Fuzzy Systems and Data Mining X
A.J. Tallón-Ballesteros (Ed.)

doi:10.3233/FAIA241421

A Study on ERP Autonomous Experiment Assisted by Digital Intelligence Fusion

Weibo Huang[1], Ziting He, Xiaopeng Xie and Haolin Chen
Guangdong University of Foreign Studies, Guangzhou, China.

Abstract. This study applied Digital Intelligence Fusion to explore an autonomous experimental model for ERP learning systems to enhance the practical experience for the undergraduates. Using AI's ability to create virtual environments, the proposed model simulated real-life scenarios to provide students with an immersive educational experience. The study employed big data analysis to build an intelligent ERP system through data warehousing and data mining techniques to improve data quality and automatically correct inaccuracies. The model performs well in capturing contextual features and generating responses that are very similar to the actual responses. This novel approach demonstrated a unique contribution to this field, enhancing ERP education through AI-driven immersive learning and big data analytics.

Keywords. ERP, AI, Digital Intelligence Fusion, Autonomous Experiment

1. Introduction

The rapid development of artificial intelligence provides new possibilities for the reform of the teaching in economics and management courses. It is necessary to address the challenges of ERP systems that are developing in sync with today's social innovation, which has led to an urgent demand for ERP educational tools [1]. The capacity of AI to construct virtual environments allows it to simulate complex scenes in the real world and provide students with an immersive learning experience that closely resembles the real situation [2]. This not only compensates for the deficiencies of traditional laboratories and AI experimental platforms in simulating the workflow and environment of enterprises and institutions, but also eliminates geographical and temporal constraints, thereby enhancing the flexibility of practical teaching [3].

In recent years, a significant number of higher education institutions, both domestic and international, have engaged in active research and development concerning the potential applications of artificial intelligence in the field of education. [4] For example, at Stanford University Business School, AI is being introduced into the fields of economics and management education with the aim of enabling students to gain a deeper understanding of complex business concepts and strategies. This is achieved by simulating business environments and decision-making scenarios. Similarly, at Wharton

[1] Corresponding Author: Weibo Huang, Center for Contemporary Education Technology, Experimental Teaching Center, Guangdong University of Foreign Studies, China; Email: 172564946@qq.com.
This study is financially supported by the Undergraduate Innovation Training Project of Guangdong University of Foreign Studies in 2024, and is also financially supported by the Experimental Teaching and Teaching Laboratory Construction Research Project of Guangdong University of Foreign Studies in 2024.

Business School, AI is being used to deliver financial and investment education, aiming to improve students' investment decision-making and risk management skills. In this context, developing an ERP autonomous experimental model can provide students with more practical learning opportunities. By combining artificial intelligence technology with the ERP autonomous experimental model, students can apply their learned knowledge in real environments, thereby strengthening their practical operation and problem-solving abilities, and helping them better adapt to the challenges and opportunities of future business environments.

2. ERP Autonomous Experiment Model Based on Big Data

The use of big data technology to construct an intelligent ERP system with a big data centre to develop an enterprise resource planning system for e-commerce companies based on data warehousing and data mining technology [5], in which the input of fundamental data is charged to the application of algorithms to quality mining and automatically rectify the data [6].

In the analysis centre, particularly in the customer management module, neural network algorithms and other such techniques will be employed to analyse the potential consumption patterns of customers, utilising the characteristics of microservice architecture. The big data centre system establishes an ERP autonomous experimentation model, as illustrated in Figure 1.

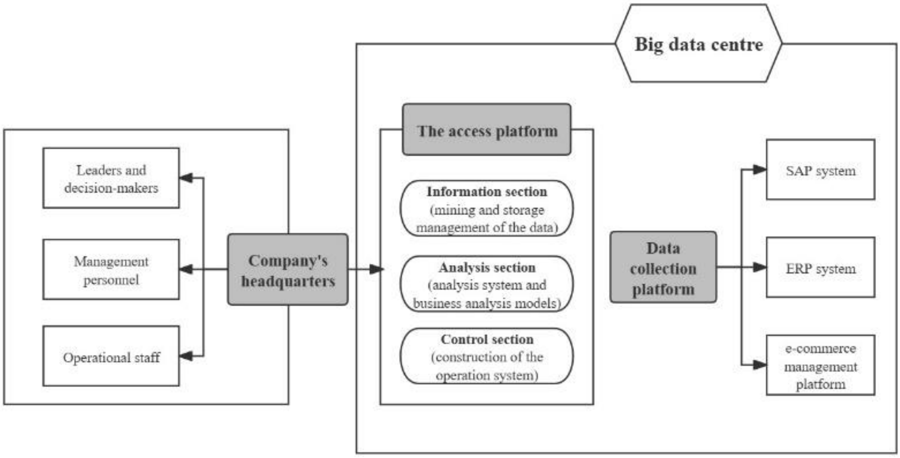

Figure 1. Autonomous ERP experiment model built through the big data centre system

The entire big data centre is managed by the company's headquarters. Leaders and decision-makers, management personnel, as well as operational staff can access the big data centre through the access platform [7]. The centre is divided into three core sections.

The information section bears responsibility for the mining and storage management of the contained data, which maintains the basic data required for the ERP system is retained [8]. Additionally, the information section will encompass a customer management module dedicated to data processing, which can be utilized by operational staff for the analysis of the data.

The role of the analysis section is to undertake analysis of the system and to process business analysis models. Operators and managers analyse the operation and processing of the basic business service modules and the modules of microservices architecture in this section, employing analytical tools to ascertain whether the business modules are fulfilling their intended function. Under the microservice architecture, ERP systems are divided into multiple independent services, each responsible for a portion of business functions, with its own database, and capable of independent deployment and expansion. Additionally, microservices may need to call each other.Furthermore, the section analyses data from the customer management module, examining the consumption requirements of each disparate computer customer, which can be represented as data within the analysis section.

The control section is responsible for the construction of the operation system of the entire big data centre. This includes eight different data categories: data security management, data quality management, data model management, master data management, data standard management, HI data management, data model management, data lifecycle management, and basic data management. Based on the data category, corresponding data processing systems are used for data control.

The data collection platform is primarily constituted by the SAP system, the ERP system, and the e-commerce management platform.

3. Artificial Intelligence Autonomous Experiment Information Push

The collaborative filtering recommender system is based on the Transformer model, which is advantageous in natural language processing and widely applied in the field of natural language processing (NLP), self-attention mechanism is its core.The input information is encoded and mathematical computations are performed to obtain the predicted values. These values are then compared with the pre-training results on the corpus to determine the final output. In this study, the GPT model (decoder structure only) in Transformer is employed. Its training process commences with unsupervised pre-training on a substantial corpus, followed by fine-tuning utilising supervised data from the current task. This step will be undertaken here with the aforementioned fine-tuned corpus. The initial stage will involve the introduction of a tokeniser, that is to say a lexical module, which will perform the automatic segmentation of the input lexicon. Subsequently, each input word is transformed into a word vector through the application of a position embedding algorithm, which typically entails a sequence of matrix and weighting operations. In order to clearly represent the question-answer position relation, a new token is added to the word list as a segmentation in this module.

For example, <CLS><sentQ>question<sentA>answer<SEP>

The token [CLS], which is similar to token [BOS] and is frequently employed in classification tasks to mark the beginning of a sequence, will be appended before each sentence in the training set. Subsequently, the token [SEP], which is analogous to the token [EOS] used to segment different sentences or sequence fragments, will be added. The position corresponding to each [CLS] is the sentence vector of each sentence.

The sentQ and sentA are labelled with the starting position of the question and answer respectively. They are placed at the beginning or end of the sentence to serve as spacers that enable the model to grasp the interval between two sentences, thereby facilitating the recognition and reinforcement of learning associated with these key markers.

In the position encoding stage, the segmentation embedding, which has been processed as described above, is summed up with the input token embedding and position embedding in Formula 1. The resultant values of the fusion calculation are then output.

$$\text{embedding} = \text{token} + \text{segment} + \text{position} \tag{1}$$

In the pre-training phase of GPT, a multilayer decoder is employed. After positional encoding of the input, the decoder performs multi-head self-attention operations on the input contextual vocabulary. This is followed by a Feed-Forward Neural (FFN) layer, which conducts linear transformations on the attention scores. Ultimately, the output of the target sentence was generated. The process of learning the text from the unsupervised labelled corpus and thus generating the predicted next sentence can be summarised in Formula 2 to Formula 5.

$$L_1(U) = \sum_i log\, P(u_i | u_{i-k}, \dots, u_{i-1}; \Theta) \tag{2}$$

$$h_0 = UW_e + W_p \tag{3}$$

$$h_l = transformer\_block(h_{l-1}) \forall i \in [1, n] \tag{4}$$

$$P(u) = softmax(h_n W_e^T) \tag{5}$$

Here, U=(u-k, …, u-1) denotes the context vector of the input participle. k is the size of the context window, n is the number of layers, P represents the conditional probability of the generated result, W_e is the participle embedding matrix, W_p is the positional embedding matrix, and the above two matrices are evolved from the basic GPT model embedding layer.

Following the completion of the aforementioned extensive corpus training, the model adjusts its parameters in accordance with the specific requirements of the prediction task. In this context, the model is designed to meet the needs of the prediction service, as outlined in this paper. The textual corpus data is drawn from the target users themselves and their surrounding users. If dataset D has been created following the application of manual labelling, each instance will comprise a sequence of input tokens $(\mathrm{x}^1, \dots, \mathrm{x}^\mathrm{m})$. These inputs are then passed through the model that has completed its initial training phase to obtain the final activation value of the Transformer module, represented by h_l^m. This activation value is subsequently fed into an additional linear output layer with parameter W_y, which allows for the continuation of the prediction process in depth, as illustrated by the following Formula 6.

$$P(y | x^1, \dots, x^m) = \mathrm{softmax}\left(h_l^m W_y\right) \tag{6}$$

At the same time, the pre-trained model, as previously described, will undergo fine-tuning through the use of language modelling (LM). The initial phase of language modelling pre-training is primarily concerned with word and word sense level processing. This implies that the language model must also possess the capacity to comprehend sentences and the relationships between them. Consequently, this LM-based fine-tuning can accelerate the convergence of the model and simultaneously enhance the generalisation of the supervised model. In other words, the model is capable of predicting the vocabulary with greater frequency, taking into account the contextual information of the input, and is also able to identify and rectify errors on occasion.

Its training process is to predict the next word based on the previous word sequence. For a long sequence of words S, as shown in Formula 7. where w denotes different parameters:

$$S = \{w_1, \dots, w_{|S|}\} \tag{7}$$

After completing the above steps, the probability value of the next maximum probability word will be predicted based on the previous inputs or existing words in the process shown in Formula 8:

$$\mathcal{L}_1(S) = \sum_{i=1}^{|S|} \log P(w_i | w_0, w_1, \dots, w_{i-1}) \tag{8}$$

For the specific LM objectives of this study, the loss function for the language modeling fine-tuning module (hereinafter referred to as Loss) is constructed by calculating the log-likelihood function using the logits derived from the normalized hidden states via the softmax function and the actual labels. The formula for log-likelihood computation is as follows:

$$\mathrm{Loss}(\mathrm{x}, \mathrm{class}) = -\frac{1}{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{n}} \log\big(\Pr(\mathrm{x}[\mathrm{j}] | \mathrm{x}[\mathrm{class}])\big) \tag{9}$$

Here, x[class] denotes the normalised probability value of x under class.

In the experimental process, the modelling process of LM is optimised by weights μ, as shown in Formula 10.

$$\mathrm{L}_3(\mathrm{D}) = \mathrm{L}_2(\mathrm{D}) + \mu(\mathrm{L}_1(\mathrm{D})) \tag{10}$$

The only additional parameters required during the fine-tuning process are the parameter W_y derived from the preceding output layer and the recently incorporated separator tags <CLS> and <SEP> as previously outlined.

After completing the transformer model pre-training and fine-tuning, the target user's raw data is used to input into the transformer and predict the output. The model generates the results, extracts the integration information, checks the keywords, and reintroduces them into the collaborative filtering recommender system. It then calculates the similarity between the transformer prediction generation results and other users, re-generates the target user's neighbourhood matrix, and obtains a new nearest-neighbour UserNeig of the target user.

The final stage is to forecast the target user's u-rating for the target user neighbourhood, newUserNeig(u), generated by the aforementioned module. This is achieved through the utilisation of a weighted average method to compute the scoring method, which then allows for the generation of a list of recommendation results. The N highest-rated items are then selected for recommendation to the user for reference.

Here, $\mathrm{P}_{\mathrm{u,j}}$ denotes the predicted rating value of the target user u for item j, $\overline{\mathrm{R}_{\mathrm{u}}}$ denotes the average rating of the target user u for the completed rated items, $\mathrm{R}_{\mathrm{v,j}}$ denotes the rating of the neighbouring user v of the target user u for item j, $\overline{\mathrm{R}_{\mathrm{v}}}$ denotes the average rating of the user v for the completed rated items, n denotes the number of the neighbours of the target user u, and $newSim(v,u)$ denotes the similarity between users v and u calculated using this method, λ is the start-up coefficient $0<\lambda\leq1$. If the target user or item is a new user or item, it will face the cold start problem. The start-up coefficient will weaken the effect of this problem and reduce the interference of the new user or item on the predicted rating value until enough interaction data is collected. θ is the personalisation coefficient ($\theta>0$), which is determined by the user's historical behaviour and preferences obtained by speculating on consumer psychology and behaviour. This coefficient increases the predicted score value more or less according to the personalisation choices of the target user.

After verifying the accuracy of the model and the practicality of the mathematical formulas, we conducted a control group experiment through model calculations to validate the model's ability to improve the effectiveness of ERP experimental teaching through digital prediction and feedback.

Step 1: Put the two types of sample datasets into the Adaboost algorithm for iterative classification of sample weights.

Step 1: Input the sample data set Z=(mi, ri), where i=1,2,... n. Execute SVM algorithm.

Step 2: Find the hyperplane that can best partition the two types of data samples, i.e., satisfy r (T'h *mi+u)1 $\geq$ 0, calculate the objective function min ?

Step 3: Determine whether the sample data can be classified well under general constraints; if not possible, add a relaxation variable T to soften the maximum interval soft margin between the maximum support vectors. This allows for a small number of classification errors to occur.

Step 4: Find the optimal parameter T best, and obtain the corresponding optimal hyperplane expression.

Step 5: Substitute into the new sample dataset Z, calculate in the solved optimal hyperplane expression, obtain the result. Does the ERP autonomous experimental model using the integration of data and intelligence have a significant effect?

The application of these mathematical principles and mechanisms of the Transformer model in the autonomous experimental mode of ERP can improve the efficiency and accuracy of data processing, especially when dealing with large amounts of business data and conducting complex business logic analysis. By using Transformer models, ERP systems can better understand and predict business processes, optimize resource allocation, and make decisions

4. Results Analysis

The ideal generative model should be able to create responses that are not only logically sound, but also exhibit rich diversity in types and scope. In other words, an efficient generative model should be able to generate both reasonable and diverse responses.

After processing the sample dataset from the second experimental teaching, we can clearly see that the model we use is more accurate, and the learning ability, practical ability, and scenario application ability of the student group using our AI autonomous experimental information push model are significantly better than those using the original ordinary teaching mode.

The improved model of this paper generated significantly more accurate responses than information retrieval-based models such as Boolean Model 25 in all three types of datasets. In addition, on the 1st-4th indicators of the BLEU metric, the model constructed in this study has shown higher efficiency and better level of accuracy compared to the current optimal model BART in NLG under the Qingyun database test, which demonstrated that the statements generated by this model were more similar to the real replies. Therefore, the GPT model constructed in this paper achieved better results in capturing contextual features. Therefore, it provided more complete and comprehensive input information for the collaborative filtering recommender system and improved the overall recommendation accuracy.

5. Conclusions

The ERP course emphasises the interconnection and mutual reinforcement between theoretical knowledge, practical application, and artificial intelligence, particularly within the context of artificial intelligence-based pedagogical approaches. Based on our improved model, we can set up, for example, an intelligent recommendation system that recommends personalized learning resources based on students' learning preferences and behavioral data, achieving personalized teaching, or a learning situation analysis system that collects and analyzes students' learning data, uses machine learning models to evaluate students' learning progress and mastery, and helps teachers provide accurate teaching feedback. The acquisition of theoretical knowledge furnishes the requisite guidance and support for practical operation and artificial intelligence. Conversely, practical operation and artificial intelligence represent the specific application and deepening of theoretical knowledge. The three elements are mutually reinforcing, facilitating students’ comprehensive engagement with the subject matter and fostering a deeper understanding of ERP courses, thereby contributing to the development of highly skilled professionals.

References

[1] Sollosy M, McInerney M. Artificial intelligence and business education: What should be taught. The International Journal of Management Education. 2022;20(3):100720, doi: 10.1016/j.ijme.2022.100720.

[2] Kumar S, Rao P, Singhania S, Verma S, Kheterpal M. Will artificial intelligence drive the advancements in higher education? A tri-phased exploration. Technological Forecasting and Social Change. 2024;201:123258, doi: 10.1016/j.techfore.2024.123258.

[3] Gangwar H, Mishra R, Kamble S. Adoption of big data analytics practices for sustainability development in the e-commerce supply chain: a mixed-method study. International Journal of Quality & Reliability Management. 2023;40(4):965-989, doi: 10.1108/IJQRM-07-2021-0224.

[4] Bandara F, Jayawickrama U, Subasinghage M, Olan F, Alamoudi H, Alharthi M. Enhancing ERP Responsiveness Through Big Data Technologies: An Empirical Investigation. Information Systems Frontiers. 2024;26:251-275, doi: 10.1007/s10796-023-10374-w.

[5] Alaskari O, Pinedo-Cuenca R, Ahmad MM. Framework for implementation of Enterprise Resource Planning (ERP) Systems in Small and Medium Enterprises (SMEs): A Case Study. Procedia Manufacturing. 2021;55:424-430, doi: 10.1016/j.promfg.2021.10.058.

[6] Chen L, Chen P, Lin Z. Artificial Intelligence in Education: A Review. IEEE Access. 2020;8:75264-75278, doi: 10.1109/ACCESS.2020.2988510.

[7] Li B. Research and development of e-commerce ERP system based on artificial intelligence technology. International Journal of Knowledge-Based Development. 2023;13(2-4):327-343, doi: 10.1504/IJKBD.2023.133333.

[8] Romero JA, Abad C. Cloud-based big data analytics integration with ERP platforms. Management Decision. 2022;60(12):3416-3437, doi: 10.1108/MD-07-2021-0872.