

*Шафорост В. В., Корнієнко К. К.,  
студенти 3 курсу спеціальності 122  
«Комп'ютерні науки»  
Хмелівський Ю.С., асистент  
кафедри інформаційних технологій.*

## **ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ТА ПРИЙНЯТТЯ РІШЕНЬ**

*Донецький національний університет імені Василя Стуса*

Завдяки прогресу в технологій обробки та зберігання даних суспільство отримало можливість зберігати та отримувати інформацію більш складним способом. Величезні об'єми інформаційних ресурсів у різних аспектах. Робота деяких організацій, реалізація підприємницьких, промислових та інших заходів тепер регулярно супроводжується процедурами реєстрації та відстеження. Документальне підтвердження всіх тонкощів його діяльності. Стало зрозуміло, що без оптимальної обробки, немає необхідності в таких потоках інформації. Особливості теперішніх вимог до зазначеної обробки полягає у наборах даних які є необмежені у обсязі та демонструють неоднорідність за своєю природою, охоплюючи як кількісні, якісні, так і текстові виміри. Також отримані результати мають бути конкретними та зрозумілими, а засоби для роботи з необробленими даними мають бути зручними для користувача.

Покращення методів обробки та зберігання інформації призвело до значного збільшення обсягу зібраної інформації, що потребує вдосконалення засобів її аналізу та обробки.

Технології аналізу отриманої інформації, що спираються на класичні статистичні підходи, мають і свої негативні риси. Певні засоби базуються на застосуванні середніх показників, що ускладнює визначення фактичного статусу досліджуваної сфери. Методи математичної статистики проявили себе насамперед як дуже корисні для перевірки апріорно сформульованих гіпотез і для проведення дослідницького обробки інформації, що є невід'ємною частиною аналітичної обробки інформації. Крім того, звичайні статистичні методи не враховують незвичні випадки. Однак певні незвичні випадки можуть становити самостійний предмет зацікавленості для дослідження, охарактеризовуючи конкретні виняткові, але значущі явища. Сама ідентифікація цих спостережень, не кажучи про їх вивчення, дослідження наслідків і детальний розгляд, може виявитися корисною для розуміння глибинної суті досліджуваних об'єктів або явищ. Згідно з сучасними дослідженнями, події такого характеру можуть стати ключовими для майбутньої поведінки та розвитку складних систем.

Вищезгадані «необроблені дані» являють собою знання, які можна застосувати в процесі ухвалення рішень [1]. У результаті виникла необхідність у

розробці нових підходів до обробки та аналізу даних, що використовуються сьогодні.

Алгоритми, які використовуються в штучному інтелекті, потребують значної кількості розрахунків. Колись це було обмежуючим фактором для інтелектуального аналізу даних, але сучасне збільшення продуктивності передових процесорів пом'якшило цю проблему. Тепер є можливість здійснити швидкий та надійний аналіз великої кількості записів за прийнятні терміни. Інтелектуальний аналіз даних є перетином кількох наук, як прикладна статистика, розпізнавання образів, штучний інтелект та теорія баз даних, що сприяють її розвитку [2].

Методи та алгоритми включають штучні нейронні мережі, дерева рішень, символічні правила, опорні векторні машини, лінійну регресію, кореляційно-регресійний аналіз, аналіз правил асоціації, наприклад алгоритм Априорі, методи обмеженого перерахування, еволюційне програмування та генетичні алгоритми, різні методи візуалізації даних та багато інших методів. Варто відмітити, що більша частина методик які використовують штучний інтелект були створені в рамках теорії штучного інтелекту. Немає одностайної згоди щодо того, які завдання потрібно віднести до інтелектуального аналізу даних. Більшість надійних джерел цитують такі методи: класифікація, кластеризація, прогнозування, асоціація, візуалізація, аналіз виявлення викидів, оцінка, аналіз взаємозв'язків і підсумовування.

Класифікація. Завдання класифікації є доволі популярним у галузі науки про дані. Унаслідок чого виконання задачі класифікації ідентифікують характеристики, що описують певні групи об'єктів у заданому наборі інформації – класи, на основі таких ознак, новий об'єкт можна віднести до певного класу. Існує різноманітність методів, які можна використовувати для вирішення завдання класифікації, такі як нейронні мережі, індукція дерева рішень і найближчий сусід [3].

Асоціація. Під час виконання завдання виявлення асоціативних правил виявляються регулярності серед взаємопов'язаних випадків у певному комплекті інформації. Розрізнення задачі асоціації від минулих завдань інтелектуального аналізу інформації полягає в його основі на одночасних подіях, а не на аналізі особливостей об'єкта для ідентифікації шаблону. Алгоритм Априорі є найбільш широко визнаним алгоритмом для розв'язання завдань аналізу правил асоціації.

Послідовність. Термін «послідовність» відноситься до послідовної асоціації. Дотримуючись послідовності, стає можливим ідентифікувати часові закономірності між транзакціями. Мета послідовності аналогічна меті асоціації, хоча й спрямована на ідентифікацію закономірностей не серед збігаючих подій, а скоріше серед подій, пов'язаних у часі, що відбуваються протягом певних проміжків часу. Це завдання також відоме як проблема виявлення узгодженого шаблону інтелектуального аналізу даних. Правило наступності передбачає, що подія Y відбудеться через заданий інтервал після події X [4].

Кластеризація. Кластеризація є логічним наслідком поняття класифікації. Це більш комбіноване завдання. Її унікальний аспект являє собою, те, що початково класифікація об'єктів не визначена. Кластеризація призводить до поділу об'єктів на групи.

Візуалізація. За допомогою візуалізації створюється графічне представлення аналізованої інформації. Графічні методи застосовуються для візуалізації, щоб показати закономірності в певній інформації. Наприклад методом візуалізації є зображення інформації у 2D та 3D вимірах.

Прогнозування. Завданням прогнозування на основі доступних даних є оцінка відсутніх або майбутніх значень числових показників, що вирішується у процесі його розв'язання. Математична статистика та нейронні мережі зазвичай використовуються для вирішення таких завдань.

Сфера застосування інтелектуального аналізу інформації є не обмеженою, тобто вона скрізь, де є якийсь даний. Проте методи інтелектуального аналізу даних викликають інтерес до себе у комерційних підприємств діяльність, яких базується на розгортанні своїх проектів та ідей на основі інформаційних джерел. Інтелектуальний аналіз даних є дуже цінним для діяльності керівників і аналітиків у повсякденному житті.

Отже інтелектуальний аналіз даних є дуже корисним для комерційних підприємств, які шукають способи отримати конкурентну перевагу, а також він є цінним допоміжним інструментом для покращення своєї повсякденної діяльності. Тому, можна зробити висновок, що інтелектуальний аналіз даних може допомогти підприємствам підвищити ефективність та прибутковість.

#### Список літератури:

1. *Методи інтелектуального аналізу даних [Електронний ресурс] – Режим доступу до ресурсу: <https://buklib.net/books/24506/>*
2. *Гороховатський В.О., Творошенко І.С. Методи інтелектуального аналізу та оброблення даних: навч. посібник. – Харків: ХНУРЕ, 2021. – 92 с.*
3. *Джеймс Г. Уїттон А Хасті Т. Тібшірані Р. Введення в статистичне навчання з прикладами на мові R. Пер. з англ. С. Е. Масціцького - М.: ДМК Пресс, 2017. - 456 с.*
4. *Штовба С.Д., Козачко О.М. Machine learning: стартовий курс : електронний навчальний посібник. Вінниця: ВНТУ, 2020. 81 с.*