Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

«ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Цифровая кафедра

Лабораторная работа №1

по дисциплине «Системы искусственного интеллекта»

Выполнил:

студент

группы АУЗ-561

Жолобов А.В.

Проверил:

преподаватель

Фокин Р.О.

Волгоград, 2023 г.

|  |  |  |
| --- | --- | --- |
| Лабораторная работа №1 | Ф.И.О. | Жолобов А.В. |
| Группа | АУЗ-561 |
| Преподаватель | Фокин Р.О. |
| Дата сдачи | 12.11.2023 |

**Цель работы:**

Целью данной работы является изучение инструментов для запуска, отладки и тестирования кода Python, а также применение Python для предобработки данных.

**Ход работы:**

Пошаговое выполнение примера из презентации:

[Лабораторная работа №1](https://docs.google.com/presentation/d/1juW4Bp0e8hBDMvnp0khGU9wQ2yRUtMx1/edit#slide=id.p1).

**Выполнение индивидуального задания:**

1. Ссылка на датасет:

[student\_alcohol\_data](https://www.kaggle.com/datasets/mahmoudshogaa/student-alcholoh-data).

1. Описание датасета.

Для выполнения первой лабораторной работы был выбран датасет, содержащий 397 строк с информацией о студентах.

В рассматриваемом датасете представлены данные по следующим параметрам.

school - школа (бинарное: 'GP' - Gabriel Pereira или 'MS' - Mousinho da Silveira)

sex - пол студента (бинарное: 'F' - женский или 'M' - мужской)

age - возраст студента (числовой: от 15 до 22 лет)

address - тип домашнего адреса студента (бинарное: 'U' - городской или 'R' - сельский)

famsize - размер семьи (бинарное: 'LE3' - меньше или равно 3 или 'GT3' - больше 3)

Pstatus - статус сожительства родителей (бинарное: 'T' - живут вместе или 'A' - раздельно)

Medu - образование матери (числовое: 0 - отсутствует, 1 - начальное образование (4-й класс), 2 - с 5-го по 9-й класс, 3 - среднее образование или 4 - высшее образование)

Fedu - образование отца (числовое: 0 - отсутствует, 1 - начальное образование (4-й класс), 2 - с 5-го по 9-й класс, 3 - среднее образование или 4 - высшее образование)

Mjob - работа матери (номинальное: 'учитель', связано с 'здравоохранением', гражданская 'служба' (например, административная или полиция), 'дома' или 'другое')

Fjob - работа отца (номинальное: 'учитель', связано с 'здравоохранением', гражданская 'служба' (например, административная или полиция), 'дома' или 'другое')

reason - причина выбора этой школы (номинальное: близость к 'дому', 'репутация' школы, предпочтение 'курса' или 'другое')

guardian - опекун студента (номинальное: 'мать', 'отец' или 'другое')

traveltime - время в пути от дома до школы (числовое: 1 - менее 15 мин., 2 - 15-30 мин., 3 - 30 мин. до 1 часа или 4 - более 1 часа)

studytime - недельное время на учебу (числовое: 1 - менее 2 часов, 2 - 2-5 часов, 3 - 5-10 часов или 4 - более 10 часов)

failures - количество неудач в прошлых классах (числовое: n, если 1 <= n < 3, иначе 4)

schoolsup - дополнительная образовательная поддержка (бинарное: да или нет)

famsup - семейная образовательная поддержка (бинарное: да или нет)

paid - дополнительные платные занятия по предмету (математика или португальский) (бинарное: да или нет)

activities - внеклассные мероприятия (бинарное: да или нет)

nursery - посещал ли детский сад (бинарное: да или нет)

higher - хочет получить высшее образование (бинарное: да или нет)

internet - наличие интернета дома (бинарное: да или нет)

romantic - наличие романтических отношений (бинарное: да или нет)

famrel - качество семейных отношений (числовое: от 1 - очень плохо до 5 - отлично)

freetime - свободное время после школы (числовое: от 1 - очень мало до 5 - очень много)

goout - время, проведенное с друзьями (числовое: от 1 - очень мало до 5 - очень много)

Dalc - потребление алкоголя в будний день (числовое: от 1 - очень мало до 5 - очень много)

Walc - потребление алкоголя в выходные (числовое: от 1 - очень мало до 5 - очень много)

health - текущее состояние здоровья (числовое: от 1 - очень плохо до 5 - очень хорошо)

absences - количество пропущенных занятий в школе (числовое: от 0 до 93)

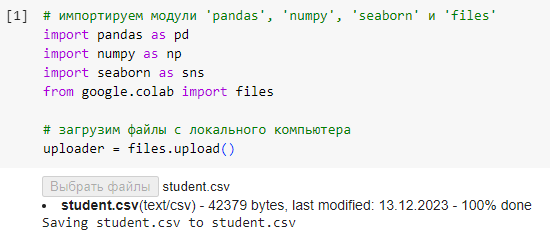
Эти оценки связаны с предметом курса, математикой или португальским:

G1 - оценка за первый период (числовое: от 0 до 20)

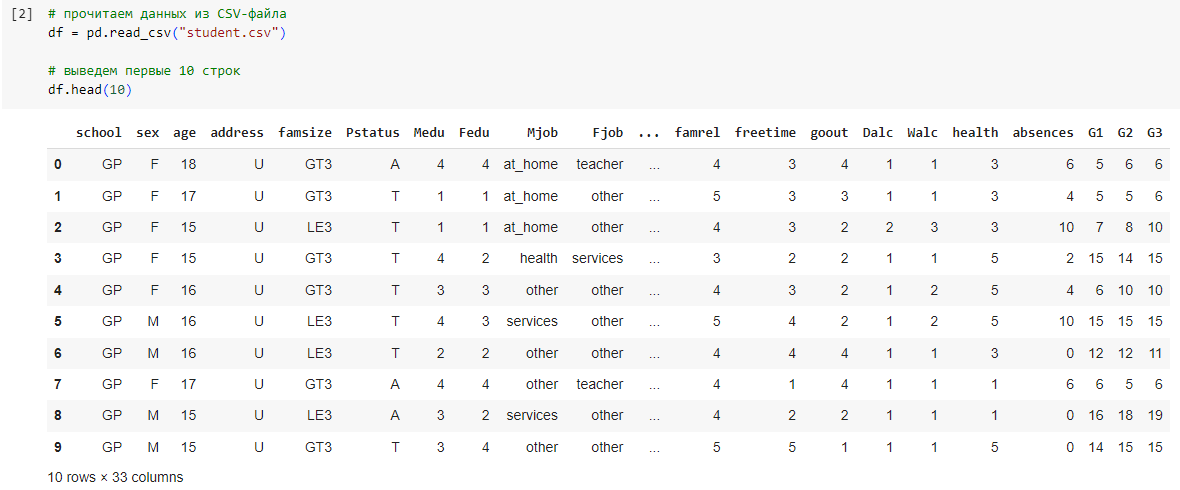
G2 - оценка за второй период (числовое: от 0 до 20)

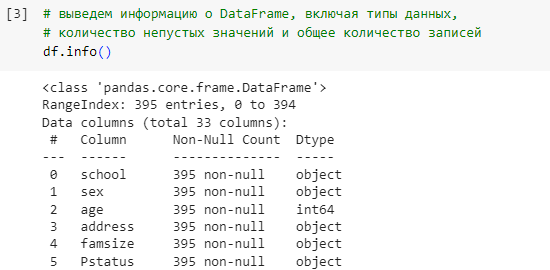
G3 - итоговая оценка (числовое: от 0 до 20, целевой показатель)

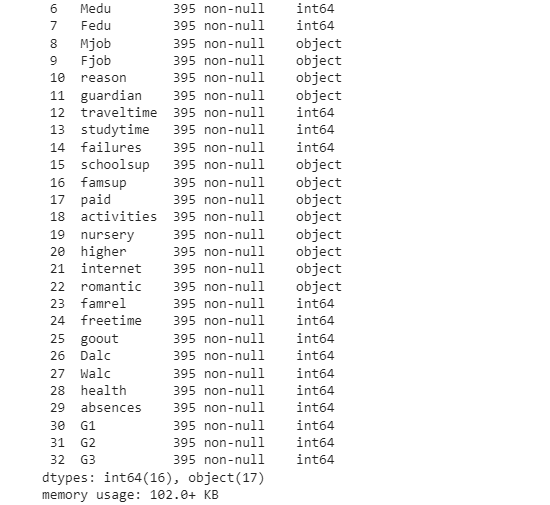
1. Загрузка датасета:



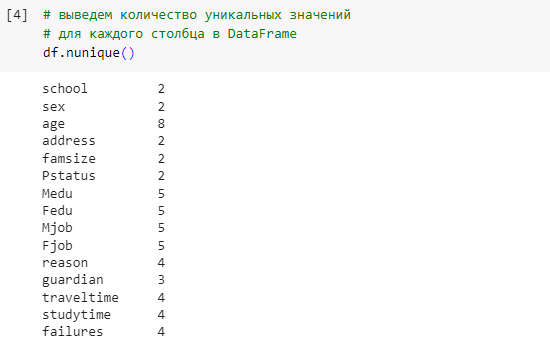
1. Предобработка данных:

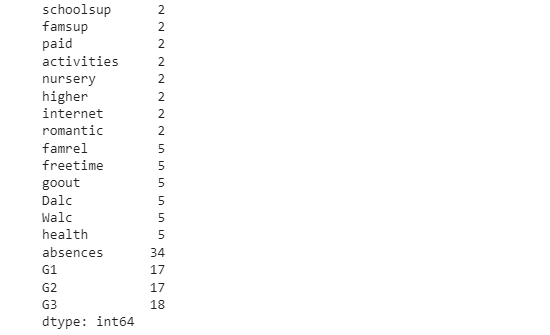




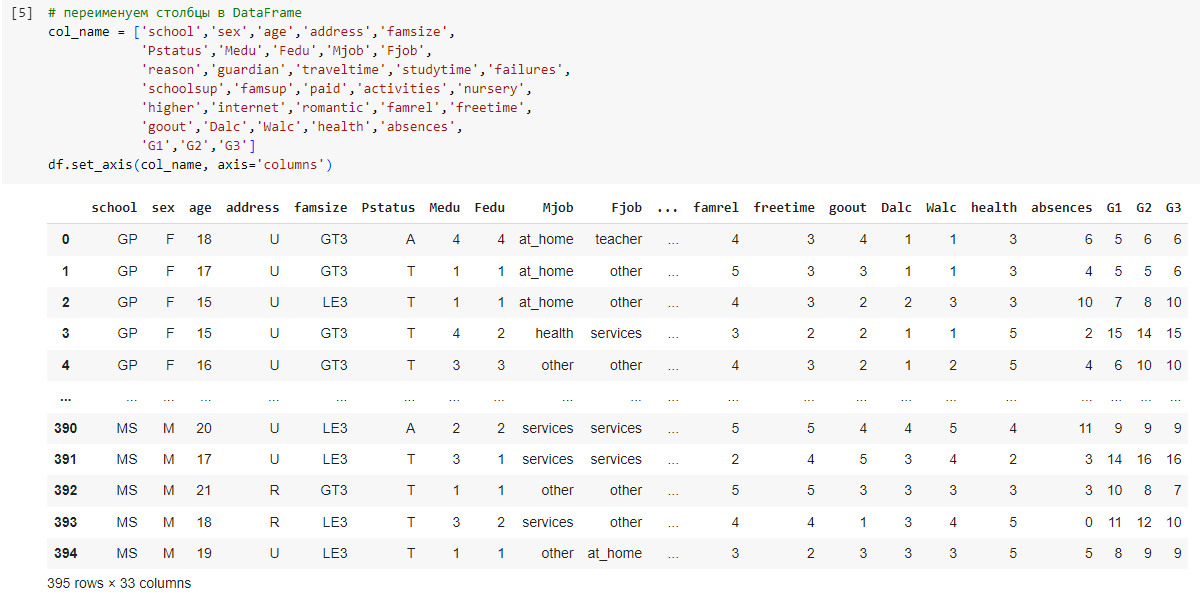


1. Подсчет уникальных значений каждого столбца:

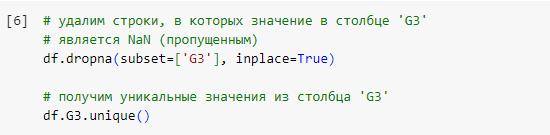




1. Переименование столбцов:

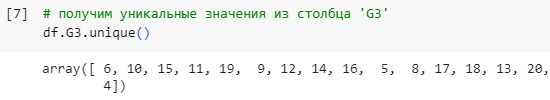


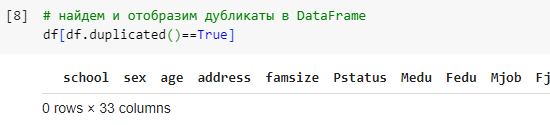
1. Проверка на пропуски:





1. Проверка на наличие дубликатов:





**Вывод:**

В ходе выполнения лабораторной работы были изучены инструменты для запуска, отладки и тестирования кода на Python, применение Python для предобработки данных, а также состоялось знакомство с PyCharm и Google Colaboratory.