Hindawi
Mathematical Problems in Engineering
Volume 2023, Article ID 9852786, 1 page
https://doi.org/10.1155/2023/9852786

Retraction

Retracted: Analysis on the Development Strategy of Private Education Based on Data Mining Algorithm

Mathematical Problems in Engineering

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Xing and D. Maia, "Analysis on the Development Strategy of Private Education Based on Data Mining Algorithm," *Mathematical Problems in Engineering*, vol. 2022, Article ID 2783398, 10 pages, 2022.

Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 2783398, 10 pages
https://doi.org/10.1155/2022/2783398

Research Article

Analysis on the Development Strategy of Private Education Based on Data Mining Algorithm

Hongjun Xing[1] and Darchia Maia[2]

[1]College of Continuing Education, Jilin Jianzhu University, Changchun, Jilin 130118, China
[2]School of Science and Technology, The University of Georgia, Tbilisi 017, Georgia

Correspondence should be addressed to Darchia Maia; dr.darchia@stu.zf.edu.ge

Received 20 April 2022; Revised 14 May 2022; Accepted 11 June 2022; Published 11 July 2022

Academic Editor: Naeem Jan

In order to improve the development effect of private education, this paper analyzes the current situation of private education combined with the data mining algorithm and explores the problems existing in the development of private education. Moreover, this paper combines the semi-parametric product estimation method with parameter estimation and applies the estimation method to model-assisted sampling estimation. This work enhances the estimate accuracy of the sample estimation and increases the field of application of the model while enhancing the classic generalized regression estimation. It also modifies the estimation accuracy on the basis of the linear assumption. The experimental study reveals that the data mining algorithm-based analysis approach for private education development provided in this work has a certain impact, and the development strategy of private education is assessed on this premise.

1. Introduction

Education is the cornerstone of national rejuvenation and social progress. Since China is a developing country, in order to improve the quality of the whole people, improve China's comprehensive national strength, cultivate knowledge innovation and scientific and technological innovation talents, and occupy a dominant position in the fierce international competition, the strategy of rejuvenating the country through science and education must be implemented. The execution of this approach necessitates the prioritisation of education above development, with basic education taking precedence. Basic education has undergone drastic reforms and attempts in terms of ideas, concepts, methods, content, management systems, and school-running systems in more than 30 years of reform and opening up in order to meet the needs of the development of the situation, and has achieved gratifying results. Human civilization, on the other hand, has entered the period of knowledge economy and the information age as we approach the twenty-first century. The advancement of time has raised the bar for the growth of abilities and the general quality of the country, requiring more individuals to obtain more and better education, as well as master more and newer knowledge and information. Furthermore, shifts in reality have posed significant problems to basic education. However, compared with the huge population base, the investment in basic education is seriously insufficient, the allocation of resources is limited, and there are large differences in education between regions. This disproportionate phenomenon has produced a very serious contradiction with the development needs of socialist modernization and with the increasingly fierce international competition for talents. For a poor country to run large-scale education, it has no choice if it does not take the road of running schools in various forms. Therefore, the reality requires that the development of China's basic education must change the situation of government-run education, form a pattern of common development of private education with government-run schools as the main body and public education, and must attract private funds to run schools and vigorously develop private education.

China's private education has not developed for a long time, and it has developed outside the relatively complete public higher education system in our country. With the gradual attention of the state and the promulgation of relevant laws and regulations, private education in China has basically enjoyed the same legal status as public education.

As an important part of higher education in our country, private colleges and universities have played a pivotal role in the period of popular education in our country. However, private colleges and universities, while reducing the pressure of education and teaching in public colleges to a certain extent, have also formed a competitive situation with public colleges and universities. Judging from the current situation, the competition between public and private colleges and universities will help public colleges and universities reduce the cost of running schools, improve the quality and efficiency of education, and establish a modern university system. But at present, there are still many deficiencies in the research of private higher education in our country. Due to the late start of private education in our country, it has only improved significantly after the reform and opening up. However, there are still very few people engaged in the research of private education. Compared with public education, private education has many shortcomings, and the research team needs to be expanded and strengthened. Today, there are many problems in private education, especially in private higher education. There are still loopholes that have not been patched, such as the government's lag in policy formulation, the unclear nature of the legal person of private colleges and universities, and the inability to fully implement legal support and incentive measures. There is research value and space. It is hoped that this research can enrich the research results of my country's private higher education to a certain extent, improve my country's private higher education system, and promote the stable and orderly development of my country's private education. "The mission of private higher education is not only to promote the popularization of higher education, but the most fundamental mission of the existence and development of private higher education in China is to revitalize the existing higher education resources and form a pattern in which public and private institutions of higher learning compete with each other and develop together."

This paper analyzes the current situation of private education combined with a data mining algorithm, explores the problems existing in the development of private education, and analyzes the development strategy of private education.

2. Related Work

Educational Equity. The educational ideal that mankind has pursued since ancient times. Literature [1] proposes: "Education can transcend all other differences of human origin and is an important equalizer for human development." For public education, it advocates freedom, neutrality, gender equality, and no sectarian color discrimination. Document [2] advocates the principle of distributive justice. The first is the principle of equality and freedom, which advocates that everyone enjoys the same rights and the most extensive freedom as others. The second is the difference principle and the principle of fairness and equality of opportunities, which advocates that in dealing with economic and political imbalances, the disadvantaged should obtain the greatest benefits, and opportunities should be opened to all under fair and equal opportunities. In terms of educational resource allocation, it is reflected in the inclination of resources to the most disadvantaged [3]. With the continuous development of research, scholars have found that distributive justice is flawed; it ignores the institutional background of formulating resource allocation and cannot deal with the problem of nonmaterial goods and resources [4]. Literature [5] puts forward the theory of relational justice, arguing that a truly fair and just social system should satisfy two conditions: one is to develop and stimulate an individual's unique ability and allow him to fully express his inner thoughts and feelings; the behavior and conditions of decision-making. In the field of education, to reflect educational equity, schools should formulate policies to accommodate disadvantaged groups, give them a voice, and be considerate and considerate of their needs and feelings.

According to mainstream research, educational fairness is primarily expressed in three reasonable criteria for allocating educational resources: (1) the idea of equality, which encompasses two aspects: equal educational rights and opportunities. (2) The concept of difference, which is used variably depending on the educated individual differences, represents disparities in educational resource allocation rather than an equal or equal distribution of educational resources. (3) Principle of compensation pay attention to the disparities in the socioeconomic situation of the educated and in the distribution of educational resources; provide some compensation to the educated with disadvantaged socioeconomic status [6].

Further in-depth study of research equity will involve research on the measurement of educational equity. According to the object and goal of educational equity, from the perspective of school finance, literature [7] selects three dimensions of educational opportunity, resources, and results to construct an educational equity measurement framework. Literature [8] measures educational equity from four aspects, specifically educational input, learner educational process, educational output, educational goals, and establishes a series of indicator systems. Literature [9] established a series of index systems for the measurement of educational equity, including the social and economic status of adult citizens, the financial status of educational resources, school quality and learning environment, access to education and training opportunities, student graduation rates, and student academic performance. Adult literacy rate, market productivity level, and other dimensions. Literature [10] constructs a multi-level measurement index system from four dimensions affecting educational equity. The four dimensions include: social, political, and cultural background; educational process, educational internal results; and educational external results.

Literature [11] pointed out that the balanced development of education should take into account regional balance, group balance, quality balance, school balance, resource balance allocation, and respect for differences in latitudes for reasonable control and distribution. From the perspective of system balance, the compulsory education system is huge and complex, including multiple factors such as management authorities, schools, teachers, students, and the market [12]. Due to the lack of functions of the education administrative department at this stage, the urgent needs of school professional development and the appeal of the interests of civil groups, and the literature [13] believes that the third party's participation in the balanced development of compulsory education has necessary practical significance. To a certain extent, the third party can supplement the lack of management functions of the education administrative department, provide support for the professional development of primary and secondary education and teaching, and meet the needs of different interest groups for the development of compulsory education.

Literature [14] believes that after the development of compulsory education gradually matures, the compulsory education policy should develop towards the realization of equality in the process of compulsory education and the results of education. In the case of insufficient financial investment in education, the establishment of high-quality private schools increases the total amount of social education resources and improves the overall level of compulsory education. However, some more developed regions take advantage of the real estate policy to set up private schools jointly with high-quality public schools, so that parents must bear high tuition fees to allow their children to enjoy high-quality compulsory education. This development model makes it easier for the superior class to obtain high-quality educational resources, accelerates the solidification of the class, and creates inequities in education. Compulsory education is a public product. If high-quality compulsory education is transferred from public schools to private schools, it will run counter to the public welfare of compulsory education. The government is obliged to control and adjust the education inequities brought about by the development of private schools [15]. Education has both public and private market attributes. The government, as a "provider" of education, but not a "producer," solves education problems by purchasing services from the market.

There are two means to adjust the allocation of production resources: government adjustment and market adjustment. As a kind of production resources, educational resources are also the same. Excessive reliance on government regulation or market regulation cannot achieve the unity of efficiency and fairness in the allocation of educational resources [16]. Excessive reliance on government regulation will lead to inefficient allocation of resources in the education market and unlimited expansion of government responsibilities. Therefore, the government should give full play to the ability of market regulation, and on the premise of ensuring the quality and fairness of education, use financial leverage to promote the development of private education, enhance the vitality of the education market, and promote the harmonious development and progress of education and society [17]. The balanced development of educational resources is a multi-dimensional, complex, and overall structure, including school balance, regional balance, group balance, resource balance allocation, quality balance, and respect for differences. In order to achieve educational equity, the government should rationally regulate and allocate educational policies, school conditions, and educational resources from the perspective of the balanced development of educational resources. The phenomenon of school choice reflects the shortage of high-quality compulsory education resources and the unbalanced allocation of educational resources [18].

3. Model-Aided Sampling Estimation Based on Semi-Parametric Product Method

This chapter largely applies the estimate approach to model-assisted sampling estimation, which is based on the semi-parametric product estimation method presented in the previous section, which combines the benefits of parametric and non-parametric estimation. While boosting the standard generalized regression estimation, it also increases sample estimation accuracy and broadens the model's application range.

The traditional model-assisted sampling estimation methods assume that the super-population model is a linear regression relationship. However, the actual survey data often rarely satisfy the linear regression relationship but present various complex changing relationships. Therefore, without limitation in this paper, we assume that the model regression function is a smooth function in the general form and establish a super-population regression model relationship as shown in formula (1), where $m(\cdot)$ is the smooth function of the auxiliary variable x_k. According to the semi-parametric estimation model expression, its corresponding super-population regression model ξ can be expressed as the following form:

$$\begin{aligned} y_k &= m(x_k) + \varepsilon_k \\ &= m(x_k, \beta) \cdot \left(\frac{m(x_k)}{m(x_k, \beta)} \right) + \varepsilon_k, \end{aligned} \tag{1}$$

here, parameter estimation is used for the first factor in the above formula $m(x_k)$, and its estimator $m(x_k, \widehat{\beta})$ is obtained.

The second factor is called the adjustment factor, denoted as $r(x_k) = (m(x_k)/m(x_k, \beta))$, which can be estimated by appropriate non-parametric methods, and its estimator is denoted as $\widehat{r}(x_k)$.

In order to simplify the problem, this section only considers the model-aided estimation problem under a single auxiliary variable x, and the case of multiple auxiliary variables can be similarly generalized. The specific expression of the $m(x_k)$ parameter term of the regression function in the definition formula (1) is a linear function; that is,

$$m(x_k, \beta) = \beta_0 + \beta_1 x_k. \tag{2}$$

According to the above definitions of $m(x_k, \beta)$ and $r(x_k)$, the specific expression of $m(x_k)$ can be obtained as

$$\begin{aligned} m(x_k) &= m(x_k, \beta) r(x_k) \\ &= (\beta_0 + \beta_1 x_k) \frac{m(x_k)}{(\beta_0 + \beta_1 x_k)}. \end{aligned} \tag{3}$$

Therefore, the new semi-parametric (SEM) super-population regression model defined in this section is specifically expressed as

$$y_k = m(x_k) + \varepsilon_k = (\beta_0 + \beta_1 x_k) \cdot r(x_k) + \varepsilon_k. \tag{4}$$

Here, for any $k \in U$, there is $E_\xi(\varepsilon_k) = 0, V_\xi(\varepsilon_k) = \sigma^2 < \infty, r(x_k) = (m(x_k)/(\beta_0 + \beta_1 x_k))$.

According to the principle of model-assisted sampling estimation, the key step in the regression estimation of the super-population model is how to obtain the estimated value of the research variable. Because the closeness of the estimated value of the research variable to the true value directly affects the accuracy of the sampling estimate. According to the basic principle of the abovementioned semi-parameter product adjustment method, we assume that the observed value $\{(y_k, x_k)\}$ (here $k \in U$) of all population units in the finite population can be obtained. This section theoretically presents the following sampling estimation steps:

First, the least squares method is used to estimate the linear part of the semi-parametric regression model. We define A and B such that $\min \sum_U [y_k - A - Bx_k]^2$, then we have

$$B = \frac{\sum_U (x_k - \overline{X})[y_k - \overline{Y}]}{\sum_U (x_k - \overline{X})^2}, \tag{5}$$

$$A = \overline{Y} - B\overline{X}. \tag{6}$$

Among them, $\overline{Y} = (\sum_U y_k/N), \overline{X} = (\sum_U x_k/N)$. According to the definition of $m(x_k, \beta)$ in formula (4), its unbiased estimator based on the finite population design can be obtained as

$$m\left(x_k, \widehat{\beta}\right) = A + Bx_k. \tag{7}$$

Then, the local polynomial method is used to fit and estimate the adjustment factor. Theoretically, any kernel estimation method can be adopted for the estimation of the adjustment term $r(x_k)$. However, considering the excellent property of the local polynomial method without boundary effects, this paper adopts the local polynomial kernel estimation. K(·) is the kernel function, h is the bandwidth, and the nonparametric adjustment term $r(x_k)$ is obtained by using p-order local polynomial regression. The estimator based on the finite population is

$$\widehat{r}_{kU} = e_1' \left(X_{Uk}' W_{Uk} X_{Uk}\right)^{-1} X_{Uk}' W_{Uk} V_U. \tag{8}$$

Among them, e_1 is a $p+1$-dimensional vector whose first item is 1 and the remaining p items are 0, V_U represents the vector $((y_1/m(x_1, \widehat{\beta})), \ldots, (y_N/m(x_N, \widehat{\beta})))$, the $N \times N$-dimensional matrix is $W_{Uk} = \mathrm{diag}\{(1/h)K((x_l - x_k)/h)\}_{l \in U}$, and the $N \times (p+1)$-dimensional matrix is $X_{Uk} = [1, x_l - x_k, \ldots, (x_l - x_k)^p]_{l \in U}$.

Finally, the linear part obtained above and the estimator of the adjustment term are multiplied, and the unbiased estimator $\widehat{m}_{kU}$ of the semi-parametric function $m(x_k)$ based on the finite population can be obtained as:

$$\begin{aligned} \widehat{m}_{kU} &= (A + Bx_k)\widehat{r}_{kU} \\ &= (A + Bx_k) e_1' \left(X_{Uk}' W_{Uk} X_{Uk}\right)^{-1} X_{Uk}' W_{Uk} V_U. \end{aligned} \tag{9}$$

Theoretically, in this estimation method, when the fitting value $m(x_k, \widehat{\beta})$ of the linear part is exactly proportional to the real semi-parametric function $m(x_k)$, the adjustment factor is a constant. When $m(x_k, \widehat{\beta})$ is compared with the real semi-parametric function $m(x_k)$, there is a model setting error, and $\widehat{r}_{kU}$ will adjust it according to the size of the error, which can reduce the variance of the estimator.

In the actual investigation, only the observed value $\{(y_k, x_k)\}$ of the sample unit (here $k \in s$) can be obtained, and the estimator $\widehat{m}_{kU}$ based on the finite population cannot be calculated. Therefore, it is necessary to construct a consistent sample-based estimator $\widehat{m}_{ks}$ to replace $\widehat{m}_{kU}$, and then the corresponding model-assisted sampling estimator is given. The specific sampling estimation steps are as follows:

First, based on the sample survey data, the Horvitz-Thompson estimator (also known as the π estimator) is used to estimate the population parameters A and B of the linear part of the model as

$$\widehat{B} = \frac{\left(\sum_{k \in s} (x_k - \widetilde{x}_s)(y_k - \widetilde{y}_s)/\pi_k\right)}{\left(\sum_{k \in s} (x_k - \widetilde{x}_s)^2/\pi_k\right)}, \tag{10}$$

$$\widehat{A} = \widetilde{y}_s - \widehat{B}\widetilde{x}_s. \tag{11}$$

Among them, $\widetilde{x}_s = ((\sum_{l \in s} x_l/\pi_l)/\widehat{N}), \widetilde{y}_s = ((\sum_{l \in s} y_l/\pi_l)/\widehat{N}), \widehat{N} = (\sum_{l \in s} 1/\pi_l)$. From this, the sample estimator of $m(x_k, \widehat{\beta})$ can be obtained as

$$m\left(x_k, \widehat{\beta}^*\right) = \widehat{A} + \widehat{B}x_k. \tag{12}$$

Secondly, according to the sample survey data, we use the local polynomial estimation combined with the π estimation method to estimate the adjustment items of the model. $V_s = ((y_1/m(x_1, \widehat{\beta}^*)), \ldots, (y_n/m(x_n, \widehat{\beta}^*)))$, $n \times (p+1)$ dimension matrix is $X_{sk} = [1, x_l - x_k, \ldots, (x_l - x_k)^p]_{l \in s}$, $n \times n$ dimension matrix is $W_{sk} = \mathrm{diag}\{(1/\pi_l h)K((x_l - x_k)/h)\}_{l \in s}$, the sample estimator of $\widehat{r}_{kU}$ can be obtained as

$$\widehat{r}_{ks} = e_1' \left(X_{sk}' W_{sk} X_{sk}\right)^{-1} X_{sk}' W_{sk} V_s. \tag{13}$$

Then, the linear part obtained above is multiplied by the sample estimator of the adjustment term and the sample estimator of the semi-parametric function $m(x_k)$ can be obtained as

$$\widehat{m}_{ks} = m\left(x_k, \widehat{\beta}^*\right) \cdot \widehat{r}_{ks}. \tag{14}$$

Finally, combined with the generalized difference estimation formula, the overall total value Y can be adjusted based on the semi-parameter product of the sample s by bringing $\widehat{m}_{ks}$ into the formula

$$\widehat{Y}_{SEM} = \sum_{s} \frac{y_k - \left(\widehat{A} + \widehat{B}x_k\right)\widehat{r}_{ks}}{\pi_k} + \widehat{A}\sum_{U}\widehat{r}_{ks} + \widehat{B}\sum_{U} x_k \widehat{r}_{ks}. \quad (15)$$

The approximate variance of the above estimator $\widehat{Y}_{\mathrm{SEM}}$ is

$$AV\left(\widehat{Y}_{\mathrm{SEM}}\right) = \sum\sum_{U} \left(\pi_{kl} - \pi_k \pi_l\right) \frac{y_k - \widehat{m}_{kU}}{\pi_k} \frac{y_l - \widehat{m}_{lU}}{\pi_l}. \quad (16)$$

Similarly, the estimator of the variance of the estimator $\widehat{Y}_{\mathrm{SEM}}$ obtained by applying the π estimation method is

$$v\left(\widehat{Y}_{SEM}\right) = \sum\sum_{s} \frac{\pi_{kl} - \pi_k \pi_l}{\pi_{kl}} \frac{y_k - \widehat{m}_{ks}}{\pi_k} \frac{y_l - \widehat{m}_{ls}}{\pi_l}. \quad (17)$$

It can be seen from the above that the super-population model on which the semi-parametric model auxiliary regression estimator $\widehat{Y}_{\mathrm{SEM}}$ is based is a smooth function without any restrictions. However, on the basis of classical linear regression estimation, the local polynomial kernel estimation approach is employed for additional modification in the model estimating process. This not only enhances the accuracy of sample estimates, but also broadens the field of application by making this set of model estimating techniques more applicable for all types of real survey data. A detailed theoretical study is conducted on the properties of the auxiliary regression estimator of the local polynomial model in the case of a single auxiliary variable. Since the form of the estimator $\widehat{Y}_{\mathrm{SEM}}$ in this paper has some similarities with it, the difference is that this paper is estimated by adjusting the model according to the product of semi-parameters. Therefore, the properties of $\widehat{Y}_{\mathrm{SEM}}$ can be studied with reference to their method and a series of asymptotic properties of semi-parametric models. In order to study the properties of this estimator, the following assumptions need to be made:

(a) For all $k \in U$, the model error term ε_k is independent of each other, the mean is 0, and the variance is $\sigma^2 < \infty$;

(b) In the super-population regression model ξ, we assume that x_k is fixed, and its independent and identical distribution function is $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$, where $f(\cdot)$ is the density function on the interval $[a_x, b_x]$, and for all $x \in [a_x, b_x]$, there is $f(x) > 0$;

(c) The function $m(\cdot)$ is continuous and it is continuously differentiable in order $p+1$, where p is the series of the local polynomial function;

(d) The kernel function $K(\cdot)$ satisfies $\int_{-1}^{1} K(u)\mathrm{d}u = 1$;

(e) When $N \longrightarrow \infty$, there is $nN^{-1} \longrightarrow \pi \in (0,1)$, and when $h \longrightarrow 0$ there is $(Nh^2/(\log\ \log N))/ \longrightarrow \infty$;

(f) For all N, $\min_{k\in U}\pi_k \geq \lambda > 0$, $\min_{k,l\in U}\pi_{kl} \geq \lambda^* > 0$, and $\lim_{N\longrightarrow\infty} \sup n \max_{k,l\in U:\ k\neq l}|\pi_{kl} - \pi_k\pi_l| < \infty$.

Based on the above assumptions, this paper presents the following two statistical properties of the semi-parametric model auxiliary estimator $\widehat{Y}_{\mathrm{SEM}}$ without proof:

This property 1 is: the estimator $\widehat{Y}_{\mathrm{SEM}}$ is an asymptotically designed unbiased and design consistent estimator of the population total Y; that is,

$$\lim_{N\longrightarrow\infty} E_p\left[\frac{\widehat{Y}_{\mathrm{SEM}} - Y}{N}\right] = 0 \,\text{with}\, \xi - \text{probability}1, \quad (18)$$

$$\lim_{N\longrightarrow\infty} E_p\left[I_{\left\{\left|\widehat{Y}_{\mathrm{SEM}} - Y\right| > \eta\right\}}\right] = 0, \text{with}\, \xi - \text{probability}1, \text{and}\, \eta > 0. \quad (19)$$

This property 2 is: the asymptotic mean squared error (AMSE) of the estimator $\widehat{Y}_{\mathrm{SEM}}$ is equal to the approximate variance of the estimator:

$$\lim_{N\longrightarrow\infty} E_p\left(\frac{\widehat{Y}_{\mathrm{SEM}} - Y}{N}\right)^2 = AV\left(\widehat{Y}_{\mathrm{SEM}}\right), \text{with}\, \xi - \text{probability}1. \quad (20)$$

In order to compare and analyze the effectiveness of the semi-parametric model-assisted regression estimator $\widehat{Y}_{\mathrm{SEM}}$ proposed in this paper, the simulation method is adopted in this section. It calculates the estimation accuracy of the semi-parametric estimation method studied in this paper (hereinafter referred to as SEM estimation) and the current mainstream generalized regression estimation method (hereinafter referred to as GREG estimation) in different situations and makes a comparative analysis. We assume that the series p of the local polynomial estimation of the non-parametric part is 1, and the kernel function $K(\cdot)$ is a Gaussian kernel. In this paper, simple random sampling without replacement is used to extract 100 units from a finite population of $N = 1000$ to form a sample s. Using this sample, the SEM estimator and GREG estimator in this paper are obtained. The estimation effect of the estimator in different situations is reflected by the relative design deviation (RB for short) and the relative mean square error (RMSE for short):

$$\mathrm{RB} = \left(\frac{[E(\widehat{Y}) - Y]}{Y}\right) \times 100\%, \quad (21)$$

$$\mathrm{RMSE} = \frac{\sqrt{\mathrm{MSE}\left(\widehat{Y}_{\mathrm{GREG}}\right)}}{\sqrt{\mathrm{MSE}\left(\widehat{Y}_{\mathrm{SEM}}\right)}}, \quad (22)$$

here, $E(\widehat{Y})$ represents the mean corresponding to the sampling estimator, Y is the total value of the population, $\mathrm{MSE}(\widehat{Y}_{\mathrm{GREG}})$ and $\mathrm{MSE}(\widehat{Y}_{\mathrm{SEM}})$ are the mean square errors of the two estimators, respectively.

This section considers the following eight different forms of regression functions:

A linear function is $m_1(x) = 1 + (x - 0.5)$.

The quadratic function is $m_2(x) = 1 + 2(x - 0.5)^2$.

The bump function is $m_3(x) = 1 + 2(x - 0.5) + \exp\left\{-200(x - 0.5)^2\right\}$.

The piecewise function is $m_4(x) = \{1 + 2(x - 0.5)\} I_{\{x\leq 0.65\}} + 0.65 I_{\{x>0.65\}}$.

The exponential function is $m_5(x) = \exp(-8x)$.

The density function is $m_6(x) = \Phi\{(1.5 - 2x)/\sigma\}$, Φ is the cumulative distribution function of the standard normal distribution.

The trigonometric function 1 is $m_7(x) = 2 + \sin(2\pi x)$.

The trigonometric function 2 is $m_8(x) = 2 + \sin(8\pi x)$.

Among them, $x \in [0, 1]$, the auxiliary variable x_k takes a uniform value in the interval [0, 1], and the error term is $\varepsilon_k \sim ii\ d.N(0, \sigma^2)$.

The simulated population is constructed according to $y_{ik} = m_i(x_k) + \varepsilon_k$, where $i = 1, \ldots, 8, k = 1, \ldots, N$, and $\sigma = 0.1$ are set, and 1000 repeated samplings are carried out under each population.

For the linear function m_1, the GREG estimator is the most perfect estimator, and its estimation accuracy is expected to be the highest. Because the assumption of the model is completely correct, that is, the assumption of the model fully conforms to the premise of the traditional generalized regression estimation. The most important thing is the comparison of the accuracy of semi-parametric model-assisted sampling estimation and traditional generalized regression estimation in the case of other functions, that is, in various cases where the function is not a linear function. For the function m_2, its overall trend is quadratic, and it is inaccurate to assume that it is a linear function for the entire domain of x_k, and it needs to be adjusted appropriately. Similarly, for functions m_3 to m_8, it is inaccurate to directly assume that they are linear functions in the entire domain, which will seriously affect their estimation accuracy. Further adjustments based on the linearity assumption will improve its estimation accuracy.

This work initially utilises the *R* programming language to do typical generalized regression estimation and SEM estimation under the confined population created by the preceding eight algorithms. When the parameter component is a linear function, it fits the picture, and a basic analysis of the fitting effect of the two is performed. The scatter plots in Figures 1–8 illustrate the fitted data distribution, and the green line indicates the generalized regression estimate fitting graph to the original function. The red line represents the fitting graph of the SEM estimation to the original function when the parameter part proposed in this paper is a linear function. The specific estimation results of the two estimation methods will be shown in Figures 1 and 2 in detail.

It can be seen from Figure 1 that, under the finite population constructed by the linear function, the fitting graph of the traditional generalized regression estimation completely coincides with the original function. Moreover, the graph corresponding to the SEM estimation in this paper deviates from the original function. However, this paper has verified that the bias decreases slightly as the sample size increases. At this point, it can be seen that the fitting effect of the traditional generalized regression estimation is better than the SEM estimation in this paper. Because the finite population constructed by the linear function fully satisfies the premise of GREG estimation, that is, there is a strict

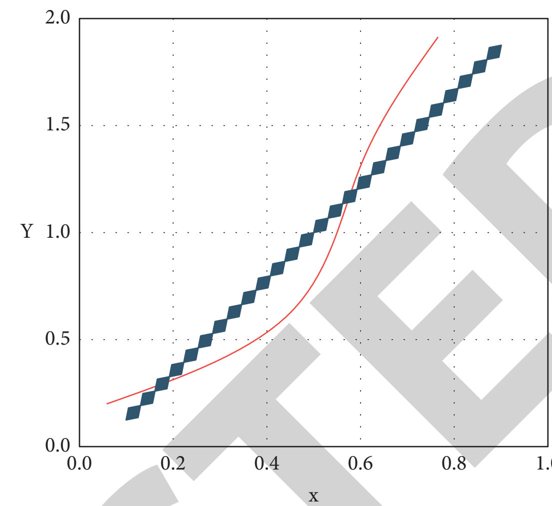

Figure 1: Linear population.

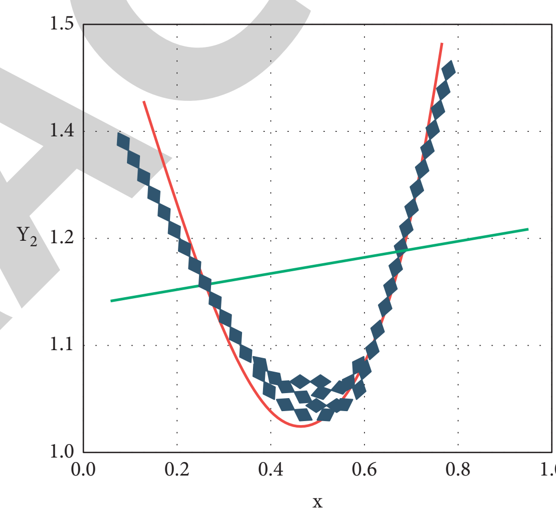

Figure 2: Overall quadratic function.

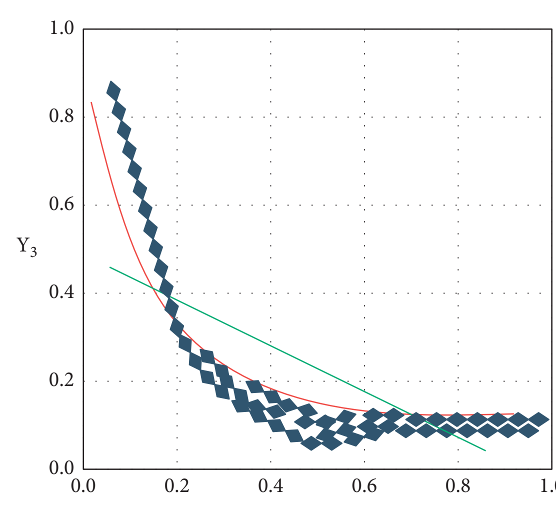

Figure 3: Exponential function total.

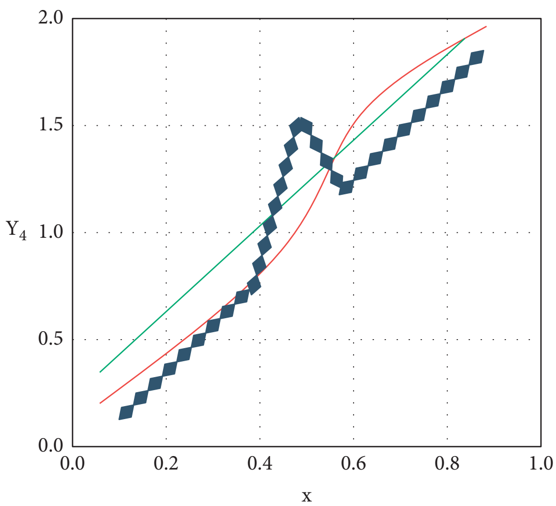

Figure 4: Overall bump function.

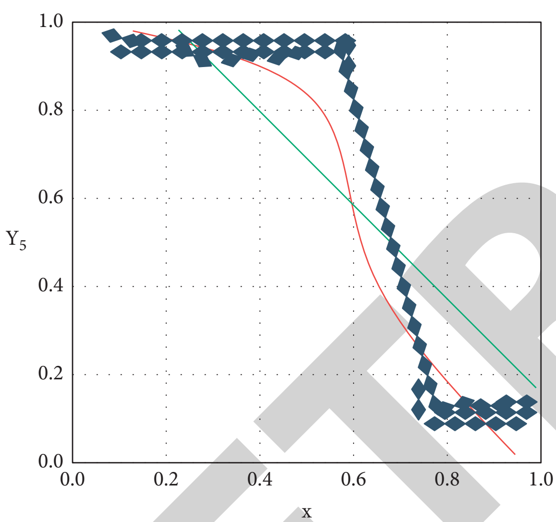

Figure 5: Overall density function.

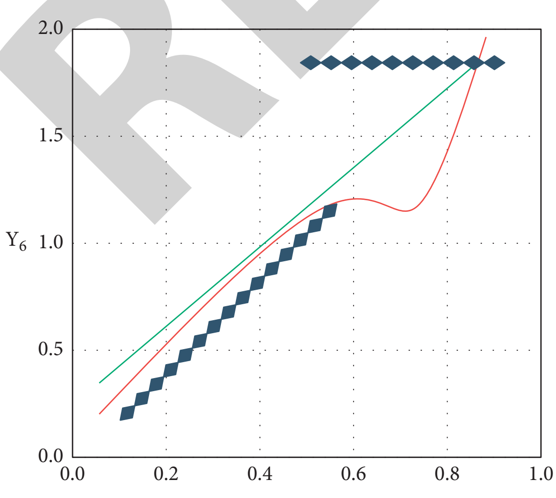

Figure 6: Overall piecewise function.

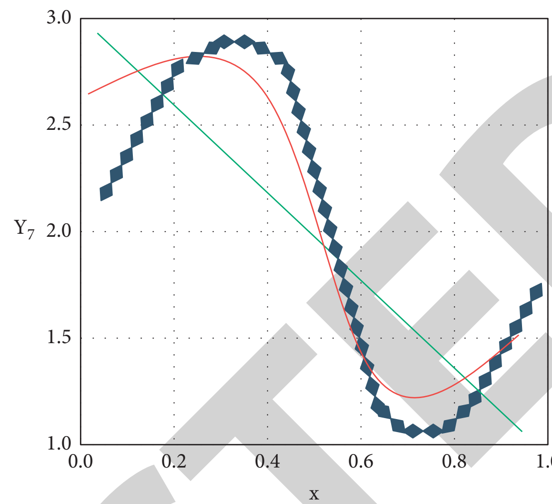

Figure 7: The population of trigonometric function 1.

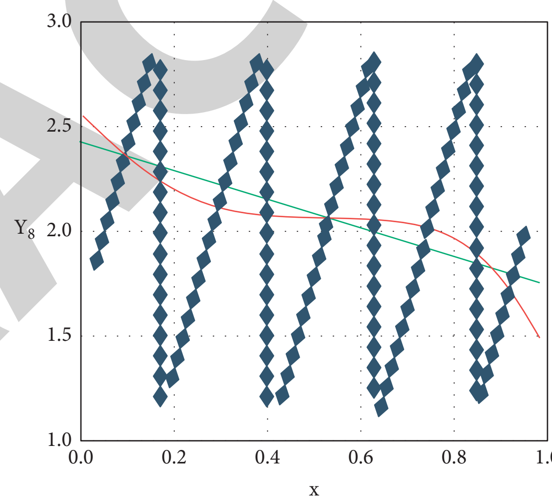

Figure 8: The population of trigonometric function 2.

linear relationship between auxiliary variables and research variables. Therefore, the generalized regression estimation method is the best estimation method.

It can be seen from Figure 2 that, under the nonlinear finite population composed of quadratic function functions, the overall trend of the fitting graph of the SEM estimation proposed in this paper when the parameter part is a linear function is similar to that of the original function. Moreover, this paper has verified that when the bandwidth and kernel function of SEM estimation are properly selected, the fitting effect of SEM estimation will be further improved with the increase of samples, and the specific situation will not be shown in detail. However, under the finite population, the fitted graph generated by GREG estimation is a horizontal straight line, and there is a large error between the original graph, which will generate a large variance and reduce the sampling estimation accuracy. From the graph, under the finite population composed of quadratic functions, the fitting effect of the SEM estimation proposed in this paper is better than that of the generalized regression estimation.

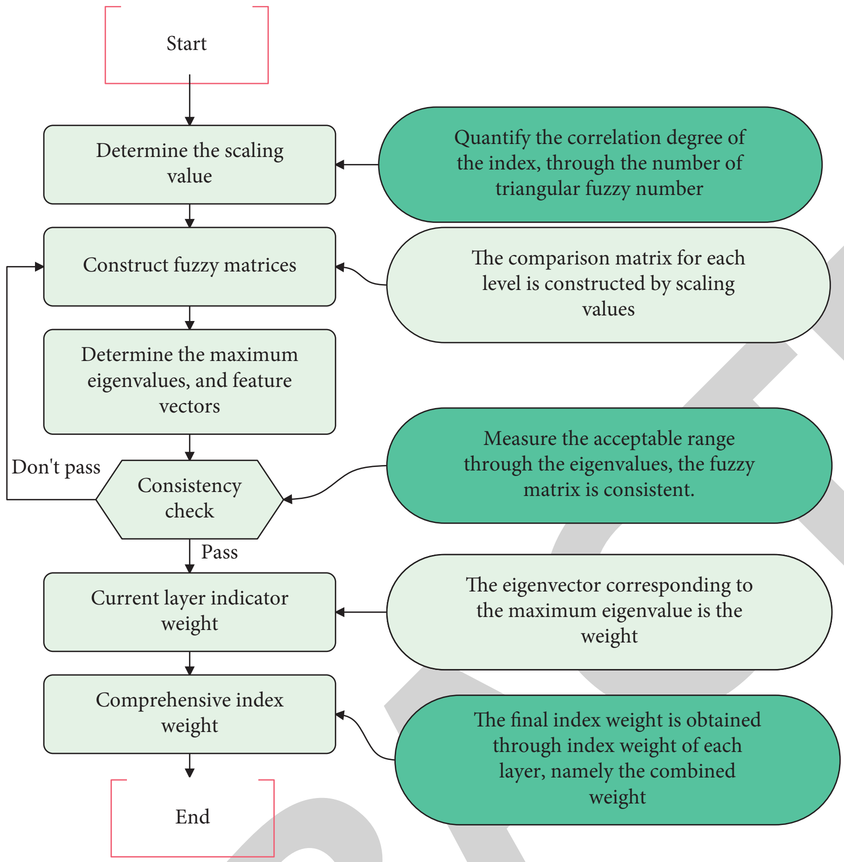

Figure 9: Flowchart of fuzzy analytic hierarchy process.

As can be seen from Figure 3, under the finite population composed of exponential functions, the fitting graph generated by traditional generalized regression estimation is a downward sloping straight line, which has a certain degree of similarity with the general trend of the original exponential function in some intervals. However, there is a large deviation between the two. In contrast, the overall trend of the fitting graph of the semi-parametric product estimation proposed in this paper is more similar to the original function than the generalized regression estimation, and this paper verifies that the fitting effect of the SEM estimation will become better with the increase of the sample size. As shown in Figures 4 and 5, under the population constructed by the bump function and the density function, the fitting effect of the SEM estimation and the GREG estimation is roughly similar to that of the finite population constituted by the exponential function shown in Figure 3. This article will not analyze them one by one. It can be seen from Figure 6 that under the finite population constructed by the piecewise function, there is no significant difference between the fitting effects of the SEM estimation and the generalized regression estimation. However, with the increase of the sample, the fitting effect of the semi-parametric model-assisted sampling estimation proposed in this paper will gradually improve.

It can be seen from Figure 7 that, under the finite population constructed by trigonometric function 1, there is a large deviation between the fitting results of the traditional generalized regression estimation and the original function. However, when the parameter part of the SEM estimation proposed in this paper is a linear function, the fitting effect of this estimation method is significantly better than that of the GREG estimation. There is a deviation, similar to the general trend of the original function, but, as the sample size becomes larger, the deviation will steadily decrease. However, the fitting results of the GREG estimate and the SEM estimation suggested in this research are considerably different from the original function under the finite population created by trigonometric function 2 given in Figure 8. With the increase of the sample size, the fitting effect of the SEM estimation will improve, but there is still a certain degree of deviation from the original function. The main reason is that the parameter part of the SEM estimation is assumed to be a linear function. The deviation from the original function is large, and the SEM estimation will have a better estimation effect when the parameter setting is appropriate.

4. Analysis on the Development Strategy of Private Education Based on Data Mining Algorithm

The full name of AHP is Fuzzy Analytic Hierarchy Process. It mainly introduces fuzzy mathematical theory on the basis of Analytic Hierarchy Process. We construct a fuzzy judgment matrix and then obtain the weight of each index to measure the quality of online courses through the fuzzy judgment matrix. Its modeling flow chart is shown in Figure 9.

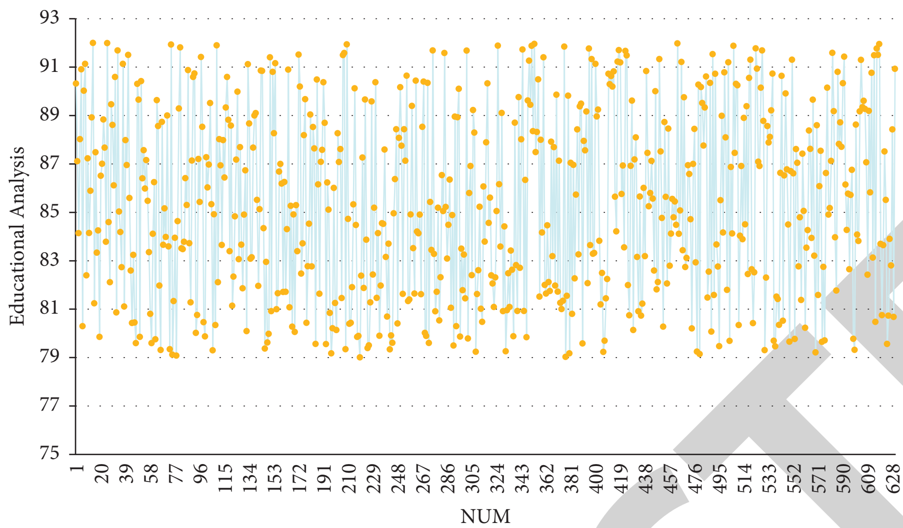

Figure 10: Statistical diagram of the effect evaluation of the analysis method for the development of private education.

Through data mining and the process shown in Figure 9, the development of private education is analyzed, and multiple sets of data are counted for evaluation. The statistical results are shown in Figure 10.

From the above research, we can see that the analysis method of private education development based on data mining algorithm proposed in this paper has certain effects. On this basis, this paper analyzes the development strategy of private education and puts forward the following suggestions.

Private education development policies need to promote the development of private schools, thereby driving the development of the compulsory education market. The development policy of private education needs to protect the interests of the legal entities of private schools, protect the interests of private teachers, and safeguard the interests of private students from a legal perspective. At the same time, it is necessary to provide financial subsidies and rewards to private schools, strengthen the supervision of private schools, and improve the quality and teaching level of private schools. Moreover, the development of private schools needs to supplement the supply of degrees for the compulsory education market, effectively alleviate the shortage of degrees in public schools, enrich the school-running model, and promote competition in the education market and optimization of resource allocation. In particular, the supply of private school places needs to play an important role in solving the problem of enrolling children who have moved with them, so as to meet the enrolment needs of school-age children in different places.

5. Conclusion

The development of private education is conducive to increasing investment in education, expanding the scale of education, and improving the quality of education. It can also provide more choices for educated people and conditionally meet social needs. At present, a number of good models have emerged in the private schools, and they have a certain reputation in the society. Moreover, private education must adapt to the large education market in China and adhere to a multi-channel, multi-type, and multi-level school-running model to meet the needs of educated people to choose freely. At the same time, it needs to form an operating system that seeks survival by service and quality and seeks development by characteristics and benefits. This study examines the present state of private education using a data mining technique, investigates the issues that have arisen in the development of private education, and evaluates the private education development plan. The findings suggest that the data mining algorithm-based analytical technique for private education development described in this work has certain impacts.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work was funded by the Jilin Province Education Science "13th Five-Year Plan" Key Project: "Promoting the development of private education policy support and supervision system research," Project approval number: ZD18069.

References

[1] J. Keating and I. Nourbakhsh, "Teaching artificial intelligence and humanity," *Communications of the ACM*, vol. 61, no. 2, pp. 29–32, 2018.

[2] J. R. Mianroodi, N. H. Siboni, and D. Raabe, "Teaching solid mechanics to artificial intelligence—a fast solver for heterogeneous materials," *Npj Computational Materials*, vol. 7, no. 1, pp. 1–10, 2021.
[3] X. Li, "The construction of intelligent English teaching model based on artificial intelligence," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 12, no. 12, pp. 35–44, 2017.
[4] S. Zou, "Designing and practice of a college English teaching platform based on artificial intelligence," *Journal of Computational and Theoretical Nanoscience*, vol. 14, no. 1, pp. 104–108, 2017.
[5] F. Kong, "Application of artificial intelligence in modern art teaching," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 13, pp. 238–251, 2020.
[6] M. Pantic, R. Zwitserloot, and R. J. Grootjans, "Teaching introductory artificial intelligence using a simple agent framework," *IEEE Transactions on Education*, vol. 48, no. 3, pp. 382–390, 2005.
[7] C. Yang, S. Huan, and Y. Yang, "A practical teaching mode for colleges supported by artificial intelligence," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 17, pp. 195–206, 2020.
[8] K. Kim and Y. Park, "A development and application of the teaching and learning model of artificial intelligence education for elementary students," *Journal of The Korean Association of Information Education*, vol. 21, no. 1, pp. 139–149, 2017.
[9] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education–where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 1–27, 2019.
[10] S. C. H. Yang, W. K. Vong, R. B. Sojitra, T. Folke, and P. Shafto, "Mitigating belief projection in explainable artificial intelligence via Bayesian teaching," *Scientific Reports*, vol. 11, no. 1, pp. 1–35, 2021.
[11] P. Ribeiro, H. Simões, and M. Ferreira, "Teaching artificial intelligence and logic programming in a competitive environment," *Informatics in Education*, vol. 8, no. 1, pp. 85–100, 2009.
[12] V. Rampton, M. Mittelman, and J. Goldhahn, "Implications of artificial intelligence for medical education," *The Lancet Digital Health*, vol. 2, no. 3, pp. e111–e112, 2020.
[13] Q. Zhou, "Reforms in teaching the course of introduction to artificial intelligence," *Solid State Technology*, vol. 64, no. 2, pp. 6331–6335, 2021.
[14] K. I. M. Haejoong and L. E. E. Sangmin, "Do less teaching, do more coaching: toward critical thinking for ethical applications of artificial intelligence," *Journal of Learning and Teaching in Digital Age*, vol. 6, no. 2, pp. 97–100, 2021.
[15] P. H. Lin, A. Wooders, J. T. Y. Wang, and W. M. Yuan, "Artificial intelligence, the missing piece of online education?" *IEEE Engineering Management Review*, vol. 46, no. 3, pp. 25–28, 2018.
[16] C. Guan, J. Mou, and Z. Jiang, "Artificial intelligence innovation in education: a twenty-year data-driven historical analysis," *International Journal of Innovation Studies*, vol. 4, no. 4, pp. 134–147, 2020.
[17] D. Ni, "The application of artificial intelligence technology in international Chinese language teaching," *Chinese Journal of Social Science and Management*, vol. 5, no. 1, pp. 101–112, 2021.
[18] R. Dilmurod and A. Fazliddin, "Prospects for the introduction of artificial intelligence technologies in higher education," *Academicia: An International Multidisciplinary Research Journal*, vol. 11, no. 2, pp. 929–934, 2021.