### Original data + encode all columns + select all columns for tranning

### 一張含有 文字, 收據 的圖片 自動產生的描述

### F class

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

F class max = 0.269

<https://machinelearningmastery.com/feature-selection-machine-learning-python/>

F\_class 3個fea

一張含有 文字 的圖片

自動產生的描述 一張含有 文字, 收據 的圖片

自動產生的描述

F\_class 5個fea

一張含有 文字 的圖片

自動產生的描述 一張含有 文字, 收據 的圖片

自動產生的描述

F\_class 8個fea

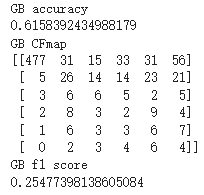
一張含有 文字 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述

F\_class 10個fea

一張含有 文字 的圖片

自動產生的描述 

F\_class 20個fea

一張含有 桌 的圖片

自動產生的描述 一張含有 文字, 收據 的圖片

自動產生的描述

### LogisticRegression

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



LogisticRegression max =0.155

https://scikit-learn.org/stable/modules/generated/sklearn.feature\_selection.SelectFromModel.html

一張含有 文字 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述

### RFE



RFE max = 0.2848

<https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html#sklearn.feature_selection.RFE>

n = 5

一張含有 文字 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述

N = 10

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

N = 12

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

N = 15

**一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述**

**N = 20**

**一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述**

嵌入式（embedded）Max = 0.269

SelectFromModel選取特征

<https://www.itread01.com/content/1527535241.html>

一張含有 文字 的圖片

自動產生的描述 一張含有 桌 的圖片

自動產生的描述

### KNeighbors

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



KNeighbors max = 0.30 -> kaggle:0.27

– K = 10

https://scikit-learn.org/stable/modules/generated/sklearn.feature\_selection.SequentialFeatureSelector.html#sklearn.feature\_selection.SequentialFeatureSelector

一張含有 桌 的圖片

自動產生的描述 一張含有 桌 的圖片

自動產生的描述

– K = 15

一張含有 桌 的圖片

自動產生的描述 一張含有 文字, 收據 的圖片

自動產生的描述

K = 18 //pred(4)

一張含有 桌 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述



K = 20

一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

### Tree-based

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



<https://scikit-learn.org/stable/modules/feature_selection.html#univariate-feature-selection>

n\_estimator = 50

一張含有 桌 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述

n\_estimator = 40

一張含有 桌 的圖片

自動產生的描述一張含有 文字, 收據 的圖片

自動產生的描述

n\_estimator = 30

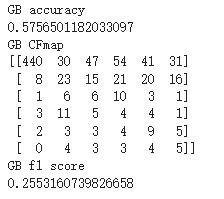
一張含有 桌 的圖片

自動產生的描述一張含有 桌 的圖片

自動產生的描述

n\_estimator = 20

一張含有 桌 的圖片

自動產生的描述 

n\_estimator = 10

一張含有 桌 的圖片

自動產生的描述 一張含有 文字, 收據 的圖片

自動產生的描述

