Innovations in Intelligent Systems and Advanced Engineering
Vol. 1, No. 2, 2025 (August)

Predictive Modeling of Infectious Disease Dynamics Using Hybrid Approach

M. Santhanalakshmi
Department of Computer Science, Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Chennai, Tamil Nadu, India
mslakshmi25@gmail.com

N. Abirami
Department of Information Technology, R.M.D. Engineering College, Kavaraipettai, Chennai, Tamil Nadu, India
abirami.it@rmd.ac.in

S. Asha
Department of Information Technology, S.A. Engineering College, Chennai, Tamil Nadu, India
ashas@saec.ac.in

D. Sailaja
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Vijayawada, Andhra Pradesh, India
dsailajaklu@gmail.com

M. Subashree
Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering, Tiruchirappalli, Tamil Nadu, India
subashreemcse@krce.ac.in

R. Aiyshwariya Devi
Department of Artificial Intelligence and Data Science, R.M.K. College of Engineering and Technology, Chennai, Tamil Nadu, India
aiyshwariyadevi@gmail.com

Presitha Aarthi M
Department of Computer Science and Engineering, Shree Sathyam College of Engineering and Technology, Salem, Tamil Nadu, India
presithacse@gmail.com

Ravindaran Maraya
Department of Indian Studies, Faculty of Arts and Social Sciences, Universiti Malaya, Kuala Lumpur, Malaysia.
ravindaranm@um.edu.my

Abstract: Modern healthcare generates vast amounts of clinical data stored in medical databases. Extracting meaningful information from this data and making informed decisions for disease diagnosis and treatment has become increasingly essential. In this study, we propose a Covid-19 infection prediction system designed to assist in preventing fatalities. The system analyzes patient records to identify effective treatments and deliver improved outcomes. The proposed framework provides physicians with a powerful tool to make data-driven decisions by utilizing historical datasets for analysis. The study employs various data mining algorithms to predict Covid-19 infections and evaluate the most effective prediction method. The predictive accuracy of Naïve Bayes, Support Vector Machine (SVM), and a hybrid approach was assessed. Experimental results reveal that the hybrid model delivers superior performance compared to the individual algorithms, achieving higher accuracy in Covid-19 prediction.

Keywords: Covid-19, Naïve bayes, support vector machine, healthcare, prediction

I. INTRODUCTION

At a time when humanity is advancing rapidly in innovation while simultaneously confronting environmental crises, the world has also faced the devastating outbreak of a new viral disease: Covid-19. The World Health Organization (WHO) labelled COVID-19 a worldwide pandemic due of its widespread and severe health effects. The spread of infection is influenced by multiple factors, including contaminated sources, population vulnerability, and viral latency [1]. Such outbreaks pose severe threats to human life and global society [2–3]. Although no precise medication exists to fully combat the epidemic, various treatment strategies—including antiviral drugs, plasma therapy, and clinical trials of potential vaccines—have been explored [4]. While these interventions provide some relief, the design and development of effective vaccines remain crucial. Social distance, mask-wearing, hand cleanliness, and tight surveillance must be used to prevent illness transmission until viable vaccinations and therapies are available [6–7].

A One Health strategy may reduce pandemic transmission at the human–animal interaction. This interdisciplinary approach combines human and veterinary knowledge to improve illness prevention and management [8]. Furthermore, advancements in technology can be leveraged to predict, track, and manage epidemic spread [9]. Among these, computational intelligence and data-driven methods play a vital role in forecasting infection trends, enabling authorities to make proactive decisions. This study predicts Covid-19 progression using deep learning and hybrid methods, By applying advanced computational techniques, it is possible to identify infection patterns, improve treatment recommendations, and ultimately reduce the impact of the pandemic.

Several studies have predicted and analysed Covid-19 transmission kinetics. This section discusses infection forecasting-related literature. A sequential Monte Carlo simulation model estimates the virus's rapid transmission rate by computing the daily mean reproduction number (R_t) given case proportions and confirmed case probabilities. When transmission is homogeneous, the outbreak risk increases significantly [10]. To control the spread, other models have been proposed that incorporate isolation measures and contact tracing. However, delays between the onset of symptoms and case detection often increase the likelihood of transmission [11].

AI, ML, DL, and Big Data analytics are used to control pandemics. These methods have been employed in areas including early infection detection, contact tracing, treatment optimization, vaccine development, and future case projection [12–13]. For example, a K-Means clustering algorithm has been used to categorize countries based on Covid-19 infection severity, helping identify nations at higher risk of future outbreaks [14]. Similarly, based on migration patterns from China, clustering methods were applied to forecast transmission rates in Singapore, providing insights into potential community spread. However, these models require large datasets to improve accuracy and generalizability [15].

AI-driven predictive models have also been compared across domains to assess their ability to control pandemic spread. Deep learning models, when trained on massive datasets, can automate diagnosis, treatment, and patient monitoring, thereby assisting healthcare professionals [16–17]. Recurrent Neural Networks (RNNs) excel in non-linear and dynamic time-series forecasting [18–20]. These studies demonstrate the versatility of computational intelligence approaches and provide insights that can be adapted for epidemic prediction.

II. METHODS AND MATERIALS

The proposed work aims to develop a **Covid-19 infection prediction framework** capable of extracting useful information from patient data based on symptoms. Due of the abundance of unstructured medical data across many illnesses, typical data mining methods often fail. The framework therefore seeks to uncover hidden relationships between symptoms and infection patterns, enabling more accurate prediction of Covid-19. Figure 1 shows the proposed hybrid covid-19 prediction architecture.

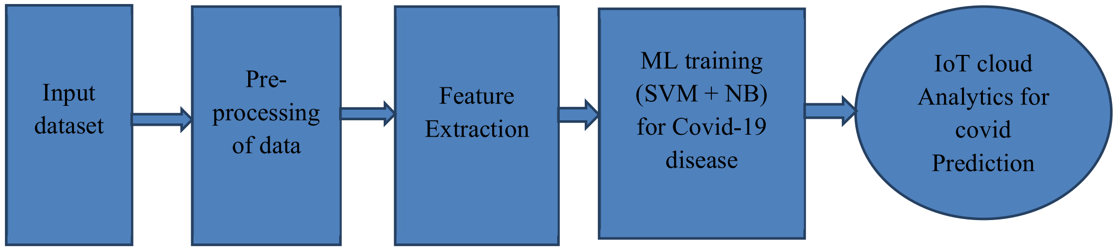

Figure 1: Proposed hybrid framework for predictive modeling of infectious disease detection

Data Source
The experimental analysis uses the **UCI dataset**, consisting of 1,080 patient records.

Preprocessing
In the preprocessing stage, a **numeric-to-nominal filter** is applied. This filter converts numeric values into nominal (categorical) attributes, ensuring compatibility with classification algorithms.

Classification
Classification predicts data instance classes using machine learning. Patterns from training data are applied to fresh records. This study uses the Naïve Bayes (NB) classifier. Percentage of accurately identified instances measures attribute quality. Validation uses 10-fold cross-validation. Naïve Bayes classifier identifies Covid-19 patients and predicts the likelihood of each characteristic for the expected class.

Hybrid Approach
To improve accuracy, a **hybrid classification approach** is proposed, combining **Support Vector Machine (SVM)** with Naïve Bayes. The SVM organizes data according to a selected set of discriminative features and uses this model as a reference table during prediction. Each entry in the table is associated with class probabilities. Meanwhile, Naïve Bayes, based on Bayes' theorem, assumes conditional independence between features and computes probability distributions accordingly.
The hybrid model works as follows:

- Attributes are divided into two groups. Some are modelled using Naïve Bayes, while others are represented in a decision table.
- A forward selection search is applied to determine which attributes should be handled by Naïve Bayes and which by the decision table.
- The outcomes from both models are combined to generate overall class probability estimates.

This approach leverages the strengths of both algorithms, producing a more robust and reliable Covid-19 infection prediction system.

III. RESULTS AND DISCUSSIONS

In this study, different algorithms such as Naïve Bayes, SVM, and a hybrid model were evaluated to compare their performance in predicting Covid-19 infections. The experiments were conducted using the WEKA tool, with attributes and results distributed as shown in the evaluation figures. The performance of the models was assessed using standard metrics: accuracy, precision, and recall.

- True Positive (TP): Number of records correctly identified as Covid-19 positive. For example, patients having the infection accurately classified as Covid-19 cases.

- False Positive (FP): Number of records incorrectly classified as positive when they were negative.

- True Negative (TN): Number of records correctly identified as negative. For instance, healthy individuals accurately classified as healthy.

- False Negative (FN): Number of records incorrectly classified as negative when, they were positive. Patients with Covid-19 mistakenly identified as healthy.

Figure 2 shows the visualization of Covid-19 patient classification.

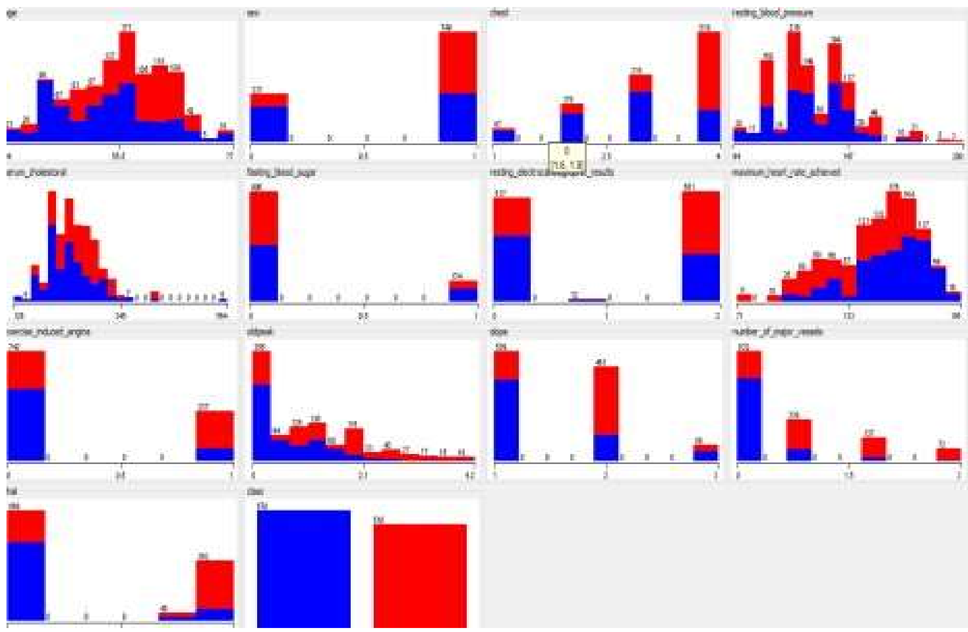

Figure 2: Covid-19 patients Visualization

From Figures 3 to 5, it is evident that the hybrid approach (SVM + Naïve Bayes) outperforms other algorithms across all performance metrics:

- Accuracy Analysis (Figure 3): The hybrid model demonstrates higher accuracy compared to Naïve Bayes, Random Forest, ANN, RNN, and standalone SVM.
- Precision Analysis (Figure 4): The hybrid model exhibits superior precision, indicating its effectiveness in reducing false positives.
- Recall Analysis (Figure 5): The hybrid approach also achieves higher recall values, meaning it successfully identifies a larger proportion of true Covid-19 cases.

Overall, the hybrid method consistently delivers better results than individual models, confirming its suitability for reliable Covid-19 infection prediction.

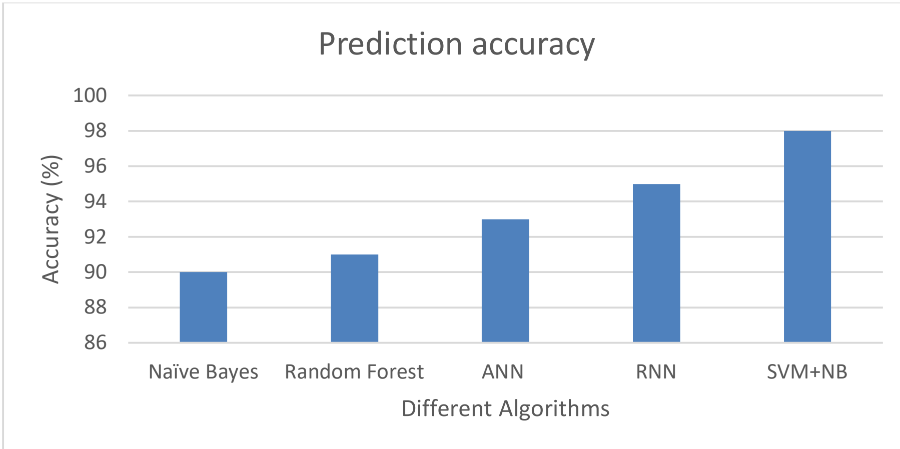

Figure 3: Accuracy analysis of proposed framework in comparison with different algorithms

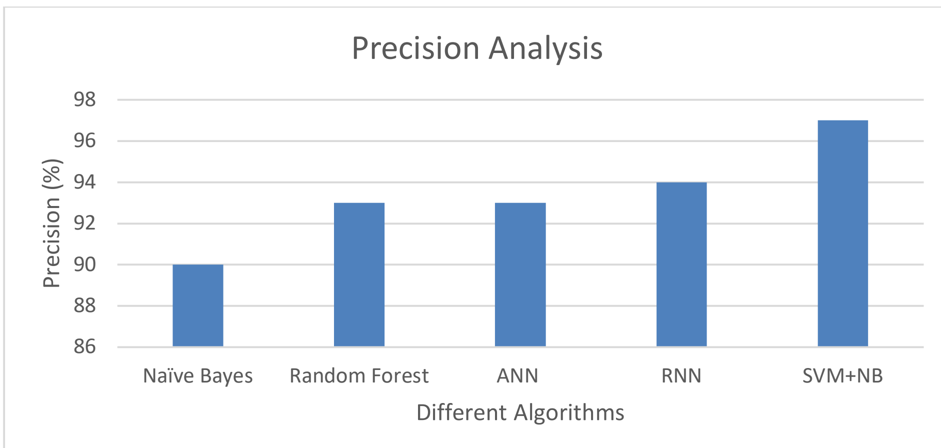

Figure 4: Precision analysis of proposed framework in comparison with other existing algorithms

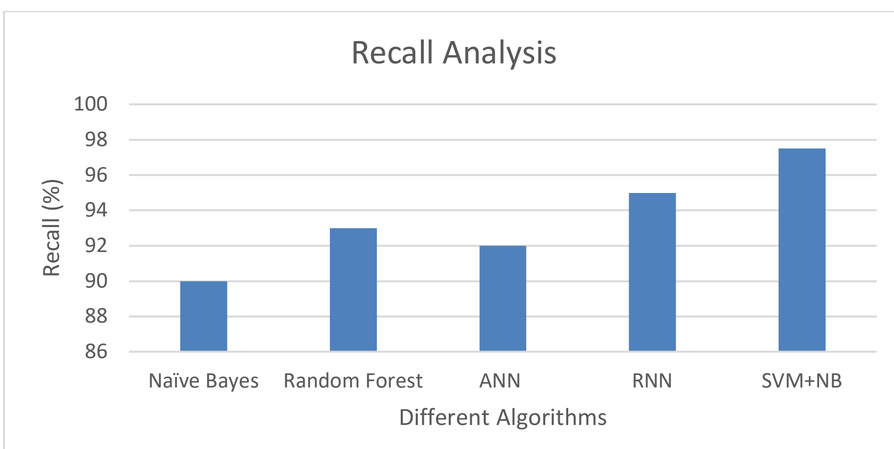

Figure 5: Recall analysis of proposed framework in comparison with other existing algorithms

IV. CONCLUSION AND FUTURE SCOPE

The fast expansion of healthcare data poses considerable obstacles in its administration and analysis. Data mining techniques may be utilised to reveal concealed trends and facilitate medical decision-making. This study presents a Covid-19 infection prediction paradigm that equips clinicians with a valuable tool for diagnosis and treatment planning. The framework was assessed utilising several classification methods, including Naïve Bayes, SVM, and a hybrid model. The examination of patient data revealed that the hybrid method attained superior prediction accuracy, precision, and recall relative to standalone algorithms. This validates the efficacy of integrating Naïve Bayes and SVM for Covid-19 forecasting. The suggested method enhances the reliability of Covid-19 diagnosis and improves patient outcomes by aiding healthcare providers in making prompt, data-driven decisions.

Future Scope: While the current study focused on patient datasets, future work can extend this framework by incorporating larger, real-time clinical datasets and applying advanced deep learning models for further performance improvement. Integration with healthcare information systems and IoT-enabled medical devices could also provide continuous monitoring and early detection, making the system more scalable and practical for real-world healthcare applications.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1]. M.A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "COVID-19 infection: origin, transmission, and characteristics of human coronaviruses," Journal of Advanced Research, no. 24, pp. 91–98, 2022.

[2]. L. Wang, Y. Wang, D. Ye, and Q. Liu, "Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence," International Journal of Antimicrobial Agents, vol. 55, no. 6, pp. 1-9, 2019.

[3]. S.K. Kumaravel, R.K. Subramani, T.K. JayarajSivakumar, R. Madurai Elavarasan, A. Manavalanagar Vetrichelvan, A. Annam, U. Subramaniam, "Investigation on the impacts of COVID-19 quarantine on society and environment: preventive measures and supportive technologies," 3 Biotech 2020; vol. 10, no. 9, 2020.

[4]. J. Zhang Jiancheng, Xie Bing, Hashimoto Kenji, "Current status of potential therapeutic candidates for the COVID-19 crisis," Covid-19BehavImmun, vol. 87, pp. 59–73, 2020.

[5]. H. Li, S.M. Liu, X.H. Yu, S.L. Tang, C.K. Tang, "Coronavirus disease 2019 (COVID-19): current status and future perspectives," International Journal Antimicrobial Agents, vol. 55, no. 5, pp. 1-9, 2020.

[6]. A. Wilder-Smith, C.J. Chiew, and V.J. Lee, "Can we contain the COVID- 19 outbreak with the same measures as for SARS?," Lancet Infect Dis 2020;20(5): e102–7.

[7]. C.C. Lai, T.P. Shih, W.C. Ko, H.J. Tang, P.R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges," International Journal of Antimicrobial Agents, vol. 55, no.3, pp. 1-9, 2020.

[8]. M.E. El Zowalaty, J.D. J¨arhult "From SARS to COVID-19: a previously unknown SARS- related coronavirus (SARS-CoV-2) of pandemic potential infecting humans – Call for a One Health approach," One Health, vol. 9, pp. 1-9. 2020.

[9]. M. Jamshidi, A. Lalbakhsh, J. Talla, Z. Peroutka, F. Hadjilooei, P. Lalbakhsh, and W. Mohyuddin, "Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment," IEEE Access, vol. 8, pp. 109581-109595, 2020.

[10]. A.J. Kucharski, T.W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R.M. Eggo, F. Sun, M. Jit, J.D. Munday, N. Davies, A. Gimma, K. van Zandvoort, H. Gibbs, J. Hellewell, C.I. Jarvis, S. Clifford, B.J. Quilty, N.I. Bosse, S., Abbott, P., Klepac, S. Flasche, "Early dynamics of transmission and control of COVID-19: a mathematical modelling study," Lancet Infect Dis, vol. 20, no. 5, pp. 553–558.

[11]. J. Hellewell , S. Abbott , A. Gimma , N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday , A.J. Kucharski, W.J. Edmunds, S. Funk , R.M. Eggo , F. Sun , S. Flasche , B.J. Quilty, N. Davies , Y. Liu , S. Clifford , P. Klepac , M. Jit , C. Diamond , H. Gibbs , K. van Zandvoort, "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," Lancet Global Health, vol. 8, no. 4, pp. 488–496, 2020.

[12]. R. Vaishya, M. Javaid, I.H. Khan, A. Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," Diabetes Metab Syndrome, vol. 14, no. 4, pp. 337–339, 2020.

[13]. W. Naud ́e, "Artificial Intelligence against COVID-19: An Early Review," IZA Institute of Labor Economics, IZA DP No. 13110, Apr. 2020. [Online]. Available: https://www.iza.org/publications/dp/13110/artificial-intelligence-against-covid-19-an-early-review; 2020.

[14]. R.M. Carrillo-Larco, and M. Castillo-Cara, "Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: an unsupervised machine learning approach," Welcome Open Research, vol. 5, no.56, 2020.

[15]. R. Pung , C.J. Chiew , B.E. Young , S. Chin , M.C. Chen , H.E. Clapham , A.R. Cook , S. Maurer- Stroh, Toh MPHS, C. Poh , M. Low , J. Lum , V.T.J. Koh, T.M. Mak , L. Cui , R.V.T.P. Lin , D. Heng , Y.S. Leo, D.C. Lye, V.J.M. Lee , K.Q. Kam, S. Kalimuddin, S.Y. Tan, J. Loh , K.C. Thoon , S. Vasoo , W.X. Khong, N.A. Suhaimi , S.J.H. Chan, E. Zhang, O. Oh , A. Ty , C. Tow , Y.X. Chua , W.L. Chaw , Y. Ng , F. Abdul-Rahman , S. Sahib , Z. Zhao, C. Tang , C. Low , E.H. Goh , G. Lim, Y. Hou , I. Roshan , J. Tan , K. Foo , K. Nandar , L. Kurupatham, P.P.Chan, P. Raj , Y. Lin , Z. Said , A. Lee , C. See, J. Markose, J. Tan, G. Chan, W. See, X. Peh, V. CCai,W.K. Chen, Z. Li, R. SooSoo,L.P. Chow, W. WeWei,. Farwin, L.W. AnAng, "vestigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures," Lancet, 2020; vol. 395, no. pp. 1039–1046, 2020.

[16]. R. Madurai Elavarasan, R. Pugazhendhi, "Restructured society and environment: a review on potential technological strategies to control the COVID-19 pandemic," Sci Total Environ 2020;725:138858.

[17]. W. Fink, V. Lipatov, and M. Konitzer, "Diagnoses by general practitioners: accuracy and reliability," Int J Forecast, vol. 25, no. 4, pp. 1-9, 2009.

[18]. Zhu Xianglei, Fu Bofeng, Yang Yaodong, Ma Yu, HaoJianye, Chen Siqi, Liu Shuang, Li Tiegang, Liu Sen, GuoWeiming, Liao Zhenyu. "Attention-based recurrent neural network for influenza epidemic prediction," BMC Bioinf 2019; vol. 20, no. (S18), pp. 1-9.

[19]. H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent Neural Networks for Time Series Forecasting: Current status and future directions," International Journal of Forecasting, vol. 37, no. 1, pp. 388–427, 2020.

[20]. S. Murugan, A. Bhardwaj, and T.R. Ganeshbabu, "Object recognition based on empirical wavelet transform," International Journal of MC Square Scientific Research, vol. 7, no. 1, pp. 74-80, 2015