1. 資料前處理

一張含有 文字 的圖片

自動產生的描述

* 處理data
  + Score: >=4的轉成1，<為0
  + Text: 去除‘.’

一張含有 文字 的圖片

自動產生的描述

* 將word做處理
  + Lemma\_stem\_text: 把詞性、單複數作轉換
  + Negation\_abbreviated\_to\_standard: 將負面詞性作轉換
  + Review\_to\_words: 將words丟進去做轉換

c. 文字探勘前處理，將文字轉換成向量，請實作 tf-idf 及 word2vec 並進行比較

Tf-idf



* 帶入tfidfVectorizer並將轉換完的結果存入array

Word2vec

一張含有 文字 的圖片

自動產生的描述

* 帶入word2vec

2. 建模：

a. 使用 Random forest進行分類

b.  請寫自行撰寫function進行k-fold cross-validation(不可使用套件)並計算Accuracy



1. 評估模型，請計算k=4的Accuracy

一張含有 桌 的圖片

自動產生的描述