Explainable AI-Blockchain System for Trustworthy Carbon Capture and Storage Predictive Analytics

Ihunanya Udodiri Ajakwe, Simeon Okechukwu Ajakwe *MIEEE*, Dong-Seong Kim *SMIEEE*
Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
ihunanya.ajakwe@gmail.com, simeonajlove@gmail.com, dskim@kumoh.ac.kr

***Abstract*—This paper introduces an innovative Explainable AI (XAI)-Blockchain system designed for predictive analytics in Carbon Capture and Storage (CCS). The proposed system leverages the transparency and immutability of blockchain technology to ensure the integrity and auditability of AI-driven predictions regarding the efficiency and safety of CCS projects' operations. By integrating XAI techniques, the system provides insights into the reasoning behind these predictions, fostering stakeholder trust and understanding. Compared with existing state-of-the-art methods in CCS analytics, the result shows that a high capture amount and favorable district code contributed to the CCS project's efficiency.**

***Index Terms*—Blockchain, Carbon Capture and Storage, Explainable AI, Predictive Analytics, Security, Trustworthy.**

I. Introduction

Carbon Capture and Storage (CCS) is a critical technology for mitigating climate change by capturing carbon dioxide (CO_2) emissions from industrial sources and storing them safely underground [1]. Effective monitoring and prediction of CCS processes are essential for ensuring their long-term viability and environmental safety. Traditional methods often rely on complex simulations and statistical models, which can be opaque and lack inherent data integrity. Few researchers have explored the application of AI and blockchain in environmental monitoring and management, such as the use of AI for predicting CO_2 leakage in storage sites [2] and optimizing capture processes by authors [3]. Also, blockchain has been applied to track carbon credits and ensure transparency in carbon markets [4].

However, the integration of explainable artificial intelligence (XAI) with blockchain for predictive analytics in CCS is relatively nascent and unexplored. Our work builds upon these individual advancements to create a novel, comprehensive solution. XAI has emerged as a crucial field, aiming to make AI decision-making processes more transparent and interpretable, while blockchain offers data integrity and security via decentralization and immutable ledgers [5]. The integration of blockchain with AI has the potential to create trustworthy and auditable cognitive systems for CCS predictive analytics.

II. Proposed System Design and Methodology

The proposed Explainable AI-Blockchain system for predictive CCS analytics comprises three key layers, as seen in Fig. 1. The DAP layer collects data from various CCS projects and operations, such as project ID, geographical location, cost, longitude, latitude, etc., and stores it. The data is then preprocessed to ensure quality and compatibility for AI model training in the XAI cognition layer. Then, the XAI Cognition layer uses classical machine learning to train the preprocessed data from DAP to predict key CCS parameters such as operational efficiency.

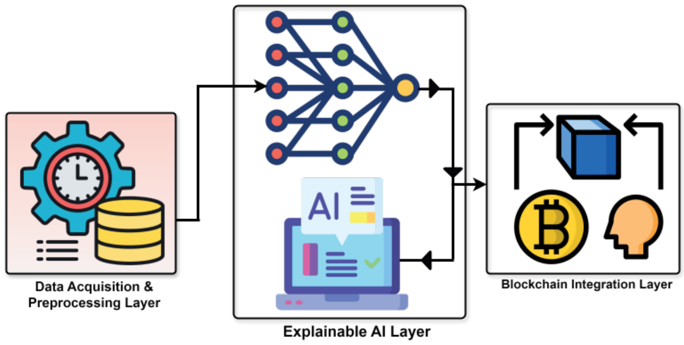

Fig. 1. System Design showing the (a) Data Acquisition & Processing (DAP) layer, (b) XAI cognition layer, and (c) Blockchain Integration (BI) Layer.

In this work, we used the Extreme Gradient Boosting (XGB) to forecast the efficiency of CCS projects as seen in equation (1). The prediction for a sample x_i in XGB is the sum of K regression trees:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{1}$$

where $\mathcal{F}$ is the space of regression trees (e.g., CARTs). Each f_k corresponds to a decision tree. The XGB objective function is to minimize as expressed in equation (2):

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2}$$

where l is a differentiable convex loss function (e.g., squared error), - $\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2$ is the regularization term, - T is the number of leaves in the tree, - w are the leaf weights, - γ and λ are regularization parameters. The XGB model trains the data and compares its performance with Linear Regression (LR) and Random Forest (RF) models. Then, XAI techniques SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), are integrated to provide explanations for the model's predictions, highlighting the factors influencing the outcomes. Finally, the BI layer uses a permissioned blockchain network to record the model parameters, prediction results, and the corresponding XAI explanations. Each transaction is timestamped and cryptographically secured to ensure data integrity and auditability.

For model training, the CCS dataset [6] from the National Energy Technology Laboratory was used, as it contains information on CCS projects worldwide.

TABLE I
MODEL PREDICTION CAPABILITY & RELIABILITY

Projects Model	Active				Inactive				Mse	Time (s)	Auc-Roc
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score			
LR	0.65	0.22	0.06	0.10	0.65	0.59	0.86	0.70	0.25	0.01	0.40
RF	0.67	0.42	0.52	0.47	0.63	0.64	0.55	0.59	0.28	0.74	0.57
XGB	**0.67**	**0.46**	**0.52**	**0.49**	**0.66**	**0.66**	**0.61**	**0.63**	**0.28**	**0.24**	**0.56**

It has 418 samples with 38 features. After feature importance, only 5 features are relevant, namely, `size or CaptureAmount/Unit base power`, `cost`, `latitude`, `longitude`, & `district code`. To avoid model overfitting, the dataset was divided into 80/20 for model training and evaluation. Simulation was performed in a Python environment with appropriate frameworks and libraries on Windows. BI layer used Hyperledger with PBFT.

III. RESULT AND PERFORMANCE EVALUATION

The results in Table I highlight the AI model's prediction of the status of CCS projects across the globe. Unlike the LR and RF models, the XGB achieved a considerable prediction of active CCS projects and inactive projects with an accuracy of 67% and 66% and F1-score of 49% and 63%. Also, the XGB model had better prediction reliability and efficiency compared to LR and RF, with a minimal prediction error of 0.28, a latency of 0.24 s, and an AUC-ROC value of 0.56 (which is a bit lower than RF but better when compared across all metrics). The comparative analysis of the reliability of the models is seen in the confusion matrix in Fig. 2 (a)-(c).

The results in Fig. 2 (d)-(e) offer information on the parameters that contributed to the status of the CCS project. The LIME result in Fig. 2 reveals that the prediction scale (0.12) is relatively low and that the two strongest contributors to raising the prediction are the high capture amount and favorable district code. However, this is offset by the very high cost and the location (longitude/latitude), which brings the prediction down. This is an invaluable insight for understanding CCS-efficiency prediction in specific scenarios.

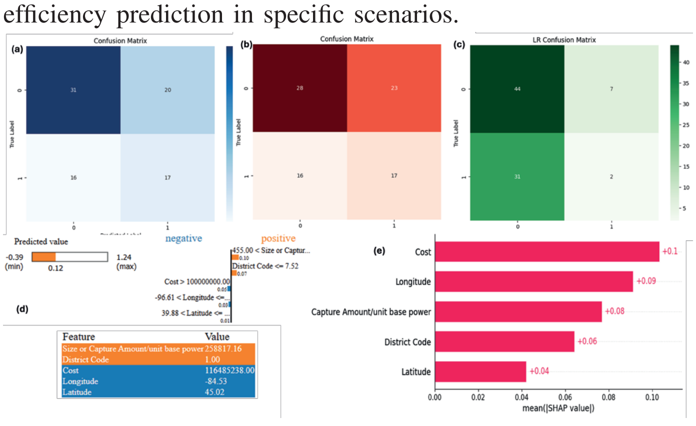

Fig. 2. (a)-(c) for Model Prediction Reliability and (d)-(e) for Model Explainability, highlighting parameters that determine the CCS project efficiency in different countries.

The SHAP analysis in Fig. 2 (e) shows that cost (0.10) is the most influential feature in predicting capture efficiency, as it significantly shifts the prediction of the model. Longitude (0.09) and the Size/Capture Amount (0.08) are also strong predictors. On the other hand, District Code and Latitude contribute less but still have a measurable influence. This implies that spatial and economic characteristics are highly correlated with how well a carbon capture system performs.

Unlike current state-of-the-art methods for CCS analytics that often rely on numerical simulations and statistical models, storing the model's prediction data and insights in a secured hyperledger offers significant improvement towards achieving zero-net carbon emission, such as clear explanations for the AI-driven predictions, increasing trust and understanding among stakeholders, immutability and auditability of the data and predictions, and reliable and trustworthy analytics for critical CCS operations.

IV. CONCLUSION AND FUTURE WORK

This paper presents a novel Explainable AI-Blockchain system for predictive carbon capture and storage analytics. By integrating the strengths of AI, XAI, and blockchain, the proposed system offers a transparent, secure, and reliable approach to predicting and understanding critical CCS parameters that determine CCS projects' efficiency. Future work will focus on developing a prototype of the proposed system and evaluating its performance on real-world CCS datasets. Further research will also explore the scalability and cost-effectiveness of the integrated blockchain solution.

ACKNOWLEDGMENT

This work was partly supported by Innovative Human Resource Development for Local Intellectualization program through the IITP grant funded by the Korean government (MSIT) (IITP-2025-RS-2020-II201612, 25%) and by Priority Research Centers Program through the NRF funded by the MEST (2018R1A6A1A03024003, 25%) and by the MSIT, Korea, under the ITRC support program (IITP-2025-RS-2024-00438430, 25%), by the IITP (Institute of Information & Communications Technology Planning & Evaluation)-ICAN (ICT Challenge and Advanced Network of HRD) grant funded by the Korean government (Ministry of Science and ICT) (IITP-2025-RS-2022-00156394, 25%).

REFERENCES

[1] I. U. Ajakwe, S. O. Ajakwe, J. M. Lee, and D.-S. Kim, "Iot-blockchain frameworks in environmental pollution monitoring and data management," in *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*, 2024, pp. 2147–2151.

[2] S. Harati, S. R. Gomari, M. A. Rahman, R. Hassan, I. Hassan, A. K. Sleiti, and M. Hamilton, "Performance analysis of various machine learning algorithms for co2 leak prediction and characterization in geo-sequestration injection wells," *Process Safety and Environmental Protection*, vol. 183, pp. 99–110, 2024.

[3] S. Manikandan, R. S. Kaviya, D. H. Shreeharan, R. Subbaiya, S. Vickram, N. Karmegam, W. Kim, and M. Govarthanan, "Artificial intelligence-driven sustainability: Enhancing carbon capture for sustainable development goals–a review," *Sustainable Development*, vol. 33, no. 2, pp. 2004–2029, 2025.

[4] A. Boumaiza and K. Maher, "Leveraging blockchain technology to enhance transparency and efficiency in carbon trading markets," *International Journal of Electrical Power & Energy Systems*, vol. 162, p. 110225, 2024.

[5] S. O. Ajakwe and D.-S. Kim, "Facets of security and safety problems and paradigms for smart aerial mobility and intelligent logistics," *IET Intelligent Transport Systems*, vol. 18, pp. 2827–2855, 2024.

[6] N. E. T. Laboratory, "Carbon capture and storage database," https://netl.doe.gov/carbon-management/carbon-storage/worldwide-ccs-database, 2024, accessed: 2025-05-01.