Data Mining Project 2

資工碩一P76084300 施逢怡

# 一、簡介

1. 專案目標:分析學生的各種資料來分類學生(好學生、表現沒這麼好的學生)

2. 資料: Student Performance Data Set

(<https://archive.ics.uci.edu/ml/datasets/student%2Bperformance>)

3. 方法: Decision tree

# 二、Design a set of rules to classify data

我認為”好學生”應該要有三種特質，分別是

1:每天溫習功課的時間至少要有2小時

2.不可以有被當的科目

3.成績要中上以上

Dataset裡面相對應的Attribute就是”studytime”, “failures”, “grade”

* studytime裡面的數值就是讀書時間(**1** - <2 hours, **2** - 2 to 5 hours, **3** - 5 to 10 hours, or **4** - >10 hours)
* failures則是被當科目的數目，n if 1<=n<3, else 4
* grade則是分為”good”和”bad”，是以他們的成績最分類

所以最後的Rules:

1. **studytime >= 2**
2. **failures == 0**
3. **grade == ”good”**