

# **DS** **DATASKAI**

Фреймворк для AI/ML-проектов

(R&D-версия)

Установка в изолированной среде (совместно с тестовым датасетом)

**Skoltech**

Веб-страница проекта: <http://dataskai.com>

## 1 Состав пакета и способ доступа

Пакет фреймворка DATASKAI для установки в изолированной среде представляет собой демонстрационную версию с тестовым датасетом (данные о пассажирах ”Титаника”).

Пакет поставляется в виде архива ZIP, разбитого на 4 части, либо в виде ISO-образа. Доступ к файлам архива возможен по запросу данных авторизации у разработчиков.

## 2 Порядок установки и запуска

Все команды необходимо выполнять из директории проекта (sample\_project).

Для успешного выполнения необходима операционная система linux с версией ядра не ниже 4.15.

Необходим python версии 3.6.9.

1) Выполнить установку необходимых зависимостей среды сборки ().

```
pip install -r toolkit_cli/requirements.txt ( )
pip install -r toolkit_cli/requirements.txt --no-index
↳ --find-links toolkit_cli/packages/ ( )
```

2) Запуск сервисов (sudo - при необходимости)

```
./scripts/server_start.sh
```

3) Загрузка данных (перед запуском подождать, пока сервисы запустятся)

```
./scripts/load_data.sh
```

4) Восстановление дампа (sudo - при необходимости)

```
docker cp data/dump/ mongo:/backup/
docker exec -i mongo mongorestore --db dataskai /backup/dataskai
```

5) Открыть ноутбук датасайнтиста. В браузере перейдите по ссылке:

```
http://0.0.0.0:8889/
```

Введите токен авторизации. Чтобы получить токен, выполнить в терминале (sudo - при необходимости):

```
docker logs dataskai-data-science-env
```

Из логов скопируйте токен (token=...)

6) Откройте ноутбук `dataskai/notebooks/titanic_example.ipynb`. Выполните все ячейки jupyter notebook. Перед выполнением последней ячейки (отправки сабмита, `submitter`) нажмите кнопку "save and checkpoint".

7) Для просмотра результатов сабмита перейдите на лидерборд. В браузере откройте ссылку:

```
http://0.0.0.0:15009/board_dataskai_sample/message/TITANIC/submits  
→ ts
```