The combination of neural network and "question matching" improves the correct rate of grassland degradation decision

Li Chunmei¹,², Pi Wei¹, Dong Suo¹ and Li Zhao¹

¹Department of Computer Technology and Application, Qinghai University, Xining, Qinghai 810016, China
²Li Chunmei’s e-mail: li_chm0422@sina.com

Abstract. This paper analysed the deficiency of grassland degradation by neural network method, and proposed a method combining neural network with “question matching”. The neural network evaluation is the main line, supplemented by “inquiry matching” evaluation, and supplemented by “degradation indicating grass species”. The combined evaluation of grassland degradation improves the correct rate of grassland degradation decision.

1. Instruction

The Sanjiangyuan region is an ecological barrier for the ecological environment security and regional sustainable development of the Yangtze River, the middle and lower reaches of the Yellow River, and Southeast Asian countries. However, under the dual influences of global climate change and human activities in recent decades, grassland vegetation degradation in the Sanjiangyuan region is seriously, the source water conservation capacity has been drastically reduced, and it has directly threatened the ecological security of the Yangtze River, the Yellow River Basin, and even Southeast Asian countries. In 2005, the state invested 7.5 billion yuan to launch the Sanjiangyuan area ecological protection project, and carried out three major constructions: ecological protection, agricultural and herdsman production and living infrastructure and ecological protection support. Since the implementation of the project in 2005, through the implementation of a series of measures such as returning grazing and returning grass, closing hillsides for afforestation, returning farmland to forests, black soil beach management, rodent control, soil and water conservation and artificial precipitation enhancement projects, the vegetation in the implementation area has been significantly restored, the water source conservation function has been realized initially, the area of the swamp and lake in the source area has expanded to varying degrees [1].

In the long-term research on Sanjiangyuan ecological environment protection, researchers collected a large amount of data on ecology, geography, geology, environmental science, sociology, economics, etc., which condensed several generations of researchers’ hard work and sweat, but because of the relationship between the disciplines of scientific research workers, they are often isolated information data, with the characteristics of the discipline and data isolation. It has long relied on expert’s experience and manual methods to evaluate grassland degradation. In the long-term research work, we solved the problem of grassland degradation degree determination and decision-making in the Sanjiangyuan region by using computer artificial intelligence technology to develop an expert system for hierarchical decision-making and treatment of alpine meadow grassland in the Sanjiangyuan alpine meadow. On the one hand, it can replace the grassland expert with computer to make expert-level decision-making, saving manpower, material resources and financial resources. On the other hand, the
expert's knowledge can be systematically summarized, stored in the computer for a long time, and the expert knowledge can be protected and passed down.

2. Neural network decision-making sanjiangyuan grassland degradation decision

For grassland degradation [2, 3, 4], there are many research methods, different angles, different entry points, and different conclusions, such as references 1, 2, and 3. This study carried out detailed research and analysis on the degradation classification of alpine meadow grassland in Sanjiangyuan area, the main factors affecting degradation, and the treatment measures under each degraded condition, and based on the alpine meadow in the Sanjiangyuan area of Qinghai Plateau, Meadow grassland grading for decision making. On the basis of the research data, the data is divided into two categories, and the first type of data is determined by using neural network. At the same time, the second type of data is supplemented by “question matching” method, and the two are organically combined to establish the management of alpine meadow in Sanjiangyuan area, provides technical support and theoretical basis for the Sanjiangyuan ecological protection project.

Neural networks [5, 6, 7, 8] have been widely promoted in many fields because of their parallelism, solving nonlinear relationships, etc. In recent years, methods for improving them have emerged in an endless stream, as in References 5, 6, 7, 8, etc. This research is not based on the improvement of the neural network itself, but on the basis of it, to its deficiencies, uses complementary methods make the neural network more perfect. The use of neural networks to determine grassland degradation has its advantages of speed and intelligence, but it also brings some details of the decision.

2.1. Evaluation of data in the "blind zone"

The title is set 17 point Times Bold, flush left, unjustified. The first letter of the title should be capitalized with the rest in lower case. It should not be indented. Leave 28 mm of space above the title and 10 mm after the title.

2.1.1. Definition of "blind zone". In the study, researchers collected and organized more than 20,000 pieces of data as a training set for neural networks. In the training set, there are five categories of data, namely, non-degraded, mildly degraded, moderately degraded, severely degraded, and extremely degraded. The grass species of each degraded category have obvious characteristics. According to these characteristics, they are summarized into five factors according to the expert's point of view, namely, the smear degree (tbdgd), the edible grass ratio (ksmbl), and the degraded indicator grass species proportion (thzszbl), soil organic matter content (tryjzhlbl), rodent damage (shqk). The five input data corresponding to each type of grassland has a distinct feature in distribution, and each type of data is roughly distributed in different regions, as shown in Figure 1 below. Experience has shown that the distribution of the five types of data is regular, as shown. The distribution of the data in the figure is the distribution of the ratio of forage to soil organic matter content. The five categories on the diagonal from the upper left to the lower right are the relation distribution of proportion of edible forage grass and soil organic matter content under the conditions of extreme degraded, severe degraded, moderate degraded, mild degraded and non-degraded respectively. It can be seen from the figure that the input data of the range indicated by the arrow cannot be evaluated, or the test result of the input data between the areas is a category that cannot be judged, "blind zone". Therefore, the data in this area, researchers call it "blind zone" data.
When "blind zone" data is used as input of neural network, the evaluation result will be wrong. Because the "blind zone" data is not within the range of training data sets and test sets of the neural network, the output of the "blind zone" data as input of the neural network is 100% wrong.

2.1.2. Description of "blind zone" data. The distribution of characteristic data of each type is regular. In soil with a high proportion of edible forage grass, the content of organic matter is naturally high; The organic matter content in soils with high patch coverage and even reduced to black soil paralysis is naturally low.

That is, there is a correlation between the five factors that affect the degree of grassland degradation. As shown in Figure 2, there is a correlation between the proportion of degraded indicator grass species and soil organic matter content. For example, the relationship between people's age and running speed is related. As people get older, their running speed increases; After reaching a certain limit, the older you get, the slower you start to run.

In this experiment, the relationship between the five grassland types and the proportion of edible forage grass, the proportion of degraded indicator grass species, and the content of soil organic matter were all correlated, as shown in Figure 3, 4, 5 below.
3. Evaluation of neural network model

3.1. Defects of neural network model evaluation

Due to the defects of the neural network model itself, the testing accuracy of the test data is not 100%, which leads to the accuracy of the neural network model is not 100%, which means that incorrect evaluation results will inevitably appear in the application. Based on this reason, researchers will conduct supplementary evaluation in the following ways to improve the accuracy of grassland evaluation.

3.2. Combination of neural network and “question matching” to improve the decision accuracy

In this study, researchers found that, in addition to the above five important factors, the degradation indicator grass species can be used again to assist in the evaluation of the degradation degree of grassland. Therefore, researchers can further determine whether there is a problem in the previous judgment by indicating whether degraded indicator grass species appears and its proportion. The specific evaluation scheme is shown in Figure 6 below.
3.2.1. "Question" whether the input data is in the blind zone to exclude areas that cannot be evaluated.

Data in the "blind zone" needs to be excluded. The distribution of the data is within a certain range to fit the turf category. For example, in the non-degraded grassland, the proportion of forage grass is bound to be a large proportion. If all the weeds are poisonous, and the grassland category is non-degraded, this cannot happen. Or, in the non-degraded type of grassland, the organic content of the grassland is very low, and there is a very dense plant community in the barren land, which is not in line with the natural law. Therefore, before entering the neural network for the first evaluation, a judgment should be made. If the data in the "blind zone" does not enter the evaluation process, directly output, your input data does not match the actual situation, please input the correct data. The evaluation process cannot be entered until the input data is "correct".

3.2.2. "Question" other "knowledge" for further judgment

3.2.2.1. Data at category junction.
In the first evaluation result of neural network, there may be a small amount of data input that is wrong, in order to avoid the occurrence of such error, this project adopts the second evaluation method and further "question" [1], the characteristics of each type of grassland can be correctly evaluated if they conform to the characteristics of this type of grassland. Otherwise, compare the "input data" of the first evaluation with the data of the "question matching" of this time, if their common characteristics of "bareland coverage" and "organic content" are similar, if on the basis of some similarity, then "question" dominant grasses and "companion species", if both are matched, the final category can be determined by combining the level of the first evaluation with which category is close. Figure 7, Figure 8 and Figure 9 below show three types of grassland characteristics: non-degraded, moderate degraded and extreme degraded.

Suppose, some data input to the neural network is (0.71, 0.19, 0.79, 0.2, 0.7), and the result of the first evaluation is (0, 0, 0, 0.7776, 0.7845), there's no way to tell if it's category 4 or 5. At this point, the input data of the neural network was traced back to the original data (the data before normalization), which was (71.5, 20, 79, 6.9, 0.72) and compared with the criteria for the secondary evaluation, the standard of bareland coverage is slightly greater than extreme degraded, and the content of organic matter is lower than extreme degraded, so again, "question" the companion situation, for example, if the herbage and the arachnids disappear, if they disappear, continue to "question" what the dominant species is, if it is twelve-year-old poisonous weeds, such as Ligularia virgaurea, Ajuga lupulina and other plants, you can interpretation this kind of grass has extreme degraded; If the content of organic matter and the coverage of the plaque land cannot be determined, continue to "question" whether its companion species are mainly containing weeds such as Oxytropis and Lancea tibetica, and if "yes", continue to "question" whether dominant grasses are mainly dominated by stoloniferous weeds, and if so, are severely degraded.

3.2.2.2. Data in category area. The data in the obvious category area, that is, the neural network model evaluates the correct category, generally speaking, will not be wrong. In case of an error, the
original data input from last time and this time can be quickly determined by using the same method. Figure 7, 8, 9 are the partial grassland features of the five categories of grassland in this project, such as three categories of grass features of non-degraded, severe degraded and extreme degraded.

![Figure 7. Characteristic of non-degraded grasses.](image)

![Figure 8. Characteristic of severe degraded grasses.](image)

![Figure 9. Characteristic of severe extreme grasses.](image)

4. Conclusion

The evaluation of grassland is a complicated process, and the conclusion will be different if the experts study from different perspectives. In this paper, taking the view of some experts as an example, the evaluation process of grassland is completed. The first evaluation of the grassland was carried out by training the neural network model with a large number of experience values from five important factors including the coverage of the patch, proportion of dominant grass species, proportion of edible forage grass, soil organic matter content and rodent damage. After that, from the perspective of organic matter content, patch coverage, companion species and dominant grass species, it was evaluated again with the input data of neural network, and finally the degradation degree of grassland was obtained, which greatly improved the accuracy of grassland decision-making.

References

[1] Li CM, Wang YM, Tian F, Zhou XK, Cui P. (2018) Cluster Computing[j].First Online: https://doi.org/10.1007/s10586-018-1698-x

[2] Liu SL, Tang YH, Zhang FW, Du YG, Lin L, Li YK, Guo XW, Li Q, Cao GM. (2017) Changes of soil organic and inorganic carbon in relation to grassland degradation in Northern Tibet[j], Ecological Research, May 2017, Volume 32, Issue 3, pp 395–404

[3] Su XK, Wu Y, Dong SK, Wen L, Li YY, Wang XX. (2015) Effects of grassland degradation and re-vegetation on carbon and nitrogen storage in the soils of the Headwater Area Nature Reserve on the Qinghai-Tibetan Plateau, China[j], Journal of Mountain Science, May 2015, Volume 12, Issue 3, pp 582–591
[4] Ma L, Yao ZS, Zheng XH, Zhang H, Wang K, Zhu B, Wang R, Zhang W, Liu CY. (2018) Increasing grassland degradation stimulates the non-growing season CO2 emissions from an alpine meadow on the Qinghai–Tibetan Plateau[j], Environmental Science and Pollution Research, September 2018, Volume 25, Issue 26, pp 26576–26591

[5] Zhou WH, Xiong SQ. (2013) Optimization of BP Neural Network Classifier Using Genetic Algorithm[j], Intelligence Computation and Evolutionary Computation, 2013, pp 599-605

[6] Hao G. (2014) Study on Prediction of Urbanization Level Based on GA-BP Neural Network[j], Proceedings of the 21st International Conference on Industrial Engineering and Engineering Management 2014, pp 521-524

[7] He F, Zhang LY. (2018) Mold breakout prediction in slab continuous casting based on combined method of GA-BP neural network and logic rules[j], The International Journal of Advanced Manufacturing Technology, April 2018, Volume 95, Issue 9–12, pp 4081–4089

[8] Cheng Y, Wu R. (2016) The Research of Aviation Dangerous Weather Forecast for Fog and Haze Based on BP Neural Network[j], Proceedings of the 5th International Conference on Electrical Engineering and Automatic Control pp 877-883