

# Đề xuất đề tài : Sử dụng thuật toán K-nearest neighbors dự đoán bệnh tiểu đường.

- **Thành viên trong nhóm:**

Trần Văn Viên	20164636
Dương Minh Quang	20163280
Lê Duy Hưng	20162009
Lê Trường Giang	10161175

- **Mô tả bài toán thực tế:**

- **Mục đích:**

- Hiểu machine learning là gì ?
- Áp dụng machine learning vào các vấn đề trong thực tế. Cụ thể ở trong phần này chúng em áp dụng trong việc kiểm tra bệnh tiểu đường.

- **Yêu cầu đạt được:**

- Từ dữ liệu của bệnh nhân chuẩn đoán được bệnh nhân có mắc bệnh tiểu đường hay không.

- **Các đặc điểm quan trọng:**

- Bộ dữ liệu tiểu đường được lấy từ UCI Machine Learning Repository. Bộ dữ liệu bao gồm 768 điểm dữ liệu với 9 thuộc tính.
- K-nearest neighbors là một trong những thuật toán supervised-learning đơn giản.

- **Sơ lược giải pháp:**

- Sử dụng thư viện pandas để đọc dữ liệu từ bộ dữ liệu.
- Chia dữ liệu làm hai phần: input và output.
- Chia dữ liệu làm hai phần: train và test.
- Xây dựng và train mô hình.
- Kiểm tra so với tập test.

- **Input:**

- Dữ liệu bệnh nhân

- **Output:**

- Bệnh nhân có mắc bệnh tiểu đường hay không.