

Data Privacy and Security HW1

How do you perturb the data?

使用 Mondrian 演算法

What is your machine learning model used in this experiment? Besides, provide some details/parameters of this model.

NearestCentroid

每個類別均以其質心表示，測試樣本被分類為具有最接近質心的類別。

<https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.NearestCentroid.html>

使用模型預設數值

Effectiveness measure: Misclassification Error, Accuracy, Precision, Recall, and AUC

分數	原本資料集	K anonymity
accuracy_score	0.61	0.32
recall_score	0.78	0.42
precision_score	0.62	0.16

```
PS C:\Users\Yuuki\Desktop\K-Anonymity-master>
accuracy_score : 0.6098290598290599
recall_score   : 0.7798448532144183
precision_score: 0.6170784227849446
PS C:\Users\Yuuki\Desktop\K-Anonymity-master>
accuracy_score : 0.3248048568950564
recall_score   : 0.41774526355280434
precision_score: 0.16303382729944352
```

Privacy level: the value of K

使用K值為3



Dataset:

國際原油價格

<https://data.gov.tw/dataset/6029>