

Конспект лекції № 9

Тема № 9. ПЕРСПЕКТИВНІ НАПРЯМИ РОЗВИТКУ СППР

Міжпредметні зв'язки: Зв'язок із елементами знань і умінь таких навчальних дисциплін як „Інформатика”, «Економічна кібернетика», «Інформаційні системи і технології в управлінні», «Управління проектами інформатизації», «Захист інформації в інформаційних системах».

Мета лекції: розкрити основні положення та зміст понять теми; розглянути основні напрями розвитку СППР, пов'язані із розвитком інформаційних технологій.

План лекції

1. Поняття штучного інтелекту. Знання та їх використання в СППР.
2. Орієнтована на знання СППР, її спрощена схема.
3. Експертна система як найпоширеніша орієнтована на знання СППР.
4. Засоби інтелектуального аналізу даних у СППР - дейтамайнінг (Data Mining).
5. Можливості інтелектуального аналізу. Недоліки технології інтелектуального аналізу даних.
6. Приклади систем дейтамайнінгу.
7. Типи процесів дейтамайнінгу.
8. Прийняття рішень з використанням аналітичних алгоритмів Data mining.

Опорні поняття: штучний інтелект, дейтамайнінг, інтелектуальний аналіз даних, нейромережа, система навчання, знання, добування даних, аналітичні алгоритми Data mining.

Інформаційні джерела:

Основна та допоміжна література:

1. Баин А.М. Современные информационные технологии систем поддержки принятия решений. М.: Форум, 2009.
2. Волошин, О. Ф. Моделі та методи прийняття рішень : навч. посіб. для студ. вищ. навч. закл. /О. Ф. Волошин, С. О. Мащенко. - 2-ге вид., перероб. та допов. - К. : Видавничополіграфічний центр "Київський університет". - 2010. - 336 с.
3. Галасюк В. В. Проблемы теории принятия экономических решений / Консалт. группа "КАУПЕРВУД"; Ин-т системных исслед. интеллект. собственности. Донецк: Наука и образование, 2000. 296 с.
4. Демиденко М.А. Системи підтримки прийняття рішень : навч. посіб. / М.А. Демиденко; Нац. гірн. ун-т. -- Електрон. текст. дані. - Д. : 2016. - 104 с. - Режим доступу: <http://nmu.org.ua>

5. Нестеренко О.В. Интеллектуальні системи підтримки прийняття рішень: навч. посібн./ О.В. Нестеренко, О.І. Савенков, О.О. Фаловський. За ред. П.І. Бідюка. - Київ: Національна академія управління. - 2016. - 188 с.
6. Олексюк О.С. Системи підтримки прийняття фінансових рішень на мікрорівні. - К.: Наукова думка, 1998. - 206 с.
7. Петровский А. Б. Системы поддержки принятия решений. / Петровский А. Романов, В. П. Интеллектуальные информационные системы в экономике [Текст] : учебное пособие / Виктор Петрович Романов ; ред. Н. П. Тихомиров ; Российская эконом. академия им. Г. В. Плеханова. - М. : Экзамен, 2003. - 496 с.
8. Петруня Ю.Є. Прийняття управлінських рішень : навчальний посібник / [Ю. Є. Петруня, Б. В. Літовченко, Т. О. Пасічник та ін.] ; за ред. Ю. Є. Петруні. - [3-тє вид., переробл. і доп.]. - Дніпропетровськ: Університет митної справи та фінансів, 2015. - 209 с.
9. Системи підтримки прийняття рішень : навч. посіб. / М.А. Демиденко; Нац. гірн. ун-т. -- Електрон. текст. дані. - Д. : 2016. - 104 с.
10. Системи підтримки прийняття рішень [Текст] : навч. посібник / О. І. Пушкар, В. М. Гіковатий, О. С. Євсєєв, Л. В. Потрашкова ; ред. О. І. Пушкар. - Харків : Інжек, 2006. - 304 с.
11. Системи підтримки прийняття рішень [Текст] : навчальний посібник для самостійного вивчення дисципліни / [уклад.: С. М. Братушка, С. М. Новак, С. О. Хайлук] ; Державний вищий навчальний заклад "Українська академія банківської справи Національного банку України". - Суми : ДВНЗ "УАБС НБУ", 2010. - 265 с.
12. Системи підтримки прийняття рішень: Навч. посіб./ О.І.Пушкар, В.М.Гіковатий, О.С.Євсєєв, Л.В.Потрашкова; За ред. О.І.Пушкаря; МОН України, Харк. нац. екон. ун-т. - Х.: ВД "ІНЖЕК", 2006. - 304 с.
13. Ситник В. Ф. Системи підтримки прийняття рішень: Навч. посіб. К.: КНЕУ, 2003. ? 624 с.
14. Томашевський О. М. Інформаційні технології та моделювання бізнес-процесів : навч. посібн. / О. М. Томашевський, Г. Г. Цегелик, М. Б. Вітер, В. І. Дудук. - К. : Центр учбової літератури, 2015. - 296 с.

Інтернет ресурси:

1. ІТ для бізнеса: Системи прийняття рішень як антикризисний інструмент: [Електрон. ресурс]. - Режим доступу: <http://www.e-executive.ru/management/itforbusiness/1951354-it-dlya-biznesa-sistemy-prinyatiya-reshenii-kak-antikrizisnyi-instrument>
2. Навч.-метод. посіб. "Системи підтримки прийняття рішень": [Електрон. ресурс]. - Режим доступу: <http://megalib.info/sistemi-pidtrimki-prijnyattya-rishen/>
3. Попов А.Л. Системи підтримки прийняття рішень: Учебное пособие: [Електрон. ресурс]. - Режим доступу: http://elar.urfu.ru/bitstream/10995/1676/5/1335843_schoolbook.pdf

4. Пошуковий сервер GOOGLE: [Електрон. ресурс]. - Режим доступу: <http://www.google.com.ua>
5. Система підтримки прийняття рішень: помічник керівника для стратегічного і оперативного управління: [Електрон. ресурс]. - Режим доступу: <http://www.epam-group.ru/about/news-and-events/in-the-news/2009/sistema-podderzhki-prinyatiya-resheniy-pomoschnik-rukovoditelya-dlya-strategicheskogo-i-operativnogo-upravleniya>
6. Системи підтримки прийняття рішень в бізнесі: [Електрон. ресурс]. - Режим доступу: http://www.sib.com.ua/arhiv_2005/6_2005/systems/systems.htm

Навчальне обладнання, ТЗН, презентація тощо: ноутбук, проектор, мультимедійна презентація.

ВИКЛАД МАТЕРІАЛУ ЛЕКЦІЇ

Питання 1. Поняття штучного інтелекту. Знання та їх використання в СППР.

У сучасному розумінні термін «*штучний інтелект*» можна трактувати як науковий напрям (дисципліну), в рамках якого ставляться та розв'язуються завдання апаратного й програмного моделювання тих видів людської діяльності, які традиційно вважаються інтелектуальними, тобто потребують певних розумових зусиль. У Державному Стандарті України ДСТУ 2938-94 *штучний інтелект* визначається як «здатність систем оброблення даних виконувати функції, що асоціюються з інтелектом людини, такі як логічне мислення, навчання та самовдосконалення».

Протягом процесу розвитку штучного інтелекту отримано низку додатків, котрі застосовуються в різних галузях науки й техніки. У даний момент застосування в бізнесі штучного інтелекту, головне, проходить у формі основаних на знаннях систем (knowledge-based systems), у яких використовуються людські знання для розв'язання проблем. Найпопулярнішим типом таких систем є експертні системи. До них належать також оброблення природною мовою, системи візуалізації, роботи, нейромережі тощо. Розглянемо їх стисло.

Експертна система — це комп'ютерна програма, у якій намагаються подати знання людини-експерта у формі евристик, це різновид ІС.

Оброблення природною мовою – це найзручніший спосіб спілкування людей із комп'ютером різ-ними розмовними мовами, наприклад це можливість комп'ютерної перевірки правопису та граматики.

Системи візуалізації – це системи, які можуть подібно людині візуально взаємодіяти зі своїм середовищем, використовувати візуальні зображення та слухові сигнали для інструктування комп'ютерів або інших пристроїв, як, наприклад, роботів.

Роботи складаються з контрольованих комп'ютером пристроїв, які імітують моторну працю людини. Вони не використовуються в системах підтримки прийняття рішень.

Нейромережі – це надзвичайно спрощені моделі нервової системи людини, що можуть імітувати такі здатності людини, як навчання, узагальнення й абстрагування. Ці здатності дають моделям можливість імітувати подібні людині поведінку та режими.

Системи з навчання містять низку операцій, які надають можливість комп'ютеру або іншому зовнішньому пристрою набувати нових знань на додаток до того, що було вже введено раніше в пам'ять фірмою-виробником або програмістами.

Апаратні засоби ШІ – це фізичні пристрої, які допомагають виконувати функції в інших додатках штучного інтелекту.

Програмні агенти – програмні продукти, що виконують завдання за дорученням користувача щодо пошуку інформації в комп'ютеризованих мережах.

Знання. Поняття «знання» (Knowledge) у системах підтримки прийняття рішень, зокрема орієнтованих на знання, має виключно важливе значення. У загальному тлумаченні термін «знання» означає те, що будь-хто щось знає та розуміє. Нагромаджені знання зберігаються в **базах знань** – це низка фактів, правил і процедур, які організуються в систему за допомогою специфічних програмних засобів, що забезпечують пошук, зберігання, перетворення й занесення в пам'ять ЕОМ структурованих одиниць знань. Знання, нагромаджені й використовувані в інтелектуальних системах, можна поділити на три типи: **декларативні, процедурні та евристичні**.

Відомо чотири основні типи моделей описання знань у базах знань:

- логічні, основою яких є формальна модель, тобто формальний опис на деякій логічній мові структури об'єкта;
- продукційні, що ґрунтуються на використанні правил (продукцій);
- фреймові, тобто форми подання знань, в основу яких покладено фрейми, кожен із яких складається зі слотів;
- моделі у вигляді семантичної мережі, тобто мережі, у вершинах якої містяться інформаційні одиниці, а дуги характеризують відношення та зв'язки між ними.

Мова, що використовується для розроблення систем на основі цих моделей, називається **мовою подання знань**. Відомі такі мови подання знань:

- логічна мова подання знань, в основу якої покладено числення предикатів першого порядку;
- продукційна мова, основними одиницями якої є продукції (правила);
- фреймова мова, в котрій для подання й маніпулювання знаннями використовується фреймова модель подання знань. Нині найвідомішими фреймовими мовами є FRL і KRL.



Рис. 9.1. Дерево додатків штучного інтелекту

Висновок. Штучний інтелект - властивість автоматизованих систем брати на себе окремі функції інтелекту людини (наприклад, вибирати і приймати оптимальні рішення на основі раніше отриманого досвіду і раціонального аналізу зовнішніх впливів).

Питання 2. Орієнтована на знання СППР, її спрощена схема.

Орієнтовані на знання СППР забезпечують менеджерів відповідними рекомендаціями. Домінуючими компонентами цих систем є «здобування» знань (knowledge discovery) та механізми їх запам'ятування. Орієнтовані на знання системи підтримки прийняття рішень зберігають і застосовують різні знання для розв'язання багатьох специфічних бізнесових проблем.

У принциповому відношенні орієнтована на знання СППР має такі самі компоненти, що і звичайна СППР, тобто елементи моделювання та керування даними, користувацький інтерфейс, засоби телекомунікацій (керування поштою та повідомленнями). Проте вона містить і додаткові компоненти – базу знань і машину (механізм) висновків (рис. 9.2).



Рис. 9.2. Спрощена схема орієнтованої на знання СППР

Машина логічного висновку – це програмне забезпечення, яке імітує виконання функцій мислення. Машина логічного висновку використовує знання, які містяться в базах знань, щоб продукувати висновки.

Основою орієнтованих на знання СППР є **механізм (середовище) розроблення**, що включає програмне забезпечення для створення та підтримки бази знань і машини висновків, а також **експерт домена (проблемної галузі)**. Експертом домена є особа, яка має знання та досвід у проблемній галузі, для якої

розробляється специфічна система. Він тісно співпрацює з інженером знань, щоб відобразити знання експерта в базі знань.

Характеристики, які є загальними для орієнтованих на знання СППР:

1. Цей вид програмного забезпечення допомагає менеджерам у розв'язуванні проблем.
2. Такі системи використовують накопичені знання у вигляді правил, фреймів або ймовірної (правдоподібної) інформації.
3. Люди взаємодіють із системою, коли виконують завдання.
4. Орієнтовані на знання СППР ґрунтуються на рекомендаціях, отриманих з людських знань, і застосовуються для розв'язання невеликого кола завдань.
5. Орієнтовані на знання СППР (експертні системи також) не «думають» на відміну від людини-експерта.

Орієнтовані на знання СППР намагаються імітувати процеси міркування людини в той час, як орієнтовані на моделі СППР мають послідовність заздалегідь визначених інструкцій для реагування на події. Для порівняння орієнтованих на знання СППР і орієнтованих на моделі СППР ми маємо запам'ятати такі формули:

Орієнтовані на знання СППР = База знань + Машина логічного висновку.

Орієнтовані на моделі СППР = Дані + Кількісні моделі.

Висновок. СППР, орієнтовані на знання (Knowledge-driven DSS) забезпечують менеджерів відповідними рекомендаціями. Домінуючими їхніми компонентами є «здобування» знань (knowledge discovery) та механізми їх запам'ятовування.

Питання 3. Експертна система як найпоширеніша орієнтована на знання СППР.

Основна концепція експертних систем (ЕС) базується на припущенні, що знання експерта можуть бути записані в комп'ютерній пам'яті і потім застосовані іншими, коли з'являється в цьому потреба.

Однак експертні системи нині мають багато обмежень і недоліків. З погляду підтримки створення рішень головний їх недолік полягає в тому, що вони **не забезпечують підтримку рішень, оскільки сама експертна система створює рішення, відтворюючи логіку людини-експерта.** Перерахуємо деякі інші обмеження та недоліки експертних систем:

- експертні системи, зазвичай, працюють тільки у вузько визначених проблемних доменах, їхній рівень розуміння середовища, в якому вони функціонують, є деякою мірою поверхневим;
- ці системи до тепер не володіють здатністю «здорового глузду», як інструментальні засоби вони, зазвичай, не здатні обмірковувати проблему багатьма способами або на різних рівнях;
- експертні системи не можуть самі навчатися;
- успішні експертні системи можуть привести до реальних змін у методиці людини щодо виконання своїх завдань.

Незважаючи на ці обмеження, багато корпорацій розробили додатки експертних систем як експериментальних, так і діючих.

На рис. 9.3 наведена схема функціонування експертної системи. Модель експертної системи, складається з чотирьох головних частин:

- 1) інтерфейсу користувача, що дає можливість користувачеві взаємодіяти з експертною системою;
- 2) бази знань, яка містить нагромаджені знання щодо розв'язування специфічної проблеми;
- 3) машини логічного висновку (inference engine), яка забезпечує здатність міркування та інтерпретує зміст баз знань;
- 4) механізму (середовища) розроблення, який використовують експерт та інженер зі знань для створення експертної системи.

Охарактеризуємо стисло ці компоненти експертної системи.

Інтерфейс користувача дає змогу менеджеру вводити інструкції (команди) та інформацію (зображені на рис. 9.3 суцільною стрілкою) в експертну систему і отримувати відповідні розв'язки чи пояснення (пунктирна стрілка).

Бази знань містять як факти, що описують проблемну галузь, так і способи подання знань, які описують те, як підбираються факти для певної логічної послідовності.

Механізм або машина висновку є частиною експертної системи, яка виконує доведення (здійснює міркування) за допомогою використання змісту баз знань у специфічній послідовності (прямій, зворотній, змішаній).

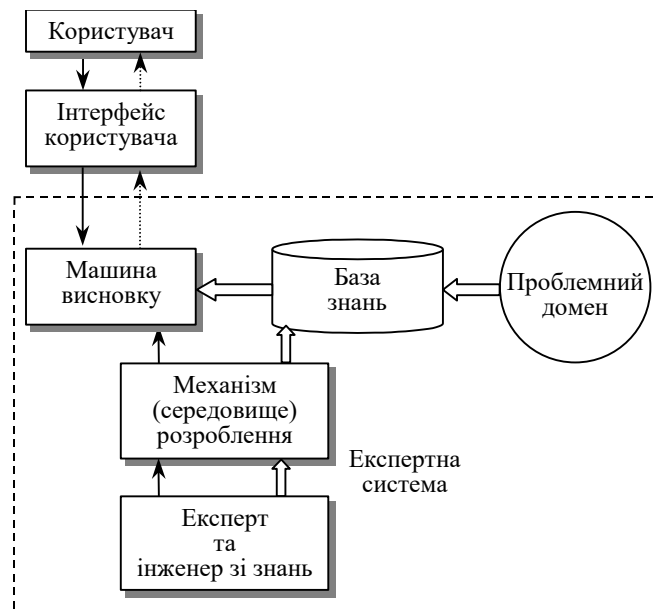


Рис. 9.3. Схема функціонування експертної системи

→ Інструкції і інформація

⋯→ Розв'язки і пояснення

⇨ Знання

Четвертою головною частиною експертної системи є механізм (середовище) розроблення, який використовується для створення експертної системи. Коли механізм висновку складається з правил, то цей процес призначений для побудови множини правил. Використовуються два основні підходи до створення ЕС: мови програмування і оболонки експертної системи.

Висновок. Експертні системи дають змогу провадити ефективну комп'ютеризацію областей, у яких знання можуть бути подані в експертній описовій формі, але використання математичних моделей утруднене або неможливе.

Питання 4. Засоби інтелектуального аналізу даних у СППР - дейтамайнінг (Data Mining).

Запровадження штрихових кодів, глобальна гіпертекстова система Інтернету тощо зробили реальною можливістю для компаній збирати великі обсяги нових даних. Однак у зв'язку з цим виникло питання про інструментальні засоби добування корисної інформації з нагромаджених обсягів «сирих» даних.

Ці засоби опісля отримали назву «Data Mining» (дейтамайнінг).

За останні роки надзвичайно зріс інтерес до дейтамайнінгу з боку ділових користувачів, котрі вирішили скористатися перевагами даної технології для отримання конкурентної переваги в бізнесі. Зростаюча зацікавленість щодо впровадження дейтамайнінгу (ДМ) у результаті закінчилася появою низки комерційних продуктів, кожен з яких має таку саму назву, описаний низкою подібних елементів, але фактично має неоднакові функціональні можливості й ґрунтується на різних особливих технічних підходах.

Очевидною перешкодою для розроблення і впровадження в корпораціях рішень з дейтамайнінгу є наявність багатьох різних підходів до нього, що мають свої певні властивості й переваги, у той час як фактично тільки кількома основними методами формуються основи більшості систем ДМ.

Дейтамайнінг (Data mining) – це тип аналітичних додатків, які підтримують рішення, розшукуючи за прихованими шаблонами (patterns) інформацію в базі даних. Цей пошук може бути зроблений або користувачем (тобто тільки за допомогою виконання запитів) або інтелектуальною програмою, яка автоматично розшукує в базах даних і знаходить важливі для користувача зразки інформації. Відповіді на інформаційні запити подаються в бажаній для користувача формі (наприклад, у вигляді діаграм, звітів тощо).

Висновок. Англomовний термін «Data mining» часто перекладається як «добування даних»; «добування знань»; «добування інформації»; «аналіз, інтерпретація і подання інформації зі сховища даних»; «вибирання інформації із масиву даних».

Важливе положення Data Mining - нетривіальність закономірностей, розшукуваних за шаблонами. Це означає, що вони повинні відображати неочевидні, несподівані (Unexpected) регулярності в даних, які становлять так звані приховані знання (Hidden Knowledge).

Питання 5. Можливості інтелектуального аналізу. Недоліки технології інтелектуального аналізу даних.

Добування даних – це процес фільтрування великих обсягів даних для того, щоб підбирати відповідну до контексту задачі інформацію. Корпорація ІВМ визначає ДМ, як «процес екстракції з великих баз даних заздалегідь невідомої, важливої інформації, що дає підстави для дій та використання її для розроблення критичних бізнесових рішень».

Інструментальні засоби добування даних **використовують різноманітні методи**, включаючи доказову аргументацію (case-based reasoning), візуалізацію даних, нечіткі запити й аналіз, нейромережі та інші. Доказову аргументацію (міркування за прецедентами) застосовують для пошуку записів, подібних до якогось певного запису чи низки записів. За допомогою візуалізації даних можна легко і швидко оглядати графічні відображення інформації в різних аспектах.

Дейтамайнінг як процес виявлення в загальних масивах даних раніше невідомих, нетривіальних, практично корисних і доступних для інтерпретації знань, необхідних для прийняття рішень у різних галузях людської діяльності, практично має нічим не обмежені сфери застосування.

Дейтамайнінг являє собою велику цінність для керівників і аналітиків у їх повсякденній діяльності. Ділові люди усвідомили, що за допомогою методів ДМ вони можуть отримати відчутні переваги в конкурентній боротьбі.

Технології дейтамайнінгу використовують велику кількість методів, частина з яких запозичена з інструментарію штучного інтелекту, а інша належить або до класичних статистичних методів, або до інноваційних методів, породжених останніми досягненнями інформаційних технологій. Вищий рівень класифікації методів дейтамайнінгу може ґрунтуватися на тому, чи зберігаються дані після дейтамайнінгу незмінними, чи вони фільтруються для подальшого використання.

На рис. 9.4 показано дерево методів дейтамайнінгу, де відображені основні види і підвиди методів, причому гілкування можна продовжити, оскільки низка методів, кластерний аналіз, нейромережі, дерева рішень мають багато різновидів.



Рис. 9.4. Дерево методів дейтамайнінгу

Висновок. Дейтамайнінг (Data mining) — це тип аналітичних додатків, які підтримують рішення, розшукуючи за прихованими шаблонами (patterns) інформацію в базі даних. Цей пошук може бути зроблений або користувачем (тобто тільки за допомогою виконання запитів) або інтелектуальною програмою, яка автоматично розшукує в базах даних і знаходить важливі для користувача зразки інформації.

Питання 6. Приклади систем дейтамайнінгу.

PolyAnalyst

Система PolyAnalyst призначена для автоматичного і напівавтоматичного аналізу числових баз даних і добування із загальних масивів даних практично корисних знань. PolyAnalyst відшукує багатofакторні залежності між змінними в базі даних, автоматично будує і тестує багатовимірні нелінійні моделі, що виражають знайдену залежність, виводить класифікаційні правила на навчальних прикладах, знаходить у даних багатовимірні кластери та будує алгоритми прийняття рішень.

Основу PolyAnalyst утворюють так звані машини досліджень (Exploration engines), тобто програмні модулі, що ґрунтуються на різних алгоритмах ДМ і призначені для автоматичного аналізу даних. Остання версія PolyAnalyst 4.3 містить 11 машин досліджень.

MineSet – візуальний інструмент аналітика

Система MineSet являє собою інструментарій для поглибленого інтелектуального аналізу даних на базі використання потужної візуальної парадигми. Характерною особливістю MineSet є комплексний підхід, що адаптує застосування не однієї, а кількох взаємодоповнюючих стратегій добування, аналізу й інтерпретації даних.

KnowledgeSTUDIO

KnowledgeSTUDIO є новою версією дейтамайнінгу корпорації з програмного забезпечення «ANGOSS». Система впроваджує найрозвинутіші методи ДМ у корпоративне середовище з тим, щоб підприємства могли досягати максимальної вигоди від своїх інвестицій у дані.

KnowledgeSTUDIO сумісна з основними статистичними пакетами програм. Система KnowledgeSTUDIO тісно інтегрується зі сховищами і вітринами даних.

У KnowledgeSTUDIO реалізована велика кількість методів дейтамайнінгу. Пропонується п'ять алгоритмів дерев рішень, три алгоритми нейромереж і алгоритм кластеризації «неконтрольованого навчання».

Висновок. Data Mining є мультидисциплінарної областю, виникла і розвивається на базі досягнень прикладної статистики, розпізнавання образів, методів штучного інтелекту, теорії баз даних та ін. Звідси велика кількість методів і алгоритмів, реалізованих у різних діючих системах Data Mining

Питання 7. Типи процесів дейтамайнінгу

Здебільшого термін «дейтамайнінг» використовується для описання автоматизованого процесу аналізу даних, в якому система сама бере ініціативу щодо генерування взірців, тобто дейтамайнінг належить до інструментальних засобів дослідницького аналізу.

З погляду орієнтації на процес є три типи процесів дейтамайнінгу (рис. 9.5): відкриття (добування) (discovery,); моделювання передбачень (predictive modeling,); аналіз аномалій (forensic analysis).

Відкриття є процесом перегляду бази даних для знаходження невидимих взірців (pattern) без наперед визначеної ідеї або гіпотези взагалі про те, що вони можуть бути. Інакше кажучи, програма бере ініціативу без попередніх міркувань стосовно того, що взірці, які цікавлять користувачів, мають насправді місце і можуть подаватися у формі доречних запитів.

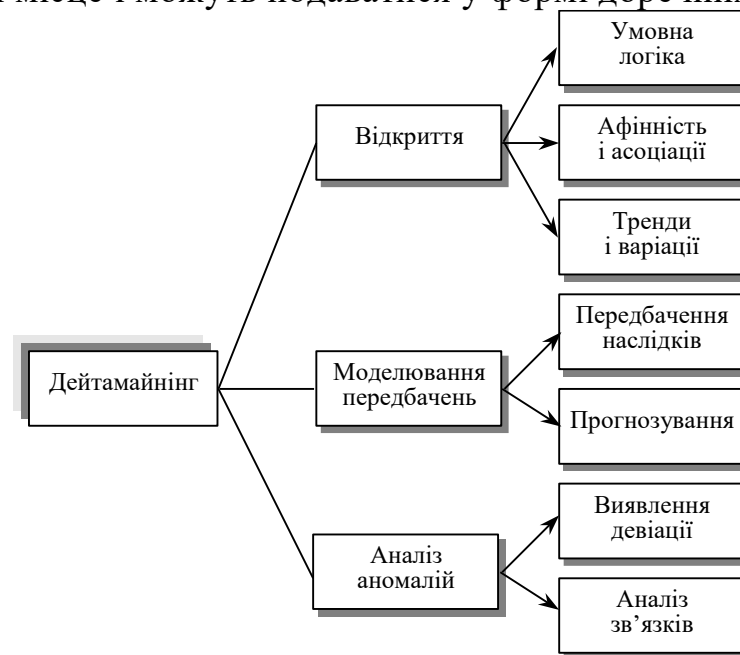


Рис. 9.5. Типи процесів дейтамайнінгу

У разі **моделювання передбачень** добуваються взірці з бази даних для їх використання, щоб передбачити майбутнє. Моделювання передбачень дає змогу користувачеві створювати записи з деякими невідомими дослідницькими значеннями, і система визначає ці невідомі значення, які ґрунтуються на попередніх шаблонах, що відкриваються з бази даних.

Аналіз аномалій (forensic analysis) є процесом застосування вибраних взірців (шаблонів) для виявлення аномалій або незвичайних елементів даних. Щоб виокремити незвичайні елементи, спершу потрібно знайти те, що є нормою, а вже потім виявляти за допомогою заданих порогових величин ті елементи, які відхиляються від звичайних.

Користувачі і дії дейтамайнінгу

Дії дейтамайнінгу, зазвичай, виконуються трьома різними типами користувачів: виконавцями, кінцевими користувачами і аналітиками. Усі

користувачі, як правило, виконують три види дії дейтамайнінгу всередині корпоративного середовища: епізодичні; стратегічні; безперервні.

Питання 8. Прийняття рішень з використанням аналітичних алгоритмів Data mining.

У більшості випадків термін “дейтамайнінг” використовується, щоб описати автоматизований процес аналізу даних, в якому система сама бере ініціативу генерувати взірці, тобто дейтамайнінг відноситься до інструментальних засобів дослідницького аналізу. З погляду орієнтації на процес є три класи активностей дейтамайнінгу: відкриття (discovery) – умовна логіка, афінність і асоціації, тренди і варіації; пророче моделювання (predictive modeling) – передбачення наслідків, прогнозування; аналіз аномалій (forensic analysis) – виявлення девітації, аналіз зв’язків.

Активності дейтамайнінгу звичайно виконуються трьома різними класами користувачів: виконавцями (executives); кінцевими користувачами (end users); аналітиками (analysts). Усі користувачі звичайно виконують три типи активностей дейтамайнінгу всередині корпоративного середовища: епізодичне; стратегічне; безперервне (постійне).

Технології дейтамайнінгу використовують велике число методів, частина з яких запозичена з інструментарію штучного інтелекту, іншу частину складають або класичні статистичні методи, або іноваційні методи, породжені останніми досягненнями інформаційної технології. Верхній рівень дихотомії технологій дейтамайнінгу може бути заснований на тому, чи зберігаються дані після дейтамайнінгу – збереження даних (Data Retention), чи вони дистилюються для подальшого використання – дистиляція шаблонів (Data Distilled).

Клас методів Data Retention включає: кластерний аналіз; метод “найближчого сусіда” (nearest neighbor), міркування за аналогією (case-based reasoning) або міркування за прецедентами (аналогічними випадками).

Клас методів Data Distilled включає: логічні методи (нечіткі запити і аналізи, правила, дерева рішень, генетичні алгоритми); візуалізацію; крос-табуляцію (агенти, довірчі мережі); методи на основі рівнянь (статистичні методи, нейромережі).

У дейтамайнінгу застосовуються також методи, які не можна віднести до жодного класу дихотомії, зокрема нелінійні регресійні методи, еволюційне програмування, алгоритми обмеженого перебору.

Особливо широко в дейтамайнінгу використовуються нейромережі, причому деякі автори вважають їх єдиними методами дейтамайнінгу – дихотомії.

Нейронна мережа (Neural Network) або просто нейромережа є програмно (інколи апаратно) реалізована система, в основу якої покладена математична модель процесу передачі й оброблення імпульсів людського мозку, що імітує механізм взаємодії нейронів (neuron) з метою опрацювання поступаючої інформації та набрання досвіду. Сьогодні програмні продукти нейромереж

пропонують десятки постачальників, зокрема HNC, Neuralware, IBM, NeoVista, TMC, SAS, ISL, Magnify та інші.

Висновок. До методів і алгоритмів інтелектуального аналізу даних відносяться наступні: штучні нейронні мережі, дерева рішень, символні правила, методи найближчого сусіда і k-найближчого сусіда, метод опорних векторів, байєсовські мережі, лінійна регресія, кореляційно-регресійний аналіз; ієрархічні методи кластерного аналізу, неієрархічні методи кластерного аналізу, в тому числі алгоритми k-середніх і k-медіани; методи пошуку асоціативних правил, у тому числі алгоритм Apriori; метод обмеженого перебору, еволюційне програмування і генетичні алгоритми, різноманітні методи візуалізації даних тощо.

Загальний висновок за темою лекції

1. СППР на основі штучного інтелекту використовує методи і алгоритми, які дозволяють комп'ютеру робити висновки і пропозиції на основі обчислень за спеціальними методами і програмами. За цією методикою комп'ютер може зробити висновок про правильність прийнятого рішення чи якогось висновку в конкретній ситуації. Комп'ютер може відстежувати і аналізувати причинно наслідкові взаємовідносини між факторами процесу прийняття рішень.
2. СППР, орієнтовані на знання (Knowledge-driven DSS) забезпечують менеджерів відповідними рекомендаціями. Домінуючими їхніми компонентами є «здобування» знань (knowledge discovery) та механізми їх запам'ятовування.
3. Експертні системи відрізняються і від інших видів програм із галузі штучного інтелекту. Експертні системи застосовуються для предметів реального світу, операції з якими зазвичай вимагають великого досвіду, накопиченого людиною. Експертні системи мають яскраво виражену практичну направленість для застосування в науковій або комерційній сфері. Однією з основних характеристик експертної системи є її швидкодія, тобто швидкість отримання результату та його достовірність (надійність). Дослідницькі програми штучного інтелекту можуть бути і не дуже швидкими, натомість, експертна система повинна за прийнятний час знайти розв'язок, що був би не гіршим за розв'язок, що може запропонувати фахівець в цій предметній області. Експертна система повинна мати можливість пояснити, чому запропоновано саме цей розв'язок і довести його обґрунтованість. Користувач повинен отримати всю інформацію, необхідну йому для того, аби переконатись в обґрунтованості запропонованого розв'язку.
4. СППР-системи - це частина "інтелектуальних ресурсів підприємства" або "засобів інтелектуального бізнес-аналізу" (Business Intelligence - BI). Важливою частиною BI-технологій є також системи інтелектуального пошуку інформації (Data Mining - DM). Data Mining - це процес виявлення в сирих даних раніше невідомих, нетривіальних, практично корисних і доступних для інтерпретації знань, необхідних для прийняття рішень у

різних сферах. Технології Data Mining представляють велику цінність для керівників і аналітиків в їх повсякденній діяльності. Ділові люди усвідомили, що за допомогою методів Data Mining вони можуть отримати відчутні переваги в конкурентній боротьбі. Data Mining є мультидисциплінарною областю, виникла і розвивається на базі досягнень прикладної статистики, розпізнавання образів, методів штучного інтелекту, теорії баз даних та ін. Звідси велика кількість методів і алгоритмів, реалізованих у різних діючих системах Data Mining.

5. До методів і алгоритмів інтелектуального аналізу даних відносяться наступні: штучні нейронні мережі, дерева рішень, символні правила, методи найближчого сусіда і k-найближчого сусіда, метод опорних векторів, байєсовські мережі, лінійна регресія, кореляційно-регресійний аналіз; ієрархічні методи кластерного аналізу, неієрархічні методи кластерного аналізу, в тому числі алгоритми k-середніх і k-медіани; методи пошуку асоціативних правил, у тому числі алгоритм Apriori; метод обмеженого перебору, еволюційне програмування і генетичні алгоритми, різноманітні методи візуалізації даних тощо.

Питання і завдання студентам для контролю знань, самостійного опрацювання матеріалу лекції, для підготовки до семінарського, практичного, лабораторного заняття за темою лекції.

1. Що таке штучний інтелект?
2. Яким чином штучний інтелект використовують в СППР?
3. Яке призначення орієнтованих на знання СППР?
4. Що таке експертна система?
5. Чим експертні системи відрізняються від інших видів систем?
6. Що таке дейтамайнінг?
7. Яке призначення дейтамайнінгу?

Укладач: _____
(підпис)

Шевчук І.Б., зав. каф., д.е.н., доцент
(ПБ, посада, науковий ступінь, вчене звання)