

ТЕХНОЛОГІЇ ІНТЕЛЕКТУАЛЬНИХ ОБЧИСЛЕНЬ

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. Величезні скупчення даних утворюють хаотичні нагромадження, які без спеціальної обробки є непотрібними, більш того, ускладнюють процес пошуку дійсно необхідної інформації. Сучасні вимоги до такої обробки наступні: дані мають необмежений обсяг і є різномірними, результати повинні бути конкретні та зрозумілі, інструменти для їхньої обробки мають бути прості у використанні. Одним із методів такої обробки є інтелектуальний аналіз даних.

Інтелектуальний аналіз даних (Data Mining) це сучасна концепція аналізу даних, яка припускає, що дані можуть бути неточними, неповними (містити пропуски), суперечливими, різномірними, непрямими, і при цьому мати гігантські обсяги. Тому розуміння даних в конкретних програмах вимагає значних інтелектуальних зусиль. В інтелектуальному аналізі даних застосовується математичний апарат для виявлення закономірностей і тенденцій, що існують в даних. Основна мета інтелектуального аналізу даних – повністю автоматичне або напівавтоматичне знаходження в зібраних даних залежностей, що представляють практичну цінність в контексті сфери застосування даної технології. Методи інтелектуального аналізу даних поділяються на три групи: пошуковозалежні, прогнозування і аналіз аномалій. Пошук залежностей полягає в перегляді бази даних з метою автоматичного виявлення залежностей. Проблема тут полягає у відборі дійсно важливих залежностей з величезного числа даних існуючих в базі даних. Прогнозування передбачає, що користувач може пред'явити системі записи з незаповненими полями і запросити відсутні значення. Аналіз аномалій – це процес пошуку підозрілих даних, які сильно відхиляються від стійких залежностей.

В основу сучасної технології ІАД покладена концепція шаблонів, що відбивають фрагменти багатоаспектних взаємин у даних. Ці шаблони є закономірностями, властивими підвибіркам даних, які можуть бути компактно виражені у зрозумілій формі. Пошук шаблонів здійснюється методами, що не обмежені рамками апріорних припущень про структуру вибірки та вид розподілів значень аналізованих показників.

Інтелектуальний аналіз даних є не тривіальністю пошукових шаблонів.

Це означає, що знайдені шаблони повинні відбивати неочевидні, несподівані закономірності в даних, складові так званих прихованих знань. Сирі дані містять прошарок знань, при грамотному трактуванні якого може бути виявлена важлива інформація.

ІАД виник та набув подальшого розвитку на базі досягнень прикладної статистики, розпізнавання образів, методів штучного інтелекту, теорії баз даних тощо. Деякі з систем інтегрують у собі відразу кілька підходів. Проте, як правило, у кожній системі є якийсь ключовий компонент, на який робиться головний акцент. Нижче приводиться класифікація зазначених методів і алгоритмів ІАД.

Предметно-орієнтовані аналітичні системи – дуже розповсюджені та різноманітні. Найбільш широким підкласом таких систем, що одержав поширення в сфері дослідження фінансових ринків, є «технічний аналіз». Він є сукупністю великої кількості методів прогнозу динаміки цін і вибору оптимальної структури інвестиційного портфеля. Статистичні пакети включають елементи ІАД, але основну увагу в них звертають на класичні методики – кореляційний, регресійний, факторний аналіз тощо.

Нейронні мережі – це великий клас систем, архітектура яких має якусь аналогію з побудовою нервової тканини з нейронів. Дерева рішень є одним із найбільш популярних підходів до рішення завдань ІАД. Вони створюють ієрархічну структуру правил типу «якщото» у вигляді дерева.

Еволюційне програмування. У даній системі гіпотези про вид залежності цільової змінної від інших змінних формулюються у вигляді програм на деякій внутрішній мові програмування.

Генетичні алгоритми. Цей підхід подібний до еволюційного програмування. Його потрібно розглядати як потужний засіб рішення різноманітних комбінаторних завдань і задач оптимізації.

Сфера застосування технологій інтелектуальних обчислень є доволі широкою – вона скрізь, де є численні дані з предметної галузі. 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.