BioWatch README – Deep Learning Week ‘25

By Team 十三幺:
Roderick Kong | Brandon Kang | Tong Jing Yen | Chua Feng Yuan | Kingsley Ser

Useful Links

Resource	Link
GitHub	GitHub Repo
Google Drive	DLW '25
Canva	DLW '25 Pitch Deck
Jupyter Notebook	ML Testing Code Gemini Wrapper

Table of Contents

1 | Hackathon Problem Statement

Track 1: Healthcare

Healthcare is shifting from reactive treatment to proactive, AI-driven prevention. Participants are challenged to develop **AI-powered solutions that for example, integrate remote health monitoring, mental well-being support, or lifestyle management.** Your solution can leverage wearable or mobile technology to track real-time health data, detect patterns, and provide intelligent alerts for early intervention. It may also incorporate AI-driven mental health support, offering personalized insights, virtual assistance, or community-driven engagement. Your system could promote healthier lifestyle choices by analyzing behaviors and recommending tailored interventions to prevent chronic diseases.

Problem Statement

Build an **AI-centric**, **seamless**, and **adaptive** solution that **empowers individuals** and **transforms healthcare accessibility**.

- **AI-centric:** be able to learn from existing datasets about indicators of health issues
- **Seamless:** able to connect between various interfaces to obtain data from user
- **Adaptive:** sufficiently reactive to manual inputs to give an update to current condition, without being overly sensitive
- **Individual empowerment:** enhance and increase user ability
- **Healthcare accessibility:** ability to obtain timely medical assistance
- **Chronic diseases:** long-term health issues (e.g. diabetes, kidney failure, high blood cholesterol)

2 | Team Problem Statement

Problem Statement

Nowadays, wearable technology has become increasingly accessible and people are having easier and cheaper access to their body vitals and metrics, through devices such as smartwatches and fitness trackers. We can leverage this increased accessibility of IoT wearable devices, and use the easily accessible data to perform data analysis for users' health.

Currently, users do not have access to early detection and warnings as well as preventive measures for diseases. Usually, when patients find out about their illnesses, it is already too late. Frequent health screenings can solve this issue, but they are expensive and time-consuming, and are usually only done on a yearly basis.

Proposed Solution

Leveraging on the myriad of data that smartwatches provide to users, we aim to create an iOS app with an accompanying watchOS app that provides data analysis for early disease prediction. The app will display **early warnings to users and preventive measures** they can take to combat these illnesses. Deep learning algorithms will be used to analyse past patient data and create a model that associates anomalous patient vitals with early disease symptoms. This model will then be run on real-time user data, analysing their historical body vitals and comparing it with their current body vitals, to determine whether there are anomalies corresponding to specific diseases.

3 | Feature List

✓	No.	Feature	Use Case	Extension	✓
☑	1	Working predictive model for various diseases (diabetes, heart disease, etc.) – Polynomial Ridge Regression with outliers removal and normalisation	Early detection for diseases based on real-time wearable data	Implement Decision Tree/Random Forest Model Implement Artificial Neural Networks/Convolutional Neural Networks (Deep Learning) Allow user to select the model to use for early detection Analyse user historical data and detect for anomalies based on personalised data	☑
☑	2	Intuitive and interactive app interface that allows user to view diseases detected early	For the user's easy reference	Add vitals in a clear and readable format	☑
☑	3	Allow user input for specialised features	For the user to input specialised features that cannot be automatically tracked	Make input process more seamless and aesthetically pleasing	☐
☐	4	Send push notifications to user when a disease is detected early	For the user to receive timely notifications when diseases are detected	-	☐
☑	5	Add a Gemini wrapper that provides the user with suggestions to combat the diseases early, and to allow the user to chat with a chatbot for more information	To provide the user with suggestions to combat detected diseases early	-	☐

☐	6	Create a companion watch app	For the user to view information on-the-go	-	☐

4 | Model Performance

Dataset	Regression	Random Forest	Deep Learning
Diabetes	**91.6% (Order 2 Ridge) 92.3% (Logistics)**	**90.4% - 91.8%**	**92.7%**
Stroke	**97.7%**	**95.1%**	**100%**
Heart Disease	**83%**	**85%**	**50%**
Sleep Disorder	**86%**	**87%**	**85%**

5 | Available Datasets

Dataset	Link
UC Irvine Machine Learning Repository	**https://archive.ics.uci.edu/**
Heart Disease	https://archive.ics.uci.edu/dataset/45/heart+disease
CDC Diabetes Health Indicators Diabetes 130-US Hospitals for Years 1999-2008	https://archive.ics.uci.edu/dataset/891/cdc+diabetes+health+indicators https://archive.ics.uci.edu/dataset/296/diabetes+130-us+hospitals+for+years+1999-2008
CMU Common Cold Study	British Cold Study (BCS): http://dx.doi.org/10.1184/RDL/1502014 Pittsburgh Cold Study 1 Data Set: http://dx.doi.org/10.1184/RDL/140904 Pittsburgh Cold Study 2 Data Set: http://dx.doi.org/10.1184/RDL/1502012 Pittsburgh Mind-Body Center Cold Study Data Set: http://dx.doi.org/10.1184/RDL/1502011 Pittsburgh Cold Study 3 Data Set: http://dx.doi.org/10.1184/RDL/1509015 Aggregated Cold Study Data Set (1-5): http://dx.doi.org/10.1184/RDL/1502013
National Poll on Healthy Aging (NPHA)	https://archive.ics.uci.edu/dataset/936/national+poll+on+healthy+aging+(npha)
Kaggle	**https://www.kaggle.com/**
Stroke Prediction Dataset	https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset
Sleep Health and Lifestyle Dataset	https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset
Dengue Prediction (Supervised)	https://www.kaggle.com/datasets/siddhvr/dengue-prediction
FitBit Fitness Tracker Data (For Simulated User Input)	https://www.kaggle.com/datasets/arashnic/fitbit

6 | Dataset Features

6.1 | Apple Health Patient Metrics

- **Basic Metrics**
    - Age
    - Sex

- **Activity**
    - Resting Energy
    - Active Energy
    - Physical Effort
    - Time in Daylight
    - Workouts
    - Exercise Minutes
    - Stand Minutes
    - Stand Hours
    - Steps
    - Walking + Running Distance
    - Walking Heart Rate Average
    - Flights Climbed

- **Body Measurements**
    - Weight
    - Height
    - BMI
    - Body Fat Percentage
    - Lean Body Mass

- **Hearing**
    - Environmental Sound Levels
    - Headphone Audio Levels
    - Environmental Sound Reduction
    - Noise Notifications

- **Heart**
    - Current Heart Rate
    - Walking Heart Rate Average
    - Resting Heart Rate
    - Heart Rate Variability

- Cardio Fitness
- Cardio Recovery
- High Heart Rate Notifications

- **Sleep**
  - Sleep Hours

6.2 | Available Dataset Features

6.2.1 | Heart Disease

Variables Table

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
age	Feature	Integer	Age		years	no
sex	Feature	Categorical	Sex			no
cp	Feature	Categorical				no
trestbps	Feature	Integer		resting blood pressure (on admission to the hospital)	mm Hg	no
chol	Feature	Integer		serum cholestoral	mg/dl	no
fbs	Feature	Categorical		fasting blood sugar > 120 mg/dl		no
restecg	Feature	Categorical				no
thalach	Feature	Integer		maximum heart rate achieved		no
exang	Feature	Categorical		exercise induced angina		no
oldpeak	Feature	Integer		ST depression induced by exercise relative to rest		no
slope	Feature	Categorical				no
ca	Feature	Integer		number of major vessels (0-3) colored by flourosopy		yes
thal	Feature	Categorical				yes
num	Target	Integer		diagnosis of heart disease		no

6.2.2 | Diabetes

Variables Table

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
ID	ID	Integer		Patient ID		no
Diabetes_binary	Target	Binary		0 = no diabetes 1 = prediabetes or diabetes		no
HighBP	Feature	Binary		0 = no high BP 1 = high BP		no
HighChol	Feature	Binary		0 = no high cholesterol 1 = high cholesterol		no
CholCheck	Feature	Binary		0 = no cholesterol check in 5 years 1 = yes cholesterol check in 5 years		no
BMI	Feature	Integer		Body Mass Index		no
Smoker	Feature	Binary		Have you smoked at least 100 cigarettes in your entire life? [Note: 5 packs = 100 cigarettes] 0 = no 1 = yes		no
Stroke	Feature	Binary		(Ever told) you had a stroke. 0 = no 1 = yes		no
HeartDiseaseorAttack	Feature	Binary		coronary heart disease (CHD) or myocardial infarction (MI) 0 = no 1 = yes		no
PhysActivity	Feature	Binary		physical activity in past 30 days - not including job 0 = no 1 = yes		no

Variables Table

Variable Name	Role	Type	Demographic	Description	Units	Missing Values
Fruits	Feature	Binary		Consume Fruit 1 or more times per day 0 = no 1 = yes		no
Veggies	Feature	Binary		Consume Vegetables 1 or more times per day 0 = no 1 = yes		no
HvyAlcoholConsump	Feature	Binary		Heavy drinkers (adult men having more than 14 drinks per week and adult women having more than 7 drinks per week) 0 = no 1 = yes		no
AnyHealthcare	Feature	Binary		Have any kind of health care coverage, including health insurance, prepaid plans such as HMO, etc. 0 = no 1 = yes		no
NoDocbcCost	Feature	Binary		Was there a time in the past 12 months when you needed to see a doctor but could not because of cost? 0 = no 1 = yes		no
GenHlth	Feature	Integer		Would you say that in general your health is: scale 1-5 1 = excellent 2 = very good 3 = good 4 = fair 5 = poor		no
MentHlth	Feature	Integer		Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good? scale 1-30 days		no
PhysHlth	Feature	Integer		Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good? scale 1-30 days		no
DiffWalk	Feature	Binary		Do you have serious difficulty walking or climbing stairs? 0 = no 1 = yes		no
Sex	Feature	Binary	Sex	0 = female 1 = male		no
Age	Feature	Integer	Age	13-level age category (_AGEG5YR see codebook) 1 = 18-24 9 = 60-64 13 = 80 or older		no
Education	Feature	Integer	Education Level	Education level (EDUCA see codebook) scale 1-6 1 = Never attended school or only kindergarten 2 = Grades 1 through 8 (Elementary) 3 = Grades 9 through 11 (Some high school) 4 = Grade 12 or GED (High school graduate) 5 = College 1 year to 3 years (Some college or technical school) 6 = College 4 years or more (College graduate)		no
Income	Feature	Integer	Income	Income scale (INCOME2 see codebook) scale 1-8 1 = less than $10,000 5 = less than $35,000 8 = $75,000 or more		no

Appendix

Chronic diseases are long-term medical conditions that often progress slowly and require ongoing management. Here are the major categories and examples of chronic diseases:

1. Cardiovascular Diseases
    - Hypertension (High Blood Pressure)
    - Coronary Artery Disease (CAD)
    - Heart Failure
    - Stroke
    - Peripheral Artery Disease (PAD)
2. Metabolic Disorders
    - Diabetes (Type 1 & Type 2)
    - Obesity
    - Metabolic Syndrome
3. Respiratory Diseases
    - Chronic Obstructive Pulmonary Disease (COPD)
    - Asthma
    - Cystic Fibrosis
    - Pulmonary Hypertension
4. Neurological Disorders
    - Alzheimer's Disease
    - Parkinson's Disease
    - Epilepsy
    - Multiple Sclerosis (MS)
    - Migraine (Chronic)
5. Autoimmune Diseases
    - Rheumatoid Arthritis (RA)
    - Lupus (Systemic Lupus Erythematosus - SLE)
    - Psoriasis/Psoriatic Arthritis
    - Inflammatory Bowel Disease (IBD) (Crohn's Disease, Ulcerative Colitis)
    - Type 1 Diabetes
6. Musculoskeletal Disorders
    - Osteoarthritis
    - Osteoporosis
    - Fibromyalgia
7. Kidney Diseases
    - Chronic Kidney Disease (CKD)
    - Polycystic Kidney Disease

8. Liver Diseases
    - Chronic Hepatitis (B & C)
    - Cirrhosis
    - Fatty Liver Disease (NAFLD/NASH)
9. Endocrine Disorders
    - Hypothyroidism
    - Hyperthyroidism
    - Cushing's Syndrome
    - Addison's Disease
10. Mental Health Disorders (Chronic Psychiatric Conditions)
    - Depression (Chronic & Recurrent)
    - Bipolar Disorder
    - Schizophrenia
    - Generalized Anxiety Disorder (GAD)
    - Obsessive-Compulsive Disorder (OCD)
11. Cancer (Various Types)
    - Breast Cancer
    - Lung Cancer
    - Prostate Cancer
    - Colorectal Cancer
    - Leukemia
    - Lymphoma
12. Skin Diseases
    - Eczema (Atopic Dermatitis)
    - Psoriasis
    - Vitiligo
13. Infectious Diseases (Chronic)
    - HIV/AIDS
    - Tuberculosis (Latent or Active Chronic TB)
    - Hepatitis B & C
14. Gastrointestinal Diseases
    - Gastroesophageal Reflux Disease (GERD)
    - Peptic Ulcer Disease (Chronic)
    - Irritable Bowel Syndrome (IBS)
15. Genetic Disorders
    - Sickle Cell Disease
    - Cystic Fibrosis
    - Huntington's Disease

16. Chronic Pain Conditions
    - Chronic Fatigue Syndrome (CFS/ME)
    - Neuropathy (Diabetic, Peripheral)
    - Trigeminal Neuralgia

1. Activity & Fitness Data

Data Type	Description	Identifier
Steps	Number of steps taken	`.stepCount`
Walking + Running Distance	Distance covered in meters/kilometers	`.distanceWalkingRunning`
Cycling Distance	Distance covered in cycling activities	`.distanceCycling`
Flights Climbed	Number of stairs climbed	`.flightsClimbed`
Active Energy Burned	Calories burned through movement	`.activeEnergyBurned`
Basal Energy Burned	Calories burned at rest	`.basalEnergyBurned`
Move Minutes	Active time during the day	`.appleExerciseTime`
Exercise Minutes	Workout-specific exercise time	`.appleStandTime`
Workout Data	Includes type, duration, energy burned, and heart rate during workouts	`.workoutType()`

---

2. Heart & Cardiovascular Data

Data Type	Description	Identifier
Heart Rate	Measured in beats per minute (BPM)	`.heartRate`
Resting Heart Rate	Average BPM while at rest	`.restingHeartRate`
Walking Heart Rate Average	Heart rate while walking	`.walkingHeartRateAverage`

Heart Rate Variability (HRV)	Standard deviation of beat intervals (SDNN)	`.heartRateVariabilitySDNN`
VO2 Max	Maximum oxygen consumption during exercise	`.vo2Max`
Blood Oxygen (SpO2)	Oxygen saturation level (%)	`.oxygenSaturation`
Blood Pressure	Systolic & diastolic pressure readings	`.bloodPressureSystolic` & `.bloodPressureDiastolic`

---

3. Sleep Tracking

Data Type	Description	Identifier
Sleep Analysis	Sleep duration and phases (awake, REM, deep sleep)	`.sleepAnalysis`

---

4. Nutrition Data

Data Type	Description	Identifier
Dietary Calories	Total calorie intake	`.dietaryEnergyConsumed`
Water Intake	Amount of water consumed (ml)	`.dietaryWater`
Macronutrients	Carbs, protein, fats, fiber, sugar	`.dietaryCarbohydrates`, `.dietaryProtein`, `.dietaryFatTotal`
Micronutrients	Vitamins & minerals (Vitamin C, Calcium, Iron, etc.)	`.dietaryVitaminC`, `.dietaryCalcium`, `.dietaryIron`

---

5. Body Measurements

Data Type	Description	Identifier
Weight	Body weight in kg/lb	`.bodyMass`
Height	Body height in cm/m/inches	`.height`
Body Mass Index (BMI)	Calculated BMI	`.bodyMassIndex`
Body Fat Percentage	Percentage of body weight from fat	`.bodyFatPercentage`

---

6. Blood & Glucose Monitoring

Data Type	Description	Identifier
Blood Glucose	Blood sugar levels (mg/dL or mmol/L)	`.bloodGlucose`
Insulin Delivery	Insulin intake (manual entry or connected device)	`.insulinDelivery`

---

7. Respiratory Data

Data Type	Description	Identifier
Respiratory Rate	Number of breaths per minute	`.respiratoryRate`
Peak Expiratory Flow Rate (PEFR)	Maximum airflow during exhale	`.peakExpiratoryFlowRate`
Forced Vital Capacity (FVC)	Maximum air exhaled after full inhalation	`.forcedVitalCapacity`

---

8. Reproductive Health

Data Type	Description	Identifier
Menstrual Cycle Tracking	Period dates, symptoms, cycle length	`.menstrualFlow`, `.ovulationTestResult`
Ovulation Test Results	Logs ovulation test results	`.ovulationTestResult`

---

9. Mindfulness & Mental Health

Data Type	Description	Identifier
Mindful Sessions	Logs meditation & mindfulness activities	`.mindfulSession`
Mood Tracking (iOS 17+ Feature)	User's mood ratings	`.moodTracking`

---

10. Environmental & Other Data

Data Type	Description	Identifier
Noise Exposure	Measures environmental sound levels (dB)	`.environmentalAudioExposure`
Headphone Audio Exposure	Measures sound exposure via headphones	`.headphoneAudioExposure`
UV Exposure	Tracks UV radiation levels	`.uvExposure`