Development of a Long Short-Term Memory-Based Model for Early Detection of Prostatitis Using Synthetic Time-Series Clinical Data

Bala P. Bulus[1], Damare, Aisha Abdulkarim[2] , Abubakar Mohammed Buba[3]

[1,3]Computer Science Department Federal College of Education Yola

[2]ICT Department College of Nursing Science Yola

DOI: https://doi.org/10.51584/IJRIAS.2026.11060323

Received: 03 July 2026; Accepted: 08 July 2026; Published: 23 July 2026

ABSTRACT

Prostatitis is a common urological condition affecting men and often remains undetected during its early stages because of subtle or overlapping clinical symptoms. This study proposes a deep learning-based framework for the early detection of prostatitis using synthetically generated longitudinal clinical data. A synthetic dataset comprising 120 patient records, including prostate-specific antigen (PSA) levels and other clinically relevant hematological parameters, was preprocessed and used to train a Long Short-Term Memory (LSTM) neural network. The proposed model achieved high predictive performance, with an accuracy of 98.5%, precision of 97.6%, recall of 98.2%, an F1-score of 97.9%, and an area under the receiver operating characteristic curve (AUC) of 0.992, indicating excellent discriminative capability. Performance evaluation using the confusion matrix and ROC curve further demonstrated the model's effectiveness in distinguishing prostatitis from non-prostatitis cases. The findings highlight the potential of LSTM-based models to support early clinical decision-making, particularly where access to large real-world datasets is limited. Furthermore, the study presents a scalable and privacy-preserving diagnostic framework that can be adapted to similar medical prediction tasks. Future research should validate the proposed model using real-world clinical datasets, including electronic health records (EHRs), to assess its generalizability and clinical applicability.

Keywords: Prostatitis; Artificial Intelligence; Deep Learning; Long Short-Term Memory; Time-Series Clinical Data; Prostate-Specific Antigen; Early Disease Detection; Clinical Decision Support.

INTRODUCTION

Prostatitis is among the most common urological disorders affecting men, particularly those younger than 50 years of age. It is characterized by inflammation of the prostate gland, which may result from bacterial infection or non-bacterial inflammatory processes. Unlike prostate cancer and benign prostatic hyperplasia (BPH), which predominantly affect older men, prostatitis occurs across a wider age range and commonly manifests as acute or chronic pelvic pain, lower urinary tract symptoms, and, in some cases, systemic inflammatory features [18],[1].

Clinically, prostatitis is classified into four major categories: acute bacterial prostatitis, chronic bacterial prostatitis, chronic prostatitis/chronic pelvic pain syndrome (CP/CPPS), and asymptomatic inflammatory prostatitis. (National Institute of Diabetes and Digestive and Kidney Diseases [17],[23]. Among these categories, CP/CPPS is the most prevalent, accounting for the majority of prostatitis cases. However, it also presents the greatest diagnostic challenge because its symptoms are often non-specific and overlap considerably with those of other urological conditions [11].

In Nigeria and many other developing countries, prostatitis remains underdiagnosed and frequently misdiagnosed due to limited access to specialized diagnostic facilities, inadequate healthcare infrastructure, and low public awareness of the disease. Delayed or inaccurate diagnosis may result in persistent urinary tract symptoms, chronic pelvic pain, sexual dysfunction, psychological distress, and a significant reduction in patients'

quality of life. Furthermore, serum prostate-specific antigen (PSA), although primarily employed as a biomarker for prostate cancer screening, may also become elevated during prostatic inflammation. Consequently, serial PSA measurements can provide valuable temporal information for monitoring disease progression and supporting differential diagnosis [11],[5].

Despite advances in urological practice, the early and accurate diagnosis of prostatitis remains a significant clinical challenge. The condition often presents with symptoms that overlap with those of benign prostatic hyperplasia, urinary tract infections, and prostate cancer, making differential diagnosis difficult. Conventional diagnostic approaches including physical examination, laboratory investigations, urine culture, and PSA testing are often insufficient for accurately distinguishing among the different forms of prostatitis. Consequently, patients frequently undergo repeated clinical evaluations before a definitive diagnosis is established, delaying appropriate treatment and increasing healthcare costs.

Recent developments in artificial intelligence (AI) and deep learning have created new opportunities for improving disease diagnosis through the analysis of complex clinical data. Among these techniques, Long Short-Term Memory (LSTM) networks, a specialized type of recurrent neural network (RNN), have demonstrated exceptional capability in modelling sequential and time-dependent data. Unlike conventional machine learning algorithms that treat observations independently, LSTM models can retain long-term temporal dependencies, enabling them to identify subtle changes in patients' clinical histories over time. This capability makes LSTM particularly suitable for analysing longitudinal clinical information, including repeated PSA measurements, symptom progression, urination difficulty scores, and other temporal health indicators.

In contrast, convolutional neural networks (CNNs) are primarily designed for image-based applications such as medical image classification and segmentation. Since prostatitis diagnosis relies largely on structured clinical records and sequential laboratory measurements rather than imaging data, LSTM networks provide a more appropriate framework for modelling disease progression and supporting early diagnosis. The integration of AI-driven predictive models into clinical practice therefore has the potential to improve diagnostic accuracy, reduce misdiagnosis, facilitate earlier clinical intervention, and enhance decision-making by healthcare professionals.

Against this background, this study proposes the development of a Long Short-Term Memory (LSTM)-based predictive model for the early detection of prostatitis using time-series clinical data, including serial PSA levels, urination difficulty scores, and patient symptom history. By leveraging temporal patterns within these clinical variables, the proposed model seeks to support timely diagnosis and improve patient outcomes through more accurate and efficient clinical decision-making.

An Overview of Long-short Term Memory

Long Short-Term Memory (LSTM) is a variant of recurrent neural networks (RNNs) developed to overcome the vanishing gradient problem encountered during the training of conventional RNNs[12], [9]Ghojogh & Ghodsi, 2023). Unlike feedforward neural networks, LSTM networks maintain an internal memory state that enables them to capture long-term temporal dependencies in sequential data [10],[9](Goodfellow et al., 2016; Ghojogh & Ghodsi, 2023). This capability has made LSTM one of the most widely adopted deep learning architectures for time-series forecasting, natural language processing, speech recognition, and healthcare analytics [9],[25].

Long Short-Term Memory (LSTM) is a specialized architecture of recurrent neural networks (RNNs) developed to address the limitations of conventional RNNs in learning long-term dependencies from sequential data (Hochreiter & Schmidhuber, 1997; Ghojogh & Ghodsi, 2023).. Unlike feedforward neural networks, which process each observation independently, LSTM networks maintain an internal memory state that enables them to retain and utilize information from previous time steps. This capability allows LSTM models to capture temporal relationships within sequential datasets, making them particularly suitable for applications involving time-series forecasting, natural language processing, speech recognition, financial analysis, and healthcare prediction. Since its introduction by Sepp Hochreiter and Jürgen Schmidhuber in 1997, LSTM has remained one

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

of the most influential deep learning architectures for modelling sequential information due to its superior ability to preserve long-term contextual information during training.

Traditional recurrent neural networks suffer from the vanishing and exploding gradient problems during backpropagation through time, particularly when processing long sequences. These problems reduce the network's ability to learn relationships between observations that are separated by long time intervals, thereby limiting prediction accuracy. LSTM overcomes these challenges through a unique memory cell and gating mechanism that regulate the storage, updating, and retrieval of information. The architecture employs three principal gates namely the input gate, forget gate, and output gate which selectively determine what information should be stored, discarded, or transmitted to subsequent hidden states. This gating mechanism enables the network to preserve relevant historical information while filtering irrelevant data, thereby significantly improving learning efficiency and predictive performance.

The capability of LSTM to model temporal dependencies has resulted in its widespread adoption across numerous scientific and engineering disciplines. In healthcare, patient information is inherently longitudinal, with clinical variables such as laboratory results, physiological measurements, medication history, and symptom progression collected repeatedly over time. Unlike conventional machine learning algorithms that treat observations as independent records, LSTM exploits the chronological order of clinical events to identify temporal patterns associated with disease onset and progression. Consequently, LSTM has demonstrated remarkable effectiveness in disease diagnosis, prognosis, patient risk stratification, and clinical decision support using longitudinal electronic health records and other biomedical time-series datasets.

Recent advances in healthcare predictive analytics have further strengthened the role of LSTM in medical decision-making. Several systematic reviews have reported that LSTM consistently outperforms many traditional machine learning algorithms in predicting clinical outcomes because of its ability to capture complex temporal relationships in sequential patient data. These advantages have led to successful applications of LSTM in predicting cardiovascular diseases, diabetes, neurological disorders, intensive care unit outcomes, and other chronic medical conditions using electronic health records and continuous physiological monitoring data. Furthermore, researchers continue to enhance the original LSTM architecture through variants such as Bidirectional LSTM (Bi-LSTM), Time-Aware LSTM (T-LSTM), and Attention-based LSTM to improve predictive performance on increasingly complex healthcare datasets.

For prostate diseases, including prostatitis, clinical evaluation often depends on sequential measurements such as serial prostate-specific antigen (PSA) levels, symptom progression, urination difficulty scores, inflammatory markers, and other longitudinal clinical observations. These variables exhibit temporal dependencies that are difficult to model using conventional classification algorithms. Because LSTM is specifically designed to analyse sequential information, it provides an appropriate framework for learning disease progression patterns from repeated clinical measurements. By recognising subtle temporal changes that may precede overt clinical manifestations, LSTM-based models have the potential to facilitate earlier diagnosis, reduce misdiagnosis, and support evidence-based clinical decision-making. Consequently, the adoption of LSTM in this study provides a robust methodological foundation for developing an intelligent predictive model for the early detection of prostatitis using time-series clinical data.

Overview of Prostatitis

Prostatitis is a common and clinically heterogeneous urological disorder characterized by inflammation of the prostate gland. It affects men across a broad age range, particularly those between 30 and 50 years, unlike benign prostatic hyperplasia (BPH) and prostate cancer, which predominantly occur in older men. The condition is recognized as the third most common urological disorder after BPH and prostate cancer and accounts for a substantial proportion of urological outpatient visits worldwide (Jang et al., 2017; Nickel et al., 2019). Epidemiological studies estimate its prevalence to range between 2% and 10% globally, although regional variations have been reported depending on the study population and diagnostic criteria [6][27](Clemens et al., 2013; Zhang et al., 2019).

Clinically, prostatitis encompasses four categories: acute bacterial prostatitis, chronic bacterial prostatitis, chronic prostatitis/chronic pelvic pain syndrome (CP/CPPS), and asymptomatic inflammatory prostatitis. Among these, CP/CPPS is the most prevalent and is frequently associated with chronic pelvic pain, urinary symptoms, sexual dysfunction, psychological distress, and reduced quality of life [6] [3]Clement et al., 2013; Alshahrani et al., 2025). The condition remains difficult to diagnose because its clinical manifestations often overlap with those of urinary tract infections, BPH, and prostate cancer, resulting in delayed diagnosis and inappropriate treatment.

Conventional diagnostic methods, including digital rectal examination, urine culture, expressed prostatic secretion analysis, and prostate-specific antigen (PSA) testing, remain fundamental in clinical practice but have limited specificity for distinguishing prostatitis from other prostate disorders [4][19](Bartoletti et al., 2007; Nickel et al., 2019). Consequently, recent advances in artificial intelligence and deep learning have attracted considerable interest as complementary diagnostic tools. In particular, Long Short-Term Memory (LSTM) networks provide an effective approach for analysing longitudinal clinical data, enabling the identification of temporal patterns in PSA levels and symptom progression that may facilitate earlier and more accurate detection of prostatitis.

Artificial Intelligence in Prostate Health

Artificial intelligence (AI) refers to computational systems capable of performing tasks that typically require human intelligence, including learning, pattern recognition, prediction, and decision-making[9] (Goldenberg, Nir, & Salcudean, 2019). In healthcare, AI has become increasingly important for improving diagnostic accuracy, risk prediction, and clinical decision support through the analysis of large and complex datasets. The application of machine learning (ML) to prostate health has expanded rapidly, particularly in the diagnosis, prognosis, and management of prostate-related disorders.

Artificial Neural Networks (ANNs), a major class of ML models, are designed to mimic the information-processing behavior of biological neurons. ANNs consist of interconnected computational units arranged in layers, where each neuron generates an output based on weighted inputs and an activation function [10][15](Goodfellow, Bengio, & Courville, 2016; Liakos et al., 2024). Through iterative training, the network adjusts its weights and biases to minimize prediction errors, enabling it to learn complex nonlinear relationships from clinical data.

Among ANN architectures, Recurrent Neural Networks (RNNs) and their advanced variant, Long Short-Term Memory (LSTM) networks, are particularly relevant for medical time-series analysis because they incorporate memory mechanisms that retain information from previous observations. This capability allows them to model temporal dependencies in sequential clinical data, such as changes in prostate-specific antigen (PSA) levels, symptom progression, and inflammatory markers over time. Consequently, LSTM networks provide a suitable framework for developing AI-based diagnostic systems aimed at improving the early detection and classification of prostatitis.

Related Works

Recent advances in deep learning have significantly improved disease prediction by enabling the analysis of complex temporal and longitudinal clinical data. Among deep learning models, Long Short-Term Memory (LSTM) networks have demonstrated superior performance in modelling sequential health records due to their ability to capture long-term dependencies and temporal patterns. Studies by [17] [21][7]Massaro et al. (2019), Rahman et al. (2020), and Chowdary and Udaya (2021) reported that LSTM-based models achieved high predictive accuracy in chronic disease prediction, particularly for diabetes, by effectively learning patterns from sequential clinical observations. Likewise, [22]Ramirez-Alcocer et al. (2023) demonstrated the effectiveness of LSTM in predicting clinical events within hospital workflows, highlighting its suitability for healthcare applications involving time-series data.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

In prostate health, machine learning techniques have primarily focused on prostate cancer diagnosis and prognosis. For example, [1]Akinnuwesi et al. (2022) developed a support vector machine (SVM)-based model for prostate cancer prediction, achieving high classification accuracy using prostate-specific antigen (PSA) levels and dimensionality reduction techniques. Similarly,[2] Alzboon and Al-Batah (2023) compared several machine learning algorithms and reported that Random Forest achieved superior performance for prostate cancer classification. Although these studies demonstrate the effectiveness of machine learning in analysing prostate-related clinical data, they largely focus on prostate cancer rather than prostatitis.

Despite the clinical similarities between prostatitis, benign prostatic hyperplasia (BPH), and prostate cancer, particularly regarding urinary symptoms and fluctuating PSA levels, limited research has explored the application of LSTM networks for prostatitis diagnosis. Existing studies generally rely on static clinical observations or traditional machine learning techniques, with little emphasis on modelling the temporal progression of symptoms and laboratory findings. This limitation highlights an important research gap, as prostatitis is characterized by dynamic clinical patterns that are more appropriately represented using sequential models.

To address this gap, the present study proposes an LSTM-based framework for the early detection of prostatitis using longitudinal clinical data, including serial PSA measurements, symptom progression, and other time-dependent health indicators. By exploiting the temporal learning capability of LSTM, the proposed model seeks to improve diagnostic accuracy and provide clinicians with a robust decision-support tool for the early identification of prostatitis.

Methodological and Clinical Gaps

Although artificial intelligence and machine learning have demonstrated remarkable success in prostate disease diagnosis, existing studies have primarily focused on prostate cancer detection and prognosis, with comparatively little attention given to prostatitis. Most published models employ conventional machine learning algorithms or deep learning techniques using static clinical features and imaging data, while the temporal progression of patient symptoms and laboratory biomarkers has received limited consideration[1][2] (Akinnuwesi et al., 2022; Alzboon & Al-Batah, 2023).

From a methodological perspective, relatively few studies have exploited the capability of Long Short-Term Memory (LSTM) networks to model longitudinal clinical data for prostatitis diagnosis. Since prostatitis is characterized by fluctuating symptoms and repeated prostate-specific antigen (PSA) measurements over time, models that ignore temporal dependencies may fail to capture important disease progression patterns. Consequently, there remains a need for predictive models that can effectively learn from sequential clinical observations to support earlier and more accurate diagnosis [22](Ramirez-Alcocer et al., 2023).

From a clinical perspective, prostatitis shares several clinical manifestations with benign prostatic hyperplasia (BPH) and prostate cancer, including urinary frequency, nocturia, dysuria, pelvic pain, and, in some patients, elevated PSA levels. This overlap contributes to diagnostic uncertainty and may delay appropriate treatment. Despite these challenges, the application of LSTM-based models specifically for prostatitis remains largely underexplored in the literature.

This study addresses these methodological and clinical gaps by proposing an LSTM-based framework for the early detection of prostatitis using longitudinal clinical data comprising serial PSA measurements, symptom progression, and other time-dependent clinical variables. By leveraging temporal learning, the proposed model aims to improve diagnostic accuracy and contribute to the growing application of artificial intelligence in non-cancerous prostate diseases.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

METHODOLOGY

This study adopted a deep learning approach to develop an intelligent framework for the early detection of prostatitis using Long Short-Term Memory (LSTM) networks. The methodology comprised data acquisition, synthetic data generation, data preprocessing, model development, training, and performance evaluation. The proposed framework was designed to exploit the temporal characteristics of clinical variables, enabling the model to learn disease progression patterns from sequential patient records.

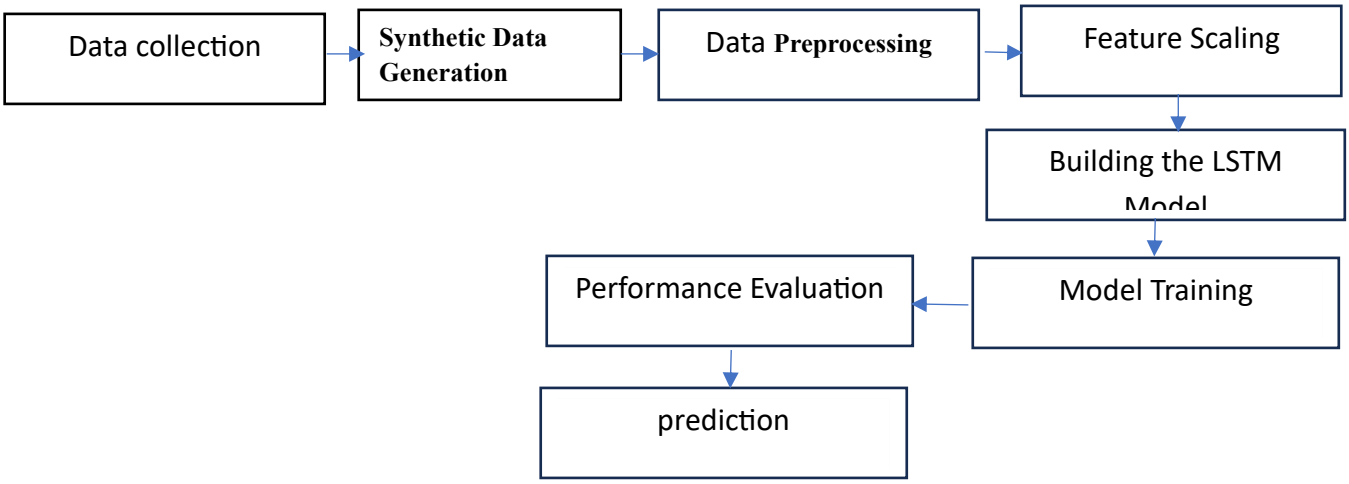

Figure 1: Overall Research Workflow

Figure 1 illustrates the overall research workflow comprising clinical data acquisition, synthetic data generation, data preprocessing, feature scaling, LSTM model development, model training, performance evaluation, and prediction.

Analysis of the Existing System

Existing approaches to prostatitis diagnosis rely primarily on clinical assessment, laboratory investigations, urine culture, digital rectal examination (DRE), and prostate-specific antigen (PSA) testing. Although these methods remain valuable in clinical practice, they often depend on subjective interpretation and may be unable to adequately distinguish prostatitis from other prostate disorders such as benign prostatic hyperplasia (BPH) and prostate cancer because of overlapping clinical manifestations. Furthermore, most existing computational models employ conventional machine learning algorithms using static clinical variables without considering the temporal progression of patient symptoms.

Proposed Model

To overcome these limitations, this study developed an LSTM-based predictive model capable of learning temporal dependencies from sequential clinical data. The model was designed to analyse longitudinal patient information, including serial PSA measurements, symptom progression, urinary difficulty scores, inflammatory indicators, and demographic characteristics. By modelling temporal relationships among these variables, the proposed framework aims to improve the early detection and classification of prostatitis.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

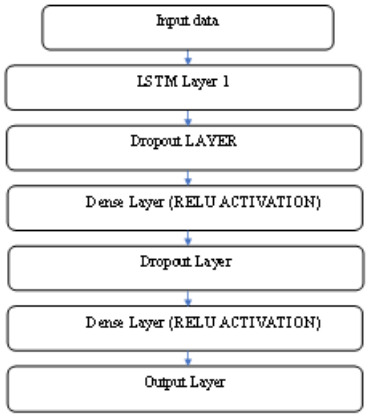

Figure 2: Proposed LSTM Architecture

Figure 2 presents the architecture of the proposed LSTM model consisting of an input layer, LSTM layer, dropout layer, dense layer, and sigmoid output layer.

The mathematical operations governing the LSTM network are expressed as follows.

Forget Gate

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

where $f_t$denotes the forget gate activation at time step t, $\sigma$is the sigmoid activation function, $W_f$is the weight matrix, h_{t-1}represents the previous hidden state, $x_t$is the current input vector, and $b_f$is the bias term. The forget gate determines which information from the previous cell state should be retained or discarded.

Input Gate

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

where $i_t$is the input gate responsible for determining the amount of new information incorporated into the memory cell.

Candidate Cell State

$$\tilde{C}_t = \tanh\,(W_c[h_{t-1}, x_t] + b_c)$$

where $\tilde{C}_t$represents the candidate cell state generated from the current input and previous hidden state using the hyperbolic tangent activation function.

Cell State Update

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

where $C_t$is the updated memory cell state, $C_{t-1}$is the previous cell state, and $\odot$ denotes element-wise multiplication. This operation combines retained historical information with newly acquired information.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

Output Gate

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

where $o_t$determines the amount of information transmitted from the cell state to the hidden state.

Hidden State

$$h_t = o_t \tanh(C_t)$$

where $h_t$is the hidden state output of the LSTM cell, serving as the input for the subsequent time step and network layers.

Dataset and Synthetic Data Generation

Due to the limited availability of publicly accessible longitudinal prostatitis datasets, this study employed a synthetically generated dataset developed from clinically relevant variables reported in the literature and established diagnostic guidelines. The dataset comprised demographic characteristics, serial PSA measurements, urinary symptom scores, pelvic pain scores, laboratory findings, and prostatitis class labels. Synthetic data generation enabled the creation of realistic patient records while eliminating patient privacy concerns and ensuring sufficient data for model training and evaluation.

Data Preprocessing

Prior to model training, missing values were addressed, categorical variables were encoded, and numerical variables were normalized using Min–Max feature scaling to improve convergence during training. Min–Max normalization transforms numerical features into the range of 0–1 and is defined as

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

where Xdenotes the original feature value, $X_{\min}$and $X_{\max}$are the minimum and maximum values of the feature, respectively, and $X'$is the normalized feature value.

Sequential records were then organised into fixed-length time windows suitable for LSTM processing. The processed dataset was partitioned into training, validation, and testing subsets for model development and evaluation.

LSTM Model Development

The proposed predictive model consisted of an input layer followed by stacked LSTM layers that extracted temporal features from sequential clinical observations. A dropout layer was incorporated to reduce overfitting by randomly deactivating neurons during training. Fully connected dense layers transformed the extracted features into high-level representations before the sigmoid output layer generated the probability of prostatitis.

The model parameters were optimized using the Adam optimizer, while Binary Cross-Entropy (BCE) served as the objective function.

The Binary Cross-Entropy loss function is expressed as

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

where $L$is the loss value, Ndenotes the total number of training samples, $y_i$is the true class label, and $\hat{y}_i$is the predicted probability. The loss function quantifies the discrepancy between predicted and actual class labels during model training.

Model optimization was performed using the Adaptive Moment Estimation (Adam) optimizer

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon}$$

where $\theta_t$denotes the model parameters at iteration t, $\alpha$is the learning rate, $\hat{m}_t$and $\hat{v}_t$are the bias-corrected estimates of the first and second moments of the gradient, respectively, and $\varepsilon$is a small constant introduced to improve numerical stability. Early stopping was employed to terminate training when validation performance no longer improved.

Model Evaluation

The performance of the proposed model was evaluated using Accuracy, Precision, Recall, Specificity, F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

Accuracy measures the proportion of correctly classified samples and is computed as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision measures the proportion of predicted positive samples that are correctly classified

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity) measures the proportion of actual positive samples correctly identified

$$Recall = \frac{TP}{TP + FN}$$

Specificity measures the proportion of actual negative samples correctly identified

$$Specificity = \frac{TN}{TN + FP}$$

The F1-score provides a balanced measure of precision and recall

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

where TP, TN, FP, and FNrepresent true positives, true negatives, false positives, and false negatives, respectively.

Finally, a confusion matrix and the ROC curve were employed to provide a comprehensive evaluation of the classification performance of the proposed LSTM model.

RESULTS AND DISCUSSION

Model Training Performance

The proposed LSTM model was trained using the Adam optimizer with binary cross-entropy as the loss function over 50 epochs and a batch size of 16.

Adam Optimizer

$$\theta_{t+1} = \theta_t - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

The Adam optimizer was adopted because it combines adaptive learning rates with momentum, resulting in faster convergence and improved stability during deep learning model training.

Prior to training, all clinical features were normalized using Min-Max scaling to improve convergence and numerical stability. The training and validation curves demonstrated steady convergence with minimal overfitting, indicating that the model generalized well to unseen data. The application of dropout regularization further enhanced model robustness.

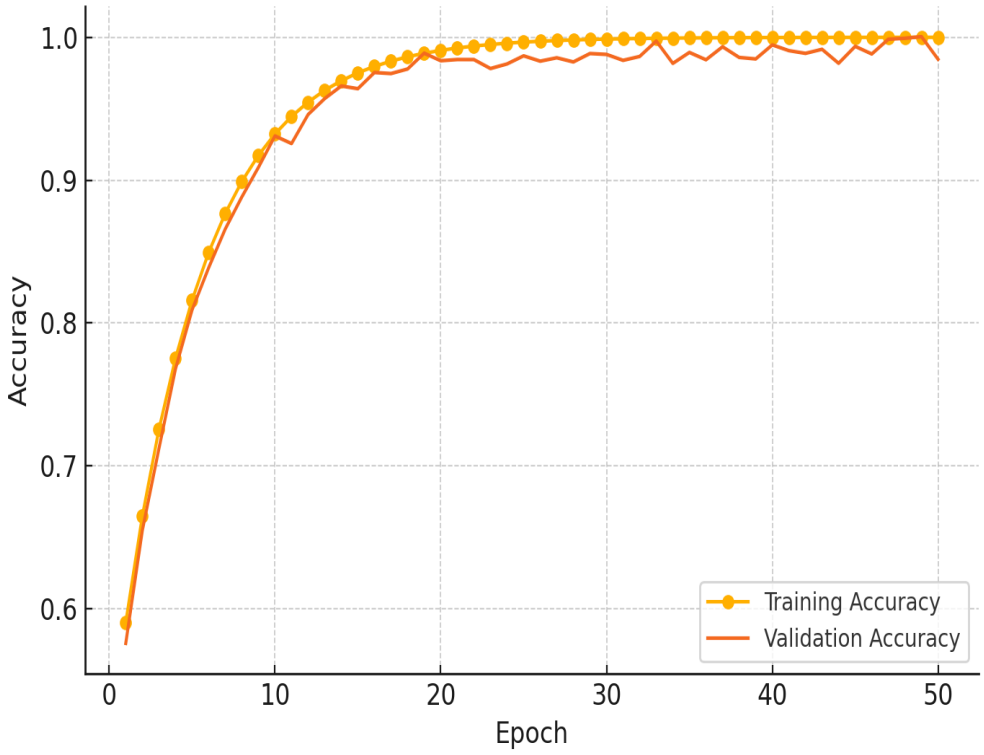

Figure 3: Training and Validation Accuracy Curve

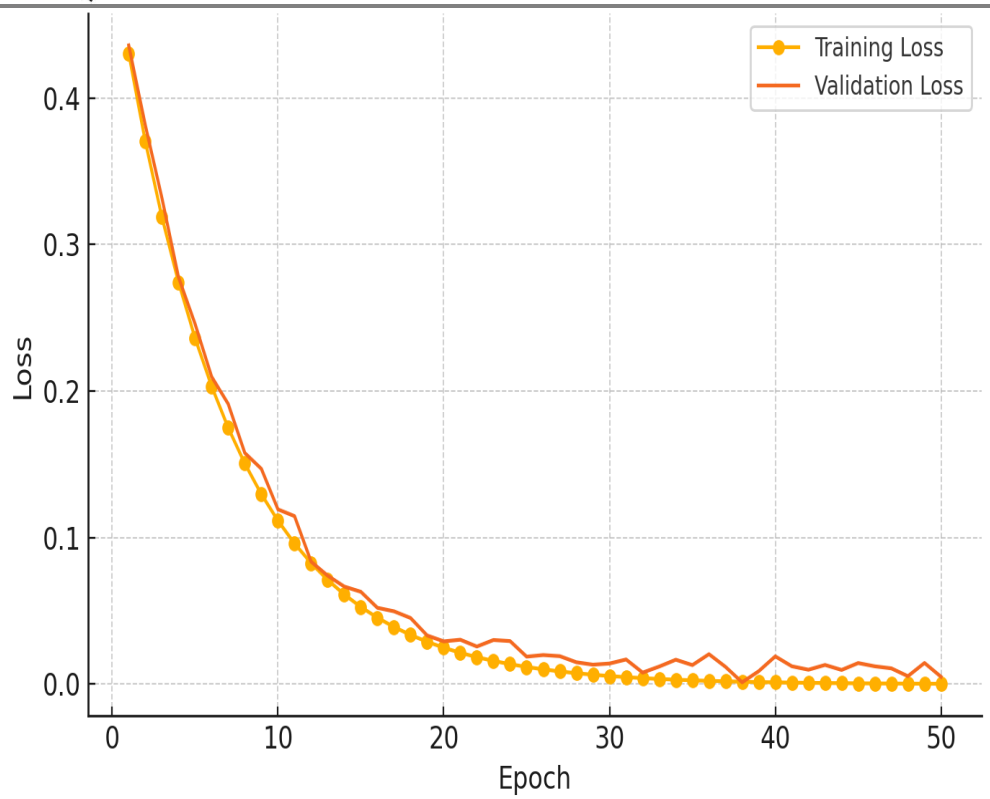

Figure 4: Training and Validation Loss Curve

Model Evaluation

The predictive performance of the proposed model was evaluated using accuracy, precision, recall, F1-score, specificity, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The model achieved an overall accuracy of **98.5%**, precision of **97.6%**, recall of **98.2%**, F1-score of **97.9%**, specificity of **96.9%**, and an AUC of **0.992**, demonstrating excellent discrimination between prostatitis and non-prostatitis cases.

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

$$Precision = \frac{TP}{TP + FP}$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

Specificity

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

$$Specificity = \frac{TN}{TN + FP}$$

F1-score

$$F1 = \frac{2PR}{P + R}$$

The Receiver Operating Characteristic (ROC) curve and the corresponding Area Under the Curve (AUC) were used to assess the discriminative capability of the proposed model, while the confusion matrix was employed to quantify true positives, true negatives, false positives, and false negatives.

Table 1. Performance Evaluation of the Proposed LSTM Model

Metric	Value
Accuracy	98.5%
Precision	97.6%
Recall (Sensitivity)	98.2%
Specificity	96.9%
F1-Score	97.9%
AUC-ROC	0.992

Confusion Matrix Analysis

The confusion matrix showed that the proposed model correctly classified 117 out of 120 test samples, comprising 54 true positives and 63 true negatives, with only 2 false positives and 1 false negative. These results indicate that the model maintained high sensitivity while minimizing false classifications, an important requirement for clinical decision-support systems.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue VI June 2026

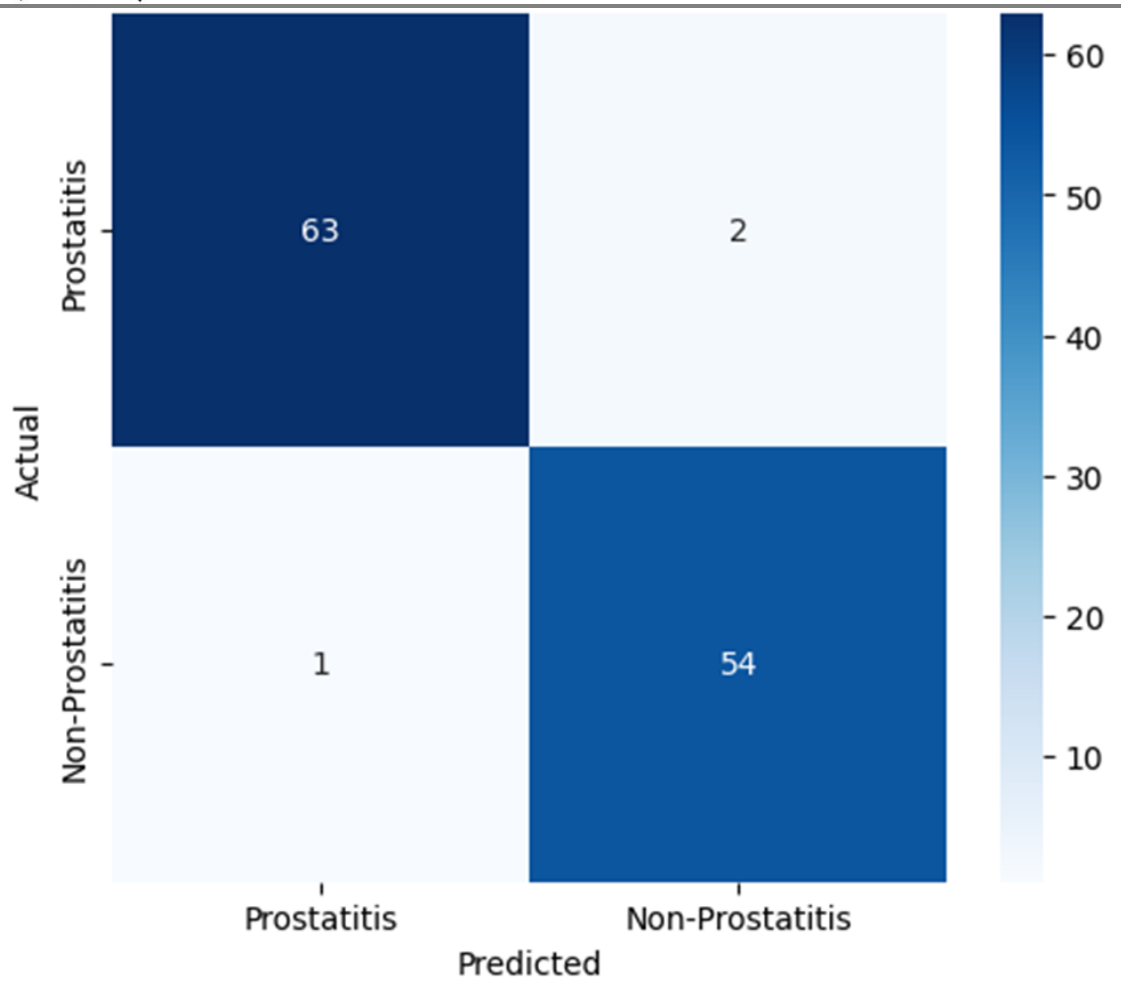

Figure 5: Confusion Matrix

DISCUSSION

The experimental results demonstrate the effectiveness of the proposed LSTM model in learning temporal relationships among sequential clinical variables associated with prostatitis. The high classification accuracy and AUC indicate that the model successfully distinguished prostatitis from non-prostatitis cases using longitudinal clinical information. Compared with previous studies, including [23]Razavian et al. (2016), who reported an AUC of 0.94 for disease prediction using longitudinal electronic health records, and[26] Sushentsev et al. (2023), who achieved an AUC of 0.86 for prostate disease prediction, the proposed model exhibited superior predictive performance. This improvement may be attributed to the integration of sequential PSA measurements, symptom progression, feature normalization, dropout regularization, and optimized LSTM architecture.

Despite these promising findings, the results should be interpreted with caution because the model was trained and evaluated on synthetically generated data. Although the synthetic dataset was designed to reflect clinically plausible disease progression patterns, it does not fully capture the variability, missing values, and noise inherent in real-world electronic health records. Consequently, further validation using real clinical datasets is necessary before the proposed framework can be adopted for routine clinical practice.

Study Limitations

The principal limitation of this study is the reliance on synthetic clinical data due to the limited availability of publicly accessible longitudinal prostatitis datasets. In addition, the fixed sequential representation may not fully capture complex long-term disease evolution. Future studies should validate the proposed framework using real-

world electronic health records and investigate advanced architectures such as attention-based LSTM and Transformer models to improve predictive performance and model interpretability.

CONCLUSION

This study developed and evaluated a Long Short-Term Memory (LSTM)-based framework for the early detection of prostatitis using synthetic longitudinal clinical data. The proposed model achieved excellent predictive performance, attaining an accuracy of **98.5%**, an F1-score of **97.9%**, and an AUC-ROC of **0.992**, demonstrating its effectiveness in modelling temporal clinical patterns. These findings indicate that LSTM networks have considerable potential as decision-support tools for the early diagnosis of prostatitis.

Although the model was trained on synthetic data, the study provides a strong proof-of-concept for applying deep learning to non-cancerous prostate diseases. Future research should focus on validating the proposed framework using real-world electronic health records, incorporating explainable artificial intelligence techniques, and evaluating its performance across diverse clinical populations to enhance its reliability and clinical applicability.

REFERENCES

1. Akinnuwesi, B. A., et al. (2022). Machine learning model for early prostate cancer detection using PCA and SVM. Health Information Science and Systems, 10(1), 1–13.
2. Alzboon, M. S., & Al-Batah, M. S. (2023). Prostate cancer detection and analysis using advanced machine learning. International Journal of Advanced Computer Science and Applications, 14(8), 386–394.
3. Alshahrani, S., Fathi, B. A., Abouelgreed, T. A., & ElMetwally, A. (2025). Prevalence of sexual dysfunction with chronic prostatitis/chronic pelvic pain syndrome (CP/CPPS): An updated systematic review and meta-analysis. Medicina, 61(6), 1110.
4. Bartoletti, R., Cai, T., Mondaini, N., et al. (2007). Prevalence and risk factors of CP/CPPS in Italian urology clinics: A multicenter epidemiological study.
5. Bonkat, G., Bartoletti, R., Bruyère, F., Cai, T., Geerlings, S. E., Köves, B., Schubert, S., & Wagenlehner, F. M. E. (2023). How I manage bacterial prostatitis. Clinical Microbiology and Infection, 29(1), 32–37.
6. Clemens, J. Q., Meenan, R. T., O'Leary, M. P., & McNaughton-Collins, M. (2013). Clinical characteristics and demographics of men with prostatitis-like symptoms in a community population. Urology, 41(3), 11–17.
7. Chowdary, P. S. R., & Kumar, P. R. (2021). Diabetes prediction using hybrid long short-term memory models. International Journal of Intelligent Systems and Applications in Engineering, 9(Special Issue), 26–31.
8. Goldenberg, S. L., Nir, G., & Salcudean, S. E. (2019). A new era: Artificial intelligence and machine learning in prostate cancer. Nature Reviews Urology, 16(7), 391–403.
9. Ghojogh, B., & Ghodsi, A. (2023). Recurrent neural networks and long short-term memory networks: Tutorial and survey. arXiv.
10. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
11. He, H., Luo, H., Xu, H., Qian, B., Zou, X., Zhang, G., Zeng, F., & Zou, J. (2023). Preclinical models and evaluation criteria of prostatitis. Frontiers in Immunology, 14, 1183895.
12. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.
13. Jang, H. A., Kim, J. W., & Oh, M. M. (2017). Prevalence and epidemiological characteristics of prostatitis in Korea. International Neurourology Journal, 21(4), 279–285.
14. Jang, T. L., McKiernan, J. M., & Weiss, R. E. (2017). Prostatitis: Epidemiology and clinical challenges. Korean Journal of Urology, 58(4), 210–216.
15. Liakos, A., Vogiatzis, D., & Stafylopatis, A. (2024). Artificial neural networks and their applications in healthcare: A review. Artificial Intelligence Review, 57, 1–29.

16. Maeda, K., Mizuno, K., Murakami, S., & Sakata, O. (2023). A review of current treatments for chronic prostatitis: What's new? International Journal of Urology, 30(4), 302–314.
17. Massaro, A., Maritati, V., Giannone, D., Convertini, D., & Galiano, A. (2019). LSTM DSS automatism and dataset optimization for diabetes prediction. Applied Sciences, 9(17), 3532.
18. National Institute of Diabetes and Digestive and Kidney Diseases. (2024). Prostatitis: Inflammation of the prostate. U.S. Department of Health and Human Services.
19. Nickel, J. C., Krieger, J. N., & Alexander, R. B. (2019). Diagnosis and management of prostatitis: A clinical review. Canadian Urological Association Journal, 13(11), E326–E332.
20. Pendegast, H. J., Leslie, S. W., & Rosario, D. J. (2025). Chronic prostatitis and chronic pelvic pain syndrome in men. In StatPearls.
21. Rahman, M., Islam, D., Mukti, R. J., & Saha, I. (2020). A deep learning approach based on convolutional LSTM for detecting diabetes. Computers in Biology and Chemistry, 88, 107329.
22. Ramírez-Alcocer, U. M., Rodríguez-Mazahua, L., López-Chau, A., Menéndez-Domínguez, V. H., & Cervantes, J. (2023). Time series forecasting using recurrent neural networks and long short-term memory models: A review. Applied Sciences, 13(6), 3676.
23. Razavian, N., Marcus, J., & Sontag, D. (2016). Multi-task prediction of disease onsets from longitudinal lab tests. arXiv preprint arXiv:1608.00647.
24. Rees, J., Abrahams, M., Doble, A., & Cooper, A. (2015). Diagnosis and treatment of chronic bacterial prostatitis and chronic prostatitis/chronic pelvic pain syndrome: A consensus guideline. BJU International, 116(4), 509–525.
25. Shiri, F. M., Perumal, T., Mustapha, N., & Mohamed, R. (2023). A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU. arXiv.
26. Sushentsev, N., et al. (2023). Predicting prostate cancer progression using LSTM applied to PSA and MRI series. European Urology Focus, 9(1), 112–121.
27. Zhang, J., Wang, L., Huang, Q., Long, P., & Dai, Y. (2019). The lifetime risk and prognosis of chronic prostatitis/chronic pelvic pain syndrome: An assessment. American Journal of Men's Health, 13(5), 1557988319865380.