RESEARCH ARTICLE OPEN ACCESS

Proficient Possessions Supervision Using Various Clustering Techniques

[1]T.Karunakaran , [2]J.John Clotilda

[1] Assistant professor, Dept.of.Computer science , Prist university Thanjavur

[2] Research Scholar, Dept.of.Computer science, Prist university Thanjavur

----------**********----------

Abstract:

Feed is the largest input in any livestock enterprise and the rapid increase in feed prices and shortage of feed resources has been one of the major constraints for farmers, livestock industries, planners and the policy makers. This calls for prudent management of available resources and application of computing techniques can be one of the possible potential approaches. India is endowed with a wide range of feed resources varying widely in their composition and utility for different livestock species. Clustering of feed resource into different groups based on the composition can help in better feed management. To evaluate and to suggest a best technique for clustering feed resources, we have evaluated three clustering techniques viz. K-means, spectral k-means and auto spectral on two different data sets containing 236 and 106 feed resources with major constituents like crude protein, crude fiber ash, fat etc., . The composition of commonly used feed resources (236 and 106) were sourced from the published articles in various scientific journals and based on the composition of the feeds, they were subjected to clustering techniques.

A total of three clustering techniques – k means, spectral k means and auto spectral were employed and the outputs were compared against the grouping done by subject matter experts.

The outputs of different clustering techniques were validated with the expert grouping using precision, recall and F-measure. Of the three methods used it was found that k-means was the best and closest to the experts classification. Clustering of feeds into different groups based on their composition will form a sound basis to select different feeds for formulating economic and nutritionally balanced diets leading to better feed management.

Keywords — **cluster, mining, data.**

----------**********----------

Introduction

This chapter presents a brief introduction to information technology in agriculture, importance of feed composition and its importance in livestock feeding . The proposed problem definition, aim and scope of related work and a brief introduction on data mining in relation to the present proposed problem.

Importance of IT in Livestock sector

Traditional research in livestock sector has been confined to improve feeding management and feed processing technologies and very little attention has been paid to the use of information technology. Data mining has become one of most active and fertile areas of research and practice due to its wide range of applications, from traditional business and engineering to today's IT technologies.

Mining techniques are showing their usefulness also in areas of agriculture and food industry. Nonetheless, compared to other domains, mining techniques have been less explored for agriculture domain applications (Mucherino et al 2009). Data mining can be a effective tool for Developing innovative applications in agriculture (Cunningham and Holmes 1999) and livestock (Angadi et al 2009).

Livestock feeding and feed composition

The availability of accurate information on the nutrient composition of feeds is essential for livestock producers so that they can formulate diets that result in healthier, more efficient animals and can use environmentally conscious farming practices to produce high-quality animal commodities. Recognizing that more than one-half of the cost of animal production can be attributed to feed costs, producers rely on appropriately managed nutrient levels in animal diets to maximize animal productivity and economic efficiency. In Indian context a wide variety of feed resources are utilized due to different livestock production systems and diverse agro climatic conditions. Based on the composition of feedstuffs, their availability and cost, least cost formulations can be worked out using the least cost packages for reducing the feeding costs and improving the profitability.

The inventory of feed resources in India is quite large due to diverse agro climatic conditions and there is need for cataloguing all the feed resources and grouping them into different groups based on their nutrient composition. This would greatly facilitate in choosing the right kind of feedstuff based on the composition and give a wider choice to choose from a range of feed resources based on their availability and cost within the same group.

The science of animal nutrition is advancing rapidly. In recent years investigators have made much progress in providing an understanding of animal metabolism, the roles of specific nutrients in normal physiological processes, and the nutritional value of feedstuffs.

The importance of the nutrient composition of feeds addressed more than a century ago continues to be recognized as a primary mechanism for optimizing animal health (Breeding et al., 1994). Recently, refined nutrient management has been recognized as an important element of sustainable agriculture practices (Honeyman, 1993). Managing animal feeding practices with nutrient composition information has proved to be beneficial for addressing the environmental concerns associated with agricultural production (Tamminga, 1992). Additionally, feed composition serves as a point of reference for the regulation of food and feed ingredients in domestic and international markets to ensure fair trade and unbiased scientific evaluation of food safety issues (Bogel and Stohr, 1994). Educators and researchers develop programs that rely heavily on information regarding the nutrient composition of feeds. Internationally, the nutrient composition of feeds influences trade and marketing decisions in quality-sensitive sectors (U.S. Department of Agriculture, Economic Research Service, 1994b) and affects the education and welfare of populations in developing countries. Finally, feed manufacturers as well as animal producers and technical consultants on farms use feed composition information on a daily basis.

Aim and scope of the work

The proposed work aims to list all the feed resources with their composition and cluster them into different groups based on their composition using the appropriate clustering techniques.

Manual grouping of the feeds will not only be cumbersome but it would be difficult as the number of parameters to be considered is more than what could be handled manually. The clustering of feed resources into different groups will facilitate the livestock industry, feed industries and trading agencies in taking better decisions in choosing the appropriate feed resources for efficient and economical livestock production. Further use of the computer applications in livestock industry would pave way for better inter disciplinary collaboration between the Information technology and the biological sciences leading to better natural resources utilization.

The proposed problem Definition

A large number of feed resources and the data on their composition are being generated worldwide over the years due to rapid increase in the demand for feed resources. More and more newer feed resources are being added to the existing feed inventory and most of these originate as byproducts from food and agro processing industries. The need for efficient and economic livestock production is driving the feed industries, farmers and researchers to try and explore newer resources. A large amount of information has been generated with regard to the feed resources and its composition. It is challenging for computer scientists and biologists to transform this data into useful information that can help in formulating better rations that are not only balanced nutritionally but are also economic. Grouping of different feed resources into different categories based on their composition will greatly facilitate in choosing a wider range of feed resources for formulating scientifically balanced diets. This would essentially help the feed industries, traders and the farmers in choosing feed resources that are locally available and are economical and avoid choosing resources that are either not available locally or are costly.

The most important component in this approach would be to classification/clustering of feed resources based on compositional similarity of the nutrients. Although traditionally feed resources are broadly classified into roughages and concentrates based on the fiber content, there is a need to have better classification of feed resources for formulating diets that are scientifically balanced as per the requirement of different species for different functions like milk, meat, eggs, wool and work.

Generally feeds with similar composition have similar feeding value. Newer feed resources can be grouped into different clusters based on their composition and they can be efficiently replace the conventional feed resources from the same group which are generally in short supply, costly or not available. Thus clustering of feed resources will have great impact on utilizing the feed resources efficiently and economically. the biological sciences leading to better natural resources utilization.

Thesis organization

This dissertation consists of 5chapters.

- Chapter 1 gives This chapter presents a brief introduction to information technology in agriculture, importance of feed composition and its importance in

livestock feeding, aim and scope of work and the problem definition

- Chapter 2 summarizes the related literature.
- Chapter 3 describes the various methodologies used in this research work.
- Chapter 4 discusses the results of clustering techniques- K mean, spectral k mean and auto spectral k mean .
- Chapter 5 Concludes the findings of the research described in this dissertation and future enhancements.

	Experimental classes					Total	% Accuracy
True Classes	0	5	13	0	8	26	50.00
	10	1	1	0	0	12	83.33
	0	0	1	0	11	12	91.67
	1	26	10	1	1	44	59.09
	0	10	1	3	0	59	76.27
	0	0	1	76	0	83	91.57
Total	11	42	28	80	20		

Data mining and Clustering

Data Mining and Knowledge Discovery in Databases (KDD) is a rapidly growing area of research and application that builds on techniques and theories from many fields, including statistics, databases, pattern recognition and learning, data visualization, uncertainty modeling, data warehousing and online analytical processing (OLAP), optimization, and high performance computing [Figure 1]. KDD is concerned with issues of scalability, the multi-step knowledge discovery process for extracting useful patterns and models from raw data stores (including data cleaning and noise modeling), and issues of making discovered patterns understandable.

Find human-interpretable patterns that describe the data.

- Classification [Predictive]
- Clustering [Descriptive]
- Association Rule Discovery [Descriptive]
- Sequential Pattern Discovery [Descriptive]

Regression [Predictive

Results of Dataset on K-means clustering Technique

The output of k mean clustering on different datasets- 236,106,342,114a,114b and114c are presented in graphical and tabular forms.

Conclusion:

Feeding is the major component of livestock production and judicious utilization of feed resources can improve the efficiency of production. In India a variety of feed resources are used for feeding different categories of livestock. These feed resource differ in their chemical composition (nutrients) and are used in different combinations to balance the ration. Feeds are grouped into different categories based on their composition and categorization of feeds is done by subject experts using the chemical composition. Due to rapid increase in the type of feed resources there is a need for proper classification using newer tools.

The present work is an attempt in this direction wherein the commonly used feed resources has been grouped using the clustering techniques and the outputs have been cross checked with expert classification to assess the efficacy

of the clustering techniques. A total of 342 observations on composition of feed stuffs have been used in different ways to study the sample size and distribution of feed resources in different groups. A total of six different datasets were tried and three clustering techniques – k means, spectral k means and auto spectral k means were used. The efficacy of clustering was compared with the expert classification was evaluated in terms of recall, precision and f measure.

The major findings of the present study are as follows

1. Clustering techniques can be used for classifying the feed resources into different groups with reasonably good accuracy.

2. Of all the three techniques k means was found to be the best and the output was close to 70% of the expert's classification.

3. The k means clustering was not affected by sample size and the distribution of different feed resources in different data sets as the outputs were almost consistent across groups.

The outputs of the present study demonstrates that the new tools – clustering can be effectively used for grouping of different feed resources without the aid of experts to an extent of 70% and thus can form a sound basis for efficient feed management.

The application of the present study is that this can be used for clustering new feed resource into a particular category which will be useful in determining the extent of usage of the new feed.

Similarly the classification helps in choosing ingredients that are locally available and are economical leading to economical ration formulations thus helping in improving the profitability. Clustering the feeds into different groups provides multiple options to the end user (farmer or feed industry) to choose from a wide range of feed resources that is best suited in terms of local availability or price advantage.

References

1. Alpert C., Kahng A., and Yao S. (1999), "Spectral partitioning with multiple eigenvectors", Discrete Applied Mathematics Vol. 90, Issues 1-3, pp. 3-26.

2. Anonymous 2008 – Annual report .Department of Animal Husbandry, Dairying & Fisheries, Ministry of Agriculture, Government of India, New Delhi

3. Arun K Pujari -"Data mining Techniques" University press.

4. Bach F. and Jordan M. (2004). Learning spectral clustering. In S. Thrun, L. Saul, and B. Scholkop (Eds.), Advances in Neural Information Processing Systems 16 (NIPS). Cambridge, MA: MIT Press

5. Carpenter G.A. and Grossberg S. (1988), "The ART of adaptive pattern recognition by a self-organizing neural network", Computer, Vol. 21, pp. 77-88

6. Chung F, (1997), "Spectral Graph Theory. CBMS Regional Conference Series in Mathematics", American Mathematical Society, 92.