عزيزي الطالب قم بالتفكير بإجابات هذه الأسئلة أثناء حل ما تبقى من أجزاء الوظيفة والعمل على المشروع وحاول أن تفكر بتفسير إجاباتك:

- 1. هل يجب إعادة تقسيم الداتا إلى مجموعة تدريب واختبار إذا كانت بالفعل مقسمة؟
  - 2. هل تحليل النصوص قبل البدء بالنمذجة رفاهية فكرية؟
- في حال لم تكن البيانات مقسمة هل نقسم الداتا أولاً أو نقوم بالمعالجة والتنظيف أولاً ثم نقسم؟
  - 4. هل تجنب تسريب المعلومات data leakage يعنى عدم معالجة مجموعة الاختبار النهائية؟
  - 5. هل النموذج الذي ستحصل عليه بعد عملية التدريب بالنسبة لخوارزمية معينة هو أفضل نموذج ممكن -بالنسبة للخوارزمية- ؟
    - 6. هل يمكنك مقارنة نتائج أكثر من خوارزمية على مجموعات مختلفة من المعطيات؟
      - 7. هل تحديد النموذج البدائي baseline model مرتبط بالمسألة أم محايد؟
        - هل تُكتب التقارير لأنها مطلوبة فقط؟