



Завдання Data Mining. Інформація та знання



Завдання Data Mining

- класифікація (classification)
- кластеризація (clustering)
- асоціація (associations)
- прогнозування (forecasting)
- аналіз відхилень або викидів (deviation detection)
- оцінювання (estimation)
- аналіз зв'язків (link analysis)
- візуалізація (visualization, graph mining)
- підведення підсумків (summarization)



Класифікація завдань Data Mining за стратегіями:

- навчання з вчителем (*класифікація, оцінка, прогнозування*)
- навчання без вчителя (*класифікація*)
- інші

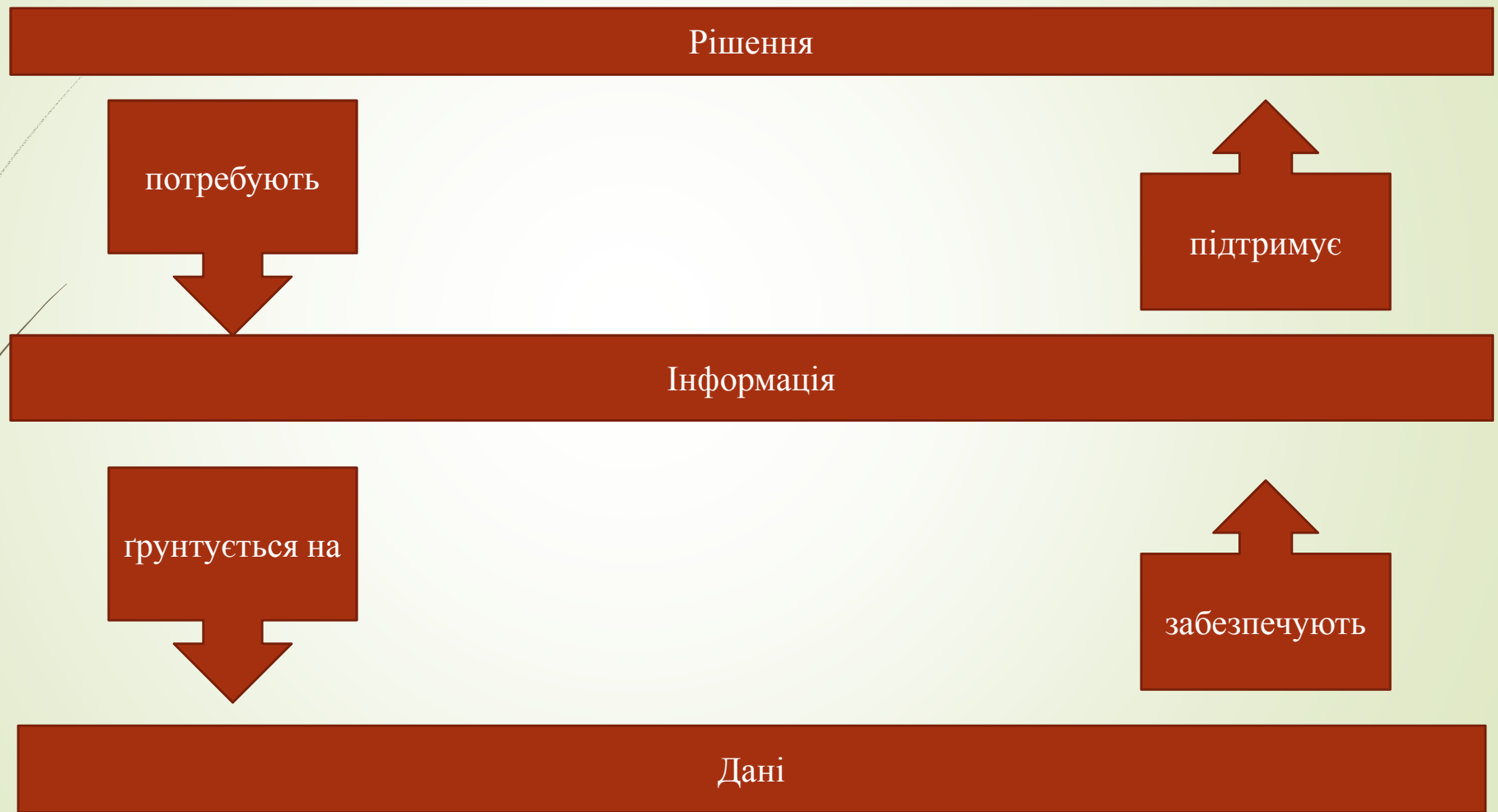
Класифікація завдань Data Mining за моделями, що використовуються:

- дескриптивні (описові)
- прогнозуючи

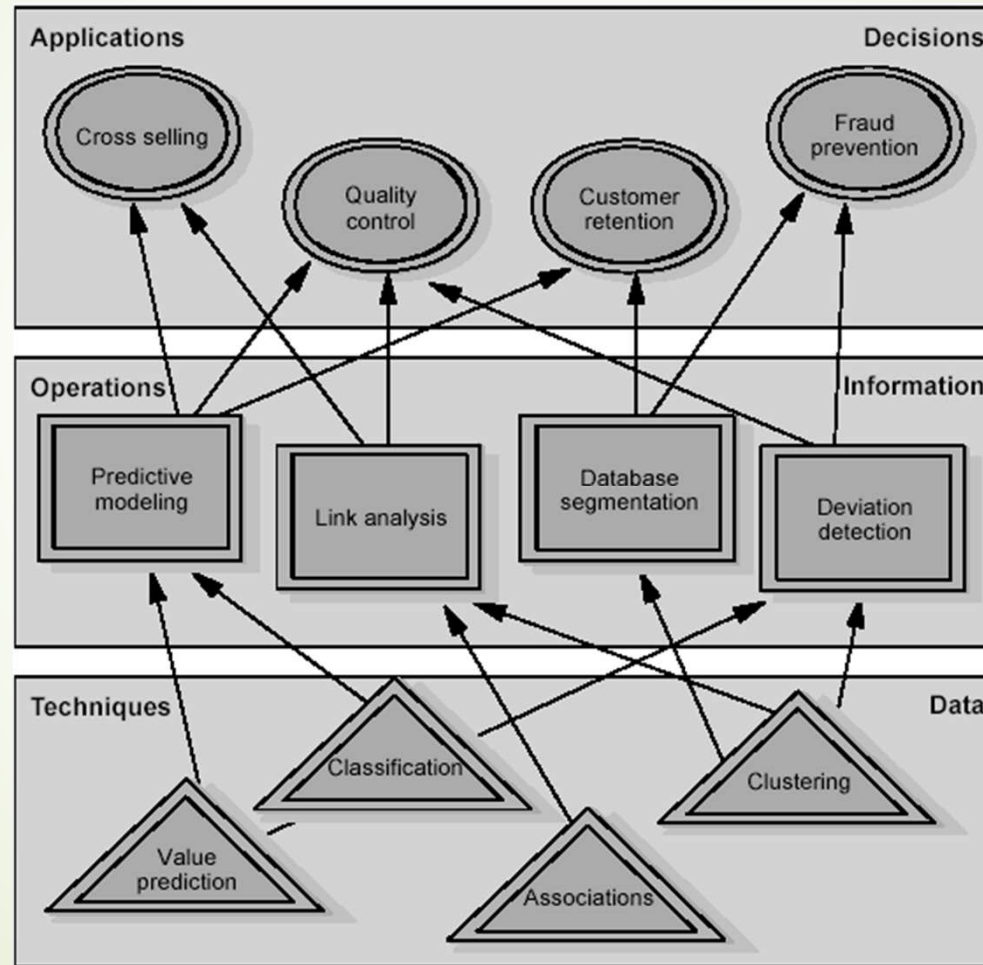
Розподіл завдань Data Mining

- ▶ автоматичне дослідження та відкриття (вільний пошук)
- ▶ прогнозування та класифікація
- ▶ пояснення та опис

Рішення, інформація, дані



Завдання, дії, додатки





Інформація

Інформація – будь-які, не відомі раніше відомості про будь-яку подію, сутність, процес тощо, що є об'єктом деяких операцій, для яких існує змістовна інтерпретація



Аналіз виключень (forensic analysis)

Дії аналізу виключень:

- виявлення відхилень (deviation detection)

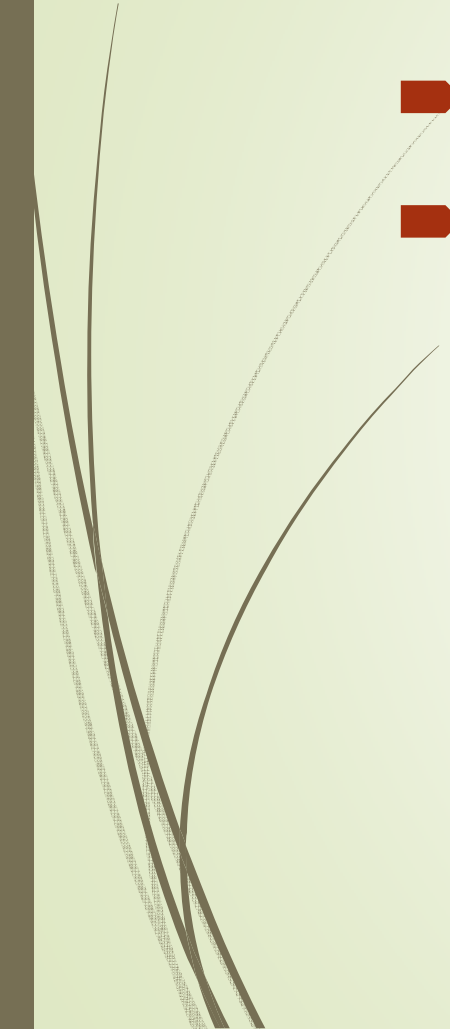


Властивості інформації:

- ▶ повнота
- ▶ достовірність
- ▶ цінність
- ▶ адекватність
- ▶ актуальність
- ▶ ясність
- ▶ доступність
- ▶ суб'єктивність



Вимоги до інформації:

- ▶ динамічний характер інформації
 - ▶ адекватність методів, що використовуються
- 



Знання

Знання – сукупність фактів, закономірностей та евристичних правил, за допомогою яких вирішується поставлене завдання



Властивості знань:

- структурованість
- зручність доступу та засвоєння
- лаконічність
- несуперечність
- процедури обробки