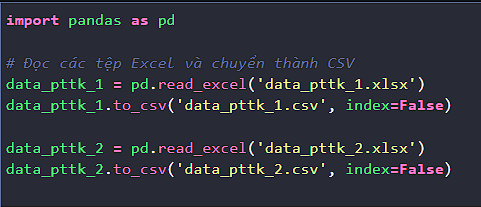
Họ và tên: Quách Xuân Phúc

MSV: B20DCCN513

Source code: [Source Code](https://drive.google.com/file/d/1iUza-2BXvPc3TO-0ITudJB9ww48zOpe4/view?usp=drive_link)

1. Sử dụng pandas để chuyển file excel sang csv



1. Gộp các file csv thành 1 file data\_pttk.csv

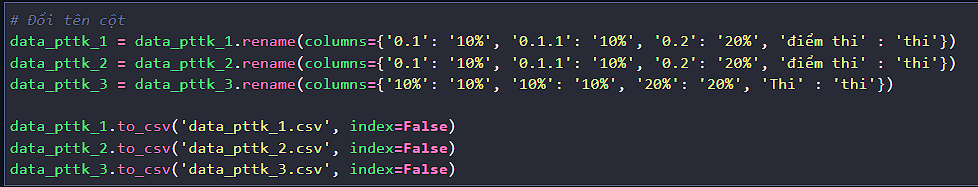
A screen shot of a computer code

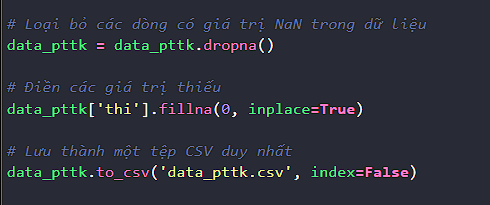
Description automatically generated

1. Thực hiện các bước tiền xử lý

A computer screen with text

Description automatically generated

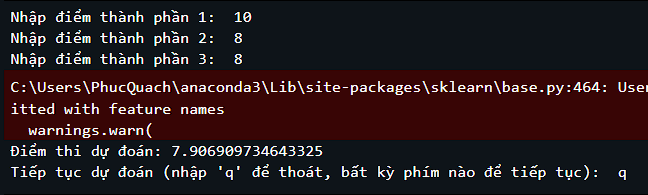




1. Sử dụng các kỹ thuật ML cơ bản (Chap 12) để dự đoán điểm khi nhập các điểm thành phần.

* Support Vector Machine

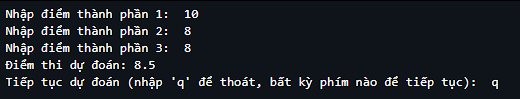




* K Nearest Neightbors

A screen shot of a computer program

Description automatically generated



1. Sử dụng Deep learning với Linear regression để dự đoán điểm thi

A screen shot of a computer program

Description automatically generated

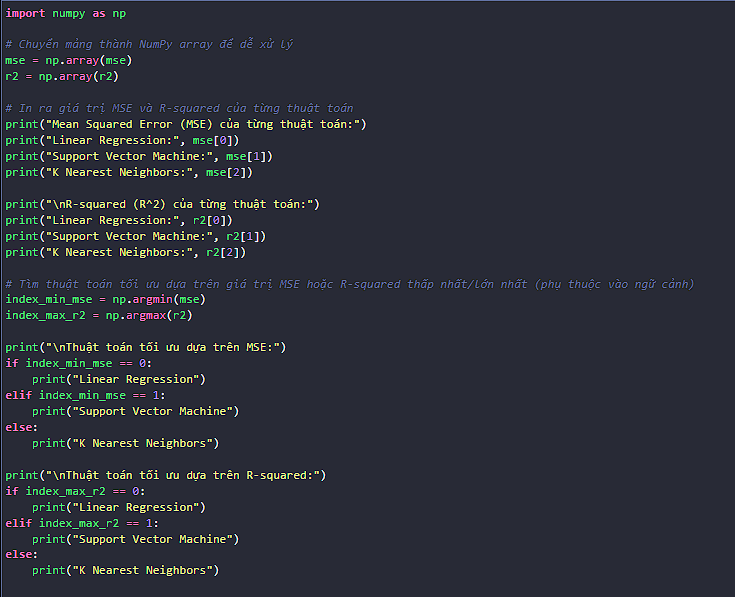
A screen shot of a computer

Description automatically generated

1. Đánh giá kết quả

A black rectangular sign with green and blue text

Description automatically generated



A computer screen shot of a black background with white text

Description automatically generated

\* Kết luận:

Linear Regression có hiệu suất tốt nhất với MSE thấp nhất trong 3 phương pháp. Ngoài ra, R-squared của Linear Regression cũng là cao nhất, cho thấy mô hình này phù hợp tốt với dữ liệu.

SVM (Support Vector Machine) có hiệu suất trung bình với MSE nằm ở mức giữa Linear Regression và K Nearest Neighbors. Tuy nhiên, R-squared của SVM cũng là cao, chỉ thấp hơn so với Linear Regression.

K Nearest Neighbors có hiệu suất kém hơn so với SVM và Linear Regression với MSE cao hơn và R-squared thấp hơn. Điều này cho thấy mô hình K Nearest Neighbors không khớp tốt với dữ liệu.

Tóm lại, Linear Regression được coi là phương pháp có hiệu suất tốt nhất trong ba phương pháp được đánh giá, theo sau là SVM và sau cùng là K Nearest Neighbors.