

**ЖА(Б)КА**

**ПРЕДСКАЗАНИЕ**  
ОБРАЗОВАНИЯ  
ЧЕЛОВЕКА  
ПО ЕГО ПОКУПКАМ  
В STEAM

**КОДИИМ**



**STEAM**

# ☰ КОМАНДА



Амира



Женя



Катя



Ася

**КОДИИМ**



Ал Ж К

# Ася

Основной ML  
разработчик  
Дата аналитик



**КОДИИМ**



# Катя

ML разработчик  
Project manager  
Дизайнер



Ал Ж

ся

**КОДИИМ**





# Женя

Дата аналитик



Ал

ТЯ СЯ

**КОДИИМ**



# Амира

Дата аналитик



**КОДИИМ**



Цель



**Цель:**  
Предсказать  
образование  
человека по его  
покупкам в Steam

**Результат:**  
Используя ML модели  
предсказали  
образование и  
оценить точность по  
F1 и Accuracy  
метрикам



# Анализ данных Steam НИУ ВШЭ

Аналог проекта

Репозиторий данных команды  
DataGeeks - Олимпиада DANO

Референс



<https://steamcommunity.com/>

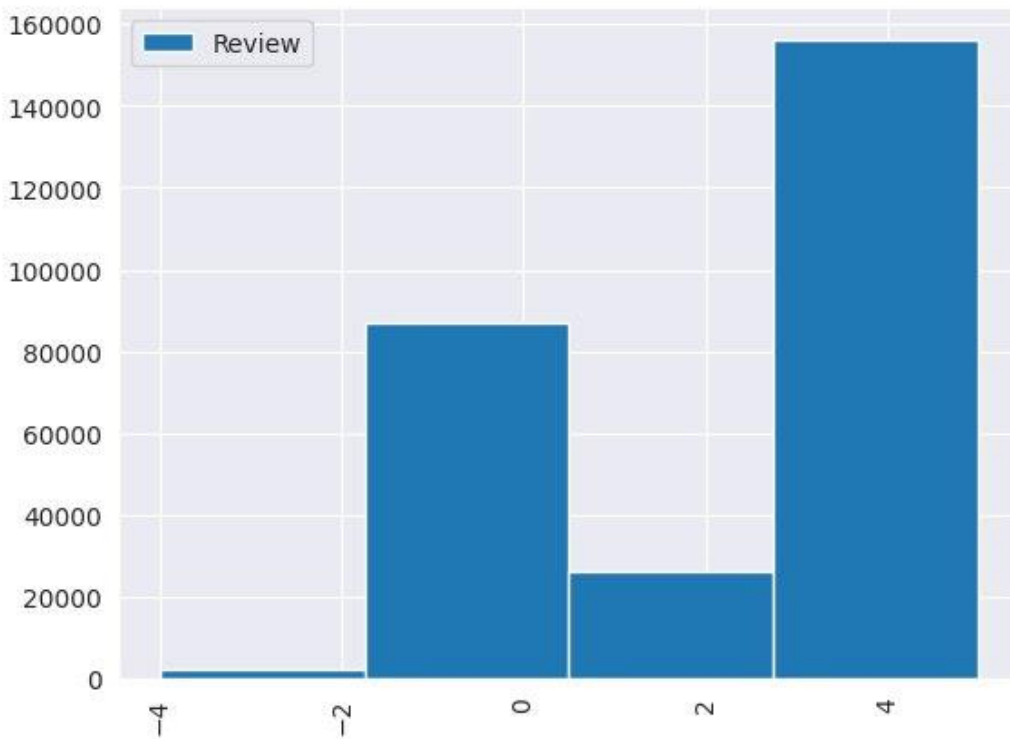






# Визуализация графиков

Графики

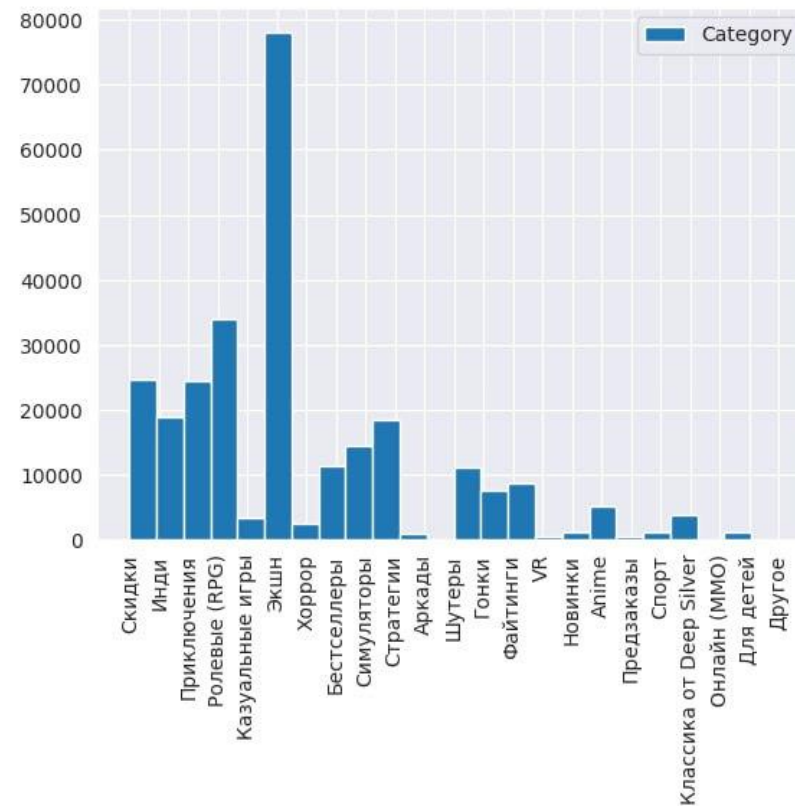
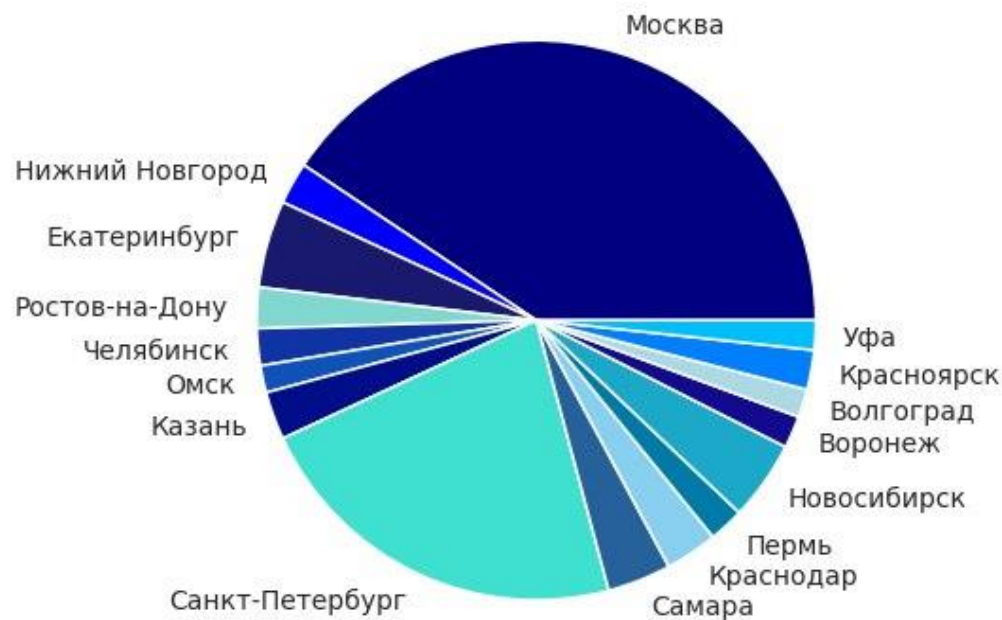




# Графики, которые у нас получились

<https://steamcommunity.com/>

Распределение по городам

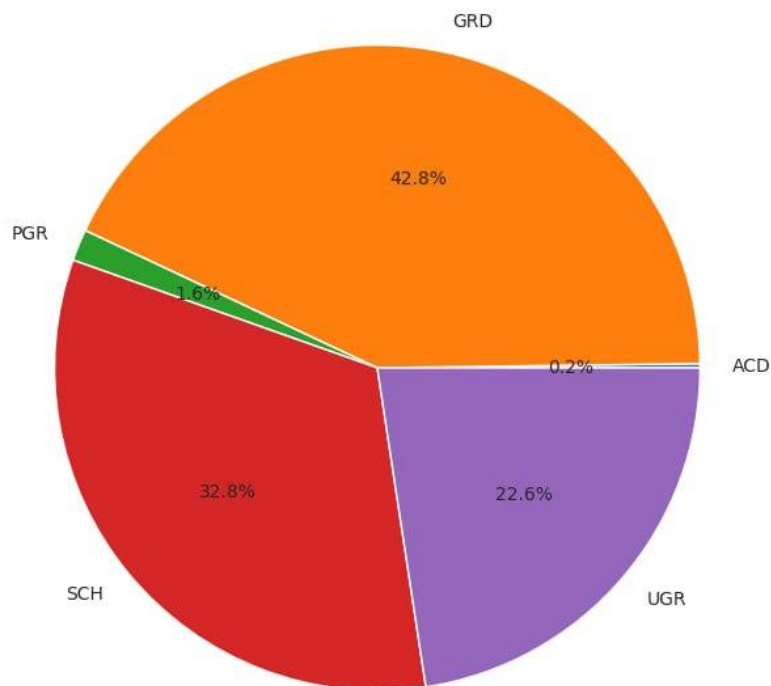




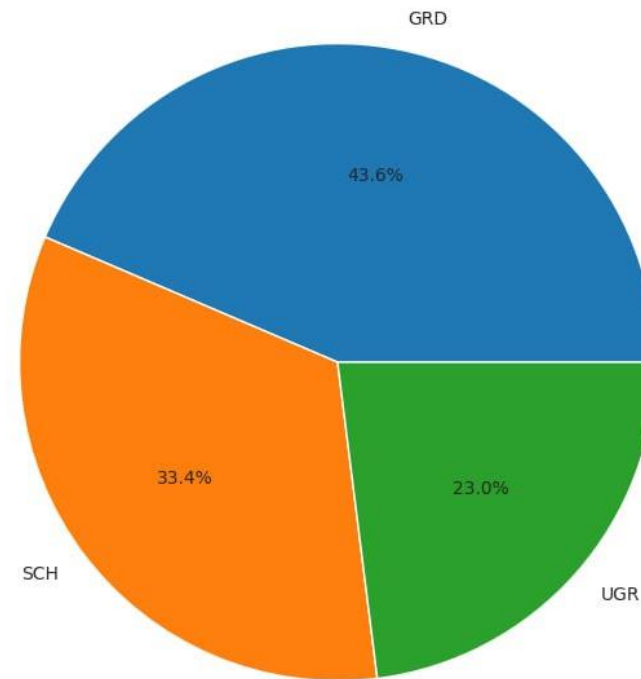
# Анализ целевой переменной

<https://steamcommunity.com/>

Уровень образования



Уровень образования





# Преобразование данных

<https://steamcommunity.com/>

Датасет



Форматирование  
(drop)



one\_hot\_encoding





ML разработка



В качестве **метрик**  
мы выбрали F1 и  
Accuracy

#### Чек лист:

Создаем моделью  
Разбиваем X и Y на train и test  
Обучаем модель  
Метрики  
Confusion matrix

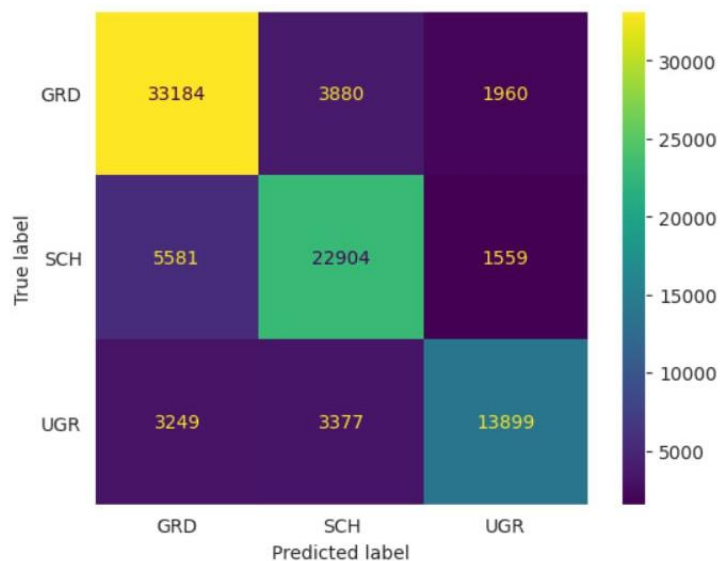


F1 = 80%  
Accuracy = 80%

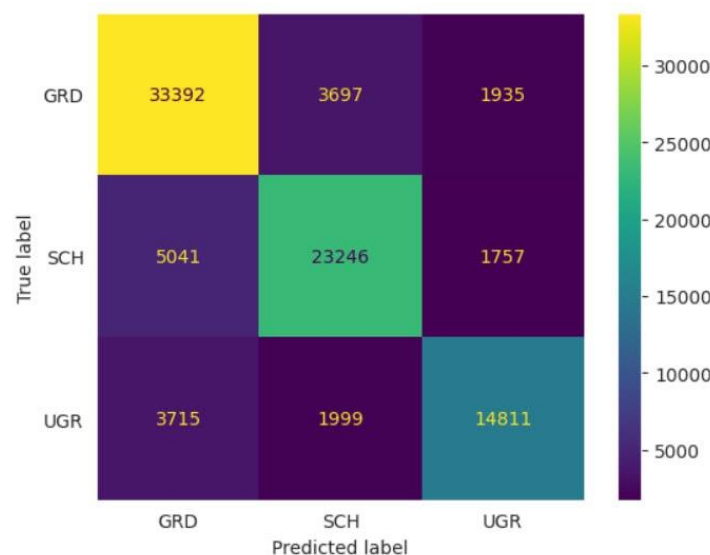


# Confusion matrix для моделей

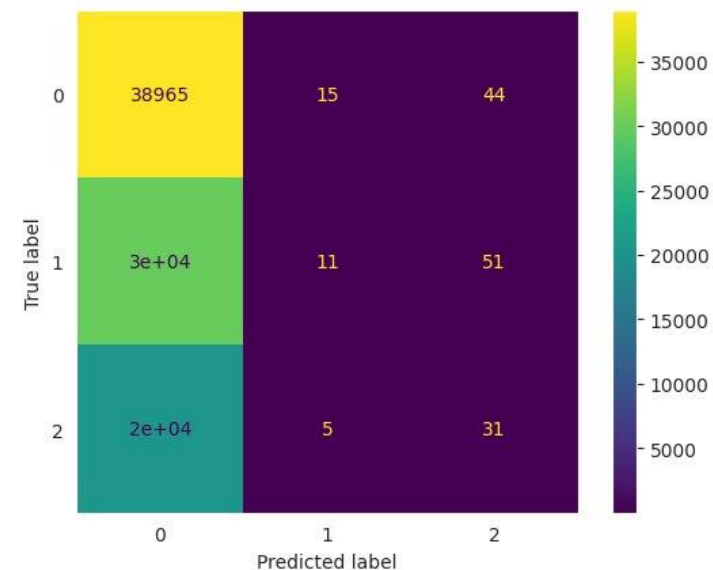
<https://steamcommunity.com/>



Decision Tree

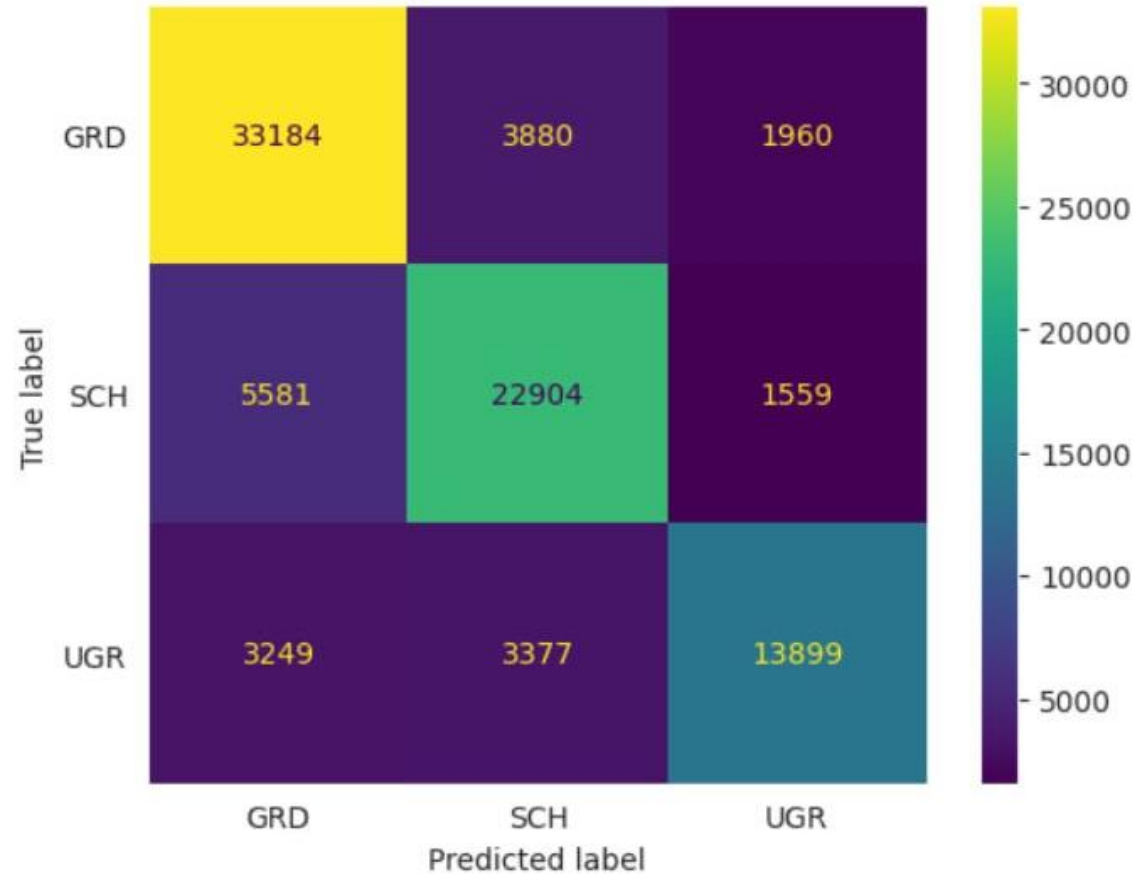


Random forest



Logistic classification

# Decision Tree

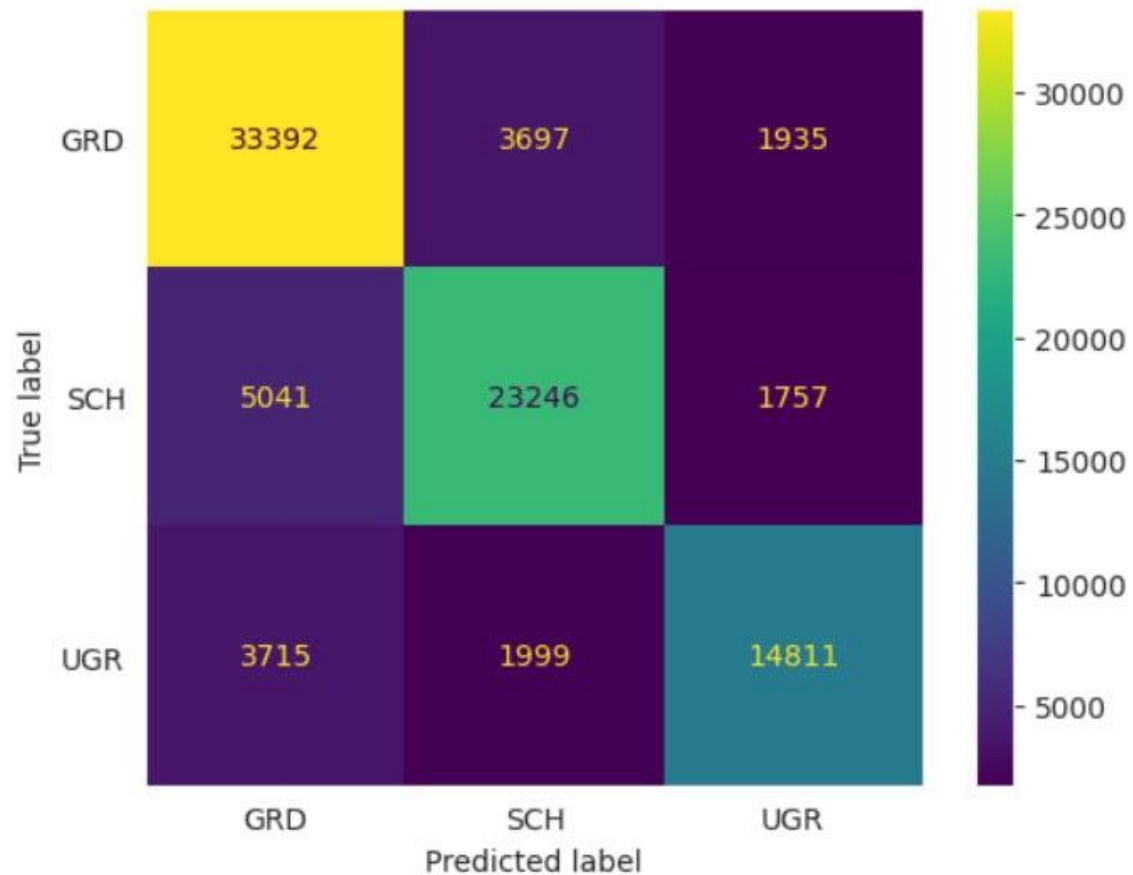


**Accuracy score:**

0.765284~

**F1:**

0.763286~



# Random forest

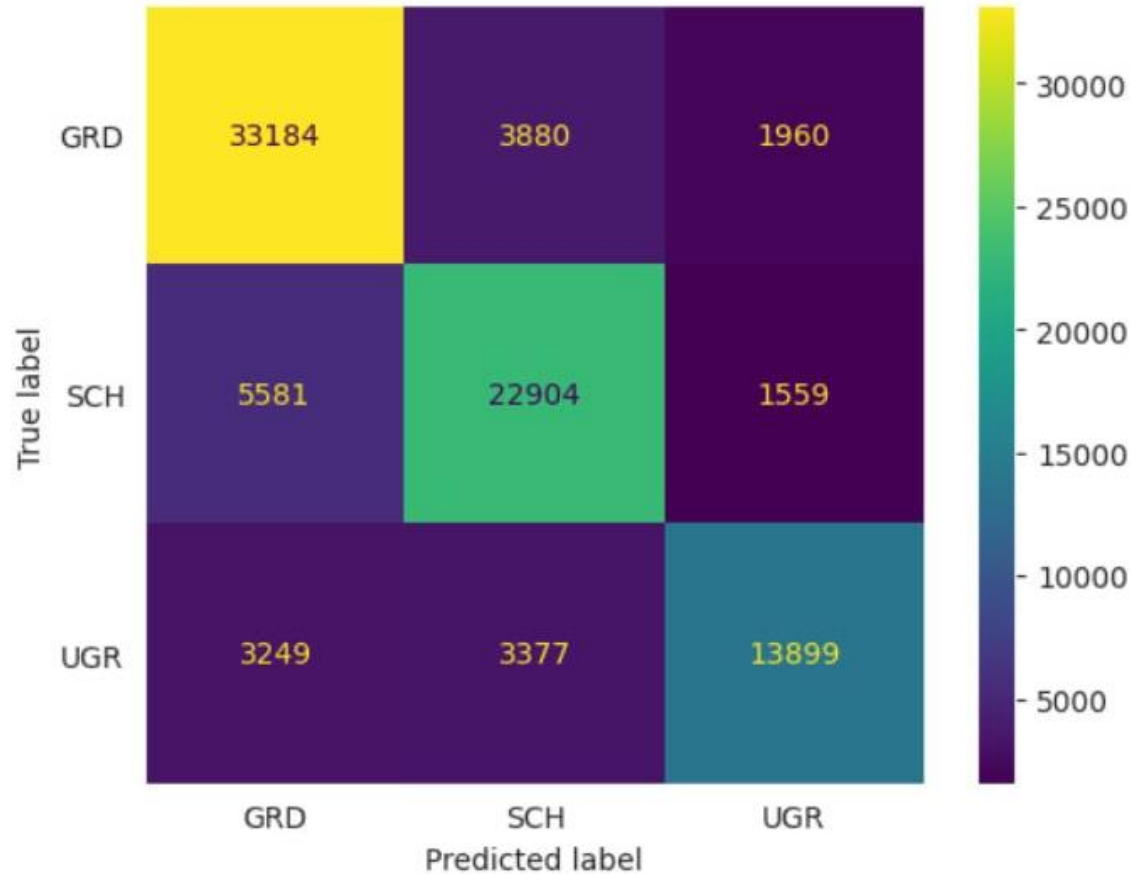
**Accuracy score:**

0/78450~

**F1:**

0.78357~

# Logistic classification



**Accuracy score:**

0.43507~

**F1:**

0.26567~



Улучшения



ME\_3428

2 Aug, 2024 @ 01:57 am

Помогите! Как я могу улучшить этот проект?



Zanyda\_228

2 Aug, 2024 @ 02:06 am

Нужно построить графики другого типа и постоянно актуализировать данные



JaanuuarY

2 Aug, 2024 @ 02:12 am

Может быть использовать другие метрики для лучшего результата?



Moe\_sostoyanie\_sejchas



2 Aug, 2024 @ 02:14 am

Я не знаю, вообще я спать хочу



# Полезные ссылки

**КОДИИМ**

  <https://steamcommunity.com/>



[Тут](#)  
ссылка на GitHub



:))