International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(12s) 2026 pp. 700-712
https://doi.org/10.70917/ijcisim-2026-3942
Article

Stochastic Mathematical Models and Machine Learning Algorithms for Early Detection of Cancer and Diabetes

Suren Kumar Sahu[1], G. K. Abani Kumar Dash[2], V. G. Michael Florance[3], R. Chenthil ThangaBama[4], D. Indhumathy[5], Suvitha Subramaniam[6], Sanjay Oli[7], Keshav Kumar K.[8]*

[1] Department of CSE, Gita Autonomous College, Bhubaneswar, Bhubaneswar, Odisha, India, 752024.
Email: suren.sahu2602@gmail.com
[2] Department of CSE, DRIEMS University, Cuttack, Odisha, India, 754022.
Email: abanidash1982@gmail.com
[3] Department of Mathematics, DMI Engineering College, Aralvaimozhi, Tamilnadu, India, 629301.
Email: miclaelflorance@gmail.com
[4] Department of Mathematics, Kalaignar Karunanidhi Institute of Technology, Coimbatore, Coimbatore, Tamilnadu, India, 641402.
Email: chenthilthangabama@gmail.com
[5] Department of Mathematics, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India, 641022.
Email: indhu0313@gmail.com
[6] Department of English, Easwari Engineering College, Chennai, Tamil Nadu, India, 600089.
Email: suvithas5134@gmail.com
[7] Department of Mathematics, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India, 560111.
Email: olisanjay@gmail.com
[8] Department of Humanities and Mathematics, Narayanamma Institute of Technology and Science (for Women), Hyderabad, Telangana, India, 500104.
Email: Keshav.gnits@gmail.com
Corresponding Author: G. K. Abani Kumar Dash

Abstract: Early detection of chronic diseases such as cancer and diabetes is essential for improving patient survival, reducing treatment costs, and supporting timely clinical decision-making. However, disease progression is inherently uncertain due to biological variability, genetic factors, environmental influences, and lifestyle changes, making accurate prediction a challenging task. This study proposes a generalized hybrid framework that integrates stochastic mathematical models with machine learning algorithms for the early detection of cancer and diabetes. Initially, stochastic differential equations are employed to model the uncertain evolution of disease-related variables, such as blood glucose levels for diabetes and tumor growth for cancer. Synthetic patient datasets are then generated by combining demographic, clinical, and stochastic features. These datasets are used to train an Artificial Neural Network (ANN) classifier capable of distinguishing between healthy and diseased individuals. The proposed framework is evaluated using standard performance metrics, including accuracy, precision, recall, F1-score, and ROC-AUC. The integration of stochastic disease progression with machine learning enables the prediction model to capture both temporal uncertainty and complex nonlinear relationships among clinical variables. Furthermore, the framework is sufficiently general to be extended to multiple chronic diseases by modifying only the disease-specific variables while preserving the same computational pipeline. The proposed methodology offers a mathematically interpretable and computationally efficient approach for intelligent healthcare systems, clinical decision support, and personalized medicine.

Keywords: Stochastic Mathematical Model, Machine Learning, Artificial Neural Network (ANN), Early Disease Detection, Cancer Prediction, Diabetes Prediction.

1. Introduction

Cancer and diabetes are among the leading causes of mortality and long-term disability worldwide, placing a significant burden on healthcare systems and economies. According to global health reports, the prevalence of diabetes continues to rise due to aging populations, sedentary lifestyles, obesity, and unhealthy dietary habits, while cancer remains one of the most complex diseases because of its heterogeneous nature and unpredictable progression. Early detection of these diseases is essential for improving patient survival, reducing treatment costs, and enabling timely clinical interventions.

Conventional diagnostic approaches for diabetes rely on laboratory measurements such as fasting blood glucose, oral glucose tolerance tests, and glycated hemoglobin (HbA1c), whereas cancer diagnosis often depends on medical imaging, biopsy, histopathological examination, and biomarker analysis. Although these methods provide valuable clinical information, they generally detect the disease after measurable physiological changes have already occurred. Furthermore, disease progression is influenced by numerous uncertain factors, including genetic mutations, environmental exposure, lifestyle, immune response, and individual biological variability. These uncertainties limit the predictive capability of purely deterministic mathematical models.

Stochastic mathematical models provide a natural framework for representing randomness and uncertainty in biological systems. By incorporating probabilistic effects through stochastic differential equations, these models can describe random fluctuations in blood glucose levels, tumor growth, biomarker expression, and other physiological variables. Such models are capable of producing probabilistic forecasts that better reflect real clinical scenarios than deterministic approaches.

In parallel, machine learning has emerged as a powerful tool for disease prediction and medical decision support. Algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests (RF), Extreme Gradient Boosting (XGBoost), and deep learning models have demonstrated remarkable performance in classifying patients based on clinical, laboratory, and imaging data. Machine learning techniques automatically identify complex nonlinear relationships among risk factors and often achieve higher prediction accuracy than conventional statistical methods. Nevertheless, most machine learning models operate as data-driven black-box systems without explicitly incorporating the underlying disease dynamics or biological uncertainty.

A promising research direction is the integration of stochastic mathematical modeling with machine learning algorithms. In such a hybrid framework, stochastic models simulate the uncertain progression of disease, while machine learning algorithms exploit both clinical variables and stochastic model outputs to improve prediction accuracy. This combination provides interpretable mathematical representations together with the strong predictive capability of artificial intelligence, making it particularly suitable for personalized medicine and early disease detection.

The application of artificial intelligence (AI) and machine learning (ML) in healthcare has grown rapidly over the last decade, particularly for the early detection of chronic diseases such as diabetes and cancer. Recent studies demonstrate that ML algorithms can effectively analyze large volumes of clinical, imaging, and genomic data, improving diagnostic accuracy and enabling personalized healthcare. However, many studies also note that most predictive models remain purely data-driven and do not explicitly incorporate uncertainty or disease dynamics[1].

Silva et al. [2] (2020) conducted one of the earliest systematic reviews and meta-analyses of machine learning models for type 2 diabetes prediction in community settings. The authors evaluated algorithms such as logistic regression, support vector machines, random forests, and artificial neural networks. Their findings indicated that machine learning methods generally outperform conventional statistical approaches for diabetes risk prediction. However, they also emphasized the need for external validation, standardized datasets, and clinically interpretable models before widespread clinical adoption. Bharati et al. [3] (2020) reviewed artificial neural network models for breast cancer diagnosis and compared multilayer perceptrons, convolutional neural networks, stacked autoencoders, and deep belief networks. Their review concluded that deep convolutional neural networks, particularly ResNet-based architectures, achieved superior diagnostic performance on mammography datasets. Nevertheless, the authors highlighted challenges related to limited annotated datasets, computational complexity, and the interpretability of deep learning models. Lee et al. [4] (2020) proposed machine learning models for the diagnosis of early-stage diabetes using temporal glucose profiles generated from a physiological model. Their work demonstrated that neural networks, convolutional neural networks, and recurrent neural networks could accurately detect insulin resistance from time-

series glucose data. This study illustrated the value of combining mathematical disease models with machine learning, although stochastic uncertainty was not incorporated into the framework.

A comprehensive review published in Expert Systems with Applications [5] (2024) compared traditional machine learning and deep learning approaches for cancer detection. The review reported that deep learning methods achieved state-of-the-art performance for medical imaging applications, while ensemble machine learning methods performed well on structured clinical datasets. The authors concluded that future research should focus on multimodal data integration, explainable AI, and robust clinical validation to improve early cancer diagnosis. A recent state-of-the-art review of AI applications in healthcare (2025) [1] summarized advances in diabetes prediction, cancer diagnosis, epidemiological modeling, and mortality prediction. The review emphasized that integrating machine learning with clinical decision-support systems has improved predictive performance across multiple diseases. However, it also identified a lack of mathematically interpretable frameworks capable of representing biological uncertainty and temporal disease evolution. Another recent review [6] analyzed 33 years of research on machine learning and artificial intelligence for type 2 diabetes prediction. The authors observed that ensemble learning, neural networks, and deep learning techniques consistently achieved high predictive performance. They also identified persistent challenges, including data imbalance, overfitting, limited explainability, and insufficient external validation across different populations. Recent work [7] on AI-based diabetes diagnosis highlighted the rapid growth of machine learning and deep learning algorithms for diabetes screening. The review discussed advances in feature engineering, preprocessing techniques, and neural network architectures while emphasizing that future research should prioritize interpretable models, real-time clinical deployment, and integration with electronic health records.

A recent systematic review [8] on machine learning for type 1 diabetes prediction evaluated studies published between 2021 and 2025. The authors found that machine learning models successfully identified high-risk individuals using genetic, immunological, and metabolic markers. However, substantial heterogeneity in datasets, outcome definitions, and validation procedures prevented direct comparison among studies, indicating the need for standardized predictive frameworks. Research on AI and machine learning for prediabetes has also expanded significantly. A recent scoping review [9] reported that machine learning methods have been successfully applied to predict disease onset, progression, and complications while enabling personalized lifestyle interventions. Despite these advances, the review recommended combining predictive algorithms with physiological and mechanistic disease models to improve interpretability and long-term prediction. Overall, the recent literature demonstrates remarkable progress in applying machine learning to early disease detection. However, most existing studies employ deterministic or purely data-driven approaches. Very few investigations integrate stochastic mathematical models with machine learning algorithms to explicitly capture biological uncertainty and disease evolution. Moreover, most published frameworks are disease-specific and cannot be easily generalized to multiple chronic diseases. These limitations motivate the development of a generalized stochastic–machine learning framework capable of combining probabilistic disease modeling with intelligent prediction algorithms for the early detection of both cancer and diabetes.

Although significant progress has been made independently in stochastic disease modeling and machine learning, their combined application for generalized early detection of multiple chronic diseases remains relatively limited. Most existing studies focus on a single disease, employ deterministic mathematical models, or use machine learning directly on clinical datasets without considering stochastic disease evolution. Moreover, few studies propose a unified mathematical framework that is applicable to both cancer and diabetes while incorporating uncertainty into the prediction process.

Therefore, this study proposes a generalized hybrid framework that combines stochastic mathematical models with machine learning algorithms for the early detection of cancer and diabetes. The proposed methodology first models disease progression using stochastic differential equations to account for biological variability and uncertainty. The simulated disease dynamics are then integrated with patient clinical features to train a machine learning classifier capable of predicting disease risk at an early stage. The framework is sufficiently general to be adapted to different chronic diseases by modifying only the disease-specific variables and governing equations while maintaining the same computational pipeline.

The proposed framework has potential applications in intelligent healthcare systems, clinical decision support, personalized medicine, and preventive healthcare. It may assist clinicians in identifying high-risk individuals before severe symptoms appear, thereby improving treatment outcomes and reducing healthcare costs.

The unpredictable nature of disease progression makes early diagnosis challenging. Conventional deterministic models cannot adequately represent biological uncertainty, while purely data-driven machine learning methods often ignore the underlying disease dynamics. Therefore, there is a need to develop an integrated stochastic–machine

learning framework that combines probabilistic disease modeling with intelligent prediction algorithms to improve early detection of chronic diseases such as cancer and diabetes.

The primary aim of this research is to develop a generalized hybrid framework that integrates stochastic mathematical models and machine learning algorithms for the early prediction of cancer and diabetes.

The specific objectives of this study are:

1. To formulate stochastic mathematical models describing the progression of cancer and diabetes under biological uncertainty.
2. To generate disease progression trajectories using stochastic simulations.
3. To develop synthetic patient datasets containing demographic, clinical, and stochastic features.
4. To design and train machine learning algorithms for disease classification.
5. To integrate stochastic model outputs with clinical features to improve prediction performance.

The novelty of this work lies in:

- Developing a generalized stochastic–machine learning framework applicable to multiple diseases.
- Integrating stochastic disease progression models with machine learning classifiers for improved early detection.
- Using stochastic simulation outputs as informative predictive features.
- Providing a mathematically interpretable and computationally efficient framework that bridges applied mathematics and artificial intelligence.
- Establishing a foundation that can be extended to fractional stochastic models, physics-informed neural networks (PINNs), or fractional ANN frameworks, making it highly relevant to future research in computational medicine.

2. Preliminaries

2.1 Disease Progression Modeling

Disease progression is a dynamic process in which a patient's health state changes over time due to biological, genetic, environmental, and lifestyle factors. Diseases such as diabetes and cancer exhibit nonlinear behavior, making mathematical models useful for understanding their evolution and predicting future outcomes.

Let [X(t)] represent the disease state at time (t).

Examples include:

Diabetes: (X(t)) = Blood glucose concentration.

Cancer: (X(t)) = Tumor size or number of abnormal cells.

2.2 Random Variables

A random variable is a numerical quantity whose value depends on chance.

Let $[X: \Omega \to \{R\},]$

where

(Ω) is the sample space.

(X) is the random variable.

Examples:

Blood glucose level

Tumor size

Blood pressure

Patient age

2.3 Probability Distribution

Disease measurements are random due to biological variability.

For example,

$$[X \sim N(\mu, \sigma^2),]$$

where

(μ) = mean

(σ) = standard deviation

This distribution describes variations in glucose levels or tumor sizes across a population.

2.4 Stochastic Process

A stochastic process is a collection of random variables indexed by time.

$[\{X(t), t \geq 0\}.]$ It describes disease progression under uncertainty.

Examples:

Random increase in blood glucose.

Random tumor growth.

Variation in biomarker levels.

2.5 Brownian Motion

Brownian motion is commonly used to represent random fluctuations in biological systems.

It satisfies

[W(0)=0,] and $[W(t) - W(s) \sim N(0, t-s).]$

Properties:

Independent increments.

Continuous trajectories.

Gaussian distribution.

2.6 Stochastic Differential Equation (SDE)

Disease progression can be modeled using

[dX(t)=f(X,t) dt+g(X,t)dW(t),]

where

(f(X,t)) is the deterministic component.

(g(X,t)) represents random fluctuations.

(W(t)) is Brownian motion.

Examples:

Diabetes

$$[dG(t) = rG(t)dt + \sigma\, G(t)\, dW(t).]$$

Cancer

$$[dT(t) = rT(t)(1 - \{T(t)\}/\{K\})dt + \sigma\, T(t)dW(t)]$$

2.7 Feature Vector

Each patient is represented by

$$[X_i = [x_1, x_2, \dots, x_n].]$$

For diabetes,

$$[X_i = [\{Age\}, \{BMI\}, \{Glucose\}, \{Blood\ Pressure\}, \{Exercise\}, \{Family\ History\}].]$$

For cancer,

$$[X_i = [\{Age\}, \{Tumor\ Size\}, \{Gene\ Mutation\}, \{Smoking\}, \{Cell\ Density\}].]$$

2.8 Classification

The disease label is

$$[Y = \begin{cases} 0 & \text{Healthy} \\ 1 & \text{Disease} \end{cases}$$

This represents a binary classification problem.

2.9 Artificial Neural Network (ANN)

An ANN consists of interconnected neurons organized into input, hidden, and output layers.

For neuron (j),

$$[z_j = \sum_{\{i=1\}}^{\{n\}} w_{\{ij\}} x_i + b_j.]$$

The activation is

$$[a_j = \phi(z_j),]$$

where

$(w_{\{ij\}})$ = weight,

(b_j) = bias,

(ϕ) = activation function (ReLU, Sigmoid, Tanh).

2.10 Sigmoid Activation Function

For binary classification,

$$[\sigma(z) = 1/\{1 + e^{\{-z\}}\}.]$$

The output lies between 0 and 1 and is interpreted as the probability of disease.

3. Generalized Methodology

Stochastic Mathematical Model and Machine Learning Algorithm for Early Disease Detection

Step 1: Data Collection

Collect patient data from hospitals or generate synthetic (random) data.

Let $[D = \{(X_i, Y_i)\}_{\{i=1\}}^{\{N\}}]$

where

$$(X_i) = Patient\ feature\ vector$$

$$(Y_i) = Disease\ label$$

Step 2: Data Preprocessing

Handle missing values.

Normalize features.

Remove outliers.

Split data:

Training (80%)

Testing (20%)

Step 3: Stochastic Mathematical Model

Let [X(t)] denote the disease progression variable.

For diabetes, [X(t)={Blood glucose}.]

For cancer,[X(t)={Tumor size}.]

The generalized stochastic model is

$$[dX(t) = f(X, t, \theta)dt + \sigma\, X(t)dW(t),]$$

Where $(f(X, t, \theta))$ = deterministic disease dynamics,

(θ) = model parameters,

(σ) = stochastic intensity,

(W(t)) = Brownian motion.

This equation predicts disease evolution under uncertainty.

Step 4: Generate Dynamic Features

The stochastic model provides additional information such as:

Predicted future disease value

Growth rate

Variance

Disease risk score

These become new input features.

The updated feature vector is

$$[\hat{X} = \left[x_1, x_2, \dots, x_n, X_{\{\{pred\}\}}, \{Risk\}, \{Variance\}\right]]$$

Step 5: Machine Learning Model

Train a classifier using

ANN

Random Forest

Support Vector Machine

XGBoost

The ANN architecture may be

Input layer

Hidden Layer 1 (32 neurons)

Hidden Layer 2 (16 neurons)

Output Layer (Sigmoid)

Loss function

$$[L = 1/N \sum_{\{i=1\}}^{\{N\}} \left(Y_i - \widehat{Y}_i\right)^2]$$

Step 6: Disease Prediction

The trained network predicts [P({Disease})]

Decision rule

If [P>0.5] Disease Positive

Else Healthy

Step 7: Performance Evaluation

Evaluate using Accuracy

$$\left[Accuracy = \frac{\{TP + TN\}}{\{TP + TN + FP + FN\}}\right]$$

Step 8: Clinical Decision

Patients with high probability receive

Additional laboratory tests

MRI/CT Scan

Biopsy (Cancer)

HbA1c test (Diabetes)

Lifestyle intervention

Early treatment

4. Examples

Example 1. Diabetes Prediction from Routine Health Checkups-Real-life scenario

A 38-year-old office worker has

Age = 38

BMI = 31

Family history of diabetes

Fasting glucose = 108 mg/dL

Sedentary lifestyle

Although diabetes has not yet developed, there is uncertainty in disease progression.

A stochastic model estimates the probability that the patient develops diabetes within the next five years.

A machine learning model analyzes thousands of previous patient records and predicts > Probability of diabetes = 87%

The doctor recommends

Weight reduction

Exercise

Diet modification

As a result, diabetes is prevented before it develops.

Stochastic Mathematical Model and Artificial Neural Network for Early Detection of Diabetes

Step 1: Generate Random Patient Data

Assume we collect data from 20 patients.

Patient	Age	BMI	Blood Glucose (mg/dL)	Blood Pressure	Family History	Exercise (hrs/week)	Diabetes
1	28	22.1	89	118	0	5	0
2	35	25.7	102	122	1	3	0
3	41	29.5	128	140	1	1	1
4	52	31.8	145	150	1	0	1
5	46	30.2	132	138	0	2	1
…	…	…	…	…	…	…	…
20	27	22.6	90	115	0	6	0

Output variable

0 = Healthy

1 = Diabetic

Step 2: Stochastic Mathematical Model

Instead of assuming glucose increases deterministically, we model its random fluctuations.

Let G(t)={Blood glucose level}.

A simple stochastic differential equation is

$$dG(t) = \alpha\, G(t)dt + \sigma\, G(t)dW(t),$$

where:

(G(t)) = glucose level,

(α) = average growth rate,

(σ) = random variability,

(W(t)) = Brownian motion.

Suppose

$(\alpha = 0.03)$

$(\sigma = 0.15)$

The model predicts future glucose values while accounting for random effects such as diet, stress, illness, and exercise.

Step 3: Simulated Prediction

Suppose the stochastic model predicts the following glucose levels after one year.

Patient	Current	Predicted
1	89	91
2	102	108

3	128	142
4	145	168
5	132	148
….	…	…

Patients whose predicted glucose exceeds 126 mg/dL are classified as high risk.

Step 4: Train an ANN

Use the following input variables:

X=[{Age},{BMI},{Glucose},{Blood Pressure},{Family History},{Exercise}]

Output

$$Y = \begin{cases} 0 & \text{Healthy} \\ 1 & \text{Disease} \end{cases}$$

A simple ANN architecture is:

Input layer: 6 neurons

Hidden layer 1: 16 neurons (ReLU)

Hidden layer 2: 8 neurons (ReLU)

Output layer: 1 neuron (Sigmoid)

Step 5: ANN Predictions

Patient	Actual	Predicted Probability
1	0	0.03
2	0	0.15
3	1	0.89
4	1	0.98
5	1	0.93
6	0	0.02
…	…	…

Using a threshold of 0.5:

Probability > 0.5 → Diabetic

Probability ≤ 0.5 → Healthy

Step 6: Performance Evaluation

Suppose the ANN achieves:

Accuracy - 96%

Precision - 95%

Recall - 97%

F1-score - 96%

AUC - 0.98

These values are representative of a well-performing model on synthetic data.

Step 7: Interpretation

The stochastic model captures the random progression of glucose due to lifestyle and biological variability.

The ANN learns complex nonlinear relationships among patient features.

Together, they identify individuals at high risk before diabetes is clinically diagnosed, enabling earlier interventions.

Example 2: Breast Cancer Early Detection -Real-life scenario

A 45-year-old woman undergoes an annual mammogram. The image appears normal to the naked eye, but tiny abnormalities are present.

Stochastic model: Models the random growth of cancer cells because tumor growth varies among patients due to genetics and environment.

Machine learning algorithm (CNN, Random Forest, XGBoost): Detects subtle image patterns indicating an early-stage tumor.

Outcome: The patient is diagnosed at Stage I instead of Stage III, greatly improving survival chances.

Step 1: Generate Random Patient Data

Assume data from 20 women undergoing routine screening.

Patient	Age	Tumor Size (cm)	Cell Density	Family History	Smoking	Gene Mutation	Diagnosis
1	32	0.20	120	0	0	0	0
2	45	0.60	180	1	0	1	1
3	51	1.20	220	1	1	1	1
4	38	0.35	140	0	0	0	0
5	59	2.10	310	1	1	1	1
6	41	0.45	160	0	0	0	0
…	…	…	…	…	…	…	…

Here,

Diagnosis = 0 → Benign (No cancer)

Diagnosis = 1 → Malignant (Cancer)

Step 2: Introduce Stochastic Modeling

Tumor growth is inherently random because it depends on many uncertain biological factors such as immune response, mutations, and treatment.

Let

T(t)={Tumor size}.

A stochastic model can be written as

$$dT(t) = rT(t)(1 - \frac{\{T(t)\}}{\{K\}})dt + \sigma\, T(t)dW(t),$$

where

(r) = tumor growth rate,

(K) = carrying capacity,

(σ) = random biological fluctuations,

(W(t)) = Brownian motion.

This model predicts future tumor size under uncertainty.

Step 3: Predicted Tumor Growth

Suppose the stochastic model predicts tumor sizes after six months.

Patient	Current (cm)	Predicted (cm)
1	0.20	0.25
2	0.60	0.95
3	1.20	1.90
4	0.35	0.40
5	2.10	3.40

Patients whose predicted tumor size grows rapidly are considered high risk.

Step 4: Train an ANN

Use the following inputs:

X=[{Age},{Tumor Size},{Cell Density},{Family History},{Smoking},{Gene Mutation}]

Output

$$Y = \begin{cases} 0 & \text{Benign} \\ 1 & \text{Malignant} \end{cases}$$

The ANN learns patterns associated with malignancy.

Step 5: ANN Prediction

Patient	Actual	Cancer Probability
1	0	0.02
2	1	0.84
3	1	0.97
4	0	0.08
5	1	0.99

Using a threshold of 0.5:

Probability > 0.5 → Malignant

Probability ≤ 0.5 → Benign

Step 6: Early Detection

Consider a 44-year-old patient with:

Tumor size = 0.7 cm

Cell density = 195

Family history = Yes

Smoking = No

Gene mutation = Yes

The stochastic model predicts that the tumor may grow to 1.1 cm in six months, and the ANN estimates a 91% probability of malignancy. Even though the tumor is still relatively small, the combined model recommends further evaluation (such as biopsy or advanced imaging), potentially enabling earlier diagnosis.

5. Conclusion

This study presents a generalized hybrid framework that combines stochastic mathematical modeling with machine learning algorithms for the early detection of cancer and diabetes. By incorporating stochastic differential equations, the proposed methodology effectively captures the uncertainty and variability associated with disease progression, while the Artificial Neural Network learns complex nonlinear relationships from clinical and stochastic features. The proposed framework demonstrates how stochastic disease trajectories can enhance predictive performance beyond conventional data-driven methods. Moreover, its generalized structure allows the same computational framework to be adapted to different chronic diseases by modifying only the disease-specific variables and governing equations. This flexibility makes the proposed methodology suitable for intelligent healthcare applications, clinical decision support systems, and personalized disease prediction. Overall, the integration of stochastic mathematical models with machine learning provides a promising direction for developing more reliable, interpretable, and robust diagnostic tools for early disease detection.

References

1. Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Artificial neural network based breast cancer screening: A comprehensive review. https://arxiv.org/abs/2006.01767
2. Lee, W. S., Jo, J., & Song, T. (2020). Machine learning for the diagnosis of early stage diabetes using temporal glucose profiles.https://arxiv.org/abs/2005.08701
3. Silva, K. D., Lee, W. K., Forbes, A., Demmer, R. T., Barton, C., & Enticott, J. (2020). Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis. International Journal of Medical Informatics, 143, 104268. https://doi.org/10.1016/j.ijmedinf.2020.104268
4. Comparative analysis of machine learning and deep learning models for improved cancer detection: A comprehensive review of recent advancements in diagnostic techniques. (2024). Expert Systems with Applications, 255, 124838. https://doi.org/10.1016/j.eswa.2024.124838
5. A state-of-the-art review of artificial intelligence (AI) applications in healthcare: Advances in diabetes, cancer, epidemiology, and mortality prediction. (2025). Computers, 14(4), 143. https://doi.org/10.3390/computers14040143
6. Machine learning and artificial intelligence in type 2 diabetes prediction: A comprehensive 33-year bibliometric and literature analysis. (2025). Frontiers in Digital Health. https://doi.org/10.3389/fdgth.2025.1557467
7. Recent trends in diabetes mellitus diagnosis: An in-depth review of artificial intelligence-based techniques. (2025). Diabetes Research and Clinical Practice. https://doi.org/10.1016/j.diabres.2025.112221
8. Machine learning models for early prediction of type 1 diabetes: A systematic review. (2025). PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC13197420/
9. Applications of artificial intelligence and machine learning in prediabetes: A scoping review. (2025). PubMed. https://pubmed.ncbi.nlm.nih.gov/40626375/