

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное  
учреждение высшего образования  
«Самарский национальный исследовательский университет имени академика  
С.П. Королева»  
Институт информатики, математики и электроники Факультет информатики  
Кафедра технической кибернетики

**Отчет по лабораторной работе №2**  
по дисциплине «Инженерия данных»

**Тема: «Интерфейс и обучение НС»**

Направление подготовки: 01.04.02 Прикладная математика и информатика:  
Магистерская программа «Наука о данных»

Профиль «Системы искусственного интеллекта»

Выполнила: Шаина М. М.,  
студентка группы 6231 – 010402D

**Самара 2023**

В рамках данной лабораторной работы предлагается построить два пайплайна:

Пайплайн, который позволяет получить предсказания для исходных данных с помощью некоторой модели.

Пайплайн, который позволяет обучить или дообучить целевую модель.

Для построения такого пайплайна воспользуемся такими инструментами, как Apache Airflow.

Цель лабораторной работы — построить пайплайн, который реализует систему "Автоматического распознавания речи" для видеофайлов.

Для работы task-а по ожиданию получения нового видео необходимо создать новое подключение к airflow. Для создания подключения переходим в Airflow по адресу <http://localhost:8080/connection/list/>.

Далее для того, чтобы можно было преобразовать наш аудиофайл в текст, а после получить из него summary, необходимо зарегистрироваться на <https://huggingface.co/> и получить токен API с правами записи для возможности отправки и получения запросов к сайту.

Для сохранения конспекта в PDF, необходимо было использовать библиотеку fpdf.

В качестве исходного видео использовался фрагмент из сериала «Теория большого взрыва» длительностью 2 минуты 6 секунд.

Для второго пайплайна была обучена тестовая модель выполняющая классификацию на датасете MNIST, опять же элементы пайплайна которые могли быть выполнены в одном файле были объединены, и выполнялись в контейнере tensorflow. Итоговый пайплайн представлен на рисунке 1.

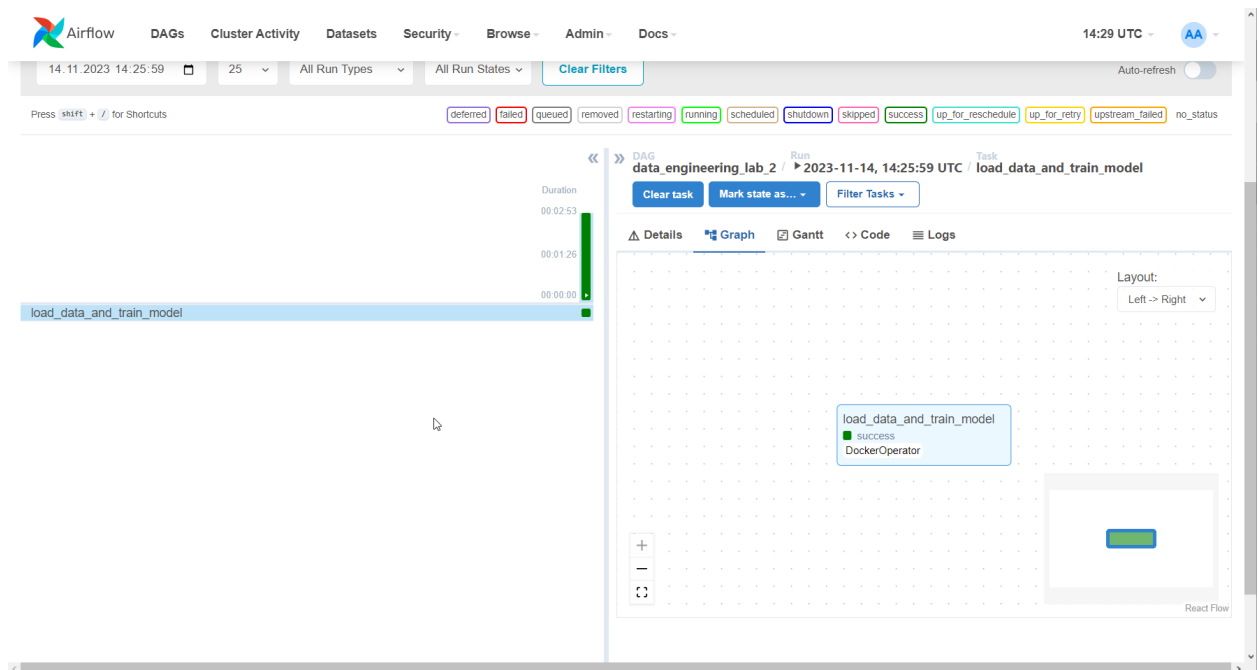


Рисунок 1 — Обучение НС