

Інформаційна технологія класифікації тональності речень

Вінницький національний технічний університет

Анотація

Описано розв'язання задачі класифікації тональності речень. Розглянуто програмний модуль автоматизованого виявлення в текстах емоційно забарвленої лексики і емоційної оцінки думок по відношенню до об'єктів на основі використання штучних нейронних мереж.

Ключові слова: тональність речень, емоційно забарвлена лексика, оцінка думок.

Abstract

The solving of the task of classification of sentences of sentences is described. The program module for the automated detection of emotionally colored vocabulary and emotional evaluation of opinions in relation to objects based on the use of artificial neural networks is considered in the texts.

Keywords: tone of sentences, emotionally colored vocabulary, judgment of thoughts.

Досить часто зустрічається завдання класифікації текстів - наприклад, визначення тональності (висловлює чи текст позитивну думку або негативну про що-небудь), або рознесення тексту по тематиках [2].

Метою аналізу тональності є визначення емоційного ставлення автора тексту до предмету, що описується у тексті, і визначення властивостей цього емоційного ставлення (позитивне, негативне, схвальне і т.д. і т.п.). Залежно від поставленого завдання нас можуть цікавити різні властивості [1].

У наш час найбільш часто використовуваними в дослідженнях методами є методи на основі машинного навчання з учителем. Суттю таких методів є те, що на першому етапі навчається машинний класифікатор (наприклад, байєсовський) на заздалегідь розмічених текстах, а потім використовують отриману модель при аналізі нових документів. Наведемо короткий алгоритм:

- 1) спочатку збирається колекція документів, на основі якої навчається машинний класифікатор;
- 2) кожен документ розкладається у вигляді вектора ознак (аспектів), за якими він буде досліджуватися;
- 3) вказується правильний тип тональності для кожного документа;
- 4) проводиться вибір алгоритму класифікації і метод для навчання класифікатора;

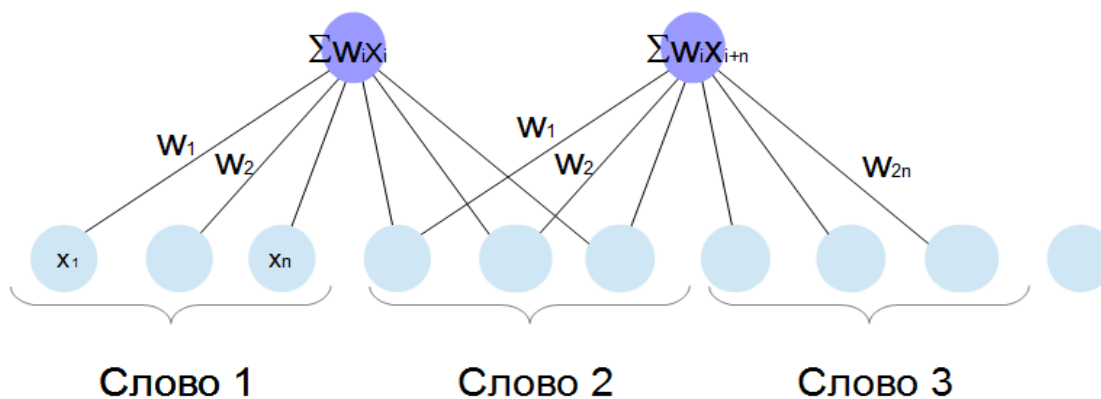


Рис. 1. Навчання нейронної мережі

5) Отриману модель використовуємо для визначення тональності документів нової колекції. [2].

Точність і якість системи аналізу тональності тексту оцінюється тим, наскільки добре вона узгоджується з думкою людини щодо емоційної оцінки досліджуваного тексту. Для цього можуть використовуватися такі показники як точність і повнота. Формула для знаходження повноти:

$$R = \frac{\text{correctly extracted opinions}}{\text{total number of opinions}}$$

де *correctly extracted opinions* — вірно певні думки, *total number of opinions* — загальна кількість думок (як знайдених системою, так і не знайдених). Точність обчислюється за формулою:

$$R = \frac{\text{correctly extracted opinions}}{\text{total number of opinions found by system}}$$

де *correctly extracted opinions* — вірно певні думки, *total number of opinions found by system* — загальна кількість думок знайдених системою. Таким чином, точність виражає кількість досліджуваних текстів, речень або документів, в оцінці яких думка системи аналізу тональності співпало з думкою експерта. При цьому, згідно з дослідженням, експерти зазвичай погоджуються в оцінках тональності конкретного тексту в 79 % випадків. Отже, програма, яка визначає тональність тексту з точністю 70 % робить це майже так само добре, як і людина.

При програмній реалізації модуля було використано інтелектуальні інформаційні технології на основі штучних нейронних мереж. Тестування показало надійну роботу розробленого інтелектуального модуля, дозволило виявити важливі залежності функціональних характеристик програми від параметрів використовуваної нейронної мережі.

Також пропонується у подальшому використовувати для класифікації тональності речень в реальному масштабі часу імпульсні нейронні мережі [3]. Це покращить точність класифікації. Крім того, імпульсні нейронні мережі мають гарні перспективи для апаратної реалізації [4] та найкраще підходять для побудови операційного ядра майбутніх нейрокомп'ютерів [5].

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Тональність речень – [Електронний ресурс]. – Режим доступу: https://uk.wikipedia.org/wiki/Тональність_речень
2. Аналіз тональності речень – [Електронний ресурс]. – Режим доступу: <https://habrahabr.ru/post/149605/>
3. В.Ф.Бардаченко, О.К.Колесницький, С.А.Василецький. Перспективи застосування імпульсних нейронних мереж з таймерним представленням інформації для розпізнавання динамічних образів// УСiМ.-2003-№6.- С. 73-82.
4. Колесницький О. К. Аналітичний огляд апаратних реалізацій спайкових нейронних мереж / О. К. Колесницький // Математичні машини і системи. – 2015. – №1, С.3-19. ISSN 1028-9763.
5. Колесницький О. К. Принципи побудови архітектури спайкових нейрокомп'ютерів / О. К. Колесницький // Вісник Вінницького політехнічного інституту. – Вінниця: УНІВЕРСУМ-Вінниця. – 2014. – №4 (115), С.70-78.

Барановський Владислав Сергійович — студент групи 2КН-18м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця.

Науковий керівник – **Колесницький Олег Костянтинович**, – доцент кафедри комп'ютерних наук ВНТУ, Вінницький національний технічний університет, м. Вінниця.

Baranovskyi Vladislav S. — Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia.

Supervisor — **Kolesnytskyj Oleg K.** — docent of the Computer Sciences Chair, Vinnytsia National Technical University, Vinnytsia.