

## ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНИХ ОБЧИСЛЕНЬ ТА УКРАЇНСЬКИЙ РИНОК

© Юрчак І.Ю., 2007

**Проаналізовано основні перспективні застосування новітніх технологій інтелектуального аналізу даних. Наведені переваги застосування новітніх алгоритмів для розв’язання реальних задач, які постають перед сучасними підприємствами. Висвітлено основні причини, що гальмують їх просування в Україні.**

**Basic perspective applications of the newest technologies of intellectual analysis of data are analysed. Advantages of application of new algorithms are pointed for the decision of the real tasks which are put before modern enterprises. Principal reasons which brake their advancement in Ukraine are reflected.**

**Вступ.** Комп’ютерні технології з організацією інтелектуальних обчислень переживають свій розквіт. Це пов’язано переважно з потоком нових ідей, що виходять з галузі комп’ютерних наук, яка утворилась на перетині штучного інтелекту, статистики та теорії баз даних. Нині відбувається стрімкий ріст числа програмних продуктів, що використовують нові технології, а також типів задач, де їх застосування надає значного економічного ефекту [4, 5, 6, 7]. Елементи автоматичної обробки і інтелектуального аналізу даних стають невід’ємною частиною концепції електронних сховищ даних та технологій інтелектуальних обчислень. Простий доступ користувача до сховища даних забезпечує отримання відповідей на задані запитання, у той час як технології інтелектуальних обчислень дозволяють побачити (“знайти”) приховані правила і закономірності у наборах даних, які користувач не може передбачити, і застосування яких може сприяти збільшенню прибутків підприємства [3].

Людський розум сам по собі не пристосований для сприйняття великих масивів різномірної інформації. Людина до того ж не здатна відслідкувати більше двох-трьох взаємозв’язків навіть у невеликих вибірках. Традиційна математична статистика, яка тривалий час претендувала на роль основного інструмента аналізу даних, часто не є ефективною для розв’язання задач з реального складного життя, оскільки вона оперує усередненими характеристиками вибірки, що можуть виявитися фіктивними величинами [1].

Більшість організацій нагромаджують під час своєї діяльності величезні обсяги даних, але єдине, що вони хочуть від них отримати – це корисну інформацію. Як можна довідатися з даних, що є вигідним для підприємства, як розмістити ресурси ефективніше або як мінімізувати втрати? Для вирішення таких проблем призначені новітні технології інтелектуального аналізу. Вони використовують складний статистичний аналіз і моделювання для знаходження моделей і відношень, що приховані у базі даних – таких моделей, що не можуть бути знайдені звичайними методами.

Модель, як і карта – це абстрактне зображення реальності, але доки модель не відповідає існуючим реальним відношенням, неможливо отримати успішні результати.

Існують два види моделей: передбачувані й описові. Перші використовують один набір даних з відомими результатами для побудови моделей, що явно передбачають результати для інших наборів, інші описують залежності в існуючих даних, що своєю чергою використовуються для прийняття вирішальних дій.

Звичайно ж компанія, що довго знаходиться на ринку і знає своїх клієнтів, вже інформована про багато моделей, які спостерігалися протягом декількох останніх періодів. Технології інтелектуального аналізу можуть не лише підтвердити ці емпіричні спостереження, але і знайти нові моделі. Спочатку це може дати невелику перевагу, але якщо зробити аналіз по кожному товару

і кожному клієнту, можна мати істотний відрив від конкурентів і значне покращання у фінансовому та ринковому становищі компанії [5].

Застосування новітніх технологій не виключає необхідності знання конкретного бізнесу, розуміння даних та аналітичних методів, що використовуються. Вони лише допомагають аналітикам у знаходженні моделей і відношень у даних, але для визначення корисності цих моделей для підприємства, їх потрібно перевіряти в реальному середовищі. Новітні технології не замінюють керівників або аналітиків, а надають їм сучасний, могутній інструмент для обробки величезних обсягів даних, оскільки, по-перше, у зв'язку з розвитком технологій записування і збереження даних сьогодні на людей вилились колосальні потоки інформаційної руди у всіляких галузях, що без продуктивної переробки можуть перетворитися в нікому не потрібне сміття. І, по-друге, засоби і методи обробки даних стають доступнішими та зручнішими, а їхні результати зрозумілими для будь-якої людини [4].

Для успішного втілення новітніх технологій необхідною умовою є наявність сховища даних. Зазначимо основні принципи побудови сховища даних. Сховище даних – це предметно-орієнтований, інтегрований, прив'язаний до часу, незмінний збір даних для підтримки процесу прийняття керівних рішень. Предметна орієнтація означає, що дані об'єднані в категорії і зберігаються відповідно до тих галузей, що вони описують, а не до застосувань, що їх використовують. Інтегрованість означає, що дані задовольняють вимоги всього підприємства, а не окремої його гілки. Цим сховище даних гарантує, що однакові звіти, які призначені для різних аналітиків, будуть містити однакові результати. Прив'язка до часу означає, що сховище можна розглядати як сукупність “історичних” даних, можна відновити картину на будь-який момент часу. Атрибут часу завжди є явно присутнім у структурах сховища даних. Незмінність означає, що, потрапивши один раз у сховище, дані вже не змінюються на відміну від оперативних систем, де дані переважно зберігаються в останній версії, оскільки постійно змінюються. У сховище дані лише долучають.

Для організації й експлуатації інформаційного сховища необхідним є спеціалізоване програмне забезпечення, оскільки сучасні засоби адміністрування сховища даних мають забезпечити ефективну взаємодію з інструментарієм інтелектуальних технологій.

Ключовою можливістю просування новітніх технологій стало стрімке падіння ціни за останнє десятиліття на пристрої збереження інформації, що істотно здешевіло і збільшило можливості збору і збереження великих обсягів інформації.

Падіння цін на процесори з одночасним збільшенням їхньої швидкодії сприяло розвитку технологій, пов'язаних з обробкою величезних масивів інформації. Внаслідок цього подолано більшість бар'єрів, що знаходяться на шляху знаходження нового знання в сховищах даних.

Клієнт-серверна архітектура також є необхідним атрибутом технології інтелектуального аналізу даних. Такий підхід надає можливості виконувати складні процедури обробки даних на високопродуктивному сервері як розробникам проектів, так і користувачам. На цьому ж сервері можуть зберігатися і по запитам клієнтів виконуватися корпоративні проекти.

### **Сфера застосування технологій інтелектуальних обчислень**

Сфера застосування технологій інтелектуальних обчислень є доволі широкою – вона скрізь, де є численні дані з предметної галузі [2]. Але насамперед такі технології призначені для комерційних підприємств, що розгортають проекти на основі інформаційних сховищ даних. Досвід багатьох підприємств доводить, що віддача від використання технологій інтелектуального аналізу даних є значною. Ділові люди усвідомили, що за допомогою новітніх технологій вони можуть мати істотні переваги в конкурентній боротьбі.

Коротко наведемо можливі бізнес-застосування технологій інтелектуального аналізу даних та обчислень [4, 5, 6, 7]:

Маркетинг даних – галузь, в якій активно застосовується інтелектуальні технології. Визначаючи коло кандидатів для розсилання цільової реклами, можна збільшити продаж, при тому зменшивши витрати на проведення такої реклами.

Телекомунікаційні компанії є лідерами в застосуванні цих технологій для визначення заходів з залучення нових та втримання існуючих клієнтів.

Страхові компанії застосовують ці технології для визначення можливих втрат клієнтів.

Компанії, що діють на фінансовому ринку, визначають ринкові і галузеві характеристики для передбачення індивідуальних і фондових переваг у найближчому майбутньому.

Супермаркети визначають, який товар продавати і як його розташувати всередині магазину для досягнення найбільшої кількості продажів.

Медичні заклади – можна визначити ефективність застосування медикаментів, хірургічних процедур і медичних тестів.

Фармацевтичні фірми використовують сховища даних з хімічних сполук для знаходження комбінацій з'єднань, які надалі можна використати як ліки для лікування різних захворювань.

Ключем до успішного застосування методів інтелектуальних обчислень є не лише вибір доцільного алгоритму, а також фаховість людини, що проводить побудову моделі та супроводжує процес моделювання. Можна виділити два основних фактори, що сприяють отриманню ефективних результатів. По-перше, це чітке і повне формулювання задачі, що підлягає розв'язанню. По-друге, це використання правильних даних, які можуть постачатися з різноманітних джерел і які вимагають перетворення та групування у визначеному порядку.

Чим більше аналітик буде експериментувати з даними, будувати моделі, оцінювати результати, тим кращим може бути результат. Робота з даними є ефективнішою, коли застосовано інтеграцію таких компонентів: візуалізацію, графічний інструментарій, засоби формування запитів, оперативну аналітичну обробку, що надає змогу зрозуміти дані й інтерпретувати результати, а також різноманітні алгоритми, які застосовують для побудови моделі.

### **Технології інтелектуальних обчислень та український ринок**

На українському ринку технології інтелектуальних обчислень роблять лише перші кроки [4]. Це можна пояснити їх високою вартістю, але, як показує історія розвитку інших галузей комп'ютерного ринку України, сам по собі цей фактор навряд чи є визначальним. Скоріше тут спостерігається дія специфічних для України негативних факторів, що різко зменшують ефективність застосування інтелектуальних технологій.

Спробуємо визначити ці фактори, проаналізувати різні класи систем інтелектуального аналізу, та зазначити їх властивості, що полегшують їхнє застосування українськими підприємствами.

Почнемо з характеристики української специфіки. Комп'ютерні системи підтримки прийняття рішень, переважно, можуть ґрунтуватися на двох підходах. Перший, традиційний, полягає в тому, що в системі фіксується досвід експерта, який використовується для вироблення, оптимального в цій ситуації, рішення. Системи інтелектуальних обчислень в основному реалізують інший підхід. Вони намагаються знайти рішення на основі аналізу історичних даних, що описують поведінку досліджуваного об'єкта, прийняті в минулому рішення, їхні результати тощо. Усі ці дані можуть містити, наприклад, часові ряди цін на різні фінансові послуги, результати фінансово-господарської діяльності підприємства, статистику продажів певних видів продукції. Зрозуміло, щоб застосування цих систем у практиці виявилось виправданим, необхідно мати досить вагому множину історичних даних – інакше прийняті на їхній основі рішення будуть безпідставними.

З цієї очевидною обставиною пов'язані головні труднощі просування технологій інтелектуальних обчислень в Україні: відмінною рисою більшості вітчизняних підприємств є порівняно невеликий термін існування. Характерний "вік" нагромаджених ними історичних даних становить 3–4 роки, і, зазвичай, інформації, що міститься в цих базах даних, виявляється недостатньо для вироблення на її основі ефективної стратегії прийняття рішень за допомогою новітніх інтелектуальних систем. Складності тут є не лише в неможливості виявлення значущих взаємозв'язків у нечисленних даних та побудови моделей на їхній основі, а і в отриманні статистично незначущих моделей та прийнятті на їхній основі неправильних рішень. Якщо даних замало, а їхня описова модель є складною, то завжди можна підігнати отриману модель під існуючі дані. Той факт, що метод відмінно працює, коли потрібно пояснити, що було в минулому, але зовсім непридатний для прийняття рішень "на майбутнє", надає сумнівів у здатності систем інтелектуальних обчислень розв'язувати реальні задачі зі сфери бізнесу і фінансів. Отже, головна проблема застосування систем видобутку знань для України – це нечисленність даних, що аналізуються, а також відсутність твердого контролю статистичної значущості отриманих результатів.

Іншою відмінною рисою української економіки, як на макрорівні, так і на рівні окремих підприємств, є її нестабільність, окрім того, вона знаходиться під впливом дії численних, здебільшого випадкових факторів. На відміну від країн Заходу, де підприємства в основному працюють у межах вже ustalеної законодавчої бази, маючи сформовані структури товарних, фінансових і інформаційних потоків, українські підприємства змушені підлаштовуватися під правила гри, що постійно змінюються. Це стосується і українських фінансових ринків, де приблизно раз у півроку відбувається істотне корегування правил роботи. Отже, людина повинна обов'язково контролювати й аналізувати результати, які продукують інтелектуальні системи, тобто побудовані моделі повинні бути прозорими та надаватися до інтерпретації, щоб гарантувати врахування всіх факторів, що впливають на рішення.

Нарешті, ще одна обставина впливає на застосування систем інтелектуальних обчислень в українських умовах. Воно пов'язано з тим, що люди, котрі відповідають за прийняття рішень у бізнесі і фінансах, переважно не є фахівцями зі статистики та штучного інтелекту і тому не можуть безпосередньо використовувати системи інтелектуального аналізу даних, що вимагають складного налаштування та попередньої обробки даних. Якщо таку систему постачають як складову частину загальної технології електронних сховищ даних, які реалізовані на підприємстві (це є розповсюдженою практикою в розвинутих країнах), то це не створює проблеми – всі налаштування і попередня обробка здійснюються автоматично. Однак українські підприємства, що використовують сховища даних для подальшого застосування технологій інтелектуального аналізу, сьогодні вкрай нечисленні. Тому важливими факторами, що визначають комерційний успіх новітніх технологій в Україні, є простота у використанні і високому ступені автоматизму.

**Висновки.** Засоби знаходження нового знання та аналізу історичних даних надають допомогу підприємствам у знаходженні прихованих залежностей в існуючих історичних даних. Отримані моделі можна використовувати як для передбачення майбутніх значень, так і для описання поточного стану. Однак засоби інтелектуальних обчислень не можуть працювати без супроводження фахівців, що добре розуміються у предметній галузі, історичних даних і загальному характері застосованих алгоритмів. Результати застосування технологій інтелектуальних обчислень мають широкий спектр, від збільшення доходів, до зменшення витрат.

Сучасні підприємці усвідомлюють, що засоби інтелектуальних обчислень – це реальний спосіб підвищення ефективності роботи підприємства. Питання не в тому, чи потрібні нові технології, а в тому, як їх застосувати у кожному конкретному випадку. Витрати на постановку задачі і супроводження інтелектуальних систем можуть на порядок перевищувати вартість окремого пакета програм. Очевидно, що варто витратити частину грошей на навчання фахівців – у підсумку це буде дешевше й ефективніше. Зростає важливість спеціалізованих консалтингових фірм, що здійснюють комплексне просування проектів, що містить формулювання задачі, аналіз методів розв'язання, вироблення рекомендацій щодо реалізації обраного алгоритму, супроводу, оптимізацію.

1. Широчин В.П. *Архитектоника мышления и нейроинтеллект*. Под ред. Ковтанюка Ю.С. – К.: Юниор, 2004. – 560 с.М. 2. Тим Джонс. *Программирование искусственного интеллекта в приложениях*. Пер. с англ. Осипова А.И. – М.: ДМК Пресс, 2006. – 312 с. 3. Юрчак І.Ю. Сучасні напрямки розвитку нейрокомп'ютерних технологій // *Вісн. Нац. ун-ту "Львівська політехніка"*. Комп'ютерні системи проектування. Теорія і практика. – № 522, 2004. – С. 21–23. 4. Компания *Forecsys: прогнозирование и анализ данных* (<http://www.forecsys.ru>). 5. Компания *NeuroProject: научные исследования и разработка программного обеспечения в области интеллектуального анализа данных* (<http://www.neuroproject.ru>). 6. Лаборатория *BaseGroup: технологии анализа данных* (<http://www.basegroup.ru>). 7. Фирма *IC: Интеллектуальный анализ данных – Управление торговлей* (<http://www.abbyu.ua>)