

Міністерство освіти і науки України
Державний вищий навчальний заклад
«Національний гірничий університет»

ЗАТВЕРДЖЕНО:
завідувач кафедри

програмного забезпечення комп'ютерних систем
(повна назва)

І.М. Удовик
(прізвище, ініціали)

(підпис)

« » 20 року

ЗАВДАННЯ

на виконання кваліфікаційної роботи магістра

спеціальності 122 Комп'ютерні науки
(код і назва спеціальності)

студенту 122М-16-1 Марчук М.В.
(група) (прізвище та ініціали)

Тема дипломної роботи Удосконалення інформаційної системи експертизи бензину на основі нечіткої моделі даних

1 ПІДСТАВИ ДЛЯ ПРОВЕДЕННЯ РОБОТИ

Наказ ректора Державного ВНЗ «НГУ» від 26.12.2017 р. № 2127 -л

2 МЕТА ТА ВИХІДНІ ДАНІ ДЛЯ ПРОВЕДЕННЯ РОБІТ

Об'єкт досліджень – процес контролю якості та ідентифікації зразків продукції.

Предмет досліджень – моделі представлення даних і методи ідентифікаційної експертизи, зокрема ІЕ виробників бензину.

Мета НДР – підвищення достовірності методів ідентифікаційної експертизи зразків продукції за рахунок розробки моделей, методів і програмних засобів, а в цілому інформаційної технології ІЕ.

Вихідні дані для проведення роботи – теоретичні й експериментальні дослідження, основи нечіткої моделі даних при вирішенні задач ідентифікаційної експертизи.

3 ОЧІКУВАНІ НАУКОВІ РЕЗУЛЬТАТИ

Актуальність даної теми зумовлена наявністю недоліків у сучасних методів ідентифікаційної експертизи: людський фактор, потреба у постійному оновленні бази даних зразків, дороге обладнання.

Наукова новизна результатів дипломної роботи визначається тим, що вперше розроблено і обґрунтовано нечітку модель представлення та обробки даних для задачі ІЕ, що дозволило ефективно виконувати класифікацію зразків продукції на основі моделювання когнітивних процесів експертизи та накопичених знань експерта.

Практична цінність результатів полягає у тому, що запропоновані в роботі моделі і методи дозволяють накопичувати та використовувати знання експерта для вирішення задач ідентифікаційної експертизи товарів.

4 ВИМОГИ ДО РЕЗУЛЬТАТІВ ВИКОНАННЯ РОБОТИ

Результати магістерської роботи повинні відповідати вимогам паспорту наукової спеціальності 05.13.06 – «Інформаційні технології».

Результати досліджень мають бути подані у вигляді, що дозволяє побачити та оцінити безпосереднє використання нечіткої моделі даних. Згідно виробничих функцій та професійних задач магістра, повинен бути розроблений програмний комплекс для вирішення задачі ідентифікаційної експертизи на основі нечіткої моделі даних.

5 ЕТАПИ ВИКОНАННЯ РОБІТ

Найменування етапів робіт	Строки виконання робіт (початок-кінець)
1	2
Аналіз теми та постановка задачі	15.09.2017 30.09.2017
Побудова нечіткої моделі представлення даних для вирішення задачі ідентифікаційної експертизи	01.10.2017 31.10.2017
Створення автоматизованої системи для вирішення задачі ідентифікаційної експертизи бензинів	01.11.2017 15.12.2017

6 РЕАЛІЗАЦІЯ РЕЗУЛЬТАТІВ ТА ЕФЕКТИВНІСТЬ

Економічний ефект від реалізації результатів роботи очікується позитивним завдяки скорочення затрат на заробітну плату людям, які виконують ідентифікаційну експертизу.

Соціальний ефект від реалізації результатів роботи очікується позитивним завдяки удосконаленню метода ідентифікаційної експертизи, що дозволяє зменшити похибку розпізнавання і зменшити загальний час виконання ІЕ бензину.

7 ДОДАТКОВІ ВИМОГИ

Відповідність оформлення ДСТУ 3008-95. Документація. Звіти у сфері науки і техніки. Структура і правила оформлення.

Завдання видав

_____ (підпис)

Корнієнко В.І.

_____ (прізвище, ініціали)

Завдання прийняв до виконання

_____ (підпис)

Марчук М.В.

_____ (прізвище, ініціали)

Дата видачі завдання: 09.09.2017р.

Термін подання дипломного проекту до ДЕК 26.01.2018

Реферат

Пояснительная записка: 85 с., 16 рис., 3 прил., 50 источников.

Объект исследования: процесс контроля качества и идентификации образцов продукции.

Цель магистерской работы: повышение достоверности методов идентификационной экспертизы образцов продукции за счет разработки моделей, методов и программных средств.

Методы исследования. Для решения поставленных задач использованы методы: анализа данных, теории распознавания образов из области вычислительного интеллекта, теории нечетких множеств, объектно ориентированного программирования.

Научная новизна полученных результатов дипломной работы определяется тем, что впервые разработана и обоснована нечеткая модель представления и обработки данных для задачи ИЭ, что позволило эффективно выполнять классификацию образцов продукции на основе моделирования когнитивных процессов экспертизы и накопленных знаний эксперта.

Практическое значение работы заключается в том, что предложенные в работе модели и методы позволяют накапливать и использовать знания эксперта для решения задач идентификационной экспертизы товаров.

Область применения. Разработанная информационная система может применяться для решения задач идентификационной экспертизы, в частности, для идентификационной экспертизы бензинов.

Значение работы и выводы. Усовершенствованная методика позволяет проектировать информационные системы со значительным сокращением как материальных затрат, так и временных, что подтверждается разработанным программным продуктом в данной магистерской работе.

Прогнозы по развитию исследований. Улучшить информационную систему, добавив метод кластеризации классов образов продуктов, с целью уменьшения процента погрешности при идентификации продуктов.

В разделе «Экономика» проведены расчеты трудоемкости разработки программного обеспечения, расходов на создание ПО и длительности его разработки, а также проведены маркетинговые исследования рынка сбыта созданного программного продукта.

Список ключевых слов: ИДЕНТИФИКАЦИОННАЯ ЭКСПЕРТИЗА, НЕФТЕПРОДУКТЫ, БЕНЗИН, C++, C#, XML, НЕЧЕТКИЙ ПОРТРЕТ, НЕЧЕТКАЯ МОДЕЛЬ, ИНФОРМАЦИОННАЯ СИСТЕМА.

Реферат

Пояснювальна записка: 85 с., 16 рис., 3 додатків., 50 джерел.

Об'єкт дослідження: процес контролю якості та ідентифікації зразків продукції.

Мета магістерської роботи: підвищення достовірності методів ідентифікаційної експертизи зразків продукції за рахунок розробки моделей, методів і програмних засобів.

Методи дослідження. Для вирішення поставлених задач використані методи: аналізу даних, теорії розпізнавання образів з області обчислювального інтелекту, теорії нечітких множин, об'єктно-орієнтоване програмування.

Наукова новизна отриманих результатів дипломної роботи визначається тим, що вперше розроблена й обґрунтована нечітка модель представлення та обробки даних для задачі ІЕ, що дозволило ефективно виконувати класифікацію зразків продукції на основі моделювання когнітивних процесів експертизи та накопичених знань експерта.

Практична цінність результатів полягає в тому, що запропоновані в роботі моделі і методи дозволяють накопичувати і використовувати знання експерта для вирішення задач ідентифікаційної експертизи товарів.

Область застосування. Розроблена інформаційна система може застосовуватися для вирішення задач ідентифікаційної експертизи, зокрема, для ідентифікаційної експертизи бензинів.

Значення роботи та висновки. Удосконалена методика дозволяє проектувати інформаційні системи зі значним скороченням як матеріальних витрат, так і тимчасових, що підтверджується розробленим програмним продуктом в даній магістерській роботі.

Прогнози щодо розвитку досліджень. Покращити інформаційну систему, додавши метод кластеризації класів образів продуктів, з метою зменшення відсотка похибки при ідентифікації продуктів.

У розділі «Економіка» проведені розрахунки трудомісткості розробки програмного забезпечення, витрат на створення ПЗ й тривалості його розробки, а також провести маркетингові дослідження ринку збуту створеного програмного продукту.

Список ключових слів: ІДЕНТИФІКАЦІЙНА ЕКСПЕРТИЗА, НАФТОПРОДУКТИ, БЕНЗИН, C++, C#, XML, НЕЧІТКИЙ ПОРТРЕТ, НЕЧІТКА МОДЕЛЬ, ІНФОРМАЦІЙНА СИСТЕМА.

The abstract

Explanatory note: 85 p., 16 fig., 3 applications, 50 sources.

Object of research: process of quality control and identification of product samples.

The purpose of the degree project: increase methods reliability of identification examination of product samples due to the development of models, methods and software.

Methods of research. To solve the problems, I used next methods: data analysis, the theory of pattern recognition from the field of computational intelligence, fuzzy sets theory, object-oriented programming.

The scientific novelty of the results is determined by the fact that the fuzzy model of data representation and processing for the IE problem was first developed and justified, which made it possible to effectively classify product samples on the basis of modeling the cognitive expertise processes and the accumulated expert knowledge.

The practical value of work lies in the fact that the models and methods proposed in the work allow the accumulation and use of the expert's knowledge for solving the identification examination problems of goods.

The scope. The developed information system can be used to solve the tasks of the identification expertise, in particular, for the identification expertise of gasolines.

The value of the work and conclusions. An improved methodology allows you to design information systems with a significant reduction in both material costs and time, which is confirmed by the developed software product in this master's work.

Projections on development research. Improve the information system by adding a method for clustering product image classes to reduce the percentage of error in identifying products.

In section "Economics" calculated the complexity of software development, the cost of creating the software and the duration of its development, and marketing studies market created by the software.

List of keywords: IDENTIFICATION EXPERTISE, PETROLEUM PRODUCTS, GASOLINE, C++, C#, XML, FUZZY PORTRAIT, FUZZY MODEL, INFORMATION SYSTEM.

Зміст

ВСТУП.....	9
РОЗДІЛ 1. АНАЛІЗ ТЕМИ ТА ПОСТАНОВКА ЗАДАЧІ.....	12
1.1. Задача ідентифікаційної експертизи.....	12
1.2. Ідентифікаційна експертиза бензину.....	18
1.3. Розпізнавання образів, як спосіб вирішення задачі ідентифікаційної експертизи.....	20
1.4. Застосування автоматизації при вирішенні задачі ідентифікаційної експертизи.....	22
1.5. Існуючі рішення задачі ідентифікаційної експертизи.....	24
1.6. Висновки.....	25
РОЗДІЛ 2. ПОБУДОВА НЕЧІТКОЇ МОДЕЛІ ПРЕДСТАВЛЕННЯ ДАНИХ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ ІДЕНТИФІКАЦІЙНОЇ ЕКСПЕРТИЗИ.....	27
2.1. Основи нечіткої логіки.....	27
2.1.1 Основні терміни нечітких множин.....	27
2.1.2 Визначення нечіткої і лінгвістичної змінних.....	30
2.2. Формальна модель нечіткого портрета.....	33
2.3. Функції приналежності.....	36
2.3.1 Метод ковзного вікна.....	36
2.3.2 Трикутна і трапецевидна функції приналежності.....	39
2.3.3 Z-подібні та S-подібні функції приналежності.....	40
2.4. Метод нечіткого виведення.....	45
2.5. Застосування нечіткої моделі для ідентифікації виробників бензину.....	47
2.6. Висновки.....	51
РОЗДІЛ 3. СТВОРЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ ІДЕНТИФІКАЦІЙНОЇ ЕКСПЕРТИЗИ БЕНЗИНІВ.....	52
3. 1. Структура автоматизованої системи ідентифікаційної експертизи бензинів.....	52
3. 2. Структура та опис бази даних.....	54
3. 3. Підсистема ідентифікації виробника палива.....	58
3. 4. Архітектура модуля ідентифікаційної експертизи.....	59
3. 5. Опис автоматизованої системи ідентифікаційної експертизи.....	61
3. 6. Висновки.....	62

РОЗДІЛ 4. ЕКОНОМІКА.....	63
4. 1. Визначення трудомісткості розробки програмного забезпечення.....	63
4. 2. Витрати на створення програмного забезпечення.....	66
4. 3. Маркетингові дослідження ринку збуту розробленого програмного продукту.....	67
4. 4. Оцінка економічної ефективності впровадження програмного забезпечення.....	69
ВИСНОВКИ.....	71
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	72
Додаток А. КОД ПРОГРАМИ.....	77
Додаток Б. ВІДГУК КЕРІВНИКА.....	84
Додаток В. РЕЦЕНЗІЯ.....	85

ВСТУП

Актуальність теми. Ідентифікаційна експертиза (ІЕ) виконується під час здійснення контролю якості зразків продукції та спрямована на боротьбу з їх фальсифікаціями. Реалізація фальсифікованої продукції призводить до значних економічних збитків і проблем безпеки життєдіяльності, тому актуальним завданням є боротьба з такими фальсифікаціями. Зокрема це стосується контролю якості бензину, який через особливості технологічних процесів виробництва і транспортування, а також інші фактори часто піддається фальсифікаціям.

Використання інформаційних технологій у галузі ІЕ набуває у світовій практиці все більшого розвитку, оскільки математичні та програмні засоби дозволяють інтенсифікувати і здешевити виконання ідентифікаційної експертизи. ІЕ товарів полягає в оцінюванні приналежності товару до конкретної класифікаційної групи. У процесі виконання ідентифікаційної експертизи виділяють такі важливі етапи: накопичення даних, аналіз даних, розробка і практичне використання методів ідентифікації товару. Сучасні інтелектуальні інформаційні технології дозволяють удосконалювати процес накопичення та обробки даних і замінювати роль експерта на етапі аналізу даних та побудови алгоритму ідентифікації.

Останнім часом все більш актуальними стають методи нечіткої логіки, які дозволяють моделювати лінгвістичні описи подій і об'єктів, процеси прийняття рішень, невизначеність у важко формалізованих системах. Нечіткі моделі успішно використовуються в моделюванні когнітивних процесів, до яких можна віднести процеси роздуму експертів. Таким чином, ідея дослідження полягає у впровадженні таких підходів для розв'язання задачі ідентифікаційної експертизи, що дозволить принципово збільшити надійність і достовірність спеціалізованих систем для її рішення.

У дипломній роботі представлені результати теоретичних та експериментальних досліджень, які полягають у розробці й обґрунтуванні

нечітких моделей і методів ідентифікації зразків продукції. Побудовано організаційну та функціональну структуру системи ідентифікаційної експертизи і на прикладі ідентифікації виробників бензинів розроблено автоматизовану систему ідентифікації бензинів.

Мета і задачі дослідження. Метою дипломного дослідження є підвищення достовірності методів ідентифікаційної експертизи зразків продукції за рахунок розробки моделей, методів і програмних засобів, а в цілому інформаційної технології ІЕ. Для досягнення мети в роботі необхідно розв'язати такі задачі:

- проаналізувати існуючі методи та інформаційні технології ідентифікаційної експертизи в галузі контролю якості продукції, у тому числі бензинів;
- розробити модель представлення даних для вирішення задачі ІЕ та методи ідентифікації зразків продукції;
- розробити методи налаштування та адаптації побудованої моделі;
- розробити автоматизовану систему ідентифікаційної експертизи (АСІЕ) на основі запропонованих моделей і методів;

Об'єктом дослідження є процес контролю якості та ідентифікації зразків продукції.

Предметом дослідження є моделі представлення даних і методи ідентифікаційної експертизи, зокрема ІЕ виробників бензину.

Методи дослідження. Для розв'язання поставлених задач використані методи: аналізу даних, теорії розпізнавання образів з області обчислювального інтелекту, теорії нечітких множин, об'єктно-орієнтованого програмування.

Наукова новизна результатів дипломної роботи визначається наступними отриманими автором теоретичними і практичними результатами:

1. Вперше розроблено і обґрунтовано нечітку модель представлення та обробки даних для задачі ІЕ, що дозволило ефективно виконувати класифікацію зразків продукції на основі моделювання когнітивних процесів експертизи та накопичених знань експерта.

2. Удосконалено метод побудови бази знань нечіткого класифікатора за рахунок нових методів автоматичної побудови функцій приналежності, що дозволило зменшити кількість нечітких правил бази знань.

Практичне значення результатів дисертаційної роботи. Запропоновані в роботі моделі і методи дозволяють накопичувати та використовувати знання експерта для вирішення задач ідентифікаційної експертизи товарів. Розроблено програмні засоби для проведення подальших досліджень з іншими предметними галузями і задачами розпізнавання образів в умовах невизначеності.

Особистий внесок магістра полягає у:

- розробці методу формування нечітких портретів;
- розробці функціональної структури нечіткої системи та методів її налаштування;
- метод формування правил.

Структура та обсяг дипломної роботи. Дана робота складається із вступу, 4 розділів, висновків, списку літератури (50 джерел) та 3 додатків. Вона містить 16 рисунків та 11 таблиць. Загальний обсяг роботи становить 85 сторінки.

РОЗДІЛ 1. АНАЛІЗ ТЕМИ ТА ПОСТАНОВКА ЗАДАЧІ

1.1. Задача ідентифікаційної експертизи

У наш час стрімко розвивається сфера виробництва. Можна відслідкувати тенденцію віддалення місць виготовлення товару від місць реалізації, що сприяє росту небезпеки фальсифікацій. Це є причиною проведення ідентифікації товару на різних етапах його обігу. Ідентифікація товарів є важливою дією при оцінці якості та встановлення їх відповідності еталонам або вимогам, які передбачаються в моральній або іншій документації. Ідентифікація - це визначення відповідності показників якості продуктів і продовольчої сировини показникам, встановленим у нормативній та технічній документації виробника продуктів чи наведеними в інформації про ці продукти, а також визначення відповідності продуктів і продовольчої сировини у звичайній загальній назві з метою сертифікації. Ідентифікація вимагає багатосторонніх досліджень товарознавцями-експертами, а також висококваліфікованими вченими-експертами. Оскільки результати ідентифікації того чи іншого виробу аналізуються і на основі цього аналізу роблять відповідні висновки, то такий процес можна назвати ідентифікаційною експертизою. Ідентифікаційна експертиза товарів проводиться для того, щоб встановити, чи належить даний товар до тієї чи іншої однорідної товарної групи. Ця ідентифікація здійснюється на основі визначення характерних індивідуальних ознак, що наведені в нормативних або супровідних документах.

Існують різні види ідентифікації в залежності від товарознавчих характеристик товарів. Їх класифікація приведена на рис. 1.1.

Асортиментна ідентифікація товарів - це встановлення тотожності і автентичності товарів їх найбільш істотним ознакам асортиментних характеристик. Даний вид ідентифікації призначений для встановлення приналежності конкретного товару до певної класифікаційної групи, виду, країни походження, торгової марки.

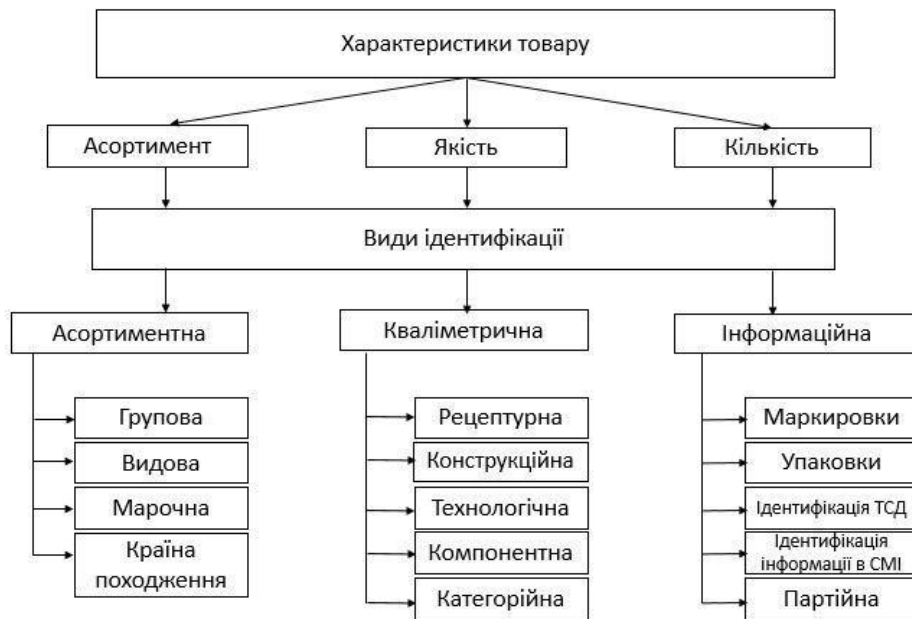


Рис. 1.1. Класифікація видів ідентифікації

Групова ідентифікація - це встановлення тотожності оцінюваного товару з товарами однорідної групи або підгрупи. В якості критеріїв ідентифікації для цього підвиду ідентифікації вибирають, найчастіше, показники функціонального призначення, сировину, а також показники хімічного складу.

Видова ідентифікація - це встановлення тотожності оцінюваного товару з товарами певного виду і підвиду. Показники та властивості видової приналежності формуються в процесі виробництва.

Марочна ідентифікація - це встановлення достовірності товару певної торгової марки і її модифікацій. В такому разі застосовують специфічні ознаки, характерні тільки для товарів конкретної торгової марки та певного виробника. Основна складність для даного виду ідентифікації в тому, що ці ознаки часто є комерційною таємницею і недоступні контролерам, експертам та іншим зацікавленим особам.

Ідентифікація країни походження товару - встановлення його приналежності до товарів, вироблених в певній країні. В якості ознак можна використовувати інформацію, зазначену в супровідних документах. Але ці

відомості можуть бути сфальсифіковані, тому по кожному виду і найменуванню товару для експерта важлива інформація по специфічних показниках, за якими можна виконати ідентифікацію.

Кваліметрична ідентифікація товарів - визначення тотожності їх споживчих властивостей і показників якості встановленим вимогам нормативних документів, та описами або зразкам.

На споживчі властивості і показники істотно впливають наступні чинники: компонентний склад сировини, рецептура, конструкція, технологічні процеси. Залежно від цих критеріїв розрізняють такі підвиди кваліметричної ідентифікації: рецептурна, конструкційна, технологічна, компонентна, категорійна.

Рецептурна ідентифікація - встановлення відповідності фактичного і певного рецептурного інгредієнту або хімічного складу. Хімічний склад багатоконпонентних харчових продуктів і тютюнових виробів залежить від кількісного і якісного складу інгредієнтів, що визначається рецептурою. Ідентифікуючою ознакою може стати масова частка кількісно переважаючих речовин, найбільш цінних інгредієнтів або речовин, які переходять з низькоякісної сировини. Наприклад, для цукристих кондитерських виробів це може бути масова частка цукру, для борошняних кондитерських виробів - цукор і жир, для варених ковбас - масова частка води, жирів і білків. Наявність в варених ковбасах вищого сорту крохмалю і грубих волокон сполучної тканини може бути ознакою використання м'ясної сировини нижчих сортів.

Конструкційна ідентифікація - встановлення тотожності конструктивних особливостей товарів вимогам, що регламентуються в технологічних інструкціях або інших документах або описах. Недотримання визначених елементів конструкції може викликати погіршення рівня якості. Наприклад, спрощення конструкції тари для зниження її маси може викликати зниження її надійності. При конструктивній ідентифікації може бути виявлена заміна дорогих високоякісних комплектуючих виробів на більш дешеві і низькоякісні. Цей підвид кваліметричної ідентифікації характерний в основному для видів

непродовольчих товарів і упаковки. Однак він може застосовуватися і для окремих продовольчих товарів. Наприклад, при оформленні тортів, а також для тютюнових виробів.

Технологічна ідентифікація - встановлення відповідності показників якості, значення яких залежать від дотримання вимог, що регламентуються технологічними інструкціями та іншими технологічними документами. Мета технологічної ідентифікації - виявлення порушень технологічного режиму виробництва і діагностики причин виникнення дефектів. Недотримання технологічного режиму викликає незначні, значні і критичні невідповідності або допустимі відхилення, які можуть побічно свідчити про ці порушення. Деякі дефекти або відхилення легко визначити органолептичним методом. Наприклад, підгоріла скоринка хліба свідчить про порушення температури при випічці.

Комплектна ідентифікація - встановлення відповідності комплекту переліку комплектуючих виробів, зазначених в експлуатаційних документах і у вкладишах, а також приналежності окремих комплектуючих виробів до конкретного комплекту. Метою такої ідентифікації є запобігання випуску та реалізації неукomплектованих товарів, що ускладнює або робить неможливим їх використання за функціональним призначенням. Для комплектної ідентифікації основним методом встановлення відповідності служить реєстраційний метод, заснований на перерахунку та обліку всіх вхідних в комплект виробів.

Компонентна ідентифікація - встановлення відповідності фактичного переліку інгредієнтів або комплектуючих виробів певного переліку, який вказаний на маркуванні або в експлуатаційних документах. Цей тип ідентифікації призначений для виявлення відсутніх окремих цінних компонентів або часткової заміни цінних компонентів на менш цінні. Крім того, при компонентній ідентифікації можуть бути виявлені невласиві або не доведені до відома споживачів та інших зацікавлених осіб компоненти. Так, для подовження термінів зберігання харчових продуктів, що швидко псуються і короткочасного зберігання (молока, пива, соків, йогуртів, варених ковбас і т. п.) додають консерванти. Однак на маркуванні відомості про їхню наявність в продуктах

відсутні, бо через негативне ставлення споживачів до них можуть бути втрачені споживчі переваги.

Категорійна ідентифікація - встановлення відповідності дійсних значень показників якості вимогам, що пред'являються до однойменних товарів певної градації якості, а також відомостями про цю градацію на маркуванні або в товаросупровідних документах.

Інформаційна ідентифікація - встановлення достовірності товарної інформації, зазначеної в товаросупровідних документах, або інших носіях інформації. Залежно від кількісних градацій і носіїв інформаційна ідентифікація може бути партійною, упаковки, маркування, товаросупровідних документів. Крім того, окремо можна виділити інформаційну ідентифікацію в засобах масової інформації, в рекламі, навчальній та науково-популярній літературі.

Ідентифікація маркування - встановлення достовірності відомостей, зазначених у маркуванні, і визначення їх відповідності інформації в товаросупровідній документації, а також зовнішнім виглядом товару.

Ідентифікація упаковки - встановлення відповідності упаковки вимогам безпеки і сумісності, що пред'являються нормативними документами, а також відомостям, зазначеним у товаросупровідних документах. Метою такої ідентифікації є запобігання якісних і кількісних втрат за рахунок використання упаковки товарів з урахуванням їх властивостей. Крім того, при ідентифікації упаковки може бути виявлено відповідність виду, розмірів і ємності упаковки вимогам, що пред'являються до кількості фасованих товарів. При ідентифікації упаковки можуть бути встановлені її сумісність і безпека для упакованих товарів, в тому числі і при тривалому зберіганні. Якщо упаковка має фіксовану масу, відхилення від цієї маси можуть перевірятися в допустимих або неприпустимих межах.

Ідентифікація товаросупровідних документів - встановлення автентичності цих документів і достовірності зазначених у них відомостей. До числа обов'язкових документів відносять товарно-транспортні накладні, посвідчення про якість, сертифікати відповідності або декларації відповідності.

Ідентифікацію товаросупровідних можуть проводити експерти під час документальної експертизи або фахівці при прийманні товарів за кількістю і якістю.

Партійна ідентифікація - діяльність з інформаційного забезпечення встановлення приналежності одиничних екземплярів товарів або сукупних пакувальних одиниць до певної товарної партії. Потреба в партійній ідентифікації виникає на етапі формування товарних партій, зберігається на всіх етапах предреалізаційної і послереалізаційної стадій і зникає після повного використання товарів на стадії утилізації.

Інформаційна ідентифікація засобів масової інформації та інших друкованих джерел - встановлення достовірності відомостей, зазначених у них, про основоположні характеристики товарів. Необхідність такої ідентифікації обумовлена тим, що в даний час в ЗМІ, а іноді і в фаховій літературі можна виявити недостовірну інформацію некомпетентних в певній галузі знань авторів.

Сформулюємо задачу ідентифікаційної експертизи в тому сенсі, в якому вона вирішується в даному дослідженні. Для товарів певного типу, які поділяються на ідентифікаційні групи в рамках відомого набору фізико-хімічних і органолептичних ознак, необхідно побудувати алгоритм, що дозволяє виконувати ідентифікацію конкретної групи для зразка, для якого така група невідома.

У загальному випадку завдання ідентифікаційної експертизи охоплюють величезну групу товарів, до якої відносяться як харчові, так і нехарчові продукти. Незважаючи на ряд загальних особливостей для виділення критеріїв ідентифікації різних груп товарів, конкретні групи товарів вимагають індивідуального підходу. Це пов'язано з тим, що різні товари відрізняються за фізико-хімічними показниками, за способом виробництва, контролю якості та ряду інших характеристик. Об'єктами ідентифікації можуть виступати як харчові продукти, наприклад апельсиновий сік, італійське вино, питна вода, так і нехарчові продукти, такі як тканини, бензини, миючі засоби та інші.

Задача визначення сорту винограду, з якого виготовлено італійське вино, представлена в сховищі завдань машинного навчання Ірвінського університету і фактично відноситься до задач ідентифікаційної експертизи. В якості ознак для вин використовують вміст спирту, фенолів, яблучної кислоти, інтенсивність кольору і деякі інші показники.

1.2. Ідентифікаційна експертиза бензину

Розглянемо особливості задачі ідентифікації виробника бензину, як однієї з показових прикладів задач ідентифікаційної експертизи.

Завдання ідентифікації виробника автомобільних бензинів виникає на підприємствах, які є проміжною ланкою між виробником автомобільного палива і АЗС (автозаправними станціями), обумовлена необхідністю боротьби з фальсифікатами і полягає в наступному. Необхідно визначити виробника і марку автомобільного бензину, зразок якого перебуває на дослідженні в лабораторії нафтопродуктів, що забезпечує контроль якості палива.

В Україні дозволені для використання автомобільні бензини відповідно до державного стандарту ДСТУ 4063-2001, в якому визначені марки А-76, А-80, А-92, А-95, А-98, а також державним стандартом на автомобільні бензини підвищеної якості ДСТУ 4839:2007, в якому визначені марки А-92-Євро, А-95-Євро, А-98-Євро.

Сучасні автомобільні бензини являють собою суміші компонентів, що отримуються в результаті переробки нафти і деяких інших технологічних процесів зі строго регламентованими фізикохімічними і експлуатаційними властивостями, і мають цілком певний хімічний склад.

Бензини відповідно до ДСТУ 4063-2001 та ДСТУ 4839:2007 повинні задовольняти вимогам і нормам, наведеним у таблиці 1.1.

Таблиця 1.1.

Фізико-хімічні показники автомобільних бензинів

Назва показника	Значення для марок ДСТУ 4839:2007			Значення для марок ДСТУ 4063-2001				
	А-92 Євро	А-95 Євро	А-98 Євро	А-76	А-80	А-92	А-95	А-98
Детонаційна стійкість:	92,0	95,0	98,0	-	80,0	92,0	95,0	98,0
- октанове число за дослідним методом, не менше ніж;								
- октанове число по моторному методу, менше ніж	82,5	85,0	88,0	76,0	76,0	82,5	85,0	88,0
Об'ємна частка вуглеводнів,%, не більше ніж:	18	18	18	-	-	-	-	-
- олефінових								
- ароматичних	35	35	35	-	-	-	-	-
Сумарний вміст ароматичних вуглеводнів, мас. %, не більш ніж	-	-	-	42	42	45	45	48
Об'ємна частка бензолу, %, не більше ніж	1,0	1,0	1,0	-	-	-	-	-
Масова частка бензолу, % мас., не більше ніж	-	-	-	5,0	5,0	5,0	5,0	5,0
Масова частка кисню, % мас., не більше ніж	2,7	2,7	2,7	2,7	2,7	2,7	2,7	2,7
Частина кисневмісних сполук, не більше ніж:	3	3	3	-	-	-	-	-
- метанол, %								
- метанол, % мас	-	-	-	3	3	3	3	3
- етанол, %	5	5	5	-	-	-	-	-
- етанол, % мас	-	-	-	5	5	5	5	5
- ізопропиловий спирт, %	10	10	10	-	-	-	-	-
- ізопропиловий спирт, % мас	-	-	-	10	10	10	10	10
- ізобутіловий спирт, %	10	10	10	-	-	-	-	-
- ізобутіловий спирт, % мас	-	-	-	10	10	10	10	10
- третбутиловий спирт, %	7	7	7	-	-	-	-	-
- третбутиловий спирт, % мас	-	-	-	7	7	7	7	7
- прості ефіри, %	15	15	15	-	-	-	-	-
- прості ефіри, % мас	-	-	-	15	15	15	15	15
інші органічні кисневмісні сполуки з температурою кінця кипіння не вище ніж 210 ° С	10	10	10	-	-	-	-	-
- об'ємна частка, %								
- масова частка, % мас	-	-	-	10	10	10	10	10

Хімічний склад бензину, який задовольняє нормативним вимогам, відрізняється в залежності від складу сировини і технології його отримання. Кожен виробник палива використовує різні компоненти при отриманні бензинів,

що дозволяє відрізнити бензини різних виробників. Різний вихідний хімічний склад сировини впливає на хімічний склад готового палива.

Зразки нафтопродуктів, які потрапляють до лабораторії контролю якості можуть бути як офіційні, які супроводжуються сертифікатом і паспортом якості, так і неофіційні, про яких найчастіше взагалі нічого невідомо. Для офіційних зразків вирішується задача контролю якості, яка в даному випадку полягає в порівнянні заявлених характеристик з отриманими експериментально. Для неофіційних зразків визначається марка палива, показники якості, а також виробник.

1.3. Розпізнавання образів, як спосіб вирішення задачі ідентифікаційної експертизи

Задача розпізнавання образів виникає в різних областях науки і техніки. Розпізнавання являє собою процес з'ясування приналежності або ідентичності невідомого об'єкта об'єктам тієї чи іншої групи. Кожному об'єкту реального світу ставиться у відповідність його модель, яка називається образом, наприклад вектор значень його ознак. Кожній групі об'єктів ставиться у відповідність узагальнений образ, тобто модель такої групи. Тому задачі розпізнавання в технічній літературі називають задачами розпізнавання образів. В задачах розпізнавання образів групу об'єктів називають класом образів, а безліч таких груп називають алфавітом класів образів. Для вирішення цієї задачі необхідно побудувати модель представлення об'єктів, визначити закономірності їх взаємодії та розробити методи прийняття рішення.

Вихідною інформацією в задачі розпізнавання є набори описів-спостережень об'єктів, які називають вибірками прецедентів. Кожне спостереження представляється вектором значень своїх ознак. Набором таких ознак є первинні дані про об'єкт, стандартизовані деяким чином.

В цілому, можна виділити три методи розпізнавання образів.

Метод перебору. У цьому випадку проводиться порівняння з базою даних, де для кожного виду об'єктів представлені різноманітні модифікації

відображення. Наприклад, для оптичного розпізнавання образів можна застосувати метод перебору виду об'єкта під різними кутами, масштабами, зміщеннями, деформаціями і т. п. Для букв потрібно перебирати шрифт, властивості шрифту і т. п. У випадку розпізнавання звукових образів, відповідно, відбувається порівняння з деякими відомими шаблонами (наприклад, слово, вимовлене кількома людьми).

Другий підхід - проводиться глибший аналіз характеристик образу. У разі оптичного розпізнавання це може бути визначення різних геометричних характеристик. Звуковий зразок в цьому випадку піддається частотному, амплітудному аналізу і т. п.

Наступний метод - використання штучних нейронних мереж. Цей метод вимагає або великої кількості прикладів задач розпізнавання при навчанні, або спеціальної структури нейронної мережі, яка враховує специфіку даного завдання. Проте, його відрізняє більш висока ефективність і продуктивність.

Проаналізувавши задачу ідентифікаційної експертизи можна встановити її прямий зв'язок із задачею розпізнавання образів. Класи образів можна поставити у відповідність ідентифікованим групам, а ознаки - показниками якості.

Вибір методу розв'язання задачі істотно залежить від предметної області, яка визначає властивості модельованих об'єктів і класів образів. Об'єктом в задачі ідентифікаційної експертизи виступає зразок продукту, отриманий відповідно до певних правил, а класами образів виступають групи продуктів, виділені у відповідність до вимог конкретної задачі.

Особливістю даної задачі є припущення, що вектори ознак деякої підмножини зразків кожної групи можуть збігатися. Тобто класи образів перетинаються. У задачі ідентифікаційної експертизи алфавіт класів образів є динамічним об'єктом. З одного боку це обумовлено тим, що на ринку з'являються нові товари, а з іншого деякі класи можуть бути не представлені в навчальній вибірці.

Результат роботи системи розпізнавання може бути представлений різними способами. Для об'єкта може бути визначено клас образів, до якого він

віднесений алгоритмом, а також може бути оцінена ймовірність, ступінь надійності або ступінь впевненості, з якою він віднесений до цього класу або кластеру. Необхідність обчислювати ймовірність або ступінь впевненості, з якою даний об'єкт відноситься до класу образів або кластеру, виникає, коли не існує однозначної відповіді, тобто класи образів перетинаються. При геометричній інтерпретації класів образів ця властивість відповідає перетину областей, що описують ці класи образів в багатовимірному просторі ознак. Такі ситуації можуть бути результатом прояву однієї з форм суб'єктивної нечіткості - неоднозначністю, розмитістю, неясністю, властивих, зокрема, завданням ідентифікаційної експертизи.

1.4. Застосування автоматизації при вирішенні задачі ідентифікаційної експертизи

Автоматизація діяльності людини з появою персональних комп'ютерів і розвиток систем управління базами даних дало поштовх розвитку автоматизованих систем зберігання та аналізу даних. У ракурсі досліджуваної проблеми системи управління базами даних можуть бути застосовані для зберігання даних про зразки товарів, а інформація по кожному випробуванню може систематизовано додаватися в базу даних, а в наслідку аналізуватися експертами при виконанні ідентифікаційної експертизи. Автоматизація процесу прийняття рішення в лабораторіях контролю якості товарів має безсумнівні переваги: зменшення часу виконання експертизи, зниження помилок, обумовлених людським фактором, створення мобільних лабораторій, оснащених приладами і системою ідентифікаційної експертизи, що дозволяє приймати рішення не тільки експертам, а й особам, що не володіє знаннями експерта, наприклад лаборантам.

В області розробки інтелектуальних комп'ютерних систем, що автоматизують процеси, пов'язані з розумовою діяльністю людини, ведеться велика кількість досліджень. Під інтелектуальною системою розуміють такі системи, які здатні навчатися, проводити логічний аналіз, синтез, здійснювати

асоціативний пошук і т.п. Виділяють два підходи при реалізації інтелектуальних систем. Перший підхід передбачає розробку програмно-апаратного комплексу, в якому тільки частина функцій реалізується методами штучного інтелекту, а сам процес інтелектуалізації системи не є основною метою. Другий підхід передбачає розробку інтелектуальної системи для вирішення деякої проблеми, причому інтелектуалізація виступає головною метою. Інтелектуальні системи можуть бути статичними та динамічними. Динамічні системи в спеціальній літературі називають системами реального часу, причому в залежності від жорсткості обмежень на швидкодію і методи зміни моделі подання знань, існує кілька їх різновидів: системи псевдореального часу, м'якого реального часу і жорсткого реального часу. Системи псевдореального часу на відміну від статичних систем отримують і обробляють дані, що надходять від зовнішніх джерел і вирішують задачу швидше, ніж відбуваються значні зміни інформації про навколишній світ. У системах м'якого реального часу допустимо, щоб час реакції на події становив більше однієї десятої або однієї секунди. До систем жорсткого реального часу ставляться більш жорсткі вимоги, вони повинні забезпечувати час реакції швидше однієї десятої секунди.

При розробці інтелектуальних інформаційних систем обробки великих обсягів промислових і біологічних даних, автоматизації металургійного, фармацевтичного, авіаційного, космічного виробництва і енергетики виникають все більш складні завдання пошуку, оптимізації і машинного навчання, які не вдалося вирішити традиційними методами штучного інтелекту і математичної статистики. Для подолання помилок традиційних методів штучного інтелекту були застосовані нечітка логіка, еволюційні і генетичні алгоритми, імунні мережі, штучні нейронні мережі. Ці методи в сукупності отримали назву м'які обчислення.

Системи нечіткого виводу були запропоновані і практично застосовані для задач управління складно формалізуємими процесами. В основі методів нечіткого управління лежить теорія нечітких множин, яка свій подальший розвиток отримала при вирішенні задач розпізнавання образів. Створити модель

для задач розпізнавання образів зі складно формалізуємою структурою класів образів дозволяють нечіткі моделі, в основі яких лежить теорія нечітких множин.

Для задачі ідентифікаційної експертизи пропонується розробити інформаційну систему з елементами обчислювального інтелекту, засновану на нечітких моделях, яка буде працювати в режимі м'якого реального часу.

1.5. Існуючі рішення задачі ідентифікаційної експертизи

Для вирішення задачі ідентифікаційної експертизи товарів використовують комплекс методів, які в кінцевому підсумку повинні забезпечити надійність і достовірність її результатів. Ці методи поділяють на три групи: органолептичні, вимірювальні, інформаційно-аналітичні, при цьому остання група заснована на аналізі товарної інформації і результатів випробувань із застосуванням органолептичних і вимірювальних методів.

Процес ідентифікації складається з трьох етапів: підготовчого, основного та заключного, кожен з яких вимагає використання своїх методів.

Підготовчий етап складається з документального та маркувального методів. Документальний метод полягає в перехресній перевірці інформації про товар у всіх супровідних документах. Маркувальний метод полягає в аналізі маркувальної інформації товару. Документальний і маркувальний методи застосовуються спільно і є обов'язковими при всіх видах ідентифікації.

Основний етап включає методи визначення значень показників та органолептичний метод. Методи визначення значень показників бувають різними. Для кожного методу розробляються як універсальні прилади, так і прилади вузької спрямованості, наприклад вимірювачі октанового числа. Органолептичні методи засновані на визначенні зовнішнього вигляду, кольору, смаку, запаху і ряду інших параметрів, які оцінюються за допомогою органів почуттів.

Заключний етап включає лише один метод – аналітико-інформаційний. Сутність аналітико-інформаційного методу полягає у відборі, узагальненні, систематизації інформації, отриманої на підготовчому і основному етапах

ідентифікації, в результаті чого отримують нову інформацію про відповідність зразка відомим ідентифікаційним групам, або про причини невідповідності таких.

Щодо існуючих рішень, то мені вдалось знайти декілька програмних комплексів. Найкращий з них, на мою думку, IDENT – система ідентифікації, розроблена в Росії, проте для її стабільної роботи необхідне постійне оновлення банка даних. Програмний комплекс не є серійним продуктом, а адаптується під конкретні цілі та умови роботи замовника. Недоліком системи є її орієнтованість на дорогий прилад - ІК-спектрометр, який використовують для оцінки основних показників якості палива.

1.6. Висновки

Предметом дослідження даної роботи є процес ідентифікаційної експертизи товару. В результаті аналізу спеціальної літератури були встановлені завдання, які необхідно вирішувати при розробці методів ідентифікації товарів. До них відноситься пошук найбільш достовірних критеріїв ідентифікації і вдосконалення експрес-методів ідентифікації, що забезпечують високий ступінь достовірності ідентифікаційної експертизи.

В основу процесу ідентифікаційної експертизи товару покладено процес аналізу показників якості. Такими є фізикохімічні показники, представлені в чисельному вигляді. Для обробки даних такого типу в області інформаційних технологій розвиваються методи аналізу даних. Сучасний рівень розвитку інформаційних технологій і комп'ютерної техніки дозволяє за рахунок їх застосування в задачах ідентифікаційної експертизи удосконалити існуючі методи ідентифікації і розробляти нові, а також підвищити достовірність такої експертизи.

В результаті аналізу процесу ідентифікаційної експертизи була сформульована конкретна постановка задачі ІЕ. Для товарів певного типу, які поділяються на ідентифікаційні групи в рамках відомого набору фізико-хімічних і органолептичних ознак, необхідно побудувати алгоритм, що дозволяє

виконувати ідентифікацію конкретної групи для зразка, для якого така група невідома. Встановлено, що задача ідентифікаційної експертизи в даній постановці може бути представлена як задача розпізнавання образів.

Для кожного конкретного товару необхідний індивідуальний підхід з урахуванням специфіки задачі, а загальним для усіх видів товару є зведення задачі групової ідентифікації до задачі розпізнавання образів. В якості методу вирішення необхідний алгоритм розпізнавання, що накладає свої обмеження на використовувані підходи.

Аналіз показав, що найкраще для вирішення поставлених задач застосувати методи обчислювального інтелекту, зокрема методи нечіткої класифікації, базовими структурами для яких є нечіткі множини.

Застосування інформаційних технологій при розробці системи ідентифікаційної експертизи в остаточному підсумку вплине на якість реалізованих в країні товарів, що позначиться на рівні життя людей. Однак, існує ряд проблем в зв'язку з розробкою таких технологій, які передбачається вирішити в даному дослідженні.

Щоб підсумувати, сформулюємо основні завдання, вирішення яких пропонується в даній роботі:

- дослідити методи нечіткої класифікації з метою їх використання для вирішення задачі ідентифікаційної експертизи;
- побудувати модель представлення даних для задачі ідентифікаційної експертизи та розробити методи ідентифікації моделі;
- розробити відкриту класифікаційну систему для вирішення задачі ідентифікаційної експертизи;
- на основі запропонованих моделей і методів розробити інформаційну систему ідентифікаційної експертизи виробника бензину.

РОЗДІЛ 2. ПОБУДОВА НЕЧІТКОЇ МОДЕЛІ ПРЕДСТАВЛЕННЯ ДАНИХ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ ІДЕНТИФІКАЦІЙНОЇ ЕКСПЕРТИЗИ

2.1. Основи нечіткої логіки

Напевно, найбільш вражаючою властивістю людського інтелекту є здатність приймати правильні рішення в умовах неповної і нечіткої інформації. Побудова моделей наближених роздумів людини і використання їх в комп'ютерних системах представляє сьогодні одну з найважливіших проблем науки.

Основи нечіткої логіки були закладені в кінці 60-х років в роботах відомого американського математика Латфі Заде. Дослідження такого роду було викликано зростаючим невдоволенням експертними системами. Хвалений "штучний інтелект", який легко справлявся із завданнями управління складними технічними комплексами, був безпорадним при найпростіших висловлюваннях повсякденного життя, наприклад "Якщо в машині перед тобою сидить недосвідчений водій - тримайся від неї подалі". Для створення дійсно інтелектуальних систем, здатних адекватно взаємодіяти з людиною, був необхідний новий математичний апарат, який переводить неоднозначні життєві твердження в мову чітких і формальних математичних формул. Першим серйозним кроком у цьому напрямку стала теорія нечітких множин, розроблена Заде. Його робота "Fuzzy Sets", опублікована в 1965 році в журналі "Information and Control", заклала основи моделювання інтелектуальної діяльності людини і стала початковим поштовхом до розвитку нової математичної теорії. Він же дав і назву для нової галузі науки - "fuzzy logic" (fuzzy - нечіткий, розмитий, м'який).

2.1.1 Основні терміни нечітких множин

Нечітка множина (fuzzy set) являє собою сукупність елементів довільної природи, щодо яких не можна з повною впевненістю стверджувати - чи належить той чи інший елемент розглянутої сукупності даній множині чи ні. Іншими

словами, нечітка множина відрізняється від звичайної тим, що для всіх або частини її елементів не існує однозначної відповіді на питання: "Чи належить той чи інший елемент розглянутій нечіткій множині?"

Формально нечітка множина визначається як безліч впорядкованих пар або кортежів виду: $\langle x, \mu_{\tilde{A}}(x) \rangle$, де $x \in X$ елементом деякої універсальної множини або універсуму X , а $\mu_{\tilde{A}}(x)$ - функція приналежності, яка ставить у відповідність кожному з елементів $x \in X$ деяке дійсне число з інтервалу $[0,1]$, тобто дана функція визначається в формі відображення:

$$\mu_{\tilde{A}}: X \rightarrow [0, 1] \quad (2.1)$$

При цьому значення $\mu_{\tilde{A}}(x) = 1$ для деякого $x \in X$ означає, що елемент x безумовно належить нечіткій множині \tilde{A} , а значення $\mu_{\tilde{A}}(x) = 0$ означає, що елемент x безумовно не належить нечіткій множині.

З усіх нечітких множин виділимо два окремих випадки, які по суті збігаються зі своїми класичними аналогами і використовуються в подальшому при визначенні інших нечітких понять. В теорії нечітких множин зберігають свій сенс деякі спеціальні класичні множини.

Порожня нечітка множина – множина, яка не містить жодного елемента. Формально визначається як нечітка множина, функція приналежності якої тотожно дорівнює нулю для всіх без винятку елементів: $\mu_{\emptyset} = 0$.

Універсум – множина, яка містить в рамках деякого контексту всі можливі елементи. Формально зручно вважати, що функція приналежності універсуму як нечіткої множини тотожно дорівнює одиниці для всіх без винятку елементів: $\mu_X = 1$.

Нечіткі множини можуть бути задані двома основними способами:

1. У формі списку з явним переліком всіх елементів і відповідним їм значень функції приналежності, що утворюють розглядаему нечітку множину. При цьому часто елементи з нульовими значеннями функції приналежності просто не вказуються в даному списку. Цей спосіб підходить коли задана нечітка множина з кінцевим дискретним носієм і невеликим числом елементів. В цьому випадку нечітку множину зручно записати у вигляді:

$\tilde{A} = \{ \langle x_1, \mu_{\tilde{A}}(x_1) \rangle, \langle x_2, \mu_{\tilde{A}}(x_2) \rangle, \dots, \langle x_n, \mu_{\tilde{A}}(x_n) \rangle \}$, де n розглядаємо число елементів нечіткої множини.

Наприклад, візьмемо в якості універсуму $X = \{1, 2, 3, \dots\}$ - безліч натуральних чисел. Тоді нечітка множина \tilde{A} , що представляє в деякому контексті "невелике натуральне число", можна задати наступним чином: $\tilde{A} = \{ \langle 1, 1.0 \rangle, \langle 2, 1.0 \rangle, \langle 3, 0.9 \rangle, \langle 4, 0.8 \rangle, \langle 5, 0.6 \rangle, \langle 6, 0.5 \rangle, \langle 7, 0.4 \rangle, \langle 8, 0.2 \rangle, \langle 9, 0.1 \rangle \}$. При цьому елементи, для яких $\mu_{\tilde{A}}(x) = 0$, відсутні в цьому списку.

2. Аналітично в формі математичного виразу для відповідної функції приналежності. Цей спосіб може бути використаний для того, щоб задати довільну нечітку множину. В цьому випадку нечітку множину зручно записувати у вигляді: $\tilde{A} = \{ \langle x, \mu_{\tilde{A}}(x) \rangle \}$, де $\mu_{\tilde{A}}$ - деяка функція, задана аналітично у формі математичного виразу $f(x)$ або графічно у формі деякої кривої.

Для формальної строгості при завданні нечітких множин необхідно явно вказувати відповідний універсум X елементів, з яких формується та чи інша конкретна нечітка множина. У загальному випадку ніяких припущень щодо елементів цієї множини не робиться. Однак з практичної точки зору доцільно обмежити універсум елементами розглянутої предметної області або розв'язуваної задачі. Оскільки при побудові нечітких моделей систем використовуються кількісні змінні, то найбільш часто в якості універсуму X використовується деяка підмножина дійсних чисел R , наприклад, множину невід'ємних дійсних чисел R^+ або натуральних чисел N .

Розглянемо приклад, який пов'язаний з розпізнаванням букв деякого алфавіту і десяткових цифр, що є досить актуальним завданням при скануванні текстових документів. Припустимо, є деяке графічне зображення, на якому представлені деякі буква і цифра (рис. 2.1).

Перше зображення породжує на множині всіх прописних літер (наприклад, українського) алфавіту $X = \{A, Б, В, \dots, Я\}$ деяку кінцеву нечітку множину $\tilde{C} = \{ \langle A, \mu_{\tilde{C}}(A) \rangle, \langle Б, \mu_{\tilde{C}}(Б) \rangle, \dots, \langle Я, \mu_{\tilde{C}}(Я) \rangle \}$. Ця нечітка множина змістовно описує відповідність зображення, представленого на рис. 2.1 (а), тій чи

іншій літері українського алфавіту. Такою множиною може бути, наприклад, наступна нечітка множина:

$$\tilde{C} = \{ \langle A, 0 \rangle, \dots, \langle И, 1.0 \rangle, \langle I, 0 \rangle, \langle \dot{I}, 0 \rangle, \langle Й, 0.9 \rangle, \langle K, 0.4 \rangle, \langle Л, 0 \rangle, \dots, \langle Я, 0 \rangle \}.$$

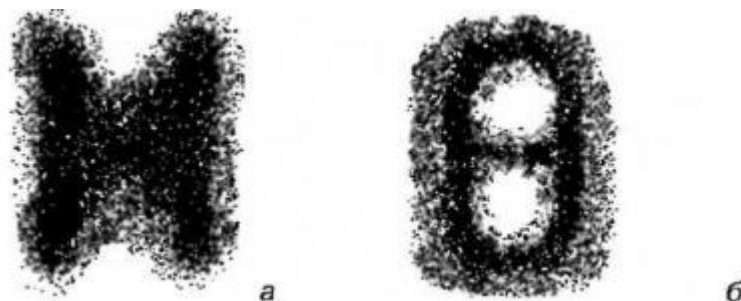


Рис. 2.1. Графічне зображення деякої літери (а) і деякої десяткової цифри (б)

Друге зображення породжує на множині всіх десяткових цифр $X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ кінцеву нечітку множину

$\tilde{D} = \{ \langle 0, \mu_{\tilde{D}}(0) \rangle, \langle 1, \mu_{\tilde{D}}(1) \rangle, \dots, \langle 9, \mu_{\tilde{D}}(9) \rangle \}$. Ця нечітка множина змістовно описує відповідність зображення, представленого на рис. 2.1 (б), тій чи іншій десятковій цифрі. В окремому випадку такою нечіткою множиною може бути, наприклад, наступна:

$$\tilde{D} = \{ \langle 0, 0.8 \rangle, \langle 1, 0 \rangle, \langle 2, 0 \rangle, \langle 3, 0.9 \rangle, \langle 4, 0 \rangle, \dots, \langle 8, 1.0 \rangle, \langle 9, 0.9 \rangle \}.$$

2.1.2 Визначення нечіткої і лінгвістичної змінних

Розглянуте раніше поняття нечіткої множини допускає різні уточнення, які доцільно використовувати для більш адекватного відображення семантики невизначеності при побудові нечітких моделей складних систем. Одним з таких уточнень є поняття лінгвістичної змінної, яке широко використовується в нечіткому управлінні для представлення вхідних і вихідних змінних керованої системи.

Нечітка змінна визначається як кортеж: $\langle a, X, \tilde{A} \rangle$, де a - найменування або назва нечіткої змінної; X - область її визначення (універсум); $\tilde{A} = \{ \langle x, \mu_{\tilde{A}}(x) \rangle \}$ - нечітка множина на X , що описує можливі значення, які може приймати нечітка змінна a . Таким чином, кажучи про нечітку змінну a , ми

завжди будемо мати на увазі деяку нечітку множину A , яка визначає її можливі значення.

Як приклад нечіткої змінної можна взяти нечітку множину B , яка характеризує "гарячу каву". У цьому випадку відповідна нечітка змінна може бути представлена наступним чином: $\langle \text{Гаряча кава}, \{x \mid 0^\circ\text{C} < x < 100^\circ\text{C}\}, B \rangle$, де $\tilde{B} = \{ \langle x, \mu_{\tilde{B}}(x) \rangle \}$ - нечітка множина з функцією приналежності $\mu_{\tilde{B}}(x)$, яка може бути задана, зокрема, графічно.

Узагальненням нечіткої змінної є так звана лінгвістична змінна. Лінгвістична змінна також визначається як кортеж: $\langle \beta, T, X, G, M \rangle$, де:

- β - найменування або назва лінгвістичної змінної;
- T - базова терм-множина лінгвістичної змінної або безліч її значень (термів), кожне з яких представляє собою найменування окремої нечіткої змінної α ;
- X - область визначення (універсум) нечітких змінних, які входять у визначення лінгвістичної змінної β ;
- G - деяка синтаксична процедура, яка описує процес утворення або генерування з множини T нових, осмислених в даному контексті значень для даної лінгвістичної змінної;
- M - семантична процедура, яка дозволяє поставити у відповідність кожному новому значенню даної лінгвістичної змінної, що отримується за допомогою процедури G , деякий осмислений зміст за допомогою формування відповідної нечіткої множини.

В якості прикладу розглянемо ситуацію зі швидкістю руху автомобільного транспорту в межах міста. Хоча правила дорожнього руху регламентують величину цієї швидкості, проте багато автолюбителів вважають за краще давати власну суб'єктивну оцінку своєї швидкості руху. При цьому використовуються такі визначення, як "мала швидкість", "середня швидкість" і "висока швидкість" руху. Очевидно, що подібна практична оцінка швидкості може ставитися до діапазону швидкостей в межах інтервалу від 0 км/год до деякої величини, яка

визначається особистими перевагами того чи іншого водія. Нехай у нашому прикладі з міркувань зручності це буде величина 100 км/ч.

Формалізація суб'єктивної оцінки швидкості руху може бути виконана за допомогою наступної лінгвістичної змінної $\langle \beta_1, T, X, G, M \rangle$, де

- β_1 - швидкість руху автомобіля;
- T - {"мала швидкість", "середня швидкість", "висока швидкість"};
- X - [0,100];
- G - процедура утворення нових термів за допомогою зв'язок логічних зв'язок "І", "АБО" і модифікаторів типу "дуже", "НЕ", "злегка" і інших. Наприклад: "мала або середня швидкість", "дуже висока швидкість" і т.п.;
- M - процедура завдання на $X = [0,100]$ нечітких змінних $\alpha_1 =$ "мала швидкість", $\alpha_2 =$ "середня швидкість", $\alpha_3 =$ "висока швидкість", а також відповідних нечітких множин для термів з $G(T)$ відповідно до правил трансляції нечітких зв'язок і модифікаторів "І", "АБО", "НІ", "дуже", "злегка".

Для розглянутого прикладу нечіткі множини A_1, A_2, A_3 , відповідні нечітким змінним: $\alpha_1 =$ "мала швидкість", $\alpha_2 =$ "середня швидкість", $\alpha_3 =$ "висока швидкість", зручно задати графічно за допомогою кусково-лінійних функцій приналежності. Один з можливих конкретних варіантів цих нечітких множин зображений на рис. 2.2. Іноді для зручності графіки функцій приналежності кількох нечітких змінних, використовуваних для завдання однієї лінгвістичної змінної, зображують на одному малюнку. Стосовно рис. 2.2 усі три графіка представлені на рис. 2.3, що дозволяє порівнювати значення функцій приналежності відповідних нечітких змінних для різних значень універсуму.

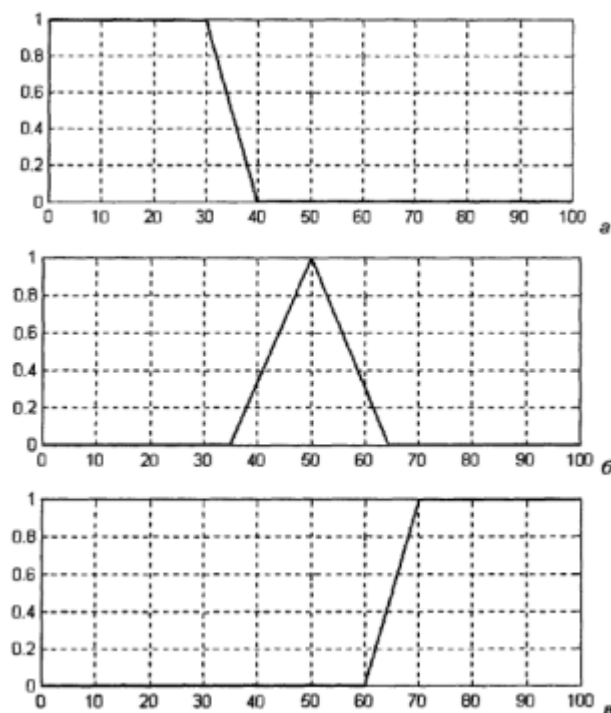


Рис. 2.2. Графіки функцій належності нечітких множин A_1, A_2, A_3 , відповідних нечітким змінним $\alpha_1 =$ "мала швидкість" (а), $\alpha_2 =$ "середня швидкість" (б), $\alpha_3 =$ "висока швидкість" (в) для лінгвістичної змінної β_1 (швидкість руху автомобіля)

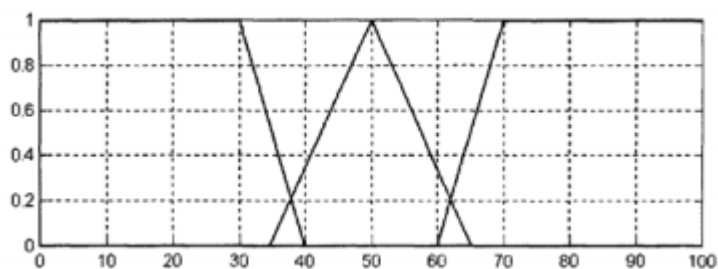


Рис. 2.3. Графіки функцій належності нечітких множин A_1, A_2, A_3 , зображені на одному малюнку

2.2. Формальна модель нечіткого портрета

Для вирішення задачі ідентифікаційної експертизи пропонується використовувати нечіткі портрети, якими представлена вхідна інформація про ідентифіковані групи. Такі портрети формуються у вигляді сукупності правил

нечітких продукцій. Факти представлені у вигляді лінгвістичних змінних, кожна з яких описує інформативну ознаку і названі відповідно до імені цієї ознаки. При побудові нечіткої моделі дотримуватимемося постановки задачі розпізнавання образів в умовах невизначеності, де об'єкт реального світу ($e \in O$) представлений у вигляді вектора інформативних ознак:

$$x = (x_1, x_2, \dots, x_m), \quad (2.2)$$

де $x_i = f_i(e)$, f_i - метод вимірювання i -го показника реального об'єкта для задачі ідентифікаційної експертизи:

$$f_i: O \rightarrow X_i, \quad (2.3)$$

де X_i - область допустимих значень ознаки, обумовлена природою досліджуваного об'єкта і методом його вимірювання, $X_i \in R$. Далі під R будемо, як завжди, розуміти безліч дійсних чисел.

Заданий алфавіт класів образів задачі розпізнавання:

$$\Omega = \{\omega_j\}_{j=1}^k, \quad (2.4)$$

де ω_j - імена класів образів, а k їх кількість.

Відомо, як виглядає кінцева множина прецедентів:

$$Z = \{(e_i, o_i)\}_{i=1}^n \quad (2.5)$$

де e_i - зразок товару, заданий вектором ознак x (2.2), а $o_i \in \Omega$ - мітка зразка.

Для представлення узагальненої інформації про класи образів будемо використовувати модель представлення даних у вигляді моделі нечіткого портрета. Дана структура використовується для представлення узагальненого образу або прототипу класу.

В ідеальній ситуації було б зручно описати кожен клас образів єдиним середнім значенням в багатовимірному просторі ознак. Для випадків, коли допускається розбиття класу на підкласи (наприклад, через сезонні зміни для задачі ідентифікації бензинів), необхідно використовувати вже кілька точок простору ознак, кожна з яких відповідає підкласу розглянутого класу образів. Однак на практиці не існує такого єдиного значення ідеального класу або підкласу. Кожній ідеальній точці відповідає безліч значень, що є кластером - в геометричній інтерпретації поняття кластера. Кожен такий кластер можна

розглядати як гранулу інформації. Поява таких інформаційних гранул обумовлена наступними причинами: похибка вимірювань, особливості технологічного процесу і людський фактор на виробництві (несумлінне ставлення, прагнення до економічної вигоди в обхід стандартів), сезонні зміни параметрів технологічного процесу, змішування зразків при зберіганні і транспортуванні.

У роботі пропонується концепція нечіткого портрета, яка враховує властивість апріорної нероздільності класів. Першочерговою метою методу є якісний опис конкретного класу образів, а не побудова розділяючого кордону між різними класами образів. Такий підхід обумовлений особливостями конкретної прикладної задачі ідентифікаційної експертизи.

Нечітким портретом S класу образів ω будемо називати пару об'єктів $\langle L, R \rangle$, де $L = \{L^g\}_{g=1}^m$ множина лінгвістичних змінних, m – їх кількість, $R = \{R^q\}_{q=1}^l$ множина правил нечітких продукцій, а l – їх кількість. Нечіткий портрет задається на множині атрибутів в m -вимірному просторі ознак.

У моделі нечіткого портрета правила нечітких продукцій записуються в наступному вигляді:

$$R^q: \text{ЯКЩО } x_1 \in A_1^q \text{ ТА } \dots \text{ ТА } x_m \in A_m^q \text{ ТО } y^q \in \omega(r^q) \quad (2.6)$$

де - x_1, x_2, \dots, x_m - інформативні ознаки задачі, відповідні компонентам вектора x , заданого в (2.2), $A_1^q, A_2^q, \dots, A_m^q$ - лінгвістичні мітки, y^q - ступінь належності об'єкта класу ω , r^q - ваговий коефіцієнт правила R^q , q - ідентифікатор правила.

Кожній інформативній ознаці товару відповідає лінгвістична змінна L_i , $i = 1, \dots, m$, де m - кількість ознак об'єкта x (2.2). Ім'я змінної L_i формулюється відповідно до назви вимірюваного параметра x_i (2.2) і обраним методом його вимірювання $f_i(e)$ (2.3), терм-множина $T_i = \{A_g^i\}_{g=1}^l$, де l - кількість правил нечітких продукцій, а універсальна множина збігається з множиною X_i з (2.3) - областю допустимих значень i -тої інформативної ознаки товару.

Нечіткою областю в m -вимірному просторі ознак назвемо m -вимірну нечітку множину. Нечіткий портрет класу образів формально описує таку нечітку область в просторі ознак задачі ідентифікаційної експертизи.

Ступінь приналежності точки простору ознак нечіткої області, заданої нечітким портретом, характеризує ступінь типовості даного об'єкта відповідному класу образів. Нечіткий портрет будується з навчальної множини Z , заданої в (2.5), причому для точок з множини Z ступінь приналежності класу чином o_i для нечіткого портрета S_i більше нуля.

Кожен нечіткий портрет окремо являє собою узагальнену інформацію про окремий клас образів, однак об'єднання нечітких портретів усіх класів образів дозволить отримати нечіткий класифікатор, який використовується для вирішення задачі ідентифікаційної експертизи.

2.3. Функції приналежності

Вибір виду функцій приналежності терм-множини лінгвістичної змінної і методу їх побудови визначає властивості нечіткого портрета, і є одним з основних питань нечіткого моделювання.

2.3.1 Метод ковзного вікна

Аналіз частоти появи значень певної ознаки для класу образів виконують методом гістограм. Необхідно з урахуванням кроку розбиття, побудувати гістограму, яку можна прямо використовувати як функцію приналежності або в разі, коли цього недостатньо, виконують ряд додаткових перетворень. У даній роботі пропонується використовувати метод ковзного вікна, який є модифікованим методом гістограм для побудови функції приналежності.

Розглянемо етап підготовки вхідних даних для формування терм-множин лінгвістичних змінних. Згрупуємо значення елементів навчальної вибірки Z , заданої в (2.5) за належністю класів образів, де $P = \{P_1, \dots, P_i, \dots, P_m\}$ - множина імен ознак, ω_j - класи образів з алфавіту класів образів (2.4). Визначимо множину

D_{ij} як множину значень ознаки P_i для класу образів ω_j : якщо $\omega_j = \{x_{j1}, x_{j2}, \dots, x_{jt}\}$, $pr_i = (x_1, x_i, \dots, x_m) = x_i$ та $pr_i\omega_j = \{pr_ix_{j1}, pr_ix_{j2}, \dots, pr_ix_{jt}\}$, то

$$D_{ij} = pr_i\omega_j \quad (2.7)$$

Кожній ознаці P_i поставимо у відповідність лінгвістичну змінну L_i , а її терми будемо будувати методом ковзного вікна.

Метод ковзного вікна для побудови функцій приналежності нечіткого портрета на вхід отримує вибірки D_{ij} , визначені у (2.7). Результатом роботи методу є нечітка множина μ_{ij} .

Так як дані досліджуваного завдання ідентифікаційної експертизи не систематизовані і кількість прикладів для кожного класу образів відрізняється, побудова гістограми пов'язана з необхідністю підбору кроку розбиття w . Щоб вирішити дану проблему при автоматичному формуванні функції приналежності, введемо додаткові параметри - крок ковзання s , та крок розбиття w , який назвемо вікном розбиття. Тоді будемо рахувати кількість елементів множини D , які потрапляють у вікно w , а вікно будемо зміщувати з кроком ковзання s і рахувати кількість елементів, що потрапили у вікно. Введемо параметри α і β , від яких буде залежати вибір кроку ковзання і ширина вікна.

Результатом роботи методу є нечітка множина MF , представлена сплайн-функцією першого порядку. Функція приналежності є кусково-лінійною функцією, а вузлові точки x_i , в яких вона визначена, обчислюються функцією $grid(D)$ по параметру s . При цьому крайові точки сітки розраховуються як $x_1 = \min x - s$ та $x_p = \max x + s$.

Далі для кожної вузлової точки виконується операція *count*, яка полягає в підрахунку кількості елементів множини D , що потрапляють у вікно $[x_i - \frac{w}{2}, x_i + \frac{w}{2}]$. Ілюстрація застосування методу ковзного вікна при побудові функції приналежності показана на рис. 2.4. Результатом роботи алгоритму є функція приналежності MF , представлена при описі алгоритму матрицею

розміру $2 \times p$, де p - кількість вузлів сітки. Далі цю функцію будемо розглядати як визначену на всій універсальній множині, будемо позначати її як $\mu_{ij}(x)$.

На форму функції приналежності μ_{ij} впливають параметри α і β . Так як на останньому етапі роботи алгоритму виконується нормування функції, $\mu_{ij} \in [0, 1]$. Відношення $\frac{\beta}{\alpha}$ визначає вид графіка і узагальнюючу здатність нечітких портретів.

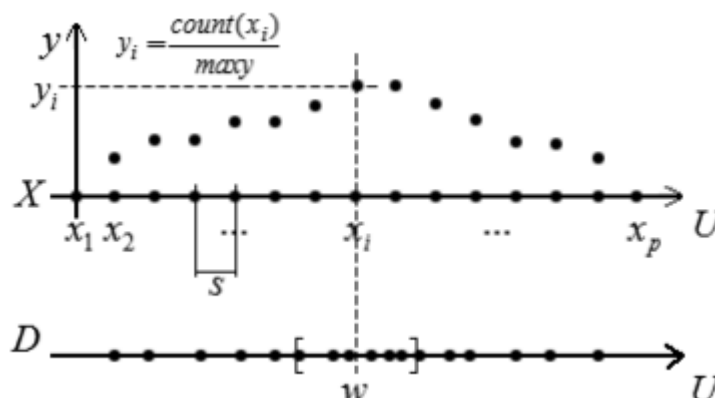


Рис. 2.4. Ілюстрація методу ковзного вікна

Відзначимо наступні особливості, що впливають на кінцеві функції приналежності нечітких портретів і властивості алгоритму розпізнавання:

- при малому числі прецедентів класу в навчальній вибірці велике число змін знаків похідної функції (узагальнююча здатність функцій погіршується, обсяг даних недостатній);
- чим менше коефіцієнт $\alpha \in (0,1]$, тим більш «жорсткими» є функції приналежності з точки зору оцінки ступенів відповідності образу класу образів;
- з ростом коефіцієнта $\beta \in [1,5]$ функції стають більш гладкими, зменшується число змін знаків їх похідних.

Отримана методом ковзного вікна функція фактично задана набором значень. Такі функції представляють великий інтерес, так як враховують всі особливості вхідних даних, що має першочергове значення для експертів.

2.3.2 Трикутна і трапецевидна функції приналежності

Розглянемо трикутну і трапецевидну функції приналежності (рис. 2.5). У нашому випадку кожна з цих функцій задана на універсумі $X = [0, 10]$, в якості якого обраний замкнутий інтервал дійсних чисел. У загальному випадку вибір універсуму може бути довільним, і не обмежений ніякими правилами.

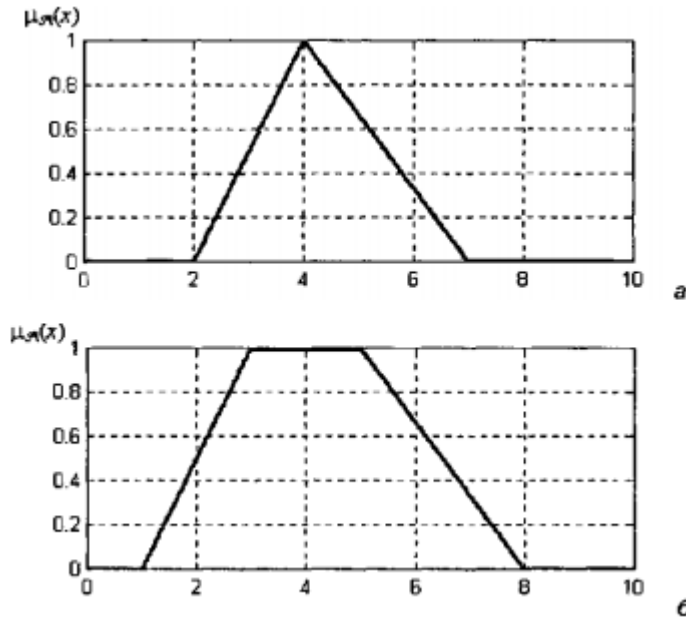


Рис. 2.5. Графіки функцій приналежності трикутної (а) і трапецієподібної (б) форми

Перша з цих функцій приналежності в загальному випадку може бути задана аналітично наступним виразом:

$$f(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{c-x}{c-b}, & b \leq c \leq x \\ 0, & c \leq x \end{cases}, \quad (2.8)$$

де a, b, c - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a \leq b \leq c$. Стосовно конкретної функції, зображеної на рис. 2.5 а, значення параметрів рівні: $a = 2, b = 4, c = 7$. Як неважко помітити, параметри a і c характеризують основу трикутника, а параметр b - його вершину.

Трапецієподібна функція приналежності в загальному випадку може бути задана аналітично наступним виразом:

$$f(x; a, b, c, d) = \left\{ \begin{array}{l} 0, \quad x \leq a \\ \frac{x-a}{b-a}, \quad a \leq x \leq b \\ 1, \quad b \leq x \leq c \\ \frac{d-x}{d-c}, \quad c \leq x \leq d \\ 0, \quad d \leq x \end{array} \right\}, \quad (2.9)$$

де a, b, c, d - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a \leq b \leq c \leq d$. Стосовно конкретної функції, зображеної на рис. 2.5 б, значення параметрів рівні: $a = 1, b = 3, c = 5, d = 8$. Як неважко помітити, параметри a і d характеризують нижню частину трапеції, а параметри b та c - верхню.

Ці функції використовуються для того, щоб задати такі властивості множин, які характеризують невизначеність типу: "приблизно дорівнює", "середнє значення", "розташований в інтервалі", "подібний до об'єкту", "схожий на предмет" та ін.

2.3.3 Z-подібні та S-подібні функції приналежності

Ці функції приналежності отримали свою назву з вигляду кривих, які представляють їх графіки. Перша з функцій цієї групи має назву Z-подібна крива або сплайн-функція і в загальному випадку може бути задана аналітично наступним чином:

$$f(x; a, b) = \left\{ \begin{array}{l} 1, \quad x \leq a \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-a}{b-a} \pi\right), \quad a \leq x \leq b \\ 0, \quad x > b \end{array} \right\}, \quad (2.10)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.6 а, при цьому значення параметрів відповідно рівні $a = 3, b = 6$. Сплайн-функція може бути також задана іншим чином:

$$f(x; a, b) = \begin{cases} 1, & x \leq a \\ 1 - 2 \left(\frac{x-a}{b-a} \right)^2, & a < x \leq \frac{a+b}{2} \\ 2 \left(\frac{b-x}{b-a} \right)^2, & \frac{a+b}{2} < x < b \\ 0, & b \leq x \end{cases}, \quad (2.11)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.6 б, при цьому значення параметрів відповідно рівні $a = 3, b = 6$.

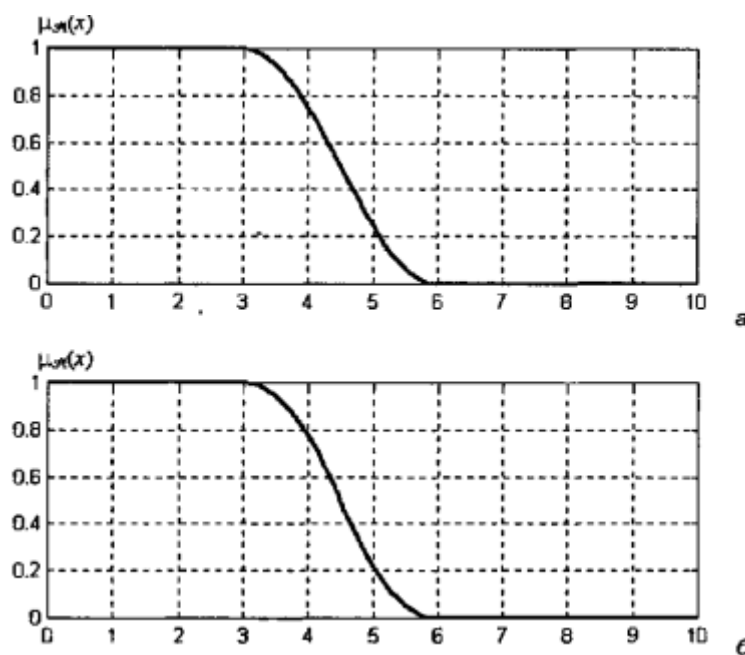


Рис. 2.6. Графіки Z-подібних функцій приналежності для значень параметрів $a = 3, b = 6$

Ці функції використовуються для представлення таких властивостей нечітких множин, які характеризуються невизначеністю типу: "мала кількість", "невелике значення", "незначна величина", "низька собівартість продукції", "низький рівень цін або доходів", "низька процентна ставка" і багатьох інших. Загальним для всіх таких ситуацій є слабка ступінь прояву тієї чи іншої якісної або кількісної ознаки. Особливість нечіткого моделювання при цьому полягає в представленні відповідних нечітких множин за допомогою незростаючих функцій приналежності.

Друга з функцій даної групи називається S -подібною кривою або сплайн-функцією і в загальному випадку може бути задана аналітично наступним чином:

$$f(x; a, b) = \begin{cases} 0, & x < a \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-b}{b-a} \pi\right), & a \leq x \leq b \\ 1, & x > b \end{cases}, \quad (2.12)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.7 а, при цьому значення параметрів відповідно рівні $a = 3, b = 6$. Сплайн-функція може бути також задана іншим чином:

$$f(x; a, b) = \begin{cases} 0, & x \leq a \\ 2 \left(\frac{x-a}{b-a}\right)^2, & a < x \leq \frac{a+b}{2} \\ 1 - 2 \left(\frac{b-x}{b-a}\right)^2, & \frac{a+b}{2} < x < b \\ 1, & b \leq x \end{cases}, \quad (2.13)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.7 б, при цьому значення параметрів відповідно рівні $a = 3, b = 6$.

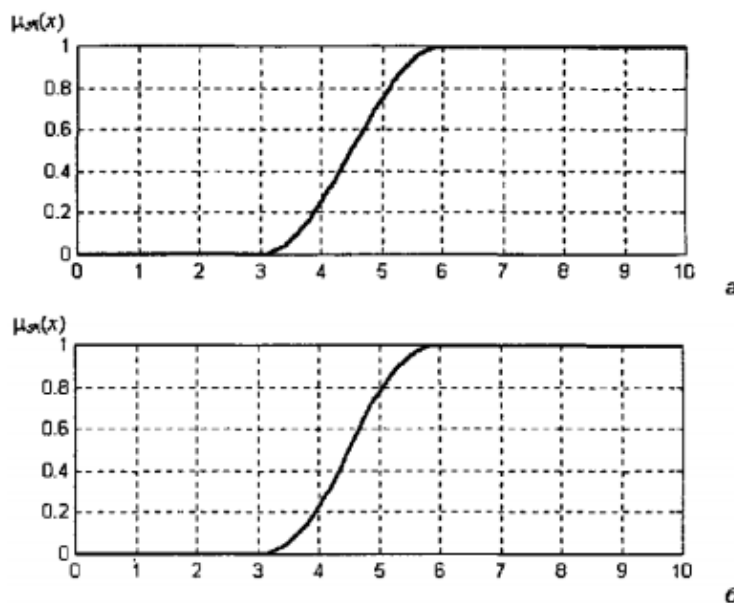


Рис. 2.7. Графіки S -подібних функцій приналежності для значень параметрів $a = 3, b = 6$

До типу S -подібних і одночасно Z -подібних функцій приналежності може бути віднесена так звана сигмоїдальна функція (сигмоїд), яка в загальному випадку задається аналітично наступним чином:

$$f(x; a, b) = \frac{1}{1 + e^{-a(x-b)}}, \quad (2.14)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$, а e - основа натуральних логарифмів, яка ініціює завдання відповідної експоненційної функції. При цьому в разі $a > 0$ може бути отримана S -подібна функція приналежності, а в разі $a < 0$ - Z -подібна функція приналежності. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.8. При цьому S -подібній функції приналежності відповідають значення параметрів $a = 3, b = 6$ (рис. 2.8 а), а Z -подібній функції приналежності відповідають значення параметрів $a = -3, b = 6$ (рис. 2.8 б).

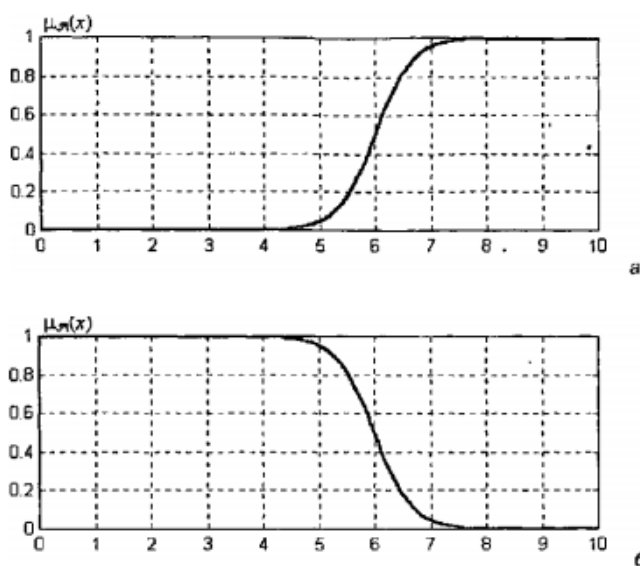


Рис. 2.8. Графіки сигмоїдальної функції приналежності для значень параметрів $a = 3, b = 6$ (а) і $a = -3, b = 6$ (б)

Розглянуті S -подібні функції використовуються для представлення таких нечітких множин, які характеризуються невизначеністю типу: "велика кількість", "велике значення", "значна величина", "високий рівень доходів і цін", "висока норма прибутку", "висока якість послуг", "високий сервіс обслуговування" і т.п. Загальним для всіх таких ситуацій є висока ступінь прояву

тієї чи іншої якісної або кількісної ознаки. Особливість нечіткого моделювання при цьому полягає в представленні відповідних нечітких множин за допомогою неспадних функцій приналежності.

Також розглянемо лінійну Z-подібну функцію (рис. 2.9 а) і лінійну S-подібну функцію (рис. 2.9 б). Перша з цих функцій в загальному випадку може бути задана аналітично наступним чином:

$$f(x; a, b) = \begin{cases} 1, & x \leq a \\ \frac{b-x}{b-a}, & a < x < b \\ 0, & b \leq x \end{cases}, \quad (2.15)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.9 а, при цьому значення параметрів відповідно рівні $a = 3, b = 6$. Друга з цих функцій в загальному випадку може бути задана аналітично наступним чином:

$$f(x; a, b) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & b \geq x \end{cases}, \quad (2.16)$$

де a, b - деякі числові параметри, які приймають довільні дійсні значення і впорядковані відношенням: $a < b$. Графік цієї функції для деякої нечіткої множини \tilde{A} і універсуму $X = [0, 10]$ зображений на рис. 2.9 б, при цьому значення параметрів відповідно рівні $a = 3, b = 6$.

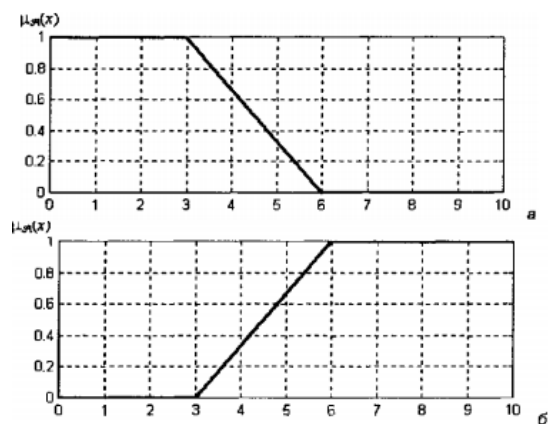


Рис. 2.9. Графіки лінійної Z-подібної функції (а) і лінійної S-подібної функції (б) приналежності для значень параметрів $a = 3, b = 6$

2.4. Метод нечіткого виведення

Говорячи про нечітку логіку, найчастіше мають на увазі системи нечіткого виведення, які широко використовуються для управління технічними пристроями і процесами.

Інформацією, яка надходить на вхід системи нечіткого виведення, є виміряні деяким чином вхідні змінні. Ці змінні відповідають реальним змінним процесу управління. Інформація, яка формується на виході системи нечіткого виведення, відповідає вихідним змінним, якими є керуючі змінні процесу управління.

Системи нечіткого виводу призначені для перетворення значень вхідних змінних процесу управління у вихідні змінні на основі використання нечітких правил продукцій. Для цього системи нечіткого виводу повинні містити базу правил нечітких продукцій і реалізовувати нечіткий висновок висновків на основі посилок або умов, представлених у формі нечітких лінгвістичних висловлювань.

Таким чином, основними етапами нечіткого виводу є:

- Формування бази правил систем нечіткого виводу.
- Фазифікація вхідних змінних.
- Агрегування підумов в нечітких правилах продукцій.
- Активізація або композиція підумов в нечітких правилах продукцій.
- Акумуляування висновків нечітких правил продукцій

База правил систем нечіткого виводу призначена для формального подання емпіричних знань або знань експертів в тій чи іншій проблемній області. У системах нечіткого виведення використовуються правила нечітких продукцій, в яких умови і укладення сформульовані в термінах нечітких лінгвістичних висловлювань розглянутих вище видів. Сукупність таких правил будемо далі називати базами правил нечітких продукцій.

В контексті нечіткої логіки під фазифікацією розуміють не тільки окремий етап виконання нечіткого виведення, а й власне процес знаходження значень

функцій приналежності нечітких множин (термів) на основі звичайних вхідних даних. Фазифікацію ще називають введенням нечіткості.

Метою етапу фазифікації є встановлення відповідності між конкретним значенням окремої вхідної змінної системи нечіткого виведення і значенням функції приналежності відповідного їй терма вхідної лінгвістичної змінної. Після завершення цього етапу для всіх вхідних змінних повинні бути визначені конкретні значення функцій приналежності по кожному з лінгвістичних термів, які використовуються в підумові бази правил системи нечіткого виведення.

Формально процедура фазифікації виконується наступним чином. До початку цього етапу передбачаються відомими конкретні значення всіх вхідних змінних системи нечіткого виведення, тобто, множина значень $V' = \{a_1, a_2, \dots, a_m\}$. У загальному випадку кожне $a_i \in X_i$, де X_i - універсум лінгвістичної змінної β_i . Ці значення можуть бути отримані або від датчиків, або деяким іншим, зовнішнім по відношенню до системи нечіткого виведення способом. Далі розглядається кожна з підумов виду " $\beta_i \in \alpha$ " правил системи нечіткого виведення, де α - деякий терм з відомою функцією приналежності $\mu(x)$. При цьому значення a_i використовується як аргумент $\mu(x)$, тим самим знаходимо кількісне значення $b'_i = \mu(a_i)$. Це значення і є результатом фазифікації підумови " $\beta_i \in \alpha$ ".

Агрегування - процедура визначення ступеня істинності умов по кожному з правил системи нечіткого виведення.

Формально процедура агрегування виконується наступним чином. До початку цього етапу передбачаються відомими значення істинності всіх підумов системи нечіткого виведення, тобто, множина значень $B = \{b'_i\}$. Далі розглядається кожна з умов правил системи нечіткого виведення. Якщо умова правила - нечітке висловлювання, то ступінь його істинності дорівнює відповідному значенню b'_i . Якщо ж умова складається з декількох підумов, причому лінгвістичні змінні в підумовах попарно не рівні один одному, то визначається ступінь істинності складного висловлювання на основі відомих значень істинності підумови.

Активізація в системах нечіткого виведення являє собою процедуру або процес знаходження ступеня істинності кожного з підумов правил нечітких продукцій. Активізація в загальному випадку багато в чому аналогічна композиції нечітких відносин, але не тотожна їй. Оскільки в системах нечіткого виведення використовуються лінгвістичні змінні, то формули для нечіткої композиції втрачають своє значення. Насправді при формуванні бази правил системи нечіткого виведення задаються вагові коефіцієнти F_i для кожного правила (передбачається, якщо ваговий коефіцієнт не заданий явно, то його значення дорівнює 1).

Формально процедура активізації виконується наступним чином. До початку цього етапу передбачаються, що відомі значення істинності всіх умов системи нечіткого виведення, тобто, множина значень $B = \{b_i''\}$ і значення вагових коефіцієнтів F_i для кожного правила. Далі розглядається кожне з висновків правил системи нечіткого виведення. Ступінь істинності правила дорівнює алгебраїчному добутку відповідного значення b_i'' на ваговий коефіцієнт F_i .

Акумуляція в системах нечіткого виведення являє собою процедуру або процес знаходження функції приналежності для кожної з вихідних лінгвістичних змінних множини $W = \{w_1, w_2, \dots, a_s\}$.

Мета акумуляції полягає в тому, щоб об'єднати або акумулювати всі ступені істинності висновків для отримання функції приналежності кожної з вихідних змінних. Причина необхідності виконання цього етапу полягає в тому, що підумови, що відносяться до однієї і тієї ж вихідної лінгвістичної змінної, належать різним правилам системи нечіткого виведення.

2.5. Застосування нечіткої моделі для ідентифікації виробників бензину

Будемо говорити про групу бензинів певної марки певного виробника як про клас образів, представлений множиною своїх представників - зразків бензину. Відома ієрархічна структура груп стандартних бензинів, представлена

на рис. 2.10. Другому рівню відповідають марки бензинів, а саме А-76, А-80, А-92, А-95 і А-98. На наступному рівні для кожної марки бензину перераховані виробники. Закодуємо назви виробників бензину двома латинськими символами: SH, PR, MZ, OR, LN.

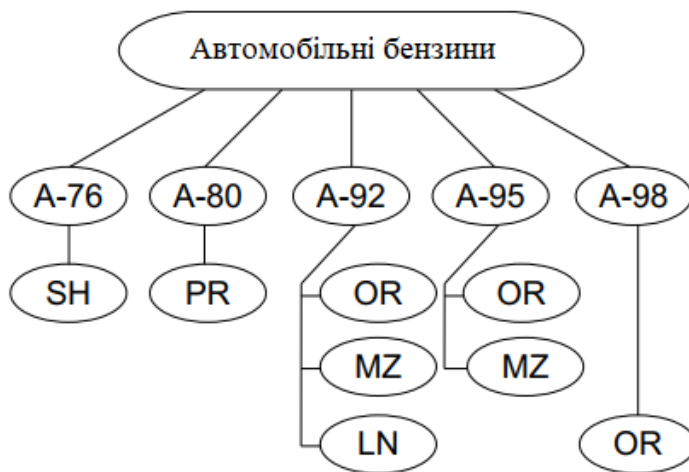


Рис. 2.10. Класифікація бензинів в задачі ідентифікації виробника

Для кожної гілки дерева запишемо послідовно назви її вузлів, крім кореневого вузла: $\{(A-76, SH), (A-80, PR), (A-92, OR), (A-92, MZ), (A-92, LN), (A-95, OR), (A-95, MZ), (A-98, OR)\}$. Для кожної такої пари задамо нову мітку, наприклад для $(A-76, SH) \rightarrow A76SH$. Таким чином, на підставі структури груп стандартних бензинів, представлені у вигляді дерева, сформуємо алфавіт класів образів: $\Omega = \{A76SH, A80PR, A92OR, A92MZ, A92LN, A95OR, A95MZ, A98OR\}$.

Виконаємо структурну ідентифікацію для даної задачі ідентифікаційної експертизи. Зазвичай в задачах ідентифікаційної експертизи в якості характеристик товару розглядають органолептичні і фізико-хімічні показники. При цьому, одна і та ж характеристика може бути виміряна різними приладами, включаючи різні екземпляри приладів одного виду. Діапазон допустимих значень ознак визначається властивостями об'єкта. Вибір набору характеристик і методу їх вимірювань виконує експерт предметної області, враховуючи при цьому наявність в лабораторії вимірювальних засобів, вартість таких вимірів, в тому числі вартість витрачених матеріалів, інформацію за нормативними

документами та ряд інших факторів. В даному випадку перелік доступних ознак визначено експертом лабораторії і приведений в таблиці 2.1.

Таблиця 2.1.

Фізико-хімічні показники зразка бензину

№	Назва показника	Одиниця виміру	Позначення	Допустимий діапазон значень
x_1	Октанове число за дослідним методом	% Ізооктану в його суміші з <i>n</i> -гептаном, при якому ця суміш еквівалентна по детонаційної стійкості досліджуваного палива в стандартних умовах дослідження	RON	[60-100]
x_2	Октанове число за моторним методом	%, так само як для октанового числа по дослідному методу	MON	[50-90]
x_3	Вміст олефінів	Об'ємні відсотки	OLF	[0-25]
x_4	Вміст алканів	Об'ємні відсотки	ALK	[20-80]
x_5	Сумарний вміст ароматичних вуглеводнів	Об'ємні відсотки	AROMA	[0-50]
x_6	Вміст бензолу	Об'ємні відсотки	AR_BNZ	[0-30]
x_7	Вміст ксилолів	Об'ємні відсотки	AR_KS	[0-50]
x_8	Вміст толуолу	Об'ємні відсотки	AR_TL	[0-50]
x_9	Вміст кисню	Масові відсотки	MOX	[0-15]
x_{10}	Метанол	Масові відсотки	MMETANOL	[0-40]
x_{11}	Етанол	Масові відсотки	METANOL	[0-40]
x_{12}	Метил-трет-бутиловий ефір	Масові відсотки	MMTBE	[0-20]
x_{13}	Етил-трет-бутиловий ефір	Масові відсотки	METBE	[0-20]
x_{14}	ТАМ ефір	Масові відсотки	MTAME	[0-20]
x_{15}	ДІП ефір	Масові відсотки	MDIPE	[0-20]
x_{16}	Трет-бутиловий спирт	Масові відсотки	MTBS	[0-20]

Кожен зразок бензину заданий інформативним вектором значень ознак $x = (x_1, \dots, x_m)$ виду (2.2), де m - кількість використовуваних інформативних ознак. Розмірність вектора інформативних ознак в даному випадку $m = 16$. У першій колонці таблиці вказані компоненти вектора x , далі вказано назву ознаки, одиницю його виміру, умовне позначення, що використовується далі в роботі, а також діапазон можливих значень, $[d_{i1}, d_{i2}]$, $i = 1, \dots, m$.

Технологічні особливості процесу транспортування і зберігання партій бензину припускають наявність залишків палива від попередніх партій в ємності при додаванні в неї нового палива. В результаті відібрані зразки можуть мати відхилення від даних сертифіката знову надійшедшої партії палива внаслідок змішування. Оцінимо відхилення прогнозу приналежності стандартного класу при аналізі зразка, який є сумішшю. При змішуванні значення показника суміші дорівнює середньоарифметичній величині цих показників для частин суміші:

$$x = \frac{P_a x_a + P_b x_b}{P_a + P_b}, \quad (2.17)$$

де P_a - маса першого продукту суміші, P_b - маса другого продукту суміші, x_a і x_b - значення показника у першого і другого продуктів суміші відповідно, а x - значення показника для суміші. Таким чином існує суттєва проблема відхилення значень ознак для проби, взятої в резервуарі на АЗС. Тоді значення показників змінюються пропорційно до складу суміші відповідно до формули (2.17).

Наступною причиною різних значень зразків, отриманих від одного виробника, є їх належність до різних партій, що, в ряді випадків, призводить до виникнення підкласів для однієї групи бензинів. При зміні режимів процесу виробництва або при зміні рецептури сумішевих бензинів часто відбувається «випадання» груп зразків по їх вуглеводневому складу, що призводить до розбиття однієї групи бензинів на підгрупи, які називаються кластерами.

В результаті аналізу було виявлено особливість вибірки – зашумленість. Паливо з вагоноцистерн є найбільш достовірним зразком конкретного виробника, так як для вагоноцистерн не допускаються залишки, тому такі зразки

розглядаються як еталонні. Тому при навчанні моделі будемо використовувати тільки дані з вагоноцистрен.

2.6. Висновки

У другому розділі в якості гіпотези висувається припущення, що задача ідентифікаційної експертизи може бути вирішена на основі нечітких моделей.

Важливою перевагою нечітких класифікаторів є інтерпретованість отриманих моделей. В основі нечіткого класифікатора лежать правила нечітких продукцій. Такі правила близькі людській інтуїції. Рішення, отримане методом нечіткої класифікації, моделює рішення, яке може бути отримано особою яка приймає рішення, в нашому випадку експертом. Таким чином, застосування нечіткого підходу в задачах ідентифікаційної експертизи дозволяє моделювати когнітивні процеси експертизи.

При виконанні ідентифікаційної експертизи в умовах динамічно мінливого оточення нечіткий підхід дозволяє вирішити задачу зменшення впливу сумнівних даних, а саме викидів.

РОЗДІЛ 3. СТВОРЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ ДЛЯ ВИРІШЕННЯ ЗАДАЧІ ІДЕНТИФІКАЦІЙНОЇ ЕКСПЕРТИЗИ БЕНЗИНІВ

3.1. Структура автоматизованої системи ідентифікаційної експертизи бензинів

Користувачами системи є співробітники лабораторії контролю якості нафтопродуктів: експерти, приймальники. Система призначена для використання в лабораторії контролю якості зразків продукції.

В автоматизованій системі ідентифікаційної експертизи передбачено виконання таких функцій:

- накопичення інформації про зразки продукції в базі даних і забезпечення доступу до них;
- контроль якості зразка продукції на підставі результатів випробувань відповідно до нормативної документації;
- ідентифікація зразка продукції у відповідність з постановкою задачі.

Далі розглянемо організаційну структуру автоматизованої системи. На вхід системи інформація надходить від зовнішніх вимірювальних датчиків. Передбачено виконання експертизи як в стаціонарній лабораторії так і в мобільних лабораторіях.

Таблиця 3.1.

Властивості значень атрибутів сутності «Зразок»

«Статус»	«Виробник»	«Власник»	«Постачальник»
Стандартний	Обов'язкове поле зі списку відомих виробників	Фіксоване значення «лабораторія»	Будь-який
Скарга	Невідомий	Фіксоване значення «лабораторія»	АЗС
Експериментальний	Не встановлений	Нема обмежень	Нема обмежень
Контрольний	Невідомий	Фіксоване значення «лабораторія»	АЗС
Розвідувальний	Невідомий	Нема обмежень	АЗС

На виході системи ми отримуємо інформацію про зразок продукції у вигляді паспорта продукту або висновку експерта. Форма представлення результату залежить від вхідної інформації про зразок, на практиці визначається за властивостями сутності «Зразок» (таблиця 3.1.). Структурна схема системи представлена на рис. 3.1.

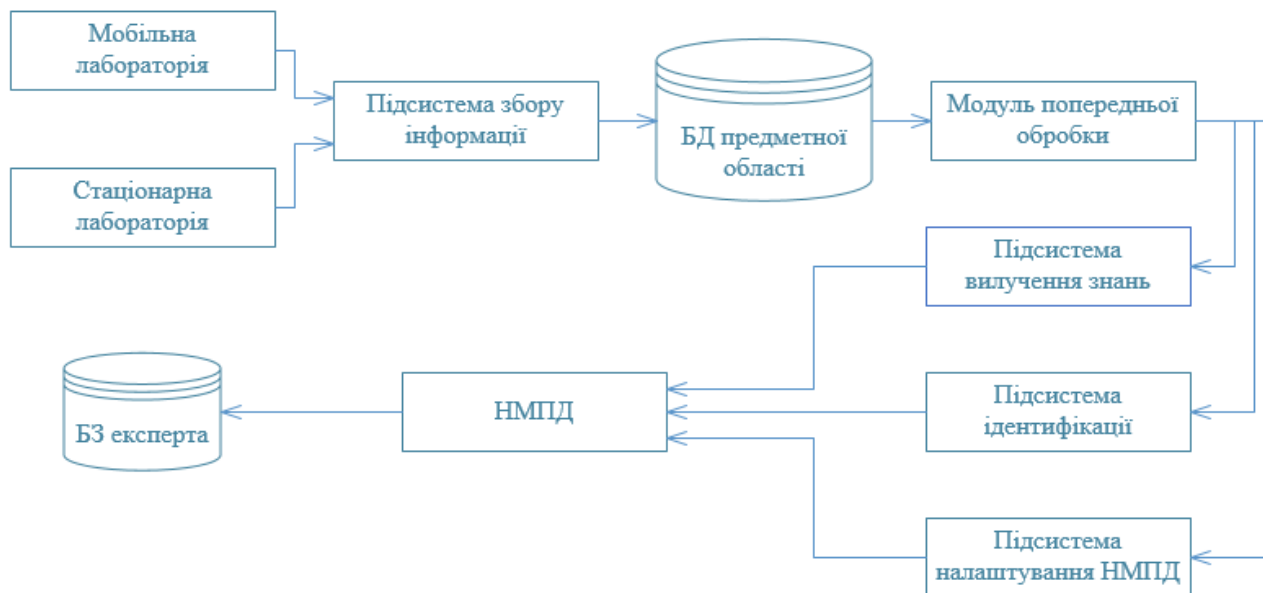


Рис. 3.1. Структурна схема АСІЕ

Розглянемо взаємозв'язок організаційної та функціональної структур. Накопиченням знань здійснює підсистема збору інформації, підсистема вилучення знань і підсистема налаштування НМПД. У базі знань, фактично, накопичуються знання експерта у вигляді нечіткої моделі, яка покладена в основу розроблюваної системи. Модуль попередньої обробки відповідає за організацію введення/виведення даних, контроль правильності вхідних значень і визначення методів подальшої обробки. Модуль адміністрування БД реалізує пошук в базі даних, пакетне вивантаження/завантаження даних, формування звітів. У підсистемі ідентифікації реалізована основна частина інформаційної технології нечіткої класифікації.

Так як основною сутністю, з якої працює система, є зразок продукції, розглянемо процес його аналізу більш детально. На приборах виконується фізико-хімічна експертиза зразка. Результатом аналізу проби є висновок

експерта, в якому або підтверджується відповідність заявленої зі зразком документації, або спростовується.

Аналіз проби виконується в три етапи: контроль даних, випробування, прийняття рішення. На першому етапі виконується оцінка відповідності зразка палива державним стандартам ДСТУ 4839:2007 або ДСТУ 4063-2001 та наділення його статусом: стандартний - нестандартний. Під випробуванням розуміється виконання експериментів на приладах, перелік яких визначений задачею. Отримані результати випробувань аналізуються на етапі прийняття рішення. Так як в розглянутій задачі класи образів перетинаються, то іноді необхідно проводити додаткові експерименти.

Ідентифікація марки бензину і виробника необхідна в трьох базових ситуаціях: вирішення сумнівів «скаржника», вирішення сумнівів оптового клієнта, прийом продукту на нафтобазу і деякі інші спірні ситуації. В кожному випадку присутні свої нюанси в прийнятті рішення, для чого в системі передбачені різні форми введення.

Знання експерта, необхідні для достовірного аналізу проби, зберігаються в базі знань, а набір даних з зовнішніх датчиків у базі даних предметної області. Розглянемо це питання детальніше.

3.2. Структура та опис бази даних

Було прийнято рішення використовувати реляційну модель бази даних для автоматизованої системи ідентифікаційної експертизи. База даних описує предметну область.

Відповідно до результатів аналізу роботи лабораторії нафтопродуктів розроблена структура БД предметної області, фізична модель якої представлена на рис. 3.2. В якості СКБД використовується SQLite, так як вона є безкоштовною, популярною і підтримується більшістю середовищ програмування. Далі проаналізуємо деякі таблиці.

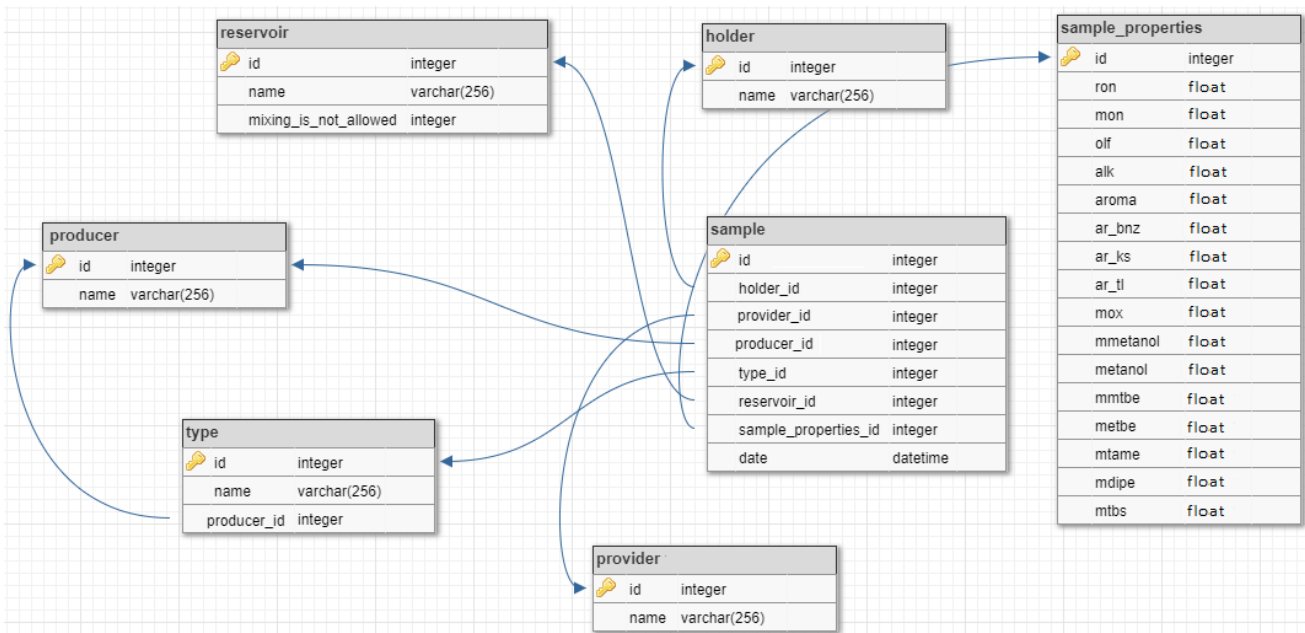


Рис. 3.2. Фізична модель бази даних предметної області

Таблиця 3.2.

Імена таблиць бази даних

Назва таблиці в БД	Сутність
holder	Власник зразка палива
provider	Постачальник палива
producer	Виробник палива
type	Марка палива
reservoir	Резервуар, з якого отриманий зразок палива
sample	Зразок палива
sample_properties	Властивості зразка палива

Таблиця 3.3.

Опис таблиці holder

Стовбець	Тип	Значення
id	int	Ідентифікатор власника
name	varchar	Ім'я власника

Таблиця 3.4.

Опис таблиці provider

Стовбець	Тип	Значення
id	int	Ідентифікатор постачальника
name	varchar	Ім'я постачальника

Таблиця 3.5.

Опис таблиці producer

Стовбець	Тип	Значення
id	int	Ідентифікатор виробника
name	varchar	Ім'я виробника

Таблиця 3.6.

Опис таблиці type

Стовбець	Тип	Значення
id	int	Ідентифікатор типу
name	varchar	Назва типу
producer_id	int	Ідентифікатор виробника

Таблиця 3.7.

Опис таблиці reservoir

Стовбець	Тип	Значення
id	int	Ідентифікатор резервуару
name	varchar	Назва резервуару
mixing_is_not_allowed	int	флаг, дорівнює 1, якщо для резервуара не допускається змішування палива різних марок від різних виробників

Таблиця 3.8.

Опис таблиці sample_properties

Стовбець	Тип	Значення
id	int	Ідентифікатор запису
ron	float	Октанове число за дослідним методом
mon	float	Октанове число за моторним методом
olf	float	Вміст олефінів
alk	float	Вміст алканів
aroma	float	Сумарний вміст ароматичних вуглеводнів
ar_bnz	float	Вміст бензолу
ar_ks	float	Вміст ксилолів
ar_tl	float	Вміст толуолу
mox	float	Вміст кисню
mmetanol	float	Метанол
metanol	float	Етанол
mmtbe	float	Метил-трет-бутиловий ефір
metbe	float	Етил-трет-бутиловий ефір
mtame	float	ТАМ ефір
mdipe	float	ДІП ефір
mtbs	float	Трет-бутиловий спирт

Таблиця 3.9.

Опис таблиці sample

Стовбець	Тип	Значення
id	int	Ідентифікатор зразка
provider_id	int	Ідентифікатор постачальника
producer_id	int	Ідентифікатор виробника
type_id	int	Ідентифікатор типу
reservoir_id	int	Ідентифікатор резервуару

sample_properties_id	int	Ідентифікатор властивостей
date	datetime	Дата надходження на експертизу

3.3. Підсистема ідентифікації виробника палива

У даній підсистемі здійснюється настройка НМПД, аналіз нових даних і повторне навчання моделі за результатами їх аналізу, а також ідентифікація виробника на основі побудованої моделі.

НМПД описується набором правил нечітких продукцій і множиною лінгвістичних змінних. Так як такі моделі добре структуровані, їх можна описати в стандартному вигляді. Для опису нечіткого класифікатора було вибрано XML-формат. XML-формат передбачає наявність ієрархічної структури і добре вирішує проблему представлення складних моделей даних в програмних системах. Для нечіткого класифікатора може бути запропонована наступна ієрархія: <Класифікатор> <Правила> <Лінгвістичні змінні>. XML-формат являє собою звід загальних синтаксичних правил, призначених для організації, зберігання структурованих даних і застосовується для організації обміну даними між програмами. Запропоновано новий підтип XML-формату, який називається «Fuzzy Classifier System», і має розширення fcs.

На рис. 3.3. наведено фрагмент fcs-документа, що описує нечіткий класифікатор для системи, що складається з п'яти правил і шести ознак. Кореневий елемент «fcs» містить два обов'язкових вузла: «model» і «fuzzyrules». Всередині структури «model» містяться загальні відомості про класифікатор: кількість класів (вузол «CN»), кількість атрибутів (вузол «AN»), імена класів образів (перераховані в дочірніх вузлах «class» вузла «classes»), імена атрибутів (перераховані в дочірніх вузлах «feature» вузла «features»), тип методу нечіткого виводу (вузли «MFtype», «AND», «OR») а також кількість правил в базі правил (вузол «RN»). Всередині структури «fuzzyrules» перераховані всі правила «rule».

Атрибути дозволяють описати деяку додаткову інформацію: дату створення («date»), назва («id», «name»), версію («version») та інші важливі для

моделі деталі. Використання XML-формату для зберігання моделі даних розпізнає дає ряд переваг:

- високий рівень інтерпретованості результатів налаштування моделі користувачем (перегляд моделі стандартними програмами для перегляду текстових даних);
- забезпечення взаємодії між різними модулями системи;
- забезпечення взаємодії між різними модулями системи;
- використання побудованої моделі в інших програмних системах, наприклад, при побудові експертних систем.

```
1 <fcs version="1.0" id="test" date="10.09.2017">
2   <model>
3     <CN description="number of classes">5</CN>
4     <classes>
5       <class id="1">A92LN</class>
6       <class id="2">A95LN</class>
7       <class id="3">A95OR</class>
8       <class id="4">A80PR</class>
9       <class id="5">A76SH</class>
10    </classes>
11    <AN des="number of features">6</AN>
12    <features>
13      <feature id="1" type="integer" begin="0" end="100">OLF</feature>
14      <feature id="2" type="integer" begin="0" end="100">AROMA</feature>
15      <feature id="3" type="integer" begin="0" end="100">AR_BNZ</feature>
16      <feature id="4" type="integer" begin="0" end="100">MTBE</feature>
17      <feature id="5" type="integer" begin="0" end="100">RON</feature>
18      <feature id="6" type="integer" begin="0" end="100">MON</feature>
19    </features>
20    <RN des="number of rules">5</RN>
21    <AND>2</AND>
22    <OR>0</OR>
23    <MFtype>MFDiscrete</MFtype>
24  </model>
25  <fuzzyrule>
26    <rule consequent="1" weight="1" type="MFDiscrete">
27      <fs feature="OLF">
28        <size>63</size>
29        <x>12,89;13,20;13,32;13,45;13,57;13,69;13,82;13,94;14,06;14,19;14,31;14,43;14,56;14,68;14,80;
30 14,93;15,05;15,17;15,30;15,42;15,54;15,67;15,79;15,91;16,04;16,16;16,28;16,41;16,53;16,65;16,78;16,90;
31 17,02;17,15;17,27;17,39;17,52;17,64;17,76;17,89;18,01;18,13;18,26;18,38;18,50;18,63;18,75;18,87;19,00;
32 19,12;19,24;19,37;19,49;19,61;19,74;19,86;19,98;20,11;20,23;20,35;20,48;20,60;21,03;</x>
33    <y>0,00;0,40;0,40;0,43;0,48;0,50;0,48;0,50;0,50;0,52;0,52;0,48;0,50;0,45;0,43;0,40;0,43;0,50
34 ;0,48;0,48;0,48;0,55;0,52;0,48;0,60;0,74;0,69;0,62;0,64;0,79;0,79;0,86;0,86;1,00;0,98;0,98;0,90;0,86;0,81;
```

Рис. 3.3. Приклад запису лінгвістичних змінних і правил нечіткої класифікації в форматі fcs

3. 4. Архітектура модуля ідентифікаційної експертизи

Даний модуль розроблений для реалізації запропонованих в дипломній роботі моделей і методів для вирішення задачі розпізнавання образів. Розроблено об'єктну модель нечіткого класифікатора. Бібліотека написана на мові C ++.

Важливою властивістю даної бібліотеки класів є її незалежність від предметної області. Бібліотека призначена для розробки автоматизованих систем розпізнавання. На рис. 3.4. можна побачити спрощену діаграму основних класів модуля.

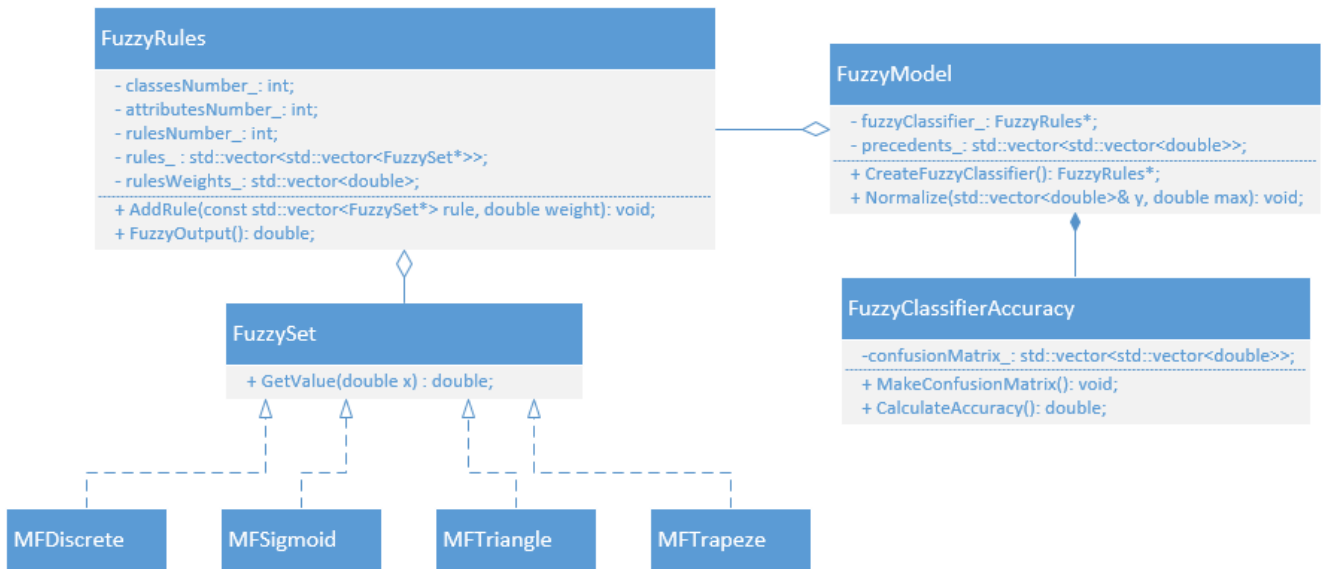


Рис. 3.4. Діаграма класів модуля ідентифікаційної експертизи

Для підтримки різних типів функцій приналежності використовується абстрактний клас FuzzySet. Клас FuzzyModel відповідає за настройку, повторне навчання та використання нечіткого класифікатора, а клас FuzzyClassifierAccuracy відповідає за контроль якості класифікатора, використовується робота з нечіткими інформаційними векторами. Екземпляри класу FuzzyRules використовуються для подання нечіткого класифікатора, причому кожне правило представлено в табличному вигляді, а лінгвістичні змінні задані в неявному вид через посилання на екземпляри класу FuzzySet.

Клас FuzzyModel призначений для формування моделі нечіткого класифікатора на основі множини прецедентів. Клас FuzzyRules представляє нечіткий класифікатор, представлений набором правил. Результат класифікації може бути представлений у вигляді інформативного вектора ступеней впевненості α . Віртуальний клас FuzzySet є батьківським класом для класів, що описують функції приналежності різних типів. Даний клас забезпечує незалежну роботу класифікатора з функціями приналежності будь-якого типу.

Класи MFDiscrete, MFSigmoid, MFTriangle, MFTrapeze реалізують сплайн першого порядку, сигмоїдальну, трикутну і трапецієдальну функції приналежності відповідно. Клас FuzzyClassifierAccuracy призначений для оцінки якості побудованого на основі моделі типу FuzzyModel класифікатора типу FuzzyRules. Оцінка відбувається порівнянням правильного результату з результатом, отриманим від нечіткого класифікатора.

3.5. Опис автоматизованої системи ідентифікаційної експертизи

Автоматизована система ідентифікаційної експертизи являє собою комп'ютерну систему, програмні компоненти якої реалізовані на мовах C++ (логіка програми) та C# (графічний інтерфейс). Для роботи з даними використовується система керування базами даних (СКБД) SQLite.

Для роботи з нечіткою моделлю розроблена бібліотека класів, яка називається Бібліотекою для нечіткого розпізнавання (Fuzzy Pattern Recognition Library).

Для зв'язку між компонентами системи використовується представлення нечіткого класифікатора в xml-форматі.

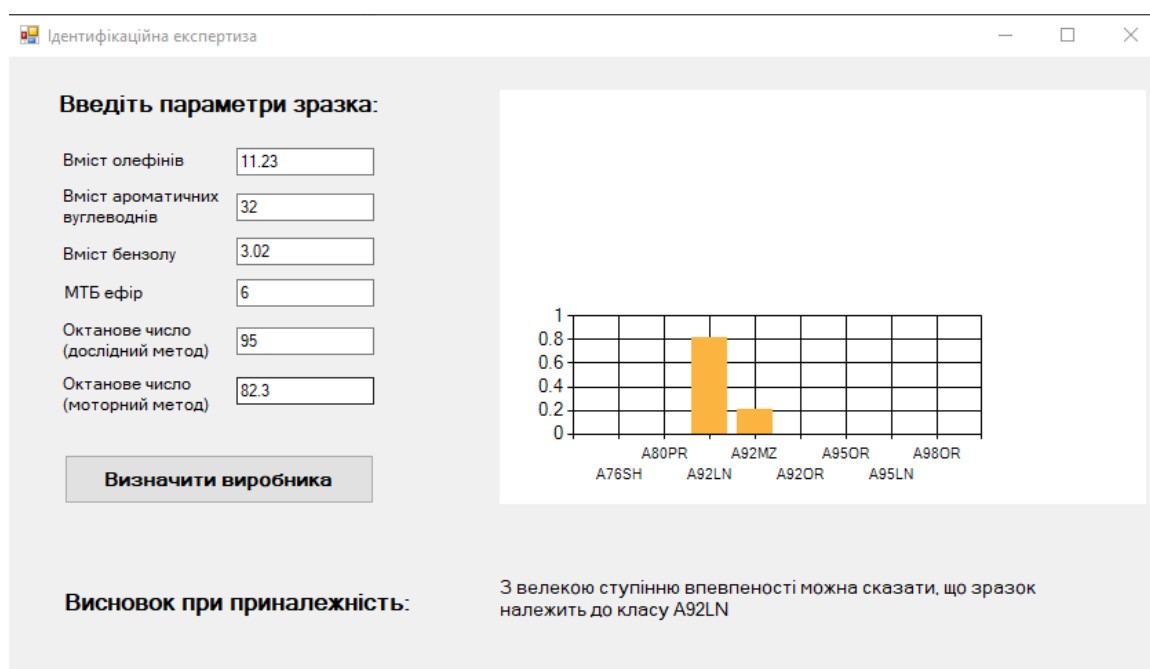


Рис. 3.5. Інтерфейс програми ідентифікаційної експертизи бензинів

Для адаптації нечіткої моделі представлення даних програмно реалізовані нечітка модель і методи її налаштування, розроблений зручний візуальний інтерфейс за допомогою мови програмування C# (рис. 3.5.).

Так як результатом роботи класифікатора є нечіткий інформаційний вектор, то для оцінки достовірності роботи системи експерту видається результат в декількох формах: візуальній, вербальній і чисельній. Візуальною формою представлення є стовпчикова діаграма. Чисельна форма представлення відповідає нечіткому порогу. Вербальне уявлення результату полягає в лінгвістичній інтерпретації значень інформаційного вектора.

3. 6. Висновки

Даний розділ присвячений побудові автоматизованої системи ідентифікаційної експертизи на прикладі бензинів. Побудована організаційна та функціональна структури для автоматизованих систем «м'якого» реального часу, призначеної для вирішення задачі ідентифікаційної експертизи продукції. Крім нафтопереробної галузі, система може бути використана, наприклад, в харчовій, хімічній та інших промисловостях.

Виконано аналіз інформаційних потоків контролю якості нафтопродуктів і побудована інформаційна модель предметної області. Це дозволило виділити структуру класів образів, яка використовується як вхідні дані задачі розпізнавання образів і побудувати реляційну базу даних інформаційної системи.

В основу методу ідентифікації зразка продукції, які використовується в інформаційній системі, покладена інформаційна технологія нечіткої класифікації. Розроблено бібліотека класів образів, яка реалізує всі математичні моделі і методи, запропоновані в дипломній роботі.

РОЗДІЛ 4. ЕКОНОМІКА

При розробці програмного забезпечення важливими етапами є визначення трудомісткості розробки ПЗ, розрахунок витрат на створення програмного продукту і тривалості його розробки.

4.1. Визначення трудомісткості розробки програмного забезпечення

- 1) Передбачуване число операторів – 1127
- 2) Коефіцієнт складності програми – 1,65
- 3) Коефіцієнт корекції програми в ході її розробки – 0,07
- 4) Годинна заробітна плата програміста, грн/год – 21
- 5) Вартість машина/години – 10

Нормування праці в процесі створення програмного забезпечення істотно ускладнено в силу творчого характеру праці програміста. Тому трудомісткість розробки програмного забезпечення може бути розрахована на основі системи моделей з різною точністю оцінки.

Трудомісткість розробки програмного забезпечення можна розрахувати за формулою:

$$t = t_o + t_u + t_a + t_n + t_{отл} + t_d, \text{ людино-годин,} \quad (4.1.1)$$

де t_o – витрати праці на підготовку й опис поставленої задачі (20 людино-годин);

t_u - витрати праці на дослідження алгоритму рішення задачі;

t_a - витрати праці на розробку блок-схеми алгоритму;

t_n - витрати праці на програмування по готовій блок-схемі;

$t_{отл}$ - витрати праці на налагодження програми на ЕОМ;

t_d - витрати праці на підготовку документації.

Складові витрати праці визначаються через умовне число операторів у ПЗ, яке розробляється.

Умовне число операторів:

$$Q = q \cdot C \cdot (1 + p), \quad (4.1.2)$$

де q – передбачуване число операторів (1127);

C – коефіцієнт складності програми (1,65);

p – коефіцієнт кореляції програми в ході її розробки (0,07).

Звідси умовне число операторів у програмі:

$$Q = 1,65 \cdot 1127 \cdot (1 + 0,07) = 1990$$

Витрати праці на вивчення опису задачі t_u визначається з урахуванням уточнення опису і кваліфікації програміста:

$$t_u = \frac{Q \cdot B}{80 \cdot k}, \quad (4.1.3)$$

де B - коефіцієнт збільшення витрат праці внаслідок недостатнього опису задачі;

k - коефіцієнт кваліфікації програміста, обумовлений від стажу роботи з даної спеціальності. При стажі роботи до 2 років, він становить 0,8.

Прийmemo збільшення витрат праці внаслідок недостатнього опису завдання не більше 50% ($B = 1,4$). З урахуванням коефіцієнта кваліфікації $K = 0,8$ отримуємо витрати праці на вивчення опису задачі:

$$t_u = \frac{1990 \cdot 1,2}{80 \cdot 0,8} = 37,3 \text{ людино-годин}$$

Витрати праці на розробку алгоритму рішення задачі:

$$t_a = \frac{Q}{20 \cdot k}, \quad (4.1.4)$$

де Q - умовне число операторів в програмі;

k - коефіцієнт кваліфікації програміста.

Підставивши відповідні значення в формулу, отримаємо:

$$t_a = \frac{1990}{23 \cdot 0,8} = 124,4 \text{ людино-годин}$$

Витрати на складання програми по готовій блок-схемі:

$$t_n = \frac{Q}{25 \cdot k}, \quad (4.1.5)$$

$$t_n = \frac{1990}{25 \cdot 0,8} = 99,5 \text{ людино-годин}$$

Витрати праці на налагодження програми на ЕОМ:

$$t_{omл} = \frac{Q}{5 \cdot k}, \quad (4.1.6)$$

$$t_{omл} = \frac{1990}{5 \cdot 0,8} = 497,5 \text{ людино-годин}$$

З урахуванням коефіцієнта запасу 1.5 отримаємо

$$t_{omл}^k = 1,5 \cdot t_{omл}, \quad (4.1.7)$$

$$t_{omл}^k = 1,5 \cdot 497,5 = 746,3 \text{ людино-годин}$$

Витрати праці на підготовку документації:

$$t_{\partial} = t_{\partial p} + t_{\partial o}, \quad (4.1.8)$$

де $t_{\partial p}$ - трудомісткість підготовки матеріалів і рукопису;

$t_{\partial o}$ - трудомісткість редагування, печатки й оформлення документації.

$$t_{\partial p} = \frac{Q}{18 \cdot k}, \quad (4.1.9)$$

$$t_{\partial o} = 0,75 \cdot t_{\partial p}, \quad (4.1.10)$$

Підставляючи відповідні значення, отримаємо:

$$t_{\partial p} = \frac{1990}{18 \cdot 0,8} = 138,2 \text{ людино-годин},$$

$$t_{\partial o} = 0,75 \cdot 138,2 = 103,7 \text{ людино-годин}$$

$$t_{\partial} = 138,2 + 103,7 = 241,9 \text{ людино-годин},$$

Повертаючись до формули (4.1.1), отримаємо повну оцінку трудомісткості розробки програмного забезпечення:

$$t = 20 + 37,3 + 124,4 + 99,5 + 497,5 + 241,9 = 1020,6 \text{ людино-годин}$$

4. 2. Витрати на створення програмного забезпечення

Витрати на створення програмного забезпечення ($K_{по}$) включають витрати на заробітну плату розробників програми ($З_{зп}$) і вартість машинного часу, необхідного для налагодження програми на ЕОМ:

$$K_{по} = З_{зп} + З_{мв}, \text{ грн.} \quad (4.2.1)$$

Заробітна плата виконавців визначається за формулою:

$$З_{зп} = t \cdot C_{пр}, \text{ грн.} \quad (4.2.2)$$

де t – загальна трудомісткість, людино-годин;

$C_{пр}$ - середня годинна заробітна плата програміста, грн/година.

З урахуванням того, що середня годинна зарплата програміста становить 21 грн/год, отримуємо:

$$З_{зп} = 1020,6 \cdot 21 = 21432,6 \text{ грн}$$

Вартість машинного часу, необхідного для налагодження програми на ЕОМ:

$$З_{мв} = t_{омл} \cdot C_{мч}, \text{ грн.}, \quad (4.2.3)$$

де $t_{омл}$ - трудомісткість налагодження програми на ЕОМ, год;

$C_{мч}$ - вартість машино-години ЕОМ, грн/год.

Підставивши в формулу (4.2.3) відповідні значення, визначимо вартість необхідного для налагодження машинного часу:

$$З_{мв} = 497,5 \cdot 10 = 4975 \text{ грн.}$$

Звідси витрати на створення програмного продукту:

$$K_{по} = 21432,6 + 4975 = 26407,6 \text{ грн.}$$

Очікуваний період розробки програмного забезпечення:

$$T = \frac{t}{B_k \cdot F_p}, \text{ міс.}, \quad (4.2.4)$$

де B_k - число виконавців (дорівнює 1),

F_p - місячний фонд робочого часу (при 40 годинному робочому тижні

$F_p = 176$ годин).

Підставивши відповідні значення, отримаємо:

$$T = \frac{1020,6}{1,176} \approx 6 \text{ міс.}$$

Вартість даного програмного забезпечення становить 26,4 тис. грн. і не вимагає додаткових витрат як при розробці програми, так і при її експлуатації. Очікуваний час розробки становить 6 місяців. Цей термін пов'язаний з досить великим числом об'єктів, часом на дослідження і розробку алгоритму вирішення поставленого завдання, програмування по готовому алгоритму, налагодження програми і підготовку документації.

4.3. Маркетингові дослідження ринку збуту розробленого програмного продукту

Ідентифікаційна експертиза (ІЕ) виникає при здійсненні контролю якості товарів і спрямована на боротьбу з їх фальсифікаціями. Реалізація фальсифікованих товарів призводить до значних економічних збитків і проблем безпеки життєдіяльності, тому актуальним завданням є боротьба з такими фальсифікаціями. У тому числі це стосується завдання контролю якості бензину, який, в силу особливостей технологічних процесів виробництва і транспортування, а також ряду інших факторів, часто піддається фальсифікації.

Метою дипломної роботи є підвищення достовірності методів ідентифікаційної експертизи зразків продукції за рахунок розробки моделей, методів і програмних засобів інформаційної технології ідентифікаційної експертизи. Для досягнення мети в роботі було вирішено наступні завдання:

- аналіз існуючих методів та інформаційних технологій ідентифікаційної експертизи в галузі контролю якості продукції, в тому числі бензинів;
- розробка моделі представлення даних для вирішення завдань ІЕ і методів ідентифікації зразків продукції;
- розробка методів настройки і адаптації побудованої моделі;
- розробка автоматизованої системи ідентифікаційної експертизи (АСІЕ) на основі запропонованих моделей і методів;

- проведення аналізу ефективності запропонованих моделей і методів ІЕ, а також АСІЕ в цілому.

Наукова новизна результатів роботи визначається наступними отриманими автором теоретичними і практичними результатами:

1. Вперше розроблена й обґрунтована нечітка модель подання та обробки даних для завдання ідентифікаційної експертизи, що дозволило ефективно виконувати класифікацію зразків продукції на основі моделювання когнітивних процесів експертизи та накопичених знань експерта;
2. Удосконалено метод побудови бази знань нечіткого класифікатора за рахунок нових методів автоматичної побудови функцій приналежності, що дозволило зменшити кількість нечітких правил бази знань;
3. Отримали подальший розвиток:
 - інформаційна технологія нечіткої класифікації на основі запропонованої нечіткої моделі, яка враховує фактор невизначеності завдання ІЕ, що дозволило вирішити проблему категоричності при віднесенні зразка до ідентифікаційної групи і підвищити достовірність методів ідентифікації зразків продукції;
 - інформаційна технологія побудови автоматизованих систем «м'якого» реального часу за рахунок запропонованої організаційної та функціональної структури автоматизованої системи ідентифікаційної експертизи, що дозволило підвищити якісні і кількісні характеристики процесу ІЕ.

Щодо існуючих рішень, то мені вдалось знайти декілька програмних комплексів. Найкращий з них, на мою думку, IDENT – система ідентифікації, розроблена в Росії, проте для її стабільної роботи необхідне постійне оновлення бази даних. Програмний комплекс не є серійним продуктом, а адаптується під конкретні цілі та умови роботи замовника. Недоліком системи є її орієнтованість на дорогий прилад - ІК-спектрометр, який використовують для оцінки основних показників якості палива.

Унікальність роботи полягає у тому, що вперше для задачі ідентифікаційної експертизи була використана нечітка модель представлення даних. Також були частково усунені недоліки конкуруючих програмних продуктів: розроблена мною програма має властивість самонавчатися, тому не потрібно постійно оновлювати базу даних.

4. 4. Оцінка економічної ефективності впровадження програмного забезпечення

Таблиця 4.1.

Розрахунок чистих грошових надходжень від розробки ПЗ

Показники, грн	За роками						Усього за 5 років	Середнє за 5 років
	0	1	2	3	4	5		
Інвестиції на ПЗ	26407	-	-	-	-	-	26407	5281
Витрати до впровадження ПЗ	-	18300	18000	18700	19400	20100	94500	18900
- на щорічну перевірку на погрішність	-	3000	3300	3600	3900	4200	18000	3600
- на щорічну перевірку держстандарту	-	4500	4500	4500	4500	4500	22500	4500
- на консультацію з іншими спеціалістами	-	4800	4000	4200	4400	4600	22000	4400
- на електроенергію	-	6000	6200	6400	6600	6800	32000	6400
Витрати після впровадження ПЗ	-	4500	4700	4850	5000	5150	24200	4840
- на щорічну перевірку на погрішність	-	500	600	650	700	750	3200	640
- на щорічну перевірку держстандарту	-	1000	1000	1000	1000	1000	5000	1000
- на консультацію з іншими спеціалістами	-	500	500	500	500	500	2500	500
- на електроенергію	-	2500	2600	2700	2800	2900	13500	2700
Економія	-	13800	13300	13850	14400	14950	70300	14060
Амортизація	-	5281	5281	5281	5281	5281	26407	5281
Чисті грошові надходження	-	8519	8019	8569	9119	9669	43893	8779
Коефіцієнт дисконтування	-	0,87	0,756	0,658	0,572	0,497	-	-
Дисконтові грошові надходження	-	7411	6062	5638	5216	4805	29133	5827

Коефіцієнти економічної ефективності

Чиста поточна вартість доходів:

$$NPU = 29133 - 26407 = 2726 \text{ грн} > 0$$

Строк окупності:

$$T = 26407/5827 = 4,5 \text{ року}$$

Індекс прибутковості:

$$ИД = 29133/26407 = 1,1$$

Показник економічної ефективності (NPU - чиста поточна вартість доходів) за роки реалізації впровадження (5 років) складе 2726 грн, тобто відповідає умовам ефективності, тому що $NPU > 0$.

Середній термін окупності капвкладень складе 4,5 року.

Індекс прибутковості за 5 років складе 1,1, тобто $ИД > 1$, проект слід прийняти.

Таким чином, показник ефективності свідчить про те, що дане впровадження є економічно вигідним.

ВИСНОВКИ

У дипломній роботі запропоновано нове вирішення актуальної науково-технічної задачі ідентифікаційної експертизи зразків товарів. Основні результати роботи мають важливе наукове-практичне значення і зводяться до такого:

1. Виконано аналіз предмета дослідження і сутності науково-практичної задачі ідентифікаційної експертизи та доведено необхідність впровадження інформаційних технологій для її рішення, зокрема для задачі ІЕ бензинів. Задача, яка була поставлена в роботі, представлена як задача розпізнавання образів. Запропоновано використання методів обчислювального інтелекту, а саме методу нечіткої класифікації, для її рішення.
2. Висунуто і підтверджено гіпотезу про ефективність нечітких методів класифікації для вирішення задачі ІЕ. Вперше розроблено нечітку модель представлення даних нечіткими портретами, а також методи її обробки, настройки і адаптації. Обґрунтовано ефективність опису класів образів нечіткими портретами, які узагальнюють інформацію, отриману з навчальних вибірок. Сукупність таких нечітких портретів формує базу знань експерта.
3. Розроблено інформаційну технологію ідентифікаційної експертизи бензину. На основі моделей і методів, запропонованих в роботі, розроблено бібліотеку програмних компонентів, яка в подальшому може бути використана для досліджень з моделлю. Запропонований метод нечіткої класифікації на основі НМПД розв'язує задачу аналізу даних і є методом формування бази знань експерта. Це дозволило накопичувати знання експертів і використовувати їх у подальшій роботі без присутності самого експерта, таким чином – моделювати когнітивні процеси експертизи.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бензини автомобільні. Технічні умови: ДСТУ 4063-2001. – К.: Держстандрат України, 2002. – 37 с.
2. Бензини автомобільні підвищеної якості. Технічні умови: ДСТУ 4839:2007– К.: Держстандрат України, 2007. – 14 с.
3. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде; [под ред. Н.Н. Моисеева]. – М.: Мир, 1976. – 167 с. Девятков В. В. Системы искусственного интеллекта. – М.: Изд-во МГТУ им. Баумана, 2001. – 352 с.
4. Журавлєв Ю. И. Распознавание. Математические методы. Программная система. Практические применения. / Ю. И. Журавлев, В. В. Рязанов, О. В. Сенько. – М.: ФАЗИС. – 2006. – 176 с.
5. Идентификация и фальсификация пищевых продуктов. Товарный справочник / Лычников Д. С., Неверов А. Н., Николаева М. А.- М.: Экономика, 1996. – 108 с.
6. Калмыков С. А. Методы интервального анализа / С. А. Калмыков, Ю. И. Шокин, З. Х. Юлдашев. – Новосибирск: Наука, 1986. – 222 с.
7. Кофман А. Введение в теорию нечетких множеств / А. Кофман; [пер. с франц]. – М.: Радио и связь, 1982. – 432 с.
8. Кратковник В. Я. Непараметрическая идентификация и сглаживание данных. Метод локальной аппроксимации / В. Я. Кратковник. – М.: Главная редакция физико-математической литературы, 1985. – 336 с.
9. Круглов В. В. Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода / В. В. Круглов, М. И. Дли. – М.: Физматлит, 2002. – 252 с.
10. Круглов В. В. Нечеткая логика и искусственные нейронные сети / В. В. Круглов, М. И. Дли, Р. Ю. Голунов. – М.: Физматлит, 2001. – 224 с.
11. Леоненков А. В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А. В. Леоненков – СПб.: БХВ-Петербург, 2003. – 736 с.: ил.

12. Нечеткие гибридные системы. Теория и практика / [И. Р. Батыршин, А. О. Недосекин, А. А. Стецко, В.Б.Тарасов и др.]; под ред. Н. Г. Ярушкиной. – М.: Физматлит, 2007 – 208 с.

13. Никифоров И. К. Нейросетевой импедансный метод и устройства идентификации и определения параметров жидких нефтепродуктов: автореф. на соиск. уч. степени канд. тех. наук. – Казань, 2005. – 19 с.

14. Николаева М.А. Идентификация и обнаружение фальсификации продовольственных товаров / М.А. Николаева, М.А. Положишникова. – М.: ИД «Форум» – ИНФРА М. – 2009. – 463 с.

15. Обработка нечеткой информации в системах принятия решений. / А. Н. Борисов, А. В. Алексеев, Г. В. Меркурьева и др. – М.: Радио и связь, 1989. – 304 с.

16. Объектно-ориентированный анализ и проектирование с примерами приложений на С++ / [Г. Буч и др.]; пер. с англ. – [3-е изд.] – М.: Вильямс, 2008. – 720 с., ил.

17. Петренко Т.Г. Пакет библиотек и среда поддержки нечетких дискретных интервальных логических систем второго типа / Т.Г. Петренко, О.С. Тимчук // Системы обработки информации. Сборник научных работ. – 2011. – № 8 (98). – С. 151-156.

18. Ротштейн А. П. Идентификация нелинейных объектов нечеткими базами знаний /А. П. Ротштейн, Д. И. Кательников // Кибернетика и системный анализ. – 1998. – № 5. – С. 53–61.

19. Румшинский Л. З. Математическая обработка результатов эксперимента / Л. З. Румшинский. – М.: Наука, 1971. – 192 с.

20. Товароведение и экспертиза промышленных товаров /.[под ред. проф. Неверова А.] – Н.: Изд. МЦФЕР, 2006. – С. 847.

21. Фомин В. Н. Математическая теория обучаемых распознающих систем / В. Н. Фомин. – Л.: Изд-во Ленингр. Ун-та. – 1976. – 236 с.

22. Фомин В. Н. Рекуррентное оценивание и адаптивная фильтрация / В. Н. Фомин. – М.: Наука. – 1984. – 288 с.

23. Хант Э. Искусственный интеллект / Э. Хант; [пер. с англ.]. – М.: Мир, 1978. – 558 с.
24. Штовба С. Д. Настройка нечеткой модели по обучающей выборке с нечетким выходом / С. Д. Штовба // Кибернетика и системный анализ. – 2007. – № 3. – 17–24 с.
25. Штовба С. Д. Обеспечение точности и прозрачности нечеткой модели Мамдани при обучении по экспериментальным данным / С. Д. Штовба // Проблемы управления и информатики. – 2007. - №4. – С. 1–13.
26. Штовба С. Д. Проектирование нечетких систем средствами MATLAB / С. Д. Штовба. – М.: Горячая линия Телеком, 2007. – 288 с.: ил.
27. Ястребова Н. Н. Построение экспертных систем на базе иерархического нечеткого вывода // Программные продукты и системы. – 2007. – № 4. – С.12–52.
28. Angelov P. Evolving Intelligent System. Methodology and Applications // P. Angelov, Dimitar P. Filev, N. Kasabov. – New Jersey: John Wiley & Sons, Inc., 2010. – 445 p.
29. Angelov P. On line learning fuzzy rule-based system structure from data streams / P. Angelov, Z. Xiaowei // In: IEEE International Conference on Fuzzy Systems, 2008. – pp. 915-922
30. Angelov P. Evolving Fuzzy Inferential Sensors for Process Industry / P. Angelov, A. Kordon, X. Zhou // 3rd Int. Workshop on Genetic & Evolving Fuzzy Syst., 4-7 March, 2008, Witten-Bomerholz. – 2008. – pp. 41-46.
31. Bezdek J. C. Fuzzy Models and Algorithms for Pattern Recognition and Image Processing / J. C. Bezdek, J. Keller, R. Krisnapuram, R. Pal. – New York: Springer Verlag, 2005. – 785 p.
32. Bezdek J. C. Numerical taxonomy with fuzzy sets / J. C. Bezdek // J. Math. Bio. – 1974. – Vol. 1. Issue 1. – P. 57–71.
33. Bezdek, J. C. What is Computational Intelligence? / J. C. Bezdek // Computational Intelligence: Imitation Life: IEEE Press, NY. – 1994. – P. 1– 12.

34. Brereton R. Chemometrics. Data Analysis for the Laboratory and Chemical Plants / R. Brereton. – 2003. – Unated Kindom: John Wiley & Sons Ltd. – 489 p.
35. Brereton R. Chemometrics for Pattern Recognition / R. Brereton. – 2009. – Unated Kindom: John Wiley & Sons Ltd. – 504 p.
36. Cordon O. A proposal on reasoning methods in fuzzy rule-based classification systems / O. Cordon, M. J. Jesus, F. Herrera // IJ of Approximate Reasoning. – NH Elsevier, 1999. – No. 20. – P. 21–45.
37. Handbook of Granular Computing / W. Pedrycz, A. Skowron, V. Kreinovich, [eds.]. – NY.: John Wiley & Sons, 2007. – 1116 p.
38. Kuncheva L. I. Combining Pattern Classifiers. Methods and Algorithms / L. I. Kuncheva. – Wiley, 2004. – 350 p.
39. Kuncheva L. I. Fuzzy Classifier Design / L. I. Kincheva. – Springer Verlag, 2000. – 315 p.
40. Konar A. Computational Intelligence. Principles, Techniques and Applications / A. Konar. – Berlin: Springer Verlag, 2005. – 708 p.
41. Negneviitsky M. Artificial intelligence: a guide to intelligent systems / M. Negniviitsky. – Pearson education Limited, 2002.
42. Statistica Data Miner [Электрон. ресурс]. – Режим доступа: URL: <http://www.statsoft.com>
43. Weka 3: Data Mining Software in Java. [Электрон. ресурс]. – Режим доступа: URL: <http://www.cs.waikato.ac.nz/ml/weka/>
44. Witten I. H. Data Mining: Practical Machine Learning Tools and Techniques / I. H. Witten, E. Frank, M. A. Hall. – Morgan Kaufmann, 2011. – 630 p.
45. Yao Y. Y. Granular computing for data mining / Y. Y. Yao // Proceedings of SPIE Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security, Dasarathy, B.V. – 2006. – pp. 1–12.
46. Zadeh L. A. Fuzzy sets / L. A. Zadeh // Information and Control .- 1965. – Vol.8. – P 338–353.

47. Zhang D. Advances in Machine Learning Application in Software Engineering / D. Zhang; eds. J. Tsai. – Hershey: Idea Group Publishing, 2007. – 384 p.

48. Zhang P. A Classifier Fusion System with Verification Module for Improving Recognition Reliability / Zhang P. // Machine Learning: articles [edt. Y. Zhang]. – 2010. – P. 61–76.

49. Zadeh L. A. From circuit theory to system theory. / L. A. Zadeh //in Facets of System Science G. L. Klir. – New York: Plenum Press. – 1991.

50. Sugeno M. A Fuzzy-Logic-Based Approach to Qualitative Modeling / Michio Sugeno, Takahiro Yasukawa // IEEE Transactions on Fuzzy Systems. – Vol. 1 – No. 1. – 1993. – P. 7–31.

КОД ПРОГРАМИ

Побудова нечіткої моделі

```

int CModelProcessing::fs_Frequency(TDoubleArray x, TDoubleArray xy[2], double
beta, double alfa)
{
    xy[0].clear(); xy[1].clear();
    //xy[0] fuzzzyset X. It`s length is not equal to sample length
    //xy[1] Mju
    //step of sliding
    double step; //sliding window width
    double wind;
    //iterators for main data
    TDoubleArray::iterator the_x;
    TDoubleArray::iterator the_x2;
    std::sort(x.begin(), x.end());
    double min = x.front(), max = x.back();
    if (max - min < 0.01)
    {
        xy[0].push_back(min);
        xy[1].push_back(1);
        return 1;
    }
    double xN = x.size();
    //-----PARAMS -----
    //for step of slidig method
    int fStep = 0;
    switch (fStep)
    {
    case 0:
    {
        //my method by coefs
        double h_middle = (max - min) / (xN - 1);
        step = beta*h_middle;
        wind = (step) / alfa;
        break;
    }
    };
    //-----SLIDING -----
    the_x = x.begin();
    double xyN = floor((max - min) / step); //number of steps
    double Point_first = (*the_x) + ((max - min) - xyN*step) / 2; //left edge of
function
    double Point, beg, end;
    xy[0].push_back(Point_first - (wind / 2));
    xy[1].push_back(0);
    int i_step = 0; // number of step
    int k, maxK = 0; //counter for points in window on step i_step
    do {
        k = 0; //the number of points in the wind
        Point = Point_first + step*i_step;
        beg = Point - wind / 2;
        end = Point + wind * 2;
        while (the_x != x.end() && (*the_x) < beg)
        {
            ++the_x;

```

```

    }
    the_x2 = the_x;
    while (the_x2 != x.end() && (*the_x2) <= end)
    {
        k++; if (k>maxK)maxK = k;
        ++the_x2;
    }
    i_step++;
    xy[0].push_back(Point);
    xy[1].push_back(k);
} while (the_x != x.end() && i_step <= xyN);
xy[0].push_back(Point_first + step*i_step + (wind / 2));
xy[1].push_back(0);
//check edges
if (xy[0].at(0)<0 && xy[1].at(0) >= 0)
{
    xy[0].at(0) = 0; xy[1].at(0) = xy[1].at(1);
    //IT is left edge with 1
};
fs_Normalize(xy[1], maxK);
return 1;
};

```

```

void CModelProcessing::fs_Normalize(TDoubleArray &y, int max)
{
    TDoubleArray::iterator the_y;
    the_y = y.begin();
    while (the_y != y.end())
    {
        (*the_y) /= max;
        ++the_y;
    };
};

```

```

CFuzzyMFSigmoid *CModelProcessing::fs_CreateSigmoid(CFuzzyMFDiscrete *FS, double param)
{
    CFuzzyMFSigmoid *FSs = new(CFuzzyMFSigmoid);
    TDoubleArray xy[2];
    xy[0] = ((CFuzzyMFDiscrete *)FS)->xy[0];
    xy[1] = ((CFuzzyMFDiscrete *)FS)->xy[1];
    int xyN = xy[0].size();
    if (xyN>0)
    {
        double porog = param;
        double TopX_left_inx, TopY_left = -1;
        double TopX_right_inx, TopY_right = -1;
        for (int i = 0; i<xyN; i++)
        {
            double znach = xy[1].at(i);
            if (znach>1 - porog)
            {
                if (TopY_left == -1)
                {
                    TopY_left = znach;
                    TopX_left_inx = i;
                    TopY_right = znach;
                    TopX_right_inx = i;
                }
                else {
                    TopY_right = znach;
                }
            }
        }
    }
};

```

```

        TopX_right_inx = i;
    }
}
}
double midLeftSide = fs_getMiddleforSigmoid(xy, 0, TopX_left_inx, -1);
double midRightSide = fs_getMiddleforSigmoid(xy, TopX_right_inx, xyN - 1, 1);
double g1, g2, b1, b2;
g2 = midLeftSide;
b2 = midRightSide;
double midLeftSide_q = xy[0].at(0) + (xy[0].at(TopX_left_inx) - xy[0].at(0)) /
2;
double midRightSide_q = xy[0].back() - (xy[0].back() - midRightSide) / 2;
g1 = -log(0.01 / 0.99) / (xy[0].at(TopX_left_inx) - midLeftSide_q);
b1 = -log(0.01 / 0.99) / (xy[0].at(TopX_right_inx) - midRightSide_q);
FSs->SetCoefs(g1, g2, b1, b2);
FSs->SetMinMax(xy[0].front(), xy[0].back());
} //end if (xyN>0)
return FSs;
};

```

```

double CModelProcessing::fs_getMiddleforSigmoid(TDoubleArray xy[2], int beg_p, int
end_p, int side)
{
    double mid;
    double x_half1 = -1;
    double x_half2 = -1;
    double y1, y2;
    y2 = xy[1].at(beg_p);
    for (int i = beg_p + 1; i <= end_p; i++)
    {
        y1 = y2;
        y2 = xy[1].at(i);
        if (y1 == 0.5 && y2 == 0.5)
        {
            x_half1 = i - 1;
            x_half2 = i;
        }
        else if (((side == -1) && (y1 <= 0.5) && (y2 >= 0.5)) || ((side == 1) &&
(y1 >= 0.5) && (y2 <= 0.5)))
        {
            double x_t = (xy[0].at(i) - xy[0].at(i - 1))*(0.5 - y1) / (y2 - y1) +
xy[0].at(i - 1);
            if (x_half1 == -1)
            {
                x_half1 = x_t;
                x_half2 = x_t;
            }
            else x_half2 = x_t;
        }
    }
    if (x_half1 != -1)
        mid = x_half1 + (x_half2 - x_half1) / 2;
    else
        mid = xy[0].at(0);
    return mid;
}

```

Алгоритм нечіткого виведення

```
int CModelProcessing::ts_GenerateSamples(int flag)
{
    int Fold;
    if (flag>1) { Fold = flag; flag = 2; };
    P_all_ind.clear();
    // int CountofSamples; //=1; //for multy samples generation
    TTwoDimDoubleArray::iterator the_P;
    TIntArray P_ind;
    TIntArray::iterator the_P_ind;
    TIntStringMap::iterator the_map;
    int classNum;
    switch (flag)
    {
    case 1: //every 3rd element in class for testing
    {TIntIntMap classElementCounter;
    the_map = NC.begin();
    while (the_map != NC.end())
    {
        classElementCounter[(*the_map).first] = 0;
        the_map++;
    };
    the_P = P_all.begin();
    while (the_P != P_all.end())
    {
        classNum = (*the_P).back();
        if (++classElementCounter[classNum] == 3)
        {
            P_ind.push_back(0); classElementCounter[classNum] = 0;
        }
        else {
            P_ind.push_back(1);
        }
        ++the_P;
    };
    P_all_ind.push_back(P_ind);
    };
    break;
    case 2:
    int C = NC.size();
    int N = P_all.size();
    for (int i = 0; i<N; i++) P_ind.push_back(1); //iáo÷àpùèà
    for (int i = 0; i<Fold; i++)
    {
        TIntArray T = P_ind;
        P_all_ind.push_back(T);
    }
    TIntIntMap classElementCounter;
    the_map = NC.begin();
    while (the_map != NC.end())
    {
        classElementCounter[(*the_map).first] = 0;
        the_map++;
    };
    for (int p = 0; p<N; p++)
    {
        classNum = P_all.at(p).back();
        P_all_ind.at(classElementCounter[classNum]).at(p) = 0; //òãñòîâûå
        if ((++classElementCounter[classNum]) == Fold)
```

```

        {
            classElementCounter[classNum] = 0;
        }
    } //end for p
    break;
};
return 1;
};

int CModelProcessing::ts_InferenceBySample(int sample_index, int flag)
{
    if (FCAccuracy != NULL) free(FCAccuracy);
    FCAccuracy = new CClassifierAccuracy(NC);
    TDoubleArray data;
    TIntDoubleMap alfa;
    int classNum, classNum_rigth;
    int CountElements = 0;
    TTwoDimDoubleArray::iterator the_x_line;
    the_x_line = P_all.begin();
    int num = 0;
    cerr << "INFERENCE ";
    while (the_x_line != P_all.end())
    {
        if (GetSampleRow(num, data, classNum_rigth, flag, sample_index))
        {
            FClassifier->Inference(data, alfa, classNum);
            FCAccuracy->AddElement(classNum_rigth, alfa, classNum);
            CountElements++;
        };
        the_x_line++; num++;
    }; //end while The_x_line
    FCAccuracy->MakeConfusionMatrix();
};

```

```

CFuzzyRules *CModelProcessing::CreateBaseClassifier(int sample_index)
{
    free(FClassifier);
    FClassifier = new CFuzzyRules(NC.size(), NA.size());
    TDoubleArray data;
    TDoubleArray MF[2];
    CFuzzySet* FS;
    TRule_line FSarray; //array of fuzzy sets pointer
    TIntStringMap::iterator the_class;
    TIntStringMap::iterator the_attr;
    int CL; AnsiString CL2;
    the_class = NC.begin();
    while (the_class != NC.end())
    {
        CL = (*the_class).first;
        CL2 = (*the_class).second;
        the_attr = NA.begin();
        FSarray.clear();
        while (the_attr != NA.end())
        {
            GetSampleCow((*the_class).first, (*the_attr).first, data, sample_index);
            if (p_USE_SLIDING_PARAM)
            {
                fs_Frequency(data, MF, p_SLIDING_BETA, p_SLIDING_ALFA);
            }
            else {

```



```

        fs_Frequency(data, MF);
    };
    FS = new CFuzzyMFDiscrete();
    ((CFuzzyMFDiscrete *)FS)->Initial(MF);
    //MEDIAN FILTRATION
    if (p_MEDIANA)
        ((CFuzzyMFDiscrete *)FS)->MedianFiltration(2, 0);
    if (p_SIGMOID)
    {
        CFuzzyMFSigmoid *FS_new = fs_CreateSigmoid((CFuzzyMFDiscrete *)FS);
        free(FS);
        FS = FS_new;
    };
    FSarray.push_back(FS);
    AnsiString info = (*the_class).second + " " + (*the_attr).second;
    ++the_attr;
}
FClassifier->AddRule(FSarray, (*the_class).first);
++the_class;
}
return FClassifier;
};

```

```

int CModelProcessing::GetSampleCow(int ClassNum, int AttributeNum, TDoubleArray &x, int
sample_index)
{
    x.clear();
    //for index variant
    bool fIndx = false; //1- if we use index array P_all_ind
    TIntArray::iterator the_indx;
    if (sample_index > 0 && !P_all_ind.empty())
    {
        fIndx = true;
        the_indx = (P_all_ind.at(sample_index - 1)).begin();
    };
    double d;
    vector <double> line;
    iterlvl2 = P_all.begin();
    while (iterlvl2 != P_all.end()) {
        line = *iterlvl2;
        if (line.back() == ClassNum)
        {
            if ((fIndx && (*the_indx) == 1) || !fIndx)
            {
                d = line.at(AttributeNum - 1);
                x.push_back(d);
            };
        }; //if
        ++iterlvl2;
        if (fIndx) { ++the_indx; };
    };
};

```

```

int CModelProcessing::GetSampleRow(int num, TDoubleArray &x, int &classNum_right, int
flag, int sample_index)
{
    x.clear();
    //for index variant
    bool fIndx = false; //1- if we use index array P_all_ind
    TIntArray::iterator the_indx;

```

```

if (sample_index>0 && !P_all_ind.empty())
{
    fIndx = true;
    the_indx = (P_all_ind.at(sample_index - 1)).begin();
};
if ((fIndx && (*(the_indx + num)) == flag) || fIndx == false)
{
    TDoubleArray::iterator the_line = (*(P_all.begin() + num)).begin();
    while (the_line != (*(P_all.begin() + num)).end() - 1)
    {
        x.push_back(*the_line);
        the_line++;
    }
    classNum_right = (*the_line);
    return 1;
}
else { return 0; };
};

int CModelProcessing::getIndexClass(int classnum)
{
    vector <double> line;
    int count = 0;
    iterlv12 = P_all.begin();
    while (iterlv12 != P_all.end()) {
        line = *iterlv12;
        if (((int)(*iterlv12).back()) == classnum)
        {
            count++;
        }
        ++iterlv12;
    }
    return count;
}

```

ВІДГУК

на дипломну роботу магістра на тему:

«Удосконалення інформаційної системи експертизи бензину на основі нечіткої моделі даних»

студента групи 122м-16-1 Марчука Миколи Валерійовича

1. Метою дипломної роботи магістра є підвищення достовірності методів ідентифікаційної експертизи зразків продукції за рахунок розробки моделей, методів і програмних засобів.

2. Актуальність даної теми обумовлена великою кількістю фальсифікацій під час проведення контролю якості зразків продукції.

3. Тема дипломної роботи безпосередньо пов'язана з об'єктом діяльності магістра спеціальності 122 «Комп'ютерні науки».

4. Наукова новизна отриманих результатів дипломної роботи визначається тим, що вперше розроблена й обґрунтована нечітка модель представлення та обробки даних для задачі ІЕ, що дозволило ефективно виконувати класифікацію зразків продукції на основі моделювання когнітивних процесів експертизи та накопичених знань експерта.

5. Оригінальність технічних рішень при розробці програмного засобу полягає в створенні універсальної бібліотеки ідентифікації товарів.

6. Практична цінність результатів полягає в тому, що запропоновані в роботі моделі і методи дозволяють накопичувати і використовувати знання експерта для вирішення задач ідентифікаційної експертизи товарів.

7. Оформлення дипломної роботи магістра виконано на сучасному рівні і відповідає вимогам, що пред'являються до робіт даної кваліфікації. Ступінь самостійності виконання досить висока.

8. Дипломна робота магістра в цілому заслуговує оцінки «_____», а сам автор - присвоєння кваліфікації «інженер з комп'ютерних систем».

Керівник дипломного
проекта магістра, д.т.н.,
проф. кафедри БІТ

В.І. Корнієнко

РЕЦЕНЗІЯ

на дипломну роботу магістра на тему:

«Удосконалення інформаційної системи експертизи бензину на основі нечіткої моделі даних»

студента групи 122м-16-1 Марчука Миколи Валерійовича

Ідентифікаційна експертиза (ІЕ) виконується під час здійснення контролю якості зразків продукції та спрямована на боротьбу з їх фальсифікаціями. Реалізація фальсифікованої продукції призводить до значних економічних збитків і проблем безпеки життєдіяльності, тому актуальним завданням є боротьба з такими фальсифікаціями. Зокрема це стосується контролю якості бензину, який через особливості технологічних процесів виробництва і транспортування, а також інші фактори часто піддається фальсифікаціям.

Використання інформаційних технологій у галузі ІЕ набуває у світовій практиці все більшого розвитку, оскільки математичні та програмні засоби дозволяють інтенсифікувати і здешевити виконання ідентифікаційної експертизи. ІЕ товарів полягає в оцінюванні приналежності товару до конкретної класифікаційної групи. У процесі виконання ідентифікаційної експертизи виділяють такі важливі етапи: накопичення даних, аналіз даних, розробка і практичне використання методів ідентифікації товару. Сучасні інтелектуальні інформаційні технології дозволяють удосконалювати процес накопичення та обробки даних і замінювати роль експерта на етапі аналізу даних та побудови алгоритму ідентифікації.

Використовувані технології розробки систем обробки інформації безпосередньо пов'язані з об'єктом діяльності магістра спеціальності 122 «Комп'ютерні науки».

Наукова новизна полягає в удосконаленні методу побудови бази знань нечіткого класифікатора за рахунок нових методів автоматичної побудови функцій приналежності.

Студент М.В. Марчук досить добре розібрався в специфіці застосування різноманітних інформаційних технологій, зокрема, мов програмування C++, C#.

Беручи до уваги вище викладене, можна зробити висновок, що дана робота цілком відповідає вимогам, що пред'являються до кваліфікаційних робіт рівня магістра.

З огляду на наукову новизну і ступінь опрацювання компонентів даної роботи, в цілому автор заслуговує оцінки «_____», а також присвоєння кваліфікації «інженер з комп'ютерних систем».