DATA COLLECTION AND ANALYSIS PLATFORM CONSTRUCTION WITH DEEP LEARNING IN THE TOBACCO PRODUCTION PROCESS

CHONG, L. M.[1] – LIU, B.[2*] – WANG, J.[2] – ZHU, M.[1] – ZHANG, X. Y.[2] – LU, Y.[2]

[1]*Hebei China Tobacco Industrial Co. Ltd, Shijiazhuang, Hebei 050000, China*
(e-mail: liming.chong@outlook.com, wuhanchengchao123@126.com)

[2]*Zhangjiakou Cigarette Factory Co. Ltd, Zhangjiakou, Hebei 075000, China*
(e-mail: ranqiang225576487@126.com, zzdxcj123@163.com, hmti54@126.com)

[*]*Corresponding author*
e-mail: bo.liu334@outlook.com

(Received 31st Jul 2025; accepted 28th Oct 2025)

Abstract. This research presents a model that combines knowledge graphs with deep neural networks to achieve quality prediction in tobacco production. In particular, a neural network model integrating a Convolutional Neural Network (CNN) and a Bidirectional Gated Recurrent Unit (BiGRU) is designed to predict key quality indicators in the cut - tobacco drying process, such as the temperature of the drying cylinder wall. The experimental results demonstrate that this model exhibits superiority in predicting the quality of the tobacco, offering corresponding guidance of tobacco production.
Keywords: *tobacco, deep learning, intelligent prediction, knowledge graph*

Introduction

As consumers' requirements for the quality and safety of tobacco products continue to rise, and the tobacco industry itself faces fierce market competition and a strict regulatory environment, improving the refined management level of tobacco production and ensuring the stability of product quality have become urgent issues for tobacco enterprises to address (Saddozai et al., 2022). The tobacco production process is a complex systematic project (Septiadi et al., 2022), involving multiple stages and numerous process parameters, such as tobacco leaf grading, fermentation, shredding, drying, and flavoring. The traditional tobacco production management approach mainly relies on manual experience and conventional statistical analysis methods (Pashovska, 2024), which are unable to promptly and accurately identify potential problems or optimization opportunities in the production process, resulting in low production efficiency and significant fluctuations in product quality.

As a structured knowledge representation method (Wang et al., 2017), the knowledge graph can model the complex relationships among processes, equipment, parameters, and quality indicators in tobacco production in the form of "entity - relationship - entity", thus providing rich domain knowledge support for quality prediction. Through the knowledge graph, the associations between processes can be clearly expressed, and semantic inputs can be provided for subsequent quality prediction models (Rossi et al., 2021).

Deep learning can automatically learn complex non - linear relationships from massive amounts of data, offering an effective means to solve the data processing and analysis problems in the tobacco production process. In the modeling of industrial production processes, deep learning can automatically learn complex non - linear relationships from a large amount of production data to establish accurate production process models. For

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

example, by constructing Long Short Term Memory (Jiang et al., 2023) to analyze and process image data in the industrial production process, product quality detection and classification can be achieved; by constructing a Recurrent Neural Network (RNN) model (Rather et al., 2015) to predict and analyze time - series data in the industrial production process, early warning and diagnosis of equipment failures can be realized, thereby enhancing production efficiency and product quality. The Support Vector Machine (Fan and Sharma, 2021) achieves classification and regression tasks by finding the optimal hyperplane and is suitable for prediction problems with small - sample datasets. The Random Forest (Meenal et al., 2021) improves the stability and accuracy of prediction by integrating multiple decision trees and is suitable for processing high - dimensional data. In addition, traditional statistical methods such as linear regression and ARIMA models also have certain applications in time - series prediction (Sirisha et al., 2022), but their performance is limited when dealing with non - linear relationships and high - dimensional data. In recent years, deep - learning - based Generative Adversarial Networks (Chen et al., 2022) and Reinforcement Learning (Huang et al., 2023) have also demonstrated potential in industrial production, enabling automated control and optimized operation of the production process.

While traditional methods such as Support Vector Machines (SVM) and Random Forests (RF), along with statistical models, have found applications in industrial settings, they face significant bottlenecks when dealing with complex processes like tobacco drying, which are characterized by strong non-linearity and high dimensionality. These limitations include: a limited capacity to capture complex couplings between process parameters; a heavy reliance on expert-driven feature engineering, which hinders generalization; and difficulty in effectively integrating unstructured domain knowledge, such as process specifications and technical standards, thereby constraining prediction accuracy and practical utility.

This research aims to utilize the knowledge graph and multiple deep neural networks, in combination with structured and unstructured data, to predict the quality of cut tobacco after the cut - tobacco drying process. Through the construction of the knowledge graph, rich domain knowledge support can be provided for the prediction model; while the multiple neural networks can fully utilize this knowledge to achieve accurate prediction of cut - tobacco quality (Dastres and Soori, 2021). By constructing a data collection and analysis platform, in - depth research on the data characteristics and inherent laws in the tobacco production process can be conducted, providing theoretical and technological support for the intelligent management of tobacco production.

Materials and methods

Construction of data collection and analysis platform for tobacco production

In tobacco production, to enhance production efficiency and quality, a data collection and analysis platform is constructed based on deep learning, primarily focusing on two aspects: knowledge graph technology and multi - neural network technology. First, a knowledge graph oriented towards the tobacco primary processing technology field is constructed. Data such as enterprise standards, citation specifications, and quality evaluations in the tobacco primary processing field are subjected to unified structured data management and representation. Then, after the completion of this knowledge graph, a domain knowledge retrieval service is provided. Subsequently, the knowledge is converted into word vector data. A multi - neural network is constructed, and the word

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

vectors are used as inputs for the word embedding layer to achieve the prediction of cut tobacco quality. Finally, an intelligent operation and maintenance system is built to predict and monitor the real - time status of the production line.

Construction of the knowledge graph for tobacco production

The triples of "entity - relationship - entity" in the knowledge graph of the tobacco production domain comprehensively and meticulously present the entire production process of tobacco from raw materials to finished products (Liu et al., 2023). In the tobacco - shred manufacturing process of tobacco production, there are close process correlations in the silk - making technology. After the blending and pre - treatment of tobacco sheets, the subsequent process is loosening and conditioning. Once the loosening and conditioning is completed, secondary flavoring follows. After secondary flavoring, the tobacco enters the cut - tobacco drying stage. The cut - tobacco drying process involves multiple important aspects. Its processing parameters include the process flow rate, and the quality indicators encompass the moisture content and temperature of the discharged material. *Table 1* presents the basic process knowledge triples of tobacco - shred production. These triples are interrelated and jointly construct the basic framework of the knowledge graph in the tobacco production domain, which facilitates in - depth understanding and management of each link in tobacco production.

Table 1. *Knowledge triples basic process of tobacco production*

Entity	Relation	Entity
General process	Branch process	Thin slice pre-treatment and blending Loosening and conditioning Primary flavoring Secondary flavoring Expansion and drying Blending and perfuming

The construction of the knowledge graph for tobacco production is shown in *Figure 1*. First, we systematically collected unstructured and semi-structured documents, including 'Primary Processing Specifications', 'Equipment Manuals', and 'Enterprise Quality Standards'. The Jieba segmentation tool was utilized with a custom dictionary of tobacco processing terminology to perform initial word segmentation (Zeng et al., 2021). Then, Core concepts relevant to the tobacco production domain were defined, including 'Process', 'Parameter', 'Equipment', and 'Quality Indicator'. Relationships between these concepts, such as 'hasParameter', 'affectsQuality', and 'precedesProcess', were explicitly formalized. The bottom - up approach mines entities, relationships, and attributes from specific data to construct a local knowledge graph (Lin et al., 2021). Finally, select a graph database to store the constructed knowledge graph, storing entities, relationships, and attributes in the form of a graph. The extracted triples were stored in a Neo4j graph database. A query service based on the Cypher language was implemented to provide domain knowledge retrieval for downstream models.

The knowledge graph constructed in the field of tobacco production technology encompasses a wide range of research contents. First and foremost, through the extraction of entities and attributes, numerous crucial pieces of information have been identified. Among them, the entities of process names include loosening and conditioning, cut -

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

tobacco drying, primary flavoring, etc.; the processing parameters involve the opening degree of the circulation damper, the moisture content of incoming materials, etc.; and the noun attributes of quality indicators comprise the moisture content of outgoing materials, the temperature of outgoing materials, etc. Meanwhile, in the construction of the knowledge graph, coreference resolution has been carried out to achieve knowledge fusion and reasoning (Ji et al., 2022). Finally, data integration has been accomplished. The knowledge graph of tobacco production technology is illustrated in *Figure 2*. Solid lines represent the sequential flow between different production processes. Dashed lines represent attribute or associative relationships, linking processes or entities with their specific parameters or quality indicators.

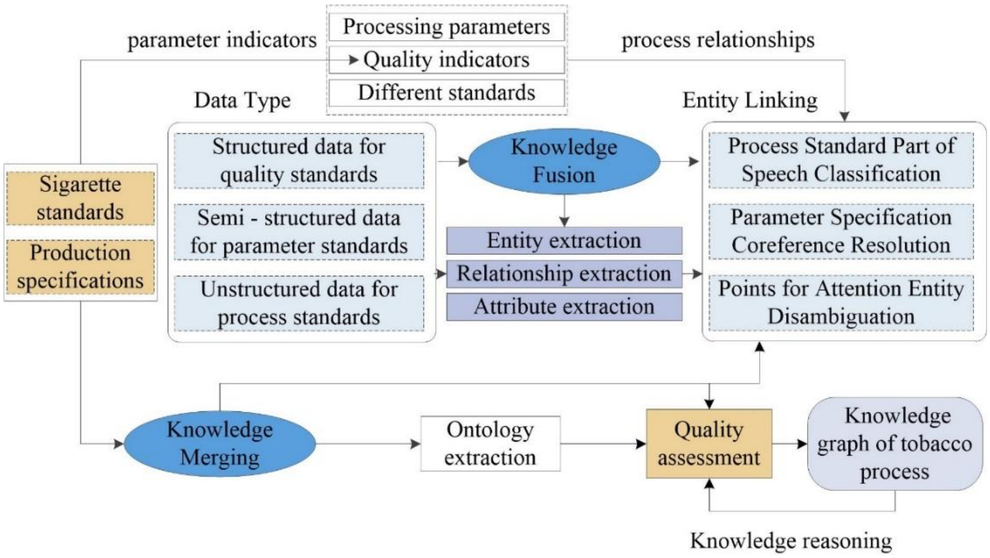

Figure 1. *Knowledge graph construction process of tobacco production*

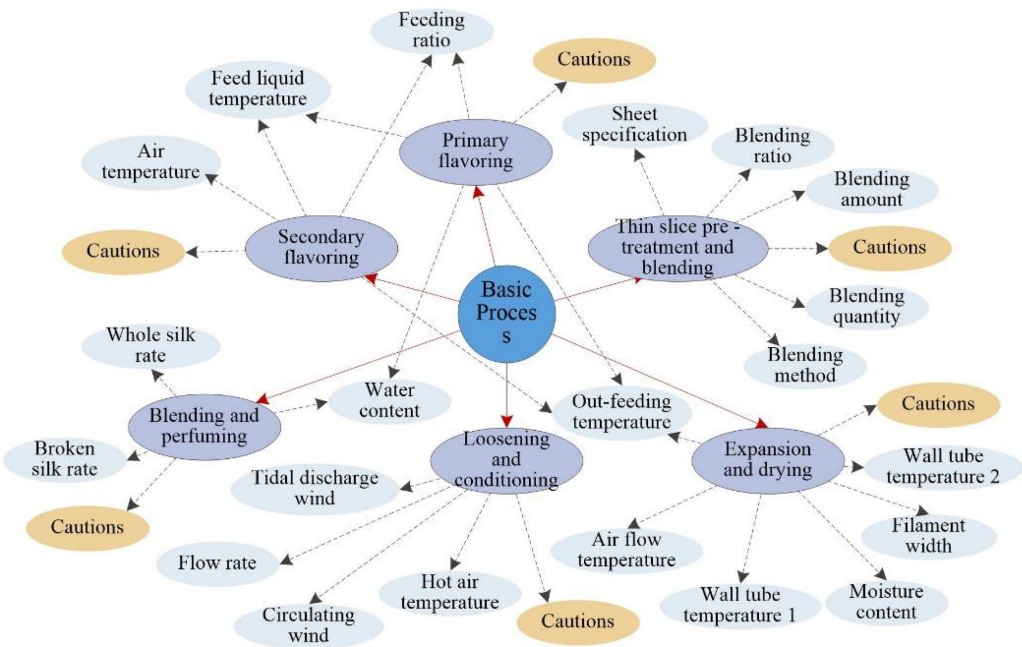

Figure 2. *Knowledge graph of tobacco production process*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

Deep learning methods for quality prediction of tobacco

Word vectorization

In the construction and application of knowledge graphs in the field of silk - making processes, word vectorization serves as a pivotal step for achieving efficient knowledge processing and analysis. This research undertakes the task of word vectorization for the knowledge graph of silk - making processes, which is essentially a process of graph embedding, knowledge representation, and learning. Considering that the enterprise standards, production specifications, and technical requirements involved in silk - making processes are all presented in textual information, this study employs the jieba word - segmentation tool to conduct a meticulous and effective initial segmentation of the professional terms in the field of silk - making processes, accurately identifying and segmenting the professional vocabulary in tobacco - making processes.

After the initial segmentation of professional terms, the word vector learning algorithm TransE and the word2vec word vector processing tool are further adopted to realize the vectorization of the knowledge graph (Do and Pham, 2021; Zhao, 2025). The TransE algorithm, based on the translational distance model, can effectively capture the semantic associations between entities and relationships in the knowledge graph.

The core idea of the TransE algorithm is that for each triple (h, r, t) in the knowledge graph (where *h* represents the head entity, *r* represents the relation, and *t* represents the tail entity), the following formula should be satisfied.

$$h + r \approx t \tag{Eq.1}$$

The collection of all triple knowledge is denoted as p, and the corresponding negative triples are denoted as p'.

$$p = (h_i, r_i, t_i) \tag{Eq.2}$$

The corresponding evaluation score function f and the maximum margin loss function L are respectively expressed as follows.

$$f(p) = \vec{h} + \vec{r} + \vec{t}_N \tag{Eq.3}$$

$$L = \frac{1}{l} \sum_{p \in P} [f(p) - f(p') + \gamma]_+ \tag{Eq.4}$$

where, γ represents the hyperparameter; l denotes the number of elements in the triple.

The word2vec word vector processing tool learns the distributed representation of vocabulary from a large amount of text data through a neural network model, mapping each technical term to a fixed-length vector. The objective of the Continuous Bag-of-Words model (CBOW) (Feng et al., 2022) is to predict the central word based on the context words. Suppose the size of the context window is C, the input context word vectors are $(w_{c-C}, ..., w_{c+C})$ respectively, and the central word is w_c.

Firstly, sum up the contextual word vectors and calculate their average to obtain the contextual representation vector $\hat{h}$.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

$$\hat{h} = \frac{1}{2C} \sum_{i=-C, i\neq 0}^{C} v_{w_{c+i}} \quad \text{(Eq.5)}$$

where, $v_{w_{c+i}}$ is the input - layer vector representation of the context word w_{c+i}. Subsequently, the probability distribution of the central word is computed via a linear transformation and the softmax function.

$$P(w_c \mid w_{c-C}, \cdots, w_{c-1}, w_{c+1}, \cdots, w_{c+C}) = \frac{\exp(u_{w_c}^T \hat{h})}{\sum_{w=1}^{V} \exp(u_w^T \hat{h})} \quad \text{(Eq.6)}$$

where, u_w represents the vector of the output layer corresponding to the word w, and V denotes the size of the vocabulary.

Multivariate deep neural network model

To achieve the prediction of the quality of the silk - making process, prediction work was carried out regarding the cylinder wall temperature in the cut - tobacco drying process of the silk - making process. We constructed a multivariate neural network model, namely CNN - BiGRU - ATT, with the results of the processing parameters and quality indicators of the drying process after being processed by the word vector embedding layer as the input. This model integrates the Convolutional Neural Network (CNN), the Bidirectional Gated Recurrent Unit (BiGRU), and the Attention mechanism, as shown in *Figure 3*. First, the convolutional layer is utilized to extract features from the input data, capturing the local feature patterns within the data. Subsequently, the bidirectional GRU layer is employed to explore the sequential relationships and long - term dependencies among the data, aiming to comprehensively understand the dynamic changes of various parameters over time in the drying process. Meanwhile, the Attention mechanism is introduced. This mechanism can adaptively identify the information and parameters that are more crucial for the prediction of the cylinder wall temperature during the prediction process, assigning different weights to different data, thus more precisely focusing on the important information. Finally, the softmax function is used to further process the processed data and output the prediction results.

The basic structure of GRU is illustrated in *Figure 4*. The BiGRU is an extended form of the GRU. On the basis of the GRU (Zhang et al., 2021), the BiGRU incorporates a reverse recurrent structure, enabling the model to leverage not only past information of the sequence but also future information for prediction and analysis.

The BiGRU consists of two independent GRU units: a forward GRU unit and a backward GRU unit. The computational process of the forward GRU at a time step t is as follows.

For the input sequence $(x_1, x_2, \ldots x_T)$, the forward GRU sequentially computes the hidden states $(\vec{h}_1, \vec{h}_2, \ldots, \vec{h}_T)$. The calculation formulas for the forward reset gate $\vec{r}_t$ and the update gate $\vec{z}_t$ are as follows.

$$\vec{r}_t = \sigma(W_r[x_t, \vec{h}_{t-1}] + b_r) \quad \text{(Eq.7)}$$

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

$$\vec{z}_t = \sigma(W_z[x_t, \vec{h}_{t-1}] + b_z) \qquad \text{(Eq.8)}$$

where, σ denotes the Sigmoid activation function, which maps the input value to the interval $(0,1)$. W_r and W_z respectively represent the weights corresponding to the reset gate and the update gate; b_r and b_z denote the bias vectors of the reset gate and the update gate. $[x_t, \vec{h}_{t-1}]$ represents the concatenation of the input vector x_t and the previous time - step hidden state $\vec{h}_{t-1}$.

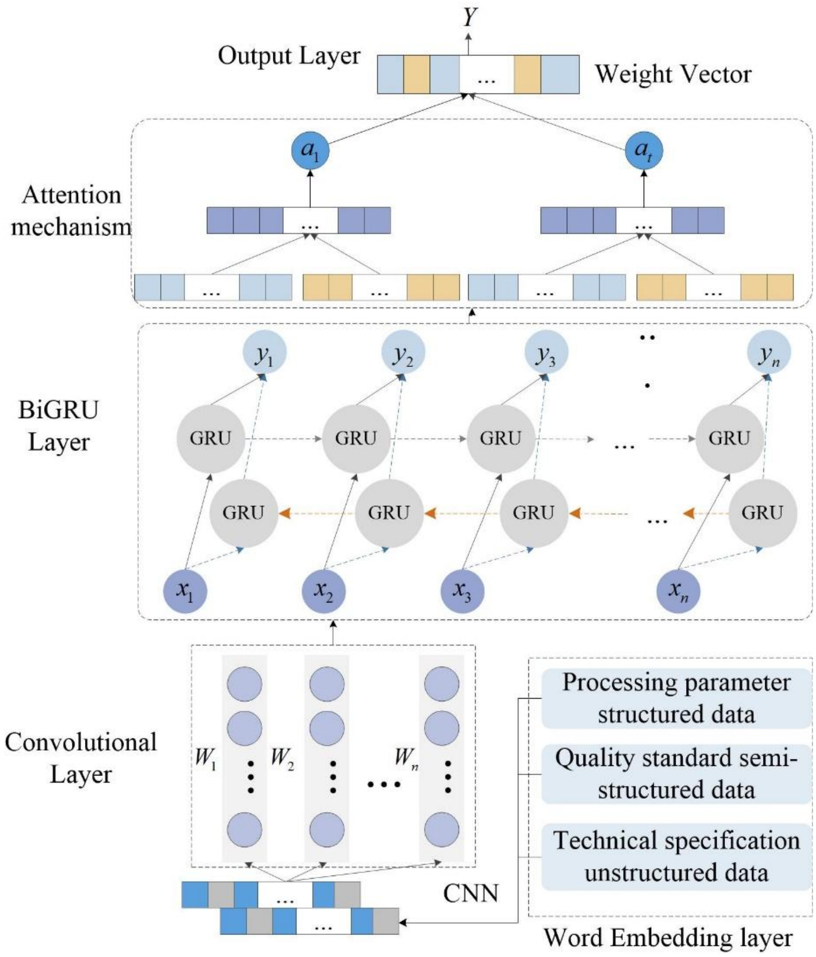

Figure 3. *Structural model of multilayer deep learning network*

The candidate hidden state $\tilde{\vec{h}}_t$ and the current hidden state $\vec{h}_t$ are calculated as follows.

$$\tilde{\vec{h}}_t = \tanh(W_{\tilde{h}}[x_t, \vec{r}_t \odot \vec{h}_{t-1}] + b_{\tilde{h}}) \qquad \text{(Eq.9)}$$

$$\vec{h}_t = (1 - \vec{z}_t) \odot \vec{h}_{t-1} + \vec{z}_t \odot \tilde{\vec{h}}_t \qquad \text{(Eq.10)}$$

where, $W_{\tilde{h}}$ is the weight matrix of the candidate hidden state, and $b_{\tilde{h}}$ is the bias vector of the candidate hidden state.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

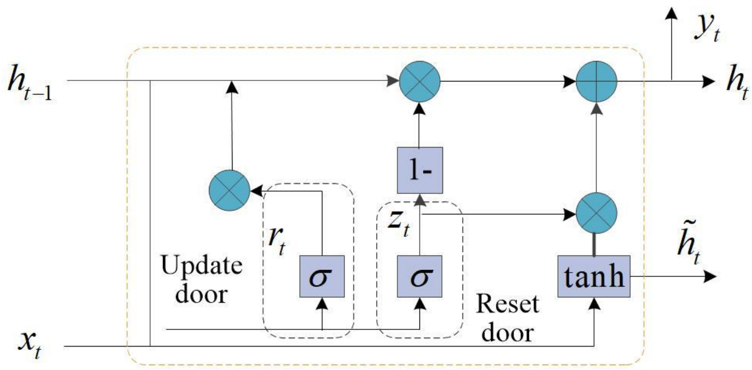

Figure 4. *Neural network structure of GRU*

The computation of the backward GRU at time step t commences from the end of the sequence. Its calculation formula is similar to that of the forward GRU, except that the order of the input sequence is reversed. At each time step, the hidden states of the forward and backward GRUs are concatenated to obtain the final hidden state h_t of the BiGRU.

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad \text{(Eq.11)}$$

In the tobacco process, the textual data contains abundant semantic information, yet it is rife with substantial noise, such as irrelevant characters, erroneous expressions, and repetitive information. To unearth the semantic information within the texts of the silk-making process and mitigate the impact of noise, this paper introduces the attention mechanism and combines it with the GRU model. The attention mechanism layer enables the model to focus on the significant parts of the text and disregard the irrelevant information (Zhang et al., 2021). Its structure and workflow are illustrated in *Figure 5*.

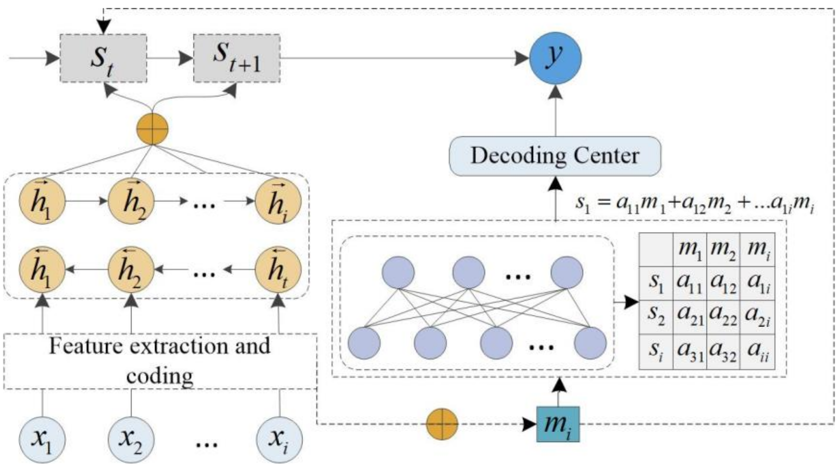

Figure 5. *Operation mechanism diagram of attention mechanism*

First, the GRU model processes the fault input sequence. The GRU model has forward and backward processing procedures, which respectively generate the output vector h_i and the backward output vector h_k. Subsequently, all the results output by the GRU

APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH 23(6):11645-11659.
http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

model are taken as the input of the attention mechanism. Through the calculation of the attention mechanism, the final output vector $\hat{h}$ is obtained, which prominently presents the semantic information of fault keywords.

In the attention mechanism, we need to calculate the attention weight of node i with respect to node k. Suppose the number of elements in the fault input sequence is T, and there are different weight matrices V, W, and U simultaneously. Thus, the intermediate variable e_{ki} is calculated.

$$e_{ki} = V\tanh(Wh_k + Uh_i + b) \tag{Eq.12}$$

To endow the attention weights with the properties of a probability distribution, it is necessary to perform normalization on the intermediate variables. Generally, the Softmax function is employed to accomplish this operation, and the formula is as follows.

$$\alpha_{ki} = \frac{\exp(e_{ki})}{\sum_{j=i}\exp(e_{ki})} \tag{Eq.13}$$

where, α_{ki} denotes the attention weight of node i with respect to node k.

After obtaining the attention weight α_{ik}, the final output vector $\hat{h}$ can be computed. Specifically, it is achieved by performing a weighted summation of the attention weights and the output vector h_i of the GRU model.

$$\hat{h} = \sum_{i=1}^{T}\alpha_{ik} \cdot h_i \tag{Eq.14}$$

The final output vector synthesizes the output information of the GRU model and, through the adjustment of attention weights, highlights the semantic information associated with the fault keywords.

Results

Dataset and evaluation index

The structured data used in this study were derived from the production records of 200 consecutive batches from the primary production line of a major Chinese cigarette factory during the 2023 calendar year. Data were automatically collected via the plant's Distributed Control System (DCS) and sensor network. To characterize the dataset, descriptive statistics for key process parameters and quality indicators are presented in *Table 2*. These data encompass the primary processing parameters and quality indicators in the cut tobacco drying process. The processing parameters include hot air temperature, circulating air flow, cylinder rotation speed, etc. These parameters directly influence the heat exchange and moisture loss of cut tobacco during the drying process. The quality indicators consist of outlet moisture content, outlet temperature, tobacco filling capacity, etc. Among them, the outlet moisture content serves as a crucial indicator for evaluating the drying effect of cut tobacco, directly impacting the subsequent processing of tobacco and the quality of finished products. The data were collected via the existing automated

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

production equipment and sensors in the cigarette factory, and the collected data were stored in the database in the form of time series.

Table 2. *Descriptive statistics of key variables (N=200)*

Variable	Unit	Mean	Std. Dev.	Min	Max
Hot Air Temperature	°C	145.2	8.7	120.0	165.0
Circulation Airflow	m^3/h	3200	210	2800	3600
Cylinder Speed	rpm	5.2	0.6	4.0	6.5
Outlet Moisture Content	%	12.4	1.1	9.8	15.2
Cylinder Wall Temperature	°C	132.5	9.8	110.0	155.0

The unstructured data such as normative reference documents, technical requirements, and operation manuals in the production of cigarette factories contain a wealth of industry knowledge and experience. To make full use of these data, this study conducts the construction of a knowledge graph. Meanwhile, with the aid of deep - learning technology and the word embedding layer, the unstructured text data are incorporated into the model. The word embedding layer maps each word in the text to a low - dimensional vector space so as to be input into the multivariate neural network model together with the structured data. To train the CNN-BiGRU-ATT neural network model, the collected data were divided into a training set, a validation set, and a test set at a ratio of 7:2:1. To evaluate the performance of the multivariate neural network model, the evaluation indicators adopted include accuracy, recall, and F1 score.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{Eq.15}$$

$$Recall = \frac{TP}{TP+FN} \tag{Eq.16}$$

$$F1 = \frac{2\times Precision \times Recall}{Precision + Recall} \tag{Eq.17}$$

In the *Equations (15-17)*, True Positives (*TP*) represent the count of samples where the cylinder wall temperature was correctly predicted to be in a 'High-Temperature' state; True Negatives (*TN*) denote samples correctly predicted as 'Normal-Temperature'; False Positives (*FP*) indicate samples incorrectly predicted as 'High-Temperature'; and False Negatives (*FN*) refer to samples incorrectly predicted as 'Normal-Temperature'.

Results and comparative experiment

Under the same dataset, compare the method (CNN-BIGRU-ATT) that fuses knowledge graphs with multi - layer neural networks with other deep learning methods. *Table 3* presents the results of predicting the temperature of the drying cylinder wall using different models.

Table 3 demonstrates the performance of different deep learning models in predicting the temperature of the drying cylinder wall during the cut - tobacco drying process. As

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

can be discerned from the table, the CNN - BiGRU - ATT model outperforms other models in terms of accuracy (87.31%), recall (81.26%), and F1 score (84.17%). In contrast, the standalone CNN model exhibits the poorest performance, with an accuracy of merely 80.88%. This result indicates that the combination of model fusion and knowledge graphs can significantly improve prediction accuracy, offering a more reliable solution for quality prediction in tobacco production.

***Table 3.** Evaluation indicators of different deep learning models fused with knowledge graphs*

Models	Accuracy	Recall	F1
CNN	80.88	73.69	77.1
BiLSTM	82.51	75.98	79.1
BiGRU	82.71	76.9	79.69
CNN-BiLSTM	85.04	78.32	81.53
CNN-BiGRU-ATT	87.31	81.26	84.17

Discussion

To assess the statistical significance of the performance differences observed in *Table 3*, a paired t-test was conducted on the accuracy scores obtained from 5-fold cross-validation. The results confirmed that the superiority of the CNN-BiGRU-ATT model over all other compared models is statistically significant (p-value < 0.05). From a practical standpoint, this enhanced accuracy and lower loss rate translate to a more reliable early warning system for abnormal temperature fluctuations in the drying cylinder. This provides operators with a longer lead time to make adjustments, thereby helping to prevent issues such as over-drying or uneven moisture content, which directly impact the final product quality and production efficiency.

Figure 6 depicts the prediction loss rates of different models during the training process. As shown in the figure, the CNN - BiGRU - ATT model has the lowest loss rate, suggesting that it can better fit the data during training and reduce prediction errors. Conversely, the standalone CNN model has the highest loss rate, indicating its weak fitting ability. The low loss rate of the CNN-BiGRU-ATT model is due to its multi - level structural design: the CNN extracts local features, the BiGRU captures long - term dependencies in time series, and the attention mechanism further reduces the impact of noise by weighted focusing on important information. This result indicates that the CNN-BiGRU-ATT model not only performs excellently in prediction accuracy but also demonstrates better stability and convergence during the training process.

To more intuitively demonstrate the efficacy of this model, the temperature of the drying cylinder wall is predicted in combination with other process parameters. The comparison results with the actual values and the relative errors are shown in *Figure 7*.

Figure 7 presents a comparison between the predicted values and the actual values of the drying cylinder wall temperature by the CNN-BiGRU-ATT model. As can be discerned from the figure, the trends of the predicted values and the actual values are substantially consistent. Although there are certain discrepancies at some time points, the overall error is relatively minor, and the prediction results are rather accurate. This indicates that the CNN-BiGRU-ATT model can effectively capture the variation patterns of the drying cylinder wall temperature. Especially after integrating the knowledge graph, the model can better comprehend the relationships between process parameters and

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

quality indicators. By introducing the attention mechanism, the model can adaptively identify key parameters, thereby enhancing the prediction accuracy. This outcome further validates the reliability of the CNN-BiGRU-ATT model in practical applications and provides effective quality prediction support for the drying process in tobacco production.

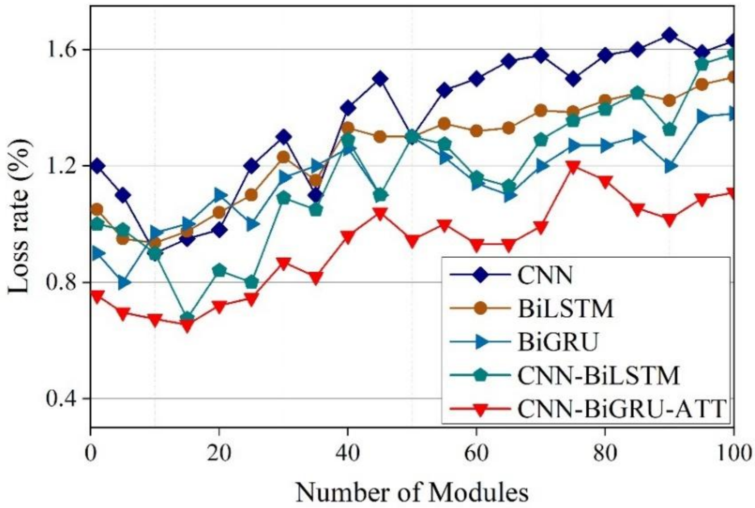

***Figure 6.** Prediction loss rates of different deep learning models integrated with knowledge*

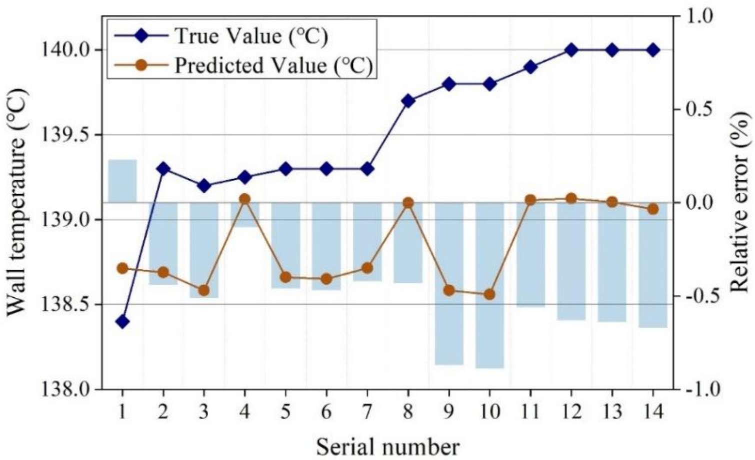

***Figure 7.** Comparison between the actual and predicted values of the wall temperature*

This study demonstrates the effective integration of a domain-specific knowledge graph with a hybrid deep learning model (CNN-BiGRU-ATT) for quality prediction in tobacco production. Our experimental results show that this approach achieves superior performance compared to several benchmark models.

The performance gain of the CNN-BiGRU-ATT model can be attributed to its architecture, which is designed to capture both local spatial features and long-term temporal dependencies—a challenge in complex industrial processes. Compared to the standalone LSTM model applied in similar agricultural and industrial contexts (Jiang et al., 2023), our hybrid model leverages CNNs for preliminary feature extraction, potentially leading to a more robust representation of input data. Furthermore, the use of

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

a Bidirectional GRU, as opposed to the unidirectional RNNs used in some prior works for time-series prediction (Rather et al., 2015), allows our model to utilize contextual information from both past and future states within the sequence, enhancing its understanding of process dynamics.

The integration of the knowledge graph provided a structured semantic framework, enriching the model with domain expertise. This aligns with the broader trend of leveraging knowledge graphs for improved reasoning in industrial applications [5, 17]. The incorporation of the attention mechanism further allowed the model to adaptively focus on the most relevant process parameters at different time steps, which is an advancement over standard sequence modeling and contributes to both performance and interpretability (Niu et al., 2021). Analysis of the attention weights revealed that the model consistently assigned higher importance to parameters like 'Hot Air Temperature' and 'Circulation Airflow' when predicting the cylinder wall temperature, which is consistent with the physical principles of the drying process.

While this study marks a step forward, the knowledge graph could be further expanded to include additional dimensions such as equipment health data and ambient environmental conditions. Exploring more advanced knowledge graph embedding techniques (Do and Pham, 2021) and their interaction with the neural network also presents a promising direction for future research to enhance the model's generalization capability and reasoning power.

Conclusion

This research successfully achieved the precise prediction of tobacco shred quality in the cut tobacco drying process by constructing a knowledge graph in the field of tobacco production and integrating it with a multi-neural network model. The experimental results indicate that the CNN-BiGRU-ATT model demonstrated outstanding performance in predicting the drying drum wall temperature, attaining an accuracy rate of 87.31%, significantly outperforming other comparative models. The introduction of the knowledge graph provided the model with abundant domain knowledge, enabling the model to better comprehend the relationship between process parameters and quality indicators. Moreover, the application of the attention mechanism allowed the model to adaptively focus on key information, further enhancing the prediction accuracy. This research not only offers an effective technical means for quality prediction in tobacco production but also provides novel insights for intelligent management in industrial production. The accurate prediction achieved by our model has direct practical implications for real-time monitoring and early fault detection in the drying process. Operators can leverage the predictions to proactively adjust key parameters, such as hot air temperature and feed rate. This capability is expected to yield significant benefits in stabilizing product quality (e.g., reducing moisture content fluctuations) and improving energy efficiency, thereby offering a viable technological pathway for the intelligent upgrading of tobacco production. Future research can further expand the coverage of the knowledge graph and explore more complex neural network structures to enhance the model's generalization ability and practical application value.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

REFERENCES

[1] Chen, P., Liu, H., Xin, R., Carval, T., Zhao, J., Xia, Y., Zhao, Z. (2022): Effectively detecting operational anomalies in large-scale IoT data infrastructures by using a GAN-based predictive model. – The Computer Journal 65(11): 2909-2925.

[2] Dastres, R., Soori, M. (2021): Artificial neural network systems. – International Journal of Imaging and Robotics 21(2): 13-25.

[3] Do, P., Pham, P. (2021): W-KG2Vec: A weighted text-enhanced meta-path-based knowledge graph embedding for similarity search. – Neural Computing and Applications 33(23): 16533-16555.

[4] Fan, M., Sharma, A. (2021): Design and implementation of construction cost prediction model based on SVM and LSSVM in industries 4.0. – International Journal of Intelligent Computing and Cybernetics 14(2): 145-157.

[5] Feng, Y., Hu, C., Kamigaito, H., Takamura, H., Okumura, M. (2022): A simple and effective usage of word clusters for CBOW model. – Journal of Natural Language Processing 29(3): 785-806.

[6] Huang, Z., Liu, H., Wu, J., Lv, C. (2023): Conditional predictive behavior planning with inverse reinforcement learning for human-like autonomous driving. – IEEE Transactions on Intelligent Transportation Systems 24(7): 7244-7258.

[7] Ji, S., Pan, S., Cambria, E., Marttinen, P., Yu, P. S. (2022): A survey on knowledge graphs: Representation, acquisition, and applications. – IEEE Transactions on Neural Networks and Learning Systems 33(2): 494-514.

[8] Jiang, S., Chen, G., Chen, D., Chen, T. (2023): Application and evaluation of an improved LSTM model in the soil moisture prediction of southeast Chinese tobacco-producing areas. – Journal of the Indian Society of Remote Sensing 51(9): 1843-1853.

[9] Lin, J., Zhao, Y., Huang, W., Liu, C., Pu, H. (2021): Domain knowledge graph-based research progress of knowledge representation. – Neural Computing and Applications 33(2): 681-690.

[10] Liu, Y., Yang, S., Xu, Y., Miao, C., Wu, M., Zhang, J. (2023): Contextualized graph attention network for recommendation with item knowledge graph. – IEEE Transactions on Knowledge and Data Engineering 35(1): 181-195.

[11] Meenal, R., Angel, P., Pamela, D., Rajasekaran, E. (2021): Weather prediction using random forest machine learning model. – Indonesian Journal of Electrical Engineering and Computer Science 22(2): 1208-1215.

[12] Niu, Z., Zhong, G., Yu, H. (2021): A review on the attention mechanism of deep learning. – Neurocomputing 452: 48-62.

[13] Pashovska, S. (2024): Current conditions in tobacco production in Macedonia-sustainability analysis, development possibilities and future perspectives. – JASRD Journal of Agriculture and Sustainable Rural Development 2(3-4): 116-127.

[14] Rather, A. M., Agarwal, A., Sastry, V. N. (2015): Recurrent neural network and a hybrid model for prediction of stock returns. – Expert Systems with Applications 42(6): 3234-3241.

[15] Rossi, A., Barbosa, D., Firmani, D., Matinata, A., Merialdo, P. (2021): Knowledge graph embedding for link prediction: A comparative analysis. – ACM Transactions on Knowledge Discovery from Data 15(2): 1-49.

[16] Saddozai, K. N., Nasrullah, M., Khan, J., Shah, S. A., Saqib, R., Zahra, N., Rasheed, M. (2022): Allocative efficiency of tobacco production in district Mardan, Pakistan. – Sarhad Journal of Agriculture 40(2): 448-453.

[17] Septiadi, D., Nursan, M., FR, A. F. U., Suparyana, P. K., Nadiawati, B. S. (2022): Analysis of agricultural input resources on chopped tobacco production of reket and seram varieties in Suralaga district east Lombok regency. – Jurnal Biologi Tropis 22(4): 1269-1278.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659

[18] Sirisha, U. M., Belavagi, M. C., Attigeri, G. (2022): Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison. – IEEE Access 10: 124715-124727.
[19] Wang, Q., Mao, Z., Wang, B., Guo, L. (2017): Knowledge graph embedding: A survey of approaches and applications. – IEEE Transactions on Knowledge and Data Engineering 29(12): 2724-2743.
[20] Zeng, K., Li, C., Hou, L., Li, J., Feng, L. (2021): A comprehensive survey of entity alignment for knowledge graphs. – AI Open 2: 1-13.
[21] Zhang, Z., Dong, Z., Lin, H., He, Z., Wang, M., He, Y., Gao. X., Gao, M. (2021): An improved bidirectional gated recurrent unit method for accurate state-of-charge estimation. – IEEE Access 9: 11252-11263.
[22] Zhao, R.-X. (2025): QTP-Net: A quantum text pre-training network for natural language processing. – arXiv preprint arXiv:2506.00321.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2306_1164511659