

Конкурсное задание



Компетенция

Машинное обучение и большие данные

Конкурсное задание включает в себя следующие разделы:

1. Формы участия в конкурсе
2. Задание для конкурса
3. Набор данных

Количество часов на выполнение задания: 4 ч.

1. ФОРМЫ УЧАСТИЯ В КОНКУРСЕ

Индивидуальный конкурс.

2. ЗАДАНИЕ ДЛЯ КОНКУРСА

ВЫБОР АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ

Выберите алгоритм машинного обучения, необходимый для проведения классификации объявлений.

Входными параметрами являются все данные, хранящиеся в наборе данных, за исключением поля jobname. Данное поле является идентификатором класса.

РЕАЛИЗАЦИЯ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ

Реализуйте описанный в предыдущей задаче алгоритм машинного обучения.

ОБУЧЕНИЕ МОДЕЛИ

Обучите модель, используя данные, хранящиеся в наборе данных (разбиение на тестовую и обучающую выборки произведите самостоятельно).

Произведите апробацию модели, используя тестовую выборку.

Вычислите процент верного распознавания = (Правильно распознанные вакансии)/(общее количество вакансий)

ПРЕДОСТАВЛЕНИЕ РЕЗУЛЬТАТОВ РАБОТЫ

Подготовьте файл readme.md, содержащий следующие пункты:

- Обоснование выбора алгоритма;
- Способ разделения выборок;
- Результаты обучения (процент верного распознавания).

Сохраните файл в папку с исходными кодами разработанного приложения с именем user[номер жеребьевки].

3. НАБОР ДАННЫХ

Исходный набор данных представлен в файле data.csv