

Индивидуальное задание

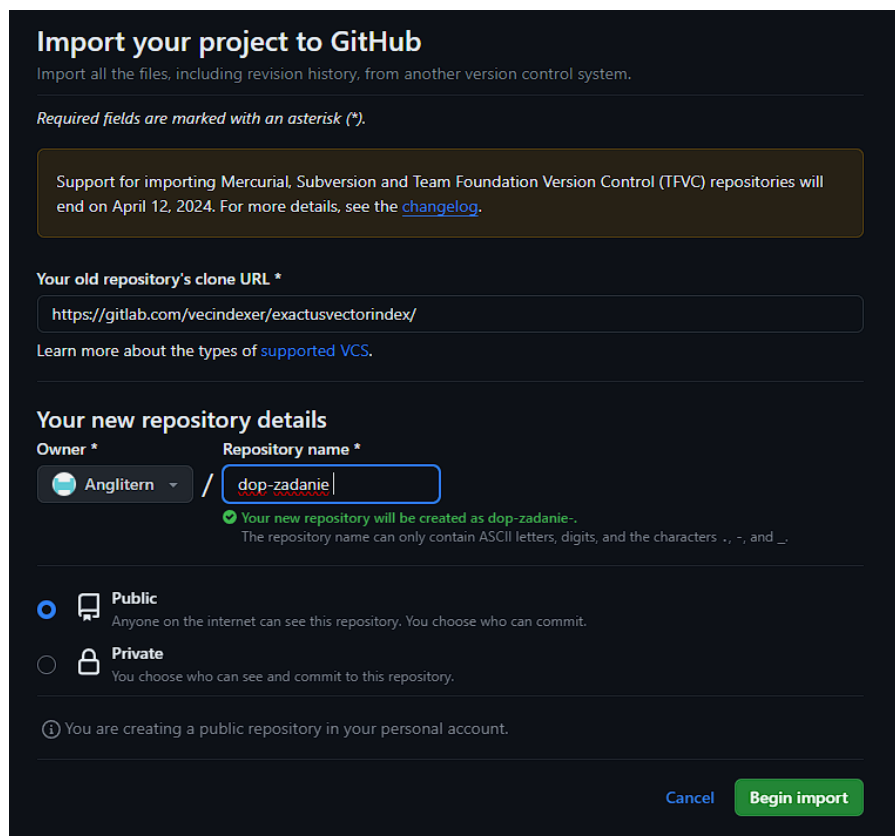
Работа с библиотекой *ExactusVectorIndex*

Шаги работы:

1. Выполнить Fork библиотеки в свою учетную запись.
2. Изучить материалы открытой библиотеки.
3. Провести установку библиотеки на свой локальный ПК.
4. Провести апробацию (например, с использованием представленных в репозитории примеров).

Порядок работы:

1. Выполнил импорт проекта из Gitlab в свою учетную запись Github как показано на рисунках 1-2.



The screenshot shows the 'Import your project to GitHub' interface. At the top, it says 'Import all the files, including revision history, from another version control system.' Below this, a note states: 'Support for importing Mercurial, Subversion and Team Foundation Version Control (TFVC) repositories will end on April 12, 2024. For more details, see the [changelog](#).' The form has two main sections. The first section, 'Your old repository's clone URL *', contains a text input field with the URL 'https://gitlab.com/vecindexer/exactusvectorindex/'. Below this is a link: 'Learn more about the types of [supported VCS](#).' The second section, 'Your new repository details', has two sub-sections. The first is 'Owner *' with a dropdown menu showing 'Anglitem'. The second is 'Repository name *' with a text input field containing 'dop-zadanie'. Below the repository name field, a green checkmark indicates: 'Your new repository will be created as dop-zadanie-. The repository name can only contain ASCII letters, digits, and the characters -, ., and _.' Below these fields are two radio button options: 'Public' (selected) and 'Private'. The 'Public' option is described as 'Anyone on the internet can see this repository. You choose who can commit.' The 'Private' option is described as 'You choose who can see and commit to this repository.' At the bottom, there is a note: 'You are creating a public repository in your personal account.' At the very bottom right, there are two buttons: 'Cancel' and 'Begin import'.

Рисунок 1 – Импорт библиотеки

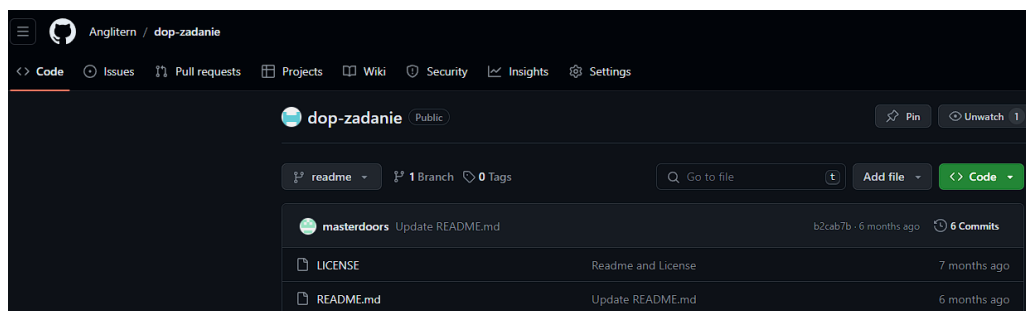


Рисунок 2 – Открытый библиотеки

2. Изучил материалы открытой библиотеки и узнал, что библиотека *ExactusVectorIndex* для распределенной индексации и поиска в сверхбольших массивах векторных представлений (эмбеддингов) с применением машинного обучения для преобразования векторных пространств и выборки данных.

Библиотека позволяет создавать:

- распределенные системы мультимодального информационного поиска в больших массивах разнотипизированных данных (текстов, изображений, мультимедиа);
- рекомендательные системы.

Библиотека *ExactusVectorIndex* может применяться в этих системах для индексации, хранения и поиска по векторным представлениям, что позволяет выполнять следующие функции: поиск ответа на вопрос, поиск похожих векторов (ближайших соседей), анализ разнородных объектов (классификация, кластеризация).

Также ознакомился с примерами по запуску и работе с библиотекой, находящимся в папке *example*, как показано на рисунке 3.

Name	Last commit	Last update
db	API and CLI synced with vecindexercpp	7 months ago
etc	u	6 months ago
examples	Update examples	6 months ago
tests	Update	6 months ago
utils	Update examples	6 months ago
vecindexerclient	Update examples	6 months ago
.gitignore	API and CLI synced with vecindexercpp	7 months ago
.gitlab-ci.yml	Cicdppl	7 months ago
LICENSE	Readme and License	7 months ago
README.md	Add examples	6 months ago
format_code.sh	Add examples	6 months ago
pytest.ini	First commit - base structure for client ...	9 months ago
requirements.txt	Change autoencoders dir	7 months ago
setup.py	Change autoencoders dir	7 months ago

Рисунок 3 – Папка *example*

3. Пробовал установить библиотеку для клиентской части, как показано на рисунке 4.

```

C:\Users\User>cd C:\nik

C:\nik>git clone https://gitlab.com/vecindexer/vecindexerclient.git
Cloning into 'vecindexerclient'...
remote: Enumerating objects: 228, done.
remote: Counting objects: 100% (228/228), done.
remote: Compressing objects: 100% (112/112), done.
remote: Total 228 (delta 106), reused 228 (delta 106), pack-reused 0
Receiving objects: 100% (228/228), 83.26 KiB | 13.88 MiB/s, done.
Resolving deltas: 100% (106/106), done.

C:\nik>cd vecindexerclient

C:\nik\vecindexerclient>python -m venv .venv
Python
C:\nik\vecindexerclient>source .venv/bin/activate
"source" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexerclient>git clone https://github.com/dvzubarev/pycbor_extra.git && cd pycbor_extra && pip install . && cd ../
Cloning into 'pycbor_extra'...
remote: Enumerating objects: 28, done.
remote: Counting objects: 100% (28/28), done.
remote: Compressing objects: 100% (15/15), done.
remote: Total 28 (delta 7), reused 28 (delta 7), pack-reused 0
Receiving objects: 100% (28/28), done.
Resolving deltas: 100% (7/7), done.
"pip" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexerclient\pycbor_extra>git clone https://github.com/Astromis/autoencoders.git
Cloning into 'autoencoders'...
remote: Enumerating objects: 21, done.
remote: Counting objects: 100% (21/21), done.
remote: Compressing objects: 100% (16/16), done.
Receiving objects: 80% (17/21) 18 (delta 2), pack-reused 0
Receiving objects: 100% (21/21), 11.35 KiB | 196.00 KiB/s, done.
Resolving deltas: 100% (2/2), done.

C:\nik\vecindexerclient\pycbor_extra>cd autoencoders && git clone https://github.com/leymir/hyperbolic-image-embeddings.git && cd hyperbolic-image-embeddings && pip install . && cd ../ && pip install . && cd ../
Cloning into 'hyperbolic-image-embeddings'...
remote: Enumerating objects: 242, done.
remote: Counting objects: 100% (127/127), done.
remote: Compressing objects: 100% (99/99), done.
Receiving objects: 100% (242/242), 1.77 MiB | 4.55 MiB/s, done.
remote: Total 242 (delta 62), reused 70 (delta 20), pack-reused 115
Resolving deltas: 100% (102/102), done.
"pip" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexerclient\pycbor_extra\autoencoders\hyperbolic-image-embeddings>pip install .
"pip" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexerclient\pycbor_extra\autoencoders\hyperbolic-image-embeddings>

```

Рисунок 4 – Установка библиотеки на ПК

Пробовал установить библиотеку для серверной части, как показано на рисунке 5.

```

C:\nik>git clone https://gitlab.com/vecindexer/vecindexercpp.git
Cloning into 'vecindexercpp'...
remote: Enumerating objects: 294, done.
remote: Counting objects: 100% (294/294), done.
remote: Compressing objects: 100% (233/233), done.
remote: Total 294 (delta 60), reused 270 (delta 43), pack-reused 0
Receiving objects: 100% (294/294), 3.45 MiB | 8.63 MiB/s, done.
Resolving deltas: 100% (60/60), done.

C:\nik>cd vecindexercpp

C:\nik\vecindexercpp>git submodule update --init
Submodule 'thirdparty/HTTPRequest' (https://github.com/elnormous/HTTPRequest) registered for path 'thirdparty/HTTPRequest'
Cloning into 'C:/nik/vecindexercpp/thirdparty/HTTPRequest'...
Submodule path 'thirdparty/HTTPRequest': checked out '703f713363172e57ea5da48a42d0f88d1e0e0d8c'

C:\nik\vecindexercpp>experimental-features = nix-command flakes
"experimental-features" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexercpp>~/.config/nix/nix.conf:
"~" не является внутренней или внешней
командой, исполняемой программой или пакетным файлом.

C:\nik\vecindexercpp>

```

Рисунок 5 – Установка библиотеки на ПК

Вывод: выполнил импорт библиотеки, изучил библиотеку *ExactusVectorIndex*, попытался установить библиотеку на свой локальный ПК и потерпел неудачу.