

## Авторська довідка

(реферату дипломної роботи магістра)

**Назва дипломної роботи магістра:** Визначення авторства документу з допомогою методів інтелектуального аналізу тексту

назви записувати нижнім регістром (як у реченні)

**Назва (англ.):** A document authorship identification using the text mining methods

переклад англійською

**Освітній ступінь :** ..... магістр

**Шифр та назва спеціальності:** ..... 125 «Кібербезпека»

**Екзаменаційна комісія:** ..... Екзаменаційна комісія №32

напр.: Екзаменаційна комісія №1

**Установа захисту:** Тернопільський національний технічний університет імені Івана Пулюя

напр.: Тернопільський національний технічний університет імені Івана Пулюя

**Дата захисту:** 26 грудня 2019 року

**Місто:** Тернопіль

**Сторінки:**

Кількість сторінок дипломної роботи: 111

Кількість сторінок реферату: 5

**УДК 004.056**

**Автор дипломної роботи**

Прізвище, ім'я, по батькові (укр.): Вівчарик Володимир Михайлович

розкривати ініціали

Прізвище, ім'я (англ.): Vivcharyk Volodymyr

використовувати паспортну транслітерацію (КМУ 2010)

Місце навчання (установа, факультет, місто, країна): ТНТУ ім. І. Пулюя, Факультет комп'ютерно-інформаційних систем і програмної інженерії, Кафедра кібербезпеки, м. Тернопіль, Україна

**Керівник**

Прізвище, ім'я, по батькові (укр.): Карпінський Микола Петрович

повністю

Прізвище, ім'я (англ.): Karpinskyi Mykola

використовувати паспортну транслітерацію (КМУ 2010)

Місце праці (установа, підрозділ, місто, країна): ТНТУ ім. І. Пулюя, Факультет комп'ютерно-інформаційних систем і програмної інженерії, Кафедра кібербезпеки, м. Тернопіль, Україна

Вчене звання, науковий ступінь, посада: доктор технічних наук, професор кафедри кібербезпеки

**Рецензент**

Прізвище, ім'я, по батькові (укр.): Кунанець Наталія Едуардівна

повністю

Прізвище, ім'я (англ.): Kunanets Nataliia

використовувати паспортну транслітерацію (КМУ 2010)

Місце праці (установа, підрозділ, місто, країна): ТНТУ, Факультет комп'ютерно-інформаційних систем і програмної інженерії, Кафедра комп'ютерних наук м. Тернопіль, Україна

Вчене звання, науковий ступінь, посада: доктор наук із соціальних комунікацій, професор кафедри комп'ютерних наук

## Ключові слова

українською: ідентифікація, визначення автора, машинне навчання, класифікація, наївний класифікатор байеса, стоп-слова, стемінг, лематизація

до 10 слів

англійською: identification, author determining problem, machine learning, classification, naive bayes classifier, stop-words, stemming, lemmatization

до 10 слів

## Анотація

українською:

В роботі досліджено основні завдання та можливі сфери застосування задачі визначення авторства деякого документу, обґрунтовано вибір моделі класифікації та програмного середовища Python для практичної реалізації методу визначення автора документу. Проведено тестування імплементованої класифікаційної моделі наївного Байєса для реальних даних, здійснено порівняння основних показників точності моделі для різного розміру простору ознак, проведено аналіз впливу на точність різних методів нормування текстових документів для задач класифікації.

англійською:

The main tasks and possible application areas of author of a document determining problem are investigated; the choice of the classification model and the Python software environment for practical implementation of the method of the document authorship identification is substantiated. Tests of the implemented naive Bayes classification model of for real data were carried out. The basic criteria of model accuracy for different size of feature vector space were compared, the influence of different methods of text documents normalization on accuracy for classification problems was analyzed.

### Бібліографічний опис:

1. Вівчарик В. Особливості методів аналізу текстів для ідентифікації авторства документу [текст] / Вівчарик В. Збірник тез VII науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» – Тернопіль (11 – 12 грудня 2019 р.), ТНТУ, 2019. – с.29.