

110590007 資工系大三 白宸安

一、why we take log when implement the Bayesian classifier?

取 log 的理由是為了讓原本要進行的乘法可以轉變成做加法就好，這樣可以方便計算。

二、the difference between Naïve Bayesian and Gaussian Naïve Bayesian classifier?

兩者的不同在於，Naïve Bayesian 是用於處理離散的資料，而 Gaussian Naïve Bayesian 則是透過 Normal distribution 的 probability density function (PDF)來處理連續的資料。實作上的主要差異是在 likelihood 的部分。

三、the difficulty you encountered :

在寫這次作業時遇到的困難點主要有：

1. 算式很細密，要注意不要寫錯。
2. 數學式子有點難搞懂。
3. 在 advanced part 時，一開始做出的 Gaussian Naïve Bayesian classifier 效果很不好。

四、Summarize how you solve the difficulty and your reflections :

1. 謹慎、仔細地寫算式，避免錯誤。
2. 去了解背後的原理才能好實作，比如 Naïve Bayesian 的 Bayes Theorem 和 f1-score 等。
3. 有使用 remove training set outliers 和修改 prior 的 weight 來優化，最後有優化出還不錯的效果。

心得：這次作業體感上比前一次簡單一點點，或許是這次有提供比較多已經做好的模板的關係。Classifier 的部分我覺得很有趣，我對這個也很有興趣，這次寫得還蠻開心的，尤其是後來發現調整 prior 的 weight 竟然可以優化這麼多感到很震驚，不太確定這是個通則還是因為這次的訓練資料剛好有這樣的特性？之後也許可以再研究看看。最後謝謝助教們幫忙批改作業！