

## ІДЕНТИФІКАЦІЯ ЗАЛЕЖНОСТЕЙ В БАЗІ ДАНИХ AUTO MPG

Вінницький національний технічний університет

### Анотація

Метою наукової роботи є ідентифікація залежностей в базі знань AUTO MPG. Розроблено набір програм для вирішення даної задачі в середовищі MATLAB. Все програмне забезпечення пройшло тестування. Результати тестування показали, що програма працює коректно. Програма розділяє дані на навчальну та тестову вибірки і на їх основі синтезує інтерпретабельну базу знань Мамдані, ідентифікує досліджувану залежність за допомогою лінійного та нелінійного регресійного аналізу. Результати виводяться у вигляді графіків і таблиць.

**Ключові слова:** MATLAB, Мамдані, RMSE, Інтерпретабельна, вибірка.

### Abstract

The purpose of scientific work is finding the dependence of shares on social networks from many attributes from the knowledge base AUTO MPG. A set of programs for solving this problem in the MATLAB environment has been developed. All software has been tested. The test results showed that the program is working correctly. The program divides the data into training and test samples and, based on them, synthesizes an interpretative knowledge base of Mamdani, identifies the studied dependence by linear and non-linear regression analysis. The results are displayed in graphs and tables.

**Keywords:** MATLAB, Mamdani, RMSE, Interpretative, Sample.

### Вступ

Проблеми прийняття рішень в складних умовах займають в даний час особливе місце в інформаційних технологіях. Математичні методи широко застосовуються для опису і аналізу складних економічних, соціальних та інших систем. Теорія оптимізації створила сукупність методів, що допомагають при використанні ЕОМ ефективно приймати рішення при відомих і фіксованих параметрах або коли параметри - випадкові величини з відомими законами розподілу.

### Результати дослідження

Для дослідження були використані дані, взяті із бібліотеки CMU StatLib. Предметом дослідження є залежність споживання палива автомобіля у міських умовах від інших 8 атрибутів.

Таблиця 1 – опис набору даних

<b>Data Set Characteristics:</b>	Multivariate	<b>Number of Instances:</b>	398	<b>Area:</b>	N/A
<b>Attribute Characteristics:</b>	Categorical, Real	<b>Number of Attributes:</b>	8	<b>Date Donated</b>	1993-07-07
<b>Associated Tasks:</b>	Regression	<b>Missing Values?</b>	Yes	<b>Number of Web Hits:</b>	387075

Цей набір даних є злегка модифікованою версією набору даних, наданої в бібліотеці StatLib. Згідно з дослідом Рос Куінлана (1993) у прогнозуванні атрибуту "mpg", 8 початкових екземплярів були видалені, оскільки вони мали невідомі значення для атрибуту "mpg".

Завдання даної курсової полягає у знаходженні споживання палива автомобіля у міських умовах від таких атрибутів, як вага автомобіля та його прискорення. У курсовій роботі були використані

наступні позначення:

X1 – вага автомобіля;

X2 – прискорення автомобіля;

Y – споживання палива (mpg – miles per gallon) .

Залежність буде ідентифікована за допомогою нечіткої бази знань Мамдані.

## СИНТЕЗ НЕЧІТКОЇ БАЗИ ЗНАНЬ МАМДАНІН

Автоматична побудова нечіткої бази знань Мамдані здійснюється за допомогою функції `genfis3()`.

За допомогою наступного коду була згенерована нечітка база знань Мамдані та виведені функції належності параметрів системи(рис 1):

```
fis=genfis3(X_tr, Y_tr, 'mamdani');  
[x,mf] = plotmf(fis,'input',1);  
subplot(2,1,1), plot(x,mf)  
xlabel('Membership Functions for Input 1')  
[x,mf] = plotmf(fis,'input',2);  
subplot(2,1,2), plot(x,mf)  
xlabel('Membership Functions for Input 2')
```

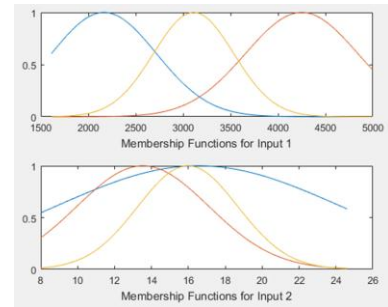


Рисунок 1 – Функції належності X1 та X2

Для цієї бази знань було розраховане RMSE за допомогою алгоритму повного перебору(рис 2)

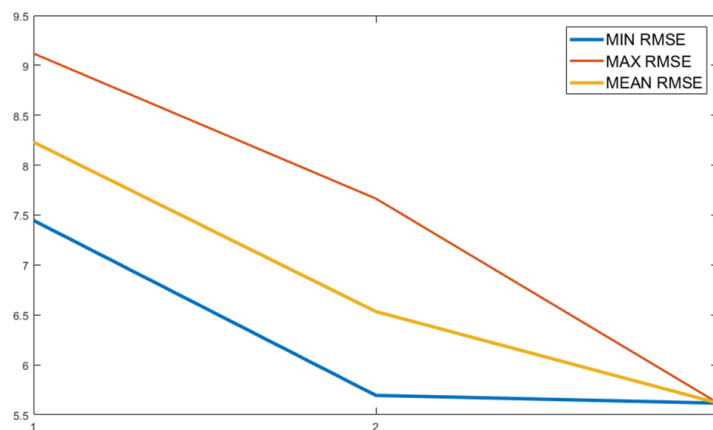


Рисунок 2 – RMSE бази знань

Інтерпретабельна база знань повинна бути зрозумілою для оточуючих. Тому її треба власноруч зробити зрозумілою. Система вказана на рис 3-4.

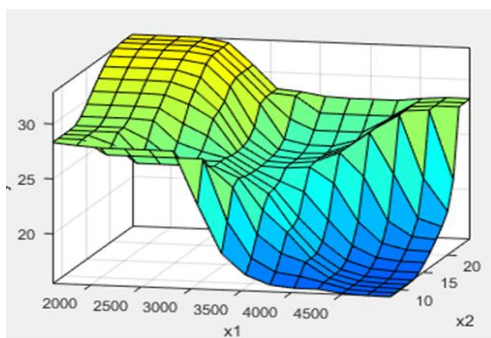


Рисунок 3 – правила

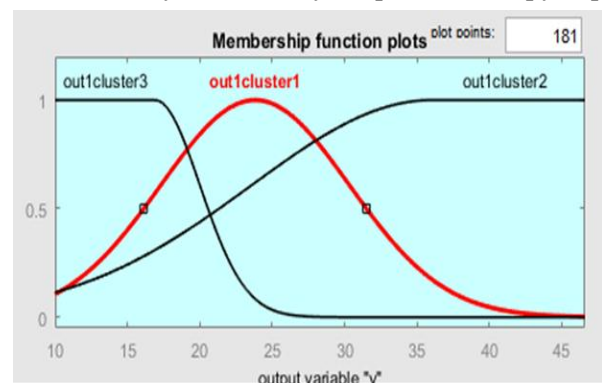


Рисунок 4 – функції належності Y

## Висновки

В курсовій роботі згідно індивідуального завдання розроблено набір програм, що виконують синтез інтерпретабельної бази знань Мамдані. В першій главі розглянута та описана база даних, яка описує витрати палива автомобілем у міських умовах по 8 атрибутам. Було здійснено завдання синтезу бази знань в середовищі MATLAB, тому що в ній функція *genfis*, що призначена для синтезу баз знань Мамдані. В другому розділі виконано розділення бази даних на навчальну та тестову вибірки та перевірено на репрезентативність, а також приведення вибірки до більш репрезентативного виду. Перевірка на репрезентативність проводилась за допомогою підрахунку математичного сподівання та дисперсії. В третьому розділі синтезовано нечітку базу знань Мамдані, а також приведено її до інтерпретабельного виду. В результаті синтезу отримано базу знань, яка складається з трьох правил.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Штовба С. Д. *Анализ критериев обучения нечеткого классификатора* / С. Д. Штовба, О. Д. Панкевич, А. В. Нагорна // *Автоматика и вычислительная техника*. – 2015. – №3. – С. 5–16
2. Штовба С. Д. *Дослідження навчання компактних нечітких баз знань типу Мамдані* / С. Д. Штовба, В. В. Мазуренко // *Штучний інтелект*. – 2011. – №4 – С. 521–529
3. Vache K. *UCI Machine Learning Repository* / K. Vache, M. Lichman. Irvine: University of California, School of Information and Computer Science. 2014. – Режим доступу: <http://archive.ics.uci.edu/ml>
4. Штовба С. Д. *Ідентифікація багатофакторних залежностей за допомогою баз знань. Лабораторний практикум : електронний навчальний посібник* / С. Д. Штовба, А. В. Галуцак – Вінниця : ВНТУ, 2015. – 96 с.
5. Штовба С.Д. *Інтелектуальні технології ідентифікації залежностей. Лабораторний практикум : електронний навчальний посібник* / Штовба С.Д., Мазуренко В.В. – Вінниця : ВНТУ, 2014. – 113 с.

**Клочак Богдан Сергійович** — студент групи АКіТ-18мс, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, Вінниця, e-mail: naruto123806@gmail.com

Науковий керівник: **Штовба Сергій Дмитрович** — доктор технічних наук, професор, професор кафедри комп'ютерних систем управління, Вінницького національного технічного університету

**Klochak Bogdan S.** — Faculty for Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, email : naruto123806@gmail.com

Supervisor: **Shtovba Sergey D.** — Doctor of Technical Sciences, Professor, Professor of Computer Control Systems Department, Vinnitsa National Technical University