

## РЕФЕРАТ

Дипломна робота: 135 с., 37 рис., 7 табл., 2 додатки та 13 джерел.

Метою даної роботи є розробка підходу до визначення рейтингів футбольних команд на основі аналізу текстів дописів соціальної мережі Twitter за допомогою алгоритму автоматизованого формування тренувальної вибірки байєсівського класифікатора.

### Результати роботи:

- запропоновано алгоритм автоматизованого формування тренувальної вибірки байєсівського класифікатора на основі латентного семантичного аналізу
- реалізовано систему аналізу текстів дописів соціальної мережі Twitter, яка опрацьовує дані «на-льоту» з використанням запропонованого алгоритму, та визначає рейтинги популярності футбольних команд у Twitter

### Новизна роботи:

- розроблено власний підхід до автоматизованого формування навчальної вибірки байєсівського класифікатора на основі латентного семантичного аналізу
- вперше модель семантичної пам'яті використовується в комбінації із алгоритмом машинного навчання як допоміжний блок формування тренувальної вибірки даних

Результати даної роботи рекомендується використовувати при розробці систем аналізу текстів як каркас до формування навчальних вибірок або цілісної автоматизованої класифікації текстів.

LSA, NB, NAÏVE BAYES, TWITTER API, DATA MINING, TEXT MINING, ТРЕНУВАЛЬНА ВИБІРКА.