

AI應用在國中學業 成就預測

以實作導向探討





利用R統計軟體轉檔成CSV

```
library(foreign)
```

```
mydata <- read.dta("C:\data\資料轉檔\data\mydata.dta")
```

```
#read.dta Read Stata Binary Files
```

```
#read.spss Read an SPSS Data File
```

```
write.table(mydata,file="C:\data\資料轉檔\csv\testing.csv",sep=";",row.names=F, na = "NA")
```



使用google colab 平台實作

<https://colab.research.google.com/>



```
from sklearn.preprocessing import OneHotEncoder  
enc = OneHotEncoder(handle_unknown='ignore')
```

Scikit-learn 機器學習-演算法及資料集合

OneHotEncoder則是適用在無序型的類別型資料，如性別、婚姻等等

(LabelEncoder有序型的類別型資料，如名次、分數等等)

handle_unknown='ignore'

對於沒有出現在訓練集中的類別，轉換時全用0替代



```
enc.fit_transform(df_basics)
enc.categories_
```

```
fit_transform
```

計算特徵數據的均值 μ 和方差 σ^2 ，再使用下面的公式進行歸一化

分類特徵轉換為分類數值

```
.categories_
```

查看各個特徵中有哪些類別別例：婚姻狀態有- 合理值、其他、同居、已婚、未婚等，配偶已過世、離婚或分居。



```
X = enc.transform(df_basics).toarray()  
X.shape
```

```
transform(df_basics).toarray()
```

將df_basics轉換成陣列

X.shape查看X矩陣

(18132, 165) 18132*165的矩陣



```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Input, Dense, Activation, Dropout
```

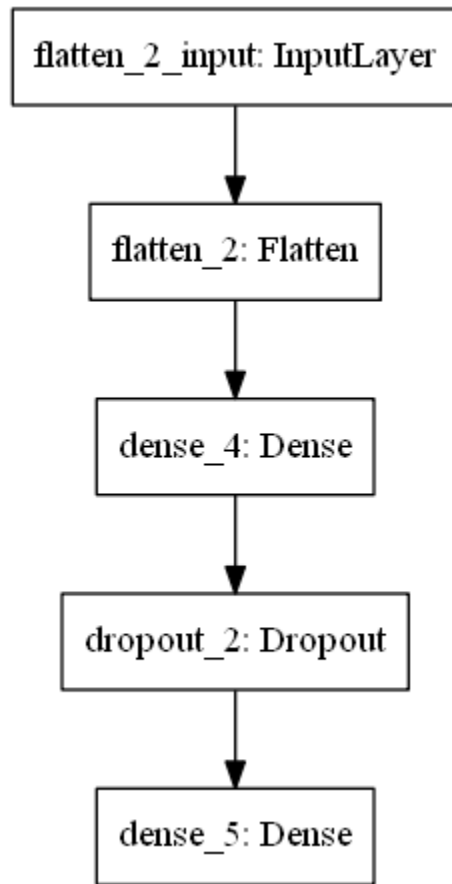
Sequential 顺序模型


Input 輸入的規格(通常只有第一層)

Dense層

全連接層(Dense)

用於將從該層之前的幾個特徵抽取層獲得的特徵表達向量映射到下一層，或者映射到最終的 **softmax**層。





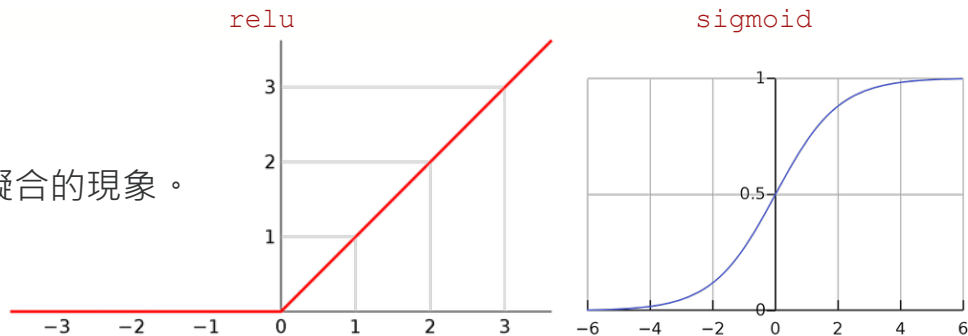
```
model.add(Dense(units=256, activation='relu', input_dim=155))
#斷開10%的神經元
model.add(Dropout(0.1))
model.add(Dense(units=256, activation='relu'))
model.add(Dropout(0.3))
model.add(Dense(units=256, activation='sigmoid'))
```

Activation 對一個層施加激活函數

若輸出值為正值，該輸出值經過ReLU函數轉換仍可正常輸出，若輸出值為負值，則經過ReLU函數轉換後，輸出值為0

Dropout層

Dropout層：去除極端值，達到避免過度擬合的現象。





```
#fit the Neural Network Model
```

```
model.compile(loss='mean_squared_error', optimizer='RMSprop')
```

```
model.fit(X, y, epochs=256, batch_size=256, verbose=1)
```

mean_squared_error

預測值與實際值的差距之平均值・誤差

RMSprop參數是一種自行修正方式・朝正確方向前進

epochs=256: 執行256次訓練週期(ee po k)

batch_size=256: 每次一批256筆資料

verbose

訓練時的視覺顯示方式。 0 = 不顯示 1 = 進度條, 2 = 每次丟一行。

