УДК 004.8
Гардиш Д.О.[1], Кліменко В.І.[2], Мазурець О.В.[3]
[1] студ., Хмельницький національний університет
[2] викл. кафедри комп'ютерних наук, Хмельницький національний університет
[3] канд. техн. наук, доц., доц. кафедри комп'ютерних наук, Хмельницький національний університет

ПІДХІД ДО АНАЛІЗУ ВІДПОВІДНОСТІ МНОЖИНИ ТЕСТОВИХ ЗАВДАНЬ СЕМАНТИЧНІЙ СТРУКТУРІ НАВЧАЛЬНИХ МАТЕРІАЛІВ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ

Навчальні матеріали можуть бути представленні різними компонентами. Найчастіше це відбувається за допомогою тексту. Він містить багато теорій, фактів, правил, прикладів для засвоєння інформації. Також інформація може бути представлена за допомогою формул чи таблиць, які зазвичай мають більш організований спосіб відображення інформації. Діаграми та графіки часто використовують для візуального відображення зв'язків між різними елементами. А ілюстрації часто застосовують задля показу конкретних об'єктів, процесів або навіть для зацікавлення [1].

Для перевірки засвоєного матеріалу використовують різні підходи. Ці підходи допомагають зрозуміти ефективність навчального процесу та здійснити аналіз навчальних досягнень. Основними методами перевірки є усна перевірка, практична діяльність, виконання завдань, спостереження, написання есе, контрольних та лабораторних робіт. Проте, найчастіше для перевірки знань використовують тестування [2].

Сьогодні, тестування вийшло на новий рівень популярності, завдяки своїй швидкості та зручності створення і проведення. Онлайн-тестування надають можливість швидко та без зайвих зусиль перевірити знання кожного. В освітньому процесі велика увага зосереджена на вдосконаленні якості та ефективності навчальних матеріалів, зокрема, тестових завдань, які є цінним інструментом для оцінки знань та розуміння учнів. Проте, варто не забувати, що звичайні методи оцінювання зазвичай обмежені тільки перевіркою правильності відповідей без урахування семантичного зв'язку та контексту [3, 4]. В результаті чого це може призвести до необ'єктивних результатів та неефективного використання навчальних матеріалів.

Метою роботи є створення методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів засобами обробки природної мови.

Метод аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів засобами обробки природної мови повинен надавати можливість зіставляти семантичну структуру з ключових слів навчальних матеріалів, обраховану за допомогою методу дисперсійної оцінки та семантичну структуру тестових завдань. На рисунку 1 схематично зображено етапи виконання методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів засобами обробки природної мови. На вхід подається обраний текстовий контент навчальних матеріалів, а також база тестових завдань, яка відповідає обраному обсягу навчального матеріалу.

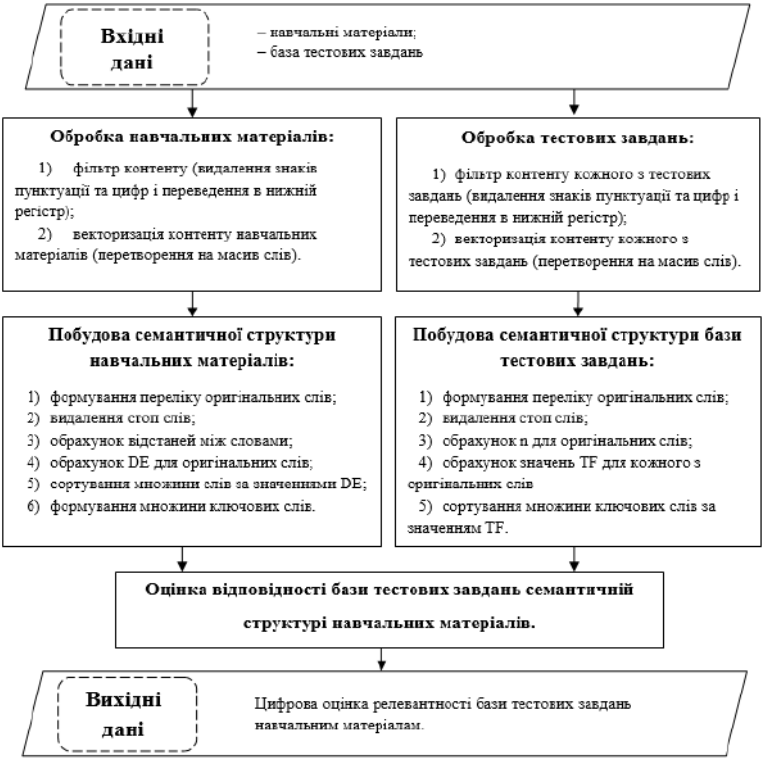

Рисунок 1 – Схема та кроки методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів

Після прийому вхідних даних, починається їх обробка. На початку виконується фільтрація навчальних матеріалів. Фільтрація включає в себе видалення знаків пунктуації та цифр, а також переведення всього тексту в нижній регістр. Після чого відбувається векторизація текстового контенту, перетворення його на масив слів. Для бази тестових завдань для початкової обробки застосовуються аналогічні етапи, що й до навчальних матеріалів.

Отримавши два переліки слів, далі виконується їх аналіз. Тобто, відбувається формування множини ключових слів для навчальних матеріалів за допомогою дисперсійного оцінювання DE та для тестових завдань за допомогою частотного аналізу TF. Для цього першим етапом відбувається формування множини оригінальних слів. Потім, з тексту видаляються стоп-слова, такі як займенники, прийменники, числівники та сполучники. Далі для навчальних матеріалів виконується обрахунок відстані між словами і визначення семантичної важливості слів за допомогою дисперсійного оцінювання DE; для бази ж тестових завдань виконується обрахунок кількостей появ оригінальних слів і визначення семантичної важливості слів за допомогою за частотного аналізу TF. Після чого, шляхом сортування множин оригінальних слів за зменшенням їх оцінок семантичної важливості та обмеженням за граничним значенням, відбувається формування множин ключових слів як для навчальних матеріалів, так і для тестових завдань.

Після цього, коли визначено ключові терміни навчальних матеріалів та ключові терміни тестових завдань, виконується етап оцінювання релевантності бази тестових завдань семантичній структурі навчальних матеріалів. Для цього перевіряється, чи усі важливі терміни, які містяться у навчальних матеріалах, наявні в тестових завданнях.

Вихідними даними методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів є числова оцінка відповідності множини тестових завдань семантичній складовій навчальних матеріалів.

У межах аналізу ефективності розробленої методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів, було проведено дослідження, у якому порівнюється результати оцінювання релевантності бази тестових завдань семантичній структурі навчальних матеріалів, одержані розробленою тестовою інформаційною системою, з оцінками експерта (рисунок 2).

Для проведення дослідження було використано дев’ять навчальних матеріалів разом із відповідними множинами тестових завдань. В результаті кожного аналізу було отримано числовий відсоток перекриття тестами контенту навчального матеріалу.

Порівнявши результати в наведеній діаграмі, можна зробити висновок, що метод працює коректно, а різниця між висновком програми та висновком

експерта складає середню різницю 8,261% (в кращому випадку 2%, а в гіршому 18%).

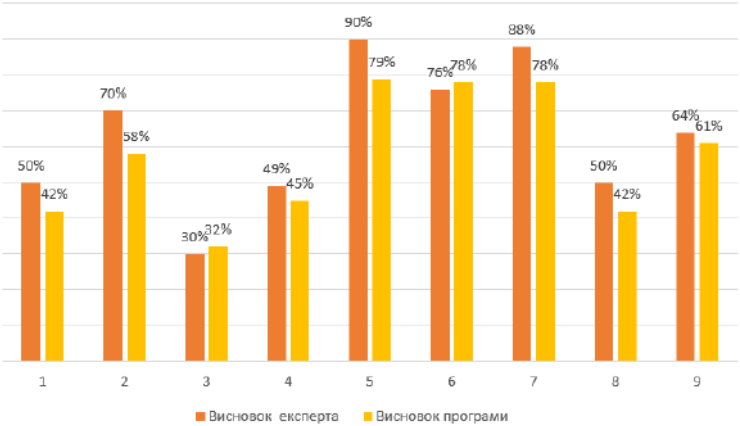

Рисунок 2 – Діаграма результатів дослідження ефективності методу аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів засобами обробки природної мови

Отже, було запропоновано підхід до аналізу відповідності множини тестових завдань семантичній структурі навчальних матеріалів засобами обробки природної мови, який було імплементовано у вигляді відповідного методу. Проведене дослідження практичного застосування розробленого методу виявило його працездатність та можливість використання для аналізу відповідності множин тестових завдань семантичній структурі навчальних матеріалів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. O. Barmak, I. Krak, O. Mazurets, S. Pavlov, A. Smolarz, W. Wojcik, Research of efficiency of information technology for creation of semantic structure of educational materials. Advances in Intelligent Systems and Computing, 2020, vol. 1020, pp. 554–569.

2. Mazurets O. Method for Using of Production Knowledge Model in Intellectual Adaptive Semantic Testing. Innovations in the Education of the Future: Integration of Humanities, Technical and Natural Sciences : International collective monograph. Faculty of Information Technology, Czech Technical University. Prague, 2023. P.340-360.

3. Гардиш Д.О., Кліменко В.І. Прикладні аспекти автоматизованого оцінювання відповідності кейса тестових завдань семантичній складовій

навчальних матеріалів. Збірник наукових праць за матеріалами XV Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2023». Хмельницький, 2023. с. 57-63.

4. Hardysh D., Klimenko V., Mazurets O. Intelligent System for Automated Assessment of Test Tasks Sets Conformity to Semantic Structure of Educational Materials. Proceedings of the XV International Scientific and Practical Conference. «Distance learning: problems, ways of development and the latest technologies». Munich, Germany. Pp. 276-282.