ИТОГОВАЯ АТТЕСТАЦИЯ

по курсу “Python-разработчик”

На тему: “Использование машинного обучения для создания проекта “Антиплагиат”

Описание проекта:

1. **Проработка бизнес-задачи проекта**. С использованием инструментариев машинного обучения, можно создать программу “Антиплагиат”, которая будет в автоматическом режиме. Ни для кого не секрет, что с приходом в нашу жизнь интернета, найти на его просторах, готовый текст(в нашем случае это реферат, курсовая, дипломная работа) не составит большого труда. Наша задача автоматизировать процесс выявления схожести поступающих на проверку работ.
2. **Понимание данных**. Так, как это всё-таки учебный проект, и у нас нет своей базы данных работ, то мы создадим свою, на основе нескольких готовых работ скачанных в сети интернет. Работы будем загружать с сайта [Библиофонд](https://www.bibliofond.ru/), это сайт с большой базой различных рефератов, курсовых работ и т.д.
3. **Подготовка данных.** После загрузки необходимых файлов в формате “.txt”, проведём необходимую подготовку данных.