



DATA MINING

Project Report

Bremen Big Data Challenge - Edition 2019

By Gari Ciodaro, Diogo Cosin, and Ralph Florent

May 17, 2019

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Laoreet non curabitur gravida arcu ac. Et molestie ac feugiat sed lectus vestibulum mattis ullamcorper velit. Turpis egestas pretium aenean pharetra. Amet justo diam vulputate ut pharetra sit. Elit scelerisque mauris pellentesque pulvinar. Mauris in aliquam sem fringilla. Id ornare arcu odio ut sem nulla. Nunc lobortis mattis aliquam faucibus purus. Eget velit aliquet sagittis id consectetur purus ut faucibus pulvinar. Nulla aliquet enim tortor auctor urna nunc id cursus. Varius quam quisque id diam vel quam elementum pulvinar.

1 Introduction

2 Background of the Data

2.1 Data Source

Provided by "The Bremen Big Data Challenge 2019" Organizers, the collected data are based on daily athletic movements [1]. Using wearable sensors above and below the knee of the individual (athletic), a dataset of 19 individuals, mainly identified as *subjects*, has been recorded. And as the competition requires, the data of 15 subjects are used as the training dataset and the remaining data as the testing dataset. The dataset is publicly available online on the official website: [BBDC](#).

2.2 Description and Format

The data comprise the following 22 movements:

- Race ('run')
- Walking ('walk')
- Standing (standing)
- Sitting ('sit')
- Get up and sit down ('sit-to-stand', 'stand-to-sit')
- Up and down stairs ('stair-up', 'stair-down')
- Jump on one or both legs ('jump-one-leg', 'jump-two-leg')
- Run left or right ('curve-left-step', 'curve-right-step')
- Turn left or right on the spot, left or right foot first ('curve-left-spin-Lfirst', 'curve-left-spin-Rfirst', 'curve-right-spin-Lfirst', 'curve-right-spin-Rfirst')
- Lateral steps to the left or right ('lateral-shuffle-left', 'lateral-shuffle-right')
- Change of direction when running to the right or left, left or right foot first ('v-cut-left-left', 'v-cut-left-right', 'v-cut-right-left', 'v-cut-right-right')

All the data are available as CSV files, or Comma-Separated Values. Starting from the training dataset identified by the *train.csv* file, the dataset is fixed and formatted as described in the following lines:

3 Data Preprocessing

4 Data Exploitation

5 Data Analysis

6 Results and Discussions

7 Conclusion

References

- [1] University of Bremen. *The Bremen Big Data Challenge 2019*. May 2019. URL: <https://bbdc.csl.uni-bremen.de/>.