

 Zawartość zarchiwizowana w dniu 2024-06-18

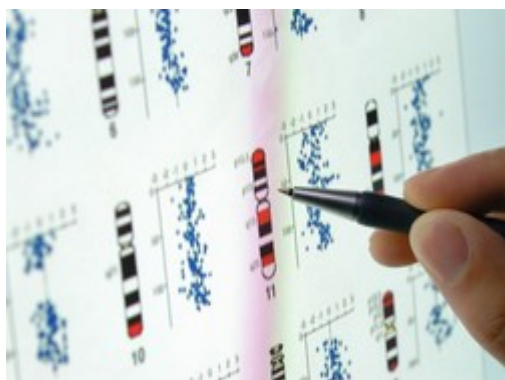


Probabilistic Models in Pseudo-Euclidean Spaces

Wyniki w skrócie

Szybkie i elastyczne systemy odpowiedzi na zapytania

Postępy w biotechnologii, a w szczególności technologie genomowe, pozwoliły na zgromadzenie ogromnej ilości danych sekwencji. Aby wydobyć istotne dane sekwencji z tak olbrzymich baz danych, konieczny jest oparty na zasadach, sformalizowany model. Taki właśnie model udało się opracować naukowcom finansowanym przez UE.



© fotohunter, Shutterstock

Naukowcy standardowo wydobywają dane, wykorzystując wyszukiwania podobieństw sekwencji w bazie danych. Jednakże, publiczne bazy danych takie jak GenBank i UniProt/SwissProt zawierają setki tysięcy sekwencji, a obecnie dostępne technologie bioinformatyczne nie zapewniają dobrej jakości procesu wydobywania danych.

Zespół PROMOS (Probabilistic models in pseudo-Euclidean spaces) znalazł rozwiązanie tego problemu w metodach bioinformatycznych. Celem naukowców było opracowanie algorytmów, które w bardzo krótkim czasie dostarczą właściwe dane sekwencji z dużych baz danych.

W pierwszym etapie badacze wykorzystali generyczne niemetryczne podobieństwa

wyniku, aby uzyskać i wdrożyć probabilistyczne modele relacyjne. Udało im się opracować platformę probabilistyczną dla metod relacyjnych w przestrzeniach pseudo euklidesowych.

Aby usprawnić uczenie się modelu oraz umożliwić szybkie wydobywanie wiedzy, naukowcy stworzyli schematy aproksymacji dla danych relacyjnych, a także model hierarchiczny oraz schematy wydobywania. To podejście specyficzne dla domeny jest skuteczne, ponieważ przekłada macierze braku podobieństwa dużej skali na aproksymowane dodatnio półokreślone macierze jądra po kosztach liniowych.

Technologię PROMOS przetestowano na kilku dużych bazach danych białek i uzyskano lepszy czas działania niż w przypadku klasycznych systemów wydobywania danych przy konkurencyjnej dokładności modelu. Metody zostały opisane w artykułach opublikowanych w kilku renomowanych czasopismach, a kilka innych prac jest jeszcze w trakcie opracowywania.

Działania prowadzone w ramach projektu oraz jego wyniki powinny znacznie przyspieszyć prace badawczo-rozwojowe w dziedzinie biotechnologii i farmacji.

Słowa kluczowe

[Dane sekwencji](#)

[bioinformatyka](#)

[PROMOS](#)

[modele probabilistyczne](#)

[przestrzenie pseudo euklidesowe](#)

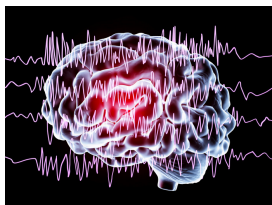
Znajdź inne artykuły w tej samej dziedzinie zastosowania



Zegary atomowe z rozbitych kryształów cyrkonu umożliwiają datowanie uderzeń asteroid

9 Lipca 2021





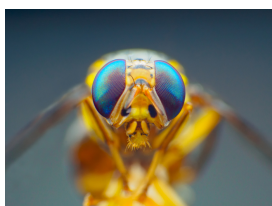
Wirtualne narzędzie do leczenia epilepsji

9 Lutego 2023



Stewia - słodsza i bardziej zrównoważona alternatywa dla cukru

6 Marca 2023



Muszki owocowe: małe owady o zaskakująco dużym rozumku

28 Marca 2022



Informacje na temat projektu

PROMOS

Identyfikator umowy o grant: 327791

Projekt został zamknięty

Data rozpoczęcia

2 Stycznia 2014

Data zakończenia

2 Kwietnia 2016

Finansowanie w ramach

Specific programme "People" implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013)

Koszt całkowity

€ 221 606,40

Wkład UE

€ 221 606,40

Koordynowany przez
THE UNIVERSITY OF
BIRMINGHAM
 United Kingdom

Ostatnia aktualizacja: 7 Lutego 2017

Permalink: <https://cordis.europa.eu/article/id/191008-fast-and-flexible-query-response-systems/pl>

European Union, 2025