

УДК 004.93'14

ЗАСТОСУВАННЯ МЕТОДІВ КЛАСТЕРНОГО АНАЛІЗУ ДЛЯ ІДЕНТИФІКАЦІЇ ГІБРИДНИХ ТВАРИН

Гітис І.В.

Науковий керівник – д.т.н., проф. Філатов В.О.

Харківський національний університет радіоелектроніки, каф. ШІ,
м. Харків, Україна

тел. +38(095) 822-07-84, e-mail: iryna.hitis@nure.ua.

The features of classification of representatives of the modern genus of the equine family and hybrids that are obtained as a result of their crossing are analyzed in the work. It is proposed to use an algorithm of cluster analysis to determine which category an animal whose morphological characteristics do not fit the description of any known breed belongs to. This approach will help simplify and automate the procedure for identifying horse hybrids for both specialists in the field of horse breeding and government officials when registering animals.

Багато століть тому коні стали невід'ємною частиною життя: вони використовувалися в сільському господарстві і як засіб пересування. Зараз у цих галузях коней замінили машини, але це не означає, що ці тварини зникли з життя людини, вони активно беруть участь у спортивних змаганнях та сфері розваг. У кожного коня є порода, яка зумовлює морфологічні ознаки, що відрізняють його від інших особин [1]. Проте у природі зустрічаються організми, які були отримані шляхом схрещення генетично відмінних форм. Це явище також може зустрічатися серед коней. Таких тварин досить важко класифікувати з приводу відсутності приналежності до жодної з існуючих порід.

Метою роботи було підвищення ефективності ведення обліку та інтелектуального аналізу властивостей зареєстрованих тварин шляхом створення бази даних та вдосконалення засобів обробки інформації.

У ході виконання роботи було створено базу даних, яка надає можливість зберігання та обробку даних про коней, що включають ім'я тварини, її породу, масть, стать, дату народження, ім'я власника та назву заводу, де було виведено коня. Для класифікації тварин, яких неможливо однозначно віднести до будь-якої породитакатегорії, було використано кластеризацію методом k-середніх. Метою метода є розбиття деякої кількості об'єктів на кластери, так щоб кожне спостереження належало до кластера з найближчим до нього середнім значенням. Незаперечними перевагами цього методу є його простота та швидкість виконання [2].

Породи коней прийнято класифікувати на три великі категорії: верхові, важковози та поні [1]. Використання кластерного аналізу допоможе віднести гібриди коней до цих груп.

Для класифікації коней були обрані наступні дані: масть, стать, вік у днях, зріст та вага. Було виконано кодування усіх символічних типів, аби дані були однорідними. Так, масть тварин була пронумерована таким чином, що найтемніша дорівнює цифрі 1, а найсвітліша – 6. У випадку зі статтю, для жіночої було обрано цифру 0, для чоловічої – 1.

Для дослідження було обрано 40 особин. У результаті виконання кластеризації методом k-середніх об'єкти було розбито на три кластери, що відповідають категоріям тварин. Фрагмент вихідних даних та результатів роботи алгоритму приведені в таблиці 1.

Таблиця 1 – Результати виконання кластеризації

Масть	Стать	Вік	Зріст	Вага	Кластер
1	1	4738	145	400	2
2	0	1511	80	165	1
2	1	2838	170	750	3
...
6	0	930	185	730	3
5	0	6229	150	380	2

До кластеру з номером 3 потрапили найважчі тварини, до кластеру з номером 1 – найлегші та найменші за зростом особини. Спираючись на отримані результати, можна обчислити центр кожного кластеру (табл. 2).

Таблиця 2 – Центри кластерів

Категорія	Кластер	Масть	Стать	Вік	Зріст	Вага
Поні	1	2,9	0,9	6312,3	110,6	169,7
Верхові	2	3,04	0,5	4065,4	159,6	479,1
Важковози	3	3,4	0,6	4160,8	176,4	772,4

Отже, використання кластерного аналізу допоможе фахівцям в області конярства та тваринництва класифікувати тварин, які є результатом схрещення різних за генотипом форм особин. Ідентифікація гібридів виконується шляхом знаходження найкоротшої відстані до центру кластеру.

Список використаних джерел:

1. Soldi A. (2007) Loshadi / per. s ital. N. V. Yarosh. Moskva : AST: Astrel.

2. Деркач О.І. Аналітична обробка текстової інформації за допомогою засобів кластеризації. Молодий вчений. 2016. № 7(34). С. 159–165.