МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ им. Н.Э. Баумана

Факультет «Информатика и системы управления» Кафедра «Систем обработки информации и управления»

ОТЧЕТ

Рубежный Контроль № <u>1</u> по дисциплине «Методы машинного обучение»

ИСПОЛНИТЕЛЬ: группа ИУ5-21М	_Лю Бэйбэй _{Фио}
ПРЕПОДАВАТЕЛЬ:	подпись ""2023_ г.
	ФИО
	подпись
	" " 202 г

Москва - 2023

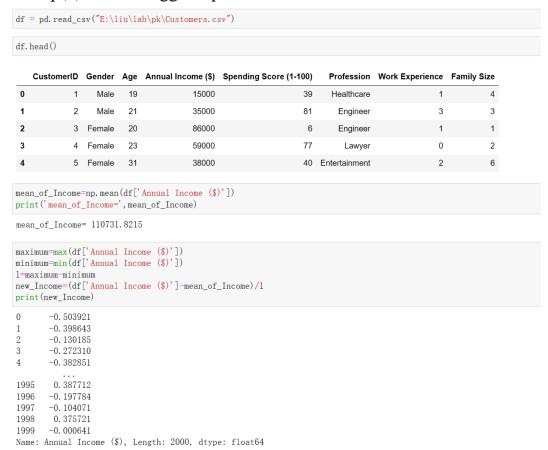
Вариант:19

Задачи:

Задача №19.

Для набора данных проведите масштабирование данных для одного (произвольного) числового признака с использованием метода "Mean Normalisation"

Набор Данных: /kaggle/input/customers-dataset/Customers.csv



Задача №39.

Для набора данных проведите процедуру отбора признаков (feature selection). Используйте класс SelectPercentile для 10% лучших признаков, и метод, основанный на взаимной информации.

Набор Данных: sklearn.datasets.load_digits

```
X, y=load_digits(return_X_y=True)
X. shape
(1797, 64)

X_new=SelectPercentile(mutual_info_classif, percentile=10).fit_transform(X, y)
X_new. shape
(1797, 7)
```

Дополнительное требование: для пары произвольных колонок данных построить график "Диаграмма рассеяния"

Набор Данных: /kaggle/input/customers-dataset/Customers.csv

```
plt.figure(figsize=(8,4))
plt.scatter(df['Age'], df['Annual Income ($)'])
plt.xlabel='Age'
plt.ylabel='Annual Income ($)'
plt.show()
```

