Physical Activity Classification



Why and where?







Literature

Webpage:

- Dataset with 6 different activities
- Accelerometer and gyroscopic data from smart phone available
- With simple 2 Layer neural network 89% Accuracy could be archieved
- With LSTM a acccuracy of 91 could be reached.

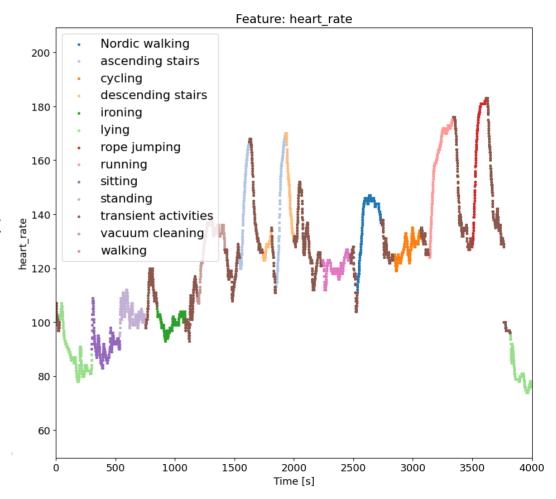
Kaggle Notebook:

- Same Dataset
- Achieved a 99% Accuracy but testing is not realistic
- All other Kaggle notebooks had the same flaw.

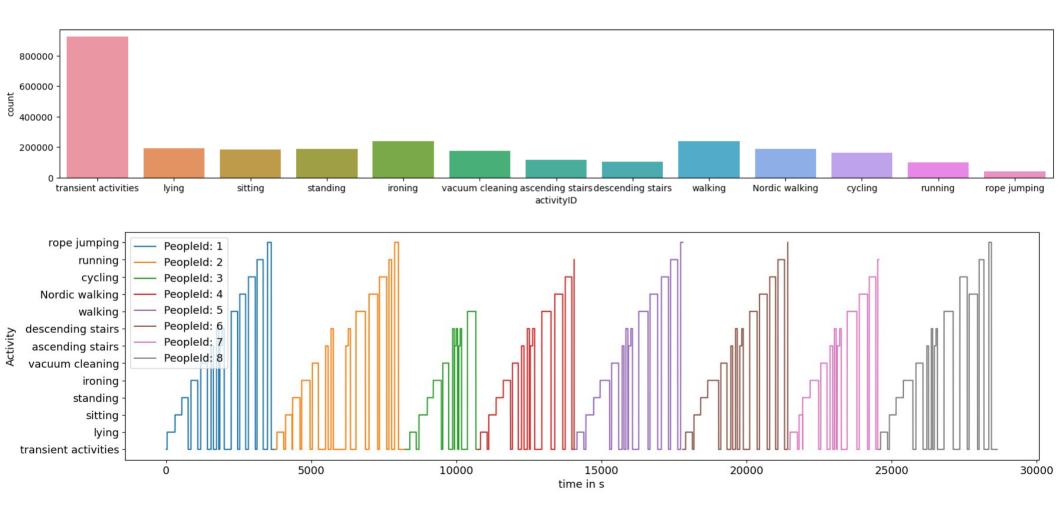
Dataset Characteristics

Dataset:

- From Kaggle
- 31 Features(heartbeat, accelerometer, gyroscope, ...)
- 13 labels(walking, lying, transient activities, etc ...)
- Data measured by 8 different people
- Around 2 million rows



Biases:



Baseline Model: Data Prep. and Evaluation

- Data undersampled, so that the activities are balanced.

- 1 person test and 7 person train and iterate over all people

- Because of undersample, accuracys may vary → do fitting 5 times and take average accuracy

Baseline Model: Results

```
# Create and train the logistic regression model
model = LogisticRegression(multi_class='multinomial', max_iter=100000)
model.fit(X_train_scaled, y_train)
```

Average accuracy of 47%

standart deviation basically 0

Test People ID: 6 Accuracy around 63%

Test People ID: 5 Accuracy around 33%

Model 1 with Simple NN: Data Prep. and Evaluation

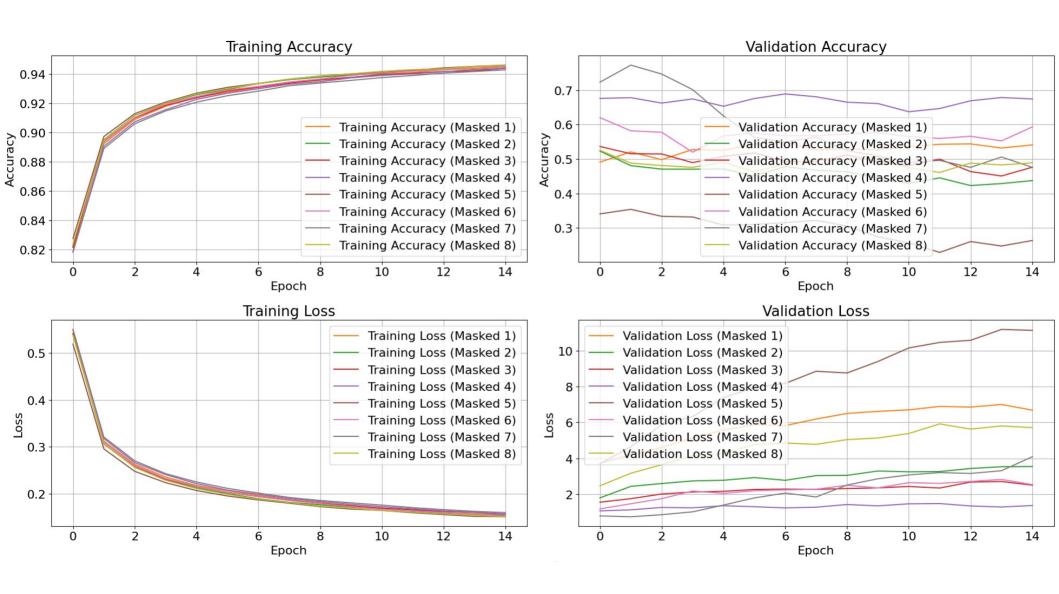
- Data undersampled, so activitys are balanced
- random shuffle to prevent sampling bias
- 1 person test and 7 person train and iterate over all people
- take the best validation accuracy from all epochs
- do fitting 5 times and take average fitting accuracy for evaluation

Model 1 with Simple NN: Results

```
# Build a neural network with Leaky ReLU and Batch Normalization
model = Sequential([
    BatchNormalization(input_shape=(X_train.shape[1],)), # Batch Normalization as the first layer
    Dense(300),
    LeakyReLU(alpha=0.01), # You can adjust the alpha parameter as needed
    Dropout(0.5),
    Dense(100),
    LeakyReLU(alpha=0.01),
    Dense(y_train.shape[1], activation='softmax')
])
```

Average accuracy of 57.1% Standart deviation of 0.2%

Already better!

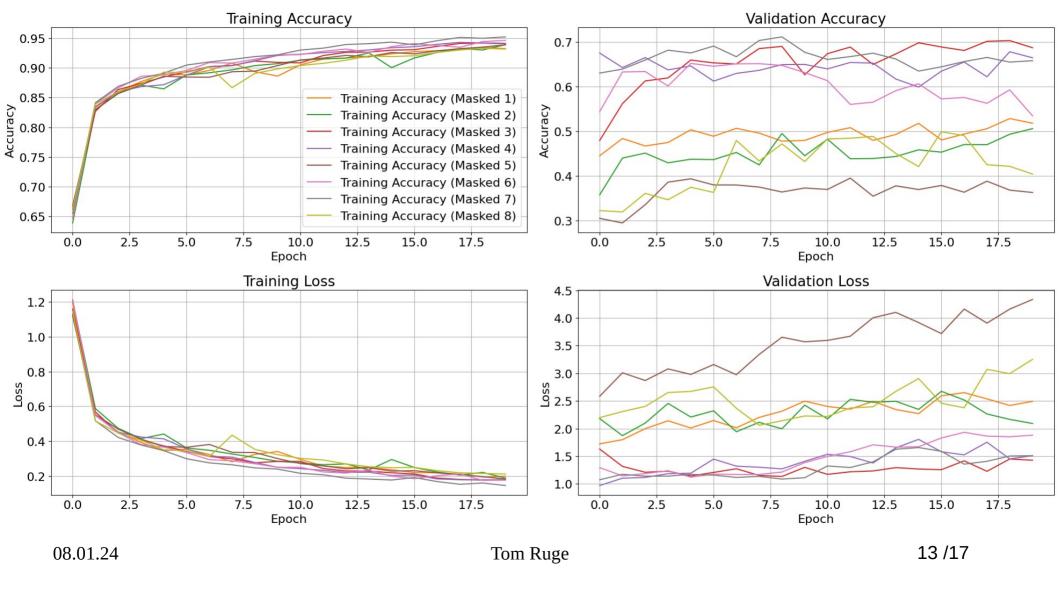


Model 2 with LSTM: Data Prep. and Evaluation

- Data undersampled, so activitys are balanced
- random shuffle of windows, to prevent sampling bias
- 1 person test and 7 person train. Iterate over all people
- take the best validation accuracy from all epochs
- do fitting 5 times and take average fitting accuracy for evaluation

Model 2 with LSTM: Results

Average accuracy of 59% Standard deviation of 1%



Results

- Both neural nets performed at least 10% better then the logistical regression.

- The first model performed with an accuracy around 57%

- The LSTM performed with an accuracy around 59% but worse then expected

Challenges and Errors

- The dataset contains many biases.

- Only 8 people

- Running time....



Conclusion and Outlook

Conclusion:

- As expected LSTM performed better then simple neural and the LSTM performed better then the sinple neural net. But the LSTM performed worse then expected.

Outlook:

- Better investigation of Hyperparameters
- reduce number of features
- average the classifications to improve accuracy.

Questions, Comments?

