

ОПТИМІЗАЦІЯ ПРОЦЕСУ ОЦІНЮВАННЯ ВІДКРИТИХ ВІДПОВІДЕЙ РЕСПОНДЕНТІВ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ

¹ Вінницький національний технічний університет

Розглянуто особливості задачі оцінювання відповідей респондентів, основні типи відповідей на питання, наголошено на перевагах використання відкритих питань та розглянуто основні проблеми оцінювання відповідей на такі питання при великій кількості респондентів з урахуванням їх особливостей. В доповіді розглянуто різні методи подання словесної відповіді у векторному вигляді. Розглянуто різні метрики обчислення близькості між окремими відповідями та запропоновано метод обчислення відстані між відповідями респондентів на основі модифікованої відстані Жаккара. Запропонована відстань враховує близькість за n -грамами. Запропоновано математичну модель процесу оцінювання у вигляді задачі оптимізації, що полягає в пошуку мінімальної послідовності відповідей респондентів, яку потрібно оцінити експерту, щоб середня відстань неоцінених відповідей до оцінених відповідей не перевищувала порогового значення. Розроблено метод розв'язання задачі оптимізації з використанням методів штучного інтелекту, а саме методів кластеризації. Описано найкращий та найгірший випадки вхідних даних для задачі. Основна ідея запропонованого методу полягає у покроковій кластеризації неоцінених відповідей респондентів та виборі характеристичної відповіді з найбільшого кластера для ручного оцінювання експертом. Сформульовано поняття характеристичної відповіді респондента та розроблено метод її вибору на основі нормального закону розподілу відстаней у межах кластера. Запропоновано автоматично оцінювати неоцінені відповіді шляхом знаходження мінімальної відстані до оцінених відповідей. Наведено результати проведених експериментальних досліджень, що полягали у виборі набору даних, методу векторизації та методу кластеризації. Наведено порівняльні таблиці та графіки використання різних методів. Проведені дослідження емпірично довели адекватність розроблених математичних моделей та методів та їх незалежність від особливостей контексту запитань і відповідей та особливостей респондентів, таких як країна походження або проживання респондента, мова спілкування, персональних особливостей. Розроблені моделі та методи можуть використовуватись в системах дистанційного навчання в якості підтримки роботи викладача.

Ключові слова: оцінювання відкритих відповідей, обробка тексту, великі дані, кластеризація, відстань Жаккара.

Вступ

В сучасному світі освітні послуги стали глобальними, мобільними та доступними майже усім. Викладачі працюють в аудиторіях та онлайн. Велику популярність здобули навчальні онлайн-платформи, на яких слухачі обирають курси, дивляться відео-матеріали, працюють у віртуальних лабораторіях тощо. Невід'ємною частиною навчання є процес оцінювання знань. Найбільш популярними видами завдань є тестові запитання з однією або багатьма правильними відповідями. Оцінювання таких відповідей відбувається простим співставленням з наперед зазначеною правильною відповіддю, воно є об'єктивним та швидким. На відміну від тестів, відкриті запитання використовуються рідше, оскільки неможливо сформулювати універсальну правильну відповідь. Також бувають запитання, за допомогою яких викладач спонукає студента до висловлення власної позиції з того чи іншого питання, та на які не може бути «неправильної» відповіді. Відкриті запитання потребують індивідуального ручного оцінювання викладачем (експертом). Складність оцінювання напряму пов'язана з кількістю опитуваних (респондентів) [1, 2].

Метою роботи є розроблення математичної моделі процесу оцінювання відкритих відповідей та методу визначення оптимальної послідовності відповідей, яку необхідно оцінити вручну, щоб на їх

основі оцінити всі інші неоцінені відповіді.

Результати дослідження

Виконаємо постановку задачі. Нехай ϵ множина відповідей респондентів на деяке питання:

$$A = \{a_1, a_2, a_3, \dots, a_n\}, \quad (1)$$

де n – кількість відповідей на запитання.

Всю множину A є супермножиною множин оцінених відповідей E та неоцінених відповідей U , таких що

$$\begin{aligned} E \cap U &= \emptyset, \\ E \cup U &= A. \end{aligned} \quad (2)$$

Елемент з множини U можна автоматично перенести у множину E , якщо існує такий елемент e_m , що відстань $d(u_i, e_m) \leq d_{min}$, де d_{min} – порогова мінімально допустима відстань між елементами, щоб можна було вважати їх ідентичними та присвоїти u_i таку саму оцінку, якою оцінено e_m .

Потрібно знайти таку впорядковану підмножину елементів множини U , яка вручну буде переведена у множину E , щоб залишок неоцінених елементів перенести у множину оцінених елементів автоматично.

Для розв'язання поставленої задачі було розроблено метод, що полягає у знаходженні кластерів серед відповідей респондентів за мінімальним пороговим значенням відстані між відповідями [3-5]. Вхідні дані:

- нехай ϵ множина об'єктів U , де кожний об'єкт представлений у вигляді вектору ознак у просторі ознак \mathbb{R}^d ;
- необхідно знайти розбиття цієї множини на k кластерів C_1, C_2, \dots, C_k , тобто
-

$$U = C_1 \cup C_2 \cup \dots \cup C_k, \quad C_i \cap C_j = \emptyset \quad \forall i \neq j. \quad (3)$$

Найкращий випадок: всі елементи множини U утворюють один кластер, та оцінка однієї відповіді з цієї множини призведе до автоматичного оцінювання всіх інших відповідей.

Найгірший випадок: елементи множини U утворюють таку кількість кластерів, що

$$|C_i| = 1 \quad \forall i.$$

Це випадок, коли всі відповіді різні, наприклад це відповіді на запитання про персональні дані респондента.

Ключовим поняттям є поняття обчислення відстані між відповідями. Оскільки відкриті питання не передбачають повного співпадіння з врахуванням порядку слів у реченні, то рекомендується використовувати відстань Жаккара [6-7]:

$$d_j(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}, \quad (4)$$

де A і B – множини слів в текстах відповідей.

Для підвищення надійності порівняння пропонується під словами розуміти n -грами тестів ($n \leq 3$) [8].

Запропонований метод було реалізовано у вигляді програми мовою Python з використанням пакету scikit-learn [9]. В якості набору даних було використано набір ASAG-Dataset [10]. Набір складається з 17 запитань, на які давали відповіді 38 студентів, загальна кількість відповідей 646.

Висновки

Проведені дослідження показали, що запропонований метод дозволяє спростити процес оцінювання відповідей респондентів на відкриті питання. Метод є мовонезалежним та дозволяє автоматично оцінювати відповіді з обраною точністю. В якості подальших досліджень планується розробка методу швидкого обчислення відстані між текстами довільної довжини.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] Burstein, J., Chodorow, M., Leacock, C. Automated Essay Evaluation: The Criterion® Online Writing Service. Lawrence Erlbaum Associates, 2004. 224 p. ISBN 978-0805846330.
- [2] Shermis, M. D., Burstein, J. C. (Eds.). Automated Essay Scoring: A Cross-disciplinary Perspective. Routledge, 2013. 256 p. ISBN 978-0-415-89602-0.
- [3] Manning, C. D., Raghavan, P., Schütze, H. Introduction to Information Retrieval. Cambridge: Cambridge University Press, 2008. 482 p. ISBN 978-0-521-86571-5.
- [4] Tan, P.-N., Steinbach, M., Kumar, V. Introduction to Data Mining. Pearson, 2006. 769 p. ISBN 978-0-321-32136-7.
- [5] Aggarwal, C. C., Reddy, C. K. (Eds.). Data Clustering: Algorithms and Applications. Chapman and Hall/CRC, 2014. 588 p. ISBN 978-1-4665-5821-2.
- [6] Roy, D., Kanaan-Izquierdo, S., Ojha, A., & Srivastava, G. Machine Learning for Text. CRC Press, 2021. 320 p. ISBN 978-0367612497.
- [7] Петренко, О. С., Коваленко, О. М. Методи машинного навчання в обробці природної мови: навчальний посібник. Київ: КНУ імені Тараса Шевченка, 2020. 240 с. ISBN 978-966-930.
- [8] Іванов, М. П. Автоматизація обробки текстової інформації: моделі та методи. Львів: Видавництво Львівської політехніки, 2018. 300 с. ISBN 978-617-607.
- [9] Müller, A. C., Guido, S. Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media, 2016. 394 p. ISBN 978-1-4493-6941-5.
- [10] DigiKlausur. ASAG Dataset [Електронний ресурс] / DigiKlausur // GitHub. – Режим доступу: <https://github.com/DigiKlausur/ASAG-Dataset>. – Назва з екрану.

Гришук Тетяна Вікторівна — канд. техн. наук, доцент, доцент кафедри комп'ютерних систем управління.

Вінницький національний технічний університет

T. V. Gryshchuk¹

Optimization of the process of assessment of open feedbacks of respondents by means of artificial intelligence

¹Vinnitsia National Technical University

In this work, the characteristics of the task of assessment of respondents' answers, the main types of answers to questions,

the advantages of using open-ended questions are emphasized, and the main problems of evaluating answers to such questions with a large number of respondents, taking into account their characteristics, are considered. The report discusses various methods of presenting a verbal response in vector form. Various metrics for calculating the proximity between individual answers are considered and a method for calculating the distance between respondents' answers based on the modified Jaccard distance is proposed. The proposed distance takes n-gram proximity into account. A mathematical model of the evaluation process is proposed in the form of an optimization problem, which consists in finding the minimum sequence of respondents' answers, which the expert needs to evaluate, so that the average distance between unrated answers and rated answers does not exceed the threshold value. A method of solving the optimization problem using artificial intelligence methods, namely clustering methods, has been developed. The best and worst case of input data for the problem are described. The main idea of the proposed method is the step-by-step clustering of unrated answers of respondents and the selection of a characteristic answer from the largest cluster for manual evaluation by an expert. The concept of the respondent's characteristic answer is formulated and the method of its selection based on the normal distribution law of distances within the cluster is developed. It is proposed to automatically evaluate unrated answers by finding the minimum distance to rated answers. The results of the conducted experimental studies, which consisted in the selection of the data set, the vectorization method, and the clustering method, are given. Comparative tables and graphs of the use of various methods are given. The conducted studies empirically proved the adequacy of the developed mathematical models and methods and their independence from the specifics of the context of questions and answers and the characteristics of the respondents, such as the country of origin or residence of the respondent, the language of communication, and personal characteristics. The developed models and methods can be used in distance learning systems as support for the teacher's work.

Keywords: assessment of open responses, text processing, big data, clustering, Jaccard distance.

Gryshchuk Tetiana Viktorivna — Cand. Sc. (Eng.), Assistant Professor, Assistant Professor of the Chair of Computer Control Systems