

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

---



## **BÁO CÁO**

## **THỰC HÀNH LAB 5**

MÔN: THU THẬP VÀ TIỀN XỬ LÝ DỮ LIỆU  
SINH VIÊN THỰC HIỆN: NGÔ ĐỨC VŨ  
MÃ SỐ SINH VIÊN: 20520950

Câu hỏi: Đọc file README.md và cho biết các thông tin sau:

- a) Chức năng (nội dung) của bộ dữ liệu là gì?
- b) Các tác giả của bộ dữ liệu?
- c) Cấu trúc các tập tin của bộ dữ liệu? Ý nghĩa từng tập tin.
- d) Các thông tin nào mà các record lưu trữ?

## TRẢ LỜI:

### 1, Chức năng của bộ dữ liệu:

Nhận biết hoạt động của con người bằng tập dữ liệu điện thoại thông minh

### 2, Các tác giả của bộ dữ liệu:

Jorge L. Reyes-Ortiz, Davide Anguita, Alessandro Ghio, Luca Oneto.

Smartlab - Non Linear Complex Systems Laboratory

DITEN - Università degli Studi di Genova.

Via Opera Pia 11A, I-16145, Genoa, Italy.

### 3, Cấu trúc của bộ dữ liệu:

Bộ dữ liệu bao gồm các file:

=====

- 'README.txt'

- 'features\_info.txt': : Hiển thị thông tin về các biến được sử dụng trên vector đặc trưng.

- 'features.txt': Danh sách thuộc tính.

- 'activity\_labels.txt': Liên kết các nhãn lớp với tên hoạt động của chúng.

- 'train/X\_train.txt': Bộ huấn luyện.

- 'train/y\_train.txt': Nhãn huấn luyện.

- 'test/X\_test.txt': Bộ test.

- 'test/y\_test.txt': Bộ nhãn.

Các tệp sau đây có sẵn cho dữ liệu thử nghiệm và huấn luyện. Mô tả của chúng là tương đương.

- 'train/subject\_train.txt': Mỗi hàng xác định chủ thể đã thực hiện hoạt động cho mỗi mẫu của số. Phạm vi của nó là từ 1 đến 30.

- 'train/Inertial Signals/total\_acc\_x\_train.txt': Tín hiệu gia tốc từ trục X của máy đo gia tốc trên điện thoại thông minh tính theo đơn vị trọng lực tiêu chuẩn 'g'. Mỗi hàng hiển thị một vector 128 phần tử. Mô tả tương tự cũng áp dụng cho các tệp 'total\_acc\_x\_train.txt' và 'total\_acc\_z\_train.txt' cho trục Y và Z.

- 'train/Inertial Signals/body\_acc\_x\_train.txt': Tín hiệu gia tốc của cơ thể thu được bằng cách lấy tổng gia tốc trừ đi trọng lực.

- 'train/Inertial Signals/body\_gyro\_x\_train.txt': Vector vận tốc góc do con quay hồi chuyển đo được cho mỗi mẫu cửa sổ. Đơn vị là radian / giây.

Lưu ý:

=====

- Các tính năng được chuẩn hóa và giới hạn trong  $[-1,1]$ .
- Mỗi vector đặc trưng là một hàng trên tệp văn bản.

#### 4, Các thông tin nào mà các record lưu trữ?

- Gia tốc ba trục từ gia tốc kế (tổng gia tốc) và gia tốc cơ thể ước tính.
- Vận tốc góc ba trục từ con quay hồi chuyển.
- Một vector với 561 đặc trưng với các biến miền thời gian và tần số.
- Nhãn hoạt động của nó.
- Định danh của đối tượng thực hiện thí nghiệm.