processes

MDPI

Article

Research on Stuck Pipe Prediction Based on Supervised and Unsupervised Ensemble Learning

Boyi Xia [1], Yiwei Wang [2], Qihao Li [1], Xianzhi Song [2], Zhaopeng Zhu [3,*], Muchen Liu [1] and Yanlong Yang [2]

[1] School of Artificial Intelligence, China University of Petroleum (Beijing), Beijing 102249, China
[2] School of Petroleum Engineering, China University of Petroleum (Beijing), Beijing 102249, China
[3] School of Mechanical and Transportation Engineering, China University of Petroleum (Beijing), Beijing 102249, China
* Correspondence: zhuzp@cup.edu.cn; Tel.: +86-010-89732176

Abstract

Stuck pipe is a common and serious accident in oil drilling processes, which may lead to huge economic losses and safety risks. In recent years, the rapid development of artificial intelligence technology has provided new ideas for stuck pipe prediction. Existing intelligent prediction studies on stuck pipe mostly focus on the optimization and application of a single unsupervised or supervised algorithm, or the research on simple ensemble learning of these two types of algorithms. This paper proposes a stuck pipe prediction method based on mechanism constraints and a deep learning ensemble model. By integrating the advantages of mechanism constraints and various time-series data processing models, this method achieves accurate prediction of stuck pipe. The method first performs preprocessing, feature engineering, and mechanism constraints on multi-parameter time-series data during drilling, then constructs three models, namely, Autoencoder, BiLSTM, and Transformer, for feature extraction and preliminary prediction, respectively. Finally, it integrates the prediction results of multiple models through a meta-model to improve prediction accuracy. The experimental results show that after introducing mechanism constraints, the accuracy of each model increases by an average of 10%. For the stuck pipe prediction task, the accuracy and precision of the proposed ensemble model reach 90.1% and 95.9%, respectively. Compared with single models, the ensemble model achieves an optimal balance between the false alarm rate and missing alarm rate, which are 7.7% and 11.0%, respectively. Its comprehensive performance is significantly better than that of single models, which can provide effective risk early warning for drilling operations.

Keywords: stuck pipe prediction; deep learning; ensemble learning; mechanism constraint

check for updates

Academic Editor: Zhibin Lin

Received: 19 August 2025
Revised: 28 September 2025
Accepted: 13 October 2025
Published: 16 October 2025

Citation: Xia, B.; Wang, Y.; Li, Q.; Song, X.; Zhu, Z.; Liu, M.; Yang, Y. Research on Stuck Pipe Prediction Based on Supervised and Unsupervised Ensemble Learning. *Processes* **2025**, *13*, 3309. https://doi.org/10.3390/pr13103309

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the process of oil drilling, stuck pipe is one of the common and highly hazardous accidents in drilling operations. Stuck pipe accidents not only lead to prolonged drilling cycles and increased costs, but in severe cases, they may even result in major losses, such as wellbore abandonment. Realizing real-time and accurate prediction of stuck pipe risks is of great significance for improving drilling efficiency and reducing operational risks [1]. Traditional stuck pipe prediction mainly relies on the experiential judgment of on-site engineers, who identify signs of stuck pipe by observing changes in drilling parameters. This method has high subjectivity and limited accuracy, making it difficult to meet the demands of complex drilling environments. With the development of artificial intelligence technology, data-driven stuck pipe prediction methods have gradually emerged [2].

In recent years, intelligent models have been widely applied in the field of stuck pipe prediction. Aditi et al. [3] compared the performance of various machine learning algorithms, pointing out the advantages of extra trees in handling high-dimensional data and extracting hidden patterns. Salvatore et al. [4] correlated precursor events of stuck pipe with stuck pipe risks through a hidden Markov model (HMM) and classified different levels of alarms. Wu Honglin et al. [5] used the Mann–Kendall method to screen key features with trend characteristics and proposed an evaluation model based on the polynomial fitting algorithm (PFA). Zhang Tao et al. [6] constructed a prediction model integrating variational mode decomposition (VMD) features and physical constraints based on near-bit measurement data. Siruvuri et al. [7] utilized a convolutional neural network model to monitor and prevent the occurrence of stuck pipe accidents. Zhu Shuo et al. [8] conducted intelligent prediction of stuck pipes based on the fusion of mechanism models and data models. Chamkalani et al. [9] employed drilling parameters and a hybrid least squares support vector machine model to predict stuck pipe phenomena during drilling processes. Yi Siqi et al. [10] proposed a random forest stuck pipe risk assessment method based on SMOTE undersampling. Ahmed et al. [11] used a sliding window machine learning approach to predict drilling of stuck pipes. Zhang Wanxing et al. [12] carried out research on prediction and prevention of complex downhole conditions, such as stuck pipe and well kick, using a generative adversarial network long short-term memory network (GAN-LSTM). Al-Baiyat and Abbas et al. [13,14] optimized prediction performance by comparing the performance of artificial neural networks and support vector machine models in stuck pipe prediction. Liu Zihao et al. [15] improved support vector machines based on ensemble learning, together with the particle swarm optimization algorithm, to optimize intelligent prediction results of stuck pipe. Zhang Tao et al. [16] optimized parameters for the support vector machine (SVM) using the SCNGO algorithm and applied the SCNGO-SVM model to stuck pipe prediction. Existing research on stuck pipe prediction has the following limitations: (1) For single-algorithm optimization, although it can capture some features, it is difficult to balance temporal dependence and anomaly detection simultaneously. (2) Simple ensemble methods are mostly limited to the fusion of supervised models and fail to integrate unsupervised feature extraction. (3) There is a lack of drilling mechanism constraints, and pure data-driven approaches are divorced from engineering practice.

Based on the above background, this paper proposes a stuck pipe prediction method based on a deep learning ensemble model. Ensemble learning can effectively improve the stability and accuracy of stuck pipe prediction by integrating the advantages of multiple models. This method designs a complete drilling data preprocessing and feature engineering process, including missing value handling, standardization, and mechanism constraints. It proposes a novel ensemble learning architecture, constructing an ensemble model integrating Autoencoder, bidirectional LSTM, and Transformer to address the limitations of single models. It also embeds drilling mechanism constraints into data preprocessing to solve the problem that pure data-driven approaches are divorced from engineering practice. By making full use of mechanism constraints and the ability of different models to process time-series data, it achieves multi-scale feature complementarity and further improves prediction performance.

2. Research Methods

The dataset used in this study comes from field measured data of 10 wells in a block of an oilfield, including real-time logging data and drilling logs. The real-time logging data cover parameters such as well depth, bit position, weight on bit, rotary speed, torque, hook height, hook load, pump stroke, and total hydrocarbon content. The drilling logs record on-site stuck pipe cases, including the time period, depth position, and cause of

stuck pipe occurrence. The above dataset provides training and test verification samples for the training of intelligent models and supports the stuck pipe prediction results.

2.1. Data Preprocessing and Feature Engineering

In the data preprocessing stage, the original logging data are cleaned, including missing value filling and standardization, to provide high-quality data for subsequent analysis.

Missing value filling uses a combination of forward filling and backward filling. During the drilling process, many parameters, such as hook height and bit position, have physical continuity and will not change abruptly in a short time. Using forward filling to process missing values reasonably retains the physical characteristics of these parameters.

Z-score normalization is used for data normalization. Different features (such as hook height and hook load) may have significant differences in dimensions and value ranges. If normalization is not performed, the model will be biased toward features with larger value ranges, leading to instability in the learning process. Standardization does not change the original distribution pattern of logging data, but only translates and scales it, accelerating the convergence speed of gradient descent during neural network training:

$$x_{standard} = \frac{x - \mu}{\sigma}, \tag{1}$$

where, x: original feature value, μ: mean of training data, and σ: standard deviation of training data.

2.2. Feature Engineering

Based on the mechanism knowledge in the drilling field, 13 key features reflecting drilling conditions are initially extracted from the original data, including well depth (DMEA), bit position, hook height, hook load, weight on bit, rotary speed, torque, standpipe pressure, casing pressure, inlet density, outlet density, inlet flow rate, and outlet flow rate. Then, combined with on-site practice and expert experience, 7 features are selected: bit position (DBTM), hook height (HKHT), hook load (HKLD), weight on bit (WOB), rotary speed (RPM), torque (TQA), and standpipe pressure (SPP). The original data all have corresponding labels, which are used for later training and result analysis.

Four evaluation indicators, namely, accuracy, precision, false alarm rate, and miss rate, are used to evaluate the results of stuck pipe prediction.

Accuracy represents the proportion of correctly predicted samples to the total number of samples, reflecting the overall prediction correctness of the model:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2}$$

where, TP: actual stuck pipe, predicted as stuck pipe, TN: actual normal, predicted as normal, FP: actual normal, predicted as stuck pipe (false alarm), and FN: actual stuck pipe, predicted as normal (missed alarm).

Precision represents the proportion of samples predicted as stuck pipe by the model that are actually stuck pipe, focusing on the accuracy of predicting stuck pipe to avoid false alarms:

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{3}$$

False alarm rate (FAR) represents the proportion of samples that are actually normal but incorrectly predicted as stuck pipe by the model, measuring the probability of false alarms:

$$FAR = \frac{FP}{FP + TN}. \tag{4}$$

Miss rate (MR) represents the proportion of samples that are actually stuck pipe but incorrectly predicted as normal by the model, measuring the probability of missed alarms:

$$\mathrm{MR} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}}. \tag{5}$$

These indicators comprehensively evaluate the model's stuck pipe prediction results from different perspectives.

2.3. Mechanism Constraints

The mechanism constraint realizes refined screening and feature enhancement of original training samples by embedding physical laws and empirical rules in the field of drilling engineering, significantly improving label quality and model learning efficiency (Figure 1).

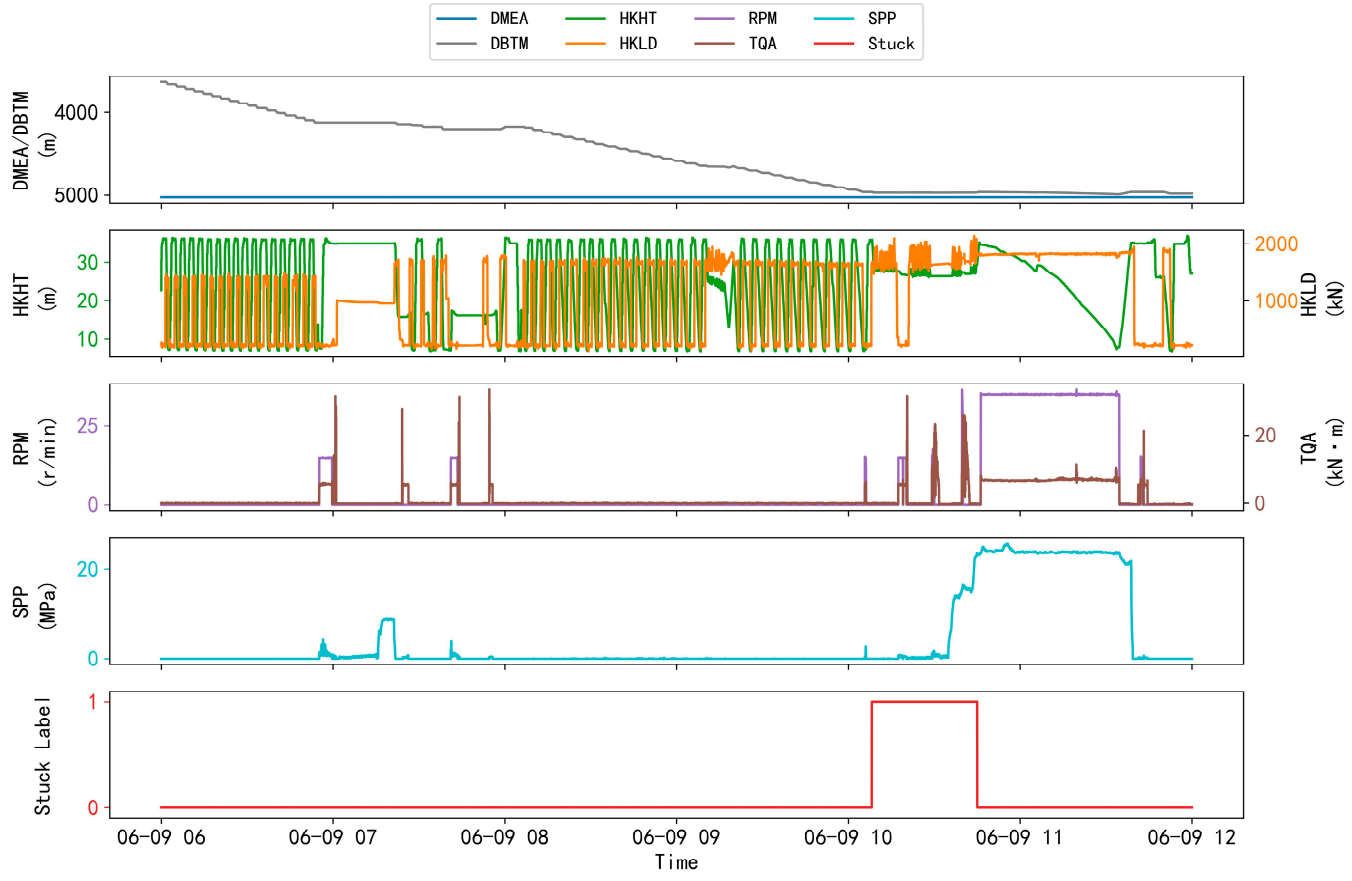


Figure 1. Original training sample and stuck pipe label.

Specifically, this mechanism optimizes the data distribution through three paths:

(1) Accurately identify and filter pseudo-stuck pipe samples. Based on the physical characteristics of drilling conditions, scenarios where parameters are abnormal but not stuck pipe are defined. For example, during drilling, due to operations such as setting and making a connection, the hook height will briefly exceed the range of 10–30 m and the hook load will rise or fall rapidly in a short time, but the load change conforms to the rules. Such samples are marked as invalid stuck pipe signals and excluded; on the contrary, for the characteristics of real stuck pipe events, such as static bit position, rotational speed approaching zero, but severe fluctuation of hook load, the weight of these samples in the training set is highlighted through

mechanism constraints, making it easier for the model to capture real stuck pipe samples (Figure 2).

(2) Systematically filter data noise in non-effective drilling stages. For non-drilling stages such as tripping in, tripping out, and setting, the invalid intervals are defined through multi-parameter joint criteria, and only samples in the normal drilling stage are retained for model training. This operation significantly reduces the interference of non-target working conditions on the model and improves the signal-to-noise ratio of effective signals in the training dataset.

(3) Establishing working condition adaptive differentiated judgment rules. Based on the physical laws of different drilling stages, scenario-specific stuck pipe identification logic is designed: in the normal drilling stage, the torque fluctuation amplitude is taken as the core criterion; in the tripping stage, focus is placed on monitoring abnormal trends in hook load to determine tripping obstacles caused by stuck pipe, thereby improving the judgment stability of the model in diverse scenarios.

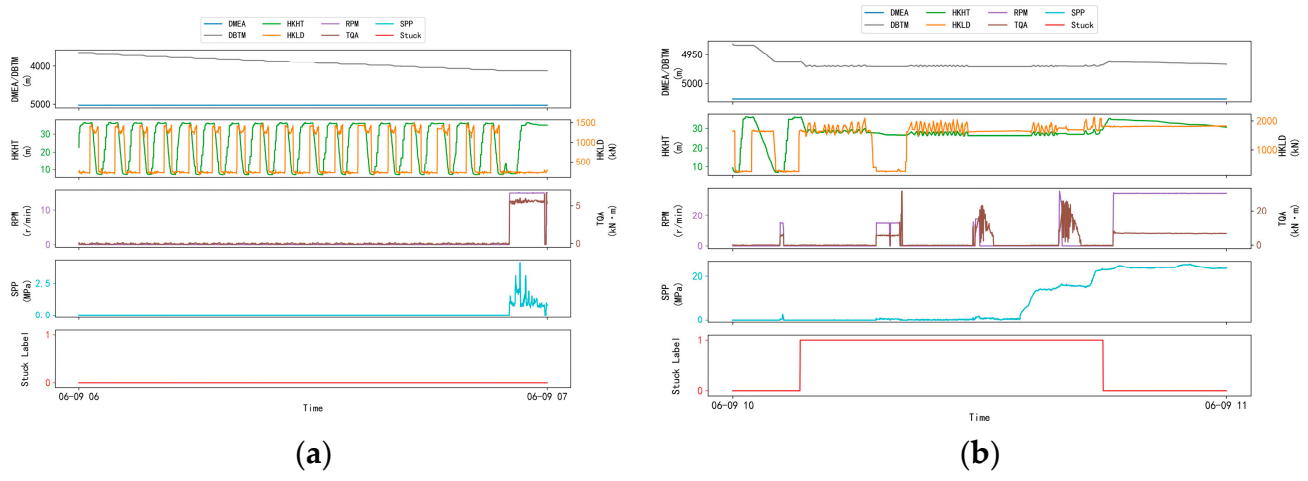


Figure 2. (**a**) Normal drilling sample and (**b**) stuck pipe sample.

The primary mechanism-based constraints are presented in Table 1.

Table 1. The primary mechanism-based constraints.

Category of Constraint Condition	Specific Judgment Logic	Involved Drilling Parameters	Key Thresholds/Rules
Data Validity Filtering	Filter invalid data and only perform pipe sticking judgment on valid data	HKLD, HKHT	HKLD > 600; $5 \leq$ HKHT ≤ 26; negative or abnormal parameters are set to 0
Stuck Pipe Judgment	Determine suspected sticking if DBTM is stable, RPM table stops, and HKLD fluctuates sharply	DBTM, RPM, HKLD	DBTM fluctuation < 0.3; maximum RPM < 1; HKLD fluctuation (max–min) > 100
Torque Anomaly Warning	Judge whether torque rises abnormally based on the sliding window of recent data	RPM, TQA	Mean RPM of the last 25 points < 2; mean TQA of the last 25 points > 10

The introduction of mechanism constraints essentially converts domain knowledge into data prior, and its core value is reflected in two aspects: first, reducing the model's dependence on large-scale labeled data, and making up for the lack of data information

in the small sample scenario of stuck pipe through physical rules; second, blocking the learning path of false correlations, preventing the model from misjudging non-stuck pipe patterns, such as normal height changes during drilling and regular load fluctuations during setting, as abnormalities, thereby fundamentally improving the physical interpretability and engineering reliability of predictions.

2.4. Intelligent Algorithm Design

In the field of intelligent prediction of stuck pipe under complex working conditions, with a single model it is often difficult to fully capture the multi-dimensional features of stuck pipe data. This study constructs an ensemble learning model system consisting of AE, BiLSTM, and Transformer, whose theoretical basis lies in the complementary representation ability of different model architectures for data features.

2.4.1. Autoencoder (AE)

The Autoencoder adopts a symmetric multi-layer fully connected network structure, which is essentially unsupervised learning based on the minimization of reconstruction error. In the drilling engineering scenario, the reconstruction error of normal working condition data is small, while that of abnormal working conditions (such as stuck pipe precursors) increases significantly, so anomalies can be detected through reconstruction error. This design can effectively identify potential stuck pipe precursors, such as sudden changes in hook load, abnormal rotational speed, and torque. The mathematical expression of the loss function minimization is:

$$\min_{\theta} \mathbb{E}_{x \sim p_{data}(x)} \left[\| \ x \ - D(E\ (x)) \ \|^2 \right] + \lambda \Omega(\theta), \tag{6}$$

where, θ: weights and biases of encoder E and decoder D; $\mathbb{E}_{x \sim p_{data}(x)}$: expectation of the input data distribution $p_{data}\ (x)$, average reconstruction error; $\| \ x - D(E(x)) \ \|^2$: squared error between the original input x and the reconstructed output $D(E(x))$ (reconstruction error); $\lambda\Omega(\theta)$: regularization term, where λ is the regularization coefficient and $\Omega(\theta)$ is the parameter complexity measure to prevent overfitting.

The core principle of the Autoencoder is shown in Figure 3.

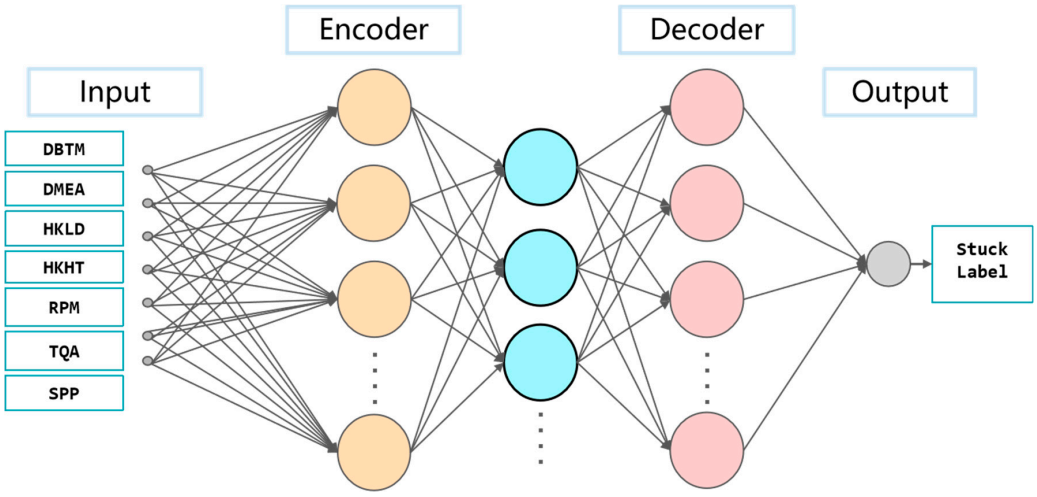


Figure 3. Principle of Autoencoder.

2.4.2. Bidirectional Long Short-Term Memory Network (BiLSTM)

LSTM is a time-series model designed to solve the long-term dependency forgetting problem of traditional recurrent neural networks (RNNs). This study adopts a bidirectional LSTM (BiLSTM) architecture to utilize both past and future time-series information and

further introduces variational dropout technology [17]. By applying random perturbations to the hidden states during training, an approximation of Bayesian inference is achieved, thereby enhancing the uncertainty quantification ability of the stuck pipe prediction model. The core principle of BiLSTM is shown in Figure 4.

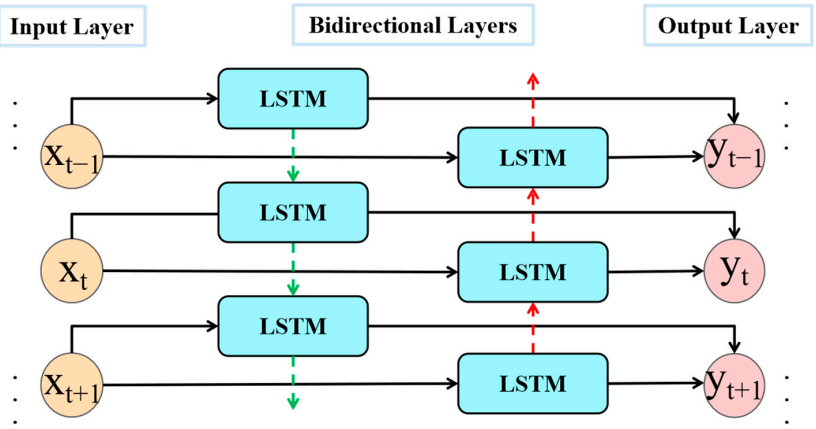


Figure 4. Principle of bidirectional LSTM.

2.4.3. Transformer Model

The multi-head attention mechanism of the Transformer architecture breaks through the dependence of time-series models on adjacent data points and can capture nonlinear correlations across arbitrary time steps. The contribution of different positions is dynamically allocated through attention weights, for example, the cross-time-step correlation between the sudden increase in torque and sudden drop in rotational speed before stuck pipe can be significantly captured. Mathematically, the self-attention mechanism can be expressed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V, \tag{7}$$

where, Q, K, V: query, key, and value matrices, respectively; QK^T: similarity matrix between query and key; $\sqrt{d_k}$: scaling factor to avoid gradient disappearance of softmax caused by large d_k; softmax: normalizes the weights to [0, 1] so that the sum of weights is 1.

Q, K, V are divided into h groups, and attention is calculated in parallel before concatenation to enhance the capture of multi-scale correlations:

$$\text{MultiHead}(Q, K, V) = \text{Concat}\left(\text{head}_1, \ldots, \text{head}_h\right) W^O \ , \tag{8}$$

where,

$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right), \tag{9}$$

where, W_i^Q, W_i^K, W_i^V: learnable parameters; $\text{Concat}(\text{head}_1, \ldots, \text{head}_h)$: concatenate the outputs of h heads; W^O: linear transformation matrix after concatenation, mapping the concatenated result to the target dimension.

This study further adopts layer normalization and residual connection technology, which theoretically ensures the stability of deep network training.

2.4.4. Ensemble Learning

Ensemble learning improves generalization ability by fusing the prediction results of multiple base models, reducing the bias and variance of a single model. The meta-model (multi-layer perceptron) in this study takes the outputs of the three base models as input

and focuses on reducing the stuck pipe miss rate and improving the stuck pipe prediction accuracy through a weighted loss function.

Among them, aiming at the imbalance of stuck pipe prediction (normal samples are far more than stuck pipe samples), higher weights are assigned to positive samples (stuck pipe):

$$\mathcal{L}_{weighted} = -\frac{1}{N}\sum_{i=1}^{N} w_i[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)], \quad (10)$$

where, N: number of samples; y_i: true label (1 for stuck pipe, 0 for normal); p_i: probability that the model predicts stuck pipe; w_i: sample weight; in this study, the loss of positive samples is amplified to force the model to pay more attention to stuck pipe identification.

The weight factor w_i is dynamically adjusted according to the sample category, and the stuck pipe samples are assigned 5 times the weight of normal samples. This design is theoretically equivalent to the constraint on the risk of missed alarms, which can effectively reduce the tail risk of the system.

From the perspective of model fusion, the diversity of base models is achieved through the following three dimensions. The Autoencoder is based on reconstruction error, BiLSTM is based on time-series evolution, and Transformer is based on the attention mechanism, forming algorithm-level complementarity. Different models naturally form multi-scale representations of original data during feature extraction, and the redundancy and complementarity of these representations enhance the robustness of the system. There are essential differences in the construction of classification hyperplanes among base models, and ensemble learning can theoretically obtain a better decision boundary by optimizing the combination of these hyperplanes (Figure 5).

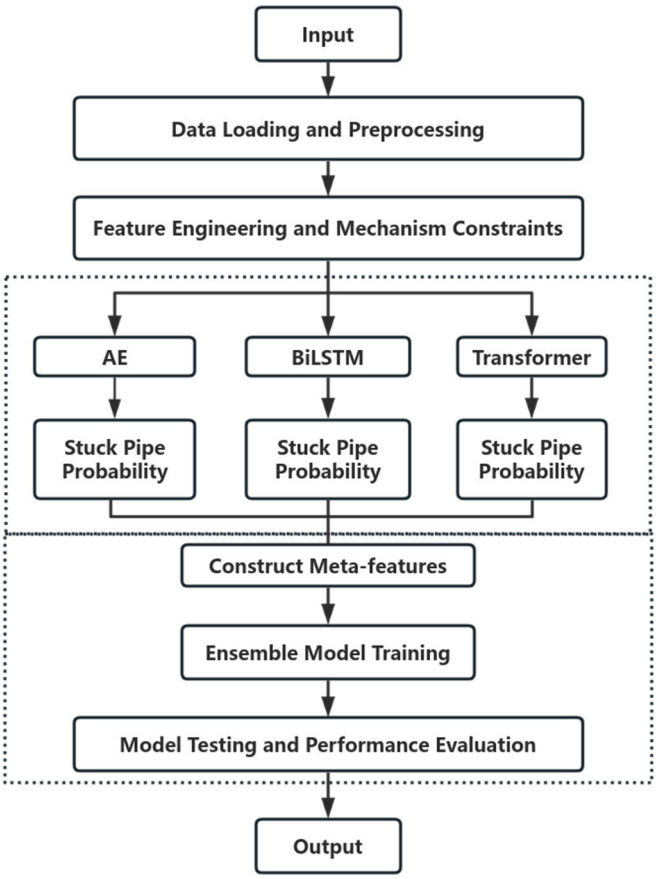


Figure 5. Overall technical flowchart.

3. Test Results and Analysis

The training and test data come from actual drilling logging data of 10 wells, with training parameters including 7 parameters: well depth, bit position, hook height, hook load, rotary speed, torque, standpipe pressure, and stuck pipe labels (1 for stuck pipe, 0 for normal). The data sampling frequency is 1–2 Hz, with a total of about 200,000 samples, and the ratio of normal samples to drilling stuck samples is approximately 10:1. The data are divided into eight wells as the training set and two wells as the test set, where the training set is used for parameter learning of the stuck pipe prediction model, and the test set is used to evaluate the generalization ability of the model. The model hyperparameters are searched for the optimal combination through grid search. The optimal hyperparameter combination is presented in Table 2.

Table 2. The optimal hyperparameter combination of the model.

Model Type	Hyperparameter Category	Parameter Value	Model Type	Hyperparameter Category	Parameter Value
BiLSTM	Number of Hidden Layers	2 layers	Transformer	Number of Encoder Layers	2 layers
	Learning Rate	0.001		Number of Attention Heads	4 heads
	Number of Neurons	64 (1st layer), 32 (2nd layer)		Feature Dimension	64
	Activation Function	Internal Gating: tanh + sigmoid; Classification Head: ReLU (hidden layer), Sigmoid (output layer)		Learning Rate	0.001
AE	Encoding Layer Dimension	Hidden Layers: 64 → 32; Bottleneck Layer: 16	Ensemble model	Number of Hidden Layers	2 layers
	Learning Rate	0.001		Number of Neurons	16 (1st layer), 8 (2nd layer)
	Training Epochs	50		Activation Function	ReLU (hidden layer), Sigmoid (output layer)
	Batch Size	32		Learning Rate	0.001

3.1. Test Results

After training and testing according to the above process, the accuracy, precision, false alarm rate, and miss rate of each model's stuck pipe prediction results are shown in Table 3.

Table 3. Test results of different models.

Model	Accuracy (Without Mechanism Constraints)	Accuracy (With Mechanism Constraints)	Precision	False Alarm Rate	Miss Rate
AE	0.751	0.869	0.945	0.099	0.146
BiLSTM	0.787	0.879	0.974	0.045	0.160
Transformer	0.747	0.862	0.905	0.188	0.113
Ensemble model	0.826	0.901	0.959	0.077	0.110

It can be concluded from Table 3 that the proposed ensemble model has the highest accuracy, and its comprehensive performance in various indicators is better than that of single models. The accuracy of the ensemble model reaches 90.1% and the precision reaches 95.9%. The comparison of false alarm rates and miss rates of different models is shown in Figure 6. Among them, the BiLSTM model has the lowest false alarm rate but a higher miss rate, indicating that it has a strong ability to distinguish normal working conditions but is insufficiently sensitive to stuck pipe signals, with significant missed alarm risks. The Transformer model has a lower miss rate than BiLSTM and AE, but the highest

false alarm rate, indicating that it has a strong ability to capture stuck pipe features but tends to misjudge normal fluctuations as stuck pipe, leading to invalid shutdowns. The AE model has an intermediate false alarm rate and miss rate, but its feature expression ability is limited, making it difficult to balance the two types of errors. The ensemble learning model achieves the optimal balance between the false alarm rate and miss rate, verifying the effectiveness of the multi-model advantage complementary strategy, which not only reduces the construction cost caused by false alarms but also reduces the safety hazards caused by missed alarms. The above results show that the ensemble model can effectively integrate the advantages of different models and improve the stuck pipe prediction performance.

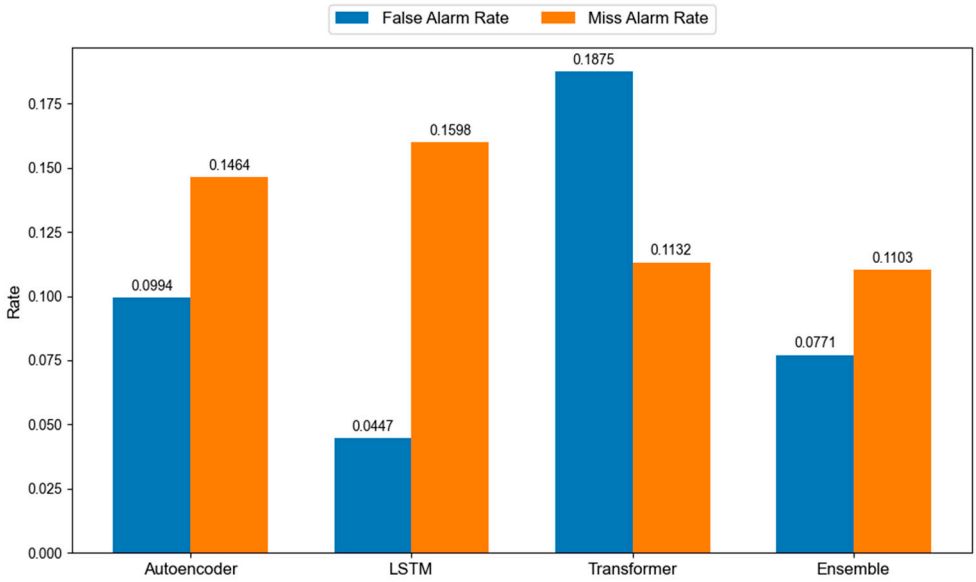


Figure 6. Comparison of false alarm rates and miss rates of different models.

Meanwhile, the McNemar test was performed between the BiLSTM (the base model with the best prediction performance) and the ensemble model (null hypothesis (H_0): there is no significant difference in prediction performance between the ensemble model and BiLSTM; alternative hypothesis (H_1): there is a significant difference in prediction performance between the two models). Based on the calculation of indicators, such as accuracy, missing alarm rate, and false alarm rate, there is a statistically significant difference in drilling stuck prediction performance between the ensemble model and BiLSTM ($\chi^2 \approx 278.15$, $p < 0.001$). The alternative hypothesis holds, and the overall prediction performance of the ensemble model is significantly superior to that of BiLSTM. Overall, the prediction effect of the ensemble model is significantly superior to that of single models.

3.2. Analysis of Prediction Result Curves

The curves of stuck pipe prediction results of each model are shown in Figure 7.

It can be seen from Figure 7 that the Autoencoder model has obvious lag in prediction during the stuck pipe period and many false alarms in special working conditions, such as tripping, reflecting the insufficient adaptability of the unsupervised model to complex working conditions. The BiLSTM model responds quickly to the start time of stuck pipe but has poor prediction stability in non-stuck pipe periods, and the problem of high miss rate is more prominent in long time series. The Transformer model captures the early signal of torque increase but has weak anti-interference ability to hook height fluctuations and many false alarms, reflecting the insufficient robustness of its attention mechanism to noise. The prediction curve of the ensemble learning model has the highest coincidence with the

real label, which not only avoids the missed alarm risk of BiLSTM but also reduces the false alarm problem of Transformer, and the prediction robustness in special working conditions, such as tripping, is significantly improved, verifying the advantage of multi-model fusion.

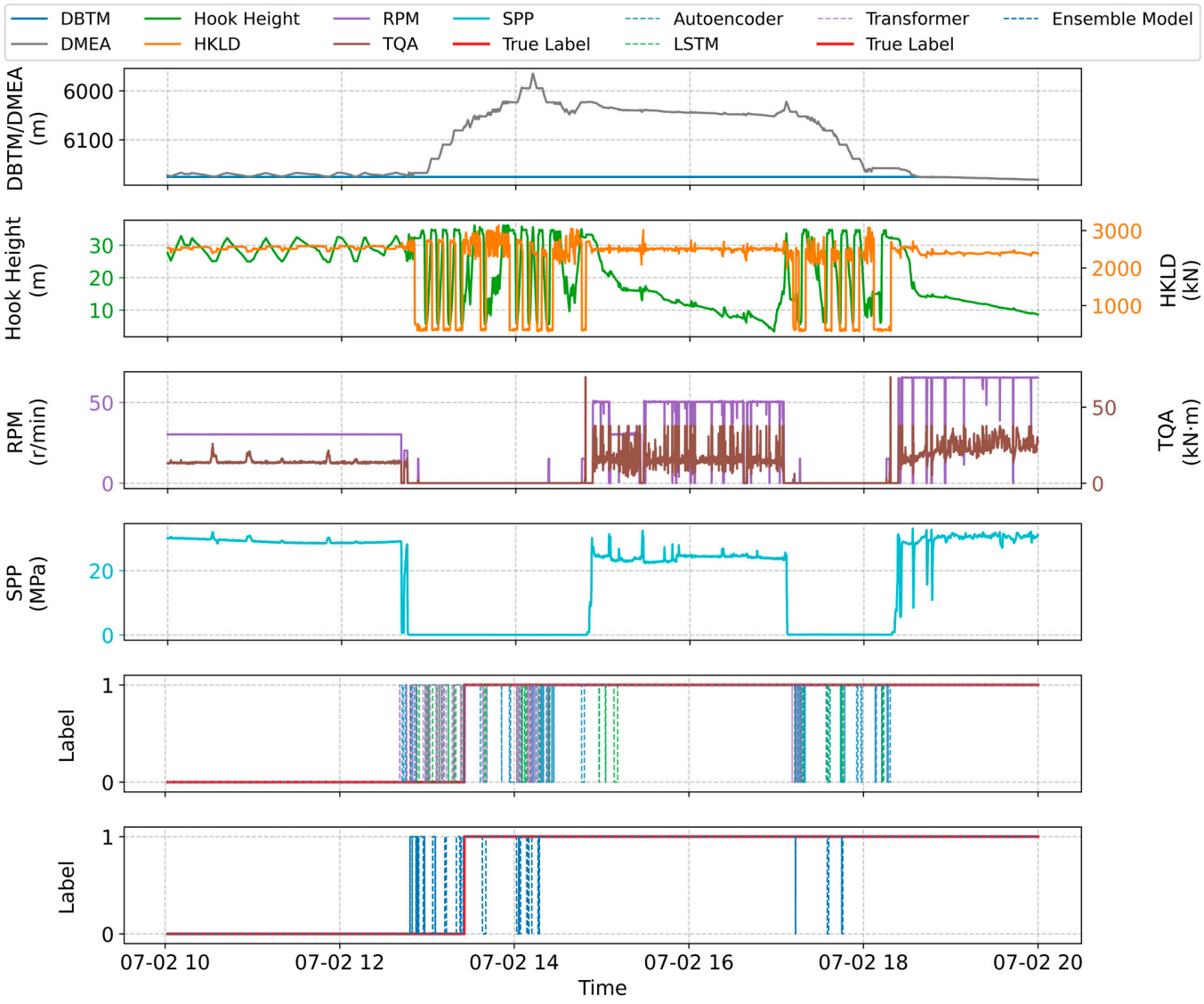


Figure 7. Comparison of stuck pipe prediction result curves of different models.

3.3. Summary

(1) The ensemble model proposed in this paper can capture different feature patterns in drilling data by integrating the advantages of Autoencoder, bidirectional BiLSTM, and Transformer. The Autoencoder provides compressed features of data, the bidirectional BiLSTM is good at capturing local time-series dependencies, and the Transformer can better capture long-distance dependencies, and their combination makes the ensemble model perform better in the stuck pipe prediction task.

(2) The drilling mechanism knowledge constraint can effectively assist the model in identifying stuck pipe signs. After adding mechanism features, the accuracy of the model is increased by 10% on average, verifying the importance of domain knowledge in model construction.

4. Conclusions and Prospects

(1) This study proposed a stuck pipe prediction method based on a deep learning ensemble model, which realizes accurate prediction of stuck pipe risks through four stages: data preprocessing, feature engineering, mechanism constraints, and ensemble prediction. Experimental results showed that the accuracy of this method on the test set reached 90.1% and the precision reached 95.9%, with the optimal balance between false alarm rate and miss rate. By integrating mechanism constraints with the ensemble model, the prediction stability was significantly better than that of single models, which can provide effective risk early warning for drilling operations.

(2) This study still has several limitations. For the mechanism constraint part of the model, manual threshold adjustment is required across different blocks or drilling platforms. The proposed ensemble learning model has certain time and computational cost limitations, making it currently difficult to integrate into drilling rig systems. The ensemble learning strategies among the various models lack a dynamic update mechanism, and the ensemble model has not yet established an automatic update process based on real-time data, which prevents the realization of real-time updates during the drilling process.

(3) Future research can be carried out in the following aspects: investigating the automatic dynamic update of mechanism constraint thresholds, studying more effective model ensemble strategies, such as dynamic weighted fusion, to further improve the adaptive ability of the model, and combining real-time drilling data to develop an online learning mechanism, so that the model can adapt to different drilling environments and continuously optimize prediction performance.

Author Contributions: Conceptualization, B.X.; methodology, Z.Z.; software, Y.W. and Q.L.; validation, B.X., X.S. and M.L.; formal analysis, Y.W.; investigation, Y.Y.; resources, B.X.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W. and X.S.; visualization, M.L.; supervision, B.X.; project administration, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Zhaopeng Zhu grant number Sponsored by CNPC Innovation Found (2022DQ02-0308).

Data Availability Statement: The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author. The data originated from actual logging data collected on-site at drilling locations.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

Full Name	Abbreviation	Full Name	Abbreviation
Bit position	DBTM	False alarm rate	FAR
Hook height	HKHT	Miss rate	MR
Hook load	HKLD	Autoencoder	AE
Weight on bit	WOB	Bidirectional long short-term memory network	BiLSTM
Rotary speed	RPM		
Torque	TQA		
Standpipe pressure	SPP		

References

1. Chen, T.; Zhang, Y.Z.; Xu, C. Research and Application of Intelligent Identification Algorithm for Drilling Conditions in Coal Mine Underground. *Saf. Coal Mines* **2025**, *56*, 242–249.
2. Li, G.; Song, X.; Tian, S.; Zhu, Z. Intelligent Drilling and Completion: A Review. *Engineering* **2022**, *18*, 33–48. [CrossRef]
3. Nautiyal, A.; Mishra, A.K. Machine learning approach for intelligent prediction of petroleum upstream stuck pipe challenge in oil and gas industry. *Env. Dev. Sustain.* **2022**, *27*, 24167–24193. [CrossRef]
4. Salvatore, D.; Marta, P.; Matteo, M.; Piroddi, L.; Spelta, A.; Zausa, F. Stuck pipe prediction from rare events in oil drilling operations. *Upstream Oil Gas Technol.* **2023**, *11*, 100096. [CrossRef]
5. Wu, H.L.; Wang, Z.B.; Si, L.; Zou, X.Y.; Gu, J.H.; Wei, D.; Tan, C. A sticking predictor construction and evaluation method for drill tools sticking prediction. *Rev. Sci. Instrum.* **2025**, *96*, 015116. [CrossRef] [PubMed]
6. Zhang, T.; Xue, Y.; Meng, Z.; Sader, M.; Zhang, W.; Li, J. Drill Sticking Prediction Based on Modal Decomposition and Physical Constraint Model of Near-Bit Data. *Processes* **2025**, *13*, 1802. [CrossRef]
7. Siruvuri, C.; Nagarakanti, S.; Samuel, R. Stuck pipe prediction and avoidance: A convolutional neural network approach. In Proceedings of the IADC/SPE Drilling Conference, SPE, Miami, FL, USA, 21–23 February 2006; p. SPE 98378-MS.
8. Zhu, S.; Song, X.Z.; Li, G.S.; Zhu, Z.P.; Yao, X.Z. Intelligent Real-time Analysis of Drill String Friction & Torque and Stuck Pipe Trend Prediction. *Oil Drill. Prod. Technol.* **2021**, *43*, 428–435.
9. Chamkalani, A.; Pordel Shahri, M.; Poordad, S. Support vector machine model: A new methodology for stuck pipe prediction. In Proceedings of the SPE Unconventional Gas Conference and Exhibition, Muscat, Oman, 28–30 January 2013; p. SPE 164003-MS.
10. Yi, S.Q.; Wei, K. Random Forest Stuck Pipe Risk Assessment Method Based on SMOTE Undersampling. *Pet. Geol. Eng.* **2023**, *37*, 100–103.
11. Ahmed, O.S.; Aman, B.M.; Zahrani, M.A.; Ajikobi, F.I. Stuck pipe early warning system utilizing moving window machine learning approach. In Proceedings of the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates, 11–14 November 2019; p. SPE 197674-MS.
12. Zhang, W.X. Prevention of Coiled Tubing Downhole Complications Based on GAN-LSTM. Master's Thesis, Yangtze University, Jingzhou, China, 2024.
13. Al-Baiyat, I.; Heinze, L. Implementing artificial neural networks and support vector machines in stuck pipe prediction. In Proceedings of the SPE Kuwait International Petroleum Conference and Exhibition, Kuwait City, Kuwait, 10–12 December 2012; p. SPE 163370-MS.
14. Abbas, A.K.; Almubarak, H.; Abbas, H.; Dawood, J. Application of machine learning approach for intelligent prediction of pipe sticking. In Proceedings of the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, United Arab Emirates, 11–14 November 2019; p. SPE 197396-MS.
15. Liu, Z.H.; Song, X.Z.; Zhu, S.; Ye, S.L.; Zhang, C.K.; Ma, B.D.; Zhu, Z.P. Intelligent Stuck Pipe Prediction Method Based on VC-SVM and Particle Swarm Optimization. *China Pet. Mach.* **2024**, *52*, 1–11.
16. Zhang, T.; Xia, P.; Li, J.; Wang, B.; Zhan, J.H. Study on SCNGO-SVM Stuck Pipe Warning Method Based on Downhole Parameters. *China Pet. Mach.* **2025**, *53*, 20–27, 36.
17. Liu, M.C.; Song, X.Z.; Li, D.Y.; Zhu, S.; Fu, L.; Zhu, Z.P.; Zhang, C.K.; Pan, T. Intelligent Prediction Model and Interpretation of Drill String Friction. *Coal Geol. Explor.* **2023**, *51*, 89–99.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.