Mining of Credit Risk Factors for New Agricultural Entities under the Big Data of Agricultural Economy in Jilin Province

Chunhui Wang

Jilin Business and Technology College, JiLin ChangChun,130507 China

wangchunhui920@126.com

Abstract. The excavation of credit risk factors for new agricultural entities plays an important role in providing comprehensive credit evaluation and promoting stable growth of the agricultural economy. The current traditional mining methods lack objectivity and accuracy, making it difficult to provide effective support for the credit evaluation of new agricultural entities. In order to improve the effectiveness of credit risk factor mining and promote the healthy development of agricultural economy, this article takes Jilin Province as the research object and combines agricultural economic big data to conduct in-depth research on the mining of credit risk factors for new agricultural entities. This article first analyzes the current situation of agricultural economy and the development of new agricultural entities in Jilin Province. Then, computer technology is used to process big data of Jilin Province's agricultural economy. Finally, based on the big data of agricultural economy, credit risk factors of new agricultural entities are excavated. And experimental analysis was conducted on credit risk factors. The results show that moral credit, business scale, market demand, and environmental risks have a key impact on the credit risk of new agricultural entities. In the analysis of impact degree, compared with the default rate in January, under the influence of these four factors, the default rate in December increased by 6.65%, 6.45%, 4.84%, and 5.54%, respectively. The conclusion indicates that agricultural economic big data can objectively explore the credit risk factors of new agricultural entities and achieve more accurate information risk assessment.

Keywords: Credit Risk Factors, Agricultural Economic Big Data, Support Vector Machine, New Agricultural Entities, Jilin Province

1 INTRODUCTION

With the continuous advancement of agricultural modernization, traditional agricultural production and operation methods in Jilin Province are constantly transforming, and new agricultural entities are constantly emerging [1-2]. While playing a crucial role in promoting the development of rural economy in Jilin Province, its credit risk is also increasingly prominent. This not only affects the credit of agricultural entities, but also

A. Haldorai et al. (eds.), *Proceedings of the 2024 3rd International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2024)*, Atlantis Highlights in Intelligent Systems 11,
https://doi.org/10.2991/978-94-6463-490-7_35

development of the agricultural economy. Effectively identifying and mining credit risk factors for new agricultural entities is an important direction for current agricultural development [3]. With the continuous improvement of digital technology, agricultural economic big data has made significant progress and its application scope is becoming wider and wider. This article takes Jilin Province as the research object and applies agricultural economic big data to explore the credit risk factors of agricultural entities, revealing their mechanisms and influencing factors. This has important practical value for improving the credit risk prevention and control level of new agricultural entities and promoting the sustainable development of agriculture in Jilin Province.

The credit risk of new agricultural entities has a significant impact on their credit level and development. Exploring the credit risk factors of new agricultural entities can reduce credit risk [4-5]. Martinez-Victoria applied a partially adjusted model to a sample of 11930 Spanish agricultural food cooperatives and agricultural food investor owned companies between 2011 and 2018, and explored the factors affecting the trade credit business of agricultural food cooperatives using a dynamic panel model and a two-step general moment method. The survey results show that external funding sources, positive cash flow, and essential operating goods are key factors affecting the trade credit of large cooperatives [6]. Based on survey data from 226 sample cooperatives in Guizhou Province, Liang Genhong used factor analysis to measure the operational credit risk of agricultural cooperatives in the planting industry. He used structural equation modeling to empirically test its influencing factors. Research has shown that the inherent factor of operator characteristics is the core factor affecting the credit risk of plantation agricultural cooperatives, which determines the development prospects of cooperatives [7]. Kumar Anil used a systematic literature review method to empirically review the identification method of credit risk factors for new agricultural entities based on machine learning. The results indicate that machine learning intelligent algorithms are suitable for credit evaluation of agricultural entities and can provide effective basis for credit of financial institutions [8]. In order to explore the impact of different factors on the credit risk of small farmers, Javed Iqbal conducted a survey using the least squares method on 200 farmers from Sargodha district. According to the research results, other variables such as age, family members, and land ownership have a positive and significant impact on the credit risk of farmers [9]. Qi Hao constructed a digital transformation index for rural financial institutions based on web crawler technology, and used data from 1700 rural financial institutions to explore the impact of their digital transformation on the credit of agricultural entities and the quality and efficiency of rural financial services. Research has found that digital transformation has significantly promoted the allocation of agricultural loans by rural financial institutions, expanded credit coverage, and optimized the proportion of credit and medium - to long-term loans [10]. The current methods for mining credit risk factors can provide some reference for credit evaluation of new agricultural entities, but there are still limitations in factor classification and objectivity in mining, making it difficult to meet the practical development needs of agricultural economy.

In order to improve the credit risk prevention and control level of new agricultural entities and promote the healthy development of agricultural economy, this article takes Jilin Province as the research object and combines agricultural economic big data to study the mining of credit risk factors of new agricultural entities. To verify the effectiveness of the method proposed in this article, an empirical survey collected 1068 samples of agricultural new subject credit in Jilin Province in 2023 as the dataset. It conducted experimental analysis from several aspects, including weight ranking, impact analysis, and prediction accuracy evaluation. At the level of weight ranking, the relative weight values of the four factors of moral credit, business scale, market demand, and environmental risk are higher. At the level of impact analysis, compared with the default rate in January, the default rate in December increased by 6.65%, 6.45%, 4.84%, and 5.54% respectively under the influence of four factors. In terms of prediction accuracy evaluation, compared to the random forest method, this method has improved the average accuracy of credit risk prediction for agricultural entities by 5.65%. In practical applications, agricultural economic big data can provide effective data support for mining credit risk factors of new agricultural entities, and help promote high-quality development of agricultural economy.

2 EXPLORATION OF CREDIT RISK FACTORS FOR NEW AGRICULTURAL ENTITIES

2.1 Research Area

2.1.1 Development of New Agricultural Entities

Jilin Province has a long agricultural tradition and plays an important role in its economic development. Especially after the implementation of the revitalization strategy for the old industrial base in Northeast China, the utilization area of arable land in Jilin Province has reached 16.41 million hectares. With policy support, the scale economy of agriculture in Jilin Province has developed rapidly. At the same time, Jilin Province also pays great attention to the development of new entities such as family farms, rural cooperative organizations, and leading enterprises in rural industrialization, and vigorously cultivates high-quality new professional farmers to drive the development of agricultural economy and promote income growth for farmers. In 2023, the agricultural economy in Jilin Province achieved rapid development. Its grain production has exceeded 80 billion kilograms for three consecutive years. It is expected that by 2024, Jilin Province can carry out a three-year action to develop and strengthen new agricultural entities, striving to cover 70% of land transfer and agricultural production custody services. The development status in 2023 is shown in Table 1:

Table 1. Development Status of Rural Economy in Jilin Province in 2023

Sequence	Item	Specifications
1	Total grain output	83.73 billion pounds
2	Value added of total grain production	An increase of 2.114 billion pounds compared to the previous year
3	Grain yield per unit	958.2 kilograms per acre
4	Value added of primary production	5% year-on-year growth

5	Value added of fixed asset investment in the first industry	62.7% year-on-year growth
6	Ranking of major grain producing provinces	First place

From Table 1, it can be seen that the agricultural economy in Jilin Province is developing well in 2023, with high growth rates in grain production and the development of the primary industry.

2.1.2 Credit Risk of New Agricultural Entities

At present, the operation of new agricultural entities in Jilin Province requires a large amount of funds, but due to the lack of a unified and acceptable credit risk factor mining system, their ability to obtain funds is insufficient, which has become a bottleneck restricting their further development. In addition, most rural credit systems in Jilin Province are not sound, resulting in small scale and chaotic management of new agricultural entities, lack of mortgage guarantees, agricultural insurance, etc., which is prone to moral hazard, adverse selection and other phenomena.

Compared with traditional farmer production, the production of new agricultural entities is more large-scale and socialized, and their modernization level is also more significant [11]. At the same time, they also face different credit risks, which include both the credit risk characteristics of ordinary small farmers and the credit risk characteristics of large-scale companies. On the one hand, the new agricultural subject production model has a positive and negative promoting effect on the risks faced by farmers in production. For example, in specific production processes, because more production factors are added, the risks faced are also more diverse. On the other hand, in the production process, new agricultural management entities have established long-term cooperative relationships, which in a sense helps to enhance the stability of their production operations. However, once a problem occurs at a certain stage of the production process, it is likely to affect the business activities of upstream and downstream cooperative entities in the supply chain.

Therefore, exploring the credit risk factors of new agricultural entities is of great significance. By understanding credit risk factors, they can be effectively identified and managed, thereby improving the accuracy of credit evaluation. This can achieve the sharing and transmission of credit information, strengthen the punishment of dishonesty, and promote the healthy development of the new agricultural subject credit system.

2.2 Mining Credit Risk Factors under Agricultural Economic Big Data

Agricultural economic big data is the mining of massive data that appears in the agricultural production process, discovering patterns from it, and improving decision-making efficiency. This article explores the credit risk factors of new agricultural entities based on big data of agricultural economy, and represents different levels of credit risk through credit risk factors.

Before mining credit risk factors, it is necessary to first analyze a large amount of agricultural economic research data, identify key factors that affect agricultural credit, and accurately evaluate the credit risk of agricultural entities. This article utilizes computer technology to effectively process various agricultural production data and store them in a database, laying a solid foundation for further analysis of credit risk factors.

The collection of agricultural economic data mainly includes data on agricultural production, prices, soil quality, policy changes, and other aspects. This article first uses computer and web crawler technology to extract statistical data released by the agricultural department and operational data of new agricultural entities. Using class libraries in programming languages to automatically crawl agricultural data sources. On the basis of data collection, clean the collected raw data. Firstly, remove duplicate data, then use the drop_duplicates() function in the Pandas library to perform data deduplication, remove missing data, and finally convert agricultural economic data into the same format; In data storage, establish a database table structure through a database management system, and store the preprocessed data in the database.

After using computer technology to process agricultural economic data, describe risk factors, which are represented as $f = g(m, n)$ in this article. m, n represent the likelihood and degree of credit risk occurrence, as shown in Figure 1:

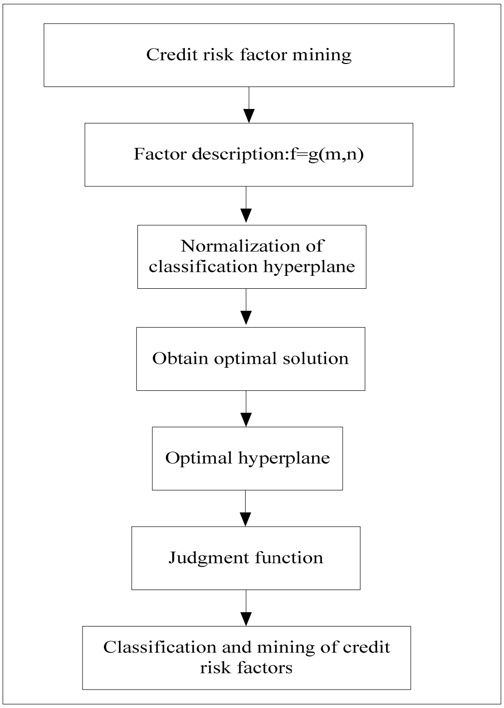

Fig. 1. Mining credit risk factors

In Figure 1, using support vector machine optimization problem, the classification hyperplane of risk factors is standardized. Let $\Delta = 1$ compress k and b according to corresponding coefficients and constrain the corresponding support vectors. This process is represented by the formula [12-13]:

$$\begin{cases} k \cdot x + b = 1, if\ y = 1 \\ k \cdot x + b = -1, if\ y = -1 \end{cases} \tag{1}$$

Set the classification hyperplane distance of the support vector on the factor to $1/\|k\|$. Factors are divided by linear programming functions with certain constraints [14]:

$$\min_{k,b} \frac{1}{2} \|k\|^2 \tag{2}$$

$$z.p. y_i(k \cdot x_i + b) \geq 1, i = 1, \cdots, r \tag{3}$$

According to optimization principles, solve linear programming functions with certain constraints. Its Lagrangian function is expressed as:

$$R = \frac{1}{2} \|k\|^2 + \sum_{i=1}^{r} \delta_i [1 - y_i(k \cdot x_i + b)] \tag{4}$$

Among them, the constraint condition is $y_i(k \cdot x_i + b) \geq 1$, and $\delta_i \geq 0$ represents its Lagrange multiplier. Using the Karush Kuhn Tucker (KKT) condition to normalize the optimization problem at the optimal solution [15]:

$$\frac{\partial}{\partial k} R = k - \sum_{i=1}^{r} \delta_i\, y_i x_i \tag{5}$$

Using the Wolf duality principle to solve optimization problems:

$$\max_{l} k(\delta) = \sum_{i=1}^{r} \delta_i - \frac{1}{2} \sum_{i,j=1}^{r} \delta_i \delta_j\, y_i x_i \tag{6}$$

$$z.p. \sum_{i=1}^{r} \delta_i\, y_i = 0, \delta_i \geq 0, i = 1, \cdots, r \tag{7}$$

By solving the above equation optimally, the corresponding optimized hyperplane can be obtained. If δ^{Δ} represents the best solution, then δ^{Δ} is generally 0; If δ^{Δ} is not 0, then the vector corresponding to it is the support vector, and the linear set of training sample vectors obtained is represented as $k^{\Delta} = {\delta^{\Delta}}_i y_i x_i$. Randomly select the support vector x_i for further determining b^{Δ}: $b^{\Delta} = y_i$-$k^{\Delta} \cdot x_i$, and finally, the obtained judgment function is shown in the formula:

$$g(x) = log\ s(k^{\Delta} \cdot x + b^{\Delta}) = log\ s(\sum_{i=1}^{r} {\delta^{\Delta}}_i y_i x_i \cdot x + b^{\Delta}) \tag{8}$$

According to the classification of credit risk factors, factor mining can be achieved and effective prediction of credit risk can be made.

3 EXPERIMENT OF CREDIT RISK FACTORS FOR NEW AGRICULTURAL ENTITIES

3.1 Experimental Data

To verify the effectiveness of mining credit risk factors for new agricultural entities under the big data of agricultural economy, this article collected 1068 samples of credit for new agricultural entities in Jilin Province in 2023 through computer and web crawlers as the dataset. Preprocess the data and classify the new agricultural management entities into three categories based on industry attributes: planting entities, animal husbandry entities, and agricultural and sideline product processing entities. Business entities that have not experienced overdue repayment or breach of contract can be recognized as non defaulting samples and recorded as 0. The operating entity that fails to repay the loan on time or experiences overdue repayment within a given time period can be identified as a default sample, recorded as 1. Among them, there are 583 samples that have not defaulted, 485 samples that have defaulted, and the default rate of the samples is about 45.4%.

3.2 Experimental Results

The factors that cause credit risk in new agricultural entities can be divided into subject factors and object factors. Among them, the subject factor depends on the moral credit, management ability, and business scale of the new agricultural subject. Object risk factors include market demand, policies and regulations, natural disasters, and environmental risks. Under the credit risk factor mining algorithm based on agricultural economic big data, the factor weights were sorted, and the results are shown in Table 2:

Table 2. Factor Weight Sorting

Sequence	Factor	Relative Weight (%)
1	Moral credit	21.03
2	Management ability	5.88
3	Business scale	17.34
4	Market demand	19.01
5	Policies and regulations	7.82
6	Natural calamities	6.11
7	Environmental risks	22.81

From Table 2, it can be seen that the relative weight values of the four factors of moral credit, business scale, market demand, and environmental risk are relatively high, reaching 21.03%, 17.34%, 19.01%, and 22.81%, respectively.

Based on the sorting results in Table 2, this article constructs a partial correlation graph of four factors: moral credit, business scale, market demand, and environmental risk, to deeply verify the impact of these four factors on the credit risk of new agricultural entities in the operating time series, as shown in Figure 2:

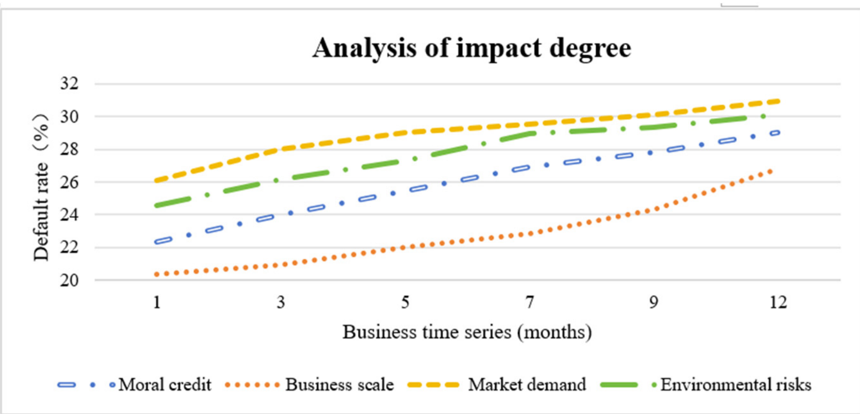

Fig. 2. Analysis results of impact degree

From Figure 2, it can be seen that with the passage of operating time, the impact of various factor variables on the credit risk of new agricultural entities gradually increases, and the overall impact of each factor shows an upward trend. From the specific analysis results, in January, under the influence of four factors: moral credit, business scale, market demand, and environmental risk, the default rates of new agricultural entities were 22.36%, 20.34%, 26.12%, and 24.55%, respectively. By December, under the influence of four factors, the default rates of new agricultural entities reached 29.01%, 26.79%, 30.96%, and 30.09%, respectively. Compared with the default rate in January, the default rate in December increased by 6.65%, 6.45%, 4.84%, and 5.54%, respectively.

To further verify the objectivity of the mining effect of credit risk factors for new agricultural entities in this article, the accuracy of the paper's method for credit risk prediction was evaluated using the true values of the data samples as a reference, and compared with the traditional Random Forest (RF) method. The final results are shown in Figure 3:

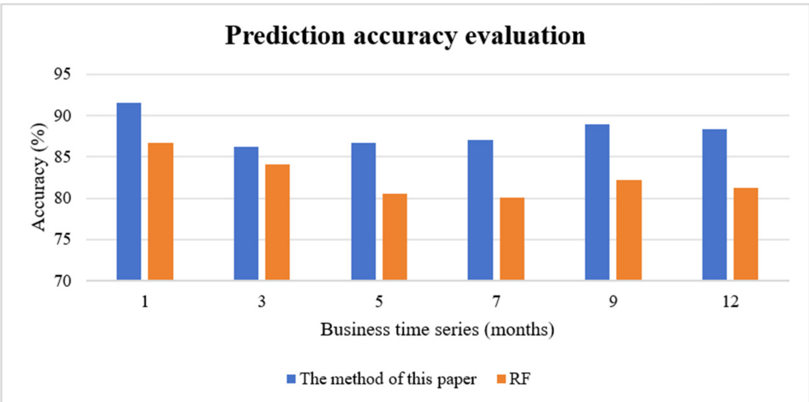

Fig. 3. Prediction accuracy evaluation results

From Figure 3, it can be seen that the accuracy of credit risk prediction for new agricultural entities based on agricultural economic big data in this article has significant advantages. Under the operating time series of new agricultural entities, the credit risk prediction accuracy of the method proposed in this article can reach up to 91.60%, with an average prediction accuracy of about 88.13%. The accuracy of credit risk prediction under traditional RF methods can reach up to 86.70%, with an average prediction accuracy of approximately 82.48%. From the specific comparison results, compared to the RF method, the method proposed in this paper has improved the average accuracy of credit risk prediction for agricultural entities by 5.65%.

4 DISCUSSIONS

In the experimental analysis, this article tested the mining effect of credit risk factors for new agricultural entities under agricultural economic big data from the aspects of weight ranking, impact degree analysis, and prediction accuracy evaluation. At the level of weight ranking, compared to other types of factors, the relative weight values of moral credit, business scale, market demand, and environmental risk are relatively higher. In the analysis of the degree of impact, under the influence of four types of factors, the default rate of new agricultural entities has shown a significant upward trend with the passage of operating time. This indicates that four factors, namely moral credit, business scale, market demand, and environmental risk, have a key impact on the credit risk of new agricultural entities. In the evaluation of prediction accuracy, compared to the RF method, the method proposed in this paper has a more ideal average accuracy result for predicting credit risks of agricultural entities. This indicates that with the support of agricultural economic big data, the paper's method can more accurately and objectively mine risk factors and achieve effective prediction of credit risks for new agricultural entities.

5 CONCLUSIONS

With the vigorous development of agriculture in Jilin Province, the credit risk issues of new agricultural entities are becoming increasingly prominent. Traditional methods for mining credit risk factors lack effective data support, making it difficult to objectively and accurately mine and evaluate credit risk factors. In order to improve the objectivity and accuracy of risk factor mining and achieve good prediction of credit risk for new agricultural entities, this article combines agricultural economic big data to conduct in-depth research on the mining of credit risk factors for new agricultural entities in Jilin Province. It not only effectively classifies and identifies key factors, but also objectively analyzes their impact, achieving accurate prediction of credit risks for new agricultural entities. From experimental analysis, it can be seen that the four factors of moral credit, business scale, market demand, and environmental risk have a significant impact on the credit risk of new agricultural entities. In the future development of new agricultural entities, it is necessary to improve from the perspective of their own moral credit and business scale, while paying attention to the impact of market demand and environmental risks, and promoting the healthy development of the agricultural economy. Although this study has certain guiding significance for the sustainable development of agricultural economy in Jilin Province, there are still limitations in this article.

TOPIC INFORMATION

Scientific Research Project of Education Department of Jilin Province in 2024. Research on Credit Enhancement Mechanism and Countermeasures of Supply Chain Financing for New Agricultural Management Subjects in Jilin Province

REFERENCES

1. Tang Mingxia, Yuan Chunxin, Zhang Huayin, Gu Yongjian, Zhou Tiejun, Zhang Haiyan. "An analysis of the needs of new agricultural business entities for agricultural financing guarantee policies to benefit farmers--A study based on Maslow's hierarchical theory of needs."Anhui Agricultural Science 50.16 (2022): 179-181.
2. Wu Shunchen, Hu Lingxiao, and Chu Yifei. "The classification and support strategy of new agricultural business entities--Based on the combing of literature and the thinking of the "sub-subject support" policy."Western Forum 29.6 (2019):53-59.
3. Wu Chenghao. "Research on the problems and countermeasures of the development of new agricultural business entities in the context of rural revitalization."Food science and Technology and Economy 44.2 (2019): 155-159.
4. Ni Xu, and Zhang Qiao. "Research on the credit evaluation index system of China's new agricultural business entities."Agricultural Outlook 15.6 (2019): 23-27.
5. Wu Yinzi. "Cultivation of new agricultural business entities: Status quo, Problems and path innovation."Modern agricultural Research 28.2 (2022): 12-14.
6. Martinez-Victoria, MCarmen, and Mariluz Mate-Sanchez-Val. "Determinants of trade credit financing: a dynamic analysis comparing agri-food cooperatives and non-cooperatives." Agricultural Finance Review 82.5 (2022): 890-913.
7. Liang, Genhong, Zhuo Zhang, Peirong Wu, Zhijie Zhang and Xiwu Shao. "Analysis of Business Risk Measurement and Factors Influencing Plantation-Based Farming Cooperatives: Evidence from Guizhou Province, China." Sustainability 16.5 (2024): 2194-2214.
8. Kumar, Anil, S. Sharma, and M. Mahdavi. "Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review." Risks 9.11 (2021): 192-206.
9. Javed, Iqbal, Ghulam Yasen, Sana Rashid, Abdul Subhan, Shoaib Akhtar, Shahbaz Ahmed, et al. "Impact of agricultural credit on farmer's income: evidence from central Punjab, Pakistan." International Journal of Agricultural Extension 11.3 (2023): 215-224.
10. Qi Hao, Zhou Yueshu, and He Lifeng. "The digital transformation of rural financial institutions and the quality and efficiency of financial services."Financial Research 50.1 (2024): 64-78.
11. Wang, Zhao, Jianhong Liu, Tongsheng Li, Jing Chao, Xupeng Gao. "Factors affecting new agricultural business entities' adoption of sustainable intensification practices in China: evidence from the main apple-producing areas in the Loess Plateau." Agronomy 11.12 (2021): 2435-2457.
12. Nalepa, Jakub, and Michal Kawulok. "Selecting training sets for support vector machines: a review." Artificial Intelligence Review 52.2 (2019): 857-900.

13. Tavara, Shirin. "Parallel computing of support vector machines: a survey." ACM Computing Surveys (CSUR) 51.6 (2019): 1-38.
14. Raj, Jennifer S., and J. Vijitha Ananthi. "Recurrent neural networks and nonlinear prediction in support vector machines." Journal of Soft Computing Paradigm (JSCP) 1.1 (2019): 33-40.
15. Avci, Cengiz, Budak, Muhammed, Yagmur, Nur, Balcik, Filiz. "Comparison between random forest and support vector machine algorithms for LULC classification." International Journal of Engineering and Geosciences 8.1 (2023): 1-10.