Research on Dynamic Fitting of Internet Enterprise Financial Statement Analysis Using BP Neural Network in the Era of Big Data Interpretation from the Perspective of Profitability Model

Xiuwen Bao (✉)
Dongfang College, Shandong University of Finance and Economics, Jinan 271000, China
hqqquet@guet.edu.cn

Abstract. The year 2021 is not only a turning year for the development of China’s internet industry, but also an eventful year. Under the situation of strong government supervision, the micro-ecology of China’s internet industry is undergoing subtle changes. Internet enterprises are completely different from manufacturing in terms of profit model and industry ecology. Although the financial statements composition of internet enterprises are simple, the risk factors are more complex and secret, the matching principle is more difficult to be compatible, and the future trend is more difficult to grasp. With the aid of flexible financial management methods and contingency theory research, flexible dynamic indicators are increased. Using the improved BP neural network to study the fitting relationship between financial income, financial cost and non-financial indicators, to the strong correlation analysis with non-financial indicators such as UV, conversion rate, per capita access depth, advertisement click rate, DC UV, etc., let the Internet The analysis of corporate financial report data is more realistic and reasonable.

Keywords: internet enterprises · non-financial indicators · dynamic fitting research · BP neural network

1 Introduction

“Big data era” not only provides new resources for enterprises, but also brings great challenges for the analysis of financial statements of internet enterprises. The business logic of “wool comes from pigs” endowed by the internet industry is completely different from that of traditional manufacturing industry. There is also a world of difference between the interpretation of financial statements and the traditional financial statement system formed in the industrial era of big machines and big equipment. The traditional financial statement framework elements, whether the balance sheet or the profit, revenue, costs and expenses in the income statement, are the products of the industrial economy era. Updating old ideas, mining new resources, better interpreting financial statements of internet enterprises, and then understanding the truth behind the “tech giants” are topics that should be studied in the era of big data. With the help of the dynamic fitting research of financial indicators and non-financial indicators, this paper puts forward the

© The Author(s) 2023
https://doi.org/10.2991/978-94-6463-034-3_61
idea of analyzing the current financial statements of internet enterprises. Based on the analysis of traditional financial indicators, combined with the characteristics of internet enterprises and big data resources, this paper optimizes the model of traditional financial statement analysis, so as to help financial personnel change from causal thinking to relevant thinking in practical work, jump out of the traditional analysis mode through complex phenomena, make accurate judgment and analysis and discover the truth behind it.

1.1 Review of Foreign Research

Barua believed that adding non-financial indicators into the performance evaluation system could reflect management behaviors more quickly and directly than using financial indicators alone [1]. Christopher research showed that using different non-financial indicators in different environments would produce different performance [2]. Eli Amir found a significant correlation between non-financial indicators and several variables such as whether an enterprise had an innovative and enterprising strategy, the degree of financial distress, the length of development cycle, quality-oriented strategic products and industry rules [3]. In Robert Hall’s “four scales” theory, four financial indicators and non-financial indicators are fitted and analyzed from the value scale (including external quality, internal quality and quality improvement procedures), operation time scale (the time period for changing raw materials into finished products), resource use scale (to measure the consumption and related costs of specific resources), and human resources scale. Rober tkaplan, Nolan Norton Institute and David Norton were engaged in a performance evaluation system, namely “future organizational performance measurement method”. Jin B pointed out that the evaluation indicators reflect competition results and analysis indicators reflect the relevant factors determining the competitiveness, both of which are indicators of enterprises competitiveness [4]. Through foreign research review, it is found that enterprise financial statement analysis is analyzed by financial and non-financial indicators. Foreign academic circles are involved in the fitting of financial indicators and non-financial indicators, but they focus on the purpose of comprehensively measuring enterprise performance, but non-financial indicators are difficult to measure. Baetge, J et al. used artificial neural network algorithm to identify threatened companies by considering both financial and non-financial indicators, and provided the idea of using artificial neural network algorithm to analyze financial indicators and non-financial indicators at the same time [5].

1.2 Review of Domestic Research

Wang Kunqi thought that current listed companies had three statements, notes to accounting statements and related schedules, which reflect the main financial information of the enterprise from different angles to statement users [6]. Liu Li believed that financial indicators were needed to maximize shareholder value, and typical financial indicators included return on equity, operating profit, cash flow, etc. [7]. Guo Fu believed that strategic objectives affected the business operation system of enterprises, resulting in indicators such as customer satisfaction and flexibility [8]. Liang Zhanfan found that the traditional financial framework reflected the backward outcome factors, and it was
impossible to evaluate the leading drivers [9]. Wang Dan external conditions and self factors affected the profit model of internet enterprises [10]. Zhang Kuidong believed that non-financial indicators were generally closely related to the enterprise strategy and related to the long-term development of the enterprise. If the non-financial information was included in the financial analysis system, it could better reflect the overall goal of the enterprise [11]. Huang Libo found that quality information disclosure enabled users of financial statements to better conduct a comprehensive and systematic analysis of enterprise operation, and then made a reasonable judgment on the development prospect of enterprises [12]. Li Youbo concluded the influence of policy-based dividend on the profitability of cross-border e-commerce [13]. Through domestic research, it is found that traditional financial indicators are the mainstream of financial analysis, but they are innovative with non-financial indicators, such as click through rate, website traffic, number of visitors, number of user registrations, registration conversion rate, etc. These non-financial indicators are of little significance to traditional enterprises, but they are important indicators to measure performance growth ability and enterprise operation ability for internet enterprises, thus affecting the correct judgment of financial users.

2 Three Challenges of Financial Statement Analysis of Internet Enterprises

2.1 The Traditional Financial Analysis Framework Shackles the Financial Statement Framework of Internet Enterprises

The business logic of “wool comes from pigs” endowed by the internet industry is completely different from that of traditional manufacturing industry. There is also a world of difference between the interpretation of financial statements and the traditional financial statement system formed in the industrial era of big machines and big equipment. The traditional financial statement framework elements, whether the balance sheet or the profit, revenue, costs and expenses in the income statement, are the products of the industrial economy era. Its main business may not generate revenue in the financial statement, and most of its revenue comes from value-added services. Under this mode, its revenue and cost are difficult to follow the principle of matching, such as Tencent, American Alphabet, Facebook, etc. There is a serious deviation between the content of financial statements and the essence of transactions. The methods and framework of information interpretation are limited. In the era of big data, it is difficult to better mine data resources to disclose information, which poses a more targeted challenge to the conception and design of financial statement framework and content.

2.2 Transformation of Financial Analysis from Causality to Correlation

While the traditional transaction mode is a B2C seller and buyer transaction, Pinduoduo uses the underlying business logic to combine bargaining, grouping and aggregation, associate buyers without causality, and combine entertainment and consumption to build a new transaction mode. The transformation of enterprise financial analysis from causality to correlation makes the business model of internet enterprises be innovated in a subversive way in practice, which makes it difficult to use the relevant theories of the nature of internet business for reference.
2.3 Change from Capital Driven to Human Driven

The traditional financial indicators and analysis framework are based on the traditional capital driven manufacturing industry. The main driving factors of current Internet enterprises are human capital, business model innovation and technological progress. The indicator analysis of traditional financial statement system established with capital as the core loses its original analytical significance, such as inventory turnover rate and asset return rate.

3 Dynamic Fitting Process of Financial Indicators and Non-financial Indicators Based on BP Neural Network

In the era of big data, BP neural network is used to realize the dynamic fitting process of Internet financial indicators and non-financial indicators, and a data-driven method is used to establish a fitting model. The process of using BP neural network to build a fitting model is divided into: preprocessing sample data; constructing BP neural network structure; training BP neural network.

The first step is to preprocess the data samples, take financial indicators and non-financial indicators as input values, and take the value to be predicted as the output value. If there is text data in it, it is encoded and converted to a numeric type, and a set of data sets \( \{x_1^i, x_2^i, \ldots, x_m^i\} \) are constructed, where \( m \) represents a total of \( m \) indicators, and \( i \) represents the i-th group. Data, and normalize the dataset.

\[
x_i^\text{normal} = \frac{x_i - x_i^\text{mean}}{x_i^\text{std}}, \quad i = 1, 2, \ldots
\]

where \( x_i^\text{mean} \) is the mean of the i-th group of data \( x_i \), and \( x_i^\text{std} \) is the standard deviation of the i-th group of data \( x_i \).

The second step is to construct the BP neural network structure, construct the input layer of the BP neural network with \( m \) financial indicators and non-financial indicators, and construct the output layer with the predicted value such as platform service fee as the output value \( y \). The activation function of the hidden layer is set to ‘relu’. As Fig. 1 shown:

The third step, before training the BP neural network, in order to prevent overfitting, use L2 regularization, and configure the loss function as:

\[
L = \sum_{i=1}^{n} (y_i^\text{label} - y_i^\text{pre})^2 + \gamma \sum_{i=1}^{n} \omega_i^2
\]

Among them, \( y_i^\text{label} \) is the label value in the i-th group of data, that is, the platform service fee, etc. \( y_i^\text{pre} \) is the predicted value of the platform service fee in the i-th group of data. \( \gamma \) is an adjustable parameter, and \( \omega_i \) represents the weight in the neural network.

During training, use Adam to optimize the gradient descent algorithm:

\[
V_{dw} = \beta_1 \cdot V_{dw} + (1 - \beta_1) \cdot dw
\]
\[ V_{db} = \beta_1 \cdot V_{db} + (1 - \beta_1) \cdot db \]  
\[ S_{dw} = \beta_2 \cdot V_{dw} + (1 - \beta_2) \cdot dw^2 \]  
\[ S_{db} = \beta_2 \cdot V_{db} + (1 - \beta_2) \cdot db^2 \]  
\[ V_{dw}^c = V_{dw}^c / (1 - \beta_1) \]  
\[ V_{db}^c = V_{db}^c / (1 - \beta_1) \]  
\[ S_{dw}^c = S_{dw}^c / (1 - \beta_2) \]  
\[ S_{db}^c = S_{db}^c / (1 - \beta_2) \]  
\[ w := w - \alpha \cdot V_{dw}^c / (\sqrt{S_{dw}^c} + \varepsilon) \]  
\[ b := b - \alpha \cdot V_{db}^c / (\sqrt{S_{db}^c} + \varepsilon) \]

At the beginning of the gradient descent iteration, \( V_{dw} = V_{db} = S_{dw} = S_{db} = 0 \), where \( w \) is the weight in the neural network and \( b \) is the bias in the neural network. \( \alpha, \beta_1, \beta_2, \varepsilon \) are hyperparameters, \( \alpha \) is the learning rate, choose \( \beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8} \).

4 Dynamic Fitting Analysis of Financial Indicators and Non-financial Indicators of Internet Enterprises

In the era of big data, big data makes the non-financial indicators of internet enterprises possible for financial statement analysis, which is becoming increasingly important. There are many tools to obtain external data, among which Google Analytics can help.
enterprises better obtain the industry reports of professional research institutions. Non-financial indicators have the characteristics of strong timeliness, rich data content and exponential growth. At this point, it is very necessary for us to extract the core non-financial indicators with financial indicators, so as to better discover the truth of the financial statements behind the “tech giants”.

4.1 Profitability Analysis of Internet Enterprises and Dynamic Fitting Factor Analysis of Financial Indicators and Non-financial Indicators

With the help of Adrian Slywotzky’s theory of profit model, the key indicator of internet enterprise ---- traffic non-financial indicator is introduced into financial analysis. Adrian Slywotzky (2002) believed that the profit model had four basic elements: profit point - platform service matrix, profit object - customer positioning and users, profit leverage - business activities, profit barrier - search technology & portal resources and profit sources - online advertising & online games (Zhen Guohong, 2007). The factors affecting the profitability of internet enterprises are shown in the Fig. 2.

4.2 Fitting Model Between Revenue and Non-financial Indicators of Networked Enterprises

In financial analysis of profitability, profit equals revenue minus cost; the key factors to generate revenue in internet enterprises are profit points, profit sources and profit leverage. The profit point corresponds to the platform service fee revenue; profit source brings platform advertising fee and game fee revenue, and profit leverage creates B2C sales revenue (Table 1).

4.2.1 Dynamic Fitting Model of Financial Indicator and Non-financial Indicator of Profit Point (Platform Service Fee)

The main factors affecting the platform service fee include sales commission and product function fee, where the size of sales commission is determined by unique visitor (UV), conversion rate and customer unit price. The value that internet companies discuss everything is traffic. The cornerstone of website traffic is Unique Visitors (UV), Page View (PV) and Visits; traffic pays attention to quality. The main indicators to measure its quality are bounce rate, online duration and access depth. It can be seen that
Table 1. Internet financial indicators

<table>
<thead>
<tr>
<th>Indicator classification</th>
<th>Indicator name</th>
</tr>
</thead>
<tbody>
<tr>
<td>solvency</td>
<td>current ratio</td>
</tr>
<tr>
<td></td>
<td>asset-liability ratio</td>
</tr>
<tr>
<td></td>
<td>amount of long-term liabilities</td>
</tr>
<tr>
<td>profitability</td>
<td>ROE</td>
</tr>
<tr>
<td></td>
<td>operating profit ratio</td>
</tr>
<tr>
<td>development ability</td>
<td>sales growth rate</td>
</tr>
<tr>
<td></td>
<td>net profit growth rate</td>
</tr>
<tr>
<td>cash flow</td>
<td>cash maturity debt ratio</td>
</tr>
</tbody>
</table>

Fig. 3. BP network of platform service fee.

non-financial indicators such as UV, PV and Visits are also key factors affecting revenue. Assuming that the impact indicators of internet enterprises are relatively stable and develop to a certain scale, the fitting model between revenue and non-financial indicators is as follows:

As shown in the Fig. 3, construct a set of input variables including sales commission, sales, commission ratio, product function fee, customer, user click, bounce rate, online duration and access depth, with platform service fee as the predicted value, the number of input layer nodes is 9, The number of hidden layer nodes is 12, and the number of output layer nodes is 1 BP neural network structure. Financial and non-financial indicators and platform service fee prediction models are established through the training of neural networks.

Internet companies have two profit objects: internet users (providers of potential income) and enterprise customers (providers of actual interests). Internet users will attract more enterprise customers. Internet companies attract more network users and increase the click through rate of the network through personalized design of various businesses, so as to improve the ranking of the company. The purpose is to get more corporate customers to invest in joining or become advertisers.
The calculation formula of customer unit price is: customer unit price = total sales (deducting discount) ÷ total number of customers, or customer unit price = total sales ÷ total number of transactions (Five people buying 10 transactions counts as five), that is, the average unit price of goods multiple the average number of goods purchased by each customer.

In Internet enterprises, the general formula is as follows: customer unit price = mobile line length × retention rate × attention rate × purchase rate × number of purchases × unit price of goods.

The conversion rate is the core of whether the website can finally make profits. Improving the conversion rate of the website is the result of the comprehensive operation strength of the website. Conversion rate refers to the ratio of the number of completed conversion actions to the total number of clicks of promotion information in a statistical cycle. The calculation formula is: conversion rate = (conversion times / clicks) × 100%.

For example, 10 users see a search promotion result, 5 of them click a promotion result and are jumped to the target URL, and 2 of them have the behavior of subsequent transformation. Then, the conversion rate of this generalization result = (2 / 5) × 100% = 40%.

4.2.2 Dynamic Fitting Model of Financial Index and Non-financial Index of Profit Source (Platform Advertising Fee)

According to the financial indicator analysis model: profit = sales × net profit margin. Platform advertising revenue and non-financial indicators are fitted and analyzed. The model is as follows: profit = sales × net profit margin = (number of buyers × customer unit price) × net profit margin = number of people entering the store × purchase conversion rate × customer unit price × net profit margin = advertising display × advertising conversion rate × purchase conversion rate × customer unit price × net profit margin = promotion display × promotion conversion rate × purchase conversion rate × customer unit price × net profit margin = search presentation × search conversion rate × purchase conversion rate × customer unit price × net profit margin.

There are many charging methods for platform advertising fees, which are also divided into three categories: by display, by action and by sales. Therefore, CPC (cost per click), CPM (cost per mille), CPA (cost per action), CPS (cost per sale), CPD (cost per download), CPT (cost per time) are important factors affecting the advertising revenue of the platform. The most commonly used billing methods in the Internet advertising market are CPM and CPC. These two billing methods are also the most favored billing methods of advertisers. They can not only take the lead in product sales, but also expose their own brands to a greater extent and shape their brand image. In recent years, new billing methods, such as OCPC and OCPM, have also emerged.

The first category: billing by display.

CPM (cost per mille): CPTM: cost per thousand impressions of targeted users (e.g. positioning based on demographic information). The difference between the two is that CPM is the impression number of all users, while CPTM is only the impression number of targeted users.

The second category: billing by action.
PPC: it is an online advertising pricing model, which pays according to the number of users who click on advertising or e-mail information.

CPL: billing is based on successful registration. When some APPs don’t know the effect, it’s the most secure billing method.

CPC (cost per click): billing is based on the number of times the advertisement is clicked.

CPA: its pricing model is based on the actions taken by each visitor to online advertising.

CPD: charge according to download.

CPR: charge according to the number of user responses.

PPL: its pricing model is based on the guided payment generated by online advertising each time.

The third category: billing by sales, which is rarely used.

CPO: charge according to each order / transaction.

CPS: billing is based on marketing effect, which generally refers to sales.

As shown in the Fig. 4, construct a set of input variables including CPC (cost per click), CPM (cost per thousand impressions), CPA (cost per action), CPS (pay per sale), CPD (pay per download) and CPT (pay per impression), takes the platform advertising fee as the predicted value, the number of nodes in the input layer is 6, the number of nodes in the hidden layer is 10, and the number of nodes in the output layer is 1 BP neural network structure. Establishment of financial indicators, non-financial indicators and platform service fee prediction models through neural network training (Table 2).

4.2.3 Dynamic Fitting Model Between Profit Leverage (B2C Sales Revenue) and Non-financial Indicators

The main profit model of B2C is the sales revenue of goods and services on its website. Enterprises sell products on B2C websites, independently develop or operate distributors, provide warehousing and logistics transportation, or attract third-party logistics. In this way, the best way to increase revenue, increase clicks and impress customers is preferential discount. The dynamic fitting model between B2C sales revenue and non-financial indicators is as follows:

As shown in the Fig. 5, construct a set of input variables with customers, conversions, clicks, 100%*moving line length, stay rate, attention rate, purchase rate, number of purchases and unit price of goods as a group, with B2C sales revenue as the predicted value, input layer nodes. The number of nodes is 9, the number of nodes in the hidden
Table 2. Interpretation of relevant indicators of flow quantity

<table>
<thead>
<tr>
<th>Indicator classification</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV</td>
<td>UV is the abbreviation of unique visitor, and the number of independent visitors to the website.</td>
</tr>
<tr>
<td>PV</td>
<td>Page view/number of web pages visited</td>
</tr>
<tr>
<td>Visits</td>
<td>Number of visits; multiple visits can be made by the same visitor in a day</td>
</tr>
<tr>
<td>Pay UV</td>
<td>Introduce traffic UV through advertising forms such as CPC / CPM / CPT.</td>
</tr>
<tr>
<td>Proportion of pay UV</td>
<td>Pay UV/UV</td>
</tr>
<tr>
<td>Direct access to UV</td>
<td>Number of visitors who directly enter the website address in the browser</td>
</tr>
<tr>
<td>Direct access to UV ratio</td>
<td>Direct access to UV/UV</td>
</tr>
<tr>
<td>Number of new visitors</td>
<td>First visit to the website</td>
</tr>
<tr>
<td>Proportion of new visitors</td>
<td>New visitors / visitors</td>
</tr>
<tr>
<td>Return visit rate of visitors in recent 30 days</td>
<td>Visitors visited in the last 30 days and the recent 30 days / number of visitors in the recent 30 days</td>
</tr>
<tr>
<td>Monthly return visit cycle of visitors</td>
<td>30 / monthly per capita visits, monthly per capita visits = all visitors in the month (multiple visits in a day count only once)/ number of visitors in the current month</td>
</tr>
<tr>
<td>Visitor acquisition cost</td>
<td>Advertising expenses / number of visitors</td>
</tr>
<tr>
<td>Number of APP activated users</td>
<td>The number of devices (including mobile devices such as mobile phones, pads, etc.) that users download and open an APP</td>
</tr>
<tr>
<td>Number of return visit users of APP in recent 30 days</td>
<td>Mobile devices accessed simultaneously in the last 30 days and the recent 30 days / number of mobile devices in the recent 30 days</td>
</tr>
</tbody>
</table>

layer is 12, and the number of nodes in the output layer is 1. Financial and non-financial indicators and platform service fee prediction models are established through the training of neural networks.

4.3 Dynamic Fitting Model Between Cost and Non-financial Indicator of Internet Enterprises

As shown in Fig. 6, the cost analysis dimensions of typical Internet enterprises are as follows: cost = sales cost + customer acquisition cost (CAC) + management cost + R & D expense. Where: customer acquisition cost = (total marketing expenses + total
sales expenses) / number of new customers, which is the basic calculation formula of customer acquisition cost. In practical application, we will have the difference between the number of new customers and the number of transactions. This algorithm does not take into account employee salary, bonus, company water and electricity, rent, etc. At present, the company’s product customer acquisition has the following ways: external channel launch, offline salesman promotion and MGM, among which MGM has the best effect. We should not only consider the promotion time cycle and the effect of launch channels, but also consider the difference of human input cost, wool party and other variables. The cost is different and the calculation method is also different.

The essence of the Internet is connection, which directly connects manufacturers and users. The power brought by the Internet is 10% digitization and 90% Internet. Among them, 10% can improve production, and 90% is the intermediate cost, including marketing expenses, channel expenses, inventory expenses, etc. “We may not be able to change 10%, but we can change 90% of the intermediate cost to zero. But whether online business or offline business, we are inseparable from the core basic problem of “customers acquisition”. Customer acquisition cost (CAC) refers to the market cost to be paid for each user, and customer lifetime value refers to the profit that a user can contribute to you in the life cycle; theoretically, as long as your customer acquisition cost is less than the customer lifetime value, you can continue to expand users until they are equal, that is, the business boundary of the channel is that customer acquisition cost equals customer lifetime value (market cost is the main cost center). Because with the
expansion of user scale, the cost of customer acquisition in the same channel will be higher and higher.

5 Conclusion

BP neural network analysis is used to analyze the fitting of financial indicators such as income, cost and expense of Internet enterprises with non-financial indicators, and the non-financial indicators are organically integrated into the basic model, which effectively makes up for the deficiency of relying on financial indicators only to evaluate the profitability of enterprises in the past, so as to make a real analysis and accurate judgment on the development status and future development trend of Internet enterprises. The main conclusions are as follows:

1. The data of non-financial indicators of Internet enterprises is large and rich. It has the characteristics of fast growth and strong timeliness. Fully extract the core non-financial indicators with financial indicators, and fully mine the required information with the help of BP neural network.
2. When analyzing the profitability of enterprises, from the perspective of financial revenue and costs, for start-up Internet enterprises, there may be huge losses or meager profits. However, from the in-depth interpretation of its relevant non-financial indicators, we will find that the loss is temporary. If the factors affecting non-financial indicators are good enough, their potential deserves the attention of investors or partners, and vice versa.

6 Research Prospect

In the era of internet sharing economy, through sharing and connecting technologies, human economy and society begin to change from material oriented to value oriented. Because value can be shared infinitely, the traditional economy brings limited customers and the sharing economy brings unlimited customers. Internet sharing economy provides a new set of wealth revenue algorithm and a new wealth production relationship chain. Where, profit equals customer base multiplier (n th-power) multiple product unit price minus total cost. In the general accounting algorithm of internet sharing economy, goods and prices are not the core. The key is the customer base and customer multiplier (N th-power), which has become the most important factor of wealth revenue. This is the biggest difference between internet sharing economy and traditional capital economy, and it is also a complete subversion of the traditional financial revenue algorithm.

References


Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.