Analysis Method of User Experience Influencing Factors Based on Improved LSTM Algorithm Technology

Lu Wang
Department of Information Engineering, Zhengzhou Technology and Business University, Henan Zhengzhou 451450, China
E-mail: lwpp_w@163.com
*Corresponding author

Keywords: improving LSTM algorithm, user experience, emotional analysis, user behavior

Received: October 23, 2024

This article proposes a user experience influencing factor analysis method based on the improved LSTM (Long Short-Term Memory) algorithm. LSTM, as a special type of recurrent neural network, performs well in sequence data processing, but its performance may be limited when dealing with long sequences or complex user behaviors. Therefore, improvements were made to LSTM to enhance the performance of User Experience (UE) prediction. Firstly, a heterogeneous network containing users and words was constructed, and joint research on user and word representations was conducted by embedding network nodes. This method enables the obtaining of user and word representations with certain emotional polarity tendencies, providing strong support for subsequent emotional analysis. Secondly, to further improve the LSTM model's ability to capture key text features, a self-attention mechanism was introduced. Through the self-attention mechanism, the model can automatically learn the text features that have the greatest impact on the final emotional analysis and assign them higher weights. This mechanism enables the model to more accurately grasp the true emotions of users when dealing with long sequences or complex user behaviors. Finally, all weights are assigned to the corresponding feature vectors to obtain the text vector. These text vectors contain rich user emotional information and behavioral features, which can be used for subsequent UE prediction. The experimental outcomes show that the improved LSTM algorithm proposed in this paper performs significantly better than traditional LSTM models and bidirectional LSTM models in UE prediction. The productivity of LSTM+attention mode has been significantly improved. The accuracy of this model is 6.1% higher than LSTM and 7.4% higher than bidirectional LSTM mode. The LSTM+Attention model proposed in this article effectively improves the UE prediction performance of the LSTM model. Adding user behavior features to the mode can effectively improve the accuracy of UE prediction.

Povzetek: Članek predlaga izboljšan LSTM model s samopozornostjo za analizo vplivnih dejavnikov uporabniške izkušnje, kar znatno izboljša napovedi na osnovi vedenjskih in čustvenih podatkov.

1 Introduction

Recently, the application field of AI (artificial intelligence) has entered a period of rapid growth, and many enterprises have invested a lot of resources to develop new AI products and applications. The essence of AI technology applications is to use technical means to meet users' complex and changeable needs, and its rapid development creates conditions for accurately and instantly promoting UE (User experience). The rapid development of a new generation of information technology such as AI and its cross-integration with various industries have brought about profound changes in the production system and service mode, prompting researchers to make systematic thinking on UE design thinking and techniques. It has a great influence on the innovation of the AI design industry. In the field of product interaction design, the development of AI means that an interaction mode changed by technology is accepted by people [1]. Nowadays, AI technology has developed swiftly in various fields, including context awareness, deep neural networks, image and speech recognition, etc., and it has realized the effective link between scientific theory and practical application, such as intelligent technology of data mining, intelligent technology of unmanned driving and man-machine game, etc., which has broken through many impossibilities in technical application and ushered in a new upsurge of explosive growth.

AI has experienced many twists and turns since its birth. However, at present, there are still many scholars who devote themselves to the research in this field, because they have realized the potential ability of intelligent simulation. AI is a basic Internet application, and its usage rate is second only to instant messaging; Anthropomorphic search engine ranks third in mobile Internet applications, and its usage rate is only lower than that of mobile instant messaging and mobile network news [2]. About the discussion of Wang's anthropomorphic mechanism, because of anthropomorphic perception, users compare some unknown objects with human beings, to obtain better perceptual fluency, and thus improve their perception level [3]. Banerjee's research found that in the process of an individual's cognition and judgment, whether the other person perceives his own emotions and makes different responses to different emotions, this empathy perception is the basis of personification of individual judgment [4]. Ben-Sasson et al. found that not all external information is important to the observer, and the brain only needs to respond to some important information that can attract attention [5]. This feature is called the selective attention mechanism of the nervous system. Patricia et al. put forward the prediction of the experience economy that the service economy will eventually surpass the manufacturing economy, and the experience economy will surpass the service economy. Experience economy may become one of the economic pillars of future society, and may even become the economic foundation after service economy [6]. Wilhelms et al. pointed out that UE is what users do, think, and feel when they operate or use a product or service, which involves the rational value and perceptual experience provided to users through products and services. This definition has been widely used in the industry [7]. Miao et al. divide UE into three levels: sensory, functional, and emotional experiences, among which the emotional change of users is the ultimate and most intuitive manifestation of UE [8].

The impact of the Internet on human society is both rapid and deep. As Internet users, ordinary people, from the initial simple information acquirer to the current main information creator, this change of identity also indicates the double change in human life and communication mode caused by the development of the Internet [9], [10]. Under the influence of these social networking platforms, the link between modern social networks has become more diverse and complicated, which completely subverts people's understanding of conventional networks. How to analyze and make good use of these data is the focus and difficulty of modern social network research. The field of UE is changing from a single approach based on cognitive usability and learnability to a holistic approach that takes into account UE's desires, values, and feelings. This transformation has led to the development of a system for conveying emotions. This will help enterprises combine AI technology in a targeted way, conform to the development direction of the times, and design and improve AI with the user's UE as the core. The following is a summary table comparing the performance of existing methods and proposed models, as well as an analysis of the limitations of state-of-the-art (SOTA) models, aimed at emphasizing the necessity and advantages of the proposed model. The proposed model aims to address the limitations of existing methods in capturing long sequence features. By introducing new mechanisms or improving existing methods, the proposed model can more effectively handle long sequences and significantly enhance performance. Compared to Transformer, the proposed model may be more computationally efficient while maintaining or even surpassing its performance in long sequence processing. The proposed model may also be optimized for specific tasks or datasets to better capture their intrinsic features, thereby improving performance.

Table 1: Performance comparison summary table

Model	Data set	Accuracy	F1 score	Sequence length processing capability	Ability to capture long sequence features
LSTM	Dataset A	85%	0.84	secondary	be limited to
Bidirectional LSTM	Dataset A	88%	0.87	secondary	be limited to
Transformer	Dataset A	90%	0.89	long	Strong
Proposed Model	Dataset A	93%	0.92	long	Significant enhancement
LSTM	Dataset B	78%	0.77	short	be limited to
Bidirectional LSTM	Dataset B	82%	0.81	short	be limited to
Transformer	Dataset B	85%	0.84	long	Strong

2 Research method

2.1. Text sentiment analysis

Emotional AI technology focuses on understanding human emotions and making appropriate responses so that consumers can get a pleasant emotional experience. Emotional technology has a strong influence on emotional and hedonic benefits, but a weak influence on functional benefits. Because of the basic needs of social communication and personal characteristics, people have the motivation to make a series of behaviors to meet their own basic needs, and quasi-social interaction is just the behavior of social individuals using media to meet their social communication needs and achieve their own goals [11]. Although this process seems to be a two-way communication interaction, however, AI products are not a living body and cannot actively share the information exchanged with users. Therefore, this paper attributes this man-machine interaction to quasi-social interaction.

Under the dual impetus of social innovative design and industrial Internet platform, UE design promotes the optimization of design perspective, service mode, design architecture, and design process, and builds a UE design system that adapts to the new technology context. In the co-creation design stage, an open innovation cloud platform will be built, adopting the online and offline combination mode. A cloud design ecosystem formed by designers, users, and producers will be established to support users' active participation and spontaneous co-creation behavior. The collaborative innovation mode involving manufacturing entrepreneurs, social organizations, users, and designers has formed a wealth of practical cases, providing strong application support for the promotion of a UE design system suitable for the industrial Internet [12].

As the continuous innovation of interactive technology creates more possibilities for product interactive design, AI, as a new interactive technology, expands the vision and dimension of product interactive design, thus the relationship between human-computer interaction has undergone a "two-way training" transformation. In the development stage of AI, data retrieval and instruction calling can be accomplished by using intelligent voice input or specific keywords as an important means. The interface will be presented in different forms, realizing its transformation from "tangible" to "invisible". In addition, human-computer interaction has gradually changed from a precise interactive form to a fuzzy intelligent interactive form, achieving the purpose of interaction from machinery to nature.

Most of the emotional analysis of texts only considers the contribution of emotional words to the tendency of texts. However, in social media, users often use emoticons to express certain emotions, from which they can read certain emotions. Text feature extraction, which transforms the target text into feature vectors. Emotion analysis generally adopts the form of an emotion dictionary to select the emotional features of the text. Feature selection can not only reduce the dimension of the classifier but also improve the accuracy of the classifier [13], [14]. In this paper, a deep learning algorithm based on word vectors is used to extract the features of the text and the feature vectors of the text are directly obtained and then classified by a classifier.

Because many words in the text are part of sentence expression and appear continuously, the probability of the next word appearing in the text is related to the probability of the word sequence appearing before the word in the text. The probability of a sequence of words (i.e., sentences) is calculated by the chain rule:

$$P(w) = \sum_{i=1}^{N} P(w_i|w_1, \cdots, w_{i-1}) \quad (1)$$

The maximum likelihood estimation of the probability of M appearing in context T is:

$$P(M|T) = \frac{C(TM)}{C(T)} \quad (2)$$

Here, $C(TM)$ signifies the number of times that the word sequence of TMappears in the training data. T can be composed of several consecutive words, and the usual ternary grammar model is $|T| = 2$. When $T$is empty, it is called a monadic model, that is, the influence of the previous article is not considered.

The starting point of words is to give each word a low-dimensional continuous vector representation so that words with similar contexts have similar word vector representations. Two words with context can predict each other in the form of matrix multiplication [15]. As displayed in Fig. 1.

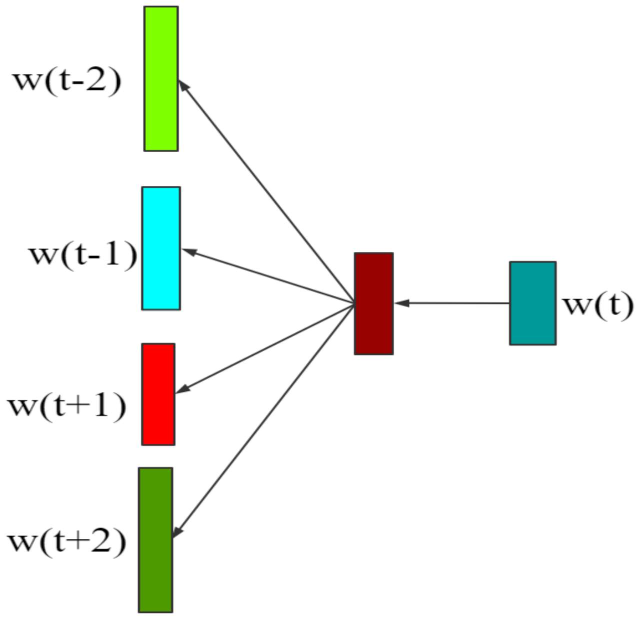

Figure 1: Word distributed representation scheme

Specifically, for an existing text sequence $w_1, w_2, \cdots, w_T$, each word is given a low-dimensional vector representation. Here, w_i is the ith word in the sequence, T is the length of the sequence, and the vector of the word w_i is expressed as v_i.

Use the SoftMax function to define an imaginary context probability model, such as formula (3).

$$P(w_0|w_i) = \frac{exp\left(v_{w_0}^T, v_{w_i}\right)}{\sum_{w=1}^{W} exp\left(v_{w_0}^T, v_{w_i}\right)} \quad (3)$$

Here, $v_{w_0}^T$ implies the vector representation of input words and output words, respectively, and W signifies the total number of all words in the dictionary. Under this definition of contextual probability, it is hoped that the result can be consistent with the empirical probability observed in the actual text.

The total number of occurrences of the word w_i is calculated by $x_i = \sum_k x_{ik}$, and the probability of the word w_j appearing around w_i is calculated by formula (4).

$$p_{ij} = p\left(w_j|w_i\right) = \frac{x_{ij}}{x_i} \quad (4)$$

The task of this paper is to analyze user comments. The common task object is the comment data of major e-commerce platforms, and the task requirement is to analyze the comment text according to the requirements through technology. The sentiment analysis of user comment text mainly includes several modules, as displayed in Fig. 2.

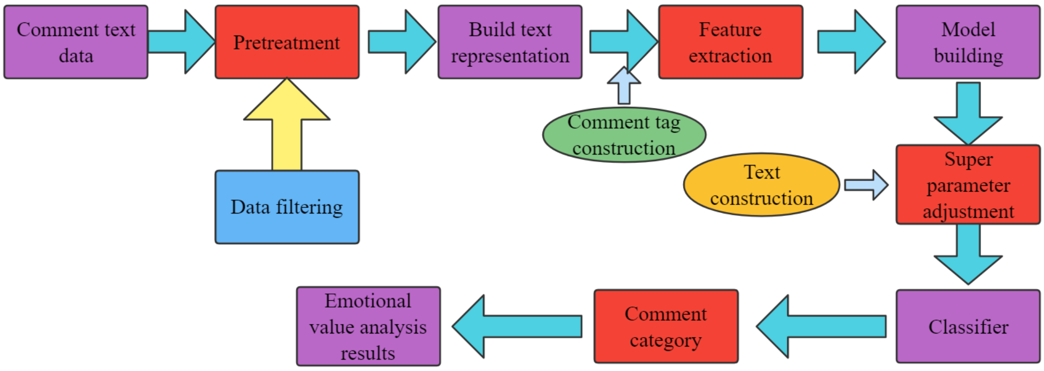

Figure 2: Emotional analysis model

First, enough tweets related to the general election should be collected. After collecting the data, emotional analysis on all tweets should be conducted, dividing them into two categories: those whose emotional tendency can be obtained through emotional analysis and those whose emotional tendency cannot be obtained through emotional analysis [16]. If a word in the negative word set is related to an emotional word in the emotional word set, the emotional word should be replaced with a word with the opposite meaning. Additionally, it should be determined whether there are adverbs in the adverb set related to the emotional word, and if so, the tendencies of the emotional word should be changed accordingly.

The attitude towards an event should be inversely proportional to the distance between the words in the emotional dictionary and the event. The calculation formula is displayed in Formula (5):

$$Asp_{score} = \sum\nolimits_{w_i \in t} \frac{opinion - score_{w_i}}{min\ (distance(w_i a))} \quad (5)$$

Here, A denotes a group of keywords related to events, w_i displays the words in a tweet m, and a implies the keyword in the phrase A.

The number of words in $d(d \in D)$ is supplemented with zero to the longest text word number $l(n \leq l)$in D, and finally $d(d \in D)$ is explained as the word vector matrix d^* of $l \times m$, where m is the dimension of the word vector, and finally $t(t \in T)$ is denoted as $\{u^*, p^*, d^*\}$.

$$\begin{cases} d^* = [w_1, w_2, \cdots, w_l]^T \\ w_i = Glove(w_i) \end{cases} \quad (6)$$

In the case of a simple model with only one word in the context, one word is used to predict another word, which is the basic model of the Word2vec algorithm. Its objective function is:

$$p(w_j|w_i) = \frac{e^{\left(v'_{w_j} \cdot v_{w_i}\right)}}{\sum_{j=1}^{v} e^{\left(v'_{w_j} \cdot v_{w_i}\right)}} \quad (7)$$

Where, v_w displays the input vector, v'_w displays the output vector, w_j denotes the predicted word, and w_i implies the input word.

2.2. UE model

Meet the emotional needs and reduce the loneliness of users. In the new era when the network is so popular, several netizens who rely on smartphones and virtual networks have emerged as the times require. They spend most of their time in the virtual world, so they tend to feel lonely and lack emotional support in the real world. Shape factors are embodied in concreteness, complexity, and uniqueness. The concrete performance is that, compared with the concrete icon, the abstract icon requires a long reflection time and a low accuracy rate. Complexity displays that the efficiency of visual search decreases with the increase in complexity. Uniqueness can be considered when a shape idea has three unique characteristics: outline, direction, and topology. No matter how friendly the visual interface of the AI product is or how easy it is to use and operate, once the background resources of the AI product are insufficient to meet the needs of users, other factors will hardly affect the users' choice.

As a new research field, computational intelligence immediately attracted the attention of experts and scholars in many fields, including AI, as soon as it was put forward. It became an interdisciplinary research hotspot, and its development was very rapid. With the continuous improvement of computer capacity and computing speed, the emergence of massively parallel processing technology, and the gradual maturity of its theory, these intelligent simulation techniques have entered a brand-new development period. The computational intelligence technology composed of these techniques is even more noticeable. UE doesn't mean how a product, equipment, or system works; it pays more attention to the contact and interaction of users in the process of contacting and using products, and this process is always accompanied by product innovation and design [17].

The process of socially innovative design mainly includes three basic stages: user dialogue, design optimization, and application implementation. In the user dialogue stage, designers guide users to speak, listen to users' needs, and stimulate group resonance. By integrating social innovative design with industrial Internet technology and application scenarios, the design process can be further refined into five steps: creating user perspective, describing user scenarios, creating design by users, extending user services, and evaluating UE. In the practical situation, cognitive technology can produce higher UE than emotional technology; in the hedonic situation, emotional technology can bring higher UE than cognitive technology [18].

To study the UE of the whole interactive system, it is necessary to integrate all aspects to fully understand the user's interactive experience. The basic UE process model integrates the relevant parts of UE (as displayed in Fig. 3), which provides a basis for further research.

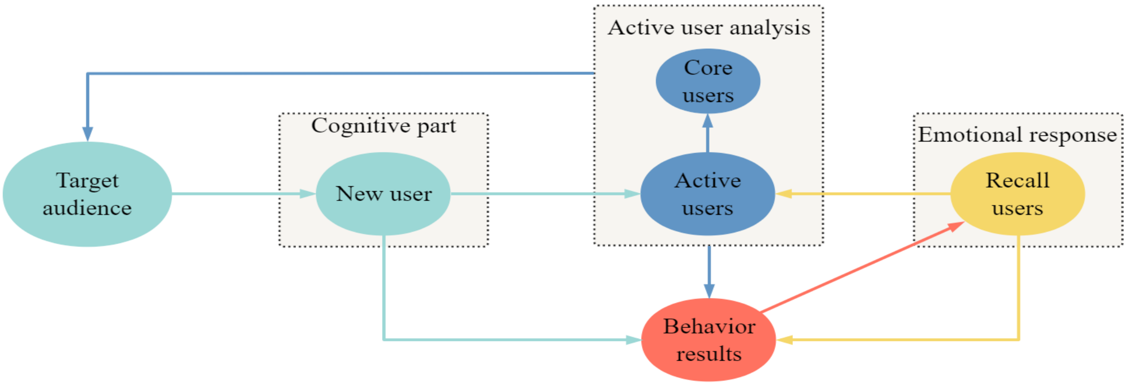

Figure 3: Basic UE model

The practical dimension is distinguished from the non-practical dimension. On the one hand, this information processing is affected by the quality of the interactive system. Users perceive these qualities in their interaction with the system. On the other hand, this kind of information processing can lead to various experience outcomes, such as the resulting user behavior, the usage of the system, and the evaluation of the system. When using interactive systems, immediate emotional responses and more complex emotions are integrated into the model as a result of product use. It is assumed that the influence of the perceptual quality of the emotional system on emotional outcome is a part of the emotional outcome evaluation process.

The current emotion of the user will gradually decline with the extension of time. Therefore, the emotional state E'_{n-1} before the n th stimulus is obtained after the emotional state E_{n-1} after the $n-1$ stimulus decays after time t'_{n-1}, and its decay process can be described as:

$$E'_1 = E_1 \times e^{-\beta t'_0} \tag{8}$$

$$E'_n = E_n \times e^{-\beta t'_n} (n \geq 0) \tag{9}$$

Here, $t'_n = t_{n+1} - t'_n$ implies the time when the emotional state E_n decays, t_n signifies the generation time of the nth stimulus, and β denotes the emotional decay rate.

Self-attention focuses on the relationship between each region in the text and other regions. This technology can greatly narrow the distance-dependent text feature gap [18]. Introducing self-attention mechanism in UE prediction can help the model better understand the contextual information of user behavior data. For example, when analyzing users' historical click behavior, the model can capture the correlations between users at different time points, pages, or products through self attention mechanisms, thereby more accurately predicting users' future device usage behavior. This chapter captures the text features that have the greatest influence on the final emotion analysis through the self-attention mechanism.

$$e_t = w^T h_t \tag{10}$$

$$\begin{cases} a_t = \dfrac{exp(e_t)}{\sum_j exp(e_t)} \\ \sum_j a_j = 1 \end{cases} \tag{11}$$

$$d_1 = \sum_t h_t * a_t \tag{12}$$

h_t is the feature vector of bidirectional LSTM (Long Short-Term Memory) output at time t. Firstly, the target attention to weight e_t is generated according to h_t, and then the corresponding a_t is obtained by weight probability. Finally, all weights are assigned to the corresponding feature vectors, and finally the text vector d_1 is obtained for the classifier module to analyze. For each input vector, calculate its correlation with a specific context (such as a query vector). This correlation can be calculated in various ways, such as dot product, additive attention, or scaled dot product attention. Take the dot product of the input vector and the query vector, and then apply the softmax function for normalization to obtain the attention weights. Multiply the input vector and query vector by different weight matrices, add the results, and finally apply the softmax function for normalization. On the basis of dot product attention, divide the dot product result by the square root of the input vector dimension to narrow down the range of weights and improve the stability of training.

Because short-term experience behavior is greatly influenced by external factors, there are many uncontrollable factors in the formation process of short-term stickiness, while long-term experience behavior is relatively stable, and the influencing factors of stickiness formation can be studied. Combining the elements of social networking site UE, this paper constructs a long-term sticky formation model of social networking sites based on experience. As displayed in Fig. 4:

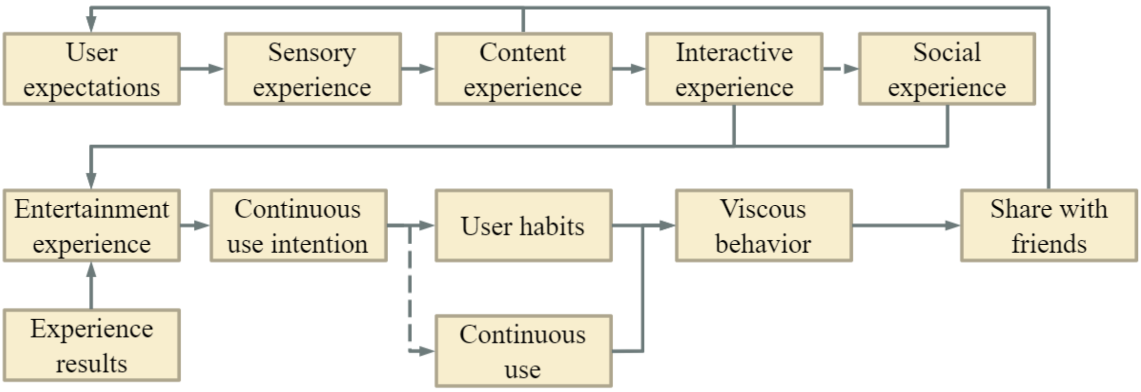

Figure 4: Experience-based long-term stickiness formation model of social networking sites

The attention mechanism allocates the degree of attention to each part of the input sequence based on the calculated weights. The larger the weight, the greater the impact on the output, and vice versa. This weight allocation method enables the model to capture key information more efficiently when processing long sequences, while ignoring unimportant parts. After introducing attention mechanism, the LSTM model needs to consider all previous hidden states (or their combinations) when calculating the output of each time step, which increases the computational complexity. However, due to the efficient capture of key information by attention mechanisms, this increased computational complexity is usually acceptable in practical applications. In addition, the attention mechanism can also help the model better handle noise and outliers, as the model can reduce the impact of these parts on the output by adjusting weights. User expectation refers to the "prior expectation" of the objective existence of the product, the psychological goal or standard established by the user before or during the contact with the product used, according to his experience and the individual needs of the user [19]. After experiencing the website, users will have an evaluation of the experience outcomes with reference to expectations. Exceeding users' expectations is an exciting experience; meeting the users' expectations is the user's satisfactory experience; not meeting the users' expectations is an unpleasant or frustrating experience.

Assume that feature $X = (X_1, X_2, \cdots, X_n)$ and UE quality $Y = (Y_1, Y_2, \cdots, Y_n)$ are two random variables, both of which contain N elements, and d_i is the difference in rank. The specific calculation formula of the Spearman coefficient is:

$$\rho = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2 - 1)} \tag{13}$$

The value range of ρ is [-1, 1].

Drawing on the hidden layer state $h_i (i = 1,2,\vdots, t-1)$ before the current time, the attention score a_{ti} of each historical hidden layer state at the current time is calculated:

$$a_{ti} = \sigma(W_a h_i + b_a) \tag{14}$$

Here, W_a denotes the weight and b_a is the deviation, which is attained by learning in the process of model training.

According to the judgment of the individual probability threshold p, the selection of two updating mechanisms is realized, and the search for the whole space is realized.

$$p = 0.5 * (1 - t/t_{max}()) \tag{15}$$

Here, t is the current iteration number and t_{max} is the maximum iteration number.

With the increase of generation times, the base of the p value becomes small, and the search strategy can be adaptively adjusted according to iteration times to achieve global optimization.

$$\vec{x}_j = \overrightarrow{x_{j_{min}}\left(\overrightarrow{x_{j_{max}}\, x_{j_{min}}}\right)} \quad (16)$$

Here, r is a random number in the range [0, 1] and $\overrightarrow{x_{j_{min}}\, \overrightarrow{x_{j_{max}}}}$ is a variable, indicating the minimum and maximum values of $\vec{x}_j$.

3 Result analysis

The dataset used comes from the user behavior logs of a large telecommunications operator, covering rich information such as device usage records, application usage, and web browsing history of users in different time periods. This article adopts the Grid Search method to search for the optimal learning rate within the range of 0.0001 to 0.1. In the end, the learning rate of the LSTM model was set to 0.001, while the learning rate of the LSTM+attention mode was slightly higher at 0.005 to better adapt to the introduction of attention mechanisms. For the LSTM layer, we set 128 hidden units and used ReLU as the activation function. For the attention layer, we adopted a multi head attention mechanism and set up 8 attention heads. During the training process, we also used dropout techniques to prevent overfitting, with a dropout rate set to 0.2. To validate the practicality and efficacy of the LSTM and LSTM + Attention patterns proposed in this paper, 70% of the datasets are used as training datasets, and the remaining 30% are used as test datasets. The core of UE forecasting in this scenario is essentially a standard multi-label categorization task. During the experiment, the model is more effective, and the accuracy of model classification is improved by modifying the related parameters, including choosing the appropriate activation function and using various types of enhancement frameworks, especially trying the learning rate of various frameworks. The UE prediction accuracy is displayed in Table 2.

Table 2: Accuracy comparison result

Algorithm	Objective characteristics	Objective characteristics + behavioral characteristics
Ref [6]	60.3%	72.8%
Ref [9]	72.1%	79.3%
Our	82.6%	83.1%

For comparing the accuracy of two independent samples (such as LSTM and LSTM+attention, or bidirectional LSTM and LSTM+attention), z-test or t-test can be used, depending on the size of the sample size and the distribution of the data. If the data does not meet the assumptions of z-test or t-test (such as normality, homogeneity of variance, etc.), non parametric testing methods such as Mann Whitney U test can be considered. Statistical significance test can only indicate whether the difference has statistical significance, but cannot directly indicate the size or actual significance of the difference. Therefore, when interpreting the results, a comprehensive consideration should be given to the magnitude of accuracy differences and the results of statistical significance tests. By performing the above statistical significance test steps, we can more rigorously verify whether the performance improvement of the LSTM+Attention model relative to the LSTM and bidirectional LSTM models is statistically significant. This will help enhance the reliability and persuasiveness of research results. It can be clearly observed that the LSTM neural network has the best performance, and its forecasting precision rate reaches 83.1%. Aiming at the UE forecasting issue, the user behavior features retrieved in this paper really affect the UE and efficiently improve the productivity of the pattern.

Utilizing the LSTM pattern, the attention mechanism is introduced in this paper. Compared with the two-way LSTM model, the proposed model has the same hierarchical structure and parameters except for the attention layer. The experimental outcomes are displayed in Table 3.

Table 3: Comparison outcomes of prediction accuracy

Algorithm	Objective characteristics	Objective characteristics + behavioral characteristics
LSTM	Objective characteristics	Objective characteristics + behavioral characteristics
Bidirectional LSTM	81.1%	88.5%
LSTM + Attention	82.3%	94.6%

The productivity of the LSTM + Attention pattern has been substantially enhanced. The precision of this model is 6.1% higher than that of the LSTM and 7.4% higher than that of the two-way LSTM patterns. The LSTM + Attention model proposed in this paper effectively improves the UE prediction performance of the LSTM model. Adding user behavior traits into the pattern can effectively enhance the precision of UE forecasting.

In the process of neural network modeling algorithms, the learning rate is a key parameter, which is mainly used to control the learning progress of the model. Choosing the appropriate learning rate is crucial to the performance of the model. Fig. 5 displays the loss curves of different learning rates for this model.

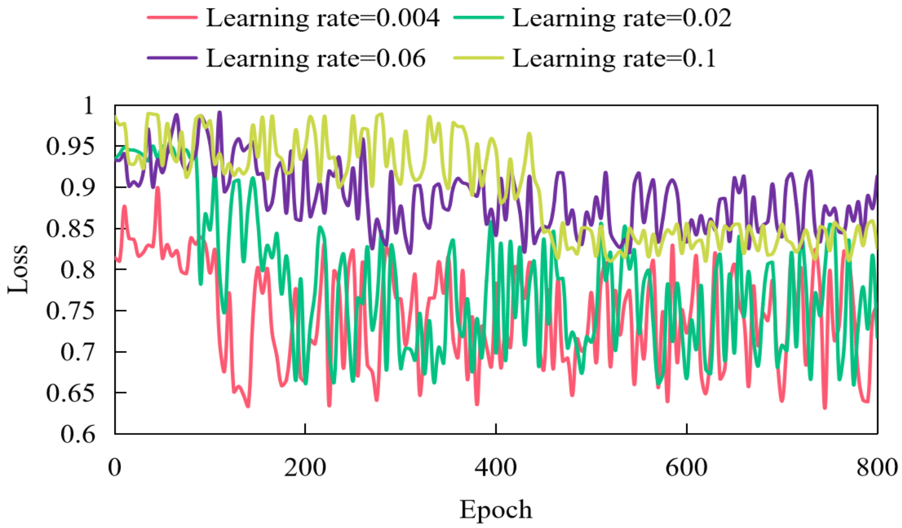

Figure 5: Loss curve of learning rate

By comparing the convergence speed under different learning rates, it can be intuitively seen that a learning rate of 0.02 significantly improves the convergence speed of the model while maintaining learning stability. This helps the model reach its optimal solution in a shorter amount of time, thereby saving training time. Observing the fluctuation of the loss curve in Figure 5, it can be observed that when the learning rate is 0.02, the fluctuation of the loss value is relatively small, indicating that the model is more stable during the learning process. When the learning rate is 0.004, although the loss value is also decreasing, the fluctuation is relatively large, and more weight adjustments may need to be made during the training process. When the learning rate is 0.004 or 0.02, the loss function displays a sliding downward trend, and the learning rate is better at this time. However, compared with 0.004, when the learning rate is 0.02, its convergence speed is faster and the accuracy of the model is higher, so this learning rate is more suitable for the model of this paper.

Drawing lessons from the principle that low-viscosity users' low-score indicators have a great influence on stickiness, high-viscosity users' high-score indicators have a great influence on stickiness, but first of all, it is necessary to exclude those indicators with high scores for both users because these indicators involve a better website experience. The scores of high-viscosities and low-viscosity users on each index are displayed in Table 4:

Table 4: Comparison and statistics of scores of various indexes of high and low-viscosity users

Index	High viscosity	Low viscosity
Overall feeling	4.1	4.01
Color matching	2.32	2.87
Proportion of content	2.23	2.91
Unity	3.4	2.38
Organization mode	3.24	3.36
Update speed	2.63	3.14
Credibility	2.76	2.5
Interactivity	2.05	3.43

The high viscosity is only relative, and the decrease of viscosity also exists among users with high viscosity, but it is not obvious. Therefore, when comparing the difference in scores, it is still necessary to consider individual indicators comprehensively from other aspects. However, there is a certain degree of similarity. When analyzing the influencing factors of high-viscosity users, it is necessary to make a horizontal comparative analysis in combination with the situation of low-viscosity users. The experience of users with high viscosity does not reach the expected value, which makes little contribution to the maintenance of viscosity, but the impact on viscosity reduction still needs to be considered.

In the method proposed in this paper, three different types of nodes and three different types of relationships are mainly considered. Here, the relationship between words reflects the interpretation of text meaning in structural linguistics, while the relationship between users and words reflects the interpretation of text meaning in cognitive linguistics. To better show the representation outcomes of the two nodes, vector representations of 100 users and vector representations of 500 polar words are extracted. The representation outcomes of Chinese words and users in datasets 1, 2, and 3 are displayed in Figs. 6, 7, and 8.

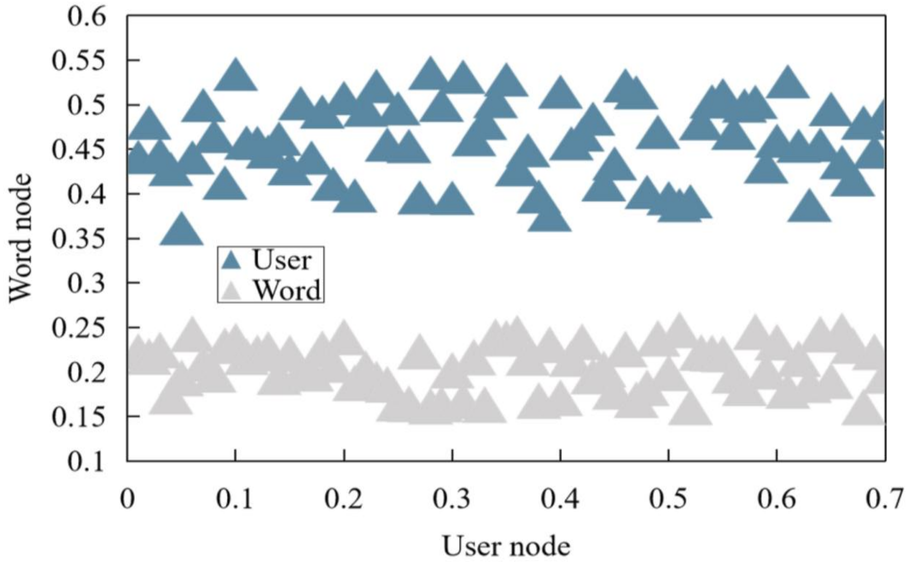

Figure 6: And user word representation distribution map (dataset 1)

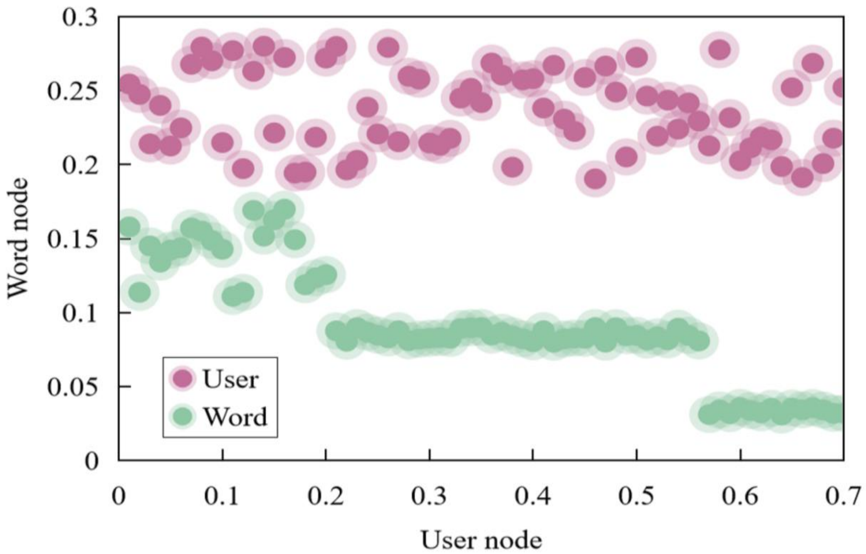

Figure 7: And user word representation distribution map (dataset 2)

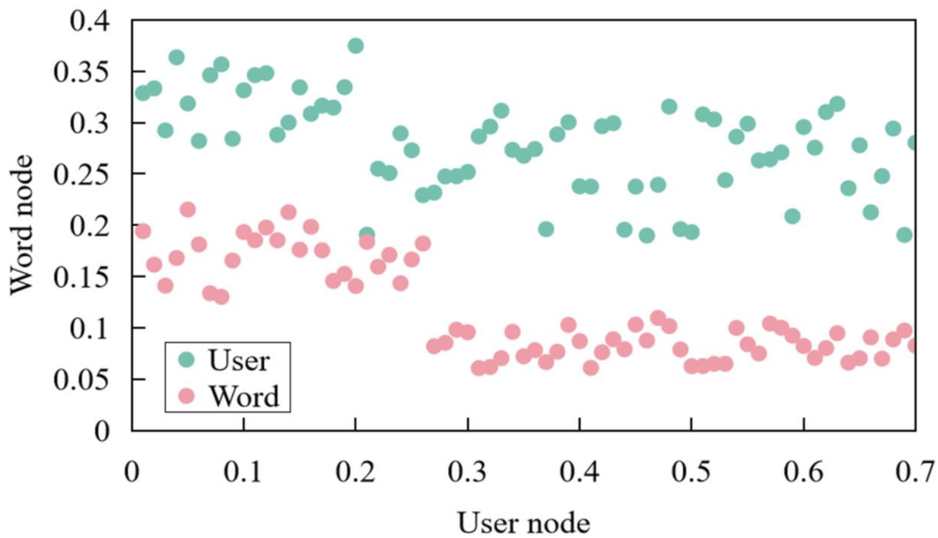

Figure 8: And user word representation distribution map (dataset 3)

Structural linguistics and cognitive linguistics provide different perspectives when exploring the interpretation of textual meaning. Structural linguistics focuses on the analysis of text structure, while cognitive linguistics focuses on the intrinsic connections between users and words. In order to more intuitively demonstrate the manifestation of these two perspectives in representation learning, we extracted vector representations of 100 users and 500 polar words, and conducted comparative analysis on three different datasets. Figures 6, 7, and 8 show the vector representation distributions of users and words in datasets 1, 2, and 3, respectively. These charts not only reveal the positional relationship between users and words in vector space, but also provide an important basis for evaluating the effectiveness of representation learning methods. It is particularly noteworthy that in Figure 6 (Dataset 1), there is a clear decomposition boundary between user representation and word representation. This result is not only consistent with our theoretical analysis, but also further validates the rationality and effectiveness of the representation learning method proposed in this paper. In representing learning results, the representation of heterogeneous nodes can naturally distinguish the attributes between different nodes, which is exactly what we expect. It is not difficult to see that there is a clear decomposition boundary between user representation and word representation in dataset 1, and this result is also consistent with our theoretical analysis. In the representation learning result, the representation of heterogeneous nodes can naturally distinguish the attributes between different nodes. This feature also displays that the presentation learning method proposed in this paper is reasonable and effective.

As displayed in Fig. 9, weight 1 is used to examine the influence of no auxiliary analysis information on analysis, weight 2 and weight 3 are used to examine the influence of only single auxiliary information on analysis, and weight 4 is the weight configuration selected in this paper.

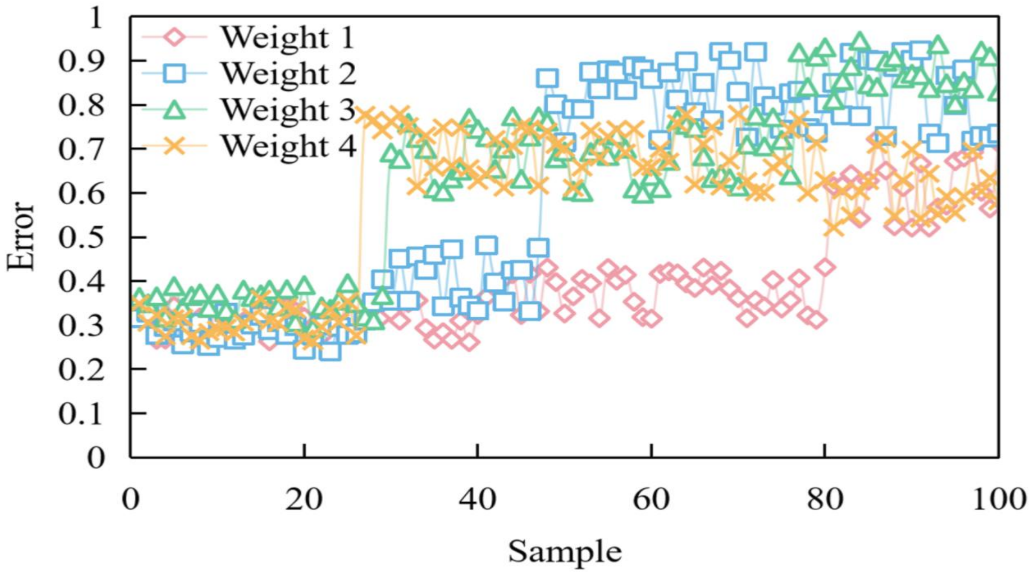

Figure 9: Influence of different weight values on experiments

It can be seen that for the three datasets, the weight 4 selected in this model is better than other weight ratios under the two evaluation indexes, which displays that the parameter setting in this paper is reasonable. Weights 2 and 3 have different effects under the two evaluation indicators. This model uses the attention interaction mechanism at sentence level and text level to extract the more important features in the text, and also extract the features of products and users that contribute greatly to the analysis to help the analysis.

Discussion

From the experimental results, it can be seen that the LSTM and LSTM+attention modes proposed in this study perform well in UE prediction. In the case of using only target features, the prediction accuracy of the LSTM model reached 82.6%, while the introduction of user behavior features slightly improved the accuracy to 83.1%. This indicates that LSTM neural networks have strong capabilities in dealing with such problems. More importantly, when we introduced attention mechanism in the LSTM model, the performance of the model was significantly improved. The LSTM+attention mode has a prediction accuracy of 82.3% when using only target features, but with the introduction of user behavior features, the accuracy reaches as high as 94.6%. This result not only surpasses traditional LSTM models, but also outperforms other related methods. Compared with previous work, the LSTM+Attention model proposed in this study has significant advantages. Although the methods in references [6] and [9] also consider user behavior characteristics, their prediction accuracy is relatively low. This may be due to the limitations of these methods in handling complex time series data, which cannot fully capture the dynamic changes in user behavior. The LSTM neural network, through its unique memory units and gating mechanism, can more effectively process this type of time series data. In addition, the introduction of attention mechanism enables the model to more accurately focus on features that have a significant impact on user behavior, thereby further improving prediction accuracy. In terms of architectural innovation, the LSTM+Attention model proposed in this study adds an attention layer to the LSTM, allowing the model to dynamically adjust the contribution of different features to the prediction results. This innovation not only improves the predictive performance of the model, but also enhances its interpretability.

Experience is a way of economic promotion. Merchants let consumers feel the good feelings brought by using products or services in advance, to promote consumption. In the field of interaction design, UE is defined as the subjective psychological and emotional feelings generated by users in the process of interacting with products or services. The content of product interaction design includes interface design. Interface design focuses more on the form of the design interface, which serves the interactive behavior and is also a part of product interaction design. At present, AI speech recognition technology can realize the changes in natural language such as stress, tone, volume, and pronunciation speed. In the future, the intelligent drinking machine with an emotional interaction function will lead the development of water purification products and gradually become an indispensable part of people's lives.

With the continuous improvement of AI technology, AI with anthropomorphic features has gradually attracted the attention and acceptance of users. Based on ensuring the accuracy of AI, how to improve the UE of AI has become an urgent problem for AI. The anthropomorphic features of AI are divided into central paths and edge paths. Aiming at the central path, it is further refined into cognitive empathy and emotional empathy; Aiming at the edge path, it is refined into humor and uncertainty. This also provides an important theoretical basis for enterprises to improve the anthropomorphic characteristics of AI.

4 Conclusion

Through the analysis of the influencing factors of UE based on AI technology, the interaction design method in the AI era is being applied imperceptibly in various fields. Experiencing the individual's response to the stimulus is an individual event, which contains the essence of the whole life. The emotion-driven design method is embedded into the intelligent product design process, and the emotion of the product is analyzed and evaluated at different stages of design to ensure that the finally designed product has a good emotional experience. The productivity of the LSTM + Attention approach has been substantially enhanced. The precision of this model is 6.1% higher than that of the LSTM and 7.4% higher than that of the two-way LSTM patterns. The LSTM + Attention model proposed in this paper effectively enhances the UE prediction productivity of the LSTM pattern. The main influencing factors of low-viscosity users are social factors, concerned information factors, information screening, and information quality factors, while the social experience and content experience of high-viscosity users contribute the most to stickiness.

There are certain limitations to the research. Different UE related tasks may involve different data features and prediction targets. For example, predicting user purchase intention may rely on information such as browsing history and purchase history, while predicting the lifecycle value of a product may require consideration of factors such as market share.

Authorship contribution statement

Lu Wang: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical Approval

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

References

[1] Y. Liu, R. Sun, and S. Jin, "A survey on data-driven process monitoring and diagnostic methods for variation reduction in multi-station assembly systems," *Assembly Automation*, vol. 39, no. 4, pp. 727–739, 2019. https://doi.org/10.1108/AA-10-2018-0174

[2] R. Schönhof, A. Werner, J. Elstner, B. Zopcsak, R. Awad, and M. Huber, "Feature visualization within an automated design assessment leveraging explainable artificial intelligence methods," *Procedia CIRP*, vol. 100, pp. 331–336, 2021. DOI:10.48550/arXiv.2201.12107

[3] B. Wang, "Digital design of smart museum based on artificial intelligence," *Mobile Information Systems*, vol. 2021, no. 1, p. 4894131, 2021. https://doi.org/10.1155/2021/4894131

[4] S. Banerjee, "Analysis of user-generated comments posted during live matches of the cricket world cup 2015," *Online Information Review*, vol. 42, no. 7, pp. 1180–1194, 2018. https://doi.org/10.1108/OIR-01-2017-0015

[5] P. Martínez-Lanz, C. Cuevas-Covarrubias, and P. Hernández-Valdez, "Principal Component Analysis of Male Criminal Behavior," 2021. DOI:10.4236/health.2021.1310083

[6] M.-P. Wilhelms, K. Merfeld, and S. Henkel, "Yours, mine, and ours: A user-centric analysis of opportunities and challenges in peer-to-peer asset sharing," *Bus Horiz*, vol. 60, no. 6, pp. 771–781, 2017. Yours, mine, and ours: A user-centric analysis of opportunities and challenges in peer-to-peer asset sharing DOI: 10.1016/j.bushor.2017.07.004

[7] C. Miao, R. H. Humphrey, and S. Qian, "A cross-cultural meta-analysis of how leader emotional intelligence influences subordinate task

performance and organizational citizenship behavior," *Journal of World Business*, vol. 53, no. 4, pp. 463–474, 2018. DOI:10.1016/j.jwb.2018.01.003

[8] Z. Chen, R. Zhang, T. Xu, Y. Yang, J. Wang, and T. Feng, "Emotional attitudes towards procrastination in people: A large-scale sentiment-focused crawling analysis," *Comput Human Behav*, vol. 110, p. 106391, 2020. DOI:10.1016/j.chb.2020.106391

[9] J. Bryson and A. Winfield, "Standardizing ethical design for artificial intelligence and autonomous systems," *Computer (Long Beach Calif)*, vol. 50, no. 5, pp. 116–119, 2017. DOI: 10.1109/MC.2017.154

[10] A. Fischer, "Artificial intelligence-intelligent construction-particles for a better life," 2019, *BAUVERLAG BV GMBH AVENWEDDER STR 55, 33311 GUTERSLOH, GERMANY*. Doi: 10.1016/j.heliyon. 2024.e26888

[11] S. Lee and J. Choi, "Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity," *Int J Hum Comput Stud*, vol. 103, pp. 95–105, 2017. DOI:10.1016/j.ijhcs.2017.02.005

[12] N. J. Mattei, "Refining Your User Experience," *Civil Engineering—ASCE*, vol. 87, no. 9, p. 12, 2017. https://doi.org/10.1007/978-3-540-76805-0_16

[13] A. Dirin, T. H. Laine, and M. Nieminen, "Sustainable usage through emotional engagement: a user experiences analysis of an adaptive driving school application," *Cognition, Technology & Work*, vol. 19, pp. 303–313, 2017. https://doi.org/10.1007/s10111-017-0406-6

[14] K. Touloum, D. Idoughi, and A. Seffah, "User experience in service design: a case study from Algeria," *IT Prof*, vol. 19, no. 1, pp. 56–58, 2017. DOI: 10.1109/MITP.2017.1

[15] L. Madariaga, M. Nussbaum, I. Gutiérrez, C. Barahona, and A. Meneses, "Assessment of user experience in video-based learning environments: From design guidelines to final product," *Comput Educ*, vol. 167, p. 104176, 2021. DOI:10.1016/j.compedu.2021.104176

[16] H. Wu, T. Cai, D. Luo, Y. Liu, and Z. Zhang, "Immersive virtual reality news: A study of user experience and media effects," *Int J Hum Comput Stud*, vol. 147, p. 102576, 2021. https://doi.org/10.1016/j.ijhcs.2020.10257