

О.В. Ольшевська, А.В. Селіванова

Одеська національна академія харчових технологій, вул. Канатна, 112, Одеса, 65039

ІМПЛЕМЕНТАЦІЯ СУЧАСНОГО ІНСТРУМЕНТАРІЮ АНАЛІЗА ДАНИХ ДЛЯ ОБРОБЛЕННЯ ПОКАЗНИКІВ ОТРИМАНИХ У СИСТЕМАХ КОНДИЦІОНУВАННЯ ПОВІТРЯ ТА ХОЛОДИЛЬНИХ УСТАНОВКАХ. ЧАСТИНА 1

В роботі розглянуто сучасні інструментарії для аналізу даних, їх основні типи та властивості, виокремлено ті, які можна використовувати при обробленні даних та показників отриманих у результаті наукових та промислових експериментів. Стаття є першою частиною, серед циклу статей присвячених теоретичним засадам сучасних методів аналізу даних і їх можливого застосування у галузі енергомашинобудування.

Ключові слова: інтелектуальний аналіз даних, системи кондиціонування повітря, температура, тиск, інструментарій.

О.В. Ольшевская, А.В. Селиванова

Одесская национальная академия пищевых технологий, ул. Канатная, 112, Одесса, 65039

ІМПЛЕМЕНТАЦІЯ СОВРЕМЕННОГО ИНСТРУМЕНТАРИЯ АНАЛИЗА ДАННЫХ ДЛЯ ОБРАБОТКИ ПОКАЗАТЕЛЕЙ ПОЛУЧЕННЫХ В СИСТЕМАХ КОНДИЦИОНИРОВАНИЯ ВОЗДУХА И ХОЛОДИЛЬНЫХ УСТАНОВКАХ. ЧАСТЬ 1

В работе рассмотрены современные инструментари для анализа данных, их основные типы и свойства, выделены те, которые можно использовать при обработанные данных и показателей полученных в результате научных и промышленных экспериментов. Статья является первой частью, среди цикла статей посвященных теоретическим основам современных методов анализа данных и их возможного применения в отрасли энергомашиностроения.

Ключевые слова: интеллектуальный анализ данных, системы кондиционирования воздуха, температура, давление, инструментарий

I. ВСТУП

Стрімкий розвиток сучасних технологій та промисловості, зокрема в галузі енергомашинобудування, «генерує» велику кількість інформації, змінних, даних та показників, які необхідно систематизувати та аналізувати, адже за будь-яким з цих показників знаходяться фінансова складова. Для оброблення даних (показників) та інформації існує велика кількість методів аналізу, більшість з яких використовується інженерами, проєктувальниками при роботі над дослідженнями.

II. ТЕОРЕТИЧНА ЧАСТИНА

Жодне дослідження не можливо уявити без аналізу даних отриманих в процесі його виконання, таким чином необхідно володіти певними методами (інструментами) аналізу даних.

Аналіз даних – це сукупність методів і засобів отримання з певним чином організованих даних інформації для прийняття рішень. Можна виділити такі етапи аналізу даних: отримання даних, обробка, аналіз та інтерпретація результатів обробки.

Одже сучасний фахівець, який працює з великою кількістю змінних мусить володіти інструментарієм для аналізу отриманої інформації.

Необхідно також зауважити, що обсяг даних наявних у дослідника ХХІ сторіччя у рази більший у порівнянні з обсягом даних дослідників ХХ сторіччя. Тому, безумовно, для аналізу даних використовують сучасні програмне забезпечення, яке дозволяє обробляти великі об'єми інформації та даних. Специфіка сучасних вимог до продуктивної переробки інформації така [2]:

- змінні мають необмежений обсяг;
- змінні є різномірними (кількісними, якісними, текстовими);
- результати повинні бути конкретний і зрозумілий;
- інструменти для обробки даних повинні бути прості у використуванні.

Для того, щоб провести автоматизований аналіз даних, використовується Data Mining.

Термін інтелектуальний аналіз даних (англ. Data Mining) означає виявлення прихованих залежностей або взаємозв'язків між змінним у великих масивах необроблених даних.

Сутність і мету Data Mining можна описати так: це технологія, призначена для пошуку у великих інформаційних масивах даних об'єктивних, корисних на практиці закономірностей [1].

Основне визначення даного інструменту у 1996 році надав його винахідник Г. П'ятецький-Шапіро:

«Data Mining - це процес виявлення у необроблених даних раніше невідомих нетривіальних, практично корисних і доступних інтерпретацій знань, необхідних для прийняття рішень у різних сферах діяльності.»

Попередником Data Mining є традиційна математична статистика, яка протягом тривалого часу була основним інструментом аналізу даних, проте її головною проблемою є усереднювання по вибірці, що приводить до операцій над фіктивними величинами.

В основу технології Data Mining покладено концепцію закономірностями, які властиві вибіркам даних, створених програмним інструментарієм і можуть бути подані у формі, зрозумілій людині.

Як і будь яка технологія Data Mining, є багатоманітним, багатофункціональним і вирішує велику кількість поставлених задач.

Задачі Data Mining [1]:

– Класифікація (Classification) - виявляються ознаки, які характеризують групи об'єктів досліджуваного набору даних - класи; за цими ознаками новий об'єкт можна віднести до того або іншого класу.

– Кластеризація (Clustering) - результатом її є поділ об'єктів на групи.

– Асоціація (Associations) - знаходять закономірності між пов'язаними подіями у наборі даних.

– Послідовність (Sequence), або послідовна асоціація (sequential association), - дає можливість знайти часові закономірності між транзакціями. Завдання послідовності подібне до асоціації, але її метою є встановлення закономірностей між подіями, пов'язаними за часом, тобто послідовність визначається високою ймовірністю ланцюжка пов'язаних за часом подій.

– Прогнозування (Forecasting) - на основі особливостей історичних даних оцінюються майбутні значення показників. Застосовуються методи математичної статистики, нейронні мережі тощо.

– Визначення відхилень (Deviation Detection), аналіз відхилень або викидів - виявлення й аналіз даних, що найбільше відрізняються від загальної чисельності даних, виявлення нехарактерних шаблонів.

– Оцінювання (Estimation) - зводиться до прогнозу безперервних значень ознак.

– Аналіз зв'язків (Link Analysis) - задача знаходження залежностей у наборі даних.

– Візуалізація (Visualization, Graph Mining) - створюється графічний образ аналізованих даних. Для вирішення задач візуалізації використовуються графічні методи, що показують наявність закономірностей в даних.

– Підбивання підсумків (Summarization) - опис конкретних груп об'єктів за допомогою аналізованого набору даних.

Всі вищезазначені задачі розділяються на описові і предикативні.

Описові, або дескриптивні (descriptive), задачі пов'язані з поліпшенням розуміння аналізованих даних. Ключовий момент у таких моделях - простота і прозорість результатів для сприйняття людиною. До такого типу задач належать кластеризація і пошук асоціативних правил.

Рішення предикативних (predictive), або прогнозуючих, задач поділяється на два етапи. На першому етапі на підставі набору даних з відомими результатами будується модель. На другому етапі вона використовується для прогнозу результатів на підставі нових наборів даних. Вимагається, щоб побудовані моделі працювали максимально коректно. До цього типу задач відносять задачі класифікації і регресії. Сюди можна віднести і задачу пошуку асоціативних правил, якщо результати її рішення можуть бути використані для прогнозу появи деяких подій.

Основною задачею Data Mining в промисловому виробництві, зокрема в енергомашинобудуванні є [2]:

– комплексний системний аналіз виробничих процесів та ситуацій;

– короткостроковий і довгостроковий прогноз розвитку виробничих процесів;

– вироблення варіантів оптимізаційних рішень роботи змінних параметрів систем устаткування;

– прогнозування якості виробу (наприклад, фруктів, овочів, м'ясо-молочної продукції) в залежності від деяких параметрів технологічного процесу обробки, зберігання та транспортування;

– виявлення прихованих тенденцій і закономірностей розвитку виробничих процесів;

– прогнозування закономірностей розвитку виробничих процесів;

– виявлення прихованих чинників впливу на той чи інший чинник виробничого процесу;

– виявлення і ідентифікація раніше невідомих взаємозв'язків між виробничими параметрами і чинниками впливу;

– візуалізацію результатів аналізу, підготовку попередніх звітів і проектів допустимих рішень з оцінками достовірності і ефективності можливих реалізацій.

Базуючись на отриманій інформації застосуємо деякі з задач на практиці

III. АНАЛІТИЧНА ЧАСТИНА

У якості масиву змінних були прийняті, данні отримані під час експлуатації камер розподільного холодильника . Робоча речовина була R717a. Дані отримувались з інтервалом в одну годину протягом одного календарного місяця.

Одними з основних параметрів для аналізу роботи будь-якої холодильної установки є її температурний режим, вартість експлуатації.

Використаємо інструмент «Аналіз ключових факторів впливу», для цього приймемо як ключо-

вий фактор температуру навколишнього середовища та визначимо її вплив на деякі параметри, наприклад, температуру конденсації, тиск конденсації, температуру кипіння та тиск в проміжній посудині (рис.1):

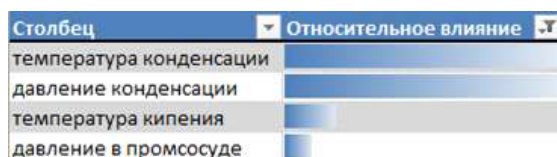


Рис.1. «Аналіз ключових факторів впливу»

Аналіз показав, точніше підтвердив те, що температура навколишнього середовища є визначним фактором для параметрів конденсації, а вплив на інші фактори не значний.

За наявності великого масиву даних інструмент «Аналіз ключових факторів впливу» є незамінним помічником для визначення ступеня впливу одного параметра на інший.

Одним з важливих та необхідних інструментів є «Прогноз», який дозволяє спрогнозувати можливі зміни у майбутньому базуючись на існуючих показниках. Цей інструмент не є суто науковим, проте він дозволяє визначити потенційно «цікавий» діапазон параметрів величини яких можуть тим, чи іншим чином вплинути на дані дослідження. Для наочності застосуємо інструмент «Прогноз» для прогнозування температури конденсації у другій половині дня (рис.2).



Рис.2. «Прогноз»

Загалом ці два інструменти відображають тільки верхівку всього можливого потенціалу та інструментарія Data Mining.

IV. ВИСНОВКИ

Важливість та необхідність застосування Data Mining, пов'язана з відсутністю «шаблонних» рішень при аналізі тієї чи іншої ситуації

ЛІТЕРАТУРА:

1. Інтелектуальні технології Data Mining і Text Mining [Електронний ресурс] – Режим доступу до ресурсу: http://pidruchniki.com/1623021247786/informatika/intelektualni_tehnologiyi_data_mining_text_mininig.
2. Grid і інтелектуальна обробка даних Data Mining [Електронний ресурс] – Режим доступу до ресурсу: <http://netallted.cad.kiev.ua/downloads/DataMining.pdf>.

O. Olshevska, A. Selivanova

Odessa National Academy of Food Technologies, 112 Kanatnaya str., Odessa, 65039, Ukraine

MODERN TOOLS DATA ANALYSIS IMPLEMENTATION FOR PROCESSING PARAMETERS OBTAINED IN AIR CONDITIONING AND REFRIGERATION UNITS. PART 1

The paper discusses the modern tools data analysis, their main types and properties, identify those which can be used in the processed data and indicators obtained from scientific and industrial experiments. The article is the first part, among a series of articles devoted to the theoretical basis of modern methods of data analysis and their possible application in the power engineering industry.

Keywords: data mining, air conditioning systems, temperature, pressure, tools

REFERENCES

1. Інтелектуальні технології Data Mining і Text Mining [Електронний ресурс] – Режим доступу до ресурсу:

- http://pidruchniki.com/1623021247786/informatika/intelektualni_tehnologiyi_data_mining_text_mininig.
2. Netallted.cad.kiev.ua, (2015). [online] Available at: <http://netallted.cad.kiev.ua/downloads/DataMining.pdf>. [Accessed 5 Nov. 2015].