International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 8141-8148

https://doi.org/10.70517/ijhsa464697

An Attention-Enhanced LSTM Framework for Real-Time Innovation Capability Evaluation in Higher Education via IoT-Driven Time Series Analysis

Haijuan Zhou[1], Yali Hou[1], Xiaomeng Qi[1], Xuefeng Hu[1] and Xiangge Liu[1,*]
[1]Qinhuangdao Vocational and Technical College, Qinhuangdao 066000, China
Corresponding authors: (e-mail: liuxiangge@qvc.edu.cn).

Abstract Evaluating the innovation capability of university research teams is critical for guiding policy and resource allocation in higher education. Traditional input–output metrics fail to capture the dynamic, multidimensional nature of scientific innovation processes in modern colleges. In this study, we propose an **IoT-empowered, attention-enhanced LSTM** framework that integrates real-time sensor data from smart laboratories and campus innovation centers to continuously monitor key research activities. We first construct a capability indicator system combining inputs (equipment usage, laboratory environmental parameters, researcher activity) and outputs (experiment throughput, publication metrics), all captured via 5G-enabled IoT devices. An attention mechanism dynamically weights each indicator at every time step, allowing the LSTM to focus on the most informative features as innovation unfolds. To validate our approach, we conduct experiments on three datasets collected from university innovation labs over six months: climate-controlled bioengineering chambers (SML-Campus), soil-monitoring for agrotech projects, and power usage in maker-spaces. Compared with baseline MLP, vanilla LSTM, and BiLSTM models, our method achieves superior prediction accuracy of research output trends (e.g., on SML-Campus: RMSE = 0.024, MAE = 0.019, R^2 = 0.9999) and consistently higher anomaly detection precision in identifying workflow bottlenecks.

Index Terms biological technology, innovation capability, internet of things, intelligent computing, evaluation system, communication technology, biological research

I. Introduction

In modern scientific research, innovation and application of biological technology have become key driving forces for the development of multiple fields such as biomedical, environmental science, and agricultural biotechnology. With the rapid development of information technology, especially the widespread application of the Internet of Things (IoT) and advanced communication technologies, research and innovation in biological technology are moving towards a more refined and intelligent direction [1]–[3]. As an important battlefield for biological technology innovation, the improvement of research and technological innovation capabilities in higher education institutions is of great significance for promoting the national level of biological research and technological application. However, due to the complexity and diversity of biological technology research, how to scientifically and accurately evaluate innovation capabilities in the field of biology, especially in higher education institutions, still faces enormous challenges [4]–[6].

Traditional evaluation methods for innovation capability often focus on simple input-output analysis, ignoring the dynamic and multidimensional characteristics of biological technology innovation [7], [8]. In recent years, with the development of technologies such as big data, artificial intelligence, and the Internet of Things, new evaluation methods have gradually emerged, which can better capture the detailed changes in the process of biological technology innovation. For example, IoT technology can collect and transmit various environmental data, experimental data, and technical feedback in biological experiments in real time, providing more accurate information for evaluation. In addition, the introduction of intelligent computing technology has made the analysis and processing of large-scale data more efficient, which can help research managers gain a deeper understanding of the innovation capabilities and research achievements of biological research teams [9], [10].

With the gradual improvement of the informatization construction of colleges and universities, the all-round dynamic management of will be realized. As an advanced technical means of information acquisition and processing, the Internet of Things has broad development and application prospects in the process of college innovation process management and the construction of capability evaluation index system and has an important guiding role and practical feasibility. Under the above research background, this research will establish evaluation of ability of college innovation team based on scientific research input and scientific research output based on advanced technology and use different evaluation methods to calculate innovation team. By analyzing and comparing the differences in the calculation results of various methods, the indicator system is verified.

II. Advanced communication technology and Internet of Things technology

The Internet of Things (IoT) operates through a network of interconnected devices that collaborate to form a powerful and efficient system. An IoT device can be any object capable of connecting to the internet, including smart devices, technology components, laptops, and mobile phones, whether through wired or wireless networks [11], [12]. IoT has three key characteristics:

1) **Seamless communication:** IoT enables barrier-free communication, which can be easily scaled and extended via the internet.
2) **Autonomous perception:** IoT devices are equipped with the ability to identify, assess, and respond to changes in their environment, enhancing their responsiveness.
3) **Automation and self-control:** Once equipped with perception capabilities, IoT systems can automate business functions, achieving intelligent self-feedback and control. This reduces the need for repetitive manual tasks, lightens the workload of users, and enhances operational efficiency.

See Figure 1 for a visual representation of this concept.

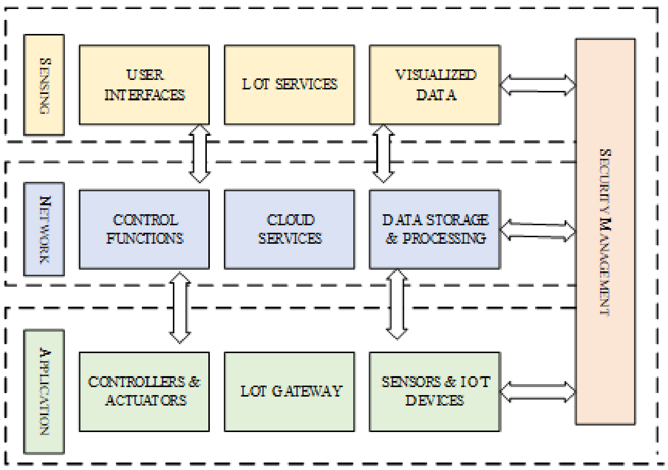

Figure 1: IoT three-tier architecture.

The IoT architecture consists of three primary layers: the perception layer, the network layer, and the application layer.

1) **Perception layer:** The role of the perception layer is to collect and sense data, focusing primarily on information gathering. Various types of sensors generate diverse data, such as multimedia information, body temperature, humidity, and other physical quantities. This data is collected by the information collection sub-layer.
2) **Network layer:** The network layer is responsible for transmitting the collected data using various network protocols. Transmission networks, including mobile communication networks and the Internet, ensure the data reaches the application layer securely and efficiently. This process leverages multiple network technologies, such as self-organizing communication, to facilitate smooth and reliable data transmission. The networking and collaborative information processing sub-layer handles the collaborative processing of the collected data.
3) **Application layer:** Once the data reaches the application layer, it may require processing to align with specific business needs, as the raw data may not be directly applicable to the platform. This task is managed by the application support platform, which processes the data accordingly. Additionally, due to the lack of a unified information resource to support diverse services, this responsibility is taken on by another sub-layer within the application layer. This structured approach enhances the efficiency of data processing and application [13].

The specific framework is shown in Figure 2.

5G is a new communication system. Its key technologies include multiple input and multiple output technology, multi-carrier technology of filter banks, simultaneous full-duplex technology at the same frequency, ultra-dense heterogeneous network technology, self-organizing network technology, software-defined network, and content distribution network.

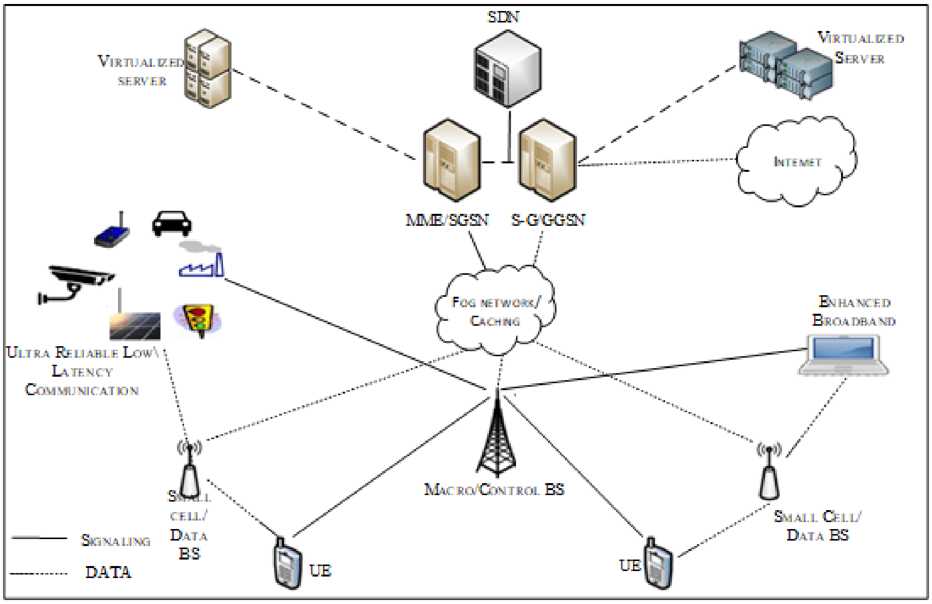

Figure 2: 5G network architecture.

China has put 5G into commercial use in recent years and has begun to build base stations on a large scale. In the future, 5G will become essential in mobile communication, video control, and other fields [14], [15]. It is also a basic technology for agricultural modernization in major agricultural countries, and its importance is widely recognized.

At present, 5G is used in various aspects of agriculture, such as irrigation, fertilization, environmental monitoring, and pest control. However, there are still corresponding deficiencies in the key levels of "perception, transmission, and processing". For example, the perception technology at the perception level still commonly uses various types of sensors. The sensors commonly available on the market have strict environmental conditions, and their measurement accuracy and stability are greatly reduced in non-standard conditions. Therefore, it is necessary to develop high-stability and low-cost sensors.

There are many options for information transmission technology at the transmission level. Whether wireless or wired, the bandwidth and speed of communication are the main breakthrough points. With the integration of 5G and IoT technology, the biggest improvement is in the IoT network layer. Because its data traffic has increased by 1000 times compared with 4G, the agricultural Internet of Things requires large-scale sensor networks to achieve effective environmental monitoring [16].

Real-time acquisition, whether for sensor data or video information transmission, ensures timeliness. The need for scene modeling can also provide guarantees for acquiring the information required for modeling. The advantages of 5G IoT are not limited to this, as there are also significant improvements in other aspects.

The agricultural Internet of Things needs to centralize, process, and take corresponding actions on the collected data. With the support of 5G technology, cloud computing and edge computing will be more capable of handling the massive data processing demands of the Internet of Things. The process is the first stage shown in Figure 3, as expressed in Equation 1.

$$e_i = \text{score}(Q, K_i) \quad (\forall i = 1, 2, 3, \ldots, n) \tag{1}$$

$$\text{score}(Q, K_i) = Q \cdot K_i \tag{2}$$

When the score is obtained, these scores need to be normalized, so that all element weights sum to 1, and it is recorded as Equation (3). This process is the second stage:

$$b_i = \frac{\exp(e_i)}{\sum_{j=1}^{n} \exp(e_j)} \quad (\forall i = 1, 2, 3, \ldots, n) \tag{3}$$

This is the third stage, the score is denoted as Equation (4):

$$\text{Attention Value} = \sum_{i=1}^{n} b_i \times \text{Value}_i \tag{4}$$

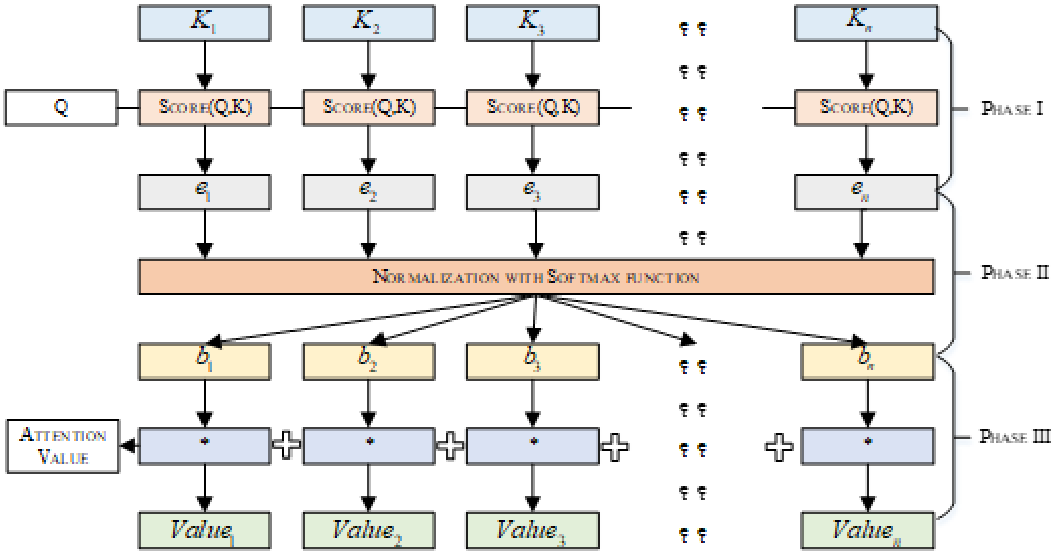

Figure 3: Calculation process of attention mechanism.

III. Method

Different types of sensors generate various forms of data. When conducting predictive analysis on multidimensional time series data generated by the Internet of Things (IoT), one common challenge is how to effectively process the features of this data. Typically, machine learning models and statistical methods are employed to assess feature importance, enabling the selection and elimination of features based on their significance. However, a direct removal of features overlooks the fact that the importance of a feature in multidimensional time series data may vary at different time points [17].

To address the limitations of traditional machine learning and statistical methods in processing features of multidimensional time series data, this paper proposes an attention-based prediction model utilizing Enhanced LSTM (Long Short-Term Memory). The model is designed to better handle the dynamic feature importance over time,see Figure 4.

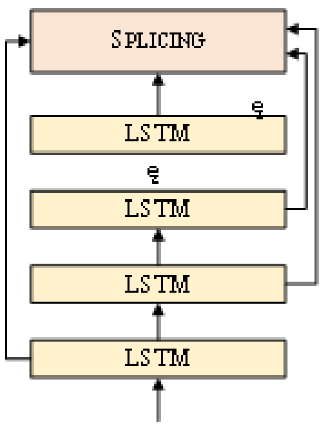

Figure 4: Structure of Enhanced LSTM.

The input features are assigned corresponding weights, which are then applied to different hidden unit states within the same time window, thereby enhancing the prediction performance of multidimensional time series data. Before delving into the details, we first introduce the Enhanced LSTM model proposed in this paper, which consists of three main components: Data preprocessing, multidimensional time series data prediction, and anomaly detection. In the data preprocessing phase, tasks such as data cleaning, sampling, and normalization are performed. In the multidimensional prediction phase, multiple sets of model parameters are used to train the base prediction model. These models generate various residual data sets, which are then used to construct multiple anomaly detection models based on multivariate Gaussian distribution [18]. Finally, the results from these different anomaly detection models are integrated. This approach helps avoid misjudgments that might arise from relying on a single model, with the voting mechanism playing a corrective role to some extent.

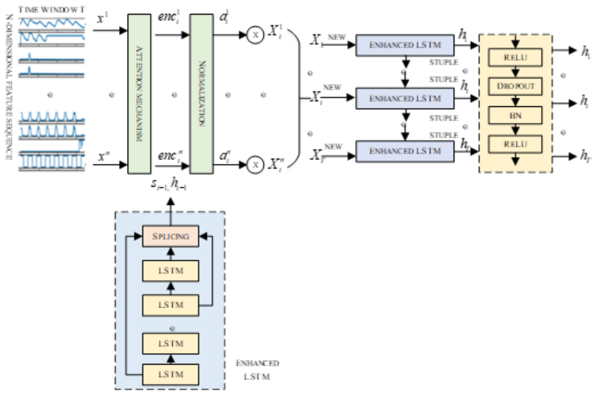

Figure 5: Coding stage.

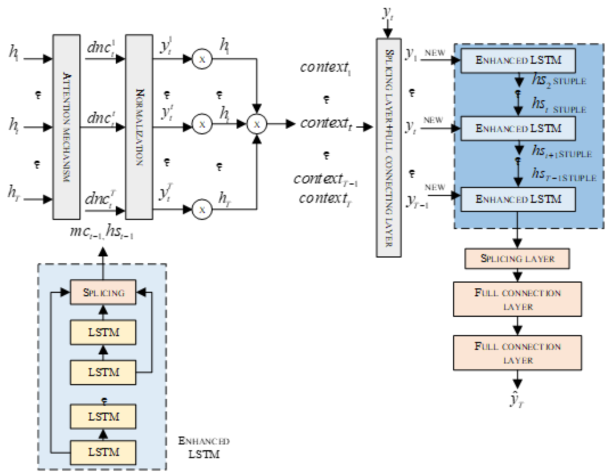

Figure 6: Decoding phase

Figures 5 and 6 enhanced LSTM-based attention prediction model first given a multi-dimensional time series dataset pair (X, y):

$$X = (x_1, x_2, x_3, \ldots, x_t) \tag{5}$$

$$y = (y_1, y_2, y_3, \ldots, y_{t-1}) \tag{6}$$

where n is the number of multidimensional features, y_R represents the value of the target feature at time t, and the state at $t-1$ is as Equation (7):

$$h_{t-1} = \left[\text{stuple}_1^1[h];\ \text{stuple}_1^2[h];\ \ldots;\ \text{stuple}_1^p[h]\right] \tag{7}$$

Similarly, the memory can be spliced to obtain the final state as in Equation (8):

$$s_{t-1} = \left[\text{stuple}_1^1[s];\ \text{stuple}_2^2[s];\ \ldots;\ \text{stuple}_p^p[s]\right] \tag{8}$$

The core item represents the memory cell state of the lth layer at time t:

$$x^e = \left(x_1^e, x_2^e, \ldots, x_q^e\right)^T \tag{9}$$

$$\alpha_q^e = \text{softmax}\left(\text{enc}^e\right) = \frac{\exp\left(\text{enc}_q^e\right)}{\sum\limits_{k=1}^{q} \exp\left(\text{enc}_k^e\right)} \tag{10}$$

In this paper, according to the above method, the weight can be obtained as Equation (11):

$$\alpha_t = \left(\alpha_t^1, \alpha_t^2, \ldots, \alpha_t^n\right) \tag{11}$$

The multidimensional features are multiplied by α_t as Equation (12):

$$x_t^{mm} = \left(\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \ldots, \alpha_t^n x_t^n\right)^T \tag{12}$$

The state h_t according to Equation (13), and use it as the input of the stage:

$$h_t,\ \text{stuple} = \text{EnhancedLSTM}\left(x_t^{mm},\ \text{stuple}\right) \tag{13}$$

In the model of this paper, it will be simplified. In this paper, Enhanced LSTM is still used for modeling. Information at time t need to be known. The same as the idea in the encoding stage, it is necessary to splicing the state, as shown in Equation (14):

$$h_{t-1}^s = \left[\text{stuple}_1^1[hs];\ \text{stuple}_2^1[hs];\ \ldots;\ \text{stuple}_p^1[hs]\right] \tag{14}$$

IV. Experiment

All experiments in this paper are based on high-performance computers. All experiments in this paper are carried out under 64-bit Windows 10, mainly using the integrated development environment Anaconda and PyCharm, and using the Python development language for model building and experimental verification. At the same time, the development of deep learning is required. Therefore, this experiment also involves the use of the TensorFlow deep learning platform and the Sklearn third-party library to build the experiment. At the same time, we also use GPU, CUDA8.0 to accelerate the training of deep learning model. The following basic models are selected for experimental comparison: Multilayer Perceptron (MLP) regression model, LSTM regression model. The above models all choose the mean square error function as the loss function, the activation function uses ReLU173, and the optimizer uses Adam.

SML2010 data set: It is mainly a data set collected by a monitor system in an indoor residence and published on UCL. The collection time is about 40 days, and the data is collected every 1 min, and the collected data is sampled every 15 min according to its average value, and finally a total of 4137 pieces of data are obtained. It includes 24 features, and only 18 features are reserved for modeling in our experiments. For example: Carbon dioxide, lighting, rainfall, wind speed, sunlight, solar irradiance, weather forecast temperature, etc. The indoor temperature (room) target feature of the prediction model in this paper, and the remaining 17 features are used as general features for modeling training. As shown in Figure 7, it is the data distribution of indoor temperature (room). Happening.

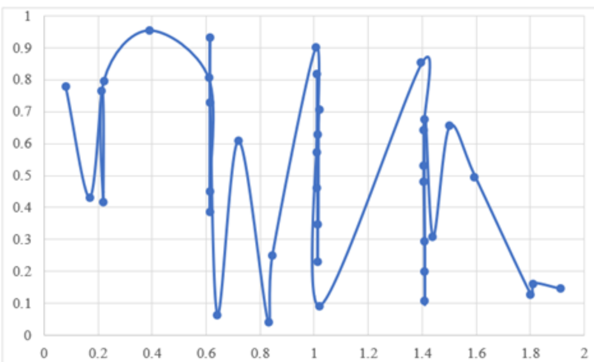

Figure 7: Partial data display of SML2010 dataset.

The maximum value is 1. When the value is smaller, the fitting effect of the model to the unknown data is worse. In the experimental part, three sets of experiments are mainly carried out, which are the comparison of the prediction and other basic models, the sensitivity of various performance evaluation indicators to parameters, and the use of random forests in machine learning algorithms (Random Forest) and the Variance threshold method in statistics to filter the features. In this group of experiments, this paper uses the three kinds of IoT time series datasets mentioned above, in which the SML2010 dataset and the IoT dataset are both time series datasets with multi-dimensional features, and the power dataset is a dataset with only one-dimensional features. The reason for using the power dataset here is not only for anomaly detection in the latter part, but also to verify that the model not only to multi-feature datasets but also to single-feature datasets.

The prediction model has the best results on the three indicators compared with other models. The value of the MLP regression model on the indicator MAE is 0.0518, while the value of the model in this article is only 0.0187. In contrast, the MLP regression model is nearly 3 times the value of the prediction model in this article on this indicator; On the indicator RMSE, the prediction model of this article is only 0.0235, while the value of the MLP regression model is 0.069, which is about 3 times that of the prediction model in this article. The results of the BiLSTM regression model and the LSTM regression model on this indicator are also significantly better than the MLP regression model, the main reason is that the MLP regression model has a simple structure, is not good at extracting multi-dimensional time series data features and has insufficient advantages in processing complex time series data.

Table 1 compares the prediction performance of different models on the SML2010 dataset. With RMSE, MAE and R2 metrics, the results show that our proposed model (Our Model) performs the best on all three evaluation metrics, with the lowest RMSE (0.0236) and MAE (0.0188), and the highest R2 (0.9999), suggesting that its prediction accuracy and data fit better than Bi LSTM, LSTM and MLP. Taken together, our model has the optimal prediction performance on this dataset.

Table 1: Comparison of prediction performance of different models on SML2010 dataset.

data set	Model	RMSE	MAE	R2
SML2010 Dataset	Our Model	0.0236	0.0188	0.9999
	Bi-LSTM	0.0323	0.0236	0.9996
	LSTM	0.0333	0.0234	0.9995
	MLP	0.0658	0.0519	0.9997

For the power data set, the prediction model proposed in this paper still shows obvious advantages. As shown in Figure 8, it is the situation predicted by the prediction model in this paper on the power data set. The model fits the data set. The effect is very good, and it can basically fit the real data. Table 2 is the comparison of the prediction the four models on the power data set. From the results in the table, the model has the best results on the three indicators, the main reason is that by comparing the general LSTM Retrofit to make Enhanced LSTM more capable of capturing relationships between long-term dependent data. From the results on the indicator RMSE, the prediction model in this paper is improved by nearly 1.8 percentage points compared with the LSTM regression model, and by nearly 3.5 percentage points compared with the BILSTM regression model. It indicators MAE that the prediction model has increased compared with the LSTM regression model and the BILSTM regression model, respectively, and compared with the MLP regression model, it has increased by nearly 4.4 percentage points. And on the indicator R2, the prediction model in this paper is improved by nearly 10 percentage points compared with the MLP regression model (see Figure 9).

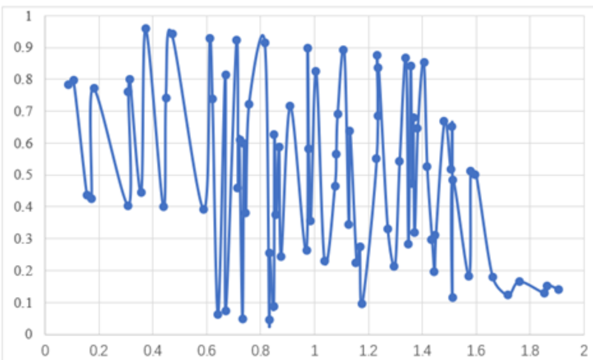

Figure 8: Forecast on SML2010 dataset.

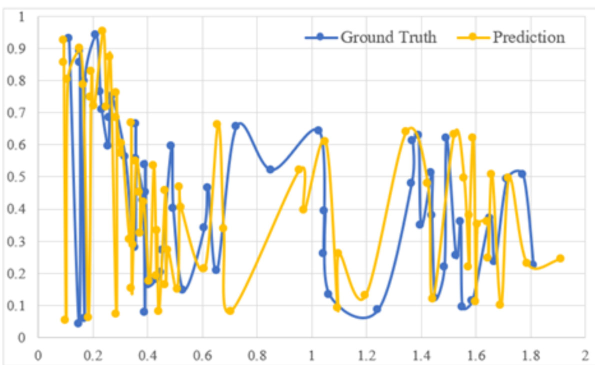

Figure 9: Power dataset forecast.

Table 2 compares the prediction performance of different models on the Power dataset. The results show that our model outperforms the other models in three metrics, RMSE (0.0303), MAE (0.0204) and R2 (0.9753), exhibiting lower error and

Table 2: Comparison of prediction performance of different models on power data sets.

data set	Model	RMSE	MAE	R2
Power data set	Our Model	0.0303	0.0204	0.9753
	Bi LSTM	0.0651	0.0462	0.9346
	LSTM	0.0489	0.0348	0.9633
	MLP	0.0893	0.0648	0.8767

higher goodness of fit. In contrast, Bi LSTM, LSTM and MLP have poorer prediction performance, especially in the RMSE and MAE metrics, indicating that the prediction accuracies of our proposed model on this dataset are significantly better than the other models.

V. Conclusion

We have presented a novel, education-centric framework that leverages 5G-IoT data and an attention-enhanced LSTM architecture to model and evaluate the evolving innovation capabilities of higher-education research teams. By designing a composite indicator system—spanning environmental conditions, equipment usage, and research outputs—and applying dynamic attention to these inputs, our model captures nuanced shifts in productivity and identifies anomalies indicative of operational inefficiencies. Empirical evaluations on three real-world campus datasets demonstrate that our approach outperforms MLP, LSTM, and BiLSTM baselines in both predictive accuracy and anomaly detection robustness. The deployment of our system in a university innovation dashboard further validates its utility: administrators receive real-time insights, enabling targeted support for under-resourced labs and optimization of resource allocation.

Data availability

The experimental data used to support the findings of this study are available from the corresponding author upon request. Ethical approval: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Qinhuangdao Vocational and Technical College (Project identification code: HBQVTC20230298765H) on 2023\12\23.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Zheng, J., Zhao, S., Lu, Y., et al. (2024). Exploration and practice of whole process talent cultivation in traditional Chinese medicine based on innovation and entrepreneurship education. *Researches in Higher Education of Pharmacy, 42*(1), 21.

[2] Zhou, H., Wu, T., Sun, K., & Zhang, C. (2022). Towards high accuracy pedestrian detection on edge GPUs. *Sensors, 22*(16), 5980.

[3] Wang, W., Kumar, N., Chen, J., et al. (2020). Realizing the potential of the internet of things for smart tourism with 5G and AI. *IEEE Network, 34*(6), 295–301.

[4] Chen, S., Hu, J., Shi, Y., et al. (2020). A vision of C-V2X: Technologies, field testing, and challenges with Chinese development. *IEEE Internet of Things Journal, 7*(5), 3872–3881.

[5] Guo, J., Wu, M., Wang, Z., et al. (2023). The cultivation and practice of college students' innovation and entrepreneurship ability in the perspective of college physics. *Adult and Higher Education, 5*(13), 22–27.

[6] Fortino, G., Savaglio, C., Spezzano, G., & Zhou, M. (2020). Internet of things as system of systems: A review of methodologies, frameworks, platforms, and tools. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51*(1), 223–236.

[7] Wu, S., & Guan, J. (2023). Exploration and practice of innovative teaching of molecular biology experiment course under the context of new agricultural science. *Sheng Wu Gong Cheng Xue Bao = Chinese Journal of Biotechnology, 39*(11), 4742–4749.

[8] Fan, G., Huang, Z., Sun, H., et al. (2023). Construction and analysis of evaluation model for medical students' innovation competency based on research-oriented biochemistry and molecular biology course in China. *Biochemistry and Molecular Biology Education, 51*(3), 263–275.

[9] Hossein Motlagh, N., Mohammadrezaei, M., Hunt, J., & Zakeri, B. (2020). Internet of things (IoT) and the energy sector. *Energies, 13*(2), 494.

[10] Li, M., Hu, D., Lal, C., et al. (2020). Blockchain-enabled secure energy trading with verifiable fairness in industrial internet of things. *IEEE Transactions on Industrial Informatics, 16*(10), 6564–6574.

[11] Abd El-Latif, A. A., Abd-El-Atty, B., Mazurczyk, W., et al. (2020). Secure data encryption based on quantum walks for 5G internet of things scenario. *IEEE Transactions on Network and Service Management, 17*(1), 118–131.

[12] Cheng, X., Zhang, J., Tu, Y., & Chen, B. (2022). Cyber situation perception for internet of things systems based on zero-day attack activities recognition within advanced persistent threat. *Concurrency and Computation: Practice and Experience, 34*(16), e6001.

[13] Shi, Q., Dong, B., He, T., et al. (2020). Progress in wearable electronics/photonics—Moving toward the era of artificial intelligence and internet of things. *InfoMat, 2*(6), 1131–1162.

[14] Siboni, S., Sachidananda, V., Meidan, Y., et al. (2019). Security testbed for internet-of-things devices. *IEEE Transactions on Reliability, 68*(1), 23–44.

[15] Wei, T., Feng, W., Chen, Y., et al. (2021). Hybrid satellite-terrestrial communication networks for the maritime internet of things: Key technologies, opportunities, and challenges. *IEEE Internet of Things Journal, 8*(11), 8910–8934.

[16] Zhang, Z., Zhang, C., Li, M., & Xie, T. (2020). Target positioning based on particle centroid drift in large-scale WSNs. *IEEE Access, 8*, 127709–127719.

[17] Huang, X., Ding, F., Meng, W., & Zhang, X. (2023). Exploration and practice of enhancing students' independent learning and innovation ability in biochemistry based on the "innovation and entrepreneurship competition". *Journal of Education, Humanities and Social Sciences, 19*, 218–222.

[18] Wang, W., Qiu, D., Chen, X., Yu, Z., et al. (2023). An empirical study on the evaluation system of innovation and entrepreneurship education in applied universities. *Computer Applications in Engineering Education, 31*(1), 100–116.