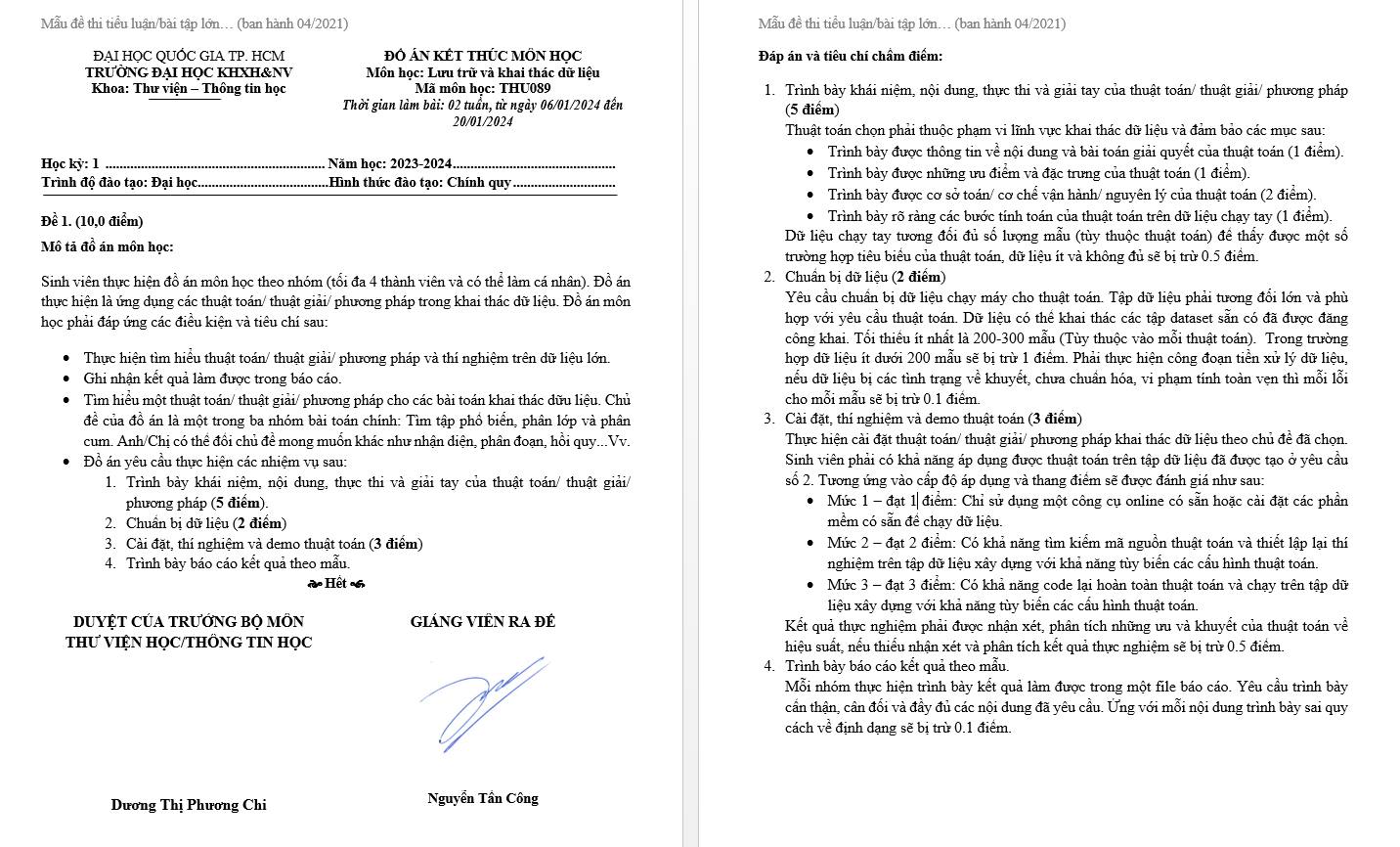
**ORGANIZATION**



Hình 1. Đề bài

1. TỔNG QUAN
2. Giới thiệu nhóm /lý do chọn dự án
3. Sơ đồ tổng quan trình bày (phân rã chức năng)
4. Công cụ thực hiện
   1. Giới thiệu Python
   2. Các thư viện sử dụng
5. Thông tin dự án của nhóm và phân chia nhiệm vụ
6. LÝ THUYẾT THUẬT TOÁN
7. Giới thiệu về data mining và bài toán phân loại/phân cụm/tìm luật kết hợp
8. Thuật toán K – nearest-neibourgh
   1. Giới thiệu về thuật toán và bài toán mà thuật toán giải quyết
      1. Tổng quan về thuật toán
      2. Vấn đề mà bài toán giải quyết ( về mặt lý thuyết và thực tế)
   2. Đặc trưng của thuật toán ( ưu điểm và nhược điểm đối với từng trường hợp cụ thể)
   3. Nguyên lý thuật toán ( trình bày sơ đồ, công thức và giải thích từng phần)
   4. Minh họa thuật toán
      1. Dữ liệu (dạng bảng)
      2. Các bước thực hiện (trình bày các bước và có vẽ sơ đồ)
      3. Kết quả
9. CHUẨN BỊ DỮ LIỆU
10. Giới thiệu dữ liệu
11. Tiền xử lý dữ liệu
    1. Làm sạch dữ liệu
    2. Biến đổi dữ liệu
12. Mô tả dữ liệu (thống kê mô tả)
13. KẾT QUẢ VÀ ĐÁNH GIÁ THUẬT TOÁN
14. Xây dựng thuật toán
15. Áp dụng thuật toán cho bài toán
16. So sánh thuật toán tự xây dựng và sử dụng thư viện
17. Thử nghiệm thuật toán
    1. Thử nghiệm với sự thay đổi của k
    2. Thử nghiệm với sự thay đổi của p
18. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN
19. Kết luận
20. Hướng phát triển
21. TÀI LIỆU THAM KHẢO