**Інформація про набір даних:**

Два набори даних пов’язані з червоним і білим варіантами португальського вина «Vinho Verde». Через проблеми конфіденційності та логістики доступні лише фізико-хімічні (вхідні) та сенсорні (вихідні) змінні (наприклад, немає даних про типи винограду, бренд вина, ціну продажу вина тощо).

Ці набори даних можна розглядати як завдання класифікації. Класи впорядковані і не збалансовані (наприклад, нормальних вин набагато більше, ніж відмінних або поганих). Алгоритми виявлення викидів можна використовувати для виявлення кількох відмінних або поганих вин.

**Інформація про атрибути:**

Вхідні змінні (на основі фізико-хімічних тестів):

1 - фіксована кислотність

2 - летюча кислотність

3 - лимонна кислота

4 - залишковий цукор

5 - хлориди

6 - вільний діоксид сірки

7 - сірчистий газ загальний

8 - щільність

9 - pH

10 - сульфати

11 - спирт

Вихідна змінна (на основі сенсорних даних):

12 - якість (оцінка від 0 до 10)

**Завдання:**

Використовуючи різні моделі задач класифікації та регресії вибрати ту, яка (на вашу думку) буде найкраще робити передбачення.

Детальніше:

1. Провести аналіз атрибутів (змінних), визначити їх тип (дискретний чи неперервний).
2. Провести попередню обробку даних (візуалізувати дані, позбутися пропущених значень, нормувати чи логарифмувати дані, тощо).
3. Дослідити зв’язки між змінними.
4. Використовувати крос-валідацію для навчання та оцінки моделі.
5. Вибрати що найменше 5 класифікаційних (регресійних) моделей, серед яких врешті знайти найточнішу.
6. Вітаються будь-які інші кроки, що сприяють покращенню якості моделі.