

- розуміння: виразити розуміння фактів і ідей шляхом організації, порівняння, перекладу, інтерпретації за допомогою опису і формулювання основних ідей;
- застосування: за допомогою отриманих знань вирішити завдання в нових ситуаціях шляхом застосування фактів, методів і правил;
- аналіз: вивчити і розбити інформацію на частини, виявляючи мотиви або причини, зробити висновки і знайти докази на підтримку узагальнення;
- синтез: побудувати структуру або шаблон з різноманітних елементів і актів задачі частини разом, щоб сформувати в цілому; зібрати інформацію разом по-іншому, комбінуючи елементи в новій моделі або пропонувати альтернативні рішення;
- оцінка: виражати і захищати думки шляхом суджень про інформацію, обґрунтованість ідей або якість роботи на основі набору критеріїв.

При формуванні моделі оцінки для визначення компетентності членів команди проекту застосовуються такі методи оцінювання, як професійне тестування, аналіз конкретних ситуацій, кейс-метод. При цьому ці методи використовуються для оцінки рівня оволодіння компетенціями «знання», «розуміння» та «застосування». Для оцінки кваліфікацій наступних трьох рівней набір методів оцінювання треба розширити за рахунок введення ділових стратегічних ігор, які показують знання, вміння і володіння навичками одночасно.

ДЖЕРЕЛА

1. Лукьянов, Д. В. Визначення ядер знань на графі компетенцій проектних менеджерів // Восточно-Европейский журнал передовых технологий. – 2012. – № 1 (10/55). – С. 26 – 28.
2. Bloom, B.S. Taxonomy of educational objectives: The classification of educational goals: Handbook I, cognitive domain. – New York: Longman, 1956.
3. Бушуев, С.Д. Управление проектами: Основы профессиональных знаний и система оценка компетентности проектных менеджеров: (National Competence Baseline, NCB UA Version 3.0) / С.Д. Бушуев, Н.С. Бушуева. – К.: ІРІДІУМ, 2006.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ІСНУЮЧИХ НАПРЯМІВ У ІНТЕЛЕКТУАЛЬНОМУ АНАЛІЗІ ДАНИХ

Щербина А.Д.¹, Шибасєв Д.С.², к.т.н., Рудніченко М.Д.,¹ к.т.н., Шибасєва Н.О.¹

¹Одеський національний політехнічний університет;

²Одеський національний морський університет

Україна, Одеса

nickolay.rud@gmail.com

У роботі наведено результати порівняльного аналізу існуючих напрямів у інтелектуальному аналізі даних, зокрема розглянуто завдання, методи та алгоритми text mining, visual mining та web mining. На базі виконаного аналізу запропоновано концепцію подальшого розвитку досліджень

Ключові слова: data mining, text mining

Вступ. Інтелектуальний аналіз даних (ІАД) - це обробка інформації і сирих даних різними методами, засобами і технологіями штучного інтелекту з метою виявлення нових знань, раніше не відомих користувачеві, які також є практично корисними, нетривіальними і доступними для інтерпретації у різних сценаріях аналізу даних[1].

В ІАД виділяють кілька окремих галузей, таких як: аналіз текстової інформації (text mining), візуальний аналіз даних (visual mining), вилучення знань з мережі Інтернету (web mining). У кожному з напрямків розглядаються свої завдання по обробці певного типу контенту, створені власні алгоритми і методи аналізу та вилучення знань.

До основних завдань аналізу текстової інформації відносять пошук кореляцій в готових текстових даних, вилучення з тексту заданих властивостей і характеристик, визначення тексту в одну із заданих категорій (класифікація), твір семантичного аналізу текстів.

Основними методами для вирішення перерахованих завдань є: класифікація, кластеризація, побудова семантичних мереж, суммаризація, тематичне індексування.

Також застосовують кілька видів алгоритмів: засновані на правилах алгоритми, в яких використовують заздалегідь підготовлені словники; алгоритми, що використовують статистику, де розраховується ймовірність появи того чи іншого слова, а також змішані алгоритми, в яких використовуються обидва підходи.

Завданнями візуального аналізу даних є представлення даних у зручній для користувача формі, особливо часто використовується в дослідженнях, коли дані зашумлені і неоднорідні, або у випадках, коли завдання і цілі досліджень остаточно не визначені.

У процесі аналізу використовуються методи геометричних перетворень: накладення вимірювань, рекурсивні шаблони, паралельні координати, матриця діаграм розкиду, орієнтовані на пікселі методи.

При цьому в даному аналізі не використовуються складні математичні обчислення і алгоритми [2]. Результати порівняльного аналізу завдань, методів і алгоритмів ІАД наведені в таблиці 1.

Таблиця 1 – Таблиця аналізу задач, методів и алгоритмів ІАД

	Text mining	Visual mining	Web mining
Задачі	Пошук кореляцій в готових текстових даних. Витяг з тексту заданих властивостей і характеристик. Віднесення тексту до однієї із заданих категорій (класифікація). Твір семантичного аналізу текстів.	Представлення даних в зручній для користувача формі.	Витяг web-контенту і web-структур. Пошук інформації. Аналіз використання та доступу до web-контенту.
Методи	Класифікація. Кластеризація. Побудова семантичних	Рекурсивні шаблони. Паралельні	Модифіковані асоціативні правила. Машинне навчання.

	мереж. Суммарізація. Тематичне індексування.	координати. Матриця діаграм розкиду. Орієнтовані на пкселі методи	Статичні методи, в тому числі і NLP. Прискорені дерева рішень. Прихована Марківська модель.
Алгоритми	Засновані на правилах алгоритми. Алгоритми, що використовують статистику. Змішані алгоритми.	Не використовуютьс я складні математичні обчислення і алгоритми	Алгоритми аналізу текстової інформації. Алгоритм argіori.

При отриманні даних з web виділяють такі завдання як витяг web-контенту і web-структур, пошук інформації, аналіз використання і доступу до web-контенту.

При цьому використовуються модифіковані асоціативні правила, машинне навчання, статичні методи, в тому числі і NLP, прискорені дерева рішень, прихована Марковська модель.

Різні алгоритми використовуються для окремих завдань, серед них для отримання web-контенту і web-структур і пошуку інформації застосовують алгоритми аналізу текстової інформації, для аналізу використання і доступу до web-контенту найчастіше застосовується алгоритм argіori, розроблений для аналізу частих наборів, але модифікованого для виявлення частих фрагментів послідовностей і переходів [3].

Найбільш актуальним є використання методів класифікації, кластеризації і змішаних алгоритмів для задач аналізу текстової інформації, матриць діаграм розкиду для візуального аналізу даних і машинного навчання при отриманні даних з web.

Висновки. Актуальним в подальшому є вивчення і детальний розгляд в рамках магістерської дисертації наведених алгоритмів і методів кожного напрямку ІАД, з метою розробки прикладного програмного забезпечення для розпізнавання текстових даних, класифікації їх за заданими критеріями і візуалізації отриманих результатів.

ДЖЕРЕЛА

1. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 2009. – 270 с.
2. Рафалович В. Data mining, или Интеллектуальный анализ. – М.: СмартБук, 2014. – 110 с.
3. Барсегян А.А., Куприянов М.С., Холод И.И. Анализ данных и процессов. — СПб.: БХВ-Петербург, 2009. — 512 с.