Інтелектуальні технології аналізу наукометричних даних та нечіткі моделі для інтегрального оцінювання кращих 100 науковців України

Володимир Красиленко, Діана Нікітович, Юлія Тріфонова

Кафедра економіки, менеджменту та інформаційних технологій, Вінницький соціально-економічний інститут Університету "Україна", УКРАЇНА,
м. Вінниця, вул. Червоноармійська, 1,
E-mail: krasilenko@mail.ru; Nikitovitch.Diana@yandex.ru

Abstract: The article informs us of the results of integral valuation of a hundred of the best Ukrainian scientists over epy indexes of scientometrical database. There was built careless model, which let us to define influence of factors and resulting rating of each scientists.

Ключові слова: інтелектуальний аналіз даних, кластерний аналіз, наукометричні бази даних, нечітка модель, нечіткі лінгвістичні змінні.

I. Вступ

В багатьох країнах на основі показників наукометричних баз даних таких як Web of Science, Scopus та інших визначають і оцінюють рейтинги науковців [1]. Для такого оцінювання за результатами попередньої кластеризації доцільно використовувати сучасні інтелектуальні інформаційні технології та нейропакети [1,2]. Їх використання дозволяє ефективно здійснювати не лише кластерний аналіз об'єктів із значною кількістю якісних та кількісних ознак, але і визначати впливовість факторів та давати інтегральні оцінки, в тому числі і на основі нечітких моделей. **Тому метою даної роботи** є інтелектуальний аналіз даних та побудова програмою Fuzzy TECH нечіткої моделі для інтегрального оцінювання 100 науковців України.

II. Виклад матеріалу та результатів дослідження

Для побудови нечіткої моделі ми використали визначені при кластеризації найбільш впливові показники і обрали вхідними змінними (ВЗ) 3 нечітких лінгвістичних змінних: "загальна кількість цитувань у базі Web of Science", "індекс Гірша" та "загальна кількість цитувань у базі Scopus", а вихідною змінною – "Оцінка". Кількість термів у змінних дорівнювала 3, а у вихідній змінній було 5 термів. Ми покажемо властивості термів ВЗ та базу знань з визначеними правилами. Тут ми наводимо лише отримані результати без опису процесів їх отримання. Нечіткий висновок на основі розробленої моделі в режимі настроювання при різних значеннях (ВЗ) показаний на рис.1, на якому

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua

наведені функції належності 5 термів вихідної змінної та прийняті моделлю в різних ситуаціях рішення (відмічені червоною стрілкою!).

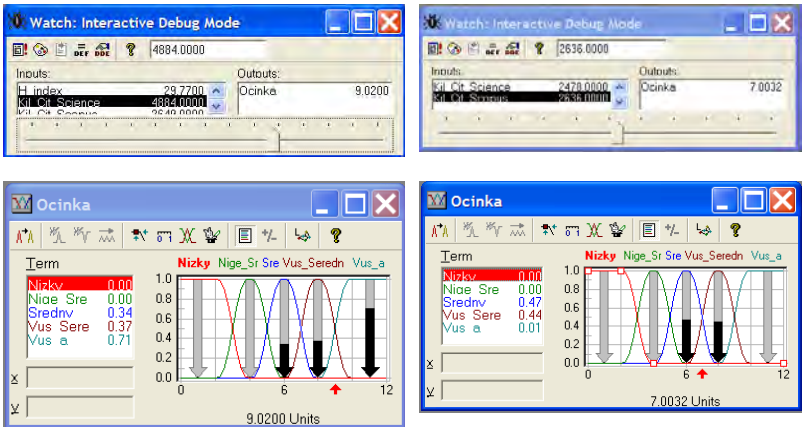

Рис. 1 Вид робочого інтерфейсу програми Fuzzy TECH в режимі настроювання

Аналіз показує, що для отримання оцінки на рівні 9 балів необхідні такі значення ВЗ: H_index – 29.7700; Kil_Cit_Science – 4884.0000; Kil-_Cit_ Scopus – 2649.0000, а для оцінки в 7.0032 – такі: H_index – 28.1100, Kil_Cit_Science – 2478.0000, Kil_Cit_Scopus – 2636.0000. В доповіді ми проаналізуємо інші ситуації та обґрунтуємо вибір трьох ВЗ шляхом аналізу впливовості за допомогою нейропакетів.

Висновок

Проведені дослідження показують, що побудована нечітка модель дозволяє не лише визначати впливовість факторів на оцінку та місце в рейтингу, але й за обмежень на значення деяких змінних шукати потрібні значення інших показників для отримання необхідної оцінки.

Література

1. Красиленко В.Г. Інтелектуальний кластерний аналіз кращих 100 науковців України за показниками наукометричних баз даних / В.Г. Красиленко, Д.В. Нікітович // Науково-методичний збірник матеріалів НПК ВСЕІ Університету «Україна». – Вінниця, ВСЕІ Університету «Україна», 2013. – С. 101–104.
2. Красиленко В.Г. Інтегральне оцінювання читабельності рефератів за допомогою нечіткої моделі / В.Г. Красиленко, Ю.М. Трифонова, С.К. Грабовляк // Науково-методичний збірник матеріалів НПК ВСЕІ Університету «Україна». – Вінниця, Вінницький соціально-економічний інститут Університету «Україна», 2012. – С. 167–169.

Lviv Polytechnic National University Institutional Repository http://ena.lp.edu.ua