The Hong Kong Polytechnic University

Department of Electrical and Electronics Engineering

EIE4430 Honours Project

2024-2025 Semester 1

Student Name: Chan Hou Ting Constant (21034774d)

Project Title: **Machine learning model to predict the risk of diabetes**

Progress Report (1/10/2024)

**Works did in past month**

I tried to implement Logistic Regression and XGBoost classifier to test the performance of the model. For the Logistic Regression, I applied K-fold cross-validation to ensure the train data and test data could be fully utilized and learned. Moreover, I used GridSearch function to get the best parameters by setting a grid of parameters in both models. During the tuning in XGBoost, I found that the accuracy is the same but the recall is decreased after tuned. Maybe needs to adjust the hyperparameters to get a better performance.

**Logistic Regression**

一張含有 文字, 螢幕擷取畫面 的圖片

自動產生的描述

**XGBoost (Before Tunning)**

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

**XGBoost (GridSearch function)**

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述

**XGBoost (After Tunning)**

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述