



**2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ
ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ**

УДК 519.876.2:681.5.015

ВИКОРИСТАННЯ ТЕХНОЛОГІЇ DATA MINING В АВТОМАТИЗОВАНІЙ СИСТЕМІ ПРОСТЕЖУВАННОСТІ ВИРОБНИЦТВА СИРОЇ СОНЯШНИКОВОЇ ОЛІЇ

Жигайло О.М.¹

¹Одеська національна академія харчових технологій, Одеса

E-mail: dr_jam2006@ukr.net

Copyright © 2014 by author and the journal "Automation technological and business - processes".

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



ONAF
Open Access

DOI: [10.15673/2312-3125.19/2014.27953](https://doi.org/10.15673/2312-3125.19/2014.27953)

Анотація

Для реалізації процесів виробництва сирової соняшникової олії використовується автоматизована система простежуваності. Її вдосконалення за рахунок застосування методик вилучення знань дасть можливість збільшити ефективність управління якістю та досягти економічного зростання підприємства.

Abstract

For realization processes of production of crude sunflower oil is used automated traceability system. Its improvement through the use of methods of knowledge extraction will increase the effectiveness of quality management and will achieve economic growth of the company.

Ключові слова

Управління якістю, система простежуваності, соняшникова олія, вилучення знань, KDD, Data Mining, кластер.

Вступ.

Для успішного представлення своєї продукції на міжнародних ринках українські підприємства харчової промисловості повинні користуватися міжнародними стандартами. Однією з головних цілей підприємств є підвищення якості та безпеки продуктів, що виробляються, так як саме ці складові знаходяться в центрі уваги урядових і промислових кіл усього світу. Досягнення цієї мети можливе за рахунок впровадження системи управління якістю (ДСТУ ISO 9001:2009) та системи управління безпечністю харчових продуктів (ДСТУ ISO 22000:2007).

У відповідності зі стандартом [1] керування та управління якістю включає в себе розробку політики та цілей у сфері якості, планування якості, управління якістю, забезпечення якості та поліпшення якості. Поліпшення якості розглядається як частина менеджменту якості, спрямована на збільшення здатності виконати вимоги до якості. Серед аспектів до яких можуть ставитися ці вимоги, для прикладу, називають такі: результативність, ефективність чи простежуваність.

Також відомо, що проектування і розробка це сукупність процесів, які перетворюють вимоги у встановлені характеристики або специфікації на продукцію, процес або систему. А простежуваність відноситься до характеристик (рис.1) і визначає можливість простежити історію, застосування або місцезнаходження того, що розглядається.



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

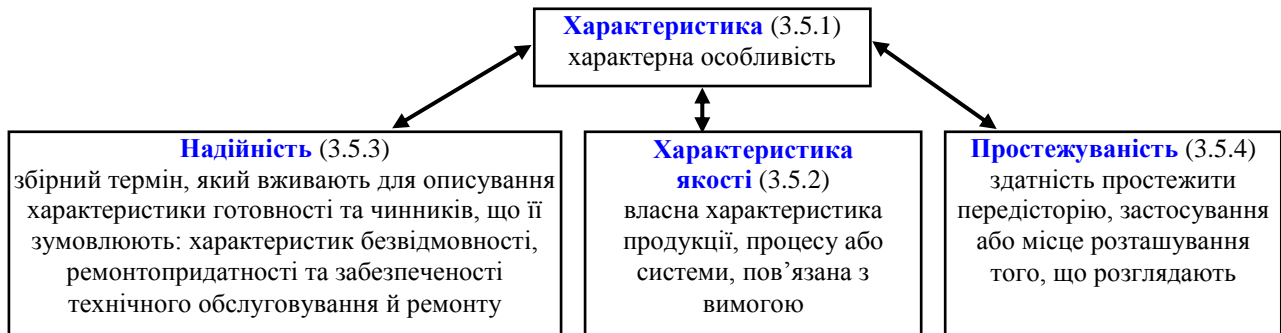


Рис. 1 – Поняття, що відносяться до характеристик

У відповідності зі стандартом [2] організація повинна встановити та застосовувати систему простежуваності, яка дає змогу ідентифікувати партії продукції та їх зв'язок з партіями сировини, протоколами щодо оброблення та постачання. Вона повинна представляти повну сукупність даних та операцій, а також бути здатна містити необхідну інформацію про продукцію та її компоненти протягом всього ланцюжка виробництва та використання продукції або окремої частини продукції. Виконання цих операцій та підготовка необхідних даних можуть бути названі процесом ідентифікації та простежуваності. Тобто можна задіяти ідеологію процесного підходу, яка дозволяє виділити на підприємстві найбільш важливі бізнес-процеси та організувати їх ефективне управління.

Відомо, що у процесі виробництва відбувається рух трьох потоків: матеріального (сировини, матеріалів, деталей, вузлів і т.д.), енергетичного (електроенергія, газ, пар, вода і т.д.) та інформаційного (данні про сировину, матеріали, деталі, енергоресурси і т.д. на паперових та інших носіях, які супроводжують матеріальний потік). Тому необхідною умовою досягнення ефективного результату в процесі ідентифікації та простежуваності по всьому виробничому циклу, починаючи з надходження сировини, матеріалів і комплектуючих виробів від постачальників і закінчуючи відвантаженням готової продукції, є забезпечення однозначного зв'язку всіх потоків у процесі виробництва продукції. Це означає, що інформаційний потік повинен бути безперервним по всьому технологічному ланцюжку, об'єктивно відображати зміни продукції в процесі виробництва, зберігати відмітні ознаки об'єкта ідентифікації.

Таким чином для ефективного управління процесом ідентифікації та простежуваності необхідно реалізувати комплекс робіт, який включає наступні взаємопов'язані напрями: технічне (технологічне) забезпечення; інформаційне забезпечення; організаційне забезпечення. Інформаційне забезпечення ідентифікації та простежуваності має включати в себе розробку і впровадження інформаційних форм (супровідних документів, інших носіїв), схем інформаційних потоків і точок реєстрації інформації, а також процедур обробки, систематизації, зберігання та реалізації даних про об'єкти ідентифікації, включаючи машинні способи обробки інформації (рис. 2).

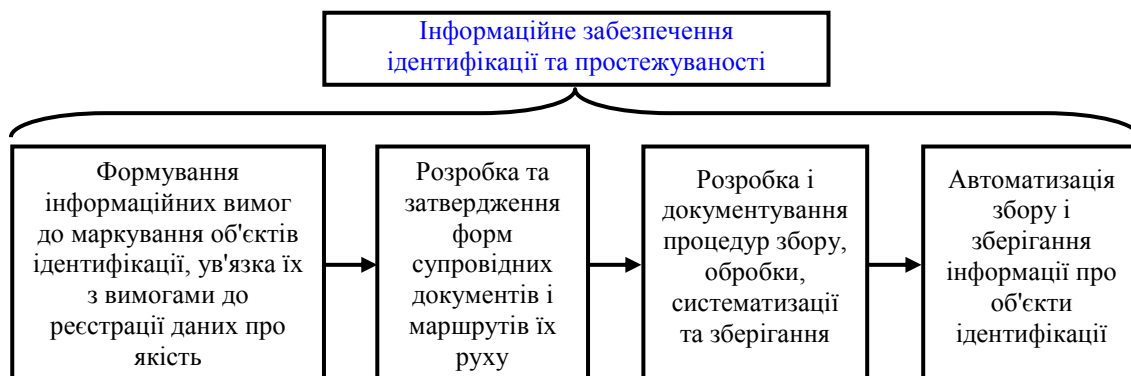


Рис. 2 – Склад робіт з інформаційного забезпечення

Мета і завдання дослідження

На даний час для підприємств продовжує бути актуальним питання збільшення рівня автоматизації процесів



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

документообігу, оперативного обліку та функціонування систем управління якістю і безпечності кормів та харчових продуктів. Тому метою статті є розробка і дослідження більш ефективних алгоритмів управління якістю виробництва сирової соняшникової олії. А завдання полягає у визначенні взаємозв'язків між вхідними та вихідними даними цього процесу для досягнення наступних цілей:

- 1) розробка алгоритмів поліпшення якості соняшникової олії за рахунок застосування автоматизованої системи простежуваності;
- 2) прогнозування економічних результатів діяльності оліє-пресового заводу.

Матеріали і методи

Проведені дослідження показали, що зараз в Україні вже існує велика кількість підприємств, які потребують впровадження автоматизованої системи простежуваності. Тому їх розробка є більш ніж доцільною.

Одним з позитивних прикладів вирішення такої задачі став процес виробництва сирової соняшникової олії на оліє-пресовому заводі (ОПЗ) ТОВ "Катеринопільський елеватор" [3]. Загальну, спрощену, схему процесу виробництва олії на ОПЗ представлено на рис. 3.

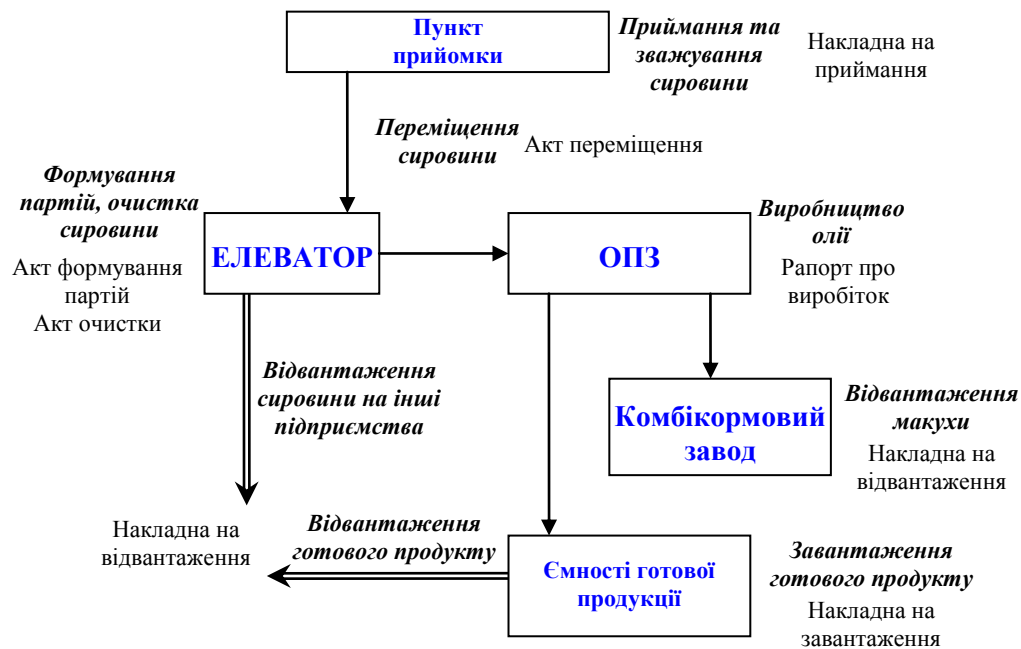


Рис. 3 – Загальна схема процесу виробництва олії

На базі цього заводу була розроблена, а потім впроваджена автоматизована система простежуваності виробництва сирової соняшникової олії. Вона є повноцінною та реалізує наступні принципи:

- а) автоматична ідентифікація;
- б) збір, запис та зберігання даних;
- в) організація інформаційних зв'язків і посилань;
- г) обмін даними.

Спрощений алгоритм роботи цієї системи в рамках одного сезону представлено на рис. 4.



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

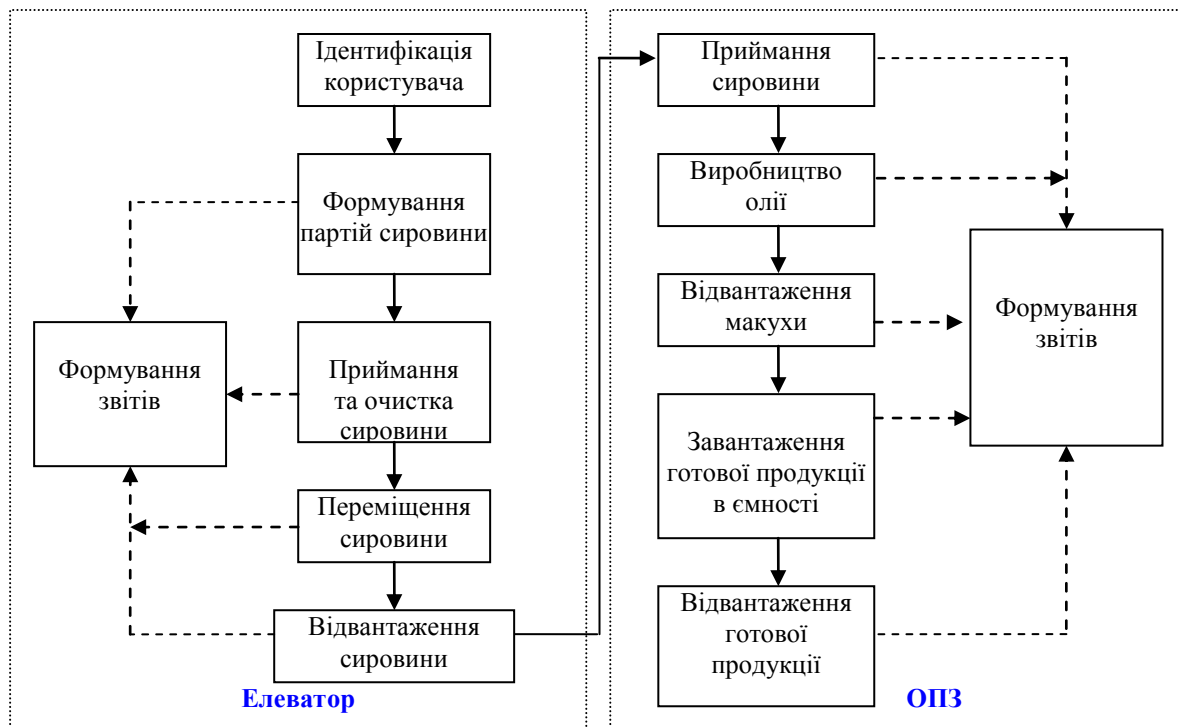


Рис. 4 – Спрощений алгоритм роботи автоматизованої системи простежуваності

Розроблене програмне забезпечення охоплює всі етапи виробництва олії на ОПЗ та надає підприємству змогу вести автоматизований облік сировини та готової продукції, формує інформативні звіти, а головне – забезпечує управління процесом простежуваності.

На цьому етапі робота системи не впливає на якість олії. Але можна зробити припущення, що вона може це робити, так як може збирати та зберігати великі об'єми даних, що накопичуються у ході вирощування насіння, його збирання та обробки. Для цього треба розуміти, які самі дані треба аналізувати та які інтелектуальні алгоритми дозволять автоматизувати управління якістю готового продукту, якщо їх додати у функціонал розробленої системи. Тому розглянемо звідки береться необхідна інформація та ті дані, що її відображають.

Соняшникову олію отримують з насіння соняшнику, що містить до 50 % (і більше) жиру в перерахунку на абсолютно суху речовину. Її виробляють шляхом пресування (гарячого або холодного) і екстракції. Залежно від ступеня очищення соняшникова олія поділяється на три види - нерафінована, гідратована і рафінована. Спосіб отримання і ступінь очищення впливають на органолептичні та фізико-хімічні показники соняшникової олії. Нерафінована соняшниковая олія відрізняється натуральним смаком і запахом, властивим підсмаженому насінню, має світло-жовтий колір і за органолептичними та фізико-хімічними показниками ділиться на три сорти - вищий, перший, другий. Різниця за сортами обумовлено кольоровістю (відповідно не більше 15, 25, 35 мг йоду), кількістю відстою (не більше 0,05; 0,1 і 0,2% по масі), кислотним числом (не більше 1,5; 2,25 і 6,0 мг КОН/г), кількістю фосфатидів (не більше 0,4; 0,6 і 0,8% в перерахунку на стеароолеоліцітін).

Масло вищого і першого сортів має бути прозорим, над осадом допускається «сітка» (наявність в прозорому маслі окремих дрібних часток подібних до воску речовин, ледве помітних неозброєним оком), у другому сорті допускається легке помутніння (наявність суцільного фону найдрібніших частинок подібних до воску речовин, незначно знижують прозорість масла). У масла вищого і першого сортів повинні бути запах і смак, властиві соняшниковій олії, без стороннього запаху, присмаку і гіркоти. У другому сорті допускаються злегка затхлий запах і присмак легкої гіркоти; масло другого сорту призначено промислової переробці.

Визначення фізико-хімічних параметрів лежить в основі контролю якості масел, дозволяє проводити ідентифікацію готової партії, достовірно визначати його природність, ступінь очищення, виключати можливість фальсифікації або невідповідності критеріям безпеки. До основних і найбільш часто контрольованих показників (крім вже вказаних) також відносяться щільність, в'язкість, показник заломлення, температура плавлення і застигання, температура спалаху, твердість, перекисне число, число омилення, вміст вітаміну Е, кількість афлатоксину В1.



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

Якість соняшникової олії залежить від якості вихідної сировини - насіння соняшнику відразу після збору врожаю, а також від умов їх зберігання, так як збирають насіння на протязі сезону, а їх переробка в соняшкову олію може здійснюватися цілий рік. Тому саме показники якості насіння підлягають визначенню та аналізу. До них можна віднести: кислотне число, вологість, оліїстість, лузжистість, маса 1000 сім'янок.

В свою чергу також відомо, що якість врожаю насіння соняшнику в значній мірі залежить від умов вирощування (приналежність до тієї чи іншої кліматичної зони) та особливостей проведення прибирання. Так при перестойі і прибиранні соняшнику у вологу погоду кислотне число може зрости в кілька разів і перевищити допустиму для харчового масла (високого і першого сорту) межу 6,0 мг КОН/г, переводячи його в розряд технічних масел (другого сорту). Ще одну серйозну проблему викликає урожай з підвищеною засміченістю. Через насіння бур'янів, які мають підвищену інтенсивність дихання, підвищується температура зернової маси і тим самим інтенсифікується протікання біохімічних процесів в ньому, зокрема гідролітичних і окислювальних. Все це, в кінцевому підсумку, призводить до зниження якості вихідної сировини і тягне за собою збільшення витрат на його додаткову обробку (очищення, сушіння) та зниження якості кінцевого продукту. Наслідком цього стає зниження економічних показників ефективності всього виробництва.

Таким чином для повної оцінки умов виробництва більш якісної олії слід збирати інформацію о місцях та кліматичних умовах вирощування соняшника, що включає в себе метеорологічні спостереження, а також о якості насіння, яке відправляють на зберігання. Ця інформація може бути представлена вхідними даними, а інформація об якості олії, що виготовляється, вихідними даними для автоматизованої системи простежуваності (рис. 4).

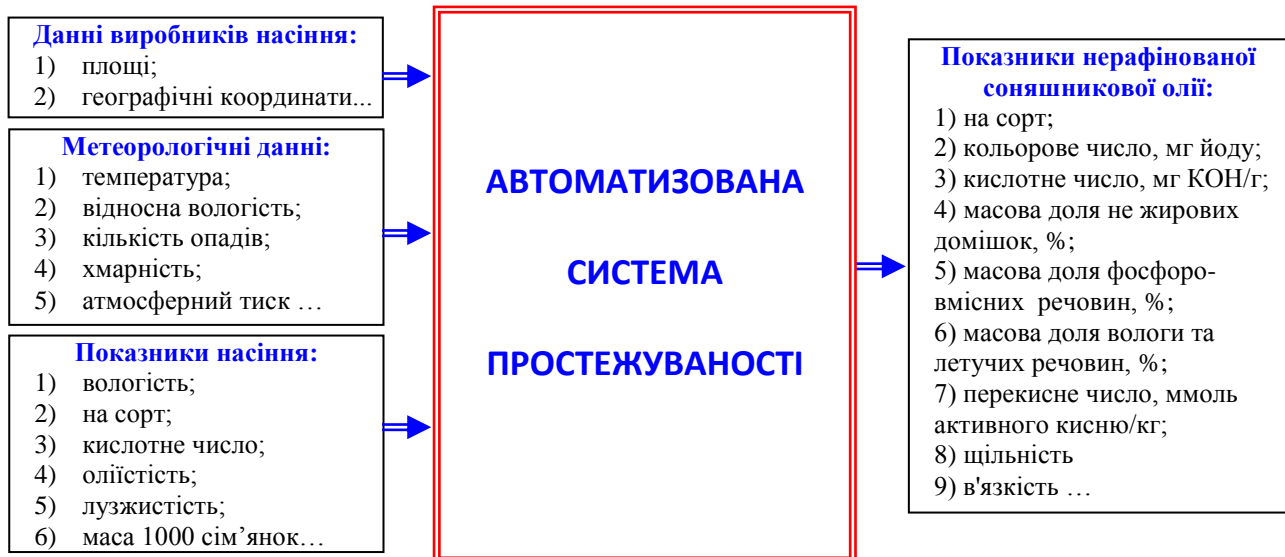


Рис. 4 – Загальна структура даних, що можуть оброблятися в автоматизованій системі простежуваності

Особливості виробництва сирої соняшникової олії, впроваджена автоматизована система простежуваності створюють хороші передумови для можливості використання інформаційного підходу до аналізу предметної області, яка нами розглядається. Цій підхід набув поширення в такій методиці вилучення знань з баз даних, як **Knowledge Discovery in Databases (KDD)** [4]. Сьогодні на її базі створюється більшість прикладних аналітичних рішень у бізнесі та багатьох інших областях.

Вона описує не конкретний алгоритм або математичний апарат, а послідовність дій, яку необхідно виконати для виявлення корисного знання. Методика не залежить від предметної області. Це набір атомарних операцій, комбінуючи які, можна отримати потрібне рішення. KDD включає в себе етапи підготовки даних, вибору інформативних ознак, очищення, побудови моделей, додаткової обробки та інтерпретації отриманих результатів. Ядром цього процесу є сукупність великого числа різних методів **Data Mining**. Вони дозволяють виявляти в «сирих» даних раніше невідомі, нетривіальні, практично корисні і доступні інтерпретації знань та закономірностей, необхідних для прийняття рішень у різних сферах людської діяльності. Існує кілька умовних класифікацій задач Data Mining. По одному з варіантів виділяють чотири базових класи:



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

1. Класифікація - це встановлення залежності дискретної вихідної змінної від вхідних змінних.
2. Регресія - це встановлення залежності безперервної вихідної змінної від вхідних змінних.
3. Кластеризація - це угруповання об'єктів (спостережень, подій) на основі даних, що описують властивості об'єктів. Об'єкти всередині кластера повинні бути схожими один на одного і відрізнятися від інших, які увійшли в інші кластери.
4. Асоціація - виявлення закономірностей між пов'язаними подіями.

Якщо розглядати оліє-пресове виробництво, то серед основних завдань використання Data Mining на ньому можна виділити такі:

1. Прогнозування якості олії залежно від деяких параметрів виробничих процесів;
2. Виявлення прихованих факторів впливу;
3. Виявлення та ідентифікація раніше невідомих взаємозв'язків між виробничими параметрами і факторами впливу;
4. Вироблення оптимізаційних рекомендацій з управління виробничими процесами;
5. Візуалізацію результатів аналізу, підготовку попередніх звітів і проектів допустимих рішень з оцінками достовірності та ефективності можливих реалізацій.

Результати дослідження та їх обговорення

Серед основних завдань використання технологій Data Mining є вирішення задачі *прогнозування*. Для цього можуть бути використані наступні методи:

- 1) кластерного аналізу для автоматичного дослідження та відкриття;
- 2) регресія, нейронні мережі, генетичні алгоритми, дерева рішень для прогнозування та класифікації.

Метою кластеризації є пошук існуючих структур. Це описова процедура, яка не робить ніяких статистичних висновків, але дає можливість провести розвідувальний аналіз і вивчити "структуру даних". В якості інструменту за допомогою якого можливо скористатися вказаними методами було обрано **Deductor Academic** [4]. Deductor це технологічна платформа для створення закінчених аналітичних рішень. У ньому зосереджені найсучасніші методи добування, маніпулювання, візуалізації даних, кластеризації, прогнозування та багато інших технологій інтелектуального аналізу даних.

Як приклад, розглянемо проведення аналізу та виявлення корисних знань з даних по прийманню насіння на Катеринопільський елеватор в 2011 році, що фрагментом представлені в табл. 1.

Таблиця 1 - Данні по прийманню насіння на елеватор

Дата	Код госп-ва	Назва господарства	№ автомобіля	Приход	Волога	Код складу	Назва складу
29.08.2011	42	Зарічанка ФГ	СА6272ВА	15880	4.00	03	Склад 3
29.08.2011	77	Правобережне ФГ	СА3987ВС	13620	8.59	05	Склад 5
29.08.2011	109	Фролов В.В.	СА6482ВС	27240	6.38	05	Склад 5
31.08.2011	42	Зарічанка ФГ	СА6272ВА	17640	4.02	03	Склад 3
31.08.2011	55	Корсунське ТОВ	СА1878ВС	17720	9.20	03	Склад 3
31.08.2011	69	Оптіма - Грейн ПП	СА5361ВС	27040	8.96	03	Склад 3
...
04.11.2011	80	ПРИЛУКИ	95368627	31660	7.38	13	Склад 13
04.11.2011	80	ПРИЛУКИ	95527016	31660	7.45	13	Склад 13
05.11.2011	95	СТАРОКОНСТАНТИ-НІВСЬКИЙ елеватор	95345575	36140	7.15	13	Склад 13
05.11.2011	95	СТАРОКОНСТАНТИ-НІВСЬКИЙ елеватор	95352563	38260	7.01	13	Склад 13

Приймання насіння відбувалося з 29 серпня по 5 листопада. Загальна кількість записів у базі даних по цьому процесу складала 1140 строк, а насіння привозили з 114 господарств. По якості насіння розглядався тільки один показник - його волога.

Проведення кластерного аналізу було здійснено за допомогою двох алгоритмів, які мають відношення до кластеризації багатомірних векторів:



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

1) алгоритмом k-середніх (k-means, f-means), що засновані на оптимізації деякої цільової функції, яка визначає оптимальне в певному сенсі розбиття множини об'єктів на кластери (рис. 5 та рис. 6);



Рис. 5 – Профілі кластерів при використанні алгоритму k-середніх

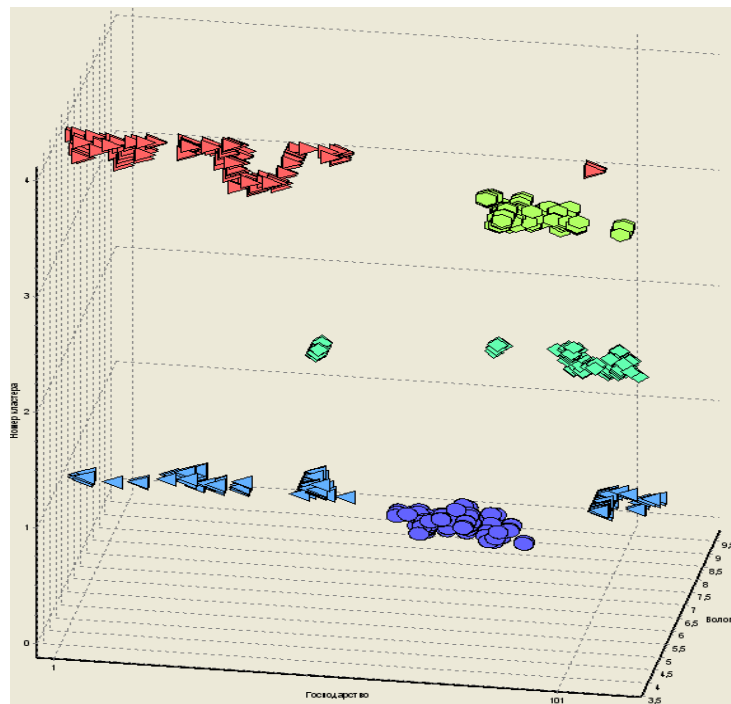


Рис. 6 – Діаграма розміщення кластерів при використанні алгоритму k-середніх

2) нейромереживим алгоритмом - карта Кохонена, що самоорганізується, в якому використовується метод навчання без вчителя, тобто результат навчання залежить тільки від структури вхідних даних (рис. 7, рис. 8 та рис. 9).



**2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ
ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ**

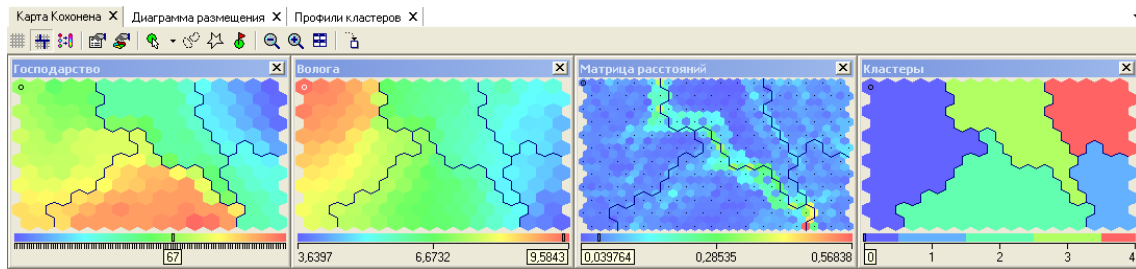


Рис. 7 – Карта Кохонена

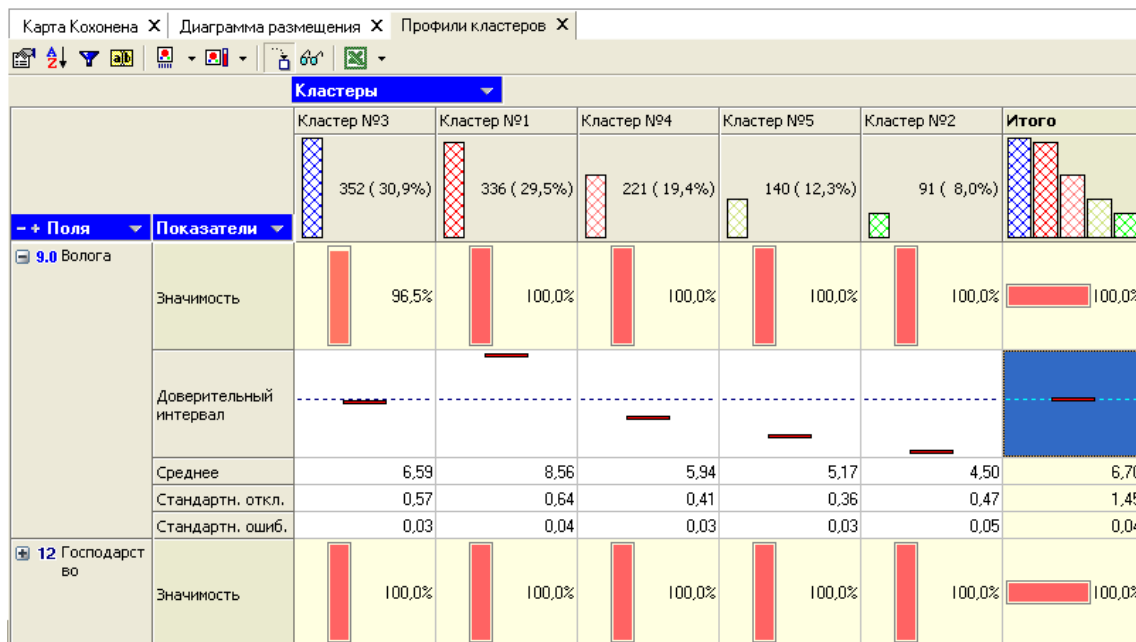


Рис. 8 – Профілі кластерів при використанні алгоритму карта Кохонена, що само організується

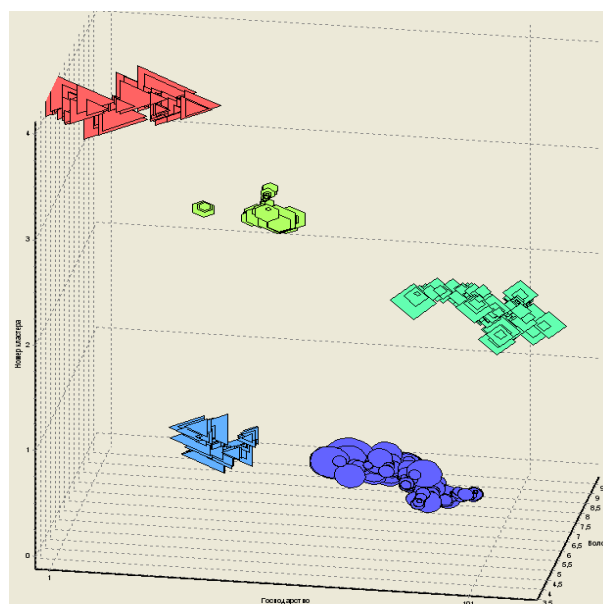


Рис. 9 – Діаграма розміщення кластерів при використанні алгоритму карта Кохонена, що само організується



2 АВТОМАТИЧНІ ТА АВТОМАТИЗОВАНІ СИСТЕМИ УПРАВЛІННЯ ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ

Висновки

1. Відстеження історії створення готового продукту передбачає існування автоматизованої системи простежуваності. Великі об'єми інформації, які може в собі містити база даних цієї системи, повинні використовуватися більш раціонально. Тому в статті запропоновано розглядати цю систему як початок для створення автоматизованої системи управління якістю продукції;
2. Розглянуте в статті виробництво сирової соняшникової олії, показало, що на якість готової продукції великий вплив справляють якісні показники сировини, а на них в свою чергу впливають умови її вирощування та особливості проведення прибирання, які насамперед обумовлені метеорологічними показниками. Тобто для більш повноцінної реалізації процесу простежуваності в автоматизовану систему слід включати максимальну кількість чинників від яких залежить характеристика готової продукції;
3. Результати кластерного аналізу показали, що представлені в прикладі дані дозволили сформувати 5 кластерів, в які об'єдналися певні господарства. Їх відмінність полягає у вологовмісту насіння, яке вони привезли. Це дозволяє здійснювати розміщення насіння по складам більш коректно та прогнозувати рівень економічних показників готової продукції.

Література

1. ДСТУ ISO 9000:2007. Системи управління якістю. Основні положення та словник термінів. (ISO 9000:2005, IDT). Київ, Держспоживстандарт України, 2008;
2. ДСТУ ISO 22005:2009. Простежуваність у кормових та харчових ланцюгах. Загальні принципи та основні вимоги щодо розроблення та запровадження системи (ISO 22005:2007, IDT). Київ, Держспоживстандарт України, 2010;
3. Трішин Ф.А., Жигайло О. М., Гусаковський В.А. Автоматизована система простежуваності // Автоматизація технологічних і бізнес-процесів. – Одеса, 2012. – № 9,10. – с. 21-25;
4. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям (+CD): Учебное пособие. 2-е изд., испр. — СПб.: Питер, 2013. - 706 с.: ил.

УДК 531.1:303.094.7

ИСПОЛЬЗОВАНИЕ ПАКЕТА SIMSCAPE ДЛЯ МОДЕЛИРОВАНИЯ МЕХАНИЧЕСКИХ СИСТЕМ

Чернышев Н.Н.¹

¹ ГВУЗ «Донецкий национальный технический университет», Донецк

Email: cnm@donntu.edu.ua

Copyright © 2014 by author and the journal "Automation technological and business - processes".

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



DOI: 10.15673/2312-3125.19/2014.27343

Аннотация

В статье разработана имитационная модель механической системы с помощью пакета прикладных программ Matlab&Simcape.