

Кластерний аналіз в задачах управління харчовим підприємством

Сергій Грибков¹, Олена Харкянен²

Кафедра інформаційних систем, Національний університет харчових технологій, Україна, м. Київ, вул. Володимирська, 68
E-mail: 1. sergio_nuft@ukr.net, 2. helen_nuft@ukr.net

Abstract – Discusses the use of one of data mining methods to support decision making in the management of pasta factory.

Keywords: data mining, cluster analysis, food enterprise, pasta factory.

З розвитком інформаційних технологій підприємства харчової промисловості все більше застосовують різноманітні засоби аналізу даних. Одним з підходів для класифікації багатовимірних процесів та явищ є кластерний аналіз. Використання множини обчислювальних процедур дозволяє виділити та згрупувати схожі об'єкти в порівняно однорідні групи підмножин – кластери.

Основними задачами кластеризації є: розробка класифікації, дослідження різноманітних корисних схем групування об'єктів, породження гіпотез на основі дослідження даних, перевірка гіпотез або дослідження для визначення чи дійсно групи, що виділені тим чи іншим способом присутні в наявних даних.

Задача кластерного аналізу полягає у тому, щоб на основі даних, які містяться у множині X , розбити множину об'єктів I на m кластерів (підмножин) $\pi_1, \pi_2, \dots, \pi_m$ таким чином щоб кожен об'єкт I_j належав одній і тільки одній підмножині розбиття і щоб об'єкти, які належать різним кластерам були різнорідні. Рішенням задачі кластерного аналізу є розбиття, яке задовольняє деякому критерію оптимальності [2].

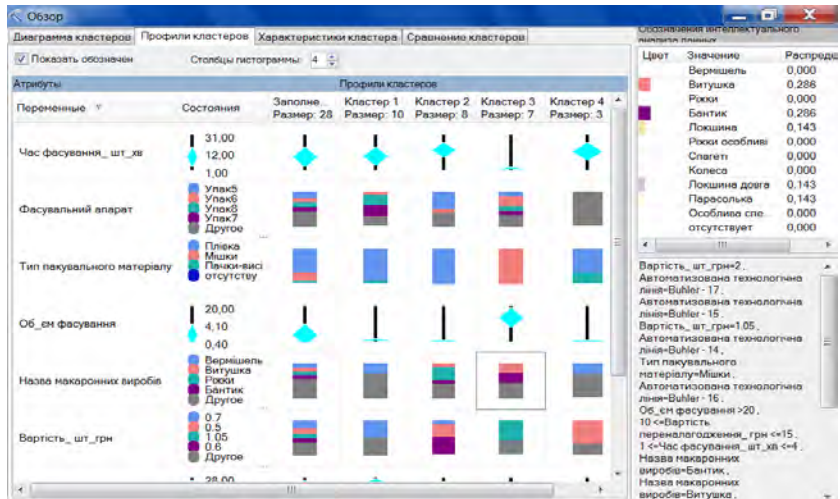
Вирішення задачі кластеризації вимагає кількісного визначення подібності або різнорідності об'єктів. Основними функціями визначення відстані між об'єктами є евклідова відстань, l_1 -норма, сюрремум-норма, l_p -норма, махаланобіса.

В практиці проведення кластерного аналізу розрізняють ієрархічні методи (метод ближнього сусіда або одиночний зв'язок, метод віддалених сусідів або повного зв'язку, метод Варда, центроїдний метод) та ітеративні методи (алгоритм k -середніх, алгоритм PAM).

Одним з лідерів харчових підприємств макаронної галузі України є ВАТ «Макаронна фабрика» (м. Київ), що виробляє понад 60 найменувань макаронних виробів високої якості на рівні світових стандартів. Для обліку господарської діяльності на підприємстві впроваджена сучасна OLTP-система, сховище даних та вітрини даних.

Пропонується доповнити існуючі інформаційні структури підприємства алгоритмами інтелектуального аналізу даних для надання

аналітикам зручних засобів підтримки при прийнятті управлінських рішень. Використавши надбудову "Клієнт інтелектуального аналізу даних", яка надає можливість застосування алгоритмів служб MS Analysis Services в середовищі MS Excel, побудована модель підтримки прийняття рішень методом кластеризації. Аналіз профілів кластерів допомагає визначити пріоритетність випуску продукції при виконанні замовлень.



Висновок

Впровадження засобів інтелектуального аналізу даних, зокрема, кластерного аналізу дозволяє на основі множини накопиченої підприємством інформації отримувати нові, корисні для ведення господарської діяльності знання використання яких забезпечить більш оперативне реагування на зміни ринково середовища.

Література

1. Барсегян А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Текст] / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. – 2-е изд., перераб. и доп. – СПб. : БХВ-Петербург, 2007. – 384 с.: ил.
2. Дюран Б. Кластерный анализ [Текст] / Б. Дюран, П. Оддел ; пер. с англ. Е. З. Демиденко ; под ред. А. Я. Боярского. – М., Статистика, 1977. – 128 с. : ил.
3. Мандель И. Д. Кластерный анализ [Текст] / И. Д. Мандель. – М.: Финансы и статистика, 1988. – 176 с.: ил.
4. Факторный, дискриминантный и кластерный анализ [Текст] / Дж.-О.Ким, Ч. У. Мьюллер, У. Р. Клекка [и др.] ; пер. с англ. ; под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.: ил.