**Министерство науки высшего образования Российской Федерации**

**федеральное государственное автономное образовательное учреждение высшего образования**

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»**

**(Университет ИТМО)**

Факультеттехнологий искусственного интеллекта

**Лабораторная работа №3**

**Выполнили:**

Гончаренко Данила Олегович, группа J4150

Стрельницкая Татьяна Викторовна, группа J4140

**Преподаватель:**

Старобыховская Анастасия Александровна

Санкт-Петербург

2024

Оглавление

[Задание 2](#_Toc180409830)

[Основные этапы 3](#_Toc180409831)

[Вывод 17](#_Toc180409832)

# Задание

В рамках данной лабораторной работы поставлена цель – проанализировать результаты после сравнения исследования скорости работы и точности для валидационной выборки изначально, после конвертации модели в onnx и после оборота в бенто.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Взять лучшую модель из лабораторной 2,

2. На валидационной выборке происследовать: скорость работы на cpu/gpu (если есть), измерить точность, собрать примеры корректных кейсов и ошибочных.

3. Сконвертировать модель в ONNX + проверить качество и скорость,

4. Обернуть в бенто по шаблону. Запустить сервис + проверить качество и скорость.

5. Проанализировать результаты. Сделать выводы и описать почему они могли получиться именно такими.

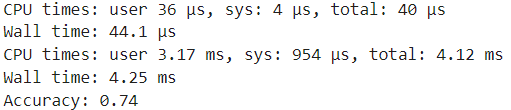
6. Написать отчёт, добавить скрины UI bento, ссылку на код, опыт и впечатление от ONNX и bentoml, а также выводы по экспериментам.

# Основные этапы

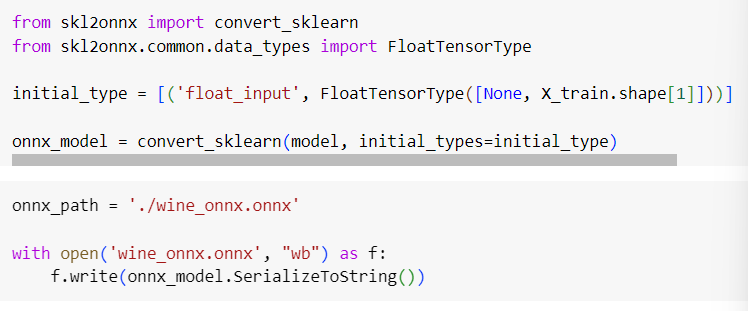
Для конвертации модели была выбрана модель логистической регрессии,

Анализ проводился на CPU.

Результаты работы без конвертации в ONNX:



Далее проводим конвертацию в ONNX.



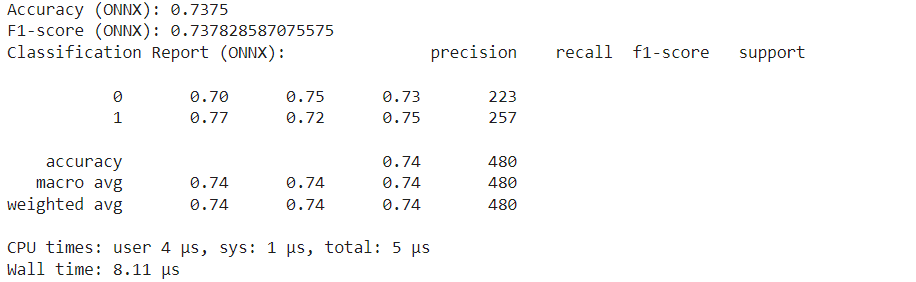
Проводим проверку конвертации:



Тестируем сконвертированную в ONNX модель:



Получаем такие результаты:



В результате конвертации модели Логистической регрессии в ONNX скорость выполнения модели повысилось с 3.16 ms у пользователя до 4 us.

# Выводы

По итогам выполненной лабораторной работы можно сделать следующие выводы:

- Логистическая регрессия применима не во всех задачах анализа данных и машинного обучения, но все же имеет большую область использования и отлично подходит для задачи бинарной классификации.

- На качество модели влияет огромное множество факторов, и не на все из них мы можем повлиять, но изменив некоторые параметры и проведя ряд экспериментов можно добиться улучшения прогнозов.

- Применение анализа главных компонент и выявление с его помощью наиболее важных для оценки качества вина параметров позволило улучшить производительность и сократить время работы. Кроме того, значение метрики accuracy увеличилось на 1%.

- Анализ и оценка выбросов в наборе данных позволяет оценить их влияние на работу модели. В рамках данной лабораторной работы было выявлено, что проблема не слишком высокой точности, очевидно, не в выбросах. А удаление сэмплов, включающих выбросы, не только не улучшило, а наоборот существенно ухудшило работу модели. Вероятно, необходимо было провести более тщательный анализ каждого из признаков или попробовать применить другие методы.

- Регуляризация данных помогает стабилизировать работу модели и избежать переобучения за счет уменьшения влияния шумов и выбросов. Однако для данной конкретной модели логистической регрессии польза регуляризации не слишком большая. Скорее всего, за счет того, что набор данных небольшой и сложность модели в целом невысокая.

- Поставленная цель достигнута, задачи решены полностью.