

Дипломна робота: 87 с., 19 рис., 7 табл., 2 додатки, 19 джерел.

Об'єкт дослідження – алгоритми машинного навчання для аналізу текстової інформації.

Предмет дослідження – застосування методів аналізу текстової інформації в задачі класифікації спаму.

Мета роботи – порівняти деякі методи машинного навчання при аналізі текстової інформації на прикладі задачі класифікації спаму.

Актуальність роботи – в наш час інтелектуальний аналіз даних є однією з найбільш перспективних наукових та прикладних областей, що активно розвивається впродовж останніх десятиліть і має практичне застосування в багатьох важливих сферах суспільного життя. Обробка текстової інформації має ключове значення в цій сфері і є важливою складовою сучасного бізнесу.

Результати роботи – при порівнянні різних методів аналізу текстової інформації було виділено недоліки та переваги кожного й проаналізовано доцільність їх використання при розв'язанні певних задач машинного навчання.

Шляхи подальшого розвитку предмету дослідження – розширення кола задач, в яких буде доцільним використання розглянутих методів. Розширення списку методів для аналізу текстової інформації та аналіз їх ефективності.

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ, АНАЛІЗ ТЕКСТОВОЇ ІНФОРМАЦІЇ, МАШИННЕ НАВЧАННЯ, КЛАСИФІКАЦІЯ ТЕКСТІВ, ЗАДАЧА КЛАСИФІКАЦІЇ, РОЗПІЗНАВАННЯ СПАМУ