

Олеся Войтович, Вероніка Островська (Вінниця)

## СИСТЕМА МОНІТОРИНГУ СОЦІАЛЬНИХ МЕРЕЖ ДЛЯ ВИЯВЛЕННЯ ЕМОЦІЙНОЇ ТОНАЛЬНОСТІ ІНФОРМАЦІЙНИХ МЕМІВ

Для аналізу інформаційних мемів в соціальних мережах для виявлення мемів під час проведення інформаційно-психологічних операцій під час інформаційної війни можна використати методи інтелектуального аналізу тексту, а саме аналіз тональності (Sentiment Analysis, SA), який визначає емоційне забарвлення повідомлення (позитивне, негативне або нейтральне) [1]. Однак тексти на природній мові часто містять більше однієї тональної оцінки. Система, яка здійснює аналіз тональності такого тексту за трьома класами, видасть нейтральну оцінку, взаємно компенсуючи позитивну і негативну складові. Це обмеження можна частково подолати, переходячи до аспектного аналізу тональності (Aspect Based Sentiment Analysis, ABSA), що має на меті визначення тональності тексту щодо до тих чи інших аспектів.

**Постановка задачі ABSA**, в порівнянні з SA [2], має такі відмінності:

- $\Phi = \{\varphi^{aspect1}, \dots, \varphi^{aspectm}\}$  – цільова функція аспектного аналізу, що представляє собою об'єднання функцій аналізу тональності по кожному з аспектів;
- $\hat{\Phi} = \{\hat{\varphi}^{aspect1}, \dots, \hat{\varphi}^{aspectm}\}$  – апроксимуюча функція аспектного аналізу тональності, яку потрібно визначити;

**Формалізація задачі.** Для аналізу тональності текстових даних пропонується використати глибоке навчання рекурентних нейронних мереж, яке не викликає складнощів із перенавчанням.

В задачі класифікації, коли кількість можливих класів більше двох, застосовують функцію Softmax (узагальнення логістичної функції для багатовимірною випадку). Функція перетворює вектор  $h^{(t)}$  розмірності  $V$  в вектор тієї ж розмірності, де кожна координата отриманого вектора представлена дійсним числом в інтервалі  $[0,1]$  і сума координат дорівнює 1. На практиці застосовується не просто Softmax-перетворення вектора  $h^{(t)}$ , а повнозв'язний шар прямого поширення, іменованій часто Softmax-шаром нейронної мережі. Таким чином, загальний математичний опис простої рекурентної моделі аналізу тональності має наступний вигляд:

$$\hat{y} = \text{Soft max}(W^{hc} h^{(t)} + b^c),$$

де  $\hat{y} \in R^3$  – вектор ймовірностей приналежності до кожного із 3 класів;  $W^{hc} \in R^{l \times m}$  і  $b^c \in R^c$  – параметри, що навчаються;  $h^{(t)}$  – значення прихованого стану мережі.

Модель рекурентної нейромережі для одноаспектного тернарного аналізу, що забезпечує аналіз текстової фрази «промова була хорошою», на основі Softmax-шару, наведена на рис. 1.

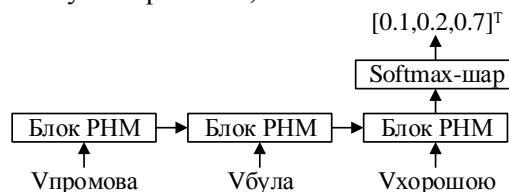


Рис 1. Модель РНМ для одноаспектного тернарного аналізу

$V_{\text{промова}}, V_{\text{була}}, V_{\text{хорошою}}$  – векторні подання слів. Блок РНМ – блок обчислення прихованого стану рекурентної нейронної мережі. На основі обчисленого вектора прихованого стану  $h$  проводиться нелінійне перетворення і класифікація в рамках Softmax-шару.

**Висновки.** Побудова моделі глибокого навчання дозволяє навчити аспектно-орієнтований класифікатор для аналізу публікацій із соціальних мереж.

## Література

1. Pang B., Lee L. Opinion mining and sentiment analysis // Foundations and trends in information retrieval. – 2008. – Vol. 2, 1-2. – Pp. 1-135.
2. Островська В. М., Войтович О. П., Куперштейн Л. М. Метод виявлення тролінгу як інформаційно-психологічної операції в кіберпросторі // Інформаційні технології та комп'ютерне моделювання; матеріали статей МНПК, Івано-Франківськ: п. Голіней О. М., 2017 - С. 212-215.