Application Research of RPA Technology in the Construction of Financial Robot Business Automation Processing Platform

Yaru Qi*

Huainan Vocational Technical College, Huainan, Anhui, China

*2461303213@qq.com

Abstract. There are still many inefficient manual processing tasks in the financial field, so it is crucial to adopt financial robots based on RPA technology. In order to automatically obtain the financial fields of web pages and intelligent classified financial data, the financial data automatic capture robot and the financial classification robot with TF-IDF algorithm are designed, and the automatic processing platform with these two robots as the core is built. The experimental results show that the autocrawler robot can accurately generate XML files with an accuracy of more than 90%. The TF-IDF algorithm also performed well in classification accuracy, recall and F1 values, at 71%, 85% and 92%, respectively. These two kinds of financial machines can effectively realize business automation processing and improve the work efficiency of financial personnel.

Keywords: RPA technology; financial robot; classification accuracy; business automation.

1 INTRODUCTION

The development of technologies such as big data and artificial intelligence are changing the way many industries work, with machines gradually replacing humans. Because of its tedious and repetitive nature, financial work often brings pressure to practitioners and leads to work mistakes [1]. In order to relieve stress and improve efficiency, it is very important to establish a financial automation processing platform. Foreign scholars use RPA technology and artificial intelligence technology to optimize the financial process, reduce the pressure on enterprises, and improve the computing efficiency [2]. Domestic scholars have also developed financial robots such as automatic billing and financial accounting to realize the digital transformation of finance. Financial robots handle tasks according to the preset program, with high precision, which can reduce the cost of the enterprise and optimize the personnel structure [3]. RPA technology as a new theory in the field of automation, this study aims to design a financial robot based on RPA technology, build a business automation platform, and increase the intelligent attributes of financial work [4].

A. Haldorai et al. (eds.), *Proceedings of the 2024 3rd International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2024)*, Atlantis Highlights in Intelligent Systems 11,
https://doi.org/10.2991/978-94-6463-490-7_37

2 PLATFORM CONSTRUCTION

2.1 Financial data processing model based on auto-crawler robot.

Increasing financial data makes manual processing inefficient [5]. To solve this, the research uses a focused web crawler for automatic data collection. Figure 1 schematically shows the process flow of the web crawler.

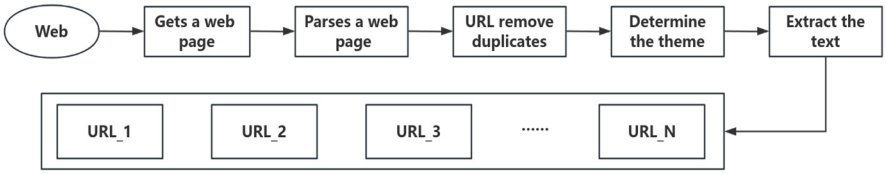

Fig. 1. Schematic diagram of the focused web-crawler process.

Crawler technology involves storing URLs, acquiring and analyzing pages, determining subject content, and extracting text. The subject content unit evaluates webpage relevance to finance. We chose XML for crawler file format due to its readability and transferability. In operation, the seed URL enters the queue and is processed by an automatic crawler robot designed for correlation determination and multi-threaded crawling. Figure 2 shows the correlation judgment process.

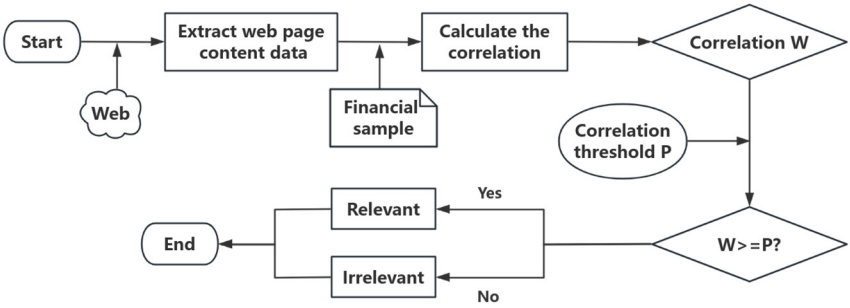

Fig. 2. Relevance judgment process.

Figure 2 flowchart defines the correlation W and lower limit P between the web page and the financial field. At WP, web data is related to financial fields; W <P, no correlation. P value setting affects the results of crawler, too large data is less, too small data is more, reduce efficiency. The production-consumer multithreaded crawler unit includes the roles of producer, consumer and both. The producer pushes the web URL to retrieve the list; the consumer analyzes the web content to determine any financial information, extracts the data set and saves it [6]. Finally, the overall structure diagram of the crawler robot is obtained, as shown in Figure 3. The crawler robot initializes the user data to the database, and extracts, classifies and merges the data on demand. Use the financial field to establish the XML file and resolve the URL list. Multi-threaded

mode extracts the web content URL, and join the list to be crawled. Finally, extract the web body data, calculate the correlation judgment, and store the results structured.

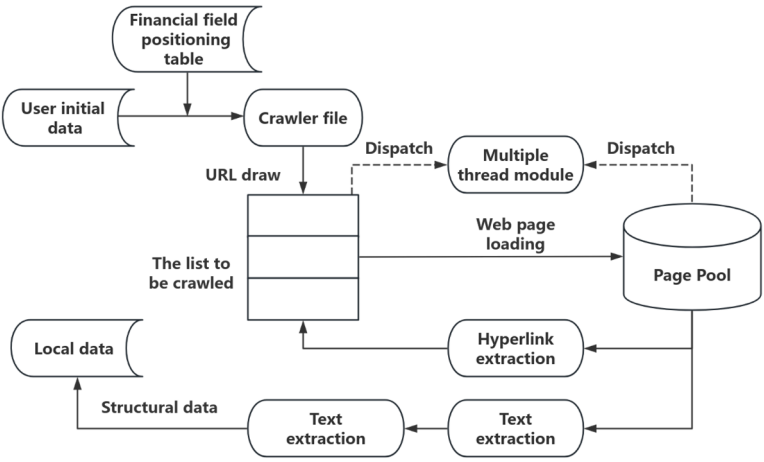

Fig. 3. Overall structure diagram of the crawler robot.

2.2 Modeling of a financial classification robot based on TF-IDF

After using the automatic crawler to obtain the financial data information of the web page, the financial text classification robot is designed to solve the problems of difficult classification and filing and heavy workload. The TF-IDF algorithm is suitable to deal with such problems, which represents the total number of feature words appearing in the text by TF and is normalized by formula 1 to evaluate the importance of the text [7].

$$TF_{i,k} = \frac{n_{i,k}}{m_i} \tag{1}$$

The IDF indicates the importance of individual feature words in the text. Its mathematical expression is shown in formula 2.

$$IDF_{i,k} = \log \frac{D}{n_k} \tag{2}$$

Formula 3 calculates the weight of feature words.

$$w_{i,k} = TF_{i,k} \times \mathrm{IDF_{i,k}} \tag{3}$$

In order to enhance text feature word extraction and sentence structure feature judgment, N-Gram model is used to deepen the similarity evaluation and feature extraction of financial text string. Assuming that the sentence S consists of L words $w_1, w_2, \ldots, w_L$, then the probability of this sentence appearing in the text is shown in formula 4.

$$P(S) = P(w_1^L) = P(w_1, w_2, \cdots, w_L) \tag{4}$$

The N-Gram model is a probabilistic language model, which treats words as random variables, and can be applied to the financial text robot classification function, as shown in Figure 4. The robot classifies based on the time condition, first extracts the financial text time, and then compares it with the preset time. Later than the condition is regarded as a new document, otherwise it needs to be filed. After completing the crawler and classification and archiving robot, the automation platform architecture, including network application end and server side, which is divided into three layers: interaction layer, business layer and data layer [8]. The interaction layer displays the operation interface and responds to user operations; the service layer and the data layer serve the server side; the service layer processes service requests and feedback results, and the data layer is responsible for data storage and other functions.

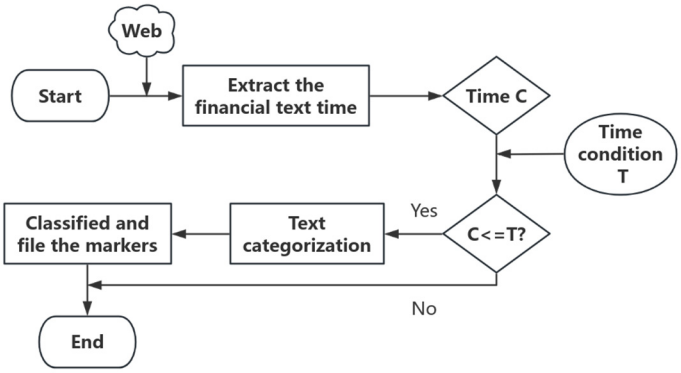

Fig. 4. Classified archiving process structure diagram.

3 ANALYSIS OF THE DATA PROCESSING RESULTS

The text correlation determination function has an important influence on the scope and amount of financial data captured by the crawler robot. The experiment crawled data from multiple websites and focused on 65 operational management samples, of which 50 were operational data and 15 were management data, aiming to capture the financial samples of operation. With a correlation threshold step size of 0.2 and a value range of [0,1], Table 1 shows the grasping results based on correlation determination.

Table 1. Grasping the results of the correlation determination of the financial text.

Financial samples			Correlation threshold	Grab the results		
ABC				A	B	C
65	50	15	0	65	50	15
65	50	15	0.2	59	50	9
65	50	15	0.4	54	50	4
65	50	15	0.6	52	50	2
65	50	15	0.8	46	46	0

65	50	15	1	0	0	0

According to Table 1, the change of correlation threshold affects the grasping results of financial samples. At the threshold of 0, the crawler grabbed freely and sampled the whole sample A with non-target C. The threshold increases, grasp less, and target B screens better. At thresholds of 0.2,0.4, and 0.6, B was 50 and C decreased to 9,4, and 2. However, if the threshold is too high as 0.8, the grasping effect decreases, and B is 46. Suitable thresholds can be accurately screened. Furthermore, the study tested the impact of XML files on crawler data transfer, divided into samples that meet and did meet XML file node data. Table 2 generates the results for the crawler file. The autocrawler robot can generate XML files well, and the accuracy rate is higher than 90% no matter why the number of XML value is set in advance. With the XML file format 25, the highest file generation accuracy was obtained with a size of 96.0%. Therefore, the study-designed automatic crawler robot can perform the text generation task very well.

Table 2. The crawler files generated the results table.

Data setting		Number of XML file generation	Accuracy rate (%)
Number of XML file formats	Number of non-XML file formats		
20	10	19	95.0
25	15	24	96.0
30	20	28	93.3
35	25	33	94.3

4 CONCLUSION

In order to solve the tedious problems of financial workers in obtaining financial data and classifying them, we studied two financial robots with RPA technology: financial data autocrawler and financial text classification robot. For the financial data autocrawler robot, the grasping effect is the best when the correlation threshold is 0.6. The robot can generate XML text with more than 90% accuracy, showing powerful text generation capabilities. Financial classification robot in the N-Gragram value of 4 had the best classification effect with an accuracy of 68.5%. When the number of feature words is set to 50%, the robot also improves the classification efficiency while maintaining a high classification accuracy, and the calculation time can only take 210 seconds. In the algorithm comparison, the financial classification robot using TF-IDF performed the best, with the classification accuracy as high as 71%, which was 5% higher than the K-means algorithm, showing its excellent financial text classification ability.

REFERENCE

1. Wang Wei. (2024). Discussion on the Application of RPA in Enterprise Management Accounting. The Chinese financier (Z1),180-181.doi:10.19294/j.cnki.cn11-4799/f.2024.z1.050.
2. Kogan, G., Kokina, J., Stampone, A., & Boyle, D. M. (2022). RPA in Accounting Risk and Internal Control: Insights from RPA Program Managers. Accounting Horizons, 1-12.
3. Ye Feiju. (2023). RPA application in the field of management accounting. Vitality (23), 157-159.
4. Qiu Lifei, Yang Guoli & Xie Danyao. (2023). Research on the application of RPA in accounting work. Fulcrum (09), 122-124.
5. Mashtakov, M. M., Shirokova, S. V., & Bolsunovskaya, M. V. (2023). Application of RPA technology in management and decision-making processes. Technoeconomics: an international journal, 2(1).
6. Jiang Yuan. (2023). Digital transformation is essential: the RPA technology innovation of the integration of finance and management accounting. Tax payment (16), 49-51.
7. Shi Qiuxia. (2023). Research on the digital transformation of RPA Robot-driven Management accounting. Accounting Research (04), 53-57.
8. Langmann, C., & Turi, D. (2023). Application Examples for RPA in Financial and Management Accounting. In Robotic Process Automation (RPA)-Digitization and Automation of Processes: Prerequisites, functionality and implementation using accounting as an example (pp. 121-126). Wiesbaden: Springer Fachmedien Wiesbaden.