Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0414U005111

Особливі позначки: відкрита

Дата реєстрації: 10-11-2014

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Островський Олексій Вікторович
2. Ostrovskyy Oleksiy

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 01.05.01

Назва наукової спеціальності: Теоретичні основи інформатики та кібернетики

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 24-10-2014

Спеціальність за освітою: 8.04030101

Місце роботи здобувача: Інститут кібернетики ім. В.М. Глушкова

Код за ЄДРПОУ: 05417176

Місцезнаходження: 03680,м.Київ,МСП, пр. Ак. Глушкова, 40

Форма власності:

Сфера управління: Національна академія наук України

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): Д 26.194.02

Повне найменування юридичної особи: Інститут кібернетики імені В.М. Глушкова Національної академії наук України

Код за ЄДРПОУ: 05417176

Місцезнаходження: проспект Академіка Глушкова, 40, м. Київ, Київська обл., 03187, Україна

Форма власності:

Сфера управління: Національна академія наук України

Ідентифікатор ROR: Не застосовується

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Інститут кібернетики ім. В.М. Глушкова

Код за ЄДРПОУ: 05417176

Місцезнаходження: 03680,м.Київ,МСП, пр. Ак. Глушкова, 40

Форма власності:

Сфера управління: Національна академія наук України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 28.17.19

Тема дисертації:

1. Методи розпізнавання на основі моделей Маркова з прихованими змінними

2. Recognition methods based on Markov models with hidden variables

Реферат/анотація:

1. Дисертаційна робота присвячена дослідженню моделей Маркова з прихованими змінними та їхніх узагальнень у застосуванні до розпізнавання функціональних фрагментів генів та визначення просторової структури білка. Побудовано математичний апарат, що дозволяє описати обидві задачі розпізнавання та адаптувати наявні критерії якості їх вирішення. Запропоновано ймовірнісні моделі на основі моделей Маркова з прихованими змінними та ланцюжків Маркова вищих порядків; обґрунтовано можливість їхнього застосування за рахунок закономірностей розподілу амінокислот у білках, синтезованих по різних нитках ДНК. На основі принципу максимуму правдоподібності створено алгоритм динамічного програмування для розв'язання задачі розпізнавання, який використовує запропоновані моделі. Розглянуто композиції цих моделей з ексклюзивною компетентністю складових та пов'язаний алгоритм розпізнавання, у якому кожен рядок розпізнається одним з складових алгоритмів у залежності від спостережуваних характеристик рядку. Для створення розбиття на області компетентності запропоновано предикати, залежні від концентрації

коротких ланцюжків нуклеотидів або амінокислот; описано алгоритми відбору предикатів на основі навчальної вибірки, що ґрунтуються на методах відбору інформативних ознак у машинному навчанні. Вирішено питання визначення оптимальних параметрів імовірнісних моделей композиції та побудови розбиття на області компетентності. У результаті обчислювального експерименту показано, що якість запропонованих моделей та їхніх композицій відповідає актуальним алгоритмам біоінформатики.

2. The thesis is devoted to hidden Markov models and their generalizations being applied to the two bioinformatics problems: recognition of gene functional fragments and forecasting protein spatial structure. The mathematical apparatus to describe both recognition problems uniformly and to adapt existing quality criteria is built. Probabilistic models based on hidden Markov models and high order Markov chains are proposed, their validity being justified based on amino acid distribution in proteins synthesized from the different DNA strands. Based on the maximum likelihood principle, a dynamic programming algorithm using proposed models is devised for solving the stated recognition problem. Model mixtures with exclusive competences are considered, as well as the corresponding recognition algorithm that processes each observed string of states using a specific constituent algorithm chosen based on observable characteristics of the string. To create the partition into competence regions, we propose predicates based on contents of short sequences of nucleotides or amino acids and describe algorithms for selecting predicates using the training set, which are based on feature selection in machine learning. The problem of reaching optimal parameters for constituent models is solved, as well as the problem of building the optimal competence regions. On conducting the computational experiment, the quality of proposed models and their compositions is shown to relate to the quality of state of the art bioinformatics algorithms.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки:

Стратегічний пріоритетний напрям інноваційної діяльності:

Підсумки дослідження:

Публікації:

Наукова (науково-технічна) продукція:

Соціально-економічна спрямованість:

Охоронні документи на ОПІВ:

Впровадження результатів дисертації:

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Гупал Анатолій Михайлович

2. Gupal Anatoliy

Кваліфікація: д.ф.-м.н., 01.05.01

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Крак Юрій Васильович
2. Крак Юрій Васильович

Кваліфікація: д.ф.-м.н., 01.05.04

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Власне Прізвище Ім'я По-батькові:

1. Пашко Сергій Володимирович
2. Пашко Сергій Володимирович

Кваліфікація: к.ф.-м.н., 01.05.01

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Рецензенти

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Сергієнко Іван Васильович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Сергієнко Іван Васильович

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.