Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0416U001806

Особливі позначки: відкрита

Дата реєстрації: 14-04-2016

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Вознюк Тарас Григорович
2. Voznyuk Taras Grygorovych

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: 24-03-2016

Спеціальність за освітою: 8.08010101

Місце роботи здобувача:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): Д 26.001.09

Повне найменування юридичної особи: Київський національний університет імені Тараса Шевченка

Код за ЄДРПОУ: 02070944

Місцезнаходження: вул. Володимирська, 60, м. Київ, Київська обл., 01033, 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: Київський національний університет імені Тараса Шевченка

Код за ЄДРПОУ: 02070944

Місцезнаходження: 01033, м. Київ, вул. Володимирська, 64

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 16.31.03

Тема дисертації:

1. Застосування семантико-синтаксичної тензорної моделі природної мови для аналізу кореферентних зв'язків у текстах

2. The use of semantic-syntactic tensor model of natural language for analysis of coreferential relations in texts

Реферат/анотація:

1. Основним результатом дисертації є розробка та математичне обґрунтування нових алгоритмів ідентифікації та аналізу кореферентних зв'язків у 18 природномовних текстах, що має істотне значення для розв'язання фундаментальної задачі комп'ютерної лінгвістики - семантичного аналізу текстів. Для цього було застосовано тензорну модель природної мови, керуючі простори синтаксичних структур речень та методи машинного навчання. В дисертаційній роботі розроблено новий алгоритм побудови керуючих просторів синтаксичних структур речень, який дозволив отримати зручне та стисле представлення моделі, зменшити розмірність тензора, отримати більш надійний та стійкий опис семантико-синтаксичних зв'язків між словами. В результаті інтеграції розроблених алгоритмів в одну з найкращих систем визначення кореферентних зв'язків Stanford Deterministic Coreference Resolution вдалось покращити за запропонованою

на конференції MUC-6 мірою точність визначення на тестовій вибірці конференції CoNLL-2011 з 61.03% до 64.45%.

2. Thesis is dedicated to improvement of systems for finding coreferential connections in the text using the tensor model of natural language. A large 100Gb text corpus was analyzed and a six-dimensional tensor of typical syntactic structures was built using pipeline data processing architecture and syntactic structures parsers. New efficient algorithms for constructing the control space of syntactic structures was developed to improve the tensor model. This made possible to save more information about the semantic and syntactic relationships in tensors of less dimension. The developed algorithms were used to solve the problem of finding coreferential connections in natural language texts. It was decided to use multi-sieve approach for Coreference resolution, because it demonstrated best result on well-known sets. The paper describes the new sieves designed for the system, which implements classifier of support vector machine for the 28-dimensional feature space. The tensor model of language and control space syntactic structures were used to calculate the feature vector used in machine learning.

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. Марченко Олександр Олександрович
2. Marchenko Alexander Alexandrovich

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. Скобелєв Володимир Геннадійович

2. Скобелєв Володимир Геннадійович

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

Власне Прізвище Ім'я По-батькові:

1. Глибовець Андрій Миколайович

2. Глибовець Андрій Миколайович

Кваліфікація: к.ф.-м.н., 01.05.03

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 Анісімов Анатолій Васильович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Анісімов Анатолій Васильович

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.