

Лабораторная работа №5

Классификация

1. Загрузить датасет "Iris" из библиотеки `scikitlearn` (`load_iris`).
2. Изучить структуру и содержание данных.
3. Провести предварительный анализ данных, включая проверку наличия пропущенных значений и корреляции между переменными.
4. Разбить данные на обучающую и тестовую выборки в соотношении 70/30.
5. Обучить модель логистической регрессии (`LogisticRegression`) на обучающей выборке.
6. Оценить качество модели на тестовой выборке с помощью метрик таких как точность, полнота, F1-мера и матрица ошибок (`accuracy_score`, `precision_score`, `recall_score`, `f1_score`, `confusion_matrix`).
7. Визуализировать результаты предсказания, сравнив исходные значения классов цветков и предсказанные значения.
8. Провести кросс-валидацию модели и оценить ее качество с помощью метрик точности, полноты, F1-меры и матрицы ошибок.
9. Попробовать улучшить качество модели путем вариации параметров модели или использования других моделей (например, k- ближайших соседей (`KNeighborsClassifier`), метода опорных векторов (`SVMClassifier`) или случайного леса (`RandomForestClassifier`)).
10. Сделать выводы о качестве модели и ее применимости для классификации типов цветков Iris.