

Overview

The PAMAP2 Physical Activity Monitoring dataset ([from here \(Links to an external site.\)](#)), contains data of 18 different physical activities (such as walking, cycling, playing soccer etc) performed by 9 subjects wearing 3 inertial measurement units (IMU) and a heart rate monitor. This data is stored in individual text files per subject. Each row in each file represents one reading and contains 54 attributes (including timestamp, activity ID, heart rate and IMU sensory data). Please see the readme.pdf file supplied with the dataset for information about the different attributes and how the dataset was collected.

Assuming the goal is to develop hardware and/or software which can determine the amount and type of physical activity carried out by an individual, what actionable insights can you derive from the dataset?

Specific Requirements

You are required to:

1. carry out thorough exploratory data analysis and appropriately handle missing or dirty data;
2. develop and test at least one hypothesis for a relationship between a single pair of attributes;
3. develop and test at least one model which uses multiple attributes to make predictions.