

The algorithm was constructed on the base of the modified categorical model and successfully tested on the base of solving of practical tasks of the increase of technological processes control efficiency in public corporation „Sumykhimprom”.

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ С ПОМОЩЬЮ MS OFFICE EXCEL 2007

Кунцев С.В., УАБД,г. Сумы

В современном бизнес-анализе получила распространение новая технология интеллектуального анализа данных - Data Mining. Она представляет собой технологию поддержки процесса принятия решений, основанную на выявлении скрытых закономерностей и систематических взаимосвязей между переменными внутри больших массивов информации [1]. Накопленные сведения автоматически обобщаются. Полученную информацию можно характеризовать как знания и применять к другим массивам данных. Обнаружение новых знаний используется для повышения эффективности бизнеса.

Методами Data Mining можно выявить определенные типы закономерностей, среди которых обычно выделяют пять стандартных типов: "ассоциацию", "последовательность", "классификацию", "кластеризацию", "временные закономерности".

С помощью методов Data Mining можно успешно решать сложные задачи в экономике. В банковском деле с их помощью выявлять мошенничество с кредитными карточками, разбивать клиентов на категории, прогнозировать клиентуру. В страховании – анализировать риски. В телекоммуникационном бизнесе – проверять лояльность клиентов.

Табличный процессор MS Office Excel 2007 обладает набором активных надстроек, которые позволяют решить сложные математические задачи. Возможности процессора можно существенно увеличить, если использовать надстройки для "интеллектуального анализа данных" [2]. В этих надстройках используются аналитические средства SQL Server 2005 Analysis Services [3].

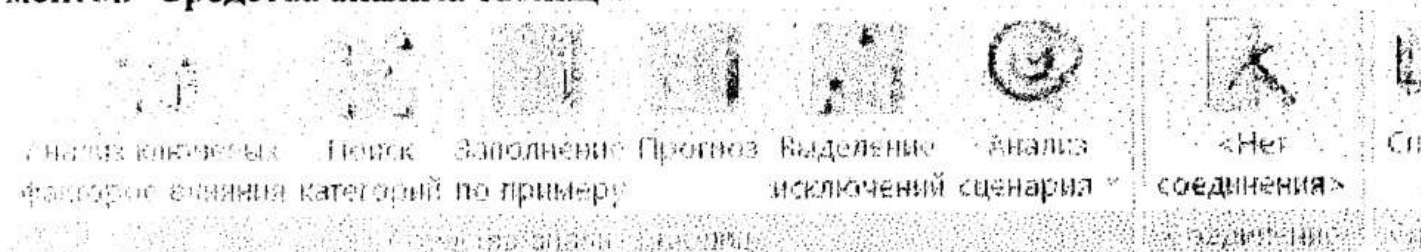
Надстройки интеллектуального анализа данных для Office 2007 помогают выявлять закономерности и тренды, существующие в сложных данных, отображать такие закономерности в диаграммах и интерактивных средствах просмотра, а затем формировать цветные сводные отчеты для презентаций и бизнес-аналитики. Можно анализировать корреляции и формировать прогнозы для данных, хранящихся в таблицах Microsoft Office Excel, или создавать и изменять модели интеллектуального анализа данных, хранящихся в экземпляре Analysis Services, а также получать результаты в графическом виде в среде Microsoft Office Visio.

Надстройки интеллектуального анализа данных для Office предоставляют мастера, которые максимально упрощают получение полезных сведений из кубов или работу с большими базами данных.

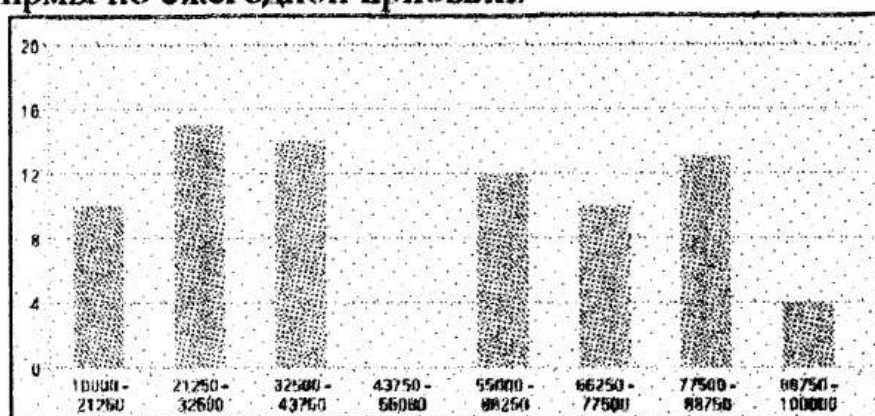
Секція інформатики

Для пользователей, занимающихся бизнес-аналитикой или интеллектуальным анализом данных, эти надстройки предоставляют мощные и простые в использовании средства для работы с моделями интеллектуального анализа данных в Analysis Services.

Чтобы MS Excel определил диапазон данных как таблицу, надо применить "Быстрый" формат и выбрать вид оформления: "Форматировать как таблицу". После щелчка мышью внутри таблицы MS Excel на экране будет отображаться пункт меню "Работа с таблицами", а под ним - панель инструментов "Средства анализа таблиц".



В процессе построения аналитических моделей много времени занимает задача выбора правильных атрибутов из источника данных и представление их в удобном формате. Группа кнопок "Подготовка данных" предоставляет инструменты для основных задач подготовки данных до начала их углубленного анализа. Кнопка "Просмотр данных" – служит для построения графика распределения дискретных и непрерывных переменных, а также для добавления группировок в исходные данные. Кнопка "Очистка данных" – служит для удаления выбросов и для изменения значений меток дискретных данных. Кнопка "Секционирование данных" – служит для разбиения исходных данных на обучающее и тестовое множество посредством случайных выборок исходных данных. Ниже представлена гистограмма, где сгруппированы клиенты фирмы по ежегодной прибыли.



1. Дюк В.А., Самойленко А.П. Data Mining: учебный курс. – Изд во «Питер», 2001. – 368 с.
2. Использование компонентов Data Mining в продуктах Office 2007. Гончаров М. Web-страница "Технологии добычи знаний и интеллектуального анализа данных" <http://www.businessdataanalytics.ru/DataMiningInOffice2007.htm>

3. Харинатх, Сивакумар, Куинн, Стивен. SQL Server 2005 Analysis Services и MDX для профессионалов. – М.: ООО "И.Д. Вильямс". - 2008. – 848 с.

ПОБУДОВА ІЄРАРХІЧНОЇ СТРУКТУРИ ДЛЯ РОЗПІЗНАВАННЯ ЕЛЕКТРОНОГРАМ ЗА ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИМ МЕТОДОМ

Крамінська Л.С., студ. 5-го курсу ММФ, асп. Мартиненко С.С., СумДУ

Машинне розпізнавання електронограм, одержаних на електронному мікроскопі в режимі мікродифракції, є актуальною задачею, оскільки дозволяє підвищити точність та оперативність експрес-аналізу хімічного складу різних матеріалів в металургії, геології, кристалографії та інше.

Більшість відомих алгоритмів розпізнавання зображень орієнтовано на розв'язання модельних задач, які виключають перетин класів і потребують статистичної стійкості та однорідності навчальної матриці, що на практиці, як правило, не виконується. Одним із шляхів вирішення цієї проблеми є застосування для розпізнавання електронограм методів інформаційно-екстремальної інтелектуальної технології (ІІТ), що ґрунтується на максимізації інформаційної спроможності системи розпізнавання шляхом введення на етапі навчання додаткових інформаційних обмежень [1].

Нехай дано вектор параметрів функціонування системи розпізнавання

$$g = \langle g_1, \dots, g_\xi, \dots, g_\Xi \rangle,$$

тоді алгоритм оптимізації параметра функціонування, наприклад, g_ξ у рамках інформаційно-екстремального методу синтезу системи розпізнавання зображень передбачає виконання циклу ітераційної процедур:

$$g_\xi = \langle \arg \left\{ \max_{G_\xi} \left\{ \dots \left\{ \max_{G_\xi} \frac{1}{M} \sum_{m=1}^M E_m \right\} \right\} \right\} \rangle,$$

де E_m - інформаційний критерій функціональної ефективності (КФЕ) навчання системи розпізнавати реалізації класу $X_m^o \in \{X_m^o \mid m = \overline{1, M}\}$

Для проведення дослідження використовувались електронограми таких металів: алюміній, золото, тодоракіт. Як КФЕ навчання системи розпізнавання використовувався ентропійний критерій Шеннона. Як алгоритм навчання системи було реалізовано інформаційно-екстремальний алгоритм паралельної оптимізації контрольних допусків на ознаки розпізнавання [1].