International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-10

https://doi.org/10.70917/ijcisim-2026-0052
Article

Dynamic update and adaptive prediction analysis of user profiles based on conditional random fields

Jin Li *

School of Information and Engineering, Swan College, Central South University of Forestry and Technology, Changsha, Hunan, 410211, China; lijinchunchun1982@163.com

Abstract: With the rapid development of mobile internet, user behavior data has experienced explosive growth, posing significant challenges to traditional static user profiling methods. Traditional methods often overlook the dynamic nature of user behavior, making it difficult to accurately capture changes in user interests. To address this issue, this paper proposes a user profiling method based on the Conditional Random Field (CRF) model, combining dynamic updates and adaptive predictive analysis. This method fully leverages the advantages of CRF in modeling long-range dependencies. By designing flexible adaptive feature templates and optimized training strategies, the user profiling can precisely track the evolution of user interests. Experimental results show that the model significantly outperforms traditional methods in terms of accuracy, precision, recall, and F1 score, particularly among highly active user groups. The method not only significantly improves the timeliness and accuracy of user profiles but also provides more reliable data support for personalized recommendation systems and precision marketing. By adapting to changes in user behavior trends, this method offers more effective support for various applications based on user profiles.

Keywords: conditional random fields; user profiling; dynamic updating; adaptive prediction; data mining

1. Introduction

1.1. Research Background and Significance

In the context of the booming digital economy, the explosive growth of user data is reshaping the underlying logic of personalized services [1-2]. Companies are no longer satisfied with static user profiles and instead seek to dynamically, real-time, and precisely capture user behavior trends. Traditional user profiling methods rely heavily on static attribute tags, making it difficult to capture shifts in user interests, behavioral diversity, and time-sensitive needs. These methods are increasingly unable to meet the real-time and predictive data requirements of scenarios such as personalized recommendations, precise ad targeting, intelligent customer service, and user retention [3-5].

With the rapid evolution of application scenarios such as social platforms, e-commerce, and content distribution systems, user behavior not only exhibits extreme diversity in frequency but also shows obvious non-stationary characteristics in time series [6-7]. These characteristics hide enormous value, and only user profiling systems that can dynamically sense changes in user behavior and achieve adaptive adjustments can truly optimize services “centered on the user.”

Academia and industry have conducted extensive exploration in user modeling, particularly the widespread application of deep learning models (such as LSTM and Transformer) in modeling temporal behavior [8-9]. However, these methods often suffer from issues such as complex structure, high training costs, and weak generalization capabilities, and exhibit certain limitations in scenarios with small samples, cold starts, and strong feature dependencies [10-11]. Conditional Random Field (CRF) models, which excel at handling long-range dependencies in sequence labeling tasks, have increasingly become a preferred choice for dynamic modeling. They can flexibly incorporate multiple feature functions while capturing the relationships between state transitions and context, making them naturally suited for the complex, multi-dimensional, and dynamically changing data structures of user profiling [12-13].

Additionally, with the recent trend toward more refined user segmentation and deeper behavioral prediction, user profiling systems are no longer limited to single static label outputs but have shifted toward continuous learning and updating [14-15]. This trend presents new modeling challenges: how can the system maintain stability while also having the ability to “quickly sense” changes? How can prediction accuracy be maintained even when user behavior fluctuates dramatically? How can the system accommodate the differing interests of highly active and less active user groups? These issues require systematic solutions.

To this end, an increasing number of researchers are attempting to introduce model structures with time-aware mechanisms to achieve dynamic modeling of user interest evolution. In this direction, CRF has a natural advantage, but it still requires algorithmic innovations in feature selection, adaptive updating, and time weight allocation to unlock its potential in user profiling [16]. Only through the deep integration of model structures and data processing strategies can predictive behavioral patterns be extracted from complex user behavior data, driving a fundamental leap in the capabilities of intelligent recommendation systems and precision marketing.

Conditional Random Fields (CRFs) and their derivative models have widespread applications, including popular fields such as user preferences, pedestrian intent, and emotion prediction. For example, [17] proposes a CRF-based personalized faceted recommendation method to address multi-faceted search problems, using a faceted classification model to predict and rank user interest levels, and recommending the highest-interest facets to users. Derivative models of CRFs also demonstrate excellent performance in adaptive prediction. Reference [18] proposes a conditional random field model based on a tripartite graph, creating a probabilistic model based on a tripartite graph to reveal user preferences. In testing experiments, the model achieved an average improvement of 15% in metrics such as diversity, recall rate, and F1 score. In pedestrian motion prediction, Reference [19] utilizes a graphical factor latent dynamic conditional random field model to capture the intrinsic relationship between pedestrian intent and action. It finds that pedestrian motion information can enhance the model's prediction accuracy of pedestrian intent. Experimental results show that the model can predict pedestrian intent 1.2 seconds before the action is executed, with a prediction accuracy exceeding 70%. Furthermore, Reference [20] employs a latent dynamic conditional random field model to enhance performance prediction. According to experimental results, this model achieves predictions 0.9 seconds in advance of actual event occurrence across datasets, leveraging long short-term memory networks on datasets with identical time series features. Reference [21] designed a sparse hidden dynamic conditional random field model to understand user intent from users' search sessions. The results show that the proposed model significantly outperforms support vector machines, conditional random fields, and Latent Latent Network (LatNet) - dynamic conditional random field models in terms of intent prediction accuracy. The multi-scale continuous conditional random field model proposed in [22] effectively addresses issues in social recommendation systems by optimizing both relationship dependencies and feature combinations. Experimental results on real-world datasets demonstrate that the proposed model outperforms traditional system filtering recommendation algorithms. Literature [23] combines continuous conditional random fields with support vector machines to model emotion prediction in dimensional space. The results show that this method outperforms previous baseline methods across all four emotional dimensions: valence, arousal, strength, and expectation.

With the development of intelligent algorithms, it is necessary to explore combining intelligent algorithms with conditional random fields to optimize the performance of single conditional random field models. Literature [24] expands the predictive capabilities of conditional random fields through mixed modeling. The study found that the conditional random field model using mixed learning algorithms outperforms single conditional random field models in various scenarios and demonstrates the predictive accuracy of mixed models in sequence labeling problems.

In summary, these studies validate the effectiveness and feasibility of conditional random field models in different prediction tasks, providing valuable references for the research topics discussed in this paper.

1.2. Innovative Aspects of This Study

The improved method proposed by this research institute focuses on the common issues of lag and rigidity in traditional user profiling modeling processes. It attempts to dynamically track and model the evolution of user interests by introducing the structural framework of conditional random fields. In constructing the model, we introduced a two-layer cross-structure. The uniqueness of this design lies in the fact that it does not simply extract features from time series but instead embeds behavioral changes and interest evolution into two layers of different scales. The lower layer uses a set of local feature functions to perceive short-term behavioral changes at a high frequency, capturing users' immediate responses at the micro level. The upper layer focuses on the stability and trend of user interests over a larger time window and models them using long-range feature functions. The parameter sharing mechanism between the two layers enables the collaborative propagation of information between them, significantly outperforming previous single-layer structures that only considered short-term states or long-term preferences. This approach can more precisely depict the temporal evolution process of user profiles.

In terms of feature template generation, this study takes a different approach by introducing an adaptive adjustment mechanism, whose core is the real-time adjustment of dynamic weighting factors. To achieve this adaptive update, the system calculates the weighted sum of the variation magnitudes of each feature based on the latest behavioral change data at a given time point, applies a nonlinear transformation in the form of a σ function, and obtains dynamic weighting factor values adapted to the current context.

At the same time, to improve the stability and convergence speed of the training process, we constructed a parameter update strategy based on historical gradients and introduced a gradient memory matrix to store the directional information accumulated during the model learning process. Compared to relying solely on current gradients, this strategy is less likely to fall into local optima when faced with complex user behavior data and further improves overall training efficiency.

In summary, systematic optimization across three distinct levels enables the model to demonstrate significantly superior dynamic adaptability and prediction accuracy compared to traditional methods during experiments, thereby achieving superior performance in user profile update tasks.

2. Research Methods

2.1. Theoretical Basis of the Study

Conditional Random Field theory, through the continuous exploration of researchers such as Lafferty, has gradually evolved into a unique discriminative probabilistic graphical model. Unlike generative models such as hidden Markov models, which require numerous assumptions about the joint distribution of observation sequences, the characteristic of conditional random fields to directly model conditional probability distributions gives them unique advantages in practical applications. This undirected graph model has opened up a new approach to the problem of sequence labeling, and the mathematical expression of conditional random variables can be traced back to the in-depth research on conditional probability logic conducted by the Giglio team.

From a mathematical perspective, a conditional random field is defined on an observation sequence X and a label sequence Y. Given an observation sequence X, the conditional probability expression for calculating the label sequence Y is:

$$P(Y \mid X) = \frac{1}{Z(X)} \exp\left(\sum_k \lambda_k f_k(Y, X) \right) \tag{1}$$

In the equation, $Z(X)$ represents the normalization factor, which ensures that the sum of probabilities is 1; f_k represents the feature function, which describes the association between the label sequence and the observation sequence; λ_k is the weight parameter. This construction method is closely related to the ring condition theory proposed by Orsak. The specific calculation method of the normalization factor $Z(X)$ is as follows:

$$Z(X) = \sum_Y \exp\left(\sum_k \lambda_k f_k(Y, X) \right) \tag{2}$$

This expression form has a similar functional structure to the conditions established by Sun Weifeng in his research on the stability of the Bergomolny equation. Model training involves two core steps: parameter estimation and feature selection. Parameter estimation uses the maximum likelihood method, which is achieved by maximizing the log-likelihood function of the training data, i.e.,:

$$L(\lambda) = \sum_{i=1}^{N} \log P(y^{(i)} \mid x^{(i)}; \lambda) \tag{3}$$

An improved quasi-Newton algorithm is introduced during the model optimization process, combining the advantages of finite memory to improve computational efficiency. In the feature selection stage, this paper designs feature functions that can capture local dependencies in sequences. These functions include state features $s(y_i, x, i)$ and transition features $t(y_{i-1}, y_i, x, i)$, whose construction methods are highly consistent with the ideas of existing related research.

Conditional random fields demonstrate significant advantages over traditional hidden Markov models, not only by breaking free from the constraints of the feature independence assumption but also by allowing the introduction of an unlimited number of feature functions, significantly enhancing the model's expressive power. To enhance the model's generalization ability, this paper introduces a regularization term [25] into the objective function. Commonly used regularization methods include L1 and L2 regularization, namely:

$$L^{'}(\lambda) = L(\lambda) - \frac{\sigma}{2} \| \lambda \|^2 \tag{4}$$

The regularization coefficient σ balances the model's fitting ability and generalization performance by adjusting the model complexity. In practical applications, feature template design, sample quality, and computational efficiency all affect model performance, opening up new avenues for feature template optimization.

2.2. Data Collection and Preprocessing

This study collected user behavior data from a certain e-commerce platform between September 2023 and February 2024. A crawler system developed using Python was used to obtain user behavior data, and historical profile information was collected synchronously through interface calls. The collected data dimensions covered user behavior trajectories in internet scenarios. The data types and formats are shown in Table 1.

Table 1. User behavior data sheet.

Data type	Collection dimension	Data format
Browsing behavior	Page dwell time, click depth, browsing path	Numerical type, sequential type
Interactive behavior	Collection mark, comment content, share count	Boolean type, text type, numerical type
Purchase behavior	Order amount, purchase frequency, repurchase rate	Numerical type
Search behavior	Key words, search time, click sequence	Text type, time type, sequence type

A differentiated processing strategy was adopted for cleaning the raw data. Gradient statistics were used to detect outliers in the time series data, and window functions were used to eliminate short-term noise fluctuations. Setting the time window size to w, the smoothing calculation formula for the time series χ_τ is:

$$\hat{x}_t = \frac{1}{w} \sum_{i=t-w+1}^{t} x_i \tag{5}$$

Three strategies are used to handle missing data. For missing rates below 5%, nearest neighbor interpolation is used. For missing rates between 5% and 15%, collaborative filling based on user similarity is used. Features with missing rates above 15% are directly discarded. Data quality control uses multi-layer filtering: business logic verification removes non-compliant data, and Bayesian anomaly detection identifies extreme values. The detection function is:

$$P(x_t \mid x_{1:t-1}) = \int P(x_t \mid \theta) P(\theta \mid x_{1:t-1}) d\theta \tag{6}$$

Combine domain rules for business rationality filtering, standardize data using the minimum-maximum normalization method, and map features to the [0,1] interval, i.e.,:

$$x_i^{'} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{7}$$

After rigorous data preprocessing, we obtained 247,893 high-quality user behavior records. The dataset was divided into a training set and a test set in an 8:2 ratio, with each record containing 42 feature dimensions. These features cover user basic attributes, behavioral characteristics, and historical profile information. The skewness coefficient of the preprocessed data is within the range of [-0.5, 0.5], laying a reliable foundation for model training.

2.3. Feature Extraction and Template Design

Accurate extraction and modeling of user behavior characteristics are key steps in dynamic profile updating. To address this challenge, we drew on the idea of weight adaptation in neural networks and designed a multi-scale feature extraction scheme. This scheme breaks down user profile characteristics into three dimensions: static attributes, dynamic behavior, and temporal evolution. It constructs a feature vector $\phi(x)$ to describe user behavior trajectories, as follows:

$$\phi(x) = [\phi_s(x), \phi_d(x,t), \phi_e(x, t-w:t)]^T \tag{8}$$

In the equation, $\phi_s(x)$ describes the user's basic attributes, $\phi_d(x,t)$ reflects real-time behavioral characteristics, and $\phi_e(x, t-w:t)$ captures the evolution of user interests through a time window w. Considering that different features contribute differently to the user profile, we introduce a dynamic weighting mechanism to construct the feature template F_t at time point t, i.e.,:

$$F_t = \sum_{k=1}^{K} \omega_k(t) \cdot f_k(x_t) \tag{9}$$

The calculation of dynamic weight $\omega_k(t)$ uses the attention mechanism, namely:

$$\omega_k(t) = softmax(\frac{Q_k(t)K_k(t)^T}{\sqrt{d}})V_k(t)^{\psi} \tag{10}$$

Based on this adaptive feature extraction framework, we conducted an in-depth study on the design methods of three types of basic feature functions. The time decay feature function $f_d(x_t) = e^{-\lambda(t_0 - t)} \cdot \chi_t$ quantifies the timeliness of user interest using an exponential decay form, while shopping behavior sequences are modeled using the state transition characteristics of Markov chains. Browsing behavior is described using cumulative statistical features of time windows. The generalization performance of feature templates is effectively controlled by introducing a regularization term, namely:

$$L(\omega) = \| F_t - F_{t-1} \|_2 + \alpha \| \omega \|_2 \tag{11}$$

The regularization coefficient α in the formula needs to be determined through cross-validation to find the optimal value.

Related research shows that this feature extraction scheme, which combines static attributes and dynamic behavior, can not only accurately characterize the gradual process of user interest, but also sensitively capture sudden shifts in interest, laying a solid foundation for subsequent user profile prediction analysis.

2.4. Model Training Optimization Methods

This study employs an improved conditional random field model to analyze user data. To address the challenge of capturing changes in user behavior, a weighted maximum likelihood estimation algorithm was designed. This algorithm adjusts sample weights based on timeliness, enabling the model to be more sensitive to recent behavioral changes.

For the training samples $(x^{(i)}, y^{(i)})$ in the time window $[t-w, t]$, the weighted log-likelihood function is constructed by calculating the weight coefficient $\gamma_i = e^{-\lambda(t-t_i)}$, i.e.,:

$$L(\theta) = \sum_{i=1}^{N} \gamma_i logP(y^{(i)} \mid x^{(i)}; \theta) \tag{12}$$

This time-decay-based weighting method effectively balances the conflict between historical data and real-time data. The study also used adaptive predictive analytics to dynamically optimize the forecast results and adjusted the forecast values using exponential smoothing techniques. The core forecast formula is:

$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)\hat{y}_{t-1} \tag{13}$$

The smoothing coefficient α is automatically updated based on the trend of the prediction error, i.e.,:

$$\alpha_t = \sigma(\beta \cdot \Delta e_t) \tag{14}$$

Here, Δe_t represents the rate of change in prediction error, β is the sensitivity parameter, and σ is the S-shaped activation function.

During the study, an online learning algorithm for error gradients was introduced, which updates the model parameters each time a new sample is obtained, i.e.,:

$$\theta^{(t+1)} = \theta^{(t)} - \eta_t \nabla L_t(\theta^{(t)}) \tag{15}$$

The learning rate η_t is adjusted according to the annealing strategy, that is:

$$\eta_t = \eta_0 / (1 + \delta t) \tag{16}$$

In the formula, η_0 is the initial learning rate, and δ is the decay coefficient.

To prevent overfitting, the L_2 regularization term is added, that is:

$$L^{'}(\theta) = L(\theta) - \frac{\lambda}{2} \| \theta \|^2 \psi \tag{17}$$

Based on the above strategy, model training is implemented to provide accurate model support for dynamic updates to user profiles.

3. Experimental Analysis

3.1. Changes in Model Training Loss

After the model is created, training begins. The training process consists of several stages. First, the network hyperparameters are reasonably configured and set. Next, sample data augmentation methods are used to improve the model's generalization ability. Then, the model is actually trained. During this period, the training process is continuously evaluated and optimized to achieve the best model performance. Before initiating the model training process, the proper configuration of hyperparameters is a critical step. This is because the selection and configuration of hyperparameters directly influence the direction of the entire training process and the final performance of the model. Unlike general parameters adjusted through network evaluation optimization during training, hyperparameters are important control variables that must be pre-set and determined before formal training begins. They are not automatically obtained by the model through learning on training data but must be manually set by researchers based on experimental objectives, system environmental conditions, expected results, and other relevant factors.

Therefore, in the research practice of this paper, before commencing actual network training operations, key hyperparameters such as learning rate, batch size, regularization strength, and the number of hidden layer nodes are carefully and comprehensively considered and set based on the specific experimental requirements and constraints of the system. This aims to maximize the optimization of model training efficiency and prediction performance, ensuring that the model can meet the requirements of specific application scenarios while also achieving ideal generalization capabilities. The specific parameter settings are shown in Table 2.

Table 2. Hyper parameter setting.

Name	Default value	Optimal value
Batch Size	36	6
Learning Rate	0.005	0.0005
Epochs	30	120
Momentum	0.88	0.99
WeightPiror	0.005	0.00005

Within the supervised learning framework, given a training dataset containing multiple samples with input feature vectors and corresponding true labels, the loss function is designed as a non-negative real-valued function that measures the distance or consistency deviation between the model's parameterized predicted output and the true labels. Specifically, for each training sample, it evaluates the “error” level of the model's prediction at the individual level. During batch processing, it calculates the loss for all samples and takes the average or other summary statistics to comprehensively reflect the model's performance across the entire dataset.

The training curve is a tool used to observe changes in the performance of deep learning models during training. During training, the model iterates through the entire dataset multiple times (each iteration is called an epoch), then adjusts the weights and parameters based on the results of each iteration. The learning curve shows how evaluation metrics such as training error and accuracy change as the number of epochs increases. The training loss curve is used to measure the model's fit to the training data after each epoch. As training progresses, the loss value gradually decreases, indicating the model's learning progress. The accuracy curve is used to monitor the model's generalization ability. By observing changes in accuracy, overfitting can be effectively prevented. Figure 1 shows the training loss curve for the model in this paper.

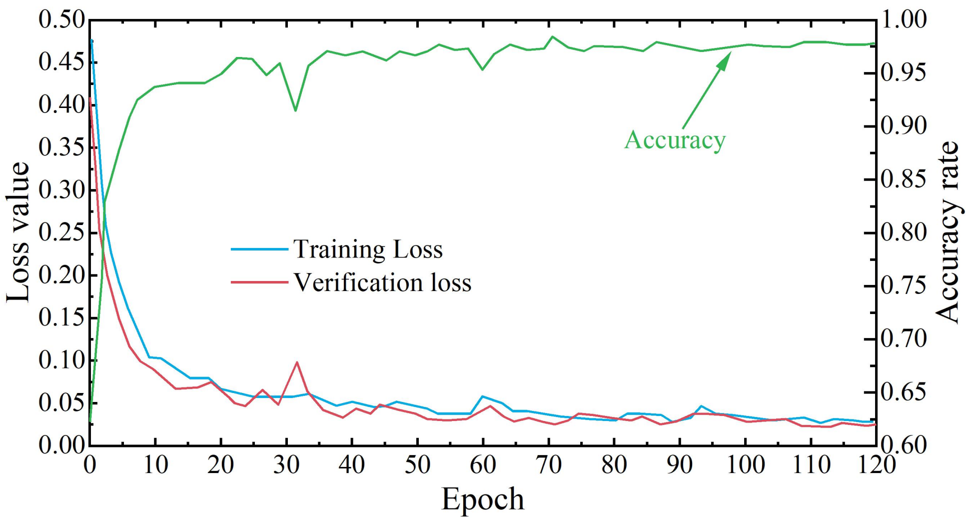

Figure 1. The training loss curve of the model in this paper.

The model completed training within 120 epochs. As shown in the figure, the training and validation losses decreased rapidly within the first 15 epochs and then stabilized, indicating that the model learned quickly in the early stages and gradually converged. The validation loss exhibits minimal fluctuations, indicating that the model performs consistently across different validation sets and possesses strong generalization capabilities. The accuracy curve rises rapidly and stabilizes around 0.975 after approximately 30 epochs, demonstrating the model's robust classification capability for dynamically updating user profiles. These experimental results confirm that the model can accurately capture trends in user interest changes and possesses strong robustness and generalization capabilities.

3.2. Dynamic Updates and Experimental Validation

In exploring the effectiveness of user profiling update methods based on conditional random fields, we designed a series of experiments using real user behavior data. By dividing the user sample set into 28 time windows on a daily basis and applying a sliding window mechanism to achieve incremental learning and dynamic updates, the system can adjust the parameter weights of the conditional random field in real time based on newly added user behavior data and update user profiles using adaptive feature templates. To comprehensively evaluate model performance, we set up four control experiments using four different methods: static classification using support vector machines, long short-term memory networks, traditional conditional random fields, and the adaptive conditional random field method proposed in this paper. Performance was compared across four dimensions: accuracy, precision, recall, and F1 score. The specific experimental data are shown in Table 3. To visually illustrate the performance trends of each model across different time windows, we plotted the accuracy rate as a function of time window, as shown in Figure 2.

Table 3. Performance comparison results of different models.

Model	Accuracy	Precision	Recall	*F*1 score
SVM	78.24	76.89	77.56	77.22
LSTM	83.67	82.91	83.45	83.18
CRF	85.42	84.78	85.13	84.95
ACRF	87.96	87.23	87.58	87.40

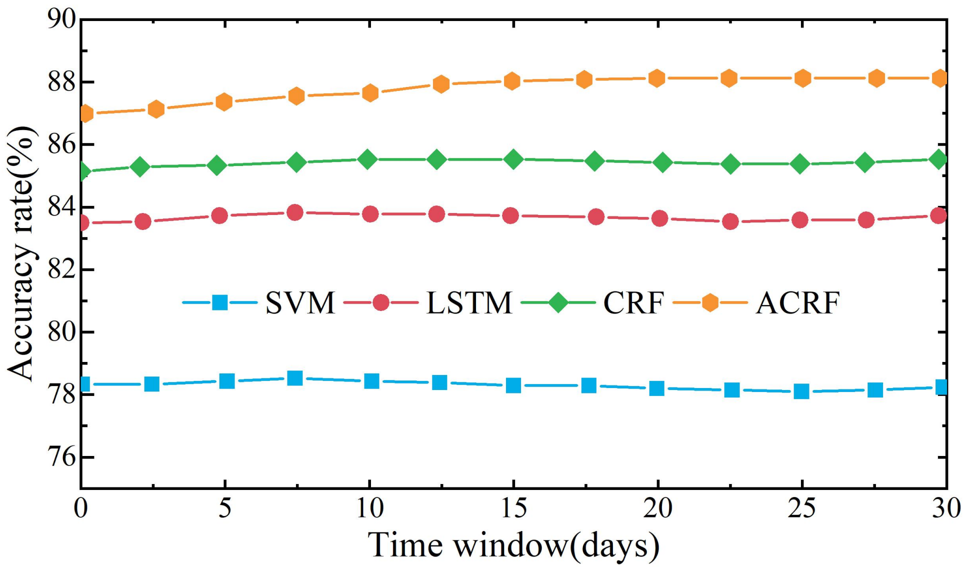

Figure 2. The trend of accuracy changing with the time window.

The experimental data fully validate the superiority of the adaptive conditional random field method proposed in this paper across all evaluation metrics. Compared to the baseline support vector machine model, this method achieves a 9.72 percentage point improvement, and outperforms the long short-term memory network model and traditional conditional random field model by 4.29 and 2.54 percentage points, respectively. This significant performance advantage stems from the enhanced sensitivity of adaptive feature templates to changes in user behavior, the strengthened influence of recent behavioral data through a time-decay-based sample weighting mechanism, and the continuous learning capability of dynamic parameter update strategies for new behavioral patterns. Through in-depth analysis of grouped experiments on different user groups with varying activity levels, it was found that the adaptive conditional random field model performed most notably in high-activity user groups, achieving a prediction accuracy rate of 89.42%. This result reflects the model's outstanding performance in handling high-frequency behavioral data and stable behavioral patterns. Additionally, when faced with sudden changes in user behavior, the model can quickly adjust and reach a new stable state within 2-3 time windows. This adaptability provides important technical support for real-world application scenarios.

4. Conclusion

This chapter summarizes the dynamic user profiling update system based on conditional random fields. The study found that although the model performs excellently in terms of accuracy and adaptability, it still faces certain challenges when dealing with scenarios such as sparse data from low-activity users and excessive feature dimensions. To alleviate cold start and overfitting issues, the study introduced time-decaying weights and regularization mechanisms, and optimized feature template design, effectively enhancing the model's generalization capability and real-time update performance.

Experimental results show that the proposed method outperforms comparison models across multiple metrics, achieving a prediction accuracy of 87.96%, which further increases to 89.42% among highly active users. Overall, the proposed dynamic update framework maintains stability while rapidly adapting to changes in user behavior, providing effective support for personalized recommendations and precise marketing, and demonstrating broad application potential.

Funding
This research was supported by the Teaching Reform Research Project of Ordinary Undergraduate Universities in Hunan Province: Design and Practice of Collaborative Learning Activities for the "Java Programming" Course Based on Student Portraits (Project Number: 202401001871).

References

1. Galbi, D. A. (2001). Some economics of personal activity and implications for the digital economy. *Reading*, *6*(4), 3.
2. Xiong, Y. (2022). Personalized marketing of agricultural products based on digital economy environment. *Academic Journal of Business & Management*, *4*(6), 36-40.
3. Wu, W., Ghazali, M., & Huspi, S. H. (2024). A Review of User Profiling Based on Social Networks. *IEEE Access*.
4. Hui, G., Mang, C., LiQing, Z., & ShiKun, X. (2024). Research on Personalized Recommendation Algorithms Based on User Profile. *Int. J. Adv. Comput. Sci. Appl*, *15*(3).
5. Ouaftouh, S., Zellou, A., & Idri, A. (2019). Social recommendation: A user profile clustering-based approach. *Concurrency and Computation: Practice and Experience*, *31*(20), e5330.
6. Wang, Y., Min, Q., & Han, S. (2016). Understanding the effects of trust and risk on individual behavior toward social media platforms: A meta-analysis of the empirical evidence. *Computers in Human Behavior*, *56*, 34-44.
7. Tang, T., Hämäläinen, M., Virolainen, A., & Makkonen, J. (2011, September). Understanding user behavior in a local social media platform by social network analysis. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 183-188).
8. Lu, W., Yi, Z., Liang, B., Rui, Y., & Ran, B. (2023). Learning car-following behaviors for a connected automated vehicle system: An improved sequence-to-sequence deep learning model. *IEEE Access*, *11*, 28076-28089.
9. Fang, H., Zhang, D., Shu, Y., & Guo, G. (2020). Deep learning for sequential recommendation: Algorithms, influential factors, and evaluations. *ACM Transactions on Information Systems (TOIS)*, *39*(1), 1-42.
10. Zhang, R., Li, W., & Mo, T. (2018). Review of deep learning. *arXiv preprint arXiv:1804.01653*.
11. Bhatt, C., Kumar, I., Vijayakumar, V., Singh, K. U., & Kumar, A. (2021). The state of the art of deep learning models in medical science and their challenges. *Multimedia Systems*, *27*(4), 599-613.
12. Yu, B., & Fan, Z. (2020). A comprehensive review of conditional random fields: variants, hybrids and applications. *Artificial Intelligence Review*, *53*(6), 4289-4333.
13. Li, Y., Li, C., Li, X., Wang, K., Rahaman, M. M., Sun, C., ... & Wang, Q. (2022). A comprehensive review of Markov random field and conditional random field approaches in pathology image analysis. *Archives of Computational Methods in Engineering*, *29*(1), 609-639.
14. Wang, N. (2021, November). Computer intelligent prediction model of internet impulsive reward based on user portrait algorithm. In *2021 IEEE 4th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)* (pp. 545-548). IEEE.
15. Wu, F., Lyu, F., Ren, J., Yang, P., Qian, K., Gao, S., & Zhang, Y. (2023). Characterizing internet card user portraits for efficient churn prediction model design. *IEEE Transactions on Mobile Computing*, *23*(2), 1735-1752.
16. Li, X. Y., Zhang, L. M., & Li, J. H. (2016). Using conditioned random field to characterize the variability of geologic profiles. *Journal of Geotechnical and Geoenvironmental Engineering*, *142*(4), 04015096.
17. Dong, Y., Chu, Q., & Ling, P. (2015, November). Personalized Facet Recommendation based on Conditional Random Fields. In *2015 3rd International Conference on Machinery, Materials and Information Technology Applications* (pp. 207-210). Atlantis Press.
18. Wang, X., Han, L., Li, J., & Yan, H. (2024). A conditional random field recommendation method based on tripartite graph. *Expert Systems with Applications*, 238, 121804.
19. Neogi, Satyajit, et al. "Context based pedestrian intention prediction using factored latent dynamic conditional random fields." *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017.
20. Neogi, S., Hoy, M., Dang, K., Yu, H., & Dauwels, J. (2020). Context model for pedestrian intention prediction using factored latent-dynamic conditional random fields. *IEEE transactions on intelligent transportation systems*, 22(11), 6821-6832.
21. Shen, Y., Yan, J., Yan, S., Ji, L., Liu, N., & Chen, Z. (2011, March). Sparse hidden-dynamics conditional random fields for user intent understanding. In *Proceedings of the 20th international conference on World wide web* (pp. 7-16).
22. Xin, X., King, I., Deng, H., & Lyu, M. R. (2009, November). A social recommendation framework based on multi-scale continuous conditional random fields. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 1247-1256).

23. Baltrušaitis, Tadas, Ntombikayise Banda, and Peter Robinson. "Dimensional affect recognition using continuous conditional random fields." *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. IEEE, 2013.
24. Kim, M. (2016). Mixtures of conditional random fields for improved structured output prediction. *IEEE transactions on neural networks and learning systems*, 28(5), 1233-1240.
25. Song, Dongjin, et al. "Efficient robust conditional random fields." *IEEE Transactions on Image Processing* 24.10 (2015): 3124-3136.