

Математичне моделювання опінії із застосування лінгвістичних змінних в задачах opinion mining

Сергій Лупенко¹, Олег Пастух², Богдан Хомів³

Кафедра комп'ютерних систем та мереж, Тернопільський національний технічний університет ім. І. Пулюя, УКРАЇНА, м. Тернопіль, вул. Руська, 56,

E-mail: 1. lupenko@ua.fm, 2. pastukh@gmail.com, 3. bogdan.khomiv@gmail.com

The report is devoted to the use the linguistic variables in opinion mining of textual information.

Ключові слова – опінія, емоційне забарвлення, об'єкт, лінгвістична змінна.

I. Вступ

На сьогоднішній день все більше і більше інтернет-користувачів залишають відгуки стосовно певних товарів чи послуг на сторінках соціальних мереж, форумів чи блогів. Під опінією розуміється позитивне чи негативне емоційне забарвлення висловлювання інтернет-користувача щодо певного об'єкта.

Як відомо [1], опінія висловлювань подається п'ятіркою

$$O = (e_j, a_{j,k}, so_{j,k,i,l}, h_i, t_l), \quad (1)$$

де e_j – цільовий об'єкт (сутність); $a_{j,k}$ – аспект/компонента/властивість цільового об'єкта e_j ; $so_{j,k,i,l}$ – значення емоційного забарвлення опінії, висловленого власником опінії h_i стосовно компоненти $a_{j,k}$ цільового об'єкта e_j в час t_l .

Проаналізувавши у доповіді такі системи, як Opinion Observer, OPINE, Review Seer, Red Opal, Web Fountain, Opinion Miner, Bing Shopping, Google Products, Quark Shop зауважено, що числовий еквівалент емоційного забарвлення $so_{j,k,i,l}$ може набувати таких значень: +1 – позитивне висловлювання, -1 – негативне висловлювання, 0 – нейтральне висловлювання, або приймати інші числові значення в залежності від вибраної шкали, що в свою чергу, не в повній мірі відображає опінію користувача щодо певного об'єкта.

II. Застосування лінгвістичних змінних

У доповіді запропоновано описати $so_{j,k,i,l}$ відповідно до (1) у вигляді лінгвістичної змінної, що дасть змогу представити емоційне забарвлення висловлювань за допомогою нечітких значень відповідно до її термножини.

Таким чином $so_{j,k,i,l}$ може бути подане у вигляді набору [2]:

$$so_{j,k,i,l} = (b, P, X, G, M), \quad (2)$$

де: b – найменування лінгвістичної змінної; P – множина її значень, що є іменами нечітких змінних, областю визначення, кожної з яких є множина X ; G – синтаксична процедура, що дає змогу оперувати елементами терм-множини T ; M – семантична процедура, що дає змогу перетворити кожне нове значення лінгвістичної змінної, утвореною процедурою G , у нечітку змінну, тобто сформувані відповідну нечітку множину.

III. Фазифікація та дефазифікація

Власник опінії H_i може виступати в ролі експерта, а емоційне забарвлення $SO_{j,k,i,l}$ можна описати у вигляді лінгвістичної змінної (2).

Нехай експерт визначає роздільну здатність фотокамери за допомогою поняття «не погана», при цьому мінімальна роздільна здатність буде дорівнювати 0 МПікселів, а максимальна – 15 МПікселів.

Формалізацію висловлювання експерта пропонується здійснювати за допомогою лінгвістичної змінної (2), де b – роздільна здатність; P – {«погана», «задовільна», «хороша»}; X – [0, 15]; G – процедура утворення нових термів за допомогою зв'язувань «і», «або» і модифікаторів типу «дуже», «не», «злегка» та ін.; M – процедура задання на X = [0, 15] нечітких підмножин $A1$ = «погана роздільна здатність», $A2$ = «задовільна роздільна здатність», $A3$ = «хороша роздільна здатність».

Таким чином можна провести фазифікацію та побудувати функції належності лінгвістичної змінної (2).

Дефазифікацію можна проводити такими методами, як: центр ваги, медіана, найбільший з максимумів, найменший з максимумів, центр максимумів.

ВИСНОВОК

У доповіді запропоновано використовувати математичний апарат нечіткої логіки із застосуванням лінгвістичних змінних для опису опінії. Такий опис дасть змогу представити емоційне забарвлення висловлювань за допомогою нечітких значень відповідно до терм-множини лінгвістичної змінної, що в свою чергу забезпечить більш повне відображення опінії користувача щодо певного об'єкта.

Література

1. Bing L. Web Data Mining. Exploring Hyperlinks, Contents and Usage Data, Second Edition : / Liu Bing. – Springer, 2011. – 622 p.
2. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику. / С.Д. Штовба // Винница: Издательство винницкого государственного технического университета, 2001. – 198 с.