# Blockchain-Enabled Smart Certificate Authentication System with Performance Prediction using DLSTM

**A. S. Sangeetha[1], S. Shunmugan[2]**

[1]Register No: 20123152292036, Research Scholar, Department of Computer Applications, S.T.Hindu College, Nagercoil, India,
Affiliated to Manonmaniam Sundaranar University, Abhishekapatti, Tirunelveli-627012, Tamilnadu, India.
Email: asn.sangeetha84@gmail.com
[2]Associate Professor, Department of Computer Science, S.T.Hindu College, Nagercoil, India,
Email: shunsthc@gmail.com
Affiliated to Manonmaniam Sundaranar University, Abhishekapatti, Tirunelveli-627012, Tamilnadu, India.

## Abstract

The integration of secure credential management and intelligent analytics is vital for modern educational ecosystems. Blockchain ensures immutable, decentralized, and tamper-resistant storage of student records while preserving privacy through cryptographic techniques, enabling secure verification without disclosing sensitive information. On the analytics front, DLSTM models are trained on temporally structured academic datasets to capture sequential patterns influencing student success. The experimental evaluation compares DLSTM with conventional machine learning models including Random Forest and Support Vector Machine (SVM). The results indicate that DLSTM achieves superior predictive performance with an accuracy of 97.7%, macro precision of 95.13%, macro recall of 96.75%, and macro F1-score of 95.89%. In comparison, Random Forest and SVM achieved accuracies of 96.2% and 93.6%, respectively, with lower recall and F1-scores.

**Keywords:** Blockchain, DLSTM, Academic, Certificate, Verification

## 1. Introduction

The verification of academic credentials [1] is critical for ensuring trust in educational and recruitment processes. With increasing global mobility in education and employment, traditional paper-based and centralized digital verification systems have become inadequate due to vulnerabilities such as data breaches, manipulation, and certificate forgery [2].

Blockchain technology, with its decentralized, immutable, and transparent ledger, offers a robust solution to these challenges [3], [4]. By eliminating intermediaries and enabling direct verification, blockchain reduces verification time, enhances trust, and safeguards data integrity. In parallel, assessing student performance has gained equal importance in academic and hiring decisions. Traditional predictive models often fail to capture temporal dependencies in student data, resulting in less accurate forecasts [5], [6].

Deep learning models, particularly Long Short-Term Memory (LSTM) networks, excel in modeling sequential patterns and time-dependent features, enabling more reliable predictions. The performance is further enhanced by the Bidirectional LSTM (Bi-LSTM) variation, which

ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 11 , Issu 10 2022

analyzes data both forward and backward. Incorporating explainable AI methods, such as SHAP (SHapley Additive exPlanations), enhances transparency by identifying the most influential factors in predictions.

This research proposes a Blockchain-Enabled Smart Certificate Authentication System integrated with a DLSTM-based Student Performance Prediction Model. The system ensures secure storage and verification of academic credentials on a blockchain network while predicting student performance using historical academic and behavioral data. This dual approach provides a comprehensive decision-support framework for universities, employers, and policymakers.

## 2. Background

### 2.1 Academic Certificate Frauds

Academic certificate frauds refer to fraudulent activities that occur during active educational processes or immediately before, during, or after the issuance of academic credentials. These include impersonation during online examinations using remote access tools, tampering with digital records in institutional databases, issuing certificates without proper evaluation, and unauthorized access to internal systems to modify marks or grades. In many cases, insiders such as administrative staff or faculty are complicit, helping students bypass academic protocols for a fee. Some frauds occur at the moment of certificate generation, where fake data is directly fed into the system to produce seemingly legitimate credentials. Additionally, online platforms and freelance services offer forged documents on demand, complete with fake verification portals that simulate authenticity. The dynamic nature of these frauds often executed within minutes using digital tools makes them difficult to trace through traditional checks. Real-time detection and prevention require robust identity verification, audit trails, blockchain-backed issuance, and AI-driven anomaly detection models to monitor suspicious behaviors in educational systems.

### 2.2 Blockchain for Certificate Verification

Blockchain technology offers a secure, transparent, and tamper-proof solution for academic certificate verification by transforming how credentials are issued, stored, and validated. Traditional verification systems rely heavily on manual processes and centralized databases, making them vulnerable to forgery, data manipulation, and delays. In contrast, blockchain enables institutions to issue digitally signed certificates that are recorded on a decentralized ledger. Once stored, these credentials become immutable. Any attempt to alter or forge the data is instantly detectable.

Verifiers such as employers, universities, or government bodies can instantly authenticate a certificate's origin, content, and validity without depending on intermediaries. Furthermore, integration with Decentralized Identifiers (DIDs) ensures that ownership and access rights remain with the individual, enabling self-sovereign identity management. By combining blockchain with secure public verification portals, the system supports real-time, trustless verification across borders. This reduces administrative overhead, enhances trust in digital education, and empowers learners with full control over their academic credentials. Projects like Blockcerts and MIT's Digital Diploma have shown promising results.

## 3. Literature Survey

### 3.1. Blockchain-Based Academic Certificate Verification and Authentication

Hasan, Rahman, and Islam proposed DistB-CVS, a distributed blockchain-based system tailored for the Bangladesh education sector, which uses smart contracts to automate certificate validation and prevent forgery while reducing verification time [7]. Similarly, Dinesh Kumar et al. presented a permissioned blockchain framework where only authorized institutions issue certificates, thus ensuring data integrity and minimizing dependency on intermediaries [8]. Ghazali and Saleh introduced a graduation certificate verification model employing cryptographic hashing and digital signatures, enabling direct verification by employers without intermediaries, although the system lacks detailed implementation specifics [9]. Li and Wu provided a comprehensive overview of blockchain-based academic certificate authentication, outlining system architecture and challenges such as scalability and privacy, highlighting blockchain's potential to streamline verification and reduce fraud [4].

Permissioned blockchains and smart contracts emerge as common solutions to control access and automate validation processes. However, challenges such as system scalability, privacy concerns, and comprehensive real-world implementations remain to be fully addressed.

### 3.2. Student Performance Prediction Using Deep Learning

Predictive analytics for student performance plays a crucial role in academic support and decision-making. Traditional methods often fall short in capturing temporal dependencies within sequential student data. Deep learning models, especially Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM) networks, have proven effective in this domain.

Yunita, Santoso, and Hasibuan [5]explored the use of deep learning models to predict student academic performance. Their study leveraged neural networks capable of modeling complex, nonlinear relationships within educational data, demonstrating that deep learning can outperform traditional machine learning methods in accuracy. By utilizing features such as prior grades, attendance, and demographic information, their model captured intricate patterns that influence student outcomes, highlighting the potential of deep learning techniques for effective academic prediction and early intervention strategies.

A system that uses blockchain to safely store student data and machine learning to forecast students' future employment positions based on their academic records was proposed by Shah et al., [10]. This system aims to provide a decentralized, tamper-proof database of student records, which can be accessed by universities, employers, and students. It highlights the benefits of combining blockchain's security with machine learnings predictive capabilities. The study also includes a SWOT analysis and discusses the implementation details, including the use of Python for blockchain representation and various machine learning models for prediction. The authors conclude that this integrated approach can enhance the credibility and accuracy of educational records and predictions.

Abu Zohair [11] explored the feasibility of using machine learning for performance prediction with limited datasets, focusing on graduate and postgraduate programs. The objective was to identify suitable classification models for predicting dissertation project grades and to determine key indicators that influence performance. The methodology included data preprocessing, attribute selection through visualization techniques such as heat maps and hierarchical clustering, and evaluation of multiple algorithms including Multilayer Perceptron, Linear Discriminant Analysis, Naïve Bayes, Support Vector Machine, and K-Nearest Neighbor. The results demonstrated that Support Vector Machine with a radial kernel achieved the highest accuracy, with 76.3 percent for dissertation grades and 69.7 percent for overall course grades. It was also observed that grades in core courses strongly correlated with dissertation outcomes, while pre-admission factors such as student age contributed significantly to prediction accuracy. The study emphasizes that reliable prediction models can be developed even with small datasets, making this approach particularly relevant for emerging or smaller educational institutions.

While blockchain-based academic credential verification and deep learning-based student performance prediction have matured individually, their integration remains limited. Combining secure certificate authentication with intelligent performance prediction can yield comprehensive academic systems that not only validate credentials but also provide actionable insights to institutions, employers, and students. Addressing this integration gap is vital for advancing trustworthy and data-driven academic ecosystems.

## 4. Proposed Framework

We propose a Smart Certificate Authentication System (SCAS) combining the strengths of Blockchain and Deep Learning.

### 4.1. Smart Certificate Authentication System using Blockchain

The integration of blockchain technologies and Student Performance Prediction is revolutionizing certificate verification systems by enhancing security, efficiency, and trustworthiness. In this framework, a user (e.g., a student) requests a certificate from an issuing authority (e.g., a university). IoT devices such as biometric scanners and exam monitoring systems collect real-time data, including attendance, assessment scores, and behavioral metrics. Upon successful verification, a smart contract on the blockchain is triggered to generate, sign, and store the certificate as an immutable digital asset. The student is notified, and a unique identifier linked to the blockchain record is issued by the university through the college. The student can obtain the certificate and his/her blockchain ID by logging into the SCAS system using their username and private key.

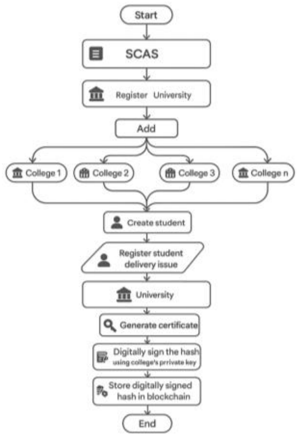


Fig.1. University Certificate Generation Process using Blockchain

Later, when a certificate holder applies for employment, they provide the employer with the blockchain-issued certificate ID and the corresponding certificate PDF. Using the certificate ID, the employer retrieves the stored certificate hash from the blockchain and verifies its digital signature using the issuing college's public key. Subsequently, the employer generates a hash from the submitted certificate PDF and compares it with the blockchain-stored hash. A matching result confirms the authenticity and integrity of the certificate, whereas a mismatch indicates potential tampering or invalidity.

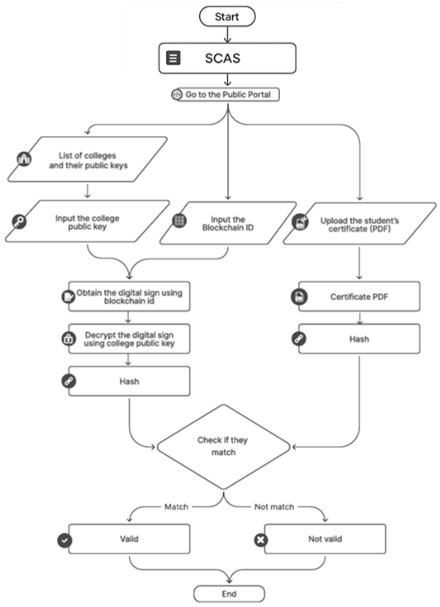


Fig.2. Certificate Verification Process

### 4.2. Student Performance Prediction using DLSTM Classifier

Deep Long Short-Term Memory (DLSTM) networks are an advanced form of LSTM, a type of Recurrent Neural Network (RNN) that overcomes vanishing gradient issues through memory cells and gating mechanisms. In DLSTM, multiple LSTM layers are stacked to capture both short- and long-term dependencies, enabling richer hierarchical pattern learning. This

ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 11 , Issu 10 2022

makes DLSTM effective for tasks like time-series forecasting and educational analytics, such as predicting student performance from historical academic and behavioral data.

### 4.3. Implementation Details

The selection of tools and platforms was based on their reliability, modularity, and compatibility with decentralized identity management and predictive analytics.

#### 4.3.1. Blockchain Layer

The blockchain backbone of the system is implemented using Hyperledger Indy, a permissioned distributed ledger platform designed specifically for decentralized identity and credential management. The system employs the Plenum Byzantine Fault Tolerant (BFT) consensus protocol to ensure fault tolerance and reliable transaction ordering. Blockchain interactions, including DID creation and credential exchange, are facilitated through the Indy-SDK using Python-based wrapper libraries.

#### 4.3.2. Artificial Intelligence Layer

Student performance prediction is carried out using a Deep Long Short-Term Memory (DLSTM) model developed with TensorFlow and Keras. The model is trained on time-series academic data, including attendance, internal assessments, coursework, and lab performance. Preprocessing, training, and evaluation tasks are performed using Scikit-learn and NumPy.

#### 4.3.3. Web and Backend Technologies

A web-based portal is developed using the Django framework to manage certificate requests, user authentication, and system interactions. The frontend interface leverages HTML to define content structure, CSS for visual presentation, and JavaScript to enable dynamic and interactive functionality. The system enables certificate verification via unique certificate ID lookup, providing real-time access to blockchain records.

#### 4.3.4. Storage Infrastructure

The metadata and transactional logs are managed using a PostgreSQL relational database for efficient querying and auditing.

#### 4.3.5. Security and Cryptographic Methods

To ensure data integrity, SHA-256 hashing is used to generate unique digital fingerprints for certificate contents prior to blockchain storage. Elliptic Curve Cryptography is used for safe key exchange and digital signatures. For session security and user authorization in the web interface, JWT (JSON Web Tokens) are implemented.

#### 4.3.6. Deployment

All system components are containerized using Docker to ensure consistency and portability across development and deployment environments. The blockchain node, Django backend, and AI containers are all orchestrated via Docker Compose.

### 5. Evaluation and Results

### 5.1 Dataset

The dataset used for student performance prediction comprises both temporal and static features, capturing academic, behavioral, and socio-demographic aspects. It was constructed using a combination of real-world data collected from Postgraduate (PG) students in Kanyakumari district, Tamil Nadu, India, and synthetically generated data to ensure sufficient volume for deep learning training. Proper consent was obtained from participants, and all personally identifiable information was anonymized. The final dataset contains 10,000 records, maintaining the statistical distribution and correlation patterns observed in the real data.

Temporal features include Semester Number (integer values from 1 to 6, representing time steps), GPA, Attendance, Assignment Score, MOOC Courses Completed, Backlogs, Extracurricular Activities, Self-study Hours, Stress Level, and Motivation Level, all normalized to a range between 0 and 1. Additionally, Sports Participation and Internship Participation are encoded as binary values (0 or 1) to indicate involvement.

Static features such as Gender, Parent's Education, and Socioeconomic Status are represented using one-hot encoding to capture categorical information without imposing ordinal relationships. This comprehensive feature set enables the model to learn both short-term academic fluctuations and long-term performance trends, while also incorporating contextual factors that may influence a student's academic trajectory

### 5.2. Evaluation Metrics

In order to comprehensively evaluate the performance of the implemented models, several key classification metrics were employed. Accuracy measures the ratio of correctly predicted instances to the total instances evaluated, serving as a simple metric for overall model effectiveness. However, accuracy alone may not be sufficient when dealing with imbalanced datasets, as it can obscure poor performance in minority classes.

To address this limitation, Macro Precision, Macro Recall, and Macro F1-score were considered. Macro Precision computes the precision for each class independently and then averages the values, ensuring that all classes contribute equally regardless of their frequency. This metric is particularly useful in assessing the reliability of the model's positive predictions for each class.

Similarly, Macro Recall measures the ability of the model to correctly identify instances belonging to each class, averaged across all classes. It provides insight into the model's sensitivity, particularly for less represented classes. The Macro F1-score serves as the harmonic mean of Macro Precision and Macro Recall, capturing a balance between the two. By averaging the F1-score across all classes, this metric ensures that the evaluation reflects consistent performance across both majority and minority classes, rather than being dominated by the most frequent categories.

## 6. Evaluation and Analysis

### 6.1. Quantitative Evaluation

The performance of all models was evaluated using Accuracy, Macro Precision, Macro Recall, and Macro F1-score. As summarized in Table 1, DLSTM achieved the highest overall performance with an accuracy of 97.7%, macro precision of 95.13%, macro recall of 96.75%, and macro F1-score of 95.89%. Random Forest followed with an accuracy of 96.2%, while SVM recorded the lowest performance at 93.6%.

**Table 1 – Model Performance Summary**

| **Model** | **Accuracy** | **Macro Precision** | **Macro Recall** | **Macro F1-score** |
|---|---|---|---|---|
| DLSTM | 0.977 | 0.951305 | 0.967534 | 0.958894 |
| Random Forest | 0.962 | 0.948692 | 0.928435 | 0.937830 |
| SVM | 0.936 | 0.927834 | 0.856783 | 0.888272 |

 

### 6.2. Confusion Matrix Analysis

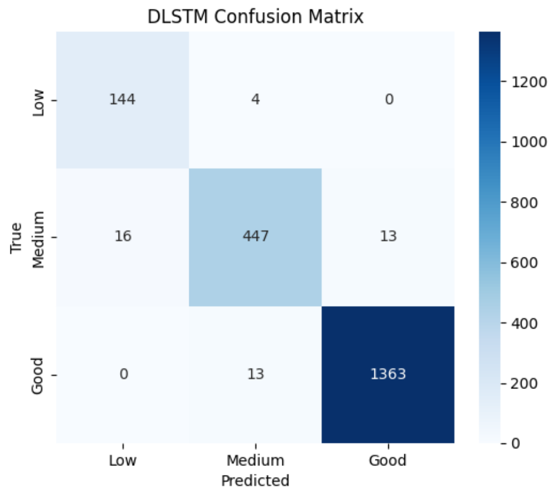


Fig.3. Confusion Matrix of DLSTM

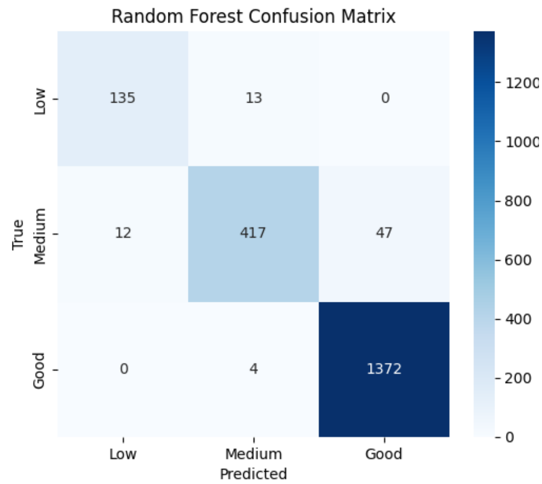


Fig.4. Confusion Matrix of Random Forest

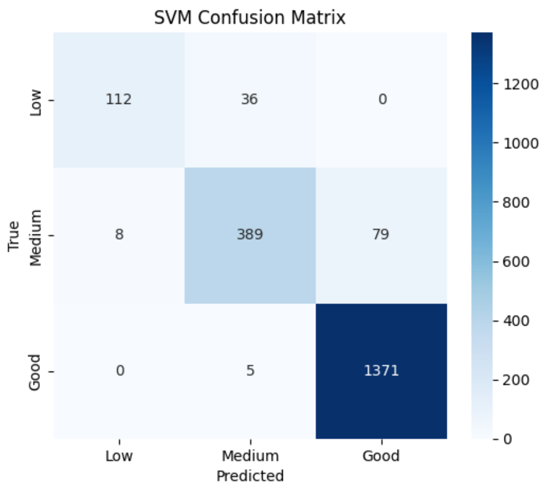


Fig.5. Confusion Matrix of SVM

The confusion matrices in Fig.3, Fig.4, Fig.5 provide a detailed view of each model's classification strengths and weaknesses. The DLSTM model in Fig.3 demonstrates consistently strong performance across all categories, with minimal misclassifications in the "Low" and "Good" classes, and a well-balanced distribution in the "Medium" category. The Random Forest model in Fig.4 also achieves high performance but exhibits slightly higher misclassification in the "Medium" category, which could negatively impact recall for borderline students. In contrast, the SVM model in Fig.5 shows a higher degree of confusion between the "Medium" and "Good" categories, contributing to its comparatively lower macro recall and F1-score.

## 7. Conclusion

This study presented a Blockchain-Enabled Smart Certificate Authentication System integrated with a DLSTM-based Student Performance Prediction Model. The blockchain component ensures secure, transparent, and tamper-proof storage and verification of academic credentials, effectively addressing the challenges of centralized verification systems prone to fraud and data manipulation. The DLSTM model, trained on a hybrid dataset comprising real and synthetic records from postgraduate students, accurately predicts student performance by capturing both temporal academic fluctuations and static socio-demographic influences.

Experimental results demonstrated that the proposed DLSTM model outperformed traditional machine learning approaches, achieving higher accuracy, precision, recall, and F1-scores. This dual-framework approach can be beneficial for universities, employers, and policymakers in streamlining recruitment, admissions, and academic planning processes. Future work will explore the incorporation of privacy-preserving techniques such as federated learning, expansion to multi-institutional datasets, and deployment on scalable blockchain infrastructures to enhance applicability in real-world scenarios.

## References

1. E. Leka and B. Selimi, "Development and evaluation of blockchain based secure application for verification and validation of academic certificates," *Annals of Emerging Technologies in Computing*, vol. 5, no. 2, pp. 22–36, 2021, doi: 10.33166/AETiC.2021.02.003.
2. A. Camilleri and G. A. F., "Blockchain in Education," *Jrc Science for Policy Report*, no. June, pp. 1–136, 2017, doi: 10.2760/60649.
3. A. Zhang and X. Lin, "Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain," *J. Med. Syst.*, vol. 42, no. 8, 2018, doi: 10.1007/s10916-018-0995-5.
4. R. Li and Y. Wu, "Blockchain based Academic Certificate Authentication System Overview."
5. A. Yunita, H. B. Santoso, and Z. A. Hasibuan, "Deep Learning for Predicting Students' Academic Performance," *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*, Oct. 2019, doi: 10.1109/ICIC47613.2019.8985721.
6. M. A. Al-Barrak and M. Al-Razgan, "Predicting Students Final GPA Using Decision Trees: A Case Study," *International Journal of Information and Education Technology*, vol. 6, no. 7, pp. 528–533, 2016, doi: 10.7763/IJIET.2016.V6.745.
7. M. Hasan, A. Rahman, and M. J. Islam, "DistB-CVS: A Distributed Secure Blockchain based Online Certificate Verification System from Bangladesh Perspective," vol. 2020, pp. 28–29, 2020, doi: 10.13140/RG.2.2.14973.59369.
8. K. Dinesh Kumar, P. Senthil, and D. S. Manoj Kumar, "Educational certificate verification system using blockchain," *International Journal of Scientific and Technology Research*, vol. 9, no. 3, pp. 82–85, 2020.

9. [9] O. Ghazali and O. S. Saleh, "A Graduation Certificate Verification Model via Utilization of the Blockchain Technology".
10. D. Shah, D. Patel, J. Adesara, P. Hingu, and M. Shah, "Integrating machine learning and blockchain to develop a system to veto the forgeries and provide efficient results in education sector," *Vis. Comput. Ind. Biomed. Art*, vol. 4, no. 1, Dec. 2021, doi: 10.1186/s42492-021-00084-y.
11. L. M. Abu Zohair, "Prediction of Student's performance by modelling small dataset size," *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 1–18, Dec. 2019, doi: 10.1186/S41239-019-0160-3/FIGURES/13.