



Факултет по математика и информатика

Интегриране на данни от ракови заболявания

Курсова Работа

Изготвил: Парашкев Катерски

Фн.8MI3400138

Специалност: Био- и медицинска информатика

01.07.2022

Въведение

RESTful API написано на Python. Използва библиотека за свързване с MongoDB, базата данни - *pymongo*. За сървъра се използва web framework-а Flask заради неговата минималистичност и леснота за използване. Има функционалност за предсказване на броя дни оставащ живот. Използва се Ridge Regression чрез библиотеката Scikit-learn.

Функционалност

- CRUD за следните данни:
 - Rnaseq (RNA sequence gene expression)
 - Cnvsnp (Copy Number Variations inferred from Single Nucleotide Polymorphism)
 - Mirseq (microRNA sequence expression)
 - Clinical data
- Комбиниране на всички данни за пациент
- Предсказване на брой дни оставащ живот

Endpoints

CRUD GET

- /clinical връща всички данни от clinical колекцията
- /clinical/<id> връща специфичен документ
- /rnaseq връща всички данни от rnaseq колекцията
- /rnaseq/<id> връща специфичен документ
- /mirseq връща всички данни от mirseq колекцията

- `/mirseq/<id>` връща специфичен документ
- `/cnvsnp` връща всички данни от `cnvsnp` колекцията
- `/cnvsnp/<id>` връща специфичен

CRUD POST (Insert data)

- Могат да се качват данни в JSON формат като обект или списък от обекти
 - `/clinical`
 - `/ranseq`
 - `/mirseq`
 - `/cnvsnp`

CRUD PUT (Update data)

- Могат да се качват полетата в JSON формат, които ще бъдат ъпдейтнати в документа, като се подаде `id` за съответния документ
 - `/clinical/<id>`
 - `/ranseq/<id>`
 - `/mirseq/<id>`
 - `/cnvsnp/<id>`

CRUD DELETE (Delete data)

- Могат да се трият документи като се подаде `id` за съответния документ
 - `/clinical/<id>`
 - `/ranseq/<id>`
 - `/mirseq/<id>`
 - `/cnvsnp/<id>`

Комбиниране на всички данни за пациент

- `/patient-data/<id>`
 - Връща всичко комбинирано
 - `/clinical`
 - `/ranseq`
 - `/mirseq`
 - `/cnvsnp`
 - Ако няма стойности съответната колекция връща стойност **null**

Предсказване на брой дни оставащ живот

- **/predict POST**
 - Очаква да се подадат данни във формат *clinical* като се използва само:
 - Години
 - Stage
 - Pathology T stage
 - Pathology N stage
 - Pathology M stage
 - Връща предсказания брой години
 - Използва Ridge Regression (Имплементирано в Scikit-learn)
- **/update-model POST**
 - Fit-ва модела спрямо данните в колекцията *clinical*