

Задание для экзамена по курсу «Машинное обучение»

Требуется построить одну из изученных на курсе моделей машинного обучения для ранней диагностики диабета на основе имеющегося набора данных. Выберите один из следующих далее вариантов решения и выполните указанные пункты в задании. Использовать следующий источник данных:

<https://www.kaggle.com/datasets/tanshihjen/early-stage-diabetes-risk-prediction>

Вариант №1

Логистическая регрессия

1. Построить модель логистической регрессии с использованием библиотеки `sklearn`..
2. Провести перекрёстную проверку полученной модели, количество блоков подобрать самостоятельно.
3. Построить ROC-кривую для полученного классификатора.

Вариант №2

Метод k ближайших соседей

1. Выполнить построение модели классификации методом k ближайших соседей, используя библиотеку `sklearn`. Количество ближайших соседей k подобрать экспериментально.
2. Провести перекрёстную проверку полученной модели, количество блоков подобрать самостоятельно.
3. Построить ROC-кривую для полученного классификатора.

Вариант №3

Прогнозирование при помощи нейронной сети

1. Построить полносвязную нейронную сеть с тремя слоями с использованием библиотеки `sklearn`. Выполнить обучение сети методом прямого распространения ошибки для прогнозирования диабета.
2. Провести оценку полученной сети методом k-блочной перекрёстной проверки и визуализацию сети.