**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ**

**КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение

высшего образования

**«Поволжский государственный университет телекоммуникаций и информатики»**

Факультет ***Информационных систем и технологий***

Направление подготовки ***09.04.02 Информационные системы и технологии***

**Отчет по лабораторной работе №3**

**По дисциплине “Системы и методы искусственного интеллекта”**

**студента 1 курса группы**

**\_\_**ИСТм-43**\_\_\_**

Емелин Максим Сергеевич

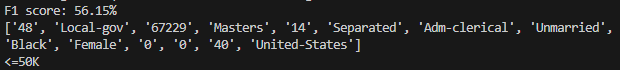
(фамилия, имя, отчество)

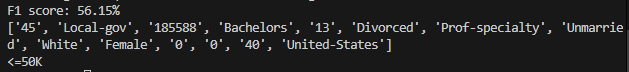
Самара, 2024

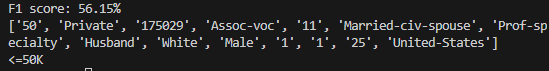
18)



19)







20)

1. **100000 данных** — F1 score = **56.15%**
2. **1000 данных** — F1 score = **33.33%**
3. **55000 данных** — F1 score = **56.15%**
4. **10000 данных** — F1 score = **45.87%**
5. **17500 данных** — F1 score = **50.44%**

**Вывод**: При увеличении числа данных (max\_datapoints) модель часто показывает лучший результат (F1 score). Меньшее количество данных может привести к значительному снижению точности (например, 1000 данных). Это подтверждает важность объема обучающих данных для улучшения модели.

21)

1. **test\_size: 0.2** — F1 score = **56.15%** (<=50K)
2. **test\_size: 0.3** — F1 score = **56.15%** (>50K)
3. **test\_size: 0.4** — F1 score = **56.15%** (>50K)
4. **test\_size: 0.5** — F1 score = **56.15%** (<=50K)
5. **test\_size: 0.9** — F1 score = **56.15%** (>50K)

**Вывод**: Изменение размера тестовой выборки **test\_size** не значительно влияет на F1 score, но может изменять пропорцию предсказанных классов (<=50K или >50K), что может быть важным для оценки модели в контексте баланса классов.