Контрольное задание

Компетенция «Машинное обучение и большие данные»

**Общее описание**

В эти три для мы с вами познакомимся с задачами машинного обучения, с которыми могут столкнуться специалисты в реальной жизни: от загрузки и предобработки данных и до вопросов комментирования кода.

**Модуль 4. Разработка REST API.**

В данном модуле участники должны разработать REST API и подготовить отчет о проделанной работе. По умолчанию, участники должны использовать модель, которую они разрабатывали в Модуле 3. В случае, если в модуле 3 не получилось разработать ни одной модели, участникам предлагается использовать синтетический датасет и использовать простую предварительно подготовленную модель из папки **training\_datasets. Использование своей модели из модуля 3 принесет в 2 раза больше баллов!**

Входные данные:

1. В архиве находятся следующие файлы: **Простая модель машинного обучения.ipynb**, **Простая модель машинного обучения.html**, **model.pickle**.

Подробное описание задания:

1. Разработайте REST API.
   1. Выберите задачу, которую вы будете решать. На выбор есть 2 задачи из модуля 3, а также задача по синтетическим данным из папки **training\_datasets**. Использование задачи из Модуля 3 обеспечит в 2 раза больше баллов.
   2. Реализуйте основной функционал API, а именно:
      1. Реализуйте метод для загрузки данных на сервер. Определите тип и формат данных, который вы будете принимать. Возвращать метод должен список загруженных файлов
      2. Реализуйте метод для предобработки данных и сохранения выборки в сжатом формате (pickle, parquet) на сервере. Возвращать метод должен отчет об успешной предобработке и объем памяти, занимаемый предобработанным набором данных.
      3. Реализуйте метод для запроса списка доступных наборов данных на сервере. Возвращать метод должен список доступных наборов данных
      4. Реализуйте метод для выбора рабочего набора данных и подгрузки его в память сервера
      5. Реализуйте метод для запроса списка доступных моделей на сервере. Возвращать метод должен список доступных моделей
      6. Реализуйте метод для выбора одной из существующих на сервере моделей в качестве рабочей. Возвращать метод должен название выбранной модели
      7. Реализуйте метод для выполнения предсказаний при помощи рабочей модели и рабочего набора данных. Возвращать метод долен список предсказаний по данному набору.
      8. Реализуйте метод для формирования графика зависимости предсказаний от одного из параметров. Метод должен принимать на вход название одного из параметров (столбцов в наборе данных), а также использовать последний полученный набор предсказаний для построения точечной диаграммы. Возвращать метод должен график (изображение) в формате numpy-массива. Для отправки по сети возможно дополнительные преобразования данных.
      9. Реализуйте метод для запроса информации о имеющихся в API методах и их кратком функционале.
      10. Реализуйте метод для запроса информации о текущей версии API, а также об авторе и организации-разработчике. Возвращать метод должен строку с информцией.
      11. Позаботьтесь о том, чтобы при передаче любому из методов данных неправильных форматов, система не падала, а возвращала сообщения об ошибке.
2. Подготовить отчет по модулю (отчет должен быть в формате HTML, отчет должен быть структурирован, описывать методы, разработанные для REST API, содержать комментарии и пояснения к каждому методу).

Выходные данные по модулю:

1. Отчет в формате HTML.
2. Файл .ipynb / .py содержащий код для REST API.

Отчет должен иметь строго определенные название:

M4\_{lastname}\_{first name}

(например, M4\_Ivanov\_Ivan)