

Лабораторна робота №3. Логістична регресія

Підгрупа №1

1. Відкрити та зчитати наданий файл з даними.
2. Визначити та вивести кількість записів, а також кількість полів та їх тип у завантаженому наборі даних.
3. Вивести перші 10 записів набору даних.
4. Додати до набору даних атрибут **HighQuality**, який дорівнює 1, якщо **quality** більше або дорівнює шести, та 0 в інших випадках.
5. Вивести перші 10 записів набору даних та видалити після цього атрибут **quality** з набору даних.
6. Перемішати набір даних та розділити його на навчальну (тренувальну) та тестову вибірку, використовуючи функцію **ShuffleSplit**.
7. Використовуючи відповідні функції бібліотеки **scikit-learn**, збудувати класифікаційну модель логістичної регресії та навчити її на тренувальній вибірці, вважаючи, що цільова характеристика визначається стовпчиком **HighQuality**, а всі інші виступають в ролі вихідних аргументів.
8. Обчислити класифікаційні метрики збудованої моделі для тренувальної та тестової вибірки. Представити результати роботи моделі на тестовій вибірці графічно.
9. З'ясувати вплив максимальної кількості ітерацій (від 5 до 5000) на результати класифікації. Результати представити графічно.
10. Проаналізувати ступінь впливу атрибутів на результат класифікації. Збудувати класифікаційну модель логістичної регресії, залишивши від 3 до 5 найбільш важливих атрибутів та порівняти її результати із моделлю з п. 7.