Research Article
Motivation Analysis of Technological Startups Business Models Based on Intelligent Data Mining and Analysis

Xuejiao Ren¹ and Xiaozhou Ding²

¹Guangdong Mechanical & Electrical Polytechnic, Guangzhou 510515, China
²Harbin Engineering University, Harbin, Heilongjiang 150001, China

Correspondence should be addressed to Xuejiao Ren; 2018010027@gdmec.edu.cn

Received 3 March 2022; Revised 7 April 2022; Accepted 27 April 2022; Published 17 May 2022

Academic Editor: Wei Liu

Copyright © 2022 Xuejiao Ren and Xiaozhou Ding. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To improve the motivation analysis effect of the technological startups business model, this study combines intelligent data mining technology to analyze related factors and proposes a method of adjusting parameters of machine-learning model based on Bayesian optimization algorithm. Moreover, this study integrates the prior information and sample information of the machine learning algorithm performance function and builds an intelligent analysis model based on the elements of the technological startups business model. Finally, this article designs experiments to evaluate the effect of intelligent data mining of the system model in this study. The research shows that the motivation analysis model of the technological startups’ business model based on intelligent data mining analysis proposed in this study meets the basic requirements of the design system of this study.

1. Introduction
The essence of innovative enterprises should be to rely on innovation to make profits and thus rely on innovation to survive and develop. Only when an enterprise can rely on innovation to make profits and thus revolve around innovation to survive and develop, it has the willingness to maintain a high innovation investment, can maintain continuous innovation capabilities, will have endless innovation results, and will undergo long-term accumulation [1]. Otherwise, if an enterprise cannot rely on innovation to make profits and cannot rely on innovation for survival and development, then the enterprise will not be able to maintain a high investment in innovation, it will not be possible to maintain continuous innovation capabilities, and it will not be possible to have endless innovation results. Once the support policy is cancelled, these enterprises will lack the motivation to carry out continuous technological innovation and return to the old path of “innovation inaction,” which makes the original innovation pilot policy investment lose its value. Then, the pilot goal of guiding enterprises to take the road of independent innovation and cultivating a large number of exemplary innovative enterprises will fail [2]. Only by analyzing and grasping the more essential characteristics of innovative enterprises can we formulate more scientific standards, provide a scientific reference for cultivating innovative enterprises and building an innovative country, and ensure the smooth implementation of the pilot work of innovative enterprises in the country’s and the national innovation strategy [3].

This study combines intelligent data mining technology to analyze the business model motivation of technological startups, which provides a theoretical reference for subsequent related research.

2. Related Work
In the early stages of business model development, since no consensus has been reached on the definition of a business model, most studies are revealing specific business model types, among which the research results of e-commerce genres are particularly significant. The e-commerce genre focuses on exploring the “new way of doing business” that appears under network conditions, especially under the
conditions of the Internet [4]. Literature [5] reveals the new economic characteristics under network conditions. First, information assets are not lost in use like physical assets, but can be copied on a large scale at a very low cost. Second, companies can provide products or services to customers through information networks, services forming a new economic scale; third, companies can provide products for cross-industry and cross-regional customers, creating new economies of scope; fourth, greatly reducing transmission costs; and fifth, the transformation of supply and demand, rapidly changing from supply-side to demand-side leading. From an economic point of view, literature [6] summarizes the new laws under the network economy, one is the high production cost, low copy cost, or zero-copy cost of information products; the second is network externalities, where the utility of information products or services increases with users. But with the rapid rise, third is the conversion cost, including learning costs, fixed investment, external utility, etc.; fourth is the lock-in effect, and due to the existence of the above factors, users may be locked in products or services with inferior effects. Literature [7] analyzes and summarizes the four major value creation sources of e-commerce. The first is efficiency, i.e., e-commerce can improve efficiency by reducing search costs, increasing the range of choices, enhancing information symmetry, simplicity, speed, and economies of scale; the second is complementarity. E-commerce can create value by enhancing the complementarity of products and services, online and offline, different technologies, and different activities; the third is novelty, where e-commerce can create new transaction structures, new transaction content, new participants, etc. The fourth is create value with the lock-in effect, where e-commerce can create value by locking customers in by increasing conversion costs, positive network effects, etc. On this basis, literature [8] believes the existing entrepreneurial theories, neither strategic theories can explain the value creation of e-commerce well, so it is necessary to introduce the concept of “business model” as the focus of corporate value research under the network economy, and believe that business model reveals that companies trade with them in order to use opportunities to create value design of content, structure, and governance.

Literature [9] laid the theoretical foundation for product platform research, expanded the most basic product platform concept, and proposed horizontal, vertical, and comprehensive derived product family derived maps, as well as the basic principles of product platform generation update. Literature [10] further expands the application scope of the platform concept and defines the concept of an industrial platform: an industrial platform is a product, service, or technology that is developed by one or several companies as the basis for other companies to create complementary products, services, and technologies. Literature [11] summarized the basic concepts and types of platforms in detail and applied them to different scenarios such as product platforms, supply chain platforms, industrial platforms, and multilateral market platforms. Platform strategy research: literature [12] believes that the platform strategy model is to attract multilateral groups to the platform to form a platform ecosystem so that multilateral groups can interact and cooperate based on the rules and space of the platform and to exert the overall network effect and achieve a win-win situation for all parties. Literature [13], Researching platform strategy: from the perspective of business ecosystem, the research shows that both buyers and sellers of platform companies will affect each other’s income, and the strategies adopted by one side of the platform will also affect the strategic choices of the other side.

Literature [14] proposed that business model innovation not only requires a systematic way of thinking but also needs to consider the degree of matching with the external environment. Enterprise business model innovation can be achieved by changing the four elements of customers, technology, infrastructure, and profit models. Literature [15] analyzes the ways and strategies of enterprise business model innovation based on the perspective of external stakeholders of the enterprise and believes that the driving force for enterprises to achieve continuous innovation of business models is the cooperative operation between enterprises. Literature [16] found that, different from the industrial age enterprises, the business model innovation of modern enterprises is based on the entire ecosystem, which is inseparable from the integration of internal resources of the organization, and the construction of the external environmental ecosystem of the enterprise cannot be ignored. Literature [17] analyzes business model innovation from the perspective of a value chain and believes that enterprise business model innovation can be understood as the change and adjustment of the original value chain by the enterprise or the change of the constituent elements of the enterprise value chain.

3. Intelligent Data Mining Algorithm

In actual machine learning tasks, the learning capabilities of different algorithm models are also different. If a model can achieve a high prediction accuracy rate, then this model is called a “strong learning algorithm”; conversely, if a model’s prediction accuracy rate is only slightly higher than random guessing, then this model is called a “weak learning algorithm.” Generally speaking, compared with weak learning algorithms, strong learning algorithms have a more complex structure and longer training time. For a complex machine learning task, training a strong learning algorithm is difficult, but training a weak learning algorithm is much simpler.

Therefore, it is natural to think that in the task of machine learning, if a “weak learning algorithm” has been obtained, then it can be upgraded to a “strong learning algorithm,” which will reduce the cost of learning while ensuring the learning effect. Researchers have proposed many algorithms for implementing the idea of boosting methods, of which the AdaBoost algorithm is the most representative one.

The AdaBoost algorithm was originally a boosting method proposed for the two-classification problem. The idea is to continuously change the weight distribution of the
data set when training weak classifiers so that different weak classifiers focus on different parts of the data set and finally combine these weak classifiers to get the final strong classifier.

The idea of AdaBoost algorithm needs to study two main issues: the first is how to change the weight distribution of the data set during the training of each weak classifier; the second is how to combine the weak classifiers to generate a strong classifier. In response to the first problem, the AdaBoost algorithm increases the weight of data samples that were judged incorrectly by the previous round of weak classifiers, and at the same time reduces the weight of data samples that are correctly judged and focuses on these judgments in the next round of weak classifier training. Wrong data: intuitively, the AdaBoost algorithm continuously trains the weak classifier to "correct" the previous errors. In theory, the error rate can eventually be minimized; for the second problem, the AdaBoost algorithm uses a weighted linear combination method. Combining weak classifiers into strong classifiers, the algorithm will reduce the weight of weak classifiers with high errors, and at the same time increase the weight of weak classifiers with high accuracy, so that it can play a greater role in the final classification results. One point is very similar to the linear model's idea of weighting features. The AdaBoost algorithm flow is shown in Algorithm 1.

\[
\begin{align*}
    w_{t+1,j} &= \begin{cases} 
    \frac{w_t e^{-\alpha_i}}{Z_t}, & G_t(x_i) = y_i, \\
    \frac{w_t e^{\alpha_i}}{Z_t}, & G_t(x_i) \neq y_j,
    \end{cases} \\
    Z_t &= \sum_{i=1}^{N} e^{-\alpha_i}.
\end{align*}
\]

It can be seen that when a data sample is incorrectly classified by a weak classifier, its weight will increase and the weight of a correctly classified data sample will decrease. Therefore, in the next round of training, the role of misclassified data samples in learning will increase, so that the next round of weak classifiers will pay more attention to learning these data samples. The function of the normalization factor \(Z_t\) is to ensure that the sum of the modified weights is always 1.

Usually, linear regression algorithms are generally used to solve regression problems, i.e., to estimate continuous value results. If you want to use a linear regression model to solve the classification problem, the most direct method is the threshold method, i.e., to classify the data sample by setting a threshold. For example, for a binary classification problem, when the linear regression model predicts a data sample greater than the inter-classification value, it is classified into 1 class, otherwise it is classified into 0 class. For standard data sets, the threshold method is feasible. However, in actual data mining, the linear regression model is susceptible to the influence of data points whose features take larger values in the data set. Therefore, when the data in the data set increases, the linear regression model is constantly adjusting the parameters to fit the new data and the interval value also needs to be adjusted continuously. Therefore, the linear regression model is not a robust model to solve the classification problem through the intermediate value method, and the threshold also has the disadvantage of insufficient interpretability.

Logistic regression model is also called log probability regression, which is a commonly used classification model to deal with binary classification problems. Aiming at the deficiencies of the threshold method, the logistic regression model restricts the predicted value of the linear regression model to a fixed value range and then performs classification by setting the threshold value.

The predictive expression of the logistic regression model is [19]

\[
Y = \begin{cases} 
1, & f(x) \geq p, \\
0, & f(x) < p.
\end{cases}
\]

Among them, \(x \in \mathbb{R}^n\) is the input feature, \(Y \in [0, 1]\) is the output category, \(p \in (0, 1)\) is the classification threshold, and the expression of the regression function \(f(x)\) is [20]

\[
f(x) = g(w \cdot x + b).
\]

Among them, \(w \in \mathbb{R}^n, b \in \mathbb{R}\) is the parameter, \(g(\cdot)\) is the sigmoid function, and the expression is

\[
g(z) = \frac{1}{1 + e^{-z}}.
\]


\[ g(x) = \frac{1}{1 + e^{-x}}. \quad (4) \]

From (4), we can see that the value range of the sigmoid function between the domain \((-\infty, +\infty)\) is \((0, 1)\). At \(x = 0\), \(g(x) = 0.5\), and as \(x\) increases, \(g(x)\) approaches 1. Compared with the threshold method classification, logistic regression has the following advantages:

(a) The model is stable and robust. It can be seen from formula (2)–(10) that logistic regression limits the predicted value of linear regression to an interval. When data points with larger values appear in the data set, the parameters of the model will not be greatly affected.

(b) The model is easy to understand and the prediction results are highly interpretable. Since the value range of the sigmoid function is \((0, 1)\), which is consistent with the value range of the probability, the value of the regression function \(f(x)\) can be regarded as the probability or confidence that the category \(Y\) corresponding to the data sample \(x\) is category 1, and the classification threshold can be set to 0.5. That is, when the probability that the category \(Y\) corresponding to the data sample \(x\) is category 1 is greater than 0.5, then \(x\) is classified into category 1, which is also consistent with actual daily experience.

In the field of machine learning, the Gaussian process is a kind of random process that exists widely in nature and can be used to represent the distribution of a function. The Gaussian process refers to a set of random variables in which any finite number of random variables in this set obeys the joint Gaussian distribution, so the distribution function of this set of random variables obeys the Gaussian process regression. The Gaussian process assumes that similar inputs will produce similar outputs, and this assumption is used to construct a statistical model of the two functions. The statistical characteristics of the traditional Gaussian distribution are determined by the mean and covariance, the statistical characteristics of the Gaussian process are completely determined by its mean function \(m: x \rightarrow R\) and covariance function \(k: x \times x \rightarrow R\). If the function \(f(x)\) obeys the Gaussian process distribution, it is expressed as

\[ f(x) \sim GP(m(x), k(x, x')). \quad (5) \]

The Gaussian process is different from the ordinary Gaussian distribution. The probability density function \(f(x)\) corresponding to any \(x\) is no longer a scalar but a normal distribution function. This function reflects the probability distribution of \(f(x)\). Usually, for the convenience of calculation, it is assumed that the mean value function of the Gaussian process is \(m(x) = 0\). For the covariance function \(k\), the exponential square function is often selected as the covariance function:

\[ k(x_i, x_j) = \exp\left(\frac{1}{2\sigma^2}||x_i - x_j||^2\right). \quad (6) \]

When the values of \(x_i\) and \(x_j\) are close, the function value is close to 1, otherwise, the function value is close to 0. This function reflects that when two sampling points are close, they have greater mutual influence and strong correlation. However, when the distance is farther, the correlation between them becomes weaker.

The process of using Gaussian process to determine the posterior distribution probability of function \(f(x)\) is as follows:

(1) The algorithm first selects \(t\) observations of function \(f\) as the training set \(D_{1:t} = \{x_n, f_n\}_{n=1}^{t}, f_n = f(x_n)\). We assume that the function \(f\) obeys the \(t\)-dimensional normal distribution, that is, \(f \sim N(0, K)\), where:

\[
K = \begin{bmatrix}
k(x_1, x_1) & k(x_1, x_2) & \ldots & k(x_1, x_t) \\
k(x_2, x_1) & k(x_2, x_2) & \ldots & k(x_2, x_t) \\
\vdots & \vdots & \ddots & \vdots \\
k(x_t, x_1) & k(x_t, x_2) & \ldots & k(x_t, x_t)
\end{bmatrix}. \tag{7}
\]

Each element in the matrix \(K\) is calculated by the covariance function \(k\). The function \(k\) measures the degree of approximation between two input values. Obviously, without considering the influence of noise, the diagonal element is \(k(x_i, x_i) = 1\).

(2) According to the established function model, the algorithm calculates the function value \(f_{n+1} = f(x_{n+1})\) corresponding to the new sampling point \(x_{n+1}\). According to the assumption of the Gaussian process, the new set composed of \(f_{1:t}\) and the function value \(f_{t+1}\) in the training set obeys the \(t + 1\)-dimensional normal distribution; that is,

\[
\begin{bmatrix} f_{1:t} \\ f_{t+1} \end{bmatrix} \sim N\left(0, \begin{bmatrix} K & k \\ k^T & k(x_{t+1}, x_{t+1}) \end{bmatrix}\right). \tag{8}
\]

Among them,
where \( f_{t+1} \) obeys one-dimensional normal distribution, that is, \( f_{t+1} \sim N(\mu_{t+1}, \sigma_{t+1}^2) \). From the joint normal distribution, it can be solved:

\[
\begin{align*}
\mu_{t+1}(x_{t+1}) &= k^T K^{-1} f_{t+1}, \\
\sigma_{t+1}^2(x_{t+1}) &= -k^T K^{-1} k + k(x_{t+1}, x_{t+1}).
\end{align*}
\]

(10)

It can be seen from the above formula that what is obtained by the Gaussian process is not a point estimate of the function value \( f_{t+1} \), but a probability distribution of all the values of \( f_{t+1} \). If the number of samples in the training set is large enough, the Gaussian process can obtain an approximate estimate of the function \( f(x) \) distribution. A more intuitive representation of the Gaussian process is shown in Figure 1.

Figure 1 shows a one-dimensional Gaussian process with two observations. The black points in the figure represent the observed data points, the black curve represents the mean value of the function value in the Gaussian process domain, and the blue area indicates the value range of the function value within one standard deviation of the mean value. It can be seen from the figure that the variance of function \( f \) is smaller when it is close to the observation data point and larger when it is far away from the observation data point.

After obtaining the posterior distribution of the function, Bayesian optimization uses the acquisition function \( u \) to derive where the function \( f \) obtains the maximum value. We assume that the higher value of the acquisition function corresponds to the larger value of the objective function \( f \). Therefore, maximizing the collection function is equivalent to maximizing the function \( f \):

\[
x^+ = \arg \max_{x \in A} u(x|D).
\]

(11)

The following will introduce several common acquisition functions when the prior distribution is Gaussian distribution. Before the introduction, we explain the following symbols and theorems:

- \( \Phi(\cdot) \) - the cumulative distribution function of the standard normal distribution;
- \( \varphi(\cdot) \) - the probability density function of the standard normal distribution;
- \( x^+ \) represents the position where the function \( f \) obtains the optimal value after obtaining \( t \) sampling points according to Bayesian optimization, where the expression of \( x^+ \) is as follows:
A continuous random variable $X$ has a density function $f(x)$ and a probability distribution function $F(x)$. Then, for any constant $a < b$, there is:

$$P(a \leq X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)\,dx.$$  \hfill (13)

We set the random variable $X$ to have a probability density function $f(x)$, if

$$\int_{-\infty}^{\infty} |x| f(x)\,dx = \infty.$$  \hfill (14)

Then,

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx,$$  \hfill (15)

is the mathematical expectation of $X$.

3.1. Probability of Improvement (PI). The function optimization strategy is to explore near the current optimal value point to find the point that is most likely to be better than the current optimal value until the number of algorithm iterations reaches the upper limit. The function expression is:

$$PI(x) = P(f(x) \geq f(x'))$$

$$= \Phi\left(\frac{\mu(x) - f(x)}{\sigma(x)}\right).$$ \hfill (16)

The PI strategy is simple, but its shortcomings are very obvious. Since the function is based on the greedy idea, that is, only considering the use of the current optimal solution, the selection of sampling points is limited to a small range, and it is easy to fall into the local optimal solution.

In order to solve the problem that the PI function is easy to fall into the local optimal solution, a parameter is usually added when using the PI function. Only when the difference between the value of the next sampling point and the current optimal value is not less than, the sampling point will be considered to replace the current optimal value point. The expanded PI function expression is

$$PI(x) = P(f(x) \geq f(x^*) + \epsilon)$$

$$= \Phi\left(\frac{\mu(x) - f(x^*) - \epsilon}{\sigma(x)}\right).$$ \hfill (17)

Figure 4: The mechanism of the effectiveness of the technological startups business model.
3.2. Expected Improvement (EI). The optimization strategy of this function is to calculate the expectation of the improvement degree of the function value when exploring the vicinity of the current optimal value point. If the improvement of the function value after the algorithm is executed once is less than the expected value, it means that the current optimal value point may be a local optimal solution, and the algorithm will search for the optimal value point in other positions in the domain. Compared with the PI function, the EI function is not easy to fall into the local optimal solution.

When using the EI function, the improvement $I$ of the function after the optimization algorithm is executed once is defined as the difference between the sampling point value and the current optimal value. If the sampling point value is less than the current optimal value, the function improvement degree is 0; that is,

$$I(x) = \max\{0, f_{t+1}(x) - f(x^*)\}.$$  \hspace{1cm} (18)

According to the EI function optimization strategy, the next sampling point of the function is the expected maximum value of the function improvement, namely,

$$x = \arg\max E[I(x)]$$
$$= \arg\max E(\max\{0, f_{t+1}(x) - f(x^*)\}).$$  \hspace{1cm} (19)

When $f_{t+1}(x) - f(x^*) \geq 0$ is, under the condition of Gaussian process prior, function $f_{t+1}(x)$ obeys a normal distribution with mean $\mu(x)$ and variance $\sigma^2(x)$. Therefore, the random variable $I$ obeys a normal distribution with mean $\mu(x) - f(x^*)$ and variance $\sigma^2(x)$, that is, the distribution function of $I$ is
The expectation of the improvement degree $I$ of the function is

$$E(I) = \int_{-\infty}^{\infty} I f(I) dI$$

$$= \int_{-\infty}^{\infty} I^{1/2} \pi \sigma(x) \exp\left(-\frac{\left(I - f(x)\right)^2}{2\sigma^2(x)}\right) dI$$

$$= \sigma(x) \Phi(Z) + \phi(Z).$$

Among them,

$$Z = \frac{\mu(x) - f(x^*)}{\sigma(x)}.$$  \hspace{1cm} (22)

Formula (21) expresses the expectation of the improvement degree of the function, that is, the expression of the EI function.

3.3. GP Upper Confidence Bound (GP-UCB). In the process of executing an optimization algorithm, this function judges whether to use the current optimal value point (corresponding to the high $\mu(x)$ interval) or explore other low confidence intervals (corresponding to the high $\sigma(x)$ interval) in the next execution. The function balances these two behaviors by setting an additional parameter $\kappa$. The function expression is defined as

$$UCB(x) = \mu(x) + \kappa \sigma(x).$$ \hspace{1cm} (23)

To determine the acquisition function, it is necessary to observe the regression performance of the three functions in the same Gaussian process. Figure 2 shows the optimization process comparison of PI, EI, and GP-UCB for the same function. The blue curve is the mean line of the predicted function value of the Gaussian process, the red dotted line is the actual value curve of the function, the red solid line is the current optimal solution, and the blue area is the value range of the function value within one standard deviation near the mean. It can be seen from Figure 2 that after 5 iterations of the algorithm, both the EI function and the GP-UCB function have found the global optimal value of the function, while the PI function has fallen into the local optimal solution. According to the experimental comparison results, we use the EI function to optimize the subsequent machine learning algorithm parameters because EI has better performance than PI on multiple local optimal solution problems. In addition, compared with GP-UCB, EI does not set other parameters, so it has more advantages than GP-UCB in terms of simplicity.

4. Motivation Analysis of the Technological Startups Business Model Based on Intelligent Data Mining Analysis

As shown in Figure 3, this study believes that the connotation of the business model of effective technological startups is to make correct strategic decisions, realize the value creation of the enterprise, bring relevant economic and social benefits to the enterprise in the efficient operation of
the entire industrial chain, and maintain the long-term sustainable and stable development of the enterprise.

Taking the constituent elements of the technological startups business model as the starting point of the research, this article further identifies the key factors that affect the effectiveness of the technological startups business model and verifies the mechanism of these influencing factors on the effectiveness of the technological startups business model. This can fully explain the feasibility of evaluating the effectiveness of the technological startups business model by using factors that affect the effectiveness of the technological startups business model as an evaluation index. Through relevant analysis, this article believes that the mechanism of commercial effectiveness is shown in Figure 4.

Commercial customer value, financial value, brand building, supplier relations, and sustainable development are all significantly related to the effectiveness of corporate technological startups business models. According to the

---

**Figure 8:** Schematic diagram of the logical structure of the platform.

**Figure 9:** Platform module structure diagram.
results of gray correlation analysis, the revised technical startups business model effectiveness mechanism is shown in Figure 5, and the following specific analysis is carried out.

This study constructs a relationship network for the evaluation of the effectiveness of the technological startups business model starting from the components of the technological startups business model, as shown in Figure 6.

The effectiveness of the technological startups business model is mainly reflected in the sustainable development of the company, and human resources are the basis for the integration of various resources of the company, ensuring the high growth of the company. Moreover, a stable cooperative relationship with suppliers is also the basic guarantee for the long-term development of an enterprise. It is mainly measured by the proportion of scientific research personnel, employee satisfaction, risk identification ability, new product development ability, and cooperation stability.

In summary, we have established an evaluation index system for the effectiveness evaluation of the technological startups business model, as shown in Figure 7.

Based on the above analysis, the logical structure of the platform as shown in Figure 8 can be designed.

The structure diagram of the platform module is shown in Figure 9.

After constructing the above model structure, the intelligent data mining effect of the system model in this study is evaluated in the form of simulation experiment, and the data mining effect is calculated, and the results shown in Table 1 are obtained.

From the above analysis, we can see that the motivation analysis model of technological startups business model based on intelligent data mining analysis has a good effect in data mining. On this basis, the motivation analysis effect of this system is evaluated and the results shown in Table 2 are obtained.
From the above research, we can see that the motivation analysis model of technological startups business model based on intelligent data mining analysis in this study meets the basic requirements of the system design in this study.

5. Conclusion
Research on new-type enterprises has important guiding significance for China’s construction of an innovative country. However, although the current research has laid a theoretical foundation in the definition of innovative enterprises, business models, intellectual property rights, value chains, and value networks, in general, they have not clearly revealed the nature and business model issues of innovative enterprises. Therefore, it is necessary to take innovative enterprises as the basic research object, and in accordance with the theoretical framework of business models, from the perspective of innovative enterprises’ value source, internal value chain, and external value network, reveal the process of innovative enterprises’ knowledge value creation around intellectual property rights. At the same time, it is necessary to explain the composition and characteristics of the external value network they are in, analyze the business model of innovative enterprises with knowledge assets as the core, and reveal the specific types of business models under different social and economic conditions to fill the theoretical gaps of innovative enterprises. This article combines intelligent data mining technology to analyze the motivations of the technological startups business model, which provides a theoretical reference for subsequent related research.

Data Availability
The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare no competing interests.

Acknowledgments
This study was sponsored by MOE (Ministry of Education in China) Youth Foundation Project of Humanities and Social Sciences, Research on the Influence Mechanism of Institutional Environment on Innovative Enterprise Innovation Performance (No. 20YJC630116); and Guangdong High-level Talent Introduction Foundation Project, A System Dynamics Simulation Analysis of the Relationship between Institutional Environment, Business Model and Innovation Performance (No. Gccrcxm-201909).

References
[1] C. Barnard, “EU employment law and the European social model: the past, the present and the future,” Current Legal Problems, vol. 67, no. 1, pp. 199–237, 2014.
[2] E. Totty, “The effect of minimum wages on employment: a factor model approach,” Economic Inquiry, vol. 55, no. 4, pp. 1712–1737, 2017.
[3] J. H. Fede, K. S. Gorman, and M. E. Cimini, “Student employment as a model for experiential learning,” Journal of Experiential Education, vol. 41, no. 1, pp. 107–124, 2018.
[4] S. Van de Walle, B. Steijn, and S. Jilke, “Extrinsic motivation, PSM and labour market characteristics: a multilevel model of public sector employment preference in 26 countries,” International Review of Administrative Sciences, vol. 81, no. 4, pp. 833–855, 2015.
[5] P. Barbieri, R. Bozon, S. Scherer, R. Grottì, and M. Lugo, “The Rise of a Latin Model? Family and fertility consequences of employment instability in Italy and Spain,” European Societies, vol. 17, no. 4, pp. 423–446, 2015.
[6] U. Huws, N. H. Spencer, and D. S. Syrdal, “Online, on call: the spread of digitally organised just-in-time working and its implications for standard employment models,” New Technology, Work and Employment, vol. 33, no. 2, pp. 113–129, 2018.
[7] J.-F. Wen and D. V. Gordon, “An empirical model of tax convexity and self-employment,” The Review of Economics and Statistics, vol. 96, no. 3, pp. 471–482, 2014.
[8] R. Nakase-Richardson, L. F. Stevens, C. Dillahunt-Aspillaga et al., “Predictors of employment outcomes in veterans with traumatic brain injury: a VA traumatic brain injury model systems study,” The Journal of Head Trauma Rehabilitation, vol. 32, no. 4, pp. 271–282, 2017.
[9] M. Bornay-Barrachina, A. López-Cabral, and R. Valle-Cabrera, “How do employment relationships enhance firm innovation? The role of human and social capital,” International Journal of Human Resource Management, vol. 28, no. 9, pp. 1363–1391, 2017.
[10] G. R. Bond and R. E. Drake, “Making the case for IPS supported employment,” Administration and Policy in Mental Health and Mental Health Services Research, vol. 41, no. 1, pp. 69–73, 2014.
[11] D. Tavani and L. Zamparelli, “Endogenous technical change, employment and distribution in the Goodwin model of the growth cycle,” Studies in Nonlinear Dynamics & Econometrics, vol. 19, no. 2, pp. 209–216, 2015.
[12] P. Frijters, D. W. Johnston, and M. A. Shields, “The effect of mental health on employment: evidence from Australian panel data,” Health Economics, vol. 23, no. 9, pp. 1058–1071, 2014.
[13] F. Portella-Carbó, “Effects of international trade on domestic employment: an application of a global multiregional input-output supermultiplier model (1995-2011),” Economic Systems Research, vol. 28, no. 1, pp. 95–117, 2016.
[14] G. R. Bond, S. J. Kim, D. R. Becker et al., “A controlled trial of supported employment for people with severe mental illness and justice involvement,” Psychiatric Services, vol. 66, no. 10, pp. 1027–1034, 2015.
[15] G. Topa, C.-M. Alcover, J. A. Moriano, and M. Depolo, “Bridge employment quality and its impact on retirement adjustment: a structural equation model with SHARE panel data,” Economic and Industrial Democracy, vol. 35, no. 2, pp. 225–244, 2014.
[16] M. Shirmohammadi, M. Beigi, and J. Stewart, “Understanding skilled migrants’ employment in the host country: a multidisciplinary review and a conceptual model,” International Journal of Human Resource Management, vol. 30, no. 1, pp. 96–121, 2019.
[17] G. P. Clarkson, “Twenty-first century employment relationships: the case for an altruistic model,” Human Resource Management, vol. 53, no. 2, pp. 253–269, 2014.
[18] C. Topsakal, “Policy on teacher employment model in Turkey: contracted teachers,” *International Journal of Educational Methodology*, vol. 5, no. 4, pp. 671–682, 2019.

[19] L. Ottomanelli, S. Barnett, L. Goetz, and R. Toscano, “Vocational rehabilitation in spinal cord injury: what vocational service activities are associated with employment program outcome?” *Topics in Spinal Cord Injury Rehabilitation*, vol. 21, no. 1, pp. 31–39, 2015.

[20] C. C. Lo and T. C. Cheng, “Race, employment disadvantages, and heavy drinking: a multilevel model,” *Journal of Psychoactive Drugs*, vol. 47, no. 3, pp. 221–229, 2015.