|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра прикладной математики (ПМ)**

**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

по дисциплине «Технологии и инструментарий машинного обучения»

**Практическое занятие № 1**

**Задание №1** **- №3**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИМБО-02-22, Ким Кирилл Сергеевич* | (подпись) | |
| Преподаватель | *Трушин Степан Михайлович, преподаватель* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_202\_\_г. | |  | |

Москва 2024 г.

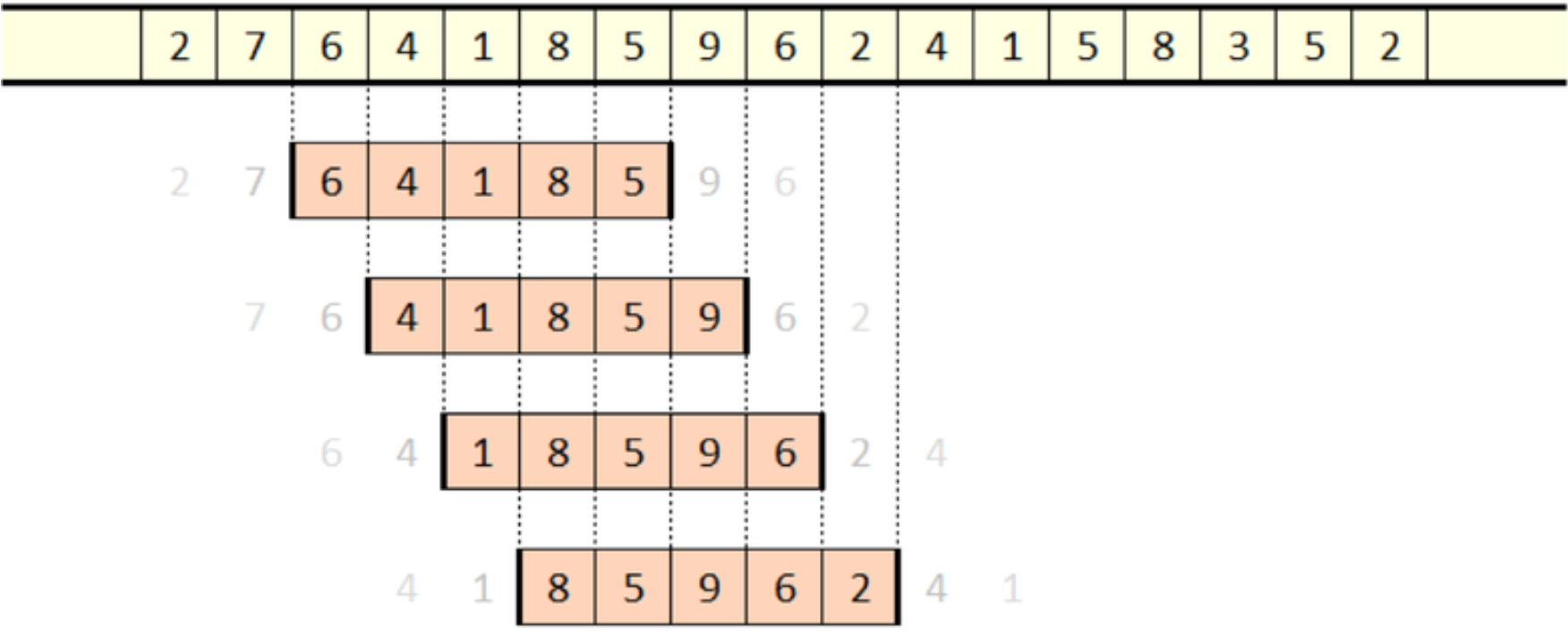
**Цель занятия:**

Целью данной работы является ознакомится с Google Colab и выполнить задания.

**Постановка задачи:**

**Задание 1 – Скользящее окно (Общий вариант)**

Реализовать функцию скользящего окна для заданного одномерного ряда данных, формирующую из него матрицу по следующему правилу как на Рисунке 1.



**Рисунок 1 – Правило**

**Задание 2 – Работа с базой данных (Общий вариант)**

В задании предлагается ответить на несколько вопросов по данным репозитория UCI Adult (ссылка https://archive.ics.uci.edu/ml/machine-learningdatabases/adult/adult.data) (качать данные не надо – они уже есть в репозитории) и ответить на следующие вопросы:

1. сколько мужчин и женщин (признак sex) представлено в этом наборе данных?
2. каков средний возраст (признак age) женщин?
3. какова доля граждан Германии (признак native-country)?
4. Постройте гистограмму распределения (bar plot) образования людей

(признак education).

1. каковы средние значения и среднеквадратичные отклонения возраста тех, кто получает менее 50K в год?
2. правда ли, что люди, которые получают больше 50k, имеют как минимум высшее образование? (признак education – Bachelors, Profschool, Assoc-acdm, Assoc-voc, Masters или Doctorate)
3. выведите статистику возраста для каждой расы (признак race) и каждого пола. Используйте groupby и describe. Каков максимальный возраст мужчин расы Amer-Indian-Eskimo?
4. среди кого больше доля зарабатывающих много (>50K): среди женатых или холостых мужчин (признак marital-status)? Женатыми считаем тех, у кого marital-status начинается с Married (Married-civspouse, Married-spouse-absent или Married-AF-spouse), остальных считаем холостыми
5. какое максимальное число часов человек работает в неделю (признак hours-per-week)? Сколько людей работают такое количество часов и каков среди них процент зарабатывающих много?
6. посчитайте среднее время работы (hours-per-week) зарабатывающих мало и много (salary) для каждой страны (native-country)

После построить 4 визуализации результатов.

**Задание 3 – Работа с базой данных (Свой вариант)**

На основе выбранного вами датасета провести его анализ по минимум 10 признакам. Датасет можно выбрать с данного сайта ([https://www.kaggle.com](https://www.kaggle.com/)).

Мой датасет (https://www.kaggle.com/datasets/thebumpkin/178-queen-album-tracks-with-spotify-data).

**Результат работы:**

Данную работу можете увидеть в блокноте Google Colab:

<https://colab.research.google.com/drive/1uGwl65y2QPXTKZImZiLmMT3bh4xzGmYB?usp=sharing>

**Список использованных источников и литературы:**

1. Ростовцев В.С. Искусственные нейронные сети,   
   Издательство "Лань", 2019. — 216 с. — URL: https://e.lanbook.com/book/122180
2. Араки М. Манга: Машинное обучение,   
   Издательство "ДМК Пресс", 2020. — 214 с. — URL: <https://e.lanbook.com/book/179473>
3. Платонов, А. В. Машинное обучение : учебное пособие для вузов / А. В. Платонов. — Москва : Издательство Юрайт, 2022. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/508804