**УДК:004.75.032**

***Стисло Тарас Романович****, старший викладач кафедри*

*інформаційних технологій, Заклад вищої освіти*

*«Університет Короля Данила», м. Івано-Франківськ* [*ORCID*](https://orcid.org/0000-0002-2377-7985)

***Лозко Любомир Михайлович,*** *магістр, Заклад вищої освіти*

*«Університет Короля Данила», м. Івано-Франківськ*

***Федорук Едуард Ігорович,*** *аспірант, ІФНТУНГ, м. Івано-Франківськ*

***Касянчук Ігор Петрович,*** *аспірант, ІФНТУНГ, м. Івано-Франківськ*

**ШАБЛОНИ ВИДОБУВАННЯ ДАНИХ В ЗАДАЧАХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ**

Електронна комерція змінила вигляд більшості бізнес-функцій конкурентних підприємств. Інтернет-технології бездоганно автоматизували процеси інтерфейсу між клієнтами та роздрібними торговцями, роздрібними торговцями та дистриб’юторами, дистриб’юторами та фабриками, а також фабриками та їхніми постачальниками. Загалом, електронна комерція та електронний бізнес (електронна комерція) уможливили онлайн-транзакції. Крім того, створення великомасштабних даних у реальному часі ніколи не було таким простим. Оскільки дані, що стосуються різних видів бізнес-транзакцій, є легкодоступними, доцільно звернутися до послуг інтелектуального аналізу даних, щоб отримати бізнес сенс із цих наборів даних**[1,2]**.

Інтелектуальний аналіз даних (DM) має своєю основною метою створення неочевидної, але корисної інформації для осіб, які приймають рішення, з дуже великих баз даних. Різноманітні механізми цього покоління включають абстракції, агрегації, узагальнення та характеристики даних. Ці форми, у свою чергу, є результатом застосування складних методів моделювання з різноманітних галузей статистики, штучного інтелекту, управління базами даних і комп’ютерної графіки.

**Аналіз підходів до видобування даних в задачах електронної комерції.**

У різних областях електронної комерції, пов’язаних із конкретними даними, підприємства-учасники можуть збільшити свою економічну віддачу, використовуючи знання, отримані з просторових баз даних. Однак на практиці просторові дані часто розподіляються на кількох сайтах. Через безпеку, конкуренцію та відсутність відповідних алгоритмів виявлення знань просторова інформація з таких фізично розосереджених місць часто не використовується належним чином. У пропонованій системі централізований сервер збирає власні просторові дані для конкретних сайтів від підписаних підприємств, а також відповідні дані з публічних і комерційних джерел і об’єднує знання для надання цінної управлінської інформації клієнтам, які підписалися. Програмне забезпечення для інтелектуального аналізу просторових даних інтерфейсує цю базу даних для отримання цікавих і нових знань із даних. Конкретні цілі включають краще розуміння просторових даних, виявлення зв’язків між просторовими та непросторовими даними, створення баз просторових знань, оптимізацію запитів і реорганізацію даних у базах просторових даних.

Проблеми, пов’язані з інтелектуальним аналізом просторових даних, включають численні шари даних, відсутні атрибути та високий рівень шуму через низьку чутливість інструментів і просторову інтерполяцію на рідко зібраних атрибутах. Щоб вирішити деякі з цих проблем, дані очищаються шляхом видалення дублікатів, видалення викидів і фільтрації через медіа- фільтр із заданим розміром вікна.

За допомогою процесу виявлення знань (KDD) створені алгоритми навчання, які виконують моделювання даних, використовуючи набори даних з різних областей. Щоб покращити здатність прогнозувати при роботі з неоднорідними просторовими даними, підхід, який використовується базується на ідентифікації просторових регіонів, що мають подібні характеристики, за допомогою алгоритму кластеризації.

Алгоритм кластеризації використовується для поділу багатовимірних даних на значущі підгрупи (кластери), щоб шаблони в межах кластера були більш схожими один на одного, ніж шаблони, що належать до різних кластерів. Збір даних на правильному рівні абстракції дуже важливий. Журнали веб-сервера спочатку призначалися для налагодження програмного забезпечення сервера. Тому вони передають дуже мало корисної інформації про транзакції, пов’язані з клієнтами. Підходи, включаючи сеанси веб-журналів, можуть дати кращі результати. Бажаною альтернативою було б, щоб сам сервер додатків реєстрував дії, пов’язані з користувачем. Це, безперечно, буде багатшим на семантику порівняно з веб-журналами без стану, і його легше підтримувати порівняно з веб-журналами з повним станом. Розробляючи форми інтерфейсу користувача, потрібно враховувати питання DM.

## Висновки.

Ми оглянули, як веб-майнінг у широкому значенні, DM у застосуванні до електронної комерції можна застосувати для покращення послуг, що надаються засобами електронної комерції. Статистика, штучний інтелект та методи бази даних є актуальними засобами для DM загалом. Зі зростанням інтересу до поняття семантичної мережі все більше сайтів використовують структуровану семантику та онтології домену як частину дизайну сайту, створення та доставки вмісту. Основним завданням для систем персоналізації наступного покоління є ефективна інтеграція інтелектуального аналізу даних в електронній комерції, семантичні знання з онтологій домену в різні частини процесу, включаючи фази підготовки даних, виявлення шаблонів і рекомендацій.

Література:

1. Alghanam, Orieb Abu, Sumaya N. Al-Khatib, and Mohammad O. Hiari. "Data mining model for predicting customer purchase behavior in e-commerce context." International Journal of Advanced Computer Science and Applications 13.2 (2022).

2. Yan, Yongsai, and Qiwei Liu. "An Overview of the Application of Data Mining Technology in E-Commerce." Academic Journal of Engineering and Technology Science 2.2 (2019): 76-88.