

Курсова робота

На тему:

"Пошук та визначення плагіату в текстах
українською мовою"

Виконав: студент 1-го року навчання
освітньої програми "Прикладна математика"
спеціальності 113 Прикладна математика
Величко Ростислав Анатолійович
Керівник: доктор технічних наук
Глибовець А.М.

Визначення плагіату

Виявлення плагіату є однією з важливих задач у сфері обробки природної мови (NLP). За допомогою NLP, можна виявляти та аналізувати структуру тексту, визначати стиль письма та інші характеристики, які можуть вказувати на наявність плагіату. NLP використовується для розуміння та обробки природної мови, яка є складною для обробки комп'ютерами через її семантичну та контекстуальну складність. Алгоритми NLP допомагають аналізувати та обробляти текстові дані та забезпечують можливість виявлення плагіату

Актуальність проблеми визначення плагіату українських текстів

Проблема плагіату в українських текстах є досить актуальною з кількох причин. По-перше, зростання кількості наукових досліджень та студентських робіт, які подаються на оцінку, призводить до збільшення кількості підроблених та скопійованих матеріалів. По-друге, в епоху цифрових технологій, дуже легко копіювати та вставляти фрагменти текстів з різних джерел, що робить виявлення плагіату більш складним завданням. Таким чином, визначення плагіату стає все більш важливою проблемою, яка потребує розробки ефективних інструментів для виявлення та запобігання таким порушенням наукової чесності.

Мета курсової роботи та її задачі

Метою роботи є розробка програми на мові Python, яка визначатиме рівень плагіату в українських текстах за допомогою семантичного моделювання (LDA та LSA)

Огляд методів визначення плагіату

Алгоритми роботи з моделями LDA та LSA

Опис структуры програми

Оцінка результатів роботи програми

Висновки