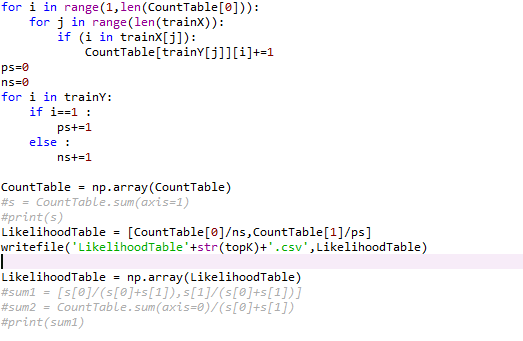
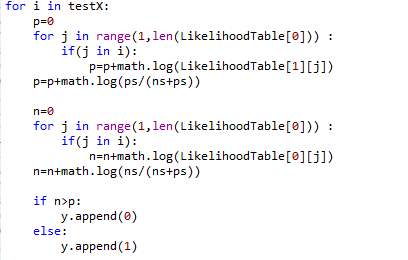
**學號:P76074517 姓名:林怡瑄**

先將大於K的字濾掉，分別統計正負評train data內每個字出現的次數(CountTable)，之後將正評與負評每個字統計的數量分別除以正評與負評所有篇數的總和得到𝑃 (𝑋 |𝑌)(LikelihoodTable) ，TOPK 的LikelihoodTable有寫在對應的csv檔裡，而所有正評/所有篇數與所有負評/所有篇數得到P(y=1)與P(y=0)的機率。

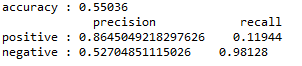


之後用條件機率的公式分別算出正跟負的機率比較大者為我預測的結果，輸出預測結果在predict檔內，之後計算accuracy, precision, recall。

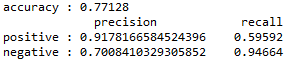


**結果截圖:**

**Top100**



**Top1000**



**Top10000**

