

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Реферат на тему:

Системи інтелектуального аналізу даних

Підготував:
Студент групи
Факультету
Перевірив:

– 2022

Зміст

Вступ.....	3
Розділ 1. Інтелектуальний аналіз даних	4
Історія та актуальність тематики.....	4
Основні поняття	6
Data mining.....	10
Деякі важливі задачі інтелектуального аналізу	13
Відбір неінформативних ознак.....	13
Критерії оцінювання інформативності ознак	13
Асоціативні правила	14
Розділ 2. Системи інтелектуального аналізу даних	15
Поняття системи інтелектуального аналізу даних	15
Архітектура та функціональні можливості ВІ.....	17
Приклади ВІ-систем.....	20
Висновки.....	21
Використана література	22

Вступ

Сучасний математичний апарат, що використовується для звичного аналізу даних, стає все складнішим в прикладній реалізації із-за збільшення об'ємів даних, з якими необхідно працювати, та ускладнення самого математичного апарату. В результаті, ми отримали цілі програмні комплекси, орієнтовані безпосередньо на роботу з даними. Однак, попри автоматизацію багатьох процесів все рівно виникають специфічні ситуації, які вимагають, як мінімум, знання стандартних канонів, з якими доводиться зустрічатися при аналізі даних.

Метою роботи є розглянути основні засади інтелектуального аналізу даних та загальну ідею систем інтелектуального аналізу даних, які ці засади реалізують.

Предметом дослідження є розгляд основних концепцій, що реалізуються в ВІ-системах.

Об'єктом дослідження є механізм комп'ютерної реалізації концепцій, закладених в сучасний інтелектуальний аналіз даних.

Розділ 1. Інтелектуальний аналіз даних

Історія та актуальність тематики

Математичні методи аналізу даних широко використовують при дослідженні різноманітних систем і процесів – природних, технічних, екологічних, економічних, соціальних тощо. З огляду на це формування відповідних знань та навичок є необхідною складовою підготовки фахівців у галузі системних наук і кібернетики, інформатики та багатьох інших галузей знань.

Про застосування статистичних методів аналізу даних вперше згадується у Книзі чисел. Основи сучасних методів аналізу даних були закладені Томасом Байєсом (байєсівський підхід, байєсівські оцінки), Даніїлом Бернуллі (застосування нормального розподілу в теорії похибок, перші таблиці нормального розподілу, поділ похибок спостережень на випадковій систематичні тощо), Карлом Гаусом (метод найменших квадратів); Андрієм Миколайовичем Колмогоровим (статистичні методи контролю за якістю, статистика Колмогорова – Смирнова, узагальнена відстань Колмогорова тощо), Адрієном Марі Лежандром (метод найменших квадратів), Вільфредо Парето (розподіл Парето, діаграма Парето), Френсисом Гальтоном (теорія кореляції), Карлом Пірсоном (теорія кореляції, критерії згоди, метод головних компонент), Чарльзом Спірменом (техніка факторного аналізу, рангова кореляція), Рональдом Фішером (метод максимальної правдоподібності, критерії згоди тощо). Помітний внесок у розвиток цих методів зробив видатний український математик Михайло Васильович Остроградський, який у середині XIX ст. сформулював основні ідеї статистичного контролю за якістю виробництва.

Сучасні методи аналізу даних були розвинені у працях Ю.П. Адлера, С.А. Айвазяна, Т. Андерсона, Й. Барда, Л.М. Большева, Б.В. Гнеденко, Н. Дрейпера, А.М. Дуброва, К. Іберли, І.А. Ібрагімова, А.Г. Івахненка, Дж. Кіфера, К.Х. Крамера, М. Кендалла, Г. Куллдорфа, Б.Ю. Лемешка, Ю.В. Лінника, Г.В. Мартинова, В.В. Налімова, М.С. Нікуліна, О.І. Орлова, І.М.

Парасюка, Е. Пітмена, Ю.В. Прохорова, Е. Пятецького-Шапиро, С.Р. Рао, Г. Сміта, А. Стюарта, Дж. Тьюкі, Г. Хоттелінга, П. Хьюбера, А. Хьютсона, О.О. Чупрова, Д.У. Юла та багатьох інших дослідників. Останнім часом значного поширення набувають нові технології й методи аналізу даних, зокрема методи інтелектуального аналізу даних (data mining), які використовують для виявлення прихованих закономірностей у великих масивах даних, та нейроінформатики, а також методики й засоби статистичного контролю за якістю на виробництві та в управлінні організаціями. Основні процедури аналізу даних найчастіше реалізують за допомогою сучасних комп'ютерних технологій. При цьому дослідники або самі будують розрахункові алгоритми й пишуть відповідні комп'ютерні програми, або використовують наявне програмне забезпечення – електронні таблиці MS Excel, спеціалізовані пакети SPSS, STATISTICA, математичні пакети загального призначення MatLab, MathCad тощо. Але навіть при застосуванні спеціалізованих пакетів досліднику необхідно володіти теоретичними основами математичних методів аналізу даних, оскільки зазвичай це передбачає необхідність вибору оптимальних алгоритмів та певних параметрів їх реалізації, іноді з декількох сотень можливих варіантів. Це зумовлює необхідність вивчення майбутніми фахівцями основних понять та алгоритмів аналізу даних.

Основні поняття

Інтелектуальний аналіз даних – науковий і технологічний напрямок, пов'язаний з обробкою інформації та виявленням в ній закономірностей і тенденцій, які можуть застосовуватися при підтримці прийняття рішень. Можливість практичної реалізації інтелектуального аналізу даних виникла завдяки розвитку і повсюдному поширенню інформаційних технологій, виникненню і розвитку методів штучного інтелекту. Великі обсяги даних, їх складність, мережева природа, динаміка і різноманітність інформації привели до вибухового зростання кількості і потужності методів і засобів інтелектуального аналізу даних.

Результат інтелектуального аналізу даних у загальному сенсі – знання щодо закономірностей і тенденцій, що повинно мати наступні властивості:

- відображати результати дослідження системи (пізнання об'єктивної реальності);
- бути виражено певним, зрозумілим людині чином (використовує загальноприйняті символи, поняття, природну мову);
- бути компактным (за формою, опису), що робить його доступним до розуміння, інтерпретації і подальшого використання.

Інтелектуальний аналіз даних найчастіше вирішує наступні чотири завдання:

1. Асоціація – виявлення залежностей між пов'язаними подіями, що вказують, що з події X слідує подія Y . Такі правила називаються асоціативними.

2. Кластеризація – це групування об'єктів (спостережень, подій) на основі даних (властивостей), що описують сутність об'єктів. Об'єкти усередині кластера повинні бути "схожими" один на одного і відрізнятися від об'єктів, що увійшли в інші кластери. Чим більше схожі об'єкти всередині кластера і чим більше відмінностей між кластерами, тим точніше кластеризація.

3. Класифікація – встановлення функціональної залежності між вхідними і дискретними вихідними змінними, що відповідають певним класам. За допомогою класифікації вирішується завдання віднесення об'єктів (спостережень, подій) до одного з заздалегідь відомих класів.

4. Регресія – встановлення функціональної залежності між вхідними і безперервними вихідними змінними. Прогнозування, зокрема, найчастіше зводиться до вирішення завдання регресії.

Технологія інтелектуального аналізу даних є міждисциплінарною областю дослідження. Вона використовує методи таких дисциплін, як теорія інформації, системи штучного інтелекту, теорія ймовірностей, математична статистика, машинне навчання.

Також до базових понять, які використовуються в інформатиці, відносяться дані, інформація і знання. Ці поняття часто використовуються як синоніми, проте між цими поняттями існують принципові відмінності.

Розглянемо таку ієрархію понять «відомості – дані – інформація – знання». Відомості є нижчим рівнем ієрархії і визначаються як частина знань, критерій істинності яких суб'єктивний у різних учасників пізнавального процесу. Під даними розуміються неупорядковані спостереження, числа, слова, звуки, зображення. Це набір дискретних, об'єктивних фактів. Термін дані походить від слова data – факт, а інформація (information) означає роз'яснення, виклад.

Дані – це сукупність відомостей, зафіксованих на певному носії в формі, придатній для постійного зберігання, передачі і обробки. Перетворення і обробка даних дозволяє отримати інформацію, тобто коли дані організовані, впорядковані, згруповані і категоризуються, вони стають інформацією.

Інформація – це сукупність даних, впорядкована з певною метою, що додає їм сенс. Інформація – це результат перетворення і аналізу даних. Відмінність інформації від даних полягає в тому, що дані – це фіксовані відомості, які зберігаються на певних носіях, а інформація з'являється в

результаті обробки даних при вирішенні конкретних завдань. Наприклад, в базах даних зберігаються різні дані, а по певному запиту система управління базою даних видає необхідну інформацію.

Знання трактується як інформація, готова до продуктивного застосування, дієва, змістовна. Знання – це зафіксована і перевірена практикою оброблена інформація, яка використовувалася і може багаторазово використовуватися для прийняття рішень. Знання являють собою сукупність оформленого досвіду, цінностей, контекстуальної інформації, експертного розуміння, що складають основу для оцінки і інтеграції нових досвіду та інформації.

Першим етапом аналізу будь-яких даних зазвичай є визначення їх типу. Основною є класифікація даних за шкалами їх вимірювання. Згідно з нею розрізняють такі типи ознак:

- Номінальні ознаки (ознаки з неупорядкованими станами, класифікаційні ознаки) – це дані, що вимірюють в номінальній шкалі (класифікаційній, шкалі найменувань). Найменування класів можуть бути виражені за допомогою чисел, але ці числа можуть використовуватися лише для відповіді на питання: належать два об'єкти до одного класу чи ні. Прикладами номінальних ознак є назви біологічних видів, назви навчальних дисциплін, кольори тощо. З погляду автоматизації аналізу даних і застосування стандартних алгоритмів доцільно обирати такі позначення класів: 0, 1, 2, ... Але з цими числами не можливо виконувати будь-які дії, крім перевірки їх рівності або нерівності.

- Порядкові ознаки (ознаки з упорядкованими станами, ординальні ознаки) – це дані, що вимірюють в порядкових шкалах. Ці дані можуть порівнюватися між собою у певному відношенні: “більше – менше”, “легше – важче”, “правіше – лівіше” тощо. Прикладами порядкових ознак є сила землетрусу, військові звання, оцінки студентів тощо. Якщо значення порядкової ознаки є числами, то вони можуть застосовуватися і для

порівняння ступеня вияву класифікаційної ознаки, але відстані між класами при цьому будуть не визначені.

- Кількісні (числові, варіаційні) ознаки – це ознаки, які вимірюють у кількісних (інтервальних, відносних, циклічних та абсолютних) шкалах вимірювань. Дії, що можуть виконуватися з числовими характеристиками даних, залежать від шкали вимірювань.

Важливими типами в даній класифікації є поділ ознак за дискретністю чи неперервністю теоретичної функції розподілу, законом розподілу тощо. Як характеристики вибірки можна використовувати точкові та інтервальні оцінки. Точковими оцінками параметрів вибірки називають такі оцінки, що визначаються одним числом. Як приклад таких оцінок можна навести середні арифметичні й медіани вибірок. При малих обсягах вибірок, а також при їх значному відхиленні від нормального закону розподілу точкові оцінки можуть істотно відхилятися від істинних значень оцінюваних параметрів. Тому поряд з ними використовують інтервальні оцінки параметрів. Інтервальні оцінки визначаються двома числами – межами інтервалу, до якого із заданою ймовірністю потрапляє оцінюваний параметр.

Наряду з вище згаданим часто вводять поняття описової статистики. Описова статистика – це набір основних статистичних (характеристик) показників емпіричної вибірки значень кількісної ознаки. Стандартні методи їх розрахунку, як правило, розроблені, виходячи із припущення, що розподіл є нормальним. Причиною цього є наявність зручного математичного апарату для обробки відповідних даних.

Data mining

Дуже важливим поняттям в інтелектуальному аналізі даних є поняття глибинного аналізу даних. Під глибинним аналізом даних (Data Mining) розуміють дослідження та виявлення комп'ютером (алгоритмами, засобами штучного інтелекту) у сирих даних прихованих структур і залежностей, які раніше не були відомі, нетривіальні, мають практичну цінність, доступні для інтерпретації людиною тощо. В контексті Data Mining розрізняють предиктивні та описові моделі.

Предиктивні моделі будуються на підставі набору даних з відомими результатами. Вони використовуються для прогнозу результатів на підставі інших наборів даних. Вимагається, щоб модель працювала максимально точно, була статистично значимою і виправданою. До них належать моделі класифікації, що описують правила або набір правил, відповідно до яких можна віднести опис будь-якого нового об'єкта до одного з класів. Такі правила будуються на підставі інформації про наявні об'єкти шляхом поділу їх на класи; моделі послідовностей, що описують функції, що дають змогу прогнозувати зміну параметрів. Вони будуються на підставі даних про зміну певного параметра за минулий період часу.

Описові (descriptive) моделі пов'язані із залежностями в наборі даних, взаємного впливу різних чинників, тобто базуються на побудові емпіричних моделей різних систем. Ключовий момент у таких моделях – легкість і прозорість для сприйняття людиною. До них належать такі види моделей:

- кластеризації – описують групи (кластери), на які можна поділити об'єкти, дані щодо яких піддаються аналізу. Грукуються об'єкти (спостереження, події) на основі даних (властивостей). Об'єкти усередині кластера мають бути подібними один до одного і відрізнятися від об'єктів, що ввійшли до складу інших кластерів;
- виключень – описують виняткові ситуації в записах, які різко відрізняються від основної множини записів;

- підсумкові (результатні) – виявлення обмежень на даних. Подібні обмеження важливі для розуміння даних масиву, тобто це нове знання, здобуте у результаті аналізу.

- асоціації – виявлення закономірностей між пов'язаними подіями.

Вибір методів Data Mining часто залежить від типу наявних даних і від того, яку інформацію потрібно дістати. До найпоширеніших методів можна віднести:

- об'єднання (association; іноді вживають термін affinity, що означає подібність, структурну близькість) — виокремлення структур, що повторюються в часовій послідовності. Цей метод визначає правила, за якими можна встановити, що один набір елементів корелює з іншим. Користуючись ним, зокрема, аналізують ринковий кошик, розробляють каталоги, здійснюють перехресний маркетинг тощо;

- аналіз часових рядів (sequence-based analysis, або sequential association), який дає змогу знаходити часові закономірності між даними (транзакціями);

- кластеризація (clustering) – групування записів, що мають однакові характеристики, наприклад за близькістю значень полів у базах даних. При цьому можуть залучатися статистичні методи або нейромережі. Кластеризація часто розглядається як перший необхідний крок для подальшого аналізу даних;

- класифікація (classification) — віднесення запису до одного із заздалегідь визначених класів;

- оцінювання (estimation);

- нечітка логіка (fuzzy logic);

- статистичні методи, що дають змогу знаходити криву, найближче розміщену до набору точок даних;

- генетичні алгоритми (genetic algorithms);

- фронтальні перетворення (fractal-based transforms);

- нейронні мережі (neural networks) — дані пропускаються через шари вузлів, «навчених» розпізнавати ті чи інші структури — використовуються для аналізу переваг і цільових ринків, а також для приваблювання замовників.

До вище перерахованих методів також можна віднести ще й візуалізацію — побудову графічного образу даних, що допомагає у процесі загального аналізу даних вбачати аномалії, структури, тренди.

Data Mining складається з кількох фаз:

- 1) виявлення закономірностей (вільний пошук);
- 2) використання виявлених закономірностей для прогнозу невідомих значень (прогностичне моделювання);
- 3) аналіз виключень — стадія призначена для виявлення і пояснення аномалій, знайдених у закономірностях.

При роботі, як правило, виділяється типовий ряд етапів:

1. Формування гіпотези;
2. Збір даних;
3. Підготовка даних (фільтрація);
4. Вибір моделі;
5. Підбір параметрів моделі і алгоритму навчання;
6. Навчання моделі (автоматичний пошук інших параметрів моделі);
7. Аналіз якості навчання, якщо незадовільний перехід на п. 5 або п. 4;
8. Аналіз виявлених закономірностей, якщо незадовільний перехід на п. 1, 4 або 5.

Деякі важливі задачі інтелектуального аналізу

Відбір неінформативних ознак

У багатьох задачах аналізу даних (наприклад, задачах розпізнавання образів чи прогнозу) для визначення значення вихідного параметра застосовуються математичні моделі. Для синтезу таких моделей використовується вхідна навчальна вибірка, що складається з досить великого набору ознак, які характеризують досліджуваний об'єкт або процес. Масиви даних великого розміру, як правило, містять надлишкові й неінформативні ознаки, які ускладнюють не тільки процес побудови та використання моделі, але й приводять до її надлишковості, що збільшує час отримання кінцевого результату. Таким чином, при вирішенні багатьох задач важливим етапом є процес редукції вхідного набору ознак.

Складність вирішення задач вибору максимально значимої комбінації ознак полягає в її комбінаторному характері. Використання повного перебору всіх можливих комбінацій при великій кількості ознак приводить до комбінаторного вибуху (у зв'язку з обчислювальною складністю даних процесів). Тому такий підхід на практиці виявляється неприйнятним, у результаті чого були розроблені методи скороченого перебору комбінацій ознак.

Критерії оцінювання інформативності ознак

При роботі з даними за допомогою систем інтелектуального аналізу даних однією із частих ситуацій є перевірка інформативності відібраних ознак. Тому для ефективного використання систем інтелектуального аналізу необхідно розібратися з тим, які є критерії оцінювання інформативності.

Критерії, що використовуються для оцінювання інформативності ознак, можуть бути класифіковані на критерії, що дозволяють оцінити індивідуальну інформативність ознак, і критерії, що оцінюють спільний вплив деякого набору ознак на вихідний параметр.

Асоціативні правила

При роботі з даними в ВІ-системі часто буває достатнім використовувати певний набір правил чи суджень, які дозволили б дещо пришвидшити процес аналізу даних. До таких суджень відносяться так звані асоціативні правила. Асоціативні правила дозволяють виявляти закономірності між пов'язаними подіями і є одним з інструментів видобування знань з масивів даних.

Розділ 2. Системи інтелектуального аналізу даних

Поняття системи інтелектуального аналізу даних

Сучасні підприємства змушені збирати й зберігати значні обсяги інформації - відомості про клієнтів та оперативні дані, що надходять в організацію з різних джерел, тощо. Обсяги і швидкість цих інформаційних потоків постійно збільшуються, тому підприємства все частіше звертаються до бізнес-аналізу як засобу, який дає змогу отримувати корисні для підприємства відомості з величезної кількості інформації, що зберігається в корпоративних базах даних.

Уперше термін Business Intelligence (BI) введено в обіг Г. Дреснером (тодішнім співробітником аналітичної компанії Gartner) у 1989 р. для визначення процесу, який включає доступ і дослідження інформації, її аналіз з метою оптимізації прийняття рішень.

BI - це інструментарій для аналізу даних, побудови звітів і запитів, які можуть допомогти користувачам обробити потужні масиви даних для того, щоб синтезувати з них значущу інформацію.

Також BI можна означити наступним чином. BI (Business Intelligence, інтелектуальний аналіз даних, бізнес-аналітика) — комп'ютерні методи і інструменти для організацій, що забезпечують переклад транзакційної ділової інформації в форму, придатну для бізнес-аналізу, а також засоби для роботи з обробленою таким чином інформацією.

Як видно, на даний момент немає однозначного трактування терміна BI. Частина визначень трактує BI як процес, інша - як результат процесу здобуття знань, тобто знань про ПрО для прийняття рішень, ще деякі - як бізнес-аналітику.

Нині BI консолідує три підходи:

- 1) процес перетворення даних в інформацію і знання про бізнес для підтримки прийняття поліпшених і неформальних рішень;
- 2) інформаційні технології (методи і засоби) збору даних, консолідації інформації і забезпечення доступу бізнес-користувачів до знань;

3) знання про бізнес, здобуті в результаті поглибленого аналізу детальних даних і консолідованої інформації.

Основа технології BI - організація розподіленого доступу кінцевих користувачів й аналіз структурованих кількісних даних та інформації про бізнес.

BI породжує ітераційний процес бізнес-користувача, що включає доступ до даних та їхній аналіз, і відповідне формування висновків, знаходження взаємозв'язків для змін на підприємстві. BI має широкий спектр користувачів на підприємстві, включаючи керівників та аналітиків.

До сучасних тенденцій BI можна віднести наступні: високий ступінь якості даних, стандартизація та консолідація інфраструктури, стратегічний підхід до інформації, подальший розвиток проблеми інтеграції корпоративних даних, навчання кінцевих користувачів, управління довідковими даними, використання Business Intelligence як робочого інструменту при здійсненні бізнес-аналізу.

BI іноді розуміють як технологію управління знаннями Knowledge Management (KM). KM забезпечує категоризацію, розвідку і семантичну обробку текстів, розширений пошук інформації тощо. Нині корпорації IBM і Microsoft реалізують стратегії інтеграції програмних засобів бізнес-аналізу й інструментів управління знаннями.

Архітектура та функціональні можливості BI

Сучасні підприємства змушені збирати й зберігати значні обсяги інформації - відомості про клієнтів та оперативні дані, що надходять в організацію з різних джерел, тощо. Обсяги і швидкість цих інформаційних потоків постійно збільшуються, тому підприємства все частіше звертаються до бізнес-аналізу як засобу, який дає змогу отримувати корисні для підприємства відомості з величезної кількості інформації, що зберігається в корпоративних базах даних.

Уперше термін Business Intelligence (BI) введено в обіг Г. Дреснером (тодішнім співробітником аналітичної компанії Gartner) у 1989 р. для визначення процесу, який включає доступ і дослідження інформації, її аналіз з метою оптимізації прийняття рішень.

BI - це інструментарій для аналізу даних, побудови звітів і запитів, які можуть допомогти користувачам обробити потужні масиви даних для того, щоб синтезувати з них значущу інформацію.

BI можна означити й по-іншому. BI – це процес прийняття рішень на основі зібраних даних. Це часто використовується менеджерами менеджерами та керівниками для реалізації дієвих рішень, у зв'язку з чим це поняття може бути взаємозамінюваним з поняттями «Business Analytics» та «Data Analytics». BI включає в себе управління продуктивністю, аналітику, прогностичне моделювання, аналіз даних тощо.

Як видно, сьогодні немає однозначного трактування терміну BI. Частина визначень трактує BI як процес, інша - як результат процесу здобуття знань, тобто знань про ПрО для прийняття рішень, ще деякі - як бізнес-аналітику.

Нині BI консолідує три підходи:

- 1) процес перетворення даних в інформацію і знання про бізнес для підтримки прийняття поліпшених і неформальних рішень;
- 2) інформаційні технології (методи і засоби) збору даних, консолідації інформації і забезпечення доступу бізнес-користувачів до знань;

3) знання про бізнес, здобуті в результаті поглибленого аналізу детальних даних і консолідованої інформації.

Основа технології ВІ - організація розподіленого доступу кінцевих користувачів й аналіз структурованих кількісних даних та інформації про бізнес.

ВІ породжує ітераційний процес бізнес-користувача, що включає доступ до даних та їхній аналіз, і відповідне формування висновків, знаходження взаємозв'язків для змін на підприємстві. ВІ має широкий спектр користувачів на підприємстві, включаючи керівників та аналітиків.

Сучасні тенденції ВІ: високий ступінь якості даних, стандартизація та консолідація інфраструктури, стратегічний підхід до інформації, подальший розвиток проблеми інтеграції корпоративних даних, навчання кінцевих користувачів, управління довідковими даними, використання Business Intelligence як робочого інструменту при здійсненні бізнес-аналізу.

Перейдемо тепер до розгляду архітектури ВІ-системи. Що слід розуміти під архітектурою ВІ-системи? Деякі джерела дають наступне означення: архітектура ВІ-системи – це термін, що використовується для опису стандартів та політик для організації даних за допомогою технік, що базуються на використанні комп'ютера, та технологій, що створюють системи інтелектуального аналізу даних для візуалізації, аналізу та звітності.

Якщо розглядати загальну схему архітектури ВІ-системи, то схематично її можна зобразити наступним чином:

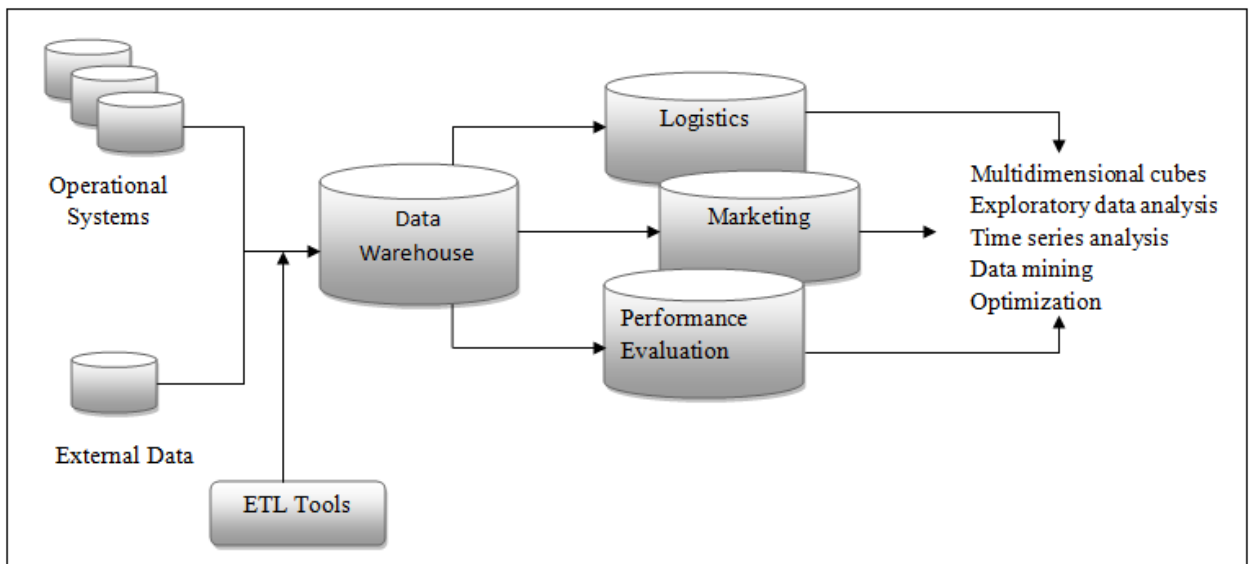


Рис. 1. Загальна архітектура ВІ-системи.

З урахуванням вище сказаного очевидно, що вказану архітектуру можна реалізовувати по-різному. Відповідно, на даний момент існує немало, переважно комерційних ВІ-систем, які забезпечують власну реалізацію як функціоналу роботи з даними, так і різний степінь реалізації таких параметрів, як безпека даних, відповідність тенденціям, тощо.

Приклади ВІ-систем

Нині серед лідерів корпоративних ВІ-платформ можна назвати IBM Cognos, BI Oracle, Microstrategy, SAP Business Objects та ін.

Згідно з дослідженнями, проведеними компанією OLAP Report, безумовним лідером світового ринку ВІ у 2006 р. стала компанія Microsoft, за нею слідувала Hyperion і Cognos. У п'ятірку лідерів увійшли Business Objects і MicroStrategy.

Зокрема, умовно можна сформувати такий список популярних ВІ-систем:

- Tableau;
- QlikView;
- SAP Business Objects;
- IBM Cognos Analytics;
- Microsoft Business Intelligence;
- Oracle Business Intelligence;
- Yellowfin;
- WebFOCUS;
- TIBCO Spotfire;
- SAS;
- Targit;
- Izenda Embedded BI & Analytics;
- MicroStrategy;
- Board та ряд інших.

Кожна із даних систем має свою специфіку та по-своєму справляється з тими чи іншими задачами. В результаті цього, їх застосування та популярність досить сильно варіюється.

Висновки

В роботі розглянуто основні поняття та засади, закладені в такий напрям, як інтелектуальний аналіз даних. Було детально розібрано основні поняття та ідеї, які є фундаментальними в даному напрямку діяльності.

Також в роботі досліджено загальні питання щодо так званих ВІ-систем, а саме чим вони являються, основні їх функції, тенденції розвитку, архітектуру та найпопулярніші приклади ВІ-систем.

Використана література

1. IT_Enterprise. Business Intelligence, BI [Електронний ресурс] / IT_Enterprise – Режим доступу до ресурсу: <https://www.it.ua/knowledge-base/technology-innovation/business-intelligence-bi>.
2. Інтелектуальні системи аналізу і звітності на підприємствах. Приклади використання інтелектуальних ВІ-платформ [Електронний ресурс] – Режим доступу до ресурсу: https://pidru4niki.com/1055120847775/informatika/intelektualni_sistemi_analizu_zvitnosti_pidpriyemstvah_prikladi_vikoristannya.
3. Ланде Д. В. Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки / Д. В. Ланде, І. Ю. Субач, Ю. Є. Бояринова. – Київ, 2018. – 297 с.
4. Бахрушин В. Є. Методи аналізу даних / В. Є. Бахрушин. – Запоріжжя, 2011. – 268 с.
5. Олійник А. О. Інтелектуальний аналіз даних / А. О. Олійник, С. О. Субботін, О. О. Олійник. – Запоріжжя, 2012. – 278 с.
6. The Role Of Data Warehousing In Your Business Intelligence Architecture [Електронний ресурс] – Режим доступу до ресурсу: <https://www.datapine.com/blog/data-warehousing-and-business-intelligence-architecture/>.