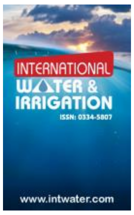


International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

Market Trend Prediction of Digital Economy Based on Machine Learning Algorithm

Chengxia Li[1*]
[1]School of Dean's office, Qinghai Open University, Xining, Qinghai, 810007, China
[1*]Corresponding author e-mail: licx0527@163.com

Abstract: The focus of digitization on helping enterprises, organisations, and governments accomplish their sustainability goals is a defining feature of the contemporary economic landscape. Machine Learning has recently been included into company models and plans along with the use of the newest technology, which has allowed the environment for business in the digital economy will get even better. The digital economy is extensively covered, extremely innovative, and highly porous. It serves as both a fresh hub for economic growth and a pivot for modernising and changing established sectors. The digital economy has the power to provide new employment opportunities, increase consumer demand, and boost investment activity. It is a crucial component in building a contemporary economic structure. Based on data pertaining to the raise of the digital economy in Zhejiang Province, China, between 2014 and 2023, this study finds that taking two samples—the bare minimum needed for node splitting—gets the best results. The research data is then analysed to create a Multi Linear based Random Forest Regression Algorithm (ML-RFR), where training data comprises 70% and testing data has 30%. The study concluded that rather than heedlessly following economic trends, the expansion of the digital economy necessitates raising the standard of the sector's digital economy, extending its level of development, and optimising its development model.

Keywords: Digital Economy, Market Trends, ML-RFR, regression, Machine Learning

1. Introduction

The term "digital economy" describes the financial activity resulting from billions of every day interactions between people, businesses, devices, data and processes that take place online. It includes every facet of the economy made possible by digital technologies, such as digital payments, online advertising, e-commerce, and more. The digital economy is going to see a major upheaval as artificial intelligence grows. On the other hand, financial crises, economic expansion, and the advancement of digital technologies have put significant pressure on national and global economies, particularly on the deficit in the budget, banking and financial services, and the corporate environment. These factors have also had an impact on the earnings and revenues of economic participants [1]. The general direction of the price of an asset or market value over a given period of time is referred to as a market trend. Generally speaking, they are classified as horizontal, upward, or downward trends based on the direction of the movement. Trend analysis is the process of identifying and analysing these patterns using indicators of technology and charts to potentially forecast price movements in the future.

The digital economy was first conceived as an idea for economic growth. It was an effective way for people to use big data to lead them in realising the fastest, most efficient way to allocate resources in order to produce high-quality economic growth [2]. A digital economy is an economic system that, in the process of developing, uses data, either directly or indirectly, to drive resources to fulfill their roles and encourage the growth of productive forces. Since the digital economy encompasses a broad spectrum of industries, advancements in manufacturing, technology, and production are possible. High-quality development is achievable in the digital economy by encouraging entrepreneurial activity, realising big data and cloud computing, as its size grows and injects new dynamic energy into economic progress [3].

One of the greatest inventions ever made is the financial market, which is essential to the functioning of the entire economic system. The technique of projecting and ascertaining the future values of a company's stock is known

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

as stock market prediction. An important turning point in every nation's thriving and expanding economy is the stock market. Since it has a significant impact on trading strategies and helps sellers and brokers make significant profits by market predicting behaviour and making the right choice about whether to hold or sell their stocks, accurate forecasting is of businessmen and financial experts are very interested. However, because the stock market is chaotic and dynamic, it can be difficult to conduct accurate forecasts. Numerous cutting-edge ML techniques have been trained and tested for the prediction of market prices during the past few decades. They have been demonstrated to be significantly more effective since the machine learning techniques accurately accomplish the task of making predictions by closely examining past data [4]. While constructing and evaluating machine learning models has been the focus on multiple studies and publications, this article aims to give a summary of the outcomes of machine learning predictions. It can assist individuals in developing a general understanding of different machine learning techniques. They offer a useful point of reference for future developments in machine learning models for stock market forecasting. In the digital economy, the creation, provision, and exchange of goods and services are significantly influenced by intangible assets like information and communication technologies [5], [6]. One area of the global economy that is expanding quickly is the digital economy. It is projected that the digital economy would contribute more than 20% of the world economy by 2025. In today's digital economy, market trends have a significant impact on market dynamics in addition to traditional economic indicators like GDP growth, inflation, and unemployment [7]. Data and information may now be exchanged in real time, and the quick movement of information components constantly improves the elements' spatial mobility. Building a numerical power is a goal that China and even other nations across the world strongly value. They advance favourable policies and relevant implementation tactics to aggressively promote their own economic growth [8].

Government and people, government and market have all been able to integrate through the use and integration of digital information technology. Government resource sharing and information exchanges will result in more effective and convenient digital government services for citizens and businesses, as well as more equitable and effective institutional conditions and development environments for the market. They will also enable more efficient distribution of resources for production factors. The government's governance structure and management approaches in the age of the digital economy can thus adjust, develop, and integrate the digital economy in order to ultimately support the efficient, steady, and long-term growth of the social economy. Zhejiang Province, China, is used as an example of the digital economic development level in this study. It primarily looks at the effects of the digital economy's growth while also anticipating how the Chinese province of Zhejiang would see its own development in this regard.

2. Literature review

The phrase "digital economy" mainly describes the latest and significant changes in the economic framework, whereby many economic sectors are now realized through procedures and processes based on the most recent digital innovations produced by automated technology data digitalization in the digital age [9]. Since this explanation involves a significant alteration of the economic mechanism rather than merely suggesting modifications to the traditional rules of economic construction, it has gained widespread acceptance among economists. This term primarily refers to the way digital technologies are embraced, applied, and utilised in traditional economics [10]. Generally speaking, as economists' primary definition of AI is the explanation of automation, robots, or machines, different types of scientific specialists may understand AI differently. However, for an engineer who views automation as the technology that enables processes or procedures to operate with little to no human intervention, this is a fairly limited definition. However, since artificial intelligence is described as "a growing resource of interactive, autonomous, self-learning agency, which enables computational art facts to perform tasks that otherwise would require human intelligence to be executed successfully" [11] still implies more than simply digitizing tasks or using computers and other digital tools for work. At the same time, technologies related to AI have been growing and changing very quickly. So, the idea that AI was only about robots, automated processes, or machines has grown even more. Now, software and hardware platforms interact more, and new

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

technologies like virtual assistants, DL machines, machine vision and image recognition, intelligent robots, and NLP are being developed. Technologies that people and businesses use every day are bringing to light the foundations of the digital economy. These technologies are slowly changing the business world and "forcing" the economy to move regarding more digital methods and procedures [12].

"Digital transformation" means incorporating modern technologies into all parts of a business's operations. Since digital technologies are becoming more popular, traditional business strategies will have to be changed into business models that make the most of the business environment. This is because the digital economy is changing the way things are done [13]. Machine Learning, Big Data, and the newest technologies are now widely utilized in enterprises and society development in order to boost revenue growth and meet sustainability goals. This means that AI is no longer an uncommon or new idea; it is one that is always changing and growing [14]. Since these factors have been considered, governments ought to allocate additional funds to completely develop AI technologies in order to spur economic growth, particularly in light of the convenience that AI brings [15]. Author has come up with a machine learning-based method for predicting stock trends [16], with a focus on making sure that the files they get have as little missing data as possible. Compared to the initial investigations, the results show that the suggested method works.

This paper developed an empirical modeling technique based on genetic programming that made it possible to forecast economic growth using survey data that was predicted [17]. In particular, the author estimates the symbolic regression using an evolutionary method. This establishes a connection between the survey expectations and the numbers and factors that serve as the standard, allowing the author to infer the mathematical function form that is comparable the variable to target. An empirically-derived a group of economic expansion markers serves as a foundation for forecasting GDP progression. Furthermore, the author assessed the effect of the 2008 crisis in financial on the anticipated higher performance of the evolution of economic activity using GDP estimates.

The approach uses the KM estimation method to calculate the survival function [18]. Additionally, This paper provided theoretical evidence for the consistency of random survival forest (RSF) and concluded that, when applied to high-dimensional data, RSF performs noticeably better than other survival analysis techniques [19]. This recommended adding NCL technology to the random forest algorithm, mostly for the training set that isn't balanced. They recommended that the data be processed using NCL technology and that the results of the processing be put into groups using the random forest algorithm. The test results make it clear that the improved RF method is being used. The forest method does a better job of classifying things [20]. Random forest method to deal with the problem of uneven dataset classification by using a price-sensitive learning algorithm. The cost-sensitive random forest technique makes a cost-sensitive decision tree by first using a different random sampling method on the training set to make many bags. Then, it picks at random some traits from each bag to get rid of. Putting together the bags makes an ensemble algorithm [21]. By examining the connection between each random forest tree's strength and correlation coefficient, Specifically, when constructing a random forest, a reduction in the random forest's generalisation error can be attained by strengthening the decision tree within the forest. However, they also discovered that the correlation between decision trees should be reduced as the strength increases [22].

3. Methodology

3.1 Application of ML On The Digital Economy

In the upcoming years, machine learning's influence on the digital economy will only increase. Companies who can effectively utilise artificial intelligence (AI) will have a strong advantage in the emerging economy. These are a few instances from the digital economy that are currently in use. Figure 1 show Industry Applications of Advanced Technologies.

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

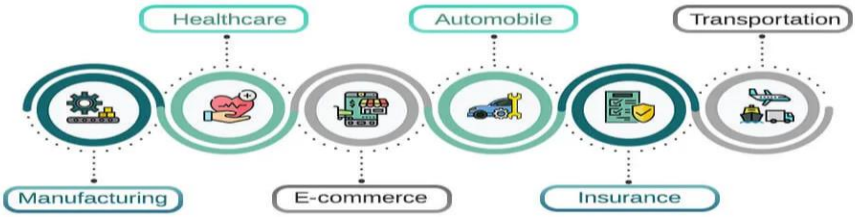


Figure 1: Industry Applications of Advanced Technologies

All things considered, AI is significantly affecting the digital economy. It is increasing production and efficiency, raising customer satisfaction, reaching a wider audience, upending established sectors, quickening innovation, stimulating economic growth, and tackling global issues. Companies who can effectively utilise artificial intelligence (AI) would have a strong advantage in the emerging economy [23].

3.2 Multi Linear based Random Forest Regression(ML-RFR) Algorithm

Figure 2 showing Proposed flow chart

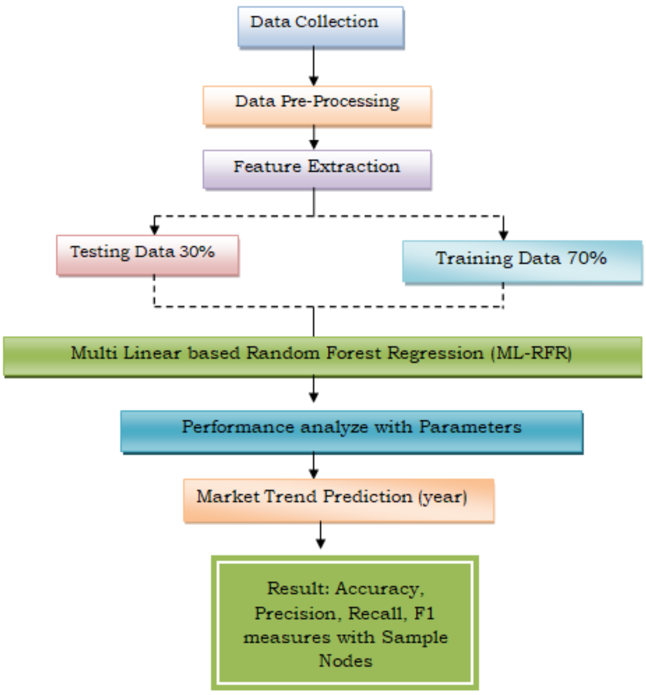


Figure 2: Proposed flow chart

A popular technique for predictive analysis that clarifies the link between continuous variables is multilinear regression. The term "linear regression" originates from the fact that it shows that the independent variable is linked in a straight line(on the X-axis) and the dependent variable (on the Y-axis). The definition of multi linear regression is when there is just one input variable (x). On the other hand, multi linear regression is employed when there are several input variables. In essence, the linear regression model shows the relationship between variables as a straight line with a preset slope. Upon establishment of the random forest, every decision tree within the forest conducts choice of grouping to determine which group is most frequently choosen, hence enabling the prediction of the class to which the new sample will belong. Repeated sampling is a possibility since the random forest will replace the rows and columns in the input data set to randomly sample it. Each decision tree needs m sample sets to be trained, assuming there are m decision trees. Using complete samples to train m trees is not recommended as this will cause whole sets to disregard the sample set, which will weaken the model's capacity for generalisation. The prediction classification is eventually obtained after n samples are extracted using replacement, m decision trees are trained using these n samples, and so on. The chain rule equation 1 is typically used to calculate many random variables.

$$p(_{x1,x2.....xn}) = p(_{x1}) U \ Px(xi|x1 \ldots\ldots xi - n) \qquad (1)$$

Because of its great operational efficiency, ease of use, and simplicity in implementation, random forests require little in the way of mathematical or statistical expertise from its users. Apart from the aforementioned benefits, random forest exhibits a robust over fitting characteristic and can be concurrently employed for the advantages of classification and regression tasks. Nevertheless, the regression problem's use of random forest yields less favourable results than the classification problem. When addressing the regression issue, it is unable to provide predictions that extend outside the range of data in the training sampling set due to its inability to produce a continuous output. It is common for data to overfit. Because random forests are better at handling high-dimensional and unbalanced data sets, their predictive power will be diminished for low-dimensional data sets.

The model will perform worse when it comes to learning if the generalisation economy error is high, and higher if it is the opposite. Equation 2 provides a definition for the generalisation error.

$$Rexp(f) = Ep[L(Y, f(X)\ p(x, y)\ dxdy] \qquad (2)$$

When considering external factors, the primary ones are the training sample's size, the amount and kind of categories, and the imbalance sample's data. In general, an algorithm's classification improves with increased separation accuracy and index formula is presented in Equation 3 below.

$$I(y_i, y_i) = \frac{(y_i \cap y_i)}{|y_i \cup y_i|} \qquad (3)$$

The sample prediction scenario yields the algorithm's overall performance on the test set. The coefficient is 1 if the actual circumstance is entirely consistent with the anticipated outcome. Equation 4 in this work treats the multi linear regression indicators as follows:

$$z_{ijt} = \frac{x_{\mathrm{ijt}} - \min(x_{ijt})}{\max_{(x_{ijt})} - min((x_{ijt}))} \qquad (4)$$

Equation (5) computes the proportion of the standardised metrics.

$$\phi_{ijt} = \frac{z_{ijt}}{\sum_{j=1}^{N} \sum_{t=1}^{T} z_{ijt}} \qquad (5)$$

Equation (6) computes the information digital economy entropy of index i.

$$e_i = -\frac{1}{1n(NXT)} \sum_{j=1}^{N} \sum_{t=1}^{T} \phi_{ijt} \ \mathrm{X}\ 1\mathrm{n}(\phi_{ijt}) \qquad (6)$$

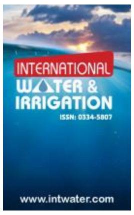
Equation (7) computes the information digital economy entropy redundancy of index i.

$$d_{i=}1\text{-}e_i \qquad (7)$$

Equation (8) computes the market predictions regression i.

$$p_i = \frac{d_i}{\sum_{i=1}^{m} d_i} \qquad (8)$$

Lastly, equation (9) computes the degree of multi linear regression development for every province for every year.



International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

$$MLRFR_{jt}=\sum_{i=1}^{M}\sum_{j=1}^{m}\sum_{t=1}^{m} z_{ijt}\ \mathrm{X}\ w_i \tag{9}$$

Equation 5,6 is based on the market prediction regression p_i and the standardised indicators z_{ijt}. The value of $MLRFR_{jt}$, which ranges from 0 to 1, represents the province j multi linear regression development level in year t.

3.3 ML-RFR Digital Economy Modeling Design

To validate the research hypotheses, we first build the regression model below to show how the digital economy affects sufficient development. Equation 10 gives the model's shape as follows:

$$MLRFR_{nt} = \beta Digital_{nt} + \gamma X_{nt} + C_n + \alpha_t + V_{nt} \tag{10}$$

The variables $MLRFR_{nt}$ and $Digital_{nt}$ _nt represent the level of multi linear based random forest regression and enough growth of province n in year t, respectively. The impact of the digital economy on the multi linear based random forest regression sufficient development is indicated by the regression coefficient β. A number of factors under control is represented by X_{nt}, while C_n and α_t indicate time- and individual-fixed effects, respectively and V_{nt} stands for random disturbance. Then, we investigate whether the intermediary variable is total factor productivity in order to better understand the potential trend of the impact of the digital economy on ML-RFR and sufficient development. Therefore, we construct a regression model of the digital economy impacting overall productivity of factor on the basis of model (10), as well as a multi linear regression model of the digital economy and overall productivity of factor jointly affecting balanced and sufficient development. Next, we determine whether an intermediary impact exists by evaluating the significant judgment and matching regression coefficient. The model's structure is shown in equations 11 and 12 as follows:

$$TFP_{nt} = \beta Digital_{nt} + \gamma X_{nt} + C_n + \alpha_t + V_{nt} \tag{11}$$

$$MLRFR_{nt} = \beta_1 Digital_{nt} + \beta_2 TFP_{nt} + \gamma X_{nt} + C_n + \alpha_t + V_{nt} \tag{12}$$

The level of overall productivity of factor in province n in year t is represented by TFP_{nt}, and the remaining variables' meanings are in line with model (10). The variables as a input are simpler to control than the output variables. Therefore, in order to increase the overall productivity of factor metric, we select the input-driven MLRFR algorithm based on the variable returns. Equation 13 represents the input-oriented MLRFR model in its current form.

$$min\left[\theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r +\right)\right] \qquad (13$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta x_i 0 \qquad \mathrm{i = 1,2 \ldots\ldots\ldots, m}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_r 0 \qquad \mathrm{r = 1,2 \ldots\ldots\ldots, s} \tag{14}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad \mathrm{j = 1,2 \ldots\ldots, n}$$

In 14th equations, the efficiency value is denoted by δ, the non-Archimedean infinitesimal is represented by ε, the input ,output variables are denoted by x and y, respectively, and the weight is represented by λ.We choose the capital stock and labour as the input variables; the fixed asset investment is known as the capital stock that has been determined using the perpetual inventory method, and the labour is the total count of employees. The regional ML-RFR serves as the output indicator. The fixed asset investment and the multilinear regression are translated to the actual value among them.

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

4. Results and Discussion

4.1 Dataset

The data used in this study pertains to Zhejiang Province, China's digital economy growth between 2014 to 2023. ARIMA predictions are used to fill in the missing values in the information that was gathered. To fill in the gaps, the following metrics are used: year-end regularity, employment in associated industries and urban units, and the rate of Internet penetration. The number of residents, the rate at which mobile phones are used, and the university's measure of digital financial inclusion to fit the growth level of the digital economy, constructs a multi linear based random Forest regression model. (In the subsequent study, the independent variables were named y1, y2...y7 in this order). The total telecommunication business is used as the dependent variable. The additional benefit of the tertiary industry is an independent variable. Multi-linear regression is a popular approach for analysing such issues since it is a conventional linear regression model. This study creatively introduces a nonlinear regression model, such as the random forest regression model, to investigate this issue from a nonlinear standpoint, we can obtain expected outcomes.

We employed Y_i, the outcome indicator, as the unifying designation for the model's input parameters. Table 1 displays their starting values: The following is the market trend of the indices of economic development that were examined:

Y1= the equity to efficiency coefficient when utilising one's own assets;

Y2= the value of equity divided by the total capital as the coefficient of financial independence;

Y3= current asset to total asset ratio;

Y4= the net profit to revenue ratio;

Y5= proportion of the economy's assets to revenue differential;

Y6= The percentage of the allocation of the quantity of borrowed money in the quantity of equity representing the ratio of attracted and equity;

Y7= the own funds coefficient of the digital economy expressed as a ratio of equity to own working capital;

Y8= ratio of total assets to profit as a measure of financial resources;

Y9= net profit to cost of sales (goods, works, services) as a percentage;

Z= at the conclusion of the reporting period, the balance sheet shows the digital economy from regular activities before taxes.

Table 1: Input Parameters Initial Values For The Dataset

Parameter	Years									
	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Y1	0.9	0.45	0.52	0.10	0.52	1.09	1.15	0.05	0.06	0.05
Y2	0.10	0.49	0.54	0.11	0.55	0.97	1.23	0.06	0.09	0.07
Y3	0.18	0.52	0.58	0.15	0..43	0.83	1.31	0.10	0.14	0.10
Y4	0.20	0.53	0.60	0.19	0.56	0.85	1.28	0.13	0.18	0.16
Y5	0.15	0.54	0.61	0.20	0.37	0.87	1.42	0.16	0.23	0.04
Y6	0.9	0.54	0.62	0.18	0.45	0.92	1.58	0.18	0.28	0.07
Y7	0.10	0.41	0.63	0.16	0.34	1.32	1.87	0.21	0.18	0.09
Y8	0.35	0.20	0.70	0.10	0.47	1.39	1.92	0.14	0.07	0.15
Y9	0.23	0.40	0.82	0.18	0.73	3.12	2.43	0.06	0.20	0.06
Z	0.15	0.45	0.69	0.15	0.49	1.10	1.46	0.08	0.22	0.07

4.2 Multi Linear based Random Forest Regression (ML-RFR)

After the study data was examined, multi linear regression findings were produced, as the table below illustrates.

Table 2: Regression Analysis For The Values

Intercept	Estimated	Std.Error	P-values <0.05
Y1	0.92	0.18	0.07
Y2	0.65	0.36	0.19
Y3	0.34	0.28	0.23
Y4	0.84	0.34	0.21
Y5	0.23	0.09	0.05
Y6	0.38	0.28	0.24
Y7	0.65	0.41	0.35
Y8	0.43	0.12	0.13
Y9	0.45	0.08	0.10

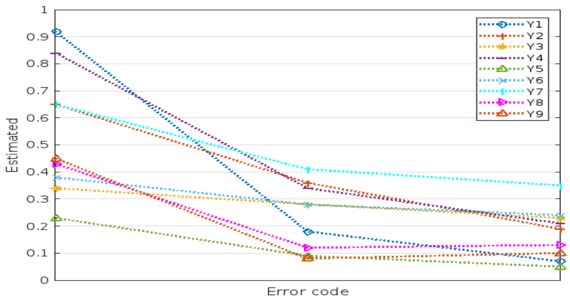


Figure 3: Multi Linear Regression Model Fitting Graph

The multi linear regression analysis results indicate that the model's value is 0.992. This indicates that the following variables are relevant to urban areas: the percentage of permanent residents at the end of the year, the rate of mobile phone penetration, employment in urban areas, university officials, the computer software and IT service sectors, and the transmission of information. 96.7% of the variations in the entire telecoms business can be explained by the financial inclusion index and the tertiary industry's added value, and the multi linear regression model's overall fitting impact is good. Above is the model fitting effect graph.

4.3 Random Forest Regression

The quantification of the features that are chosen will have an impact on the outcome when determining the optimal split point in a decision tree. Significance indicators of the computed categories must be used to choose suitable characteristics. The logarithm value, the value of the square root and the overall amount of characteristics, and the quantity of samples required to divide all the extraction of features nodes are common selection techniques that impact the height of the tree and define the degree of splitting. Analysis of the node selection process is done by experimentation. In Figure 3, the experiments are displayed.

The findings in Figure 4 demonstrate how the each leaf node's purity in the tree is impacted by the minimal quantity of samples required to be added when a new node is formed. The node that remains after splitting will be eliminated if the samples it contains are smaller than this amount. A number that is too high will prevent nodes from discriminating from one another. When two samples are taken the bare minimum needed for two, three, or four node splitting the best results are obtained, guaranteeing that the decision tree is split entirely.

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

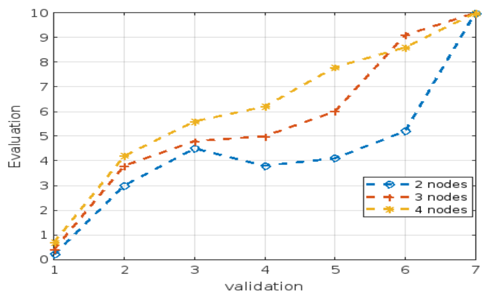


Figure 4: The Limited Quantity Of Samples Required For 2, 3, 4 Node Split

4.4 ML-RFR Market Trend Prediction Value

Table 3: Result For ML-RFR Market Trend Prediction Value

Year	Observed	Predicted	Actual value
2014	4.5	4.0	5.1
2015	3.4	3.8	3.6
2016	6.8	7.4	7.8
2017	5.8	6.0	5.4
2018	5.2	4.8	5.0
2019	6.2	6.8	6.5
2020	7.8	7.5	8.0
2021	8.3	7.8	8.2
2022	6.3	6.0	6.5
2023	5.7	5.4	5.9
Mean	6.7	6.5	6.5
Median	8.2	7.2	7.9

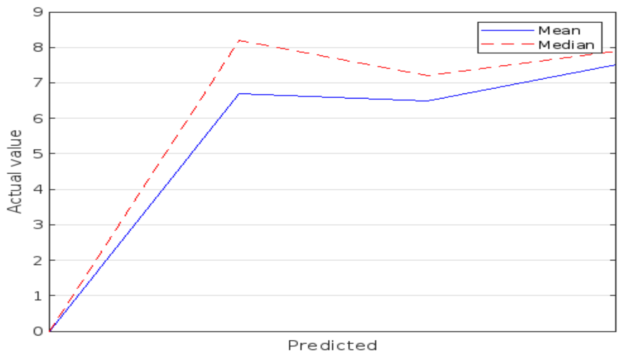


Figure 5: Prediction And Actual Value Of Digital Economy In Market Trends

Table 3 and Figure 5 show the observed forecast based on the year values. Their analysis shows that there is no association between the regression values and that the real regression has a normal distribution.

4.5 Accuracy

Table 4: Algorithm Comparison in Accuracy

Features	Years (2014-2023) Accuracy (%)				
	SVM	PCA	ARIMA	LSTM	ML-RFR [Proposed]
Y1	67.4	56.0	73.6	83.9	92.9
Y2	57.2	53.5	67.5	78.2	88.1

Y3	65.2	68.9	70.2	68.3	91.8
Y4	49.90	58.8	65.3	79.1	94.5
Y5	74.3	69.2	78.4	88.2	97.8

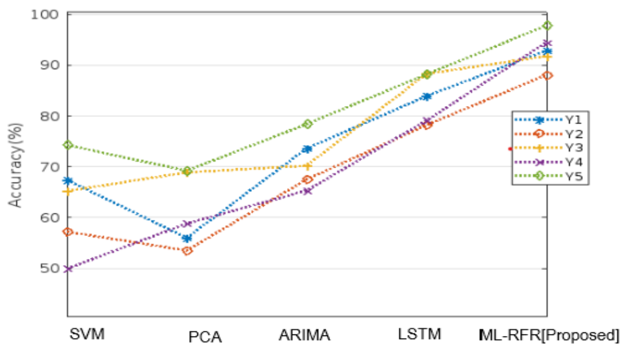


Figure 6: Accuracy Of Current And Suggested Techniques

Figure 6 illustrates the proposed method's accuracy. The accuracy of the gadget is a measure of how closely the economies of the evaluations of a quantity match the value that genuinely matches the given number. Comparing the suggested technique's predictions about the digital economy with consumer consumption to the existing methods are SVM [24], PCA [25], ARIMA [26], and LSTM [27] methodologies, it is discovered that the recommended strategy ML-RFR is more accurate. The proportion of the total that represents the accuracy level is reported. The suggested solution uses a machine learning algorithm to predict market trends related to the digital economy. In comparison, the proposed method ML-RFR achieves higher 97.8% accuracy, while SVM only reaches 74%, PCA only reaches 69%, ARIMA only reaches 78.4% and LSTM only reaches 83%.

4.6 Precision

Table 5: Outcome Value of Precision

Methods	Precision (%)
SVM	74.8
PCA	69.4
ARIMA	78.2
LSTM	88.2
ML-RFR [Proposed]	96.4

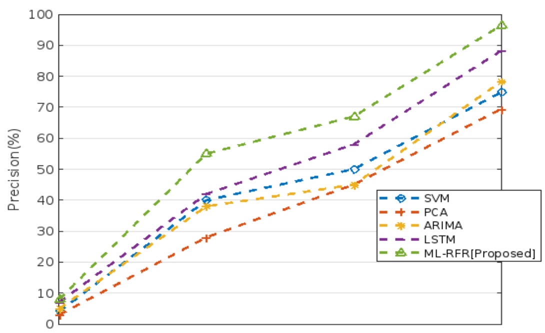


Figure 7: Precision Of Proposed and Existing Methods

International Water and Irrigation
Volume 44 Issue 2 (2025)
ISSN: 0334-5807
Pages 137 - 150

Compared to the current methods, the suggested method's precision is significantly improved (see Figure 7). By comparison, the precision achieved by the suggested method, ML-RFR, is 96.4%, whereas SVM, PCA, ARIMA, and LSTM only manage 74.8%, 69.4%, 78.2%, and 88.2%, respectively. This is due to the following level of market trend in the digital economy of precision. The recommended system performs at the highest level as a consequence. Table 5 displays the precision of the proposed approach.

4.7 Recall

Table 6: Outcome value of Recall

Methods	Recall (%)
SVM	72.6
PCA	68.3
ARIMA	78.1
LSTM	88.1
ML-RFR [Proposed]	96.0

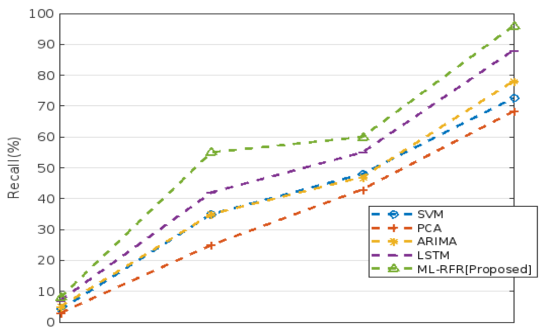


Figure 8: Recall of current and suggested techniques

Recall can also refer to the sensitivity or true positive rate. When compared to the existing methods, the proposed solution has the better recall degree. Table 6 and Figure 8 present the results of the recall. In comparison to the existing techniques, SVM (72.6%), PCA (68.3%), ARIMA (78.1%), and LSTM (88.1%), our suggested method, ML-RFR (96%), fared better. As such, it is recommended that the suggested techniques be used. ML-RFR will use the description to forecast market trends for the digital economy based on machine learning.

4.8 F1 Measures

Table 7: Outcome of F1 measures

Methods	F1 measures (%)
SVM	71.9
PCA	67.1
ARIMA	77.8
LSTM	85.0
ML-RFR [Proposed]	97.3

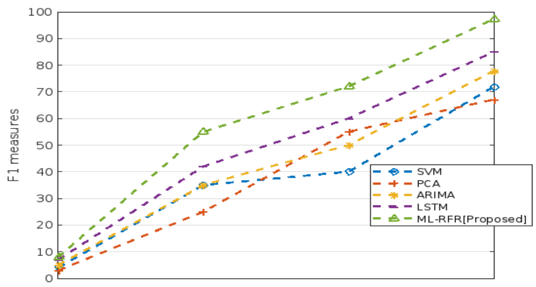


Figure 9: F1 Measures of Proposed And Existing Methods

5. Conclusion

This research examines the market trend prediction of the digital economy using machine learning (ML), focusing on factors such as industry employment quality, industry development level, industry development mode, and high-quality employment. Ultimately, it draws the following conclusions: This analysis uses data related to the digital economy's development in Zhejiang Province from 2014 to 2023. The ML techniques determine the Multi Linear-based Random Forest Regression Algorithm (ML-RFR) coefficient. With respect to the model Y1, Y2, Y3,..., Y9 coefficient, The regression coefficient's p-value indicated its significance level. The digital economy prediction years from 2014 to 2023 yielded positive dynamic factor indicators. We obtain the best results when we take two samples, the very minimum required for node splitting. When we choose the bare minimum of two samples for 2, 3, and 4 node splitting, the decision tree fully splits and achieves an average value of 88, reflecting the impact of random forests with varying parameters. This yields the best results. The F1 assessment's mean value was 97.3%. This article evaluated and contrasted the innovative machine learning algorithms and approaches used in finance, focusing on the digital economy's ability to forecast market trends. We have discussed a variety of machine learning algorithms and strategies based on input types, goals, and parameters. The suggested method's ML-RFR algorithms and methodologies have gained popularity in the digital economy because of their improved accuracy, precision, recall, and F1 measures obtained. Numerous have demonstrated the significant impact of the digital economy on upcoming trends. Therefore, it would be fascinating to incorporate a combination of technical and fundamental assessments into future work on state-of-the-art machine learning to increase prediction efficiency.

Reference

[1] C. Dirican, “The Impacts of Robotics, Artificial Intelligence on Business and Economics,” Procedia - Social and Behavioral Sciences, Jul. 2015, vol. 195, no. 1877-0428, pp. 564–573, doi: https://doi.org/10.1016/j.sbspro.2015.06.134.

[2] W. Zhu and J. Chen, “The spatial analysis of digital economy and urban development: A case study in Hangzhou, China,” Cities, Apr. 2022, vol. 123, p. 103563, doi: https://doi.org/10.1016/j.cities.2022.103563.

[3] K. Ji, X. Liu, and J. Xu, “Digital Economy and the Sustainable Development of China’s Manufacturing Industry: From the Perspective of Industry Performance and Green Development,” Sustainability, Mar. 2023, vol. 15, no. 6, p. 5121.doi: https://doi.org/10.3390/su15065121.

[4] I. Parmar et al., “Stock Market Prediction Using Machine Learning,” 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Dec. 2018, doi: https://doi.org/10.1109/icsccc.2018.8703332.

[5] Z. Weng, J. Duan, J. Zhou, and L. Zhao, “A Literature Review of Digital Economy in China: Trends, Drivers, and Implications,” Advances in economics, management and political sciences, Jun. 2024, vol. 89, no. 1, pp. 144–151, doi: https://doi.org/10.54254/2754-1169/89/20231417.

[6] Edi-Cristian Dumitra, M. Gândea, and Radu Alexandru Budu, “AI – The New Player in Digital Economy,” Sciendo eBooks, Oct. 2023, pp. 877–886. doi: https://doi.org/10.2478/9788367405546-081.

[7] T. V. Cherian, Getzi Jeba Leelipushpam Paulraj, Joyce Beryl Princess, and Immanuel Johnraja Jebadurai, “A comparative analysis of machine learning and deep learning techniques for aspect-based sentiment analysis,” Elsevier eBooks, Jan. 2024, pp. 23–37. doi: https://doi.org/10.1016/b978-0-443-22009-8.00006-9.

[8] A. Li, “Research on the Influence of Digital Economy Development on China’s Inter-provincial Path Dependence Effect,” Advances in computer science research, Jan. 2023, pp. 218–225, doi: https://doi.org/10.2991/978-94-6463-042-8_33.

[9] L. D. Williams, “Concepts of Digital Economy and Industry 4.0 in Intelligent and information systems,” International Journal of Intelligent Networks, 2021, vol. 2, pp. 122–129, doi: https://doi.org/10.1016/j.ijin.2021.09.002.

[10] S. Yang and J. He, “Analysis of Digital Economy Development Based on AHP-Entropy Weight Method,” Journal of Sensors, Jun. 2022, vol. 2022, pp. 1–8, doi: https://doi.org/10.1155/2022/7642682.

[11] Y. Lu and Y. Zhou, “A Short Review on the Economics of Artificial Intelligence,” SSRN Electronic Journal, 2019, doi: https://doi.org/10.2139/ssrn.3433527.

[12] L. Wang and L. Zhao, “Digital Economy Meets Artificial Intelligence: Forecasting Economic Conditions Based on Big Data Analytics,” Mobile Information Systems, Oct. 2022, vol. 2022, pp. 1–9. doi: https://doi.org/10.1155/2022/7014874.

[13] R. F. Reier Forradellas and L. M. Garay Gallastegui, “Digital Transformation and Artificial Intelligence Applied to Business: Legal Regulations, Economic Impact and Perspective,” Laws, Aug. 2021, vol. 10, no. 3, p. 70. doi: https://doi.org/10.3390/laws10030070.

[14] H. Hang and Z. Chen, “How to realize the full potentials of artificial intelligence (AI) in digital economy? A literature review,” Journal of Digital Economy, Dec. 2022, vol. 1, no. 3, doi: https://doi.org/10.1016/j.jdec.2022.11.003.

[15] Y. HE, “The Effect of Artificial Intelligence on Economic Growth: Evidence from Cross-Province Panel Data,” Korean Artificial Intelligence, Dec. 2019, vol. 7, no. 2, pp. 9–12, doi: https://doi.org/10.24225/kjai.2019.7.2.9.

[16] P. Hao and J. Guo, “Constrained center and range joint model for interval-valued symbolic data regression,” Computational Statistics & Data Analysis, Dec. 2017, vol. 116, pp. 106–138, doi: https://doi.org/10.1016/j.csda.2017.06.005.

[17] O. Claveria, E. Monte, and S. Torra, “A Genetic Programming Approach for Economic Forecasting with Survey Expectations,” Applied Sciences, Jun. 2022, vol. 12, no. 13, p. 6661, doi: https://doi.org/10.3390/app12136661.

[18] X. Ge, Z. Zhou, X. Zhu, Y. Wu, and Y. Zhou, “The Impacts of Digital Economy on Balanced and Sufficient Development in China: A Regression and Spatial Panel Data Approach,” Axioms, Jan. 2023, vol. 12, no. 2, pp. 113–113, doi: https://doi.org/10.3390/axioms12020113.

[19] H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, “Random survival forests,” The Annals of Applied Statistics, Sep. 2008, vol. 2, no. 3, pp. 841–860, doi: https://doi.org/10.1214/08-aoas169.

[20] W. Gao and Z. Ding, “Construction of Digital Marketing Recommendation Model Based on Random Forest Algorithm,” Security and Communication Networks, Aug. 2022, vol. 2022, pp. 1–9, doi: https://doi.org/10.1155/2022/1871060

[21] F. Huang, "Personalized Marketing Recommendation System of New Media Short Video Based on Deep Neural Network Data Fusion," Journal of Sensors, Nov. 2021, vol. 2021, pp. 1–10, doi: https://doi.org/10.1155/2021/3638071.

[22] Y. Dexiang, M. Shengdong, Y. Liu, G. Jijian, and L. Chaolung, "An Improved Deep-Learning-Based Financial Market Forecasting Model in the Digital Economy," Mathematics, Jan. 2023, vol. 11, no. 6, p. 1466, doi: https://doi.org/10.3390/math11061466.

[23] V. Božić, "The Impact of Artificial Intelligence on Developing Digital Economy," doi: https://doi.org/10.13140/RG.2.2.29025.89443.

[24] [24] J. Fu, X. Zhou, and G. Mei, "Internet Digital Economy Development Forecast Based on Artificial Intelligence and SVM-KNN Network Detection," Computational intelligence and neuroscience, Jun. 2022, vol. 2022, pp. 1–11, doi: https://doi.org/10.1155/2022/5792694.

[25] R. Luan and P. Xu, "Risk Prediction of the Development of the Digital Economy Industry Based on a Machine Learning Model in the Context of Rural Revitalization," Information resources management journal, May 2024, vol. 37, no. 1, pp. 1–21, doi: https://doi.org/10.4018/irmj.343095.

[26] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks," Future Internet, Aug. 2023, vol. 15, no. 8, p. 255, doi: https://doi.org/10.3390/fi15080255.

[27] G. Xu, S. Peng, C. Li, and X. Chen, "Synergistic Evolution of China's Green Economy and Digital Economy Based on LSTM-GM and Grey Absolute Correlation," Sustainability, Sep. 2023, vol. 15, no. 19, pp. 14156–14156, doi: https://doi.org/10.3390/su151914156.