Prediction of the individual with over-weighted based on the center of mass acceleration data during the gait cycle.

Project Final Report

Team Members: Yu-Pin Liang

Abstract

Obesity is a prevalent issue in the U.S. The prevalence of obesity is increasing over the decade. Many underlying risks can lead to a critical health problem. Moreover, obesity is also a potential risk factor for cardiovascular disease (CVD) and diabetes. Recent advances in wearable motion sensor technologies offer an opportunity to translate laboratory findings to the clinical environment with similar gait balance measures. Inertial measurement units (IMU) combine accelerometers, gyroscopes, and magnetometers into a single sensor and could be used to estimate COM acceleration. The use of wearable sensors to detect mobility impairments has grown rapidly and provides a time-efficient and user-friendly measurement of gait and balance performance. The IMU is placed at COM's proxy location. This project aims to investigate whether the gait pattern would change in an individual with obesity. We hypothesized that obesity could be detected by using the COM acceleration data from self-selected speed walking.

1 Introduction

Obesity is a prevalent issue in the U.S. The prevalence of obesity is increasing over the decade. There are over 32.2 % of the adult were obese. Many underlying risks can lead to a critical health problem—for instance, the mortality rate increases due to chronic disease. Moreover, obesity is also a potential risk factor for cardiovascular disease (CVD) and diabetes.

Previously, whole-body COM is traditionally calculated as a weighted sum of linked rigid body segments captured with a whole-body reflective marker set and camera-based motion analysis equipment, which significantly restricts testing activities to laboratory settings. Recent advances in wearable motion sensor technologies offer an opportunity to translate laboratory findings to the clinical environment with similar gait balance measures. Inertial measurement units (IMU) combine accelerometers, gyroscopes, and magnetometers into a single sensor and could be used to estimate COM acceleration. The use of wearable sensors to detect mobility impairments has grown rapidly and provides a time-efficient and user-friendly measurement of gait and balance performance. The IMU is placed at a fixed body landmark, such as the 5th lumbar vertebra (L5), as a COM's proxy location. However, such a fixed landmark location does not account for any instantaneous changes in body segment alignment that could result in the relocation of whole-body COM.

Whether the subject is obsessed or not is a yes/no question. The model with the function of classifier can be used in this kind of problem. For example, a support vector machine (SVM) is a perfect candidate in the binary classification scenario. The SVM is also an algorithm of a supervised machine learning model. The training data can be labeled in a specific way to answer a particular research question. After the SVM model building by the training dataset, it has the ability to predict the category from the testing dataset as well as the unknown label.

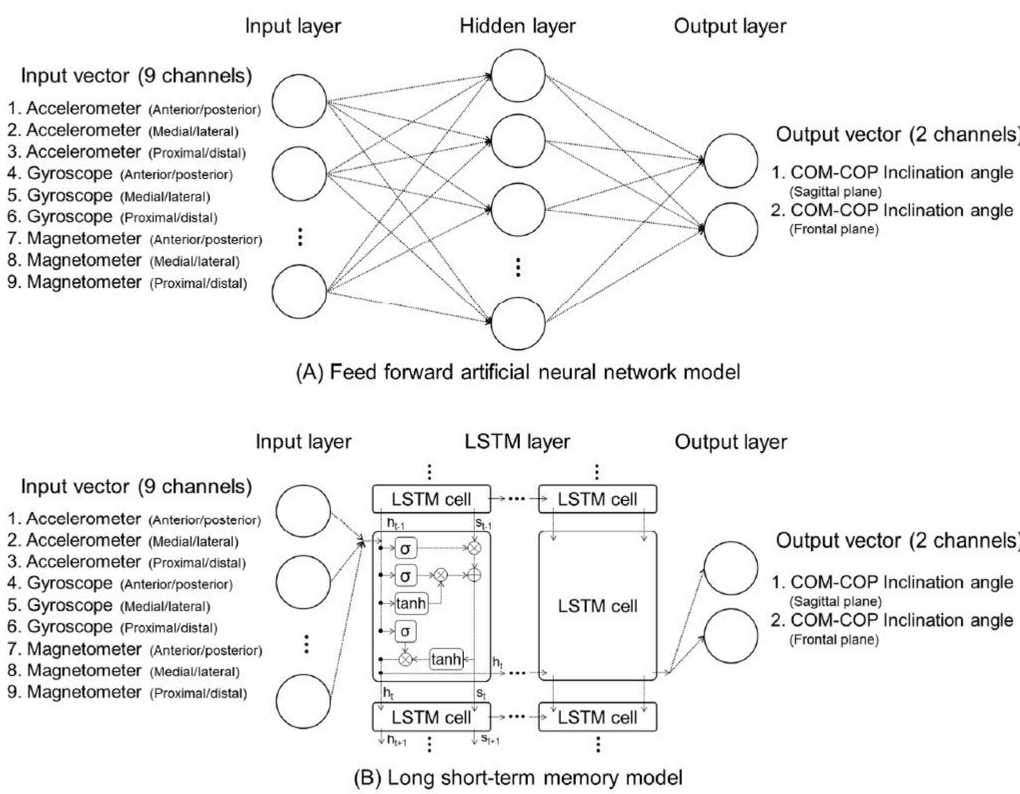
However, the SVM is not a model that considers the data as the time-series problem. The data from the IMU is the data with sequential characteristics. The Recurrent Neural Network is able to perform the data prediction by using the time-series data. The Long Short-Term Memory can memorize the prediction outcome from the last layer and synthesize the prediction result for the next layer. There are four gates in an LSTM neuron: input modulation gate, input gate, output gate, and forget gate.

This project aims to investigate whether the gait pattern would change in an individual with obesity. We hypothesized that obesity could be detected by using the COM acceleration data from self-selected speed walking.

2 Related Work

In Biomechanics, many researchers were still studying multiple types of sports activities, for example, running, jumping, sprinting, and cutting. As the human age growing, the kinematic and kinetic change due to the metabolic system, especially the muscle mass, will also decrease. Previously, the change of the muscle mass due to aging was identified by the traditional inferential statistics approach, which is time-consuming and insufficient sensitivity. Machine learning models can help scientists in Biomechanics to decrease the data processing time and increase the sensitivity of the changes.

Gait stability is also an important factor need to be addressed to prevent falling. The center of mass-center of pressure) the inclination angle is a good indicator to understand the balance ability of a person. The ANN also plays an important role in transferring this complex concept to a single IMU. This study stated that if the IMU could detect the change of this inclination angle, it represented that we don't need to calculate the kinematic parameters by going through the complicated process. The ML model, which is a feed-forward ANN and long-short term memory (LSTM) network.[5]



3 dataset

The colab was used to run the SVM model, so the data was imported to the github. There are 100 time stamp for each cell. x,y,z feature stack in to one cell.

4 SVM

Compare the accuracy between SVM model (without consider time-stamp)with LSTM (consider time-stamp) Data processing Sampling rate : IMU data 128 Hz Filtered with a 2nd order, zero-lag, and low-pass Butterworth filter with a 12 Hz cutoff frequency. 80 % of Training data, 20 % of Testing Data

```python
import numpy as np
import sklearn, sklearn.datasets, sklearn.utils, sklearn.model_selection, sklearn.svm
import pandas as pd
from sklearn import metrics
from sklearn.metrics import classification_report


url = 'https://raw.githubusercontent.com/yu-pin-liang/COMS-574-ML_finalProject/main/dataset.csv'
data = pd.read_csv(url,sep='delimiter', header=None,engine='python',delimiter = ',')  # Dataset
X = np.array(data)

y  = X[:,101]
X = X[:,0:100]

X, y = sklearn.utils.shuffle(X, y, random_state=1)
X_train, X_test, y_train, y_test = sklearn.model_selection.train_test_split(X, y, test_size=0.2, random_state=1)

clf = sklearn.svm.SVC(kernel='rbf',random_state=1)


clf.fit(X_train, y_train)


#print (X_train, X_test, y_train, y_test)

y_pred = clf.predict(X_test)






print("Accuracy:",metrics.accuracy_score(y_test, y_pred))



print(classification_report(y_test, y_pred))
```

The figure show the algorithm for the SVM model. The sklearn library was used and built for the training data.

5 LSTM

```
inputSize = 3;
numHiddenUnits = 100;
numClasses = 2;

layers = [ ...
    sequenceInputLayer(inputSize)
    bilstmLayer(numHiddenUnits,'OutputMode','last')
    fullyConnectedLayer(numClasses)
    softmaxLayer
    classificationLayer]

maxEpochs = 20;
miniBatchSize = 6;

options = trainingOptions('adam', ...
    'ExecutionEnvironment','cpu', ...
    'GradientThreshold',1, ...
    'MaxEpochs',maxEpochs, ...
    'MiniBatchSize',miniBatchSize, ...
    'SequenceLength','longest', ...
    'Shuffle','never', ...
    'Verbose',0, ...
    'Plots','training-progress');
net = trainNetwork(XTrain,YTrain,layers,options);



numObservationsTest = numel(XTest);
for i=1:numObservationsTest
    sequence = XTest{i};
    sequenceLengthsTest(i) = size(sequence,2);
end
[sequenceLengthsTest,idx] = sort(sequenceLengthsTest);
XTest = XTest(idx);
YTest = YTest(idx);



miniBatchSize = 6;
YPred = classify(net,XTest, ...
    'MiniBatchSize',miniBatchSize, ...
    'SequenceLength','longest');

acc = sum(YPred == YTest)./numel(YTest)
```

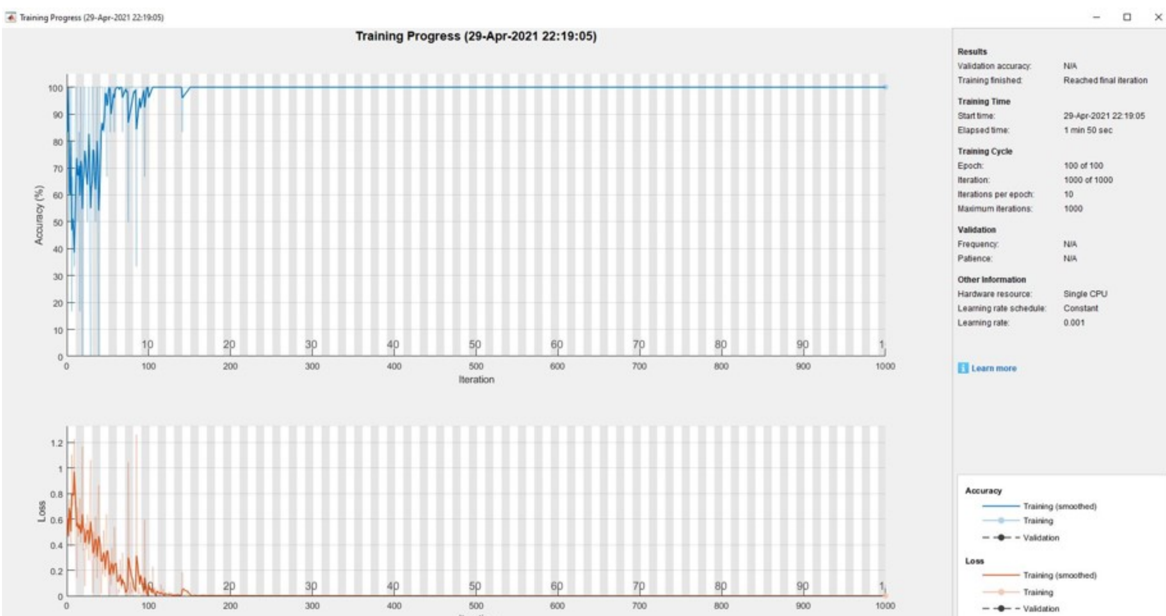
The matlab was used to build the LSTM model. The reason of switching to the matlab is because the experimental data variable are storing by using the matlab format. The Learning rate was set to 0.001, and number of Maximal epochs was set to 100. The Hardware is used in CPU. The maxEpochs is also set to 100 and miniBatchSize is set to 6.

6 Experimental Results

```
print("Accuracy:",metrics.accuracy_score(y_test, y_pred))


print(classification_report(y_test, y_pred))
```

```
Accuracy: 0.8571428571428571
```




```

acc = sum(YPred == YTest)./numel(YTest)

```

Command Window

```
     5   ''   Classification Output   crossentropyex

acc =

    0.9333
```

The accuracy for the SVM model is 0.85, and the accuracy for the LSTM model is 0.93

7 Conclusion

LSTM had higher accuracy rate compare with the SVM. Both of them have high accuracy. In the future plan, more features need to be added to make model increased the accuracy. For example, the data from the gyroscopes, and magnetometers. Moreover, change the model to the model which can predict the time-series result (predict the muscle force at each moment), not only just a classifier model.

8 References

1. Fukuchi RK, Eskofier BM, Duarte M, Ferber R. Support vector machines for detecting age-related changes in running kinematics. J Biomech. 2011 Feb 3;44(3):540-2. doi: 10.1016/j.jbiomech.2010.09.031. Epub 2010 Oct 25. PMID: 20980005.

2. Woldaregay, A. Z., Årsand, E., Walderhaug, S., Albers, D., Mamykina, L., Botsis, T., Hartvigsen, G. Data-driven modeling and prediction of blood glucose dynamics: Machine learning a pplications in type 1 diabetes. Artificial Intelligence in Medicine. 2019; 98, 109-134.

3. Tkatchenko, A. Machine learning for chemical discovery. Nature Communications, 2020; 11(1).

4. Rundo, Trenta, Stallo, D., Battiato. Machine Learning for Quantitative Finance Applications: A Survey. Applied Sciences, 2019;9(24), 5574. doi:10.3390/app9245574

5.Choi A, Jung H, Mun JH. Single Inertial Sensor-Based Neural Networks to Estimate COM-COP Inclination Angle During Walking. Sensors (Basel). 2019 Jul 5;19(13):2974. doi: 10.3390/s19132974. PMID: 31284482; PMCID: PMC6651410.