

Role of Artificial Intelligence Robots in Data Processing and Privacy Protection

Pei Sun^{a*}, Gang Wang^b, Zuohu Chen^c, Zhenguo Peng^d, Yuntao Dou^{1,e}

Gansu Tongxing Intelligent Technology Development Co., LTD., 730050, Gansu, China

Data processing and data privacy are two pressing social, ethical and technological issues that have emerged as a result of the transformational influence of the digital era. The focus of this paper is to provide an overview of the role of AI robots with a focus on the use of LSTM to predict Internet traffic and how to implement the LSTM/Differential Privacy combination to harness AI for creating Internet traffic prediction models. As indicated by the tests that have been performed to evaluate the implementation of Differential Privacy, while it may impact the overall quality of the model's output, a reduction in the model's prediction quality, the model greatly increases the model's level of data security and protection of users' privacy. Additionally, by means of multiple layers of neuronal configurations, the model is able to generalize better when an additional Dropout layer is added and the Adam optimizer is utilized to avoid overfitting. Results of the experiments illustrated that the model is able to make high-quality predictions even with Differential Privacy included in its design, specifically for both datasets with a RMSE of 87.93 (DS1), 88.46 (DS2) and MAE of 64.25 (DS1), 65.73 (DS2), thus demonstrating the ability of AI to strike a balance between data and privacy value. The research results indicate that a stacked LSTM model can accurately and effectively model temporal characteristics in Internet traffic allowing for enhanced network resource optimization and security alerting.

Keywords: data processing; privacy protection; artificial intelligence; long short-term memory; differential privacy

1. INTRODUCTION

The current digital age relies heavily on data as the driving force for economic and social growth. Because of the significant growth of both the amount of data collected and the speed at which new technologies are being developed and utilized for processing that data, there has come time for businesses to begin reviewing their policies and processes around how they collect and use data and the need for protecting user data (both personal and corporate) to ensure safe and adequate utilization [1]. The rapid growth of both Internet and IoT technologies has resulted in a dramatic increase in the volume of sensitive data being collected by

businesses and organizations of all sizes. The utilization of data for purposes such as improving operational efficiencies and market competitiveness, in addition to improving public service and scientific research, is evidence of the importance of utilizing data processing technology to optimize the use of data by organizations [2]. However, with the increased utilization of data also comes an increased potential for the risk of unauthorized access to data or privacy breaches. The disclosure of personal, private data can give rise to numerous negative outcomes, such as identity theft, destruction of property, and undermining of trust within society. Therefore, protecting individuals' personal privacy throughout the data processing lifecycle is a primary concern across society. Using privacy protection technologies is one of the main ways to resolve this issue. Protecting personal privacy involves more than simply encrypting or anonymizing personal data: it is a methodical approach that addresses numerous

*Corresponding author. ^aEmail: sunp95185@163.com, ^bwgang1616@163.com, ^cChengzHU198909@163.com, ^dpengzg1284@163.com, ^eDouyt0991@163.com

factors associated with data collection, storage, transmission and utilization [3–4]. Through the application of privacy protection technology, the personal privacy rights of users will be safeguarded and the confidentiality and integrity of data will be maintained during the processing operation [5].

In addition to being physical robots, artificial intelligence robots can be defined more generally as any automated or software system that has an intelligence-processing capability. The large amounts of data generated by businesses can be gathered automatically by artificial intelligence robots. Artificial intelligence robots can enhance both the speed and accuracy of processing this large volume of data. Using machine learning and deep learning algorithms, artificial intelligence robots can extract meaningful insights from large sets of data, perform sophisticated analyses and make predictions, and assist with decision making. During data processing, artificial intelligence robots also use advanced technologies to protect users' privacy and ensure that users' information is kept secure. Thus, it is important to study the role of artificial intelligence robots in the areas of data processing and privacy protection in order to encourage innovation and ensure social welfare.

The aim of the current study is to examine how AI technology is being deployed for the processing of large datasets and protecting individuals' privacy. Specifically, case studies are employed where a short-term memory network is used to predict Internet Traffic Based on Traffic History. The primary goal of this research is to examine ways to efficiently analyze and process the ever-increasing amounts of data without jeopardizing the safety of the data or the individual's privacy. The results suggest that the Internet traffic prediction model is adept at predicting the time series characteristics of Internet traffic and achieving predictions of Internet Traffic with a high level of accuracy, thus facilitating a higher level of quality of service, providing early warning of threats to Network Security, and optimizing the use of Network Resources. However, given that we are now in the era of Big Data, it is becoming increasingly important to develop solutions that will allow for an optimal balance between the use of data and the protection of individual privacy. The experimentation with the inclusion of differential privacy techniques resulted in a slight decrease in the degree of predictive performance of the prediction model; however, it was not sufficient to affect the overall degree of predictive power of the SLSTM model. This result indicates that AI robots will be of tremendous value in the processing of large amounts of data as well as in the protection of individuals' privacy, thus creating a win-win scenario for achieving the value of data while protecting the personal privacy rights of individuals.

2. RELATED WORK

Data processing refers to the entire process from data collection, organization, transformation to analysis, with the aim of extracting useful information and insights from data [6–7]. Privacy protection involves ensuring that personal or sensitive information is not leaked, abused, or accessed without authorization during data processing. Data processing can

improve decision-making efficiency and quality, and promote technological progress and business development; privacy protection ensures the information security of individuals and organizations, maintains public trust, and promotes the legality and compliance of data use [8]. The combination of the two can fully utilize the value of data and prevent potential privacy risks. Mahajan and Reddy and other scholars aimed to address the difficulty in processing patient genetic data under the Healthcare 4.0 standard due to security risks [9]. A Healthcare 4.0 framework leveraging lightweight encryption and blockchain technology was proposed for securely managing microarray gene expression data, thereby guaranteeing the confidentiality and integrity of genomic data storage and transmission. The results showed that this method improved computational efficiency and security, reducing encryption and decryption time by 325 milliseconds and 20 milliseconds, respectively. Jia et al. proposed a distributed ledger-based collaborative learning framework in the Industrial Internet of Things (IIoT) aimed at mitigating privacy and sensitivity concerns associated with sharing significant volumes of data produced by diverse industrial IoT endpoints, and further developed a privacy-preserving data aggregation mechanism derived from this framework [10]. Extensive experimental results indicated that the model had better performance on the selected indicators. However, it did not mention its applicability in different scenarios. Singh et al. proposed intelligent healthcare privacy protection security architecture based on blockchain and federated learning to address the issue of user IoT data privacy protection in smart healthcare environments in smart cities [11]. The experimental results showed that the proposed intelligent healthcare privacy protection security architecture based on blockchain and federated learning had good performance, but lacked applicability and performance evaluation in different scenarios.

Artificial intelligence (AI) refers to the technology whereby computer systems perform tasks by simulating human intelligent behavior [12–13]. The role of AI is significant, as it provides solutions to complex problems through automation and intelligence, leading to increased efficiency and productivity, and creating new avenues for innovation and development in every industry, thus improving the quality of human life and providing socio-economic benefits [14]. AI has excellent potential application value in many areas including, but not limited to, medical diagnosis, autonomous vehicles, financial analysis and intelligent customer service [15–16]. According to Chen et al., AI has been implemented in education since the introduction of computer-based pedagogy, then the shift to web-based intelligent education systems and then again with the introduction of computer systems, humanoid robots, and web-based chatbots for teaching and student management [17]. Using Machine Learning and Adaptive Technology, educational institutions can create personalized, custom-made courses and learning materials, which have significantly improved students' learning experiences and educational outcomes. Additionally, Alowais et al. describe how AI is poised to change clinical practice by improving patient care and quality of life [18]. A review of the literature reveals that AI is being used in the healthcare industry to improve diagnostic accuracy, to

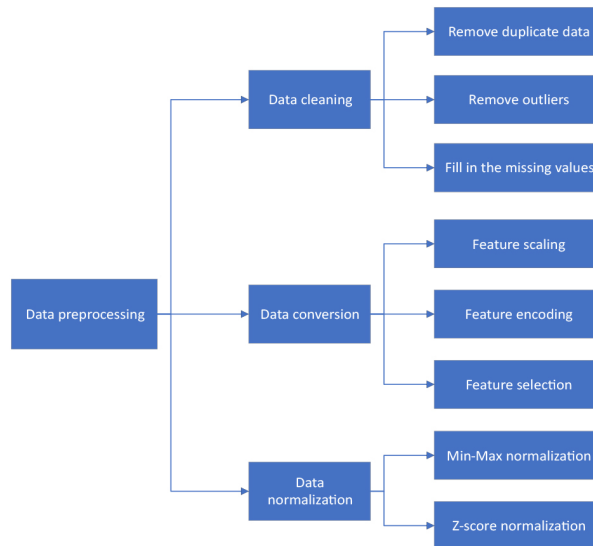


Figure 1 Data preprocessing.

provide better treatment decisions, and to support clinical laboratory testing. However, data privacy remains an issue. Yüksel et al.'s research on AI application in engineering and design found that using AI techniques increased the speed, accuracy, thoroughness, and efficiency of the engineering design process [19]. Sohail noted that due to advances in scientific computing and research, small datasets and low-dimensional problems are no longer difficult to optimize [20]. However, large, random, and multidimensional datasets result in complex optimization problems. To overcome this issue, he proposed a genetic algorithm model based on the assumption of natural selection, which he demonstrated in his research, as an inexpensive and effective means of optimizing large datasets in the field of computational biology.

3. APPLICATION OF ARTIFICIAL INTELLIGENCE ROBOTS IN DATA PROCESSING

Artificial intelligence (AI) is now considered the primary driver of the future of data processing through the effective and intelligent use of data [21–22]. With the massive amount and complexity of data growing exponentially, the traditional methods of data processing will not be able to keep up with the demand. AI has disrupted all of the traditional methods of data processing with its ability to learn and predict and, at the same time, develop models to support decision making [23–24]. This section creates a model for predicting internet traffic based on long short-term memory (LSTM) to reinforce the importance of ai in efficient data processing.

3.1 Data preprocessing

In the AI domain, data preparation—encompassing the transformation and refinement of data before model ingestion—is a vital stage in the pipeline. As shown in Figure 1, data processing primarily consists of three steps: data conversion, data cleaning, and data normalization.

In Figure 1, data cleaning ensures data integrity and accuracy by removing duplicates, processing outliers, and filling missing values. Data transformation optimizes data structure through feature scaling, feature encoding, and feature selection to meet the needs of machine learning. Feature scaling improves model performance; feature encoding converts non-numeric features into digital form; feature selection selects relevant features to simplify the model. The role of data normalization is to adjust the data distribution and ensure that all features are within a uniform and appropriate range for easy algorithm understanding and processing.

3.2 LSTM network

One particular kind of recurrent neural network (RNN) is the LSTM network [25–26]. LSTM outperforms conventional RNNs in extended sequence training by employing a gating system that consists of an input gate, an output gate, and a forget gate to prevent gradient vanishing and gradient explosion issues. Long-term temporal dependencies can thus be better captured and remembered by the model [27–28]. Figure 2 shows how the LSTM is structured.

In Figure 2, the memory unit runs through the entire sequence and stores key information, which can be considered as the backbone of information transmission. LSTM performs addition and subtraction operations on it through a gate mechanism, allowing it to retain or discard information. The forget gate determines how much information the memory unit should discard based on the current input x_t and the previous hidden state h_{t-1} through a sigmoid function. In Internet traffic prediction, input data includes historical traffic data and key time characteristics. The specific formula is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

The input gate, which is made up of candidate memory and input gate gating, controls the way that fresh data updates the memory unit. The input gate gate i_t determines how much

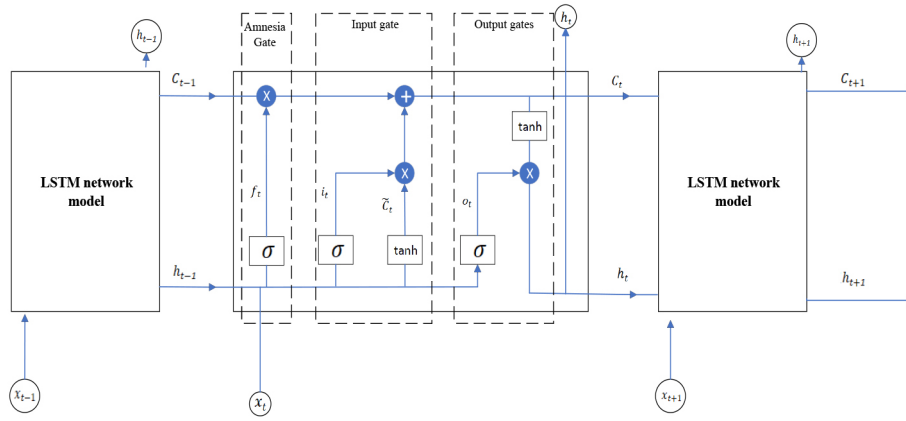


Figure 2 LSTM structure diagram.

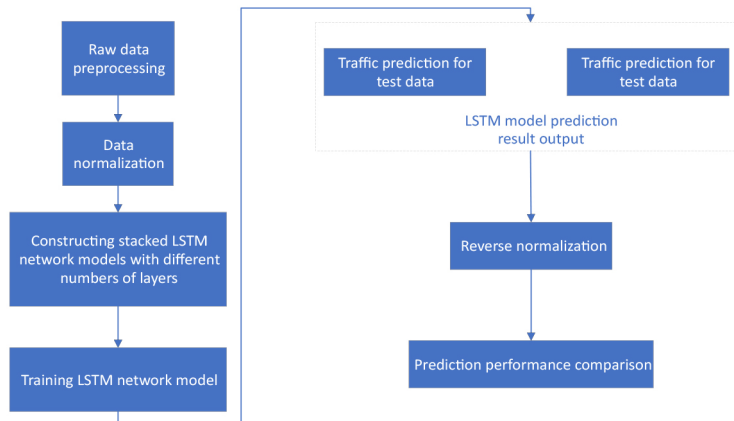


Figure 3 Stacked LSTM network internet traffic prediction flow chart.

new information C_{t+1} is written through the sigmoid function. Candidate memory $\tilde{C}_t$ generates candidate memories through the tanh function. The specific formulas are:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

Finally, the new memory unit state C_t is determined by both the forget gate and the input gate, and the formula is:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

The output gate o_t controls how much information memory unit state C_t outputs to the next hidden state. The formula is as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

The new hidden state h_t is determined by modulating the output gate of the memory unit state through the tanh function:

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

In practical applications, the representation ability of single-layer LSTM networks is relatively weak and may not be sufficient to capture high-order features and relationships in complex time series. Moreover, when dealing with very long time series or sequences with strong dependencies, there may be insufficient memory capacity.

3.3 Stacked LSTM network

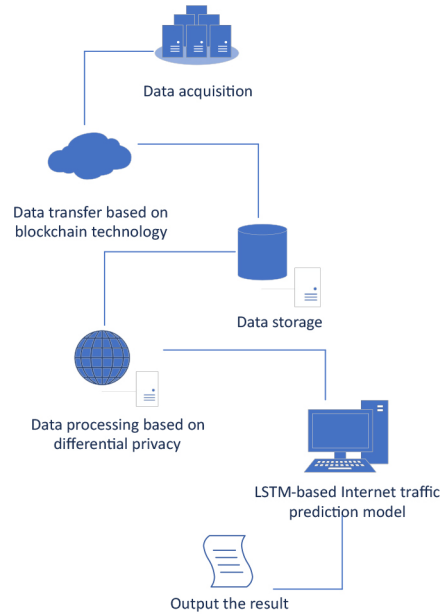
The stacked LSTM network consists of multiple LSTM layers, and the output of each layer serves as the input of the next layer. In each layer of LSTM, the information of the input sequence is processed by gating mechanisms and memory units to generate a more abstract and advanced representation. The flow of stacked LSTM networks used to predict Internet traffic in this study is shown in Figure 3.

Firstly, the raw data is preprocessed and the Min-Max scaling function is used to normalize the raw data to the $[0, 1]$ interval. Then, four different stacked LSTM models are established to predict the data. After training, these models are used to predict traffic on test data. Finally, the predicted results are reconstructed and the predictive performance of each model is compared and analyzed. Table 1 displays the relevant parameter settings.

To mitigate overfitting, this study incorporates a dropout layer configured with a dropout probability of 0.2 following each LSTM layer. Through iterative updates of the model parameters using the training dataset, the Adam optimization algorithm adaptively adjusts the learning rate to suit varying parameter landscapes. These strategies enhance the network's generalization capability and boost its predictive accuracy on unseen test data.

Table 1 Parameter settings for stacked LSTM network model.

Parameter name	Parameter settings			
LSTM layers (layer)	1	2	3	4
Number of hidden layer neurons	256	(256,128)	(256,128,64)	(256,128,64,32)
Dropout			0.2	
Optimizer			Adam	
The number of data samples captured for a single training			256	

**Figure 4** Overall Internet traffic prediction model.

3.4 Data storage based on blockchain

The technology known as blockchain is a way to store data securely in a distributed manner and to keep all of the data intact and unchangeable by means of encryption and consensus algorithms. The blockchain works best with high trust and high traceability use cases for storing data, and uses distributed networks to help create a greater degree of transparency and to secure the information being exchanged. The decentralized and tamperproof design of the blockchain provides the ideal foundation for sharing data and optimizing models. Storing data in many nodes in a network increases security and decreases the likelihood of an attack, thus increasing the availability and longevity of the data.

3.5 Applying differential privacy

Differential privacy, a mathematical technique, seeks to protect the private nature of individual data while allowing for the effective statistical analysis of datasets. The key idea is to introduce some amount of random (and controlled) “noise” into the individual data records so that the output from performing analysis on the complete dataset with and without a specific individual cannot be distinguished, thereby masking the individual’s identity. Even if the analysis results are leaked, the existence of any individual cannot be

disclosed, and the goal of privacy protection is achieved. The specific steps for applying differential privacy technology are as follows: first, the data is collected and preprocessed. After obtaining the data, it is cleaned and preprocessed, including removing or blurring the user’s personal identification information. Then, the data is aggregated by time window to reduce the sensitivity of individual user data. An appropriate differential privacy mechanism is selected, and a reasonable noise addition strategy is set according to the preset privacy budget. For the differential privacy mechanism M of function f , it satisfies:

$$\Pr [M(f(D)) \leq e^\epsilon \Pr [M(f(D'))]] \quad (7)$$

Where, D and D' are two datasets that differ by one data individual, and ϵ is the privacy budget that controls the degree of privacy leakage. The added noise makes it difficult to determine whether specific individuals exist in the dataset even after the attacker obtains the result $M(f(D))$. The dataset is then divided into training and test sets. The test set is used to determine the LSTM model’s prediction performance after it has been trained on the training set. In order to guarantee that LSTM models can handle and forecast noisy data in real-world applications, it is crucial to employ training data with differential privacy processing.

Figure 4 displays the Internet traffic forecast model that applies blockchain and differential privacy technologies. By striking a balance between data use and privacy protection, differential privacy technology only successfully protects

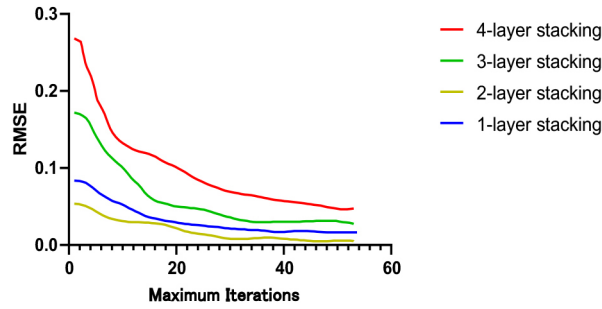


Figure 5 Iterative training results of LSTM with different layers.

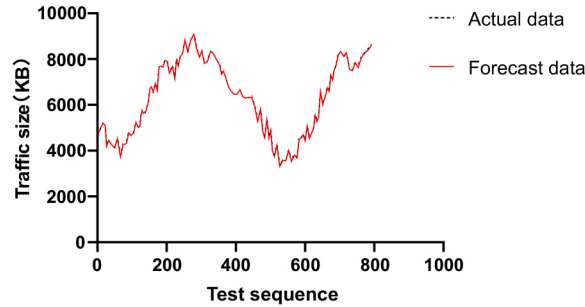


Figure 6 Internet traffic prediction results of two-layer stacked LSTM model (ISP-Data).

users' sensitive information, but also guarantees that the model can gather sufficient data to make accurate traffic predictions based on the principle of protecting privacy when developing the Internet traffic prediction model of LSTM networks. This enhances the model's credibility and usefulness.

4. EXPERIMENTS

This experiment uses Python 3.7.11 and TensorFlow Keras for deep learning experiments. The computer is configured with an AMD Ryzen 7 5800X processor, a base frequency of 3.8 GHz, an NVIDIA GeForce RTX 3080 graphics processor, equipped with 10GB GDDR6X memory, and a 1TB hard drive size. The mean absolute error (MAE) and root mean square error (RMSE) are the metrics utilized to assess the findings in this paper. The following are the calculation formulas:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (9)$$

Where, y_i represents the real value of Internet traffic data, and $\hat{y}_i$ represents the predicted value of Internet traffic data. The data is divided into two parts, which are respectively from ISP (Internet Service Provider) Internet traffic (represented by ISP-Data) in eight European cities and Kaggle's Internet Traffic Volume Data Set (represented by

ITV-Data). ISP-Data contains 8000 groups of network data, and 6000 groups of network data are extracted from ITV-Data. 90% of the dataset in the article serves as training data for the model, and the remaining 10% serves as testing data. By modifying the stacked layers, the experimental approach assesses the model's performance and identifies the optimal prediction model. Next, a comparison of the stacked LSTM model with a historical average (HA) and a deep belief network (DBN) is performed.

The experimental results of LSTM models with different layers are shown in Figure 5.

From Figure 5, it can be learnt that the stacked LSTM model performs well in the iterative training process, and the RMSE of the first ten iterations shows a significant downward trend. The RMSE of the 11-20 iterations slows down, and after 20 iterations, it tends to be stable and less volatile. Moreover, the RMSE of the stacked LSTM network with 2 layers is the smallest, indicating that the model has an optimal solution. As a result, while predicting Internet traffic, the stacked LSTM model's number of layers is set to 2, and its number of iterations is set to 20. As displayed in Figure 6, this configuration is utilized to forecast the Internet traffic of ISP-Data data. The findings are compared with the actual findings. Figure 7 displays a comparison of the actual ITV-Data results with the expected results.

As shown in Figures 6 and 7, the predicted outputs of the 2-layer LSTM model applied to both datasets closely align with the ground-truth values, demonstrating high predictive accuracy.

To further evaluate the forecasting capability of the stacked LSTM architecture, this study benchmarks the optimal stacked LSTM model against a Deep Belief Network (DBN)

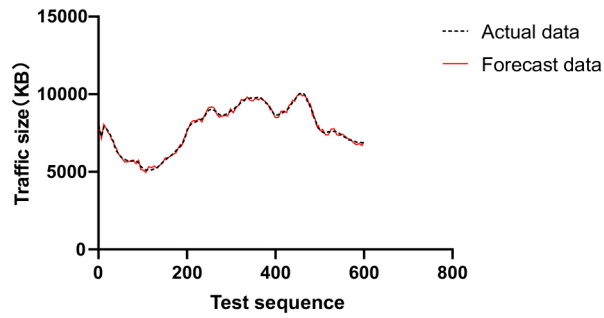


Figure 7 Internet traffic prediction results of two-layer stacked LSTM model (ITV-Data).

Table 2 Comparison of prediction results of different prediction models.

Method	RMSE		MAE	
	ISP-Data	ITV-Data	ISP-Data	ITV-Data
DBN	101.07	99.23	66.74	66.15
HA	124.74	125.17	82.79	82.66
2-layer stacked LSTM model	85.53	86.37	64.79	65.02

Table 3 Experimental results of two models without applying differential privacy and two models with the application of differential privacy.

Method	RMSE		MAE	
	ISP-Data	ITV-Data	ISP-Data	ITV-Data
DBN model without differential privacy	100.23	100.79	67.84	67.91
DBN model with differential privacy	101.15	102.28	68.53	68.69
LSTM model without differential privacy	86.74	87.31	63.93	64.14
LSTM model with differential privacy	87.93	88.46	64.25	65.73

and a Historical Average (HA) baseline using the ISP-Data and ITV-Data datasets. The comparative results are summarized in Table 2.

From Table 2, it can be seen that among the three prediction methods, HA performs the worst, followed by DBN. Compared to the other two techniques, the strategy adopted in this article performs better, suggesting that the prediction method has exceptional performance and high reliability of the prediction findings.

To evaluate the effect of implementing differential privacy on the model's prediction accuracy, this study conducts experiments on the ISP-Data dataset and ITV-Data dataset by integrating differential privacy into the stacked LSTM model and DBN model. The results are compared with those of the two models that do not have differential privacy. Table 3 displays the final experimental results.

As shown in Table 3, the DBN model's RMSE and MAE increase by 0.92 and 1.49 on two datasets, respectively, when differential privacy technology is applied, while the LSTM model's RMSE and MAE increase by 1.19 and 1.15, and by 0.32 and 1.59 on two datasets, respectively. Predictive performance is marginally reduced overall when differential privacy technology is used, although this effect is not statistically significant. In LSTM models in particular, although the increase in MAE is relatively large on the ITV-Data dataset, considering the original cardinality of prediction error, this increase is still within an acceptable range. This shows that in applications such as Internet traffic prediction, even if differential privacy technology is applied,

the model can still maintain a relatively stable prediction performance, thus protecting data privacy to a certain extent without seriously damaging the practicality of the model.

5. CONCLUSIONS

This study showed the better data processing skills of AI robots and achieved outstanding results in the field of Internet traffic prediction using stacked long short-term memory networks, a state-of-the-art AI technology. The advancement and promise of AI technology in the area of privacy protection are demonstrated by the creative combination of differential privacy algorithms, which can increase prediction accuracy while maintaining data security and user privacy. The experimental data shows that, without implementing differential privacy protection, the RMSE values of the stacked LSTM model for Internet traffic prediction of the two data sets are 86.74 and 87.31, and the MAE values are 63.93 and 64.14, indicating the powerful effectiveness of AI in data processing. The incorporation of Differential Privacy mechanisms increases RMSE and MAE performance metrics from base levels to 1.19 to 1.15 and 0.32 to 1.59 respectively. Ultimately, this trade-off has proven to secure substantial gains in privacy protection over those associated with the original predictive performance. Going forward, through optimal model design, increased privacy parameterization, enhanced data processing abilities, and the use of the latest algorithms, we can achieve both

increased predictive accuracy and continuing improvements in privacy protection, increased computational processing speed and greater resource utilization efficiency, and the parallel development of AI technologies in terms of both security and efficiency.

REFERENCES

1. Yaqoob, I., Salah, K., Jayaraman, R., & Al-Hammadi, Y. (2022). Blockchain for healthcare data management: opportunities, challenges, and future recommendations. *Neural Computing and Applications*, 34(7), 1–16. <https://doi.org/10.1007/s00521-020-05519-w>
2. Liang, W., Yang, Y., Yang, C., Hu, Y., Xie, S., Li, K. C., & Cao, J. (2022). PDPChain: A consortium blockchain-based privacy protection scheme for personal data. *IEEE Transactions on Reliability*, 72(2), 586–598. <https://doi.org/10.1109/TR.2022.3190932>
3. Liu, C., Le Roux, L., Körner, C., Tabaste, O., Lacan, F., & Bigot, S. (2020). Digital twin-enabled collaborative data management for metal additive manufacturing systems. *Journal of Manufacturing Systems*, 62, 857–874. <https://doi.org/10.1016/j.jmsy.2020.05.010>
4. Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1), 91–99. <https://doi.org/10.1016/j.gltp.2022.04.020>
5. Diaz Linares, I., Pardo, A., Patch, E., Dehghantanha, A., & Choo, K. K. R. (2022). IoT privacy, security and forensics challenges: an unmanned aerial vehicle (uav) case study. *Handbook of Big Data Analytics and Forensics*, 7–39. https://doi.org/10.1007/978-3-030-74753-4_2
6. Lăzăroi, G., Andronie, M., Iatagan, M., Geamănu, M., Ștefănescu, R., & Dijmărescu, I. (2022). Deep learning-assisted smart process planning, robotic wireless sensor networks, and geospatial big data management algorithms in the internet of manufacturing things. *ISPRS International Journal of Geo-Information*, 11(5), 277. <https://doi.org/10.3390/ijgi11050277>
7. Demichev, V., Szyrwił, L., Yu, F., Teo, G. C., Rosenberger, G., Niewianda, A., ... & Ralser, M. (2022). dia-PASEF data analysis using FragPipe and DIA-NN for deep proteomics of low sample amounts. *Nature Communications*, 13(1), 3944. <https://doi.org/10.1038/s41467-022-31492-0>
8. Zhao, Y., & Chen, J. (2022). A survey on differential privacy for unstructured data content. *ACM Computing Surveys (CSUR)*, 54(10s), 1–28. <https://doi.org/10.1145/3490237>
9. Mahajan, H., & Reddy, K. T. V. (2024). Secure gene profile data processing using lightweight cryptography and blockchain. *Cluster Computing*, 27(3), 2785–2803. <https://doi.org/10.1007/s10586-023-04123-6>
10. Jia, B., Zhang, X., Liu, J., Zhang, Y., Huang, K., & Liang, Y. (2021). Blockchain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT. *IEEE Transactions on Industrial Informatics*, 18(6), 4049–4058. <https://doi.org/10.1109/TII.2021.3085960>
11. Singh, S., Rathore, S., Alfarraj, O., Tolba, A., & Yoon, B. (2022). A framework for privacy-preservation of IoT healthcare data using Federated Learning and blockchain technology. *Future Generation Computer Systems*, 129, 380–388. <https://doi.org/10.1016/j.future.2021.11.028>
12. Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., ... & Aspuru-Guzik, A. (2022). On scientific understanding with artificial intelligence. *Nature Reviews Physics*, 4(12), 761–769. <https://doi.org/10.1038/s42254-022-00518-3>
13. Ahmed, I., Jeon, G., & Piccialli, F. (2022). From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where. *IEEE Transactions on Industrial Informatics*, 18(8), 5031–5042. <https://doi.org/10.1109/TII.2022.3146552>
14. Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., ... & Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972), 47–60. <https://doi.org/10.1038/s41586-023-06221-2>
15. Khan, B., Fatima, H., Qureshi, A., Kumar, S., Hanan, A., Hussain, J., & Abdullah, S. (2023). Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. *Biomedical Materials & Devices*, 1(2), 731–738. <https://doi.org/10.1007/s44174-023-00063-2>
16. Gour, S., Kasera, R. K., Acharjee, T., & Niranjan Singh, W. (2026). Privacy and security of smart irrigation data using IOTA distributed ledger. *Cyber-Physical Systems*, 12(1), 1–30.
17. Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
18. Alowais, S. A., Alghamdi, S. S., Alsuehaby, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., ... & Albekairy, A. M. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education*, 23(1), 689. <https://doi.org/10.1186/s12909-023-04698-z>
19. Yüksel, N., Börklü, H. R., Sezer, H. K., & Canyurt, O. E. (2023). Review of artificial intelligence applications in engineering design perspective. *Engineering Applications of Artificial Intelligence*, 118, 105697. <https://doi.org/10.1016/j.engappai.2022.105697>
20. Sohail, A. (2023). Genetic algorithms in the fields of artificial intelligence and data sciences. *Annals of Data Science*, 10(4), 1007–1018. <https://doi.org/10.1007/s40745-021-00354-9>
21. Reddy, S. (2022). Explainability and artificial intelligence in medicine. *The Lancet Digital Health*, 4(4), e214–e215. [https://doi.org/10.1016/S2589-7500\(22\)00029-2](https://doi.org/10.1016/S2589-7500(22)00029-2)
22. Kirmani, A. R. (2023). Artificial intelligence-enabled science poetry. *ACS Energy Letters*, 8(1), 574–576. <https://doi.org/10.1021/acsenenergylett.2c02758>
23. Ossowska, A., Kusiak, A., & Swietlik, D. (2022). Artificial intelligence in dentistry—Narrative review. *International Journal of Environmental Research and Public Health*, 19(6), 3449. <https://doi.org/10.3390/ijerph19063449>
24. Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
25. Song, T., Liu, C., Wu, R., Jin, Y., & Jiang, D. (2022). A hierarchical scheme for remaining useful life prediction with long short-term memory networks. *Neurocomputing*, 487, 22–33. <https://doi.org/10.1016/j.neucom.2022.02.032>
26. Granata, F., Di Nunno, F., & de Marinis, G. (2022). Stacked machine learning algorithms and bidirectional long short-term memory networks for multi-step ahead streamflow forecasting: A comparative study. *Journal of Hydrology*, 613(1–4), 128431. <https://doi.org/10.1016/j.jhydrol.2022.128431>
27. Lui, C. F., Liu, Y., & Xie, M. (2022). A supervised bidirectional long short-term memory network for data-driven dynamic soft sensor modeling. *IEEE Transactions on Instrumentation and Measurement*, 71, 1–13. <https://doi.org/10.1109/TIM.2022.3152856>
28. Hsu, C. Y., Lu, Y. W., & Yan, J. H. (2022). Temporal convolution-based long-short term memory network with

- attention mechanism for remaining useful life prediction. *IEEE Transactions on Semiconductor Manufacturing*, 35(2), 220–228. <https://doi.org/10.1109/TSM.2022.3164578>
29. Liu, Y., He, W., Zhu, J., Yang, R. & Yang, X. (2026). Intelligent decision-making system for power grid dispatching based on generative AI. *International Journal of Engineering Intelligent Systems*, 34(1), 5–16.
30. Hou, Y. (2026). Accounting information cloud data integrity verification algorithm based on data mining. *International Journal of Engineering Intelligent Systems*, 34(1), 73–78.

