Human Activity Recognition from Accelerometer Data

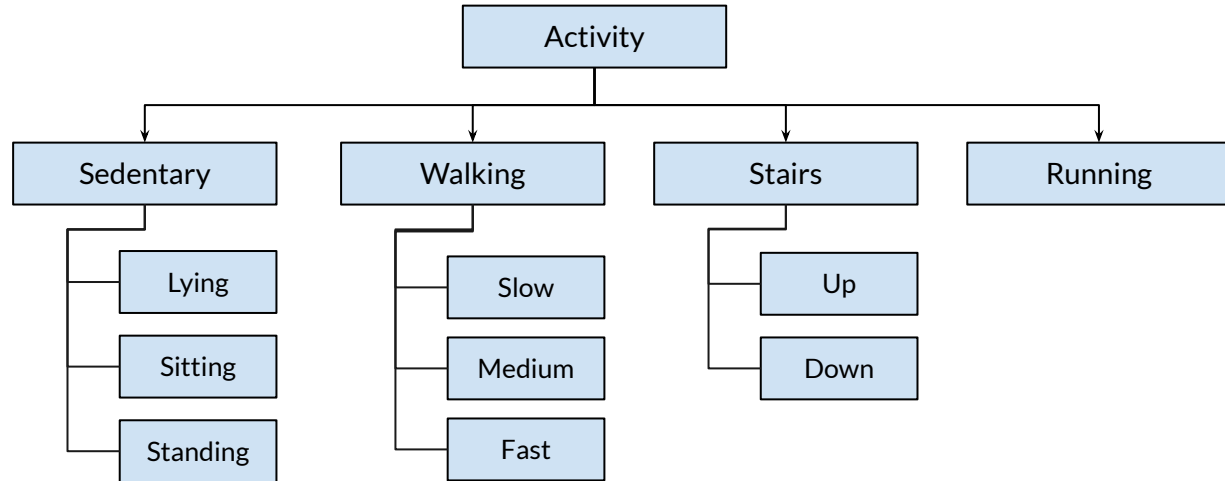Data Science Initiative, Brown University

Sagar Raichandani

GitHub

Project Goal

Objective

To create a subject independent classifier from time-series accelerometer data that identifies physical activity.
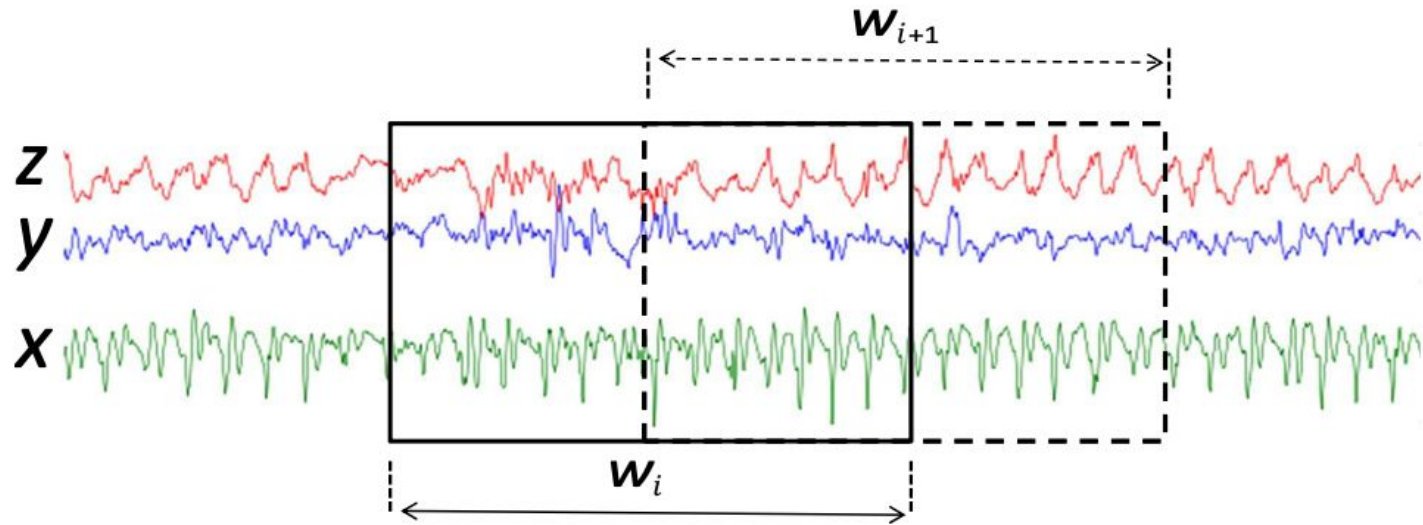
Data Description

| Feature	Description
Time	Time of signal measurement from tri-axial accelerometers.
x	Acceleration in g's along x-axis [-8g to +8g range]
y	Acceleration in g's along y-axis [-8g to +8g range]
z	Acceleration in g's along z-axis [-8g to +8g range]
Class	Activity class -- sedentary, walking, stairs, running

Raw acceleration data along 3-D coordinates
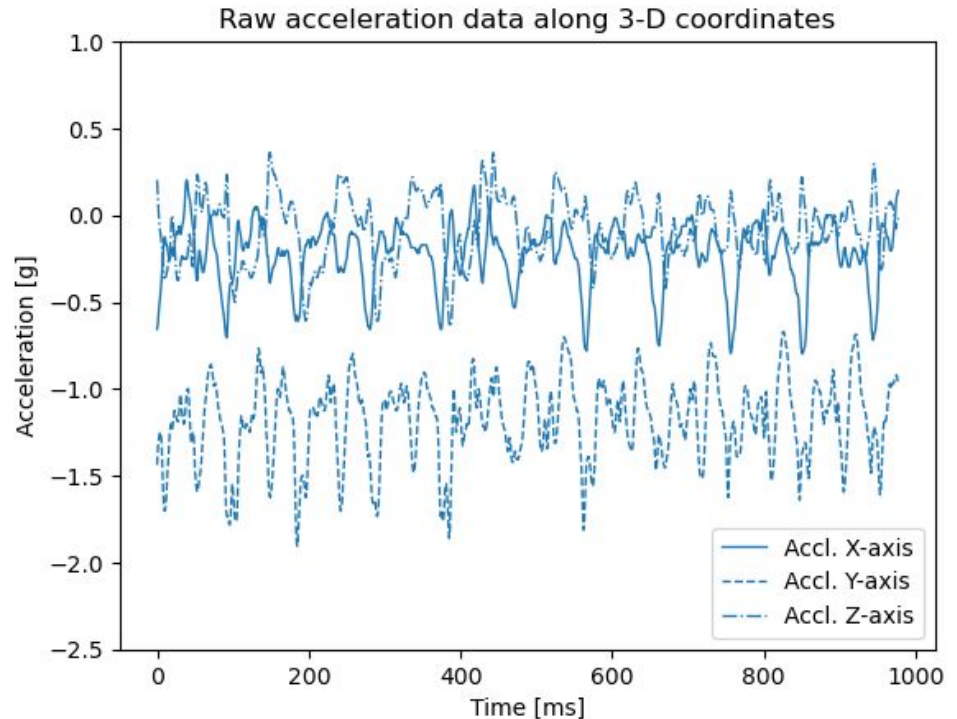
Windowing

Feature Engineering

Class Balance

Class Ratios
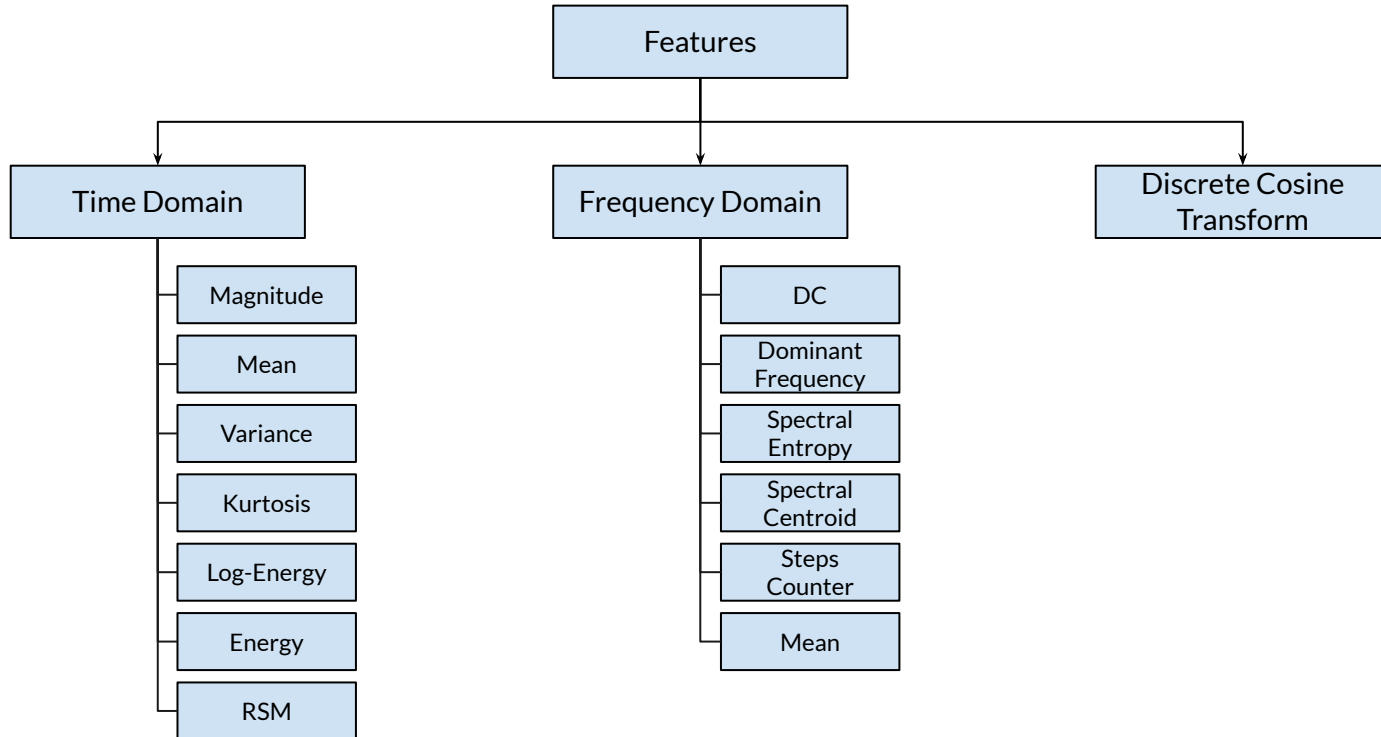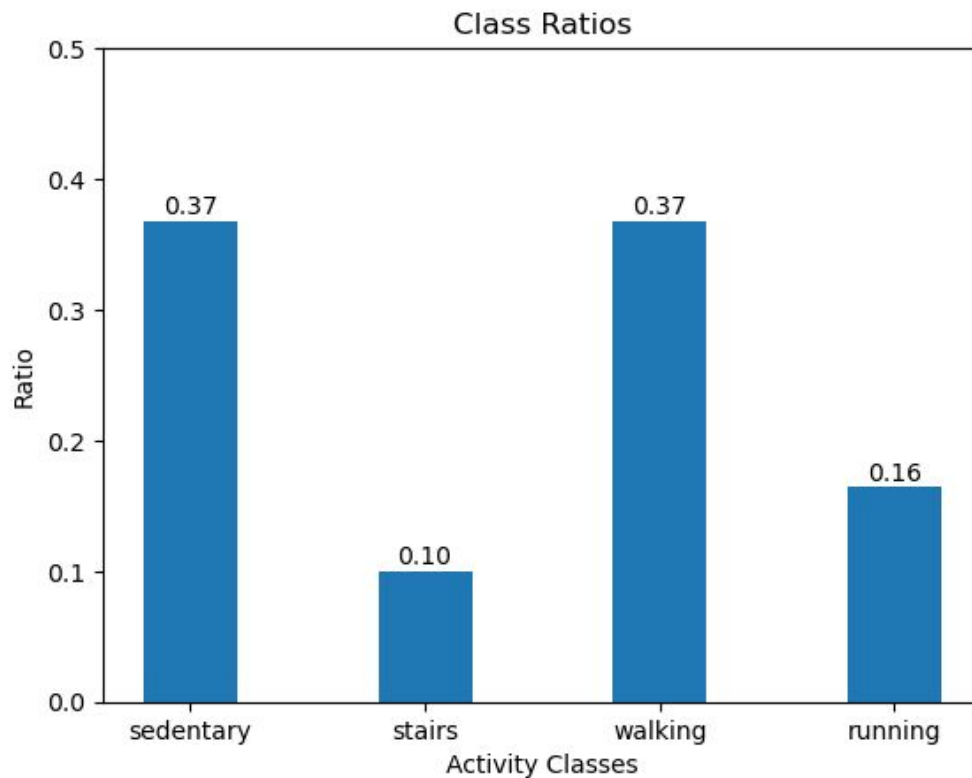
Acceleration increases with increase in activity intensity
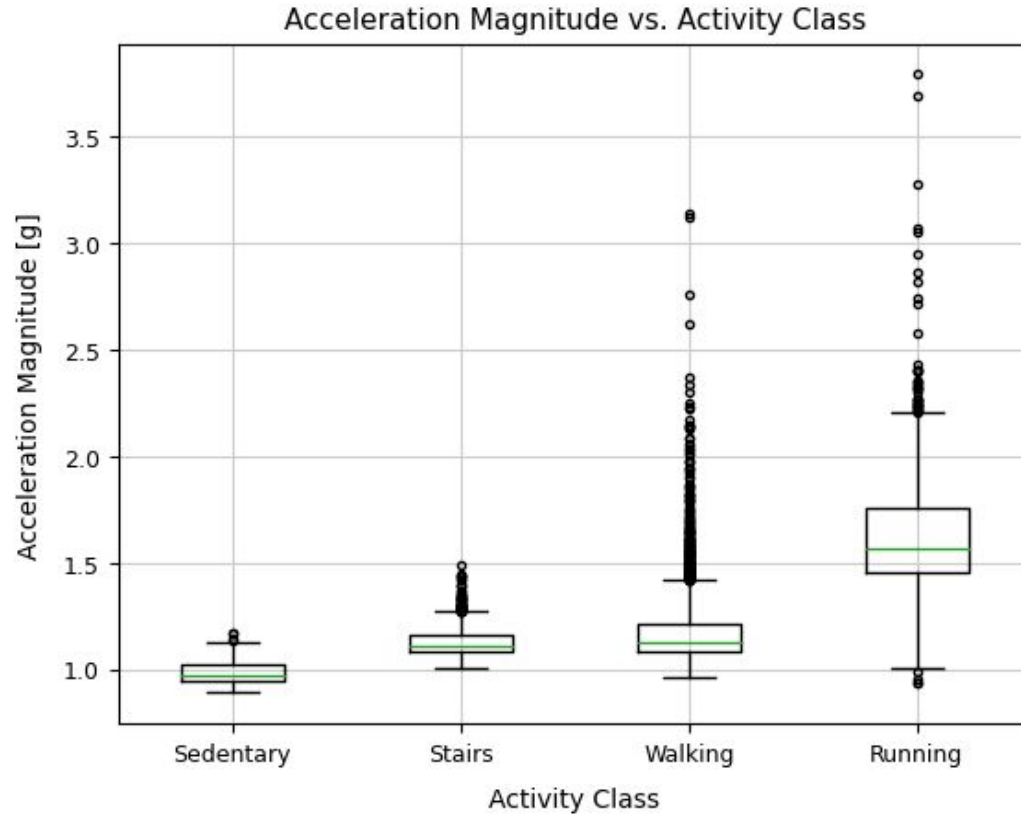
Dominant frequency varies with activity class

Category-Specific Histogram of Dominant Frequency
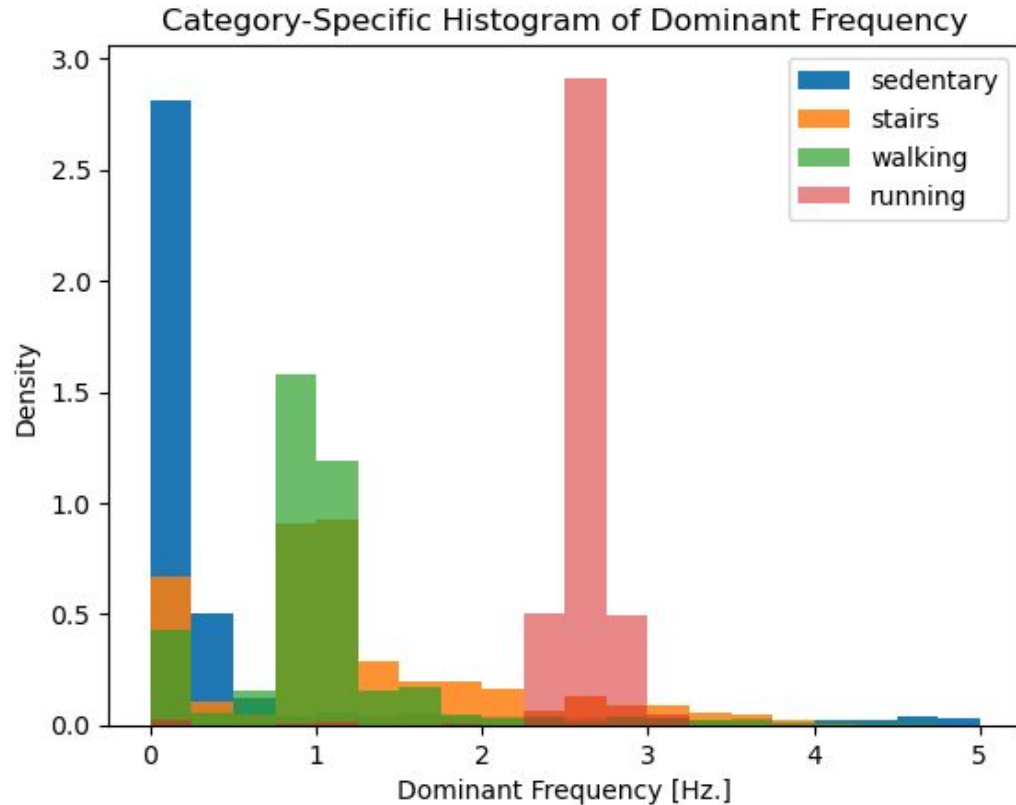
Log-Energy vs. Number of Steps
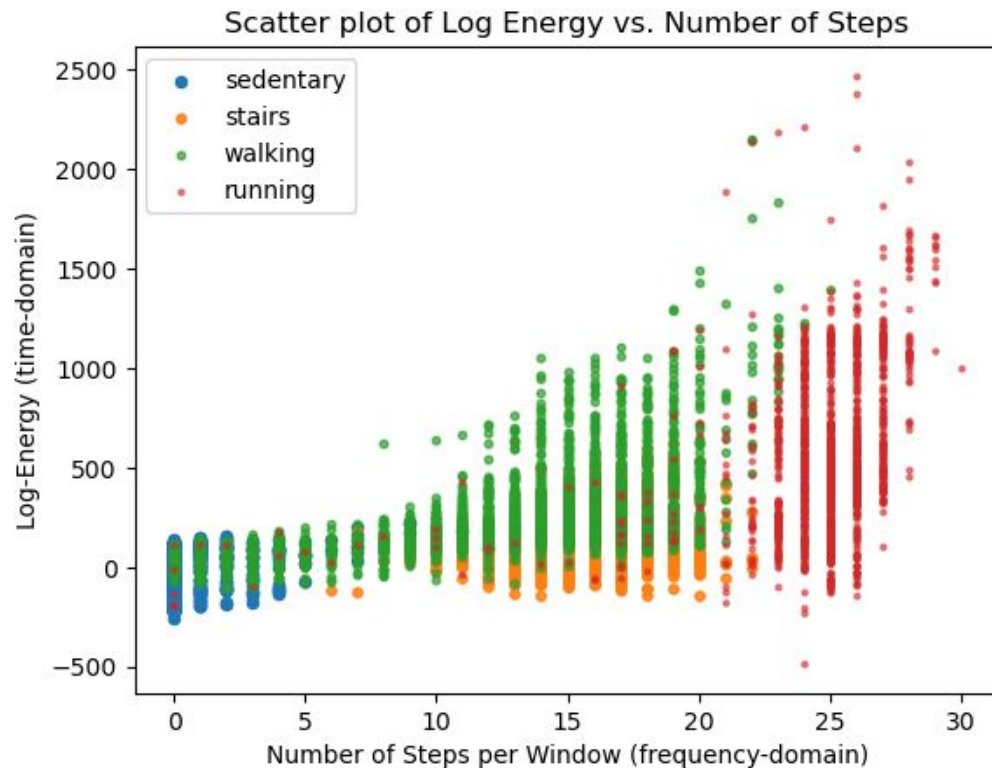
Splitting using Group Shuffle and Stratified K Fold

Splitting

Group Structure

Imbalanced Class

Ordered

Considerations

Test Set (~20%)
Group shuffle split with 1 split and 6 groups

Validation Set (~20%)
Stratified group k fold with 5 splits and shuffle true

Train Set (~60%)
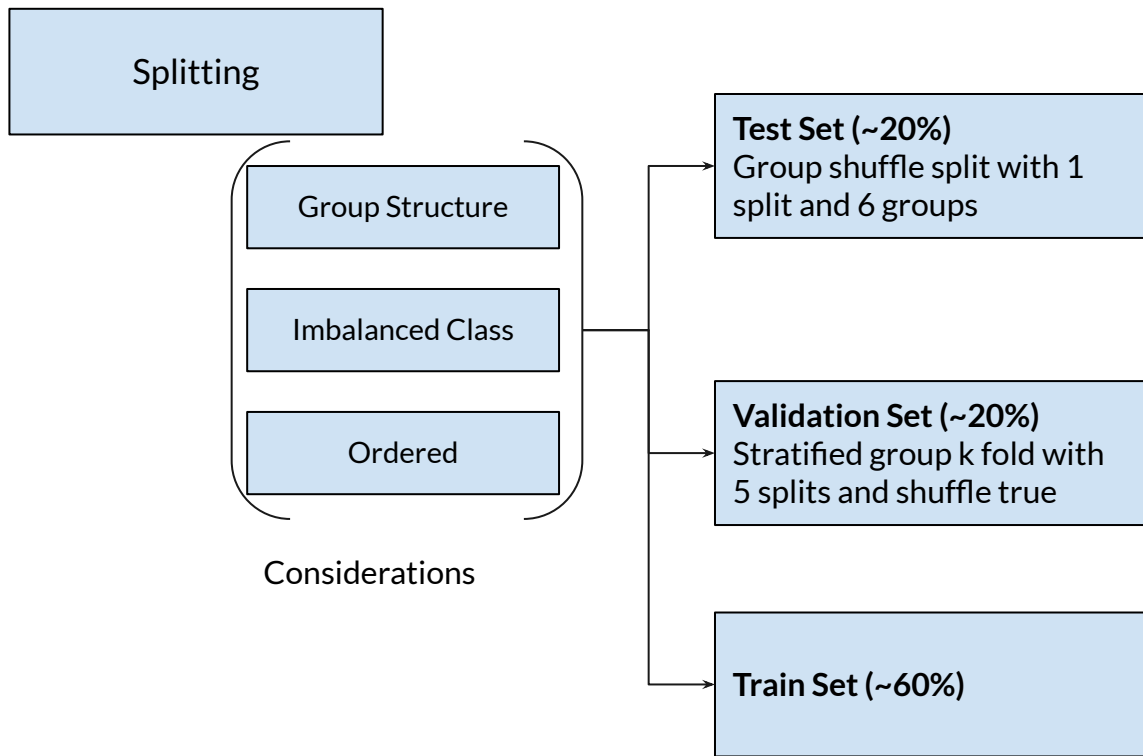
Impact of Preprocessing

Data State

	Raw Data	After Windowing & Feature Engineering		After Preprocessing
Size	4.8M rows x 4 features	9.4K rows x 45 features	Apply Standard Scaler on X_train	9.4K rows x 45 features
Missing Values	-	-		-
Feature types	0 categorical 3 continuous	0 categorical 45 continuous		0 categorical 45 continuous

Next Steps

- Add features from Discrete Cosine Transform

- EDA for interaction of DCT features with time and frequency domain features

- Identify the appropriate evaluation metrics

- Compare models applicable for multi-class classification
    - Decision Trees
    - Support Vector Machines
    - K-nearest neighbors

- Hyperparameter tuning, ideal window length, window overlap, etc.

Thank You!

Questions/Comments?