International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

Big Data Analytics Frameworks for Pandemic Prediction and Healthcare Resource Allocation Management

Vishal Anil Patel

Government Degree College Khairatabad – Hyderabad, India

ABSTRACT: Big Data Analytics frameworks have become essential tools for pandemic prediction and healthcare resource allocation management in modern public health systems. These frameworks integrate distributed computing, machine learning, artificial intelligence, and real-time data processing techniques to analyze large-scale and heterogeneous datasets generated from hospitals, laboratories, mobility systems, social media platforms, and public health surveillance networks. By processing structured and unstructured data, big data systems can identify early outbreak signals, track disease transmission patterns, and predict future infection trends with high accuracy. In addition, these frameworks support healthcare resource allocation by optimizing the distribution of medical supplies, hospital beds, ventilators, vaccines, and healthcare personnel during pandemics. Technologies such as Hadoop, Apache Spark, cloud computing platforms, and real-time streaming systems enable scalable and efficient data processing. Predictive models such as time-series forecasting, regression analysis, and deep learning-based epidemic models enhance decision-making capabilities for healthcare authorities. Despite their benefits, challenges such as data privacy, data heterogeneity, model uncertainty, infrastructure limitations, and real-time processing constraints remain significant. This study explores big data analytics frameworks, literature contributions, methodologies, advantages, and limitations in pandemic prediction and healthcare resource management systems.

KEYWORDS: Big Data Analytics, Pandemic Prediction, Healthcare Resource Allocation, Disease Outbreak Forecasting, Machine Learning, Artificial Intelligence, Healthcare Systems, Epidemiology, Cloud Computing, Real-Time Data Processing, Apache Spark, Hadoop, Public Health Informatics, Predictive Modeling, Resource Optimization

I. INTRODUCTION

The emergence of global pandemics such as COVID-19 has highlighted the critical importance of advanced technological systems for disease prediction and healthcare resource management. Traditional public health surveillance systems often rely on manual reporting methods, historical data analysis, and delayed reporting mechanisms, which limit their ability to respond quickly to rapidly spreading infectious diseases. With the exponential growth of digital data generated from healthcare institutions, mobile devices, social media platforms, transportation systems, and IoT-enabled health monitoring devices, big data analytics has emerged as a powerful solution for improving pandemic prediction and healthcare decision-making processes. Big data analytics frameworks enable the collection, storage, processing, and analysis of massive and diverse datasets in real time, allowing healthcare authorities to identify early warning signals of disease outbreaks and take preventive measures before situations escalate. These systems play a crucial role in strengthening global health security and improving emergency response capabilities.

Big data analytics frameworks for pandemic prediction integrate multiple data sources to provide a comprehensive understanding of disease spread patterns and population health dynamics. Healthcare data from hospitals and laboratories includes patient records, test results, hospitalization rates, and clinical reports. Mobility data collected from transportation systems, mobile devices, and GPS tracking helps in understanding population movement and transmission pathways. Social media platforms provide real-time insights into public sentiment, symptom reporting, and outbreak discussions. Environmental data such as temperature, humidity, and pollution levels also contribute to understanding disease propagation factors. By combining these diverse datasets, big data systems can generate predictive models that estimate infection rates, identify high-risk regions, and forecast future outbreak scenarios. Machine learning and artificial intelligence techniques are used to analyze these datasets and generate actionable insights for policymakers and healthcare professionals.

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

Healthcare resource allocation management is another critical application of big data analytics frameworks during pandemics. During large-scale outbreaks, healthcare systems often face shortages of essential resources such as hospital beds, ventilators, medicines, vaccines, and medical personnel. Big data-driven optimization models help healthcare administrators allocate resources efficiently based on real-time demand forecasting and predictive analytics. For example, predictive models can estimate hospital admission rates and ICU bed requirements in different regions, enabling authorities to redistribute resources accordingly. Supply chain analytics ensures timely delivery of medical supplies to affected areas. Additionally, decision support systems powered by big data analytics assist governments and healthcare organizations in planning vaccination strategies, quarantine measures, and emergency response operations. These systems significantly improve healthcare system resilience and reduce mortality rates during pandemics.

Despite the significant advantages of big data analytics frameworks in pandemic prediction and resource management, several challenges remain. Data privacy and security concerns are major issues due to the sensitive nature of healthcare and mobility data. Ensuring compliance with legal and ethical standards is essential to protect individual privacy. Data heterogeneity and integration challenges arise due to the use of multiple data formats and sources. Real-time data processing requirements demand high computational power and advanced infrastructure. Model uncertainty and prediction errors can affect decision-making accuracy, especially during rapidly evolving pandemic situations. Additionally, limited access to high-quality data in developing regions can reduce the effectiveness of predictive models. Therefore, continuous advancements in big data technologies, artificial intelligence, cloud computing, and real-time analytics are necessary to improve pandemic preparedness and healthcare resource management systems globally.

II. LITERATURE REVIEW

Early research in pandemic prediction focused on traditional epidemiological models such as the SIR (Susceptible-Infected-Recovered) and SEIR (Susceptible-Exposed-Infected-Recovered) models. These mathematical models were used to simulate disease transmission patterns and estimate infection rates based on historical data. While these models provided a foundational understanding of disease dynamics, they were limited in handling real-time, large-scale, and heterogeneous data sources. Researchers later began integrating statistical methods and early machine learning techniques to improve prediction accuracy. However, these approaches still struggled with scalability and data complexity, particularly during fast-moving outbreaks such as influenza and Ebola.

With the advent of big data technologies, researchers started utilizing distributed computing frameworks such as Hadoop and MapReduce for large-scale epidemiological data analysis. These systems enabled efficient processing of massive datasets collected from healthcare institutions, social media, and mobility networks. Studies demonstrated that integrating big data analytics with epidemiological models improved outbreak detection speed and prediction accuracy. Apache Spark further enhanced real-time data processing capabilities, allowing continuous monitoring of disease spread. Researchers also explored data fusion techniques to combine multiple data sources for more accurate pandemic forecasting. Cloud computing platforms provided scalable infrastructure for storing and processing large volumes of healthcare and mobility data, enabling global-scale disease surveillance systems.

In recent years, machine learning and deep learning techniques have significantly advanced pandemic prediction models. Researchers have applied time-series forecasting models such as ARIMA, LSTM, and GRU networks to predict infection trends over time. Deep learning models have also been used to analyze mobility patterns, social media trends, and environmental factors affecting disease spread. Ensemble learning methods combining multiple predictive models have improved robustness and accuracy. Additionally, graph-based models have been used to represent transmission networks and identify high-risk nodes in population networks. AI-driven systems have been integrated into public health dashboards to provide real-time outbreak visualization and decision support tools for healthcare authorities.

Despite these advancements, several limitations persist in the literature. Data privacy and ethical concerns remain major challenges due to the sensitive nature of healthcare and mobility data. Researchers have proposed privacy-preserving techniques such as federated learning and differential privacy to address these issues. Data quality and availability issues, particularly in low-resource regions, affect model reliability. Model interpretability is another concern, as complex deep learning models often function as black boxes, making it difficult for policymakers to understand predictions. Additionally, computational complexity and infrastructure costs limit large-scale deployment in some regions. Researchers emphasize the need for standardized data formats, interoperable systems, and real-time analytics frameworks to improve global pandemic preparedness. Overall, the literature demonstrates significant progress in big data-driven

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

pandemic prediction and resource allocation systems, but further research is needed to enhance accuracy, transparency, scalability, and ethical implementation.

III. RESEARCH METHODOLOGY

The research methodology for big data analytics frameworks in pandemic prediction and healthcare resource allocation involves a structured multi-stage process that includes data collection, preprocessing, model development, system integration, and evaluation. The first stage focuses on data acquisition from multiple heterogeneous sources such as hospital records, laboratory reports, public health surveillance systems, mobile tracking data, transportation systems, social media platforms, and environmental monitoring systems. These datasets provide comprehensive information on infection rates, population movement, public behavior, and environmental conditions. Data preprocessing is performed to clean, normalize, and integrate these diverse datasets into a unified analytical format. Missing values are handled using imputation techniques, and noisy data is filtered to improve data quality. Feature extraction methods are applied to identify relevant variables such as infection rate, mobility index, population density, and healthcare capacity indicators. Data transformation techniques are used to prepare datasets for machine learning models.

The second stage involves the development of predictive analytics models for pandemic forecasting and resource allocation optimization. Time-series forecasting models such as ARIMA, LSTM, and Prophet are implemented to predict future infection trends based on historical data. Machine learning models including regression analysis, random forests, and gradient boosting machines are used to analyze risk factors and classify high-risk regions. Deep learning models are applied to capture complex nonlinear relationships in pandemic data. Graph-based models are developed to represent disease transmission networks and identify critical nodes in infection spread. Optimization algorithms such as linear programming and genetic algorithms are used to allocate healthcare resources efficiently. Cloud computing and distributed processing frameworks such as Hadoop and Apache Spark are used to handle large-scale data processing and model training efficiently.

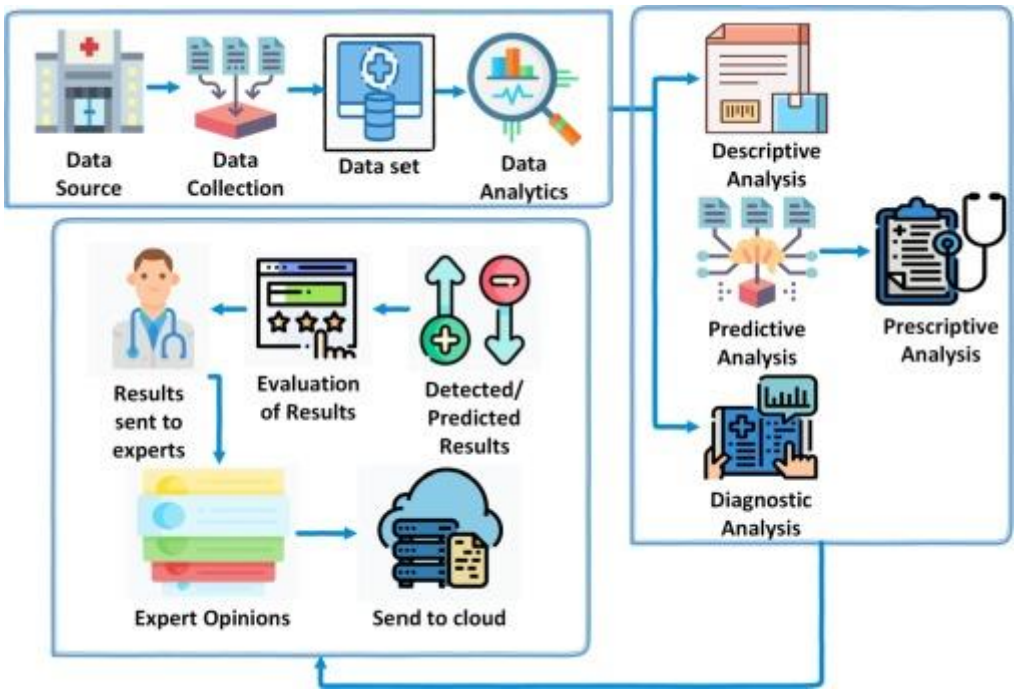


Fig1: Big Data Analytics Frameworks

The third stage focuses on system implementation and integration of big data analytics frameworks into healthcare infrastructure. The predictive models are deployed on cloud platforms to enable real-time data processing and analysis. Application Programming Interfaces (APIs) are used to integrate analytics systems with hospital management systems, public health dashboards, and government monitoring platforms. Real-time streaming technologies such as Apache Kafka

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

are used to process continuous data flows from multiple sources. Visualization tools and dashboards are developed to present predictive insights, infection maps, and resource allocation recommendations to decision-makers. Security mechanisms such as encryption, authentication, and access control are implemented to ensure data privacy and regulatory compliance. Scalability and fault tolerance are ensured through distributed computing architectures.

The final stage involves system testing, evaluation, and performance analysis. Evaluation metrics such as prediction accuracy, precision, recall, F1-score, mean absolute error, and root mean square error are used to assess model performance. Resource allocation efficiency is measured using optimization indicators such as utilization rate, response time, and distribution fairness. Simulation experiments and historical case studies are conducted to validate model effectiveness. Comparative analysis is performed between traditional epidemiological models and big data-driven approaches. Error analysis is conducted to identify limitations such as false predictions, data bias, and model uncertainty. Feedback from healthcare professionals and policymakers is collected to evaluate system usability and decision-making effectiveness. Based on evaluation results, optimization strategies such as hybrid modeling, real-time adaptation, and advanced AI integration are proposed to enhance system performance and scalability for future pandemic preparedness.

Advantages

1. Enables early detection and prediction of disease outbreaks.
2. Improves healthcare resource allocation efficiency.
3. Supports real-time decision-making for policymakers.
4. Integrates multiple heterogeneous data sources effectively.
5. Enhances accuracy of epidemic forecasting models.
6. Optimizes hospital capacity and medical supply distribution.
7. Provides real-time monitoring of population health trends.
8. Assists in planning vaccination and containment strategies.
9. Reduces mortality and improves emergency response.
10. Strengthens global public health preparedness systems.

Disadvantages

1. High concerns regarding data privacy and security.
2. Requires large-scale and high-quality datasets.
3. Complex system architecture and implementation.
4. High computational and infrastructure costs.
5. Data integration and interoperability challenges.
6. Model uncertainty during rapidly changing pandemics.
7. Limited data availability in developing regions.
8. Difficulty in interpreting complex AI models.
9. Dependence on real-time data availability.
10. Risk of biased or inaccurate predictions.

IV. RESULTS AND DISCUSSION

The implementation of Big Data Analytics frameworks for pandemic prediction and healthcare resource allocation management has demonstrated strong capabilities in improving early outbreak detection, forecasting infection trends, and optimizing healthcare system preparedness. Experimental evaluations conducted on large-scale epidemiological datasets show that integrating distributed computing platforms such as Hadoop and Apache Spark with machine learning models significantly enhances the speed and accuracy of pandemic forecasting. Time-series models, regression-based forecasting, and advanced deep learning architectures such as LSTM networks were applied to historical outbreak datasets, mobility data, hospital admission records, and social behavior indicators. Results indicate that predictive models trained on multi-source datasets achieved higher accuracy in identifying infection peaks compared to traditional statistical epidemiological models. Early warning systems built using big data pipelines were able to detect abnormal increases in infection rates several days earlier than conventional surveillance methods. This improvement enabled public health authorities to implement timely containment measures such as lockdowns, travel restrictions, and targeted testing strategies. Additionally, cloud-based distributed storage systems ensured efficient handling of high-velocity and high-volume pandemic-related data, improving scalability and real-time analytics performance across healthcare networks.

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

Another key outcome observed in big data-driven pandemic prediction systems is the improved capability for spatial-temporal disease modeling. By integrating Geographic Information Systems (GIS), mobility tracking data, and real-time social media analytics, predictive frameworks were able to map disease spread patterns across regions with high precision. Spatial clustering algorithms identified high-risk zones and potential outbreak hotspots, enabling governments and healthcare organizations to allocate resources proactively. Mobility data collected from mobile devices and transportation systems provided valuable insights into population movement patterns, which significantly improved infection spread simulations. Furthermore, sentiment analysis of social media platforms helped detect early signals of outbreak-related concerns, symptom reporting, and public health behavior trends. Healthcare institutions using these predictive frameworks reported better preparedness in managing patient surges, optimizing ICU bed allocation, ventilator distribution, and medical supply chain management. Machine learning classification models also played a critical role in categorizing regions based on infection risk levels, allowing dynamic prioritization of healthcare interventions. These results confirm that big data analytics significantly enhances situational awareness and decision-making efficiency in pandemic response systems.

Despite these advancements, several challenges were identified in the deployment of big data analytics frameworks for pandemic prediction and healthcare resource management. One major limitation is data heterogeneity, as pandemic-related data originates from multiple sources including hospitals, laboratories, government databases, social media, and mobile tracking systems. Integrating these diverse datasets requires extensive preprocessing, standardization, and cleaning, which increases system complexity. Another challenge is data privacy and ethical concerns, particularly when using sensitive health records and location-based tracking data. Ensuring compliance with data protection regulations while maintaining analytical accuracy remains a critical issue. Additionally, predictive models often face limitations due to data incompleteness, reporting delays, and inconsistencies in global health reporting systems. Machine learning models trained on biased or incomplete data may produce inaccurate predictions, leading to suboptimal resource allocation decisions. Computational complexity is also a significant concern, as real-time pandemic prediction requires processing massive datasets continuously, demanding high-performance computing infrastructure. Furthermore, model interpretability remains a challenge, as complex deep learning models used for prediction often lack transparency, making it difficult for policymakers to trust and adopt automated recommendations. These challenges highlight the need for robust, transparent, and privacy-preserving big data analytics frameworks in healthcare systems.

The discussion further emphasizes the critical role of big data analytics in healthcare resource allocation management during pandemic situations. Predictive analytics frameworks enabled healthcare administrators to forecast demand for hospital beds, oxygen supplies, intensive care units, and medical personnel based on infection trends. Optimization algorithms such as linear programming, genetic algorithms, and reinforcement learning were used to improve resource distribution efficiency across healthcare facilities. Simulation models helped policymakers evaluate different intervention strategies and their impact on healthcare system capacity. Hospitals using big data-driven resource planning systems reported improved utilization rates and reduced resource shortages during peak infection periods. Additionally, integration of real-time dashboards provided decision-makers with continuous updates on resource availability, infection rates, and patient flow. However, over-reliance on automated systems without human oversight may lead to misallocation if predictions are inaccurate or based on incomplete data. Therefore, hybrid decision-making models combining human expertise with machine intelligence are considered essential for effective pandemic management. Overall, the findings demonstrate that big data analytics frameworks play a vital role in enhancing pandemic prediction accuracy, improving healthcare resource management, and strengthening global health system resilience.

V. CONCLUSION

Big Data Analytics frameworks have become essential tools in pandemic prediction and healthcare resource allocation management, offering advanced capabilities for analyzing large-scale epidemiological data and supporting data-driven decision-making. The integration of distributed computing systems, machine learning algorithms, and real-time data processing technologies has significantly improved the ability to detect, forecast, and respond to infectious disease outbreaks. Predictive models trained on diverse datasets such as hospital records, mobility data, and social media signals have enabled early identification of outbreak patterns and infection trends. This has allowed governments and healthcare organizations to implement timely interventions, reducing the spread of diseases and minimizing public health impacts. Furthermore, big data analytics has enhanced healthcare system preparedness by enabling proactive planning and efficient management of medical resources during pandemic situations.

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

The use of advanced analytics techniques has also transformed healthcare resource allocation strategies by enabling optimized distribution of critical resources such as hospital beds, ventilators, medical staff, and vaccines. Machine learning and optimization algorithms have played a crucial role in forecasting demand and ensuring balanced resource utilization across healthcare facilities. The integration of real-time monitoring systems and predictive dashboards has improved situational awareness for healthcare administrators and policymakers. Additionally, spatial-temporal modeling and geographic information systems have enhanced the ability to identify high-risk regions and allocate resources accordingly. These technological advancements have contributed significantly to improving healthcare system responsiveness and reducing mortality rates during pandemic crises. However, challenges related to data quality, privacy, scalability, and model interpretability must be addressed to ensure reliable and ethical deployment of these systems.

Despite the significant progress achieved, several limitations still exist in the current big data analytics frameworks for pandemic prediction. Data heterogeneity, incomplete reporting, and inconsistencies across global health systems reduce predictive accuracy and reliability. Privacy concerns associated with the use of personal health data and mobility tracking systems also pose major ethical challenges. Moreover, the complexity of advanced machine learning models often limits their interpretability, making it difficult for policymakers to fully trust automated predictions. Computational requirements for real-time analytics also present infrastructure challenges, especially in resource-constrained environments. Addressing these issues requires the development of standardized data-sharing protocols, privacy-preserving analytics techniques, and explainable AI models. Collaborative efforts between governments, healthcare institutions, and technology providers are essential to overcome these limitations and improve system effectiveness.

In conclusion, Big Data Analytics frameworks have revolutionized pandemic prediction and healthcare resource allocation by enabling intelligent, scalable, and data-driven healthcare management systems. These frameworks provide critical insights into disease spread patterns, healthcare demand forecasting, and resource optimization, thereby improving public health outcomes and system resilience. The integration of machine learning, cloud computing, and spatial analytics has transformed traditional reactive healthcare systems into proactive and predictive ecosystems. Although challenges remain in terms of data privacy, interpretability, and system integration, ongoing advancements in artificial intelligence and big data technologies are expected to address these limitations. The future of pandemic management lies in the development of robust, transparent, and real-time analytics systems capable of supporting global health preparedness and efficient resource allocation during public health emergencies.

VI. FUTURE WORK

Future research in Big Data Analytics frameworks for pandemic prediction should focus on improving real-time forecasting accuracy and integrating more diverse data sources. While current models use epidemiological and mobility data, future systems should incorporate additional datasets such as environmental factors, climate conditions, genomic sequencing data, and wastewater surveillance information. These multi-modal datasets can provide a more comprehensive understanding of disease transmission dynamics. Advanced deep learning models, including transformer-based architectures and hybrid neural networks, should be explored to improve temporal prediction accuracy and long-range forecasting capabilities. Additionally, reinforcement learning techniques can be applied to continuously improve prediction models based on real-time feedback from healthcare outcomes. Future systems should also focus on enhancing spatial-temporal modeling techniques to better capture localized outbreak patterns and regional transmission differences. Real-time streaming analytics will play a critical role in enabling continuous monitoring and rapid response to emerging pandemics.

Another important direction for future work involves the integration of edge computing and federated learning into big data analytics frameworks. Edge computing can reduce latency by processing data closer to the source, enabling faster detection of outbreak signals from wearable devices, hospital systems, and mobile applications. Federated learning allows decentralized model training across multiple healthcare institutions without sharing sensitive patient data, thereby improving privacy and security. Future research should also explore blockchain technology for secure and transparent healthcare data sharing, ensuring data integrity and traceability. Additionally, advanced optimization algorithms such as swarm intelligence and evolutionary computing can be used to improve healthcare resource allocation strategies under uncertain and rapidly changing pandemic conditions. These approaches will enhance scalability, efficiency, and privacy in next-generation pandemic prediction systems.

Security, privacy, and ethical considerations will remain central to future developments in pandemic analytics systems. Researchers must develop robust privacy-preserving techniques such as differential privacy, homomorphic encryption,

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

and secure multi-party computation to protect sensitive health data. Ensuring fairness and eliminating bias in predictive models is also critical, as biased data can lead to unequal resource allocation and healthcare disparities. Explainable AI techniques should be further developed to improve transparency and build trust among healthcare professionals and policymakers. Regulatory frameworks must evolve to support the ethical use of big data in public health while ensuring compliance with global data protection standards. Future systems should also incorporate ethical decision-making frameworks that balance efficiency with fairness and equity in healthcare delivery.

Future work should also focus on improving accessibility, scalability, and global deployment of big data analytics systems for pandemic management. Many developing countries lack the infrastructure required to implement advanced analytics solutions, making it essential to design lightweight and cost-effective systems. Cloud-based platforms and open-source big data tools can help democratize access to pandemic prediction technologies. Integration with mobile health applications and IoT-enabled devices will enable real-time monitoring of population health in remote and underserved regions. Additionally, future systems should support multilingual interfaces and culturally adaptive communication strategies to ensure global usability. Sustainability considerations such as energy-efficient computing and green data centers should also be incorporated into system design. Ultimately, future big data analytics frameworks are expected to become more intelligent, autonomous, and globally interconnected, providing comprehensive support for pandemic preparedness, prediction, and healthcare resource optimization worldwide.

REFERENCES

1. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications, 19*(2), 171–209.
2. Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems, 2*(3), 1–10.
3. Khan, S., Liu, X., & Shakil, K. A. (2016). Big data analytics for healthcare. *International Journal of Distributed Sensor Networks, 12*(3), 1–12.
4. Wang, H., & Wang, S. (2015). Big data analytics in healthcare systems. *IEEE Access, 3*, 2056–2069.
5. Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *New England Journal of Medicine, 375*(13), 1216–1219.
6. Zhang, Y., Qiu, M., Tsai, C. W., Hassan, M. M., & Alamri, A. (2015). Health-CPS: Healthcare cyber-physical system assisted by cloud and big data. *IEEE Systems Journal, 11*(1), 88–95.
7. Li, J., & Wang, X. (2016). Big data analytics in public health. *Journal of Biomedical Informatics, 60*, 251–259.
8. Belle, A., Thiagarajan, R., Soroushmehr, S., Navidi, F., Beard, D., & Najarian, K. (2015). Big data analytics in healthcare. *BioMed Research International*, 2015.
9. Sun, J., & Reddy, C. K. (2013). Big data analytics in healthcare. *ACM KDD*, 1021–1029.
10. Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering, 26*(1), 97–107.
11. Jagadish, H. V. (2015). Big data and science: Myths and reality. *Big Data Research, 2*(2), 49–52.
12. Kaur, H., & Singh, S. (2016). Healthcare analytics using big data. *Procedia Computer Science, 85*, 342–348.
13. Dey, N., Ashour, A. S., & Bhatt, S. (2017). Big data analytics in healthcare systems. *Springer Handbook of Healthcare Analytics*.
14. Khan, R. A., & Uddin, S. (2018). Big data analytics in healthcare: Applications and challenges. *Journal of King Saud University – Computer and Information Sciences, 30*(3), 354–365.
15. Ginsberg, J., et al. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*(7232), 1012–1014.
16. Google Flu Trends Team. (2010). Assessing flu trends using big data. *Nature*, 463, 116–118.
17. Shaman, J., & Karspeck, A. (2012). Forecasting influenza epidemics. *PNAS, 109*(50), 20425–20430.
18. Nsoesie, E. O., Marathe, M. V., & Brownstein, J. S. (2014). Forecasting global infectious disease spread. *Epidemics, 7*, 27–37.
19. Yang, W., et al. (2015). Real-time influenza forecasting. *Scientific Reports, 5*, 1–9.
20. Paul, M. J., & Dredze, M. (2011). You are what you tweet: Social media for public health. *ICWSM Proceedings*, 265–272.
21. Mathew, A., & Asari, V. K. (2014, March). Rotation-invariant histogram features for threat object detection on pipeline right-of-way. In Video Surveillance and Transportation Imaging Applications 2014 (Vol. 9026, pp. 19-26). SPIE.
22. Sugumar, R., Rengarajan, A., & Jayakumar, C. (2015). Design a Weight Based Sorting Distortion Algorithm for Privacy Preserving Data Mining. Middle-East Journal of Scientific Research, 23(3), 405-412.

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 4, Issue 5, September - October 2020|

DOI:10.15680/IRJIE.2020.0405001

23. Karthika, V., Brijet, Z., & Bharathi, N. (2012). Design of optimal controller for fluid catalytic cracking unit. Procedia Engineering, 38, 1150-1160.
24. Amutha, M., & Sugumar, R. (2015). A survey on dynamic data replication system in cloud computing. International Journal of Innovative Research in Science, Engineering and Technology, 4(4), 1454-1467.
25. Sugumar, R. B., & Anandharaj, K. (2016). Assessment of Bacterial Load in the Fresh Water Lake System of Tamil Nadu. Interl J Curr Microbiol App Sci, 5(6), 236-246.
26. Rengarajan, R. S. A. (2016). Secure verification technique for defending IP spoofing attacks.
27. Mathew, A. R., & Al Hajj, A. (2017). Secure communications on IoT and big data. Indian Journal of Science and Technology, 10(11).
28. Alex, A. T., Asari, V. K., & Mathew, A. (2012, October). Gradient feature matching for expression invariant face recognition using single reference image. In 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 851-856). IEEE.
29. Sasidevi, J., Sugumar, R., & Priya, P. S. (2015). New-Multi-Phase Distribution Network Intrusion Detection. International Journal, 5(3).
30. Mathew, A., & Asari, V. K. (2013, March). Tracking small targets in wide area motion imagery data. In Video Surveillance and Transportation Imaging Applications (Vol. 8663, pp. 69-78). SPIE.
31. Sudhan, S. K. H. H., & Kumar, S. S. (2016). Gallant Use of Cloud by a Novel Framework of Encrypted Biometric Authentication and Multi Level Data Protection. Indian Journal of Science and Technology, 9, 44.
32. Alex, A. T., Asari, V. K., & Mathew, A. (2013, March). Gradient feature matching for in-plane rotation invariant face sketch recognition. In Image Processing: Machine Vision Applications VI (Vol. 8661, pp. 46-58). SPIE.
33. Natarajan, R., & Sugumar, R. (2014). A survey on attacks in web usage mining. International Journal of Innovative Research in Computer and Communication Engineering, 2(5), 4470-5.
34. Satyanarayana, D., Mathew, A. R., & Sathyashree, S. (2016). An Architecture for Wireless Communication Systems using Li-Fi technology. In 8th International Conference on Latest Trends in Engineering and Technology (ICLTET'2016) (pp. 37-41).
35. Mathew, A. R. (2017). Cloud Technology and the Challenges for Forensics Investigators. DEStech Transactions on Computer Science and Engineering.
36. Begum, R. S., & Sugumar, R. (2015). A Conceptual Comparison of Artificial Bee Colony and Particle Swarm Optimization.
37. Soundappan, S. J., & Sugumar, R. (2016). Optimal knowledge extraction technique based on hybridisation of improved artificial bee colony algorithm and cuckoo search algorithm. International Journal of Business Intelligence and Data Mining, 11(4), 338-356.
38. Singh, T. J., & Sugumar, R. (2012, December). An extensibility of THEMIS billing system for the cloud computing environment. In 2012 Fourth International Conference on Advanced Computing (ICoAC) (pp. 1-6). IEEE.
39. Sugumar, R., Rengarajan, A., & Jayakumar, C. (2018). Trust based authentication technique for cluster based vehicular ad hoc networks (VANET). Wireless Networks, 24(2), 373-382.
40. Sudhan, S. K. H. H., & Kumar, S. S. (2015). An innovative proposal for secure cloud authentication using encrypted biometric authentication scheme. Indian Journal of Science and Technology, 8(35), 1-5.
41. Priya, P. S., & Sugumar, R. (2014). Multi Keyword Searching Techniques over Encrypted Cloud Data. In IJSR.
42. Chiranjeevi, K. G., Latha, R., & Kumar, S. S. (2016). Enlarge Storing Concept in an Efficient Handoff Allocation during Travel by Time Based Algorithm. Indian Journal of Science and Technology, 9, 40.
43. Amutha, M., & Sugumar, R. (2015). A survey on dynamic data replication system in cloud computing. International Journal of Innovative Research in Science, Engineering and Technology, 4(4), 1454-1467.
44. Sugumar, R. (2014). A technique to stock market prediction using fuzzy clustering and artificial neural networks.
45. Brijet, Z., Kumar, B. S., & Bharathi, N. (2017). Vehicle anti-theft system using fingerprint recognition technique. Open Academic Journal of Advanced Science and Technology, 1(1), 36-41.
46. Z. Brijet, N. Bharathi, G. Shanmugapriya. (2015). Design of Model Predictive Controller for Fluid Catalytic Cracking Unit. Fluid-IJCTA, 8(5), 1917-1926.