

❖ 2) Đọc file README.md và cho biết các thông tin sau:

➤ a) Chức năng (nội dung) của bộ dữ liệu là gì?

- Bộ dữ liệu được thu thập từ 1 nhóm tình nguyện viên gồm 30 người trong độ tuổi từ 19-48. Mỗi người thực hiện sáu hoạt động (WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING) khi đang đeo đồng hồ thông minh (Samsung Galaxy S II) ở thắt lưng. Các thí nghiệm được quay video để dán nhãn dữ liệu theo cách thủ công. Sử dụng gia tốc kế và con quay hồi chuyển được nhúng đã thu được gia tốc tuyến tính 3 trục và vận tốc góc 3 trục ở tốc độ không đổi 50Hz. Tập dữ liệu thu được được chia ngẫu nhiên thành hai tập, trong đó 70% tình nguyện viên được chọn để tạo dữ liệu huấn luyện và 30% dữ liệu thử nghiệm.

➤ b) Các tác giả của bộ dữ liệu?

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

➤ c) Cấu trúc các tập tin của bộ dữ liệu? Ý nghĩa từng tập tin.

- 'README.txt': Thông tin tổng quát của bộ dữ liệu.
- 'features_info.txt': Thông tin về các cột (thuộc tính) của dữ liệu.
- 'features.txt': Tên các cột (thuộc tính) của dữ liệu
- 'activity_labels.txt': Nhãn các hoạt động đã được đánh số hay các giá trị của cột dữ liệu cần dự đoán.
- 'train/X_train.txt': Dữ liệu về các thông số dùng để huấn luyện.
- 'train/y_train.txt': Dữ liệu và các nhãn hoạt động đã được đánh số dùng để huấn luyện.
- 'test/X_test.txt': Dữ liệu về các thông số dùng để thử nghiệm.
- 'test/y_test.txt': Dữ liệu và các nhãn hoạt động đã được đánh số dùng để thử nghiệm.

Các tệp dưới đây đều miêu tả chung cho cả tập train và test :

- 'train/subject_train.txt': Mỗi hàng dữ liệu là số hiệu của tình nguyện viên. Được đánh số từ 1 đến 30.
- 'train/Inertial Signals/total_acc_x_train.txt': Tín hiệu gia tốc từ trục X của gia tốc kế điện thoại thông minh 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 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 trừ đi trọng lực khỏi tổng gia tốc.
- 'train/Inertial Signals/body_gyro_x_train.txt': Vector vận tốc góc được đo bằng con quay hồi chuyển cho từng mẫu. Đơn vị là radian/giây.

➤ d) Các thông tin nào mà các record lưu trữ?

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