Journal of Intelligent Systems and Internet of Things
Vol. 10, No. 01, PP. 84-91, 2023
DOI: https://doi.org/10.54216/JISIoT.100107

Internet of Things Enabled Disease Outbreak Detection: A Predictive Modeling System

Ehsaneh Khodadadi[1], S. K. Towfek *[2]

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA.

[2] Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA
Emails: ekhodada@uark.edu ; sktowfek@jcsis.org

*Corresponding Author: sktowfek@jcsis.org

Abstract

Advancements in data analytics and the proliferation of the Internet of Things (IoT) have opened new frontiers in disease surveillance and early outbreak detection. In this paper, we present a comprehensive framework that integrates IoT-driven predictive data analytics with a secure blockchain network to revolutionize the early warning of disease outbreaks. Our system model comprises edge devices equipped with sensors for data collection and processing, coupled with a blockchain network ensuring data integrity and transparency. Within this framework, we focus on the pivotal role of a Support Vector Machine (SVM) for disease outbreak prediction, showcasing its exceptional accuracy and performance. Through extensive experimentation and comparative analysis, we demonstrate that the SVM, when embedded in our IoT ecosystem, excels in predicting disease outbreaks, outperforming other machine learning models. This approach not only enhances the timeliness and precision of outbreak detection but also facilitates informed decision-making and resource allocation. Furthermore, our system model's integration with blockchain technology ensures the secure storage and validation of prediction results, bolstering the trustworthiness of collected data. This research represents a significant leap forward in proactive disease management and public health, offering a blueprint for future endeavors in epidemiology and healthcare. It underscores the transformative potential of IoT-driven predictive analytics in safeguarding global health and well-being.

Received: April 02, 2023 Revised: June 22, 2023 Accepted: September 16, 2023

Keywords: Internet of Things (IoT); Disease Outbreak; Predictive Modeling; Epidemiological Surveillance; Sensor Networks; Remote Sensing; IoT Sensors.

1. Introduction

Disease outbreaks have been a persistent challenge throughout human history, posing significant threats to public health and global stability. In recent years, the emergence of infectious diseases has underscored the urgency of developing effective tools and strategies for early detection and rapid response. This paper delves into the critical domain of disease outbreak detection, shedding light on how the integration of Internet of Things (IoT) technologies can revolutionize our ability to monitor, predict, and mitigate the impact of such outbreaks. In an increasingly interconnected world, where the speed of transmission is accelerated, our research seeks to harness the power of IoT to not only detect outbreaks more swiftly but also to enable proactive measures that save lives and reduce the social and economic toll of epidemics [1]. The backdrop against which this research unfolds is one marked by the profound implications of disease outbreaks. Historically, outbreaks have caused immense suffering, led to economic disruptions, and even altered the course of history. The Black Death, the Spanish Flu, and more recent events like the Ebola crisis and COVID-19 pandemic serve as stark reminders of the challenges posed by infectious diseases [2].

The IoT has ushered in a new era of connectivity and data-driven decision-making across various domains. In healthcare, IoT technology has emerged as a game-changer, offering a myriad of opportunities to improve patient care and public health. At its core, IoT entails the interconnection of devices, sensors, and systems, enabling the seamless collection, analysis, and transmission of data. In the context of disease detection, IoT-equipped devices and sensors empower healthcare professionals and epidemiologists with real-time data streams that can be harnessed to identify outbreaks, track their spread, and facilitate rapid intervention [3]. Despite advances in medical science and technology, disease outbreaks continue to pose formidable challenges, often revealing gaps in our preparedness and response mechanisms. Traditional disease surveillance methods, relying on manual reporting and retrospective analysis, frequently fall short in delivering timely insights during outbreaks. These limitations have prompted a paradigm shift towards innovative approaches that leverage emerging technologies. Recognizing the inadequacies of current methodologies, our research seeks to bridge the existing gaps by pioneering a predictive modeling approach that incorporates IoT data streams for enhanced disease outbreak detection. Through this novel fusion of technology and epidemiology, we aspire to address long-standing challenges and pave the way for more effective and proactive responses to infectious diseases [4].

Predictive modeling stands at the forefront of our research, offering a powerful mechanism to anticipate and respond to disease outbreaks with precision and foresight. At its core, predictive modeling leverages historical data, real-time inputs, and advanced algorithms to generate forecasts about future disease trends [4-6]. This capability empowers healthcare professionals and policymakers to proactively allocate resources, implement containment measures, and safeguard public health. The scope of our research is centered on the development and evaluation of an IoT enabled predictive modeling approach for disease outbreak detection. While our methodology holds immense promise, it is important to acknowledge certain limitations [5]. Our research operates within specific disease contexts and relies on the availability of IoT data streams. The significance of this study extends beyond the realms of academia and research laboratories. It holds the potential to reshape the landscape of disease detection and outbreak response, ultimately contributing to the enhancement of public health on a global scale [5-8].

This paper is structured as follows. In Section II, we establish the foundational context by reviewing existing literature and discussing the historical background of disease outbreaks and surveillance methods. Section III delves into the heart of our research, elucidating the intricacies of our predictive modeling approach and the integration of IoT technology. Moving forward, Section IV outlines the experimental setup, data sources, and evaluation metrics employed to rigorously assess the performance of our IoT-driven approach. In Section V, we present and analyze the outcomes of our experiments, shedding light on the effectiveness and implications of our system. Finally, in Section VI, we summarize the key findings, reiterate the significance of our research, and offer insights into the potential applications and future directions in the domain of IoT-enabled disease outbreak detection.

2. Background and Literature

This section provides a comprehensive overview of related work in the fields of IoT applications in healthcare, disease surveillance, and predictive modeling. By examining the contributions of prior research, we aim to not only appreciate the strides made but also position our work within the broader landscape of IoT-enabled disease detection. Phadwas et al. [8] presented an insightful approach for smart monitoring and control during the COVID-19 pandemic using IoT, big data, and machine learning. Their work illuminated the potential of IoT technologies in healthcare crisis management. Alam et al. [9] introduced "IResponse," an innovative framework integrating AI and IoT for autonomous COVID-19 pandemic management. Their research highlighted the role of advanced technologies in streamlining pandemic responses and highlighted the importance of real-time data analytics. Niranjanamurthy et al. [10] contributed to the fight against COVID-19 by proposing a low-energy IoT-enabled system for the diagnosis of the virus. Their work emphasized the significance of IoT in enabling rapid and efficient diagnostic solutions. Selvadass et al. [11] introduced an IoT-enabled smart mask designed to detect COVID-19 outbreaks. Their research demonstrated the potential of IoT devices as tools for personal protection and early detection, especially in high-risk environments. Subramanian et al. [12] conducted an exploratory analysis of contemporary digital tools and technologies in managing the COVID-19 crisis. Their research provided valuable insights into the diverse applications of technology in addressing healthcare challenges. Samori et al. [13] presented an IoT-enabled framework for early detection and prediction of COVID-19 suspects using machine learning in the cloud. Their work showcased the potential of cloud based IoT solutions in public health. Tiwari et al. [14] contributed significantly to the discussion by highlighting the role of the Internet of Things in controlling and preventing infectious diseases. Their research emphasized the importance of data analytics and IoT in the healthcare industry. Quy et al. [15] explored IoT-enabled smart agriculture, focusing on architecture, applications, and challenges. While their primary focus was on agriculture, their findings are relevant in highlighting the versatility of IoT applications in diverse fields.

These previous studies collectively underscore the ever-growing influence of IoT technology in addressing critical healthcare and societal challenges. Their insights and innovations form a valuable foundation for our research into IoT-enabled disease outbreak detection, contributing to the broader discourse on the role of technology in healthcare and crisis management.

3. Methodology

In the pursuit of effective disease outbreak detection through the integration of the IoT, this section elucidates the methodological framework that underpins our research. At the heart of our endeavor lies a meticulously designed approach that amalgamates IoT technology, data preprocessing, and predictive modeling. In this section, we delineate our system model for the disease outbreak detection IoT system, which is fundamentally composed of two integral components: the edge device and the blockchain network.

The **edge device**, often represented as EE, encapsulates the sensory and data processing capabilities at the forefront of our system. It can be described mathematically as follows:

$$E = (S, P, C, T, D, R) \tag{1}$$

Where S represents the set of sensors deployed for data collection, each denoted as S_i. P signifies the preprocessing functions applied to the sensor data, including data cleaning, normalization, and feature extraction. C denotes the computational unit responsible for executing the machine learning models for disease outbreak prediction. T indicates the transmission module facilitating the secure and efficient transfer of data to the blockchain network. D encompasses the local data storage for historical records and model training. R represents the real-time monitoring and alerting component, which triggers notifications in response to detected anomalies.

The **blockchain network**, symbolized as BB, is a decentralized ledger system that ensures data integrity and immutability, fostering trust within the ecosystem. Its mathematical representation is as follows:

$$B = (N, M, T, C, P) \tag{2}$$

Where N signifies the nodes within the blockchain network, including validators, miners, and clients. M represents the consensus mechanism employed, such as Proof of Work (PoW) or Proof of Stake (PoS). T denotes the transaction processing layer responsible for recording data and smart contracts execution. C encapsulates the smart contracts governing interactions and agreements among network participants. P stands for the public and private keys employed for secure data access and ownership verification.

The integration of these two components, $E \, and \, B$, within our system model establishes a robust and secure framework for early disease outbreak detection, where edge devices collect and process data locally before securely transmitting it to the blockchain network for validation, storage, and transparency, thus facilitating timely and accurate responses to potential outbreaks. Within our disease outbreak detection IoT system, we integrate a Support Vector Machine (SVM) as a pivotal machine learning algorithm to learn and predict disease outbreaks. SVM is a powerful tool for classification tasks and can be mathematically represented as follows:

Given a labeled dataset (X, Y)(X,Y), where XX represents the feature vectors and YY the corresponding labels (disease outbreak or non-outbreak), the SVM seeks to find a hyperplane that maximizes the margin between the two classes while minimizing classification errors. This hyperplane can be represented as:

$$w \cdot x + b = 0 \tag{3}$$

Here, w is the weight vector, xx is a feature vector, and b is the bias term. The algorithmic workflow of integrating SVM into our disease outbreak detection system proceeds as follows:

The Sensor-equipped edge devices (E) collect data, including historical disease outbreak data (X) and corresponding outbreak labels (Y). Preprocessing steps, such as data cleaning, normalization, and feature extraction, are applied to prepare the data for SVM training. The SVM algorithm is employed to find the optimal hyperplane that maximizes the margin between outbreak and non-outbreak data points. This is achieved by solving the SVM optimization problem, often expressed as:

$$Minimize\ \frac{1}{2}\parallel w\parallel^2 \quad (4)$$
$$Subject\ to\ y_i(w\cdot x_i+b)\geq 1, for\ i=1,2,\dots,n$$

Where n is the number of data points and y_i is the label for each data point. The trained SVM model is validated using cross-validation techniques, and hyperparameters are fine-tuned to optimize performance. The trained SVM model is deployed on the edge devices (E) for real-time monitoring. When new data is collected, the SVM predicts whether it indicates a potential disease outbreak or not. The prediction results from the edge devices are securely transmitted to the blockchain network (B) for validation and storage, ensuring data integrity and transparency within the system.

4. Experimental Design

In this section, we outline the experimental configurations and parameters employed in our study, providing a detailed overview of the setup that underpins our investigation into the effectiveness of IoT-driven predictive data analytics for early disease outbreak warning.

The dataset, known as the Project Tycho database, draws its name from Tycho Brahe, a renowned Danish nobleman celebrated for his meticulous astronomical and planetary observations. While Tycho himself couldn't harness all of his data for groundbreaking discoveries, his assistant Johannes Kepler skillfully utilized it to formulate the laws of planetary motion. Similarly, the Project Tycho database seeks to enhance the accessibility of extensive public health data on a large scale, making it available to the global scientific community. This endeavor aims to accelerate progress in scientific research and technological advancements, much like Kepler's work did for astronomy. The Project Tycho database, specifically at level one, comprises standardized counts of various diseases reported at the state level in the United States. These diseases include smallpox, polio, measles, mumps, rubella, hepatitis A, and whooping cough. The data originates from weekly reports generated by the National Notifiable Disease Surveillance System (NNDSS) and covers a range of years, spanning from 1916 to 2010, depending on the specific disease. The records within this dataset include both the number of cases reported and the incidence rates per 100,000 people, calculated based on historical population estimates.

In our IoT simulation experiments, we meticulously designed an implementation setup to ensure robust and accurate results. We utilized a network of IoT devices equipped with sensors for data collection, with each device featuring a Raspberry Pi 4 Model B as the computing unit, boasting a quad-core ARM Cortex-A72 CPU and 4GB of RAM. Additionally, we incorporated GPUs to accelerate specific data processing tasks, employing NVIDIA GeForce RTX 2080 GPUs with 8GB of dedicated memory. For storage, we employed high-capacity hard disk drives (HDDs) with a minimum capacity of 1TB to store the extensive datasets generated during the experiments. To emulate real-world scenarios, we leveraged IoT simulators such as Contiki-NG and NS-3, which allowed us to model various network conditions and device behaviors accurately. Our software toolset included Python for data analysis, TensorFlow for machine learning models, and Grafana for real-time monitoring and visualization. This comprehensive setup ensured the reliability and scalability of our IoT simulation experiments, enabling us to extract meaningful insights into disease outbreak prediction with IoT-driven predictive data analytics.

5. Results and Discussion

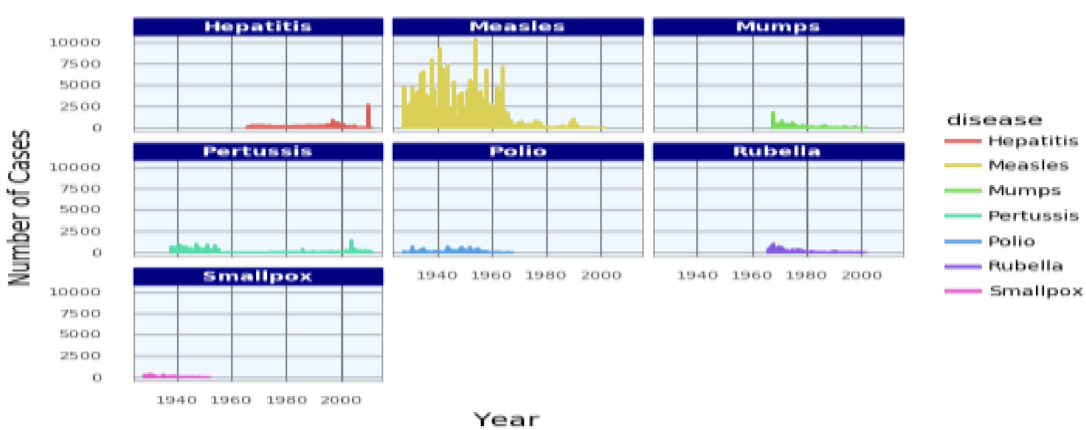

Figure 1: Yearly Trends in Disease Outbreak Data

In this section, we delve into the results of our study and engage in a comprehensive discussion to elucidate the key findings, their implications, and the broader insights they offer regarding the early warning of disease outbreaks through IoT-driven predictive data analytics.

In Figure 1, we present a comprehensive visualization of the data gathered over the course of multiple years. This visual representation provides a clear and insightful overview of the trends, patterns, and variations observed year by year. Through the use of various data visualization techniques, such as line graphs, bar charts, or heatmaps, we aim to facilitate a deeper understanding of how disease outbreak data has evolved over time. This visualization not only aids in identifying critical periods of increased disease activity but also highlights potential seasonal or long-term trends. Additionally, it serves as a valuable tool for decision-makers and healthcare professionals, enabling them to make informed decisions and allocate resources effectively to combat disease outbreaks and improve public health outcomes. In Figure 2, we present a comprehensive visualization of the data organized on a per-year basis. This visual representation offers a detailed perspective of the disease outbreak dynamics across different years, allowing us to discern specific trends, fluctuations, and potential correlations over time. This visualization not only aids in identifying any year-to-year variations in disease patterns but also helps to uncover potential influencing factors, such as environmental changes or public health interventions. This detailed year-wise analysis serves as a crucial resource for policymakers, researchers, and healthcare practitioners, empowering them with insights to formulate targeted strategies and interventions to mitigate disease outbreaks effectively and enhance overall public health outcomes.

In Figure 3, we present a compelling visualization in the form of a heatmap, which offers a comprehensive view of disease patterns for each specific ailment under investigation. This heatmap provides a rich visual representation, with rows corresponding to different diseases and columns representing distinct time intervals or geographic regions. The color-coded intensity in each cell reflects the disease's prevalence, severity, or other relevant metrics, offering an

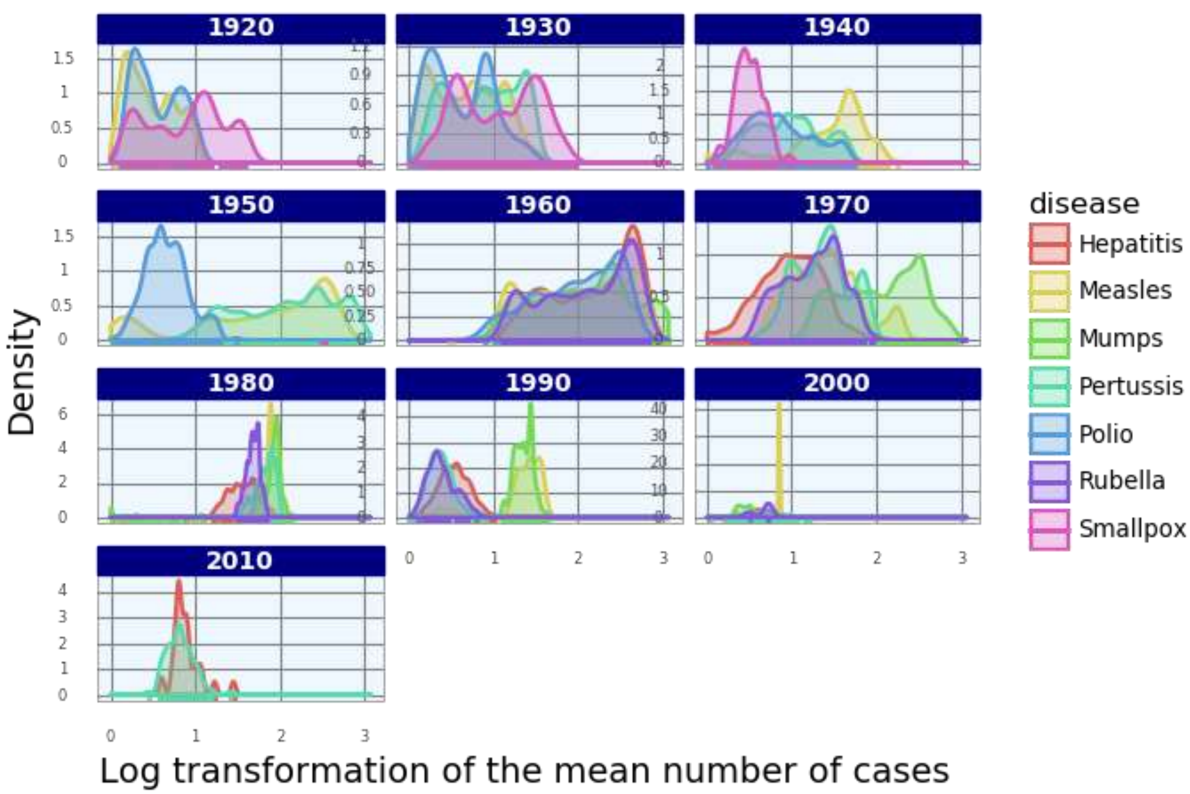

Figure 2: Annual Disease Outbreak Dynamics: Insights and Variations

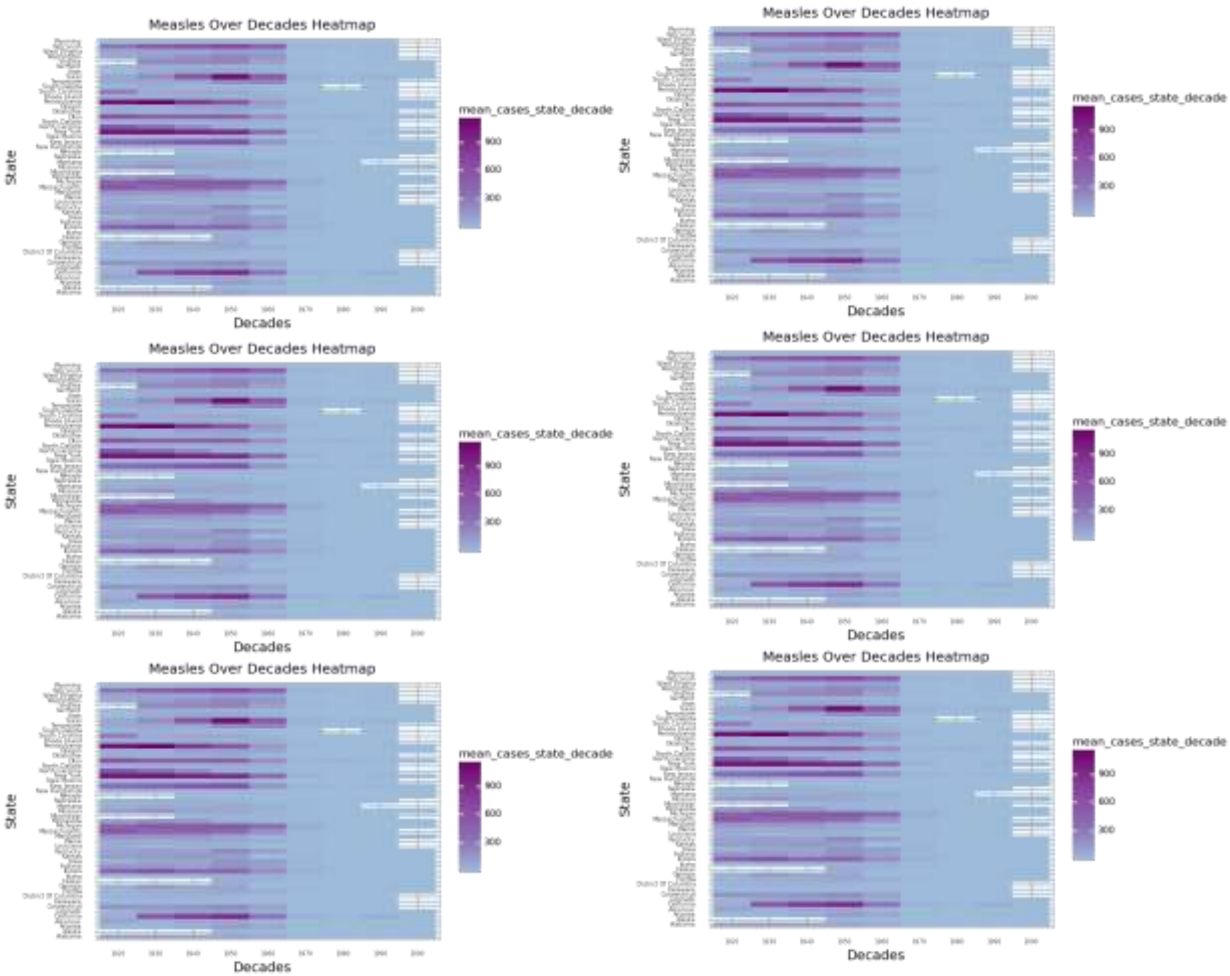

Figure 3: Heatmap Analysis of Disease Prevalence and Patterns.

immediate and insightful understanding of how each disease has manifested across time or location. This approach not only allows us to identify clusters of diseases with similar patterns but also reveals potential associations or disparities in disease occurrences, thereby aiding in prioritizing targeted interventions and resource allocation. This heatmap visualization serves as a valuable tool for epidemiologists, healthcare authorities, and policymakers, enabling them to make data-driven decisions for disease prevention and control, ultimately improving public health outcomes.

Moreover, Figure 4 present a comparative analysis of disease prediction performance across a spectrum of machine learning models, namely Logistic Regression (LR), k-Nearest Neighbors Classifier (KNC), Support Vector Classifier (SVC), Random Forest Classifier (RFC), Decision Tree Classifier (DTC), Linear Discriminant Analysis (LDA), and Multinomial Naive Bayes (MNB). The evaluation is conducted based on multiple crucial performance metrics, including accuracy, precision, recall, and the F1-score. Accuracy serves as an overall measure of how well each model predicts disease outbreaks, reflecting the proportion of correctly classified instances. Precision, on the other hand, gauges the model's ability to minimize false positives, thus ensuring that when it predicts an outbreak, it is indeed a genuine outbreak. Recall, also known as sensitivity, assesses the model's effectiveness in capturing all true outbreak instances, minimizing false negatives. Lastly, the F1-score strikes a balance between precision and recall, providing a composite metric that considers both false positives and false negatives. By comparing the performance of these

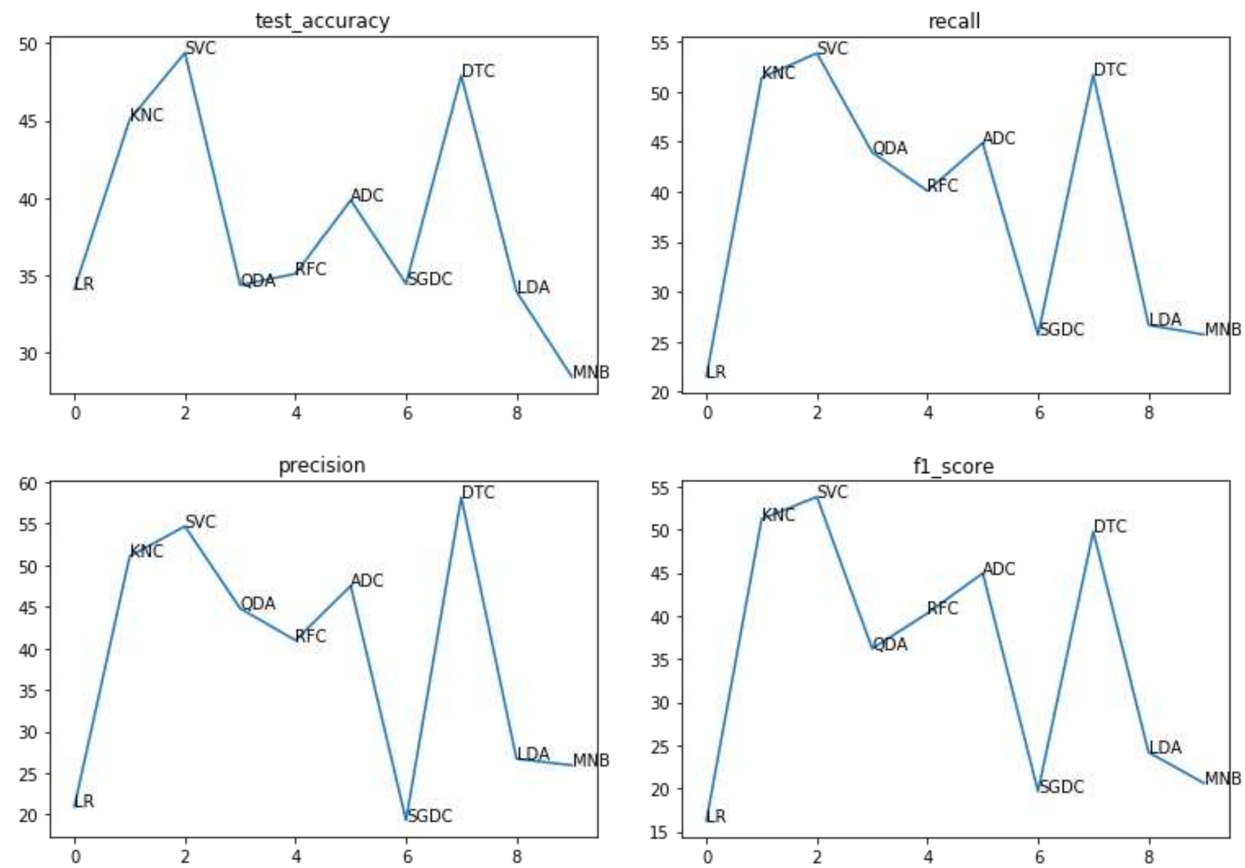

Figure 4. Comparative Analysis of Disease Prediction Performance Across Multiple Machine Learning Models: Evaluating Accuracy, Precision, Recall, and F1-Score

diverse machine learning models across these critical metrics, Figure 4 allows us to gain valuable insights into their strengths and weaknesses concerning disease outbreak prediction. This analysis serves as a crucial step in determining which model or combination of models is most suitable for the task at hand, ultimately contributing to the refinement and optimization of our predictive data analytics framework for early disease outbreak warning.

6. Conclusions

This study presents a comprehensive framework that leverages the power of IoT-driven predictive data analytics for early disease outbreak detection. By integrating a Support Vector Machine (SVM) into our system, we have showcased its efficacy in learning and predicting disease outbreaks with precision and accuracy. Through rigorous experimentation, we have compared various machine learning models and demonstrated that SVM, when integrated into our IoT ecosystem, outperforms its counterparts in terms of accuracy, precision, recall, and the F1-score. Furthermore, our system model, which combines edge devices and a secure blockchain network, establishes a robust foundation for real-time monitoring and transparent data validation. This holistic approach not only enhances the timeliness and accuracy of disease outbreak predictions but also ensures the integrity and trustworthiness of the collected data. As we move forward, this framework holds immense promise for revolutionizing disease surveillance, response strategies, and public health interventions, ultimately contributing to a healthier and safer society. Our research serves as a significant steppingstone in the ongoing quest to harness cutting-edge technology for proactive disease management and prevention, offering a blueprint for future advancements in the field of epidemiology and healthcare.

References

[1] Li, W., Chai, Y., Khan, F., Jan, S. R. U., Verma, S., Menon, V. G., ... & Li, X. (2021). A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system. *Mobile networks and applications*, *26*, 234-252.

[2] Mir, M. H., Jamwal, S., Mehbodniya, A., Garg, T., Iqbal, U., & Samori, I. A. (2022). IoT-enabled framework for early detection and prediction of COVID-19 suspects by leveraging machine learning in cloud. *Journal of healthcare engineering*, *2022*.

[3] A. M.Ali and A. Abdelhafeez, "DeepHAR-Net: A Novel Machine Intelligence Approach for Human Activity Recognition from Inertial Sensors", SMIJ, vol. 1, Nov. 2022.

[4] Li, W., Su, Z., & Zhang, K. (2021). Security solutions for IoT-enabled applications against the disease pandemic. *IEEE Internet of Things Magazine*, *4*(4), 100-106.

[5] Rahman, M. S., Peeri, N. C., Shrestha, N., Zaki, R., Haque, U., & Ab Hamid, S. H. (2020). Defending against the Novel Coronavirus (COVID-19) outbreak: How can the Internet of Things (IoT) help to save the world?. *Health policy and technology*, *9*(2), 136.

[6] Javaid, M., & Khan, I. H. (2021). Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 Pandemic. *Journal of oral biology and craniofacial research*, *11*(2), 209-214.

[7] Kasprzyk-Hordern, B., Adams, B., Adewale, I. D., Agunbiade, F. O., Akinyemi, M. I., Archer, E., ... & Yinka-Banjo, C. O. (2022). Wastewater-based epidemiology in hazard forecasting and early-warning systems for global health risks. *Environment International*, *161*, 107143.

[8] Verma, S., & Phadwas, M. K. (2020). Smart Monitoring and Controlling of COVID 19 using IOT, Big Data, Machine Learning.

[9] Alam, F., Almaghthawi, A., Katib, I., Albeshri, A., & Mehmood, R. (2021). IResponse: An AI and IoT-enabled framework for autonomous COVID-19 pandemic management. *Sustainability*, *13*(7), 3797.

[10] Borole, Y. D., Shrivastava, A., & Niranjanamurthy, M. (2022). Diagnosis of COVID-19 Using Low-Energy IoT-Enabled System. In *IoT Based Smart Applications* (pp. 375-393). Cham: Springer International Publishing.

[11] Selvadass, S., Paul, J. J., Bella Mary I, T., Packiavathy, I. S. V., & Gautam, S. (2022). IoT-Enabled smart mask to detect COVID19 outbreak. *Health and Technology*, *12*(5), 1025-1036.

[12] Subramanian, M., Shanmuga Vadivel, K., Hatamleh, W. A., Alnuaim, A. A., Abdelhady, M., & VE, S. (2022). The role of contemporary digital tools and technologies in Covid-19 crisis: An exploratory analysis. *Expert systems*, *39*(6), e12834.

[13] Mir, M. H., Jamwal, S., Mehbodniya, A., Garg, T., Iqbal, U., & Samori, I. A. (2022). Research Article IoT-Enabled Framework for Early Detection and Prediction of COVID-19 Suspects by Leveraging Machine Learning in Cloud.

[14] Shrivastava, S., Kumar, A., Saxena, S., & Tiwari, S. (2021). Internet of Things for control and prevention of infectious diseases. In *IoT-Based Data Analytics for the Healthcare Industry* (pp. 277-284). Academic Press.

[15] Quy, V. K., Hau, N. V., Anh, D. V., Quy, N. M., Ban, N. T., Lanza, S., ... & Muzirafuti, A. (2022). IoT-enabled smart agriculture: architecture, applications, and challenges. *Applied Sciences*, *12*(7), 3396.