

**Приложение 5**

**« » \_\_\_\_\_\_\_\_2021**

**Директор МОБУ СОШ№10 им.Д.Г.Новопашина**

**Шилова Н.А.**

**КОНКУРСНОЕ ЗАДАНИЕ**

***ДЛЯ*** ***МУНИЦИПАЛЬНОГО ЧЕМПИОНАТА***

**компетенции**

**№ F5 «МАШИННОЕ ОБУЧЕНИЕ И БОЛЬШИЕ ДАННЫЕ»**

*Конкурсное задание включает в себя следующие разделы:*

[1. Форма участия в конкурсе: 2](#_Toc66870131)

[2. Общее время на выполнение задания: 2](#_Toc66870132)

[3. Модули задания и необходимое время 2](#_Toc66870134)

[4. Критерии оценки. 3](#_Toc66870135)

1. **Форма участия в конкурсе**: Индивидуальный конкурс
2. **Общее время на выполнение задания:** 6 ч.

## Модули с описанием работ

Модуль 1: Предобработка данных и визуализация зависимостей

Данные представлены в одном из форматов: txt, pdf, doc, docx, xls, xslx, xml, csv. Необходимо преобразовать исходный набор данных к виду, пригодному для загрузки. Далее необходимо провести предобработку данных, например, разбить сложные атрибуты, очистить от технической информации, удалить спецсимволы. В обработанном наборе данных выделяются значимые атрибуты для дальнейшего анализа и построения модели машинного обучения.

Примерные этапы задания в модуле 1:

* 1. Предобработка данных и выделение значимых атрибутов;
  2. Разбиение сложных атрибутов;
  3. Дополнение недостающими данными;
  4. Предварительный анализ данных;
  5. Визуализация зависимостей данных.

Модуль 2: Построение модели машинного обучения

В этом модуле продолжается работа с данными, подготовленными в предыдущем модуле. На основе предобработанных данных предстоит построить модель машинного обучения. Как правило это модель кластеризации, классификации или регрессии. Необходимо определить точность работы модели и предпринять какие-либо действия для повышения точности.

Примерные этапы задания в модуле 2:

* 1. Визуализация зависимостей данных;
  2. Выбор модели классификации;
  3. Обучение модели классификации.

## Критерии оценки

**Таблица 1.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **№ п/п** | **Модуль, в котором используется критерий** | **Критерий** | **Время**  **выполнен ия**  **Модуля** | **Баллы** | | |
| **Судейс кие** | **Объек тивн**  **ые** | **Общие** |
| 1. | Модуль 1.  Предобработка данных и визуализация зависимостей | A. Подготовка репозитория  хранения данных | 3 часа | 0 | 7,5 | 7,5 |
| B. Предобработка больших данных | 1,5 | 3,5 | 5 |
| 2. | C.Подготовка обучающей и тестирующей  выборок | 2 | 8,5 | 10,5 |
| 3. | Модуль 2.  Построение модели машинного обучения | C.Подготовка обучающей и тестирующей  выборок | 3 часа | 0 | 4,5 | 4,5 |
| 4. | D. Разработка математического  аппарата | 2 | 17 | 19 |
| 5. | F. Документирование | 0 | 3,5 | 3,5 |
| **Итого** | | | |  |  |  |