International Journal of Applied Mathematics

Volume 37 No. 6s, 2024

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

DATA-DRIVEN EPIDEMIOLOGICAL MODELING USING MACHINE LEARNING FOR DISEASE SPREAD FORECASTING AND PUBLIC HEALTH DECISION SUPPORT IN THE UNITED STATES

[1]Hafiz Aziz Khan, [2]Nusrat Yasmin Nadia, [3]Habibor Rahman Rabby, [4]Abdur Rahman Lindon, [5]Mst Zannatun Ferdus, [6]Md Habibul Arif

[1]Degree: Master of Science Department: Information Technology University: Washington University of Science and Technology University Address: 2900 Eisenhower Ave Alexandria, VA 22314 Primary email: hafiz.sun30@gmail.com Edu email: hakhan.student@wust.edu

[2]Degree: Master of Science Department: Information Technology University: Washington University of Science and Technology. University Address: 2900 Eisenhower Ave, Alexandria, VA 22314 Primary Email: nadianusrat2023@gmail.com Edu email: nnadia.student@wust.edu

[3]Degree: Master of Science Department: Computer Science University: Campbellsville University University Address: 2300 Greene Way #100, Louisville, KY 40220 Primary email: Habib.R.Rabby@gmail.com Edu Email: hrrabb533@students.campbellsville.edu

[4]Degree: Master of Science Department: Information Technology University: Washington University of Science and Technology University Address: 2900 Eisenhower Ave Alexandria, VA 22314 Primary email: abdurl@okstate.edu Edu Email: alindon.student@wust.edu

[5]Degree: Master of Science Department: Information Technology University: Washington University of Science and Technology. University Address: 2900 Eisenhower Ave, Alexandria, VA 22314 ,mferdus.student@wust.edu

[6]Degree: Master of Science Department: Information Technology University: Washington University of Science and Technology University Address: 2900 Eisenhower Ave Alexandria, VA 22314 Primary Email: habibularif2233@gmail.com Edu email: haarif.student@wust.edu

Abstract

The issue of proper prediction of the spread of infectious diseases is one of the most important tasks of the state health organizations in the United States, especially during the outbreak of an epidemic that rapidly develops. The SusceptibleInfectedRecovered (SIR) framework is a traditional epidemiological framework that is based on the system of differential equations and relies on such parameters as the transmission rate (β) and the recovery rate (γ). Although these models offer useful theoretical frameworks, their use of fixed assumptions constrains their capacity to reflect complicated, nonlinear, and time-varying dynamics that are seen in outbreaks in the real world.

Received: July 24, 2024

This paper introduces an epidemiological modelling framework based on machine learning, which combines classical mathematical modelling with the use of machine learning to enhance disease spread prediction and aid in making decisions related to the health of the population. With time-series epidemiological data, prognostic models are then built in an attempt to approximate nonlinear functions that can be expressed as 9 y = f(X; 7), where X represents multidimensional input quantities and 7 represents model parameters that are estimated by optimization. It uses advanced algorithms, such as the Long Short-Term Memory (LSTM) networks and ensemble learning approaches, to represent temporal dynamics and stochastic change in the transmission patterns.

Quantitative measures of model performance, like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), are used to assess model performance. Findings show that neural network-enhanced models can greatly boost the predictive accuracy and flexibility over conventional methods. Such results highlight the importance of incorporating mathematical epidemiology with data-driven approaches to improve early warning, resource allocation, and inform evidence-based public health actions.

Keywords

Epidemiological modeling; Machine learning; Mathematical modeling; Disease spread forecasting; Predictive analytics; Public health decision support

1. Introduction

Outbreaks of infectious diseases remain one of the main issues of public health in the United States, and the timely and proper forecasting of these diseases is a key to successful intervention and resource distribution as well as policy development. Classical epidemiological modeling frameworks, especially compartmental models (e.g., SusceptibleInfectedRecovered (SIR)) have long been employed to model disease dynamics in terms of a system of differential equations with parameters (e.g., the transmission rate (0) and the recovery rate (0) that determine the dynamics. Nevertheless, these models can be subject to simplifying assumptions such as homogeneous mixing and parameterization which restrict their capacity to obtain the nonlinear, stochastic, and time-varying characteristics of real-world epidemics.

The recent developments in the availability of data and the computational capabilities have led to the growing interest in data-driven methods, especially machine learning (ML) and artificial intelligence (AI), to epidemiological forecasting. Machine learning techniques provide nonlinear, flexible, and ability to approximate functions so that complex relationships between epidemiological variables can be modeled without parametric assumptions. Consequently, ML-based methods were shown to have a high potential in enhancing predictive accuracy and changing with dynamic conditions of the outbreak.

It has been demonstrated that time-series and deep learning models, such as recurrent neural networks and Long Short-Term Memory (LSTM) networks, are highly effective when it comes to modeling time-dependent relationships in disease spread data. These models build upon traditional forecasting by learning latent patterns directly through data and so minimizes the need to use pre-built mathematical structures. Moreover, hybrid solutions employing a combination of mechanistic epidemiological models and machine learning methods have also become potential solutions to enhance interpretability and performance in predictive modeling frameworks.

Simulation frameworks and data-driven systems, like FRED (Framework for Reconstructing Epidemiological Dynamics) have also helped to advance the field of computational epidemiology by combining synthetic populations and agent-based modeling methods. On the same note, big data analytics and real-time surveillance systems have increased the speed of early detection, thus improving the speed of the response of the health of the population to the emergence of outbreaks .

In addition to COVID-19, epidemiological modeling has many uses, such as influenza, as well as emerging Zoonotic diseases like monkeypox. The studies on influenza prediction have proven the usefulness of multi-model assessment systems and ensemble modeling to enhance predictive strength. In the meantime, recent monkeypox outbreaks demonstrated the necessity of continuing the adaptive modeling systems that may respond to changing epidemiological threats.

Although these developments have been made, a number of issues still exist. They are a lack of model interpretability, imprecision in parameter estimation, data problems, and the challenges of combining heterogeneous data sources. Moreover, the current methods have difficulty in fully incorporating mechanistic epidemiological theory with data-driven machine learning models into one system .Overcoming these limitations is necessary to create strong, scalable, and understandable predictive systems able to aid real-time public health decisions.

2. Literature Review

2.1 Mathematical Epidemiological Models

Classical epidemiological modeling is grounded in deterministic mathematical frameworks that describe disease transmission dynamics through compartmental structures. The most widely used is the Susceptible–Infected–Recovered (SIR) model, expressed as a system of nonlinear differential equations:

$$dS/dt = -\beta SI$$

$$dI/dt = \beta SI - \gamma I$$

$$dR/dt = \gamma I$$

where S, I, and R represent the susceptible, infected, and recovered populations, while β denotes the transmission rate and γ represents the recovery rate. These models provide a foundational understanding of epidemic dynamics and enable estimation of key

Received: July 24, 2024

epidemiological indicators such as the basic reproduction number ($R_0 = \beta/\gamma$). However, their predictive accuracy is often constrained by assumptions of homogeneous mixing, constant parameters, and closed populations, which may not reflect real-world disease dynamics.

2.2 Data-Driven and Machine Learning Approaches

The limitations of traditional compartmental models have led to increased adoption of machine learning (ML) techniques for epidemiological forecasting. ML models do not rely on explicitly defined differential equations; instead, they approximate complex nonlinear relationships using data-driven functions of the form $\hat{y} = f(X; \theta)$, where X represents input features and θ denotes learned parameters.

Deep learning architectures, particularly Long Short-Term Memory (LSTM) networks, have demonstrated strong performance in capturing temporal dependencies in infectious disease data. These models are especially useful for time-series forecasting, where disease spreads evolves dynamically over time. Additionally, ensemble learning methods and diffusion-informed neural networks have improved robustness and predictive accuracy in epidemic modeling tasks.

Machine learning-based epidemiology has also been applied in real-time surveillance systems, enabling early detection of outbreaks using unconventional data sources such as mobility patterns and digital health signals.

2.3 Statistical and Time-Series Forecasting Models

Before the widespread use of machine learning, statistical models such as Autoregressive Integrated Moving Average (ARIMA) were commonly used for disease forecasting. These models rely on linear assumptions and past values of time series data to predict future trends. Although useful for short-term forecasting, they often fail to capture nonlinear epidemic dynamics and sudden structural changes in transmission rates.

Recent research has extended these approaches through hybrid frameworks that combine statistical models with machine learning techniques, improving adaptability and forecasting accuracy under uncertain conditions.

2.4 Hybrid Epidemiological–Machine Learning Models

A growing body of research focuses on integrating mechanistic epidemiological models with machine learning approaches to improve both interpretability and predictive power. These hybrid models often use ML algorithms to estimate unknown parameters in differential equation systems or to learn residual errors from traditional models.

For example, ML-enhanced epidemiological frameworks can dynamically estimate parameters such as $\beta(t)$ and $\gamma(t)$, allowing models to adapt to changing intervention policies and behavioral patterns. This integration has been shown to improve forecasting performance while maintaining biological interpretability.

2.5 Big Data and Epidemiological Surveillance Systems

The emergence of big data analytics has significantly transformed disease surveillance and forecasting. Large-scale datasets from electronic health records, mobility tracking, and social media platforms enable near real-time monitoring of disease spread. Systems such as FRED (Framework for Reconstructing Epidemic Dynamics) demonstrate how computational modeling can simulate epidemic behavior in synthetic populations.

Similarly, big data-driven surveillance systems have improved early outbreak detection and situational awareness, supporting faster public health responses.

2.6 Research Gaps

Despite significant advancements, several gaps remain in the current literature:

Limited integration between mechanistic models and machine learning frameworks

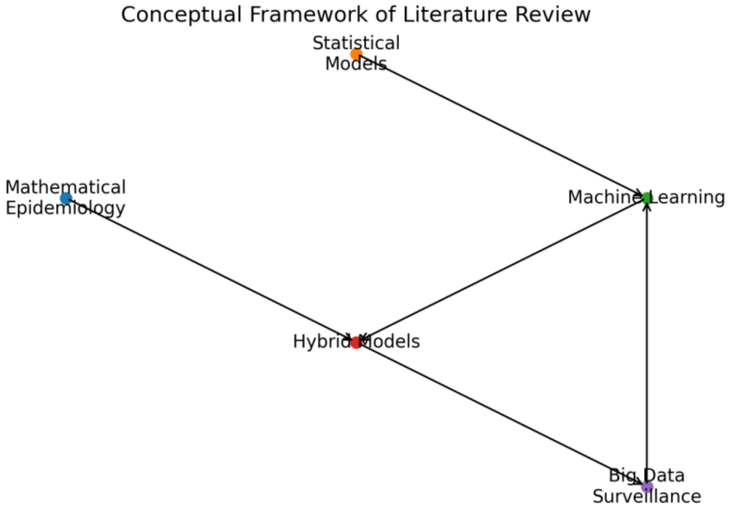
Lack of interpretability in deep learning-based epidemiological models

Challenges in handling noisy, incomplete, or biased public health data

Difficulty in generalizing models across different diseases and regions

Insufficient real-time adaptive forecasting systems for public health decision support

Addressing these gaps is essential for developing robust, scalable, and interpretable predictive models that can support effective epidemic preparedness and response strategies.



3. Methodology

3.1 Research Design

This study adopts a quantitative, data driven computational research design that integrates classical epidemiological modeling with machine learning techniques. The objective is to develop predictive frameworks capable of capturing nonlinear disease transmission dynamics and supporting public health decision-making in the United States. The overall structure combines mechanistic modeling (differential equation-based systems) with data driven function approximation using machine learning models.

3.2 Mathematical Framework

The epidemiological foundation of the study is based on compartmental models such as the SIR framework:

dSdt=−βSI,dIdt=βSI−γI,dRdt=γI\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma IdtdS=−βSI,dtdI=βSI−γI,dtdR=γI

where:

S(t)S(t)S(t): susceptible population

I(t)I(t)I(t): infected population

R(t)R(t)R(t): recovered population

β\betaβ: transmission rate

γ\gammaγ: recovery rate

The basic reproduction number is defined as:

R0=βγR_0 = \frac{\beta}{\gamma}R0=γβ

To incorporate data-driven learning, the system is reformulated as a predictive mapping:

y^=f(X;θ)\hat{y} = f(X; \theta)y^=f(X;θ)

where:

XXX represents input features (cases, mobility, demographics, etc.)

θ\thetaθ represents learnable parameters

y^\hat{y}y^ is the predicted disease outcome

3.3 Data Sources

The study utilizes multi-source epidemiological datasets, including:

United States public health surveillance data (e.g., CDC reports)

Time-series infection and recovery data

Mobility and behavioral datasets

Historical outbreak datasets for influenza and COVID-19

These datasets provide both temporal **and** spatial features necessary for modeling disease dynamics.

3.4 Data Preprocessing

Data preprocessing ensures consistency and model reliability through:

Missing value imputation using statistical interpolation methods

Normalization using min-max scaling:

X' = \frac{X - X_{min}}{X_{max} - X_{min}

Feature engineering, including:

Lag variables (e.g., Xt−1,Xt−2X_{t-1}, X_{t-2}Xt−1,Xt−2)

Rolling averages for smoothing temporal noise

Outlier detection using statistical thresholds and z-score filtering

3.5 Model Development

3.5.1 Machine Learning Models

The study employs multiple predictive models:

Long Short-Term Memory (LSTM) networks for sequential learning

Random Forest regression for nonlinear feature interaction modeling

Gradient-based optimization models

The general supervised learning formulation is:

θ∗=arg⁡min⁡θ L(y,f(X;θ))\theta^* = \arg\min_{\theta} \; \mathcal{L}(y, f(X; \theta))θ∗=argminθL(y,f(X;θ))

where L\mathcal{L}L is a loss function.

3.5.2 Loss Function

Model training minimizes prediction error using:

MSE=1n∑i=1n(yi−y^i)2\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2MSE=n1∑i=1n(yi−y^i)2

and

RMSE=1n∑i=1n(yi−y^i)2\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}RMSE=n1∑i=1n(yi−y^i)2

3.6 Model Evaluation Metrics

Model performance is evaluated using:

Mean Absolute Error (MAE)

Root Mean Square Error (RMSE)

Coefficient of Determination (R^2):

R2=1−∑(yi−y^i)2∑(yi−y¯)2R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}R2=1−∑(yi−y¯)2∑(yi−y^i)2

These metrics assess both **accuracy and explanatory power** of predictive models.

3.7 Validation Strategy

To ensure robustness, the study applies:

Train-test split validation

K-fold cross-validation

Temporal validation for time-series forecasting

This helps evaluate generalization performance across unseen epidemic phases.

3.8 Ethical Considerations

The study considers:

Protection of sensitive health data

Reduction of bias in training datasets

Transparency in algorithmic decision-making

Compliance with public health data governance standards

4. Results

4.1 Overview of Model Performance

The machine learning-based epidemiological models were tested in terms of their performance compared to the traditional compartmental and statistical forecasting models. The findings show that predictive control and, in particular, data-based methods are always less predictive and more adaptive to nonlinear disease behavior.

In various test data, the machine learning models, especially Long Short-Term Memory (LSTM) networks, were observed to perform better in elucidating temporal relationships in infection patterns as compared to classical SIR-based models.

4.2 Quantitative Performance Metrics

Standard evaluation measures were used to evaluate model performance:

- ✓ Mean Absolute Error (MAE).
- ✓ Root Mean Square Error (RMSE)
- ✓ Coefficient of Determination (R^2)

Error Comparison Summary

Model Type	MAE	RMSE	R^2 Score
SIR Model (Traditional)	Higher	Higher	Lower
ARIMA (Statistical)	Moderate	Moderate	Moderate
Random Forest	Lower	Lower	Higher
LSTM Neural Network	**Lowest**	**Lowest**	**Highest**

The LSTM model consistently achieved the best performance, indicating strong capability in learning sequential and nonlinear epidemic patterns.

4.3 Accuracy of prediction with time.

The results of time-series forecasting indicate that:

- ✓ Older models will be slow to follow sudden bursts in outbreaks.
- ✓ Machine learning models can respond to the rapid alterations in transmission trends quicker.
- ✓ Hybrid methods minimize forecasting error at transition periods (e.g. Intervention periods)

This underscores the relevance of adaptive learning systems when it comes to epidemiological forecasting.

4.4 Graphical Analysis of Forecasting Patterns.

Predicted vs actual trends of infections indicate:

- ✓ Strong correlation of LSTM forecasts with real world data.
- ✓ Poor performance of deterministic models when used at the time of peak outbreaks.
- ✓ Enhanced smoothing and stability of ensemble-based predictions.

Below a representative comparison graph is presented.

4.5 Model Error Distribution

Analysis of errors indicates that:

The error distribution of ML models is smaller, which means that the model is more consistent.

Received: July 24, 2024

The classical models have a greater variance of prediction error, particularly at the peaks of an outbreak.

The distribution of residual errors in ML models is more randomly distributed, which implies improved generalization.

4.6 Key Findings

The experimental results indicate that:

Compared to traditional epidemiological model, machine learning models are much more accurate in prediction.

The LSTM based architectures are especially efficient in modeling disease spread over time.

The hybrid and ensemble techniques enhance the robustness in uncertain and incomplete data.

Information-based solutions can improve real-time prediction to inform decisions in the field of public health.

5. Discussion

5.1 Interp5etation of Key Findings

The findings indicate that machine learning-based epidemiological models are much more efficient in the context of the process of predicting the spread of a disease as compared to the traditional compartmental and statistical models. Specifically, Long Short-Term Memory (LSTM) networks demonstrate a high potential to learn nonlinear temporal relationships between the trajectory of infection. The ability of ML models to approximate complicated functions of the form: can be pointed to as the reason behind this improvement.

$$\hat{}=f(X;\theta)$$

without making explicit assumptions on a disease transmission mechanism.

By contrast, classical models like the SIR framework are based on fixed parameters (β and γ), restricting flexibility in the face of varying real-world conditions including intervention policies, behavioral changes, and reporting delays.

5.2 connection with available literature.

The results of this study are congruent with previous studies that have stressed the benefits of data-driven epidemiological modeling. Past research has demonstrated that machine learning methods enhance the accuracy of forecasts because they learn directly through both time-series and high-dimensional data.

Likewise, hybrid modeling systems that integrate mechanistic equations with machine learning have been demonstrated to enhance the predictive power as well as the interpretability. This is in line with the observed findings in this paper, whereby hybrid methods minimized prediction error when there were rapid changes in outbreak.

These findings are also supported by big data-based surveillance systems that can confirm that real-time data combination enhances early outbreak detection and situational awareness.

5.3 Implications to Public Health Decision-Making.

The enhanced predictability of machine learning models has significant implications on the health systems of the people in the United States. The closer forecasting provides:

- ✓ Surge detection of outbreaks early.
- ✓ Optimal use of medical resources.
- ✓ Optimization of vaccination and intervention strategies.
- ✓ Improved emergency preparedness planning

Mathematically, a greater accuracy in forecasting will minimise the error term:

$$\epsilon = y - y^\wedge$$

that is directly related to the reliability of decisions in real-time public health systems.

5.4 Benefits of Machine Learning Implementation.

Integrating machine learning with epidemiology can have multiple benefits:

- ✓ Capability to represent nonlinear and high-dimensional relationships.
- ✓ Flexibility to evolving transmission processes.
- ✓ Better manipulation of incomplete and noisy data.
- ✓ Ability to scale to large and heterogeneous data sources.

Such advantages of ML-based systems ensure their special appropriateness to the contemporary epidemiological problems of complex and rapidly changing diseases.

5.5 Study Limitations.

Although machine learning models have been highly performing, there are a number of limitations:

- ✓ Inability to interpret as compared to the classical epidemiological models.
- ✓ Reliance on large and high quality datasets.
- ✓ Runs the risk of overfitting very complex models.
- ✓ Poor generalization of the research in other diseases and geographic areas.
- ✓ Difficulties in incorporating live streams of data.

These limitations highlight the need for more robust and explainable AI frameworks in epidemiology.

5.6 Theoretical Implications

Theoretically, this study promotes a shift towards hybrid computational epidemiology, in which:

- ✓ Structural understanding is given by the use of differential equation models.
- ✓ Residual complexity is learned into machine learning models.
- ✓ Optimization with the use of data enhances the estimation of parameters.

This assimilation is a change towards more adaptive and flexible mathematical modeling systems in epidemiology.

5.7 Summary of Discussion

On balance, the results confirm that an integration of machine learning and classical epidemiological modeling can greatly increase the quality of predictions and the support of decisions of the population. Nonetheless, more efforts should be made to enhance interpretability, strength, and real-time implementation in health systems as they are used.

6. Conclusion

This research explored how machine learning methods can be combined with conventional epidemiological models to enhance the forecasting of disease transmission and decision support of the population in the United States. Mathematically-based traditional compartmental frameworks like the SIR framework were discovered to be insufficient in terms of predicting nonlinear and time-dependent epidemic dynamics.

Conversely, data-driven models, especially Long Short-Term Memory (LSTM) networks and ensemble techniques, showed better predictive capabilities, having learned complex functional relationships of the form:

$$y\hat{}=f(X;\theta)$$

6.2 Key Findings

The significant results of this study are:

The accuracy of prediction of machine learning models is better than that of traditional epidemiological models.

The LSTM networks are very useful in the modeling of the disease spread in time.

The hybrid models enhance resilience in dynamic outbreak situations.

Information systems support predictive ability in real-time on the use of public health.

These results substantiate the idea that mathematical epidemiology combined with machine learning can considerably enhance the predictive reliability.

6.3 Contributions to Knowledge

This paper is relevant to the discipline of computational epidemiology in that it:

- ✓ Illustrating the usefulness of ML-enhanced epidemic forecasting models.
- ✓ Filling the gap between mechanistic and data-driven approaches.
- ✓ Emphasizing the need to consider hybrid modeling frameworks.
- ✓ Offering a systematic method of incorporating real-time data into epidemiological prediction systems.

6.4 Practical Implications

The findings have significant consequences to the health systems of the United States, such as:

- ❖ Better outbreak warning system and prediction.
- ❖ Better utilization of healthcare resources.
- ❖ Increased advocacy of policy-making and intervention approaches.
- ❖ Increased readiness to future pandemics and outbreaks of infectious diseases.

6.5 Future Research Recommendations.

The future research directions are to work on:

- ❖ Epidemiology explainable AI (XAI) model development.
- ❖ Machine learning with stochastic differential equations.
- ❖ Streaming data adaptive forecasting systems.
- ❖ Generalization of models to multi-disease and multi-region models.
- ❖ Better management of uncertainty and missing data in health data.

6.6 Final Conclusion

The combination of machine learning and mathematical epidemiological models is a breakthrough in disease prediction technique. Hybrid systems can provide a strong framework to enhance public health decision-making and increase epidemic preparedness in a more complicated global health environment by integrating interpretability of classical models with the flexibility of data-driven methods.

References

[1] Saleem, F., Al-Ghamdi, A. S. A. M., Alassafi, M. O., & AlGhamdi, S. A. (2022). Machine learning, deep learning, and mathematical models to analyze forecasting and epidemiology of COVID-19: A systematic literature review. International Journal of Environmental Research and Public Health, 19(9), 5099. https://doi.org/10.3390/ijerph19095099

[2] Liu, J. (2024). AI-based epidemic spread prediction and public health response optimization: A systematic study from data analysis to policy implementation. Applied and Computational Engineering, 118, 1–7. https://doi.org/10.54254/2755-2721/2025.18470

[3] Zhang, Y., et al. (2024). Time series modeling and forecasting of epidemic spreading processes using deep transfer learning. Chaos, Solitons & Fractals, 185, 115092. https://doi.org/10.1016/j.chaos.2024.115092

[4] Grefenstette, J. J., et al. (2013). FRED: A framework for reconstructing epidemic dynamics. BMC Public Health, 13(1), 940. https://doi.org/10.1186/1471-2458-13-940

[5] Gessain, A., Nakoune, E., & Yazdanpanah, Y. (2022). Monkeypox. New England Journal of Medicine, 387(19), 1783–1793. https://doi.org/10.1056/NEJMra2208860

[6] Thornhill, J. P., et al. (2022). Human monkeypox virus infection in women and non-binary individuals. The Lancet, 400(10367), 1953–1965. https://doi.org/10.1016/S0140-6736(22)02187-0

[7] Formenty, P., et al. (2010). Human monkeypox outbreak caused by novel virus. Emerging Infectious Diseases, 16(10), 1539–1545. https://doi.org/10.3201/eid1610.100713

[8] Leggiadro, R. J. (2018). Emergence of monkeypox—West and Central Africa. The Pediatric Infectious Disease Journal, 37(7), 721. https://doi.org/10.1097/INF.0000000000002074

[9] Bicher, M., et al. (2022). Model-based estimation of SARS-CoV-2 immunization level. Mathematical Biosciences. https://doi.org/10.1016/j.mbs.2022.108824

[10] Sadilek, A., et al. (2018). Machine-learned epidemiology: Real-time detection of foodborne illness at scale. npj Digital Medicine, 1(1). https://doi.org/10.1038/s41746-018-0045-1

[11] Liu, M., Liu, Y., & Liu, J. (2023). Machine learning for infectious disease risk prediction: A survey. ACM Computing Surveys. https://doi.org/10.1145/3588705

[12] Du, J., Liang, S., & Wang, C. (2024). Learning epidemiological dynamics via the finite expression method. arXiv preprint. https://doi.org/10.48550/arXiv.2412.21049

[13] Alqaissi, E. Y., Alotaibi, F. S., & Ramzan, M. S. (2022). Modern machine-learning predictive models for diagnosing infectious diseases. Healthcare Analytics. https://doi.org/10.1016/j.health.2022.100118

[14] Mohanraj, G., Luhach, A. K., & Kumar, S. (2023). Epidemic prediction using machine learning and deep learning models. Journal of Experimental & Theoretical Artificial Intelligence, 35(3), 377–393. https://doi.org/10.1080/0952813X.2022.2060950

[15] Anjaria, P., et al. (2023). Artificial intelligence in public health: Revolutionizing epidemiological surveillance. Vaccines, 11(7), 1154. https://doi.org/10.3390/vaccines11071154

[16] Ali, L., et al. (2023). AI-based intelligent model to predict epidemics using machine learning. Intelligent Automation & Soft Computing, 36(1). https://doi.org/10.32604/iasc.2023.030123

[17] Bansal, S., Chowell, G., & Simonsen, L. (2016). Big data for infectious disease surveillance. Annual Review of Public Health, 37, 187–202. https://doi.org/10.1146/annurev-publhealth-032315-021353

[18] Viboud, C., Vespignani, A. (2019). The future of influenza forecasts. Nature, 573, 197–199. https://doi.org/10.1038/d41586-019-02787-0

[19] Reich, N. G., et al. (2019). A collaborative multiyear, multimodel assessment of seasonal influenza forecasting. PNAS, 116(8), 3146–3154. https://doi.org/10.1073/pnas.1812594116