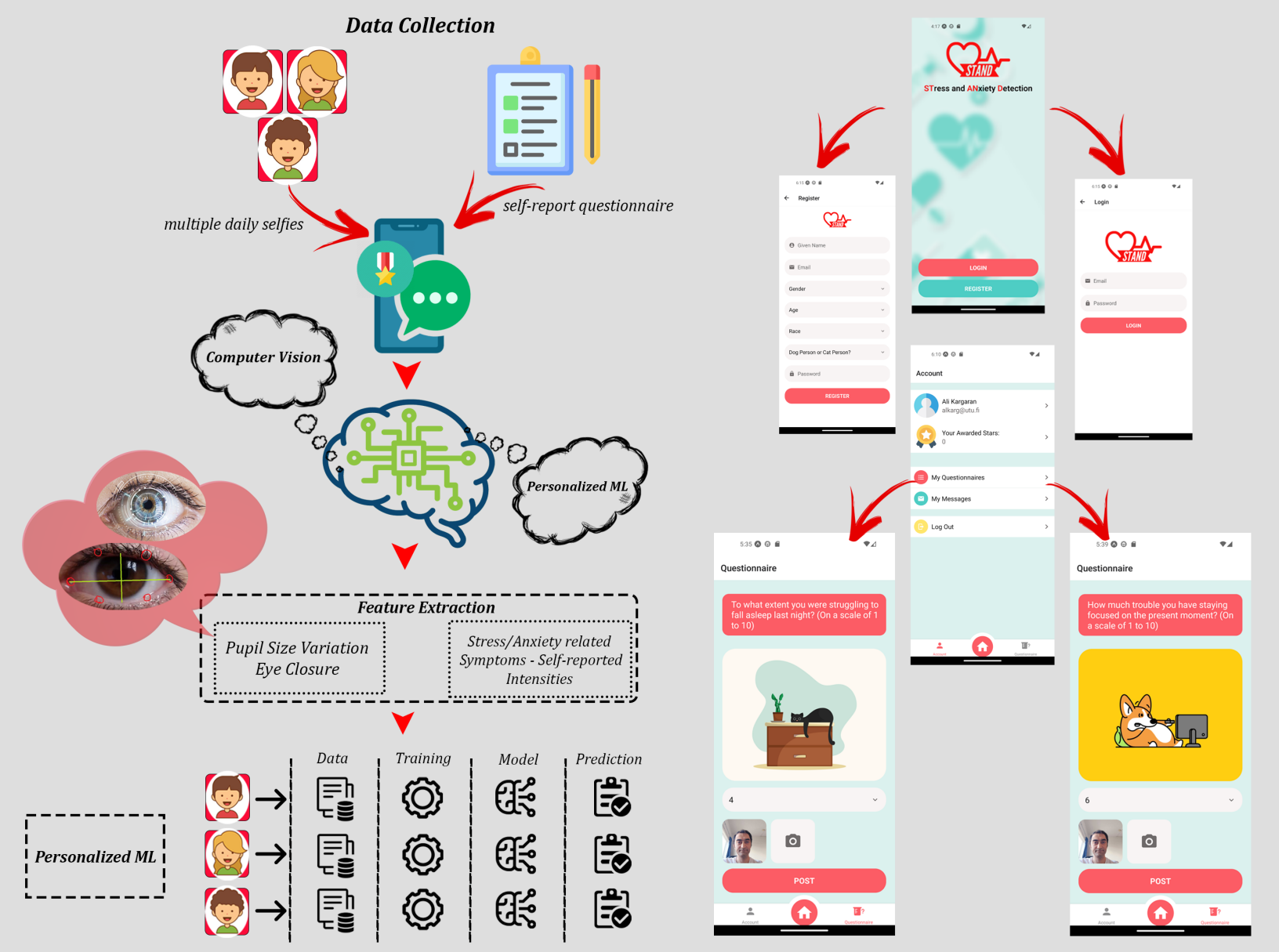
Alzheimer's and Parkinson's Screening



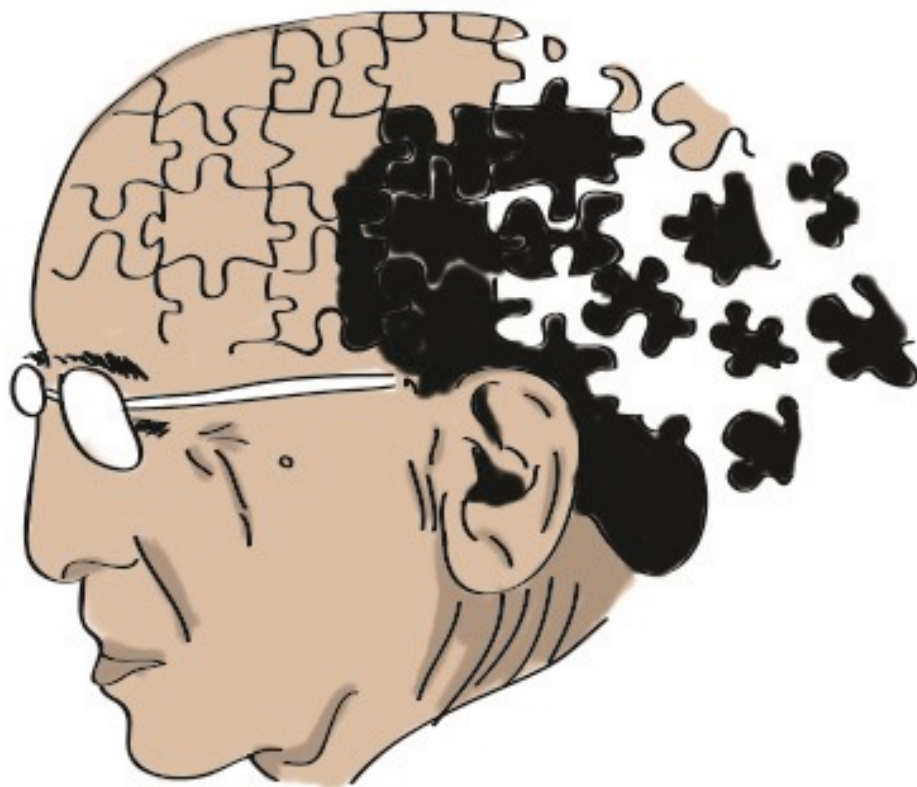
Jerry Boster



Personalized Mood Analysis



Classification of Dementia using NLP

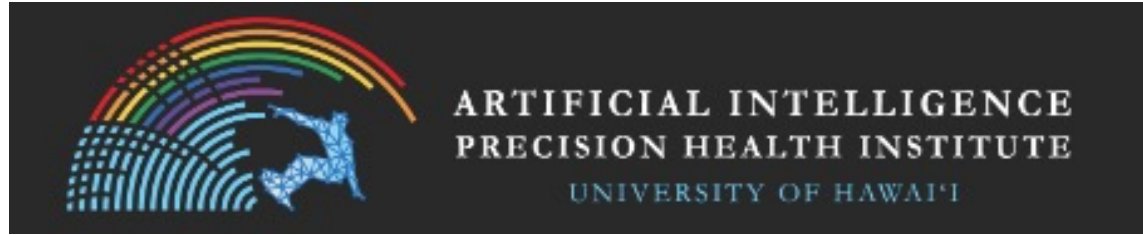


TalkBank



DementiaBank

Cancer Diagnostics and Screening



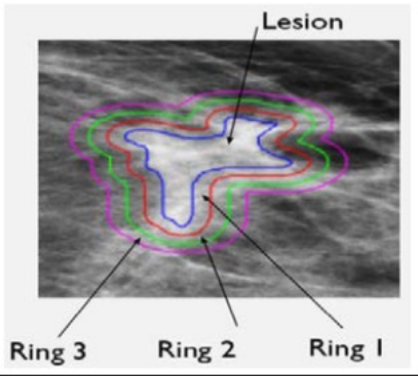
Dr. Shepherd



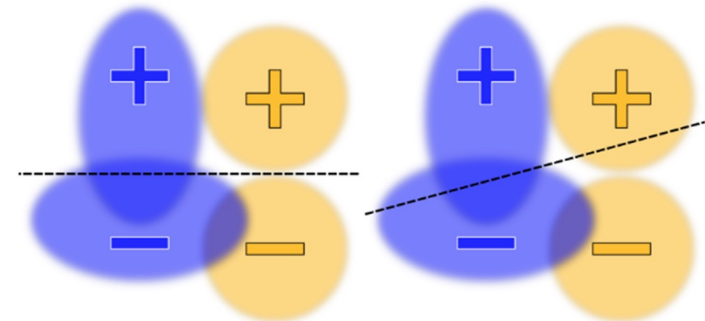
Dr. Sadowski

Research

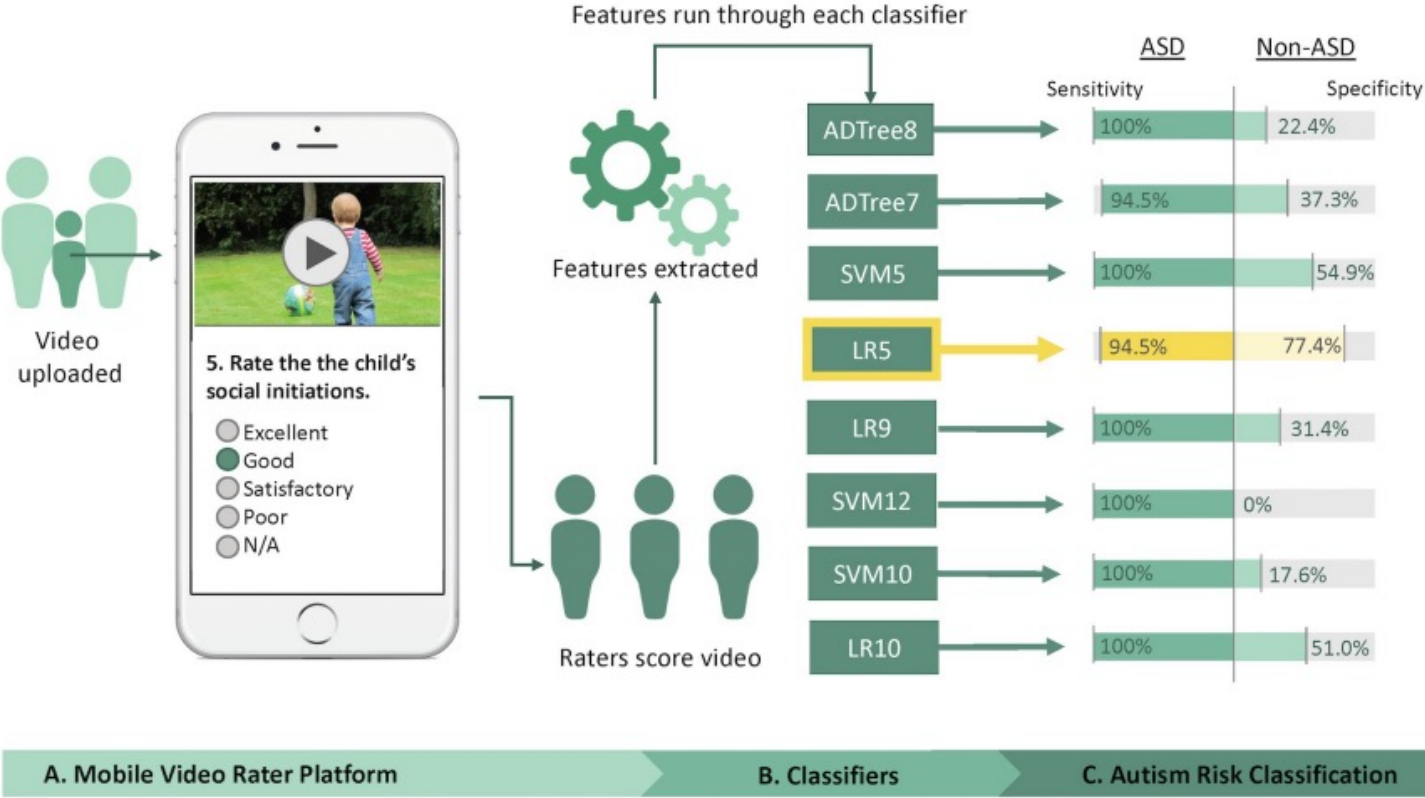
Our studies use artificial intelligence to improve and quantify the use of digital imaging in the diagnosis of diseases.



3CB non-overlapping ROIs: Lesion (blue), Ring 1 (red), Ring 2 (green), Ring 3 (purple). Each ring is 2mm in width. (2020). Leong L, Giger M, Drukker K, Kerlikowske K, Joe B, Greenwood H, et al. *Three compartment breast machine learning model for improving computer-aided detection.*



Digital Autism Diagnostics



Dr. Wall

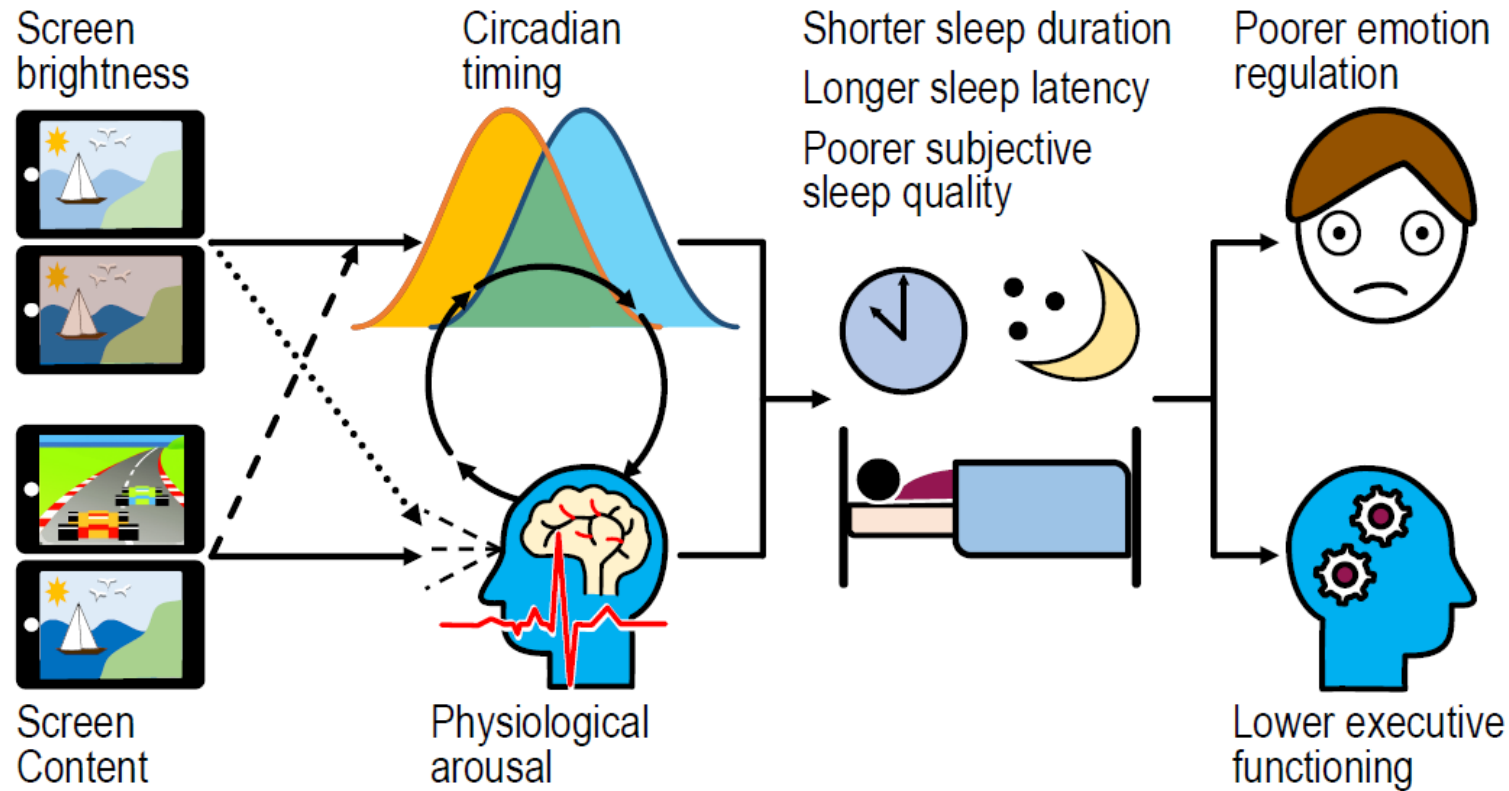


BMIR

Stanford Center for
Biomedical Informatics Research

WE CONNECT DATA TO HEALTH

Pediatric Sleep Behavior Analysis



Dr. Moreno



Dr. O'Connor



Machine Learning on NIH Datasets



The future of health begins with you.

The *All of Us* Research Program is inviting one million people across the U.S. to help build one of the most diverse health databases in history. We welcome participants from all backgrounds. Researchers will use the data to learn how our biology, lifestyle, and environment affect health. This may one day help them find ways to treat and prevent disease.

Machine Learning on Hawaii-Specific Datasets



The Multiethnic Cohort Study

Understanding ethnic differences in cancer to prevent it in all populations

BREAST HEALTH QUESTIONNAIRE STANDARDIZATION

This questionnaire aims to standardize the breast health information being collected at mammography clinics across Hawai'i and the US Affiliated Pacific Islands (USAPI). Many imaging sites are collecting inadequate information to accurately...

Opportunities to Work with Me

Undergraduate Students

- ICS 499 credits
- UROP funding (a few thousand dollars per semester)
- Honors thesis

Graduate Students

- ICS 699/700 credits (MS thesis Plan A)
- PhD advisorship
- If all goes well after 1 year of working for credits, there is a possibility of RA/GAships (limited spots; pending grant funding)



HAWAI'I DIGITAL HEALTHCARE LAB

peterwashington.github.io

Peter Yiğitcan Washington

Assistant Professor

Information & Computer Sciences

University of Hawai'i at Mānoa

pyw@hawaii.edu

Outline for Today

Putting It All Together: How ChatGPT Works

Quick Advertisement: Machine Learning for Healthcare Research

Final Exam Review

Course Review

Which of the following algorithms is used for unsupervised learning?

A. K-Nearest Neighbors

B. Support Vector Machines

C. Random Forest

D. K-Means Clustering

Which of the following algorithms is used for unsupervised learning?

A. K-Nearest Neighbors

B. Support Vector Machines

C. Random Forest

D. K-Means Clustering

Which of the following best describes self-attention in the context of deep learning?

- A. A technique that allows a model to focus on different parts of an input sequence
- B. A regularization technique that encourages the weights of a model to be small
- C. A technique for reducing the dimensionality of a dataset by projecting it onto a lower-dimensional space

Which of the following best describes self-attention in the context of deep learning?

- A. A technique that allows a model to focus on different parts of an input sequence**
- B. A regularization technique that encourages the weights of a model to be small
- C. A technique for reducing the dimensionality of a dataset by projecting it onto a lower-dimensional space

Which of the following best describes the role of the Q-network in deep Q-learning?

- A. The Q-network learns the optimal policy for the agent based on the current state of the environment
- B. The Q-network approximates the expected reward of taking a particular action in a particular state
- C. The Q-network is used to generate trajectories of states and actions for the agent to follow

Which of the following best describes the role of the Q-network in deep Q-learning?

- A. The Q-network learns the optimal policy for the agent based on the current state of the environment
- B. The Q-network approximates the expected reward of taking a particular action in a particular state**
- C. The Q-network is used to generate trajectories of states and actions for the agent to follow

Which of the following tradeoffs between types of neural network architectures is true?

- A. MLPs are best suited for image recognition tasks, while CNNs are best suited for natural language processing tasks
- B. RNNs are best suited for sequence modeling tasks, while Transformers are best suited for image recognition tasks
- C. CNNs are best suited for tasks that involve local spatial relationships, while RNNs are best suited for tasks that involve temporal dependencies
- D. Transformers are best suited for tasks that involve structured data, while MLPs are best suited for tasks that involve unstructured data

Which of the following tradeoffs between types of neural network architectures is true?

- A. MLPs are best suited for image recognition tasks, while CNNs are best suited for natural language processing tasks
- B. RNNs are best suited for sequence modeling tasks, while Transformers are best suited for image recognition tasks
- C. CNNs are best suited for tasks that involve local spatial relationships, while RNNs are best suited for tasks that involve temporal dependencies**
- D. Transformers are best suited for tasks that involve structured data, while MLPs are best suited for tasks that involve unstructured data

Suppose you are training a neural network for a classification task with a large dataset. During training, you notice that the model is overfitting to the training data, as the training accuracy is much higher than the validation accuracy. Which of the following is a potential solution to reduce overfitting?

- A. Increase the number of epochs for training.
- B. Decrease the batch size for training.
- C. Use data augmentation techniques to generate more training data.
- D. All of the above are possible solutions.

Suppose you are training a neural network for a classification task with a large dataset. During training, you notice that the model is overfitting to the training data, as the training accuracy is much higher than the validation accuracy. Which of the following is a potential solution to reduce overfitting?

- A. Increase the number of epochs for training.
- B. Decrease the batch size for training.
- C. Use data augmentation techniques to generate more training data.**
- D. All of the above are possible solutions.

Suppose you are training a deep neural network on a large dataset. After several training epochs, you notice that the network's training loss continues to decrease, but the validation loss plateaus and then starts to increase. Which of the following is a possible explanation for this behavior?

- A. The network is underfitting the data, and adding more layers or neurons may improve performance.
- B. The network is overfitting the data, and reducing the model complexity or applying regularization may improve performance.
- C. The learning rate is too high, and reducing it may improve performance.

Suppose you are training a deep neural network on a large dataset. After several training epochs, you notice that the network's training loss continues to decrease, but the validation loss plateaus and then starts to increase. Which of the following is a possible explanation for this behavior?

A. The network is underfitting the data, and adding more layers or neurons may improve performance.

B. The network is overfitting the data, and reducing the model complexity or applying regularization may improve performance.

C. The learning rate is too high, and reducing it may improve performance.

Suppose we have a $1000 \times 1000 \times 3$ dimension input image (width x height x channel). We apply a convolutional layer with 50 5×5 kernels. What is the dimension of the resulting tensor (width x height x channel) if we have $\text{stride}=2$ and no padding?

- A. $499 \times 499 \times 50$
- B. $498 \times 498 \times 50$**
- C. $500 \times 500 \times 50$
- D. $502 \times 502 \times 50$

Outline for Today

Putting It All Together: How ChatGPT Works

Quick Advertisement: Machine Learning for Healthcare Research

Final Exam Review

Course Review

Core Ingredients of Machine Learning

- How will you represent your input data / **features**?
- What are you trying to optimize for / what is your **loss function**?
- What type of **model** are you going to use?
- How will you **evaluate** your model for success?

Feature Representations

- Dimensionality Reduction
- Feature Representation Engineering
- Feature Selection

Loss Functions

- Mean Square Error (MSE) → Regression
- Cross Entropy → Classification
- Custom loss functions for new applications (reinforcement learning, generative models, representation learning, etc)

Models

- Classical ML models
 - Simple data (tabular data / CSV files)

- Deep learning
 - Complex data (images, video, text, sound, etc)

Evaluation

- MAE / MSE / RMSE → Regression
- Accuracy, Precision, Recall, Specificity, AUROC, F1-Score, ... → Classification
- Specialized evaluation metrics for specialized applications

ML is an **art** and a **science**

Applications of ML

2020

Healthcare



Finance & Insurance



Transportation



Construction



Retail & Warehousing



Govt. & City Planning



Media & Entertainment



Education



Manufacturing



Legal



Mining



CBINSIGHTS

Energy



Telecom



Real Estate



CROSS-INDUSTRY TECH

AI Processors



NLP, NLG, & Computer Vision



Sales & CRM



AI Model Development



Cybersecurity



BI & Ops Intel



Other R&D



DevOps & Model Monitoring



Go out there and change the world



Congratulations!

ICS/DATA 435 and ICS 635

Spring 2023