



Интегриране на данни от ракови заболявания



Парашкев Катерски



Въведение

- Python
- Pymongo
- Scikit-learn



Flask

web development,
one drop at a time

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

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 документ

CRUD POST (Insert data)

- Могат да се качвам 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
 тържени 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 компинирано
 - `/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*

Благодаря за
вниманието