Муниципальное автономное общеобразовательное учреждение

«Лицей инновационных технологий»

**Пояснительная записка к решению задания отборочного этапа хакатона «IT-Education HACK» по направлению «Искусственный интеллект».**

Выполнили: Кучун Дмитрий, Свиридов Георгий, Павлюченко Артем, Павлов Александр

Руководитель проекта: Назина Александра Александровна, учитель

информатики и ИКТ.

Хабаровск, 2023 г.

Оглавление

**Элементы оглавления не найдены.**

# Введение

# Уровень 1

## Задание 1

Задание 1.1: Создание двух массивов с соответствующими названиями (cat\_features, num\_features). В cat\_features находятся названия категориальных признаков, а в num\_features числовых соответственно.

Задание 1.2: Была создана функция mistakes, в котором создает DataFrame с помощью библиотеки pandas, которая принимает данные из файла students\_data.csv (исходный файл таблицы). После чего на основе условий элементов каждого признака функция проверяет правильность исходных данных. В результате работы функция возвращает количество ошибок, учитывая также пропуски, но именно количество опечаток с ошибками выводится, используя функцию misses (задание 2.1).

Примечание: При нахождении ошибок и опечаток мы выводили каждую ошибку, относящуюся к тому или иному признаку. На основе полученных данных мы меняли регистр или символ на правильный, используя функцию replace().

Задание 1.3: Для определения ошибок и опечаток в таблице элементы некоторых признаков были приведены к единому типу (float, int)

## Задание 2

Задание 2.1: Создается функция misses, которая принимает как аргумент двумерный массив data с исходными данными из файла students\_data.csv. Далее она считает общее количество пропусков и возвращает это значение, а также она с помощью словаря main\_dict возвращает признаки, в которых были найдены пропуски.

Задание 2.2: Как уже было сказано, функция misses возвращает признаки, в которых были пропущены значения. Ими оказались: Dalc, Walc, cheating, famrel, romantic. В Dalc и Walc мы решили заполнять пропуски минимально возможным значением, потому что, если человек не заполнил этот признак, значит значение этого показателя равно минимальному - единице. В famrel прослеживается такая же логика. А в cheating и romantic работает так: если человек не указал в этой графе ничего, значит значение равно no. Все исправления пропусков прописаны в функции correction, которая далее помогает решать другие задания.

## Задание 3

Задание 3.1: Пропущено

Задание 3.2: