Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0420U102411

Особливі позначки: відкрита

Дата реєстрації: 25-12-2020

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Олійник Юрій Олександрович

2. Oliinyk Yurii Oleksandrovych

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 05.13.06

Назва наукової спеціальності: Інформаційні технології

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 11-12-2020

Спеціальність за освітою: 7.080401 Інформаційні управляючі системи та технології

Місце роботи здобувача: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Перемоги, буд. 37, м. Київ, Київська обл., 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): Д 67.052.01

Повне найменування юридичної особи: Херсонський національний технічний університет

Код за ЄДРПОУ: 05480298

Місцезнаходження: Бериславське шосе, 24, м. Херсон, Херсонська обл., 73008, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

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

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Перемоги, буд. 37, м. Київ, Київська обл., 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 28.23.24

Тема дисертації:

1. Інформаційна технологія аналізу текстових потоків даних

2. Text stream analysis information technology

Реферат/анотація:

1. Робота присвячена розробленню інформаційної технології аналізу текстових потоків даних. Об'єкт дослідження - процеси інтелектуального аналізу інформації у текстових потоках даних. Мета дослідження - підвищення якості аналізу та швидкості обробки потоків текстової інформації українською мовою за рахунок розроблення нових методів та інформаційної технології аналізу текстових потоків даних. Методи дослідження ґрунтуються на засадах комп'ютерної лінгвістики, теорії ймовірностей, математичній статистиці та інших методах інтелектуального аналізу даних. Достовірність та обґрунтованість отриманих результатів обумовлені коректним використанням математичного апарату, а також підтверджуються результатами обчислювальних експериментів. Здійснено аналіз наявних підходів і методів аналізу текстових даних, проведено аналіз особливостей аналізу даних у потоках даних. Формалізовано модель потоку текстових даних на основі ковзного вікна, що дозволило розширити можливості застосування методів інтелектуального аналізу та машинного навчання даних для аналізу текстових потоків даних. Уперше

розроблено метод виявлення аномалій у текстових потоках даних на основі розширеного алгоритму Isolation Forest, моделі текстових потоків даних на основі ковзного вікна та доданої фази попередньої обробки текстових даних та реферування. Удосконалено метод визначення тональності тексту в потоках текстової інформації, що передбачає комбінації алгоритмів Gradient Boosting, Rule Based та застосування моделі потоку текстових даних на основі ковзного вікна. Уперше запропоновано інформаційну технологію аналізу текстових потоків даних, яка використовує розроблені методи: метод виявлення аномалій та метод визначення тональності, що на відміну від наявних технологій забезпечує інтелектуальну обробку українськомовних текстів, підтримує розроблену модель ковзного вікна, а також високопродуктивні обчислення та візуалізацію обробки потоків даних у режимі онлайн. Галузь використання – в освітньому процесі у закладах освіти та науки, в установах та на підприємствах України.

2. The research is devoted to developing of text stream analysis information technology. Object of research is processes of intelligent analysis of text stream data. The purpose of the thesis is to increase the quality of analysis and performance of Ukrainian-language text data streams due development of new methods and text stream analysis information technologies. Research methods based on principles of computer linguistics, probability theory, mathematical statistics and data mining methods. The reliability and validity of the obtained results based on correct using of mathematical apparatus and confirmed by results of computational experiments. Modern approaches and methods analysis are performed, text streams analysis characteristics are detected. A text data stream model based on a sliding window was formalized, which allowed to expand the possibilities of using text mining methods and machine learning methods for text data streams. For the first time, anomalies detecting method in text data streams was proposed. Developed method based on an Isolation Forest algorithm, which in difference to the existing ones, supports text data stream model based on a sliding window, preprocessing and summarization phase of text data. Text stream sentiment analyses method based on combination of Gradient Boosting, Rule Based algorithms, and using text stream model based on a sliding window has been improved. For the first time, a text stream analysis information technology has been proposed, which supports the anomalies detection method and text stream sentiment analyses method. The information technology as opposed to existing ones supports Ukrainian-language texts data mining processing; developed text stream model based on a sliding window, high-performance computing and online visualization. Scope is processes in educational and scientific institutions, Ukrainian organizations and enterprises.

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. Томашевський Валентин Миколайович

2. Tomashevskyi Valentyn Mykolaiovych

Кваліфікація: д. т. н., 05.13.06

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. Фісун Микола Тихонович

2. Fisun Mykola T.

Кваліфікація: д. т. н., 05.13.06

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. Бісікало Олег Володимирович

2. Bisikalo Oleh V.

Кваліфікація: д. т. н., 05.13.06

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 Рудакова Ганна Володимирівна

Власне Прізвище Ім'я По-батькові головуючого на засіданні Рудакова Ганна Володимирівна

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності Юрченко Т.А.