

1. 總結一下這個項目的目標以及機器學習如何有助於實現它。作為答案的一部分，提供有關數據集的一些背景知識以及如何使用它來回答項目問題。當你得到它時，數據中是否有任何異常值，你是如何處理這些異常值的？[關鍵字：“EDA”，“Outlier”]
2. 您最終在 POI 的偵測中使用了哪些特徵 (features)，以及您使用哪些方法來選擇它們？你有沒有做任何 data scaling？為什麼或者為什麼不？您應該嘗試自己設計一些新的特徵 - 解釋您嘗試製作該特徵的原因及其背後的基本原理。(您不一定要在最終分析中使用它，只需對其進行設計和測試) 在 feature selection 步驟中，如果您使用了決策樹之類的算法，請同時提供您 feature importance 的數值 [關鍵字：“create feature”，“feature selection”，“normalization”]
3. 您最終使用了什麼算法？您嘗試了哪些？算法之間的模型性能如何不同？[關鍵字：“modeling”]
4. 調整算法的超參數 (hyper-parameter) 是什麼意思，如果你做得不好怎麼辦？你是如何調整特定算法的參數的？你調整了什麼參數？（有些算法沒有你需要調整的參數 - 如果你選擇的那個是這種情況，請確定並簡要說明你將如何為不是你最終選擇的模型或不同的模型做到這一點。利用參數調整，例如決策樹分類器）。[關鍵字：“Hyper-parameter tuning”]
5. 什麼是驗證，如果你做錯了，你可以犯下的經典錯誤是什麼？您是如何驗證分析的？[關鍵字：“Hyper-parameter tuning”]
6. 給出至少 2 個評估指標和每個評估指標的平均表現。解釋您選擇該指標的原因。[關鍵字：“Evaluation metrics”]