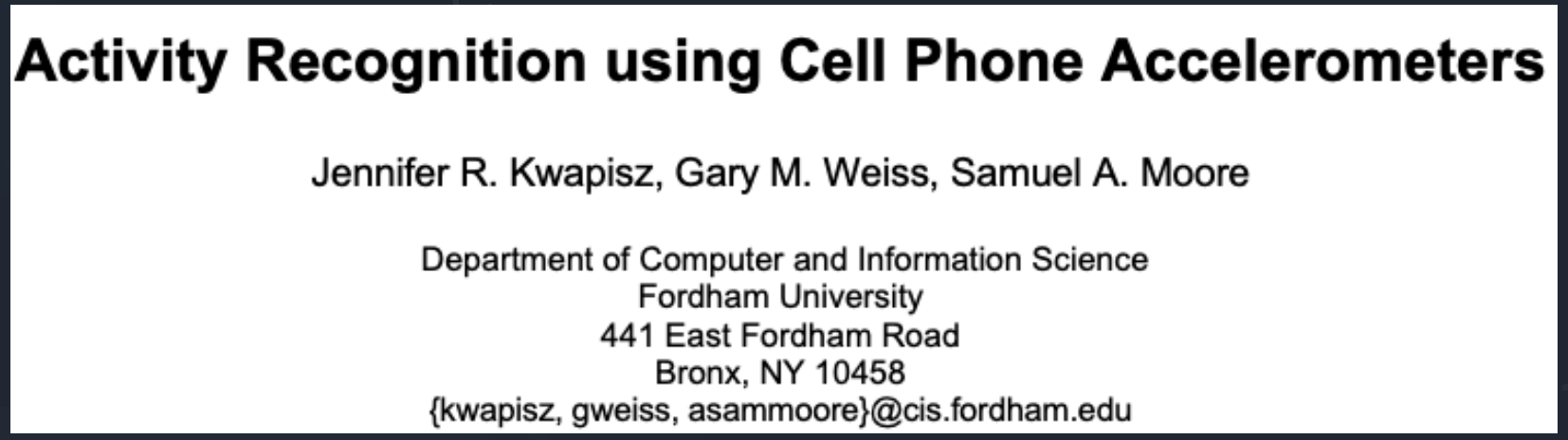
Physical Activity Classified by Tri-Axial Accelerometer

Dominique Barnes
Data Science Institute
October 18th, 2023
https://github.com/dbarnes16/DATA1030-Midterm.git

Introduction

Dataset

- Kaggle dataset: Activity-Detection-IMU-Sensor
- Wireless Sensor Data Mining (WISDM) Lab
  - Fordham University
- Data collected in a controlled setting



Activity Recognition using Cell Phone Accelerometers

Jennifer R. Kwapisz, Gary M. Weiss, Samuel A. Moore

Department of Computer and Information Science
Fordham University
441 East Fordham Road
Bronx, NY 10458
{kwapisz, gweiss, asammoore}@cis.fordham.edu

Introduction

Dataset

- 29 volunteer subjects
- Carried a phone in their front pocket
- Phone application recorded:
  - User's name
  - Start and stop of data collection
  - Activity label
  - Accelerometer data

Introduction

Goal

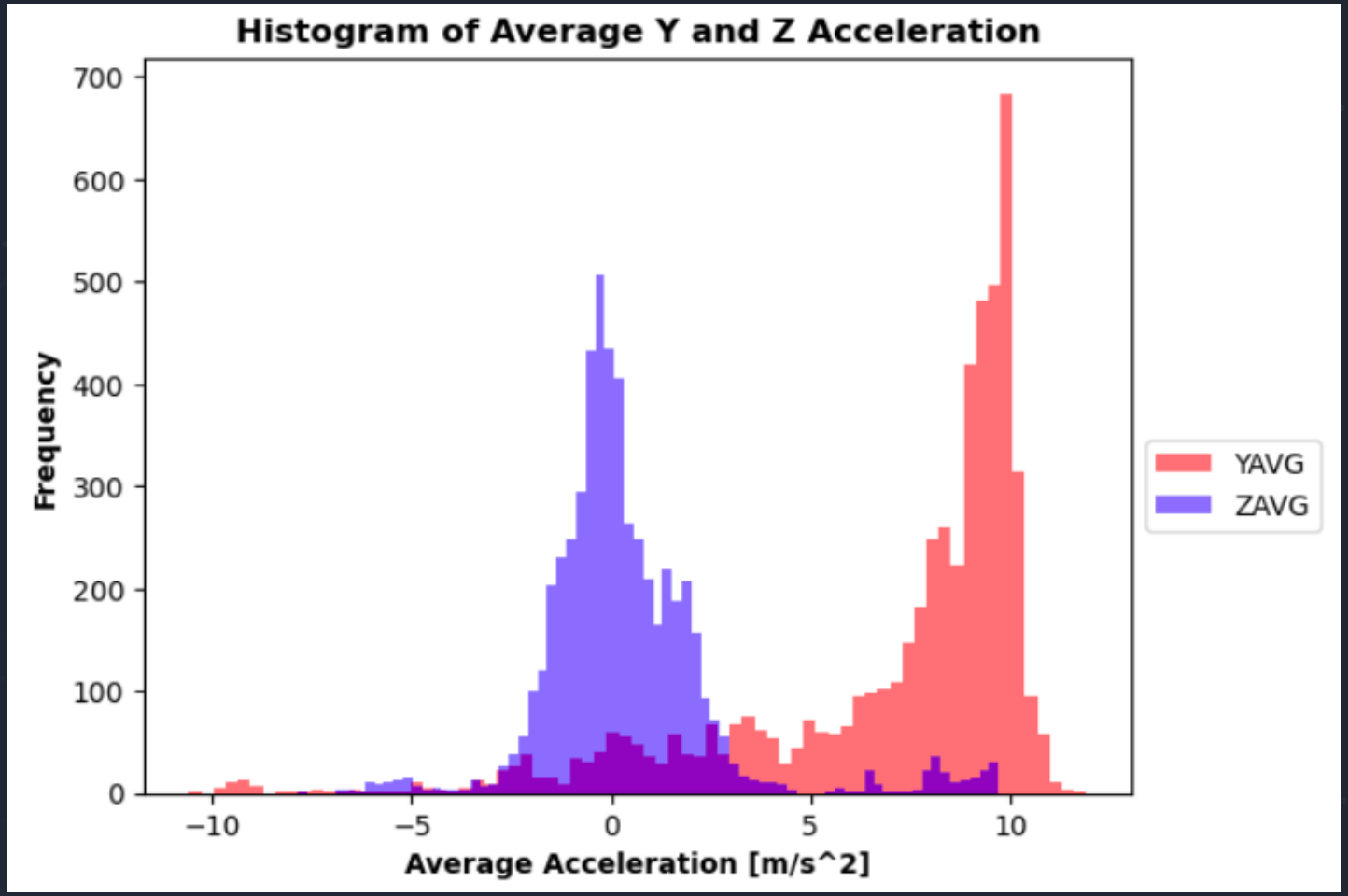
- **Classify** the user's motion by the phone-based accelerometers

Introduction
Classification

- **Target Variable**:
  - Physical activity class (categorical)

Exploratory Data Analysis

Histograms of Average Acceleration



Exploratory Data Analysis

Average Acceleration

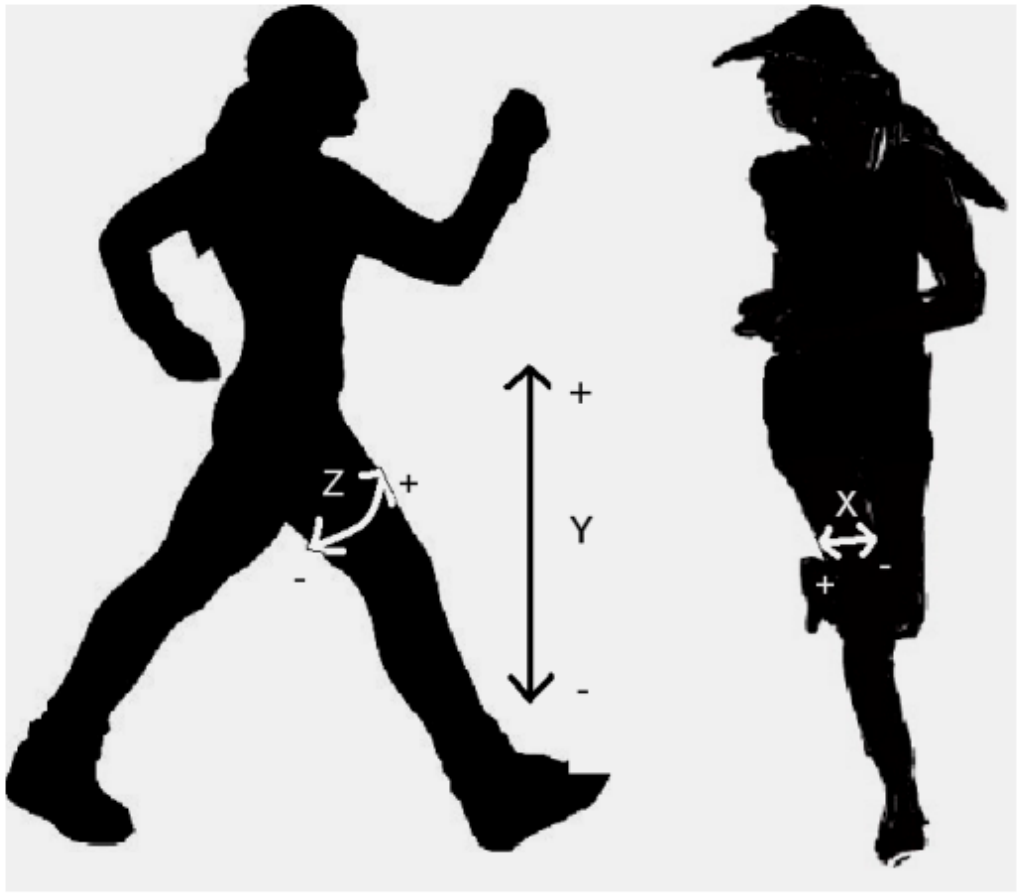
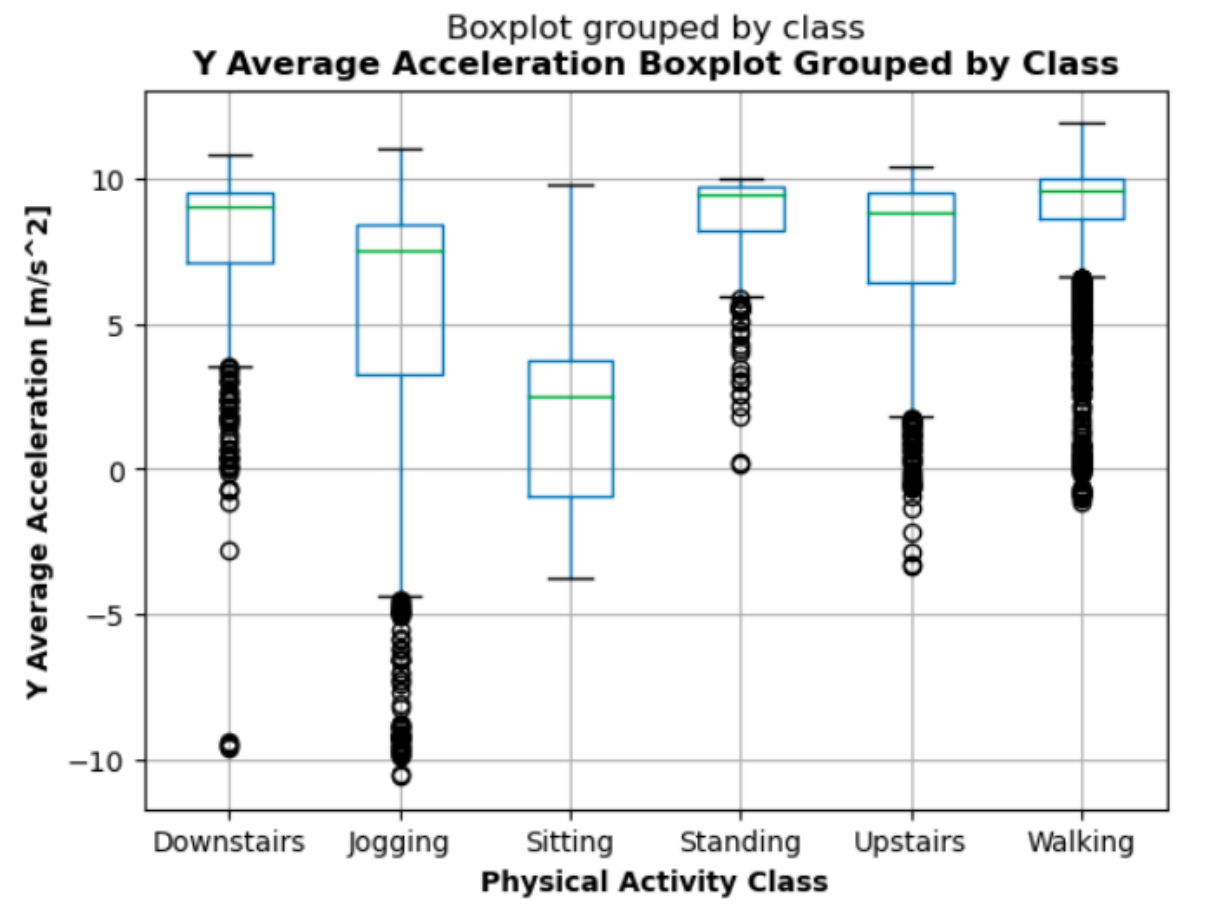


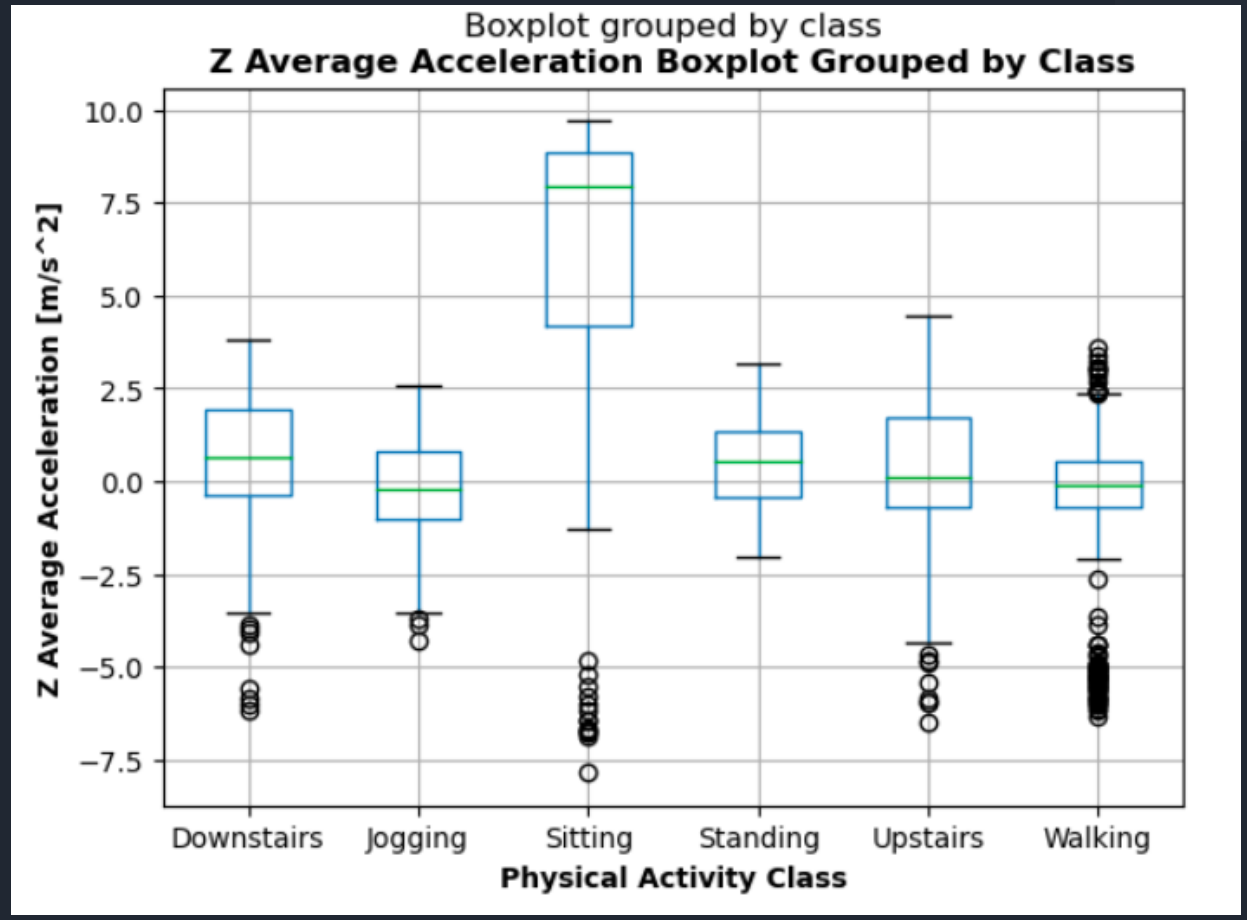
Figure 1: Axes of Motion Relative to User

- The Z-axis captures the forward and backward movement of the leg
- The Y-axis captures the upward and downward motion of the leg
- The X-axis captures the horizontal movement of the leg

Exploratory Data Analysis

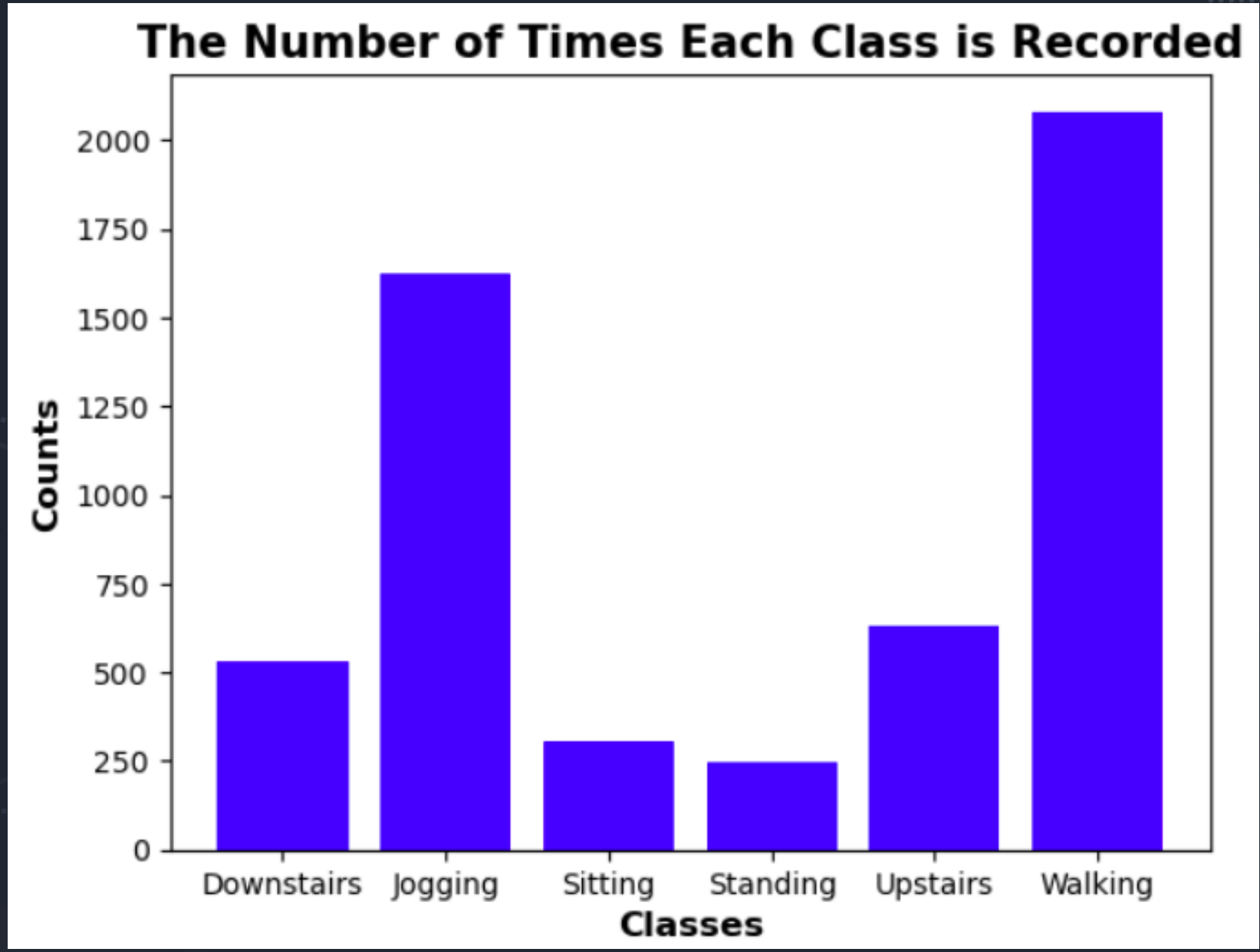
Y and Z Average Acceleration





Exploratory Data Analysis

Data Imbalance



Splitting

Is the data IID?

No. It is group-structured

What will the model predict?

Physical activity class on unseen **users**

Group KFold

Missing Values

- No imputation or column/row dropping was done for the feature matrix
- Fraction of points with missing values: 0.0875
- Fraction of missing values in features:

Feature	Missing Value Fraction
X PEAK	0.0703
Y PEAK	0.0242
Z PEAK	0.0190

Preprocessing

Standard Scaler

X0	float64
X1	float64
X2	float64
X3	float64
X4	float64
X5	float64
X6	float64
X7	float64
X8	float64
X9	float64
Y0	float64
Y1	float64
Y2	float64
Y3	float64
Y4	float64
Y5	float64
Y6	float64
Y7	float64
Y8	float64
Y9	float64
Z0	float64
Z1	float64
Z2	float64
Z3	float64
Z4	float64
Z5	float64

Z6	float64
Z7	float64
Z8	float64
Z9	float64
XAVG	int64
YAVG	float64
ZAVG	float64
XPEAK	float64
YPEAK	float64
ZPEAK	float64
XABSOLDEV	float64
YABSOLDEV	float64
ZABSOLDEV	float64
XSTANDDEV	float64
YSTANDDEV	float64
ZSTANDDEV	float64
RESULTANT	float64

- X_train.shape
  - (3260, 44)
- X_train_prep.shape
  - (3260, 44)

Questions?

Dominique Barnes

https://github.com/dbarnes16/DATA1030-Midterm.git