

**Research paper**

**Journal of Artificial Intelligence and Data Mining (JAIDM)** Journal homepage: http://jad.shahroodut.ac.ir



# **An Intelligent Blockchain-Based System Configuration for Screening, Monitoring, and Tracing of Pandemics**

Ali Shabrandi<sup>1\*</sup>, Ali Rajabzadeh ghatari<sup>2</sup>, Mahmoud Dehghan Nayeri<sup>3</sup>, Nader Tavakoli<sup>4</sup> and Sahar Mirzaei<sup>5</sup>

*1,\* PhD candidate of Industrial Management, Faculty of Management and Economics, Tarbiat Modares University, 1411944961, Tehran, Iran. Email: a.shabrandi@modares.ac.ir, corresponding author, Phone: +989129365477*

<sup>7</sup> Department of Industrial Management, Faculty of Management and Economics, Tarbiat Modares University.

3 *Department of Industrial Management, Faculty of Management and Economics, Tarbiat Modares University.*

*<sup>4</sup> Department of Emergency Medicine, Trauma and Injury Research Center, Iran University of Medical Sciences.*

5 *Graduate of Environmental Health Engineering, Iran University of Medical Sciences, Tehran, Iran.*

#### **Article Info Abstract**

#### **Article History:**

*Received 05 April 2024 Revised 25 May 2024 Accepted 29 July 2024*

*DOI:10.22044/jadm.2024.14379.2541*

**Keywords:**

*Pandemic, Blockchain, Artificial Intelligence, Configuration, System*

*\*Corresponding author: [A.Shabrandi@Modares.ac.ir](mailto:A.Shabrandi@Modares.ac.ir) (A.Shabrandi).*

This study proposes a high-level design and configuration for an intelligent dual (hybrid and private) blockchain-based system. The configuration includes the type of network, level of decentralization, nodes, and roles, block structure information, authority control, and smart contracts and intended to address the two main categories of challenges –operation management and data management– through three intelligent modules across the pandemic stages. In the prehospital stage, an intelligent infection prediction system is proposed that utilizes in-house data to address the lack of a simple, efficient, agile, and low-cost screening method for identifying potentially infected individuals promptly and preventing the overload of patients entering hospitals. In the in-hospital stage, an intelligent prediction system is proposed to predict infection severity and hospital Length of Stay (LoS) to identify high-risk patients, prioritize them for receiving care services, and facilitate better resource allocation. In the post-hospital stage, an intelligent prediction system is proposed to predict the reinfection and readmission rates, to help reduce the burden on the healthcare system and provide personalized care and follow-up for higher-risk patients. In addition, the distribution of limited Personal protective equipment (PPE) is made fair using private blockchain (BC) and smart contracts. These modules were developed using Python and utilized to evaluate the performance of state-of-the-art machine learning (ML) techniques through 10-fold cross-validation at each stage. The most critical features were plotted and analyzed using SHapely Adaptive exPlanations (SHAP). Finally, we explored the implications of our system for both research and practice and provided recommendations for future enhancements.

#### **1. Introduction**

Biological disasters have become a global threat over the last century, causing severe health and socioeconomic impacts that surpass other types of disasters [1]. Disaster management aims to mitigate hazard effects, provide adequate assistance to the affected population, and enable prompt and effective recovery [2]. Unlike other

disasters, pandemics are a unique type of natural disaster that have global impacts and long-term economic consequences that cross national boundaries. This necessitates an interdisciplinary and intersectoral approach to integrated systems that assist health, economic, and social perspectives, and balance the conflicting objectives among these domains. Furthermore, the dynamic nature of a pandemic requires real-time data collection and analysis to inform decision making and response efforts.

During a pandemic, the data needed to answer questions varies by stage [3]. For instance, in the pre-pandemic stage, there is a limited number of infections that focus on gathering data on epidemiological parameters. As the pandemic progresses, tracking the virus spread, monitoring healthcare capacity, and assessing vaccination campaign efficacy by gathering more data will become possible. These data can help inform public health strategies and interventions, thereby allowing authorities to make informed decisions. Additionally, analyzing the data collected throughout the pandemic can provide valuable insights into the long-term effects of the virus and guide future preparedness efforts.

A new coronavirus emerged in Wuhan, China, in December 2019, caused by the novel "Severe Acute Respiratory Syndrome Coronavirus 2, " which was later named COVID-19 by the World Health Organization (WHO) in February 2020. This epidemic quickly spread from China to other countries, and in less than three months after the

outbreak started, the disease became a global pandemic [4]. The COVID-19 sudden outbreak imposes an overwhelming burden on countries' medical systems through a surge in the need for hospital beds and a lack of medical equipment [5, 6] and has opened a new opportunity for extensive research in various fields.

Figure 1 presents three stages that a patient with COVID-19 may encounter during the pandemic, beginning with the pre-hospital stage. At this initial stage, the patient may exhibit symptoms of the disease and undergo screening tests. If test results are positive or inconclusive, the patient may be referred to a hospital for further diagnosis and treatment. If the indicators suggest that hospitalization is necessary, the patient may be assigned to one of five levels of hospitalization based on the severity of the disease. Following the completion of treatment and discharge from the hospital, the patient enters the post-hospital stage, during which recovery is monitored. During this stage, the patient may face the risk of readmission to the hospital due to complications or relapse or may fully recover and subsequently contract the virus again at a later time.



**Abbreviation: UOB:** Under Observation, **GU:** General Unit, **ICU:** Intensive Care Unit, **INT:** Intubated, **IHM:** In-Hospital Mortality

#### **Figure 1. Scope of our study**

Our research is motivated by the existence of two significant categories of challenges in the management of COVID-19 pandemic, which are applicable to the management of other pandemics in general. Table 1 provides a summary of the challenges identified in the literature and their corresponding proposed solutions. The first category of challenges relates to the operation management of the pandemic response. In the prehospital stage, the main challenge is the lack of a simple, efficient, agile, and low-cost screening method to identify potentially infected people. This is due to the intrinsic characteristics of the coronavirus including high infection rates [7], different virus mutations [8], and the possibility of reinfection [9], which necessitate such a system. The lack of early diagnosis results in an increase in hospitalized patients and subsequently puts pressure on the healthcare system [10, 11]. Moreover, overlap of the initial symptoms of this disease with those of other common illnesses complicates this problem [12]. At the beginning of the pandemic, the lack of test-kit, the timeconsuming diagnosis and the need for specialized

laboratories limited the use of these kits use in the COVID-19 pandemic [13]. It is beneficial to have a publicly accessible online system that intentionally directs individuals to appropriate facilities during the early stages of an outbreak by examining the symptoms and medical histories of those who are suspected of having COVID-19.

## **Table 1. Main COVID-19 pandemic management**





As evidence, a recent study showed that different pre-hospital pathways of COVID-19 patients have a significant impact on risk factors that could affect the outcomes of hospitalizations, resuscitations, and deaths [14]. In the in-hospital stage, the main challenge is the inability to anticipate the risks associated with disease severity and LoS in the hospital. This reduces the possibility of devising a suitable treatment plan and lowers the preparedness of decision-making centers in the event of hospital admission surges. Therefore, a system for allocating limited resources to patients is needed to monitor patients and allocate them based on relevant indicators, which in turn can help reduce mortality [15]. In the post-hospital stage, the main challenge is the lack of a tracing system for discharged patients in terms of symptoms and disease improvement processes, especially for those at higher risk. The uncertainty of immunity levels [16] and the potential for reinfection with this disease [17] exacerbate this issue. The lack of such a system reduces the effectiveness of treatment plans and impedes the collection and

analysis of data on the long-term effects and outcomes of COVID-19. This hinders the development of evidence-based guidelines and policies for the management and prevention of COVID-19. Additionally, the frequency and types of post-COVID symptoms, the risk factors and outcomes associated with post-COVID symptoms, such as return to emergency services, hospital readmission, post-discharge death and preventive and rehabilitative strategies for post COVID-19 patients are the most important inputs for macro policies in the control and management of epidemics. Therefore, the scope of designing a system that addresses the management of the COVID-19 pandemic extends beyond hospital monitoring and encompasses the period following discharge and subsequent follow-ups.

The effective utilization of PPE by hospitals on the frontlines of pandemics and continuous exposure to infectious diseases is crucial for preventing and controlling the spread of viruses. The COVID-19 pandemic highlighted the importance of PPE, as many countries have reported a shortage of PPE in hospitals due to the absence of a reliable system to provide accurate data on the demand and supply of PPE. In some instances, medical professionals were forced to use tape to mend torn masks to avoid contracting COVID-19 [18]. BCT can facilitate verifiable payment settlements to ensure that the PPE quota allocated by Ministry of Health and Medical Education (MOHE) reaches the end customer. To promote transparency in the PPE supply network, the BC can propose the use of a private BC and smart contract that involves all relevant parties, including consumers, manufacturers, and distributors.

The second category of challenges during the COVID-19 pandemic concerns data management, which involves the collection, maintenance, and sharing of COVID-19 data. Traditional systems have limitations, such as a single point of failure [19] and a lack of data security and privacy [20, 21]. Moreover, the inaccurate statistics of patients at the onset of the pandemic [22] highlight the need for data sharing [23] and transparency [24] to support the research community and the governments in making appropriate policies to cope with the unprecedented demand for information and knowledge [23]. Therefore, it is vital to adopt new technologies that enable secure and collaborative data sharing among various stakeholders, such as governments, epidemiologists, researchers, bioengineers, funding agencies, and physicians. This can facilitate the development of vaccines, drugs, procedures, treatment methods, and effective

prevention and treatment strategies for COVID-19 [25, 26]. However, traditional healthcare data management systems face significant challenges in ensuring privacy, security, immutability, transparency, traceability, auditing, data authenticity, flexible access, and trust [27]. Many of these systems rely on a centralized architecture, which organizes data in a hierarchical structure and stores it in a central repository [19]. This makes them vulnerable to the single point of failure risk in case of natural disasters [28]. To alleviate the pressure on the health and treatment system and provide optimal care for patients, it is essential to implement systems that can enable early, efficient, and prognostic diagnosis of COVID-19. These systems should also respect data confidentiality and provide data on the patient's condition to the relevant parties. Furthermore, robust and intelligent systems and collaboration among various stakeholders can ensure accurate and timely dissemination of knowledge. This can assist the data users, such as researchers, doctors, and policymakers, in devising and applying effective solutions for the prevention and treatment of COVID-19.

The main objective of this study is to tackle the challenges mentioned above and to design and configure a high-level system based on BCT and AI, named BlockCOV. This system aims to support the pre-hospital screening of potential patients, predict the severity and LoS of hospitalized patients, and trace discharged patients to predict the reinfection and readmission rates. This research endeavors to advance the field of application of expert system by examining the potential applications of BC and ML technologies in pandemic management. It is expected that the results of this research will facilitate the utilization of the proposed system in future pandemics. To achieve this objective, this study addresses the following questions.

**RQ1:** How can a system be designed and configured to assist at different stages of the COVID-19 pandemic?

**RQ2:** How can BC and ML technologies be utilized to assist in the management of the COVID-19 pandemic?

**RQ3:** What are the main demographics, symptoms, and comorbidities used to predict infection?

**RQ4:** What are the main demographics, symptoms, and comorbidities used to predict severity?

**RQ5:** What are the main risk factors to predict LoS in hospitalized patients?

This research seeks to expand the knowledge of the predictive risk factors for COVID-19 infection, severity, and LoS in hospitals, with a particular focus on Iran.

The structure of this paper is as follows: Section 2 reviews the literature on AI, BCT, and COVID-19, providing a brief overview of related works and an introduction to BC and its fundamental properties. Section 3 constitutes the core of our study, where the proposed BC-based system design and deployment process are detailed. This section also analyzes the suitability of BC for COVID-19 pandemic management through a decision model, delving into the conceptual model, network architecture, and configurations of BlockCOV. Section 4 presents the implementation of intelligent recommendation modules, along with the results and analysis of experiments, addressing research questions RQ3 to RQ5. Section 5 discusses the findings and implications of this study, offering conclusions and potential directions for future research.

#### **2. Literature Review 2.1. AI and COVID-19**

Intelligent systems, including AI and MLalgorithms, have been widely applied and evaluated in various domains [29-32], specifically in healthcare [33, 34]. This is evidenced by the publication of approximately 80 review paper on this topic since 2020. AI techniques play a significant role in the early-stage screening [35] and rapid diagnosis (pre-hospital) [36, 37], severity classification (in-hospital) [38, 39], LoS prediction, and tracing (post-hospital) [40] of COVID-19 patients. Moreover, they can support public health professionals in making complex decisions [41].

## **2.2. AI and Pre-hospital Stage**

In the pre-hospital stage, an effective screening scheme leads to the rapid diagnosis of COVID-19, thereby reducing the burden on healthcare systems. Most previous models were based on clinical data and thus were not effective for the rapid screening of COVID-19 in the general population. In the literature, COVID-19 infection prediction models have used different input data such as demographic, X-ray [38], CT scans [39, 42-46], symptoms [47], laboratory tests [48, 49], comorbidities and/or a combination of these features [35, 50]. However, each prediction method has its drawbacks. For example, CT-based models require expensive equipment and professional staff, expose patients to unnecessary irradiation [51], and result in overwhelming use of the limited resources of the health system.

Our study focused on the use of in-house and selfreportable data to predict COVID-19 infection. This involves the use of demographic data, symptoms, and comorbidities, which are nonclinical data, as input features for the prediction model. Few studies have addressed this, but they are essential for the frontline response to COVID-19. Guhathakurata et al. [52] used a Support Vector Machine (SVM) model to predict COVID-19 infection based on 16 symptom-related features. They reported a high accuracy of 98.73, but their study had some limitations. For example, they did not include demographic variables, describe how they trained and tested the SVM model, explain how they chose the hyperparameters, and used a small and imbalanced dataset for the symptom data. These factors may affect the validity and reliability of the SVM model. In a related study, Guhathakurata et al. in a subsequent paper [53] attempted to improve the previous study by expanding the coverage of the input features for COVID-19 prediction. They employed artificial neural networks (ANNs) based on demographic data (four features), symptoms (nine features) and comorbidities (three features) to diagnose COVID-19. They used a dataset of approximately 10,000 samples, and reported that the long short-term memory (LSTM) model achieved the highest accuracy of 98.9. However, their approach is a black-box method that lacks transparency, interpretability, and explainability for users and stakeholders. Malik et al. [54] applied five ML algorithms using demographic and symptom data to diagnose COVID-19 in patients. They found that Naive Bayes (NB) and Decision Tree (DT) were the best performing methods, achieving an accuracy of 93.70 each. However, their study was limited by the use of a small and imbalanced dataset that may not capture the variability of COVID-19 symptoms or characteristics of the general population.

## **2.3. AI and In-hospital Stage**

From the perspective of the COVID-19 pandemic, two criteria are crucial for the hospital utilization of COVID-19 patients. The first criterion is the prediction of COVID-19 severity. In the literature, the assessment of COVID-19 severity has employed several methods [55], which can be compared into two dimensions based on the type of variables used and the definition of severity. In this study, we utilized demographic data, symptoms, and comorbidities, as they are more agile and useful for rapid responses to the pandemic. Other types of data, such as electronic medical records (EMR), laboratory tests, imaging, prescriptions, and vital signs, may also be employed, but they are more time-consuming and less conducive to prompt interventions. The second dimension encompasses three main definitions of severity: those established by national or global health organizations like WHO or Centers for Disease Control (CDC), research groups, prior publications, or a combination of different clinical events. Defining the severity of COVID-19 and its associated risks is crucial for proactive clinical decision-making and resource allocation [56]. However, this definition is not fixed, but evolves over the course of the pandemic. Therefore, we used clinical events as the severity degree, and it will be discussed in section 3.2.2. Such prediction models can facilitate the stratification of patients based on their risk level and inform decisions on whether outpatient treatment is adequate or hospital admission is warranted. Moreover, highrisk patients can benefit from more advanced and expensive diagnostic tests, such as chest CT, instead of conventional X-rays, or a more extensive blood count.

The second criterion for hospital utilization of COVID-19 patients is their LoS of COVID-19 patients in the hospital. The LoS criterion determines how long a patient with COVID-19 must stay in the hospital and when they can safely be discharged. This is measured as the total duration of the patient's hospitalization over a specific time frame, which includes all consecutive admissions and discharges.

According to the current literature, LoS models can be broadly categorized into two main types: classification models and regression models. The former is typically employed to forecast categorical results, while the latter predict numerical values. Numerous studies have used ML algorithms to predict survival and calculate the LoS of patients [57]. According to Ebinger et al. [58], the LoS of COVID-19 patients was predicted by three ML models trained on the EHRs of 966 patients from a large US academic and medical center, with an accuracy of 0.765. In the ICU of Saudi Arabia, the Random Forest model achieved the highest accuracy (94.16) in predicting the LoS of COVID-19 patients, as reported in a previous study [59]. Another study in Iran compared seven ML techniques and found that the SVM algorithm performed best on the laboratory data of 1225 COVID-19 patients, with an average accuracy of 99.5, average specificity of 99.7, and average sensitivity of 99.4 [60]. A systematic study revealed that factors such as age, sex, and chronic

comorbidities, such as hypertension and diabetes, had a significant impact on the risk of death and LoS in COVID-19 patients [61]. In [62, 63], a datadriven methodology using ML algorithms was used to predict the LoS of admission for COVID-19 patients. In this study, we applied ML algorithms to predict LoS, determined the most important features, and compared them to those in previous studies.

## **2.4. AI and Post-hospital Stage**

There are many studies on the role of AI in the COVID-19 pandemic, and we refer the reader to the latest related review articles. [64-74].

## **2.5. Intelligent BC and COVID-19**

Blockchain technology (BCT) originated from virtual currency transactions [75] and diversified its applications in various fields. In recent years, the health sector has attracted attention, particularly due to the growing interest in BCenabled programs, which have altered its direction [76-80]. Given rapid advancements in healthcare, the industry has recognized BC as a flexible technology with potential benefits [81, 82]. The medical community's adoption of innovative methods at the intersection of healthcare and BC has led to a focus on identifying the root causes of the current healthcare systems and exploring potential solutions based on BCT. This has resulted in numerous countries' dedicated efforts towards transforming the entire healthcare ecosystem. For example, IBM Watson agreed to a two-year contract with the US Food and Drug Administration to implement BCT to securely share patient data [83]. The US Centers for Disease Control and Prevention are also testing BC capabilities, such as time-stamping, peer-to-peer reporting, and feature processing, for real-time detection of disease outbreaks [84].

Recent studies have explored the use of BC in information management systems [85], especially in the healthcare context [27, 86, 87]. One of the most serious challenges in healthcare is data management, which faces problems such as lack of diagnostic data, interoperability, and inability to maintain confidentiality and security of patient health records [20, 21]. BCT offers a promising solution for these challenges by enabling secure and transparent data management. As a result, proposed solutions have been created to address the shortcomings of healthcare information technology systems in both the public and private sectors. For instance, Stafford et al. [88] claimed that organizations that adopt BC can ensure rapid interoperability between healthcare, user-centered

medical research, and the prevention and detection of counterfeit drugs. BC can also help increase the accuracy of disease diagnoses in cases where security and privacy are challenging for the healthcare systems [89].

Literature on intelligent systems often concentrates on limited stages of pandemic management. Studies [90, 91] have proposed a system design that utilizes BC, AI, and drones to control the spread of COVID-19. The system capitalizes on the benefits of BC, such as security, transparency, and decentralization, to ensure the dependability and reliability of data collected by drones. AI is integrated to provide drones with image processing, face recognition, and object detection capabilities. The system also employs AI to analyze data and offers real-time feedback and guidance to authorities and the public. Another study [92] presented a communication scheme that leverages BC and AI to enable multiswarm drones to address COVID-19 situations. Study [93] proposed a smart healthcare system that integrates BC and AI to monitor and detect COVID-19 in biomedical images. This system can be used for self-testing, diagnosis, and data sharing. It employs deep learning models to analyze chest X-ray images and classify them as COVID-19 positive or negative. The system also utilizes BCT to store and verify diagnosis results and patient information. Some research on COVID-19 pandemic management aims to monitor patients who are in the hospital or in isolation using BC and ML

methods. For example, [94] presented a framework of oxygen level monitoring and severity calculation for COVID-19 patients using a private BC on Hyperledger Fabric. The study also uses AI to analyze the oxygen saturation data and generate a severity score and a probabilistic decision of being a COVID-19 patient. Similarly, [95] employed AI to analyze the blood oxygen saturation data and generate a severity score and a probabilistic diagnosis of COVID-19. The paper also utilized BC to store and verify the diagnosis results and patient information in a secure and decentralized manner.

## **2.6 BC and Its Fundamental Properties**

BC, a distributed database or ledger, is an appendonly store of time-stamped transactions maintained across many machines (nodes) in a peer-to-peer network. The BC structure consists of a linked list of blocks that contain an ordered set of transactions. The connection between a block and its predecessor is often secured using cryptographic hashing. Figure 2 shows the overall structure of BC. BC gathers transactional data and

stores them in blocks. Once a block is full, the data is encrypted to produce a hash, which is a hexadecimal number. The hash is then added to the header of the next block and encrypted with the other data in the block. Consequently, a chain of connected blocks is produced. Next, transactions follow a specific process depending on the BC nodes involved. BC nodes are users with different roles across the BC network, for example, to initiate or validate a transaction. Typically, the transaction is sent to the network and waits until the miner nodes pick it up and the mining process starts. User nodes are responsible for verifying the transactions in the chain. The process of agreeing with the validity of transactions in the chain is called consensus. Many consensus algorithms have been proposed in the literature, each with unique performance and security features [96, 97]. BCT has several characteristics that make it highly prevalent. Each of these characteristics is briefly explored below:



**Figure 2. Blockchain overview**

### **2.6.1. Decentralization**

Decentralization is a process in which power and authority are distributed away from a centralized place or authority, allowing everyone in the network to access data and conduct transactions directly with end-users without the need for a trusted third party. This is in contrast to a centralized network, where all transactions are handled by a centrally trusted authority, making it easier for hackers to tamper with data on a single node and limiting access to transaction history only to registered individuals. BCT uses decentralization principles to address these issues and improve data security.

## **2.6.2. Transparency and Privacy**

BC is a Distributed Ledger Technology (DLT), characterized by its decentralized and secure nature. It comprises a distributed database shared among nodes in a peer-to-peer network. Each node maintains a copy of the BC ledger, which is updated as it continues to complete its network-

wide responsibility to validate transactions. Transparency is an essential characteristic of BCT because every transaction is visible to every node in the network. However, the use of pseudonyms in BC transactions ensures that the privacy of the user is protected even if the transaction is made using a public address. In other words, the real identity of the user remains concealed, whereas the transaction history is transparent and visible to all nodes in the network.

## **2.6.3. Immutability**

BC offers a framework for measuring truth [98]. Immutability, a fundamental characteristic of BC, refers to its capacity to maintain a permanent, irrevocable, and unchangeable history of transactions. This property has the potential to enhance the trustworthiness and reliability of data that companies use and exchange on a daily basis while also streamlining the auditing process. The irreversible nature of BC transactions is due to the hashing process of the blocks in the chain. The hashing process of a new block always incorporates metadata from the hash results of the preceding block. It is impossible to manipulate or delete data after it has been validated and placed in the BC because subsequent blocks in the chain would reject the attempted modification owing to invalid hashes. In other words, if data is tampered with, the BC will fail, and the cause will be obvious. This property is not observed in traditional databases, where information can be easily updated or erased.

### **2.6.4. Peer to Peer Network**

The peer-to-peer network property in BC is a key characteristic that enables direct interaction and data transmission between nodes or participants in the BC network without reliance on a central authority or intermediary. This property is critical for the decentralization, distribution, and security of BC networks. Additionally, it enhances the scalability, performance, and robustness of the network by accommodating a large volume of transactions and nodes without impeding the system speed or reliability. The realization of the peer-to-peer network property in the BC is achieved through the use of cryptographic protocols, consensus algorithms, and distributed ledger technology, which guarantees the validity, integrity, and immutability of the data stored and exchanged on the BC.

#### **2.6.5. Distributed Ledger**

A distributed ledger is a characteristic of a BC that enables data to be stored and synchronized across multiple nodes or computers in a network. Each node maintains a copy of the ledger that records transactions, such as the exchange of assets or data, among the network participants. This enhances the transparency, security, and decentralization of the ledger, as no single authority or intermediary can control or manipulate data. BC is a specific form of DLT, but not all distributed ledgers follow the BC model. BC is a specific form of DLT, but not all distributed ledgers follow the BC model. BC employs a data structure that links transactions in chronological order using cryptographic hashes. This ensures the immutability of the ledger, as any alteration in one block would invalidate the hashes of subsequent blocks. Other forms of distributed ledgers can adopt different data structures or consensus mechanisms to achieve comparable objectives.

#### **2.6.6. Irreversibility**

Irreversibility is a property in which a process, once it occurs, cannot be reversed. Hashing is a complex process that generates a unique fixedlength output for any input that cannot be inverted. For example, a private key cannot be generated by using a public key. Furthermore, a small change in the input can result in a completely different output. Therefore, minor modifications are not an option for this system. To compromise the network, it is necessary to alter every piece of data stored at every node. Moreover, because the hash function is one-way, it is impossible to reconstruct the original data from the hash results. Therefore, to modify or erase a transaction in the BC, all subsequent blocks must be changed, which is practically impossible because of the high computational power and consensus required.

#### **2.6.7. Anonymity**

Anonymity means that the real identity of the user is hidden from the public, whereas the transaction history is transparent and visible to all nodes in the network. Anonymity is achieved by employing pseudonymous addresses, which are unique strings of characters that represent a user's identity in the BC. These addresses are generated through cryptographic processes and are not directly linked to a person's real-world identity. As a result, users can conduct transactions without disclosing their personal information or revealing their true identities. However, anonymity is not absolute on the BC, as advanced techniques can be used to trace transactions and infer the identity of users. For example, some BC networks use proof-of-work algorithms that require users to reveal their public keys, which can be used to track their activities.

Some BC networks use zero-knowledge proof systems, which enable verification without disclosing data, to enhance user anonymity.

#### **2.6.8. Auditability**

Auditability is a characteristic of the BC that enables the data stored in the BC to be tamperresistant and verifiable. This ensures that transactions recorded in the BC are transparent, traceable, and accountable. Auditability is achieved by using cryptographic techniques, such as digital signatures, hashes, and proofs, to link transactions in chronological order and prevent any unauthorized changes. Auditability also allows network participants to verify the validity and integrity of data without relying on a central authority or intermediary. Auditability is one of the main advantages of using BCT in various applications, such as financial services, public registries, provenance, and regulation. Audits can take different forms depending on their purpose and scope. They may include financial audits, compliance and regulatory audits, or any combination thereof, and BCT can facilitate any of them. In the BC, a digital distributed ledger records and validates all transactions that occur in a network using a digital timestamp. Consequently, past records can be audited and traced by accessing any node in the network [99].

#### **3. Proposed Method**

Figure 3 illustrates the idealized process and steps for designing and deploying a BC-based system.



#### **Figure 3. A proposed blockchain-based system design and deployment process**

This involves describing the problem statement and requirements in Sections 1 and 2, by an evaluation of the suitability of BC and a presentation of high-level conceptual design and BC configuration. However, the detailed design of the BC and its development, testing, and deployment are not the focus of this study (Steps 5–7). While a platform has been selected for this research among the existing ones, some design

parameters are not mentioned here, as they are dependent on the platform chosen which is explaned in Section 3.3.

### **3.1. BC Suitability**

**No. Answers**

The initial consideration in determining whether to implement a BC-based solution to tackle a particular problem is to evaluate the appropriateness of employing the BC. A recent study [100] revealed that 62% of the share of BC suitability in BC evaluation studies highlights the significance of this evaluation step. Given the intricate internal structure of BC and its numerous configurations and variants [101], it is crucial to assess whether the BC aligns with the requirements and characteristics of the specific use case [102]. Although a wide range of BCs have emerged since the introduction of Bitcoin in 2008 [75], there are few models available in the literature to guide the evaluation of BC suitability for use cases. Consequently, there is limited knowledge and many common misconceptions in this area [103]. Existing approaches that assist decision-makers in adopting BC are divided into three main categories: decision flowcharts, conceptual frameworks, and decision models [100]. Decision flowcharts consist

of a series of questions that help the user obtain a recommended decision based on the fundamentals of BCT. These flowcharts can be used efficiently by decision makers from various backgrounds and do not require an in-depth technical understanding of BCT. By reviewing the literature on the suitability of BC, 15 papers were found. The primary outcomes of these decision-making flowcharts are Yes or No answers regarding the suitability of the BC [104-107], determining the type of BC [108-112] and alternative solutions to BC such as Central Database, Shared Central Database, Distributed DB, or Distributed Ledger [102, 113-115]. The series of inquiries utilized in these flowcharts encompasses a spectrum of topics, ranging from the fundamental properties of BC to more technical aspects, such as consensus algorithms and on-chain vs. off-chain strategies. In this study, we were inspired by two of the most cited works, Wüst et al. [102] and Lo et al. [109], and we proposed a customized COVID-19 decision flowchart. Figure 4. shows the flowchart and Table 2 provides questions and answers to evaluate the suitability of applying BC to the COVID-19 pandemic problem.

#### **Table 2. BC suitability inquiries**

- **1 Yes.** To analyze the patients' data in the three stages of the pandemic, including screening, hospitalization, and follow-up after discharge, it is necessary to store the patients' pathways throughout the treatment process so that the patient can be tracked effectively and securely whether an individual state infected or not needs to be stored. If data storage is not required, BC will not provide any additional benefits to the existing technical solutions.
- **2 Yes.** Multiple parties involve in the COVID-19 pandemic. The BC network includes participants from various sectors and roles such as primary care centers, public and private hospitals, laboratories, patients, physicians, research institutes, government authorities, and the MOHE. These participants can read and/or write data on the BC depending on their permissions and responsibilities.
- **3 No.** There is no need for an entity such as the government or the MOHE to execute a certain operation or alter the policy or configuration of an operation.
- **4 Yes.** All writers on the BC have unique identifications. Primary and secondary care facilities are assigned a unique ID and every individual, including physicians and patients, has a unique National Code. If a system requires only a single writer, a BC will not offer any extra assurances compared to a conventional database, which would likely be a more suitable choice, especially from a performance perspective.
- **5 No.** Not all writers on the BC are necessarily trusted, as they may have different incentives, interests, and perspectives.
- **6 Yes.** The trusted authority that verifies and provides data and information on the BC is distributed among multiple nodes and is not controlled by a single entity. A decentralized trusted authority does not rely on a single point of trust or failure, but uses a network of nodes that can validate and verify the data using cryptographic techniques.
- **7 Yes.** The operation of the BC system is distributed among multiple nodes. A decentralized operation does not rely on a single point of control or failure but uses a network of nodes that can execute and process transactions on the BC.
- **8 Yes.** Immutability ensures that the data stored in the BC network cannot be altered or tampered with by any malicious actor, and prevents data misuse or political censorship. Immutability can enhance the trustworthiness and transparency of the data, which can be useful for adopting appropriate policies against various pandemics, such as tracking the spread of the virus, verifying test results, allocating medical resources, and monitoring the patient's condition and progress.
- **9 No.** It does not require a high performance, such as PayPal. However, BC may not reach the same level of performance as centralized systems such as PayPal, advancements in technology and optimization techniques can help improve its scalability and speed over time.



**Figure 4. COVID-19 blockchain suitability flow chart**

#### **3.2. BC Configuration**

Since the advent of bitcoin, many configurations and variants have emerged [101]. Therefore, it is essential to consider systematically the features and configurations of BCs and their impact on quality attributes for the entire system that is to be designed. We have categorized the configurations of BC-based system into the following categories: a) level of decentralization and Authority control, b) Network participants' definition, c) block structure, d) smart contracts.

## **3.2.1. Level of Decentralization and Authority Control**

In the BC, decentralization refers to the transfer of supervision and decision-making from a centralized organization to a distributed network. Decentralized networks aim to reduce the degree of trust that participants need to place in each other and prevent their ability to exercise authority or command over each other in ways that degrade the efficiency of the network. Depending on the level

of decentralization required, there are four types of BCs.

**Public BC:** It operates on open-source principles. It allows anyone to join and participate in the network as a user, developer, or community member, without imposing any barriers to entry. The participants collectively maintain the integrity and validity of the ledger, which records all transactions and state changes. A public BC is transparent, as every participant can access and verify the transaction history and the current state of the ledger. It is also decentralized as no central authority or intermediary controls or regulates the network. Moreover, a public BC is inclusive and resilient because it enables global participation and resists censorship and shutdown. Well-known examples of public BCs are Bitcoin and Ethereum.

**Private BC:** It is a private asset of an organization or individual. Unlike the public BC, a central administrator manages all critical aspects of the network. Participants can join such a private network through authentic and verified invitations. Validation is also necessary, either by the network operator(s) or a clearly defined set of protocol rules implemented by the network. Unlike public BCs, private BCs do not disclose transactions or related information to the public. The most popular private BC platform is Hyperledger Fabric.

**Consortium BC:** It is governed by a selected group of participants (e.g., several organizations) that form a consortium or federation. These BCs are semi-decentralized and involve multiple organizations working together. They allow each organization to have a representative node that participates in the consensus process. The main advantage of this type of BC over a private BC is that it relies on (and is managed by) multiple organizations instead of one. Users in a consortium can operate or run a node, conduct transactions, and audit the BC. The most popular consortium BC platform is Ripple.

**Hybrid BC:** This type of BC combines the advantages of other types of BCs to address their limitations and provides an effective solution for reliable data sharing, access management, etc. A hybrid BC is a DL that balances controlled access and freedom. The hybrid BC architecture is characterized by the fact that it is not open to everyone, but still offers BC features such as integrity, transparency, and security. The most popular hybrid BC platform is Dragonchain.

BlockCOV has two separate BCs: a hybrid BC for managing pandemic-related data and a private BC for the fair distribution of PPE. The participants in the BC network are referred to as nodes and are responsible for ensuring the accuracy and reliability of the data stored in the BC throughout the pandemic. In general, each BC has three types of nodes with distinct functions.

**Full nodes** are vital in decentralized BC networks because they store and verify the entire public data of a BC. They authenticate and store every block but only retain recent data. Full nodes support the network by validating blocks and verifying public blocks and states. By contrast, **light nodes** are designed for devices with limited capacity, such as embedded devices or mobile phones, and only store a block header that confirms the validity of previous transactions. The block header contains important information such as a timestamp and a unique number (nonce). However, **archive nodes** store the entire history of a BC, including all previous states. They have a snapshot of the BC ecosystem in each block since its creation. Table 3 presents the proposed entities, node types, and roles in the BlockCOV ecosystem. A schematic of the theoretical framework and its constituent elements is shown in Figure 5.





It is essential that all three types of nodes be included in the BlockCOV network. The MOHE in both BCs serves as an authority node, functioning as a master node that controls access and grants permissions to other nodes seeking access to the network. Additionally, MOHE can operate as a validator node that ratifies transactions and blocks within the network. As an authority node, MOHE is expected to store the complete transaction history of the network. MOHE is responsible for authenticating the nodes and acting as an authority node, which means that it simultaneously functions as a full and archive node. This node can quickly retrieve data from a database. Deciding the number of archive nodes required has some tradeoff between implementation cost and the feasible level of fraud and attacks, and is not covered in this study.

The second node type comprises participant nodes that upload data to the network. These nodes include primary healthcare providers, hospitals, and individuals, and they function as full nodes capable of sending and receiving transactions on the network and verifying the validity of transactions and blocks. Participant nodes also possess the ability to join and leave private subnetworks within the network. To access the network, these full nodes require a public key and private key combination. These nodes sign the information records they upload using their private keys, and BlockCOV uses this signature to confirm that the information is submitted by authorized nodes within the network. In addition to managing user accounts, MOHE assumes a supervisory role in reviewing records submitted by network members. If there are any issues with the data, MOHE investigates the associated members. During the pre-hospital stage, the screening data serve as the primary input to the BC network and can be transmitted to the network by both primary healthcare providers and individuals who have already been authenticated for online screening in the form of self-screening. At the in-hospital stage, private and public hospitals submit treatment process data to the BC network. The details of the required data are discussed in Section 3.3. Additionally, in the post-hospital stage, patients send their follow-up data after being discharged from hospitals, enabling tracing of their conditions. These follow-ups can be reminded by the smart contract-based notification of the BlockCOV.

The third type of node in the network is a light node. Light nodes are a type of node that are available for public use by individuals, research centers, and insurance companies. These nodes are designed for viewing transactions only upon request to the MOHE and synchronizing with the latest state of the network. Despite their limited functionality, light nodes are also able to participate in the transaction validation process within the network. Unlike full nodes, light nodes do not require authorization to access the network and can join without requesting permission. Their primary role is to verify the data blocks of other users, ensuring the integrity and accuracy of the information being shared. Insurance companies, research centers, and individuals are light nodes that do not require sending data to the BC network but can be responsible for validating transactions in the network.

For the fair distribution of PPE, we propose a private BC architecture characterized by a limited scope of transactions. Specifically, the only transactions permitted on this BC are those involving the transfer of amounts between consumers and distributors, as well as between distributors and manufacturers. This topic is further discussed in Section 3.4.3. For the COVID-19 management network, we propose a hybrid BC architecture that combines a public BC with a private BC to address the challenges mentioned in Table 1. We can store and verify both public and

private data on the network depending on the type and source of the data. For example, we can store and verify basic information such as patient name, national code, sex, and age on a private BC, as these data are sensitive and personal and should not be exposed to the public. Encryption can be used to protect data from unauthorized access and disclosure and hashing can be used to check their integrity and validity. Smart contracts can also be used to control the access and visibility of data by defining the rules and policies of the network. Only authorized participants such as the MOHE, primary health providers, and hospitals can access and decrypt the data. On the other hand, we can store and verify the symptoms, comorbidities, clinical data, dates and times of admission and discharge, and follow-up data on the public BC, as these data are useful and relevant to the public interest and can help with the screening, monitoring, and tracing of COVID-19. Digital signatures can be used to verify the identity and authenticity of data sources, and consensus protocols can be used to secure and verify transactions and data in the network. Anyone can access and validate these data, but only authenticated participants, such as primary health centers, hospitals, and patients, can send them to the network. Therefore, the hybrid BC is chosen because it provides transparency and trust for public health data and transactions, which can improve the awareness and compliance of people and authorities. It allows for the selective disclosure and access control of private health data and transactions, which can protect the privacy and confidentiality of patients and organizations.

It enables faster and cheaper transactions and data processing, as the private BC can reduce the network congestion and transaction fees of the public BC.

The proposed modules are designed and applied to provide suggestions and recommendations for use, and are comprised of three modules, as depicted in Figure 5(c). These modules include infection prediction, severity and LoS prediction, reinfection and readmission rate prediction. These modules are implemented using a ML model and trained using historical data available on the BC. The input data for these modules is sourced from three separate data chains (Figure 5e) namely, the screening, monitoring, and tracing chains, maintained by the participants in the network. In the screening chain, individuals who are suspected of having COVID-19 are screened and, if necessary, referred to a hospital for further investigation. In the monitoring chain, patients who require more care are hospitalized, monitored, and provided with



**Figure 5. Conceptual model of BlockCOV. (a): Hybrid BC for COVID-19 (b) Private BC for PPE (c) proposed modules (d) possible function types in the network (e) supported chains (f) proposed platform for implementation**

medical equipment, based on the severity of their illness and hospital LoS. In the tracing chain, patients who have been discharged from the hospital are followed-up. Stakeholders in the network use smart contracts to add or query

information records. These smart contracts can be deployed in the BlockCOV system and are compatible with the XDC EVM BC. The details of each chain are as follows.

### **3.2.2. Screening Chain**

The objective of screening is to identify, diagnose, isolate, and treat cases of COVID-19 in its early stages. Figure 6 illustrates the screening chain process, which consists of two stages: remote/online screening and clinical screening. Due to the rapid and extensive transmission of the pandemic virus at an early stage, it is recommended that initial remote screening be conducted. Individuals can use online systems, application programs, or dedicated call centers to undergo COVID-19 screening. Upon completion of the questionnaire, the individual's data, including demographics, symptoms, comorbidities, and history of close contact with infected individuals, are recorded in the system. Based on this information, if an individual is predicted to have COVID-19, their severity will be predicted. If they belong to the high-risk group, they will be referred to hospitals for clinical screening. Following the clinical procedure, the final diagnosis will be uploaded to the system. Depending on the patient's condition, they may be hospitalized, quarantined, or advised to rest at home. Hospitalized patients enter the second chain.

## **3.2.3. Monitoring Chain**

The monitoring chain comprises two types of entries: those referred from the previous chain, and those who go directly to the hospital. From a pandemic management perspective, the purpose of the monitoring chain is to predict severity and LoS to prepare appropriate treatment plans and resource allocation. The severity of COVID-19 is determined using clinical event data as criteria, as defined in Table 4 and depicted in Figure 7. This process begins with the date and time of admission to the hospital until discharge or IHM. Discharged patients then enter the third chain of the BlockCOV system.

## **3.2.4. Tracing Chain**

Individuals discharged from the previous chain enter this chain. Figure 8 presents the possible posthospital outcomes for patients who are treated for COVID-19, which can be categorized into four states: out-of-hospital mortality (OHM), reinfection  $(A \rightarrow C \rightarrow D)$ , readmission due to incomplete recovery (path  $A\rightarrow B\rightarrow F$ ), and readmission after complete recovery  $(A\rightarrow C\rightarrow D\rightarrow E\rightarrow F)$ . We focused on readmission and reinfection, as these are crucial factors to be considered in pandemic management. These outcomes can influence the effectiveness and costeffectiveness of public health interventions and

policies aimed at reducing the burden and severity of COVID-19 [116]. The risk of reinfection and readmission due to COVID-19 poses a significant challenge to the healthcare system, as it can lead to an increased demand for critical resources such as hospital beds, ventilators, and personal protective equipment. This strain on the healthcare system can compromise the quality of care, increase the risk of virus transmission, and make it more difficult to implement effective containment and mitigation strategies such as testing, tracing, isolation, and vaccination. Therefore, it is essential to closely trace and report these outcomes, and utilize data and evidence to inform and improve COVID-19 prevention and management policies and strategies.





## **3.3. Block Structure Information**

One of the emerging BC platforms, XDC, which has gained significant recognition [117], serves as the system design framework in our study. As a pioneering hybrid BC, XDC facilitates fast, secure, and cost-effective transactions to establish hybrid relay bridges and attain spontaneous block finality, thereby enhancing transaction security and fostering transparency among stakeholders. XDC employs a consensus algorithm known as DPoS, which is a delegated proof-of-stake mechanism that designates network validators by means of coin-holders delegating their votes. XDC stands

out for its exceptional energy efficiency, surpassing the energy consumption of PoW mining by a factor of ten and improving upon the energy efficiency of PoS mechanisms [118]. In contemporary BC systems, the duty of constructing blocks has been delegated to the chosen platform, leaving aside the role of miners as previously observed in older BC architectures. This research focuses solely on block information. The blocks in our proposed architecture primarily comprise three types of transactions, which are loaded by primary healthcare, individuals, or hospitals. Figures 9 and 10 depict the arrangement of data fields and structures within block transactions. Each transaction contains five fundamental data elements, including the time stamp, sender,

recipient, amount, and content of data records. The sender, one of the three primary entities in the BC network, is identified through the timestamp and the transaction time. At present, the receiver and amount fields are set to NULL, and the records are stored in the "data" field. The records are organized in a hashed table structure, similar to a dictionary that contains unique keys and their corresponding values. The keys to these hash tables represent various classes and records, such as screening, hospital monitoring, and follow-up. The values of these hash tables are objects that consist of record information and other nested objects. In a hash table format, the records can be stored in an unformatted manner within the BC.



**Figure 6. Screening chain process**



**Figure 8. Reinfection and Readmission Pathway**













**Figure 10. Block information structure**

## **3.4. Smart Contracts in BlockCOV**

Smart contracts, which are self-executable codes deployed on a BC, automatically execute when triggered [119]. These contracts can be used to build trust in a no-trust contractual environment and enforce terms and conditions when certain conditions are met [120]. Proposed for the administration of the COVID-19 pandemic on the XDC platform-based BC is the use of smart contracts to further implement the monitoring function [121]. Smart contracts offer the advantage of executing exactly as planned, without the possibility of failure, censorship, fraud, or thirdparty interference. The XDC platform facilitates the creation and deployment of decentralized apps (dApps) with a runtime that can execute smart contract code, making it simple for developers. Our COVID-19 pandemic management system on the BC can autonomously monitor system status and provide inquiry features by developing suitable smart contracts.

## **3.4.1 COVID-19 Inquiry**

The process of obtaining information regarding individuals' infection records in BlockCOV is carried out through the utilization of query functions and smart contracts, as shown in Figure 11(a), and can be accomplished in two ways. If the individual seeking the information is the patient, the query will display comprehensive data such as the date and name of screening, hospitalization center, and post-discharge follow-up data. However, if the inquiry is being made by a light node, such as a transportation ticket issuing center, the query will only display the individual's most recent infection status and the date of the infection.

## **3.4.2 COVID-19 Follow-up**

During the pandemic stages, pre- and post-hospital notification systems can assist in pandemic agile intervention. At the pre-hospital stage, some individuals have suspicious symptoms, but they are not severe enough to visit the screening or healthcare centers. Sending notification messages to invite these individuals to perform online screening is useful. These discriminations are made either through online self-screening or through healthcare call centers. The second place that needs notification is the follow-up process after patient discharge from the hospital, where the patient's condition must be monitored. Regular short-term symptomatic monitoring contributes to patient recovery. In addition, the possibility of reinfection with COVID-19 can be monitored. This notification system can be properly managed by

utilizing the BlockCOV described in this study, combined with smart contracts. MOHE must execute a *detect ()* function called by the web middleware, as shown in Figure 11(b). This function is used to query the list of people that need to be considered, as mentioned above. A notification is broadcast to the network and its status is updated by the follow-up function of F/U\_Notify *()*. If after a certain period of time follow-up is not done by the person to whom the notification is sent, the call center function is executed, which is followed up by phone and then publishes the information on the network using the add function.

## **3.4.3. BlockCoin**

The pandemic's significant obstacle was PPE's increased demand, causing severe shortages, primarily affecting frontline workers and medical staff. Various factors, such as stockpiling, panic buying, and misuse, had a profound impact on the supply chain, putting numerous lives in jeopardy [122]. Typically, purchasers of PPE such as hospitals acquire these products from distributors, who in turn pay manufacturers in cash for the PPE. The utilization of a virtual currency is indispensable for the effective facilitation of value exchanges within the BlockCOV ecosystem. The XRC20 token standard delineates the prerequisites for implementing token contracts on XDC networks, encompassing the requisite functions and events. Figure 12 (a) depicts the functions and events of XRC20 token. Traditional transactions involve the exchange of cash among manufacturers, distributors, and hospitals. However, the introduction of the XRC20 interface presents an alternative method of conducting payments through the transfer of token balances. Our proposed design envisions the utilization of XDC coins to replace cash transactions between manufacturers and distributors, as well as between the distributor and the hospital. The XDC coin is always set at a 1:1 conversion ratio with legal tender, ensuring that its value is equivalent to that of legal money. The process begins with consumers logging into the web middleware (Figure 12(b)) and purchasing the necessary PPE from the distributor.

Upon the completion of a successful sale, the BlockCOV system prompts the execution of the *pay()* function to finalize the transaction. The *pay()* function initially calls the *TransferFrom()* function to transfer the BlockCoin from the consumer's address to the distributor's address, with the amount of BlockCoin being equivalent to the selling price of the PPE to the end user (e.g., 20 BlockCoins). This is followed by a second invocation of the *TransferFrom()* function, which transfers the BlockCoin from the distributor's address to the manufacturer's address. In this instance, the amount of BlockCoin transferred is the selling price less the distributor's income (e.g., 3 BlockCoins), and the manufacturer's income is the remaining XDC coins (e.g., 17 BlockCoins).

## **4. Evaluation Results**

BlockCOV incorporates three intelligent suggestion modules: infection prediction, severity and LoS prediction, and reinfection and readmission predictions. The development of this intelligent prediction system was guided by CRISP-DM methodology [123], which took 19 months because of the difficulty in accessing government data. This highlights the challenges faced when developing data-driven intelligent systems, particularly when working with sensitive

and confidential data. Three integrated datasets are required to implement proposed modules: screening dataset, hospital admission dataset, and reinfection/readmission tracing dataset for discharged patients. Data were collected from 117 public and private hospitals in Iran between February 1, 2020, and September 30, 2020. Figure 13 shows the flow of raw data and its preprocessing across different pandemic stages. From a total of 201,911 rows of raw data obtained, 66,143 clean records were entered into the infection prediction module. Among 18,740 patients admitted to the hospital, only 8,639 rows of data provided suitable conditions for use in the second module. Unfortunately, post-hospital tracing data are not systematically available for HIS in Iran. The model behind the third module is explained in Section 3.2.3. However, we were unable to implement this part because of a lack of data. The cleaned datasets applied in state-of-the-art classifiers were selected based on a literature review. The hyperparameters of these classifiers were tuned using a grid search and the models were built using Python. The accuracies of the models were compared, and the most important features at each stage were determined using SHapely Adaptive explanations.



**Figure 11. (a) Flow of COVID-19 infection query (b) Flow of COVID-19 follow up notification**



**Figure 12. (a) XRC20 token function and events (b) PPE payment by BlockCoin**

**4.1. Infection Prediction Module** In this module, our goal is to achieve an accurate prediction of COVID-19 infection based on nonclinical data using ML classification algorithms. We conducted a comprehensive review of the literature, both generally and in the context of COVID-19, and selected 11 state-of-the-art classification algorithms for the analysis. The dataset used in this study consisted of 66,143 records, including one target variable and 38 input features such as basic information (6), symptoms (17), semi-clinical data (1), and comorbidities (14). A detailed description of these features and the dataset characteristics is provided in Appendix A1. To fine-tune hyperparameters of the classifiers, we used a Grid Search. The hyperparameter names and tuned values are given in Appendix A2. We applied the classifiers to the dataset using Python software and evaluated their accuracy scores through 10 fold cross-validation. The results, which can be found in Table 5, indicate that XGBoost achieved the highest accuracy of 73.94%.

Figure 14 shows the absolute SHAP values for the top 20 factors that influence the outcome. In respond the RQ3, the SHAP values suggest that "*contact with infected people*" is the most critical factor in predicting COVID-19 infection, with a substantial impact on the model output. The second and third most important features are "*cough*" and "*age*," respectively, which also have moderate impacts on the model output. These features reflect the common symptoms and risk factors of COVID-19, as cough is a frequent sign of respiratory infection, and older age is associated with higher mortality and severity of COVID-19. Other features with weak impacts on the model output include "*muscle pain*," "*fever*," "*cardiovascular disease*," "*PO2*," "*respiratory disease*," "*cancer*," "*male*," "*diabetes*," "*kidney disorders*," "*anorexia*," "*headache*," "*stomach*," "*loss of consciousness*," "*nausea*," "*diarrhea*," and "*drug addiction*." These features represent either less common or mild symptoms of COVID-19 or underlying comorbidities that may increase susceptibility or complications of COVID-19. On the other hand, the sum of 21 other features has a weak impact on the model output, indicating that these features are either irrelevant or protective against COVID-19 infection. These features may include demographics, symptoms, and comorbidities that are not directly associated with COVID-19.

## **4.2. Severity and LoS Prediction Module**

The infected people can be referred to hospitals for further measures. If a patient's condition is such that they need to be hospitalized, there must be a mechanism to predict the severity and LoS of the disease. The data collected from BlockCOV can enable the prediction of LoS and the risk level of severity by providing the date and time of start and end of receiving services in the hospital.

In this section, we present the results of our experiments for predicting the severity and LoS of COVID-19 using 11 state-of-the-art machinelearning algorithms. In this module, only hospitalized patients were included, and their infection with COVID-19 was concluded based on the final diagnosis at the end of the treatment based on clinical investigation. We used a cleaned dataset of 8,639 patients with various symptoms and comorbidities and classified them into five categories: UOB (41.38%), GU (36.75%), ICU (0.59%), INT (1.41%), and DD (19.86%). A detailed description of these features and the dataset characteristics is provided in Appendix A3. To fine-tune hyperparameters of the classifiers, we used a Grid Search. The hyperparameter names and tuned values are given in Appendix A4.

We applied the ML algorithms to train and test our models on oversampled data and used a 10-fold cross-validation technique to evaluate the performance of each algorithm. The results of the cross-validation are presented in Table 6.

As shown in Table 6, the XGBoost algorithm achieved a high mean accuracy performance of 74.81% among the other algorithms and has superior performance over the other methods. From the above results, we can conclude that the XGBoost algorithm is the most effective and robust method for predicting the severity of COVID-19 using the selected features. These methods can provide valuable insights and guidance for clinical decision-making and the management of the pandemic.

Figure 15 shows a chart of the top 20 most important features based on SHAP values towards outputs for predicting COVID-19 severity in hospitalized patients. These features include the various symptoms, risk factors, and comorbidities of COVID-19. This reveals that some features are more important for predicting certain outputs than others are. The top six features are *"Age," "Respiratory distress,"* "*contact with infected people", "PO2," "Cough," "Fever"*. In response to RQ4, these features are deemed to be the most relevant for predicting the severity of COVID-19 in hospitalized patients. We also observed that some features had different impacts on different classes. For example, "*Age*," "*Respiratory distress*," and "*diabetes*" are the most important features for predicting the most severe outcomes (ICU, INT, and DD), whereas *PO2*, "*contact with infected people*," and "*fever"* are the most important features for predicting less severe

outcomes (UOB and GN). Some features, such as "*cough*," "*muscle pain*," and "*sex*," are important for predicting all classes but with different degrees of impact. Our analysis revealed the complex relationship between the features and classes and provided insights into the factors that influence COVID-19 severity in hospitalized patients.

From a hospital administration perspective, LoS negatively impacts bed turnover rates, impedes the efficient flow of patients, and results in bed shortages, which complicates both the strategic and operational management of the hospital [124]. As illustrated in Figure 13, 8,639 cleaned IDs were utilized to predict the LoS. The LoS distributions are detailed in Appendix A5. The LoS was divided into 13 classes and XGBoost model was trained and tested using a 10-fold cross-validation, and achieved an accuracy of 77.31%. Table 7 presents

the performance of each LoS class. The performance of each LoS class was evaluated using precision, recall, f1-score, and support metrics. The results showed that the XGBoost model performed well for most classes, especially for the extreme classes  $(0, 1, 2, 14-17, and 18+)$ , which had high precision and recall values. In response to RQ5, figure 16 shows a chart of the top 10 most important features based on SHAP values towards outputs for predicting LoS.



**Figure 13. The flow of raw and cleaned data used in proposed intelligent modules**





**Table 6. Accuracy score (%) of classifiers using 10-fold cross-validation.**





**Figure 16. Graph of XGBoost SHAP feature importance.**

## **Table 7. XGBoost classification performance**



## **5. Discussion**

## **5.1. Main Findings**

Prior research on pandemic management systems can typically be assessed in terms of data privacy, security, and efficiency in data sharing or the purpose and roles of these systems in assisting with pandemic management. For instance, a study by [125] utilized AI and BCT to focus on contact tracing as a crucial aspect of pandemic control. Their proposed system comprised three integrated modules that addressed contact tracing, COVID-19 status recording to serve as a digital pass, and the prevention of infected and suspected users from accessing public places. Additionally, their system categorized zones based on infection levels to help users avoid highly contaminated areas. While contact tracing has proven to be an efficient strategy in fighting COVID-19, quickly screening individuals in the early stages of an outbreak is crucial to preventing the overload of treatment systems. They used BC as a secure and transparent platform for storing and sharing the data related to the pandemic, such as the infection status, the contact history, the public places access.

The significance of appropriate modeling as a determining factor in pandemic decision-making is widely recognized, yet there remains a dearth of pertinent research in this domain. However, our proposed system addresses this issue by proposing non-clinical infection prediction without the need for individuals to visit a screening center or rely on in-house data. While our model's accuracy may not rival those relying on images to detect COVID-19, it is worth noting that systems utilizing clinical data are often characterized by their lack of agility, high costs, and limited practical implications during the early stages of a pandemic. For instance, in [126], deep learning was employed to develop a novel architecture for identifying the virus in radiological images, achieving an accuracy of 96% in classifying COVID-19 images.

The spread of misinformation is a significant problem that has detrimental impacts on individuals, public health, and governments worldwide [127]. Caceres et al. [128] highlighted the challenges of misinformation in the era of social media and its potential to undermine efforts to address public health crises. BCT has the potential to store and verify data in a decentralized and distributed manner, without relying on a central authority or intermediary. By incorporating BC into the COVID-19 pandemic management network, it may be possible to prevent misinformation by providing secure and transparent means of tracking and sharing data related to various stages of the pandemic. This

technology also offers a collaborative and participatory platform for MOHE, hospitals, patients, individuals, research centers, and other stakeholders involved in pandemic management, including health workers, policymakers, researchers, and the public. This can lead to improved communication and coordination among stakeholders and sectors as well as increased transparency and accountability in decisionmaking processes. In addition, it encourages the creation and dissemination of accurate and reliable information that adheres to community-driven standards and norms for pandemic management.

When discussing pandemic management, it is essential to consider three distinct stages. Table 8 presents a comparison of related works. Nine studies were selected and compared across several aspects, including the pandemic stage covered, dataset type and size, intelligent system proposition, agility method usage and performance indicators. Although numerous papers are related to our study, we selected a representative sample to illustrate how our work differs from others.

Contrary to our work, most studies focus on a single stage. For example, in the in-hospital stage, a recent review study [55] identified 314 eligible articles. Of these, 152 (48.4%) presented *mortality* as the outcome, 66 (21.0%) focused on severity and/or critical illness, 35 (11.1%) combined *ICU admission* and *mortality*, 17 (5.4%) assessed *ICU admission* only, 6 (1.9%) examined *mechanical ventilation* only, and 38 (12.1%) assessed *multiple combined outcomes*.

According to Table 8, Studies [33] and [44] focused solely on the pre-hospital stage, with datasets containing 4,434 and 5,434 records, respectively, compared to our dataset of 66,143 records. Variations in data collection methods and reporting standards across regions further complicate the comparison and aggregation of COVID-19 data. Additionally, the limited availability of large datasets hinders comprehensive analysis and modeling of the pandemic. While there is no consensus in the literature regarding a sufficient number of samples [34], small-sample studies are more susceptible to minor analytical errors, resulting in false-negative results [35, 36]. Researchers are advised to conduct large-scale studies to produce statistically realistic effects due to their higher statistical power. Results from large studies are statistically more reliable than those from small studies because of the reduced risk of increased effect size and lower Type I error [37].

Our literature review categorizes self-reportable input features into three main groups: basic information (demographic data), symptoms, and past or current diseases. The extent to which these features are covered in each study is a key difference between this and previous studies. For example, [33] used only one demographic variable, whereas our work utilized six basic variables. In another comparison, [33] considered only one past/current disease, while our model included fourteen items.

Studies [129] and [63] focused only on the inhospital stage to predict the severity of mortality, ICU admission, and mortality, respectively. In contrast, our model categorizes in-hospital risk into five levels: UOB, GU, ICU, INT, and IHM. Despite the uncertainties in COVID-19 data, our model achieved suitable performance indicators (74.81% accuracy in the pre-hospital stage and 77.3% accuracy in the in-hospital stage) relative to studies using the same dataset.

The second part of our comparison focuses on studies that propose intelligent systems for managing COVID-19. Although the number of such papers is limited, three notable instances are presented in Table 8. Research on intelligent systems often concentrates on specific stages of pandemic management. For instance, studies [90, 91] have introduced a system design that BC, AI, and drones to control the spread of COVID-19. This system leverages the advantages of BC, such as enhanced security, transparency, and decentralization, to ensure the reliability and integrity of data collected by drones. AI is incorporated to equip drones with capabilities such as image processing, face recognition, and object detection. Additionally, AI is utilized to analyze the data and provide real-time feedback and guidance to both authorities and the public.

Another study [92] proposed a communication framework that combines BC and AI to enable multiswarm drones to address COVID-19 scenarios. Study [93] introduced a smart healthcare system that integrates BC and AI to monitor and detect COVID-19 in biomedical images. This system facilitates self-testing, diagnosis, and data sharing by employing deep learning models to analyze chest X-ray images and classify them as either COVID-19 positive or negative. Furthermore, the system utilizes blockchain technology to store and verify diagnostic results and patient information.

		<b>Scope</b>					
	Ref. Year	Pre-h.	In-h.	Pos-h.	<b>Propose</b> system	<b>Agile</b> method	<b>Description</b>
$[130]$	2021	×	×	✓	No	No	Only focused on post-hospital stage of pandemic
[131]	2021	×	×	✓	No	No	Only focused on post-hospital stage of pandemic
[129]	2022	×	✓	×	No	N <sub>o</sub>	AUC=0.975. Only mortality and ICU admission severity level are considered. Clinical data are used that are not agile approach.
[63]	2023	×	✓	×	No	No	AUC= 0.84. Only Mortality severity level and LoS are considered. Small dataset (1,291 records).
$[44]$	2022	✓	×	×	N <sub>o</sub>	Yes	$AUC = 0.98$ . Unreal dataset, small dataset $(5,434)$ , Only one demographics, 8 symptoms, 5 past/current disease are considered.
$[33]$	2022	✓	×	×	No	Yes	$AUC = 0.65$ . Small dataset (4,434 records), only one Past/current disease variable is considered.
[90]	2021	✓	×	×	Yes	N <sub>o</sub>	AI is integrated to provide drones with image processing, face recognition, and object detection capabilities.
[91]	2021	✓	×	×	Yes	No	It presented a communication scheme that leverages BC and AI to enable multiswarm drones to address COVID-19 situations
[93]	2021	✓	×	×	Yes	No	It proposed a smart healthcare system that integrates BC and AI to monitor and detect COVID-19 in biomedical images
This study	2024			✓	Yes	Yes	Propose intelligent system to cover all stages. Used large dataset $(66, 143$ records) includes variables: basic information $(6)$ , symptoms $(17)$ , semi-clinical data $(1)$ , and comorbidities $(14)$ . Performance indicators resulted in 74.81% accuracy in prehospital stage and 77.3% in-hospital stage accuracy.

**Table 8. Comparison of related COVID-19 research papers**

## **5.2. Implications for Research and Practice**

The proposed BlockCOV has implications for research in healthcare, information systems, and

Artificial Intelligence. First, we did not delve into some technical BC designs, such as determining the TPS, on-chain vs. off-chain data strategy, its

consideration, and transaction costs. Therefore, before developing and deploying the proposed system, technical design parameters should be studied to validate and evaluate the feasibility, functionality, and performance of the proposed system in a simulated or real-world setting, and to identify and resolve any potential issues or limitations. Second, we propose an XDC platform for deploying the proposed system. The XDC platform offers robust and efficient infrastructure for deploying the proposed system. Its advanced features and scalability make it ideal. Different decision support systems such as MCDM and AHP can be used to define the selection methodology among multiple platforms. Third, to assess the effectiveness of the proposed system compared to other existing or alternative solutions for pandemic management, such as centralized or distributed databases, cloud computing, or other BC platforms, a systematic literature review, meta-analysis, or benchmarking study must be conducted. This involves evaluating the strengths and weaknesses of each solution, and identifying the best practices and lessons learned from each approach. Fourth, the system can be extended to address other types of pandemics, such as influenza, Ebola, and Zika, and other public health emergencies, such as natural disasters, bioterrorism, and environmental pollution. This would require adapting the system's design and configuration to suit the specific characteristics, challenges, and needs of each scenario and incorporating the latest scientific and technological developments in the relevant domains.

The proposed BlockCOV has several practical implications. First, adapting the system to the changing dynamics and needs of the pandemic, such as new variants, vaccines, policies, or behaviors. This requires updating and fine-tuning the system's parameters, models, and algorithms to reflect the latest data and evidence and to provide accurate and timely predictions and recommendations. Second is the impact and outcomes of the system on pandemic management and public health, such as the infection rate, mortality rate, recovery rate, resource utilization, cost-effectiveness, and quality of life. This would require conducting longitudinal and comparative studies using appropriate indicators and metrics to measure and analyze the system's performance and value, and to identify the best practices and areas for improvement.

## **5.3. Conclusions**

COVID-19 has brought to light the need for a reenvisioned pandemic management framework that emphasizes data-driven systems in robust technological infrastructure. BCT holds great promise as a means of overcoming the challenge of insufficient valid and globally shared datasets in public health, allowing for informed decisionmaking by public health leaders rather than leaving such decisions to politicians, as has been observed during the pandemic. Our study leveraged AI and BC to address the main challenges of pandemic management across different stages. We designed and configured an intelligent system to tackle challenges in the pre-hospital, in-hospital, and post-hospital stages from a patient-centric perspective. BlockCOV also addressed the issue of limited PPE distribution, ensuring fair distribution through smart contracts. Our proposed model demonstrates how early-stage screening requirements, infection severity prediction, prediction of length of hospital stay, reinfection probability, and readmission can be modeled. The smart contract concept also shows how virtual coins can be used for fair distribution of PPE during a pandemic. Experimental results for early infection, LoS, and severity prediction can be achieved using AI, providing valuable recommendations for stakeholders, particularly policymakers.

Although the COVID-19 pandemic may have nearly ended, this research remains valuable for several reasons. First, it contributes to the advancement of knowledge and innovation in healthcare, information systems, and AI by proposing a novel and comprehensive system that combines BC and AI to address the challenges of pandemic management. Second, this research provides a practical solution that can be implemented and adapted to different contexts and scenarios, such as other types of pandemics, public health emergencies, or routine healthcare services, to improve the quality and efficiency of the healthcare sector and public health outcomes. Third, the research was implemented in the context of the COVID-19 case in Iran, which can be beneficial for other countries and regions facing similar or different problems and opportunities in dealing with pandemics and other health issues. Fourth, this research discusses the implications of the proposed system and suggests future directions for improvement and further research, which can inspire and guide other researchers and practitioners interested in this topic.

#### **Acknowledgements**

This research is humbly dedicated to the memory of our esteemed professor, Adel Azar, who tragically passed away at the outset of our endeavors. May his legacy and teachings endure in the hearts and



minds of all who had the privilege of learning from him.

#### **Note**

This research work was endorsed by th[e](https://ethics.research.ac.ir/ProposalCertificateEn.php?id=229801&Print=true&NoPrintHeader=true&NoPrintFooter=true&NoPrintPageBorder=true&LetterPrint=true) [Research](https://ethics.research.ac.ir/ProposalCertificateEn.php?id=229801&Print=true&NoPrintHeader=true&NoPrintFooter=true&NoPrintPageBorder=true&LetterPrint=true)  [Ethics Committees of Tarbiat Modares University,](https://ethics.research.ac.ir/ProposalCertificateEn.php?id=229801&Print=true&NoPrintHeader=true&NoPrintFooter=true&NoPrintPageBorder=true&LetterPrint=true) Tehran, Iran.

### **References**

[1] Ammirato, S., R. Linzalone, and A.M. Felicetti, Knowledge management in pandemics. A critical literature review. *Knowledge Management Research & Practice*, vol.19, no. 4, pp. 415-426, 2021.

[2] Cavallo, A. and V. Ireland, Prepared for the Global Assessment Report on Disaster Risk Reduction 2015. *International Journal of Disaster Risk Reduction*, 2014.

[3] Adiga, A., et al., Data-Driven Modeling for Different Stages of Pandemic Response. *Journal of the Indian Institute of Science*, vol. 100, no. 4, pp. 901-915, 2020.

[4] WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020 [cited 2023 June. 1, 2023].

[5] Zhai, Y., et al. From isolation to coordination: how can telemedicine help combat the COVID-19 outbreak. in Digital Health and Medical Analytics: *Second International Conference, DHA 2020, Beijing, China,* Revised Selected Papers 2. 2021. Springer, July 25, 2020.

[6] Looi, M.-K., Covid-19: Is a second wave hitting Europe? bmj, 2020. 371.

[7] Abebe, E.C., et al., The newly emerged COVID-19 disease: a systemic review. *Virology journal*, vol. 17, no. 1, pp. 1-8, 2020.

[8] Muhammad, G. and M.S. Hossain, COVID-19 and non-COVID-19 classification using multi-layers fusion from lung ultrasound images. *Information Fusion*, vol. 72, pp. 80-88, 2021.

[9] Dos Santos, L.A., et al., Recurrent COVID-19 including evidence of reinfection and enhanced severity in thirty Brazilian healthcare workers. *Journal of Infection*, vol. 82, no. 3, pp. 399-406, 2021.

[10] Li, Y., et al., Stability issues of RT‐PCR testing of SARS-CoV-2 for hospitalized patients clinically diagnosed with COVID‐19. *Journal of medical virology*, vol. 92, no. 7, pp. 903-908, 2020.

[11] Wang, S., et al., A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *European Respiratory Journal*, vol. 56, no. 2, 2020.

[12] Banerjee, A., et al., Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population. *International immunopharmacology*, vol. 86, pp. 106705, 2020.

[13] Arpaci, I., et al., Predicting the COVID-19 infection with fourteen clinical features using machine learning classification algorithms. *Multimedia Tools and Applications*, vol. 80, pp. 11943-11957, 2021.

[14] Grannec, F., et al., Pre-Hospital Management of Patients with COVID-19 and the Impact on Hospitalization. *Medicina 2023*, vol. 59, pp. 1440, 2023.

[15] Son, K.-B., T.-j. Lee, and S.-s. Hwang, Disease severity classification and COVID-19 outcomes, Republic of Korea. *Bulletin of the World Health Organization*, vol. 99, no. 1, pp. 62, 2021.

[16] Breathnach, A.S., et al., Prior COVID-19 significantly reduces the risk of subsequent infection, but reinfections are seen after eight months*. The Journal of Infection*, vol. 82, no.4, pp. 11, 2021.

[17] Gupta, V., et al., A symptomatic Reinfection in 2 Healthcare Workers from India with Genetically Distinct Severe Acute Respiratory Syndrome Coronavirus 2. *Clinical Infectious Diseases,* vol.73, no.9, 2020.

[18] Nagesh, S. and S. Chakraborty, Saving the frontline health workforce amidst the COVID-19 crisis: challenges and recommendations. *Journal of global health*, vol. 10, no.1, 2020.

[19] Zhu, P., et al., Enhancing Traceability of Infectious Diseases: A Blockchain-Based Approach. *Information Processing and Management*, vol. 58, no. 4, 2021.

[20] Pandey, P. and R. Litoriya, Securing and authenticating healthcare records through blockchain technology. *Cryptologia*, Vol. 44, no. 4, pp. 341-356, 2020.

[21] Ismail, L., H. Materwala, and S. Zeadally, Lightweight blockchain for healthcare. *IEEE Access*, vol. 7, pp. 149935-149951, 2019.

[22] Lau, H., et al., Internationally lost COVID-19 cases. *Journal of Microbiology, Immunology and Infection*, vol. 53, no. 3, pp. 454-458, 2020.

[23] Gao, F., et al., Management and Data Sharing of COVID-19 Pandemic Information. *Biopreservation and Biobanking*, vol. 18, no. 6, pp. 570-580, 2020.

[24] Rasheed, J., et al., COVID-19 in the Age of Artificial Intelligence: A Comprehensive Review. Interdisciplinary Sciences: *Computational Life Sciences*, vol. 13, pp. 1-23, 2021.

[25] WHO. Coronavirus Disease 2019 (COVID-19): Situation report—30. 2020.

[26] Coppola, L., et al., Biobanking in health care: evolution and future directions. *Journal of translational medicine*, vol. 17, no.1, pp. 1-18, 2019.

[27] Yaqoob, I., et al., Blockchain for healthcare data management: opportunities, challenges, and future recommendations. *Neural Computing and Applications, 2020*: pp. 1-16.

[28] Spreeuwenberg, P., M. Kroneman, and J. Paget, Reassessing the global mortality burden of the 1918 influenza pandemic. *American journal of epidemiology*, vol. 187, no. 12, pp. 2561-2567, 2018.

[29] Al-Emran, M., Hierarchical reinforcement learning: a survey. *International journal of computing and digital systems*, vol. 4, no. 02, 2015.

[30] Arpaci, I., et al., Analysis of twitter data using evolutionary clustering during the COVID-19 pandemic. *Computers, Materials & Continua*, vol. 65, no. 1, pp. 193-204, 2020.

[31] Mhamdi, C., M. Al-Emran, and S.A. Salloum, Text mining and analytics: A case study from news channels posts on Facebook, in Intelligent Natural Language Processing: Trends and Applications. pp. 399-415. *Springer*, 2018.

[32] Zaza, S. and M. Al-Emran. Mining and exploration of credit cards data in UAE. in 2015, *Fifth International Conference on e-Learning (econf)*, IEEE, 2015.

[33] Bian, J. and F. Modave, The rapid growth of intelligent systems in health and health care. 2020, *SAGE Publications Sage UK: London, England*. pp. 5- 7.

[34] Nguyen, T. T., Nguyen, Q. V. H., Nguyen, D. T., Yang, S., Eklund, P. W., Huynh-The, T., ... & Hsu, E. B. (2020). Artificial intelligence in the battle against coronavirus (COVID-19): a survey and future research directions. *arXiv preprint arXiv:2008.07343.*

[35] Shabrandi, A., et al., Fast COVID-19 Infection Prediction with In-House Data Using Machine Learning Classification Algorithms: A Case Study of Iran. *Journal of AI and Data Mining*, pp. 573-585, 2024.

[36] Aslan, M.F., et al., CNN-based transfer learning– BiLSTM network: A novel approach for COVID-19 infection detection. *Applied Soft Computing*, vol. 98 pp. 106912, 2021.

[37] Wang, B., et al., AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system. *Applied Soft Computing*, vol. 98, pp. 106897, 2021.

[38] Chandra, T.B., et al., Coronavirus disease (COVID-19) detection in Chest X-Ray images using majority voting-based classifier ensemble. *Expert systems with applications*, vol. 165, pp. 113909, 2021.

[39] Lassau, N., et al., Integrating deep learning CTscan model, biological and clinical variables to predict severity of COVID-19 patients. *Nature communications*, vol. 12, no.1, pp. 1-11, 2021.

[40] Afrash, M.R., et al., Predicting hospital readmission risk in patients with COVID-19: a machine learning approach. *Informatics in medicine unlocked*, vol. 30, pp. 100908, 2022.

[41] Long, J.B. and J.M. Ehrenfeld, The role of augmented intelligence (AI) in detecting and preventing the spread of novel coronavirus. 2020, Springer.

[42] Gozes, O., et al., Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis*. arXiv:2003.05037*, 2020.

[43] Jin, C., et al., Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. *Nature communications*, vol.11, no. 1, pp. 1-14, 2020.

[44] Punn, N.S. and S. Agarwal, Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks. *Applied Intelligence*, vol. 51, no. 5, pp. 2689-2702, 2021.

[45] Song, Y., et al., Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 18, no. 6, pp. 2775- 2780, 2021.

[46] Wang, S., et al., A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *European radiology*, vol. 31, pp. 1-9, 2021.

[47] Tostmann, A., et al., Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020. *Eurosurveillance*, vol. 25, no. 16, pp. 2000508, 2020.

[48] Punn, N.S., S.K. Sonbhadra, and S. Agarwal, COVID-19 epidemic analysis using machine learning and deep learning algorithms. *MedRxiv*, 2020.

[49] Feng, C., et al., A novel triage tool of artificial intelligence assisted diagnosis aid system for suspected COVID-19 pneumonia in fever clinics. *MedRxiv*, pp. 2020.03. 19.20039099, 2021.

[50] Mei, X., et al., Artificial intelligence–enabled rapid diagnosis of patients with COVID-19. *Nature medicine*, vol. 26, no. 8, pp. 1224-1228, 2020.

[51] Shaverdian, N., et al., Need for caution in the diagnosis of radiation pneumonitis during the covid-19 pandemic. *Advances in radiation oncology*, vol. 5, no. 4, pp. 617-620, 2020.

[52] Guhathakurata, S., et al., A novel approach to predict COVID-19 using support vector machine, in Data Science for COVID-19. *Elsevier*. pp. 351-364, 2021.

[53] Guhathakurata, S., et al., A new approach to predict COVID-19 using artificial neural networks, in Cyberphysical systems, *Elsevier*, pp. 139-160, 2022.

[54] Malik, M., et al., Determination of COVID-19 Patients Using Machine Learning Algorithms. *Intelligent Automation & Soft Computing*, vol. 31, no. 1, 2022.

[55] Buttia, C., et al., Prognostic models in COVID-19 infection that predict severity: a systematic review. *European journal of epidemiology*, vol. 38, no. 4, pp. 355-372, 2023.

[56] Liu, L., et al., Predicting COVID-19 Severity: Challenges in Reproducibility and Deployment of Machine Learning Methods. *International Journal of Medical Informatics*, pp. 105210, 2023.

[57] Mahboub, B., et al., Prediction of COVID-19 hospital length of stay and risk of death using artificial intelligence-based modeling. *Frontiers in medicine*, vol. 8, pp. 592336, 2021.

[58] Ebinger, J., et al., A machine learning algorithm predicts duration of hospitalization in COVID-19 patients. *Intelligence-based medicine*, vol. 5, pp. 100035, 2021.

[59] Alabbad, D.A., et al., Machine learning model for predicting the length of stay in the intensive care unit for COVID-19 patients in the eastern province of Saudi Arabia. *Informatics in Medicine Unlocked*, vol. 30, pp. 100937, 2022.

[60] Afrash, M., et al., Predictive modeling of hospital length of stay in COVID-19 patients using machine learning algorithms. *Journal of Medicinal and Chemical Sciences*, vol. 4, no. 5, pp. 525-537, 2021.

[61] Jutzeler, C.R., et al., Comorbidities, clinical signs and symptoms, laboratory findings, imaging features, treatment strategies, and outcomes in adult and pediatric patients with COVID-19: A systematic review and meta-analysis. *Travel medicine and infectious disease*, vol. 37, pp. 101825, 2020.

[62] Saadatmand, S., et al., Using machine learning in prediction of ICU admission, mortality, and length of stay in the early stage of admission of COVID-19 patients. *Annals of Operations Research*, vol. 328, no. 1, pp. 1043-1071, 2023.

[63] Amiri, P., et al., Prediction of mortality risk and duration of hospitalization of COVID-19 patients with chronic comorbidities based on machine learning algorithms. *Digital Health*, vol. 9, pp. 20552076231170493, 2023.

[64] Alafif, T., et al., Machine and Deep Learning towards COVID-19 Diagnosis and Treatment: Survey, Challenges, and Future Directions. *International Journal of Environmental Research and Public Health*, vol. 18, no. 3, pp. 1117. 2021.

[65] Nayak, J., et al., Intelligent system for COVID-19 prognosis: a state-of-the-art survey. *Applied Intelligence*: pp. 1-31.

[66] Syeda, H.B., et al., Role of machine learning techniques to tackle the COVID-19 crisis: Systematic

review. *JMIR medical informatics*, vol. 9, no. 1, pp. e23811, 2021.

[67] Lalmuanawma, S., J. Hussain, and L. Chhakchhuak, Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals*, pp. 110059, 2020.

[68] Tayarani-N, M.-H., Applications of artificial intelligence in battling against Covid-19: a literature review. *Chaos, Solitons & Fractals*, pp. 110338, 2020.

[69] El-Bouzaidi, Y.E.I. and O. Abdoun, Advances in Artificial Intelligence for Accurate and Timely Diagnosis of COVID-19: A Comprehensive Review of Medical Imaging Analysis. *Scientific African*, pp. e01961, 2023.

[70] Burnazovic, E., et al., Application of Artificial intelligence in COVID-19-related geriatric care: A scoping review. Archives of Gerontology and Geriatrics, vol. 116, pp. 105129, 2024.

[71] Khan, M., et al., Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review. *Expert Systems with Applications*, vol. 185, pp. 115695, 2021.

[72] Comito, C. and C. Pizzuti, Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review. *Artificial Intelligence in Medicine*, vol. 128: pp. 102286, 2022.

[73] Tayarani N, M.-H., Applications of artificial intelligence in battling against covid-19: A literature review. *Chaos, Solitons & Fractals*, vol. 142, pp. 110338, 2021.

[74] Yi, J., et al., Review on the COVID-19 pandemic prevention and control system based on AI. *Engineering Applications of Artificial Intelligence*, vol. 114: pp. 105184, 2022.

[75] Nakamoto, S., Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*, 2008.

[76] Sullivan, C. and E. Burger, E-residency and blockchain. *Computer Law & Security Review*, vol. 33, no. 4, pp. 470-481, 2017.

[77] Yüksel, B., A. Küpçü, and Ö. Özkasap, Research issues for privacy and security of electronic health services. *Future Generation Computer Systems*, vol. 68, pp. 1-13, 2017.

[78] Kshetri, N., Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy*, vol. 41, no. 10, pp. 1027- 1038, 2017.

[79] Beninger, P. and M.A. Ibara, Pharmacovigilance and Biomedical Informatics: A Model for Future Development. *Clinical Therapeutics*, vol. 38, no. 12, pp. 2514-2525, 2016.

[80] Au, M.H., et al., A general framework for secure sharing of personal health records in cloud system. *Journal of Computer and System Sciences*, vol. 90, pp. 46-62, 2017.

[81] Wang, S., et al., Blockchain-powered parallel healthcare systems based on the ACP approach. IEEE *Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 942-950, 2018.

[82] Andrew, J., et al., Blockchain for healthcare systems: Architecture, security challenges, trends and future directions. *Journal of Network and Computer Applications*, pp. 103633, 2023.

[83] IBM. How the FDA is piloting blockchain for the pharmaceutical supply chain. 2020 [cited 2021 3 Apr 2021]; Available from: https://www.ibm.com/blogs/blockchain/2020/05/howthe-fda-is-piloting-blockchain-for-the-pharmaceuticalsupply-chain/.

[84] N., A. Blockchain in healthcare: use cases for your practice. 2020 [cited 2021 3 Apr 2021]; Available from: https://www.itransition.com/blog/blockchain-inhealthcare.

[85] Yong, B., et al., An intelligent blockchain-based system for safe vaccine supply and supervision. International *Journal of Information Management*, vol. 52, pp. 102024, 2020.

[86] Hussien, H.M., et al., Blockchain technology in the healthcare industry: Trends and opportunities. *Journal of Industrial Information Integration*, vol. 22: pp. 100217, 2021.

[87] He, W., Z.J. Zhang, and W. Li, Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic. *International Journal of Information Management*, vol. 57, pp. 102287, 2021.

[88] Stafford, T.F. and H. Treiblmaier, Characteristics of a Blockchain Ecosystem for Secure and Sharable Electronic Medical Records. *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1340- 1362, 2020.

[89] Zhang, A. and X. Lin, Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain. *Journal of Medical Systems*, vol. 42, no. 8, pp. 140, 2018.

[90] Islam, A., et al., A Blockchain-Based Artificial Intelligence-Empowered Contagious Pandemic Situation Supervision Scheme Using Internet of Drone Things. *IEEE Wireless Communications*, 28(4), pp. 166- 173, 2021.

[91] Kumar, A., et al., A drone-based networked system and methods for combating coronavirus disease (COVID-19) pandemic. *Future Generation Computer Systems*, vol. 115, pp. 1-19, 2021.

[92] Gupta, R., et al., Blockchain-envisioned softwarized multi-swarming UAVs to tackle COVID-I9 situations. *IEEE network*, vol. 35, no. 2, pp. 160-167, 2020.

[93] Karaarslan, E. and D. Aydın, 2 - An artificial intelligence–based decision support and resource management system for COVID-19 pandemic, in Data Science for COVID-19, U. Kose, et al., Editors. 2021, *Academic Press*. pp. 25-49.

[94] Wadud, M.A.H., et al. A Patient Centric Agent Assisted Private Blockchain on Hyperledger Fabric for Managing Remote Patient Monitoring. in 2020 *11th International Conference on Electrical and Computer Engineering (ICECE)*, 2020.

[95] Saha, R., et al., Internet of Things Framework for Oxygen Saturation Monitoring in COVID-19 Environment. *IEEE Internet of Things Journal*, vol. 9, no. 5l, pp. 3631-3641, 2022.

[96] Shrimali, B. and H.B. Patel, Blockchain state-ofthe-art: architecture, use cases, consensus, challenges and opportunities. *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 9, pp. 6793-6807, 2022.

[97] Xie, M., et al., A survey on blockchain consensus mechanism: research overview, current advances and future directions. *International Journal of Intelligent Computing and Cybernetics*, vol. 16, no. 2, pp. 314-340, 2023.

[98] Shrier, D., W. Wu, and A. Pentland, Blockchain & infrastructure (identity, data security). *Massachusetts Institute of Technology-Connection Science*, vol. 1, no. 3, pp. 1-19, 2016.

[99] Mitani, T. and A. Otsuka, Traceability in permissioned blockchain. *IEEE Access*, vol. 8, pp. 21573-21588, 2020.

[100] Almeshal, T.A. and A.A. Alhogail, Blockchain for businesses: a scoping review of suitability evaluations frameworks. *IEEE Access*, vol. 9, pp. 155425-155442, 2021.

[101] Xu, X., et al. The blockchain as a software connector. in 2016 13th *Working IEEE/IFIP Conference on Software Architecture (WICSA)*, IEEE, 2016,

[102] Lo, S.K., et al. Evaluating suitability of applying blockchain. in 2017, *22nd international conference on engineering of complex computer systems* (ICECCS), IEEE, 2017.

[103] Schneider, B. and W. Azan. Perceptions and Misconceptions of Blockchain: The Potential of Applying Threshold Concept Theory. in 2022 IEEE *6th International Conference on Logistics Operations Management (GOL)*, IEEE, 2022.

[104] Betzwieser, B., Franzbonenkamp, S., Riasanow, T., Böhm, M., Kienegger, H., & Krcmar, H. (2019). A decision model for the implementation of blockchain solutions.

[105] Simaremare, A., et al. Suitability study of Blockchain application in electric utility company business processes. in IOP Conference Series: Materials Science and Engineering. *IOP Publishing*, 2021.

[106] Masood, F., & Faridi, A. R. A multi-criteria decision-making approach to analyse the viability of blockchain in software development projects. *Journal of Intelligent & Fuzzy Systems*, vol. 44, no. 1, pp: 113-124, 2023.

[107] Spencer-Hicken, S., C. Schutte, and P. Vlok, Blockchain feasibility assessment: A quantitative approach. *Frontiers in Blockchain*, vol. 6, pp. 1067039, 2023.

[108] Peck, M.E., Blockchain world-Do you need a blockchain? This chart will tell you if the technology can solve your problem. *IEEE Spectrum*, vol. 54, no. 10, pp. 38-60, 2017.

[109] Wüst, K. and A. Gervais. Do you need a blockchain? in 2018 crypto valley conference on blockchain technology (CVCBT). *IEEE*, 2018.

[110] Emmadi, N., et al. Practical deployability of permissioned blockchains. in Business Information Systems Workshops: *BIS 2018 International Workshops, Berlin, Germany*, July 18–20, 2018, Revised Papers 21. 2019. Springer.

[111] Hunhevicz, J.J. and D.M. Hall, Do you need a blockchain in construction? Use case categories and decision framework for DLT design options. *Advanced Engineering Informatics*, vol. 45, pp. 101094, 2020.

[112] Pedersen, A.B., M. Risius, and R. Beck, A tenstep decision path to determine when to use blockchain technologies. *MIS Quarterly Executive*, vol. 18, no. 2, pp. 99-115.

[113] Hassija, V., et al., Framework for determining the suitability of blockchain: Criteria and issues to consider. Transactions on Emerging Telecommunications Technologies, vol. 32, no. 10, pp. e4334, 2021.

[114] El Madhoun, N., J. Hatin, and E. Bertin. Going beyond the blockchain hype: In which cases are blockchains useful for its applications? in 2019 *3rd Cyber Security in Networking Conference (CSNet)*. IEEE, 2019.

[115] Koens, T. and E. Poll. What blockchain alternative do you need? in Data Privacy Management, Cryptocurrencies and Blockchain Technology: *ESORICS 2018 International Workshops, DPM 2018 and CBT 2018, Barcelona, Spain*, September 6-7, 2018, Proceedings 13. 2018. Springer.

[116] Loo, W.K., et al., Systematic review on COVID-19 readmission and risk factors: future of machine learning in COVID-19 readmission studies. *Frontiers in public health*, vol. 10, pp. 898254l, 2022.

[117] Basile, D., et al., Digitalizing circular economy through Blockchains: The blockchain circular economy index. Journal of Industrial and Production Engineering, vol. 40, no. 4, pp. 233-245, 2023.

[118] XDC. XDC Network. 2023 [cited 2023] 11/24/2023]; Available from: https://xdc.org/.

[119] Kushwaha, S.S., et al., Ethereum smart contract analysis tools: A systematic review. *IEEE Access*, vol. 10, pp. 57037-57062, 2022.

[120] Grishchenko, I., M. Maffei, and C. Schneidewind. A semantic framework for the security analysis of ethereum smart contracts. *in International Conference on Principles of Security and Trust*. 2018. Springer.

[121] Christidis, K. and M. Devetsikiotis, Blockchains and smart contracts for the internet of things. *IEEE* Access, 2016. vol. 4, pp. 2292-2303.

[122] Burki, T., Global shortage of personal protective equipment. *The Lancet Infectious Diseases*, vol. 20, no. 7, pp. 785-786, 2020.

[123] Wirth, R. and J. Hipp. CRISP-DM: Towards a standard process model for data mining. *in Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. 2000. Manchester.

[124] Wen, Y., et al., Time-to-event modeling for hospital length of stay prediction for COVID-19 patients. *Machine Learning with Applications*, vol. 9, pp. 100365, 2022.

[125] Alansari, S.A., et al., Efficient and privacypreserving infection control system for COVID-19-like pandemics using blockchain. I*EEE Internet of Things Journal*, vol. 9, no. 4, pp. 2744-2760, 2021.

[126] Ahmed, I., A. Chehri, and G. Jeon, Artificial Intelligence and Blockchain Enabled Smart Healthcare System for Monitoring and Detection of COVID-19 in Biomedical Images. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, pp. 1-10, 2023.

[127] Ahmad, R.W., et al., Blockchain and COVID-19 pandemic: Applications and challenges. *Cluster Computing*, pp. 1-26, 2023.

[128] Caceres, M.M.F., et al., The impact of misinformation on the COVID-19 pandemic. *AIMS public health*, vol. 9, no. 2, pp. 262, 2022.

[129] Ceci, Flavio Maria, Marco Fiore, Francesca Gavaruzzi, Antonio Angeloni, Marco Lucarelli, Carolina Scagnolari, Enea Bonci et al. "Early routine biomarkers of SARS-CoV-2 morbidity and mortality: outcomes from an emergency section." *Di, agnostics,*  vol. 12, no. 1, p. 176, 2022.

[130] Piri, Seyed Mohammad, Maryam Edalatfar, Sina Shool, Mohammad Naser Jalalian, and Soheil Tavakolpour. "A systematic review on the recurrence of SARS-CoV-2 virus: frequency, risk factors, and possible explanations." *Infectious Diseases*, vol. 53, no. 5, pp. 315-324, 2021.

[131] Dhillon, Rubaid Azhar, Mohammad Aadil Qamar, Jaleed Ahmed Gilani, Omar Irfan, Usama Waqar, Mir Ibrahim Sajid, and Syed Faisal Mahmood. "The mystery of COVID-19 reinfections: A global systematic review and meta-analysis." *Annals of Medicine and Surgery*, vol. 72, pp.103130, 2021.

## **پیکربندی سیستم هوشمند بلاکچین-محور برای غربالگری، نظارت و ردیابی بیماریهای همهگیر**

علی شبرندی<sup>5</sup>\*، علی رجب زاده قطری<sup>۲</sup>، محمود دهقان نیری<sup>۳</sup>، نادر توکلی ٔ و سحر میرزایی<sup>۵</sup>

1 دانشجوی دکتری مدیریت صنعتی، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس، تهران، ایران.

<sup>۲</sup> گروه مدیریت صنعتی، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس.

3 گروه مدیریت صنعتی، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس.

4 گروه اورژانس، مرکز تحقیقات تروما و آسیب، دانشگاه علوم پزشکی ایران.

5 دانش آموخته مهندسی بهداشت محیط، دانشگاه علوم پزشکی ایران، تهران، ایران.

ارسال 2524/54/55؛ بازنگری 2524/55/25؛ پذیرش 2524/50/22

#### **چکیده:**

این مطالعه پیکربندی سـطح بالا یک ســیســتم هوشــمند مبتنی بر بلاکچین دوگانه (هیبرید و خصــوصــی) را برای مدیریت همهگیری کووید-۱۹ ارائه میدهد. این پیکربندی شـامل نوع شـبکه، سطح تمرکززدایی، گرهها و نقشها، اطلاعات ساختار بلوک، کنترل اختیارات و قراردادهای هوشمند است و دو دســته اصــلی چالش - مدیریت عملیات و مدیریت داده – را از طریق ســه ماژول هوشــمند در مراحل مختلف همهگیری کووید-١٩ مورد مخاطب قرار میدهد. در مرحله پیش بیمارسـتانی، یک سـیسـتم هوشمند پیشبینی ابتلای به کووید-۱۹ را ارائه میدهد که این سیستم از دادههای خانگی برای رفع فقدان یک روش غربالگری سـاده، کارآمد، چابک و کمهزینه برای شـناسـایی سـریع افراد بالقوه آلوده اسـتفاده میکند. در مرحله داخل بیمارستانی، یک سطیست پیشبینی هوشمند برای پیشبینی شدت ابتلا و مدت اقامت بیمار در بیمارستان برای شناسایی بیماران پرخطر، اولویتبندی آنها برای دریافت خدمات مراقبتی و تسـهیل تخصـیص بهتر منابع پیشـنهاد شـده است. در مرحله پسابیمارستانی، یک سیستم پیشبینی میزاس پیشبینی میزان ابتلای و بسـتری مجدد، برای کمک به کاهش بار سـیسـتم مراقبتهای بهداشـتی در معرض خطر پیشـنهاد شـده اسـت. علاوه بر این، توزیع منصـفانه تجهیزات حفاظت شـخصـی محدود را با اسـتفاده از بلاکچین خصـوصـی و قراردادهای هوشـمند ارائه مینماید. این ماژول&ا با اسـتفاده از پایتون توسعه یافتهاند و برای ارزیابی عملکرد، تکنیکهای پیشـرفتهی یادگیری ماشین از طریق اعتبارسنجی متقابل در هر مرحله استفاده میشوند. مهم ترین ویژگیها با اسـتفاده از مقادیر (SHAP) ترسـیم و تحلیل شدند. در نهایت، دلالتهای سیستم خود را برای تحقیق و عمل بررسی و توصیههایی برای پیشرفتهای آینده ارائه شده است.

**کلمات کلیدی:** همهگیری، بلاکچین، هوش مصنوعی، پیکربندی، سیست .