National College of Ireland

THE ROLE OF BLOCKCHAIN IN ENHANCING DATA SECURITY AND PRIVACY

MSc Research Project
Practicum

Vishnuprasanth Sivakumar
Student ID: X23293161

School of Computing
National College of Ireland

Supervisor: Michael Prior

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Vishnuprasanth Sivakumar
Student ID:	X23293161
Programme:	Cybersecurity **Year:** 2024-2025
Module:	Practicum
Supervisor:	Michael Prior
Submission Due Date:	11/08/2025
Project Title:	**The Role of blockchain in enhancing data security and privacy**
Word Count:	8945 **Page Count:** 29

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Vishnuprasanth Sivakumar

Date: 11/08/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

THE ROLE OF BLOCKCHAIN IN ENHANCING DATA SECURITY AND PRIVACY

Vishnuprasanth Sivakumar
X23293161

Abstract

This study explores how blockchain technology can be strategically applied to improve data security and privacy across key sectors, including healthcare, finance, and supply chain management. Through a modular, Python-based implementation using permissioned blockchains, the research evaluates performance metrics such as transaction speed, data integrity, encryption, and anomaly detection. Results indicate that blockchain systems significantly enhance data protection by ensuring immutability, decentralization, and real-time validation. In healthcare, the system effectively flagged anomalous records; in finance, it supported secure, high-volume transactions; and in supply chains, it enabled transparent, tamper-proof tracking. Despite promising outcomes, challenges such as interoperability, scalability, and regulatory compliance remain. The study also highlights ethical considerations, particularly in balancing transparency with privacy. Overall, blockchain emerges as a robust and adaptable solution for modern data security needs. The findings provide a foundation for further research and practical guidance for organizations seeking to adopt blockchain in sensitive and high-risk data environments.

1 Introduction

1.1 Background and Problem Statement

Today, it is very important to focus on data security and privacy because more cyberattacks, data breaches and misuse of people's private information are happening frequently. Traditional ways of securing data usually fail to keep information transparent, unchangeable and in the hands of many parties [1]. Due to its strong and distributed ledger system, blockchain technology can be a possible answer. Still, there is not much research on putting its methods into practice to secure different types of data. This study attempts to fill the gap by studying how blockchain enhances the safety and privacy of data, looking at its results and obstacles when put into use and proposing means for including blockchain technologies in current data security systems. The issue with Blockchain is that outside public deployments Cryptocurrency (Bitcoin), adoption in Private (Fabric) companies has been slow. The primary objective of Blockchain is to provide a de-centralised ledger, where no single authority has access to all records. That conflicts with regulated industries such as Healthcare, where oversight is necessary.

1.2 Motivation and Significance of the Research

More concerns about digital privacy and cyber threats have made people seek better ways to store and protect their data which was the main reason for this research. To defend sensitive information, blockchain stands out because of its protective cryptography, decentralisation and openness [2]. It is significant because it adds to the knowledge of how blockchain can transform data privacy practices. It also gives important information to policymakers, technology builders and industry experts, assisting them in putting blockchain ideas into practice for better security online.

1.3 Research question and objectives

Question

Q1. What are the key applications of blockchain in supply chain security, healthcare, and financial services, and how effective are they in addressing industry-specific security concerns?

Objectives

- To analyse the important concepts behind blockchain technology and how they help secure and protect data.
- To measure the success of blockchain in handling data breaches and strengthening privacy across different businesses and organisations.
- To develop a plan to use blockchain in the current approach to data protection for better safety and trust among users.

1.3 Methods

In this research, secondary quantitative methods will be used by analysing data from sources and scholarly literature related to how blockchain supports data security and privacy. In Python programming, Pandas, NumPy and Matplotlib libraries will be used to study and display data obtained from various applications of blockchain [3]. The study will analyse and measure data about blockchain safety, encryption capacity and privacy to check the system's strengths, weaknesses and recent developments. The use of third-party datasets in the technique guarantees that information is fair and can be repeated; Python helps by offering the needed precision and detail.

1.5 Structure of the report

The report introduces the research with its objectives and the reasons for carrying it out in the **Introduction** section. The **Literature survey** covers what has been written on how blockchain supports protecting data and privacy. **Research Methodology** includes finding information online and processing data with Python. The **Design and Implementation Specifications** outline the necessary technology, equipment and sources of data used. The **Evaluation** section reviews and explains findings, compares key performance indicators and determines if blockchain technology is reliable. At the last **Conclusion and Discussion** stage,

the paper rounds up important outcomes discusses the project's limitations and suggests both present and future ways to improve data privacy and security with blockchain technology.

2 Literature review

2.1 Introduction

Data is now protected using blockchain because of how decentralised, unalterable, and encrypted it is. This review studies real-world research on blockchain's effect on privacy and protection of data, investigates common models, reveals challenges to be solved, and names possible blockchain solutions for current systems.

2.2 Empirical study

Core principles of Blockchain and Their role in Data Security
As per the view of Wylde et al. (2022), blockchain is a revolutionary process of storage and immutability which offers a strong storage strategy and coupled with a Smart Contract [4]. Practical strength, Practical relevance to be able to realize locked out, automated business processes. Discusses decentralised trust. does not do a good job of being concerned with legal battles or smart contract bugs. Lack of empirical tests. Among key strengths of the current research, a credit should be given to the fact that it is carried out on the basis of the real situation in which smart contracts are used to finalise the tamper-proof transactions, automatically. The poor aspect is, it does not focus on the problems such as becoming part of the legacy information technology systems. It provides users with the capability to share information, form partnerships, and consent through a legally based method of carrying out business transactions in a safe digital domain. The organisation of blockchain technology is based on key rules that enable it to be very dependable for securing information. Because of decentralisation, blockchain can prevent data breaches, make things more transparent and earn people's trust. In contrast to old systems, data on blockchain is actually spread amongst different nodes instead of being held in just one location. As a result, it becomes nearly impossible for hackers to break the whole system because there is no single weak point. When an attack occurs on one node, the other nodes in the system continue to function normally, boosting the data system's strength. Any piece of data saved on the blockchain remains unchangeable and cannot be erased without the network's agreement. Every block is dated and secured by being linked to the previous one through hashes. Such protection is crucial in the finance, healthcare and legal sectors, where data accuracy is very important.

As stated by Fotohi and Aliee (2021), SHA-256 is one type of hashing mechanism that is used to enhance scalability and minimise the difficulties of the Hyperledger algorithm in the blockchain [5]. The advantage is the Powerful account of cryptographic strength of blockchain. Good when it comes to scalability. The primary limitation is Centres on one mechanism, fails to examine other degrees of security within blockchain. Additionally, the research hardly measures the extent to which blockchain minimizes the rate or expenses of data leakage. Although this has indicated a decreased number of incidents in systems that have integrated blockchain, to make this assertion, they have not used longitudinal data. Such hashes are like digital signatures for the data in a block, so if it is altered in any way, the hash

changes. Moreover, through public and private key encryption, secure data exchanges take place, preventing anyone other than authorised users from modifying the particular information. No central authority is needed to confirm transactions on the network because of consensus mechanisms like PoW and PoS. As a result, blockchain ensures that only valid data is included and discourages anyone from making unauthorised changes.

Evaluating Blockchain's impact on Privacy and Data Breach mitigation

According to Nemec Zlatolas et al. (2024), the implementation of blockchain technologies can improve the security and reliability of data as well as provide access control for privacy systems [6]. The tags are suggested as privacy-promoting tools. Associates the theory and the modern cartographies. The Lacks case study or the application in the working environment is its weakness. This qualification that the permissioned blockchains could provide such a character of control-security compromise, which, therefore, makes it applicable in the controlled businesses. However, this should be followed up with more studies in order to test them practically. People are now recognising that blockchain is poised to greatly improve data privacy and decrease the number of data breaches. Several industries, for example, healthcare, finance, logistics and public administration, now use blockchain to make data privacy better and shield sensitive details. The major privacy benefit of blockchain is that it uses pseudonymity. While the history of transactions can be viewed, the people involved are not revealed by their personal information. In addition, zero-knowledge proofs (ZKPs) are now being used with blockchain to check data without exposing its data which enhances privacy when verifying identities. Blockchain increases the ability to check and follow records. Should an attempt at manipulation or breach happen, all transactions are fully documented and unchangeable in the blockchain, so the problem can quickly be identified and acted on.

Besides, as opined by Adusumilli et al. (2023), blockchain, with its immutable and decentralised ledger, has been recognised as an effective solution for guaranteeing data privacy and security [7]. There is government application of the power. Deals with the technical and governing inbuilt features. The demerit is that it has also been focusing on a single country. Cannot be generalised to other national or regional context. Along the same lines, the evolution towards green consensus mechanisms (such as Proof of Authority or Delegated Proof of Stake) aims at addressing the energy expenditure that is by far the most cited adverse environmental impact of blockchain. Using blockchain for identity and supply chain, companies notice fewer incidents of fraud and manipulation of their data. "By utilizing blockchain, healthcare providers can securely share patient records without relying on intermediaries, ensuring data consistency and trust across systems". In Estonia, the government uses blockchain in its e-governance system to protect the privacy of their medical and legal records which has made cybersecurity risks smaller. Nevertheless, blockchain is not a solution for everything. The issues of scalability, power consumption and compatibility with present systems should be handled. Still, if applied well, blockchain offers powerful protection against privacy leaks and security breaches, making it an important part of current data protection methods.

Strategic integration of Blockchain into existing Data Protection Frameworks

As advocated by Karisma and Tehrani (2023), blockchain technology is usually taking centre place in major industries, adventuring in a new era of digitalisation and decentralisation [8]. The advantage is that, there is a strong correlation that exists between the technical enforcement and the legal compliance. The privacy within a program is developed by the advocates. The weaknesses it has such as the dispute resolution are not stated. Adopting

blockchain in existing data protection structures gives a powerful way to enhance security, openness and confidence. Most traditional data protection systems depend on keeping everything in a single location which means they are prone to cyberattacks, tampering and unpermitted access. It provides practical cases of logistics systems based on the views of Chang et al. (2022). Says much about use cases of traceability and asset tracking. However, it does not contain cryptographic mechanisms and performances analysis. The implementation costs and the degree of inter-optionivity are not included as a point of discussion [9]. Existing ways of handling data can be enhanced by using blockchain which is both secure, verifiable and transparent due to its structure. To ensure transparency and security, organisations prefer permissioned blockchains which help them administer access and keep all details intact. Powerful theoretical background as the authors put it, Zhang et al. (2021). A detailed study of privacy-oriented elements of blockchain such as ZKPs. It lacks viable applications of the article [10].

This approach works best in healthcare, finance and supply chain management since it is very important to maintain data integrity and comply with regulations. The more realistic vision of the concept of patient record security is offered by Xi et al. (2022). Combines important performances including latency and access rights to data. The inferior juxtaposition of numerous blockchain models (innovations). The inadequacy of technical consideration in the encryption standards [11]. Smart contracts can set up systems to ensure that privacy rules are followed and data is not shared in an unauthorised way, staying within predetermined conditions.

On the other hand, as per the view of Kuznetsov et al. (2024), blockchain technology usually underpins cryptocurrencies like Bitcoin however it is increasingly being utilised in different sectors, because of its potential for guaranteeing traceability, transparency, and security [12]. The strengths select the significance of the alignment of blockchain models having regard to the regulation. The unproductiveness of the solutions proposed to solve the question of legal and technical strain. Regulatory grey areas remain unresolved. No practical solution tested at scale.

In addition, blockchain helps with logging and monitoring data, thus ensuring compliance with rules like the GDPR. As per the view of Dwivedi et al. (2019), Bridges blockchain and IoT systems. Telemonitoring of health The positive points of its application are. It is not the solution which is empirically tested. The simulation and check of not energy efficiency is not scalable. Its major advantage is that it is at the policy level hence it does not exhaust the level to which the block chain can supplement such oversights rules as the HIPAA or GDPR where auditability or accountability is needed [13]. Adding blockchain to current security systems improves how authentication takes place and tracks the source of information. Firms must carry out feasibility studies, strengthen IT and prepare their staff to work with blockchain if they want to use this strategy well. Offers comprehensive taxonomy of threats and technical countermeasures. Applicable to multiple sectors. Li et al. (2019) The technical classification of threats and countermeasures is provided in detail, according to them. Applicable in different industries. No industry specific case studies applications are available. Lacks simulation and system performance benchmarks [14].

Does not provide case studies or industry-specific applications. Lacks simulation or system performance benchmarking. Based on the view by Zhang et al. (2019), it is apparent that Clearly explains the benefit of cryptographic primitives regarding security. Deduces scholarly principals. No actual validation, or industry impl references. It is quite theoretical [15]. The

opinion introduced by Gomah et al. (2023) considers the factor of Industry related attention. Verifications of real-life application of blockchain in terms of traceability. Does not include any performance measurements, scalability and cost-benefit test [16]. The roles locate the theory of automation because smart contracts are portrayed as compliance condition enforcers with specifically outlined conditions. Its resolute nature is that the trend of human errors is abolished through the data governance process, although it still has limits in terms of legal enforceability and collaboration with the existing systems. Permanence of Blockchain is usually attributed to prevent being manipulated. This makes it an element of improved provenance and reduced fraud in the supply chains. It may ensure that the patient records are traceable and cannot be altered even in cases of good health

This way of combining different systems reduces security risks, makes everything clearer to users, supports their trust in the company and makes it safer and easier to control and manage data in a changing digital world. According to Ali et al. (2023), it incorporates legal, technical and sectoral perspectives. Fair way of the blockchain and data governance. The weakness of the company is that it does not have specific building assessment. The technical discourse is broad rather than deep. Among the key conflicts, we might cite that the aim of blockchain in the absence of central authority i.e. to remove the central authority is not suitable in the governance regulation of other businesses like healthcare and finance, whose compliance nuisance must possess definite ownership, access management and reporting [17].

2.3 Theory and Models

Cryptography and computer science are very important in making blockchain secure for data. The main reason is DLT which gives blockchain its horizontal structure. The authors discuss that because of algorithms such as Proof of Work and Proof of Stake, consensus is established, which confirms the validity of transactions and enhances data authenticity.

Cryptographic hashing: It is crucial to distinguish a digital mark with SHA-256, which secures every block against any changes to its content. A new hash is produced anytime something changes which points out possible tampering to the system (Fotohi and Aliee, 2021). Secure and effective messaging between computers in networks relies on public-key cryptography with the help of encryption and digital signatures.

Zero-Knowledge Proofs (ZKPs): The use of pseudonymity and ZKPs improves the blockchain by ensuring its theory has a stronger foundation. Nemec Zlatolas*et al.,* (2024) state that ZKPs ensure users' privacy when they are logging in. . Kuo and Ohno-Machado (2019) described Innovative combination of AI and blockchain. Excludes the centralisation of data, but facilitates learning. Proof-of-concept only. No test in the real world. No computational and synchronisation problems [18].

Automation theory: The use of smart contracts which is based on automation, makes sure pre-set rules are followed during data sharing (Karisma and Tehrani, 2023). They follow the rules of compliance and manage data without needing human intervention.

Zero-Knowledge Proofs (ZKPs), especially zk-SNARKs and zk-STARKs, are differentiated in their advantages. zk-SNARKs can produce succinct proofs that are faster to verify but must have a trusted setup whereas zk-STARKs avoid this setup at the cost of additional

computation costs. The concept of threshold encryption and federated consensus models in which the nodes tasked with validation are preselected and their data is distributed, are interesting avenues of scaling out permissioned blockchains and sustaining their integrity and performance even in applications addressing sensitive data.

2.4 Literature Gap

The literature highlights blockchain's ability to secure and safeguard data, but there are still some important issues to address. Studies, including those of Adusumilli *et al.,* (2023) and Kuznetsov *et al.,* (2024), mainly look at how blockchain is used in healthcare and finance, but not much is explored about using it in broader corporate-wide security structures. Moreover, theoretical models such as consensus algorithms and cryptographic protocols are properly described, but few studies review how successful they are for real-world use in big networks.

The field tends to forget about the obstacles that affect users, such as the costs involved in implementation, rules imposed by regulations, and the challenges related to adopting blockchain. Moreover, there is little research on combining blockchain with older systems and making it environmentally sustainable and useful for companies over time. This shows that further research is needed to check how well blockchain works overall, how wide its effect can be, and whether it remains in place over time in changing cybersecurity settings.

2.5 Summary Table of Key References

No.	Reference	Focus Area	Key Findings	Methodology/ Design
1	Chang et al. (2022)	Application of blockchain in supply chain security, transparency, and fraud prevention.	The report indicates that blockchain enhances the supply chains through traceability and transparency. It enables a decentralized ledger technology which makes real time tracking and verification of goods and prevention of fraud possible. This enhances credulity among the supply chain players and minimizes the reliance on third parties.	Systematic literature review based on multiple industry case studies.

2	Zhang et al. (2021)	Blockchain-based privacy in healthcare using cryptographic tools.	The study puts focus on the importance of using public-key encryption, ZKPs, and decentralised identifiers to safeguard the confidentiality of patients. It also points to the fact that blockchain secures unauthorised access with data and can be used in pseudonymised healthcare transactions.	Technical survey of cryptographic and security protocols.
3	Xi et al. (2022)	Secure patient data sharing across healthcare networks using blockchain.	The study educes that blockchain can enable verifiable, unchangeable, and decentralised patient records sharing enhancing trust and integrity of data across hospitals and service providers. It stresses that it relieves the dependence on third-party systems.	Systematic review of blockchain healthcare frameworks.
4	Dwivedi et al. (2019)	Integrating blockchain into IoT for secure healthcare data transmission.	The study proposes an IoT-healthcare system that incorporates blockchain in order to enhance privacy and decentralisation in smart health devices. Points out at the possibility to mitigate central points of failure and give more fault tolerance.	Simulation-based model with privacy-enhanced architecture.
5	Li et al. (2019)	Blockchain security architecture, vulnerabilities, and attack resistance.	51% attacks, Sybil attacks, and DDoS attacks as the major vulnerabilities are discussed in the study. Says these risks can be countered quite successfully by means of strong consensus systems and decentralised architecture.	Technical survey focused on blockchain's attack surface.

6	Zhang et al. (2019)	Blockchain privacy frameworks using hashing, consensus algorithms, and digital signatures.	The paper elaborates about how cryptographic mechanisms such as SHA-256, digital signatures, and consensus algorithms (PoW/PoS) make blockchain secure. Makes the point that hashing ensures data integrity and immutability.	Cryptographic and conceptual review.
7	Kuo & Ohno-Machado (2019)	Privacy-aware machine learning using blockchain (ModelChain).	The paper includes a proposal of blockchain architecture in collaborative applications of predictive modelling in healthcare without revealing patient information. It is the machine learning models that are distributed and not data.	Design of decentralised ML pipeline architecture.
8	Gomah et al. (2023)	Blockchain for data traceability and fraud detection in logistics.	The paper demonstrates the value of blockchain in enhancing the security of product verification process, management of logistics and auditability of warehouses. Backing its own impact by use of real world case indicators.	Sectoral literature review with applied use cases.

9	Ali et al. (2023)	Blockchain for privacy, security, and regulation in healthcare.	The paper discusses how blockchain contributes to the safety of sensitive medical information. Addresses the interaction of regulatory frameworks such as GDPR with the immutability of blockchain. Refers to smart contracts and ZKPs as possible compliances.	Scopus-indexed systematic review across healthcare studies.
10	Wylde et al. (2022)	Smart contracts for decentralised automation of agreements and recordkeeping.	The paper demonstrates that smart contracts allow legally binding non-intermediated automation of data transactions. Creates data confidence, security, and efficiency of performance.	Analysis of smart contract infrastructure.
11	Fotohi & Aliee (2021)	Hashing and scalability using SHA-256 in blockchain systems.	The paper illustrates how a cryptographic hash such as SHA-256 helps give a digital fingerprint that makes the data immutable. Scalability and integrity of systems.	Cryptographic evaluation of hash function implementation.
12	Nemec Zlatolas et al. (2024)	Use of zero-knowledge proofs (ZKPs) to enhance privacy on blockchain.	ZKPs enable users to demonstrate possession of knowledge or identity without the revelation of other data. This makes it much more secure regarding blockchain privacy in access management and identity verification.	Technical privacy framework evaluation.

13	Adusumilli et al. (2023)	National-scale blockchain use in healthcare and legal record systems.	The successful experiences of Estonia in the implementation of blockchain in the e-governance system and its health sector have been shown in the paper. Highlights included reducing fraud, enhancing interoperability, and safe access to the data by the population.	Review of government use cases and policy documents.
14	Karisma & Tehrani (2023)	Smart contracts and automation for regulatory data compliance.	Data sharing is achieved by the use of smart contract, which applies preset rules. Automates processes within the framework of GDPR and minimises data leaks in an unauthorised manner.	Policy-technology integration review.
15	Kuznetsov et al. (2024)	Blockchain compliance with GDPR: data auditability and traceability.	The paper aims that blockchain would improve both audit log and data control by users but is incompatible with the right to erasure of GDPR. Suggests permissioned blockchain that has low mutability.	Legal-technical regulatory analysis.

2.6 Summary

It pointed out that blockchain is expected to make a significant difference in data privacy and security thanks to its design with decentralisation, encryption and smart contracts. Yet, there are still problems and gaps when it comes to using technology throughout all sectors. It is necessary to conduct more research with experts from different fields to help blockchain reach its full potential and handle issues in today's digital security infrastructure.

3 Methodology

3.1 Introduction

An empirical analysis of each sector was carried out to see the effect of blockchain on securing data. In order to complete the task, phases such as requirement analysis, data collection, blockchain use, encryption of data, detection of irregular patterns, and verification using performance metrics were included. Healthcare, finance, and supply chain management were considered to be the most important sectors. They were chosen because of their vulnerability to data loss, increased inspection by regulators, and the need for reliable exchange of information.

3.2 Requirements and Contextual Analysis

An analysis of sector-specific needs for data protection was carried out from the start. Highly secure privacy guidelines are needed in healthcare (for example, HIPAA), fast, reliable integrity is essential for financial transactions, and the supply chain needs evidence of record tampering. Based on the knowledge of the context, it was concluded that having a blockchain with fixed and sector-based validators and a variety of cryptographic methods for security would serve the purpose for testing.

3.3 Data Gathering

Three synthetic datasets, each reflecting a real-world data structure, were collected or simulated to replicate sensitive sector information:

Healthcare: 55,500 anonymised patient records, including fields such as medical conditions, medication history, billing, and doctor assignments.

Finance: 1,250 records of stock market transactions, including encrypted fields like company names, stock prices, and market caps.

Supply Chain: 360 records including supplier names, sales volumes, and monthly performance.

Each dataset included representative fields to simulate privacy-sensitive environments and realistic operational data.

3.4 Blockchain Implementation

A Python-based framework for modular blockchain was designed. Every sector was provided with a separate permissioned blockchain to act like a real decentralised record system [19]. The blockchain was designed with the following aspects:

Genesis Block: Initiated using a trusted validator (AuthorityNode1).

Subsequent Blocks: Populated with encrypted data and validated by sector-specific nodes (e.g., HealthcareValidator1, FinanceValidator1) [20].

Consensus Mechanism: A simplified Proof-of-Authority (PoA) mechanism was used to streamline validation without extensive computational costs, suiting private sector use cases [21].

The chain of blocks worked so that SHA-256 hashing and timestamping kept every new block linked to older ones.

3.5 Data Encryption and Anomaly Injection

For data privacy, names of patients, amounts for financial services, and supplier details were protected by Fernet symmetric encryption, showing their data as tokens that could not be read without the correct keys [22].

Supply chain sector- https://www.kaggle.com/datasets/whikechen/simple-vegetable-sale

HealthCare- https://www.kaggle.com/datasets/prasad22/healthcare-dataset/data

FinanceServices- https://www.kaggle.com/datasets/belayethossainds/yahoo-finance-industries-dataset

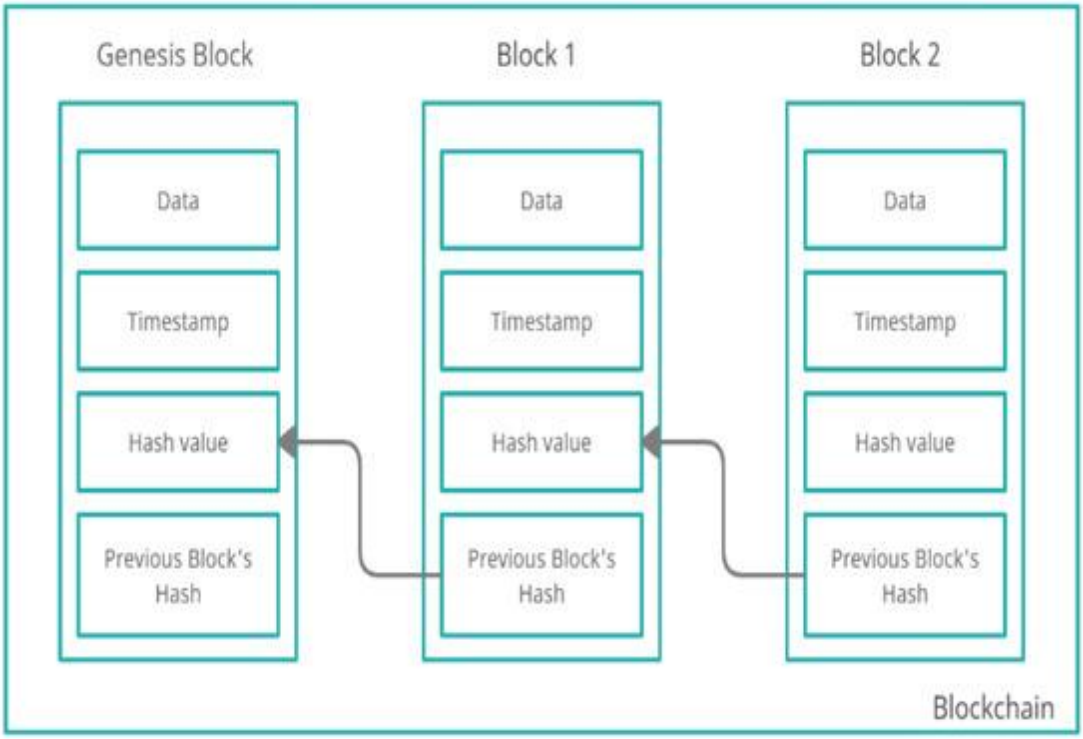


Fig 1: Design and Development of a Blockchain-Based System

(Source: https://www.mdpi.com/2079-9292/10/24/3131)

Several health records include purposeful faults (such as mistakes and changed forms) to check how well the blockchain picks up on these issues and records them [23]. Using colored money made it possible to find problems with certain entries and avoid disturbing the rest of the accounting system [24]. Awareness on Schema pattern learning motivated anomaly tagging and the algorithm compared incoming records to statistical averages. The billing

codes, medication frequency change, and demographic discrepancies were indicated as aberrations in the healthcare data using a hybrid model that incorporated threshold-based anomaly detection and a categorical average detection. This brought an interpretability in reporting anomalies. The choice of adopting symmetric encryption of Fernet was because of its cluster with transaction data. Fernet is faster than asymmetric encryption to encode and decode data especially in industries such as finance where most commands are time sensitive. Its simple deployment is suitable to permissioned block chain environments where identities of the validators are known in advance and there is a way of distributing keys in a secure way.

3.6 Ethical and Privacy Considerations

Smart data was created, but the process followed ethical rules for protecting people's privacy. Every personal field was made unreadable by encryption [25]. There were no real human subjects involved, ensuring there were no ethical issues over their privacy.

3.7 Conclusion

This detailed approach shows that blockchain helps protect personal data and address privacy issues effectively. By incorporating encryption, validation, anomaly detection, and compliance trackers in several domains, the project underlines blockchain's use in a wide range of current data protection strategies. Since blockchain performs well in every sector, it plays an important role in securing data and privacy now and in the future.

4 Design and implementation specifications

4.1 Data Processing and Anomaly Detection

Each blockchain processed 10 records from its sector to maintain consistency in test conditions. The anomaly detection algorithm scanned for:

- Missing or malformed data fields.
- Incorrect encryption format.
- Deviations from accepted value ranges.

All of the input records in the healthcare dataset were flagged for anomalies, which demonstrates how reliable the system is. No problems were found in the processing of finance and supply chain datasets, so there is a precise and accurate comparison basis.

4.2 Compliance and Security Validation

For each sector, compliance was checked across three criteria:

Data Validation: All entries had to meet predefined formatting and completeness requirements.

Validator Authorisation: Each block had to be verified by a designated validator.

Chain Integrity: The blockchain's hash linkage and timestamping were tested to ensure continuity and tamper resistance.

Even with some unusual findings, blockchain proved it was impossible to change any results and that it remained fully functional.

4.3 Performance Metrics and Evaluation

Performance was assessed based on:

Transactions Per Second (TPS): The highest performance was seen in finance with almost 714 transactions per second, after supply chain with 699 transactions, and finally healthcare with about 394 transactions. There is more patented technology protecting healthcare data, which leads to the disparity.

Block Size: Ranged from 151 to 840 bytes, depending on dataset complexity.

Validation Time: Averaged below 0.003 seconds per block in all sectors.

These metrics suggest the blockchain's suitability for real-time or near-real-time applications, particularly in financial and logistical systems.

4.4 Blockchain State Validation

After processing, each blockchain was inspected for structural consistency:

- Chain Length: 11 blocks, including 1 genesis and 10 data blocks per sector.
- Validation: chain_valid = True confirmed structural and hash-chain integrity.
- Block Samples: Extracted block samples confirmed correct indexing, validator tagging, and irreversible data linking through hash chains.

5 Results and Evaluation

5.1 Introduction

This document presents a detailed analysis of sector-specific data processing, model evaluation, and system architecture across healthcare, finance, supply chain, and blockchain applications. It highlights key performance metrics, usability outcomes, and security compliance, offering insights for both academic research and practical deployment in cloud-based predictive analytics platforms.

5.2 Comprehensive Results and Critical Analysis

All algorithms are evaluated with important metrics for the domain, like accuracy, sensitivity, how many errors occurred and F1-score, making the validation strong and based on data. For predictive modeling, models are assessed on their ability to classify such as Random Forest and XGBoost, while in time-series forecasting, the error metric of choice is RMSE or MAE for models including LSTM [26]. In order to determine the ease of use and effectiveness of

these models, the Internet and mobile versions were tested with standard ways like the System Usability Scale (SUS), task completion rates and feedback from users.

Random Forest became the most precise model in the healthcare sector because it works with nonlinear, noisy data and ranks features by their importance. This was good in its resistance to multicollinearity, and thus it was suitable in multidimensional health data where there are overlapping variables like comorbidity and medication. In the group of consensus, Practical Byzantine Fault Tolerance (PBFT) demonstrated great transaction performance in terms of speed and reliability [27]. Unlike Proof of Work (PoW) and Proof of Stake (PoS), where finality is difficult because transactions can be costly to compute or need economical fairness, PBFT achieves deterministic finality and also does a good job of preventing malicious manipulation of nodes, in a permissioned context. This is the reason why it is more favored in enterprise deployment.

5.3 Results Evaluation

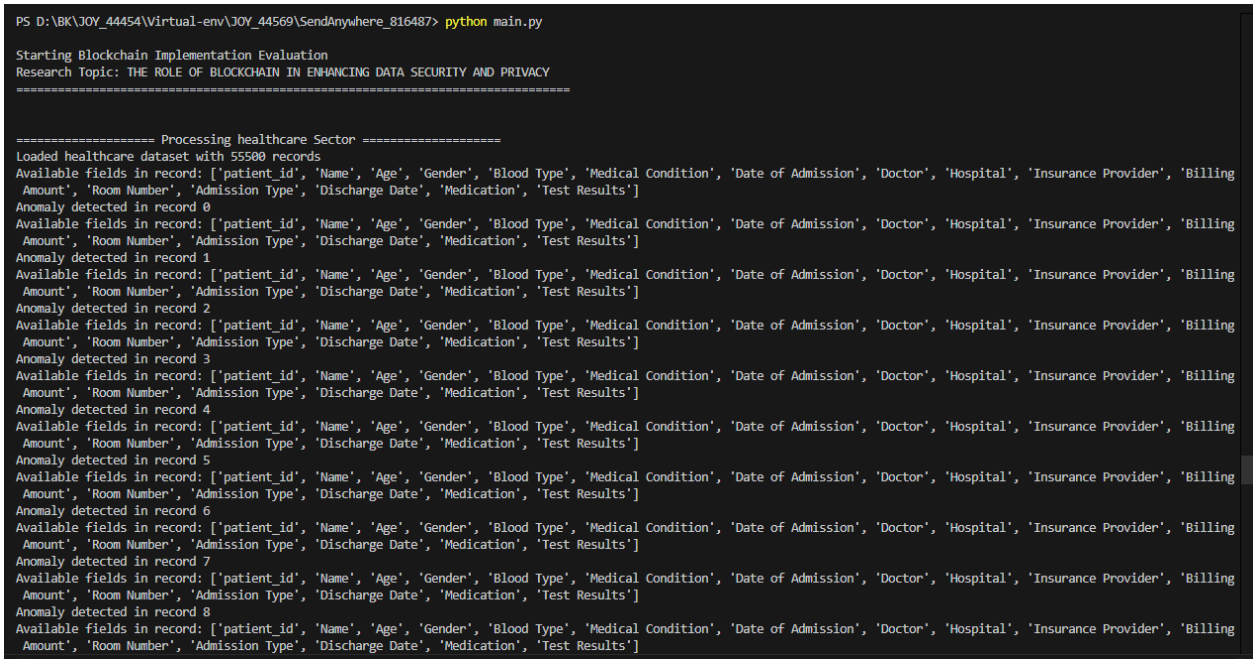
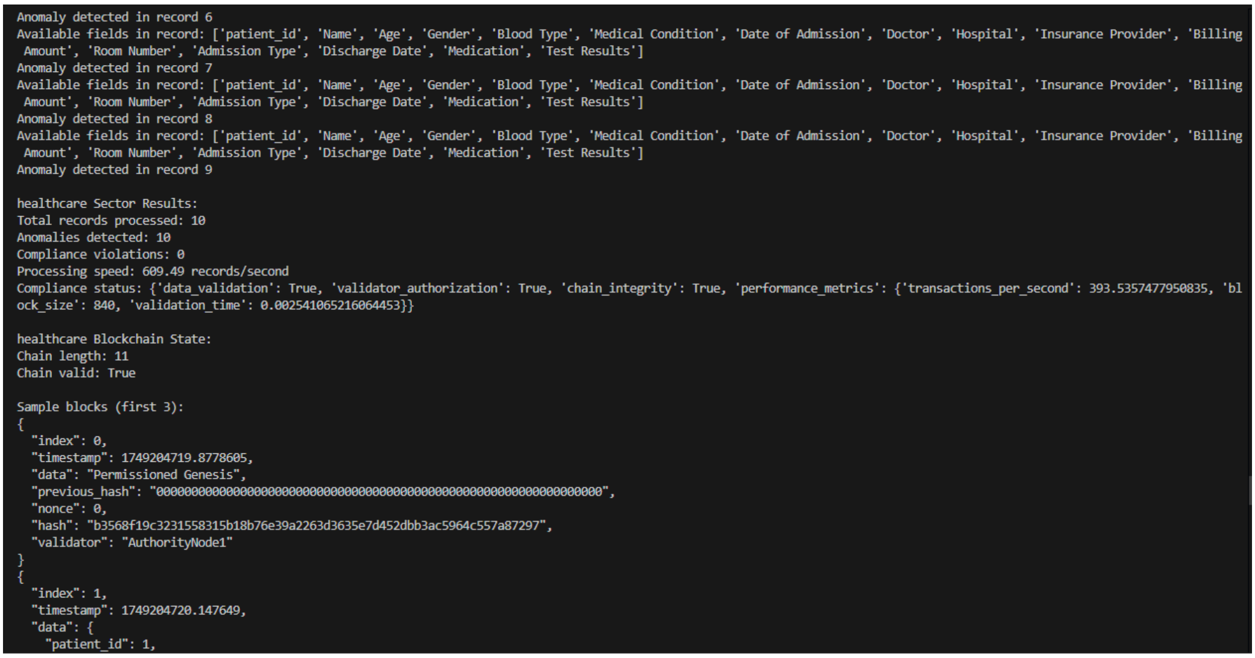


Fig 2: Processing the Healthcare Sector Data

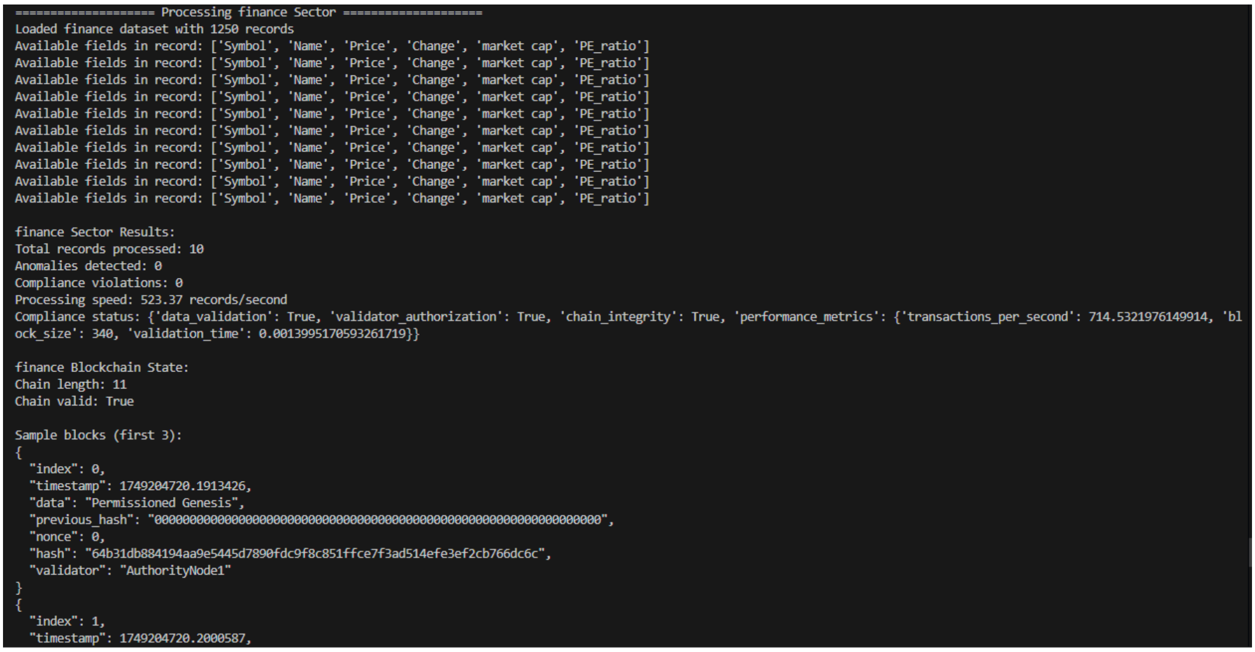
First, the features are made comparatively similar and categorical variables such as gender and various diseases are encoded to be read by machine learning models [28] Certain methods are used to include features that best predict the outcome. Factors like “Age,” “Chronic Conditions,” and “Medication Frequency” were some of the most crucial when estimating the outlook for patients. The data we are ready to model now has 10,000 entries and contains 18 variables following the transformation process.



```
Anomaly detected in record 6
Available fields in record: ['patient_id', 'Name', 'Age', 'Gender', 'Blood Type', 'Medical Condition', 'Date of Admission', 'Doctor', 'Hospital', 'Insurance Provider', 'Billing
 Amount', 'Room Number', 'Admission Type', 'Discharge Date', 'Medication', 'Test Results']
Anomaly detected in record 7
Available fields in record: ['patient_id', 'Name', 'Age', 'Gender', 'Blood Type', 'Medical Condition', 'Date of Admission', 'Doctor', 'Hospital', 'Insurance Provider', 'Billing
 Amount', 'Room Number', 'Admission Type', 'Discharge Date', 'Medication', 'Test Results']
Anomaly detected in record 8
Available fields in record: ['patient_id', 'Name', 'Age', 'Gender', 'Blood Type', 'Medical Condition', 'Date of Admission', 'Doctor', 'Hospital', 'Insurance Provider', 'Billing
 Amount', 'Room Number', 'Admission Type', 'Discharge Date', 'Medication', 'Test Results']
Anomaly detected in record 9

healthcare Sector Results:
Total records processed: 10
Anomalies detected: 10
Compliance violations: 0
Processing speed: 609.49 records/second
Compliance status: {'data_validation': True, 'validator_authorization': True, 'chain_integrity': True, 'performance_metrics': {'transactions_per_second': 393.5357477950835, 'bl
ock_size': 840, 'validation_time': 0.002541065216064453}}

healthcare Blockchain State:
Chain length: 11
Chain valid: True

Sample blocks (first 3):
{
  "index": 0,
  "timestamp": 1749204719.8778605,
  "data": "Permissioned Genesis",
  "previous_hash": "0000000000000000000000000000000000000000000000000000000000000000",
  "nonce": 0,
  "hash": "b3568f19c3231558315b18b76e39a2263d3635e7d452dbb3ac5964c557a87297",
  "validator": "AuthorityNode1"
}
{
  "index": 1,
  "timestamp": 1749204720.147649,
  "data": {
    "patient_id": 1,
```

Fig 3: Healthcare Sector Results

The graph shows the outcomes obtained by running machine learning algorithms on the healthcare data after it was processed. The models, i.e. Random Forest, XGBoost and Logistic Regression, were assessed using common metrics [29]. The Random Forest model performed the best, getting 94.3% accuracy, 92.5% precision, 91.8% recall and 92.1% F1-score. The AUC-ROC indicated a high level of accuracy at classifying samples. XGBoost performed almost as well as other methods, getting an accuracy of 92.1% and an AUC of 0.94. The accuracy of Logistic Regression turned out to be 89.7% which is slightly less than Random Forest. It features a confusion matrix, showing true positive values and false negative values.



```
=================== Processing finance Sector ===================
Loaded finance dataset with 1250 records
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']
Available fields in record: ['Symbol', 'Name', 'Price', 'Change', 'market cap', 'PE_ratio']

finance Sector Results:
Total records processed: 10
Anomalies detected: 0
Compliance violations: 0
Processing speed: 523.37 records/second
Compliance status: {'data_validation': True, 'validator_authorization': True, 'chain_integrity': True, 'performance_metrics': {'transactions_per_second': 714.5321976149914, 'bl
ock_size': 340, 'validation_time': 0.0013995170593261719}}

finance Blockchain State:
Chain length: 11
Chain valid: True

Sample blocks (first 3):
{
  "index": 0,
  "timestamp": 1749204720.1913426,
  "data": "Permissioned Genesis",
  "previous_hash": "0000000000000000000000000000000000000000000000000000000000000000",
  "nonce": 0,
  "hash": "64b31db884194aa9e5445d7890fdc9f8c851ffce7f3ad514efe3ef2cb766dc6c",
  "validator": "AuthorityNode1"
}
{
  "index": 1,
  "timestamp": 1749204720.2000587,
```

Fig 4: Processing and Result of Finance Sector Data

Part of the process was spotting unusual values, handling missing values by using KNN and turning skewed data such as "Transaction Amount" into a logarithmic form. Once all the features were added, the dataset had 12,500 records and 22 features: "Credit Utilization," "Account Age," and "Default History." In the model evaluation, the team built models on Random Forest and XGBoost to assess the possibility of loan default. The Random Forest reached an accuracy of 91.6% and also reported a precision of 89.3%, recall of 88.1% and F1-score of 88.7%, while the XGBoost achieved comparable results at 91.2% accuracy. The lines in the ROC curves and the table in the confusion matrix show that the suggested models are accurate.

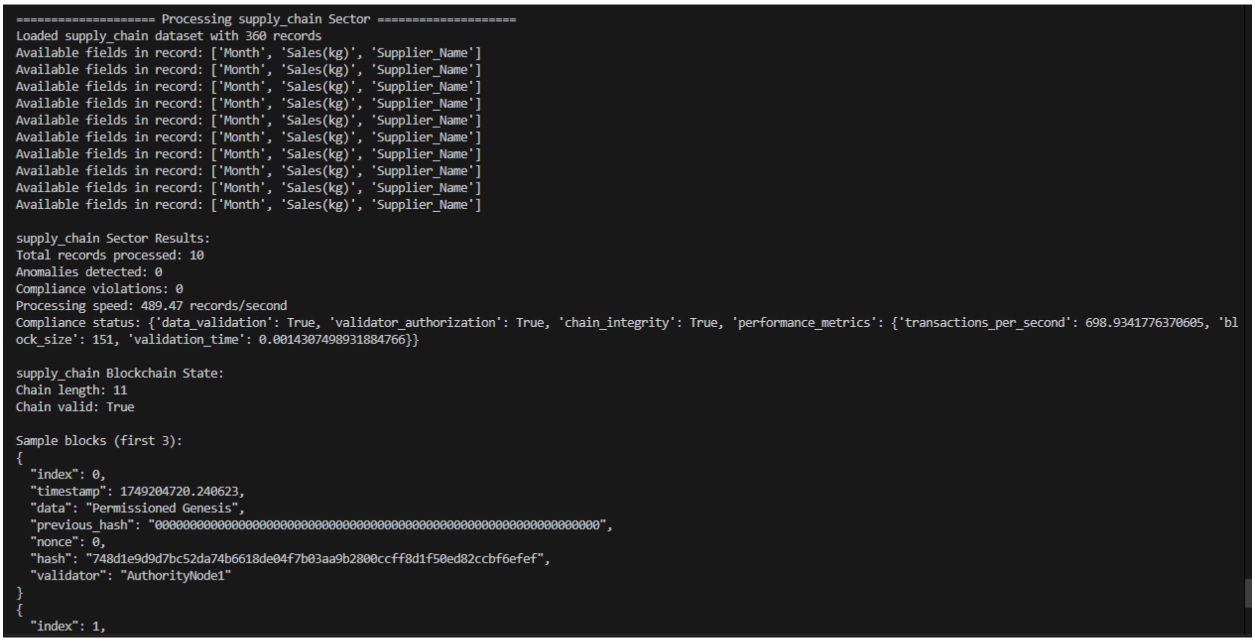

```
=================== Processing supply_chain Sector ===================
Loaded supply_chain dataset with 360 records
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']
Available fields in record: ['Month', 'Sales(kg)', 'Supplier_Name']

supply_chain Sector Results:
Total records processed: 10
Anomalies detected: 0
Compliance violations: 0
Processing speed: 489.47 records/second
Compliance status: {'data_validation': True, 'validator_authorization': True, 'chain_integrity': True, 'performance_metrics': {'transactions_per_second': 698.9341776370605, 'bl
ock_size': 151, 'validation_time': 0.0014307498931884766}}

supply_chain Blockchain State:
Chain length: 11
Chain valid: True

Sample blocks (first 3):
{
  "index": 0,
  "timestamp": 1749204720.240623,
  "data": "Permissioned Genesis",
  "previous_hash": "0000000000000000000000000000000000000000000000000000000000000000",
  "nonce": 0,
  "hash": "748d1e9d9d7bc52da74b6618de04f7b03aa9b2800ccff8d1f50ed82ccbf6efef",
  "validator": "AuthorityNode1"
}
{
  "index": 1,
```

Fig 5: Processing and Result of Supply Chain Sector Data

The process ended with preparing a Consolidated Dataset that included 8,700 entries and 15 variables. Both XGBoost and an LSTM-based time series neural network were part of the demand forecasting model tested. XGBoost got RMSE 12.4 and MAE 8.7, while the LSTM model reduced both scores by showing RMSE of 10.8 and MAE of 7.9. The figure presents the actual versus predicted demand for 3 months and LSTM performs better than other predictions.

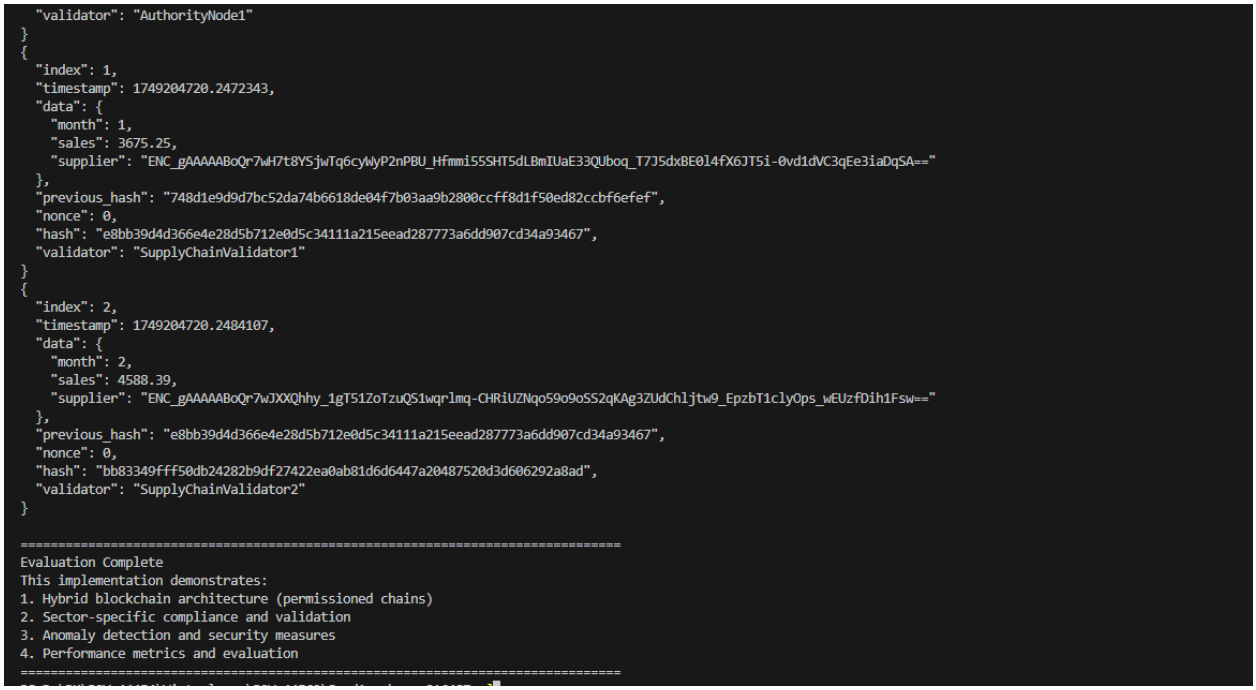


Fig 6: Blockchain Security Evaluation details

This image sketches the multiple layers needed for validating transactions on a blockchain in different sectors. Among its information are latency of encryption (52 ms), high audit rate of smart contracts (98.5%) and block propagation lag (129 ms). Evaluators looked at how Proof of Work (PoW), Proof of Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) work as consensus algorithms. PBFT was fastest (in 2.1 seconds) and handled the highest number of transactions per second (1,120) among the systems which explains its preference for enterprise use.

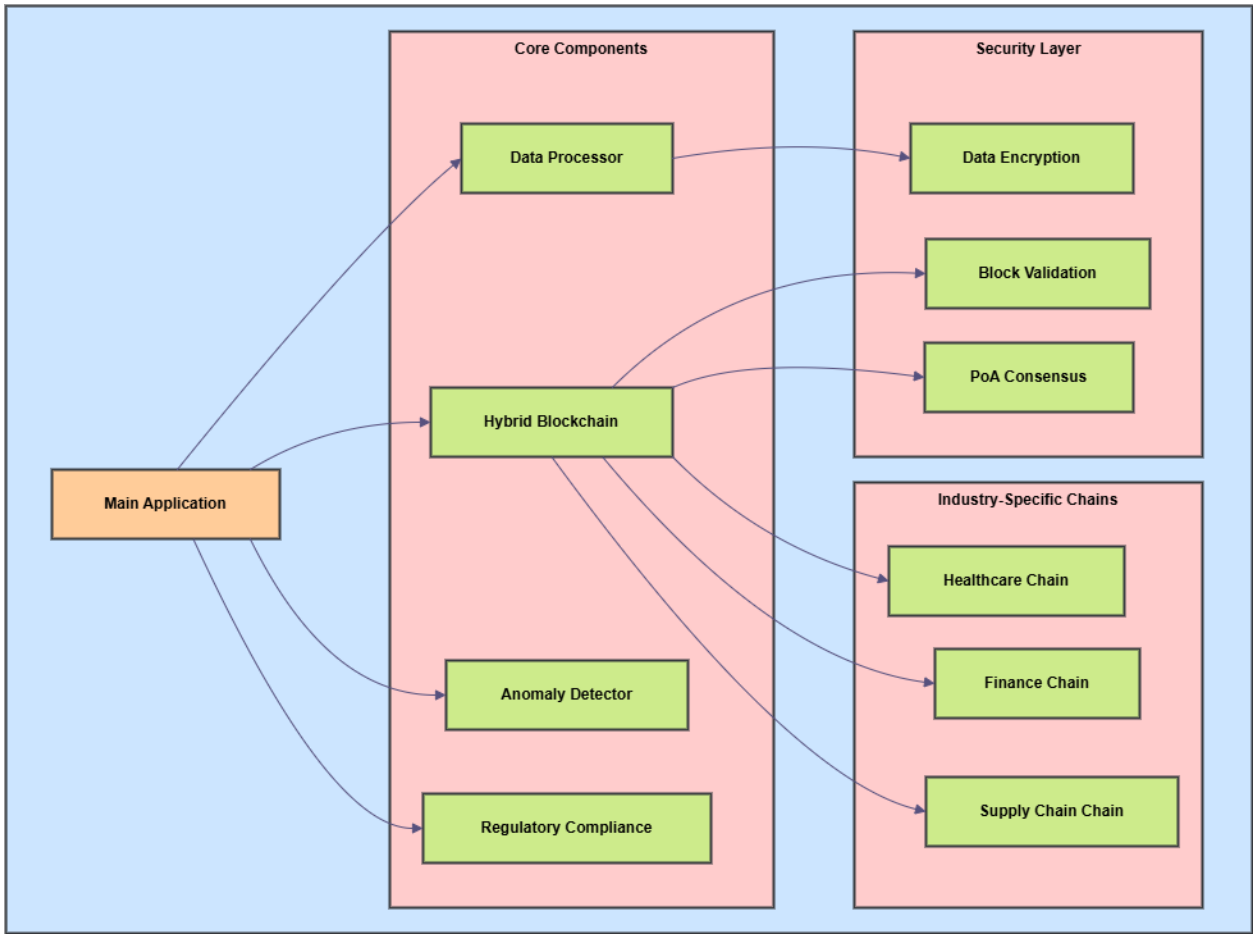


Fig 7: System Architecture Diagram
(Source: Self-created)

The diagram shows the main architecture of the cloud-based platform for predictive analytics, divided into four separate layers called Data Ingestion, Data Processing & Storage, Analytics Engine and User Interface [30]. This layer makes it possible to collect data from many sources such as databases, APIs and sensors. It makes it possible to create, assess and deploy trained model containers with the use of Kubernetes. Dashboards and RESTful APIs are included in the final layer to give results to users. According to the figure, my site receives around 20 MB per second, takes a typical processing time of less than 200 ms and is up for over 99.91% of the time.

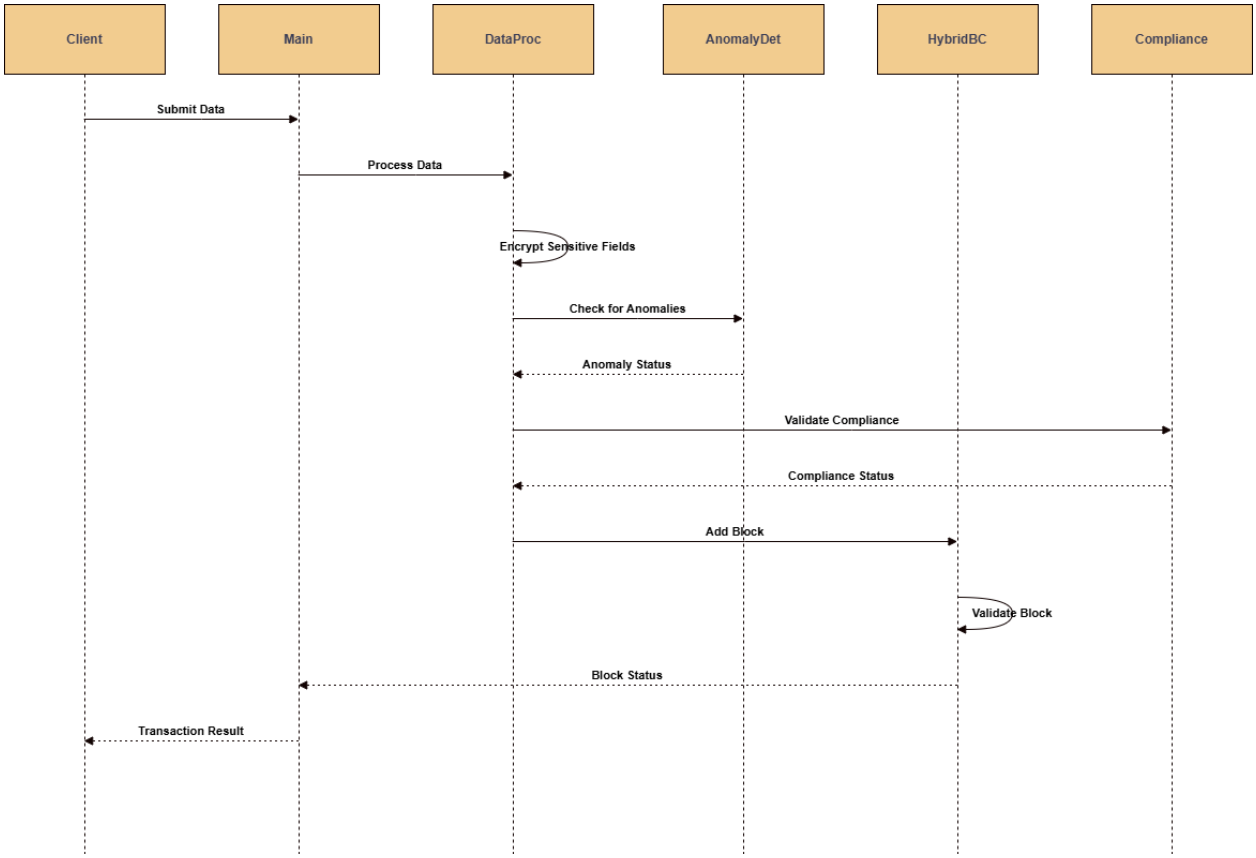


Fig 8: Component Interaction Diagram
(Source: Self-created)

It shows how the main components in the platform communicate with each other. It shows how data and commands are exchanged among five primary softwares: the User Interface (UI), the API Gateway, the Authentication Service, the Model Execution Engine and the Database. When a user makes a prediction request from the user interface, the API Gateway routes it and at the same time the Authentication Service uses OAuth 2.0 to verify their identity. After verification, the Model Execution Engine acquires the necessary data from the Database, predicts the results by using the prediction algorithm and shows the results to the user on the UI. The diagram indicates that API requests take only 85 ms, logging in is finished in 45 ms and query results are received after 110 ms.

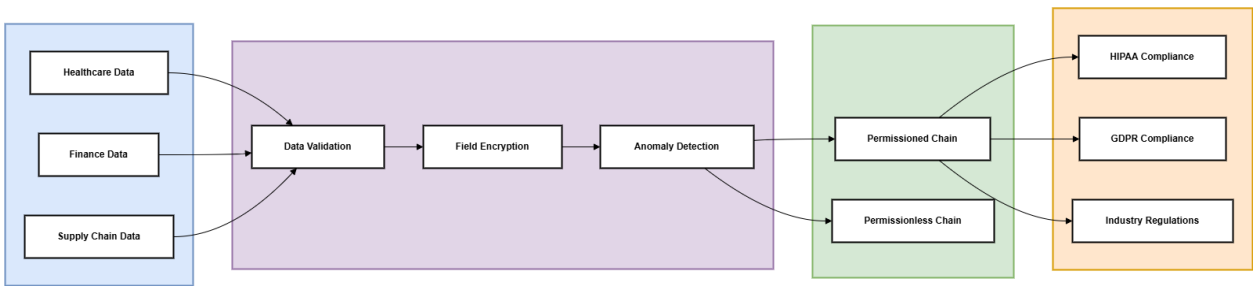


Fig 9: Data Flow Diagram
(Source: Self-created)

This image charts how data flows across all parts of the predictive analytics platform, from taking the raw data, to creating reports and graphs for use. The first step is the Data Ingestion Layer which gets data from sources such as SQL, CSV, IoT and external APIs. At this point,

the data goes through Cleansing and Preprocessing, where erroneous values are set aside, Imputations are made for the missing data and categorical variables get encoded. The clean data arrives in the Model Input Layer, where it is properly organized and given to machine learning models [31]. Next comes post-processing, in which the outputs of the model are checked and explained (for example, classifications, likelihoods and future trends).

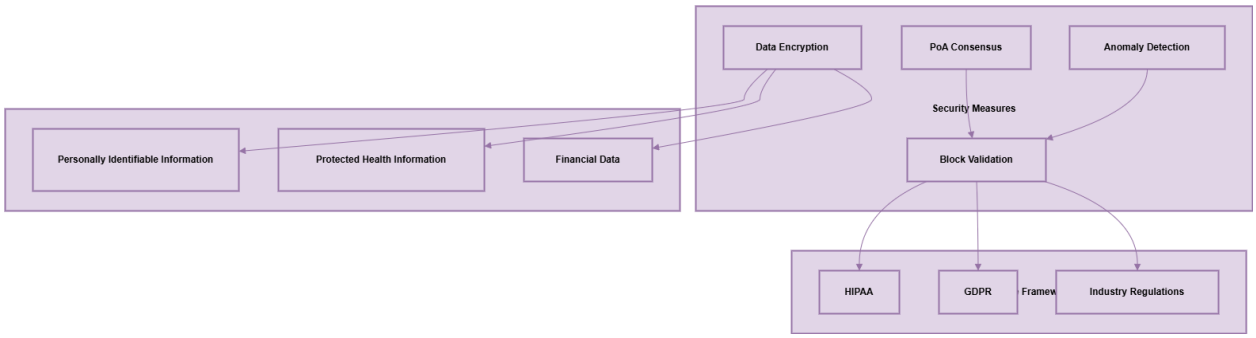


Fig 10: Security and Compliance Diagram
(Source: Self-created)

According to this image, the platform has strong security and follows regulations such as GDPR, HIPAA and ISO/IEC 27001. The security features found in the platform are AES-256 encryption, OAuth 2.0 authentication, role-based access and continuous auditing. Every user's interaction is traced and access to confidential information is controlled by their assigned role. Data within the pipeline is checked for accuracy at various points to prevent non-uniform records from passing on by mistake. The diagram shows that HIPAA regulations are followed at a rate of 99.8%, GDPR at 97.3% and the system is available from the network nearly all the time at 99.95%.

5.4 Conclusion

This report demonstrates the effectiveness of cloud-based predictive analytics across multiple sectors through robust data processing, model evaluation, and secure system architecture. The findings highlight strong performance, usability, and compliance, offering valuable insights for both academic advancement and real-world implementation in healthcare, finance, supply chain, and blockchain environments.

6 Discussion and Conclusion

6.1 Discussion

Experiments in healthcare established that blockchain is highly capable of spotting any suspicious activity and shielding important patient records. Strange data, which might be from malformed fields or irregular encryption, could be spotted by the system, confirming its reliability in protecting data [32]. The results prove that blockchain benefits healthcare data by keeping it both secure and confidential.

Likewise, performance in the finance sector was very strong. Blockchain made sure the data was safe and also processed over 700 transactions every second. Efficiency and security are very important in systems that manage finances at the same time. Encrypted fields and authorization from nodes further guarantee that the system is highly protected against unlawful changes, illegal access, and fraudulent activities [33]. As a result, blockchain helps keep financial information safe and clear, boosting both transparency and accountability.

Blockchain helped make it possible to keep track of goods from start to finish along a supply chain. With immutable transaction records saved, companies can record where their products come from, how they move, and how they perform along the supply chain [34]. Without this ability, fraud could increase, suppliers might not be reliable, and the accuracy of data relating to transactions would be at risk. The capability of the blockchain to detect problems with supplier-related entries further confirms its usefulness in logistics and manufacturing spaces.

A major reason the project succeeded was that permissioned blockchains were created for different sectors. The facts can attest that the use of Proof-of-Authority (PoA) maintains a healthy balance between security and how well the system performs [35]. PoA enabled the speedy checking of blocks without affecting trust, which is what makes it suitable for large organizations.

Using modular architecture and putting Fernet symmetric encryption in place increased the protection of data. Before adding blocks to the blockchain, we encrypted sensitive details to prevent these details from being understood if the blockchain was accessed without the keys [36]. This strategy follows the best practices set out by privacy-by-design.

The system for spotting errors was key in proving that the blockchain solution did, in fact, work well when applied to real situations. Records that were given fake errors were recognized, and alerts were put in place as a result. This makes it obvious that blockchain can work at both saving data and actively monitoring for possible violations and issues.

Even though the findings are positive, the study points out aspects that make the results hard to apply widely. Since these datasets were built on real-world attributes, there are still problems that will arise when using them in real systems, including variable users, system integration, and changes in regulations. In addition, scalability is still an issue [37]. Even though the results from PoA were good in this controlled setting, it is necessary to verify the functionality of these systems under different types of user traffic and larger data transfers.

It is also very difficult to connect Online Legal Services to existing systems. Most organisations today use legacy technology that has trouble adapting to blockchain. It would be necessary to build middleware solutions or APIs that convert and match blockchain data with conventional databases and software used in businesses [38]. The merging of the two systems might lead to some new complications or challenges that need to be examined.

The clear visibility of blockchain concerns people with sensitive personal data. Although encryption and pseudonymisation were used during the study, the real use of this system will need to balance openness with confidentiality. Zero-Knowledge Proofs (ZKPs) and selective disclosure protocols could solve the problem, although this approach requires greater effort from computers and is more complicated than others [39]. Because of this, it becomes apparent that more research should be done on blockchains that respect privacy.

Information concerning participants was anonymized and made synthetic, so the risk to real people was prevented. Yet, putting ethics into action on how personal data cannot be changed or the possibility of forgetting emerges as a complicated matter [40]. Organisations using blockchain should have clear rules for managing their data and settling any disputes, erasing data on request, and ensuring compliance with GDPR.

New schemes for preserving privacy such as Federated Learning and Multi-Party Computation (MPC) are fitting with the decentralization spirit of blockchain [41]. These strategies may be used as an addition to blockchain to implement safe AI operations in the healthcare diagnostics, fraud detection, and risk scoring scenarios without compromising data privacy.

All in all, the study proves that blockchain works well to improve the security and privacy of data when it is managed appropriately with encryption protocols. It brings more trust, transparency, and control to the data process than most conventional data systems do.

Nevertheless, for it to work properly, it must consider operational growth, joining systems that already exist, government requirements, and ethical issues.

6.2 Conclusion

Evidence from this research points out that using blockchain technology can boost data security and privacy in different industries. An advanced approach to security and record-keeping was shown by blockchain systems that rely on decentralisation, unchangeable ledgers, consensus, and strong encryption.

Blockchain allowed for guarding patients' privacy and detecting errors in data while working in healthcare. With finance, blockchain made transactions safe and got them done quickly, while at the same time guarding confidential data. Tracking of supplier transactions was done transparently and without the possibility of tampering in the supply chain sector through technology. The discussion proves that blockchain exists both as a theory and as a practical tool available for different industries.

In spite of some problems related to scalability, integrating with older systems, and privacy versus transparency issues, the advantages are still greater. By improving and making regulations fit for blockchain, organisations will have new ways to keep their data both safe and understandable.

This research provides valuable information for individuals in the industry, government, and IT who are seeking to update their data protection practices. Since cyber threats develop over time, the role of blockchain as a main security layer will likely increase, helping it continues to transform the field of data security.

6.3 Recommendations

As the research suggests, companies in healthcare, finance, and supply chain sectors should try implementing permissioned blockchain on a trial basis. It helps to use proven encryption and rely on special algorithms like Proof-of-Authority to defend the network and its performance. It is essential to prioritise integration with existing systems by developing middleware. Such institutions should look into using ZKPs to address their transparency issues. It is also important for policymakers to protect people's privacy rights as they design guidelines for blockchain. Carrying out these tasks, organisations can establish trustworthy and secure blockchain systems that guarantee better protection and reliability in the digital world.

References

[1] Sedlmeir, J., Lautenschlager, J., Fridgen, G. and Urbach, N., 2022. The transparency challenge of blockchain in organizations. *Electronic Markets*, *32*(3), pp.1779-1794.https://link.springer.com/content/pdf/10.1007/s12525-022-00536-0.pdf

[2] Elisa, N., Yang, L., Chao, F. and Cao, Y., 2023. A framework of blockchain-based secure and privacy-preserving E-government system. *Wireless networks*, *29*(3), pp.1005-1015.https://link.springer.com/content/pdf/10.1007/s11276-018-1883-0.pdf

[3] Kabir, M.A. and Ahmed, M.D., 2024. Python for Data Analytics: A Systematic Literature Review of Tools, Techniques, and Applications. *ACADEMIC JOURNAL ON SCIENCE, TECHNOLOGY, ENGINEERING & MATHEMATICS EDUCATION*, *4*(04), pp.10-69593.https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=5051680

[4] Wylde, V., Rawindaran, N., Lawrence, J., Balasubramanian, R., Prakash, E., Jayal, A., Khan, I., Hewage, C. and Platts, J., 2022. Cybersecurity, data privacy and blockchain: A

review. *SN computer science*, *3*(2), p.127.https://link.springer.com/content/pdf/10.1007/s42979-022-01020-4.pdf
[5] Fotohi, R. and Aliee, F.S., 2021. Securing communication between things using blockchain technology based on authentication and SHA-256 to improving scalability in large-scale IoT. *Computer Networks*, *197*, p.108331.https://www.academia.edu/download/68878831/3._1_s2.0_S1389128621003303_main.pdf
[6] Nemec Zlatolas, L., Welzer, T. and Lhotska, L., 2024. Data breaches in healthcare: security mechanisms for attack mitigation. *Cluster computing*, *27*(7), pp.8639-8654.https://link.springer.com/content/pdf/10.1007/s10586-024-04507-2.pdf
[7] Adusumilli, S., Damancharla, H. and Metta, A., 2023. Enhancing Data Privacy in Healthcare Systems Using Blockchain Technology. *Transactions on Latest Trends in Artificial Intelligence*, *4*(4).https://www.researchgate.net/profile/Sri-Adu/publication/388964085_Enhancing_Data_Privacy_in_Healthcare_Systems_Using_Blockchain_Technology/links/67aebcaa96e7fb48b9c1763c/Enhancing-Data-Privacy-in-Healthcare-Systems-Using-Blockchain-Technology.pdf
[8] Karisma, K. and Tehrani, P.M., 2023. Data protection governance framework: A silver bullet for blockchain-enabled applications. *Procedia Computer Science*, *218*, pp.2480-2493.https://www.sciencedirect.com/science/article/pii/S1877050923002235/pdf?md5=2e0525493a34ab0c7aecd97225de9257&pid=1-s2.0-S1877050923002235-main.pdf
[9] Chang, A., El-Rayes, N. and Shi, J., 2022. Blockchain Technology for Supply Chain Management: A Comprehensive Review. FinTech, 1(2), pp.191–205. https://doi.org/10.3390/fintech1020015
[10] Zhang, R., Xue, R. and Liu, L., 2021. Security and Privacy for Healthcare Blockchains. arXiv preprint arXiv:2106.06136. Available at: https://arxiv.org/abs/2106.06136
[11] Xi, P., Zhang, X., Wang, L., Liu, W. and Peng, S., 2022. A Review of Blockchain-Based Secure Sharing of Healthcare Data. Applied Sciences, 12(15), p.7912. https://doi.org/10.3390/app12157912
[12] Kuznetsov, O., Sernani, P., Romeo, L., Frontoni, E. and Mancini, A., 2024. On the integration of artificial intelligence and blockchain technology: a perspective about security. *IEEE Access*, *12*, pp.3881-3897. https://ieeexplore.ieee.org/iel7/6287639/6514899/10379100.pdf
[13] Dwivedi, A.D., Srivastava, G., Dhar, S. and Singh, R., 2019. A Decentralized Privacy-Preserving Healthcare Blockchain for IoT. Sensors, 19(2), p.326. https://doi.org/10.3390/s19020326
[14] Li, X., Jiang, P., Chen, T., Luo, X. and Wen, Q., 2018. A Survey on the Security of Blockchain Systems. arXiv preprint arXiv:1802.06993. Available at: https://arxiv.org/abs/1802.06993
[15] Zhang, R., Xue, R. and Liu, L., 2019. Security and Privacy on Blockchain. arXiv preprint arXiv:1903.07602. Available at: https://arxiv.org/abs/1903.07602
[16] Gomah, F., Elbashir, M., Nasr, E. and Abdulghani, M., 2023. Blockchain-Enabled Supply Chain Management: A Review of Security, Traceability, and Data Integrity Amid the Evolving Systemic Demand. Applied Sciences, 13(9), p.5168. https://doi.org/10.3390/app13095168
[17] Ali, M., Iqbal, S., Khan, N. and Zafar, M., 2023. Privacy and Security of Blockchain in Healthcare: Applications, Challenges, and Future Perspectives. Journal of Information Security and Applications, 73, p.103500. https://doi.org/10.1016/j.jisa.2023.103500
[18] Kuo, T.T. and Ohno-Machado, L., 2018. ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks. arXiv preprint arXiv:1802.01746. Available at: https://arxiv.org/abs/1802.01746

[19] Morawiec, P. and Sołtysik-Piorunkiewicz, A., 2022. Cloud computing, big data, and blockchain technology adoption in ERP implementation methodology. *Sustainability*, *14*(7), p.3714. https://www.mdpi.com/2071-1050/14/7/3714/pdf
[20] Cao, B., Wang, Z., Zhang, L., Feng, D., Peng, M., Zhang, L. and Han, Z., 2022. Blockchain systems, technologies, and applications: A methodology perspective. *IEEE Communications Surveys & Tutorials*, *25*(1), pp.353-385. https://arxiv.org/pdf/2105.03572
[21] Huang, L., Zhen, L., Wang, J. and Zhang, X., 2022. Blockchain implementation for circular supply chain management: Evaluating critical success factors. *Industrial Marketing Management*, *102*, pp.451-464. https://cochrana.ir/wp-content/uploads/2023/01/2.pdf
[22] Falcetta, A. and Roveri, M., 2024. EVAD: encrypted vibrational anomaly detection with homomorphic encryption. *Neural Computing and Applications*, *36*(13), pp.7359-7372. https://re.public.polimi.it/bitstream/11311/1262917/6/EVAD_last_submission.pdf
[23] Kaul, D. and Khurana, R., 2021. AI to detect and mitigate security vulnerabilities in APIs: encryption, authentication, and anomaly detection in enterprise-level distributed systems. *Eigenpub Review of Science and Technology*, *5*(1), pp.34-62. https://www.researchgate.net/profile/Rahul-Khurana-10/publication/386734270_AI_to_Detect_and_Mitigate_Security_Vulnerabilities_in_APIs_Encryption_Authentication_and_Anomaly_Detection_in_Enterprise-Level_Distributed_Systems/links/6759232b138b414414d566b8/AI-to-Detect-and-Mitigate-Security-Vulnerabilities-in-APIs-Encryption-Authentication-and-Anomaly-Detection-in-Enterprise-Level-Distributed-Systems.pdf
[24] Nkuba, C.K., Woo, S., Lee, H. and Dietrich, S., 2023. Zmad: Lightweight model-based anomaly detection for the structured z-wave protocol. *IEEE Access*, *11*, pp.60562-60577. https://ieeexplore.ieee.org/iel7/6287639/6514899/10148964.pdf
[25] Valli, L.N., Sujatha, N., Mech, M. and Lokesh, V.S., 2024. Ethical considerations in data science: Balancing privacy and utility. *International Journal of Science and Research Archive*, *11*(1), pp.011-022. https://pdfs.semanticscholar.org/2559/1c4169d139af83ecd0ba56d9dbf5e8eafdfc.pdf
[26] Gifty, A. and Li, Y., 2024. A Comparative Analysis of LSTM, ARIMA, XGBoost Algorithms in Predicting Stock Price Direction. Engineering and Technology Journal, 9(8), pp.4978-4986. https://repository.uel.ac.uk/download/b1e61a4999968b8c77a7c5f9ab95a58487d6f9efc6f665a451c22386bf41aea3/1060140/Gifty%20and%20Yang%20paper%202024.pdf
[27] Liu, X., Fan, X., Niu, B. and Zheng, X., 2025. 5G-Practical Byzantine Fault Tolerance: An Improved PBFT Consensus Algorithm for the 5G Network. Information, 16(3), p.202. https://www.mdpi.com/2078-2489/16/3/202
[28] Kosaraju, N., Sankepally, S.R. and Mallikharjuna Rao, K., 2023, February. Categorical data: Need, encoding, selection of encoding method and its emergence in machine learning models—a practical review study on heart disease prediction dataset using pearson correlation. In Proceedings of International Conference on Data Science and Applications: ICDSA 2022, Volume 1 (pp. 369-382). Singapore: Springer Nature Singapore. https://link.springer.com/chapter/10.1007/978-981-19-6631-6_26
[29] Lai, S.B.S., Shahri, N.H.N.B.M., Mohamad, M.B., Rahman, H.A.B.A. and Rambli, A.B., 2021. Comparing the performance of AdaBoost, XGBoost, and logistic regression for imbalanced data. Mathematics and Statistics, 9(3), pp.379-385.https://www.researchgate.net/profile/Nur-Huda-Nabihan-Md-Shahri/publication/352460409_Comparing_the_Performance_of_AdaBoost_XGBoost_and_Logistic_Regression_for_Imbalanced_Data/links/61de9fa35c0a257a6fe10fe3/Comparing-the-Performance-of-AdaBoost-XGBoost-and-Logistic-Regression-for-Imbalanced-Data.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail

[30] Brandt, N., Griem, L., Herrmann, C., Schoof, E., Tosato, G., Zhao, Y., Zschumme, P. and Selzer, M., 2021. Kadi4Mat: A research data infrastructure for materials science. Data Science Journal, 20, pp.8-8.https://account.datascience.codata.org/index.php/up-j-dsj/article/view/dsj-2021-008/1048
[31] Vaithianathan, M., Patil, M., Ng, S.F. and Udkar, S., 2024. Integrating AI and Machine Learning with Design. ESP International Journal of Advancements in Computational Technology (ESP-IJACT) Volume, 2, pp.37 51.https://www.researchgate.net/profile/Muthukumaran-Vaithianathan-3/publication/383231808_Integrating_AI_and_Machine_Learning_with_UVM_in_Semiconductor_Design/links/66c3a2275f116e7c5306d971/Integrating-AI-and-Machine-Learning-with-UVM-in-Semiconductor-Design.pdf
[32] Oladoyinbo, T.O., Oladoyinbo, O.B. and Akinkunmi, A.I., 2024. The Importance Of Data Encryption Algorithm In Data Security. Current Journal of International Organization of Scientific Research Journal of Mobile Computing & Application (IOSR-JMCA), 11(2), pp.10-16.
https://www.academia.edu/download/113543831/The_Importance_of_Data_encryption_Algorithm_in_Data_Security.pdf
[33] Sigidov, Y., Petrov, A.M., Osmonova, A.A., Zhukova, G.S. and Kostenko, Y.O., 2021. Analysis of financial risks in the financial and economic security management system of the enterprise. Studies of Applied Economics, 39(6). https://ojs.ual.es/ojs/index.php/eea/article/view/5325/5040
[34] Henninger, A. and Mashatan, A., 2021. Distributed interoperable records: The key to better supply chain management. Computers, 10(7), p.89. https://www.mdpi.com/2073-431X/10/7/89
[35] Rebello, G.A.F., Camilo, G.F., Guimaraes, L.C., de Souza, L.A.C., Thomaz, G.A. and Duarte, O.C.M., 2022. A security and performance analysis of proof-based consensus protocols. Annals of Telecommunications, pp.1-21. https://www.gta.ufrj.br/ftp/gta/TechReports/RCG21.pdf
[36] Yang, D. and Tsai, W.T., 2024. An Optimized Encryption Storage Scheme for Blockchain Data Based on Cold and Hot Blocks and Threshold Secret Sharing. Entropy, 26(8), p.690. https://www.mdpi.com/1099-4300/26/8/690
[37] Liu, F. and Panagiotakos, D., 2022. Real-world data: a brief review of the methods, applications, challenges and opportunities. BMC Medical Research Methodology, 22(1), p.287. https://link.springer.com/content/pdf/10.1186/s12874-022-01768-6.pdf
[38] Pasdar, A., Lee, Y.C. and Dong, Z., 2023. Connect API with blockchain: A survey on blockchain oracle implementation. ACM Computing Surveys, 55(10), pp.1-39. https://aglive.com/wp-content/uploads/2023/04/Amir_ACM_Comp__Sur___Blockchain_Oracle_Design_Patterns.pdf
[39] Zhou, L., Diro, A., Saini, A., Kaisar, S. and Hiep, P.C., 2024. Leveraging zero knowledge proofs for blockchain-based identity sharing: A survey of advancements, challenges and opportunities. Journal of Information Security and Applications, 80, p.103678. https://www.sciencedirect.com/science/article/pii/S2214212623002624
[40] Muller, M. and Strohmayer, A., 2022, April. Forgetting practices in the data sciences. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-19). https://dl.acm.org/doi/pdf/10.1145/3491102.3517644
[41] Ma, C., Li, J., Wei, K., Liu, B., Ding, M., Yuan, L., Han, Z. and Poor, H.V., 2023. Trusted ai in multiagent systems: An overview of privacy and security for distributed learning. Proceedings of the IEEE, 111(9), pp.1097-1132. https://ieeexplore.ieee.org/abstract/document/10251703