Contribution Analysis of Factors Affecting the Growth of Chinese Construction Enterprises Based on the XGBOOST Algorithm

Yanbing Hou[1, #], Chi Qin[2, #, *]

[1] Department of Mathematics, Sichuan University, 610044, Chengdu, China

[2] Guangxi Financial Vocational College, 530000, Guangxi, China

* Corresponding Author Email: DetachmentQC@163.com

[#]These authors contributed equally.

Abstract. The cornerstone of China's economy is the building sector. It is critical to research boosting the core competitiveness of China's construction companies in the context of challenging domestic development challenges and complex international competition. As a result, this paper aims to establish 15 influencing factors that impact the growth of construction enterprises from both internal and external dimensions and uses the XGBOOST machine learning method to analyze the contribution of influencing factors based on the macro and micro multi-faceted analysis perspective. China Vanke Group is used as the research object in this essay. The findings demonstrate that the regression model's fitting degree and goodness of fit are at their maximum, and confidence is at its highest when the number of base learners is 750. The component that significantly impacts how fast businesses grow is the total operating expense growth rate. For construction businesses to improve their core competitiveness and achieve sustainable development, this article suggests specific approaches. This study helps foster the high-quality growth of businesses, producing an improved outcome and ensuring the industrial chain's stability.

Keywords: XGBOOST; Chinese construction enterprises; influencing Factors of Enterprise Growth.

1. Introduction

China's economic and social development now emphasizes high-quality development. When it comes to transforming and upgrading construction enterprises, the need for high-quality and efficient development is extremely urgent, and the enhancement of core competitiveness is the key to success. These challenges are brought on by the pressure of domestic reform and the complex international competitive situation. The policy and economic environment have been optimized in recent years, allowing China's construction businesses to dramatically strengthen their core competitiveness by integrating superior resources more quickly; Overall, nevertheless, it continues to face development challenges like poor supply quality, a lack of high-end supply, and inadequate innovation potential. Numerous studies are being conducted in this area at the moment. It also mainly focuses on identifying and analyzing the global influences shaping construction firms' expansion. Therefore, a fundamental issue facing China's construction firms is how to improve their core competitiveness while achieving high-quality development[1].

In terms of qualitative analysis, (1) LU et al. [2] studied the core competitive factors of Chinese construction enterprises based on literature analysis and expert consultation, discovered and improved the system and mechanism of scientific and technological innovation, and proposed that it is essential to promote China's construction standards at the enterprise level, strengthen human resource management and optimize the talent cultivation mechanism.

(2) From the perspective of psychology, FANG et al. [3] studied the influence mechanism of behavioral finance on working capital management of Chinese construction enterprises based on the basic theory and analysis framework of the herding effect, mental accounting, and representative principle error. Proposed: China's construction enterprises should moderately improve the degree of

democratization of operation and management decisions, and enterprises need to carefully analyze the demand for funds to avoid units from the actual situation of the enterprise's irrational behavior.

(3) In the enterprise growth cycle, ZOU et al.[4] studied the factors affecting the growth of Chinese construction enterprises based on the life cycle theory and found that the government must create a favorable external environment by improving technological innovation and learning ability, cultivating technological innovation ability, improving entrepreneurial ability, building learning employees, strengthening capital management and operation, paying attention to the appointment of talents, establishing a sound project management system, implementing a strict internal control system, establishing a sound project quality management system, and shaping and learning excellent corporate culture.

In quantitative terms. (1) Based on the C-D Douglas production function and SWOT method, ZHOU et al.[5]studied the theoretical basis of the core competitiveness and sustainable development of enterprises and found that it is very important to promote enterprise restructuring, reform the management system, establish a professional subcontracting system, increase market development efforts, develop a variety of business models, strengthen the construction of enterprise culture, strive to improve the quality of employees and accelerate the information construction of construction enterprises.

(2) Based on [6] variation function, Gini coefficient, and Theil index, DENG et al. studied the growth difference and influencing factors of the construction industry and found that it is imperative to accelerate local economic development, ensure steady GDP growth, increase regional openness and residents ' income, curb inflation and accelerate urbanization.

(3) Based on Kaiser-Meyer-Olkin (KMO) test and Bartlett Test of Sphericity (Bartlett Test of Sphericity), LIU et al.[7] conducted a reliability study on the influencing factors of the growth of Chinese intelligent building enterprises and found that it is imperative to attach importance to the construction of enterprise talent training system, give full play to the training and management functions of industry associations for employees, and improve the market access and clearing system for intelligent building enterprises and related products.

The aforementioned research techniques have numerous restrictions on the feature variables and make it challenging to fit complicated nonlinear issues, which limits the accuracy of the findings. The issue of model over-fitting may result in a "multidimensional disaster" in selecting feature indicators and damage the model's capacity for generalization. This study quantifies the internal and external elements that impact China Vanke Group's growth and uses descriptive statistical analysis to assess the model's correctness using *MSE* and R^2. This article utilizes big data and machine learning to select the characteristic indicators, taking into account relevant statistical characteristics like the difference and regularity of the relevant professional knowledge and experimental data so that the final results also have strong adaptability outside the sample and can more effectively achieve the prediction effect.

2. Construction of the indicator system of the growing influence of construction enterprises

2.1. Growth of construction enterprises

In an enterprise's life cycle, enterprise growth is the process of dynamic development, which is the development trend influenced by both internal and external factors. Enterprise development encompasses improving enterprise quality and expanding enterprise scale, and construction enterprises are no exception. The accumulation of human and non-human resources [8], such as staffing of construction enterprises and the scale of real estate development and sales, are the main manifestations of the scale expansion of construction enterprises. While the quality improvement of construction companies primarily relates to the progress of the enterprise's ownership, technical and industrial structures may be defined explicitly as enterprise innovation and the capacity to react to the external environment and others. Currently, the growth rate is mainly used to quantify the growth of enterprises. Common indicators include operating revenue growth rate, profit, full scale, return on

net assets, capital preservation, and appreciation rate. When an organization's growth is quantified, the development of the business can be objectively reflected. This can help investors make decisions and give business leaders development ideas so that high-quality companies can flourish over time. To analyze the trend of scale expansion of construction enterprises, this paper chooses the growth rate of total assets to quantify the growth of enterprises, which is denoted as Y, and the formula is as follows

$$Y = \frac{X_{n} - X_{n-1}}{X_{n-1}} \tag{1}$$

X_n represents the current period-end value of total assets, and X_{n-1} represents the year-end value of total assets.

2.2. Construction of index system of factors affecting the growth of construction enterprises

2.2.1 Internal factors

Internal factors (II) primarily refer to an enterprise's capacity for growth, frequently measured by its financial indicators. As an enterprise's capacity improves over time, its growth will become more robust. The internal factors influencing the growth of construction companies include profitability, solvency, operating capacity, development potential, enterprise scale, enterprise risk, and investment value. As a profitable enterprise[9], the profitability of construction enterprises directly affects the development of the industry. In this paper, the net profit margin of total enterprise assets is used to reflect the ability of enterprise assets to obtain profits and reflect the ability of enterprises to create profits. Total enterprise assets are all owned or controlled by the enterprise, which can effectively reflect the enterprise's scale and production capacity. This study uses the equity ratio, which is the ratio of total business liabilities to real owner equity, to assess the company's long-term viability. This paper also uses the total asset turnover ratio, sustainable growth rate, comprehensive leverage, price-earnings ratio, and total operating cost growth rate to reflect the enterprise's investment value and market performance as well as the ability to operate the business.

2.2.2 External factors

External factors (OI) refer to the external environment for business growth. An excellent external environment can accelerate business growth. In contrast, a bad external environment can cause a blow to business growth, and external factors can be broken down into three levels: market perspective, national perspective, and international perspective. When seen from the standpoint of the market, a firm's growth is more challenging, the more competitive the market is, the quicker it expands during times of high consumer demand, and the better the loan position, the better the external market views the company. From a national viewpoint, national policies affect businesses differently at different periods; the more favorable the economic climate, the better the prospects and the more effectively the enterprises develop. The global environment frequently restricts the growth of businesses, and foreign commerce plays a significant role in this.

In summary, from the external factors, we choose corporate share price to indicate market competition, fixed asset investment price index to indicate consumer demand, bank loan amount to enterprises to indicate corporate lending, real estate development investment amount to reflect national policy, SSE index to reflect the economic environment, and exchange rate (RMB to USD exchange rate) to reflect international form.

In summary, the index system constructed in this paper is shown in Table 1.

Table 1. Table of the index system of enterprise growth influencing factors four

Tier 1 Indicators	Secondary indicators	Indicator Unit	REF
Internal factors (II)	The net profit margin on total corporate assets (II_1)	%	[10]
	Equity ratio（ II_2 ）	%	[11]
	Total Asset Turnover （II_3 ）	%	[11]
	Integrated Leverage（ II_4 ）	%	[12]
	Sustainable Growth Rate （ II_5 ）	%	[13]
	P/E ratio (II_6)	%	[14]
	Total assets (II_7)	Billion	
	Total operating cost growth rate (II_8)	%	
External factors (OI)	Corporate share price (annual individual share transaction amount) (OI_1)	Billion	
	Corporate taxation (OI_2)	Billion	
	Exchange rate (RMB to USD) (OI_3)	Yuan	[15]
	Investment in real estate development (OI_4)	Billion	[16]
	Bank loans to businesses (OI_5)	Billion	
	Fixed asset investment price index - previous year = 100 year-on-year (OI_6)	&	
	SSE (OI_7)	Point	

3. XGBOOST Method

XGBOOST（Extreme Gradient Boosting）is based on the iterative cumulative gradient lift decision tree algorithm. That is, starting from a naive model, a new fitting model is established according to the errors of observed values in the sample set, and it is added to the existing model to form the set model iteratively. The performance of the integrated model is often better than that of the single model. Compared with the traditional enhanced tree method, XGBOOST introduces rule terms into the objective function to prevent overfitting of the model and adopts second-order Taylor expansion for the loss function to further optimize the efficiency and accuracy of the model. The optimized distributed gradient enhancement library has the characteristics of high efficiency, flexibility, and portability. XGBOOST is a large parallel tree tool. It's the fastest and best open source boosting tree toolkit, more than 10 times faster than regular toolkit. In addition, XGBOOST also supports running in various distributed environments such as Kubernetes, Hadoop, SGE, MPI and Dask, which is a good solution to large-scale data problems in the construction industry[17].

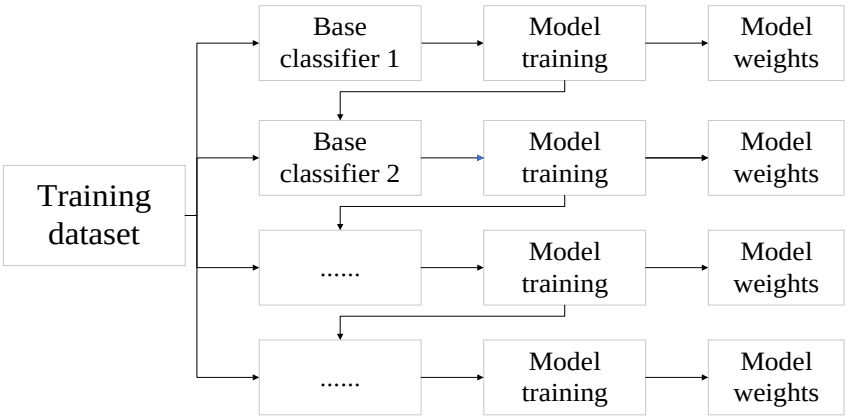


Figure 1. xgboost flow chart

The core formula for XGBOOST calculation is as follows:

STEP1: In building the XGBOOST model, we need to build N regression trees so that the predicted value of the tree group is as close as possible to the true value (accuracy) and has the greatest generalization ability. That is, given several training samples consisting of input vector x and output vector y（x_1,y_1）,(x_2,y_2)...(x_n,y_n)，The target finds an approximation function $F(\vec{x})$, Such that the loss function $L(y, F(\vec{x}))$ minimizes the loss value. Where $F(\vec{x})$ is our prediction.

A typical definition of $L(y, F(x))$ is:

$$L(y, F(\vec{x})) = \frac{1}{2}(y - F(\vec{x}))^2 \tag{2}$$

The optimal function is assumed to be $F^*(\vec{x})$ namely:

$$F^*(\vec{x}) = \arg\min E_{(x,y)}[L(y, F(\vec{x}))] \tag{3}$$

Suppose $F(x)$ is the weighted sum of a family of basis functions $f_i(x)$:

$$F(\vec{x}) = \sum_{i=1}^{M} \gamma_i f_i(x) + const \tag{4}$$

The $F(\vec{x})$ here is weighted by a weak classifier.

STEP2: When $F(\vec{x})$ is the first, we define it as a constant function, as follows:

$$F_0(\vec{x}) = \arg\min \sum_{i=1}^{n} L(\gamma_i, \gamma) \tag{5}$$

When we ask for the m-th $F(\vec{x})$ the first *m-1* classifier has already been selected, so as long as we find the m-th optimal weak classifier and add it to the first *m-1*, it becomes the m-th $F(\vec{x})$ and greedy thinking is used here[9]

$$F_m(\vec{x}) = F_{m-1}(\vec{x}) + \arg\min_{f \in H} \sum_{i=1}^{n} L(y_i, F_{m-1}(\vec{x}_i) + F(\vec{x}_i)) \tag{6}$$

$$\begin{aligned} &\hat{y}_i^{(0)} = 0 \\ &\hat{y}_i^{(1)} = f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\ &\hat{y}_i^{(2)} = f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\ &..... \\ &\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \end{aligned} \tag{7}$$

From this, the objective function is derived:

$$L^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{8}$$

STEP3: XGBOOST uses the Taylor expansion formula and expands to the second order, where the Taylor expansion formula is:

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2 \tag{9}$$

Let $\hat{y}_i^{(t-1)}$ in the objective function be x in the formula, let $f_t(x_i)$ be Δx, let g_i be the first derivative of $f(x)$, and let h_i be the second derivation, as shown in the figure below:

$$g_i \overset{def}{==} \frac{\partial L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, h_i \overset{def}{==} \frac{\partial^2 L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}} \tag{10}$$

where *C* is constant, so the objective function becomes:

$$J(f_t) \approx \sum_{i=1}^{n}\left[L\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^{\,2}(x_i) \right] + \Omega(f_t) + C \tag{11}$$

STEP4: The output of the regression tree is a numerical value, in XGBOOST, the regression tree is actually *f(x)*, and its output value is actually the weight, that is, the value of its leaf node is the weight w_1, w_2,..., then that means $f_i(x)$ can be equivalent to w_i, so that it becomes a general routine of finding the loss function, deriving w, and then iterating in turn.

And simplify the formula. $\Omega(f_t)$ is a regular term, and its actual formula is as follows:

$$\Omega(f_t) = \Upsilon T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^{\,2} \tag{12}$$

The objective function becomes:

$$\begin{aligned} J(f_t) &\approx \sum_{i-1}^{n}\left[L\left(y_i, \hat{y}_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) + C \\ &= \sum_{i-1}^{n}\left[g_i w_{q(x_i)} + \frac{1}{2} h_i w_{q(xi)}^2 \right] + \gamma \cdot T + \lambda \cdot \frac{1}{2}\sum_{j-1}^{T} w_j^2 + C \end{aligned} \tag{13}$$

Here g_i is actually the gradient, $g_i \overset{def}{==} \frac{\partial L(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}$, and $L\left(y_i, \hat{y}_i^{(t-1)}\right)$ becomes a constant added to C because the former; The *t-1* weak classifier has been determined, then the predicted value has already been determined, so it is constant. The formula is as follows:

$$\sum_{j=1}^{T}\left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2}\left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma \cdot T + C \tag{14}$$

Define this equation:

$$J(f_t) == \sum_{j=1}^{T}\left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2}\left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma \cdot T + C \tag{15}$$

$$G_j \overset{def}{==} \sum_{i \in i} g_i, H_j \overset{def}{==} \sum_{i \in j} h_i \tag{16}$$

Bring (15) into equation (14) and we will get it:

$$J(f_t) = \sum_{j=1}^{T}\left[G_j w_j + \frac{1}{2}(H_j + \lambda) w_j^2 \right] + \gamma \cdot T + C \tag{17}$$

Partial derivation of the weight *w*:

$$\frac{\partial L(f_t)}{\partial w_j} = G_j + (H_j + \lambda) w_j \overset{\text{cause}}{==} 0 \Rightarrow w_j = -\frac{G_j}{H_j + \lambda} \tag{18}$$

The formula w is substituted back into the objective function, and the objective function is finally obtained:

$$J(f_t) = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j^2}{H_j+\lambda} + \gamma \cdot T \tag{19}$$

Finally, the difference in the loss function before and after the split is calculated:

$$Gain = \frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} - \frac{(G_L+G_R)^2}{H_L+H_R+\lambda} - \gamma \tag{20}$$

4. The setting of inspection indicators

An algorithm test is a step to check and review the results obtained by the program, that is, to prove the relationship between data input and output, as well as the proof formula and calculation method in the algorithm, to make the final data result from more real. In this paper,*MSE* and $R^2$are used to test and analyze XGBOOST. The average sum of squared errors of the corresponding points between the predicted data and the original data is used to ensure that the quantified sample data and have a certain degree of distribution fitting comparison and verification, to select the relatively optimal mathematical model. In parameter estimation, mean square error refers to the expected value of the square variance between the estimated value of the parameter and the true value of the parameter, which is denoted as *MSE*, and its formula is[18].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{21}$$

MSE is a measure that reflects the degree of difference between the estimator and the estimated quantity. Mean square error range [0,+∞), when the predicted value is completely consistent with the real value is equal to 0, that is, the perfect model; the larger the error, the larger the value. An easy to calculate and intuitive indicator used to measure correlation. The predicted value explains what proportion of the y_i variance of the variable is and measures how well the predicted value fits the true value. In popular understanding, assuming the variance of y_i is 1 unit,R^2 represents how much the variance of the residual y_i is reduced after the use of the model.

$$R^2 = 1 - \frac{(\hat{y}_i - y_i)^2}{(y_i - \bar{y})^2} \tag{22}$$

When R^2=1: is the best case, all the predicted values are equal to the true values. R^2=0: It is possible to simply predict that all y values are equal to the mean value of y, that is, all $\hat{y}_i$ values are equal to $\bar{y}$ (that is, the mean value of real y values), but there are other possibilities. R^2<0: the prediction ability of the model is poor, which is worse than simply predicting that all y values are equal to the mean value of y ". So that means that maybe you're using the wrong model, or maybe the model is assuming that there's no lower limit for R^2 because you can make predictions that are off by any degree. Therefore, the range of R^2 is $(-\infty,1]$[19].

5. Empirical analysis

5.1. Selection of research subjects and data collection

Vanke Group, founded in 1984, has become the top-ranked real estate company in China after more than 30 years of development. In 2022, the company's total revenue reached 337.7 billion yuan,

an increase of 24.38% year-on-year, and its business has expanded to more than 50 cities. Vanke Group was selected for research and analysis for two reasons: (1) Vanke Group has an internal sound system and a more stable development process. (2) By thoroughly analyzing the contributions of its development-affecting variables, Vanke Group, the leading enterprise in China's domestic real estate market, may better understand the development pulse of China's leading construction enterprise. As a result, understanding this business is crucial.

In this paper, based on the impact indicator system constructed in Chapter 2, annual data of the case companies from 2001 to 2021 were collected from the China Stock Market & Accounting Research Database (CSMAR), and descriptive statistics of the indicators were conducted on the data, which are shown in Table 2.

Table 2. Descriptive statistics of the index system of factors influencing enterprise growth

Indicator/Unit	Max	Min	Mean	Mid
II_1/%	0.065	0.020	0.043	0.041
II_2/%	5.487	1.074	3.024	3.367
II_3/%	0.736	0.221	0.393	0.334
II_4/%	1.343	1.028	1.094	1.070
II_5/%	0.200	0.073	0.139	0.142
II_6/%	0.949	-0.039	0.273	0.242
II_7/%	37.271	4.834	12.68	10.225
II_8/billion	19386.381	64.829	5723.875	2962.084
OI_1/billion	8080.087	50.725	2655.306	2412.700
OI_2/billion	329.052	1.795	108.921	77.787
OI_3/¥	8.277	6.097	7.070	6.835
OI_4/billion	141442.950	4984.050	58836.485	48259.400
OI_5/billion	1543.223	1.600	436.293	247.900
OI_6/%	106.500	99.600	102.271	101.900
OI_7/point	5521.490	1220.930	2754.360	2815.130
Y/%	1.206	0.037	0.347	0.279

5.2. Parameter calculation

This paper is based on SPSSPRO software, bringing in the relevant data of the constructed index system in 5.1, using the constructed XGBOOST algorithm, by adjusting the number of model parameters base learners, controlling the range of the number of base learners from 50 to 1000 and setting the step size to 50, using R^2, *MSE* as the criterion to discriminate the reasonableness of the parameters, the specific results are shown in Table 3.

Table 3. Table of XGBOOST regression discriminant parameters

Number of base learners	R^2	MSE	Number of base learners	R^2	MSE
50	1.747	-0.436	100	-0.03	1.717
150	1.564	0.064	200	1.616	-0.01
250	0.286	0.706	300	0.345	-3.715
350	0.645	0.484	400	0.182	0.175
450	0.245	0.005	500	3.154	-0.716
550	0.158	-0.03	600	1.315	0.020
650	1.471	0.065	700	0.261	0.766
750	0.144	0.861	800	1.434	0.033
850	0.210	0.156	900	1.445	0.006
950	1.440	0.032	1000	1.452	-0.033

The results in Table 3 show that when the number of base learners is 750, the *MSE* value is only 0.144, which indicates that the degree of difference between the predicted and actual values of the test set of the regression model is low. The R^2 value is as high as 0.861, indicating that the regression model's fit is high. At this time, the algorithm's accuracy converges, the error reaches the lowest, the goodness of fit becomes the highest, and the confidence level is the strongest.

5.3. Analysis and Recommendations

Based on the parameter calculation and adjustment of 5.2, the discriminant index shows better results when the number of base learners is 750. The model fits better, which can better explain the relationship between the case group's dependent variable of total assets growth rate and the independent variables in the index system of enterprise development influence factors. Each variable's impact on the dependent variable may be quantified, and the precise outcomes are displayed in Figure 2.

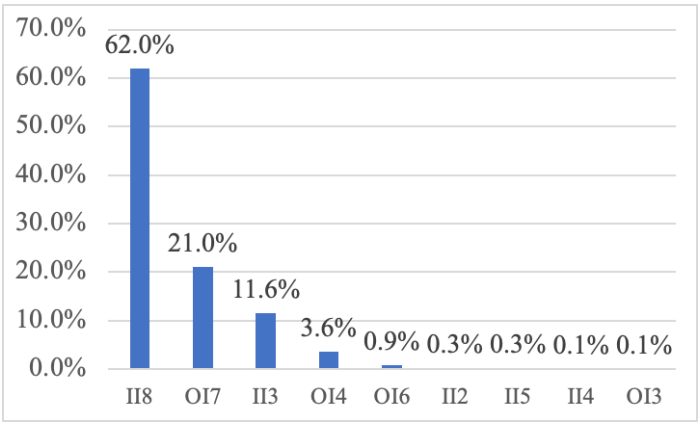


Figure 2. Plot of importance index of independent variables to dependent variable characteristics

From Figure 2, the factors that have an impact on the growth rate of total assets (Y) of the case companies are the total operating cost growth rate (II_8), SSE (OI_7), total asset turnover (II_3), investment in real estate development (OI_4), fixed asset investment (OI_6), equity ratio (II_2), sustainable growth rate (II_5), integrated leverage (II_4), the exchange rate (RMB to USD) (OI_3). II8 has the highest importance at 62%, followed by OI_7 and II_3 at 21% and 11.60%, respectively. OI_4 also influences the constructed model with a volume of 3.6%. In contrast, II_2, II_5, II_4, and OI_3 have a lower influence rate on the case companies with a sum of characteristic importance of 0.8%.

II_8 is the sum of direct and other business costs. Case enterprises mainly include the costs invested in the development and construction, and operation of real estate goods sold. Construction enterprises have incredibly high costs, which are essential to enterprise production. OI_7 displays the country's current economic situation by reflecting the movement of listed companies' equities on the Shanghai Stock Exchange. Construction companies are influenced by the economy. When things are going well financially, more people are eager to invest in and support the growth of real estate businesses. II_3 is the enterprise's sales revenue and average total assets in a certain period; the index reflects the amount of revenue that can be created by the enterprise's assets and portrays the enterprise's operating capacity, which is one of the crucial factors in the sales process of the enterprise. OI_4 reflects the strength of the state's investment in construction enterprises; real estate enterprises and people's lives are closely related, and the state's support for construction enterprises will effectively promote the development of enterprises.

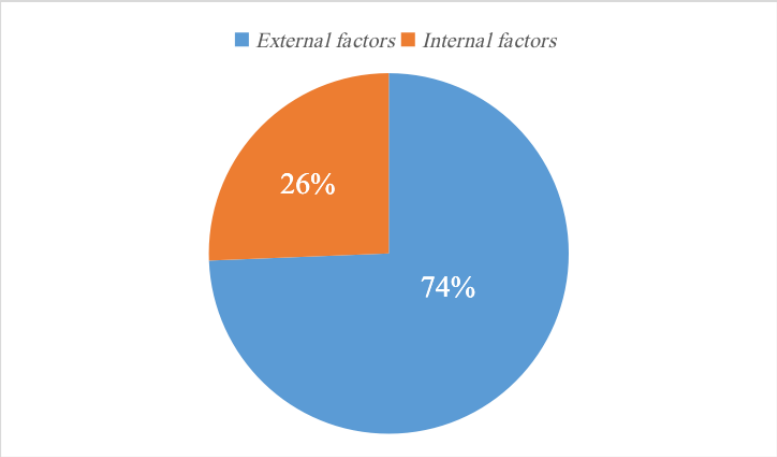


Figure 3. Internal factors external factors share chart

The percentages of internal and external factors are shown in Figure 3,Which displays the proportions of internal and external elements. 74% of the components are internal, compared to 26% of the factors that are external. This graph demonstrates that when it comes to businesses, internal variables have a bigger influence on their expansion. The internal environment of the business affects its quality. The operating costs, operating capacity, solvency, development potential, and risk level of the case companies have a greater influence on the company's growth, which is related to the status of the construction enterprises as an essential pillar in the national economy, occupying a significant market share. Therefore, for enterprises in the industry, in addition to the external environment to help them to develop, they need to improve their ability to obtain better growth.

This paper gives relevant suggestions and opinions by analyzing the growth rate of total assets and the index system of influencing factors of the case companies.

(1)Improving enterprise strength, pursuing information transformation, and high-quality development are the primary priorities for traditional construction companies. Internal business factors than external ones more impact traditional construction businesses and strengthening an organization's core competitiveness is the key to prosperit. Human resource management, technological innovation, corporate brand culture, and industry chain integration are good places to start. In the information era, launch a brand-new company model for development. Businesses should actively grow the digital sector while keeping an eye on production and operating expenses, operational capacity, and solvency risk levels.

(2)To encourage the construction industry's peaceful development, the state and government must fully acknowledge the construction sector's importance to the national economy and the various issues it is currently facing. They must also regulate the macro market, increase social investment in real estate in a particular economic context, and introduce pertinent policies. To ensure the stability of the industrial chain, it should also consider the real estate sector's protracted business cycle and the requirement to enhance the multi-level financing system.

6. Conclusion

This paper takes China's construction enterprises (China Vanke Group) as the research object. Based on literature research, the sustainable development of enterprises and the theory of core competitiveness of enterprises are discussed. The machine learning method is used to analyze the current situation of China's construction enterprises. The internal and external factors influencing the growth of construction organizations in each growth stage are identified and quantified based on the theoretical review. The most significant elements affecting the growth of construction enterprises in each growth stage are found by analyzing the influence of numerous factors on the evolution of Chinese construction enterprises in different growth phases and descriptive statistical analysis. Additionally, several countermeasures and recommendations were made to support the expansion of Chinese construction firms. These are the conclusions:

(1) The model's *MSE* and R^2 are at their respective highest points, and its confidence is at its maximum point when the number of base machine learning is 750.

(2) It may be inferred from the machine learning results through the characteristic importance index diagram between the independent and dependent variables that 74% of the factors influencing an enterprise's growth are internal, while only 26% are external.

(3) Generally, the price-earnings ratio of China's construction companies, followed by the trading index of the exchange where construction enterprises are located, significantly influences enterprises' growth.

To ensure the integrity, reason, and efficacy of budget management, businesses should thoroughly examine the financial requirements of each unit. Open up a new business model in the information age, boost their core competitiveness, and innovate in human resource management, science, and technology, corporate brand culture, industry chain integration, etc. Ultimately, the shift of the digital path for information To ensure the stability of the enterprise's current assets, we should establish a multi-level financing system, control the cost of accounts receivable, strictly regulate the credit limit required for credit sales, and adjust and reduce the detrimental effects of extreme financial behavior on the quality of short-term borrowing and accounts receivable management of Chinese construction enterprises.

References

[1] National Bureau of Statistics of China. Statistical Communiqué of the People's Republic of China on National Economic and Social Development in 2021 February 28, 2022

[2] Lu Chunfang, Wu Jun, Wang Mengjun, Li Xinsheng, Liu Bo, Liang Chao, Hu Xiaodong. Research on the Core Competitiveness of Chinese Construction Enterprises under the Background of High-Quality Development [J / OL]. in 2021 August 06, 2022

[3] Fang Xiaobing, Hu Siyue, ' Behavioral Finance and Working Capital Management of Construction Enterprises: Analysis of Impact Mechanism, '(C) 1994-2022 China Academic Journal Electronic Publishing House.All rights reserved.

[4] Zou Guihua, ' Research on the Influencing Factors of the Growth of Construction Enterprises in China Based on Life Cycle Theory, ' (C) 1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.

[5] Comprehensive, ' Research on Sustainable Development of Construction Enterprises Based on Core Competitiveness, '(C) 1994-2022 P. R. China Academic Journal Electronic Publishing House. All rights reserved.

[6] Qiao Lirong, ' The Impact of Financial Sharing Service Center on Financial Management of Construction Enterprises, ' (C) 1994-2022 P. R. China Academic Journal Electctronic Publishing House. All rights reserved.

[7] Liu, ' Research on the Influencing Factors of the Growth of Intelligent Building Enterprises in China ' (C) 1994-2022 P.R.China Academic Journal Electctronic Publishing House. All rights reserved.

[8] LiuYao, Shengdao. Enterprisegrowth: definition and measurement [J]. Soft Science, 2011, 25(02): 141-144.

[9] Fan Jing. Analysis of internal influencing factors of development performance of leading agricultural industrialized enterprises based on financial perspective [J]. Business Economics, 2021(07): 125–129.

[10] Cen Chengde. Empirical research on the growth of listed companies [J]. Business Research 2002.

[11] Park, Chun-Hua, Wang, Jia-Yu. Hypothesis and empirical analysis of factors influencing the efficiency of internal control in listed companies [J]. Green Finance and Accounting, 2017(03): 35-43.

[12] Liu Hui. Analysis and Discussion on the Current Situation and Countermeasures of Financial Management in Construction Enterprises [J]. Chinese and Foreign Enterprise Culture, 2021(07): 15–16.

[13] Yang Chen. An empirical study on the growth of listed companies [J]. Journal of Chifeng College (Natural Science Edition), 2017, 33(10): 88–90.

[14] Shen Zhichao. Analysis of growth of listed companies in China [J]. Modern Enterprise, 2015(07): 54–55.

[15] Zhang Juan,Wang Guangjun. Trends of RMB exchange rate changes and the impact on the market[J]. China Procurement Development Report, 2008(00):128-132.

[16] Xiang Yong, Zheng Mao, Dai Tianhui. Research on the dynamic factors and influencing mechanism of high-quality development of China's construction industry [J]. Building economy, 2019, 40(12): 15-20.

[17] Friedman J, Hastie T, Tibshirani R, et al. (2000). "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)." The annals of statistics, 28(2), 337–407.

[18] Friedman JH (2001). "Greedy function approximation: a gradient boosting machine." Annals of Statistics, pp. 1189–1232.

[19] Xgboost: eXtreme Gradient Boosting Tianqi Chen, Tong He Package Version: 0.6-4 January 4, 2017