



UNIVERSITÀ DEGLI STUDI DI
MILANO-BICOCCA

F1801Q145

MODELLI PROBABILISTICI PER LE DECISIONI

HAR Bayesian Network

Studenti:

Artifoni Mattia

Brena Luca

Bottoni Federico

Matricole:

807466

808216

806944

Giugno 2019

Indice

| | | |
|----------|--|----------|
| 1 | Abstract | 2 |
| 2 | Introduzione | 2 |
| 2.1 | Dominio di riferimento | 2 |
| 2.2 | Ipotesi e assunzioni | 3 |
| 3 | Scelte di design | 3 |
| 3.1 | Analisi statistica e qualitativa | 3 |
| 3.2 | Normalizzazione | 3 |
| 3.3 | Discretizzazione | 3 |
| 4 | I modelli di rete | 3 |
| 4.1 | pgmpy | 3 |
| 4.2 | Il modello correlato | 3 |
| 5 | Risultati e conclusioni | 4 |

1 Abstract

During the last 5 years, research on Human Activity Recognition (HAR)[1] has reported on systems showing good overall recognition performance. As a consequence, HAR has been considered as a potential technology for e-health systems. Here, we propose a machine learning based HAR classifier[3]. We also provide a full experimental description that contains the HAR wearable devices setup and a public domain dataset comprising 165,633 samples. We consider 5 activity classes, gathered from 4 subjects wearing accelerometers mounted on their waist, left thigh, right arm, and right ankle. As basic input features to our classifier we use 12 attributes derived from a time window of 150ms. Finally, the classifier uses a committee AdaBoost that combines ten Decision Trees. The observed classifier accuracy is 99.4

2 Introduzione

Il progetto ha l'obiettivo di creare un modello di Rete Bayesiana capace di predire il tipo di azione che sta effettuando un ipotetico individuo che indossa il "HAR wearable devices setup", una particolare sistema indossabile composto da 4 accelerometri che permette di analizzare i vettori accelerazione dei sensori in questione. Viene fornito dal progetto di riferimento[1] un dataset contenente dati sufficienti per effettuare training e testing del modello

2.1 Dominio di riferimento

La natura dei dati utilizzati è definita nel paper[3] del progetto di provenienza. La singola entry del dataset rappresenta uno snapshot acquisito dai sensori e consiste in:

- user: stringa rappresentante l'username dell'individuo in oggetto
- geneder: stringa relativa al suo genere
- age: intero dell'età
- how_tall:in_meters: decimale che esprime l'altezza
- weight: intero riguardo al peso
- body_mass_index: decimale rappresentante la distribuzione del peso
- xi: intero che esprime la componente x del vettore accelerazione nel sensore i-esimo

- y_i : intero che esprime la componente y del vettore accelerazione nel sensore i -esimo
- z_i : intero che esprime la componente z del vettore accelerazione nel sensore i -esimo
- `class`: stringa, il target del modello sviluppato: rappresenta l'azione/situazione dell'individuo al momento dello snapshot, può assumere il valore di "walking", "standing", "standingup", "sitting" e "sittingdown"

2.2 Ipotesi e assunzioni

3 Scelte di design

3.1 Analisi statistica e qualitativa

3.2 Normalizzazione

3.3 Discretizzazione

4 I modelli di rete

4.1 pgmpy

Il software scelto è pgmpy[2] di Python, una libreria che permette di modellare le dipendenze in modo agile, stimare le CPT delle variabili sfruttando dei metodi che accettano il dataset ed effettuare inferenze dichiarando la variabile di query e le evidenze.

Utilizzando la libreria ci siamo resi conto di come sia performante utilizzando modelli semplici e correlati da pochi record, tuttavia appena è avvenuta l'esecuzione della stima delle CPT nel primo modello completo ideato abbiamo riscontrato le prime difficoltà: il modello decisamente complesso viene bloccato da Python date le eccessive combinazioni possibili. Abbiamo quindi tentato di semplificare la rete.

4.2 Il modello correlato

5 Risultati e conclusioni

Riferimenti bibliografici

- [1] <http://groupware.les.inf.puc-rio.br/har>.
- [2] <http://pgmpy.org/>.
- [3] Katia Vega Eduardo Velloso Ruy Milidiú Wallace Ugulino, Débora Cardador and Hugo Fuks. Wearable computing: Accelerometers' data classification of body postures and movements, 2012.