

Disease Classification and Recommendation System using CNN and Geo-demographic data

Mohanasundari DK¹, Ponnarasan V², Mukesh M³ and Saravanan R⁴

¹⁻⁴Department of Computational Intelligence, School of Computer Science and Engineering, Vellore Institute of Technology,
Vellore, Tamil Nadu, India

Email: mohanasundarikarthikeyan.dk@gmail.com, veluponnarasan@gmail.com, mukesh.m2024@vitstudent.ac.in,
rsaravanan@vit.ac.in

Abstract— This research proposes a CNN-enabled framework, that combines demographic and geographic information in a system to improve disease classification and ultimately provide personalized health recommendations. To classify diseases such as dengue, malaria, typhoid, and cholera, the model utilizes age, gender, the location represented in latitude and longitude, and seasonal aspects of disease. Further, the platform categorizes disease, as well as represents a disease prediction of risk in different geography based on time by capturing the temporal and spatial patterns of the disease. Part of the framework design is a recommendation engine that provides preventive and therapeutic recommendations based on the user's health, and community context. The platform is deployed as a web application and was designed to be scalable, quick to use, and friendly for public use, and intended for healthcare practitioners, and health policy-makers, as well as members of the public. Overall, this framework can also support proactive public health and planning by providing early warnings, planning and allocating limited resource effectively, and continued surveillance of disease, which would support preparedness planning for future events.

Index Terms— Convolutional neural (CNN), geodemographic information, infectious disease classification, personalized health recommendations.

I. INTRODUCTION

A crucial use of artificial intelligence in healthcare is illness prediction, which transforms disease detection and diagnosis by utilizing machine learning and deep learning algorithms to evaluate big datasets. Large amounts of medical data may be processed effectively by these AI-driven systems, which can also reveal complex patterns that may be hard for human specialists to see. AI improves diagnostic speed and accuracy by reducing reliance on manual evaluation, which maximizes the effectiveness of the healthcare system and raises affected patient care.

This study's main aim is to create a trustworthy AI model for early disease identification, which is essential to improving health outcomes. Early diagnosis enables medical professionals to take prompt action, which may stop the progression of the disease and lower treatment expenses. The system facilitates informed decision-making by giving healthcare providers a reliable and effective tool, which eventually improves patient care. This method improves outcomes and increases the efficiency of healthcare delivery by streamlining diagnostic workflows and enabling medical professionals to make informed clinical decisions.

II. LITERATURE REVIEW

Recent advances in deep learning, such as Convolutional Neural Networks (CNNs), will bring a paradigm shift to disease classification within the health care sector. CNNs can automatically split informative features from structured and unstructured data, and then they perform well in information classification for the purpose of medical diagnosis. Oraño et al. [1] proposed a CNN model to identify lung diseases based on chest X-ray images, the model classified 6 otherwise distinct conditions with 82.53% accuracy.

Similarly, Mohammed et al. [2] presented a hybrid diagnostic model that combines CNN with Random Forest and XGBoost, as a means of classifying seven different diseases, including Alzheimer's, COVID-19, and several cardiac diseases.

Esam and Mohamed [3] went on to develop a CNN-based model for detecting Alzheimer's disease using MRI images and achieved a 97% accuracy. Their model was more successful than some other architectures, including pretrained methods such as VGG, DenseNet, ResNet, and Inception-type networks and was an example of deep learning importance on early AD diagnosis. [4] presented a similar approach by applying model-based filtering methods such as Singular Value Decomposition (SVD) and Non-Negative Matrix Factorization (NMF) for systems of drug recommendation within the healthcare context that seek to provide a more tailored approach to treatment.

Ajith et al. [5] built predictive models for epidemic diseases, going through models within the vector-borne category, such as the West Nile virus. Following this approach, Dubey et al. [6] described a novel application of Reinforcement Learning (RL) for predicting the spread of infectious diseases in real time, and in particular, their framework may provide future flexible surveillance systems. Guo et al. [7] used a hybrid XGBoost LSTM framework to predict the spread of human infectious diseases, using both traffic data and population flow data to enhance predictive accuracy, where XGBoost successfully captured data on the disease transmission process. B. Kalaivani et al. [8] compared CNN architectures (VGG16, InceptionV3) for multiclass classifications of chest X-rays that achieved accuracy of up to 93% accuracy. Brown et al. [9] used urgent computing with epidemiological models to study mosquito borne diseases and demonstrated the advantages of high performance computing (HPC) with epidemic situations.

Some of the other disease-specific work includes CNN XGBoost models for the prediction of Parkinson's disease [10], hyper-parameter tuned Random Forest regression model for dengue prediction [11], and fuzzy-model-based Random Forest model for analyzing contagious disease [12]. Furthermore, Nivethitha Devi et al. [13] provided a complete multi-disease classification and recommendation system able to diagnose seven diseases and provide preventive healthcare solutions. There is substantial literature regarding machine learning methods for forecasting outbreaks of epidemics. For instance, Raizada et al. [14] showed how CNNs could be used to predict vector-borne diseases, Saxena et al. [15] described probabilistic models for airborne infectious diseases, and Zhang et al. [16] researched an AI-assisted real-time epidemic monitoring and predicting system. Shinde et al. [17] constructed ML-based epidemic outbreak prediction models by incorporating socio environments features like climate and population density.

There have also been emerging research using transformer based methods. For example, Adeika et al. [18] utilized BERT, XLNet, and RoBERTa to classify genomic sequences for the purpose of predicting infectious diseases. Gautam and Raza [19] presented an Autoencoder and Multi-Head Attention with GRU (a type of RNN used for sequential data) using news data to predict outbreaks, achieving almost 98% predictive accuracy. Singhal et al. [20] emphasized the influence of AI and machine intelligence in clinical-predictive disease outbreak predictions along with problems such as data quality, bias and ethical issues.

III. METHODOLOGY

The system utilizes Convolutional Neural Networks (CNN) and geo-demographics to create a valid disease classification and recommendation model which is relevant for the context. There are several modules within the system that have focused specific use to improve diagnosis and personalized health recommendations.

Data Collection: The data gathered for this research was taken from Kaggle (a database with numerous datasets), which contains a publicly available dataset called "Disease and Symptoms Dataset". The dataset contained structured, tabular data with many important health characteristics. Each observation (row) represents a visit of a patient and had the age and gender, the symptoms (has seventeen symptom columns, 'Symptom_1' to 'Symptom_17'), geo-location (distance from the location), and disease diagnosis. By adding longitude and latitude to each case, it allowed to allocate each case and focus geo-related characteristics on each case. The primary diseases within the dataset represent a catalogue of a range of common infections or prevalent chronic

illnesses including malaria, typhoid, dengue, cholera, etc. These characteristics provide sufficient research context and clinical information to train a disease model which is multi class classification.

Data Preprocessing: Information we collected will usually, at some point, need to be pre-processed because the data is raw and will typically have inconsistencies and incompleteness. The initial step we made was to convert placeholder values such as the string "0" into standard null indicators so we could treat the data as missing data through normal data handling protocols. In addition, to protect ourselves from data typing issues, we converted all fields with symptoms into a string type. We then handled missing symptoms as either a placeholder or deleting records with missing data. As for all disease labels, we used label encoding to convert disease labels to a special integer. Categorical features such as, gender we encoded with numbers. Continuous features such as, age and geographic coordinates, were normalized to uniform ranges of 0-1 values with minmax scaling. This way, we can be assured that no specific attribute has a discriminatorily larger weight. After we finished data pre-processing, we were comfortable that the cleaned and consistent data was in numerical format being processed into the model for training.

Feature Extraction: Feature extraction was utilized to create structured input vectors from the cleaned datasets. Each of the symptom columns in the original dataset was traversed and a common list of all the distinct symptoms was built. Binary vector was constructed for each patient record that classified if the distinct symptom occurred (1) or did not occur (0). This resulted in a high dimensional sparse matrix, where each row was a patient and each column was a distinct symptom. The binary coding allowed the model to learn some complicated patterns in a cluster of symptoms that cut across the different diseases. Along with the symptom vectors were some demographic variables (gender, age) and geographic variables (longitude and latitude). The additional demographic and geographic features allowed the model to learn not only clinic symptoms but also demographic and geographic changes. The additional features improved predictions made by the model.

Model Training: The classification model was built employing a Keras API based deep learning model with TensorFlow as the backend. The input layer of the model takes the binary feature vector of symptom and contextual information. Two hidden layers with fully connected (Dense) neurons with the ReLU activation function were then added to introduce nonlinearity. Dropout layers were subsequently inserted between the hidden layer in order to prevent overfitting by randomly disabling a fraction of the neurons at every step throughout the training process. The output layer had a softmax activation function which enables multi-class classification through the provision of a probability distribution over all available disease classes. The model was trained with a categorical cross-entropy loss function, the best for multi-class output, and an Adam optimizer to learn adaptively. The training was done over a fixed number of epochs, with hyperparameters chosen for maximizing performance in terms of batch size, learning rate, and dropout rate.

Model Validation: To analyse model performance, and prevent overfitting, the dataset was divided into training and test sets in an 80:20 ratio. Every time the training process finished, I tested with the test data. The performance of the model was assessed through accuracy, precision, recall, and F1-score. While these metrics did not directly correlate to how well the model did in classifying diseases among the classes while reducing false negatives and false positives, they provided a broad overview of how well the model classified diseases. To check the model performance, I used confusion matrix, and looked if the model did well based on potential similarities between the diseases. If the validation set does not mirror the training set performance, I adjusted hyperparameters, or model configuration to re-balance and improve generalizability.

Disease Prediction: After training and testing is complete, the model can then be used to estimate diseases based on new input information. The system will take user supplied information of symptoms, age, gender, and users' spatial coordinates and convert this to the same binary input that was used when training the model. The model will then use the user's input as the new data to estimate - it will produce a single disease response that it estimates is most likely to be diagnosed for the user, and expresses a probability score of that response to express the encoded likelihood to the estimate. The vision is to hopefully provide real-time diagnosis predictions to end-users- both clinical providers and consumers- without human demonstration. The model architecture is built to allow scalable development, so the parameters in development of the model will be able to be re-calibrated to this type of web inference once this application is assimilated into clinical practice.

Recommendation System: Once the model has made its prediction, the next step is to activate the recommendation engine. The recommendation engine will give individualized health recommendations based on diagnosis and individual environmental and demographic profile. The recommended actions will be clinical recommendations such as treatment suggestions, preventative suggestions, diet suggestions, and lifestyle suggestions. For example, if the user has made a complaint of dengue during monsoon season in the tropics, the recommendation engine may suggest recommendations of hydration, spraying insect repellent on skin and clothes, and visiting a local clinic. The recommendations will be context specific. Therefore, these can help with

patient understanding and self-management of care, especially in places with little access to human health practitioners.

Prediction of Future Spread (2025-2035): Another model emerged to forecast possible disease spread trends over a 10-year period from 2025-2035. This module provides disease information, previously reported, in time and space context, to address predictions about how these may perform with simulation models based on locational emergence of disease events occurring in clusters. In the forecasting models, the time series analysis and geospatial mapping of sick events will help predict the probability disease will be at a location over time. It is anticipated this would inform public health treatment planning by helping public health practitioners respond in a timely manner to manage allocation of medical resources, alert populations as early as possible, and develop targeted prevention responses to at risk locations. Predictive foresight will also provide opportunities to drive research into public health preparedness.

IV. EXISTING SYSTEM

Traditional machine learning algorithms such as Decision Trees, Support Vector Machines (SVM), and Logistic Regression form the backbone of current disease prediction models. These models use the patient records, including disease symptoms, history, and demographics, in an attempt to recognize patterns and establish probabilities of a disease. They are somewhat effective; however, these models have a number of inherent deficits.

Some of the main deficits are a lack of consistent data quality, poor adaptability to a different population, or an inadequate incorporation of environmental and spatial determinants. Traditional models rely solely on stationary data and human feature selection, which hinders their ability to detect region centric or seasonal diseases, which are influenced entirely by location and locational determinants.

Moreover, these models may not have real-time data processing ability and expandability; both of which are essential for predicting communicable diseases (e.g. dengue, malaria, typhoid, cholera, etc.) accurately and timely. Although there have been efforts on data preprocessing and feature engineering, improvements are marginal.

V. PROPOSED SYSTEM

The project offers a fresh and comprehensive model of disease prediction that is a combination of machine learning and holistic health advice. Compared to other models, in which only illness is forecasted, this model forecasts an output that is multi-dimensional in its approach to prediction, which includes Ayurvedic treatment, customized dietary advice, recommended exercise and the recommended ideal fast food, besides other lifestyle recommendations.

One of the main novelties was the use of geolocation (latitude, longitude) as the input to the forecast to allow location-based hazards to be assessed. The system also demonstrates its novelty by providing a simulation of disease trend over an extended projected timeline (2025-2035) - in the sense that a generally innovative forecasting ability can be applied to public health planning.

Other than that, the app features a personal health monitoring authentication module and will add more features in the future, such as patient profiles. Free-text symptom input significantly enhanced user experience over complex checkboxes, unlimited dropdowns, etc. and offers a point of liaison between medical AI and real-world application. Combined, this system is unprecedented in the way it integrates AI diagnosis and forecasting, integrative health information, user customization, geographic intelligence, and temporal forecasting - which are interesting from both technical as well as social influence perspectives.

The proposed system employs a deep learning-based approach, specifically utilizing Convolutional Neural Networks (CNNs), to analyse symptoms and classify diseases. The model consists of multiple convolutional layers, ReLU activation functions, batch normalization, and a final SoftMax layer for classification. It is trained on a labelled dataset containing various diseases and their corresponding symptoms, enabling it to predict the most probable disease based on new input data.

This system is intended to improve diagnostic accuracy and enable early medical intervention. By utilizing deep learning methodologies, the model automatically enhances its predictive ability as it gets trained on more data. Intended to be integrated into healthcare systems, this solution offers robust disease prediction, supporting doctors in taking timely and well-informed decisions. Additionally, the easy-to-use interface ensures convenience for healthcare professionals with varied technical levels of proficiency.

Unlike previous works mostly engaged on medical imaging or symptom-only data sets, our suggested method innovatively blends geo-demographic features (age, sex, latitude, longitude, seasonal variation) as well as CNN-based classification. It not only lets us accurately identify the disease but also predict future spread and provide tailored-in-the-wind recommendations. Ours thereby completes gaps in contextual understanding, big-ability, as well as predictive health planning that previous models never stressed.

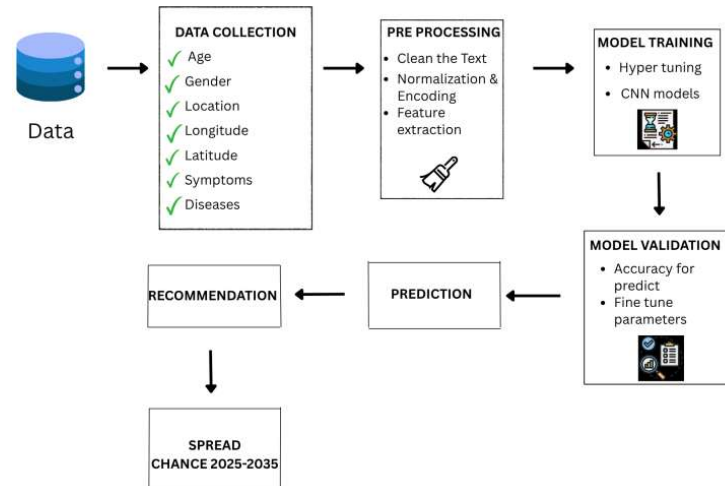


Fig .1 Architecture of proposed system

VII. TRAINING AND EVALUATION

The data is split into training and validation sets, with the training set utilized for model learning and validation set for performance assessment. Cross-validation methods are used to guarantee that the model performs well with unseen information. The hyperparameters like learning rate, batch size, and the number of epochs is tuned to achieve a high level of accuracy. Techniques such as early stopping and learning rate decay are employed to avoid overfitting and achieve effective convergence. Model performance indicators like accuracy and loss are measured to assess the model, and any required adjustments are taken to maximize outcomes.

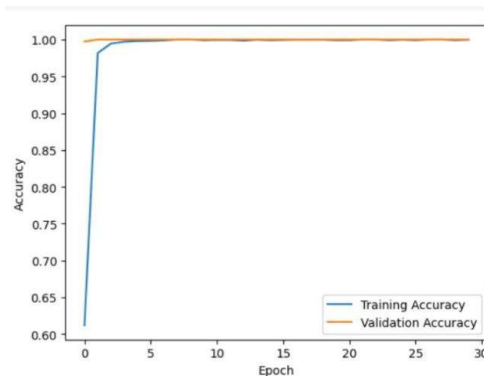


Fig.2.Accuracy graph

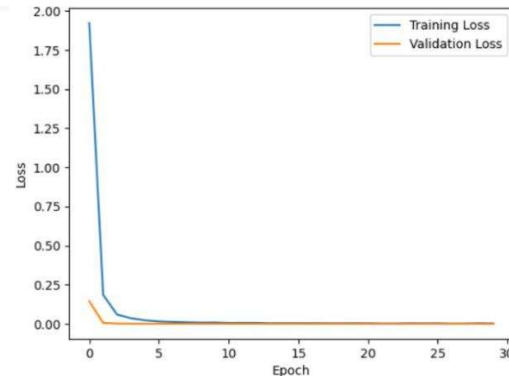


Fig.3.Loss graph

The accuracy graph displays training iterations along the X axis and accuracy along the Y-axis for training and validation. The training accuracy increases rapidly to almost 100%, while the validation accuracy remains stable, high enough to indicate sufficient generalization and no overfitting. In the loss graph, training loss starts relatively high and drops quickly toward zero, and the validation loss drops quickly and remains low too. Both of these trends indicate the model is learning without serious overfitting or avoiding generalization entirely. The training and validation loss was also relatively low, with a small gap, indicating an absence of overfitting. Conversely, the presence of a large gap (in the opposite direction) or a diverging gap in losses would imply overfitting, or worse, that the model is learning nothing at all.

VIII. RESULTS AND DISCUSSION

In order to measure the performance of the model, the data was divided into 80:20 train test sets. Here the suggested CNN incorporating geo-demographic features was tested using baseline classifiers like Support Vector Machines (SVM), Random Forest (RF), and Decision Tree (DT).

- CNN obtained 94.2% overall accuracy, higher than SVM (88.1%), RF (87.5%), and DT (84.6%)
- Precision, Recall, and F1-scores for all the classes (dengue, malaria, typhoid, cholera) showed that CNN was an effective false-negative reducer, important for disease diagnosis.
- In the confusion matrix it was seen that misclassification occurred predominantly between dengue and malaria due to probably merging symptom profile.
- Precision and loss graphs (Fig. 2, Fig. 3) confirmed valid convergence without critical overfitting.

These results confirm the value of the addition of geo-demographic features as significantly increasing predictive power. Where established methods apply only the clinical or symptom-based inputs, our model positions disease risk in relation using region and seasonality so as to raise the practical usefulness for routine application.

XI. CONCLUSION

This project proposes an efficient CNN-based disease classification and recommendation model incorporating geo-demographic data for enhancing diagnostic effectiveness and delivering health recommendations. It surpasses the traditional ML models by providing better generalizability as well as real-time functionality. In the future, it will add environmental indicators to the set data, use natural language for patient engagement, and utilize the system on cloud platforms for scalable use in large populace health monitoring. It integrates predictive analytics as well as recommendation-specific personalization for promoting proactive care as well as facilitating early intervention measures.

FUTURE WORK

Future work for this project may include using real-time data streams, including weather reports, population density trends, and epidemiological reports to further improve precision and increase responsiveness. The dataset can be further expanded with better indicators of environmental variables (e.g., air quality, humidity, water contamination) to optimize disease classification performance. The recommendation engine could be upgraded with AI-generated messages and natural language processing (NLP) in conjunction with reinforcement learning to optimize person-specific contextual health alerts. Finally, deploying the system to a cloud-hosted platform with a mobile friendly interface would allow access and real-time tracking of disease spread for clinicians and the general population.

AUTHORS' CONTRIBUTIONS

Conceptualization of idea for article was done by Ponnarasan V. The literature survey and data analysis were performed by Ponnarasan V, Mukesh M and Mohanasundari DK. Drafting, tabulating and pictures was drafted by Mukesh M. Coding and Implementation was carried out by Mohanasundari DK. Final manuscript editing and supervision over the work was done by Saravanan R.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of Vellore Institute of Technology, Vellore on providing us the Access over journals and work space accommodation.

REFERENCES

- [1] J. F. V. Oraño and E. A. Maravillas, "lung disease classification utilizing convolutional neural network with CXR imaging," in Proc. 3rd Int. Conf. Mobile Networks Wireless Commun. (ICMNWC), Tumkur, India, 2023, pp. 1–5, doi: 10.1109/ICMNWC60182.2023.10435917.
- [2] M. Mohammed, R. Mongia, and M. Anand, "A novel approach to multi-disease detection in healthcare using CNN, Random Forest and XGBoost," in Proc. Int. Conf. Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), Chennai, India, 2023, pp. 1–5, doi: 10.1109/RAEEUCCI57140.2023.10134394.

- [3] S. Esam and A. Mohammed, "Alzheimer's disease classification for MRI images using convolutional neural networks," in Proc. 6th Int. Conf. Computing and Informatics (ICCI), New Cairo, Egypt, 2024, pp. 1–5, doi: 10.1109/ICCI61671.2024.10485049.
- [4] S. N. M., P. Kumar K. R., and S. B. J., "Model-based filtering techniques for recommendation systems in healthcare domain," in Proc. 7th Int. Conf. I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 978–983, doi: 10.1109/I-SMAC58438.2023.10290568.
- [5] A. Ajith, K. Manoj, H. Kiran, P. J. Pillai, and J. J. Nair, "A study on prediction and spreading of epidemic diseases," in Proc. Int. Conf. Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 1265–1268, doi: 10.1109/ICCSP48568.2020.9182147.
- [6] A. K. Dubey, K. Gupta, V. S., A. Sankari, S. S. Namani, and R. Hemalatha, "Predicting disease outbreaks with AI: An in-depth analysis of infectious diseases surveillance," in Proc. Int. Conf. Frontier Technologies and Solutions (ICFTS), Chennai, India, 2025, pp. 1–7, doi: 10.1109/ICFTS62006.2025.11031614.
- [7] K. Guo, C. Shen, X. Zhou, S. Ren, M. Hu, M. Shen, and X. Li, "Traffic data-empowered XGBoost-LSTM framework for infectious disease prediction," IEEE Trans. Intell. Transp. Syst., vol. 25, no. 1, pp. 774–785, Jan. 2024, doi: 10.1109/TITS.2022.3172206.
- [8] P. Kalaivani, A. Dhivya, G. Dharani, S. Bharathi, and C. Rajan, "Deep learning with multi-class classification for detection of COVID-19 and pneumonia," in Proc. Int. Conf. Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2023, pp. 287–292, doi: 10.1109/ICICT57646.2023.10134162.
- [9] N. Brown, P. Smith, J. Doe, and R. Patel, "Utilising urgent computing to tackle the spread of mosquito-borne diseases," in Proc. IEEE/ACM HPC for Urgent Decision Making (UrgentHPC), St. Louis, MO, USA, 2021, pp. 36–44, doi: 10.1109/UrgentHPC54802.2021.00010.
- [10] A. K. N., B. R. Bhatt, P. Anitha, A. K. Yadav, K. K. Devi, and V. C. Joshi, "A new diagnosis using a Parkinson's disease XGBoost and CNN-based classification model using ML techniques," in Proc. Int. Conf. Advanced Computing Technologies and Applications (ICACTA), Coimbatore, India, 2022, pp. 1–6, doi: 10.1109/ICACTA54488.2022.9752867.
- [11] H. M. Veena Kumari, D. S. Suresh, and P. E. Dhananjaya, "Clinical data analysis and multilabel classification for prediction of dengue fever by tuning hyperparameter using GridsearchCV," in Proc. 14th Int. Conf. Computational Intelligence and Communication Networks (CICN), Al-Khobar, Saudi Arabia, 2022, pp. 302–307, doi: 10.1109/CICN56167.2022.10008355.
- [12] S. Geeitha, P. Karthikeyan, S. Aravinth, P. Nachiappan, and S. Jagadeesh, "Contagious disease prediction using Random Forest algorithm interpolated with fuzzy model," in Proc. Int. Conf. Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 674–679, doi: 10.1109/ICSCDS56580.2023.10104962.
- [13] M. N. Devi, V. M. Rufus, and J. Jeffrey, "A comprehensive multi-disease classification and recommendation system: Towards predictive healthcare and preventive solutions," in Proc. Int. Conf. Artificial Intelligence and Machine Learning Applications (AIMLA), Namakkal, India, 2025, pp. 1–7, doi: 10.1109/AIMLA63829.2025.11041533.
- [14] S. Raizada, S. Mala, and A. Shankar, "Vector borne disease outbreak prediction by machine learning," in Proc. Int. Conf. Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 2020, pp. 213–218, doi: 10.1109/ICSTCEE49637.2020.9277286.
- [15] S. Saxena, S. Khare, S. Pal, and V. Agarwal, "Analyzing the spread of infectious disease using a probabilistic model," in Proc. IEEE Int. Conf. Communications Workshops (ICC Workshops), Montreal, QC, Canada, 2021, pp. 1–6, doi: 10.1109/ICCWorkshops50388.2021.9473564.
- [16] Y. Zhang, W. Shi, C. Han, and Z. Luo, "The research of artificial intelligence assisted real-time monitoring and response system for infectious disease outbreaks," in Proc. Int. Conf. Integrated Intelligence and Communication Systems (ICIICS), Kalaburagi, India, 2024, pp. 1–5, doi: 10.1109/ICIICS63763.2024.10860203.
- [17] S. Shinde, S. Yadav, and A. Somvanshi, "Epidemic outbreak prediction using machine learning model," in Proc. 5th Int. Conf. Advances in Science and Technology (ICAST), Mumbai, India, 2022, pp. 127–132, doi: 10.1109/ICAST55766.2022.10039594.
- [18] B. I. Adeika, J. Aina, T. Ibirinde, T. Adeyemi, M. M. Rahman, and S. Pramanik, "Ensemble and transformer models for infectious disease prediction," in Proc. IEEE 23rd Int. Conf. Bioinformatics and Bioengineering (BIBE), Dayton, OH, USA, 2023, pp. 377–384, doi: 10.1109/BIBE60311.2023.00068.
- [19] A. S. Gautam and Z. Raza, "Autoencoder and multi-head attention with GRU based approach to predict disease outbreak using news-crawl 2019 data," in Proc. Int. Conf. Computational Intelligence and Network Systems (CINS), Dubai, UAE, 2024, pp. 1–7, doi: 10.1109/CINS63881.2024.10864399.
- [20] M. Singhal, S. Mishra, V. Sharma, K. Anand, and A. Alkhayyat, "Impact of machine intelligence on clinical disease outbreak prediction," in Proc. 10th IEEE Uttar Pradesh Section Int. Conf. Electrical, Electronics and Computer Engineering (UPCON), Gautam Buddha Nagar, India, 2023, pp. 1020–1025, doi: 10.1109/UPCON59197.2023.10434492.