**MAI391 Computer project**

Mỗi nhóm tự chọn 1 thuật toán trong Machine Learning (trong hoặc ngoài giáo trình), giải thích lý thuyết ngắn gọn bằng slides, code thuật toán và demo trên data thực.

Các nhóm có thể lựa chọn thực hiện làm theo các câu hỏi có sẵn như dưới đây, với số điểm tối đa tương ứng.

**Hỗn hợp (tối đa 6 điểm).**

Sinh viên tìm hiểu các thư viện có sẵn hoặc viết code ngắn cho *3 trong 4 bài toán* dưới đây:

1. Tìm giá trị riêng, vector riêng. Phân tích SVD của 1 ma trận. Xấp xỉ ma trận bởi các ma trận rank-1.
2. Backpropagation để tích đạo hàm của 1 hàm nhiều biến (sẽ được cho trước, tùy từng nhóm).
3. Gradient descent algorithm.
4. Các bài toán linear and quadratic programming sử dụng CVXOPT.

**Linear regression (tối đa 8 điểm).**

1. Tìm một data có D features và target y (D >= 2).
2. Viết code để tìm Linear regression model to fit this data with all D features:

y ≈ θ0 + θ1\*Feature1+…+ θD\*FeatureD.

1. Chọn ra 1 feature (trong số D features), gọi là X. Fit data bao gồm feature X và target y bằng một vài models là các đa thức bậc 2, 3, 4,.... So sánh RMSE (trên test data) của các models này để chọn ra best model.

*Nhóm tự đặt thêm các câu hỏi thú vị sẽ có cơ hội được điểm cao hơn.*

**PCA (tối đa 8 điểm).**

Code theo các bước đã trình bày trên lớp, và demo trên một data cụ thể. Sv có thể dùng sẵn thư viện về eigenvalues, eigenvectors.

*Để được điểm cao hơn, nhóm cần gắn với một ứng dụng cụ thể của PCA, ví dụ image compression.*