

# Competition

☰ Tags

## Nguồn tài liệu:

Tại đây.

- Thông tin, hướng dẫn sử dụng data:
  - Information\_data\_competition.docx
- **Dữ liệu đầu vào X:**
  - group1
  - group2
  - group3
- **Dữ liệu đầu ra y:**
  - filtered\_subjects\_with\_age.tsv

---

## Dữ liệu:

Dữ liệu đã được tiền xử lý và giảm kích thước từ 306 kênh xuống còn 52 kênh. Tần số lấy mẫu là 250 Hz.

---

## Nhiệm vụ:

Chọn ngẫu nhiên 40 đối tượng từ dữ liệu được cung cấp. Đây sẽ là tập kiểm tra.

Sử dụng các đối tượng còn lại để huấn luyện một bộ hồi quy. Bộ hồi quy này sẽ dự đoán tuổi của đối tượng.

Kiểm tra bộ hồi quy tối ưu của bạn trên tập dữ liệu kiểm tra và tính toán MAE của tuổi dự đoán (xem công thức bên dưới).

---

## Tài liệu cần nộp:

Một tệp zip bao gồm tất cả các mục sau:

a) Mã Python của bạn. **Mã cần được COMMENT bằng tiếng Anh.**

b) Tập `environment.yml` dành cho conda.

c) Một báo cáo ngắn bằng tiếng Anh, bao gồm:

- Tóm tắt chiến lược bạn sử dụng để dự đoán tuổi não từ dữ liệu chuỗi thời gian.
- Phương pháp: mô tả các mô hình bạn sử dụng và chiến lược chống overfitting.
- Kết quả của phương pháp bạn dùng. Bao gồm giá trị MAE (Mean Absolute Error) bạn tìm được trên tập kiểm tra.
- Thảo luận và kết luận: Kết quả của bạn có tốt không? Những cách nào có thể cải thiện trong tương lai?
- Tham khảo các tài liệu hoặc bài báo bạn đã sử dụng.

**Cách tính điểm:**

- 5 điểm: code chạy được với file config
- 5 điểm: chất lượng tài liệu và độ dễ đọc của code.
- 10 điểm: độ chính xác của dự đoán.
- 20 điểm: chất lượng của báo cáo.
- Điểm cộng tiềm năng: thảo luận về công việc trong tương lai tốt đến đâu.

Tổng điểm sẽ chiếm 40% điểm cuối cùng của bạn cho khóa học này.

**Khi chấm điểm, Mr. Guy Nagel sẽ xem xét rằng bạn có ít thời gian hơn để hoàn thành nhiệm vụ này so với kế hoạch ban đầu.**