

Môn: **Máy học**

Bài kiểm tra #2: Học không giám sát

Thời gian: 120 phút

Tập tin ***Football_Players.xls*** chứa dữ liệu về những cầu thủ bóng đá trên thế giới. Sử dụng ngôn ngữ Python (**tập tin Notebook .ipnb**) để thực hiện những yêu cầu sau:

1.1 Xây dựng mô hình gom cụm theo 2 phương pháp k-Means và HAC.

1.2 Đánh giá các mô hình gom cụm đã xây dựng để chọn mô hình tốt nhất dựa trên:

- các siêu tham số: khoảng cách, linkage
- các tiêu chí: Silhouette score, Rand index

Lưu ý:

- Nén tất cả bài làm vào 1 file .RAR (hay .ZIP) với cách đặt tên:

<Tên>, <Họ>.RAR

VD: **Anh, TranTuan.RAR** (*không tạo các thư mục con*)

- Bài làm sẽ bị trừ điểm nếu không thực hiện đầy đủ những yêu cầu nêu trên.