

*Бізнес-задача та обґрунтування вибору датасету.*

Датасет: [Wine Quality Dataset](#)

Кількість записів: 4898 (біле вино) + 1599 (червоне).

Кількість ознак: 11 фізико-хімічних параметрів (кислотність, вміст цукру, рН, сірка, алкоголь тощо).

Цільова змінна: quality – оцінка якості вина від 0 до 10 балів.

1. Регресійна підзадача:

- Мета – передбачити числове значення якості вина.
- Вхідні змінні: alcohol, volatile acidity, citric acid, sulphates тощо.
- Вихідна змінна: quality.
- Методи: поліноміальна регресія (для виявлення нелінійних залежностей) і Elastic Net (для відбору найважливіших ознак та запобігання переобученню).
- Очікуваний результат – модель, здатна точно прогнозувати якість на основі фізико-хімічних характеристик.

2. Класифікаційна підзадача:

- Мета – класифікувати вино за рівнем якості: 3–4 – низька; 5–6 – середня; 7–8 – висока.
- Метод – логістична регресія.
- Очікуваний результат – автоматичне віднесення вина до категорії якості, що дозволить сегментувати продукцію для різних ринкових ніш.