

# Welcome to TMD dataset

The first free dataset about  
Transportation Mode Detection

What is



## Transportation Mode Detection?

Identify user's transportation modes through observations of the user, or observation of the environment, is a growing topic of research, with many applications in the field of **Internet of Things (IoT)**. **Transportation mode detection** can provide context information useful to offer appropriate services based on user's needs and possibilities of interaction.

---



## What is the **goal** of TMD?

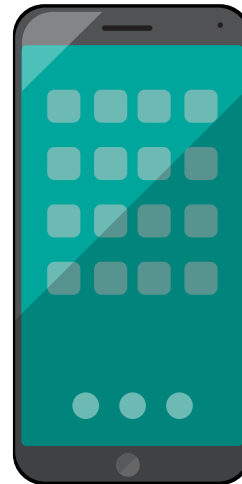
User transportation mode recognition can be considered as a **HAR** task (Human Activity Recognition). Its goal is to identify which kind of transportation - walking, driving etc.- a person is using. Transportation mode recognition can provide context information to enhance applications and provide a better user experience, it can be crucial for many

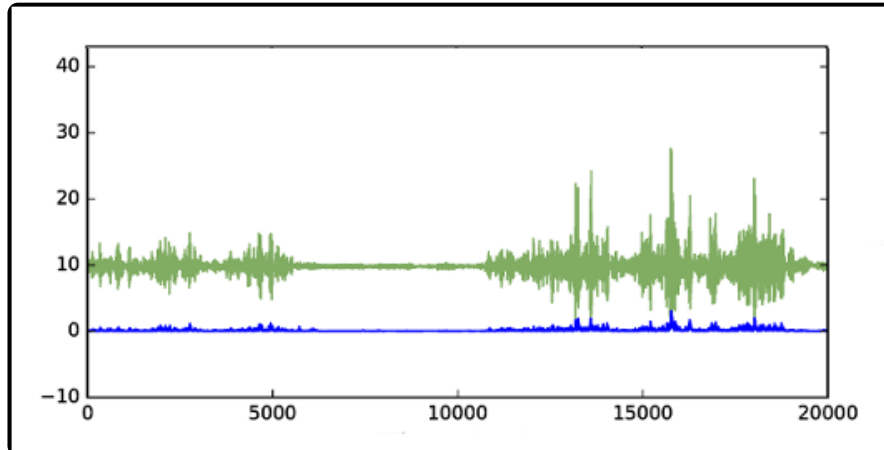
different applications, such as *device profiling, monitoring road and traffic condition, Healthcare, Traveling support* etc..

---

## Context acquisition via **smartphone's sensors**

A sensor measures different physical quantities and provides corresponding raw sensor readings which are a source of information about the user and their environment. Due to advances in sensor technology, sensors are getting more powerful, cheaper and smaller in size. Almost all mobile phones currently include sensors that allow the capture of important context information. For this reason, one of the key sensors employed by context-aware applications is the mobile phone, that has become a central part of users lives.





## Data collection

We collect sensors data from thirteen volunteer subjects, ten male, and three female. The set of classes we classify is composed by

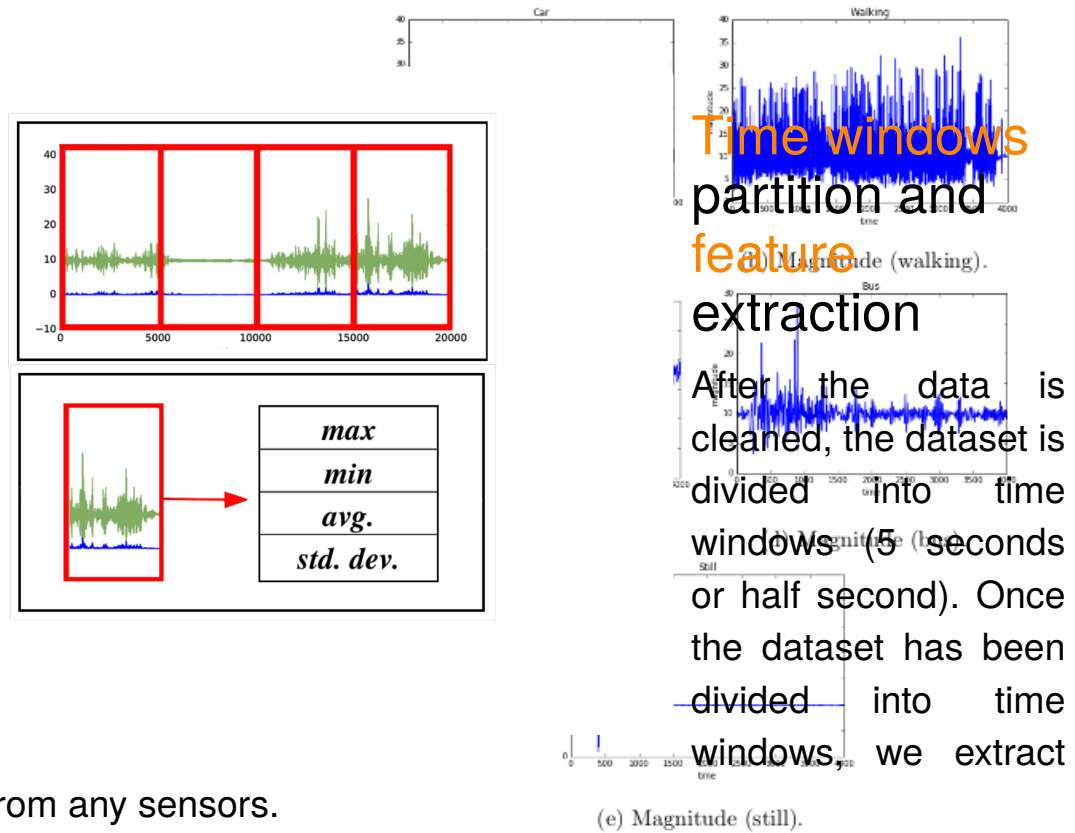
**walking, car, still, train and bus.**

In total, our dataset is composed of 226 labeled files (in continuous updating) representing the same number of activities corresponding to more than 31 hours of data: 26% of data is annotated as **walking**, 25% as driving a **car**, 24% as standing **still**, 20% as being on **train**, and 5% as being on **bus**.

## Sensors data **cleaned** and processed sensors based on **magnetometer orientation**

Initial data pre-processing phase: data cleaning operations are performed, such as delete measure from the sensors to exclude, make the values of the sound and speed sensors positive etc...

Furthermore some sensors, like ambiental (sound, light and pressure) and proximity, returns a single data value as the result of sense, this can be directly used in dataset. Instead, all the other return more than one values that are related to the coordinate system used, so their values are strongly related to **orientation**. For almost all we can use an orientation-independent metric, **magnitude**.

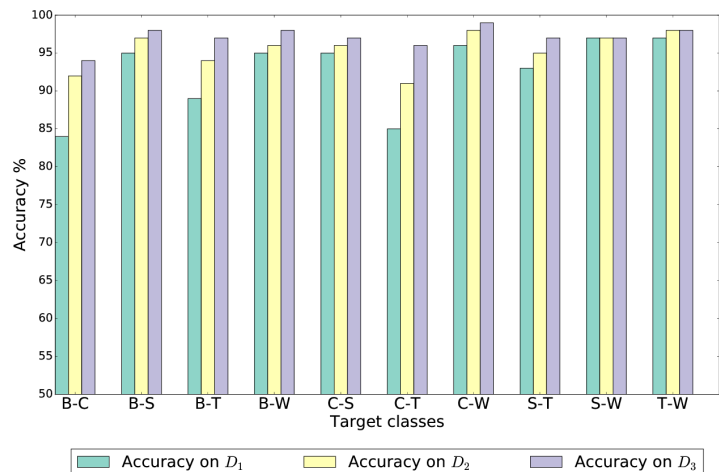


## Evaluation with different machine learning techniques

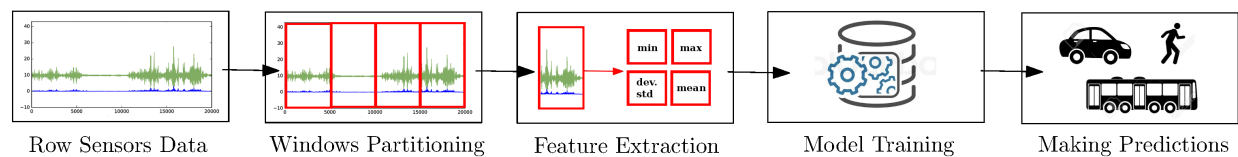
We build three different sensors set and for each set, we build four model with four different **classification algorithms**

(Decision Tree,

Neural Network, Random Forest and Support Vector Machines) and we got as best accuracy **96%**. Then we perform other experimentation such as the **class-to-class classification**, in which we perform a similar analysis but considering only two classes to discriminate between.



## Complete process



Copyright © Università di Bologna - 2017