Assignment 3\_Taeyeun Oum

# Introduction

The goal of this assignment was to perform a binary classification take and show final classification accuracies by using 10-fold cross-validation for different cases such as different training epochs. All students were required to use the data called processed.cleveland.data at <http://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/>.

# Realization

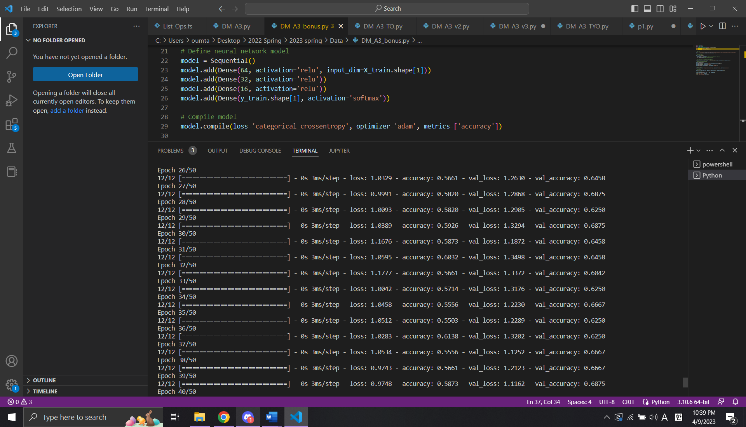
The classification was realized using python running on Visual Studio Code. The data was pre-processed by using the SVC functions from the sklearn.svm python library based on the final column of the data. Zeroes were converted into Healthy(+1), whereas all non-zero values were converted into Sick(-1). The following image displays the result.

A screenshot of a computer

Description automatically generated

# bonus

A neural network instead of a SVM model was utilized for classification. In my code, I first loaded and preprocessed the data, including dropping any rows with missing values and converting the class labels to 0-based indexing and one-hot encoding. Then I split the data into training and testing sets. Finally, I defined the neural network model using Keras' Sequential API. The model consists of 4 dense layers, with ReLU activation in the hidden layers and softmax activation in the output layer for multiclass classification. The model was compiled using categorical cross-entropy loss and the Adam optimizer.



Overall, the result showed lower accuracies of the pre-processed data compared to the one produced using the SVM model by approximately 0.06 lower.