# TRƯỜNG ĐẠI HỌC Y TẾ CÔNG CỘNG

KHOA KHOA HỌC CƠ BẢN

# BÀI TẬP CUỐI KHÓA HỌC

## Môn học: Học sâu

## Đối tượng: CN KHDL1

Sinh viên Cử nhân Khoa học Dữ liệu

Hình thức: Sinh viên làm bài tập theo nhóm: 3-4 sinh viên/ nhóm và trả lời câu hỏi trực tiếp trên lớp.

Thời hạn nộp bài: 10 ngày làm việc

## 1. Giới thiệu

Trong suốt khóa học Khoa học Dữ liệu, chúng ta đã cùng nhau khám phá nền tảng lý thuyết và thực hành các công cụ thiết yếu, từ Thống kê mô tả, Tiền xử lý dữ liệu (Data Preprocessing), đến xây dựng và đánh giá các Mô hình Học máy (Machine Learning Models). Bài tập cuối khóa này không chỉ là một cột mốc để đánh giá sự tiếp thu kiến thức mà còn là cơ hội để các anh/chị vận dụng tổng hợp các kỹ năng đã học vào một vấn đề thực tế, mô phỏng quy trình làm việc của một Nhà Khoa học Dữ liệu (Data Scientist).

## 2. Mục tiêu học tập

Sau khi hoàn thành bài tập này, sinh viên có khả năng:

* Phân Tích và Xác Định Vấn Đề (Problem Framing and Analysis)
* Xử Lý và Khám Phá Dữ Liệu (Data Handling and Exploratory Data Analysis - EDA)
* Dựng và Đánh Giá Mô Hình (Model Building and Evaluation)
* Truyền Đạt và Đưa ra Kiến Nghị (Communication and Recommendation)

## 3. Yêu cầu sản phẩm đầu ra

Sản phẩm cuối cùng của nhóm là một báo cáo chi tiết dưới dạng file Word (.docx) và kèm theo slide thuyết trình (nếu có), theo các yêu cầu sau:

* **Định dạng:** File Word (.docx).
  + Tiêu đề: font chữ Times New Roman hoặc Times New Roman, cỡ chữ 13, in đậm
  + Nội dung: font chữ Times New Roman cỡ chữ 13, thường, cách dòng 1,5
  + Khổ giấy A4; lề trái, phải 3 cm; trên dưới 2 cm, lề trên dưới 2 cm
  + Đóng bìa, không đóng giấy bóng kính.
  + Trang bìa (phụ lục 1)
* **Cấu trúc nội dung:** Nội dung trình bày theo các đề mục rõ ràng, mạch lạc, sử dụng heading phù hợp để phân cấp thông tin. Diễn giải khoa học, logic và có tính thuyết phục.
* **Minh họa:** Bổ sung hình ảnh, sơ đồ, biểu đồ minh họa phù hợp để làm rõ các khái niệm, quy trình hoặc giải pháp đề xuất. Các hình ảnh, sơ đồ phải có chú thích rõ ràng khi sử dụng
* **Trích dẫn tài liệu tham khảo:** Trích dẫn tất cả các tài liệu tham khảo (sách, bài báo khoa học, website, quy định pháp luật, v.v.) theo chuẩn Vancouver ở cuối báo cáo.

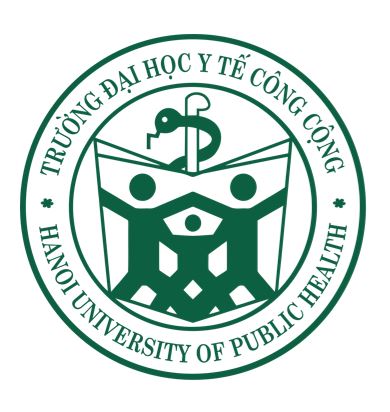
## 4. Đề xuất chủ đề bài tập

Sinh viên có thể lựa chọn một trong các chủ đề dưới đây hoặc đề xuất chủ đề khác phù hợp với nội dung môn học và được giảng viên chấp thuận. Các chủ đề này khuyến khích sinh viên đi sâu vào các vấn đề thực tiễn và áp dụng kiến thức đã học.

|  |  |  |
| --- | --- | --- |
| Nhóm 1 | Xây dựng mô hình hỗ trợ chẩn đoán bệnh tim |  |
|  | <https://www.kaggle.com/datasets/data855/heart-disease/data> |
|  | <https://www.kaggle.com/code/shahrinalam/predicting-presence-of-heart-diseases-using-ml-alg> |
| Nhóm 2 | Xây dựng mô hình hỗ trợ chẩn đoán bệnh ung thư vú |
|  | <https://www.kaggle.com/datasets/reihanenamdari/breast-cancer> |
|  | <https://www.kaggle.com/code/chanchal24/breast-cancer-prediction-with-7-models-98-acc> |
|  | <https://www.kaggle.com/code/buddhiniw/breast-cancer-p> |
| Nhóm 3 | Xây dựng mô hình hỗ trợ chẩn đoán bệnh tiểu đường |
|  | <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database> |
|  | <https://www.kaggle.com/code/shrutimechlearn/step-by-step-diabetes-classification-knn-detailed> |
| Nhóm 4 | Xây dựng mô hình dự đoán bệnh để cung cấp chẩn đoán chính xác và sớm dựa trên triệu chứng của bệnh nhân. |
|  | [Disease Prediction Using Machine Learning](https://www.geeksforgeeks.org/machine-learning/disease-prediction-using-machine-learning/) | | |  |  |
|  | link: https://www.geeksforgeeks.org/machine-learning/disease-prediction-using-machine-learning/ |
| Nhóm 5 | Xây dựng mô hình để phân loại bệnh lao qua X-quang ngực |
|  | [Tuberculosis (TB) Chest X-ray Database](https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset) | | |  |  |
|  | Link: https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset |
| Nhóm 6 | Xây dựng mô hình chuẩn đoán bệnh Parkinson liên quan rối loạn thần kinh tiến triển ảnh hưởng đến chuyển động |
|  | [Parkinson Disease Prediction using Machine Learning](https://www.geeksforgeeks.org/machine-learning/parkinson-disease-prediction-using-machine-learning-python/) |  |
|  | Link: https://www.geeksforgeeks.org/machine-learning/parkinson-disease-prediction-using-machine-learning-python/ |
| Nhóm 7 | Xây dựng mô hình phát hiện ung thư da sử dụng một bộ dữ liệu chứa hình ảnh cho hai loại là ác tính hoặc lành tính. |
|  | <https://www.geeksforgeeks.org/deep-learning/skin-cancer-detection-using-tensorflow/> |
| Nhóm 8 | Xây dựng mô hình dự đoán bệnh tim sử dụng mạng nơ-ron nhân tạo |  |
|  | <https://www.geeksforgeeks.org/deep-learning/heart-disease-prediction-using-ann/> |

**TRƯỜNG ĐẠI HỌC Y TÉ CÔNG CỘNG**

**KHOA CÁC KHOA HỌC CƠ BẢN**



**BÀI TẬP HẾT MÔN**

**HỌC SÂU**

**Lớp CNCQKHDL1-1A**

**NHÓM SỐ:**

**Chủ đề:**

**Thành viên: (ghi rõ mã sinh viên và họ và tên)**

**Hà Nội, 2025**

**KHUNG CHẤM ĐIỂM BÀI VIẾT VÀ BÁO CÁO**

**Nhóm:**

**Chủ đề:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TT** | **TIÊU CHUẨN ĐÁNH GIÁ** | **KHUNG ĐIỂM** | **ĐIỂM CHẤM** | | |
| **Nhóm** | **Cá nhân** | **Tổng** |
| **A** | **ĐIỂM CHUNG CỦA NHÓM** |  |  |  |  |
| **1** | **Trình bày, bố cục** | **10** | .. |  |  |
| Trình bày gọn gàng theo đúng hướng dẫn. | 5 |  |  |  |
| Sử dụng ngôn ngữ dễ hiểu, ít lỗi chính tả | 5 |  |  |  |
| **2** | **Nội dung trình bày** | **55** | .. |  |  |
| Đặt vấn đề ngắn gọn, dễ hiểu, | 5 |  |  |  |
| Nêu rõ được vấn đề/chủ đề của bài viết. | 5 |  |  |  |
| Nguồn thông tin phù hợp, có số liệu, thông tin minh chứng đầy đủ. | 5 |  |  |  |
| Đánh giá toàn diện các biện pháp an ninh và an toàn trong hệ thống thông tin y tế | 15 |  |  |  |
| Áp dụng chính xác các quy định pháp lý có liên quan | 15 |  |  |  |
| Giải pháp sáng tạo và thực tiễn | 5 |  |  |  |
| Các tài liệu tham khảo được liệt kê đầy đủ/đúng cách. | 5 |  |  |  |
|  | ***Tổng phần A*** | **65** | *..* |  |  |
| **B** | **ĐIỂM MỖI THÀNH VIÊN** | **35** |  |  |  |
|  | *Trả lời câu hỏi liên quan đến nội dung bài tập rõ ràng, mạch lạc và thuyết phục\** |  |  |  |  |
|  | *SV1: họ và tên* | 35 |  | *..* | *..* |
|  | *SV2: họ và tên* | 35 |  | *..* | *..* |
|  | *SV3: họ và tên* | 35 |  | *..* | *..* |
|  | *SV4: họ và tên* | 35 |  | *..* | *..* |
|  | *SV5: họ và tên* | 35 |  | *..* | *..* |

*\* Ghi rõ tên mỗi sinh viên*

**Nhận xét:**

**Ngày tháng năm 202**

**Giảng viên chấm**