Fraud Detection in Medical Insurance Claim Systems using Machine Learning

[1]*Mr. Krishna Annaboina (Assistant professor) Computer Science and Engineering (Internet Of Things) Guru Nanak Institutions Technical Campus Telangana, India anneboina.krishna@gmail.com*

[2]*Samala Prasoona (U.G. Student) Computer Science and Engineering (Internet Of Things) Guru Nanak Institutions Technical Campus Telangana, India samalaprasoona03@gmail.com*

[3]*Chada Ashritha (U.G. Student) Computer Science and Engineering (Internet Of Things) Guru Nanak Institutions Technical Campus Telangana, India ashrithachada2004@gmail.com*

[4]*Pesara Chakradhar Reddy (U.G. Student) Computer Science and Engineering (Internet Of Things) Guru Nanak Institutions Technical Campus Telangana, India chakradharreddy9542@gmail.com*

Abstract –

Fraud detection in medical insurance claim systems is crucial for preserving healthcare service integrity and minimizing financial losses. This study explores the application of Support Vector Machines (SVM) enhanced by GridSearchCV for hyperparameter optimization, aiming to detect fraudulent claims effectively. The research methodology involves preprocessing a comprehensive medical insurance claims dataset, focusing on extensive feature selection and engineering to improve model performance. GridSearchCV is utilized to conduct an exhaustive search over specified parameter ranges, identifying the optimal hyperparameters for the SVM model. To evaluate the model's effectiveness, metrics such as accuracy, precision, recall, and F1-score are employed. The results indicate that the optimized SVM model significantly improves the detection of fraudulent claims, outperforming baseline models. This study underscores the efficacy of integrating SVM with GridSearchCV in developing robust fraud detection systems, contributing to more reliable and efficient processing of medical insurance claims.

Keywords – *Fraud detection, Machine learning, Support Vector Machines (SVM),GridSearchCV, Hyperparameter optimization, Model accuracy, Evaluation metrics, Accuracy, Precision.*

I. INTRODUCTION

The healthcare industry generates vast amounts of data, including patient records, insurance claims, and clinical information. Detecting fraudulent claims, which cause significant financial losses, is a critical challenge. Manual review of claims is impractical due to its time and cost demands, making automated systems essential.This research proposes using Support Vector Machines (SVM), a robust classification algorithm, to identify fraudulent claims efficiently. To optimize the model's accuracy, GridSearchCV is employed to fine-tune hyperparameters. This automated approach accelerates claim evaluation, reduces manual investigation needs, and enhances fraud detection precision, minimizing financial losses.Globally, the insurance sector, including emerging markets like India, faces similar challenges, with rapid growth increasing fraud risks. This machine learning framework can improve operational efficiency and safeguard the financial integrity of insurers worldwide.

PROBLEM STATEMENT:

The increasing prevalence of fraudulent medical insurance claims creates substantial financial losses and undermines the efficiency and fairness of healthcare systems. Fraud contributes to higher premiums and reduced trust among legitimate stakeholders. Traditional detection methods, such as manual audits and rule-based systems, are inadequate for managing the growing scale and sophistication of fraud.Addressing this issue requires an automated, scalable, and accurate fraud detection system capable of analyzing large, complex datasets, detecting patterns, and adapting to evolving fraud tactics. Machine learning offers a promising solution by enabling anomaly detection, fraud classification, and real-time processing. Furthermore, ensuring data security and privacy remains essential to building a robust framework that reduces financial losses and strengthens the integrity of the medical insurance ecosystem

OBJECTIVE:

The primary objective of the project is to develop an efficient, scalable, and accurate fraud detection system for medical insurance claims using machine learning. This system aims to identify and mitigate fraudulent activities in insurance claims by leveraging advanced machine learning techniques for anomaly detection, classification, and predictive analysis.

Specific objectives include:

1. **Fraud Detection:** To accurately detect and classify fraudulent claims using machine learning models with high precision, recall, and overall performance.

2. **Data Security:** To ensure the privacy and security of sensitive medical and financial data during processing and analysis.

3. **Real-Time Application:** To design a framework capable of real-time fraud detection for immediate decision-making.

4. **Cost Efficiency:** To reduce financial losses for insurance providers by minimizing false claims while maintaining operational efficiency.

5. **Adaptability:** To build a system that adapts to evolving fraud tactics and remains effective against novel fraud schemes.

6. **Comprehensive Insights:** To provide actionable insights into fraudulent patterns and scenarios, enabling better prevention strategies.

7. **Stakeholder Benefits:** To enhance trust and transparency among all stakeholders, including insurance companies, healthcare providers, and policyholders, by improving the integrity of the claims process.

By achieving these objectives, the project seeks to improve the overall effectiveness and reliability of fraud detection systems in the medical insurance domain.

EVALUATION:

The evaluation of a machine learning-based fraud detection system for medical insurance claims involves a comprehensive analysis of its performance, robustness, scalability, real-world applicability, and impact on business operations. Key performance metrics such as accuracy, precision, recall, F1 score, and AUC-ROC curves are used to assess the system's ability to reliably classify fraudulent and legitimate claims. Robustness is evaluated by analyzing the system's capability to handle data preprocessing tasks, including cleaning raw data, managing missing information, and addressing imbalanced datasets. Scalability and computational efficiency are tested by simulating high transaction volumes and assessing the system's training and prediction times on large datasets. To ensure real-world applicability, the model is validated on unseen datasets to determine its adaptability to new and evolving fraud patterns. Security and privacy are critical considerations, requiring the system to comply with data protection regulations and safeguard sensitive medical and financial information. The system's impact on business operations is measured by quantifying cost savings, enhancing operational efficiency, and reducing manual verification efforts. User experience is assessed based on stakeholder feedback regarding usability, transparency, and seamless integration into existing workflows. Continuous updates and testing are essential to maintaining the system's relevance, incorporating feedback, and adapting to emerging fraud tactics. Overall, the success of the project is defined by achieving high-performance metrics, reducing fraudulent claims and financial losses, and gaining positive feedback from end-users while ensuring long-term effectiveness and sustainability.

II. LITERATURE SURVEY

The application of machine learning (ML) and blockchain technologies in fraud detection for medical insurance claims has introduced innovative approaches to enhance detection accuracy and system resilience**. Najmeddine Dhieb et al. (2022)** demonstrated the effectiveness of **XGBoost** in fraud detection, achieving a 7% improvement in accuracy compared to decision tree models. Their study highlights the potential for developing AI-based solutions tailored to specific insurance domains, opening avenues for applications beyond medical insurance. **Similarly, Sudeep Tanwar et al. (2021)** integrated ML techniques, including **Support Vector Machines (SVM),** Clustering, Bagging, and **Convolutional Neural Networks (CNN),** with **Blockchain Technology (BT).** This integration not only improved fraud detection capabilities but also enhanced resistance to attacks. Future research aims to address challenges such as infrastructure availability, quantum resilience, and data privacy concerns.

Other studies focus on combining traditional and advanced models for better performance**. Mathias Bartl** and Simone **Krummaker (2021)** explored the use of **Decision Trees (DT)** alongside **Random Forests (RF)** to predict fraud in insurance claims. Their findings suggest that integrating conventional actuarial methods, such as Chain-Ladder or Bornhuetter- Ferguson, with ML models can improve prediction accuracy. Similarly, **Leila Ismail and Sherali Zeadally (2021)** proposed a **blockchain-based framework**, Block-HI, for fraud detection in health insurance. Their research emphasized the need to compare blockchain-based systems with traditional manual methods while focusing on enhancing interoperability among healthcare data systems.

Sequence mining has also been utilized in healthcare claims fraud detection. **Irum Matloob et al. (2020)** achieved up to 85% accuracy using **sequence mining** but noted challenges in ensuring data privacy and managing time-intensive preprocessing. **G. Kowshalya** and **Dr. M. Nandhini (2021)** employed **Naïve Bayes, J48,** and **Random Forest algorithms**, with Random Forest outperforming the others in fraud prediction. Future studies may investigate correlations

between claims and premium amounts to further refine classification algorithms for real-world datasets.

The integration of blockchain technology in healthcare fraud detection is gaining traction. **Shuai Wang et al. (2020)** proposed a **Blockchain-powered Parallel Healthcare System (PHS),** combining artificial system modeling and computational experiments for improved diagnosis and fraud detection. Their research emphasized the importance of consortium blockchains to enhance security, scalability, and integrity. Similarly, **Sabyasachi Chakraborty et al. (2021)** incorporated blockchain with Internet of Things (IoT) technologies, leveraging wearable devices and bio-sensors for decentralized health data collection. Their future work focuses on improving the secure governance of data generated by such devices.

Blockchain benchmarking frameworks have also been explored. **Tien Tuan Anh Dinh et al. (2020)**

developed BLOCKBENCH to evaluate blockchain platforms' performance in data processing, suggesting improvements in blockchain design through database principles. **Valentina Gatteschi et al. (2021)** examined blockchain and smart contract applications in insurance to lower operational costs, increase transparency, and improve customer satisfaction. Their findings suggest the potential for extending blockchain solutions to other industries with similar operational challenges.

Smart contracts are also gaining attention in insurance automation. **Mayank Raikwar et al. (2021)** proposed blockchain-based frameworks to automate processes, reduce administrative expenses, and enhance fraud detection. They recommended extending blockchain encryption capabilities to improve security. **Riya Roy** and **Thomas George K (2022)** compared ML algorithms such as Decision Tree, Random Forest, and Naïve Bayes for auto insurance fraud detection. Decision Tree and Random Forest showed superior performance, with future research aimed at testing additional algorithms to enhance precision and recall.

Finally, **Xueping Liang et al. (2020)** proposed a Hyperledger Fabric-based blockchain system for sharing health data, demonstrating its scalability and efficiency in handling large datasets. Their work suggests potential applications for broader healthcare fraud detection scenarios. **Fei Tang et al. (2020**) introduced an authentication mechanism for blockchain-based **Electronic Health Records (EHRs),** achieving lower computational and communication costs. Future studies may explore alternative algorithms to further enhance the security of blockchain-based healthcare solutions.

This body of research demonstrates the significant advancements and challenges in leveraging ML and blockchain technologies for fraud detection, paving the way for more secure, efficient, and accurate systems in the healthcare and insurance industries.

DATASETS:

The datasets used in this project include a combination of publicly available, synthetic, and proprietary data sources to ensure robust training and evaluation of the fraud detection model:

1. **Medicare Claims Data**: This publicly available dataset, provided by the Centers for Medicare & Medicaid Services (CMS), includes comprehensive healthcare provider and claims information. It helps identify fraudulent activities through unusual billing patterns and discrepancies in claims.
2. **Health Insurance Fraud Detection Dataset (Kaggle)**: A widely used dataset in the machine learning community, this dataset includes anonymized claim details, patient demographics, and labeled instances of fraudulent claims.

3. **Synthetic Datasets**: Synthetic data is generated to simulate real-world claim scenarios, addressing challenges like data scarcity or imbalance. It ensures the model is exposed to a wide variety of fraud cases.

4. **Proprietary Insurance Datasets**: Collected and maintained by insurance companies, these datasets contain historical claims data with annotations for fraudulent and legitimate claims. Access to this data is controlled to comply with privacy regulations such as HIPAA.

5. **Domain-Specific Datasets**: Additional datasets, such as hospital billing records or state-level health insurance claim data, are used to provide a granular view of claim behaviors in specific regions or medical specialties.

The datasets typically include features like claim IDs, provider details, patient demographics, diagnosis and treatment codes, claim amounts, and fraud labels. This diverse data source ensures that the machine learning model is well-trained to detect patterns indicative of fraudulent claims while addressing privacy and ethical considerations.

III. METHODOLOGY

1. Data Collection and Preprocessing

Data Sources: Data is collected from multiple sources, including insurance company databases, publicly available datasets, or synthetic datasets containing relevant medical insurance claim details, such as claim amounts, patient demographics, and provider information.
Data Cleaning: Missing or incomplete data is addressed using imputation techniques, such as mean imputation or interpolation. Outliers are handled through statistical methods like interquartile range (IQR) or Z-score thresholds to prevent extreme values from distorting the model. Consistency in data formats, including date structures and categorical variables, is ensured.
Data Transformation: Categorical variables (e.g., gender, insurance type) are converted into numerical representations using methods like one-hot encoding or label encoding. Numerical features (e.g., claim amounts) are normalized or standardized to improve model convergence and performance.

2. Feature Engineering

Feature Selection: Relevant features for detecting fraud are identified using domain expertise and statistical methods, including correlation analysis, chi-squared tests, and recursive feature elimination. Key features include claim amounts, claim frequency, claimant demographics, policy details, and provider characteristics. **Feature Creation**: Additional features are engineered to detect potential fraud, such as the ratio of claimed amounts to policy limits to identify unusually high claims. Another feature could represent the time interval between successive claims, highlighting patterns that might indicate fraudulent activities.

3. Model Development

SVM Implementation: A Support Vector Machine (SVM) classifier is implemented using libraries like scikit-learn in Python. The SVM model is used for binary classification (fraud vs. non-fraud) or multi-class classification in cases where multiple fraud types are present. **Kernel Selection**: Different kernel functions (e.g., linear, polynomial, radial basis function [RBF]) are tested to determine the best-performing option for the dataset. The choice of kernel depends on the data's characteristics, with RBF being a common choice for non-linear decision boundaries.

4. Hyperparameter Tuning with GridSearchCV

Parameter Grid: A grid of hyperparameters is defined, including regularization parameter (C), kernel coefficient (gamma) for non-linear kernels, and kernel type. Additional parameters like degree (for polynomial kernels) and coef0 (for polynomial and sigmoid kernels) are also explored.
Cross-Validation: GridSearchCV performs an exhaustive search across the parameter grid, using cross-validation to evaluate performance for each combination. The optimal parameters are selected based on cross-validated accuracy to ensure robust generalization to unseen data**.**

5. Model Evaluation

Evaluation Metrics: The optimized SVM model is assessed using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. These metrics provide insights into the model’s ability to identify fraudulent claims while minimizing false positives and negatives. **Confusion Matrix**: A confusion matrix is generated to evaluate the model’s performance in identifying true positives, false positives, true negatives, and false negatives. This analysis helps determine how effectively the model differentiates between legitimate and fraudulent claims.

6. Implementation and Deployment

User Interface: A user-friendly interface is designed for insurance adjusters and other stakeholders. The interface allows users to input claim data, receive fraud predictions, and view detailed analyses, such as the likelihood of fraud and key contributing features. This integration streamlines workflows and enhances fraud detection efficiency

Deployment: The system is deployed on a scalable platform capable of processing high volumes of claim data. Security measures are implemented to protect sensitive medical and financial information, and the system is designed for easy maintenance and updates as new data and fraud patterns emerge.

TESTING AND VALIDATION:

Testing and validation are essential to ensuring the accuracy, reliability, and robustness of the fraud detection system. The project adopts a comprehensive multi-step evaluation strategy:

1. **Data Splitting:** The dataset is partitioned into training, validation, and testing subsets, typically in a 70-20-10 ratio. This splitting strategy helps prevent overfitting and ensures the model's ability to generalize effectively to new, unseen data.

2. **Performance Metrics:** The system's performance is assessed using evaluation metrics such as accuracy, precision, recall, F1-score, and the area under the Receiver Operating Characteristic curve (AUC-ROC). These metrics provide a comprehensive evaluation of the model's capability to identify fraudulent claims accurately and minimize errors.

3. **Cross-Validation:** To ensure consistent performance across various data samples, k- fold cross-validation is employed. This technique divides the dataset into multiple folds, iteratively training and validating the model on different subsets to evaluate its stability and robustness.

4. **Error Analysis:** Misclassified instances are thoroughly examined to pinpoint the model's limitations, such as its inability to detect specific types of fraud. This detailed analysis guides further refinements to improve the system's ability to handle diverse fraud scenarios.

5. **Comparison:** The results of the proposed model are benchmarked against traditional fraud detection approaches and other machine learning algorithms. This comparative analysis highlights the advantages, improvements, and unique contributions of the developed system in addressing fraud detection challenges effectively.

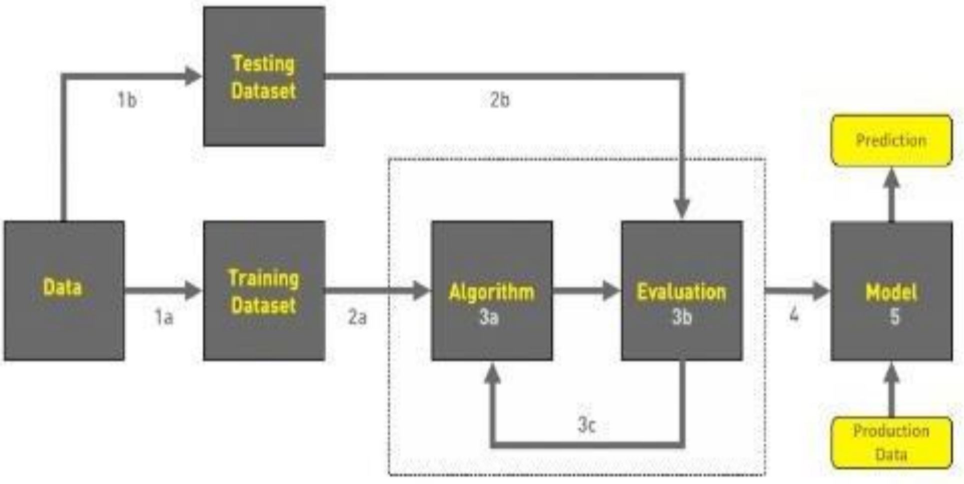

Fig:System Architecture

FUTURE SCOPE

The project lays a strong foundation for advancing fraud detection in medical insurance claims. Future improvements include:

1. **Integration with Real-Time Systems**: Incorporating the model into real-time claim processing workflows for immediate fraud detection.

2. **Advanced ML Techniques**: Exploring deep learning techniques like autoencoders, graph neural networks, and ensemble models for improved accuracy.

3. **Explainability and Interpretability**: Developing explainable AI (XAI) frameworks to interpret model predictions for regulatory and operational transparency.

4. **Scalability**: Enhancing the model to handle large-scale data with efficient computational resource utilization.

5. **Blockchain Integration**: Implementing blockchain-based secure systems to maintain transparent and tamper-proof records.

6. **Broader Fraud Scenarios**: Expanding the system to include other types of insurance fraud, such as auto and property claims.

IV. DISCUSSION AND RESULT

Fraud detection in medical insurance claim systems leveraging **machine learning (ML)** has witnessed remarkable progress, with numerous studies showcasing promising results. Various ML techniques, including **Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM)**, and **XGBoost**, have been extensively applied to identify fraudulent claims. For instance, **XGBoost** has demonstrated significant efficacy, achieving a 7% improvement in accuracy over traditional decision tree models. This underscores the capability of gradient boosting algorithms to detect intricate patterns in large and complex datasets. Similarly, Random Forest and Naïve Bayes have been employed in different studies, with Random Forest frequently outperforming other methods in terms of accuracy for fraud detection tasks.

The findings from these studies suggest that ML algorithms can effectively process large volumes of historical claims data, extracting meaningful features to differentiate legitimate claims from fraudulent ones. For instance, **Naïve Bayes** has proven particularly effective in detecting fraudulent claims associated with premium-related data, while **Support Vector Machines** have been shown to excel at categorizing fraudulent patterns in claims datasets. Additionally, the integration of ML with **blockchain technology** has bolstered fraud detection systems by enhancing their security and resilience. This combination provides a decentralized and tamper-proof platform for managing sensitive medical data, thereby ensuring greater transparency and system integrity.

Despite these advancements, challenges persist in implementing these techniques. Data quality remains a critical factor, as incomplete or redundant data can significantly impair the performance of ML models. To address this, data preprocessing and feature engineering are employed to improve data quality before model training. Furthermore, while approaches such as sequence mining and analysis of **patient time-series data** have yielded promising outcomes in detecting fraudulent claims, the unstructured nature and complexity of healthcare data continue to pose challenges. Limitations such as restricted access to real-world datasets and concerns about privacy also hinder the effectiveness of fraud detection models.

Nevertheless, the application of ML in fraud detection within medical insurance claims demonstrates significant potential. These technologies can identify suspicious patterns that might otherwise elude traditional manual review methods. Ongoing advancements in **deep learning, neural networks**, and **blockchain-based frameworks** are expected to enhance fraud detection capabilities further. These developments aim to enable real-time, accurate identification of fraudulent activities while improving the scalability and efficiency of such systems. Future research is likely to emphasize enhancing model interpretability, addressing data privacy challenges, and developing hybrid approaches that integrate multiple ML algorithms with blockchain to create more robust fraud detection solutions.

V. CONCLUSION

In conclusion, addressing healthcare insurance fraud is critical for safeguarding financial resources and ensuring the proper allocation of healthcare services. The current application of **Support Vector Machine (SVM)** coupled with **Grid Search Cross-Validation (CV)** has demonstrated strong potential in detecting fraudulent activities within the system. By fine- tuning model parameters through cross-validation, SVM exhibits significant capability in identifying intricate fraud patterns and differentiating them from legitimate claims. This approach provides a robust foundation for building more sophisticated fraud detection systems in the future.

Looking forward, the integration of advanced **machine learning (ML)** techniques, such as deep learning, ensemble methods, and reinforcement learning, has the potential to enhance the effectiveness of fraud detection frameworks. These advanced algorithms can leverage historical data to improve predictive accuracy, making it easier to identify evolving fraud schemes. Additionally, the inclusion of diverse data sources, such as **electronic health records (EHRs),** patient behavior analytics, and transaction histories, can enable a more comprehensive and nuanced analysis of claims data, further improving detection capabilities.

The proposed system would greatly benefit from the development of a **detailed fraud taxonomy** that classifies various types of fraudulent activities, such as upcoding, phantom billing, and identity theft. By categorizing fraud into specific types, the system can adopt tailored detection approaches for each scenario, thereby enhancing both accuracy and efficiency. Moreover, employing a diverse range of **ML techniques**, including neural networks, Random Forests, and **XGBoost,** can facilitate a comparative analysis of model performance for different types of fraud, ensuring the use of the most suitable algorithm for each detection task.

Furthermore, integrating **blockchain technology** into the system can add a critical layer of security, safeguarding the integrity and authenticity of claims data. Blockchain's immutable and decentralized ledger system can ensure transparency and protect sensitive information from unauthorized modifications or tampering. This integration will also foster greater trust among stakeholders by enhancing accountability and reducing opportunities for fraudulent manipulation of data.

By combining the predictive strength of machine learning with the security features of **blockchain technology**, this hybrid approach can effectively mitigate financial losses due to fraud while simultaneously reinforcing the credibility and transparency of the healthcare insurance system. This automated, secure, and efficient system has the potential to significantly reduce administrative overhead, improve operational workflows, and ensure that healthcare resources are directed towards legitimate claims. These advancements will not only benefit insurers and healthcare providers but also improve outcomes for patients and regulators. Ultimately, this comprehensive framework will enhance the sustainability of healthcare services while fostering trust and efficiency across the ecosystem.

VI. REFERENCES

[1] N. Dhieb, H. Ghazzai, H. Besbes, and Y. Massoud, "A Secure AI-Driven Architecture for Automated Insurance Systems: Fraud Detection and Risk Measurement," *IEEE Access*, vol. 8, pp. 58546–58558, 2020. doi: 10.1109/ACCESS.2020.2983300.

[2] S. Tanwar, Q. Bhatia, P. Patel, A. Kumari, P. K. Singh, and W. C. Hong, "Machine Learning Adoption in Blockchain-Based Smart Applications: The Challenges, and a Way Forward," *IEEE Access*, vol. 8, pp. 474–448, 2020. doi: 10.1109/ACCESS.2019.2961372.

[3] M. Bärtl and S. Krummaker, "Prediction of Claims in Export Credit Finance: A Comparison of Four Machine Learning Techniques," *Risks*, vol. 8, no. 1, 2020. doi: 10.3390/risks8010022.

[4] L. Ismail and S. Zeadally, "Healthcare Insurance Frauds: Taxonomy and Blockchain-Based Detection Framework (Block-HI)," *IT Prof.*, vol. 23, no. 4, pp. 36–43, 2021. doi: 10.1109/MITP.2021.3071534.

[5] I. Matloob, S. A. Khan, and H. U. Rahman, "Sequence Mining and Prediction-Based Healthcare Fraud Detection Methodology," *IEEE Access*, vol. 8, pp. 143256–143273, 2020. doi: 10.1109/ACCESS.2020.3013962.

[6] G. Kowshalya and M. Nandhini, "Predicting Fraudulent Claims in Automobile Insurance," *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, no. Icicct, pp. 1338–1343, 2018. doi: 10.1109/ICICCT.2018.8473034.

[7] S. Wang et al., "Blockchain-Powered Parallel Healthcare Systems Based on the ACP Approach," *IEEE Trans. Comput. Soc. Syst.*, vol. 5, no. 4, pp. 942–950, 2018. doi: 10.1109/TCSS.2018.2865526.

[8] S. Chakraborty, S. Aich, and H. C. Kim, "A Secure Healthcare System Design Framework Using Blockchain Technology," *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2019-February, pp. 260–264, 2019. doi: 10.23919/ICACT.2019.8701983.

[9] T. T. A. Dinh, R. Liu, M. Zhang, G. Chen, B. C. Ooi, and J. Wang, "Untangling Blockchain: A Data Processing View of Blockchain Systems," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 7, pp. 1366–1385, 2018. doi: 10.1109/TKDE.2017.2781227.

[10] W. Kozlow, M. J. Demeure, L. M. Welniak, and J. L. Shaker, "Acute Extracapsular Parathyroid Hemorrhage: Case Report and Review of the Literature," *Endocr. Pract.*, vol. 7, no. 1, pp. 32–36, 2001. doi: 10.4158/ep.7.1.32.

[11] M. Raikwar, S. Mazumdar, S. Ruj, S. Sen Gupta, A. Chattopadhyay, and K. Lam, "2018 9th IFIP International Conference on New Technologies, Mobility and Security, NTMS 2018 - Proceedings," *2018 9th IFIP Int. Conf. New Technol. Mobil. Secur. NTMS 2018 - Proc.*, vol. 2018-January, 2018.

[12] R. Roy and K. T. George, "Detecting Insurance Claims Fraud Using Machine Learning Techniques," *Proc. IEEE Int. Conf. Circuit, Power Comput. Technol. ICCPCT 2017*, 2017. doi: 10.1109/ICCPCT.2017.8074258.

[13] X. Liang, J. Zhao, S. Shetty, J. Liu, and D. Li, "Integrating Blockchain for Data Sharing and Collaboration in Mobile Healthcare Applications," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2017-October, pp. 1–5, 2018. doi: 10.1109/PIMRC.2017.8292361.

[14] F. Tang, S. Ma, Y. Xiang, and C. Lin, "An Efficient Authentication Scheme for Blockchain- Based Electronic Health Records," *IEEE Access*, vol. 7, pp. 41678–41689, 2019. doi: 10.1109/ACCESS.2019.2904300.