

## **Наименование работы:**

Задача классификации и её метрики качества

## **Цель работы:**

Ознакомиться с задачей регрессии, методами её решения и оценки на практических примерах

## **Задачи:**

- a) В соответствии с вариантом выбрать набор данных для задачи классификации;
- b) Обучить на выбранном наборе данных несколько различных моделей:
  - 1. логистическая регрессия;
  - 2. наивный байесовский классификатор;
  - 3. метод К-ближайших соседей;
  - 4. метод опорных векторов;
  - 5. дерево решений;
- c) С помощью обученных моделей сделать прогноз для тестовых данных;
- d) Оценить прогноз каждой из обученных моделей по следующим метрикам:
  - 1. точность;
  - 2. полнота;
  - 3. f1-мера;
- e) Предоставить полученные в ходе обучения параметры моделей:
  - 1. веса;
  - 2. граф дерева решений;
- f) Определить наиболее подходящую метрику для оценки качества классификации и обосновать свой выбор.

Для реализации использовать язык программирования Python 3.x и библиотеки Pandas, Numpy, Sklearn

## **Оформление результатов:**

Результаты лабораторной работы оформляются в виде отчета в формате PDF.

## **Структура отчета:**

- 1. Титульный лист;
- 2. Основная часть;
- 3. Заключение.

**В основной части** приводится описание выполнения каждой из поставленных задач в виде текста и скриншотов программного кода

**В заключении** приводятся практически значимые выводы по проделанной работе

Таблица 1 - наборы данных

№	Датасет	Описание
1	Walmart Recruiting: Trip Type Classification <a href="https://www.kaggle.com/c/walmart-recruiting-trip-type-classification/data">https://www.kaggle.com/c/walmart-recruiting-trip-type-classification/data</a>	Данные о посещениях магазина Walmart покупателями. Целевой признак: тип посещения магазина.
2	IEEE-CIS Fraud Detection <a href="https://www.kaggle.com/c/ieee-fraud-detection/data">https://www.kaggle.com/c/ieee-fraud-detection/data</a>	Данные об онлайн-транзакциях. Целевой признак: является ли транзакция мошеннической
3	Home Credit Default Risk <a href="https://www.kaggle.com/c/home-credit-default-risk/data">https://www.kaggle.com/c/home-credit-default-risk/data</a>	Данные о заемщиках банка. Целевой признак: способен ли заёмщик выплатить кредит

